Springer Series in Statistics

Advisors:
D. Brillinger, S. Fienberg, J. Gani, J. Hartigan,
J. Kiefer, K. Krickeberg

Springer Series in Statistics

Measures of Association for Cross-Classifications
Leo A. Goodman and **William H. Kruskal**
1979 / 146 pp. / 12 illus. / cloth / ISBN 0-387-**90443**-3

Statistical Decision Theory: Foundations, Concepts, and Methods
James O. Berger
1980 / 425 pp. / 20 illus. / cloth / ISBN 0-387-**90471**-9

Simultaneous Statistical Inference, Second Edition
Rupert G. Miller, Jr.
1981 / 299 pp. / 25 illus. / cloth / ISBN 0-387-**90548**-0

Point Processes and Queues: Martingale Dynamics
Pierre Brémaud
1981 / 354 pp. / 31 illus. / cloth / ISBN 0-387-**90536**-7

Non-negative Matrices and Markov Chains, Second Edition
E. Seneta
1981 / 279 pp. / cloth / ISBN 0-387-**90598**-7

Computing in Statistical Science through APL
Francis John Anscombe
1981 / 426 pp. / 70 illus. / cloth / ISBN 0-387-**90549**-9

Concepts of Nonparametric Theory
John W. Pratt and **Jean D. Gibbons**
1981 / 462 pp. / 23 illus. / cloth / ISBN 0-387-**90582**-0

Estimation of Dependences based on Empirical Data
Vladimir Vapnik
1982 / xvi, 399 pp. / 22 illus. / cloth / ISBN 0-387-**90733**-5

Theory of Statistical Experiments
H. Heyer
1982 / x, 289 pp. / cloth / ISBN 0-387-**90785**-8

Applied Statistics: A Handbook of Techniques
Lothar Sachs
1982 / xxvii, 706 pp. / 59 illus. / cloth / ISBN 0-387-**90558**-8

Extremes and Related Properties of Random Sequences and Processes
M.R. Leadbetter, Georg Lindgren, and **Holger Rootzén**
1983 / xii, 336 pp. / 28 illus. / cloth / ISBN 0-387-**90731**-9

M. R. Leadbetter
Georg Lindgren
Holger Rootzén

EXTREMES
AND RELATED PROPERTIES OF RANDOM SEQUENCES AND PROCESSES

With 28 Illustrations

Springer-Verlag
New York Heidelberg Berlin

M. R. Leadbetter
Department of Statistics
The University of North Carolina
Chapel Hill, NC 27514
U.S.A.

Georg Lindgren
Department of Mathematical Statistics
University of Lund
Box 725
S-220 07 Lund
Sweden

Holger Rootzén
Institute of Mathematical Statistics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark

AMS Classifications (1980): 60G10, 60G15, 60G55, 60F05, 62N99

Library of Congress Cataloging in Publication Data
Leadbetter, M. R.
Extremes and related properties of random sequences and processes.

(Springer series in statistics)
1. Stochastic processes. 2. Extreme value theory.
I. Lindgren, Georg, 1940– II. Rootzén, Holger.
III. Title. IV. Series.
QA274.L4 1982 519.2 82-16802

Softcover reprint of the hardcover 1st edition 1983

Typeset by Composition House Ltd., Salisbury, England.

9 8 7 6 5 4 3 2 1

ISBN-13: 978-1-4612-5451-5 e-ISBN-13: 978-1-4612-5449-2
DOI: 10.1007/978-1-4612-5449-2

Preface

Classical Extreme Value Theory—the asymptotic distributional theory for maxima of independent, identically distributed random variables—may be regarded as roughly half a century old, even though its roots reach further back into mathematical antiquity. During this period of time it has found significant application—exemplified best perhaps by the book *Statistics of Extremes* by E. J. Gumbel—as well as a rather complete theoretical development.

More recently, beginning with the work of G. S. Watson, S. M. Berman, R. M. Loynes, and H. Cramér, there has been a developing interest in the extension of the theory to include, first, dependent sequences and then continuous parameter stationary processes. The early activity proceeded in two directions—the extension of general theory to certain dependent sequences (e.g., Watson and Loynes), and the beginning of a detailed theory for stationary sequences (Berman) and continuous parameter processes (Cramér) in the normal case.

In recent years both lines of development have been actively pursued. It has proved possible to unify the two directions and to give a rather complete and satisfying general theory along the classical lines, including the known results for stationary normal sequences and processes as special cases. A principal aim of this work is to present this theory in as complete and up-to-date a form as possible, alongside a reasonably comprehensive discussion of the classical case. The treatment is thus unified with regard to both the classical and dependent cases, and also in respect to consideration of normal and more general stationary sequences and processes.

Closely related to the properties of extremes are those of exceedances and upcrossings of high levels, by sequences and continuous parameter processes. By regarding such exceedances and upcrossings as point processes, one may obtain some quite general results demonstrating convergence to Poisson and related point processes. A number of interesting results follow concerning the asymptotic behaviour of the magnitude and location of such quantities as the kth largest maxima (or local maxima, in the continous setting). These and a number of other related topics have been taken up, especially for continuous parameter cases.

The volume is organized in four parts. Part I provides a reasonably comprehensive account of the central distributional results of classical extreme value theory—surrounding the Extremal Types Theorem. We have attempted to make this quite straightforward, using relatively elementary methods, and to highlight the main ideas on which the later extensions to dependent cases are based.

Part II contains the basic extension of the classical theory applying to stationary sequences and to some important nonstationary cases. The main key to this work is the appropriate restriction of dependence between widely separated members of the sequence, so that the classical limits still hold. Normal sequences are particularly emphasized and provide illuminating examples of the roles played by the various assumptions.

In Part III we turn to continuous parameter cases. The emphasis in this part is on stationary normal processes, which, for clarity, we treat in some detail before giving the general theory surrounding the Extremal Types Theorem. In addition to extremal theory, this part concerns properties of local maxima, point processes of upcrossings, models for local behaviour, and related topics.

Finally, Part IV contains specific applications of (and small extensions to) the theory for particular, real situations. Since the theory largely predicts the same extremal behaviour as in the classical case, there is limited usefulness in providing data which simply illustrate this well. Rather, we have tried to grapple with typical practical issues and problems which arise in putting theory to work. We have not attempted systematic case studies, but have primarily selected examples which involve interesting facets, and raise issues that demand thoughtful consideration.

Many of the results given here have appeared in print in various forms, but a number are hitherto unpublished. Most of the contents of this work may be easily understood by a reader who has taken a (non-measure-theoretic) introductory graduate probability course. Possible exceptions include the material on point process convergence (the details being given in an appendix), but even for this we feel that a reader should be able to obtain a good intuitive understanding from the text.

It is indeed a pleasure to acknowledge the support of the U.S. Office of Naval Research, the Swedish Natural Science Research Council, the Danish Natural Science Research Council, and the Swedish Institute of Applied Mathematics, in much of the research leading to this work. We are also most grateful to Drs. Jacques de Maré and Jan Lanke for their help and suggestions on various aspects of this project and to Dr. Olav Kallenberg for his illuminating comments. We thank a number of readers including J. Castellana, B. Collings, R. Frimmel, N. Gerr, P. Hougaard, D. Kikuchi, I. McKeague, E. Murphree, and I. Skovgaard for their suggestions for improvement in clarity of the text; Ruth Bahr, Anitha Bergdahl, Betty Blake, Dagmar Jensen, Ingalill Karlsson, Anna Morén, Beatrice Tuma, and Ingrid Westerberg for their splendid typing of this volume and its earlier manuscript versions, and Sten Lindgren for preparing several of the figures.

Contents

PART I

CLASSICAL THEORY OF EXTREMES

Classical extreme value theory is concerned substantially with distributional properties of the maximum

$$M_n = \max(\xi_1, \xi_2, \ldots, \xi_n)$$

of n independent and identically distributed random variables, as n becomes large. In Part I we have attempted to give a relatively comprehensive account of the central distributional results of the classical theory, using the simplest available proofs, and emphasizing their general features which lead to subsequent extensions to dependent situations.

Two results of basic importance are proved in Chapter 1. The first is the fundamental result—here called the Extremal Types Theorem—which exhibits the possible limiting forms for the distribution of M_n under linear normalizations. More specifically, this basic classical result states that if for some sequences of normalizing constants $a_n > 0$, b_n, $a_n(M_n - b_n)$ has a nondegenerate limiting distribution function $G(x)$, then G must have one of just three possible "forms". The three "extreme value distributions" involved were discovered by Fréchet, and Fisher and Tippett, and discussed more completely later by Gnedenko. Here we use more recent proofs, substantially simplified by the use of techniques of de Haan.

The second basic result given in Chapter 1 is almost trivial in the independent context, and gives a simple necessary and sufficient condition under which $P\{M_n \le u_n\}$ converges, for a given sequence of constants $\{u_n\}$. This result plays an important role here and also in dependent cases, where it is by no means as trivial but still holds under appropriate conditions. Its importance will be seen in Chapter 1 in the development of the classical theory given there for the domains of attraction to the three extreme value types. The theory is illustrated by several examples from each of the possible limiting types and the chapter is concluded with a brief corresponding discussion of minima.

The theme of Chapter 2 is the corresponding limiting distributions for the kth largest $M_n^{(k)}$ of $\xi_1, \ldots, \xi_n$, where k may be fixed or tend to infinity with n. The case for fixed k (when $M_n^{(k)}$ is an "extreme order statistic") is of primary concern and is discussed by means of asymptotic Poisson properties of the exceedances of high levels by the sequence $\xi_1, \xi_2, \ldots$. These properties, which here involve simply the convergence of binomial to Poisson distributions will recur in more interesting and sophisticated forms in the later parts of the volume. Rather efficient and transparent estimates for the rate of convergence in the limit theorems are also presented in this chapter.

Finally, some description is given of the available theory for cases when $k = k_n$ tends to infinity with n (involving "central" and "intermediate" order statistics). This discussion is included for completeness only and will not be developed in the subsequent dependent context.

CHAPTER 1
Asymptotic Distributions of Extremes

This chapter is primarily concerned with the central result of classical extreme value theory—the Extremal Types Theorem—which specifies the possible forms for the limiting distribution of maxima in sequences of i.i.d. random variables. In the derivation, the possible limiting distributions are identified with a class having a certain stability property—the so-called *max-stable* distributions. It is further shown that this class consists precisely of the three families known (loosely) as *the* three *extreme value distributions*.

1.1. Introduction and Framework

Let $\xi_1, \xi_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s) and write M_n for the *maximum* of the first n, i.e.

$$M_n = \max(\xi_1, \xi_2, \ldots, \xi_n). \tag{1.1.1}$$

Then much of "classical" extreme value theory deals with the distribution of M_n, and especially with its properties as $n \to \infty$. All results obtained for maxima of course lead to analogous results for minima through the obvious relation $m_n = \min(\xi_1, \ldots, \xi_n) = -\max(-\xi_1, \ldots, -\xi_n)$. We shall therefore only briefly discuss minima explicitly in this work, except where its joint distribution with M_n is considered.

There is, of course, no difficulty in writing down the distribution function (d.f.) of M_n exactly in this situation; it is

$$P\{M_n \le x\} = P\{\xi_1 \le x, \xi_2 \le x, \ldots, \xi_n \le x\} = F^n(x), \tag{1.1.2}$$

where F denotes the common d.f. of the ξ_i. Much of the *Statistics of Extremes* (as is the title of Gumbel's book (1958)) deals with the distribution of M_n in a variety of useful cases and with a multitude of related questions (for example, concerning other order statistics, range of values, and so on).

In view of such a satisfactory edifice of theory in finite terms, one may question the desirability of probing for asymptotic results. One reason for such a study appears to us to especially justify it. In simple central limit theory, one obtains an asymptotic normal distribution for the *sum* of many i.i.d. random variables whatever their common original d.f. Indeed, one does not have to know the d.f. too precisely to apply the asymptotic theory. A similar situation holds in extreme value theory, and in fact, a nondegenerate asymptotic distribution of M_n (normalized) must belong to one of just three possible general families, regardless of the original d.f. F. Further, it is not necessary to know the detailed nature of F, to know which limiting form (if any) it gives rise to, i.e. to which "domain of attraction" it belongs. In fact, this is determined just by the behaviour of the tail of $F(x)$ for large x, and so a good deal may be said about the asymptotic properties of the maximum based on rather limited knowledge of the properties of F.

The central result—here referred to as The Extremal Types Theorem—was discovered first by Fisher and Tippett (1928) and later proved in complete generality by Gnedenko (1943). We shall prove this result in the i.i.d. context (Theorem 1.4.2) using a more recent, simple approach due to de Haan (1976) and later extend it to dependent situations in Chapters 3 and 13.

We shall be concerned with conditions under which, for suitable normalizing constants $a_n > 0, b_n$,

$$P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x) \tag{1.1.3}$$

(by which we mean that convergence occurs at continuity points of G—though we shall see later that the G's of interest are all continuous). In particular, we shall be interested in determining which d.f.'s G may appear as such a limit. It will be shown that the possible nondegenerate d.f.'s G which may occur as limits in (1.1.3) form precisely the class of *max-stable* distributions discussed in Section 1.3. We shall see further that every max-stable distribution G has (up to location and scale changes) one of the following three parametric forms—commonly (and somewhat loosely) called *the* three *Extreme Value Distributions.*

$$\text{Type I:}\quad G(x) = \exp(-e^{-x}), \qquad -\infty < x < \infty;$$

$$\text{Type II:}\quad G(x) = \begin{cases} 0, & x \le 0, \\ \exp(-x^{-\alpha}), \quad \text{for some } \alpha > 0, & x > 0; \end{cases}$$

$$\text{Type III:}\quad G(x) = \begin{cases} \exp(-(-x)^{\alpha}), \quad \text{for some } \alpha > 0, & x \le 0, \\ 1, & x > 0. \end{cases}$$

A central result to be used in the development of the theory is a general theorem of Khintchine concerning convergence of distribution functions. Although it is obtainable from other sources, we shall, for completeness, prove this important result in Section 1.2, along with some useful results concerning inverses of monotone functions. In Section 1.3 the d.f.'s G which may occur as limits in (1.1.3) are identified with the max-stable distributions (defined in that section), prior to the further identification with the extreme value distributions in Section 1.4, which will complete the proof of the Extremal Types Theorem.

By (1.1.2), (1.1.3) may be written as

$$F^n(a_n^{-1}x + b_n) \xrightarrow{w} G(x), \tag{1.1.4}$$

where again the notation $\xrightarrow{w}$ denotes convergence at continuity points of the limiting function. If (1.1.4) holds for some sequences $\{a_n > 0\}$, $\{b_n\}$, we shall say that F belongs to the (i.i.d.) *domain of attraction* (*for maxima*) of G and write $F \in D(G)$. Necessary and sufficient criteria are known, to determine which (if any) of the possible limiting distributions applies; that is, conditions under which $F \in D(G)$. These will be stated in Section 1.6, along with proofs of their sufficiency. (The proofs of the necessity of the conditions are somewhat lengthy and not germane to our purposes here and hence are omitted.) We also give some simple and useful sufficient conditions due to von Mises, which apply when the d.f. F has a density function—an obviously common case.

The discussion in Section 1.6 will be substantially based on a very simple convergence result treated in Section 1.5. This result gives conditions for the convergence of $P\{M_n \leq u_n\}$, where $\{u_n\}$ is any sequence of real constants. (In the case where (1.1.3) applies, we have such convergence for all members of the family of sequences $\{u_n = x/a_n + b_n\}$, as x ranges over all real numbers.)

As hinted above (and discussed more fully later), it may turn out that for a given d.f. F, there is *no* extreme value d.f. G such that $F \in D(G)$. This simply means that the maximum M_n does not have a nondegenerate limiting distribution under any *linear* normalization (a common example being the Poisson distribution, as will be seen later). On the other hand, limits of $P\{M_n \leq u_n\}$ may well be possible for interesting sequences u_n not necessarily of the form $x/a_n + b_n$ or even dependent on a parameter x.

This simple convergence result does, as noted, play an important role in connection with domains of attraction. However, it is also very important in the further theoretical development, both for the i.i.d. case considered in this and the next chapter and for the dependent sequences of later chapters.

Finally, Section 1.7 contains further examples and comments concerning particular cases, and Section 1.8 contains a brief discussion of minima.

1.2. Inverse Functions and Khintchine's Convergence Theorem

First, it will be convenient to obtain some useful results for inverses of monotone functions. Such inverses may be defined in a variety of ways, of which the following will best suit our purposes.

If $\psi(x)$ is a nondecreasing right continuous function, we define an inverse function ψ^{-1} on the interval $(\inf\{\psi(x)\}, \sup\{\psi(x)\})$ by

$$\psi^{-1}(y) = \inf\{x; \psi(x) \geq y\}.$$

(Note that the domain of ψ^{-1} is written as an open interval, but may be closed at either end if $\inf\{\psi(x)\}$ or $\sup\{\psi(x)\}$ is attained at finite x.)

Lemma 1.2.1. (i) *For ψ as above, if $a > 0$, b and c are constants, and $H(x) = \psi(ax + b) - c$, then $H^{-1}(y) = a^{-1}(\psi^{-1}(y + c) - b)$.*
(ii) *For ψ as above, if ψ^{-1} is continuous, then $\psi^{-1}(\psi(x)) = x$.*
(iii) *If G is a nondegenerate d.f., there exist $y_1 < y_2$ such that $G^{-1}(y_1) < G^{-1}(y_2)$ are well defined (and finite).*

PROOF. (i) We have

$$\begin{aligned} H^{-1}(y) &= \inf\{x; \psi(ax + b) - c \geq y\} \\ &= a^{-1}(\inf\{(ax + b); \psi(ax + b) \geq y + c\} - b) \\ &= a^{-1}(\psi^{-1}(y + c) - b), \end{aligned}$$

as required.

(ii) From the definition of ψ^{-1}, it is clear that $\psi^{-1}(\psi(x)) \leq x$. If strict inequality holds for some x, the definition of ψ^{-1} shows the existence of $z < x$ with $\psi(z) \geq \psi(x)$, and hence $\psi(z) = \psi(x)$ since ψ is nondecreasing. For $y = \psi(z) = \psi(x)$ we have $\psi^{-1}(y) \leq z$, whereas for $y > \psi(z) = \psi(x)$ we have $\psi^{-1}(y) \geq x$, contradicting the continuity of ψ^{-1}. Hence $\psi^{-1}(\psi(x)) = x$, as asserted.

(iii) If G is nondegenerate, there exist $x_1' < x_2'$ such that $0 < G(x_1') = y_1 < G(x_2') = y_2 \leq 1$. Clearly $x_1 = G^{-1}(y_1)$ and $x_2 = G^{-1}(y_2)$ are both well defined. Also $G^{-1}(y_2) \geq x_1'$ and equality would require $G(z) \geq y_2$ for all $z > x_1$ so that $G(x_1') = \lim_{\varepsilon \downarrow 0} G(x_1' + \varepsilon) = G(x_1' +) \geq y_2$, contradicting $G(x_1') = y_1$. Thus $G^{-1}(y_2) > x_1' \geq x_1 = G^{-1}(y_1)$, as required. □

Corollary 1.2.2. *If G is a nondegenerate d.f. and $a > 0$, $\alpha > 0$, b, and β are constants such that $G(ax + b) = G(\alpha x + \beta)$ for all x, then $a = \alpha$ and $b = \beta$.*

PROOF. Choose $y_1 < y_2$ and $-\infty < x_1 < x_2 < \infty$ by (iii) of the lemma so that $x_1 = G^{-1}(y_1)$, $x_2 = G^{-1}(y_2)$. Taking inverses of $G(ax + b)$ and $G(\alpha x + \beta)$ by (i) of the lemma, we have

$$a^{-1}(G^{-1}(y) - b) = \alpha^{-1}(G^{-1}(y) - \beta)$$

for all y. Applying this to y_1 and y_2 in turn, we obtain

$$a^{-1}(x_1 - b) = \alpha^{-1}(x_1 - \beta) \quad \text{and} \quad a^{-1}(x_2 - b) = \alpha^{-1}(x_2 - \beta),$$

from which it follows simply that $a = \alpha$ and $b = \beta$. □

We now obtain the promised general result of Khintchine.

Theorem 1.2.3 (Khintchine). *Let $\{F_n\}$ be a sequence of d.f.'s and G a nondegenerate d.f. Let $a_n > 0$ and b_n be constants such that*

$$F_n(a_n x + b_n) \xrightarrow{w} G(x). \tag{1.2.1}$$

Then for some nondegenerate d.f. G_* and constants $\alpha_n > 0$, β_n,

$$F_n(\alpha_n x + \beta_n) \xrightarrow{w} G_*(x) \tag{1.2.2}$$

if and only if

$$a_n^{-1}\alpha_n \to a \quad \text{and} \quad a_n^{-1}(\beta_n - b_n) \to b \tag{1.2.3}$$

for some $a > 0$ and b, and then

$$G_*(x) = G(ax + b). \tag{1.2.4}$$

PROOF. By writing $\alpha_n' = a_n^{-1}\alpha_n$, $\beta_n' = a_n^{-1}(\beta_n - b_n)$, and $F_n'(x) = F_n(a_n x + b_n)$, we may rewrite (1.2.1), (1.2.2), and (1.2.3) as

$$F_n'(x) \xrightarrow{w} G(x), \tag{1.2.1$'$}$$

$$F_n'(\alpha_n' x + \beta_n') \xrightarrow{w} G_*(x), \tag{1.2.2$'$}$$

$$\alpha_n' \to a \quad \text{and} \quad \beta_n' \to b \quad \text{for some } a > 0, b. \tag{1.2.3$'$}$$

If (1.2.1)$'$ and (1.2.3)$'$ hold, then obviously so does (1.2.2)$'$, with $G_*(x) = G(ax + b)$. Thus (1.2.1) and (1.2.3) imply (1.2.2) and (1.2.4).

The proof of the lemma will be complete if we show that (1.2.1)$'$ and (1.2.2)$'$ imply (1.2.3)$'$, for then (1.2.4) will also hold, as above.

Since G_* is assumed nondegenerate, there are two distinct points x' and x'' (which may be taken to be continuity points of G_*) such that $0 < G_*(x') < 1$, $0 < G_*(x'') < 1$.

The sequence $\{\alpha_n' x' + \beta_n'\}$ must be bounded. For if not, a sequence $\{n_k\}$ could be chosen so that $\alpha_{n_k}' x' + \beta_{n_k}' \to \pm\infty$, which by (1.2.1)$'$ (since G is a d.f.) would clearly imply that the limit of $F_{n_k}'(\alpha_{n_k}' x' + \beta_{n_k}')$ is zero or one—contradicting (1.2.2)$'$ for $x = x'$. Hence $\{\alpha_n' x' + \beta_n'\}$ is bounded, and similarly so is $\{\alpha_n' x'' + \beta_n'\}$, which together show that the sequences $\{\alpha_n'\}$ and $\{\beta_n'\}$ are each bounded.

Thus there are constants a and b and a sequence $\{n_k\}$ of integers such that $\alpha_{n_k}' \to a$ and $\beta_{n_k}' \to b$, and it follows as above that

$$F_{n_k}'(\alpha_{n_k}' x + \beta_{n_k}') \xrightarrow{w} G(ax + b), \tag{1.2.5}$$

whence since by (1.2.2)$'$, $G(ax + b) = G_*(x)$, a d.f., we must have $a > 0$. On the other hand, if another sequence $\{m_k\}$ of integers gave $\alpha_{m_k}' \to a' > 0$ and $\beta_{m_k}' \to b'$, we would have $G(a'x + b') = G_*(x) = G(ax + b)$, and hence $a' = a$ and $b' = b$ by Corollary 1.2.2. Thus $\alpha_n' \to a$ and $\beta_n' \to b$, as required to complete the proof. □

1.3. Max-Stable Distributions

We now identify the d.f.'s G which are possible limiting laws for maxima of i.i.d. sequences, i.e. which may appear in (1.1.3), with the class of so-called *max-stable* distributions. Specifically, we shall say that a nondegenerate d.f. G is *max-stable* if, for each $n = 2, 3, \ldots$, there are constants $a_n > 0$ and b_n such that $G^n(a_n x + b_n) = G(x)$.

Theorem 1.3.1. (i) *A nondegenerate d.f. G is max-stable if and only if there is a sequence $\{F_n\}$ of d.f.'s and constants $a_n > 0$ and b_n such that*

$$F_n(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G^{1/k}(x) \quad \text{as } n \to \infty \tag{1.3.1}$$

for each $k = 1, 2, \ldots$.

(ii) *In particular, if G is nondegenerate, $D(G)$ is nonempty if and only if G is max-stable. Then also $G \in D(G)$. Thus the class of nondegenerate d.f.'s G which appear as limit laws in* (1.1.3) (*for i.i.d.* $\xi_1, \xi_2, \ldots$) *coincides with the class of max-stable d.f.'s.*

Proof. (i) If G is nondegenerate, so is $G^{1/k}$ for each k, and if (1.3.1) holds for each k, Theorem 1.2.3 (with a_n^{-1} for a_n) implies that $G^{1/k}(x) = G(\alpha_k x + \beta_k)$ for some $\alpha_k > 0$ and β_k, so that G is max-stable. Conversely, if G is max-stable and $F_n = G^n$, we have $G^n(a_n^{-1}x + b_n) = G(x)$ for some $a_n > 0$ and b_n, and

$$F_n(a_{nk}^{-1}x + b_{nk}) = (G^{nk}(a_{nk}^{-1}x + b_{nk}))^{1/k} = (G(x))^{1/k},$$

so that (1.3.1) follows trivially.

(ii) If G is max-stable, $G^n(a_n x + b_n) = G(x)$ for some $a_n > 0$ and b_n, so (letting $n \to \infty$) we see that $G \in D(G)$. Conversely, if $D(G)$ is nonempty, $F \in D(G)$, say, with $F^n(a_n^{-1}x + b_n) \xrightarrow{w} G(x)$. Hence $F^{nk}(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G(x)$ or $F^n(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G^{1/k}(x)$. Thus (1.3.1) holds with $F_n = F^n$, and hence by (i), G is max-stable. □

Corollary 1.3.2. *If G is max-stable, there exist real functions $a(s) > 0$ and $b(s)$ defined for $s > 0$ such that*

$$G^s(a(s)x + b(s)) = G(x), \quad \text{all real } x, s > 0. \tag{1.3.2}$$

Proof. Since G is max-stable, there exist $a_n > 0$, b_n such that

$$G^n(a_n x + b_n) = G(x), \tag{1.3.3}$$

and hence (letting [ ] denote integer part)

$$G^{[ns]}(a_{[ns]}x + b_{[ns]}) = G(x).$$

But this is easily seen (e.g. by taking logarithms) to give

$$G^n(a_{[ns]}x + b_{[ns]}) \xrightarrow{w} G^{1/s}(x). \tag{1.3.4}$$

In view of the limit (1.3.4) and the (trivial) limit (1.3.3), and since $G^{1/s}$ is nondegenerate, Theorem 1.2.3 applies with $\alpha_n = a_{[ns]}$ and $\beta_n = b_{[ns]}$ to show that $G(a(s)x + b(s)) = G^{1/s}(x)$ for some $a(s) > 0$ and $b(s)$, as required. □

It is sometimes convenient to use a more technical sense for the word "type" than the descriptive use employed thus far. Specifically, we may say that two d.f.'s G_1, G_2 are of *the same type* if

$$G_2(x) = G_1(ax + b)$$

for some constants $a > 0, b$. Then the above definition of max-stable distributions may be rephrased as: "A nondegenerate d.f. G is max-stable if, for each $n = 2, 3, \ldots$, the d.f. G^n is of the same type as G."

Further, Theorem 1.2.3 shows that if $\{F_n\}$ is a sequence of d.f.'s with $F_n(a_n x + b_n) \xrightarrow{w} G_1$, $F_n(\alpha_n x + \beta_n) \xrightarrow{w} G_2$ ($a_n > 0$, $\alpha_n > 0$), then G_1 and G_2 are of the same type, provided they are nondegenerate. Clearly the d.f.'s may be divided into equivalence classes (which we call *types*) by saying that G_1 and G_2 are equivalent if $G_2(x) = G_1(ax + b)$ for some $a > 0, b$.

If G_1 and G_2 are d.f.'s of the same type ($G_2(x) = G_1(ax + b)$) and $F \in D(G_1)$, i.e. $F^n(a_n x + b_n) \xrightarrow{w} G_1$ for some $a_n > 0$, b_n, then (1.2.3) is satisfied with $\alpha_n = a_n a$, $\beta_n = b_n + a_n b$, so that $F^n(\alpha_n x + \beta_n) \xrightarrow{w} G_2(x)$ by Theorem 1.2.3, and hence $F \in D(G_2)$. Thus if G_1 and G_2 are of the same type, $D(G_1) = D(G_2)$. Similarly, we may see from Theorem 1.2.3 that if F belongs to both $D(G_1)$ and $D(G_2)$, then G_1 and G_2 are of the same type. Hence $D(G_1)$ and $D(G_2)$ are identical if G_1 and G_2 are of the same type, and disjoint otherwise. That is, the domain of attraction of a d.f. G depends only on the type of G.

1.4. Extremal Types Theorem

Our final task in obtaining the possible limit laws for maxima (in the sense of (1.1.3)) is to show that the max-stable distributions are simply the extreme value distributions listed in Section 1.1. More precisely, we show that a d.f. is max-stable if and only if it is of the same type as one of the extreme value distributions listed.

A d.f. of the same type as $\exp(-e^{-x})$ (i.e. $\exp\{-e^{-(ax+b)}\}$ for some $a > 0, b$) will be said to be of Type I. Similarly, we will say that a d.f. is of Type II (or Type III) if it has the form $G(ax + b)$, where G is the Type II (or Type III) extreme value distribution listed in Section 1.1. Since the parameter α may change, Types II and III distributions are really families of types within our technical meaning of "type", but obviously no confusion will arise from the customary habit of referring to "three extreme value types". The following theorem contains the principal identifications desired.

Theorem 1.4.1. *Every max-stable distribution is of extreme value type, i.e. equal to $G(ax + b)$ for some $a > 0$ and b, where for*

$$\textit{Type I:}\quad G(x) = \exp(-e^{-x}), \qquad -\infty < x < \infty;$$

$$\textit{Type II:}\quad G(x) = \begin{cases} 0, & x \le 0, \\ \exp(-x^{-\alpha}), \ \textit{for some } \alpha > 0, & x > 0; \end{cases}$$

$$\textit{Type III:}\quad G(x) = \begin{cases} \exp(-(-x)^{\alpha}), \ \textit{for some } \alpha > 0, & x \le 0, \\ 1, & x > 0. \end{cases}$$

Conversely, each distribution of extreme value type is max-stable.

PROOF. The converse is clear since, e.g. for Type I,

$$(\exp\{-e^{-(ax+b)}\})^n = \exp\{-e^{-(ax+b-\log n)}\},$$

with similar expressions for Types II and III.

To prove the direct assertion, we follow essentially the proof of de Haan (1976). If G is max-stable, then (1.3.2) holds for all $s > 0$ and all x. If $0 < G(x) < 1$, (1.3.2) gives

$$-s \log G(a(s)x + b(s)) = -\log G(x),$$

so that

$$-\log(-\log G(a(s)x + b(s))) - \log s = -\log(-\log G(x)).$$

Now it is easily seen from the max-stable property with $n = 2$, that G can have no jump at any finite (upper or lower) endpoint. Thus the nondecreasing function $\psi(x) = -\log(-\log G(x))$ is such that $\inf\{\psi(x)\} = -\infty$, $\sup\{\psi(x)\} = +\infty$ and hence has an inverse function $U(y)$ defined for all real y. Further,

$$\psi(a(s)x + b(s)) - \log s = \psi(x),$$

so that by Lemma 1.2.1(i),

$$\frac{U(y + \log s) - b(s)}{a(s)} = U(y).$$

Subtracting this for $y = 0$ we have

$$\frac{U(y + \log s) - U(\log s)}{a(s)} = U(y) - U(0),$$

and by writing $z = \log s$, $\tilde{a}(z) = a(e^z)$, and $\tilde{U}(y) = U(y) - U(0)$,

$$\tilde{U}(y + z) - \tilde{U}(y) = \tilde{U}(y)\tilde{a}(z) \tag{1.4.1}$$

for all real y, z.

Interchanging y and z and subtracting, we obtain

$$\tilde{U}(y)(1 - \tilde{a}(z)) = \tilde{U}(z)(1 - \tilde{a}(y)). \tag{1.4.2}$$

Two cases are possible, (a) and (b) as follows.

(a) $\tilde{a}(z) = 1$ for all z when (1.4.1) gives

$$\tilde{U}(y + z) = \tilde{U}(y) + \tilde{U}(z).$$

The only monotone increasing solution to this is well known to be simply $\tilde{U}(y) = \rho y$ for some $\rho > 0$, so that $U(y) - U(0) = \rho y$ or

$$\psi^{-1}(y) = U(y) = \rho y + v, \qquad v = U(0).$$

Since this is continuous, Lemma 1.2.1(ii) gives

$$x = \psi^{-1}(\psi(x)) = \rho\psi(x) + v$$

or $\psi(x) = (x - v)/\rho$, so that $G(x) = \exp\{-e^{-(x-v)/\rho}\}$ when $0 < G(x) < 1$.

As noted above G can have no jump at any finite (upper or lower) endpoint and hence has the above form for all x, thus being of Type I.

(b) $\tilde{a}(z) \neq 1$ for some z when (1.4.2) gives

$$\tilde{U}(y) = \frac{\tilde{U}(z)}{1 - \tilde{a}(z)}(1 - \tilde{a}(y)) = c(1 - \tilde{a}(y)), \quad \text{say}, \tag{1.4.3}$$

where $c = \tilde{U}(z)/(1 - \tilde{a}(z)) \neq 0$ (since $\tilde{U}(z) = 0$ would imply $\tilde{U}(y) = 0$ for all y, and hence $U(y) = U(0)$, constant).

From (1.4.1), we thus obtain

$$c(1 - \tilde{a}(y + z)) - c(1 - \tilde{a}(z)) = c(1 - \tilde{a}(y))\tilde{a}(z),$$

which gives $\tilde{a}(y + z) = \tilde{a}(y)\tilde{a}(z)$. But $\tilde{a}$ is monotone (from (1.4.3)), and the only monotone nonconstant solutions of this functional equation have the form $\tilde{a}(y) = e^{\rho y}$ for $\rho \neq 0$. Hence (1.4.3) yields

$$\psi^{-1}(y) = U(y) = v + c(1 - e^{\rho y})$$

(where $v = U(0)$). Since $-\log(-\log G(x))$ is increasing, so is U, so that we must have $c < 0$ if $\rho > 0$ and $c > 0$ if $\rho < 0$. By Lemma 1.2.1(ii),

$$x = \psi^{-1}(\psi(x)) = v + c(1 - e^{\rho\psi(x)}) = v + c(1 - (-\log G(x))^{-\rho}),$$

giving, where $0 < G(x) < 1$,

$$G(x) = \exp\left\{-\left(1 - \frac{x - v}{c}\right)^{-1/\rho}\right\}.$$

Again, from continuity of G at any finite endpoints, we see that G is of Type II or Type III, with $\alpha = +1/\rho$ or $-1/\rho$, according as $\rho > 0$ ($c < 0$) or $\rho < 0$ ($c > 0$). □

The main result now follows immediately.

Theorem 1.4.2 (Extremal Types Theorem). *Let* $M_n = \max(\xi_1, \xi_2, \ldots, \xi_n)$, *where* ξ_i *are i.i.d. random variables. If for some constants* $a_n > 0$, b_n, *we have*

$$P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x) \tag{1.4.4}$$

for some nondegenerate G, then G is one of the three extreme value types listed above. Conversely, each d.f. G of extreme value type may appear as a limit in (1.4.4) *and, in fact, appears when G itself is the d.f. of each ξ_i.*

PROOF. If (1.4.4) holds, Theorem 1.3.1 shows that G is max-stable and, hence by Theorem 1.4.1, is of extreme value type. Conversely, if G is of extreme value type, it is max-stable by Theorem 1.4.1, and Theorem 1.3.1(ii) shows that $G \in D(G)$, yielding the conclusions stated. □

Looking ahead, if $\xi_1, \xi_2, \ldots$ are not necessarily independent, but $M_n = \max(\xi_1, \ldots, \xi_n)$ has an asymptotic distribution G in the sense of (1.4.4), then (1.3.1) holds with $k = 1$, where F_n is the d.f. of M_n. If one can show that if (1.3.1) holds for $k = 1$ then it holds for all k, it will follow that G is max-stable by Theorem 1.3.1(i) and hence that G is an extreme value type. Thus our approach when we consider dependent cases will be simply to show that, under appropriate assumptions, the truth of (1.3.1) for $k = 1$ implies its truth for all k, from which the Extremal Types Theorem will again follow.

Returning now to the i.i.d. case, we note again that Theorem 1.4.2 *assumes* that $a_n(M_n - b_n)$ has a nondegenerate limiting d.f. G and then proves that G must have one of the three stated forms. It is easy to construct i.i.d. sequences $\{\xi_n\}$ for which *no* such G exists. To see a simple example, it will be convenient here (and subsequently) to use the notation x_F for the right endpoint of a d.f. F, i.e.

$$x_F = \sup\{x; F(x) < 1\} \qquad (\leq \infty).$$

That is, $F(x) < 1$ for all $x < x_F$ and $F(x) = 1$ for all $x \geq x_F$.

Suppose that each ξ_n has d.f. F which is such that $x_F < \infty$ and that F has a jump at x_F, i.e. $F(x_F-) < 1 = F(x_F)$. Then it follows readily (as will be seen in Corollary 1.5.2) that if $\{u_n\}$ is any sequence and $P\{M_n \leq u_n\} \to \rho$, then $\rho = 0$ or 1. Thus if $P\{a_n(M_n - b_n) \leq x\} \to G(x)$, it follows, by taking $u_n = x/a_n + b_n$, that $G(x) = 0$ or 1 for each x, so that G is degenerate.

Other even more common examples, such as the case where each ξ_n is Poisson, will be discussed in Section 1.7.

1.5. Convergence of $P\{M_n \leq u_n\}$

We have been considering convergence of probabilities of the form $P\{a_n(M_n - b_n) \leq x\}$, which may be rewritten as $P\{M_n \leq u_n\}$ where $u_n = u_n(x) = x/a_n + b_n$. The convergence was required for every x. On the other hand, it is also of interest to consider sequences $\{u_n\}$ which may simply not depend on any parameter x or may be more complicated functions than the linear one considered above. The following theorem is almost trivial in the i.i.d. context, but nevertheless is very useful (as we shall see in the next section), and will extend in important ways to apply to dependent (stationary) sequences and to continuous time processes.

Theorem 1.5.1. *Let $\{\xi_n\}$ be an i.i.d. sequence. Let $0 \le \tau \le \infty$ and suppose that $\{u_n\}$ is a sequence of real numbers such that*

$$n(1 - F(u_n)) \to \tau \quad \text{as } n \to \infty. \tag{1.5.1}$$

Then

$$P\{M_n \le u_n\} \to e^{-\tau} \quad \text{as } n \to \infty. \tag{1.5.2}$$

Conversely, if (1.5.2) *holds for some τ, $0 \le \tau \le \infty$, then so does* (1.5.1).

PROOF. Suppose first that $0 \le \tau < \infty$. If (1.5.1) holds, then

$$P\{M_n \le u_n\} = F^n(u_n) = \{1 - (1 - F(u_n))\}^n \tag{1.5.3}$$

may be written as $(1 - \tau/n + o(1/n))^n$, so that (1.5.2) follows at once.

Conversely, if (1.5.2) holds ($0 \le \tau < \infty$), we must have $1 - F(u_n) \to 0$. (For, otherwise, $1 - F(u_{n_k})$ would be bounded away from zero for some subsequence $\{n_k\}$, leading, from (1.5.3), to the conclusion $P\{M_n \le u_n\} \to 0$.) By taking logarithms in (1.5.2) and (1.5.3), we have

$$n \log\{1 - (1 - F(u_n))\} \to -\tau,$$

so that $n(1 - F(u_n))(1 + o(1)) \to \tau$, giving the result.

Finally, if $\tau = \infty$ and (1.5.1) holds but (1.5.2) does not (i.e. $P\{M_n \le u_n\} \not\to 0$), there must be a subsequence $\{n_k\}$ such that $P\{M_{n_k} \le u_{n_k}\} \to e^{-\tau'}$ as $k \to \infty$ for some $\tau' < \infty$. But, as above, (1.5.2) implies (1.5.1), with n_k replacing n so that $n_k(1 - F(u_{n_k})) \to \tau' < \infty$, contradicting the assumption that (1.5.1) holds with $\tau = \infty$. Similarly, (1.5.2) implies (1.5.1) when $\tau = \infty$. □

The following simple results follow as corollaries. We use the notation $x_F = \sup\{x; F(x) < 1\}$ introduced previously for the right endpoint of a d.f. F.

Corollary 1.5.2. (i) *$M_n \to x_F$ ($\le \infty$) with probability one as $n \to \infty$.*

(ii) *If $x_F < \infty$ and $F(x_F-) < 1$ (i.e. if F has a jump at its right endpoint), and if for a sequence $\{u_n\}$, $P\{M_n \le u_n\} \to \rho$ as $n \to \infty$, then $\rho = 0$ or 1.*

PROOF. If $\lambda < x_F$ ($\le \infty$), $1 - F(\lambda) > 0$ so that (1.5.1) holds with $u_n \equiv \lambda$, $\tau = \infty$, and thus (1.5.2) gives $\lim P\{M_n \le \lambda\} = 0$. Since clearly $P\{M_n > x_F\} = 0$ for all n, it follows that $M_n \to x_F$ in probability. Since $\{M_n\}$ is monotone, it converges a.s., and hence $M_n \to x_F$ a.s. Hence (i) follows.

Suppose now that $x_F < \infty$ and $F(x_F-) < 1$. Let $\{u_n\}$ be a sequence such that $P\{M_n \le u_n\} \to \rho$. Since $0 \le \rho \le 1$, we may write $\rho = e^{-\tau}$ with

$0 \le \tau \le \infty$, and by Theorem 1.5.1 obtain $n(1 - F(u_n)) \to \tau$. If $u_n < x_F$ for infinitely many values of n, and since for those values $1 - F(u_n) \ge 1 - F(x_F -) > 0$, we must have $\tau = \infty$. The only other possibility is that $u_n \ge x_F$ for all sufficiently large values of n, giving $n(1 - F(u_n)) = 0$ so that $\tau = 0$. Thus $\tau = \infty$ or 0, and hence $\rho = 0$ or 1, proving (ii). □

We take up the general question of domains of attraction of the extreme value distributions (using Theorem 1.5.1 in an essential way) in the next section. For the purposes of this volume, however, normal sequences occupy a most important position. We therefore show here how Theorem 1.5.1 may be used directly to obtain the (Type I) limit law for i.i.d. normal sequences. As will be seen, the application is straightforward, even though some calculations are involved. Throughout this and all subsequent chapters, Φ, ϕ will denote the standard normal distribution function and density, respectively. We shall repeatedly have occasion to use the well-known relation for the tail of Φ, stated here for easy reference:

$$1 - \Phi(u) \sim \frac{\phi(u)}{u} \quad \text{as } u \to \infty. \tag{1.5.4}$$

Theorem 1.5.3. *If $\{\xi_n\}$ is an i.i.d. (standard) normal sequence of r.v.'s, then the asymptotic distribution of $M_n = \max(\xi_1, \dots, \xi_n)$ is of Type I. Specifically,*

$$P\{a_n(M_n - b_n) \le x\} \to \exp(-e^{-x}), \tag{1.5.5}$$

where

$$a_n = (2 \log n)^{1/2}$$

and

$$b_n = (2 \log n)^{1/2} - \tfrac{1}{2}(2 \log n)^{-1/2}(\log \log n + \log 4\pi).$$

Proof. Write $\tau = e^{-x}$ in (1.5.1). Then we may take $1 - \Phi(u_n) = (1/n)e^{-x}$. Since $1 - \Phi(u_n) \sim \phi(u_n)/u_n$ we have $(1/n)e^{-x}u_n/\phi(u_n) \to 1$, and hence $-\log n - x + \log u_n - \log \phi(u_n) \to 0$ or

$$-\log n - x + \log u_n + \tfrac{1}{2} \log 2\pi + \frac{u_n^2}{2} \to 0. \tag{1.5.6}$$

It follows at once that $u_n^2/(2 \log n) \to 1$, and hence

$$2 \log u_n - \log 2 - \log \log n \to 0$$

or

$$\log u_n = \tfrac{1}{2}(\log 2 + \log \log n) + o(1).$$

Putting this in (1.5.6), we obtain

$$\frac{u_n^2}{2} = x + \log n - \tfrac{1}{2}\log 4\pi - \tfrac{1}{2}\log\log n + o(1)$$

or

$$u_n^2 = 2\log n\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{\log n} + o\left(\frac{1}{\log n}\right)\right\},$$

and hence

$$u_n = (2\log n)^{1/2}\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{2\log n} + o\left(\frac{1}{\log n}\right)\right\}$$

so that

$$u_n = \frac{x}{a_n} + b_n + o((\log n)^{-1/2}) = \frac{x}{a_n} + b_n + o(a_n^{-1}).$$

Hence, since by (1.5.2) we have $P\{M_n \le u_n\} \to \exp(-e^{-x})$ where $\tau = e^{-x}$,

$$P\left\{M_n \le \frac{x}{a_n} + b_n + o(a_n^{-1})\right\} \to \exp(-e^{-x})$$

or

$$P\{a_n(M_n - b_n) + o(1) \le x\} \to \exp(-e^{-x}),$$

from which (1.5.5) follows, as required. □

While there are some computational details in this derivation, there is no difficulty of any kind, and (1.5.5) thus follows in a very simple way from the (itself simple) result (1.5.2). The same arguments can, and will later, be adapted to a continuous time context.

1.6. General Theory of Domains of Attraction

It is, of course, important to know which (if any) of the three types of limit law applies when each r.v. ξ_n has a given d.f. F. Various necessary and sufficient conditions are known—involving the "tail behaviour" $1 - F(x)$ as x increases—for each type of limit. We shall state these and prove their sufficiency, omitting the proofs of necessity since, as already mentioned, these are somewhat lengthy and less germane to our purposes here. Proofs may be found, e.g., in Gnedenko (1943) or de Haan (1976).

Before presenting the general theorems, we give some very simple and useful sufficient conditions which apply when the d.f. F has a density function f. These are due to von Mises, and simple proofs are given in de Haan (1976). Here we reproduce one of the proofs as a sample and refer the reader to de Haan (1976) for the details of the others.

Theorem 1.6.1. *Suppose that the d.f. F of the r.v.'s of the i.i.d. sequence $\{\xi_n\}$ is absolutely continuous with density f. Then sufficient conditions for F to belong to each of the three possible domains of attraction are:*

Type I: *f has a negative derivative f' for all x in some interval (x_0, x_F), $(x_F \leq \infty)$, $f(x) = 0$ for $x \geq x_F$, and*

$$\lim_{t \uparrow x_F} \frac{f'(t)(1 - F(t))}{f^2(t)} = -1;$$

Type II: *$f(x) > 0$ for all $x \geq x_0$ finite, and*

$$\lim_{t \to \infty} \frac{tf(t)}{1 - F(t)} = \alpha > 0;$$

Type III: *$f(x) > 0$ for all x in some finite interval (x_0, x_F), $f(x) = 0$ for $x > x_F$, and*

$$\lim_{t \uparrow x_F} \frac{(x_F - t)f(t)}{1 - F(t)} = \alpha > 0.$$

Proof for Type II Case. As noted, complete proofs may be found in de Haan (1976), and we give the Type II case here as a sample. Assume, then, that $\lim_{t \to \infty} tf(t)/(1 - F(t)) = \alpha > 0$, where $f(x) > 0$ for $x \geq x_0$. Writing $\alpha(t) = tf(t)/(1 - F(t))$, it is immediate that, for $x_2 \geq x_1 \geq x_0$,

$$\int_{x_1}^{x_2} \frac{\alpha(t)}{t} dt = -\log(1 - F(x_2)) + \log(1 - F(x_1))$$

so that

$$1 - F(x_2) = (1 - F(x_1)) \exp\left(-\int_{x_1}^{x_2} \frac{\alpha(t)}{t} dt\right).$$

Clearly there exists γ_n such that $1 - F(\gamma_n) = n^{-1}$, and (by writing $x_1 = \gamma_n$, $x_2 = \gamma_n x$ or vice versa according as $x \geq 1$ or $x < 1$) we obtain

$$n(1 - F(\gamma_n x)) = \exp\left(-\int_{\gamma_n}^{\gamma_n x} \frac{\alpha(t)}{t} dt\right) = \exp\left(-\int_1^x \frac{\alpha(\gamma_n s)}{s} ds\right),$$

which since $\gamma_n \to \infty$, converges to $e^{-\alpha \log x} = x^{-\alpha}$ as $n \to \infty$. Hence for $x > 0$, by Theorem 1.5.1,

$$P\{M_n \leq \gamma_n x\} \to \exp(-x^{-\alpha}) \quad \text{as } n \to \infty.$$

For $x \leq 0$ since $P\{M_n \leq \gamma_n x\} \leq P\{M_n \leq \gamma_n y\}$ for $y > 0$ and $P\{M_n \leq \gamma_n y\} \to \exp(-y^{-\alpha})$, we see by letting $y \to 0$ that $\lim_{n \to \infty} P\{M_n \leq \gamma_n x\} = 0$. Hence the Type II limit holds with the normalizing constants $a_n = \gamma_n^{-1}$, $b_n = 0$. □

The above proof for the Type II case relied on the existence of a sequence $\{\gamma_n\}$ such that $1 - F(\gamma_n) \sim 1/n$, and the normalizing constants a_n, b_n in the relation $P\{a_n(M_n - b_n) \le x\} \to G(x)$ were then obtained in terms of γ_n. (In this case, we had $a_n = \gamma_n^{-1}$, $b_n = 0$.) For an arbitrary d.f. F, such a γ_n does not necessarily exist. However, if $F \in D(G)$ for one of the extreme value distributions, then a γ_n with this property may be found. For if

$$P\{a_n(M_n - b_n) \le x\} \to G(x),$$

since G is continuous x may be chosen so that $G(x) = e^{-1}$ so that

$$P\{M_n \le \gamma_n\} \to e^{-1}$$

with $\gamma_n = x/a_n + b_n$. Hence $1 - F(\gamma_n) \sim 1/n$ by Theorem 1.5.1.

The general criteria whose sufficiency is to be proved will also rely on the existence of such a γ_n, which follows in each case, as we shall see, from the assumptions made. Again, in the following results we write $x_F = \sup\{x; F(x) < 1\}$ for any d.f. F.

Theorem 1.6.2. *Necessary and sufficient conditions for the d.f. F of the r.v.'s of the i.i.d. sequence $\{\xi_n\}$ to belong to each of the three types are (in order of increasing complexity):*

Type II: $x_F = \infty$ *and* $\lim_{t\to\infty} (1 - F(tx))/(1 - F(t)) = x^{-\alpha}, \alpha > 0,$ *for each* $x > 0$;

Type III: $x_F < \infty$ *and* $\lim_{h\downarrow 0} (1 - F(x_F - xh))/(1 - F(x_F - h)) = x^{\alpha}$, $\alpha > 0$, *for each* $x > 0$;

Type I: *There exists some strictly positive function $g(t)$ such that*

$$\lim_{t\uparrow x_F} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}$$

for all real x.

It may in fact be shown that $\int_0^\infty (1 - F(u))\,du < \infty$ when a Type I limit holds, and one appropriate choice of g is given by $g(t) = \int_t^{x_F} (1 - F(u))\,du/(1 - F(t))$ for $t < x_F$.

Proofs of Sufficiency. To highlight the simplicity of the proof, we assume first the existence of a sequence $\{\gamma_n\}$ (which may be taken nondecreasing in n) in each case such that $n(1 - F(\gamma_n)) \to 1$. (This will be proved—also very simply—below.) The γ_n constants will, of course, differ for the differing types. Clearly $\gamma_n \to x_F$ and $\gamma_n < x_F$ for all sufficiently large n.

If F satisfies the Type II criterion we have, writing γ_n for t, for each $x > 0$,

$$n(1 - F(\gamma_n x)) \sim n(1 - F(\gamma_n))x^{-\alpha} \to x^{-\alpha}$$

so that Theorem 1.5.1 yields, for $x > 0$,

$$P\{M_n \le \gamma_n x\} \to \exp\{-x^{-\alpha}\}.$$

Since $\gamma_n > 0$ (when n is large, at least) and the right-hand side tends to zero as $x \downarrow 0$, it also follows that $P\{M_n \le 0\} \to 0$, and for $x < 0$, that $P\{M_n \le \gamma_n x\} \le P\{M_n \le 0\} \to 0$. Thus $P\{M_n \le \gamma_n x\} \to G(x)$, where G is the Type II representative d.f. listed in Theorem 1.4.1. But this may be restated as

$$P\{a_n(M_n - b_n) \le x\} \to G(x), \tag{1.6.1}$$

where $a_n = \gamma_n^{-1}$ and $b_n = 0$ so that the Type II limit follows.

The Type III limit follows in a closely similar way by writing $h_n = x_F - \gamma_n$ ($\downarrow 0$) so that, for $x > 0$,

$$\lim_{n\to\infty} n(1 - F\{x_F - x(x_F - \gamma_n)\}) = x^\alpha,$$

and hence (replacing x by $-x$) for $x < 0$,

$$\lim_{n\to\infty} n(1 - F\{x_F + x(x_F - \gamma_n)\}) = (-x)^\alpha.$$

Using Theorem 1.5.1 again, this shows at once that the Type III limit applies with constants in (1.6.1) given by

$$a_n = (x_F - \gamma_n)^{-1}, \qquad b_n = x_F.$$

The Type I limit also follows along the same lines since, when F satisfies that criterion, we have, for all x, writing $t = \gamma_n \uparrow x_F$ ($\le \infty$),

$$\lim_{n\to\infty} n(1 - F\{\gamma_n + xg(\gamma_n)\}) = e^{-x},$$

giving (again by Theorem 1.5.1) the Type I limit with $a_n = (g(\gamma_n))^{-1}$, $b_n = \gamma_n$.

Finally, we must show the existence of the (nondecreasing) sequence $\{\gamma_n\}$ satisfying $\lim_{n\to\infty} n(1 - F(\gamma_n)) = 1$. For γ_n, we may take any nondecreasing sequence such that

$$F(\gamma_n -) \le 1 - \frac{1}{n} \le F(\gamma_n)$$

(such as the sequence $\gamma_n = F^{-1}(1 - 1/n) = \inf\{x; F(x) \ge 1 - 1/n\}$). For such a sequence, $n(1 - F(\gamma_n)) \le 1$ so that, trivially, $\limsup n(1 - F(\gamma_n)) \le 1$. Thus it only remains to show that in each case $\liminf n(1 - F(\gamma_n)) \ge 1$, which will follow since $n(1 - F(\gamma_n -)) \ge 1$ if we show that

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n -)} \ge 1. \tag{1.6.2}$$

For a d.f. F satisfying the listed Type II criterion, the left-hand side of (1.6.2) is, for any $x < 1$, no smaller than

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n x)} = x^\alpha,$$

from which (1.6.2) follows by letting $x \to 1$.

A similar argument holds for a d.f. F satisfying the Type III criterion, the left-hand side of (1.6.2) being no smaller (for $x > 1$, $h_n = x_F - \gamma_n$) than

$$\liminf_{n\to\infty} \frac{1 - F(x_F - h_n)}{1 - F(x_F - xh_n)} = x^{-\alpha},$$

which tends to 1 as $x \to 1$, giving (1.6.2).

Finally for the Type I case, the left-hand side of (1.6.2) is no smaller (if $x < 0$) than

$$\liminf_{n\to\infty} \frac{1 - F(\gamma_n)}{1 - F(\gamma_n + xg(\gamma_n))} = e^x,$$

which tends to 1 as $x \to 0$, so that again (1.6.2) holds. □

Corollary 1.6.3. *The constants a_n, b_n in the convergence $P\{a_n(M_n - b_n) \le x\} \to G(x)$ may be taken in each case above to be:*

$$\begin{aligned} \textit{Type II}: &\quad a_n = \gamma_n^{-1},\ b_n = 0;\\ \textit{Type III}: &\quad a_n = (x_F - \gamma_n)^{-1},\ b_n = x_F;\\ \textit{Type I}: &\quad a_n = [g(\gamma_n)]^{-1},\ b_n = \gamma_n, \end{aligned}$$

with $\gamma_n = F^{-1}(1 - 1/n) = \inf\{x; F(x) \ge 1 - 1/n\}$.

PROOF. These relationships appear in the course of the proof of the theorem above. □

It is perhaps worth noting that the criteria given above apply to any d.f. in each domain of attraction, regardless of whether the limit appears as the specific distribution $G(x)$ listed in Theorem 1.4.1 representing a type, or any other d.f. $G(ax + b)$ of that type. Indeed, if the limit appears as $G(ax + b)$, $G(x)$ is also a limit with a simple change of normalizing constants. For if

$$P\{a_n(M_n - b_n) \le x\} \to G(ax + b),$$

then clearly

$$P\{\alpha_n(M_n - \beta_n) \le x\} \to G(x),$$

with $\alpha_n = aa_n$, $\beta_n = b_n - b/(aa_n)$.

1.7. Examples

We shall give examples of distributions belonging to each of the three domains of attraction and then cases where no nondegenerate limiting distribution exists. For reference, the constants a_n, b_n will appear in the usual form (1.1.3), viz.

$$P\{a_n(M_n - b_n) \le x\} \to G(x). \tag{1.7.1}$$

The examples in the first group (1.7.1–1.7.5) all involve the Type I domain.

Example 1.7.1 (Normal distribution). As shown in Theorem 1.5.3, the (standard) normal distribution belongs to the Type I domain of attraction with constants

$$\left.\begin{aligned} a_n &= (2\log n)^{1/2} \\ b_n &= (2\log n)^{1/2} - \tfrac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi) \end{aligned}\right\} \tag{1.7.2}$$

Note that it is very easy to determine that a Type I domain of attraction applies by means of Theorem 1.6.1, though this does not provide the constants. Corollary 1.6.3 does give the constants, but the work involved in transforming these to the above values differs very little from the calculations of Theorem 1.5.3. □

Example 1.7.2 (Exponential distribution). In the case of an exponential distribution with unit parameter, we have

$$F(x) = 1 - e^{-x}.$$

Theorem 1.6.1 is easily applied to demonstrate a Type I domain. However, it is very simple to show this (and obtain the constants) directly. For if $\tau > 0$, we may choose u_n so that $1 - F(u_n) = \tau/n$ simply by taking

$$u_n = -\log\frac{\tau}{n} = -\log\tau + \log n,$$

so that by Theorem 1.5.1,

$$P\{M_n \le -\log\tau + \log n\} \to e^{-\tau}.$$

By writing $\tau = e^{-x}$, we obtain (1.7.1) with

$$a_n = 1, \qquad b_n = \log n, \qquad G(x) = \exp\{-e^{-x}\}. \qquad \square$$

Example 1.7.3 (Type I extreme value distribution itself). As previously noted, each extreme value distribution belongs to its own domain of attraction (Theorem 1.3.1). The constants are easily obtained using the max-stable property. If $F(x) = \exp(-e^{-x})$, then

$$F^n(x) = \exp(-e^{-x+\log n}),$$

so that

$$F^n(x + \log n) = F(x).$$

By taking

$$a_n = 1 \quad \text{and} \quad b_n = \log n,$$

we have, for all n,

$$P\{a_n(M_n - b_n) \le x\} = F^n(x + \log n) = F(x),$$

so that the Type I limit holds with these constants. □

Example 1.7.4 (A monotone transformation (lognormal) of the normal distribution). If f is a monotone increasing function and $\xi_i' = f(\xi_i)$, then clearly

$$M_n' = \max(\xi_1', \ldots, \xi_n') = f(M_n).$$

If $\{\xi_i\}$ is any i.i.d. sequence such that (1.7.1) holds, we have

$$P\left\{M_n \leq \frac{x}{a_n} + b_n\right\} \to G(x)$$

so that

$$P\left\{M_n' \leq f\left(\frac{x}{a_n} + b_n\right)\right\} \to G(x).$$

In some cases, f may be expanded and linear terms only retained to give the same limiting d.f. G for M_n' (with changed constants, e.g. $a_n' = a_n/f'(b_n)$, $b_n' = f(b_n)$). For example, if the ξ_i are normal, then a_n, b_n are given by (1.7.2) above. Taking $f(x) = e^x$, we obtain lognormal ξ_i' and

$$P\left\{M_n' \leq \exp\left(\frac{x}{a_n} + b_n\right)\right\} \to G(x) = \exp(-e^{-x}),$$

giving

$$P\left\{e^{-b_n}M_n' \leq 1 + \frac{x}{a_n} + o\left(\frac{1}{a_n}\right)\right\} \to G(x),$$

from which it follows (since $a_n \to \infty$) that

$$P\{a_n e^{-b_n}(M_n' - e^{b_n}) \leq x\} \to G(x),$$

so that M_n' has the Type I limit with the constants

$$a_n' = a_n e^{-b_n}, \qquad b_n' = e^{b_n}. \qquad \square$$

The above Type I examples have all involved cases where the distributions have infinite upper endpoints, i.e. $x_F = \infty$. It is easy to construct cases where the Type I distribution occurs with $x_F < \infty$, as the following example shows.

Example 1.7.5 ($F(x) = 1 - e^{1/x}$ for $x < 0$, $F(x) = 1$ for $x \geq 0$). By Theorem 1.5.1, if $\{u_n\}$ is such that $ne^{1/u_n} \to \tau > 0$, it follows that

$$P\{M_n \leq u_n\} \to e^{-\tau}.$$

By writing $\tau = e^{-x}$ $(-\infty < x < \infty)$ and taking $u_n = (\log \tau - \log n)^{-1}$, it follows that

$$P\{M_n \leq -(\log n + x)^{-1}\} \to \exp(-e^{-x}),$$

from which it is readily checked that

$$P\left\{(\log n)^2\left[M_n + \frac{1}{\log n}\right] \le x + o(1)\right\} \to \exp(-e^{-x}),$$

giving a Type I limit with

$$a_n = (\log n)^2, \qquad b_n = -(\log n)^{-1}. \qquad \square$$

Turning now to Type II cases in the next examples, we see first that it is very simple to "manufacture" such cases by using the criterion of Theorem 1.6.2.

Example 1.7.6 (Pareto distribution). Let $F(x) = 1 - Kx^{-\alpha}$, $\alpha > 0$, $K > 0$, $x \ge K^{1/\alpha}$. Then $(1 - F(tx))/(1 - F(t)) = x^{-\alpha}$ for each $x > 0$, when t is sufficiently large so that Theorem 1.6.2 shows that a Type II limit applies. Indeed, for $u_n = (Kn/\tau)^{1/\alpha}$ we have $1 - F(u_n) = \tau/n$, so that Theorem 1.5.1 gives

$$P\left\{M_n \le \left(\frac{Kn}{\tau}\right)^{1/\alpha}\right\} \to e^{-\tau}.$$

Putting $\tau = x^{-\alpha}$ for $x \ge 0$, we have, on rearranging,

$$P\{(Kn)^{-1/\alpha}M_n \le x\} \to \exp(-x^{-\alpha}),$$

so that a Type II limit holds with

$$a_n = (Kn)^{-1/\alpha}, \qquad b_n = 0. \qquad \square$$

Example 1.7.7 (Type II extreme value distribution). As with the Type I case, we know from Theorem 1.3.1 that a Type II d.f.

$$F(x) = \begin{cases} \exp(-x^{-\alpha}), & x \ge 0, \qquad \alpha > 0, \\ 0, & x < 0, \end{cases}$$

belongs to its own domain of attraction. Since clearly $F^n(n^{1/\alpha}x) = F(x)$, it follows that

$$P\{n^{-1/\alpha}M_n \le x\} = F(x)$$

for all n, and hence the Type II limit holds with

$$a_n = n^{-1/\alpha}, \qquad b_n = 0. \qquad \square$$

Example 1.7.8 (Cauchy distribution). For the standard Cauchy distribution,

$$F(x) = \tfrac{1}{2} + \frac{1}{\pi}\tan^{-1} x$$

so that

$$\frac{1 - F(tx)}{1 - F(t)} = \frac{\pi/2 - \tan^{-1} tx}{\pi/2 - \tan^{-1} t}.$$

But it is readily checked (e.g. by writing $\tan^{-1} t = \pi/2 - \theta$) that $\lim_{t\to\infty} t(\pi/2 - \tan^{-1} t) = \lim_{\theta\to 0} \theta \cot \theta = 1$ with the same limit when t is replaced by tx, so that $(1 - F(tx))/(1 - F(t)) \to x^{-1}$ as $t \to \infty$, demonstrating that a Type II limit applies by Theorem 1.6.2. We may simply obtain the constants from Corollary 1.6.3 as $a_n = \gamma_n^{-1}$, $b_n = 0$, where $1 - F(\gamma_n) = 1/n$, from which $\gamma_n = \tan(\pi/2 - \pi/n) = \cot \pi/n$. Hence appropriate constants are

$$a_n = \gamma_n^{-1} = \tan \frac{\pi}{n} \sim \frac{\pi}{n}, \qquad b_n = 0. \qquad \square$$

For Type III cases, the distribution must, by Theorem 1.6.2, have a finite upper endpoint x_F. The simplest such example is the uniform distribution, as we now see.

Example 1.7.9 (Uniform distribution on (0, 1)). Here $F(x) = x$, $0 \le x \le 1$. For $\tau > 0$ and $u_n = 1 - \tau/n$, we have $1 - F(u_n) = \tau/n$ for $n \ge \tau$, so that by Theorem 1.5.1,

$$P\left\{M_n \le 1 - \frac{\tau}{n}\right\} \to e^{-\tau}.$$

Hence for $x < 0$ and $\tau = -x$ we have

$$P\{n(M_n - 1) \le x\} \to e^{x},$$

which shows a Type III limit with $\alpha = 1$,

$$a_n = n, \qquad b_n = 1. \qquad \square$$

The uniform distribution is a particular case of the following obvious class of distributions with Type III limit.

Example 1.7.10. (Polynomial growth at a finite endpoint). Let $x_F < \infty$, $K > 0$, and

$$F(x) = 1 - K(x_F - x)^{\alpha}, \qquad x_F - K^{-1/\alpha} \le x \le x_F.$$

It follows at once from Theorem 1.6.2 that a Type III limit $\exp(-(-x)^{\alpha})$ holds and, directly as in Example 1.7.9 or by Corollary 1.6.3, that we may take

$$a_n = (nK)^{1/\alpha}, \qquad b_n = x_F. \qquad \square$$

Since the asymptotic behaviour of the tail $1 - F(x)$ determines which (if any) domain of attraction a d.f. F belongs to, one may get a different limiting type, or none at all, by truncating F on the right—regardless of its form left of the truncation point. This is illustrated by the following simple example.

Example 1.7.11 (Truncated exponential distribution). We saw in Example 1.7.2 that an exponential distribution on $(0, \infty)$ belongs to the Type I domain of attraction. If we truncate at a finite value x_F, writing $F(x) = K(1 - e^{-x})$ for $0 \le x \le x_F$ (with $K = (1 - e^{-x_F})^{-1}$), it is readily seen from Theorem 1.5.1, writing $\tau = -x$, for $x < 0$, that

$$P\left\{M_n \le x_F + \frac{x(e^{x_F} - 1)}{n} + o\left(\frac{1}{n}\right)\right\} \to e^x,$$

giving a Type III limit with

$$a_n = \frac{n}{(e^{x_F} - 1)}, \qquad b_n = x_F.$$

Example 1.7.12. (Type III extreme value d.f.). Here

$$F(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \le 0, \qquad \alpha > 0, \\ 1, & x > 0. \end{cases}$$

As in the Type I and II cases, F belongs to its own domain of attraction. Also, for each n, $F^n(n^{-1/\alpha}x) = F(x)$ so that

$$P\{n^{1/\alpha}M_n \le x\} = F(x)$$

for all n, demonstrating the Type III limit with

$$a_n = n^{1/\alpha}, \qquad b_n = 0. \qquad \square$$

These examples illustrate the complete range of possibilities for limiting distributions. The constants b_n may—as expected—take positive, negative, or zero values according to the particular cases involved. One might feel that the (strictly positive) scaling constants a_n should either tend (universally) to infinity or (universally) to zero. However, it is interesting to note that both such limits may occur, as well as other cases, such as that in Example 1.7.2 where a_n is constant.

We turn now to consider cases in which nondegenerate limiting distributions for the maximum do not exist under any linear normalization. As noted in Section 1.4, this is certainly the case if the common d.f. F of the members of the i.i.d. sequence $\{\xi_n\}$ has a finite right endpoint x_F and a jump at that point. Indeed, Corollary 1.5.2 shows that then if for any sequence $\{u_n\}$, $P\{M_n \le u_n\} \to \rho$, then ρ is either zero or one.

A perhaps more common situation in which this same result applies occurs for certain discrete distributions (such as the Poisson and geometric), as we shall now see from the following theorem.

Theorem 1.7.13. *Let $\{\xi_n\}$ be an i.i.d. sequence of r.v.'s with common d.f. F. Then, if $0 < \tau < \infty$, there exists a sequence $\{u_n\}$ satisfying* (1.5.1) (*i.e.* $1 - F(u_n) \sim \tau/n$) *if and only if*

$$\frac{1 - F(x)}{1 - F(x-)} \to 1 \quad \text{as } x \to x_F, \tag{1.7.3}$$

or equivalently, if and only if

$$\frac{p(x)}{1 - F(x-)} \to 0, \tag{1.7.4}$$

where $p(x) = F(x) - F(x-)$.

Hence, by Theorem 1.5.1, *if* $0 < \rho < 1$, *there is a sequence* $\{u_n\}$ *such that* $P\{M_n \le u_n\} \to \rho$ *if and only if* (1.7.3) (*or* (1.7.4)) *holds. For* $\rho = 0$ *or* 1, *such a sequence may always be found.*

Proof. We suppose that (1.5.1) holds for some τ, $0 < \tau < \infty$ but that, say (1.7.4), does not. Then there exists $\varepsilon > 0$ and a sequence $\{x_n\}$ such that $x_n \to x_F$ and

$$p(x_n) \ge 2\varepsilon(1 - F(x_n-)). \tag{1.7.5}$$

Now choose a sequence of integers $\{n_j\}$ so that $1 - \tau/n_j$ is "close" to the midpoint of the jump of F at x_j, i.e. such that

$$1 - \frac{\tau}{n_j} \le \frac{F(x_j-) + F(x_j)}{2} \le 1 - \frac{\tau}{n_j + 1}.$$

Clearly we have either

(i) $u_{n_j} < x_j$ for infinitely many values of j, or
(ii) $u_{n_j} \ge x_j$ for infinitely many j-values.

If alternative (i) holds, then for such j,

$$n_j(1 - F(u_{n_j})) \ge n_j(1 - F(x_j-)). \tag{1.7.6}$$

Now, clearly

$$\begin{aligned} n_j(1 - F(x_j-)) &= \tau + n_j\left[\left(1 - \frac{\tau}{n_j}\right) - \frac{F(x_j) + F(x_j-)}{2} + \frac{p(x_j)}{2}\right] \\ &\ge \tau + \frac{n_j p(x_j)}{2} - n_j\left(\frac{\tau}{n_j} - \frac{\tau}{n_j + 1}\right) \\ &\ge \tau + \varepsilon n_j(1 - F(x_j-)) - \frac{\tau}{n_j + 1} \end{aligned}$$

by (1.7.5) so that

$$(1 - \varepsilon)n_j(1 - F(x_j-)) \ge \tau - \frac{\tau}{n_j + 1}.$$

Since clearly $n_j \to \infty$, it follows that (since $0 < \tau < \infty$ by assumption)

$$\limsup_{j\to\infty} n_j(1 - F(x_j-)) > \tau,$$

and hence by (1.7.6),

$$\limsup_{j\to\infty} n_j(1 - F(u_{n_j})) > \tau,$$

which contradicts (1.5.1). The calculations in case (ii) ($u_{n_j} \geq x_j$ for infinitely many j) are very similar, with only the obvious changes.

Conversely, suppose that (1.7.3) holds and let $\{u_n\}$ be any sequence such that $F(u_n-) \leq 1 - \tau/n \leq F(u_n)$ (e.g. $u_n = F^{-1}(1 - \tau/n)$), from which a simple rearrangement yields

$$\frac{1 - F(u_n)}{1 - F(u_n-)}\tau \leq n(1 - F(u_n)) \leq \tau$$

from which (1.5.1) follows since clearly $u_n \to x_F$ as $n \to \infty$. □

It is of interest to note from the theorem that the existence of a sequence $\{u_n\}$ such that (1.5.1) holds for some τ, $0 < \tau < \infty$ (or such that $P\{M_n \leq u_n\} \to \rho$ for some ρ, $0 < \rho < 1$) implies the existence of such a sequence for every such τ (or ρ).

If the r.v.'s $\{\xi_n\}$ are integer valued, and $x_F = \infty$, (1.7.3) becomes $(1 - F(n))/(1 - F(n - 1)) \to 1$ as $n \to \infty$, whereas (1.7.4) is just $p_n/(1 - F(n - 1)) \to 0$ where $p_n = P\{\xi_1 = n\}$. If either of these two (equivalent) conditions is violated, it follows at once (writing $u_n = x/a_n + b_n$) that there can be no nondegenerate limit for M_n. In particular, the Poisson and geometric distributions are examples of such cases, as we now see.

Example 1.7.14 (Poisson distribution). In this case, $p_r = e^{-\lambda}\lambda^r/r!$ for $\lambda > 0$, $r = 0, 1, 2, \ldots$, so that

$$\frac{p_n}{1 - F(n - 1)} = \frac{\lambda^n/n!}{\sum_{r=n}^{\infty} \frac{\lambda^r}{r!}} = \frac{1}{1 + \sum_{r=n+1}^{\infty} \frac{n!}{r!}\lambda^{r-n}}.$$

The sum in the denominator may be rewritten as

$$\sum_{s=1}^{\infty} \frac{\lambda^s}{(n + 1)(n + 2)\cdots(n + s)} \leq \sum_{s=1}^{\infty} \left(\frac{\lambda}{n}\right)^s = \frac{\lambda/n}{1 - \lambda/n}$$

(when $n > \lambda$) and thus tends to zero as $n \to \infty$ so that $p_n/(1 - F(n - 1)) \to 1$. Hence Theorem 1.7.13 shows that no limiting distribution exists and, indeed, that no limit of the form $P\{M_n \leq u_n\} \to \rho$ exists other than for $\rho = 0$ or 1, whatever the sequence of constants $\{u_n\}$. □

Example 1.7.15 (Geometric distribution). Here $p_r = (1 - p)^{r-1}p$, $0 < p < 1$, for $r = 1, 2, \ldots$ so that

$$\frac{p_n}{1 - F(n - 1)} = \frac{(1 - p)^{n-1}}{\sum_{r=n}^{\infty}(1 - p)^{r-1}} = p,$$

which again shows that no limit $P\{M_n \le u_n\} \to \rho$ exists except for $\rho = 0$ or 1 and that there is no nondegenerate limiting distribution for the maximum in the geometric case. □

Finally, we re-emphasize that the existence of a limit other than zero or one for $P\{M_n \le u_n\}$ (and, in particular, of nondegenerate limiting distributions) depends precisely on whether or not $1 - F(u_n) \sim \tau/n$ for some τ, $0 < \tau < \infty$. If F is continuous, a sequence u_n may be chosen (by taking $u_n = F^{-1}(1 - \tau/n)$) for any given τ, and then such a limit exists (though not necessarily leading to a limiting nondegenerate d.f. for M_n under linear normalization).

If F is not continuous, the existence of any $\{u_n\}$-sequence satisfying $1 - F(u_n) \sim \tau/n$ is not guaranteed and fails if the jumps continue to be so "large" that there is no number u_n such that $F(u_n)$ is a "good approximation to $1 - \tau/n$". Specifically it is necessary and sufficient for the existence of such a $\{u_n\}-$ sequence that the jumps $p(x) = F(x) - F(x-)$ satisfy (1.7.4), viz. $p(x)/(1 - F(x-)) \to 0$ as $x \to x_F$, i.e. $p_n/(1 - F(n-1)) \to 0$ in the case of integer-valued r.v.'s. For positive integer-valued r.v.'s, this has the interesting interpretation as the convergence to zero of the "conditional failure" or "hazard" rate. For $p_n/(1 - F(n-1))$ is clearly the conditional probability $P\{\xi_i = n \mid \xi_i \ge n\}$ that the "lifetime" ξ_i of an item surviving to time n, will then terminate at n. Note that the hazard rate is actually the constant p for the geometric distribution considered, and converges to 1 in the Poisson case.

1.8. Minima

As noted, the minimum

$$m_n = \min\{\xi_1, \ldots, \xi_n\}$$

is simply given as

$$m_n = -\max\{-\xi_1, \ldots, -\xi_n\}$$

so that limiting results for minima can clearly be obtained from those for maxima. This will be useful below in obtaining the possible limiting distributions for minima. However, some rather obvious facts follow at least as simply directly, as is the case for the following analogue of Theorem 1.5.1.

Theorem 1.8.1. *Let $\{\xi_n\}$ be an i.i.d. sequence. Let $0 \le \eta \le \infty$ and suppose that $\{v_n\}$ is a sequence of real numbers such that*

$$nF(v_n) \to \eta \quad \text{as } n \to \infty. \tag{1.8.1}$$

Then

$$P\{m_n > v_n\} \to e^{-\eta} \quad \text{as } n \to \infty. \tag{1.8.2}$$

Conversely, if (1.8.2) *holds for some η, $0 \le \eta \le \infty$, then so does* (1.8.1).

Proof. This result is proved in an exactly analogous fashion to Theorem 1.5.1, simply noting that $P\{m_n > v_n\} = (1 - F(v_n))^n$. □

It is also easy to see that the events $\{M_n \le u_n\}$ and $\{m_n > v_n\}$ are asymptotically independent (and hence so are the events $\{M_n \le u_n\}$, $\{m_n \le v_n\}$) if the sequences $\{u_n\}$, $\{v_n\}$ satisfy (1.5.1) and (1.8.1), respectively.

Theorem 1.8.2. *Suppose that the sequences* $\{u_n\}$, $\{v_n\}$ *satisfy* (1.5.1) *and* (1.8.1), *respectively. Then*

$$P\{M_n \le u_n, m_n > v_n\} \to e^{-(\tau+\eta)}, \tag{1.8.3}$$

so that by Theorem 1.5.1,

$$\begin{aligned} P\{M_n \le u_n, m_n \le v_n\} &= P\{M_n \le u_n\} - P\{M_n \le u_n, m_n > v_n\} \\ &\to e^{-\tau}(1 - e^{-\eta}). \end{aligned} \tag{1.8.4}$$

Also by Theorems 1.5.1 *and* 1.8.1 *we have* $P\{M_n \le u_n, m_n > v_n\} - P\{M_n \le u_n\}P\{m_n > v_n\} \to 0$, *with the same relationship holding with* $\{m_n > v_n\}$ *replaced by* $\{m_n \le v_n\}$.

Proof. The probability in (1.8.3) is just

$$\begin{aligned} P\{v_n < \xi_i \le u_n \text{ for } i = 1, 2, \ldots, n\} &= (F(u_n) - F(v_n))^n \\ &= \{1 - F(v_n) - (1 - F(u_n))\}^n \\ &= \left[1 - \frac{\tau + \eta}{n} + o\left(\frac{1}{n}\right)\right]^n \\ &\to e^{-(\tau+\eta)} \end{aligned}$$

if $0 \le \tau, \eta < \infty$ so that the result follows. The cases where τ or η is infinite are dealt with simply since if, e.g. $\tau = \infty$, $P\{M_n \le u_n, m_n > v_n\} \le P\{M_n \le u_n\} \to 0$. □

As a corollary, we find at once that if M_n, m_n have limiting distributions under linear normalizations, then their limiting joint distribution is just the product of these.

Theorem 1.8.3. *Suppose for some constants* $\{a_n > 0\}$, $\{b_n\}$ *and* $\{\alpha_n > 0\}$, $\{\beta_n\}$ *and d.f.'s* G, H,

$$P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x), \tag{1.8.5}$$

$$P\{\alpha_n(m_n - \beta_n) \le x\} \xrightarrow{w} H(x). \tag{1.8.6}$$

Then

$$P\{a_n(M_n - b_n) \le x, \alpha_n(m_n - \beta_n) \le y\} \xrightarrow{w} G(x)H(y). \tag{1.8.7}$$

Proof. This follows from Theorem 1.8.2 by identifying u_n, v_n with $x/a_n + b_n$ and $y/\alpha_n + \beta_n$, τ, η with $-\log G(x)$ and $-\log(1 - H(y))$, respectively. □

The final question of this section concerns the nondegenerate distributions H which are possible in (1.8.6). As suggested above, these may be obtained from the known results for maxima given by the Extremal Types Theorem. The possible limiting laws form the class of *min-stable* distributions, i.e. are the d.f.'s F such that, for each $n = 2, 3, \ldots$, there are constants $a_n > 0$, b_n such that

$$(1 - F(a_n x + b_n))^n = 1 - F(x). \tag{1.8.8}$$

Theorem 1.8.4 (i) (Extremal Types for Minima). *Let* $m_n = \min(\xi_1, \xi_2, \ldots, \xi_n)$, *where* ξ_i *are i.i.d. random variables. If for some constants* $\alpha_n > 0$ *and* β_n, *we have*

$$P\{\alpha_n(m_n - \beta_n) \le x\} \xrightarrow{w} H(x) \tag{1.8.9}$$

for some nondegenerate H, *then* H *is one of the three following extremal types for minima*:

$$\text{Type I:}\quad H(x) = 1 - \exp(-e^x), \qquad -\infty < x < \infty;$$

$$\text{Type II:}\quad H(x) = \begin{cases} 1 - \exp\{-(-x)^{-\alpha}\}, & \alpha > 0, \quad x < 0, \\ 1, & \qquad\qquad x \ge 0; \end{cases}$$

$$\text{Type III:}\quad H(x) = \begin{cases} 0, & \qquad\qquad x < 0, \\ 1 - \exp(-x^\alpha), & \alpha > 0, \quad x \ge 0. \end{cases}$$

(*As with maxima, transformations* $ax + b$ ($a > 0$) *of the argument are permitted for each type.*)

(ii) *The min-stable distributions are those given in* (i) *above.*

Proof. (i) Suppose that (1.8.9) holds so that, writing

$$M_n' = \max(-\xi_1, -\xi_2, \ldots, -\xi_n) = -m_n,$$

$$\begin{aligned} P\{\alpha_n(M_n' + \beta_n) < x\} &= 1 - P\{\alpha_n(M_n' + \beta_n) \ge x\} \\ &= 1 - P\{\alpha_n(m_n - \beta_n) \le -x\} \\ &\to 1 - H(-x) = G(x), \quad \text{say,} \end{aligned}$$

where convergence occurs at all points x of continuity of G. But for such x and $\varepsilon > 0$ such that G is also continuous at $x + \varepsilon$, since

$$P\{\alpha_n(M_n' + \beta_n) < x\} \le P\{\alpha_n(M_n' + \beta_n) \le x\} \le P\{\alpha_n(M_n' + \beta_n) < x + \varepsilon\},$$

we have, letting $n \to \infty$ and then $\varepsilon \to 0$,

$$P\{\alpha_n(M_n' + \beta_n) \le x\} \to G(x),$$

so that G is one of the three (maximal) extreme value d.f.'s. Since

$$H(x) = 1 - G(-x),$$

the three forms listed above follow from the three possible forms for G given by the Extremal Types Theorem.

(ii) If F is min-stable and (1.8.8) holds, then the d.f.

$$G(x) = \lim_{\varepsilon \downarrow 0} 1 - F(-x - \varepsilon) = 1 - F(-x-)$$

satisfies $1 - F(-x) \leq G(x) \leq 1 - F(-x - \varepsilon)$ for all $\varepsilon > 0$, and hence

$$G^n(a_n x - b_n) \leq (1 - F(-a_n(x + \varepsilon) + b_n))^n = 1 - F(-x - \varepsilon) \leq G(x + \varepsilon),$$

and

$$G^n(a_n x - b_n) \geq (1 - F(-a_n x + b_n))^n = 1 - F(-x) \geq G(x - \varepsilon)$$

for any $\varepsilon > 0$. Since $G(x)$ and $G^n(a_n x - b_n)$ are right continuous it follows that $G^n(a_n x - b_n) = G(x)$, so that G is max-stable and, by Theorem 1.4.1, is one of the three extreme value distributions for maxima. This proves part (ii). □

Note that the Type III limit for minima is simply the Weibull distribution—containing the exponential as a special case when $\alpha = 1$. Note also that since the d.f. of $-\xi_i$ is clearly $1 - F(-x-)$, the criteria for domains of attraction for maxima may be readily adapted for minima (replacing the "tail" $1 - F(x)$ by $F(-x-)$ or, indeed, by $F(-x)$). The condition for a Type II limit, for example, then reads (for a distribution F which is not bounded on the left):

$$\lim_{t \to -\infty} \frac{F(tx)}{F(t)} = x^{-\alpha}, \qquad \alpha > 0, \quad \text{for each } x > 0,$$

with corresponding modifications in the other cases.

CHAPTER 2

Exceedances of Levels and kth Largest Maxima

In this chapter, we investigate properties of the exceedances of levels $\{u_n\}$ by $\xi_1, \xi_2, \ldots$, i.e. the points i for which $\xi_i > u_n$, and as consequences, obtain limiting distributional results for the kth largest value among $\xi_1, \ldots, \xi_n$. In particular, when the familiar assumption $n(1 - F(u_n)) \to \tau$ $(0 < \tau < \infty)$ holds (Equation (1.5.1)), it will be shown that the exceedances take on a Poisson character as n becomes large. This leads to the limiting distributions for the kth largest values for any *fixed* rank $k = 1, 2, \ldots$ (the kth "extreme order statistics") and to their limiting joint distributions.

It is obviously of interest to gain some sense of how fast the convergence properties of this chapter and Chapter 1 take place, and questions of this kind will be discussed in Section 2.4.

For our purposes, the Poisson results arising from (1.5.1) are of most interest—especially in dependent cases to be considered in subsequent chapters. However, we also briefly indicate other cases (when $\tau = \infty$) leading to normal distributions for numbers of exceedances, and thence to limiting distributions for kth largest values when the rank $k = k_n$ may depend on the sample size n.

2.1. Poisson Properties of Exceedances

We now look at the choice of u_n which makes (1.5.1) hold in a slightly different light. Let us regard u_n as a "level" (typically becoming higher with n) and say that an *exceedance* of the level u_n by the sequence occurs at "time" i if $\xi_i > u_n$. The probability of such an exceedance is clearly $1 - F(u_n)$, and hence the mean number of exceedances by $\xi_1, \ldots, \xi_n$ is $n(1 - F(u_n)) \to \tau$.

That is, the choice of u_n is made so that the mean number of exceedances by $\xi_1, \ldots, \xi_n$ is approximately constant. We shall pursue this theme further now in developing Poisson properties of the exceedances. In the following, S_n will denote the number of exceedances of a level u_n by $\xi_1, \ldots, \xi_n$.

Theorem 2.1.1. *If $\{\xi_n\}$ is an i.i.d. sequence, $0 \leq \tau \leq \infty$, and if $\{u_n\}$ satisfies* (1.5.1), *i.e.* $n(1 - F(u_n)) \to \tau$, *then for* $k = 0, 1, 2, \ldots,$

$$P\{S_n \leq k\} \to e^{-\tau} \sum_{s=0}^{k} \frac{\tau^s}{s!} \tag{2.1.1}$$

(the right-hand side being taken as zero when $\tau = \infty$).

Conversely, if (2.1.1) *holds for any one fixed k, then* (1.5.1) *holds (and* (2.1.1) *thus holds for all k).*

Proof. We shall show that if S_n is *any* binomial r.v. with parameters n, p_n, and $0 \leq \tau \leq \infty$, then (2.1.1) holds if and only if $np_n \to \tau$. The result will then follow in this particular case, with $p_n = 1 - F(u_n)$.

If S_n is binomial and $np_n \to \tau$, then (2.1.1) follows at once from the standard Poisson limit for the binomial distribution when $0 < \tau < \infty$, and simply when $\tau = 0$ since then $P\{S_n \leq k\} \geq P\{S_n = 0\} = (1 - p_n)^n = (1 - o(1/n))^n$ so that $P\{S_n \leq k\} \to 1$. When $\tau = \infty$, we have, for any $\theta > 0$,

$$P\{S_n \leq k\} \leq \sum_{s=0}^{k} \binom{n}{s} \left(\frac{\theta}{n}\right)^s \left(1 - \frac{\theta}{n}\right)^{n-s}$$

when $np_n \geq \theta$ (the right-hand side being decreasing in θ), so that

$$\limsup_{n\to\infty} P\{S_n \leq k\} \leq e^{-\theta} \sum_{s=0}^{k} \frac{\theta^s}{s!} \to 0 \quad \text{as} \quad \theta \to \infty,$$

giving $\lim_{n\to\infty} P\{S_n \leq k\} = 0$.

Conversely, if (2.1.1) holds for some k but $np_n \not\to \tau$, there exists $\tau' \neq \tau$, $0 \leq \tau' \leq \infty$, and a subsequence $\{n_l\}$ such that $n_l p_{n_l} \to \tau'$. The same argument as above shows that $P\{S_{n_l} \leq k\} \to e^{-\tau'} \sum_{s=0}^{k} \tau'^s/s!$ as $l \to \infty$, which contradicts (2.1.1) since the function $e^{-x} \sum_{s=0}^{k} x^s/s!$ is strictly decreasing in $0 \leq x \leq \infty$. □

Note that if the "time scale" is changed by a factor of n so that when $\xi_i > u_n$, an exceedance is plotted at i/n (rather than at i), S_n is then the number of such plotted exceedances in the unit interval $(0, 1]$ and has a limiting Poisson distribution. Similarly, the number plotted in any (bounded) set B has a limiting Poisson distribution, and the numbers in two or more disjoint sets are clearly independent. This suggests that the exceedances of u_n, if plotted at points i/n rather than at i, behave like a Poisson process on the positive real line when n is large. This is pointed out now for its interest, but will be taken up much more explicitly in subsequent chapters, for dependent cases.

Note also that it would be natural to say that an *upcrossing* of u_n occurs at i if $\xi_i \leq u_n < \xi_{i+1}$. Then the random variable S_n used above is asymptotically the same as the number of *upcrossings* of the level u_n between 1 and n. Thus we may obtain a Poisson limit for the number of such upcrossings. Such Poisson properties of upcrossings will play an important role when we consider continuous time processes.

2.2. Asymptotic Distribution of kth Largest Values

Write, now and subsequently, $M_n^{(k)}$ for the kth largest among $\xi_1, \ldots, \xi_n$ and, as above, S_n for the number of $\xi_1, \ldots, \xi_n$ which exceed u_n. It is readily checked that the events $\{M_n^{(k)} \leq u_n\}$, $\{S_n < k\}$ are identical since, if the former occurs, the kth largest of $\xi_1, \ldots, \xi_n$ does not exceed u_n, and hence no more than $k - 1$ or $\xi_1, \ldots, \xi_n$ exceeds u_n, so that $S_n < k$ and vice versa. Thus

$$P\{M_n^{(k)} \leq u_n\} = P\{S_n < k\}. \tag{2.2.1}$$

By using this relationship, Theorem 2.1.1 may be immediately restated as follows.

Theorem 2.2.1. *Let* $\{\xi_n\}$ *be an i.i.d. sequence. If* $\{u_n\}$ *satisfies* (1.5.1) *for some* τ, $0 \leq \tau \leq \infty$, *then*

$$P\{M_n^{(k)} \leq u_n\} \to e^{-\tau} \sum_{s=0}^{k-1} \frac{\tau^s}{s!}, \tag{2.2.2}$$

$k = 1, 2, \ldots$.

Conversely, if (2.2.2) *holds for some fixed k, then* (1.5.1) *holds and so does* (2.2.2) *for all* $k = 1, 2, \ldots$.

We may further restate this result to give the asymptotic distribution of $M_n^{(k)}$ in terms of that for M_n $(= M_n^{(1)})$.

Theorem 2.2.2. *Suppose that*

$$P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x) \tag{2.2.3}$$

for some nondegenerate (and hence Type I, II, or III) d.f. G. Then, for each $k = 1, 2, \ldots$,

$$P\{a_n(M_n^{(k)} - b_n) \leq x\} \xrightarrow{w} G(x) \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!}, \tag{2.2.4}$$

where $G(x) > 0$ *(and zero where* $G(x) = 0$*).*

Conversely, if for some fixed k,

$$P\{a_n(M_n^{(k)} - b_n) \leq x\} \xrightarrow{w} H(x) \tag{2.2.5}$$

for some nondegenerate H, then H(x) must be of the form on the right-hand side of (2.2.4), *where* (2.2.3) *holds with the same* G, a_n, b_n. *(Hence* (2.2.4) *holds for all k.)*

PROOF. If (2.2.3) holds, then (2.2.2) holds with $k = 1, \tau = -\log G(x) \le \infty$, so that by Theorem 2.2.1, (2.2.2) holds for all k, i.e. (2.2.4) follows.

Conversely, if (2.2.5) holds for some fixed k and x is given, we may clearly find τ, $0 \le \tau \le \infty$, such that

$$H(x) = e^{-\tau} \sum_{s=0}^{k-1} \frac{\tau^s}{s!},$$

since this function decreases continuously from 1 to 0 as τ increases. Thus (2.2.2) holds for this k and hence, by Theorem 2.2.1, for all k including $k = 1$, which gives (2.2.3) with $G(x) = e^{-\tau}$ (nondegeneracy of G being clear). $\square$

We see that for i.i.d. random variables any limiting distribution of the kth largest, $M_n^{(k)}$, has the form (2.2.4) based on the same d.f. G as applied to the maximum, and moreover, that the normalizing constants are the same for all k including $k = 1$ (the maximum itself). Thus we have a complete description of the possible nondegenerate limiting laws.

2.3. Joint Asymptotic Distribution of the Largest Maxima

The asymptotic distribution of the kth largest maximum was obtained above by considering the number of exceedances of a level u_n by $\xi_1, \ldots, \xi_n$. Similar arguments can, and will presently, be adapted to prove convergence of the joint distribution of several large maxima.

Let the levels $u_n^{(1)} \ge \cdots \ge u_n^{(r)}$ satisfy

$$\begin{gathered} n(1 - F(u_n^{(1)})) \to \tau_1, \\ \vdots \\ n(1 - F(u_n^{(r)})) \to \tau_r, \end{gathered} \tag{2.3.1}$$

where $0 \le \tau_1 \le \tau_2 \le \cdots \le \tau_r \le \infty$, and define $S_n^{(k)}$ to be the number of exceedances of $u_n^{(k)}$ by $\xi_1, \ldots, \xi_n$.

Theorem 2.3.1. *Suppose that* $\{\xi_n\}$ *is an i.i.d. sequence and that* $\{u_n^{(k)}\}$, $k = 1, \ldots, r$, *satisfy* (2.3.1). *Then, for* $k_1 \ge 0, \ldots, k_r \ge 0$,

$$P\{S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \ldots, S_n^{(r)} = k_1 + \cdots + k_r\}$$

$$\to \frac{\tau_1^{k_1}}{k_1!} \cdot \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!} \cdot \ldots \cdot \frac{(\tau_r - \tau_{r-1})^{k_r}}{k_r!} e^{-\tau_r} \tag{2.3.2}$$

as $n \to \infty$ (the right-hand side being interpreted as zero if $\tau_r = \infty$, regardless of whether any other $\tau_i = \infty$).

PROOF. Writing $p_{n,k} = 1 - F(u_n^{(k)})$ for the probability that ξ_1 exceeds $u_n^{(k)}$, it is easy to see that the left-hand side of (2.3.2) equals

$$\binom{n}{k_1} p_{n,1}^{k_1} \binom{n-k_1}{k_2} (p_{n,2} - p_{n,1})^{k_2} \cdots$$
$$\cdot \binom{n - k_1 - \cdots - k_{r-1}}{k_r} (p_{n,r} - p_{n,r-1})^{k_r} \cdot (1 - p_{n,r})^{n-k_1-\cdots-k_r}. \quad (2.3.3)$$

From (2.3.1), it follows in turn that (for $\tau_r < \infty$)

$$\binom{n}{k_1} p_{n,1}^{k_1} = \frac{n \cdot \cdots \cdot (n - k_1 + 1) p_{n,1}^{k_1}}{k_1!} \to \frac{\tau_1^{k_1}}{k_1!},$$

$$\binom{n - k_1 - \cdots - k_{l-1}}{k_l} (p_{n,l} - p_{n,l-1})^{k_l}$$
$$= \frac{(n - k_1 - \cdots - k_{l-1}) \cdot \cdots \cdot (n - k_1 - \cdots - k_l + 1)(p_{n,l} - p_{n,l-1})^{k_l}}{k_l!}$$
$$\to \frac{(\tau_l - \tau_{l-1})^{k_l}}{k_l!}$$

for $2 \le l \le r$, and that

$$(1 - p_{n,r})^{n-k_1-\cdots-k_r} \to e^{-\tau_r},$$

and thus (2.3.2) is an immediate consequence of (2.3.1) and (2.3.3) when τ_r, and hence all τ_i, are finite. On the other hand, if $\tau_r = \infty$, (2.3.2) does not exceed $P\{S_n^{(r)} = k_1 + \cdots + k_r\}$, which tends to zero by Theorem 2.1.1, so that again (2.3.2) holds. □

Clearly, reasoning as in (2.2.1),

$$P\{M_n^{(1)} \le u_n^{(1)}, \ldots, M_n^{(r)} \le u_n^{(r)}\}$$
$$= P\{S_n^{(1)} = 0, S_n^{(2)} \le 1, \ldots, S_n^{(r)} \le r - 1\}, \quad (2.3.4)$$

and thus the joint asymptotic distribution of the r largest maxima can be obtained directly from Theorem 2.3.1. In particular, if the distribution of $a_n(M_n^{(1)} - b_n)$ converges, then it follows not only that $a_n(M_n^{(k)} - b_n)$ converges in distribution for $k = 2, 3, \ldots$ as was seen above, but also that the joint distribution of $a_n(M_n^{(1)} - b_n), \ldots, a_n(M_n^{(r)} - b_n)$ converges. This is, of course, completely straightforward, but since the form of the limiting distribution becomes somewhat complicated if more than two maxima are considered, we state the result only for the two largest maxima.

Theorem 2.3.2. *Suppose that*

$$P\{a_n(M_n^{(1)} - b_n) \le x\} \xrightarrow{w} G(x) \tag{2.3.5}$$

for some nondegenerate (and hence Type I, II, or III) d.f. G. Then, for $x_1 > x_2$,

$$\begin{aligned} P\{a_n(M_n^{(1)} - b_n) \le x_1, a_n(M_n^{(2)} - b_n) \le x_2\} \\ \xrightarrow{w} G(x_2)\{\log G(x_1) - \log G(x_2) + 1\}, \end{aligned} \tag{2.3.6}$$

when $G(x_2) > 0$ (and to zero when $G(x_2) = 0$).

PROOF. We have to prove that

$$P\{M_n^{(1)} \le u_n^{(1)}, M_n^{(2)} \le u_n^{(2)}\}$$

converges if $u_n^{(1)} = x_1/a_n + b_n$ and $u_n^{(2)} = x_2/a_n + b_n$. If (2.3.5) holds, then by Theorem 1.5.1, $n(1 - F(u_n^{(1)})) \to \tau_1$, $n(1 - F(u_n^{(2)})) \to \tau_2$, where $\tau_1 = -\log G(x_1)$, $\tau_2 = -\log G(x_2)$ $(0 \le \tau_1 \le \tau_2 \le \infty)$. Hence, by Theorem 2.3.1,

$$\begin{aligned} P\{S_n^{(1)} = 0, S_n^{(2)} \le 1\} \\ &= P\{S_n^{(1)} = 0, S_n^{(2)} = 0\} + P\{S_n^{(1)} = 0, S_n^{(2)} = 1\} \\ &\to e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2} = e^{-\tau_2}(\tau_2 - \tau_1 + 1), \end{aligned}$$

which by (2.3.4) proves (2.3.6). □

2.4. Rate of Convergence

Every convergence result is accompanied by the question of rate of convergence. We shall in this section examine the basic limit theorems for maxima (Theorems 1.4.2, 1.5.1, 1.5.3), the Poisson limit theorem (Theorem 2.1.1), and the convergence theorems for kth maxima (Theorems 2.2.1 and 2.2.2) from this point of view. Let $\{u_n\}$ be a sequence of levels, where u_n may or may not be one of a family $u_n = u_n(x) = x/a_n + b_n$, and let as before F be the common d.f. of the i.i.d. sequence $\xi_1, \xi_2, \ldots$. If $\{u_n\}$ satisfies the hypothesis of Theorem 1.5.1, i.e. if $n(1 - F(u_n)) \to \tau$, then, writing $\tau_n = n(1 - F(u_n))$,

$$P\{M_n \le u_n\} = \left(1 - \frac{\tau_n}{n}\right)^n \to e^{-\tau}. \tag{2.4.1}$$

If the d.f. F is continuous, one can always obtain equality in (2.4.1) for any τ, n (by taking $u_n = F^{-1}(e^{-\tau/n})$), but often $\{u_n\}$ will be determined by other considerations, e.g. it may be of the form $u_n(x) = x/a_n + b_n$, in which case it is not possible to have $(1 - \tau_n(x)/n)^n \equiv e^{-\tau(x)}$, for $\tau_n(x) = n(1 - F(u_n(x)))$, unless F is max-stable.

It is instructive to consider separately the two approximations

$$\left(1 - \frac{\tau_n}{n}\right)^n \approx e^{-\tau_n} \tag{2.4.2}$$

and

$$e^{-\tau_n} \approx e^{-\tau}, \tag{2.4.3}$$

which together make up (2.4.1). As will be seen shortly, there is a quite satisfying uniform bound of the order $1/n$ on the approximation in (2.4.2). As for (2.4.3), a Taylor expansion gives a precise pointwise estimate of the approximation, but unfortunately it does not seem possible to find a useful uniform bound on (2.4.3) for $\tau_n = \tau_n(x)$, $\tau = \tau(x)$, $-\infty < x < \infty$. The essential part of proving these results is contained in the following lemma.

Lemma 2.4.1. (i) *If* $0 \le x \le n$ *then*

$$\begin{aligned} 0 \le e^{-x} - \left(1 - \frac{x}{n}\right)^n &\le \frac{x^2 e^{-x}}{2} \cdot \frac{1}{n-1} \\ &\le 2e^{-2} \cdot \frac{1}{n-1} \\ &\le 0.3 \cdot \frac{1}{n-1} \quad \text{for } n = 1, 2, \ldots, \end{aligned} \tag{2.4.4}$$

and further

$$e^{-x} - \left(1 - \frac{x}{n}\right)^n = \frac{x^2 e^{-x}}{2}\frac{1}{n}\left(1 + O\left(\frac{1}{n}\right)\right) \quad \text{as } n \to \infty, \tag{2.4.5}$$

uniformly for x in bounded intervals.

(ii) *If* $x - y \le \log 2$ *then*

$$e^{-y} - e^{-x} = e^{-x}\{(x-y) + \theta(x-y)^2\}, \tag{2.4.6}$$

with $0 < \theta < 1$.

Proof. (i) The first inequality in (2.4.4) is an immediate consequence of the inequality $e^{-x/n} \ge 1 - x/n$, and since the third and fourth inequalities are obvious only the second one remains to be established. We shall instead prove the equivalent inequality

$$0 \le \frac{x^2}{2(n-1)} - 1 + e^x\left(1 - \frac{x}{n}\right)^n = f(x), \quad \text{say}. \tag{2.4.7}$$

Now,

$$f'(x) = \frac{x}{n-1}\left\{1 - e^x\left(1 - \frac{x}{n}\right)^{n-1}\left(1 - \frac{1}{n}\right)\right\},$$

and the expression within brackets assumes its minimum in $[0, n]$ for $x = 1$. Since $1 - e(1 - 1/n)^n \geq 0$ according to the first inequality in (2.4.4), this shows that $f'(x) \geq 0$ for $0 \leq x \leq n$, and since $f(0) = 0$, (2.4.7) follows.

Next, let $f_n(x) = 1 - \exp\{x + n\log(1 - x/n)\}$. It is then straightforward to check that $f_n(0) = f_n'(0) = 0$, $f_n''(0) = 1/n$, and that $f_n'''(x) = O(1/n^2)$, uniformly for x in bounded intervals. Hence, Taylor expansion gives that

$$e^{-x} - \left(1 - \frac{x}{n}\right)^n = f_n(x)e^{-x} = \frac{x^2e^{-x}}{2}\frac{1}{n}\left(1 + O\left(\frac{1}{n}\right)\right),$$

uniformly for x in bounded intervals, which proves (2.4.5).

(ii) Again, by Taylor's formula,

$$\begin{aligned} e^{-y} - e^{-x} &= e^{-x}\{e^{x-y} - 1\} \\ &= e^{-x}\{x - y + \tfrac{1}{2}(x - y)^2 e^{\theta'(x-y)}\}, \end{aligned}$$

with $0 < \theta' < 1$, which proves (2.4.6) since $0 < \exp\{\theta'(x - y)\}/2 < 1$ for $x - y \leq \log 2$. □

First-order bounds on the rate of convergence now follow simply. The reader is also referred to Dziubdziela (1978) and Galambos (1978) for related results.

Theorem 2.4.2. *Let $\{\xi_n\}$ be an i.i.d. sequence, put $\tau_n = n(1 - F(u_n))$, and write*

$$\Delta_n = \left(1 - \frac{\tau_n}{n}\right)^n - e^{-\tau_n}, \qquad \Delta_n' = e^{-\tau_n} - e^{-\tau},$$

so that

$$P\{M_n \leq u_n\} - e^{-\tau} = \Delta_n + \Delta_n'.$$

Then

$$0 \leq -\Delta_n \leq \frac{\tau_n^2 e^{-\tau_n}}{2}\cdot\frac{1}{n-1} \leq 0.3\cdot\frac{1}{n-1},$$

where the first bound is asymptotically sharp, in the sense that if $\tau_n \to \tau$ then $\Delta_n \sim -(\tau^2 e^{-\tau}/2)/n$. Furthermore, for $\tau - \tau_n \leq \log 2$,

$$\Delta_n' = e^{-\tau}\{(\tau - \tau_n) + \theta(\tau - \tau_n)^2\},$$

with $0 < \theta < 1$.

PROOF. Since $P\{M_n \leq u_n\} = (1 - \tau_n/n)^n$ this follows at once from Lemma 2.4.1, after noting that $0 \leq \tau_n = n(1 - F(u_n)) \leq n$. □

If $\tau_n \to \tau$, so that (2.4.1) holds, for $u_n = u_n(x) = x/a_n + b_n$, $\tau_n = \tau_n(x)$, $\tau = \tau(x)$, then by Theorem 1.2.3, (2.4.1) holds also if a_n and b_n are replaced by different constants α_n and β_n such that $\alpha_n/a_n \to 1$, $(\beta_n - b_n)a_n \to 0$. However, the speed of convergence to zero of Δ_n' (and thus of $P\{M_n \le u_n\}$ to $e^{-\tau}$) clearly can be quite different for different choices of the constants, and one may be interested in finding good choices of a_n and b_n from this point of view. Further, even if the "best" a_n's and b_n's are used the rates of convergence can be completely different for different distributions.

As was noted already by Fisher and Tippett, extremes from the normal distribution converge remarkably slowly to their limiting form. In fact, for a_n and b_n as in Theorem 1.5.3, and with $u_n = u_n(x) = x/a_n + b_n$, we have

$$\begin{aligned}\frac{u_n^2}{2} = x + \log n - \log(2\log n)^{1/2} \\ - \log\sqrt{2\pi} + \frac{(\log\log n)^2}{16\log n} \\ + \frac{o((\log\log n)^2)}{\log n},\end{aligned}$$

and hence, using

$$\frac{\phi(x)}{x}(1 - x^{-2}) \le 1 - \Phi(x) \le \frac{\phi(x)}{x} \quad \text{for } x \ge 0,$$

$$\begin{aligned}\tau_n &= n(1 - \Phi(u_n)) = \frac{n}{u_n}\phi(u_n)(1 + O(u_n^{-2})) \\ &= \frac{n}{\{2\log n\}^{1/2}} \cdot \frac{1}{\sqrt{2\pi}} \exp\{-x - \log n + \log(2\log n)^{1/2} \\ &\quad + \log\sqrt{2\pi} - (\log\log n)^2(16\log n)^{-1}(1 + o(1))\} \\ &\quad \times \left(1 + O\left(\frac{\log\log n}{\log n}\right)\right) \\ &= e^{-x}\{1 - (\log\log n)^2(16\log n)^{-1}(1 + o(1))\}.\end{aligned}$$

Thus, for $\tau(x) = e^{-x}$,

$$\tau(x) - \tau_n(x) \sim \frac{e^{-x}}{16}\frac{(\log\log n)^2}{\log n},$$

and hence, as the error Δ_n in Theorem 2.4.2 is of the order n^{-1}, for i.i.d. standard normal r.v.'s $\{\xi_n\}$,

$$\begin{aligned}P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x}) \\ \sim \frac{\exp(-e^{-x})e^{-x}}{16}\frac{(\log\log n)^2}{\log n},\end{aligned} \tag{2.4.8}$$

so the convergence in Theorem 1.5.3 is extremely slow. Of course this might depend on the particular choice of a_n and b_n used above. P. Hall (1979) investigated this problem further, and by elementary, but rather complicated, calculations proved that if a_n and b_n are chosen as solutions to

$$\frac{n\phi(a_n)}{a_n} = 1, \qquad b_n = a_n, \tag{2.4.9}$$

then

$$\frac{C_1}{\log n} < \sup_{-\infty < x < \infty} |P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})| \le \frac{C_2}{\log n}, \tag{2.4.10}$$

for some strictly positive constants C_1, C_2, with $C_2 \le 3$. He further proved that one cannot improve on this rate of convergence by choosing a_n and b_n in some other way. Thus, even if the constants a_n and b_n used in Theorem 1.5.3 do not give the optimal rate of convergence, not very much can be gained by using different a_n and b_n.

Figures 2.4.1 and 2.4.2 illustrate the rate of convergence for normal r.v.'s. It can be seen that notwithstanding the slow rate of convergence, the differences $P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})$ are fairly small even for small n. The problem is just that the fit improves little with increasing n. This is explained by the form of the right-hand side of (2.4.8). The first factor, $\exp(-e^{-x})e^{-x}/16$, is rather small, with a maximum of 0.023 for $x = 0$, while the second factor, $(\log\log n)^2/\log n$, is virtually constant for moderate values of n. For example, for $10^2 \le n \le 10^{10}$ it varies between 0.43 and 0.54. Thus, the first-order approximation to $P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})$ changes very little for n in this range, but on the other hand it is small (≤ 0.013). In fact, there is some improvement of the approximation for n increasing from, say, 10^3 to 10^6, but this is due to higher-order effects. Further, for moderate values of n it does not make much difference in the maximal error whether one defines a_n, b_n by (2.4.9) or as in Theorem 1.5.3.

Next, we shall turn to the rate of convergence in Theorem 2.1.1, i.e. to the question of how well the distribution of $S_n = \sum_{i=1}^n \chi_{\{\xi_i > u_n\}}$ is approximated by a Poisson distribution with mean τ, where χ is the indicator function, equal to 1 if the event within brackets occurs, and zero otherwise.

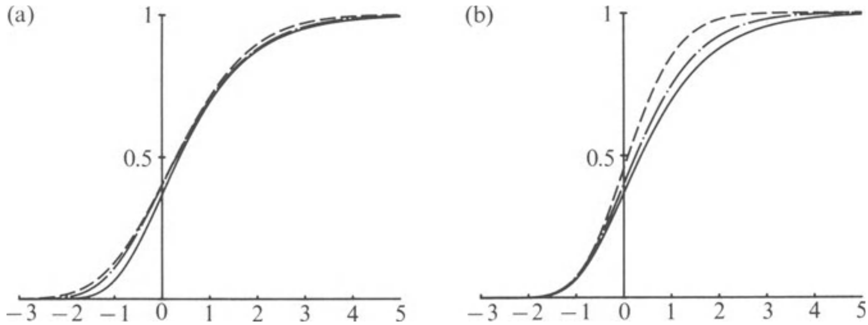


Figure 2.4.1. Plot of $\exp(-e^{-x})$ (——) and of $P\{a_n(M_n - b_n) \le x\}$ for $n = 10$ (---) and $n = 10^3$ (-··-); (a) for a_n, b_n given by (1.7.2), (b) for a_n, b_n given by (2.4.9).

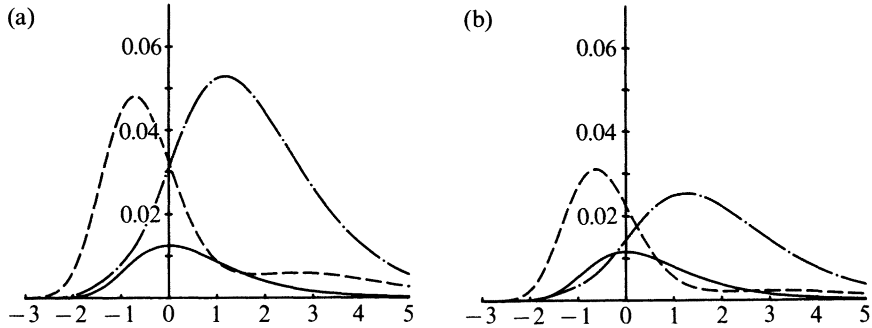


Figure 2.4.2. Plot of $(\exp(-e^{-x})e^{-x}/16)(\log \log n)^2/\log n$ (——) and of

$$P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})$$

for a_n, b_n given by (1.7.2) (– – –) and as given by (2.4.9) (–·–); (a) $n = 10^3$, (b) $n = 10^6$.

Suppose X and Y are positive, integer-valued r.v.'s. We then define the *variation distance* d between their d.f.'s as

$$d(X, Y) = \frac{1}{2} \sum_{k=0}^{\infty} |P\{X = k\} - P\{Y = k\}|.$$

Clearly $d(X, Y) \ge 0$, with $d(X, Y) = 0$ if and only if X and Y have the same distribution, $d(X, Y) = d(Y, X)$, and

$$d(X, Z) \le d(X, Y) + d(Y, Z), \tag{2.4.11}$$

so d is a metric on the space of distributions on the positive integers. Further, it is easily seen that d is a metric for convergence in distribution of positive, integer-valued r.v.'s, i.e. $d(X_n, X) \to 0$ if and only if X_n tends to X in distribution. We shall use a simple and very elegant approach due to Serfling (1978) in deriving variation distance bounds for the convergence in Theorem 2.1.1. We first note some basic properties of the variation distance which are of use in the proofs below, and which also show the interest and usefulness of this distance. Clearly, for h denoting a real-valued function,

$$d(X, Y) = \tfrac{1}{2} \sup_{|h| \le 1} |E(h(X)) - E(h(Y))|. \tag{2.4.12}$$

Further, writing $a^+ = \max(a, 0)$, $a^- = -\min(a, 0)$,

$$0 = \sum_{k=0}^{\infty} (P\{X = k\} - P\{Y = k\}) = \sum_{k=0}^{\infty} (P\{X = k\} - P\{Y = k\})^+ - \sum_{k=0}^{\infty} (P\{X = k\} - P\{Y = k\})^-,$$

and hence

$$\begin{aligned} d(X, Y) &= \sum_{k=0}^{\infty} (P\{X = k\} - P\{Y = k\})^+ \\ &= \sum_{k=0}^{\infty} (P\{X = k\} - P\{Y = k\})^-, \end{aligned}$$

from which it follows easily (e.g. taking $A = \{k; P\{X = k\} \geq P\{Y = k\}\}$), that

$$d(X, Y) = \sup_A |P\{X \in A\} - P\{Y \in A\}| \geq \sup_{k=1,2,\ldots} |P\{X < k\} - P\{Y < k\}|, \tag{2.4.13}$$

where the supremum is over all sets A of integers. We note in passing that since $\{M_n^{(k)} \leq u_n\} = \{S_n < k\}$ the latter inequality shows that variation bounds for S_n directly lead to corresponding bounds for $M_n^{(k)}$.

It will be convenient to extend the notation by writing $d(X, G)$ for the variation distance between the d.f. of X and the d.f. G and $d(F, G)$ for the distance between the d.f.'s F and G. Further we shall denote the d.f. of a Poisson r.v. with mean τ by $P(\tau)$ and that of a binomial r.v. with parameters n and p by $B(n, p)$. The following lemma is due to Serfling (1978).

Lemma 2.4.3. (i) *Suppose X and Y are defined on the same probability space. Then*

$$d(X, Y) \leq P(X \neq Y).$$

(ii) *If $X_1, \ldots, X_n$ are mutually independent and $Y_1, \ldots, Y_n$, similarly, are mutually independent, then*

$$d\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) \leq \sum_{i=1}^n d(X_i, Y_i).$$

(iii) $$d(P(\tau_1), P(\tau_2)) \leq |\tau_1 - \tau_2|.$$

(iv) $$d\left(B\left(n, \frac{\tau}{n}\right), P(\tau)\right) \leq \frac{\tau^2}{n}.$$

Proof. (i) follows easily from (2.4.13), since

$$|P\{X \in A\} - P\{Y \in A\}| \leq P\{X \in A, Y \notin A\} + P\{X \notin A, Y \in A\} \leq P\{X \neq Y\}.$$

Next, suppose the hypothesis of (ii) is satisfied and let h be a function with $|h(x)| \leq 1$ for all x, so that by (2.4.12) $|E(h(X_1)) - E(h(Y_1))|/2 \leq d(X_1, Y_1)$ and, similarly, $|E(h(X_1 + k)) - E(h(Y_1 + k))|/2 \leq d(X_1, Y_1)$ for $k = 1, 2, \ldots$. Thus, since X_1 is independent of X_2 and Y_1 is independent of Y_2,

$$\begin{aligned}
&\frac{|E(h(X_1 + X_2)) - E(h(Y_1 + Y_2))|}{2} \\
&= \frac{1}{2}\left|\sum_{k=0}^{\infty} \{E(h(X_1 + k))P\{X_2 = k\} - E(h(Y_1 + k))P\{Y_2 = k\}\}\right| \\
&\leq \frac{1}{2}\sum_{k=0}^{\infty} |E(h(X_1 + k)) - E(h(Y_1 + k))|P\{X_2 = k\} \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty} E(h(Y_1 + k))|P\{X_2 = k\} - P\{Y_2 = k\}| \\
&\leq \sum_{k=0}^{\infty} d(X_1, Y_1)P\{X_2 = k\} + \frac{1}{2}\sum_{k=0}^{\infty} |P\{X_2 = k\} - P\{Y_2 = k\}| \\
&= d(X_1, Y_1) + d(X_2, Y_2).
\end{aligned}$$

By (2.4.12) this proves that

$$d(X_1 + X_2, Y_1 + Y_2) \le d(X_1, Y_1) + d(X_2, Y_2),$$

and the general case then follows by induction.

To prove (iii), suppose first $\tau_1 \ge \tau_2$, and let X and Y have independent Poisson distributions with means $\tau_1 - \tau_2$ and τ_2, respectively. Then by (i),

$$\begin{aligned} d(P(\tau_1), P(\tau_2)) &= d(X + Y, Y) \le P\{X > 0\} \\ &= 1 - \exp\{-(\tau_1 - \tau_2)\} \le \tau_1 - \tau_2. \end{aligned}$$

The same argument for $\tau_1 < \tau_2$ concludes the proof of (iii).

We shall first prove (iv) for $n = 1$. Let X be binomial with parameters 1 and τ and let Y have a Poisson distribution with mean τ. Then

$$\begin{aligned} |P\{X = 0\} - P\{Y = 0\}| &= e^{-\tau} - 1 + \tau, \\ |P\{X = 1\} - P\{Y = 1\}| &= \tau - \tau e^{-\tau}, \\ \sum_{k=2}^{\infty} |P\{X = k\} - P\{Y = k\}| &= P\{Y \ge 2\} = 1 - e^{-\tau} - \tau e^{-\tau} \end{aligned}$$

and thus, using the definition of d.

$$d(X, Y) = \tfrac{1}{2}(2\tau - 2\tau e^{-\tau}) = \tau(1 - e^{-\tau}) \le \tau^2.$$

Next, let the r.v.'s $X_1, \ldots, X_n$, $Y_1, \ldots, Y_n$ be mutually independent, with $X_1, \ldots, X_n$ binomial with parameters 1, τ/n, and with $Y_1, \ldots, Y_n$ having Poisson distributions with means τ/n. Then, by (ii),

$$\begin{aligned} d\left(B\left(n, \frac{\tau}{n}\right), P(\tau)\right) &= d\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Y_i\right) \\ &\le \sum_{i=1}^{n} d(X_i, Y_i) \\ &\le n\left(\frac{\tau}{n}\right)^2 = \frac{\tau^2}{n}. \end{aligned}$$

□

As in Theorem 2.4.2, the error in approximating the distribution of S_n, the number of exceedances of u_n by $\xi_1, \ldots, \xi_n$, by a Poisson distribution with mean τ can be split up into two parts; the first error term is of the order $1/n$ and arises from approximating the binomial distribution of S_n by a Poisson distribution with mean τ_n, and the second part comes from replacing τ_n by τ.

Theorem 2.4.4. *Suppose $\{\xi_n\}$ is an i.i.d. sequence with d.f. F, put*

$$\tau_n = n(1 - F(u_n)),$$

and let $S_n = \sum_{i=1}^{n} \chi_{\{\xi_i > u_n\}}$. Then

$$d(S_n, P(\tau_n)) \le \frac{\tau_n^2}{n}$$

and

$$d(S_n, P(\tau)) \le \frac{\tau_n^2}{n} + |\tau_n - \tau|.$$

PROOF. Since S_n is binomial with parameters n and τ_n/n, this follows from Lemma 2.4.3(iii) and (iv), and the fact that

$$d(S_n, P(\tau)) \le d(S_n, P(\tau_n)) + d(P(\tau_n), P(\tau)). \qquad \square$$

Corollary 2.4.5. *With hypothesis and notation as in Theorem 2.4.4, let $M_n^{(k)}$ be the kth largest among $\xi_1, \ldots, \xi_n$. Then*

$$\left| P\{M_n^{(k)} \le u_n\} - \sum_{s=0}^{k-1} e^{-\tau_n} \frac{\tau_n^s}{s!} \right| \le \frac{\tau_n^2}{n}$$

and

$$\left| P\{M_n^{(k)} \le u_n\} - \sum_{s=0}^{k-1} e^{-\tau} \frac{\tau^s}{s!} \right| \le \frac{\tau_n^2}{n} + |\tau_n - \tau|.$$

PROOF. These follow at once from the theorem and the relation $\{S_n < k\} = \{M_n^{(k)} \le u_n\}$, using (2.4.13). $\square$

By comparing the corollary with Theorem 2.4.2 it is seen that it gives the right order of convergence but that the constants in the bounds are too large.

2.5. Increasing Ranks

The results of Sections 2.2 and 2.3 apply to the kth largest $M_n^{(k)}$ of $\xi_1, \xi_2, \ldots, \xi_n$ when k is fixed. We refer to this as the case of *fixed ranks* (or *extreme* order statistics). It is also of interest to consider cases where $k = k_n \to \infty$ as $n \to \infty$, and we shall call this the case of *increasing* ranks. Two particular rates of increase are of special interest:

(i) $k_n/n \to \theta$ $(0 < \theta < 1)$, which we shall call the case of *central* ranks;
(ii) $k_n \to \infty$ but $k_n/n \to 0$, which will be referred to as the *intermediate* rank case.

For the consideration of fixed ranks, it was useful to define levels $\{u_n\}$ satisfying (1.5.1), i.e. $n(1 - F(u_n)) \to \tau$. In the case where $k_n \to \infty$, we shall find that the appropriate restrictions are that $nF(u_n)(1 - F(u_n)) \to \infty$, and writing $p_n = 1 - F(u_n)$,

$$\frac{k_n - np_n}{(np_n(1 - p_n))^{1/2}} \to \tau \tag{2.5.1}$$

for a fixed constant τ, or equivalently (as we shall see),

$$n - k_n \to \infty, \qquad \frac{k_n - np_n}{(k_n(1 - k_n/n))^{1/2}} \to \tau. \tag{2.5.2}$$

Theorem 1.5.1 now has the following counterpart (in which S_n is again the numbers of exceedances of a level u_n by $\xi_1, \xi_2, \ldots, \xi_n$).

Theorem 2.5.1. *With the above notation, let $k_n \to \infty$, write $p_n = 1 - F(u_n)$, and suppose $np_n(1 - p_n) \to \infty$. If $\{u_n\}$ satisfies (2.5.1), then*

$$P\{S_n \leq k_n\} \to \Phi(\tau) \quad \text{as } n \to \infty. \tag{2.5.3}$$

Conversely, if (2.5.3) holds so does (2.5.1).

In the above statements, (2.5.1) can be replaced by the equivalent condition (2.5.2).

PROOF. We may write $S_n = \sum_1^n \chi_i$, where $\chi_i = 1$ or 0 according as $\xi_i > u_n$ or $\xi_i \leq u_n$. The χ_i are thus i.i.d. with $P\{\chi_i = 1\} = p_n = 1 - P\{\chi_i = 0\}$. It follows from the Berry–Esseen bound that, for some constant C,

$$\left| P\{S_n \leq k_n\} - \Phi\left(\frac{k_n - np_n}{(np_n(1 - p_n))^{1/2}}\right) \right| \leq \frac{C}{(np_n(1 - p_n))^{1/2}},$$

which tends to zero since $np_n(1 - p_n) \to \infty$. The main result follows since

$$\Phi\left(\frac{k_n - np_n}{(np_n(1 - p_n))^{1/2}}\right) \to \Phi(\tau)$$

if and only if $(k_n - np_n)/(np_n(1 - p_n))^{1/2} \to \tau$ (Φ and its inverse function both being continuous).

Finally, that (2.5.1) implies (2.5.2) follows by writing

$$k_n = np_n + \tau(np_n(1 - p_n))^{1/2}(1 + o(1))$$

and noting that this implies $k_n \sim np_n$ and $(n - k_n) \sim n(1 - p_n)$. Similarly, (2.5.2) implies (2.5.1). □

Corresponding to Theorems 2.2.1 and 2.2.2, we thus have the following results.

Theorem 2.5.2. *With the above notation, suppose that $k_n \to \infty, np_n(1 - p_n) \to \infty$ ($p_n = 1 - F(u_n)$). If (2.5.1) or (2.5.2) holds, then*

$$P\{M_n^{(k_n)} \leq u_n\} \to \Phi(\tau). \tag{2.5.4}$$

Conversely, if (2.5.4) holds so do (2.5.1) and (2.5.2).

PROOF. If (2.5.1) holds, it also holds with k_n replaced by $k_n - 1$, so that by Theorem 2.5.1,

$$P\{S_n < k_n\} = P\{S_n \leq k_n - 1\} \to \Phi(\tau),$$

and hence (2.5.4) follows from (2.2.1). The converse follows along the same lines. □

Theorem 2.5.3. *Again, with the above notation, suppose that* (2.5.1) *or* (2.5.2) *holds with* $u_n = u_n(x) = x/a_n + b_n$ $(\tau = \tau(x))$ *for some sequences* $\{a_n > 0\}$, $\{b_n\}$. *Then*

$$P\{a_n(M_n^{(k_n)} - b_n) \leq x\} \xrightarrow{w} H(x), \tag{2.5.5}$$

where $H(x) = \Phi(\tau(x))$. *Conversely, if* (2.5.5) *holds for some nondegenerate d.f.* H, *then we have* $H(x) = \Phi(\tau(x))$, *where* (2.5.1) *and* (2.5.2) *hold with* $u_n = x/a_n + b_n$, $\tau = \tau(x)$.

2.6. Central Ranks

The case of central ranks, where $k_n/n \to \theta$ $(0 < \theta < 1)$, has been studied in Smirnov (1952). While we shall have little to say about this in later chapters, for the sake of completeness, a few basic facts for the i.i.d. sequence will be discussed here. First, we note that it is possible for two sequences $\{k_n\}$, $\{k_n'\}$ with $\lim k_n/n = \lim k_n'/n$ to lead to *different* nondegenerate limiting d.f.'s for $M_n^{(k_n)}$, $M_n^{(k_n')}$. Specifically, as shown in Smirnov (1952), we may have $k_n/n \to \theta$, $k_n'/n \to \theta$ and

$$P\{a_n(M_n^{(k_n)} - b_n) \leq x\} \xrightarrow{w} H(x), \tag{2.6.1}$$

$$P\{a_n'(M_n^{(k_n')} - b_n') \leq x\} \xrightarrow{w} H'(x), \tag{2.6.2}$$

where $a_n > 0, b_n, a_n' > 0, b_n'$ are constants and $H(x)$, $H'(x)$ are nondegenerate d.f.'s of different type. However, this is not possible if

$$\sqrt{n}\left(\frac{k_n}{n} - \theta\right) \to 0, \tag{2.6.3}$$

as the following result shows.

Lemma 2.6.1. *Suppose that* (2.6.1) *and* (2.6.2) *hold, where* H, H' *are nondegenerate and* k_n, k_n' *both satisfy* (2.6.3). *Then* H *and* H' *are of the same type, i.e.* $H'(x) = H(ax + b)$ *for some* $a > 0, b$.

Proof. If the terms of the i.i.d. sequence $\xi_1, \xi_2, \ldots$ have d.f. F, then by Theorem 2.5.3,

$$\frac{k_n - n(1 - F(x/a_n + b_n))}{(k_n(1 - k_n/n))^{1/2}} \to \tau(x), \tag{2.6.4}$$

where $H(x) = \Phi(\tau(x))$. By (2.6.3), we then have

$$\sqrt{n}\,\frac{\theta - (1 - F(x/a_n + b_n))}{(\theta(1 - \theta))^{1/2}} \to \tau(x).$$

Again by (2.6.3), with k_n replaced by k'_n, we must therefore have that (2.6.4) holds with k'_n replacing k_n, and hence by Theorem 2.5.3, that

$$P\{a_n(M_n^{(k'_n)} - b_n) \le x\} \to \Phi(\tau(x)) = H(x).$$

But if H_n is the d.f. of $M_n^{(k'_n)}$, this says that $H_n(x/a_n + b_n) \xrightarrow{w} H(x)$, whereas also $H_n(x/a'_n + b'_n) \xrightarrow{w} H'(x)$ by (2.6.2). Thus by Theorem 1.2.3, H and H' are of the same type, as required. □

It turns out that, for sequences $\{k_n\}$ satisfying (2.6.3), just four forms of limiting distributions H satisfying (2.6.1) are possible for $M_n^{(k_n)}$. For completeness, we state this here as a theorem—and refer to Smirnov (1952) for proof.

Theorem 2.6.2. *If the central rank sequence* $\{k_n\}$ *satisfies* (2.6.3), *the only possible nondegenerate d.f.'s H for which* (2.6.1) *holds are*

1. $H(x) = \begin{cases} 0, & x < 0, \\ \Phi(cx^\alpha), & x \ge 0, \end{cases} \qquad c > 0, \alpha > 0;$

2. $H(x) = \begin{cases} \Phi(-c|x|^\alpha), & x < 0, \\ 1, & x \ge 0; \end{cases} \qquad c > 0, \alpha > 0,$

3. $H(x) = \begin{cases} \Phi(-c_1|x|^\alpha), & x < 0, \\ \Phi(c_2 x^\alpha), & x \ge 0, \end{cases} \qquad c_1 > 0, c_2 > 0, \alpha > 0;$

4. $H(x) = \begin{cases} 0, & x < -1, \\ \frac{1}{2}, & -1 \le x < 1, \\ 1, & x \ge 1. \end{cases}$

It may be noted that only the third of these distributional types is continuous—in contrast to the situation for the fixed rank cases.

If the restriction (2.6.3) is removed, the situation becomes more complicated, and the range of possible limit distributions is much larger. These questions, as well as domains of attraction, are discussed in two papers by Balkema and de Haan (1978a, b).

2.7. Intermediate Ranks

By an *intermediate rank sequence*, we mean a sequence $\{k_n\}$ such that $k_n \to \infty$ but $k_n = o(n)$. The general theory for increasing ranks applies, with some slight simplification. For example, we may rephrase (2.5.2) as

$$p_n = \frac{k_n}{n} - \tau\frac{k_n^{1/2}}{n} + o\left(\frac{k_n^{1/2}}{n}\right).$$

The following result of Wu (1966) gives the possible normalized d.f.'s of $M_n^{(k_n)}$ when k_n is nondecreasing.

Theorem 2.7.1. *If $\xi_1, \xi_2, \ldots$ are i.i.d. and $\{k_n\}$ is a nondecreasing intermediate rank sequence, and if there are constants $a_n > 0$ and b_n such that*

$$P\{a_n(M_n^{(k_n)} - b_n) \le x\} \xrightarrow{w} H(x)$$

for a nondegenerate d.f. H, then H has one of the three forms

1. $H(x) = \begin{cases} \Phi(-a \log |x|), & x < 0, \quad a > 0, \\ 1, & x \ge 0; \end{cases}$

2. $H(x) = \begin{cases} 0, & x \le 0, \quad a > 0, \\ \Phi(a \log x), & x > 0; \end{cases}$

3. $H(x) = \Phi(x), \qquad -\infty < x < \infty.$

This theorem is rather satisfying, though it does not specify the domains of attraction of the three limiting forms. Some results in this direction have been obtained in Chibisov (1964), Smirnov (1967), and Wu (1966). However, these are highly dependent on the rank sequence $\{k_n\}$. For example, a class of rank sequences $\{k_n\}$ such that $k_n \sim l^2 n^\theta$ $(0 < \theta < 1)$ are studied in Chibisov (1964). If F is any d.f., it is known that there is at most one pair of (l, θ) such that F belongs to the domain of attraction of Type 1, and the same statement holds for Type 2. In addition, there are rank sequences such that only the normal law, Type 3, is a possible limit, and moreover, there are distributions attracted to it for every intermediate rank sequence $\{k_n\}$.

As for central ranks, we shall not be concerned with intermediate ranks for the dependent cases considered in the sequel. A reader interested in the behaviour of intermediate ranks for dependent sequences is referred to Watts (1977), (1980), Watts *et al.* (1982), and references in these works.

PART II

EXTREMAL PROPERTIES OF DEPENDENT SEQUENCES

In Chapters 3–6, which comprise Part II, we focus on the effects of dependence among $\xi_1, \ldots, \xi_n$, on the classical extremal results. For the most part it will be assumed that the ξ's, while dependent, are identically distributed, and in fact form a strictly stationary sequence. However, some important non-stationary cases will also be briefly considered.

The task of Chapter 3 is to generalize basic results concerning the maximum M_n, to apply to stationary sequences. As will be seen, the generalization follows in a rather complete way under certain natural restrictions limiting the dependence structure of the sequence. In particular, it is shown that under these restrictions the limit laws in such dependent cases are precisely the same as the classical ones, and indeed, in a given dependent case, that the same limiting law applies as would if the sequence were *independent*, with the same marginal distribution. Some results and examples of non-normal sequences where this is not true are also given.

In Chapter 4 this theory is applied to the case of stationary *normal* sequences. It is found there that the dependence conditions required are satisfied under very weak restrictions on the covariances associated with the sequence.

Chapter 5 is concerned with the topic of Chapter 2—namely, the properties of $M_n^{(k)}$, the kth largest of the ξ_i. The discussion is approached through a consideration of the "point process of exceedances of a level u_n" by the sequence $\xi_1, \xi_2, \ldots$. This provides what we consider to be a helpful and illuminating viewpoint. In particular, a simple convergence theorem shows the Poisson nature of the exceedances of high levels, leading to the desired generalizations of the classical results for $M_n^{(k)}$.

Two topics complementing the theory for normal sequences are dealt with in Chapter 6. In the first the previous extremal results are shown (appropriately modified) to apply to a class of nonstationary normal sequences. This, in particular, provides the asymptotic distribution of the maximum of a stationary normal sequence to which an appropriate trend, or seasonal component, has been added. The second topic concerns stationary normal sequences under very strong dependence. There is no complete theory for this case but a variety of different limiting results are exhibited.

CHAPTER 3
Maxima of Stationary Sequences

In this chapter, we extend the classical extreme value theory of Chapter 1 to apply to a wide class of dependent (stationary) sequences. The stationary sequences involved will be those exhibiting a dependence structure which is not "too strong". Specifically, a distributional type of mixing condition—weaker than the usual forms of dependence restriction such as strong mixing—will be used as a basic assumption in the development of the theory.

3.1. Dependence Restrictions for Stationary Sequences

There are various ways in which the notion of an i.i.d. sequence may be generalized by permitting dependence, or allowing the ξ_n to have different distributions, or both. For example, an obvious generalization is to consider a sequence which is Markov of some order. Though the consideration of Markov sequences can lead to fruitful results for extremes, it is not the direction we shall pursue here.

We shall keep the assumption that the ξ_n have a common distribution; in fact, it will be natural to consider *stationary* sequences, i.e. sequences such that the distributions of $(\xi_{j_1}, \ldots, \xi_{j_n})$ and $(\xi_{j_1+m}, \ldots, \xi_{j_n+m})$ are identical for any choice of $n, j_1, \ldots, j_n$, and m. Then we shall assume that the dependence between ξ_i and ξ_j falls off in some specified way as $|i-j|$ increases. This is different from the Markov property where, in essence, the past, $\{\xi_i; i < n\}$, and the future, $\{\xi_j; j > n\}$, are independent given the present, ξ_n.

The simplest example of the type of restriction we consider is that of m-dependence, which requires that ξ_i and ξ_j be actually *independent* if $|i - j| > m$.

A more commonly used dependence restriction of this type for stationary sequences is that of *strong mixing* (introduced first by Rosenblatt (1956)). Specifically, the sequence $\{\xi_n\}$ is said to satisfy the strong mixing assumption if there is a function $g(k)$, the "mixing function", tending to zero as $k \to \infty$, and such that

$$|P(A \cap B) - P(A)P(B)| < g(k)$$

when $A \in \mathscr{F}(\xi_1, \ldots, \xi_p)$ and $B \in \mathscr{F}(\xi_{p+k+1}, \xi_{p+k+2}, \ldots)$ for any p and k; $\mathscr{F}(\cdot)$ denotes the σ-field generated by the indicated random variables. Thus when a sequence is mixing, any event A "based on the past up to time p" is "nearly independent" of any event B "based on the future from time $p + k + 1$ onwards" when k is large. Note that this mixing condition is uniform in the sense that $g(k)$ does not depend on the particular A and B involved.

The *correlation* between ξ_i and ξ_j is also a (partial) measure of their dependence. Hence another dependence restriction of the same type is $|\operatorname{Corr}(\xi_i, \xi_j)| \le g(|i - j|)$, where $g(k) \to 0$ as $k \to \infty$. Obviously such a restriction will be most useful if the ξ_n form a normal sequence.

Various results from extreme value theory have been extended to apply under each of the more general restrictions mentioned above. For example, Watson (1954) generalized (1.5.2) to apply to m-dependent stationary sequences. Loynes (1965) considered quite a number of results (including (1.5.2) and the Extremal Types Theorem under the strong mixing assumption for stationary sequences. Berman (1964b) used some simple correlation restrictions to obtain (1.5.5) for stationary normal sequences.

It is obvious that the results of Loynes (1965) and Berman (1964b) are related—similar methods being useful in each—but the precise connections are not immediately apparent, due to the different dependence restrictions used. Berman's correlation restrictions are very weak assumptions, leading to sharp results for *normal* sequences. The mixing condition used by Loynes, while being useful in many contexts, is obviously rather restrictive. In this chapter, we shall propose a much weaker condition of "mixing type", which first appeared in Leadbetter (1974), and under which, for example, the results of Loynes (1965) will still be true. Further, this condition will be satisfied for stationary normal sequences under Berman's correlation conditions (as we shall see in the next chapter). Hence the relationships between the various results are clarified.

3.2. Distributional Mixing

In weakening the mixing condition, one notes that the events of interest in extreme value theory are typically those of the form $\{\xi_i \le u\}$ or their intersections. For example, the event $\{M_n \le u\}$ is just $\{\xi_1 \le u, \xi_2 \le u, \ldots, \xi_n \le u\}$. Hence one may be led to propose a condition like mixing but only required to hold for events of this type. For example, one such natural condition would be the following, which we shall call Condition D. For brevity, we will write $F_{i_1 \ldots i_n}(u)$ for $F_{i_1 \ldots i_n}(u, \ldots, u)$ if $F_{i_1 \ldots i_n}(x_1, \ldots, x_n)$ denotes the joint d.f. of $\xi_{i_1}, \ldots, \xi_{i_n}$.

The condition *D will be said to hold if for any integers $i_1 < \cdots < i_p$ and $j_1 < \cdots < j_{p'}$ for which $j_1 - i_p \ge l$, and any real u,*

$$|F_{i_1 \ldots i_p, j_1 \ldots j_{p'}}(u) - F_{i_1 \ldots i_p}(u) F_{j_1 \ldots j_{p'}}(u)| \le g(l), \tag{3.2.1}$$

where $g(l) \to 0$ as $l \to \infty$.

We shall see that the Extremal Types Theorem—and a number of other results—hold under D. However, while D is a significant reduction of the requirements imposed by mixing, it is possible to do better yet. We shall consider a condition, to be called $D(u_n)$, which will involve a requirement like (3.2.1) but applying only to a certain sequence of values $\{u_n\}$ and not necessarily to all u-values. More precisely, if $\{u_n\}$ is a given real sequence, we define the condition $D(u_n)$ as follows.

The condition *$D(u_n)$ will be said to hold if for any integers*

$$1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_{p'} \le n$$

for which $j_1 - i_p \ge l$, we have

$$|F_{i_1 \ldots i_p, j_1 \ldots j_{p'}}(u_n) - F_{i_1 \ldots i_p}(u_n) F_{j_1 \ldots j_{p'}}(u_n)| \le \alpha_{n,l}, \tag{3.2.2}$$

where $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for some sequence $l_n = o(n)$.

The modifications of the condition $D(u_n)$, indicated in the following lemma, are sometimes convenient. In the following [ ] denotes integer part.

Lemma 3.2.1. (i) *The $\alpha_{n,l}$ appearing in $D(u_n)$ may be taken to be nonincreasing in l for each n.*

(ii) *For such $\alpha_{n,l}$ taken nonincreasing in l for each fixed n, the condition $\alpha_{n,l_n} \to 0$ as $n \to \infty$, $l_n = o(n)$, may be rewritten as*

$$\alpha_{n,[n\lambda]} \to 0 \quad \textit{for each } \lambda > 0. \tag{3.2.3}$$

Proof. For (i), we simply note that $\alpha_{n,l}$ may be replaced by the maximum of the left-hand side of (3.2.2) over all allowed choices of i's and j's to obtain a possibly smaller $\alpha_{n,l}$ which is nonincreasing in l for each n and still satisfies $\alpha_{n,l_n} \to 0$ as $n \to \infty$.

For (ii), it is trivially seen that if $\alpha_{n,l_n} \to 0$ for some $l_n = o(n)$, then (3.2.3) holds.

The converse may be shown by noting that (3.2.3) implies the existence of an increasing sequence of constants m_k such that $\alpha_{n,[n/k]} < k^{-1}$ for $n \geq m_k$. If k_n is defined by $k_n = r$ for $m_r \leq n < m_{r+1}$, $r \geq 1$, then $m_{k_n} \leq n$ so that $\alpha_{n,[n/k_n]} \leq k_n^{-1} \to 0$, and the sequence $\{l_n\}$ may be taken to be $\{[n/k_n]\}$. $\square$

Strong mixing implies D, which in turn implies $D(u_n)$ for any sequence $\{u_n\}$. Also, $D(u_n)$ is satisfied, for appropriately chosen $\{u_n\}$, by stationary normal sequences under weak conditions, whereas strong mixing need not be.

The following lemma demonstrates how the condition $D(u_n)$ gives the "degree of independence" appropriate for the discussion of extremes in the subsequent sections. If E is any set of integers, $M(E)$ will denote $\max\{\xi_j; j \in E\}$ (and, of course, $M(E) = M_n$ if $E = \{1, \ldots, n\}$). It will be convenient to let an "interval" mean any finite set E of consecutive integers $\{j_1, \ldots, j_2\}$, say; its *length* will be taken to be $j_2 - j_1 + 1$. If $F = \{k_1, \ldots, k_2\}$ is another interval with $k_1 > j_2$, we shall say that E and F are *separated by* $k_1 - j_2$.

Throughout, $\{\xi_n\}$ will be a stationary sequence.

Lemma 3.2.2. *Suppose $D(u_n)$ holds for some sequence $\{u_n\}$. Let n, r, and k be fixed integers and $E_1, \ldots, E_r$ subintervals of $\{1, \ldots, n\}$ such that E_i and E_j are separated by at least k when $i \neq j$. Then*

$$\left| P\left(\bigcap_{j=1}^{r} \{M(E_j) \leq u_n\}\right) - \prod_{j=1}^{r} P\{M(E_j) \leq u_n\} \right| \leq (r-1)\alpha_{n,k}.$$

Proof. This is easily shown inductively. For brevity, write $A_j = \{M(E_j) \leq u_n\}$. Let $E_j = \{k_j, \ldots, l_j\}$, where (by renumbering if necessary) $k_1 \leq l_1 < k_2 \leq \cdots \leq l_r$. Then

$$\begin{aligned} |P(A_1 \cap A_2) - P(A_1)P(A_2)| &= |F_{k_1 \ldots l_1, k_2 \ldots l_2}(u_n) - F_{k_1 \ldots l_1}(u_n)F_{k_2 \ldots l_2}(u_n)| \\ &\leq \alpha_{n,k} \end{aligned}$$

since $k_2 - l_1 \geq k$. Similarly,

$$\begin{aligned} &|P(A_1 \cap A_2 \cap A_3) - P(A_1)P(A_2)P(A_3)| \\ &\quad \leq |P(A_1 \cap A_2 \cap A_3) - P(A_1 \cap A_2)P(A_3)| + |P(A_1 \cap A_2) \\ &\qquad - P(A_1)P(A_2)|P(A_3) \\ &\quad \leq 2\alpha_{n,k} \end{aligned}$$

since $E_1 \cup E_2 \subseteq \{k_1, \ldots, l_2\}$ and $k_3 - l_2 \geq k$. Proceeding in this way, we obtain the result. $\square$

This lemma thus shows a degree of independence for maxima on separated intervals, which will be basic to the proof of the Extremal Types Theorem for stationary sequences.

3.3. Extremal Types Theorem for Stationary Sequences

Our purpose in this section is to show that the Extremal Types Theorem applies also to stationary sequences under appropriate conditions. That is, if $M_n = \max(\xi_1, \ldots, \xi_n)$ where $\{\xi_i\}$ is stationary, and if

$$P\{a_n(M_n - b_n) \leq x\} \to G(x) \tag{3.3.1}$$

for some constants $\{a_n > 0\}$, $\{b_n\}$ and nondegenerate G, we wish to show (under conditions to be stated) that G is an extreme value distribution, or equivalently (by Theorem 1.4.1), that it is max-stable. By Theorem 1.3.1, this will follow if

$$P\{a_{nk}(M_n - b_{nk}) \leq x\} \xrightarrow{w} G^{1/k}(x) \tag{3.3.2}$$

for each $k = 1, 2, \ldots$. Since the case $k = 1$ in (3.3.2) is just (3.3.1), it is sufficient (as noted after Theorem 1.4.2) to show that if (3.3.2) holds for $k = 1$, then it also holds for $k = 2, 3, \ldots$. This will clearly be the case if, for each $k = 2, 3, \ldots$,

$$P\left\{M_{nk} \leq \frac{x}{a_{nk}} + b_{nk}\right\} - P^k\left\{M_n \leq \frac{x}{a_{nk}} + b_{nk}\right\} \to 0 \tag{3.3.3}$$

as $n \to \infty$. Hence it is sufficient to show that (3.3.3) holds to obtain the desired generalization of the Extremal Types Theorem.

The method used is to divide the interval $\{1, \ldots, n\}$ into k intervals of length $[n/k]$ and use "approximate independence" of the maxima on each via Lemma 3.2.2 to give a result which implies (3.3.3). To apply Lemma 3.2.2, we must shorten each of these intervals to separate them. This leads to the following construction—used, for example, in Loynes (1965) and given here in a slightly more general form for later use also.

Let k be a fixed integer, and for any positive integer n, write $n' = [n/k]$ (the integer part of n/k). Thus we have $n'k \leq n < (n' + 1)k$. Divide the first $n'k$ integers into $2k$ consecutive intervals, as follows. For large n, let m be an integer, $k < m < n'$, and write

$$I_1 = \{1, 2, \ldots, n' - m\}, \; I_1^* = \{n' - m + 1, \ldots, n'\},$$

$I_2, I_2^*, \ldots, I_k, I_k^*$ being defined similarly, alternatively having length $n' - m$ and m. Finally, write

$$I_{k+1} = \{(k - 1)n' + m + 1, \ldots, kn'\}, \qquad I_{k+1}^* = \{kn' + 1, \ldots, kn' + m\}.$$

(Note that I_{k+1}, I_{k+1}^* are defined differently from I_j, I_j^* for $j \leq k$.)

The main steps of the approximation are contained in the following lemma. These are, broadly, to show first that the "small" intervals I_j^* can be essentially disregarded and then to apply Lemma 3.2.2 to the (now separate) intervals $I_1, \ldots, I_k$. In the following, $\{u_n\}$ is any given sequence (not necessarily of the form $x/a_n + b_n$).

Lemma 3.3.1. *With the above notation, and assuming $D(u_n)$ holds,*

(i)
$$0 \le P\left(\bigcap_{j=1}^{k} \{M(I_j) \le u_n\}\right) - P\{M_n \le u_n\}$$
$$\le (k+1)P\{M(I_1) \le u_n < M(I_1^*)\},$$

(ii)
$$\left|P\left(\bigcap_{j=1}^{k} \{M(I_j) \le u_n\}\right) - P^k\{M(I_1) \le u_n\}\right| \le (k-1)\alpha_{n,m},$$

(iii) $|P^k\{M(I_1) \le u_n\} - P^k\{M_{n'} \le u_n\}| \le kP\{M(I_1) \le u_n < M(I_1^*)\}.$

Hence, by combining (i), (ii), *and* (iii),

$$|P\{M_n \le u_n\} - P^k\{M_{n'} \le u_n\}| \le (2k+1)P\{M(I_1) \le u_n < M(I_1^*)\} + (k-1)\alpha_{n,m}. \quad (3.3.4)$$

PROOF. The result (i) follows at once since $\bigcap_{j=1}^{k}\{M(I_j) \le u_n\} \supset \{M_n \le u_n\}$, and their difference implies $M(I_j) \le u_n < M(I_j^*)$ for some $j \le k$, or otherwise $\xi_j \le u_n$ for $1 \le j \le kn'$ but $\xi_j > u_n$ for some $j = kn'+1, \ldots, k(n'+1)$, which in turn implies $M(I_{k+1}) \le u_n < M(I_{k+1}^*)$ (since $m > k$ and hence $k(n'+1) < kn' + m$). Since the probabilities of the events $M(I_j) \le u_n < M(I_j^*)$ are independent of j by stationarity, (i) follows.

The inequality (ii) follows from Lemma 3.2.2 with I_j for E_j, noting that $P\{M(I_j) \le u_n\}$ is independent of j.

To obtain (iii), we note that

$$0 \le P\{M(I_1) \le u_n\} - P\{M_{n'} \le u_n\} = P\{M(I_1) \le u_n < M(I_1^*)\}.$$

The result then follows, writing $y = P\{M(I_1) \le u_n\}$ and $x = P\{M_{n'} \le u_n\}$, from the obvious inequalities

$$0 \le y^k - x^k \le k(y-x) \quad \text{when } 0 \le x \le y \le 1. \qquad \square$$

We now dominate the right-hand side of (3.3.4) to obtain the desired approximation.

Lemma 3.3.2. *If $D(u_n)$ holds, $r \ge 1$ is any fixed integer, and if $n \ge (2r+1)mk$, then with the same notation as in Lemma* 3.3.1,

$$P\{M(I_1) \le u_n < M(I_1^*)\} \le \frac{1}{r} + 2r\alpha_{n,m}. \quad (3.3.5)$$

It then follows from Lemma 3.3.1 *that*

$$P\{M_n \le u_n\} - P^k\{M_{[n/k]} \le u_n\} \to 0 \quad \text{as } n \to \infty. \quad (3.3.6)$$

PROOF. Since $n' \ge (2r+1)m$, we may choose intervals $E_1, \ldots, E_r$, each of length m, from $I_1 = \{1, 2, \ldots, n'-m\}$ $(n' = [n/k])$, so that they are separated from each other and from I_1^* by at least m $(k < m < n'$ again). Then

$$P\{M(I_1) \le u_n < M(I_1^*)\} \le P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}, \{M(I_1^*) > u_n\}\right)$$
$$= P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}\right) - P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}, \{M(I_1^*) \le u_n\}\right).$$

By stationarity, $P\{M(E_s) \le u_n\} = P\{M(I_1^*) \le u_n\} = p$, say, and by Lemma 3.2.2, the two terms on the right differ from p^r and p^{r+1} (in absolute magnitude) by no more than $(r - 1)\alpha_{n,m}$ and $r\alpha_{n,m}$, respectively. Hence

$$P\{M(I_1) \le u_n < M(I_1^*)\} \le p^r - p^{r+1} + 2r\alpha_{n,m},$$

from which (3.3.5) follows since $p^r - p^{r+1} \le 1/(r + 1)$, for $0 \le p \le 1$.

Finally by (3.3.4) and (3.3.5), taking $m = l_n$ according to (3.2.2) ($l_n = o(n)$),

$$\limsup_{n\to\infty} |P\{M_n \le u_n\} - P^k\{M_{n'} \le u_n\}|$$
$$\le \frac{2k+1}{r} + (k - 1 + 2r(2k + 1)) \limsup_{n\to\infty} \alpha_{n,l_n} = \frac{2k+1}{r},$$

from which it follows (by letting $r \to \infty$ on the right) that the left-hand side is zero. Thus (3.3.6) is proved. □

It may be noted that if

$$\limsup_{n\to\infty} n(1 - F(u_n)) < \infty$$

(which, e.g. is the case if (1.5.1) holds with $\tau < \infty$), then (3.3.6) is an even more direct consequence of Lemma 3.3.1. In fact, taking $m = l_n = o(n)$, we then have

$$\begin{aligned} P\{M(I_1) \le u_n < M(I_1^*)\} &\le l_n P\{\xi_1 > u_n\} \\ &= \frac{l_n}{n} n(1 - F(u_n)) \\ &\to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Thus (3.3.6) follows at once from $D(u_n)$ and (3.3.4).

The Extremal Types Theorem now follows easily under general conditions.

Theorem 3.3.3. *Let $\{\xi_n\}$ be a stationary sequence and $a_n > 0$ and b_n given constants such that $P\{a_n(M_n - b_n) \le x\}$ converges to a nondegenerate d.f. $G(x)$. Suppose that $D(u_n)$ is satisfied for $u_n = x/a_n + b_n$ for each real x. Then $G(x)$ has one of the three extreme value forms listed in Theorem* 1.4.1.

Proof. Writing $u_n = x/a_n + b_n$ and using (3.3.6) (with nk in place of n), we obtain (3.3.3). As remarked in connection with (3.3.3), (3.3.2) holds for all k since it holds by assumption for $k = 1$. Hence if F_n is the d.f. of M_n, by Theorem 1.3.1 G is max-stable and thus an extreme value type by Theorem 1.4.1. □

Corollary 3.3.4. *The result remains true if the condition that $D(u_n)$ be satisfied for each $u_n = x/a_n + b_n$ is replaced by the requirement that Condition D holds. (For then $D(u_n)$ is satisfied by any sequence, in particular by $u_n = x/a_n + b_n$ for each x.)*

It is intuitively plausible that the same criteria for domains of attraction should apply to the marginal d.f. F, as in the classical i.i.d. case. We shall show that this is, in fact, true (under a further assumption at least) as a simple consequence of the results of the next section. It may be noted that for extremes of continuous parameter processes (dealt with in Chapter 13), such an assertion does not hold exactly as for sequences, but that the application of the criteria must be modified in a simple (and rather interesting) way.

3.4. Convergence of $P\{M_n \le u_n\}$ Under Dependence

The results so far have been concerned with the *possible* forms of limiting extreme value distributions. We now turn to the *existence* of such a limit, in that we formulate conditions under which (1.5.1) and (1.5.2) are equivalent for stationary sequences, i.e. conditions under which

$$P\{M_n \le u_n\} \to e^{-\tau} \tag{3.4.1}$$

is equivalent to

$$n(1 - F(u_n)) = nP\{\xi_1 > u_n\} \to \tau \quad \text{as } n \to \infty. \tag{3.4.2}$$

As noted above, it will follow directly from these results (as shown explicitly in the next section) that the classical (i.i.d.) criteria for domains of attraction may be applied for such appropriate dependent sequences.

It may be seen from the derivation below that if (3.4.2) holds, then Condition $D(u_n)$ is sufficient to guarantee that $\liminf P\{M_n \le u_n\} \ge e^{-\tau}$. However, we need a further assumption to obtain the opposite inequality for the upper limit. Various forms of such an assumption may be used. Here we content ourselves with the following simple variant of conditions used in Watson (1954) and Loynes (1965); we refer to this as $D'(u_n)$.

The condition $D'(u_n)$ *will be said to hold for the stationary sequence* $\{\xi_j\}$ *and sequence* $\{u_n\}$ *of constants if*

$$\limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P\{\xi_1 > u_n, \xi_j > u_n\} \to 0 \quad \textit{as } k \to \infty \tag{3.4.3}$$

(*where* [ ] *denotes the integer part*).

Note that under (3.4.2), the level u_n in (3.4.3) is such that there are on the average approximately τ exceedances of u_n among $\xi_1, \ldots, \xi_n$, and thus τ/k among $\xi_1, \ldots, \xi_{[n/k]}$. The condition $D'(u_n)$ bounds the probability of more than one exceedance among $\xi_1, \ldots, \xi_{[n/k]}$. This will eventually ensure that there are no multiple points in the point process of exceedances which, of course, is necessary in obtaining (as we shall later) a simple Poisson limit for this point process.

Our main result generalizes Theorem 1.5.1 to apply to stationary sequences under $D(u_n)$, $D'(u_n)$. A form of the "only if" part of this theorem was first proved by R. Davis (1979).

Theorem 3.4.1. *Let $\{u_n\}$ be constants such that $D(u_n)$, $D'(u_n)$ hold for the stationary sequence $\{\xi_n\}$. Let $0 \le \tau < \infty$. Then* (3.4.1) *and* (3.4.2) *are equivalent, i.e.* $P\{M_n \le u_n\} \to e^{-\tau}$ *if and only if* $n(1 - F(u_n)) \to \tau$.

PROOF. Fix k, and for each n, write n' $(=n'_{n,k}) = [n/k]$. Since

$$\{M_{n'} > u_n\} = \bigcup_{j=1}^{n'} \{\xi_j > u_n\},$$

we have

$$\sum_{j=1}^{n'} P\{\xi_j > u_n\} - \sum_{1 \le i < j \le n'} P\{\xi_i > u_n, \xi_j > u_n\} \le P\{M_{n'} > u_n\}$$
$$\le \sum_{j=1}^{n'} P\{\xi_j > u_n\}.$$

Using stationarity, it follows simply that

$$1 - n'(1 - F(u_n)) \le P\{M_{n'} \le u_n\} \le 1 - n'(1 - F(u_n)) + S_n, \quad (3.4.4)$$

where $S_n = S_{n,k} = n' \sum_{j=2}^{n'} P\{\xi_1 > u_n, \xi_j > u_n\}$. Since $n' = [n/k]$, condition $D'(u_n)$ gives $\limsup_{n\to\infty} S_n = k^{-1}o(1) = o(k^{-1})$ as $k \to \infty$.

Suppose now that (3.4.2) holds. Then $n'(1 - F(u_n)) \to \tau/k$ so that $n \to \infty$ in (3.4.4) gives

$$1 - \frac{\tau}{k} \le \liminf_{n\to\infty} P\{M_{n'} \le u_n\} \le \limsup_{n\to\infty} P\{M_{n'} \le u_n\} \le 1 - \frac{\tau}{k} + o\left(\frac{1}{k}\right).$$

By taking the kth power of each term and using (3.3.6), we have

$$\left(1 - \frac{\tau}{k}\right)^k \le \liminf_{n\to\infty} P\{M_n \le u_n\} \le \limsup_{n\to\infty} P\{M_n \le u_n\} \le \left(1 - \frac{\tau}{k} + o\left(\frac{1}{k}\right)\right)^k.$$

Letting $k \to \infty$, we see that $\lim_{n\to\infty} P\{M_n \le u_n\}$ exists and equals $e^{-\tau}$, as required to show (3.4.1).

Conversely, if (3.4.1) holds, i.e. $P\{M_n \le u_n\} \to e^{-\tau}$ as $n \to \infty$, we have from (3.4.4) (again with $n' = [n/k]$),

$$1 - P\{M_{n'} \le u_n\} \le n'(1 - F(u_n)) \le 1 - P\{M_{n'} \le u_n\} + S_n. \quad (3.4.5)$$

But since $P\{M_n \le u_n\} \to e^{-\tau}$, (3.3.6) shows that $P\{M_{n'} \le u_n\} \to e^{-\tau/k}$, so that letting $n \to \infty$ in (3.4.5), we obtain, since $n' \sim n/k$,

$$1 - e^{-\tau/k} \le \frac{1}{k} \liminf_{n\to\infty} n(1 - F(u_n)) \le \frac{1}{k} \limsup_{n\to\infty} n(1 - F(u_n))$$
$$\le 1 - e^{-\tau/k} + o\left(\frac{1}{k}\right),$$

from which (multiplying by k and letting $k \to \infty$) we see that $n(1 - F(u_n)) \to \tau$ so that (3.4.2) holds. □

If $n(1 - F(u_n)) \to \infty$, the condition $D'(u_n)$ is not satisfied even for i.i.d. sequences, as is easily seen. However, the result also applies when $\tau = \infty$ if we modify the $D(u_n)$, $D'(u_n)$ conditions in a natural manner for such sequences. We show this in the following corollary.

Corollary 3.4.2. *The same conclusions hold with $\tau = \infty$ (i.e. $P\{M_n \le u_n\} \to 0$ if and only if $n(1 - F(u_n)) \to \infty$) if the requirements that $D(u_n)$, $D'(u_n)$ hold are replaced by the condition that, for arbitrarily large τ ($< \infty$), there exists a sequence $\{v_n\}$ such that $n(1 - F(v_n)) \to \tau$ and such that $D(v_n)$, $D'(v_n)$ hold.*

PROOF. Fix $\tau < \infty$. If $n(1 - F(u_n)) \to \infty$, then clearly $u_n \le v_n$ for sufficiently large n so that

$$\limsup_{n\to\infty} P\{M_n \le u_n\} \le \lim_{n\to\infty} P\{M_n \le v_n\} = e^{-\tau}$$

by the theorem. Since this holds for arbitrarily large τ, by letting $\tau \to \infty$ we see that $P\{M_n \le u_n\} \to 0$.

Conversely, if $P\{M_n \le u_n\} \to 0$, fix $\tau > 0$ and let v_n be chosen (satisfying $D(v_n)$, $D'(v_n)$) such that $n(1 - F(v_n)) \to \tau$. Then $P\{M_n \le v_n\} \to e^{-\tau} > 0$ so that clearly $v_n \ge u_n$ for sufficiently large n, giving $n(1 - F(u_n)) \ge n(1 - F(v_n)) \to \tau$. Since this holds for arbitrarily large $\tau > 0$, we must have $n(1 - F(u_n)) \to \infty$, as desired. □

Note that if $D(u_n)$, $D'(u_n)$ hold and it is assumed that (3.4.2) holds just for some subsequence $\{n_j\}$ of integers, i.e. $n_j(1 - F(u_{n_j})) \to \tau$ as $j \to \infty$, then the same proof shows that (3.4.1) holds for that subsequence, i.e. $P\{M_{n_j} \le u_{n_j}\} \to e^{-\tau}$. This observation may be used to give an alternative proof of the statement that (3.4.1) implies (3.4.2) in this theorem by assuming the existence of some subsequence n_j for which $n_j(1 - F(u_{n_j})) \to \tau' \ne \tau$, $0 \le \tau' \le \infty$.

3.5. Associated Independent Sequences and Domains of Attraction

In order to discuss the domains of attraction for dependent sequences in the Extremal Types Theorem, it is useful to introduce an i.i.d. sequence $\{\hat{\xi}_n\}$ having the same common d.f. F as each member of the stationary sequence $\{\xi_n\}$. The sequence $\{\hat{\xi}_n\}$ will be termed (following Loynes (1965)) the "independent sequence associated with $\{\xi_n\}$", and we write $\hat{M}_n = \max(\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_n)$. The following result is then a direct corollary of Theorem 3.4.1.

Theorem 3.5.1. *Let $D(u_n)$, $D'(u_n)$ be satisfied for the stationary sequence $\{\xi_n\}$. Then $P\{M_n \le u_n\} \to \theta > 0$ if and only if $P\{\hat{M}_n \le u_n\} \to \theta$. The same holds with $\theta = 0$ if the conditions $D(u_n)$, $D'(u_n)$ are replaced by the requirement that for arbitrarily large $\tau < \infty$ there exists $\{v_n\}$ satisfying $n(1 - F(v_n)) \to \tau$ and such that $D(v_n)$, $D'(v_n)$ hold.*

PROOF. The condition $P\{\hat{M}_n \le u_n\} \to \theta$ may be rewritten as $P\{\hat{M}_n \le u_n\} \to e^{-\tau}$ with $\tau = -\log \theta$, and by Theorem 1.5.1, holds if and only if $1 - F(u_n) \sim \tau/n$. By Theorem 3.4.1, the same is true for $P\{M_n \le u_n\}$, so that the result follows when $\theta > 0$. When $\theta = 0$ the result follows similarly using Corollary 3.4.2. □

We may also deduce at once that the limiting distribution of $a_n(M_n - b_n)$ is the same as that which would apply if the ξ_n were i.i.d., i.e. it is the same as that of $a_n(\hat{M}_n - b_n)$ under conditions $D(u_n)$ and $D'(u_n)$. Part of this result was proved in Loynes (1965) under conditions which include strong mixing.

Theorem 3.5.2. *Suppose that $D(u_n)$, $D'(u_n)$ are satisfied for the stationary sequence $\{\xi_n\}$, when $u_n = x/a_n + b_n$ for each x ($\{a_n > 0\}$, $\{b_n\}$ being given sequences of constants). Then $P\{a_n(M_n - b_n) \le x\} \to G(x)$ for some non-degenerate G if and only if $P\{a_n(\hat{M}_n - b_n) \le x\} \to G(x)$.*

PROOF. If $G(x) > 0$, the equivalence follows from Theorem 3.5.1, with $\theta = G(x)$.

In the case where $\theta = 0$, the continuity of G (necessarily an extreme value distribution) shows that, if $0 < \tau < \infty$, there exists x_0 such that $G(x_0) = e^{-\tau}$. $D(v_n)$, $D'(v_n)$ hold for $v_n = x_0/a_n + b_n$ and $P\{M_n \le v_n\} \to e^{-\tau}$ or $P\{\hat{M}_n \le v_n\} \to e^{-\tau}$ depending on the assumption made, so that Theorem 3.4.1 or Theorem 1.5.1 shows that $n(1 - F(v_n)) \to \tau$. Thus the case $\theta = 0$ follows from the second statement of Theorem 3.5.1, on writing $u_n = x/a_n + b_n$. □

Note that the case $\theta = 0$ may alternatively be obtained from that for $\theta > 0$ by using the continuity of G at its left-hand endpoint, where this is finite, in an obvious way. However, the above proof using the previous results for $\tau = \infty$, seems natural and instructive.

In view of this result, the same criteria may be used to determine the domains of attraction (under D and D' conditions), as in the classical i.i.d. case. Further, the same constants may be used in the normalization, as if the sequence were i.i.d. This will be illustrated for normal sequences in the next chapter, where we verify the D and D' conditions.

3.6. Maxima Over Arbitrary Intervals

It will be necessary in the development of the Poisson theory of exceedances (in Chapter 5) to use asymptotic results for the maximum of ξ_i for i belonging to "intervals" (i.e. sets of consecutive integers) whose lengths are asymptotically proportional to n. First we give two lemmas which are useful here and elsewhere, showing how a sequence $\{u_n\}$ may be replaced by an "appropriately close" sequence $\{v_n\}$ in considering $P\{M_n \le u_n\}$, and $D(u_n)$. The first of these results gives the replacement results, and the second provides a specific useful case of such replacement.

Lemma 3.6.1. *Let $\{\xi_n\}$ be a stationary sequence and $\{u_n\}$, $\{v_n\}$ real sequences such that $n(F(u_n) - F(v_n)) \to 0$ as $n \to \infty$ (which holds, in particular, if $n(1 - F(u_n)) \to \tau$, $n(1 - F(v_n)) \to \tau$ for some τ, $0 \le \tau < \infty$). Then*

(i) *if I_n is an interval containing v_n integers, where $v_n = 0(n)$, then $P\{M(I_n) \le u_n\} - P\{M(I_n) \le v_n\} \to 0$ so that for $0 \le \rho \le 1$, $P\{M(I_n) \le u_n\} \to \rho$ if and only if $P\{M(I_n) \le v_n\} \to \rho$,*
(ii) *$D(u_n)$ holds if and only if $D(v_n)$ holds.*

PROOF. Let K be a constant, $m \le Kn, k_1, \ldots, k_m$ distinct integers in $(1, 2, \ldots, Kn)$. Then with the standard notation,

$$|F_{k_1 \ldots k_m}(u_n) - F_{k_1 \ldots k_m}(v_n)| = \left| P\left(\bigcap_{i=1}^{m}\{\xi_{k_i} \le u_n\}\right) - P\left(\bigcap_{i=1}^{m}\{\xi_{k_i} \le v_n\}\right)\right|.$$

If $u_n \ge v_n$ the right-hand side is

$$P\left(\bigcap_{i=1}^{m}\{\xi_{k_i} \le u_n\}\right) - P\left(\bigcap_{i=1}^{m}\{\xi_{k_i} \le v_n\}\right) \le P\left(\bigcup_{i=1}^{m}\{v_n < \xi_{k_i} \le u_n\}\right)$$
$$\le Kn(F(u_n) - F(v_n)).$$

This and the corresponding obvious modification when $u_n < v_n$ yield

$$|F_{k_1 \ldots k_m}(u_n) - F_{k_1 \ldots k_m}(v_n)| \le Kn|F(u_n) - F(v_n)| \to 0 \quad \text{as } n \to \infty. \tag{3.6.1}$$

In the above calculations, m and the k_i may depend on n. In particular by taking $m = v_n$ $(\le Kn)$ and $k_i = i$, $1 \le i \le v_n$, it follows that $P\{M_{v_n} \le u_n\} - P\{M_{v_n} \le v_n\} \to 0$, from which (i) also follows by stationarity.

To prove (ii) suppose that $D(u_n)$ holds and let $1 \le i_1 \le \cdots \le i_p \le j_1 \le \cdots \le j_{p'} \le n$, with $j_1 - i_p \ge l$. Then, with an obvious compression of notation (writing $\mathbf{i} = (i_1, \ldots, i_p)$, $\mathbf{j} = (j_1, \ldots, j_{p'})$),

$$|F_{\mathbf{ij}}(u_n) - F_{\mathbf{i}}(u_n)F_{\mathbf{j}}(u_n)| \le \alpha_{n,l},$$

where $\alpha_{n,l_n} \to 0$ for some $l_n = o(n)$. Now

$$\begin{aligned} |F_{\mathbf{ij}}(v_n) - F_{\mathbf{i}}(v_n)F_{\mathbf{j}}(v_n)| \le{} & |F_{\mathbf{ij}}(v_n) - F_{\mathbf{ij}}(u_n)| + |F_{\mathbf{ij}}(u_n) - F_{\mathbf{i}}(u_n)F_{\mathbf{j}}(u_n)| \\ & + F_{\mathbf{i}}(u_n)|F_{\mathbf{j}}(u_n) - F_{\mathbf{j}}(v_n)| \\ & + F_{\mathbf{j}}(v_n)|F_{\mathbf{i}}(u_n) - F_{\mathbf{i}}(v_n)|. \end{aligned}$$

By taking $K = 1$, and applying (3.6.1) three times, successively identifying $(k_1, \ldots, k_m)$ with $(i_1, \ldots, i_p)$, $(j_1, \ldots, j_{p'})$, and $(i_1, \ldots, i_p, j_1, \ldots, j_{p'})$,

$$|F_{\mathbf{ij}}(v_n) - F_{\mathbf{i}}(v_n)F_{\mathbf{j}}(v_n)| \le \alpha^*_{n,l},$$

where $\alpha^*_{n,l} = \alpha_{n,l} + \lambda_n$, and $\lambda_n \to 0$ as $n \to \infty$ Hence $\alpha^*_{n,l_n} \to 0$ as $n \to \infty$, so that $D(v_n)$ follows. □

The next result shows that if u_n is defined to satisfy (3.4.2) for some τ, then v_n may be naturally defined so that $n(1 - F(v_n)) \to \tau'$ for another given τ', and moreover if $D(u_n)$ (or $D'(u_n)$) holds, then so does $D(v_n)$ (or $D'(v_n)$) when $\tau' < \tau$.

Lemma 3.6.2. *Suppose that $\{u_n\}$ satisfies* (3.4.2) *for a fixed $\tau > 0$, i.e. $n(1 - F(u_n)) \to \tau$, and for a fixed $\theta > 0$ define*

$$v_n = u_{[n/\theta]}. \tag{3.6.2}$$

Then

(i) *$\{v_n\}$ satisfies*

$$n(1 - F(v_n)) \to \theta\tau, \tag{3.6.3}$$

(ii) *if $\theta < 1$ and $D(u_n)$ holds, so does $D(v_n)$,*
(iii) *if $\theta < 1$ and $D'(u_n)$ holds, so does $D'(v_n)$,*
(iv) *if $\{w_n\}$ is a sequence satisfying $n(1 - F(w_n)) \to \tau' < \tau$ and $D(u_n)$ holds, so does $D(w_n)$.*

PROOF. (i) By (3.4.2) $n(1 - F(v_n)) = n(1 - F(u_{[n/\theta]})) \sim n\tau/[n/\theta] \to \theta\tau$ as required.

(ii) If $D(u_n)$ holds and $1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_{p'} \le n, j_1 - i_p \ge l$,

$$|F_{\mathbf{ij}}(v_n) - F_{\mathbf{i}}(v_n)F_{\mathbf{j}}(v_n)| = |F_{\mathbf{ij}}(u_{[n/\theta]}) - F_{\mathbf{i}}(u_{[n/\theta]})F_{\mathbf{j}}(u_{[n/\theta]})|$$

which does not exceed $\alpha^*_{n,l} = \alpha_{[n/\theta],l}$, since $j_{p'} \le n \le [n/\theta]$. If $\alpha_{n,l_n} \to 0$ then $\alpha^*_{n,l^*_n} \to 0$ with $l^*_n = l_{[n/\theta]} = o(n)$, so that (ii) follows.

Part (iii) also follows simply, since for $\theta \le 1$,

$$\begin{aligned} n\sum_{j=2}^{[n/k]} P\{\xi_1 > v_n, \xi_j > v_n\} &= n\sum_{j=2}^{[n/k]} P\{\xi_1 > u_{[n/\theta]}, \xi_j > u_{[n/\theta]}\} \\ &\le [n/\theta] \sum_{j=2}^{[[n/\theta]/k]} P\{\xi_1 > u_{[n/\theta]}, \xi_j > u_{[n/\theta]}\}. \end{aligned}$$

By $D'(u_n)$ the upper limit of this expression over n, or $[n/\theta]$, tends to zero as $k \to \infty$, so that (iii) follows.

(iv) Write $\tau' = \theta\tau$ so that $\theta < 1$, and by (ii) $D(v_n)$ holds. Then $n(1 - F(v_n)) \to \theta\tau$ and $n(1 - F(w_n)) \to \theta\tau$ so that $D(w_n)$ holds by Lemma 3.6.1(ii). □

Our first main result is an easy corollary of Theorem 3.4.1

Theorem 3.6.3. *Let $\{\xi_n\}$ be a stationary sequence and $\{u_n\}$, $\{v_n\}$ real sequences such that $n(1 - F(u_n)) \to \tau$, $n(1 - F(v_n)) \to \theta\tau$ as $n \to \infty$, where $\tau > 0, \theta > 0$ are fixed constants. Suppose that $D(v_n)$, $D'(v_n)$ hold. Then if $\{I_n\}$ is a sequence of intervals with ν_n members, and $\nu_n \sim \theta n$, we have*

$$P\{M(I_n) \le u_n\} \to e^{-\theta\tau} \quad \textit{as } n \to \infty. \tag{3.6.4}$$

PROOF. By stationarity it is sufficient to show that $P\{M_{\nu_n} \le u_n\} \to e^{-\theta\tau}$. Now it follows from Theorem 3.4.1 that $P\{M_n \le v_n\} \to e^{-\theta\tau}$ and hence it is only necessary to show that

$$P\{M_{\nu_n} \le u_n\} - P\{M_{\nu_n} \le v_{\nu_n}\} \to 0. \tag{3.6.5}$$

This follows simply by the same argument as in Lemma 3.6.1(i), or alternatively from that result by writing $w_n = u_{[n/\theta]}$, and noting that $n(1 - F(w_n)) \to \theta\tau$ so that Lemma 3.6.1(i) gives

$$P\{M_n \le w_n\} - P\{M_n \le v_n\} \to 0 \quad \text{as } n \to \infty.$$

Replacing n by ν_n and using the fact that $w_{\nu_n} = u_{[\nu_n/\theta]}$ and Lemma 3.6.1(i) again gives (3.6.5) and hence the desired result. □

Of course, if $\{u_n\}$ is given, and $n(1 - F(u_n)) \to \tau$, $\{v_n\}$ may be chosen as $v_n = u_{[n/\theta]}$ and will satisfy $n(1 - F(v_n)) \to \theta\tau$ by Lemma 3.6.2(i). This leads to the following simpler sufficient conditions for (3.6.4) when $\theta < 1$.

Corollary 3.6.4. *Let $\{\xi_n\}$ be a stationary sequence and $\{u_n\}$ a real sequence such that $n(1 - F(u_n)) \to \tau$, a fixed constant, and such that $D(u_n)$, $D'(u_n)$ hold. Then if $0 < \theta < 1$, (3.6.4) holds for a sequence $\{I_n\}$ of intervals with $\nu_n \sim \theta n$ members.*

PROOF. As noted above, define $v_n = u_{[n/\theta]}$. By Lemma 3.6.2, v_n satisfies the conditions required in Theorem 3.6.3 so that the result follows. □

The case $\tau = \infty$ which was not included in the above theorem, is also simply dealt with as follows.

Theorem 3.6.5. *Suppose that $n(1 - F(u_n)) \to \infty$, and for arbitrarily large τ, $0 < \tau < \infty$, there exists a sequence $\{v_n\}$ such that $n(1 - F(v_n)) \to \tau$ and such that $D(v_n)$, $D'(v_n)$ hold. Then for $\theta > 0$, $\nu_n \sim \theta n$ and I_n as above, having ν_n members,*

$$P\{M(I_n) \le u_n\} \to 0 \quad \textit{as } n \to \infty.$$

PROOF. Again by stationarity, it is sufficient to show that $P\{M_{\nu_n} \le u_n\} \to 0$. But for arbitrarily large $\tau > 0$, there is a sequence $\{v_n\}$ such that $n(1 - F(v_n)) \to \tau$, and hence since $\nu_n \sim \theta n$, $n(1 - F(v_{\nu_n})) \to \tau/\theta$, which implies that $u_n < v_{\nu_n}$ for sufficiently large n. Hence

$$P\{M_{\nu_n} \le u_n\} \le P\{M_{\nu_n} \le v_{\nu_n}\} \to e^{-\tau}$$

by Theorem 3.4.1, and letting $\tau \to \infty$ shows that $P\{M_{\nu_n} \le u_n\} \to 0$, as desired. □

The conditions of Theorem 3.6.5 are in fact more natural than might appear at first sight. If $\tau = \infty$, i.e. if $n(1 - F(u_n)) \to \infty$, then $D'(u_n)$ may not hold even for i.i.d. random variables. Hence some modification of at least the $D'(u_n)$ condition is desirable in that case. But for finite τ, Lemma 3.6.2 shows that if $n(1 - F(u_n)) \to \tau$ and $D(u_n)$, $D'(u_n)$ hold, then for any $\tau' < \tau$ there is a sequence $\{v_n\}$ such that $n(1 - F(v_n)) \to \tau'$ and $D(v_n)$, $D'(v_n)$ hold. It is the translation of this property into the case where $\tau = \infty$, that constitutes the conditions of Theorem 3.6.5.

From these results, we may see simply that, under D and D' conditions, the asymptotic distribution of $M(I_n)$ is of the same type as that of M_n, when I_n has $v_n \sim \theta n$ members.

Theorem 3.6.6. *Let $\{\xi_n\}$ be a stationary sequence, let $a_n > 0$ and b_n be constants, and suppose that*

$$P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x) \quad \text{as } n \to \infty$$

for a nondegenerate d.f. G. Suppose $D(u_n)$, $D'(u_n)$ hold for all u_n of the form $x/a_n + b_n$, and let I_n be an interval containing $v_n \sim \theta n$ integers for some $\theta > 0$. Then

$$P\{a_n(M(I_n) - b_n) \le x\} \to G^\theta(x) \quad \text{as } n \to \infty.$$

Proof. Consider a point x with $G(x) > 0$ and write $\tau = -\log G(x)$, $u_n = x/a_n + b_n$. Then $n(1 - F(u_n)) \to \tau$, by Theorem 3.4.1. Now since G is continuous (being an extreme value d.f.), there exists y such that $G(y) = e^{-\theta\tau}$. Write $v_n = y/a_n + b_n$. By assumption $D(v_n)$, $D'(v_n)$ hold, and since $P\{M_n \le v_n\} \to G(y) = e^{-\theta\tau}$, Theorem 3.4.1 shows that $n(1 - F(v_n)) \to \theta\tau$. Thus the conditions of Theorem 3.6.3 are satisfied, so that $P\{M(I_n) \le u_n\} \to e^{-\theta\tau} = G^\theta(x)$, which at once yields the desired conclusion, when $G(x) > 0$.

The case $G(x) = 0$ follows from the continuity of G and the fact that $P\{a_n(M(I_n) - b_n) \le x\} \le P\{a_n(M(I_n) - b_n) \le y\}$ for any $y > x$ with $G(y) > 0$. □

Finally, note that since G is an extreme value d.f., it is max-stable, and Corollary 1.3.2 shows that G^θ is of the same type as G, i.e. $G^\theta(x) = G(ax + b)$ for some $a > 0$, b. This of course gives the not unexpected result that the limit for $M(I_n)$ is of the same type as that for M_n.

3.7. On the Roles of the Conditions $D(u_n)$, $D'(u_n)$

It will be evident from the preceding sections that the conditions $D(u_n)$, $D'(u_n)$ together imply the central distributional results of extreme value theory, for stationary sequences. As mentioned already it will be shown in Chapter 5 that these conditions ensure that the exceedances of the "level" u_n by $\xi_1, \ldots, \xi_n$ take on the character of a Poisson process when n is large—leading to even further asymptotic distributional results, for kth largest values as well as for the maximum. In this, the condition $D(u_n)$ provides the independence associated with the occurrence of events in a Poisson process and, as already noted, $D'(u_n)$ limits the possibility of clustering of exceedances so that multiple events are excluded in the limit.

Apart from these intuitive comments, it is, of course, of interest to see specifically what kinds of behaviour can occur when the conditions are relaxed in some way. Trivial examples (e.g. taking all ξ_n to be the same) show that the total omission of dependence restriction leads to quite arbitrary asymptotic distributions for the maximum. More interesting examples (cf. Chapter 6) exhibit nontrivial cases in which an asymptotic distribution for the maximum exists but is not of extreme value type, when the "dependence decay" is so slow that $D(u_n)$ does not hold.

Of course, if $D(u_n)$ holds for the appropriate sequences $\{u_n\}$, the Extremal Types Theorem shows that any nondegenerate limiting distribution for the maximum of the stationary sequence $\{\xi_n\}$, must be one of the extreme value types. Still more interesting questions concern the role of $D'(u_n)$, and the extent to which the previous results can be modified to apply when only $D(u_n)$ is assumed. We shall in fact find the intriguing result that in many (perhaps most) cases of interest, the existence of a limiting distribution for the maximum $\hat{M}_n$ in the associated independent sequence, implies a limiting distribution of the same type for M_n itself—indeed with the same norming constants $\{a_n\}$, $\{b_n\}$. (Or the limits may be taken to be identical by an obvious, simple change of one set of the norming constants.)

Examples will be quoted from the literature (and an interesting class of such examples discussed in more detail in the next section) to illustrate the possible range of behaviour when $D'(u_n)$ is not assumed. However, we first give some general results, modifying those of the previous sections.

It has been noted already in Section 3.4 that (3.4.2) and $D(u_n)$ alone are sufficient to guarantee that $\liminf P\{M_n \le u_n\} \ge e^{-\tau}$ even though the full limit (3.4.1) will not hold in general without $D'(u_n)$. This fact will be useful to us in proving a modified form of Theorem 3.4.1. In this result it will be shown that if, for each $\tau > 0$, $\{u_n(\tau)\}$ is a sequence satisfying (3.4.2), and such that $D(u_n)$ holds when $u_n = u_n(\tau)$ and if $P\{M_n \le u_n(\tau)\}$ converges for at least one $\tau > 0$, then $P\{M_n \le u_n(\tau)\} \to e^{-\theta\tau}$ for all $\tau > 0$, for some fixed θ with $0 \le \theta \le 1$. This result was proved under the further assumption that $\lim P\{M_n \le u_n(\tau)\}$ exists for all $\tau > 0$, by Loynes (1965) for strongly mixing sequences, and by Chernick (1981a) under $D(u_n)$. The proof here is along similar lines to those of Loynes, and of Chernick.

Theorem 3.7.1. *Suppose $u_n(\tau)$ is defined for $\tau > 0$ and is such that $n(1 - F(u_n(\tau))) \to \tau$, and that $D(u_n(\tau))$ holds for each such τ. Then there exist constants θ, θ', $0 \le \theta \le \theta' \le 1$ such that $\limsup_{n\to\infty} P\{M_n \le u_n(\tau)\} = e^{-\theta\tau}$, $\liminf_{n\to\infty} P\{M_n \le u_n(\tau)\} = e^{-\theta'\tau}$. Hence if $P\{M_n \le u_n(\tau)\}$ converges for some $\tau > 0$, then $\theta = \theta'$ and $P\{M_n \le u_n(\tau)\} \to e^{-\theta\tau}$ for all $\tau > 0$.*

Proof. It follows from Lemma 3.3.2 that for a fixed integer k,

$$P\{M_n \le u_n(\tau)\} - P^k\{M_{n'} \le u_n(\tau)\} \to 0 \quad \text{as } n \to \infty,$$

where $n' = [n/k]$. Hence if $\limsup_{n\to\infty} P\{M_n \le u_n(\tau)\} = \psi(\tau)$, it follows that

$$\limsup_{n\to\infty} P\{M_{n'} \le u_n(\tau)\} = \psi^{1/k}(\tau). \tag{3.7.1}$$

Now it may be simply seen by considering cases where $u_n(\tau) \ge u_{n'}(\tau/k)$, $u_n(\tau) < u_{n'}(\tau/k)$ (cf. proof of Lemma 3.6.1(i)) that

$$\begin{aligned}
\left| P\{M_{n'} \le u_n(\tau)\} - P\left\{M_{n'} \le u_{n'}\left(\frac{\tau}{k}\right)\right\}\right| &\le n'\left| F(u_n(\tau)) - F\left(u_{n'}\left(\frac{\tau}{k}\right)\right)\right| \\
&= n'\left|\left(1 - F\left(u_{n'}\left(\frac{\tau}{k}\right)\right)\right) - (1 - F(u_n(\tau)))\right| \\
&= n'\left|\frac{\tau/k}{n'}(1 + o(1)) - \frac{\tau}{n}(1 + o(1))\right| \\
&= o(1),
\end{aligned}$$

since $n' \sim n/k$. Since clearly $\limsup_{n\to\infty} P\{M_{n'} \le u_{n'}(\tau/k)\} = \psi(\tau/k)$ it follows that $\limsup_{n\to\infty} P\{M_{n'} \le u_n(\tau)\} = \psi(\tau/k)$ which with (3.7.1) shows that $\psi(\tau/k) = \psi^{1/k}(\tau)$ for all $\tau > 0$, $k = 1, 2 \ldots$. Now, if $\tau' < \tau$ it is clear that $u_n(\tau') > u_n(\tau)$ when n is sufficiently large, so that $\psi(\tau)$ is nonincreasing, and is strictly positive since as observed before the theorem, $\liminf_{n\to\infty} P\{M_n \le u_n(\tau)\} \ge e^{-\tau}$. But it is well known that the only such solution to the functional equation is $\psi(\tau) = e^{-\theta\tau}$ where $\theta \ge 0$. Since as above $\psi(\tau) \ge e^{-\tau}$ it follows that $0 \le \theta \le 1$.

It follows similarly that $\liminf_{n\to\infty} P\{M_n \le u_n(\tau)\} \to e^{-\theta'\tau}$ where $0 \le \theta' \le 1$, and clearly $\theta' \ge \theta$, completing the proof of the theorem, since the final statement is obvious. □

Later we shall cite examples which show that every value of θ, $0 \le \theta \le 1$, may occur in this theorem. The case $\theta = 0$ is "degenerate" in that it leads to $P\{M_n \le u_n(\tau)\} \to 1$ for each τ, but its existence does have some, at least marginal, bearing on the range of limiting types, as will be seen.

In order to simplify statements, it will be convenient to have a name for the property shown in the above theorem. Specifically we shall say that the process $\{\xi_n\}$ *has extremal index* θ $(0 \le \theta \le 1)$ if (with the usual notation) for each $\tau > 0$

(i) there exists $u_n(\tau)$ such that $n(1 - F(u_n(\tau))) \to \tau$,
(ii) $P\{M_n \le u_n(\tau)\} \to e^{-\theta\tau}$.

By Lemma 3.6.2 if (i) holds for one fixed $\tau > 0$ it holds for all $\tau > 0$, and by Theorem 1.7.13 this is equivalent to (1.7.3). By Theorem 3.7.1, if (i) holds and $D(u_n(\tau))$ is satisfied for each τ, and if $P\{M_n \le u_n(\tau)\}$ converges for some $\tau > 0$, then (ii) holds with some θ, $0 \le \theta \le 1$ for all $\tau > 0$ and thus $\{\xi_n\}$ has an extremal index.

The next result modifies Theorem 3.5.1 when $D'(u_n)$ does not hold and generalizes a theorem of O'Brien (1974c) proved there under strong mixing assumptions. Here and subsequently in the section we continue to use the previous notation—for example, with $\hat{M}_n$ denoting the maximum of the first n terms of the independent sequence $\{\hat{\xi}_n\}$ associated with $\{\xi_n\}$.

Theorem 3.7.2. *Let the stationary sequence $\{\xi_n\}$ have extremal index θ. Let $\{v_n\}$ be a sequence of numbers, and $0 \le \rho \le 1$. Then*

(i) *for* $\theta > 0$,

$$\text{if } P\{\hat{M}_n \le v_n\} \to \rho \quad \text{then } P\{M_n \le v_n\} \to \rho^\theta, \quad \text{and conversely;}$$

(ii) *for* $\theta = 0$,

(a) *if* $\liminf_{n\to\infty} P\{\hat{M}_n \le v_n\} > 0$ *then* $P\{M_n \le v_n\} \to 1$,

(b) *if* $\limsup_{n\to\infty} P\{M_n \le v_n\} < 1$ *then* $P\{\hat{M}_n \le v_n\} \to 0$.

PROOF. (i) Suppose $\theta > 0$ and $P\{\hat{M}_n \le v_n\} \to \rho$ where $0 < \rho \le 1$. Choose $\tau > 0$ such that $e^{-\tau} < \rho$. Then

$$P\{\hat{M}_n \le u_n(\tau)\} \to e^{-\tau}, \qquad P\{\hat{M}_n \le v_n\} \to \rho > e^{-\tau},$$

so that we must have $v_n > u_n(\tau)$ for all sufficiently large n, and hence

$$\liminf_{n\to\infty} P\{M_n \le v_n\} \ge \lim_{n\to\infty} P\{M_n \le u_n(\tau)\} = e^{-\theta\tau}.$$

Since this is true for any τ such that $e^{-\tau} < \rho$ it follows that

$$\liminf_{n\to\infty} P\{M_n \le v_n\} \ge \rho^\theta.$$

(In particular, it follows that if $\rho = 1$, then $P\{M_n \le v_n\} \to 1 = \rho^\theta$ as desired.)

Similarly by taking $e^{-\tau} > \rho$ it is readily shown that $\limsup P\{M_n \le v_n\} \le \rho^\theta$ when $0 \le \rho < 1$. Hence $P\{M_n \le v_n\} \to 0$ when $\rho = 0$, as desired and for $0 < \rho < 1$, $P\{M_n \le v_n\} \to \rho^\theta$ by combining the inequalities for the upper and lower limits.

The proof that $P\{M_n \le v_n\} \to \rho^\theta$ implies $P\{\hat{M}_n \le v_n\} \to \rho$ is entirely similar, so that (i) follows.

For (ii) we assume that $\theta = 0$ so that $P\{M_n \le u_n(\tau)\} \to 1$ as $n \to \infty$, for each $\tau > 0$. If $\liminf P\{\hat{M}_n \le v_n\} = \rho > 0$, and τ is chosen with $e^{-\tau} < \rho$, then since $P\{\hat{M}_n \le u_n(\tau)\} \to e^{-\tau}$ it follows that $v_n > u_n(\tau)$ for all sufficiently large n, and hence

$$\liminf_{n\to\infty} P\{M_n \le v_n\} \ge \lim_{n\to\infty} P\{M_n \le u_n(\tau)\} = 1,$$

from which (a) follows.

On the other hand, if $\limsup P\{M_n \le v_n\} < 1$, it is readily seen from the fact that $P\{M_n \le u_n(\tau)\} \to 1$ that, for any $\tau > 0$, $v_n < u_n(\tau)$ when τ is sufficiently large, so that

$$\limsup_{n\to\infty} P\{\hat{M}_n \le v_n\} \le \lim_{n\to\infty} P\{\hat{M}_n \le u_n(\tau)\} = e^{-\tau}.$$

Since this holds for all τ, (b) follows by letting $\tau \to \infty$. □

As a corollary we may give general conditions under which the existence of an asymptotic distribution for $\hat{M}_n$ implies that M_n has an asymptotic distribution and conversely.

Corollary 3.7.3. *Let the stationary sequence $\{\xi_n\}$ have extremal index $\theta > 0$. Then M_n has a nondegenerate limiting distribution if and only if $\hat{M}_n$ does and these are then of the same type. Further the same normalization may be used, or one set of norming constants may be altered to give precisely the same limiting d.f.*

PROOF. If $P\{a_n(\hat{M}_n - b_n) \le x\} \to G(x)$, nondegenerate, then part (i) of the theorem shows (with $v_n = x/a_n + b_n$) that $P\{a_n(M_n - b_n) \le x\} \to G^\theta(x)$. Since G is max-stable, G^θ is of the same type as G (cf. Corollary 1.3.2). The converse follows similarly, noting that if $P\{a_n(M_n - b_n) \le x\} \to H(x)$ nondegenerate, then $P\{a_n(\hat{M}_n - b_n) \le x\} \to H^{1/\theta}(x)$, and $H^{1/\theta}$ must, as a limit for the maxima of an i.i.d. sequence, be max-stable.

Finally if $P\{a_n(\hat{M}_n - b_n) \le x\} \to G(x)$ then, by the above,

$$P\{a_n(M_n - b_n) \le x\} \to G(ax + b)$$

where $G^\theta(x) = G(ax + b)$ for some $a > 0$, b. Hence also

$$P\{\alpha_n(M_n - \beta_n) \le x\} \to G(x),$$

where $\alpha_n = aa_n$, $\beta_n = b_n - b/(aa_n)$. □

It follows, in particular, from this corollary (and Theorem 1.2.3) that when $\theta > 0$, if $\hat{M}_n$ and M_n both have nondegenerate limiting distributions, then these must be of the same type. If $\theta = 0$, however, it may be shown by an argument of Davis (1981) as in the following corollary, that $\hat{M}_n$ and M_n cannot both have nondegenerate limiting distributions *based on the same norming constants*. Of course if $D(u_n(\tau))$ does not hold, the trivial examples already referred to, provide cases where M_n can have a rather arbitrary distribution while $\hat{M}_n$ has a limiting distribution of extreme value type. We cite an example below where $D(u_n(\tau))$ does hold, but $\theta = 0$. Further, a recent example by de Haan (cf. Leadbetter (1982)) shows that when $\theta = 0$ it is possible for $\hat{M}_n$ and M_n to have limiting distributions (of different type) based on *different* norming constants.

Corollary 3.7.4. *Let the stationary sequence $\{\xi_n\}$ satisfy $D(u_n(\tau))$ when $u_n(\tau)$ satisfies $n(1 - F(u_n(\tau))) \to \tau$, $0 < \tau < \infty$, and let $\{\xi_n\}$ have extremal index $\theta = 0$. Then M_n and $\hat{M}_n$ cannot both have nondegenerate limiting distributions based on the same norming constants, i.e. it is not possible to have both*

$$P\{a_n(\hat{M}_n - b_n) \le x\} \to G(x), \quad P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} H(x)$$

for nondegenerate G, H.

PROOF. Suppose that the above convergence for both $\hat{M}_n$ and M_n does in fact occur. Then writing $v_n = x/a_n + b_n$, it follows from Theorem 3.7.2(ii) that $H(x) = 1$ whenever $G(x) > 0$. Hence it is readily seen that $x_0 = \inf\{x; G(x) > 0\}$ is finite and $H(x) = 1$ for $x \ge x_0$.

Since G is of extreme value type with finite left endpoint it must be of Type II, i.e. $G(x) = \psi\{a(x - x_0)\}$ where $\psi(x) = 0$ for $x \le 0$ and $\psi(x) = \exp(-x^{-\alpha})$ for $x > 0$, some $\alpha > 0$. Further by Corollary 1.6.3, if $\gamma_n = u_n(1)$,

$$P\{\hat{M}_n \le \gamma_n x\} \to \psi(x).$$

But also $P\{\hat{M}_n \le x/a_n + b_n\} \to \psi\{a(x - x_0)\}$ so that Theorem 1.2.3 gives

$$\gamma_n^{-1}a_n^{-1} \to a, \qquad \gamma_n^{-1}(b_n - 0) \to -ax_0.$$

Since $P\{M_n \le x/a_n + b_n\} \xrightarrow{w} H(x)$, a further application of Theorem 1.2.3 yields

$$P\{M_n \le \gamma_n x\} \xrightarrow{w} H\left(\frac{x}{a} + x_0\right).$$

Now since $D(u_n)$ must hold when $u_n = x/a_n + b_n$, H is of extreme value type and hence continuous so that $P\{M_n \le 0\} \to H(x_0) = 1$. But

$$P\{M_n \le 0\} \le P\{\xi_1 \le 0\} = F(0)$$

and $F(0) < 1$ since $F^n(0) = P\{\hat{M}_n \le 0\} \to \psi(0) = 0$. This contradicts the limit $P\{M_n \le 0\} \to 1$ and completes the proof of the corollary. □

We conclude this section with some examples from the literature showing something of the range of possible asymptotic behaviour for M_n. These examples are substantially concerned with cases where the index θ is less than 1. The more usual case in practice (where $D'(u_n)$ holds and $\theta = 1$) is illustrated in much more detail in Chapter 4 where normal sequences are considered.

Example 3.7.5. This example due to Chernick (1981a) concerns a strictly stationary first-order autoregressive sequence

$$\xi_n = \frac{1}{r}\xi_{n-1} + \varepsilon_n,$$

where $r \ge 2$ is an integer, $\{\varepsilon_n\}$ are i.i.d. and uniformly distributed on $\{0, 1/r, \ldots, (r-1)/r\}$, ε_n being independent of ξ_{n-1}, and ξ_n having a uniform distribution on the interval $[0, 1]$.

By the uniformity of ξ_n, $u_n(\tau)$ may be defined as $1 - \tau/n$. Chernick shows that with $u_n = u_n(\tau)$, $\tau > 0$, $D(u_n)$ holds, but $D'(u_n)$ fails. He then shows by direct argument that, for $x > 0$,

$$P\left\{M_n \le 1 - \frac{x}{n}\right\} \to \exp\left(-\frac{r-1}{r}x\right). \tag{3.7.2}$$

Replacement of x by $-x$ shows the Type III limit for M_n with the same norming constants as in the i.i.d. case (Example 1.7.9). This, of course, illustrates Corollary 3.7.3. Further, setting $u_n(\tau) = 1 - \tau/n$, (3.7.2) may be written as

$$P\{M_n \le u_n(\tau)\} \to e^{-\theta\tau},$$

showing that the extremal index $\theta = (r - 1)/r$. Since $r \geq 2$, $0 < \theta < 1$, and as r takes on the values $2, 3, \dots$, the index θ takes a sequence of values in the range $(0, 1)$. (However it seems that this example cannot be easily extended to other values of θ.) It is interesting to note that if $\varepsilon_{n+1} = (r - 1)/r$ (which happens with probability $1/r$) then $\xi_{n+1} > \xi_n$ and hence however large (i.e. close to 1) ξ_n is, there is a fixed probability $1/r$ that ξ_{n+1} is larger and then a probability $1/r^2$ that ξ_{n+2} is even larger and so on. Thus large values tend to occur in clusters, causing successive exceedances of u_n to be "too related" to allow $D'(u_n)$, $\theta = 1$, or the Poisson properties of exceedances discussed in Chapter 5. □

Example 3.7.6. Denzel and O'Brien (1975) consider a "chain dependent" process $\{\xi_n\}$ defined by means of an ergodic Markov chain $\{J_n; n \geq 0\}$ with the positive integers as states and connected by the requirement

$$\begin{aligned} P\{J_n = j, \xi_n \leq x \mid \xi_1, \dots, \xi_{n-1}, J_1, \dots, J_{n-2}, J_{n-1} = i\} \\ = P\{J_n = j, \xi_n \leq x \mid J_{n-1} = i\} = P_{ij} H_i(x), \end{aligned}$$

where P_{ij} are the transition probabilities for the chain and $H_i(x)$ is, for each i, a nondegenerate d.f. By appropriate choices of parameters they thus exhibit processes $\{\xi_n\}$ which are strongly mixing, have marginal d.f. $H(x) = \sum_1^\infty \pi_i H_i(x)$ (in which π_i are the stationary chain probabilities) and for which any value of the extremal index θ in $(0, 1]$ may be realized. Further a modification is given, choosing $P_{i,i+1} = (i + 1)(i + 3)^{-1}$ producing an example with $\theta = 0$. In this latter case $H(x) = 1 - ([x] + 2)^{-1}$, so that $\hat{M}_n$ has a nondegenerate, Type II limiting distribution. The level $u_n(\tau)$ is given explicitly as the smallest integer greater than or equal to $(n/\tau - 2)$ and

$$P\{M_n \leq u_n(\tau)\} \to 1,$$

but it is not obvious whether M_n has any sort of limiting distribution (compatible with Corollary 3.7.4). □

It is also possible to give examples where $D(u_n)$ hold (and even strong mixing) but where there is no extremal index.

Example 3.7.7. O'Brien (1974c) considers a sequence $\{\xi_n\}$ where each r.v. ξ_n is uniformly distributed over the interval $[0,1]$, $\xi_1, \xi_3, \xi_5, \dots$ being independent, and ξ_{2n} being defined as a function of ξ_{2n-1} for each n. In this way he obtains an obviously strongly mixing sequence, and further exhibits a sequence $\{v_n\}$ for which $P\{M_n \leq v_n\}$ converges to $e^{-1/2}$, but $P\{\hat{M}_n \leq v_n\}$ does not converge at all. It thus follows (Theorem 3.7.2) that $\{\xi_n\}$ has no extremal index and hence (Theorem 3.7.1) that $P\{M_n \leq u_n(\tau)\}$ does not converge for any $\tau > 0$.

By modifying the example a strongly mixing sequence $\{\xi_n\}$ is exhibited for which $P\{a_n(M_n - b_n) \leq x\}$ converges to a nondegenerate distribution (hence of extreme value type) but $P\{a_n(\hat{M}_n - b_n) \leq x\}$ does not converge.

That is $P\{M_n \le v_n\}$ converges for each member of the family $v_n = x/a_n + b_n$. This, of course, again reflects the fact that no extremal index exists in this case. □

The almost pathological nature of the available examples for which $D(u_n)$ holds but an extremal index is zero or does not exist, suggests that the cases of most practical interest are those for which an extremal index θ does exist and is nonzero. We turn now to another interesting class of such examples, for which $0 < \theta \le 1$. Examples of a different type with $0 < \theta < 1$ have also been given recently by de Haan (cf. Leadbetter (1982)).

3.8. Maxima of Moving Averages of Stable Variables

In this section we shall consider extreme values of i.i.d. stable (or "sum-stable" as opposed to max-stable) random variables with characteristic function given by (3.8.1) below, and of dependent sequences which are simply obtained from them, viz. as moving averages. This will give a conceptually easy illustration to one way in which dependence influences the behaviour of extremes, and also provide examples of processes with extremal index θ, for any $\theta \in (0, 1]$.

Because of the heavy tails of non-normal stable distributions, extreme values are caused by individual large summands, each creating a small cluster of extremes in the moving average process, and this explains why $D'(u_n)$ does not hold and extremal indices less than one are possible. In contrast, it can be seen that extreme values of moving averages of normal sequences are caused by rare combinations of moderately large summands, and that this is enough for $D'(u_n)$ to hold and for the extremal index to be equal to one.

Stationary normal sequences, of course, provide the most important example of sequences which are not independent, and as already mentioned they are treated in some generality in Chapter 4, and, later in Chapter 6. However, the stable sequences also seem to be interesting in their own right. In particular, they have the same linear structure as normal processes; arbitrary linear combinations of stable variables are stable. For the reader familiar with these concepts, it may also be mentioned that ARMA-processes with stable innovations are a special case of stable moving averages, as is easily seen by inverting the autoregressive part of the ARMA-process.

By definition, a random variable is (strictly) *stable* (γ, α, β) if it has the characteristic function

$$\psi(u) = \exp\left\{-\gamma^\alpha |u|^\alpha \left(1 - i\beta h(u, \alpha)\frac{u}{|u|}\right)\right\}, \tag{3.8.1}$$

with $0 \le \gamma$, $0 < \alpha \le 2$, $|\beta| \le 1$ and with $h(u, \alpha) = \tan(\pi\alpha/2)$ for $\alpha \ne 1$, $h(u, 1) = 2\pi^{-1} \log|u|$. Here γ is a scale parameter, α is called the *index* of the distribution, and β is the *symmetry* parameter. If $\beta = 0$ the distribution is symmetric, while if $|\beta| = 1$ and $\alpha < 2$ the distribution is said to be completely *asymmetric*. For $\alpha < 1$, the completely asymmetric stable distributions are concentrated on the positive real line if $\beta = 1$, and on the negative real line if $\beta = -1$.

If $\alpha = 2$ the distribution is clearly normal, and thus we shall in this section consider the case $0 < \alpha < 2$. If $\gamma = \alpha = 1$ and $\beta = 0$, then ψ is the characteristic function of a Cauchy distribution with density

$$f(x) = \frac{1}{\pi}\frac{1}{1 + x^2}, \qquad -\infty < x < \infty,$$

which was studied in Example 1.7.8. For $\gamma = 1, \alpha = \frac{1}{2}, \beta = 1$, the distribution has a density

$$f(x) = \frac{1}{\sqrt{2\pi}\, x^{3/2}} e^{-1/(2x)}, \qquad x > 0,$$

(and $f(x) = 0, x \le 0$), but apart from these cases, no simple expressions for the densities of stable distributions are known.

We shall derive here the limiting distribution of the maximum M_n of a moving average process $\{\xi_t\}$, defined by

$$\xi_t = \sum_{i=-\infty}^{\infty} c_i \zeta_{t-i},$$

where the ζ_t's are independent and stable $(1, \alpha, \beta)$ and where the constants $\{c_i\}$ are assumed to satisfy

$$\sum_{i=-\infty}^{\infty} |c_i|^\alpha < \infty \quad \text{and in addition, for } \alpha = 1, \beta \ne 0,$$

$$\sum_{i=-\infty}^{\infty} c_i \log|c_i| \text{ is convergent.} \tag{3.8.2}$$

It follows from (3.8.1) that $\sum_{i=-\infty}^{\infty} c_i\zeta_{t-i}$ converges in distribution if and only if (3.8.2) holds. Moreover, since the summands are independent, convergence in distribution of the sum implies convergence almost surely, and hence (3.8.2) is necessary and sufficient for ξ_t to be well defined as an a.s. convergent sum.

The only facts about the stable distributions we shall need in addition to (3.8.1) are the following estimates of the tails of $F_{\alpha\beta}$, the stable $(1, \alpha, \beta)$ distribution. Let $k_\alpha = \pi^{-1}\Gamma(\alpha)\sin(\alpha\pi/2)$. Then (cf. Bergström (1953))

$$\begin{aligned} 1 - F_{\alpha 1}(z) &\sim 2k_\alpha z^{-\alpha} && \text{as } z \to \infty, \\ F_{\alpha 1}(z) &= o(|z|^{-\alpha}) && \text{as } z \to -\infty, \end{aligned} \tag{3.8.3}$$

and for some suitable constants K_α,

$$F_{\alpha\beta}(-z) + (1 - F_{\alpha\beta}(z)) \le K_\alpha z^{-\alpha}, \qquad z > 0. \tag{3.8.4}$$

For completeness we shall, without proof, state an elementary result about convergence of distributions, which will be used below.

Lemma 3.8.1. *Let $\{\eta_n\}_{n=1}^{\infty}$, and $\{\eta_n^{(k)}\}_{n=1}^{\infty}$, $k = 1, 2, \ldots$ be r.v.'s and let F be a distribution function. Suppose that, for any $\varepsilon > 0$,*

$$\limsup_{n\to\infty} P\{|\eta_n - \eta_n^{(k)}| > \varepsilon\} \to 0 \quad \text{as } k \to \infty,$$

and that for all large k,

$$P\{\eta_n^{(k)} \le x\} \xrightarrow{w} F(x) \quad \text{as } n \to \infty.$$

Then

$$P\{\eta_n \le x\} \xrightarrow{w} F(x) \quad \text{as } n \to \infty.$$

Let

$$c_+ = \max_{-\infty < i < \infty} c_i^+, \qquad c_- = \max_{-\infty < i < \infty} c_i^-,$$

where $c_i^+ = \max(0, c_i)$, $c_i^- = \max(0, -c_i)$, and let $\tilde{M}_n = \max\{\zeta_1, \ldots, \zeta_n\}$.

The essential idea of the derivation of the limiting distribution of M_n is that a large value of a ξ_t is caused by just one of the summands ζ_{t-i} being large. In the particular case when the ζ's are completely asymmetric, with $\beta = 1$, this means that M_n and $c_+ \tilde{M}_n$ are asymptotically of the same size. This is proved in the next lemma for finite moving averages, together with some further useful estimates.

Lemma 3.8.2. *Let the ζ's be stable $(1, \alpha, \beta)$ and write*

$$\xi_t^{(k)} = \sum_{|i|<k} c_i \zeta_{t-i}, \qquad M_n^{(k)} = \max(\xi_1^{(k)}, \ldots, \xi_n^{(k)}).$$

(i) *Then, for $\varepsilon > 0$,*

$$\limsup_{n\to\infty} P\{n^{-1/\alpha} \max(|\xi_1 - \xi_1^{(k)}|, \ldots, |\xi_n - \xi_n^{(k)}|) > \varepsilon\} \to 0 \quad \text{as } k \to \infty.$$

If in addition $\beta = 1$, $c_+ > 0$, and $k > |k_0|$, where k_0 is such that $c_+ = c_{k_0}$,

(ii) *then, for $\varepsilon > 0$,*

$$P\{n^{-1/\alpha} |M_n^{(k)} - c_+ \tilde{M}_n| > \varepsilon\} \to 0 \quad \text{as } n \to \infty,$$

and

(iii)
$$P\{n^{-1/\alpha} M_n^{(k)} \le x\} \to \exp\{-k_\alpha c_+^\alpha 2x^{-\alpha}\} \quad \text{as } n \to \infty.$$

Proof. (i) Suppose first that $\alpha \ne 1$. Then it follows from (3.8.1) that

$$\xi_1 - \xi_1^{(k)} = \sum_{|i|\ge k} c_i \zeta_{1-i}$$

is stable with some (in general different) symmetry parameter and with scale parameter $(\sum_{|i|\ge k} |c_i|^\alpha)^{1/\alpha}$ and index α. Hence, by (3.8.4)

$$\begin{aligned} P\{n^{-1/\alpha}|\xi_1 - \xi_1^{(k)}| > \varepsilon\} &= P\left\{\left(\sum_{|i|\ge k} |c_i|^\alpha\right)^{-1/\alpha} |\xi_1 - \xi_1^{(k)}| \right. \\ &\qquad \left. > \left(\sum_{|i|\ge k} |c_i|^\alpha\right)^{-1/\alpha} n^{1/\alpha} \varepsilon\right\} \\ &\le K_\alpha \frac{\sum_{|i|\ge k} |c_i|^\alpha}{n\varepsilon^\alpha} \end{aligned}$$

and thus, by stationarity,

$$P\{n^{-1/\alpha}\max(|\xi_1 - \xi_1^{(k)}|, \ldots, |\xi_n - \xi_n^{(k)}|) > \varepsilon\} \le K_\alpha \varepsilon^{-\alpha} \sum_{|i|\ge k} |c_i|^\alpha,$$

which tends to zero as $k \to \infty$, and thus (i) holds if $a \ne 1$. If $\alpha = 1$ then instead

$$\xi_1 - \xi_1^{(k)} - \frac{2\beta}{\pi} \sum_{|i|\ge k} c_i \log|c_i|$$

is stable with scale parameter $\sum_{|i|\ge k}|c_i|$ and index 1. Hence

$$\begin{aligned} P\{n^{-1}|\xi_1 - \xi_1^{(k)}| > \varepsilon\} &\le P\left\{\left(\sum_{|i|\ge k}|c_i|\right)^{-1}\left|\xi_1 - \xi_1^{(k)} - \frac{2\beta}{\pi}\sum_{|i|\ge k} c_i \log|c_i|\right|\right. \\ &\qquad \left. > \left(\sum_{|i|\ge k}|c_i|\right)^{-1}\left(n\varepsilon - \left|\frac{2\beta}{\pi}\sum_{|i|\ge k} c_i\log|c_i|\right|\right)\right\} \\ &\le K_\alpha \frac{\sum_{|i|\ge k}|c_i|}{n\left(\varepsilon - \left|\frac{2\beta}{n\pi}\sum_{|i|\ge k} c_i \log|c_i|\right|\right)} \end{aligned}$$

for large n, by (3.8.4), and in the same way as above it then follows that (i) holds also for $\alpha = 1$.

(ii) Let $\varepsilon_n = n^{1/\alpha}\varepsilon/(4k(c_+ + c_-))$ and define events

$A_n = \{\zeta_i > \varepsilon_n, \zeta_j > \varepsilon_n$ for some i, j wih $1 - k < i < j < n + k, |j - i| < 2k\}$,

$B_n = \{\zeta_i < -\varepsilon_n$ for some i with $1 - k < i < n + k\}$,

and

$C_n = \{|\zeta_i| > \varepsilon_n$ for some i with $1 - k < i < 1 + k$ or $n - k < i < n + k\}$.

We shall show that

$$\{n^{-1/\alpha}|M_n^{(k)} - c_+ \tilde{M}_n| > \varepsilon\} \subset A_n \cup B_n \cup C_n \tag{3.8.5}$$

and that $P(A_n \cup B_n \cup C_n) \to 0$ as $n \to \infty$.

Now if neither A_n nor C_n occurs, for $1 \le t \le n$ the largest of $\zeta_{t-k+1}, \ldots, \zeta_{t+k-1}$ does not exceed $\tilde{M}_n + \varepsilon_n$ and the others are not greater than ε_n, and then

$$\sum_{|i|<k} c_i^+ \zeta_{t-i} \le c_+ \tilde{M}_n + (2k-1)c_+\varepsilon_n \le c_+\tilde{M}_n + \frac{n^{1/\alpha}\varepsilon}{2}, \qquad t = 1, \ldots, n.$$

Similarly, if neither B_n nor C_n occurs, $\zeta_i \ge -\varepsilon_n$ for $1 - k < i < n + k$, and hence

$$-\sum_{|i|<k} c_i^- \zeta_{t-i} \le (2k-1)c_-\varepsilon_n \le \frac{n^{1/\alpha}\varepsilon}{2}, \qquad t = 1, \ldots, n.$$

Thus, on $(A_n \cup B_n \cup C_n)^c$,

$$\xi_t^{(k)} = \sum_{|i|<k} c_i^+ \zeta_{t-i} - \sum_{|i|<k} c_i^- \zeta_{t-i} \le c_+\tilde{M}_n + n^{1/\alpha}\varepsilon, \qquad t = 1, \ldots, n,$$

i.e.

$$M_n^{(k)} \le c_+ \tilde{M}_n + n^{1/\alpha}\varepsilon. \tag{3.8.6}$$

Next, suppose t_0, k_0 are such that $\zeta_{t_0} = \tilde{M}_n$, $c_{k_0} = c_+$ and $1 \le t_0 \le n$, $|k_0| < k$. If B_n does not occur, then

$$\sum_{\substack{i \neq k_0 \\ |i| < k}} c_i^+ \zeta_{t_0 + k_0 - i} \ge -(2k-2)c_+ \varepsilon_n \ge -\frac{n^{1/\alpha}\varepsilon}{2}.$$

Further, on $A_n^c \cap C_n^c$,

$$-\sum_{\substack{i \neq k_0 \\ |i| < k}} c_i^- \zeta_{t_0 + k_0 - i} \ge -(2k-2)c_- \varepsilon_n \ge -\frac{n^{1/\alpha}\varepsilon}{2},$$

since if $\zeta_{t_0} > \varepsilon_n$ then A_n^c implies that $\zeta_{t_0+k_0-k+1}, \ldots, \zeta_{t_0+k_0+k-1}$ are smaller than ε_n, and if instead $\zeta_{t_0} = \tilde{M}_n \le \varepsilon_n$, this holds trivially on C_n^c. Thus

$$\begin{aligned}
\xi_{t_0+k_0}^{(k)} &= c_{k_0}\zeta_{t_0} + \sum_{\substack{i \neq k_0 \\ |i| < k}} c_i \zeta_{t_0+k_0-i} \\
&= c_+ \tilde{M}_n + \sum_{\substack{i \neq k_0 \\ |i| < k}} c_i^+ \zeta_{t_0+k_0-i} - \sum_{\substack{i \neq k_0 \\ |i| < k}} c_i^- \zeta_{t_0+k_0-i} \\
&\ge c_+ \tilde{M}_n - n^{1/\alpha}\varepsilon,
\end{aligned}$$

on $(A_n \cup B_n \cup C_n)^c$, and, if in addition $1 \le t_0 + k_0 \le n$, then

$$M_n \ge c_+ \tilde{M}_n - n^{1/\alpha}\varepsilon. \tag{3.8.7}$$

However, if instead $t_0 + k_0 < 1$ or $t_0 + k_0 > n$ and neither B_n nor C_n occurs then $|\zeta_i| \le \varepsilon_n$ for $1 - k < i < n + k$, and then $|M_n^{(k)}| \le (c_+ + c_-)(2k-1)\varepsilon_n$ and $|\tilde{M}_n| \le \varepsilon_n$, and it follows that

$$M_n^{(k)} - c_+ \tilde{M}_n \ge -(2k-1)(c_+ + c_-)\varepsilon_n - c_+ \varepsilon_n \ge -n^{1/\alpha}\varepsilon,$$

so (3.8.7) holds on $(A_n \cup B_n \cup C_n)^c$ also in this case. Together (3.8.6) and (3.8.7) prove (3.8.5).

Now, by the independence of $\zeta_i, \zeta_j, i \neq j$,

$$\begin{aligned}
P(A_n) &\le \sum_{i=2-k}^{n+k-2} \sum_{j=i+1}^{i+2k-1} P\{\zeta_i > \varepsilon_n, \zeta_j > \varepsilon_n\} \\
&\le (n + 2k - 3)(2k-1)P\{\zeta_1 > \varepsilon_n\}^2 \to 0 \quad \text{as } n \to \infty,
\end{aligned}$$

since $P\{\zeta_1 > \varepsilon_n\} = O(1/n)$ by the first part of (3.8.3) and the definition of ε_n. Further, this time by the second part of (3.8.3),

$$P(B_n) \le (n + 2k - 2)P\{\zeta_1 < -\varepsilon_n\} \to 0 \quad \text{as } n \to \infty,$$

and finally

$$P(C_n) \le (4k-2)P\{|\zeta_1| > \varepsilon_n\} \to 0 \quad \text{as } n \to \infty,$$

since $\varepsilon_n \to \infty$. Thus, by (3.8.5), it follows that

$$P\{n^{-1/\alpha}|M_n^{(k)} - c_+ \tilde{M}_n| > \varepsilon\} \le P(A_n) + P(B_n) + P(C_n) \to 0 \quad \text{as } n \to \infty.$$

(iii) By (3.8.3)

$$P\{\zeta_1 > n^{1/\alpha}x/c_+\} \sim 2k_\alpha c_+ x^{-\alpha} n^{-1},$$

and thus Theorem 1.5.1 shows that

$$P\{n^{-1/\alpha} c_+ \tilde{M}_n \le x\} \to \exp(-2k_\alpha c_+^\alpha x^{-\alpha}),$$

and (iii) now follows from (ii) and Lemma 3.8.1. □

In the last part of the lemma we have obtained the limiting distribution of M_n for the case when $\beta = 1$ and the moving average is finite. The general result then follows easily, as we shall now see.

Theorem 3.8.3. *Suppose* $\{\xi_t = \sum_i c_i \zeta_{t-i}\}_{t=1}^\infty$ *is a moving average of stable* $(1, \alpha, \beta)$ *variables* $\{\zeta_t\}$, *with the constants* $\{c_i\}_{-\infty}^\infty$ *satisfying* (3.8.2). *Then*

$$P\{n^{-1/\alpha} M_n \le x\} \to \begin{cases} \exp\{-k_\alpha(c_+^\alpha (1+\beta) + c_-^\alpha(1-\beta))x^{-\alpha}\}, & \text{if } x > 0, \\ 0, & \text{if } x \le 0. \end{cases} \tag{3.8.8}$$

Proof. By Lemmas 3.8.1 and 3.8.2 it is sufficient to show that

$$P\{n^{-1/\alpha} M_n^{(k)} \le x\} \to \exp\{-k_\alpha(c_+^\alpha(1+\beta) + c_-^\alpha(1-\beta))x^{-\alpha}\} \quad \text{as } n \to \infty, \tag{3.8.9}$$

for k large (and $x > 0$). Suppose $\alpha \ne 1$, and let $\{\zeta_i', \zeta_i''\}_{-\infty}^\infty$ be independent and stable $(1, \alpha, 1)$. It is then immediate from (3.8.1) that $((1+\beta)/2)^{1/\alpha}\zeta_i' - ((1-\beta)/2)^{1/\alpha}\zeta_i''$ is stable $(1, \alpha, \beta)$, and thus, defining

$$\xi_t' = \sum_{|i|<k} c_i \left(\frac{1+\beta}{2}\right)^{1/\alpha} \zeta_{t-i}', \qquad \xi_t'' = \sum_{|i|<k} \left[-c_i\left(\frac{1-\beta}{2}\right)^{1/\alpha}\right] \zeta_{t-i}'',$$

the sequences $\{\xi_t^{(k)}\}_1^\infty$ and $\{\xi_t' + \xi_t''\}_1^\infty$ have the same distribution. Hence, (3.8.9) is equivalent to

$$\begin{aligned} P\{n^{-1/\alpha}&\max(\xi_1' + \xi_1'', \ldots, \xi_n' + \xi_n'') \le x\} \\ &\to \exp\{-k_\alpha(c_+^\alpha(1+\beta) + c_-^\alpha(1-\beta))x^{-\alpha}\}. \end{aligned} \tag{3.8.10}$$

Let $M_n' = \max(\xi_1', \ldots, \xi_n')$ and $M_n'' = \max(\xi_1'', \ldots, \xi_n'')$. If $k \ge |k_0|$ for k_0 satisfying $\max(c_{-k_0+1}^+, \ldots, c_{k_0-1}^+) = c_+$, $\max(c_{-k_0+1}^-, \ldots, c_{k_0-1}^-) = c_-$, then, since M_n' and M_n'' are independent,

$$\begin{aligned} P\{n^{-1/\alpha}\max(M_n', M_n'') \le x\} &= P\{n^{1/\alpha}M_n' \le x\}P\{n^{1/\alpha}M_n'' < x\} \\ &\to \exp\{-k_\alpha c_+^\alpha(1+\beta)x^{-\alpha} - k_\alpha c_-^\alpha(1-\beta)x^{-\alpha}\}, \end{aligned} \tag{3.8.11}$$

by Lemma 3.8.2(iii). Further, it can be seen that, for $\varepsilon > 0$,

$$\begin{aligned} P\{|\max(\xi_1' + \xi_1'', \ldots, \xi_n' + \xi_n'') &- \max(M_n', M_n'')| > n^{1/\alpha}\varepsilon\} \\ &\le P\{|\xi_t'| > n^{1/\alpha}\varepsilon, |\xi_t''| > n^{1/\alpha}\varepsilon \text{ for some } t \text{ with } 1 \le t \le n\} \\ &\quad + P\{\max(M_n', M_n'') \le n^{1/\alpha}\varepsilon\}. \end{aligned} \tag{3.8.12}$$

Clearly

$$P\{|\xi_t'| > n^{1/\alpha}\varepsilon, |\xi_t''| > n^{1/\alpha}\varepsilon \text{ for some } t \text{ with } 1 \le t \le n\}$$
$$\le n P\{|\xi_1'| > n^{1/\alpha}\varepsilon\} P\{|\xi_1''| > n^{1/\alpha}\varepsilon\} \to 0 \quad \text{as } n \to \infty,$$

since $P\{|\xi_1'| > n^{1/\alpha}\varepsilon\} = 0(1/n)$ and $P\{|\xi_1''| > n^{1/\alpha}\varepsilon\} = 0(1/n)$ by (3.8.4), and further

$$\limsup_{n\to\infty} P\{\max(M_n', M_n'' \le n^{1/\alpha}\varepsilon\} \to 0 \quad \text{as } \varepsilon \to 0,$$

by (3.8.11). Thus it follows from (3.8.12) that

$$\limsup_{n\to\infty} P\{n^{-1/\alpha}|\max(\xi_1' + \xi_1'', \dots, \xi_n' + \xi_n'') - \max(M_n', M_n'')| > \varepsilon\} \to 0$$
$$\text{as } k \to \infty.$$

By Lemma 3.8.1 and (3.8.11) this proves (3.8.10) and hence (3.8.9) holds for all large k, which proves the theorem for $\alpha \neq 1$.

If $\alpha = 1$ then, with $\{\zeta_t', \zeta_t''\}$ as above, for $|\beta| < 1$,

$$\left(\frac{1+\beta}{2}\right)\zeta_t' - \left(\frac{1-\beta}{2}\right)\zeta_t'' + \frac{1}{\pi}((1+\beta)\log\left(\frac{1+\beta}{2}\right) - (1-\beta)\log\left(\frac{1-\beta}{2}\right)$$

is stable $(1, 1, \beta)$. However, to deal with the additional constant term requires only trivial changes in the proof above, of a similar kind as in the proof of Lemma 3.8.2(i) and we omit the details. □

The reader is referred to Rootzén (1978) for further results on the extremal behaviour of $\{\xi_t\}$, including counterparts to the point process convergence to be proved in Chapter 5 for processes with extremal index one, and quite detailed information on the behaviour of sample paths near extremes, as well as the corresponding results for continuous parameter processes. Further, it may be noted that essentially it is the tail behaviour of the ζ's which determines the behaviour of extremes of $\{\xi_t\}$, and that similar results hold for other distributions of ζ's, which are in the domain of attraction of the Type II max-stable distribution.

Finally, by (3.8.1), if $\alpha \neq 1$, ξ_t is stable $((\sum |c_i|^\alpha)^{1/\alpha}, \alpha, \beta \sum c_i |c_i|^{\alpha-1}/\sum |c_i|^\alpha)$, and hence, if $\hat{M}_n$ is the maximum of the associated independent process,

$$P\{n^{-1/\alpha}\hat{M}_n \le x\} \to \exp\left\{-k_\alpha \sum_{i=-\infty}^{\infty} (|c_i| + \beta c_i)|c_i|^{\alpha-1}x^{-\alpha}\right\},$$

by Theorem 3.8.3 (with $c_0 = (\sum |c_i|^\alpha)^{1/\alpha}$ and $c_i = 0$, $i \neq 0$). By comparing with the limit for $n^{-1/\alpha} M_n$, it follows that $\{\xi_t\}$ has an extremal index

$$\frac{c_+^\alpha(1+\beta) + c_-^\alpha(1-\beta)}{\sum_{i=-\infty}^{\infty} (|c_i| + \beta c_i)|c_i|^{\alpha-1}},$$

which clearly can take any value in (0, 1].

CHAPTER 4

Normal Sequences

Normality occupies a central place in probability and statistical theory, and a most important class of stationary sequences consists of those which are normal. Their importance is enhanced by the fact that their joint normal distributions are determined by the mean and the covariance structure of the sequence. In this chapter we investigate the extremal properties of stationary normal sequences. In particular covariance conditions will be obtained for the convergence of maxima to a Type I limit, both directly and by application of the general theory of Chapter 3.

4.1. Stationary Normal Sequences and Covariance Conditions

A sequence $\{\xi_n\}$ of r.v.'s is said to be *normal* if for any choice of $n, i_1, \ldots, i_n$, the joint distribution of $\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_n}$ is an n-dimensional normal distribution. These finite-dimensional distributions are clearly determined by the means of the individual ξ_n and the covariances between the pairs ξ_n, ξ_m for all n, m.

For *stationary* normal sequences the mean and variance of ξ_n do not depend on n and may—without loss of generality—be conveniently taken as zero and one, respectively, yielding a "standard normal sequence." We shall assume—usually without comment—that the sequences considered have been so standardized.

Further by stationarity, the covariances between pairs such as ξ_n, ξ_m depend only on the difference between m and n (and indeed only on its absolute value) so that we write

$$\mathrm{Cov}(\xi_n, \xi_m) = r_{n-m} = r_{|n-m|},$$

where $\{r_n\}$ is termed the covariance sequence of the process. By the standardization it follows that $r_0 = \text{Var}(\xi_n) = 1$ and $r_n = \text{Cov}(\xi_j, \xi_{j+n}) = E(\xi_j \xi_{j+n})$.

Clearly all of the finite-dimensional distributions $F_{i_1 \ldots i_n}(x_1, \ldots, x_n) = P\{\xi_{i_1} \le x_1, \ldots, \xi_{i_n} \le x_n\}$ of a standard normal sequence, are determined by the covariance sequence $\{r_n\}$. Of course a covariance sequence cannot be arbitrarily specified, since the covariance matrix of any group $\xi_{i_1}, \ldots, \xi_{i_n}$ must be non-negative definite.

As noted in Chapter 3, Berman (1964b) has given simple conditions on $\{r_n\}$ to ensure that (1.5.5) holds, i.e. that $M_n = \max(\xi_1, \ldots, \xi_n)$ has a limiting distribution of the double exponential type,

$$P\{a_n(M_n - b_n) \le x\} \to \exp(-e^{-x}) \text{ as } n \to \infty,$$

with the same constants as in Theorem 1.5.3, viz.

$$\begin{aligned} a_n &= (2 \log n)^{1/2}, \\ b_n &= a_n - (2a_n)^{-1}\{\log \log n + \log 4\pi\}. \end{aligned}$$

One of Berman's results is that it suffices that

$$r_n \log n \to 0 \quad \text{as } n \to \infty, \tag{4.1.1}$$

and we will devote the first part of this chapter to a proof of this. However, (4.1.1) can be replaced by an appropriate Cesàro convergence,

$$\frac{1}{n} \sum_{k=1}^{n} |r_k| \log k \, e^{\gamma |r_k| \log k} \to 0 \quad \text{as } n \to \infty,$$

for some $\gamma > 2$, and this and some other conditions will be discussed at the end of the chapter.

It has been shown by Mittal and Ylvisaker (1975) that (4.1.1), and therefore also the Cesàro convergence, is rather close to being necessary, in that if $r_n \log n \to \gamma > 0$ a different limit applies. In fact, this strong dependence destroys the asymptotic independence of extremes in disjoint intervals (cf. Lemmas 3.2.2, 3.3.1, and 3.3.2). We return to these matters in more detail in Chapter 6.

In this chapter we obtain the relevant convergence results for normal sequences, including the Type I limit, under various covariance assumptions, starting with the transparent condition (4.1.1). This will first be done (in Section 4.3) by a direct comparison with an i.i.d. normal sequence without reference to the $D(u_n)$, $D'(u_n)$ conditions. It will then be shown, with very little further effort, that these conditions hold, so that the results also follow from the general theory of Chapter 3. For subsequent results, e.g., in Chapter 5, it will be more convenient to rely solely on the general theory via $D(u_n)$, $D'(u_n)$ for applications to normal sequences but here both approaches are simple and instructive. Finally, the rates at which convergence occurs will be discussed in Section 4.6.

4.2. Normal Comparison Lemma

Our main task is to show that if (4.1.1) holds, then $D(u_n)$ and $D'(u_n)$ are satisfied for appropriate sequences $\{u_n\}$. The main tool for this purpose is a widely useful result—here called the Normal Comparison Lemma—which bounds the difference between two (standardized) n-dimensional d.f.'s, by a convenient function of their covariances. This result has been developed in various ways by Slepian (1962), Berman (1964b, 1971a), and Cramér (see Cramér and Leadbetter (1967)). The lemma is given here in some generality for use in later chapters, even though only a simple special form is needed at this point.

Theorem 4.2.1 (Normal Comparison Lemma). *Suppose $\xi_1, \ldots, \xi_n$ are standard normal variables with covariance matrix $\Lambda^1 = (\Lambda_{ij}^1)$, and $\eta_1, \ldots, \eta_n$ similarly with covariance matrix $\Lambda^0 = (\Lambda_{ij}^0)$, and let $\rho_{ij} = \max(|\Lambda_{ij}^1|, |\Lambda_{ij}^0|)$. Further, let $u_1, \ldots, u_n$ be real numbers. Then*

$$P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} - P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\}$$
$$\le \frac{1}{2\pi} \sum_{1 \le i < j \le n} (\Lambda_{ij}^1 - \Lambda_{ij}^0)^+ (1 - \rho_{ij}^2)^{-1/2} \exp\left(-\frac{\frac{1}{2}(u_i^2 + u_j^2)}{1 + \rho_{ij}}\right), \quad (4.2.1)$$

where $(x)^+ = \max(0, x)$.

In particular, if $\max_{i \ne j} |\rho_{ij}| = \delta < 1$, then

$$P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} - P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\}$$
$$\le K \sum_{1 \le i < j \le n} (\Lambda_{ij}^1 - \Lambda_{ij}^0)^+ \exp\left(-\frac{\frac{1}{2}(u_i^2 + u_j^2)}{1 + \rho_{ij}}\right) \quad (4.2.2)$$

for some constant K, depending only on δ. Further

$$|P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} - P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\}|$$
$$\le \frac{1}{2\pi} \sum_{1 \le i < j \le n} |\Lambda_{ij}^1 - \Lambda_{ij}^0| (1 - \rho_{ij}^2)^{-1/2} \exp\left(-\frac{\frac{1}{2}(u_i^2 + u_j^2)}{1 + \rho_{ij}}\right), \quad (4.2.3)$$

where the factor $(1/2\pi)(1 - \rho_{ij}^2)^{-1/2}$ can be replaced by K if $\max_{i \ne j} \rho_{ij} = \delta < 1$.

Proof. We shall suppose that Λ^1 and Λ^0 are positive definite (as opposed to semi-definite) and hence that $(\xi_1, \ldots, \xi_n)$ and $(\eta_1, \ldots, \eta_n)$ have joint densities f_1 and f_0, respectively. (The semi-definite case is easily dealt with by considering $\xi_i + \varepsilon_i$ and $\eta_i + \varepsilon_i$, where the ε_i are independent normal variables with mean zero and then letting $\operatorname{Var}(\varepsilon_i) \to 0$, using continuity.) Clearly, with $\mathbf{u} = (u_1, \ldots, u_n)$,

$$P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} = \int_{-\infty}^{\mathbf{u}} \cdots \int f_1(y_1, \ldots, y_n)\, d\mathbf{y},$$

$$P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\} = \int_{-\infty}^{\mathbf{u}} \cdots \int f_0(y_1, \ldots, y_n)\, d\mathbf{y},$$

where f_1, f_0 are the normal density functions based on the covariance matrices Λ^1, Λ^0, the integration ranges being $\{\mathbf{y}; y_j \le u_j, j = 1, \dots, n\}$.

If we write $\Lambda_h = h\Lambda^1 + (1-h)\Lambda^0$, $(0 \le h \le 1)$, the matrix Λ_h is positive definite with units down the main diagonal and elements $h\Lambda_{ij}^1 + (1-h)\Lambda_{ij}^0$ for $i \ne j$. Let f_h be the n-dimensional normal density based on Λ_h, and

$$F(h) = \int_{-\infty}^{\mathbf{u}} \cdots \int f_h(y_1, \dots, y_n)\, d\mathbf{y}.$$

The left-hand side of (4.2.1) is then easily recognized as $F(1) - F(0)$. Now

$$F(1) - F(0) = \int_0^1 F'(h)\, dh,$$

where

$$F'(h) = \int_{-\infty}^{\mathbf{u}} \cdots \int \frac{\partial f_h(y_1, \dots, y_n)}{\partial h}\, d\mathbf{y}.$$

The density f_h depends on h only through the elements Λ_{ij}^h of Λ_h (regarding f_h as a function of Λ_{ij}^h for $i \le j$, say). We have $\Lambda_{ii}^h = 1$ independent of h, while for $i < j$, $\Lambda_{ij}^h = h\Lambda_{ij}^1 + (1-h)\Lambda_{ij}^0$ so that $\partial\Lambda_{ij}^h/\partial h = \Lambda_{ij}^1 - \Lambda_{ij}^0$. Thus

$$F'(h) = \sum_{i \le j} \int_{-\infty}^{\mathbf{u}} \cdots \int \frac{\partial f_h}{\partial \Lambda_{ij}^h} \cdot \frac{\partial \Lambda_{ij}^h}{\partial h}\, d\mathbf{y} = \sum_{i<j} (\Lambda_{ij}^1 - \Lambda_{ij}^0) \int_{-\infty}^{\mathbf{u}} \cdots \int \frac{\partial f_h}{\partial \Lambda_{ij}^h}\, d\mathbf{y}.$$

Now a useful property of the multidimensional normal density is that its derivative with respect to a covariance Λ_{ij} is the same as the second mixed derivative with respect to the corresponding variables y_i, y_j (cf. Cramér and Leadbetter (1967, p. 26)), i.e.

$$\frac{\partial f_h}{\partial \Lambda_{ij}} = \frac{\partial^2 f_h}{\partial y_i\, \partial y_j}.$$

Thus

$$F'(h) = \sum_{i<j} (\Lambda_{ij}^1 - \Lambda_{ij}^0) \int_{-\infty}^{\mathbf{u}} \cdots \int \frac{\partial^2 f_h}{\partial y_i\, \partial y_j}\, d\mathbf{y}.$$

The y_i and y_j integrations may be done at once to give

$$F'(h) = \sum_{i<j} (\Lambda_{ij}^1 - \Lambda_{ij}^0) \int_{-\infty}^{\mathbf{u}'} \cdots \int f_h(y_i = u_i, y_j = u_j)\, d\mathbf{y}' \qquad (4.2.4)$$

where $f_h(y_i = u_i, y_j = u_j)$ denotes the function of $n-2$ variables formed by putting $y_i = u_i$, $y_j = u_j$, the integration being over the remaining variables.

Further, we can dominate the last integral by letting the variables run from $-\infty$ to $+\infty$. But

$$\int_{-\infty}^{\infty} \cdots \int f_h(y_i = u_i, y_j = u_j)\, d\mathbf{y}'$$

is just the bivariate density, evaluated at (u_i, u_j), of two standard normal random variables with correlation Λ_{ij}^h, and may therefore be written

$$\frac{1}{2\pi(1-(\Lambda_{ij}^h)^2)^{1/2}}\exp\left\{-\frac{1}{2(1-(\Lambda_{ij}^h)^2)}(u_i^2-2\Lambda_{ij}^h u_i u_j+u_j^2)\right\}.$$

Now, since $|\Lambda_{ij}^h| = |h\Lambda_{ij}^1 + (1-h)\Lambda_{ij}^0| \le \max(\Lambda_{ij}^1, \Lambda_{ij}^0) = \rho_{ij}$, it may be easily shown that the above expression does not exceed

$$\frac{1}{2\pi(1-\rho_{ij}^2)^{1/2}}\exp\left(-\frac{\frac{1}{2}(u_i^2+u_j^2)}{1+\rho_{ij}}\right).$$

(Note that $(u^2 - 2\rho uv + v^2)/(1-|\rho|) \ge (u^2 - 2|\rho||u||v| + v^2)/(1-|\rho|)$ which is a minimum when $\rho = 0$.) Eliminating any negative terms in (4.2.4) we thus obtain

$$F'(h) \le \frac{1}{2\pi}\sum_{i<j}(\Lambda_{ij}^1-\Lambda_{ij}^0)^+(1-\rho_{ij}^2)^{-1/2}\exp\left(-\frac{\frac{1}{2}(u_i^2+u_j^2)}{1+\rho_{ij}}\right)$$

and since $F(1) - F(0) = \int_0^1 F'(h)\,dh$, this proves (4.2.1).

The conclusion (4.2.2) follows at once from (4.2.1) since $(1-\rho_{ij}^2)^{-1/2} \le (1-\delta^2)^{-1/2}$, and (4.2.3) follows by using (4.2.1) as stated and with the roles of ξ_j, η_j interchanged, noting that $(\Lambda_{ij}^1 - \Lambda_{ij}^0)^+$ and $(\Lambda_{ij}^0 - \Lambda_{ij}^1)^+$ are each no greater than $|\Lambda_{ij}^1 - \Lambda_{ij}^0|$. □

Some obvious but useful corollaries are stated here for later reference. The first of these gives a slight simplification of the bounds which is sometimes convenient.

Corollary 4.2.2. *With the notation of the theorem write* $u = \min(u_1, u_2, \ldots, u_n)$. *Then the factor* $\exp(-\frac{1}{2}(u_i^2+u_j^2)/(1+\rho_{ij}))$ *may be replaced by* $\exp(-u^2/(1+\rho_{ij}))$ *in each of* (4.2.1), (4.2.2), *and* (4.2.3). *In particular,* (4.2.2) *and* (4.2.3) *become, respectively,*

$$P\{\xi_j \le u_j \text{ for } j=1,\ldots,n\} - P\{\eta_j \le u_j \text{ for } j=1,\ldots,n\}$$
$$\le K\sum_{1\le i<j\le n}(\Lambda_{ij}^1-\Lambda_{ij}^0)^+\exp\left(\frac{-u^2}{1+\rho_{ij}}\right), \tag{4.2.5}$$

$$|P\{\xi_j \le u_j \text{ for } j=1,\ldots,n\} - P\{\eta_j \le u_j \text{ for } j=1,\ldots,n\}|$$
$$\le K\sum_{1\le i<j\le n}|\Lambda_{ij}^1-\Lambda_{ij}^0|\exp\left(\frac{-u^2}{1+\rho_{ij}}\right). \tag{4.2.6}$$

Proof. These results follow immediately since clearly $\frac{1}{2}(u_i^2+u_j^2) \ge u^2$. □

The next corollary shows that if the covariances of the ξ's are dominated by those of the η's, then the ξ's are stochastically larger than the η's and the maximum of the ξ's is stochastically larger than that of the η's.

Corollary 4.2.3. *Let $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n$ be standard normal r.v.'s with* $\mathrm{Cov}(\xi_i, \xi_j) \le \mathrm{Cov}(\eta_i, \eta_j)$ *for each i, j. Then, for any $u_1, \ldots, u_n$,*

$$P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} \le P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\}. \quad (4.2.7)$$

In particular

$$P\{\max(\xi_1, \ldots, \xi_n) \le u\} \le P\{\max(\eta_1, \ldots, \eta_n) \le u\} \quad (4.2.8)$$

for all u.

Proof. This follows at once from (4.2.2) since $(\Lambda_{ij}^1 - \Lambda_{ij}^0)^+ = 0$ when $\Lambda_{ij}^1 \le \Lambda_{ij}^0$ as assumed. □

The final corollary explores the consequences of taking the η_j's to be independent.

Corollary 4.2.4. *Let $\xi_1, \ldots, \xi_n$ be jointly normal (standardized) r.v.'s with* $\mathrm{Cov}(\xi_i, \xi_j) = \Lambda_{ij}$, *such that* $\delta = \max_{i \ne j} |\Lambda_{ij}| < 1$. *Then for any real u and integers* $1 \le l_1 < \cdots < l_s \le n$,

$$|P\{\xi_{l_j} \le u \text{ for } j = 1, \ldots, s\} - \Phi(u)^s| \le K \sum_{1 \le i < j \le s} |r_{ij}| \exp\left(-\frac{u^2}{1 + |r_{ij}|}\right), \quad (4.2.9)$$

where $r_{ij} = \Lambda_{l_i l_j}$ is the correlation between ξ_{l_i} and ξ_{l_j}, and K is a constant (depending on δ).

If, furthermore, $\{\xi_n\}$ is stationary with $r_i = \mathrm{Cov}(\xi_1, \xi_{1+i})$, $1 \le l_1 < \cdots < l_s \le n$, $|r_i| \le \delta < 1$ for each $i = l_j - l_k$ then

$$|P\{\xi_{l_j} \le u \text{ for } j = 1, \ldots, s\} - \Phi(u)^s| \le Kn \sum_{i=1}^{n} |r_i| \exp\left(-\frac{u^2}{1 + |r_i|}\right). \quad (4.2.10)$$

In particular, taking s = n, it follows that, for any u,

$$|P\{M_n \le u\} - \Phi(u)^n| \le Kn \sum_{i=1}^{n} |r_i| \exp\left(-\frac{u^2}{1 + |r_i|}\right). \quad (4.2.11)$$

Proof. Let $\eta_1, \ldots, \eta_n$ be independent standard normal r.v.'s, and $\Lambda_{ij}^1 = \Lambda_{ij}$, $\Lambda_{ij}^0 = 1$ or 0 according as $i = j$ or $i \ne j$. Then (4.2.9) follows from (4.2.6), and (4.2.10) is an immediate consequence of (4.2.9). □

Note that if the right-hand side of (4.2.10) is small, then (taking $s = n$), the events $\{\xi_j \le u\}$ are almost independent for $1 \le j \le n$. Further by (4.2.11) the d.f. of the maximum is then close to the value $\Phi(u)^n$ which it would take if the ξ_j were independent, and in fact (4.2.11) provides an explicit bound for the difference. Clearly in both of these observations the approximation improves as u increases. By allowing u to depend on n, $u = u_n \to \infty$ as

$n \to \infty$, the limiting behavior of $P\{M_n \le u_n\}$ may thus be discussed by use of the known i.i.d. results and the Type I limit obtained. This will be described in the next section. Similarly the explicit bound in (4.2.11) may be used in conjunction with whatever knowledge one has about rates of convergence in the i.i,d. case, to provide such knowledge for stationary sequences. This will be taken up in Section 4.6—based on a somewhat sharper version of (4.2.11).

4.3. Extremal Theory for Normal Sequences—Direct Approach

In this section it will be shown how the standard extremal results for stationary normal sequences may be obtained directly from the corresponding i.i.d. theory by use of (4.2.11). As noted earlier, these results may also be obtained simply from the general theory of Chapter 3 by verifying $D(u_n)$, $D'(u_n)$—an approach which will be taken in the next section. The central parts of the basic result are contained in the following lemma.

Lemma 4.3.1. *Let the covariances* $\{r_n\}$ *satisfy* $\sup_{n\ge 1} |r_n| = \delta < 1$.

(i) *Let* $\{u_n\}$ *be a sequence of constants such that*

$$n \sum_{j=1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1+|r_j|}\right) \to 0 \quad \textit{as } n \to \infty. \tag{4.3.1}$$

Then, if $0 \le \tau \le \infty$,

$$P\{M_n \le u_n\} \to e^{-\tau} \tag{4.3.2}$$

if and only if

$$n(1 - \Phi(u_n)) \to \tau. \tag{4.3.3}$$

(ii) *Suppose that for each* $0 < \tau < \infty$, *the constants* $v_n = v_n(\tau)$ *defined by* $1 - \Phi(v_n) = \tau/n$ *satisfy* (4.3.1) *with* v_n *replacing* u_n. *Then if* $\{u_n\}$ *is a sequence of constants,* $n(1 - \Phi(u_n)) \to \infty$ *if and only if* $P\{M_n \le u_n\} \to 0$.

Proof. Suppose (4.3.1) holds. The conditions of Corollary 4.2.4 hold so that by (4.3.1) and (4.2.11), $P\{M_n \le u_n\} \to e^{-\tau}$ if and only if $\Phi^n(u_n) \to e^{-\tau}$. But $\Phi^n(u_n)$ is simply the probability that the maximum of n i.i.d. normal r.v.'s does not exceed u_n which, by Theorem 1.5.1, converges to $e^{-\tau}$ if and only if $n(1 - \Phi(u_n)) \to \tau$, hence proving (i).

To prove (ii), note that it follows from (i) that $P\{M_n \le v_n\} \to e^{-\tau}$ ($v_n = v_n(\tau)$). Now if $n(1 - \Phi(u_n)) \to \infty$ we must have $u_n < v_n$ for sufficiently large n so that $\limsup_{n\to\infty} P\{M_n \le u_n\} \le \lim_{n\to\infty} P\{M_n \le v_n\} = e^{-\tau}$. Since this holds for arbitrary τ it follows that $P\{M_n \le u_n\} \to 0$. Conversely if $P\{M_n \le u_n\} \to 0$, since $P\{M_n \le v_n\} \to e^{-\tau}$ we must again have $u_n < v_n$ for sufficiently large n, so that then $n(1 - \Phi(u_n)) \geq n(1 - \Phi(v_n)) = \tau$, from which it follows that $n(1 - \Phi(u_n)) \to \infty$ since τ is arbitrary. □

Note that (ii) of the theorem is another version of (i) in the case $\tau = \infty$. It may be simply checked that the version (ii) is a stronger result than that in (i) when $\tau = \infty$.

The next result is a technical lemma given by Berman (1964b), showing that the covariance condition (4.1.1) ($r_n \log n \to 0$) implies (4.3.1) when u_n is such that $n(1 - \Phi(u_n))$ is bounded—so that $u_n \to \infty$ in a manner which is not too slow. This lemma will be used here, and also in the next section in verifying $D(u_n)$, $D'(u_n)$ to obtain the same results.

Before stating the lemma, we note the easily proved fact that if the covariance $r_n \to 0$ as $n \to \infty$ then $|r_n|$ cannot equal 1 for any $n \neq 0$. (For if $|r_n| = 1$ for some $n \neq 0$, it follows that ξ_1 and ξ_{n+1} are linearly related, as are also ξ_{n+1} and ξ_{2n+1}, and hence so are ξ_1, ξ_{2n+1}, so that $|r_{2n}| = 1$. In this way it follows that $|r_{kn}| = 1$ for all k, contradicting the requirement that $r_n \to 0$.) Hence it is easy to see that $|r_n|$ is actually bounded away from 1 for all $n \geq 1$, i.e. $\sup_{n \geq 1} |r_n| = \delta < 1$.

Lemma 4.3.2. *Suppose that* (4.1.1) *holds (i.e.* $r_n \log n \to 0$), *and that* $\{u_n\}$ *is a sequence of constants such that* $n(1 - \Phi(u_n))$ *is bounded. Then* (4.3.1) *holds, viz.*

$$n \sum_{j=1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1 + |r_j|} \right) \to 0 \quad \text{as } n \to \infty.$$

Proof. It appears to be technically somewhat simpler to prove this result if $n(1 - \Phi(u_n))$ actually converges to a finite limit, and we shall do so. The result will then follow as stated, since if $n(1 - \Phi(u_n)) \leq K$, and v_n is defined by $n(1 - \Phi(v_n)) = K$, then (4.3.1) will hold with v_n replacing u_n. But since clearly $u_n \geq v_n$ it follows at once that (4.3.1) holds with u_n, as asserted.

Suppose, then, that $n(1 - \Phi(u_n))$ converges to a finite limit, $n(1 - \Phi(u_n)) \to \tau$, say. By using (1.5.4), we see that

$$\text{(i)} \ \exp\left(-\frac{u_n^2}{2} \right) \sim \frac{K u_n}{n}, \tag{4.3.4}$$

$$\text{(ii)} \ u_n \sim (2 \log n)^{1/2},$$

using K as a constant whose value may change from line to line. As above, let $\delta = \sup_{n \geq 1} |r_n| < 1$, and let α be a constant such that $0 < \alpha < (1 - \delta)/(1 + \delta)$.

Split the sum in (4.3.1) into two parts, the first for $1 \leq j \leq [n^\alpha]$ and the second for $[n^\alpha] < j \leq n$. The first sum is dominated by

$$n n^\alpha \exp\left(-\frac{u_n^2}{1 + \delta} \right) = n^{1+\alpha} \left(\exp\left(-\frac{u_n^2}{2} \right) \right)^{2/(1+\delta)} \leq K n^{1+\alpha} (u_n/n)^{2/(1+\delta)}$$

$$\leq K n^{1+\alpha-2/(1+\delta)} (\log n)^{1/(1+\delta)}$$

(where (4.3.4), (i) and (ii), have been used). This tends to zero since $1 + \alpha - 2/(1 + \delta) < 0$ from the choice of α.

To deal with the second part we define

$$\delta_n = \sup_{m \geq n} |r_m|$$

and note that

$$\delta_n \log n \leq \sup_{m \geq n} |r_m| \log m \to 0 \quad \text{as } n \to \infty.$$

Now, writing $p = [n^\alpha]$, we have for the second part of (4.3.1), for large n,

$$\begin{aligned} n \sum_{j=p+1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1 + |r_j|}\right) &\leq n\delta_p \exp(-u_n^2) \sum_{j=p+1}^{n} \exp\left(\frac{u_n^2 |r_j|}{1 + |r_j|}\right) \\ &\leq n^2 \delta_p \exp(-u_n^2) \exp(u_n^2 \delta_p) \\ &\leq K\delta_p u_n^2 \exp(\delta_p u_n^2) \end{aligned}$$

by (4.3.4), (i). But by (4.3.4), (ii),

$$\delta_p u_n^2 \sim 2\delta_{[n^\alpha]} \log n = \frac{2}{\alpha} \delta_{[n^\alpha]} \log n^\alpha,$$

which tends to zero. Thus the exponential term above tends to one and the remaining product tends to zero, so that the desired result follows. □

The main distributional results for stationary normal sequences are now summarized in the following theorem.

Theorem 4.3.3. *Let $\{\xi_n\}$ be a (standardized) stationary normal sequence with covariances $\{r_n\}$ satisfying the condition $r_n \log n \to 0$. Then*

(i) *for $0 \leq \tau \leq \infty$, $P\{M_n \leq u_n\} \to e^{-\tau}$ if and only if $n(1 - \Phi(u_n)) \to \tau$,*
(ii) *$P\{a_n(M_n - b_n) \leq x\} \to \exp(-e^{-x})$, so that the Type I limit holds, where a_n and b_n have precisely the same values as in the i.i.d. case, being given by (1.7.2).*

PROOF. If $n(1 - \Phi(u_n)) \to \tau < \infty$, Lemma 4.3.2 shows that (4.3.1) holds and hence $P\{M_n \leq u_n\} \to e^{-\tau}$ by Lemma 4.3.1.

Conversely suppose that $0 \leq \tau < \infty$ and $P\{M_n \leq u_n\} \to e^{-\tau}$. For any given $\tau' < \infty$, define v_n by $1 - \Phi(v_n) = \tau'/n$. Then by what has just been proved, $P\{M_n \leq v_n\} \to e^{-\tau'}$. First choose $\tau' > \tau$. Then clearly $v_n < u_n$ for sufficiently large n so that

$$\limsup_{n \to \infty} n(1 - \Phi(u_n)) \leq \lim_{n \to \infty} n(1 - \Phi(v_n)) = \tau'.$$

Since this holds for all $\tau' > \tau$ it follows that $\limsup_{n \to \infty} n(1 - \Phi(u_n)) \leq \tau$. This gives (i) for $\tau = 0$. For $\tau > 0$ the reverse inequality for $\liminf_{n \to \infty} n(1 - \Phi(u_n))$ follows similarly by taking $\tau' < \tau$, so that $n(1 - \Phi(u_n)) \to \tau$.

Hence (i) is proved except for the case $\tau = \infty$. But this follows immediately from Lemma 4.3.1(ii), since if $1 - \Phi(v_n) = \tau/n$, (4.3.1) holds with v_n replacing u_n, by Lemma 4.3.2.

The Type I limit in (ii) also follows simply. For if $u_n = x/a_n + b_n$ with a_n, b_n as given by (1.7.2), it follows from Theorem 1.5.3 that $P\{\hat{M}_n \le u_n\} \to e^{-\tau}$ where $\tau = e^{-x}$, and $\hat{M}_n$ is the maximum of n standard normal i.d.d. random variables. Hence Theorem 1.5.1 shows that $n(1 - \Phi(u_n)) \to \tau$, so that $P\{M_n \le u_n\} \to e^{-\tau}$ by (i), and an obvious rephrasing gives (ii). □

4.4. The Conditions $D(u_n)$, $D'(u_n)$ for Normal Sequences

In this section we use the covariance condition (4.1.1) to obtain $D(u_n)$, $D'(u_n)$ for stationary normal sequences. The extremal results obtained in Section 4.3 will then also follow at once from the general theory of Chapter 3. Further, while it was natural to give the simple special derivation of the extremal results for normal sequences, it will be much more convenient to obtain subsequent results in the normal case by specializing the general theory. This will continue to be based on the dependence restrictions $D(u_n)$, $D'(u_n)$. In the following discussion the notation established above for the (standardized) stationary normal process will again be used without comment.

Lemma 4.4.1. *Let $\{u_n\}$ be a sequence of constants.*

(i) *If* $\sup_{n \ge 1} |r_n| < 1$ *and* (4.3.1) *holds then so does* $D(u_n)$.
(ii) *If, in addition* $n(1 - \Phi(u_n))$ *is bounded then* $D'(u_n)$ *holds.*
(iii) *If* $r_n \log n \to 0$ *and* $n(1 - \Phi(u_n))$ *is bounded, both* $D(u_n)$ *and* $D'(u_n)$ *hold.*

Proof. It follows from Corollary 4.2.4 (Eqn. (4.2.10)) that if $1 \le l_1 < \cdots < l_s \le n$, then the joint d.f. $F_{l_1 \ldots l_s}$ of $\xi_{l_1}, \ldots, \xi_{l_s}$ satisfies

$$|F_{l_1 \ldots l_s}(u_n) - \Phi^s(u_n)| \le Kn \sum_{j=1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1 + |r_j|}\right). \tag{4.4.1}$$

Suppose now that $1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_{p'} \le n$. Identifying $\{l_1, \ldots, l_s\}$ in turn with $\{i_1, \ldots, i_p, j_1, \ldots, j_{p'}\}$, $\{i_1, \ldots, i_p\}$ and $\{j_1, \ldots, j_{p'}\}$ we thus have

$$|F_{i_1 \ldots i_p, j_1 \ldots j_{p'}}(u_n) - F_{i_1 \ldots i_p}(u_n) F_{j_1 \ldots j_{p'}}(u_n)| \le 3Kn \sum_{j=1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1 + |r_j|}\right),$$

which tends to zero by (4.3.1). Thus $D(u_n)$ is satisfied. (In fact $\lim_{n \to \infty} \alpha_{n,l} = 0$ for each l.) Hence (i) is proved.

To prove (ii) take $s = 2$, $l_1 = 1$, $l_2 = j$ in (4.2.9) to obtain

$$|P\{\xi_1 \le u_n, \xi_j \le u_n\} - \Phi^2(u_n)| \le K|r_{j-1}| \exp\left(-\frac{u_n^2}{1 + |r_{j-1}|}\right),$$

whence, by simple manipulation (ξ_1 and ξ_j being each standard normal),

$$|P\{\xi_1 > u_n, \xi_j > u_n\} - (1 - \Phi(u_n))^2| \le K|r_{j-1}| \exp\left(-\frac{u_n^2}{1 + |r_{j-1}|}\right).$$

Thus if $n(1 - \Phi(u_n)) \le M$, then

$$n\sum_{j=2}^{[n/k]} P\{\xi_1 > u_n, \xi_j > u_n\} \le \frac{M^2}{k} + Kn\sum_{j=1}^{n} |r_j| \exp\left(-\frac{u_n^2}{1 + |r_j|}\right),$$

from which $D'(u_n)$ follows by (4.3.1), so that (ii) holds.

Finally if $r_n \log n \to 0$ and $n(1 - \Phi(u_n))$ is bounded, then Lemma 4.3.2 shows that (4.3.1) holds so that the conditions of both (i) and (ii) above are satisfied and hence $D(u_n)$, $D'(u_n)$ hold. □

The main extremal results for normal sequences given in Theorem 4.3.3 may now be seen to follow simply from the general theory of Chapter 3. For (i) of Theorem 4.3.3 is immediate when $0 \le \tau < \infty$ from Theorem 3.4.1 and Lemma 4.4.1(iii). The case $\tau = \infty$ follows at once from Corollary 3.4.2 since if v_n is defined by $1 - \Phi(v_n) = \tau/n$, $0 < \tau < \infty$, $D(v_n)$, $D'(v_n)$ hold by Lemma 4.4.1(iii) again. Finally, Theorem 4.3.3(ii) follows at once from Theorem 3.5.2 and Lemma 4.4.1(iii). (The requirement of Lemma 4.4.1(iii) that $n(1 - \Phi(u_n))$ be bounded, with $u_n = x/a_n + b_n$, is obvious from the calculations in, e.g. Theorem 1.5.3, or may be easily checked directly by considering the terms of $\log(n\phi(u_n)/u_n)$.)

4.5. Weaker Dependence Assumptions

For practical purposes the condition (4.1.1) that $r_n \log n \to 0$ is as useful and as general as is likely to be needed. In fact this condition is rather close to what is necessary for the maximum of a stationary normal sequence to behave like that of the associated independent sequence.

As we have seen, it is the convergence (4.3.1) which makes it possible to prove $D(u_n)$ and $D'(u_n)$, and one could therefore be tempted to use (4.3.1) as an indeed very weak condition. Since it is not very transparent, depending as it does on the level u_n, other conditions, which also restrict the size of r_n for large n, have been proposed occasionally. In Berman (1964b) it is shown that (4.1.1) can be replaced by

$$\sum_{n=0}^{\infty} r_n^2 < \infty \tag{4.5.1}$$

which is a special case of

$$\sum_{n=0}^{\infty} r_n^p < \infty \quad \text{for some } p > 0. \tag{4.5.2}$$

There is no implication between (4.1.1) and condition (4.5.2), but they both imply the following weak condition,

$$\frac{1}{n}\sum_{k=1}^{n}|r_k|\log k\exp(\gamma|r_k|\log k)\to 0\quad\text{as } n\to\infty \tag{4.5.3}$$

for some $\gamma > 2$, as is proved in Leadbetter *et al.* (1978) and below after Theorem 4.5.2. We now show that (4.5.3) may be used instead of (4.1.1) to obtain the relevant $D(u_n)$, $D'(u_n)$ conditions and the limit theorems for M_n.

Lemma 4.5.1. *If* $r_n \to 0$ *as* $n \to \infty$, $\{r_n\}$ *satisfies* (4.5.3), *and if* $n(1 - F(u_n))$ *is bounded, then* (4.3.1) *holds.*

PROOF. As in Lemma 4.3.2 we may (and do) assume that $n(1 - F(u_n))$ actually converges to a finite limit τ. Using the notation in the proof of Lemma 4.3.2, let $\delta = \sup_{n\geq 1}|r_n| < 1$, take $\beta = 2/\gamma$ and let α be a constant such that $0 < \alpha < \min(\beta, (1-\delta)/(1+\delta))$.

Split the sum in (4.3.1) into three parts, the first for $1 \leq j \leq [n^\alpha]$, the second for $[n^\alpha] < j \leq [n^\beta]$ and the third for $[n^\beta] < j \leq n$. The first sum tends to zero as in Lemma 4.3.2.

Writing $\delta_n = \sup_{m\geq n}|r_m|$, $p = [n^\alpha]$, and $q = [n^\beta]$ and using (4.3.4), i.e.

$$\exp\left(-\frac{u_n^2}{2}\right) \sim \frac{Ku_n}{n} \sim \frac{K(2\log n)^{1/2}}{n},$$

we have for the second part of (4.3.1),

$$\begin{aligned} n\sum_{k=p+1}^{q}|r_k|\exp\left(-\frac{u_n^2}{1+|r_k|}\right) &\leq n^{1+\beta}\exp(-u_n^2)\exp(u_n^2\delta_p)\\ &\leq Kn^{\beta-1}u_n^2\exp(\delta_p u_n^2)\\ &\leq Kn^{\beta-1}u_n^2 n^{2\delta_p}\end{aligned}$$

which obviously tends to zero, since $\beta < 1$, and $\delta_p \log p \to 0$.

Finally, for the last part of (4.3.1) we have, again using (4.3.4),

$$\begin{aligned} n\sum_{k=q+1}^{n}|r_k|\exp\left(-\frac{u_n^2}{1+|r_k|}\right) &\leq Kn\sum_{k=q+1}^{n}|r_k|\left(\frac{u_n}{n}\right)^{2/(1+|r_k|)}\\ &\leq Kn^{-1}\log n\sum_{k=q+1}^{n}|r_k|\exp(2|r_k|\log n).\end{aligned}$$

For $k > q$ we have $\log k \geq \beta \log n$, and hence this expression is not larger than

$$\begin{aligned} &Kn^{-1}\sum_{k=q+1}^{n}|r_k|\log k\exp(2\beta^{-1}|r_k|\log k)\\ &\qquad\leq Kn^{-1}\sum_{k=1}^{n}|r_k|\log k\exp(\gamma|r_k|\log k).\end{aligned}$$

By (4.5.3) this tends to zero as $n \to \infty$, which concludes the proof of (4.3.1). □

The main results proved under (4.1.1) may now be extended as follows.

Theorem 4.5.2. *Let $\{\xi_n\}$ be a (standardized) stationary normal sequence, with covariances $r_n \to 0$ as $n \to \infty$, and satisfying* (4.5.3) *for some $\gamma > 2$. Then*

(i) *if $\{u_n\}$ is a sequence of constants such that $n(1 - \Phi(u_n))$ is bounded, then $D(u_n)$, $D'(u_n)$ both hold,*
(ii) *for $0 \le \tau \le \infty$, $P\{M_n \le u_n\} \to e^{-\tau}$ if and only if $n(1 - \Phi(u_n)) \to \tau$,*
(iii) *$M_n = \max(\xi_1, \dots, \xi_n)$ has the Type I limiting distribution*

$$P\{a_n(M_n - b_n) \le x\} \to \exp(-e^{-x})$$

where a_n and b_n have the same values as in the i.i.d. case, being given by (1.7.2).

Proof. The same arguments used in Lemma 4.4.1 and Theorem 4.3.3 may be applied with the obvious modifications. □

A few remarks may be helpful here to provide insight into condition (4.5.3). Define, for each positive x, the set $\theta_n(x) = \{k; 1 \le k \le n, |r_k| \log k > x\}$ and let $v_n(x)$ be the number of elements in $\theta_n(x)$. Consider the following condition (which we shall see is slightly stronger than (4.5.3)),

$$\begin{aligned} &n^{-1}\sum_{k=1}^{n} |r_k| \log k \to 0 \qquad \text{as } n \to \infty, \quad \text{and} \\ &v_n(K) = O(n^\eta) \quad \text{for some } K > 0, \eta < 1, \end{aligned} \tag{4.5.4)$'$}$$

and the equivalent condition

$$\begin{aligned} v_n(\varepsilon) &= o(n) \qquad \text{for all } \varepsilon > 0, \quad \text{and} \\ v_n(K) &= O(n^\eta) \quad \text{for some } K > 0, \eta < 1. \end{aligned} \tag{4.5.4)$''$}$$

Obviously (4.1.1) implies (4.5.4)$'$. Further, if

$$\sum_{k=1}^{\infty} |r_k|^p < \infty$$

for some $p > 0$ then, since

$$\sum_{k=1}^{\infty} |r_k|^p \ge \sum_{k \in \theta_n(x)} |r_k|^p \ge v_n(x)\left(\frac{x}{\log n}\right)^p,$$

it follows that $v_n(x) = O((\log n)p)$. In particular, we see that also (4.5.1) and (4.5.2) imply (4.5.4)$''$, so that both (4.1.1) and (4.5.2) are stronger than (4.5.4)$'$ and (4.5.4)$''$. The following lemma states that (4.5.4)$'$ or (4.5.4)$''$ imply (4.5.3) and consequently that both (4.1.1) and (4.5.2) imply (4.5.3).

Lemma 4.5.3. *If $r_n \to 0$ as $n \to \infty$, then* (4.5.4)′ *and* (4.5.4)″ *are equivalent and both imply* (4.5.3).

PROOF. It is easily seen that (4.5.4)′ and (4.5.4)″ are equivalent so we need only show that (4.5.4)′ implies (4.5.3). We have

$$n^{-1}\sum_{k=1}^{n} |r_k| \log k \exp(\gamma|r_k| \log k) = n^{-1}\sum_{\substack{k=1\\ k\notin\theta_n(K)}}^{n} |r_k| \log k \exp(\gamma|r_k| \log k)$$
$$+ n^{-1}\sum_{k\in\theta_n(K)} |r_k| \log k \exp(\gamma|r_k| \log k), \tag{4.5.5}$$

and proceed to estimate the sums on the right separately, assuming that (4.5.4)′ holds. Now

$$n^{-1}\sum_{\substack{k=1\\ k\notin\theta_n(K)}}^{n} |r_k| \log k \exp(\gamma|r_k| \log k) \le \exp(\gamma K) n^{-1}\sum_{k=1}^{n} |r_k| \log k \to 0,$$
$$n \to \infty,$$

by the first part of (4.5.4)′. Since we assume that $r_n \to 0$, there is an integer N such that $\gamma|r_k| < (1-\eta)/2$ for $k \ge N$. Hence

$$n^{-1}\sum_{\substack{k\in\theta_n(K)\\ k\ge N}} |r_k| \log k \exp(\gamma|r_k| \log k) \le n^{-1} v_n(K) \log n\, n^{(1-\eta)/2},$$

which tends to zero as $n \to \infty$, by the second part of (4.5.4)′. Since N is fixed, $n^{-1}\sum_{k=1}^{N} |r_k| \log k \exp(\gamma|r_k| \log k) \to 0$, and it follows that also the second term of the right-hand side of (4.5.5) tends to zero, and thus that (4.5.3) is satisfied. □

4.6. Rate of Convergence

As was seen in Section 2.4 for independent standard normal r.v.'s, the convergence of $P\{M_n \le u_n\} = \Phi(u_n)^n$ is quite slow if, e.g. $u_n = u_n(x) = x/a_n + b_n$. In particular, it is of the order $(\log\log n)^2/\log n$ if a_n and b_n are chosen as in Theorem 1.5.3, and of the order $1/\log n$ if the "best" a_n's and b_n's are used. One should, of course, not expect more rapid convergence in the dependent case. In fact, the proof of Theorem 4.3.3 essentially depends on first comparing $P\{M_n \le u_n\}$ with $\Phi(u_n)^n$ and then using the limit theorem for independent normal variables to conclude that $\Phi(u_n)^n \to e^{-\tau}$, for $\tau = e^{-x}$. We shall presently show that, if $\sum_{t=1}^{n} |r_t|$ does not increase too rapidly, the error in the approximation of $P\{M_n \le u\}$ by $\Phi(u)^n$ tends to zero as a certain power of n. It is then immediate that the speed of convergence of $P\{M_n \le u_n\}$ to $e^{-\tau}$ is determined by the approximation of $\Phi(u_n)^n$ by $e^{-\tau}$, and hence is the same as in the independent case.

However, since $\Phi(u_n)^n$ can be easily calculated with high accuracy, the bound on the size of $P\{M_n \le u\} - \Phi(u)^n$ which we shall derive, also has substantial interest in itself; it shows how $P\{M_n \le u\}$, for M_n the maximum of dependent normal variables, can be approximated with reasonable accuracy.

As a starting point for the estimation we shall use the equation (4.2.4), specialized to the case $\Lambda_{ij}^1 = r_{j-i}$, $\Lambda_{ij}^0 = 0$, $i \neq j$ (and hence with $\int_0^1 F'(h)\,dh = P\{M_n \le u\} - \Phi(u)^n$), which gives the equation

$$P\{M_n \le u\} - \Phi(u)^n = \int_0^1 \sum_{1 \le i < j \le n} r_{j-i} \int_{-\infty}^{\mathbf{u}'} \cdots \int f_h(y_i = y_j = u)\, d\mathbf{y}'\, dh. \tag{4.6.1}$$

Here $f_h(y_i = y_j = u)$ is the function obtained by putting $y_i = y_j = u$ in the density function of n normal r.v.'s with means zero, variances one, and covariances hr_k, and the remaining variables are integrated over $(-\infty, u]$. It is useful to write this expression in a slightly different form. Let

$$\begin{aligned}\phi_\rho(u) &= \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho u^2 + u^2)\right\} \\ &= \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(-\frac{u^2}{1+\rho}\right)\end{aligned}$$

be the standard bivariate normal density taken at the point (u, u), and let

$$f_h(\mathbf{y}' \mid y_i = y_j = u) = f_h(y_i = y_j = u)/\phi_{hr_{j-i}}(u)$$

be the conditional density, given that the ith and jth variables are equal to u, in the n-dimensional normal distribution introduced above. The equation (4.6.1) can then be written as

$$\begin{aligned}&P\{M_n \le u\} - \Phi(u)^n \\ &\quad = \sum_{1 \le i < j \le n} r_{j-i} \int_0^1 \phi_{hr_{j-i}}(u) \int_{-\infty}^{\mathbf{u}'} \cdots \int f_h(\mathbf{y}' \mid y_i = y_j = u)\, d\mathbf{y}'\, dh,\end{aligned} \tag{4.6.2}$$

and we clearly have that

$$0 \le \int_{-\infty}^{\mathbf{u}'} \cdots \int f_h(\mathbf{y}' \mid y_i = y_j = u)\, d\mathbf{y}' \le 1, \tag{4.6.3}$$

where the second inequality was one of the main steps in the proof of Theorem 4.2.1.

Before proceeding to the estimations, we shall introduce some notation and constants for later use. The constants are fairly involved, and could easily be simplified by making less accurate approximations, but that would make them less suited for numerical use. Let $\rho = \sup\{0, r_1, r_2, \ldots\}$ and, in case $\rho > 0$, let v be the number of t's such that $r_t = \rho$. We will throughout this section, without further comment, assume that the supremum is attained, so that $v \geq 1$. In particular, this is the case if $r_t \to 0$ as $t \to \infty$, and then also $v < \infty$. If $\rho = 0$ let $v \leq \infty$ be the number of nonzero r_t's. For the second-order terms, define ρ' to be the supremum for $t \geq 1$ of the r_t's which satisfy $r_t \neq \rho$, if this quantity is positive, and zero otherwise, and let

$$\varepsilon = 2\left(\frac{1}{1+\rho'} - \frac{1}{1+\rho}\right) = \frac{2(\rho - \rho')}{(1+\rho)(1+\rho')}. \tag{4.6.4}$$

Next define

$$c'(\rho) = \frac{(1+\rho)^{3/2}}{(1-\rho)^{1/2}}, \qquad c''(\rho) = \frac{(2-\rho)(1+\rho)}{1-\rho}, \qquad c(\rho) = c'(\rho)(4\pi)^{-\rho/(1+\rho)}, \tag{4.6.5}$$

and put $\delta = \sup\{|r_t|; t \geq 1, r_t \neq \rho\}$.

The main factor, R_n, in the bounds has a slightly different appearance in the two cases (i) $\rho > 0$ or $\rho = 0$, $v < \infty$, and (ii) $\rho = 0$, $v = \infty$, and in addition depends on a constant K, which will be introduced below. Define

$$R_n = c(\rho, K, v)n^{-(1-\rho)/(1+\rho)}\left(\log\frac{n}{K}\right)^{-\rho/(1+\rho)}(1+\rho_n)$$
$$\text{if } \rho > 0 \text{ or } \rho = 0,\ v < \infty, \tag{4.6.6$'$}$$

$$R_n = c(K, \delta)\left(\frac{1}{n}\right)\log\left(\frac{n}{K}\right)\sum_{t=0}^{n}|r_t| \quad \text{if } \rho = 0,\ v = \infty. \tag{4.6.6$''$}$$

Here

$$c(\rho, K, v) = c(\rho)K^{2/(1+\rho)}v, \qquad c(K, \delta) = 4K^2(1-\delta^2)^{-1/2}, \tag{4.6.7}$$

and ρ_n is defined by $\rho_n = 0$ for $\rho = 0$, and

$$\rho_n = C\sum{}'|r_t|n^{-\varepsilon}\left(\log\frac{n}{K}\right)^{1+\varepsilon/2}, \tag{4.6.8}$$

if $\rho > 0$, with

$$C = K^{\varepsilon}2^{(2-\rho')/(1+\rho')}(4\pi)^{\varepsilon/2}(1-\rho)^{1/2}(1+\rho)^{-3/2-1/(1+\rho')}(1-\delta^2)^{-1/2}v^{-1},$$

and with Σ' signifying that the summation is over all t in $\{1, 2, \ldots, n\}$ for which $r_t \neq \rho$.

Lemma 4.6.1. *Let* $u > 0$, *suppose* $r \neq 0$, $|r| < 1$, *write* $\rho = \max(0, r)$, *and let* c, c', c'' *be given by* (4.6.5). *Then*

(i) $$\frac{1}{2\pi u^2 r}\left\{c'(r)\exp\left(-\frac{u^2}{1+r}\right) - e^{-u^2}\right\}\Big/\left\{1 + \frac{c''(\rho)}{u^2}\right\}$$
$$\leq \int_0^1 \phi_{hr}(u)\,dh \leq \frac{1}{2\pi u^2 r}\left\{c'(r)\exp\left(-\frac{u^2}{1+r}\right) - e^{-u^2}\right\}.$$

Suppose that furthermore $u \geq 1$. Then

(ii) $$0 \leq \int_0^1 \phi_{hr}(u)\, dh \leq 2^{(2+\rho)/(1+\rho)} c(\rho)|r|^{-1}\{(1 - \Phi(u))/u^\rho\}^{2/(1+\rho)},$$

and, if $r \leq \rho'$ for some constant $0 \leq \rho' < 1$, then

(iii) $$0 \leq \int_0^1 \phi_{hr}(u)\, dh$$
$$\leq 2^{(2+\rho')/(1+\rho')}(4\pi)^{-\rho'/(1+\rho')}(1 - r^2)^{-1/2}\{(1 - \Phi(u))u\}^{2/(1+\rho')}.$$

PROOF. By partial integration

$$2\pi \int_0^1 \phi_{hr}(u)\, dh = \int_0^1 (1 - h^2r^2)^{-1/2} \exp\left(-\frac{u^2}{1 + hr}\right) dh$$
$$= \frac{1}{u^2 r}\left\{c'(r) \exp\left(-\frac{u^2}{1 + r}\right) - \exp(-u^2)\right\}$$
$$-\frac{1}{u^2}\int_0^1 \frac{(2 - hr)(1 + hr)^{1/2}}{(1 - hr)^{3/2}} \exp\left(-\frac{u^2}{1 + hr}\right) dh, \quad (4.6.9)$$

and the second inequality in (i) follows at once, since the last integral in (4.6.9) is positive. Moreover,

$$(2 - hr)(1 + hr)^{1/2}(1 - hr)^{-3/2} \leq c''(\rho)(1 - h^2r^2)^{-1/2},$$

as is easily checked, and hence

$$\int_0^1 \frac{(2 - hr)(1 + hr)^{1/2}}{(1 - hr)^{3/2}} \exp\left(-\frac{u^2}{1 + hr}\right) dh \leq 2\pi c''(\rho) \int_0^1 \phi_{hr}(u)\, dh. \quad (4.6.10)$$

Inserting (4.6.10) into (4.6.9) we obtain

$$\left\{1 + \frac{c''(\rho)}{u^2}\right\} \int_0^1 \phi_{hr}(u)\, dh \geq \frac{1}{2\pi u^2 r}\left\{c'(r) \exp\left(-\frac{u^2}{1 + r}\right) - e^{-u^2}\right\},$$

which proves the first inequality in (i).

To prove (ii) we will use the inequalities

$$\sqrt{2\pi}(1 - \Phi(u)) > \frac{\exp(-u^2/2)}{u}\frac{u^2}{1 + u^2} \geq \frac{\exp(-u^2/2)}{2u} \quad (4.6.11)$$

for $u \geq 1$. Thus, if $r = \rho > 0$, by part (i),

$$\int_0^1 \phi_{hr}(u)\, dh \leq \frac{c'(r)}{2\pi u^2 r} \exp\left(-\frac{u^2}{1 + r}\right)$$
$$\leq 2^{(2+\rho)/(1+\rho)}(4\pi)^{-\rho/(1+\rho)}|r|^{-1}c'(\rho)\{(1 - \Phi(u))/u^\rho\}^{2/(1+\rho)},$$

and similarly, for $r < 0$, $\rho = 0$,

$$\int_0^1 \phi_{hr}(u)\, dh \leq \frac{1}{2\pi u^2 |r|} \exp(-u^2) \leq \frac{4c'(0)}{|r|}(1 - \Phi(u))^2,$$

and hence (ii) holds in either case. Finally, it is immediate that, for $u \geq 1$,

$$\int_0^1 \phi_{hr}(u)\, dh \leq \frac{1}{2\pi(1-r^2)^{1/2}} \exp\left(-\frac{u^2}{1+\rho'}\right)$$
$$\leq 2^{(2+\rho')/(1+\rho')}(4\pi)^{-\rho'/(1+\rho')}(1-r^2)^{-1/2}\{(1-\Phi(u))u\}^{2/(1+\rho')},$$

by (4.6.11), which proves (iii). □

The main lemma now follows easily. We will only consider a restricted range of u-values (which may even be empty for small n). The remaining range of u's of interest to us is easier to treat, as shown in the proof of Theorem 4.6.3 below.

Lemma 4.6.2. *Suppose that for some constant* $K > 0$,

$$n(1-\Phi(u)) \leq K, \tag{4.6.12}$$

and that $1 \leq u \leq 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$. *Then*

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \phi_{hr_t}(u)\, dh \leq 4R_n, \tag{4.6.13}$$

for R_n *given by* (4.6.6)′ *and* (4.6.6)″, *in the separate cases.*

Proof. First, by (4.6.11) and (4.6.12),

$$\frac{n}{\sqrt{2\pi}} \frac{\exp(-u^2/2)}{2u} \leq K,$$

i.e.

$$\log\frac{n}{K} \leq \frac{u^2}{2}\left\{1 + \frac{1}{u^2}\log 8\pi u^2\right\} \leq 2u^2, \tag{4.6.14}$$

for $u \geq 1$.

Now, suppose that $\rho > 0$. Using Lemma 4.6.1(ii) to bound summands with $r_t = \rho$ and Lemma 4.6.1(iii) for the remaining summands, we have that

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \phi_{hr_t}(u)\, dh \leq n2^{(2+\rho)/(1+\rho)}c(\rho)\{(1-\Phi(u))/u^\rho\}^{2/(1+\rho)}v$$
$$+ n2^{(2+\rho')/(1+\rho')}(4\pi)^{-\rho'/(1+\rho')}\sum{}' \frac{|r_t|}{(1-r_t^2)^{1/2}}$$
$$\times \{(1-\Phi(u))u\}^{2/(1+\rho')}, \tag{4.6.15}$$

where $\sum'$ denotes summation over all t in $\{1, \ldots, n\}$ such that $r_t \neq \rho$. Since $n(1-\Phi(u)) \leq K$ and $(\frac{1}{2}\log n/K)^{1/2} \leq u \leq 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$ by assumption and (4.6.14), we have that

$$\left\{\frac{1-\Phi(u)}{u^\rho}\right\}^{2/(1+\rho)} \leq 2^{\rho/(1+\rho)}\left(\frac{K}{n}\right)^{2/(1+\rho)}\left(\log\frac{n}{K}\right)^{-\rho/(1+\rho)}$$

and that

$$\{(1-\Phi(u))u\}^{2/(1+\rho')} \le 2^{2/(1+\rho')}(1+\rho)^{-1/(1+\rho')}\left(\frac{K}{n}\right)^{2/(1+\rho')}\left(\log\frac{n}{K}\right)^{1/(1+\rho')}$$

Inserting this into (4.6.15) we obtain, with $\delta = \sup\{|r_t|;\ t \ge 1,\ r_t \ne \rho\}$, $\varepsilon = 2(\rho-\rho')(1+\rho)^{-1}(1+\rho')^{-1}$, and C as in (4.6.8),

$$n\sum_{t=1}^{n}|r_t|\int_0^1 \phi_{hr_t}(u)\,dh \le 4c(\rho)K^{2/(1+\rho)}\,n^{-(1-\rho)/(1+\rho)}\left(\log\frac{n}{K}\right)^{-\rho/(1+\rho)} \quad \nu$$

$$\times\left\{1 + C\sum{}' |r_t| n^{-\varepsilon}\left(\log\frac{n}{K}\right)^{1+\varepsilon/2}\right\},$$

and comparing with (4.6.6)′, this proves (4.6.13) for the case $\rho > 0$.

Next, suppose $\rho = 0$, $\nu < \infty$, so that by Lemma 4.6.1(ii),

$$n\sum_{t=1}^{n}|r_t|\int_0^1 \phi_{hr_t}(u)\,dh \le 4n\sum_{\substack{t=1\\ r_t\ne 0}}^{n}\frac{|r_t|}{|r_t|}(1-\Phi(u))^2 \le \frac{4K^2\nu}{n},$$

which shows that (4.6.13) holds also in this case.

Finally, suppose $\rho = 0$, and $\nu \le \infty$. Then, using Lemma 4.6.1(iii), similar calculations show that

$$n\sum_{t=1}^{n}|r_t|\int_0^1 \phi_{hr_t}(u)\,dh \le 4n\sum_{t=1}^{n}\frac{|r_t|}{(1-r_t^2)^{1/2}}\{(1-\Phi(u))u\}^2$$

$$\le 16K^2(1-\delta^2)^{-1/2}\sum_{t=1}^{n}|r_t|\frac{1}{n}\log\frac{n}{K},$$

proving (4.6.13) for the case $\rho = 0$, $\nu = \infty$. □

The rate of convergence to zero of $P\{M_n \le u\} - \Phi(u)^n$ now follows easily from (4.6.2), (4.6.3), and Lemma 4.6.2. To obtain efficient bounds we will, as in Lemma 4.6.2, restrict the domain of variation of u by requiring that $n(1-\Phi(u)) \le K$, for some fixed $K > 0$, or equivalently that $u \ge u_n$, where u_n is the solution to the equation $n(1-\Phi(u_n)) = K$. According to Theorem 1.5.1, this implies that if $\hat{M}_n$ is the maximum of the first n variables in an independent standard normal sequence, then

$$P\{\hat{M}_n \le u_n\} \to e^{-K} \quad \text{as } n \to \infty, \tag{4.6.16}$$

and conversely that if $\{u_n\}$ satisfies (4.6.16) then $n(1-\Phi(u_n)) \to K$. Moreover, if $P\{M_n \le u_n\} - P\{\hat{M}_n \le u_n\} = P\{M_n \le u_n\} - \Phi(u_n)^n \to 0$, then, of course, the same equivalence holds for $\hat{M}_n$ replaced by M_n.

Thus, since the bounds for the rate of convergence will be proved for $u \ge u_n$, they will apply to the upper part of the range of variation of $P\{M_n \le u\}$, and by taking K large, an arbitrarily large part of this range is covered, at the cost of a poorer bound.

Theorem 4.6.3. *Let $\{\xi_t\}$ be a (standardized) stationary normal sequence, suppose that $u \geq 1$, and that (4.6.12) holds, i.e. that*

$$n(1 - \Phi(u)) \leq K,$$

for some constant K with $n/K \geq e$. Then

$$|P\{M_n \leq u\} - \Phi(u)^n| \leq 4R_n, \tag{4.6.17}$$

with R_n given by (4.6.6). More explicitly, writing $\Delta_n = |P\{M_n \leq u\} - \Phi(u)^n|$, if $\rho = \max\{0, r_1, r_2, \ldots\} > 0$, or $\rho = 0$ and $v = \#\{t \geq 1; r_t \neq 0\} < \infty$, then

$$\Delta_n \leq cn^{-(1-\rho)/(1+\rho)}\left(\log \frac{n}{K}\right)^{-\rho/(1+\rho)}(1 + \rho_n),$$

with $c = 4c(\rho, K, v)$ and ρ_n given by (4.6.7), (4.6.8), and if $\rho = 0$, $v = \infty$, then

$$\Delta_n \leq c\frac{1}{n}\log\frac{n}{K}\sum_{t=0}^{n}|r_t|,$$

with $c = 4c(K, \delta)$ given by (4.6.7).

PROOF. By (4.6.2) and (4.6.3)

$$\begin{aligned}
|P\{M_n \leq u\} - \Phi(u)^n| &\leq \sum_{1 \leq s < t \leq n} |r_{s-t}| \int_0^1 \phi_{hr_{s-t}}(u)\,dh \\
&\leq n\sum_{t=1}^{n} |r_t| \int_0^1 \phi_{hr_t}(u)\,dh,
\end{aligned}$$

and it follows from Lemma 4.6.2 that (4.6.17) holds for u satisfying (4.6.12) and $1 \leq u \leq 2(1 + \rho)^{-1/2}(\log n/K)^{1/2}$.

To complete the proof we will show that (4.6.17), rather trivially, is satisfied also for $u > 2(1 + \rho)^{-1/2}(\log n/K)^{1/2}$. In fact,

$$\begin{aligned}
|P\{M_n \leq u\} - \Phi(u)^n| &= |P\{M_n > u\} - (1 - \Phi(u)^n)| \\
&\leq P\{M_n > u\} + (1 - \Phi(u)^n) \\
&\leq 2n(1 - \Phi(u)),
\end{aligned} \tag{4.6.18}$$

by Boole's inequality. Since $1 - \Phi(u) \leq \phi(u)/u$, we have for

$$u \geq 2(1 + \rho)^{-1/2}\left(\log \frac{n}{K}\right)^{1/2} \geq 1,$$

that

$$\begin{aligned}
1 - \Phi(u) &\leq (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}\left(2(1 + \rho)^{-1/2}\left(\log\frac{n}{K}\right)^{1/2}\right)^2\right\}\left(\log\frac{n}{K}\right)^{-1/2} \\
&\leq (2\pi)^{-1/2}\left(\frac{K}{n}\right)^{2/(1+\rho)}\left(\log\frac{n}{K}\right)^{-1/2},
\end{aligned}$$

and hence, by (4.6.18),

$$|P\{M_n \le u\} - \Phi(u)^n| \le 2(2\pi)^{-1/2} n \left(\frac{K}{n}\right)^{2/(1+\rho)} \left(\log \frac{n}{K}\right)^{-1/2} \le 4R_n,$$

by straightforward calculation. □

Thus, by the theorem, if $\rho > 0$ or $\rho = 0$, $v < \infty$ and $\sum_{t=1}^{n} |r_t|$ does not grow too rapidly, the rate of convergence is at least of the order

$$n^{-(1-\rho)/(1+\rho)}(\log n)^{-\rho/(1+\rho)}.$$

This is the right order, at least if $\{\xi_t\}$ is m-dependent. In fact, if $r_t = 0$ for $|t| > m$ for some positive integer m, and $n(1 - \Phi(u_n)) \to K > 0$, then if $\rho > 0$,

$$P\{M_n \le u_n\} - \Phi(u_n)^n \sim e^{-K} R_n, \tag{4.6.19}$$

and if $\rho = 0$, then

$$P\{M_n \le u_n\} - \Phi(u_n)^n \sim -e^{-K} R_n; \tag{4.6.20}$$

see Rootzén (1982). Comparing (4.6.19) and (4.6.20) with the bounds of Theorem 4.6.3, we see that for $\rho > 0$ or $\rho = 0$, $v < \infty$, the bounds asymptotically are too large by a factor $4e^K$. Here, the factor 4 is due to inaccuracies in the estimates (4.6.11) and (4.6.14), and can be easily reduced by restricting the range of u further, while the factor e^K is due to the estimate (4.6.3).

If $\rho = 0$, $v = \infty$, and $\sum_{t=0}^{\infty} |r_t| < \infty$, the bound given by Theorem 4.6.3 is of the order

$$\frac{1}{n} \log n.$$

It seems unlikely that this is the correct order, but the loss does not seem important, since clearly the rate of convergence cannot be better than $1/n$, in general.

Finally, as mentioned above, it is an easy corollary to the theorem that the convergence to the limiting double exponential distribution is as slow as for an independent normal sequence. For example, if a_n, b_n and C_1, C_2 are given by (2.4.9) and (2.4.10), then, under a suitable condition on the growth of $\sum_{t=1}^{n} |r_t|$,

$$\begin{aligned} 0 < C_1 &\le \liminf_{n\to\infty} \left\{ \sup_x \log n |P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})| \right\} \\ &\le \limsup_{n\to\infty} \left\{ \sup_x \log n |P\{a_n(M_n - b_n) \le x\} - \exp(-e^{-x})| \right\} \\ &\le C_2. \end{aligned}$$

and the order $1/\log n$ of convergence cannot be improved by choosing other normalizing constants than a_n, b_n. In particular, for

$$a_n = (2 \log n)^{1/2},$$

$$b_n = a_n - (2a_n)^{-1}\{\log \log n + \log 4\pi\},$$

the approximation is

$$P\{a_n(M_n - b_n) \leq x\} - \exp(-e^{-x}) \sim \tfrac{1}{16} e^{-x} \exp(-e^{-x}) \frac{(\log \log n)^2}{\log n}.$$

The reader is referred to Rootzén (1982) for the exact condition on $\sum_{t=1}^{n} |r_t|$ and the proofs of these assertions, and for further aspects of the rate of convergence of extremes of stationary normal sequences.

CHAPTER 5
Convergence of the Point Process of Exceedances, and the Distribution of kth Largest Maxima

In this chapter we return to the general situation and notation of Chapter 3 and consider the points j (regarded as "time instants") at which the general stationary sequence $\{\xi_j\}$ exceeds some given level u. These times of exceedance are stochastic in nature and may be viewed as a point process. Since exceedances of very high levels will be rare, one may suspect that this point process will take on a Poisson character at such levels. An explicit theorem along these lines will be proved and the asymptotic distributions of kth largest values (order statistics) obtained as corollaries. Generalizations of this theorem yield further results concerning joint distributions of kth largest values. The formal definition and simple properties of point processes which will be needed are given in the appendix.

5.1. Point Processes of Exceedances

If u is a given "level" we say that the (stationary) sequence $\{\xi_n\}$ has an *exceedance* of u at j if $\xi_j > u$. Such j may be regarded as "instants of time", and the exceedances therefore as events occurring randomly in time, i.e. as a *point process* (cf. Appendix).

We shall be concerned with such exceedances for (typically) increasing levels and will define such a point process, N_n, say, for each of a sequence $\{u_n\}$ of levels. Since the u_n will be typically high for large n, the exceedances will tend to be rarer and we shall find it convenient to normalize the "time" axis to keep the expected number of exceedances approximately constant. For our purposes the simple scale change by the factor n will suffice. Specifically we define for each n a process $\eta_n(t)$ at the points $t = j/n, j = 1, 2, \ldots$ by

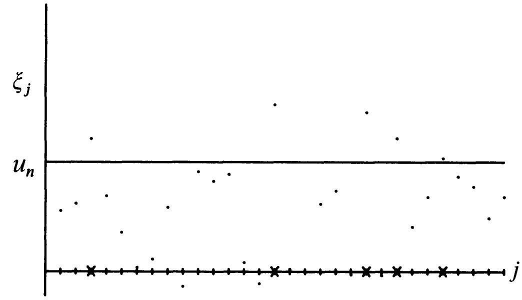


Figure 5.1.1. Point process of exceedances.

$\eta_n(j/n) = \xi_j$. Then η_n has an exceedance of u_n at j/n whenever $\{\xi_k\}$ has an exceedance at j. Hence, while exceedances of u_n may be lost as u_n increases, this will be balanced by the fact that the points j/n become more dense. Indeed, the expected number of exceedances by η_n in the interval $(0, 1]$ is clearly $nP\{\xi_1 > u_n\}$ which tends to a finite value τ, if u_n is chosen by (3.4.2).

Our first task will be to show that (under $D(u_n)$, $D'(u_n)$ conditions) the exceedances of u_n by η_n become Poisson in character as n increases (actually in the full sense of distributional convergence for point processes described in the appendix). In particular, this will mean that the number, $N_n(B)$, say, of exceedances of u_n by η_n in the (Borel) set B, will have an asymptotic Poisson distribution. From this we may simply obtain the asymptotic distribution of the kth largest among $\xi_1, \ldots, \xi_n$, and thus generalize Theorems 2.2.1 and 2.2.2. The Poisson result will be proved in the next section and the distributional corollaries in Section 5.3.

It will also be of interest to generalize Theorems 2.3.1 and 2.3.2, giving joint distributions of kth largest values. This will require an extension of our convergence result to involve exceedances of several levels simultaneously, as will be seen in subsequent sections.

5.2. Poisson Convergence of High-Level Exceedances

In the following theorem we shall first consider exceedances of u_n by η_n on the unit interval $(0, 1]$ rather than the whole positive axis, since we can then use less restrictive assumptions, and still obtain the corollaries concerning the distributions of kth maxima.

Theorem 5.2.1. (i) *Let $\tau > 0$ be fixed and suppose that $D(u_n)$, $D'(u_n)$ hold for the stationary sequence $\{\xi_n\}$ with $u_n = u_n(\tau)$ satisfying* (3.4.2). *Let $\eta_n(j/n) = \xi_j$, $j = 1, 2, \ldots$; $n = 1, 2, \ldots$, and let N_n be the point process on the unit interval* $(0, 1]$ *consisting of the exceedances of u_n by η_n in that interval, (i.e. the points j/n, $1 \le j \le n$, for which $\eta_n(j/n) = \xi_j > u_n$). Then N_n converges in distribution to a Poisson process N on* $(0, 1]$ *with parameter τ, as $n \to \infty$.*

(ii) *Suppose that, for each* $\tau > 0$, *there exists a sequence* $\{u_n(\tau)\}$ *satisfying* (3.4.2), *and that* $D(u_n(\tau))$, $D'(u_n(\tau))$ *hold for all* $\tau > 0$. *Then for any fixed* τ, *the result of* (i) *holds for the entire positive axis in place of the unit interval, i.e. the point process* N_n *of exeedances of* $u_n(\tau)$ *by* η_n, *converges to a Poisson process* N *on* $(0, \infty)$ *with parameter* τ.

PROOF. By Theorem A.1 to prove part (i) it is sufficient to show that

(a) $E(N_n((c, d])) \to E(N((c, d])) = \tau(d - c)$ as $n \to \infty$ for all $0 < c < d \le 1$, and
(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\} = \exp(-\tau m(B))$ (m being Lebesgue measure) for all B of the form $\bigcup_1^r (c_j, d_j]$, $0 < c_1 < d_1 < c_2 < \cdots < c_r < d_r \le 1$.

Here (a) is immediate since

$$E(N_n((c, d])) = ([nd] - [nc])(1 - F(u_n)) \sim n(d - c)\tau/n = \tau(d - c).$$

To show (b) we note that for $0 < c < d \le 1$,

$$P\{N_n((c, d]) = 0\} = P\{M(I_n) \le u_n\},$$

where $I_n = \{[nc] + 1, \ldots, [nd]\}$. Now I_n contains v_n integers where $v_n = [nd] - [nc] \sim n(d - c)$ as $n \to \infty$. Thus, by Corollary 3.6.4 with $\theta = d - c < 1$,

$$P\{N_n((c, d]) = 0\} \to \exp(-\tau(d - c)) \quad \text{as } n \to \infty. \tag{5.2.1}$$

Now, let $B = \bigcup_1^r (c_i, d_i]$, where $0 < c_1 < d_1 < c_2 < d_2 < \cdots < c_r < d_r \le 1$. Then, writing E_j for the set of integers $\{[nc_j] + 1, [nc_j] + 2, \ldots, [nd_j]\}$, it is readily checked that

$$\begin{aligned} P\{N_n(B) = 0\} &= P\left(\bigcap_{j=1}^r \{M(E_j) \le u_n\}\right) \\ &= \prod_{j=1}^r P\{N_n((c_j, d_j]) = 0\} \\ &\quad + \left\{P\left(\bigcap_{j=1}^r \{M(E_j) \le u_n\}\right) - \prod_{j=1}^r P\{M(E_j) \le u_n\}\right\}. \end{aligned}$$

By (5.2.1), the first term converges, as $n \to \infty$, to $\prod_{j=1}^r \exp(-\tau(d_j - c_j)) = \exp(-\tau m(B))$ (where m denotes Lebesgue measure). On the other hand, by Lemma 3.2.2 it is readily seen that the modulus of the remaining difference of terms does not exceed $(r - 1)\alpha_{n,[n\lambda]}$ where $\lambda = \min_{1 \le j \le r-1}(c_{j+1} - d_j)$. But by $D(u_n)$, (cf. (3.2.3)), $\alpha_{n,[n\lambda]} \to 0$ as $n \to \infty$ so that (b) follows. Hence (i) of the theorem holds.

The conclusion (ii) follows by exactly the same proof except that we use Theorem 3.6.3 instead of Corollary 3.6.4, taking v_n to be $u_n(\theta\tau)$ (thus satisfying $D(v_n)$, $D'(v_n)$, $n(1 - F(v_n)) \to \theta\tau$ by assumption) where now we may have $\theta > 1$. Correspondingly c, d and c_i, d_i are no longer restricted to be no greater than 1. □

It is of interest to note that the conclusion of (i) applies to any interval of unit length, so that the exceedances in *any* such interval become Poisson in character. But if the assumption of (ii) is not made, it may possibly not happen that the exceedances become Poisson on the *entire* axis (or on an interval of greater than unit length).

Corollary 5.2.2. *Under the conditions of* (i) *of the theorem, if* $B \subset (0, 1]$ *is any Borel set whose boundary has Lebesgue measure zero,* $(m(\partial B) = 0)$, *then*

$$P\{N_n(B) = r\} \to \frac{\exp(-\tau m(B))(\tau m(B))^r}{r!}, \qquad r = 0, 1, 2, \ldots$$

The joint distribution of any finite number of variables $N_n(B_1), \ldots, N_n(B_k)$ *corresponding to disjoint* B_i, *(with* $m(\partial B_i) = 0$ *for each i) converges to the product of corresponding Poisson probabilities.*

PROOF. This follows at once since $(N_n(B_1), \ldots, N_n(B_k)) \xrightarrow{d} (N(B_1), \ldots, N(B_k))$ (as noted in the appendix) when $N_n \xrightarrow{d} N$. □

It should be noted that the above results obviously apply very simply to stationary normal sequences satisfying appropriate covariance conditions (e.g. (4.1.1)).

5.3. Asymptotic Distribution of *k*th Largest Values

The following results may now be obtained from Corollary 5.2.2 generalizing the conclusions of Theorems 2.2.1 and 2.2.2.

Theorem 5.3.1. *Let* $M_n^{(k)}$ *denote the kth largest of* $\xi_1, \ldots, \xi_n$, $(M_n^{(1)} = M_n)$, *where k is a fixed integer. Let* $\{u_n\}$ *be a real sequence and suppose that* $D(u_n)$, $D'(u_n)$ *hold. If* (3.4.2) *holds for some fixed finite* $\tau \geq 0$, *then*

$$P\{M_n^{(k)} \leq u_n\} \to e^{-\tau} \sum_{s=0}^{k-1} \frac{\tau^s}{s!} \quad \text{as } n \to \infty. \tag{5.3.1}$$

Conversely if (5.3.1) *holds for some integer k, so does* (3.4.2), *and hence* (5.3.1) *holds for all k.*

PROOF. As previously, we identify the event $\{M_n^{(k)} \leq u_n\}$ with the event that no more than $(k - 1)$ of $\xi_1, \ldots, \xi_n$ exceed u_n, i.e. with $\{N_n((0, 1]) \leq k - 1\}$, so that

$$P\{M_n^{(k)} \leq u_n\} = \sum_{s=0}^{k-1} P\{N_n((0, 1]) = s\}. \tag{5.3.2}$$

If (3.4.2) holds the limit on the right of (5.3.1) follows at once by Corollary 5.2.2.

Conversely if (5.3.1) holds but (3.4.2) does not, there is some $\tau' \neq \tau$ $0 \leq \tau' \leq \infty$, and a sequence $\{n_j\}$ such that $n_j(1 - F(u_{n_j})) \to \tau'$. Now a brief examination of the proof of Theorem 5.2.1 (cf. also the remark at the end of Section 3.4) shows that if (3.4.2) is not assumed for all n but just for a sequence $\{n_j\}$ then N_{n_j} has a Poisson limit. If $\tau' < \infty$, replacing τ by τ' in the argument above, we thus have

$$P\{M_{n_j}^{(k)} \leq u_{n_j}\} \to e^{-\tau'} \sum_{s=0}^{k-1} \frac{\tau'^s}{s!}.$$

But this contradicts (5.3.1) since the function $e^{-x} \sum_{s=0}^{k-1} x^s/s!$ is strictly decreasing in $x \geq 0$ and hence 1 – 1. Thus $\tau' < \infty$ is not possible. But $\tau' = \infty$ cannot hold either since as in (3.4.4) we would have

$$P\{M_{[n/k]} \leq u_n\} \leq 1 - [n/k](1 - F(u_n)) + S_{n,k},$$

which would be negative for large n by the finiteness of $\limsup_{n\to\infty} S_{n,k}$ implied by $D'(u_n)$, at least for some appropriately chosen large k. Hence (3.4.2) holds as asserted. □

The case $k = 1$ of this theorem is just Theorem 3.4.1 again. Of course Theorem 3.4.1 is used in the proof of Theorem 5.2.1 and hence of Theorem 5.3.1. The following corollary covers the case $\tau = \infty$.

Corollary 5.3.2. *Suppose that, for arbitrarily large τ, there exists a sequence $\{v_n = v_n(\tau)\}$ satisfying $D(v_n)$, $D'(v_n)$ and such that $n(1 - F(v_n)) \to \tau$. If, for a sequence $\{u_n\}$, $n(1 - F(u_n)) \to \infty$, then $P\{M_n^{(k)} \leq u_n\} \to 0$ for all k.*

Conversely if $P\{M_n^{(k)} \leq u_n\} \to 0$ for some k, then $n(1 - F(u_n)) \to \infty$ and $P\{M_n^{(k)} \leq u_n\} \to 0$ for all k.

Proof. If $n(1 - F(u_n)) \to \infty$, $\tau < \infty$ and v_n is chosen as above then $P\{M_n^{(k)} \leq v_n\} \to e^{-\tau}$ by the theorem. But clearly $v_n > u_n$ for sufficiently large n, so that

$$\limsup_{n\to\infty} P\{M_n^{(k)} \leq u_n\} \leq \lim_{n\to\infty} P\{M_n^{(k)} \leq v_n\} = e^{-\tau}.$$

Since this holds for arbitrarily large $\tau < \infty$, $P\{M_n^{(k)} \leq u_n\} \to 0$ as asserted.

Conversely, if $P\{M_n^{(k)} \leq u_n\} \to 0$ for some k, it follows, since $M_n \geq M_n^{(k)}$, that $P\{M_n \leq u_n\} \to 0$ so that $n(1 - F(u_n)) \to \infty$ by Corollary 3.4.2. □

The following obvious corollary also holds.

Corollary 5.3.3. *Theorem* 5.3.1 *holds if the assumption* (*or conclusion*) *that* $\{u_n\}$ *satisfy* (3.4.2), *is replaced by either of the assumptions* (*conclusions*)

$$P\{M_n \leq u_n\} \to e^{-\tau}, \qquad P\{\hat{M}_n \leq u_n\} \to e^{-\tau}$$

(*where $\hat{M}_n$ as usual denotes the maxima for the associated independent sequence*). *Correspondingly the assumption* (*or conclusion*) $n(1 - F(u_n)) \to \infty$

in Corollary 5.3.2 *may be replaced by either* $P\{M_n \le u_n\} \to 0$ *or* $P\{\hat{M}_n \le u_n\} \to 0$.

Proof. The statements regarding Theorem 5.3.1 follow at once from Theorems 3.4.1 and 1.5.1. Those for Corollary 5.3.2 follow from Corollary 3.4.2 and Theorem 1.5.1. □

Theorem 5.3.4. *Let* $a_n > 0$, b_n *be constants for* $n = 1, 2, \ldots$, *and* G *a nondegenerate d.f., and suppose that* $D(u_n)$, $D'(u_n)$ *hold for all* $u_n = x/a_n + b_n$, $-\infty < x < \infty$. *If*

$$P\{a_n(M_n - b_n) \le x\} \to G(x), \tag{5.3.3}$$

then for each $k = 1, 2, \ldots$

$$P\{a_n(M_n^{(k)} - b_n) \le x\} \to G(x) \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!} \tag{5.3.4}$$

where $G(x) > 0$ (*zero where* $G(x) = 0$).

Conversely if (5.3.4) *holds for some* k, *so does* (5.3.3) *and hence* (5.3.4) *holds for all* k. *Further, the result remains true if* $\hat{M}_n$ *replaces* M_n *in* (5.3.3).

Proof. For $G(x) > 0$ the result follows from the first part of Corollary 5.3.3 by writing $u_n = x/a_n + b_n$, $\tau = -\log G(x)$. The case $G(x) = 0$ will follow from the second part of Corollary 5.3.3 provided it can be shown that for arbitrarily large τ there is a sequence $\{v_n\}$ satisfying $D(v_n)$, $D'(v_n)$ such that $n(1 - F(v_n)) \to \tau$. But since G is continuous (being an extreme value d.f.) there exists x_0 such that $G(x_0) = e^{-\tau}$ from which it is easily seen that $x_0/a_n + b_n$ provides an appropriate choice of v_n. □

5.4. Independence of Maxima in Disjoint Intervals

It would clearly be natural to extend the theorems of Chapter 3 to deal with the joint behaviour of maxima in disjoint intervals. We shall do so here—demonstrating asymptotic independence under an appropriate generalization of the $D(u_n)$ condition, and then use this to obtain a Poisson result for exceedances of several levels considered jointly. This, in turn, will lead to the asymptotic joint distributions of various quantities of interest, such as two or more $M_n^{(k)}$, and their locations, as $n \to \infty$.

As in (2.3.1), consider r levels $u_n^{(k)}$ satisfying

$$n(1 - F(u_n^{(k)})) \to \tau_k, \tag{5.4.1}$$

where $u_n^{(1)} \ge u_n^{(2)} \ge \cdots \ge u_n^{(r)}$ as in Chapter 2 and consequently $\tau_1 \le \tau_2 \le \cdots \le \tau_r$.

It is intuitively clear that we shall need to extend the $D(u_n)$ condition to involve the r values $u_n^{(k)}$, and we do so as follows.

The condition $D_r(\mathbf{u}_n)$ *will be said to hold for the stationary sequence* $\{\xi_j\}$ *if, for each choice of* $\mathbf{i} = (i_1, \ldots, i_p)$, $\mathbf{j} = (j_1, \ldots, j_{p'})$, $1 \le i_1 < i_2 < \cdots < i_p < j_1 < j_2 < \cdots < j_{p'} \le n, j_1 - i_p \ge l$, *we have* (*using obvious notation*)

$$|F_{\mathbf{i},\mathbf{j}}(\mathbf{v}, \mathbf{w}) - F_{\mathbf{i}}(\mathbf{v})F_{\mathbf{j}}(\mathbf{w})| \le \alpha_{n,l}, \tag{5.4.2}$$

where $\mathbf{v} = (v_1, \ldots, v_p)$, $\mathbf{w} = (w_1, \ldots, w_{p'})$, *the* v_i *and* w_j *each being any choice of the r values* $u_n^{(1)}, \ldots, u_n^{(r)}$, *and where* $\alpha_{n,l_n} \to 0$ *for some sequence* $l_n = o(n)$.

The condition $D_r(\mathbf{u}_n)$ extends $D(u_n)$ in an obvious way (and clearly implies $D(u_n^{(k)})$ for each k) and will be convenient even though its full strength will not quite be needed for our purposes. It will not be necessary to define an extended $D'(u_n)$ condition, since we shall simply need to assume that $D'(u_n^{(k)})$ holds separately for each $k = 1, 2, \ldots, r$.

The next result extends Lemma 3.2.2 (with slight changes of notation).

Lemma 5.4.1. *With the above notation, if* $D_r(\mathbf{u}_n)$ *holds, if* n, s, k *are fixed integers, and* $E_1, \ldots, E_s$ *subintervals of* $\{1, \ldots, n\}$ *such that* E_i *and* E_j *are separated by at least* l *when* $i \ne j$, *then*

$$\left| P\left(\bigcap_{j=1}^{s} \{M(E_j) \le u_{n,j}\}\right) - \prod_{j=1}^{s} P\{M(E_j) \le u_{n,j}\} \right| \le (s-1)\alpha_{n,l}$$

where for each j, $u_{n,j}$ *is any one of* $u_n^{(1)}, \ldots, u_n^{(r)}$.

Proof. This is proved in exactly the same manner as Lemma 3.2.2 and the details will therefore not be repeated here. □

In the following discussion we shall consider a fixed number s of disjoint subintervals $J_1, J_2, \ldots, J_s$ of $\{1, \ldots, n\}$ such that J_k $(= J_{n,k})$ has $v_{n,k} \sim \theta_k n$ elements, where θ_k are fixed positive constants with $\sum_{k=1}^{s} \theta_k \le 1$. By slightly strengthening the assumptions, we may also allow $\sum_{k=1}^{s} \theta_k > 1$, and let $J_1, J_2, \ldots, J_s$ be more arbitrary finite disjoint intervals of positive integers. Note that the intervals J_k do increase in size with n, but remain disjoint, and their total number s is fixed.

The following results then hold. In the proofs, details will be omitted where they duplicate arguments given in Chapter 3.

Theorem 5.4.2. (i) *Let* $J_1, J_2, \ldots, J_s$ *be disjoint subintervals of* $\{1, 2, \ldots, n\}$ *as defined above,* J_k *having* $v_{n,k} \sim \theta_k n$ *members, for fixed positive* $\theta_1, \theta_2, \ldots, \theta_s$, $(\sum_1^s \theta_k \le 1)$. *Suppose that the stationary sequence* $\{\xi_j\}$ *satisfies* $D_r(\mathbf{u}_n)$ *where the levels* $u_n^{(1)} \ge u_n^{(2)} \ge \cdots \ge u_n^{(r)}$ *satisfy* (5.4.1). *Then*

$$P\left(\bigcap_{k=1}^{s} \{M(J_k) \le u_{n,k}\}\right) - \prod_{k=1}^{s} P\{M(J_k) \le u_{n,k}\} \to 0 \quad \text{as } n \to \infty \tag{5.4.3}$$

for any choice of $u_{n,k}$ *from* $u_n^{(1)}, \ldots, u_n^{(r)}$ *for each* k.

(ii) *For fixed s, m, let* $J_1, J_2, \ldots, J_s$ *be disjoint subintervals of the positive integers* $1, 2, \ldots, mn$, *where* $J_k(=J_{n,k})$ *has* $v_{n,k} \sim \theta_k n$ *members,* $\theta_1, \ldots, \theta_s$ *being fixed constants* ($\sum \theta_k \le m$). *Let* $u_n(\tau)$ *satisfy* $(1 - F(u_n(\tau)) \sim \tau/n$ *for each* $\tau > 0$, *let* $\tau_1 \le \tau_2 \le \cdots \le \tau_r$ *be fixed, and suppose that* $D_r(\mathbf{u}_n)$ *holds for* $\mathbf{u}_n = (u_n(m\tau_1), \ldots, u_n(m\tau_r))$. *Then* (5.4.3) *holds.*

Proof. Let I_k denote the first $v_{n,k} - l_n$ elements of J_k, and I_k^* the remaining l_n, where l_n is chosen as in $D_r(\mathbf{u}_n)$. (These I_k, I_k^* are different from those in Chapter 3, but are so named since they play a similar role.) By familiar calculation we have

$$0 \le P\left(\bigcap_{k=1}^{s} \{M(I_k) \le u_{n,k}\}\right) - P\left(\bigcap_{k=1}^{s} \{M(J_k) \le u_{n,k}\}\right)$$

$$\le \sum_{k=1}^{s} P\{M(I_k^*) > u_{n,k}\} \le s\rho_n, \tag{5.4.4}$$

where $\rho_n = \max_{1 \le k \le s} P\{M(I_k^*)) > u_{n,k}\}$. Further

$$\left| P\left(\bigcap_{k=1}^{s} \{M(I_k) \le u_{n,k}\}\right) - \prod_{k=1}^{s} P\{M(I_k) \le u_{n,k}\} \right| \le (s-1)\alpha_{n,l_n} \tag{5.4.5}$$

by Lemma 5.4.1, and

$$0 \le \prod_{k=1}^{s} P\{M(I_k) \le u_{n,k}\} - \prod_{k=1}^{s} P\{M(J_k) \le u_{n,k}\}$$

$$\le \prod_{k=1}^{s} (P\{M(J_k) \le u_{n,k}\} + \rho_n) - \prod_{k=1}^{s} P\{M(J_k) \le u_{n,k}\}$$

$$\le (1 + \rho_n)^s - 1 \tag{5.4.6}$$

since $\prod_{k=1}^{s} (y_k + \rho_n) - \prod_{k=1}^{s} y_k$ is increasing in each y_k when $\rho_n > 0$. Now

$$\rho_n = \max_{1 \le k \le s} P\{M(I_k^*) \ge u_{n,k}\} \le \max_{1 \le k \le s} l_n(1 - F(u_{n,k}))$$

which tends to zero by (5.4.1) since $l_n = o(n)$. Hence, by (5.4.4), (5.4.5), and (5.4.6), the left-hand side of (5.4.3) is dominated in absolute value by $s\rho_n + (s-1)\alpha_{n,l_n} + (1 + \rho_n)^s - 1$ which tends to zero, completing the proof of part (i) of the theorem.

Part (ii) follows along similar lines and only the required modifications will be indicated. First I_k is defined to be the first $v_{n,k} - l_{nm}$ elements of J_k, and I_k^* the remaining l_{nm}. If $u_{n,k} = u_n^{(j)} = u_n(\tau_j)$, write $v_{n,k} = u_{nm}(m\tau_j)$. Then (5.4.4) holds with $v_{n,k}$ replacing $u_{n,k}$ as does (5.4.5) on replacing also α_{n,l_n} by $\alpha_{nm,l_{nm}}$ (using Lemma 5.4.1 with mn for n). It then follows as above that

$$P\left(\bigcap_{k=1}^{s} \{M(J_k) \le v_{n,k}\}\right) - \prod_{k=1}^{s} P\{M(J_k) \le v_{n,k}\} \to 0. \tag{5.4.7}$$

Now again, if $u_{n,k} = u_n^{(j)} = u_n(\tau_j)$ we have $v_{n,k} = u_{nm}(m\tau_j)$ and

$$n(1 - F(u_{n,k})) \to \tau_j$$

$$n(1 - F(v_{n,k})) \sim \frac{n(m\tau_j)}{nm} \to \tau_j$$

so that Lemma 3.6.1(i) gives

$$P\{M(J_k) \le v_{n,k}\} - P\{M(J_k) \le u_{n,k}\} \to 0.$$

By writing

$$A_k = \{M(J_k) \le v_{n,k}\}, \qquad B_k = \{M(J_k) \le u_{n,k}\}$$

and using the obvious inequalities

$$\begin{aligned} |P(\textstyle\bigcap A_i) - P(\bigcap B_i)| &\le P(\textstyle\bigcap A_i - \bigcap B_i) + P(\bigcap B_i - \bigcap A_i) \\ &\le P\{\textstyle\bigcup (A_i - B_i)\} + P\{\bigcup (B_i - A_i)\} \end{aligned}$$

we may approximate the first term in (5.4.7) by replacing $v_{n,k}$ by $u_{n,k}$. A slight extension of a corresponding calculation in (i) above shows a similar approximation for the second term of (5.4.7), from which it follows that (5.4.7) tends to zero when $u_{n,k}$ replaces $v_{n,k}$, so that (5.4.3) follows, as desired. □

Note that the proof of this theorem is somewhat simpler than that, e.g. in Lemmas 3.3.1 and 3.3.2. This occurs because we assume (5.4.1) whereas the corresponding assumption was not made there. We could dispense with (5.4.1) here also with a corresponding increase in complexity, but since we assume (5.4.1) in the sequel, we use it here also.

Corollary 5.4.3. (i) *If, in addition to the assumptions of part* (i) *of the theorem, we suppose that $D'(u_n^{(k)})$ holds for each $k = 1, 2, \ldots, r$, then (for $\sum_1^s \theta_k \le 1$),*

$$P\{M(J_k) \le u_{n,k}, k = 1, 2, \ldots, s\} \to \exp\left(-\sum_{k=1}^{s} \theta_k \tau_k'\right),$$

where τ_k' is that one of $\tau_1, \ldots, \tau_r$ corresponding to $u_{n,k}$, i.e. such that $n(1 - F(u_{n,k})) \to \tau_k'$.

(ii) *If in addition to the assumptions of part* (ii) *of the theorem, $D'(v_n)$ holds with $v_n = u_n(\theta_k \tau_k')$, $1 \le k \le s$, then the conclusion holds for these arbitrary positive constants θ_k.*

PROOF. (i) follows by Corollary 3.6.4 which shows that

$$P\{M(J_k) \le u_{n,k}\} \to \exp(-\theta_k \tau_k'), \qquad 1 \le k \le s.$$

For (ii), the same limit holds for each k by Theorem 3.6.3. For (noting $u_{n,k} = u_n(\tau_k')$), $D'(v_n)$ holds with $v_n = u_n(\theta_k \tau_k')$, as also does $D(v_n)$, since the assumption $D_r(\mathbf{u}_n)$ made implies $D(u_n(m\tau_k'))$ which in turn (since $\theta_k \le m$) implies $D(u_n(\theta_k \tau_k'))$ by Lemma 3.6.2(iv). □

It is easy to check that $D_r(\mathbf{u}_n)$ holds for normal sequences under the standard covariance conditions and hence Corollary 5.4.3 may be applied.

Theorem 5.4.4. *Let $\{\xi_n\}$ be a stationary normal sequence with zero means, unit variances and covariance sequence $\{r_n\}$. Suppose $\{u_n(\tau)\}$ satisfies $n(1-\Phi(u_n(\tau))) \to \tau$ for each $\tau > 0$, let $0 \le \tau_1 < \tau_2 < \cdots < \tau_r$. Suppose that $r_n \to 0$ as $n \to \infty$ and*

$$n \sum_{j=1}^{n} |r_j| \exp\left(-\frac{(u_n(m\tau_r))^2}{1+|r_j|}\right) \to 0 \quad \text{as } n \to \infty, \tag{5.4.8}$$

for some $m \ge 1$ (which will hold, in particular, for any such m if $r_n \log n \to 0$, by Lemma 4.3.2). Then $D_r(\mathbf{u}_n)$ holds with $\mathbf{u}_n = (u_n(m\tau_1), \ldots, u_n(m\tau_r))$ as does $D'(u_n(\tau))$ when $0 < \tau < m\tau_r$. It follows that if $\theta_1, \ldots, \theta_s$ are positive constants with $\sum_{i=1}^{s} \theta_i \le m$, then

$$P\{M(J_k) \le u_{n,k}, k = 1, 2, \ldots, s\} \to \exp\left(-\sum_{k=1}^{s} \theta_k \tau'_k\right) \tag{5.4.9}$$

where J_k are as in Theorem 5.4.2(ii) and $u_{n,k}$, τ'_k as in Corollary 5.4.3.

Proof. With the notation of (5.4.2) we may identify the ξ_j of Theorem 4.2.1 with $\xi_{i_1}, \ldots, \xi_{i_p}, \xi_{j_1}, \ldots, \xi_{j_{p'}}$, here and the η_j of that theorem with $\xi'_{i_1}, \ldots, \xi'_{i_p}$, $\xi'_{j_1}, \ldots, \xi'_{j_{p'}}$, such that $\xi'_{i_1}, \ldots, \xi'_{i_p}$ have the same joint distribution as $\xi_{i_1}, \ldots, \xi_{i_p}$, but are independent of $\xi'_{j_1}, \ldots, \xi'_{j_{p'}}$ which in turn have the same joint distribution as $\xi_{j_1}, \ldots, \xi_{j_{p'}}$. Then Corollary 4.2.2 gives

$$|F_{\mathbf{i},\mathbf{j}}(\mathbf{v}, \mathbf{w}) - F_{\mathbf{i}}(\mathbf{v})F_{\mathbf{j}}(\mathbf{w})| \le K \sum_{\substack{1 \le s \le p \\ 1 \le t \le p'}} |r_{i_s - j_t}| \exp\left(-\frac{u_n^2}{1+|r_{i_s-j_t}|}\right),$$

where $u_n = \min(v_1, \ldots, v_p, w_1, \ldots, w_{p'})$ and $v_1, \ldots, v_p, w_1, \ldots, w_{p'}$ are chosen from $u_n(m\tau_1), \ldots, u_n(m\tau_r)$. Replacing u_n by $u_n(m\tau_r)$ ($\le u_n$ for sufficiently large n) and using the fact that for each j there are at most n terms containing r_j, we obtain

$$|F_{\mathbf{i},\mathbf{j}}(\mathbf{v}, \mathbf{w}) - F_{\mathbf{i}}(\mathbf{v})F_{\mathbf{j}}(\mathbf{w})| \le Kn \sum_{j=1}^{n} |r_j| \exp\left(-\frac{(u_n(m\tau_r))^2}{1+|r_j|}\right)$$

which tends to zero by (5.4.8) so that $D_r(\mathbf{u}_n)$ holds as claimed ($\alpha_{n,l}$ being independent of l, in fact).

Now if $0 \le \tau \le m\tau_r$, $u_n(\tau) \ge u_n(m\tau_r)$ for sufficiently large n so that (5.4.8) holds with τ replacing $m\tau_r$ and hence Lemma 4.4.1(ii) shows that $D'(u_n(\tau))$ holds as required.

Finally if $\theta_k > \theta$, $\sum_1^s \theta_k \le m$, $D'(v_n)$ holds with $v_n = u_n(\theta_k \tau'_k)$ by the previous statement of the theorem since $\tau = \theta_k \tau'_k \le m\tau_r$. Since we have shown above that $D_r(\mathbf{u}_n)$ holds with $\mathbf{u}_n = (u_n(m\tau_1), \ldots, u_n(m\tau_r))$, (5.4.9) follows from Corollary 5.4.3(ii). □

The following result, generalizing Theorem 3.6.6, also follows from Corollary 5.4.3.

Theorem 5.4.5. *Let $\{\xi_n\}$ be a stationary sequence, $a_n > 0$, b_n, constants, and suppose that*

$$P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x) \quad \text{as } n \to \infty$$

for some nondegenerate d.f. G. Suppose that $D_r(\mathbf{u}_n)$, $D'(u_n^{(k)})$ hold for all sequences of the form $u_n^{(k)} = x_k/a_n + b_n$, and let $J_k = J_{n,k}$, $k = 1, 2, \ldots, r$ be disjoint subintervals of $\{1, \ldots, mn\}$, J_k containing $v_{n,k}$ integes where $v_{k,n} \sim \theta_k n$, m being a fixed integer and $\theta_1 > 0, \ldots, \theta_r > 0$, $\sum_1^r \theta_k \le m$. Then

$$P\{a_n(M(J_k) - b_n) \le x_k, k = 1, 2, \ldots, r\} \to \prod_{k=1}^{r} G^{\theta_k}(x_k).$$

Proof. This follows from Corollary 5.4.3 in a similar way to the proof of Theorem 3.6.6, identifying τ_k with $-\log G(x_k)$, $u_{n,k} = x_k/a_n + b_n$, $u_n(\theta_k \tau_k) = y_k/a_n + b_n$ where $\theta_k \tau_k = -\log G(y_k)$, and similarly for $u_n(m\tau_k)$. $\square$

5.5. Exceedances of Multiple Levels

It is natural to consider exceedances of the levels $u_n^{(1)}, \ldots, u_n^{(r)}$ by η_n, ($\eta_n(j/n) = \xi_j$ as before), as a vector of point processes. While this may be achieved abstractly, we shall here, as an obvious aid to intuition represent them as occurring along fixed horizontal lines $L_1, \ldots, L_r$ in the plane—exceedances of $u_n^{(k)}$ being represented as points on L_k. This will show the structure imposed by the fact that an exceedance of $u_n^{(k)}$ is automatically an exceedance of $u_n^{(k+1)}, \ldots, u_n^{(r)}$, as illustrated in Figure 5.5.1. In the figure, (a) shows the levels and values of η_n from which one can see the exceeded levels, while (b) marks the points of exceedance of each level along the fixed lines $L_1, \ldots, L_r$.

To pursue this a little further, the diagram (b) represents the exceedances of all the levels as points in the plane. That is, we may regard them as a *point process in the plane* if we wish. To be sure, all the points lie only on certain horizontal lines, and a point on any L_k has points directly below it on all lower L_k, but nevertheless the positions are stochastic, and do define a two-dimensional point process, which we denote by N_n.

We may apply convergence theory to this sequence $\{N_n\}$ of point processes and obtain the joint distributional results as a consequence. The position of the lines $L_1, \ldots, L_r$ does not matter as long as they are fixed and in the indicated order. From our previous theory each one-dimensional point process, on a given L_k, $N_n^{(k)}$, say, will become Poisson under appropriate conditions. The two-dimensional process indicated will not become Poisson in the plane as is intuitively clear, in view of the structure described. However, the exceedances $N_n^{(k)}$ on L_k form successively more "severely thinned" versions of

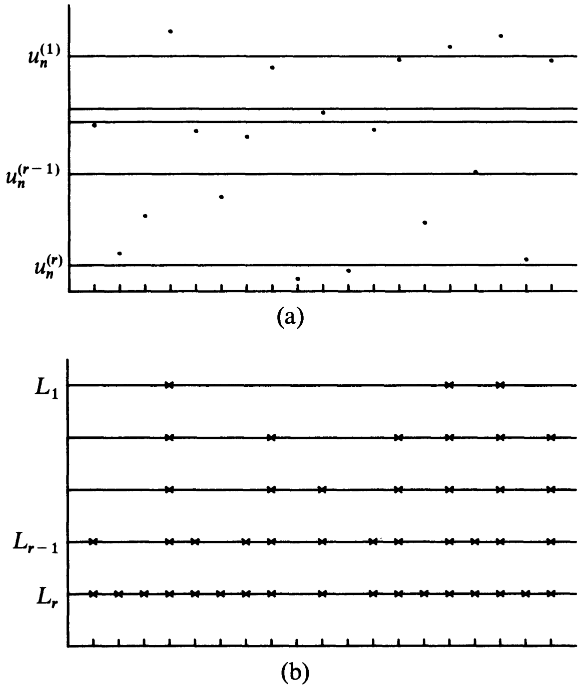


Figure 5.5.1. (a) Levels and values of $\eta_n(t)$. (b) Representation in plane (fixed L_k).

$N_n^{(r)}$ as k decreases. Of course, these are not thinnings formed by *independent* removal of events, except in the limit where the Poisson process $N^{(k)}$ on L_k may be obtained from $N^{(k+1)}$ on L_{k+1} by independent thinning, as will become apparent.

More specifically, we define the point process N in the plane, which will turn out to be the appropriate limiting point process, as follows.

Let $\{\sigma_{1j}; j = 1, 2, \ldots\}$ be the points of a Poisson process with parameter τ_r on L_r. Let $\beta_j, j = 1, 2, \ldots,$ be i.i.d. random variables, independent also of the Poisson process on L_r, taking values 1,2, ..., r with probabilities

$$P\{\beta_j = s\} = \begin{cases} (\tau_{r-s+1} - \tau_{r-s})/\tau_r, & s = 1, 2, \ldots, r-1, \\ \tau_1/\tau_r, & s = r, \end{cases}$$

i.e. $P\{\beta_j \geq s\} = \tau_{r-s+1}/\tau_r$ for $s = 1, 2, \ldots, r$.

For each j, place points $\sigma_{2j}, \sigma_{3j}, \ldots, \sigma_{\beta_j j}$ on the $\beta_j - 1$ lines L_{r-1}, $L_{r-2}, \ldots, L_{r-\beta_j+1}$, vertically above σ_{1j}, to complete the point process N. Clearly the probability that a point appears on L_{r-1} above σ_{1j} is just $P\{\beta_j \geq 2\} = \tau_{r-1}/\tau_r$ and the deletions are independent, so that $N^{(r-1)}$ is obtained as an independent thinning of the Poisson process $N^{(r)}$. Hence $N^{(r-1)}$ is a Poisson process (cf. Appendix) with intensity $\tau_r(\tau_{r-1}/\tau_r) = \tau_{r-1}$, as expected. Similarly $N^{(k)}$ is obtained as an independent thinning of $N^{(k+1)}$ with deletion probability $1 - \tau_k/\tau_{k+1}$, all $N^{(k)}$ being Poisson.

We may now give the main result.

Theorem 5.5.1. (i) *Suppose that $D_r(\mathbf{u}_n)$ holds, and that $D'(u_n^{(k)})$ holds for $1 \le k \le r$, where the $u_n^{(k)}$ satisfy* (5.4.1). *Then the point process N_n of exceedances of the levels $u_n^{(1)}, \ldots, u_n^{(r)}$ (represented as above on the lines $L_1, \ldots, L_r$) converges in distribution to the limiting point process N, as point processes on* $(0, 1] \times R$.

(ii) *If further for $0 \le \tau < \infty$, $u_n(\tau)$ satisfies $n(1 - F(u_n(\tau))) \to \tau$, $D_r(\mathbf{u}_n)$ holds for $\mathbf{u}_n = (u_n(m\tau_1), \ldots, u_n(m\tau_r))$, all $m \ge 1$, and $D'(u_n(\tau))$ holds for all $\tau > 0$, then N_n converges to N, as point processes on the entire right half plane, i.e. on* $(0, \infty) \times R$.

Proof. Again by Theorem A.1 it is sufficient to show that

(a) $E(N_n(B)) \to E(N(B))$ for all sets B of the form $(c, d] \times (\gamma, \delta]$, $\gamma < \delta$, $0 < c < d$, where $d \le 1$ or $d < \infty$, respectively, and

(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for all sets B which are finite unions of disjoint sets of this form.

Again (a) follows readily. If $B = (c, d] \times (\gamma, \delta]$ intersects any of the lines, let these be $L_s, L_{s+1}, \ldots, L_t$, $(1 \le s \le t \le r)$. Then $N_n(B) = \sum_{k=s}^t N_n^{(k)}((c, d])$, $N(B) = \sum_{k=s}^t N^{(k)}((c, d])$, and the number of points j/n in $(c, d]$ is $[nd] - [nc]$ so that

$$\begin{aligned} E(N_n(B)) &= ([nd] - [nc]) \sum_{k=s}^{t} (1 - F(u_n^{(k)})) \\ &\sim n(d - c) \sum_{k=s}^{t} \left(\frac{\tau_k}{n} + o\left(\frac{1}{n}\right)\right) \to (d - c) \sum_{k=s}^{t} \tau_k \end{aligned}$$

which is clearly just $E(N(B))$.

To show (b) we must prove that $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for sets B of the form $B = \bigcup_1^m C_k$ with disjoint $C_k = (c_k, d_k] \times (\gamma_k, \delta_k]$. Clearly, we may discard any set C_k which does not intersect any of the lines $L_1, \ldots, L_r$. By considering intersections and differences of the intervals $(c_k, d_k]$, we may change these and write B in the form $\bigcup_{k=1}^s (c_k, d_k] \times E_k$, where $(c_k, d_k]$ are disjoint and E_k is a finite union of semiclosed intervals. Thus

$$\{N_n(B) = 0\} = \bigcap_{k=1}^{s} \{N_n(F_k) = 0\}, \tag{5.5.1}$$

where $F_k = (c_k, d_k] \times E_k$. If the lowest L_j intersecting F_k is L_{m_k}, then, by the thinning property, clearly

$$\{N_n(F_k) = 0\} = \{N_n^{(m_k)}((c_k, d_k]) = 0\}. \tag{5.5.2}$$

But this is just the event $\{M([c_k n], [d_k n]) \le u_{m_k}\}$, so that Corollary 5.4.3, parts (i) and (ii), respectively, gives

$$P\{N_n(B) = 0\} \to \exp\left(-\sum_{k=1}^{s} (d_k - c_k)\tau_{m_k}\right) = P\{N(B) = 0\},$$

since (5.5.1) and (5.5.2) clearly also hold with N instead of N_n, and the result follows. □

Corollary 5.5.2. *Let $\{\xi_n\}$ satisfy the conditions of Theorem 5.5.1(i) or (ii), and let $B_1, \ldots, B_s$ be Borel subsets of the unit interval, or the positive real line, respectively, whose boundaries have zero Lebesgue measure. Then for integers $m_j^{(k)}$,*

$$P\{N_n^{(k)}(B_j) = m_j^{(k)}, j = 1, \ldots, s, k = 1, \ldots, r\}$$
$$\to P\{N^{(k)}(B_j) = m_j^{(k)}, j = 1, \ldots, s, k = 1, \ldots, r\}. \quad (5.5.3)$$

Proof. Let B_{jk} be a rectangle in the plane with base B_j and such that L_k intersects its interior, but is disjoint from all other L_j. Then the left-hand side of (5.5.3) may be written as

$$P\{N_n(B_{jk}) = m_j^{(k)}, j = 1, \ldots, s, k = 1, \ldots, r\},$$

which by the appendix converges to the same quantity with N instead of N_n, i.e. to the right-hand side of (5.5.3). □

5.6. Joint Asymptotic Distribution of the Largest Maxima

We may apply the above results to obtain asymptotic joint distributions for a finite number of the kth largest maxima $M_n^{(k)}$, together with their locations if we wish. Such results may be obtained by considering appropriate continuous functionals of the sequence N_n, but here we take a more elementary approach, in giving examples of typical results. First we generalize Theorems 2.3.1 and 2.3.2.

Theorem 5.6.1. *Let the levels $u_n^{(k)}$, $1 \le k \le r$, satisfy (5.4.1) with $u_n^{(1)} \ge u_n^{(2)} \ge \cdots \ge u_n^{(r)}$, and suppose that the stationary sequence $\{\xi_n\}$ satisfies $D_r(\mathbf{u}_n)$, and $D'(u_n^{(k)})$ for $1 \le k \le r$. Let $S_n^{(k)}$ denote the number of exceedances of $u_n^{(k)}$ by $\xi_1, \ldots, \xi_n$. Then, for $k_1 \ge 0, \ldots, k_r \ge 0$,*

$$P\{S_n^{(1)} = k_1,\ S_n^{(2)} = k_1 + k_2, \ldots, S_n^{(r)} = k_1 + \cdots + k_r\}$$
$$\to \frac{\tau_1^{k_1}}{k_1!}\frac{(\tau_2 - \tau_1)^{k_2}}{k_2!}\cdots\frac{(\tau_r - \tau_{r-1})^{k_r}}{k_r!}e^{-\tau_r}, \quad (5.6.1)$$

as $n \to \infty$.

Proof. With the previous notation $S_n^{(j)} = N_n^{(j)}((0, 1])$ and so by Corollary 5.5.2 the left-hand side of (5.6.1) converges to

$$P\{S^{(1)} = k_1, S^{(2)} = k_1 + k_2, \ldots, S^{(r)} = k_1 + \cdots + k_r\}, \quad (5.6.2)$$

where $S^{(j)} = N^{(j)}((0, 1])$. But this is the probability that precisely $k_1 + k_2 + \cdots + k_r$ events occur in the unit interval for the Poisson process on the line L_r and that k_1 of the corresponding β's take the value r, k_2 take the value $r-1$, and so on. But the independence properties of the β's show that, conditional on a given total number $k_1 + k_2 + \cdots + k_r$, the numbers taking the values $r, r-1, \ldots, 1$ have a multinomial distribution based on the respective probabilities $\tau_1/\tau_r, (\tau_2 - \tau_1)/\tau_r, \ldots, (\tau_r - \tau_{r-1})/\tau_r$. Hence (5.6.2) is

$$\frac{(k_1 + k_2 \ldots + k_r)!}{k_1!k_2!\ldots k_r!}\left(\frac{\tau_1}{\tau_r}\right)^{k_1}\left(\frac{\tau_2 - \tau_1}{\tau_r}\right)^{k_2}\cdots\left(\frac{\tau_r - \tau_{r-1}}{\tau_r}\right)^{k_r}$$
$$\times P\{N^{(r)}((0, 1]) = k_1 + k_2 + \cdots + k_r\}$$

which gives (5.6.1) since

$$P\{N^{(r)}((0, 1]) = k_1 + k_2 + \cdots + k_r\} = \exp(-\tau_r)\frac{\tau_r^{k_1+k_2+\cdots+k_r}}{(k_1 + k_2 + \cdots + k_r)!}. \qquad \square$$

Of course this agrees with the result of Theorem 2.3.1. The next result (which generalizes Theorem 2.3.2 again) is given to exemplify the applicability of the Poisson theory.

Theorem 5.6.2. *Suppose that*

$$P\{a_n(M_n^{(1)} - b_n) \le x\} \xrightarrow{w} G(x) \qquad (5.6.3)$$

for some nondegenerate d.f. G and that $D_2(\mathbf{u}_n)$, $D'(u_n^{(k)})$ hold whenever $u_n^{(k)} = x_k/a_n + b_n$, $k = 1, 2$. Then the conclusion of Theorem 2.3.2 holds, i.e. for $x_1 > x_2$,

$$P\{a_n(M_n^{(1)} - b_n) \le x_1, a_n(M_n^{(2)} - b_n) \le x_2\}$$
$$\to G(x_2)(\log G(x_1) - \log G(x_2) + 1) \qquad (5.6.4)$$

when $G(x_2) > 0$ (zero when $G(x_2) = 0$).

Proof. If $u_n^{(k)} = x_k/a_n + b_n$, by (5.6.3) and the assumptions of the theorem it follows from Theorem 3.4.1 that $n(1 - F(u_n^{(k)})) \to \tau_k$ where $\tau_k = -\log G(x_k)$ if $G(x_2) > 0$ (and hence $G(x_1) > 0$). But clearly

$$P\{M_n^{(1)} \le u_n^{(1)}, M_n^{(2)} \le u_n^{(2)}\} = P\{S_n^{(2)} = 0\} + P\{S_n^{(1)} = 0, S_n^{(2)} = 1\},$$

where $S_n^{(i)}$ is the number of exceedances of $u_n^{(i)}$ by $\xi_1, \ldots, \xi_n$. By Theorem 5.6.1 we see that the limit of the above probabilities is

$$e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2} = e^{-\tau_2}(\tau_2 - \tau_1 + 1),$$

which is the desired result when $G(x_2) > 0$. The case $G(x_2) = 0$ may be dealt with directly by taking $\tau_2 = \infty$ or perhaps most simply from the continuity of G by dominating the left-hand side of (5.6.4) by its value with y replacing x_2 where $G(y) > 0$. $\square$

As a final example, we obtain the limiting joint distribution of the second maximum and its location (taking the leftmost if two values are equal).

Theorem 5.6.3. *Suppose that* (5.6.3) *holds and that* $D_4(\mathbf{u}_n)$, $D'(u_n^{(k)})$ *hold for all* $u_n^{(k)} = x_k/a_n + b_n$, $k = 1, 2, 3, 4$. *Then if* $L_n^{(2)}$, $M_n^{(2)}$ *are the location and height of the second largest of* $\xi_1, \dots, \xi_n$, *respectively,*

$$P\left\{\frac{1}{n} L_n^{(2)} \le t,\ a_n(M_n^{(2)} - b_n) \le x\right\} \to tG(x)(1 - \log G(x)), \tag{5.6.5}$$

x *real*, $0 < t \le 1$. *That is, the location and height are asymptotically independent, the location being asymptotically uniform.*

Proof. As in the previous theorem, we see that (5.4.1) holds with $\tau_k = -\log G(x_k)$. Write I, J for intervals $\{1, 2, \dots, [nt]\}$, $\{[nt] + 1, \dots, n\}$, respectively, $M^{(1)}(I)$, $M^{(2)}(I)$, $M^{(1)}(J)$, $M^{(2)}(J)$ for the maxima and second largest ξ_j in the intervals I, J, and let $H_n(x_1, x_2, x_3, x_4)$ be the joint d.f. of the normalized r.v.'s

$$X_n^{(1)} = a_n(M_n^{(1)}(I) - b_n), \qquad X_n^{(2)} = a_n(M_n^{(2)}(I) - b_n),$$
$$Y_n^{(1)} = a_n(M_n^{(1)}(J) - b_n), \qquad Y_n^{(2)} = a_n(M_n^{(2)}(J) - b_n).$$

That is, with $x_1 > x_2$ and $x_3 > x_4$,

$$\begin{aligned} H_n(x_1, x_2, x_3, x_4) = P\{&M_n^{(1)}(I) \le u_n^{(1)}, M_n^{(2)}(I) \le u_n^{(2)},\\ &M_n^{(1)}(J) \le u_n^{(3)}, M_n^{(2)}(J) \le u_n^{(4)}\},\end{aligned}$$

where $u_n^{(k)} = x_k/a_n + b_n$ as above. Alternatively, we see that

$$\begin{aligned} H_n(x_1, x_2, x_3, x_4) = P\{&N_n^{(1)}(I') = 0, N_n^{(2)}(I') \le 1,\\ &N_n^{(3)}(J') = 0, N_n^{(4)}(J') \le 1\},\end{aligned}$$

where $I' = (0, t]$ and $J' = (t, 1]$, so that an obvious application of Corollary 5.5.2 with $B_1 = I'$, $B_2 = J'$ gives

$$\begin{aligned} \lim_{n \to \infty} H_n(x_1, x_2, x_3, x_4) &= P\{N^{(1)}(I') = 0, N^{(2)}(I') \le 1\}\\ &\quad \times P\{N^{(3)}(J') = 0, N^{(4)}(J') \le 1\}\\ &= e^{-t\tau_2}(t(\tau_2 - \tau_1) + 1)e^{-(1-t)\tau_4}((1 - t)(\tau_4 - \tau_3) + 1)\\ &= H_t(x_1, x_2)H_{(1-t)}(x_3, x_4) = H(x_1, x_2, x_3, x_4), \end{aligned} \tag{5.6.6}$$

say, where

$$H_t(x_1, x_2) = G^t(x_2)(\log G^t(x_1) - \log G^t(x_2) + 1),$$

for $x_1 > x_2$. Now clearly

$$\begin{aligned} P&\left\{\frac{1}{n} L_n^{(2)} \le t,\ a_n(M_n^{(2)} - b_n) \le x_2\right\}\\ &= P\{M_n^{(2)}(I) \le u_n^{(2)}, M_n^{(2)}(I) \ge M_n^{(1)}(J)\}\\ &\quad + P\{M_n^{(1)}(I) \le u_n^{(2)}, M_n^{(1)}(J) > M_n^{(1)}(I) \ge M_n^{(2)}(J)\}\\ &= P\{X_n^{(2)} \le x_2, X_n^{(2)} \ge Y_n^{(1)}\} + P\{X_n^{(1)} \le x_2, Y_n^{(1)} > X_n^{(1)} \ge Y_n^{(2)}\}. \end{aligned} \tag{5.6.7}$$

But, by the above calculation $(X_n^{(1)}, X_n^{(2)}, Y_n^{(1)}, Y_n^{(2)})$ converges in distribution to (X_1, X_2, Y_1, Y_2), whose joint d.f. is H, which is clearly absolutely continuous since G is (being an extreme value d.f.). Hence, since the boundaries of sets in R^4 such as $\{(w_1, w_2, w_3, w_4); w_2 \leq x_2, w_2 > w_3\}$ clearly have zero Lebesgue measure, it follows that the sum of probabilities in (5.6.7) converges to

$$P\{X_2 \leq x_2, X_2 \geq Y_1\} + P\{X_1 \leq x_2, Y_1 > X_1 \geq Y_2\}, \qquad (5.6.8)$$

and therefore the left-hand side of (5.6.5) converges to (5.6.8). This may be evaluated using the joint distribution H of X_1, X_2, Y_1, Y_2 given by (5.6.6). However, it is simpler to note that we would obtain the same result (5.6.8) if the sequence were i.i.d. But (5.6.5) is simply evaluated for an i.i.d. sequence by noting the independence of $L_n^{(2)}$ and $M_n^{(2)}$ and the fact that $L_n^{(2)}$ is then uniform on $(1, 2, \ldots, n)$, giving the limit stated in (5.6.5), as is readily shown. □

5.7. Complete Poisson Convergence

In the previous point process convergence results, we obtained a limiting point process in the plane, formed from the exceedances of a fixed number r of increasingly high levels. The limiting process was not Poisson in the plane, though composed of r successively more severely thinned Poisson processes on r lines.

On the other hand, we may regard the sample sequence $\{\xi_n\}$ itself—after suitable transformations of both coordinates—as a point process in the plane and, by somewhat strengthening the assumptions, show that this converges to a Poisson process in the plane. This procedure has been used for independent r.v.'s by, e.g. Pickands (1971) and Resnick (1975) and subsequently for stationary sequences by R. J. Adler (1978), who used the linear normalization of process values provided by the constants a_n, b_n appearing in the asymptotic distribution of M_n. Here we shall consider a slightly more general case.

Specifically, with the standard notation, suppose that $u_n(\tau)$ is defined for $n = 1, 2, \ldots; \tau > 0$ to be continuous and strictly decreasing in τ for each n, and satisfying (1.5.1), viz.

$$n(1 - F(u_n(\tau))) \to \tau. \qquad (5.7.1)$$

For example, $u_n(\tau) = a_n^{-1}G^{-1}(e^{-\tau}) + b_n$ with the usual notation when M_n has a limiting d.f. G.

Here we will use N_n to denote the point process in the plane consisting of the points $(j/n, u_n^{-1}(\xi_j))$, $j = 1, 2, \ldots$, where u_n^{-1} denotes the inverse function of $u_n(\tau)$, (defined on the range of the r.v.'s ξ_j).

Theorem 5.7.1. *Suppose $u_n(\tau)$ are defined as above satisfying* (5.7.1) *and that $D'(u_n)$ holds for each $u_n = u_n(\tau)$, and that $D_r(\mathbf{u}_n)$ holds for each $r = 1, 2, \ldots$; $\mathbf{u}_n = (u_n(\tau_1), \ldots, u_n(\tau_r))$ for each choice of $\tau_1, \ldots, \tau_r$. Then the point processes N_n, consisting of the points $(j/n, u_n^{-1}(\xi_j))$, converge to a Poisson process N on $(0, \infty) \times (0, \infty)$, having Lebesgue measure as its intensity.*

PROOF. This follows relatively simply by using the previous r-level exceedance theory. It will be convenient, and here permissible, to use rectangles which are closed at the bottom rather than the top, so that we need to show

(a) $E(N_n(B)) \to E(N(B))$ for all sets B of the form $(c, d] \times [\gamma, \delta)$, $0 < c < d$, $0 < \gamma < \delta$, and
(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for sets B which are finite unions of sets of this form.

Here, (a) follows simply since if $B = (c, d] \times [\gamma, \delta)$,

$$\begin{aligned} E(N_n(B)) &= ([nd] - [nc])P\{\gamma \le u_n^{-1}(\xi_1) < \delta\} \\ &\sim n(d - c)P\{u_n(\delta) < \xi_1 \le u_n(\gamma)\} \\ &= n(d - c)\{F(u_n(\gamma)) - F(u_n(\delta))\} \\ &\sim n(d - c)(\delta - \gamma)/n, \end{aligned}$$

while $E(N(B)) = (d - c)(\delta - \gamma)$.

To show (b) we note that any finite disjoint union of such rectangles may be written in the form $\bigcup_j (E_j \times F_j)$, where $E_j = (c_j, d_j]$ are disjoint and F_j is a finite disjoint union $\bigcup_k [\gamma_{j,k}, \delta_{j,k})$, (cf. the proof of Theorem 5.5.1). Suppose first that there is just one set E_j, i.e. $B = \bigcup_{k=1}^m E \times F_k$, say, where we write $F_k = [\tau_{2k-1}, \tau_{2k})$, $k = 1, \ldots, m$, and where we may clearly take $\tau_1 < \tau_2 < \ldots < \tau_r$ $(r = 2m)$.

Now $N_n(B) = 0$ means that, for each k, there is no $j/n \in E$ for which $u_n^{-1}(\xi_j) \in F_k$, i.e. such that $u_n(\tau_{2k}) < \xi_j \le u_n(\tau_{2k-1})$. But this is equivalent to the statement that for $j/n \in E$, ξ_j exceeds $u_n(\tau_{2k-1})$ as many times as it exceeds $u_n(\tau_{2k})$. That is, writing $N_n^{(k)}(E)$ for the number of exceedances of $u_n(\tau_k)$ by ξ_j for $j/n \in E$,

$$\{N_n(B) = 0\} = \bigcap_{k=1}^{r/2} \{N_n^{(2k-1)}(E) = N_n^{(2k)}(E)\}. \tag{5.7.2}$$

But $N_n^{(k)}$ is precisely the same as in Theorem 5.5.1 and its corollary, and their conditions are clearly satisfied, so that by Corollary 5.5.2(ii) we have

$$(N_n^{(1)}(E), N_n^{(2)}(E), \ldots, N_n^{(r)}(E)) \xrightarrow{d} (N^{(1)}(E), N^{(2)}(E), \ldots, N^{(r)}(E)), \tag{5.7.3}$$

where $N^{(1)}, \ldots, N^{(r)}$ are the r succesively thinned Poisson processes on r fixed lines as defined prior to Theorem 5.5.1, but since all the r.v.'s $N_n^{(k)}(E)$, $N^{(k)}(E)$ are integer valued, it is an obvious exercise in distributional convergence in R^r to show from (5.7.3) that the probability of pairwise equality in (5.7.2) converges to the same probability with $N^{(k)}$ replacing $N_n^{(k)}$. Thus

$$P\{N_n(B) = 0\} \to P\left(\bigcap_{k=1}^{r/2} \{N^{(2k-1)}(E) = N^{(2k)}(E)\}\right).$$

From the discussion prior to Theorem 5.5.1 we see that the events in braces on the right occur if the β_j corresponding to each Poisson event on the line L_r is even, i.e. $\beta_j = 2, 4, 6, \ldots, r$. Since $P\{\beta_j = s\} = (\tau_{r-s+1} - \tau_{r-s})/\tau_r$ if $s \leq r - 1$, and τ_1/τ_r if $s = r$, we have, writing

$$\sigma = (\tau_{r-1} - \tau_{r-2}) + (\tau_{r-3} - \tau_{r-4}) + \cdots + (\tau_3 - \tau_2) + \tau_1,$$

$$\begin{aligned} P\{N_n(B) = 0\} &\to \sum_{j=0}^{\infty} \exp(-\tau_r(d - c)) \frac{(\tau_r(d - c))^j}{j!} \left(\frac{\sigma}{\tau_r}\right)^j \\ &= \exp((d - c)(\sigma - \tau_r)) = e^{-m(B)}, \end{aligned}$$

where m denotes Lebesgue measure, since

$$(d - c)(\sigma - \tau_r) = \sum_{k=1}^{r/2} (d - c)(\tau_{2k-1} - \tau_{2k}) = -\sum_{k=1}^{r/2} m(E \times F_k) = -m(B).$$

Hence (b) follows when $B = \bigcup_k E \times F_k$. When $B = \bigcup_j (E_j \times F_j)$ the same proof applies—using the full statement of Corollary 5.5.2 with slightly more notational complexity since more τ_k's may be needed corresponding to the additional E_j's. □

In the theorem above it is required, of course that $P\{M_n \leq u_n(\tau)\} \to e^{-\tau}$. If there is a d.f. $G(x)$ and τ may be chosen as a function $\tau(x)$ such that $P\{M_n \leq u_n(\tau(x))\} \to G(x)$, then we would have $\tau(x) = -\log G(x)$ and $P\{M_n \leq v_n(x)\} \to G(x)$, with $v_n(x) = u_n(\tau(x))$. In such a case it would be natural to consider the point process formed from points $(j/n, v_n^{-1}(\xi_j))$ instead of $(j/n, u_n^{-1}(\xi_j))$. In particular, when a linear normalization leads to an asymptotic distribution, i.e. $P\{a_n(M_n - b_n) \leq x\} \to G(x)$ we have $v_n(x) = x/a_n + b_n$ and it is natural to consider the point process N_n' consisting of points $(j/n, a_n(\xi_j - b_n))$. This is the case considered in Adler (1978) where it is shown that a (nonhomogeneous) Poisson limit holds. Here we obtain this result as a corollary of Theorem 5.7.1.

Theorem 5.7.2. *Suppose that* (5.3.3) *holds, i.e.* $P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x)$ *for some nondegenerate d.f. G, and let* $x_0 = \inf\{x; G(x) > 0\}$. *Suppose that* $D'(u_n)$ *holds for all sequences* $u_n = x/a_n + b_n$, *and that* $D_r(\mathbf{u}_n)$ *holds for all* $r = 1, 2, \ldots,$ *and all sequences* $u_n^{(k)} = x_k/a_n + b_n$, $1 \leq k \leq r$, *for arbitrary choices of the* x_k. *Then if* N_n' *denotes the point process in the plane with points at* $(j/n, a_n(\xi_j - b_n))$ *we have* $N_n' \xrightarrow{d} N'$ *on* $(0, \infty) \times (x_0, \infty)$, *where* N' *is a Poisson process whose intensity measure is the product of Lebesgue measure and that defined by the increasing function* $\log G(y)$.

PROOF. By Theorem 3.4.1 the conditions of Theorem 5.7.1 hold, and hence $N_n \xrightarrow{d} N$, with the notation of Theorem 5.7.1 But if N_n has an atom at (s, t), N_n' has an atom at $(s, \tau^{-1}(t))$ where $\tau(x) = -\log G(x)$. Hence by Theorem A.3(i) $N_n' \xrightarrow{d} N'$ where N' is obtained from the Poisson process N by replacing atoms at points (s, t) by atoms at points $(s, \tau^{-1}(t))$, and by Theorem A.2(ii), this is also a Poisson process, with intensity measure λ' defined by

$$\lambda'((c, d] \times (\gamma, \delta]) = (d - c)(\tau(\gamma) - \tau(\delta)) = (d - c)(\log G(\delta) - \log G(\gamma))$$

from which the result follows. (Note that since G is an extreme value distribution, τ is continuous and strictly decreasing where G is nonzero.) □

Note that, unlike earlier results, the lower boundary $(0, \infty) \times \{x_0\}$ of $S = (0, \infty) \times (x_0, \infty)$ must be excluded. For otherwise, if x_0 is finite, $(0, 1] \times (x_0, x_0 + 1]$ would be bounded (in the technical meaning of p. 306) but $N((0, 1] \times (x_0 x_0 + 1])$ would be infinite a.s., so that N would not be a point process. A similar remark applies in Theorem 5.8.1 below. Finally we note that all the results which follow from the multilevel result (Theorem 5.5.1) may be obtained from the last two theorems—however, the D-assumptions made are correspondingly more stringent.

5.8. Record Times and Extremal Processes

There is a sizeable literature on record times of i.i.d. sequences and on so-called extremal processes. As an illustration to the results of the previous section we shall consider the asymptotic distribution of record times in dependent processes, and make a brief comment about extremal processes.

By definition, ξ_1 is a record of the sequence $\xi_1, \xi_2, \ldots$, and, for $j \geq 2$, ξ_j is a record if $\xi_j > M_{j-1}$. The *record times* are then $\tau_1 = 1$, and for $k \geq 2$,

$$\tau_k = \inf\{j > \tau_{k-1};\ \xi_j > M_{j-1}\}.$$

We shall start by noting some properties of record times of i.i.d. sequences. One interesting fact for i.i.d. ξ_n's is that the distribution of $\{\tau_k\}$ does not depend on the marginal d.f. F for continuous F's. In fact, the record times of ξ_1, $\xi_2, \ldots$ are then the same as those of $F(\xi_1), F(\xi_2), \ldots$ and these variables are obviously i.i.d. with distributions uniform on $(0, 1]$. Further, for $\{\xi_n\}$ i.i.d. with continuous marginal d.f., and for A_j the event that j is a record time, clearly

$$P(A_j) = P\{\xi_j > M_{j-1}\} = \frac{1}{j} \tag{5.8.1}$$

and further, for $m > j \geq 1$,

$$\sum_{k=1}^{\infty} P\{\tau_{k+1} > m, \tau_k \leq j\} = P\{M_m = M_j\} = \frac{j}{m} \to 0 \quad \text{as } m \to \infty, \tag{5.8.2}$$

for fixed j. In particular, (5.8.2) shows that for $k = 1, 2, \ldots$

$$P\{\tau_{k+1} > m\} \leq P\{\tau_k > j\} + P\{\tau_{k+1} > m, \tau_k \leq j\} \to P\{\tau_k > j\} \quad \text{as } m \to \infty,$$

and hence, for any $j \geq 1$, $\limsup_{m\to\infty} P\{\tau_{k+1} > m\} \leq P\{\tau_k > j\}$. Since $\tau_1 = 1$ it follows by induction that $\tau_2, \tau_3, \ldots$ are all finite with probability one.

Let R_n be the point process on the unit interval (0, 1] consisting of the points τ_k/n, for $0 < \tau_k/n \leq 1$, i.e. if $\xi_1, \xi_2, \ldots, \xi_n$ has a record at time j, then R_n has a point at j/n. By (5.8.1), for $\{\xi_n\}$ i.i.d. with continuous d.f., for $0 < c \leq d \leq 1$,

$$E(R_n((c, d])) = \sum_{j=[nc]+1}^{[nd]} \frac{1}{j} \sim \int_c^d \frac{1}{t}\, dt = \log \frac{d}{c}, \tag{5.8.3}$$

and it can be proved simply that asymptotically R_n is a Poisson process with intensity $1/t$. Several further results about record times for i.i.d. sequences can be obtained by straightforward calculations. The paper by Glick (1978) contains a quite readable elementary exposition of these and related matters.

Obviously these results need not hold for dependent sequences but, as shall be seen, for many such sequences the asymptotic distribution of record times is the same as for independent sequences. This can be proved by elementary means, in the same way as Theorems 5.6.2 and 5.6.3, but here we shall use instead the general approach to point process convergences, as outlined in the appendix, as an illustration to the power of that approach.

Theorem 5.8.1. *Suppose that the hypothesis of Theorem* 5.7.1 *is satisfied. Then the point process R_n consisting of the points τ_k/n converges, as $n \to \infty$, to a Poisson process R with intensity l/t on* (0, 1]. *In particular, if $0 < c < d \leq 1$,*

$$P\{R_n((c, d]) = k\} \to \frac{(\log(d/c))^k}{k!} \cdot \frac{c}{d} \quad \text{as } n \to \infty,$$

for $k = 0, 1, \ldots$.

Proof. By definition, convergence of R_n on (0, 1] is equivalent to convergence of the joint distribution of $R_n((c_1, d_1]), \ldots, R_n((c_k, d_k])$ for $0 < c_1 < d_1 < c_2 < \cdots < d_k \leq 1$, which in turn follows if R_n converges to R on $(\varepsilon, 1]$, for $0 < \varepsilon \leq 1$. Hence, to conclude that R_n converges to R on (0, 1] it is sufficient to prove convergence on $(\varepsilon, 1]$, for each $\varepsilon > 0$.

Let N_n and N be as in Theorem 5.7.1. Then, on $(\varepsilon, 1]$, the measure $N_n(\cdot)$ on R^2 clearly determines the measure $R_n(\cdot)$ on R^1, i.e. $R_n = h(N_n)$, where h maps measures on R^2 to measures on R^1. Suppose the integer-valued measure ν on $(0, 1] \times (0, \infty)$ is simple and has the property that for some constant $\gamma > 0$, $\nu((0, \varepsilon) \times (0, \gamma)) > 0$, $\nu((0, 1] \times (0, \gamma)) < \infty$ and $\nu((0, 1] \times \{x\}) \leq 1$ for all $0 < x < \gamma$. It is then immediate from Proposition A.4 that h is continuous at ν, cf. Figure 5.8.1. (Note that u_n^{-1} is decreasing so the records of $\{\xi_j\}$ correspond to the successive *minima* of $u_n^{-1}(\xi_j)$.) Since N a.s. has the properties required of ν, h is a.s. N-continuous (as defined after Theorem A.4) and thus

$$R_n = h(N_n) \xrightarrow{d} h(N) \tag{5.8.4}$$

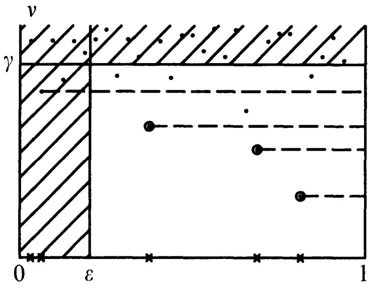


Figure 5.8.1. The dots are the points of v and the crosses the points of $h(v)$. The dashed lines illustrate, for some important points, that no two points of v are on the same horizontal line above γ.

on $(\varepsilon, 1]$. To complete the proof we only have to show that $h(N)$ has the distribution specified for R, i.e. that it is a Poisson process with intensity $1/t$.

One easy way to do this is to show directly that $R_n \to R$ in some special case, since it then follows from the uniqueness of limits of distributions that $h(N)$ has the required distribution. Thus, we shall use Theorem A.1 to show that $R_n \xrightarrow{d} R$ if $\{\xi_n\}$ is i.i.d. with continuous marginal distribution. In fact, A.1(a) then follows at once, by (5.8.3). Further, it is easily checked that, for $\varepsilon < c_1 < d_1 < c_2 < \cdots < d_k \leq 1$, $R_n((c_1, d_1]), \ldots, R_n((c_k, d_k])$ are independent, as are, by definition, $R((c_1, d_1]), \ldots, R((c_k, d_k])$. Hence, to prove A.1(b) it suffices to note that by (5.8.2)

$$P\{R_n((c, d]) = 0\} = P\{M_{[nc]} = M_{[nd]}\} = \frac{[nc]}{[nd]} \sim \frac{c}{d} = \exp\left(-\int_c^d \frac{1}{t}\,dt\right),$$

for $\varepsilon < c < d \leq 1$. Thus $R_n \xrightarrow{d} R$ in the special case of i.i.d. variables, and by (5.8.4), hence also in general, under the hypothesis of the theorem. □

Finally it is natural to consider the simultaneous distribution of record times and the values of the records, which in turn lead to the process $\tilde{M}(t) = \max_{1 \leq k \leq t} \xi_k$. The study of such processes and their convergence, after suitable scaling, to so-called *extremal processes*, was initiated by Dwass (1964) and Lamperti (1964), and has led to interesting results about $\tilde{M}(t)$ and about the limiting extremal processes. However, as was noted by Resnick (1975), convergence to extremal processes is easily derived from complete Poisson convergence by similar considerations as in the proof of Theorem 5.8.1. Since no new ideas are involved in doing so, we shall not pursue this further.

CHAPTER 6

Nonstationary, and Strongly Dependent Normal Sequences

While our primary concern in this volume is with stationary processes, the results and methods may be used to apply simply to some important nonstationary cases. In particular, this is so for nonstationary normal sequences having a wide variety of possible mean and correlation structures, which is the situation considered first in this chapter.

One important such application occurs when the process consists of a stationary normal sequence together with an added deterministic part, such as a seasonal component or trend. Cases of this kind have been discussed under certain conditions by Horowitz (1980), and as part of a more general consideration of stationary sequences which are not necessarily normal, by Hüsler (1981). Ideas from both of these works will be used in our development here.

In discussing stationary sequences, it was found that the classical limits still hold under quite a slow rate of dependence decay (e.g. $D(u_n)$). It is also of interest to determine the effect on the extremal results by permitting a very persistent dependence structure. It will be shown, following Mittal and Ylvisaker (1975), that for stationary normal sequences with such very strong dependence, limits other than the three extreme value types may occur. In particular, these cases show that the weak covariance conditions used in the previous chapters are almost the best possible for the limiting distribution of the maximum to be a Type I extreme value distribution.

6.1. Nonstationary Normal Sequences

Let $\{\xi_n\}$ be a normal sequence (in general nonstationary) with arbitrary means and variances, and correlations $r_{ij} = \mathrm{corr}(\xi_i, \xi_j)$. We shall be concerned with conditions under which

$$P\left(\bigcap_{i=1}^{n}\{\xi_i \le u_{ni}\}\right) \to e^{-\tau}, \tag{6.1.1}$$

where $0 \le \tau < \infty$, u_{ni} being constants defined for $1 \le i \le n$, $n \ge 1$. In this section we consider general forms for the constants u_{ni} and obtain a result (Theorem 6.1.3) giving conditions under which (6.1.1) holds. This will be specialized in Section 6.2 by specific choice of u_{ni} to yield results concerning asymptotic distributions of maxima. A still more general form of Theorem 6.1.3 will be given in Section 6.3. This result was proved by Hüsler (1981) for stationary normal sequences as a particular case of a consideration of stationary sequences under extended $D(u_n)$, $D'(u_n)$ types of conditions. Our proof for the nonstationary normal sequences considered here uses a version of an interesting and somewhat delicate estimation employed by Hüsler (1981).

Clearly by standardizing each ξ_i (and correspondingly replacing u_{ni} by $v_{ni} = (u_{ni} - E(\xi_i))/\mathrm{Var}(\xi_i)^{1/2})$ we may assume in (6.1.1) that each ξ_i has zero mean, unit variance, and the same given correlation structure. Of course, in applications any conditions assumed on the u_{ni} must be translated to apply to the original forms of the u_{ni}. Thus, unless stated otherwise, we assume then that each ξ_i has zero mean and unit variance.

In most of the first three sections we assume that the correlations r_{ij} satisfy $|r_{ij}| < \rho_{|i-j|}$ when $i \ne j$, for some sequence $\{\rho_n\}$ such that $\rho_n < 1$ for $n \ge 1$ and $\rho_n \log n \to 0$—an obvious generalization of the condition $r_n \log n \to 0$ used for stationary sequences. Our main results then concern conditions on the $\{u_{ni}\}$ under which the ξ_i behave like an independent sequence in so far as the probability $P(\bigcap_1^n\{\xi_i \le u_{ni}\})$ is concerned, in the specific sense that

$$P\left(\bigcap_{i=1}^{n}\{\xi_i \le u_{ni}\}\right) - \prod_{i=1}^{n}\Phi(u_{ni}) \to 0. \tag{6.1.2}$$

From such results it will then follow simply that if also

$$\sum_{i=1}^{n}(1 - \Phi(u_{ni})) \to \tau \tag{6.1.3}$$

then (6.1.1) holds, generalizing a conclusion of Theorem 4.3.3 in the stationary case.

The proofs of these results hinge on calculations similar to those of Lemma 4.3.2 and will be given by means of several lemmas. In this section we require, in proving (6.1.2), that the $\{u_{ni}\}$ should be such that $\sum_1^n (1 - \Phi(u_{ni}))$ is bounded and that $\min_{1 \le i \le n} u_{ni}$ tends to infinity as fast as some multiple of $(\log n)^{1/2}$—which will be the case in our applications to maxima in the next section. In Section 6.3 the more delicate arguments using ideas from Hüsler (1981) will be employed to extend the result to sequences for which $\min_{1 \le i \le n} u_{ni}$ tends to infinity but not necessarily even as fast as a multiple of $(\log n)^{1/2}$.

First we give a simple preliminary lemma generalizing Theorem 1.5.1.

Lemma 6.1.1. *Let* $\{u_{ni},\ 1 \le i \le n,\ n = 1, 2, \ldots\}$ *be constants such that* $\lambda_n = \min_{1 \le i \le n} u_{ni} \to \infty$. *Then, for* $0 \le \tau \le \infty$,

$$\prod_{i=1}^{n} \Phi(u_{ni}) \to e^{-\tau} \quad \text{as } n \to \infty \tag{6.1.4}$$

if and only if

$$\sum_{i=1}^{n} (1 - \Phi(u_{ni})) \to \tau \quad \text{as } n \to \infty. \tag{6.1.5}$$

Proof. By using the fact that $\log(1 - x) = -x + \psi(x)$ where, for small $x > 0$, $|\psi(x)| < Ax^2$ for some $A > 0$, it is simply seen that

$$\sum_{i=1}^{n} \log\{1 - (1 - \Phi(u_{ni}))\} = -\sum_{i=1}^{n} (1 - \Phi(u_{ni})) + \sum_{i=1}^{n} \psi(1 - \Phi(u_{ni})),$$

where

$$\left|\sum_{i=1}^{n} \psi(1 - \Phi(u_{ni}))\right| \le A \sum_{i=1}^{n} (1 - \Phi(u_{ni}))^2 \le A(1 - \Phi(\lambda_n)) \sum_{i=1}^{n} (1 - \Phi(u_{ni}))$$

so that clearly

$$\log \prod_{i=1}^{n} \Phi(u_{ni}) = -\left\{\sum_{i=1}^{n} (1 - \Phi(u_{ni}))\right\}\{1 + o(1)\}$$

from which both implications of the lemma follow simply. □

In the following results the notation already established in this section will be used without comment.

Lemma 6.1.2. *Let* $\{u_{ni}\}$ *be such that* $\sum_{i=1}^{n} (1 - \Phi(u_{ni}))$ *is bounded, and* $\lambda_n = \min_{1 \le i \le n} u_{ni} \to \infty$. *Suppose that the correlations* r_{ij} *satisfy* $|r_{ij}| \le \delta$ *for* $i \ne j$, *where* $\delta < 1$ *is a constant. Then*

$$S_n^{(1)} = \sum_{\substack{1 \le i < j \le n \\ |i-j| \le \theta_n}} |r_{ij}| \exp\left(-\frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + |r_{ij}|}\right) \to 0 \quad \text{as } n \to \infty, \tag{6.1.6}$$

where $\theta_n = e^{\eta\lambda_n^2}$, *for any* $\eta < \eta_0 = \frac{1}{2}(1 - \delta)/(1 + \delta)$

Proof. Clearly

$$S_n^{(1)} \le \delta \sum_{\substack{1 \le i < j \le n \\ |i-j| \le \theta_n}} \exp\left(-\frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + \delta}\right).$$

The exponential term does not exceed

$$\exp\left(-\frac{\min(u_{ni}^2, u_{nj}^2)}{1+\delta}\right) \le \exp\left(-\frac{u_{ni}^2}{1+\delta}\right) + \exp\left(-\frac{u_{nj}^2}{1+\delta}\right)$$

giving, for a suitable constant K

$$S_n^{(1)} \le 2\delta\theta_n \sum_{i=1}^{n} \exp\left(-\frac{u_{ni}^2}{1+\delta}\right) = 2\delta\theta_n \sum_{i=1}^{n} u_{ni}^{-1} \exp\left(-\frac{u_{ni}^2}{2}\right) u_{ni} \exp(-\eta_0 u_{ni}^2)$$

$$\le K\theta_n \lambda_n \exp(-\eta_0 \lambda_n^2) \sum_{i=1}^{n} (1 - \Phi(u_{ni}))$$

since $u_{ni} \ge \lambda_n \to \infty$, so that $u_{ni}^{-1} \exp(-u_{ni}^2/2)/(1 - \Phi(u_{ni}))$ is (for $\lambda_n > 0$) uniformly bounded in i, and $x \exp(-\eta_0 x^2)$ decreases for sufficiently large x. But $\sum_1^n (1 - \Phi(u_{ni}))$ is bounded, and it thus follows that

$$S_n^{(1)} \le K\lambda_n \exp(-(\eta_0 - \eta)\lambda_n^2) \to 0 \quad \text{as } n \to \infty,$$

since $\eta < \eta_0$. □

The main result of this section may now be readily proved.

Theorem 6.1.3. *Suppose that the correlations r_{ij} of the normal sequence $\{\xi_n\}$ are such that $|r_{ij}| \le \rho_{|i-j|}$ for $i \ne j$ where $\rho_n < 1$ for all $n \ge 1$ and $\rho_n \log n \to 0$ as $n \to \infty$. Let the constants $\{u_{ni}\}$ be such that $\sum_{i=1}^n (1 - \Phi(u_{ni}))$ is bounded and $\lambda_n = \min_{1 \le i \le n} u_{ni} \ge c(\log n)^{1/2}$ for some $c > 0$. Then* (6.1.2) *holds. If further,* (6.1.3) *holds for some $\tau \ge 0$, then so does* (6.1.1), *i.e. $P(\bigcap_{i=1}^n \{\xi_i \le u_{ni}\}) \to e^{-\tau}$.*

Proof. The assumptions clearly imply those of Lemma 6.1.2, so that $S_n^{(1)}$ (given by (6.1.6)) tends to zero as $n \to \infty$. Write

$$S_n^{(2)} = \sum_{\substack{1 \le i < j \le n \\ |i-j| > \theta_n}} |r_{ij}| \exp\left(-\frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + |r_{ij}|}\right).$$

Since $\theta_n = \exp(\eta\lambda_n^2) \ge \exp(c^2\eta \log n) = n^\alpha$, $\alpha = c^2\eta > 0$, it is clear that $S_n^{(2)}$ does not exceed the same sum with n^α replacing θ_n, and δ_{n^α} replacing $|r_{ij}|$ where $\delta_x = \sup_{j \ge x} \rho_j$. This gives, writing $p = n^\alpha$,

$$S_n^{(2)} \le \delta_p \left\{ \sum_{i=1}^{n} \exp\left(-\frac{\frac{1}{2}u_{ni}^2}{1+\delta_p}\right) \right\}^2.$$

Now define u_n by $1 - \Phi(u_n) = 1/n$ and split the sum in braces into two parts for $u_{ni} \ge u_n$, and $u_{ni} < u_n$, giving simply

$$S_n^{(2)} \le \delta_p \left\{ n \exp\left(-\frac{\frac{1}{2}u_n^2}{1+\delta_p}\right) + \sum_{u_{ni} < u_n} \exp(-\tfrac{1}{2}u_{ni}^2(1 - \delta_p)) \right\}^2$$

$$\le \delta_p u_n^2 \exp(u_n^2 \delta_p) \left\{ \frac{n}{u_n} \exp\left(-\frac{u_n^2}{2}\right) + \sum_{i=1}^{n} u_{ni}^{-1} \exp\left(-\frac{u_{ni}^2}{2}\right) \right\}^2$$

which tends to zero since $\delta_p u_n^2 \sim 2\delta_{n^\alpha} \log n = (2/\alpha)\delta_{n^\alpha} \log n^\alpha \to 0$, and the term in braces does not exceed $K\{n(1 - \Phi(u_n)) + \sum_1^n (1 - \Phi(u_{ni}))\}$ which is bounded. Since $S_n^{(1)} \to 0$, $S_n^{(2)} \to 0$ it follows that

$$\sum_{1 \le i < j \le n} |r_{ij}| \exp\left(- \frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + |r_{ij}|} \right) \to 0 \tag{6.1.7}$$

and (6.1.2) follows at once from Theorem 4.2.1 (Eqn. (4.2.3)) on taking η_j to be independent standard normal variables and making the obvious identifications, (using the fact that $\sup_{i \ne j} |r_{ij}| < 1$ under the assumption made).

Finally, if further (6.1.3) holds for some $\tau \ge 0$, then Lemma 6.1.1 shows that $\prod_1^n \Phi(u_{ni}) \to e^{-\tau}$ so that (6.1.1) follows by (6.1.2). □

6.2. Asymptotic Distribution of the Maximum

The results of the last section may be used to give the actual asymptotic distribution of the maximum in many cases of interest. We shall be especially concerned with stationary normal sequences to which a known (or in practice, perhaps, well-estimated) trend or seasonal component has been added. In such a case we will see how the classical normalizing constants are modified to include the added component. This is along the broad lines of a result by Horowitz (1980), (with a necessary correction to a normalizing constant given there) but under the present types of correlation condition. However, as will be further seen, the results apply without change to a non-stationary normal sequence with constant, say unit, variance but arbitrary correlations r_{ij} under the usual conditions $|r_{ij}| < \rho_{|i-j|}$ with $\rho_n \log n \to 0$.

We shall be concerned then with the maxima $M_n = \max\{\eta_1, \ldots, \eta_n\}$ from a normal sequence $\{\eta_n, n = 1, 2, \ldots\}$ given by $\eta_i = \xi_i + m_i$ where $\{\xi_n\}$ is the normal sequence defined in the previous section ($E(\xi_i) = 0$, $\mathrm{Var}(\xi_i) = 1$, $\mathrm{Cov}(\xi_i, \xi_j) = r_{ij}$), and $\{m_n\}$ are added deterministic components. Hence of course $E(\eta_i) = m_i$, $\mathrm{Var}(\eta_i) = 1$, $\mathrm{Cov}(\eta_i, \eta_j) = r_{ij}$. We shall assume that the constants m_i are such that

$$\beta_n = \max_{1 \le i \le n} |m_i| = o((\log n)^{1/2}) \quad \text{as } n \to \infty. \tag{6.2.1}$$

This restriction is quite mild in practice and includes the most important case of bounded m_i's considered by Horowitz (1980) as well as a variety of unbounded trends. Of course, constants tending more rapidly to infinity could be considered but may yield degenerate results. It should be noted that the restriction (6.2.1) may be weakened in an appropriate way to restrict $\max m_i$ rather than $\max |m_i|$, since it may be seen that η_i with small m_i are less likely to provide the maximum value and may be disregarded.

However, for simplicity of statement and calculation it is convenient to require (6.2.1) as stated, and this will be sufficient for practical purposes.

It will be shown that the usual limit law still holds provided the classical constant b_n is replaced by $b_n + m_n^*$ where m_n^* is chosen so that $|m_n^*| \le \beta_n$ and

$$\frac{1}{n}\sum_{i=1}^{n} \exp(a_n^*(m_i - m_n^*) - \tfrac{1}{2}(m_i - m_n^*)^2) \to 1 \quad \text{as } n \to \infty, \tag{6.2.2}$$

in which $a_n^* = a_n - \log\log n/2a_n$, $(a_n = (2\log n)^{1/2})$. The solution m_n^* to (6.2.2), (putting the left-hand side equal to 1, say) could be a difficult numerical problem, though for large n a solution with $|m_n^*| \le \beta_n$ clearly exists under (6.2.1) (e.g. $\psi(x) = n^{-1}\sum_1^n e^{a_n^*(m_i - x)(1 - 1/2a_n^{*-1}(m_i - x))}$ satisfies $\psi(\beta_n) \le 1 \le \psi(-\beta_n)$ when $\beta_n/a_n^* < 1$.) In some cases we will find an explicit form for m_n^*—e.g. when the m_i are bounded. (The simple form for m_n^* given by Horowitz (1980) is not in general correct even for bounded m_i, though it does apply under further appropriate conditions. With this notation we now give the main result.

Theorem 6.2.1. *Let $\{\eta_n\}$ be defined as above by $\eta_i = \xi_i + m_i$ where $\{\xi_n\}$ is a normal sequence with zero means, unit variances and correlations r_{ij} such that $|r_{ij}| < \rho_{|i-j|}$ for $i \neq j$ with $\rho_n < 1$ and $\rho_n \log n \to 0$. Let $\{m_i\}$ satisfy* (6.2.1) *and m_n^* be defined by* (6.2.2). *Then the maximum $M_n = \max\{\eta_1, \ldots, \eta_n\}$ satisfies*

$$P\{a_n(M_n - b_n - m_n^*) \le x\} \to \exp(-e^{-x}), \tag{6.2.3}$$

a_n and b_n being given by (1.7.2).

PROOF. Write $u_{ni} = u_n + m_n^* - m_i$ where $u_n = x/a_n + b_n$. Then the left-hand side of (6.2.3) may be written as

$$P\left(\bigcap_{i=1}^{n} \{\eta_i \le u_n + m_n^*\}\right) = P\left(\bigcap_{i=1}^{n} \{\xi_i \le u_{ni}\}\right).$$

Since $|m_n^*| < \beta_n$ for sufficiently large n, and $u_n \sim (2\log n)^{1/2}$ (cf. (4.3.2) (ii)), it follows that $\min_{1\le i\le n} u_{ni} = (2\log n)^{1/2}(1 + o(1))$. Thus if it is shown that (6.1.3) holds with $\tau = e^{-x}$ the result will follow from Theorem 6.1.3. To see that (6.1.3) holds we note that since $\min_{1\le i\le n} u_{ni} \to \infty$ as $n \to \infty$,

$$\begin{aligned}\sum_{i=1}^{n}(1 - \Phi(u_{ni})) &\sim \frac{1}{\sqrt{2\pi}}\sum_{i=1}^{n}\frac{\exp(-u_{ni}^2/2)}{u_{ni}}\\ &\sim \frac{1}{\sqrt{2\pi}}\left(\frac{\exp(-u_n^2/2)}{u_n}\right)\sum_{i=1}^{n}\exp(u_n(m_i - m_n^*) - \tfrac{1}{2}(m_i - m_n^*)^2)\end{aligned} \tag{6.2.4}$$

since

$$\left|\frac{u_{ni}}{u_n} - 1\right| = \left|\frac{(m_n^* - m_i)}{u_n}\right| \le \frac{K\beta_n}{(\log n)^{1/2}} \to 0$$

uniformly in $i \le n$. But also

$$\begin{aligned} |u_n(m_i - m_n^*) - a_n^*(m_i - m_n^*)| &\le 2|u_n - a_n^*|\beta_n \\ &\le \frac{K\beta_n}{(\log n)^{1/2}} \end{aligned}$$

using the explicit forms of a_n, b_n from (1.7.2) in $u_n = x/a_n + b_n$. Hence by (6.2.1) and (6.2.4),

$$\begin{aligned} \sum_{i=1}^{n}(1 - \Phi(u_{ni})) &\sim n(1 - \Phi(u_n))\frac{1}{n}\sum_{i=1}^{n}\exp(a_n^*(m_i - m_n^*) - \tfrac{1}{2}(m_i - m_n^*)^2) \\ &\to \tau = e^{-x} \end{aligned}$$

by (6.2.2) since $n(1 - \Phi(u_n)) \to \tau$. Hence (6.1.3) holds and the proof of the theorem is complete. □

As noted above, it is sometimes possible to obtain an explicit expression for m_n^*. For example, suppose that the m_i are bounded and $\max m_i = \beta$. Suppose that v_n of $m_1, \ldots, m_n$ are equal to β, where $v_n \sim n$. Then m_n^* may be taken to be $a_n^{-1}\log(n^{-1}\sum_1^n e^{m_i a_n}) = m_n^{*(1)}$, say. For it is readily checked that $m_n^{*(1)} = \beta + o(a_n^{-1})$ and that the choice $m_n^* = m_n^{*(1)}$ satisfies (6.2.2). This is the formula for m_n^* given by Horowitz (1980) which does in fact hold in this and similar cases.

However, if it is just assumed that the m_i are bounded the choice of m_n^* as $m_n^{*(1)}$ no longer satisfies (6.2.2) and indeed does not suffice in (6.2.3). A simple example of such a case may be constructed by taking v_n of $m_1, \ldots, m_n$ to be $+1$ and the remaining $(n - v_n)$ to be -1 where $v_n \sim ne^{-a_n}$. It is then simply seen that $m_n^* = o(a_n^{-1})$, the limit of the left-hand side of (6.2.2) is $e^{-1/2}$, and indeed that $\sum_{i=1}^n (1 - \Phi(u_n + m_n^* - m_i)) \to e^{-x-1/2}$ so that it is readily seen that (6.2.3) does not hold. (The modification to m_n^* required in this case is of course obvious.)

In cases where the m_i are bounded (or, more generally, when $\beta_n = o((\log n)^{1/6})$) an appropriate form for m_n^* turns out to be

$$m_n^* = m_n^{*(2)} = m_n^{*(1)} + a_n^{-1}\log k_n^{(1)},$$

where

$$k_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n}\exp(a_n(m_i - m_n^{*(1)}) - \tfrac{1}{2}(m_i - m_n^{*(1)})^2).$$

Repetitions of this procedure in the obvious way can give explicit expressions for m_n^* in cases where the m_i grow more rapidly (but are still subject to $\beta_n = o((\log n)^{1/2})$).

6.3. Convergence of $P(\bigcap_{i=1}^n \{\xi_i \le u_{ni}\})$ Under Weakest Conditions on $\{u_{ni}\}$

We return now to the generalization of Theorem 6.1.3 by removing the restriction $\min_{1 \le i \le n} u_{ni} > c(\log n)^{1/2}$, requiring only that $\min_{1 \le i \le n} u_{ni} \to \infty$. As noted, this was done by Hüsler (1981) for stationary normal sequences as a particular case of a theorem for stationary (but not necessarily normal) sequences. A key technique of that work involved the combining of the u_{ni} into groups whose members are of comparable size. Here we use a somewhat simplified version of that technique in obtaining our result for nonstationary normal sequences. This involves grouping the u_{ni}, $1 \le i \le n$, in sets of comparable values in the following way. Write $c_1 = \lambda_n = \min_{1 \le i \le n} u_{ni}$ and partition the integers $(1, \dots, n)$ into disjoint subsets $J_1, \dots, J_L$ as follows. Define

$$J_1 = \{i; c_1 \le u_{ni} \le 2c_1\}, \qquad d_1 = \max_{i \in J_1} u_{ni}, \qquad c_2 = \min\{u_{ni}; u_{ni} > d_1\},$$

$$J_2 = \{i; c_2 \le u_{ni} \le 2c_2\}, \qquad d_2 = \max_{i \in J_2} u_{ni}, \qquad c_3 = \min\{u_{ni}; u_{ni} > d_2\},$$

and so on, until a set J_L is obtained with $\max_{1 \le i \le n} u_{ni} \in J_L$. Thus J_k is a nonempty subset of the integers $(1, \dots, n)$ such that the minimum and maximum values of u_{ni} for $i \in J_k$ are c_k, d_k, respectively, where $d_k \le 2c_k$. Clearly also $c_{k+1} > 2c_k$ for each k. Finally by way of notation write

$$P_k = \sum_{i \in J_k} (1 - \Phi(u_{ni})), \qquad k = 1, 2, \dots, L.$$

Of course, L, c_k, d_k, P_k, and the sets J_k all depend on n, but this dependence is suppressed in the notation, which will be used without comment for the remainder of this section. The first lemma shows that sets J_k, making a relatively small contribution to $\sum (1 - \Phi(u_{ni}))$, can be discarded.

Lemma 6.3.1. *With the above notation write $A = \{k; P_k \ge \exp(-c_k^2/4), 1 \le k \le L\}$ and suppose that $\lambda_n = \min_{1 \le i \le n} u_{ni} \to \infty$ and $\sum_{i=1}^n (1 - \Phi(u_{ni}))$ is bounded. Then*

$$0 \le P\left(\bigcap \{\xi_i \le u_{ni}\}; i \in \bigcup_{k \in A} J_k\right) - P\left(\bigcap_{i=1}^n \{\xi_i \le u_{ni}\}\right) \to 0 \quad \textit{as } n \to \infty, \tag{6.3.1}$$

and

$$0 \le \prod\left\{\Phi(u_{ni}); i \in \bigcup_{k \in A} J_k\right\} - \prod_{i=1}^n \Phi(u_{ni}) \to 0 \quad \textit{as } n \to \infty. \tag{6.3.2}$$

PROOF. The difference in probabilities in (6.3.1) is clearly non-negative and dominated by

$$P\left(\bigcup_{k\notin A}\bigcup_{i\in J_k}\{\xi_i > u_{ni}\}\right) \le \sum_{k\notin A} P_k. \tag{6.3.3}$$

But if $k \notin A$ and $k > 1$, since $c_k > 2c_{k-1}$,

$$P_k < \exp\left(-\frac{c_k^2}{4}\right) < \exp(-c_{k-1}^2) = c_{k-1}\exp(-\tfrac{1}{2}c_{k-1}^2)\frac{\exp(-\frac{1}{2}c_{k-1}^2)}{c_{k-1}}$$
$$\le K\lambda_n \exp(-\tfrac{1}{2}\lambda_n^2)(1-\Phi(c_{k-1})) \le K\lambda_n \exp(-\tfrac{1}{2}\lambda_n^2)P_{k-1}$$

since $xe^{-x^2/2}$ decreases for large x, and $1-\Phi(c_{k-1})$ is one of the terms contributing to P_{k-1}. But if $1 \notin A$, $P_1 \le \exp(-\lambda_n^2/4)$ so that

$$\sum_{k\notin A} P_k \le K\lambda_n \exp(-\tfrac{1}{2}\lambda_n^2)\sum_{k=1}^{L} P_k + \exp\left(-\frac{\lambda_n^2}{4}\right) \to 0$$

since $\sum_{k=1}^{L} P_k = \sum_{i=1}^{n}(1-\Phi(u_{ni}))$ is bounded. Hence (6.3.1) follows from (6.3.3). The second conclusion (6.3.2) is also immediate from the fact that (6.3.1) applies in particular to independent normal random variables—no assumption having been made about the correlation structure in this lemma. □

In view of this lemma it is sufficient, in obtaining (6.1.2), to show that

$$P\left(\bigcap\{\xi_i \le u_{ni}\}; i \in \bigcup_{k\in A} J_k\right) - \prod\left\{\Phi(u_{ni}); i \in \bigcup_{k\in A} J_k\right\} \to 0. \tag{6.3.4}$$

This will be shown by means of the next two lemmas.

Lemma 6.3.2. *Suppose that the correlations r_{ij} satisfy $|r_{ij}| \le \rho_{|i-j|}$ where $\rho_n \to 0$ as $n \to \infty$. Suppose that $\lambda_n = \min_{1\le i\le n} u_{ni} \to \infty$ and that $\sum_{i=1}^{n}(1-\Phi(u_{ni}))$ is bounded. Then for $k, m \in A$, $k \le m$,*

$$S_{k,m} = \sum_{\substack{\theta_n < |i-j| < \gamma_{n,m}\\ i\in J_k,\, j\in J_m,\, i<j}} |r_{ij}|\exp\left(-\frac{\frac{1}{2}(u_{ni}^2+u_{nj}^2)}{1+|r_{ij}|}\right) \le K\exp\left(-\frac{\lambda_n^2}{16}\right)P_k P_m,$$

where $\theta_n = e^{\eta\lambda_n^2}$ for η as in Lemma 6.1.2, and $\gamma_{n,m} = e^{c_m^2/8}$.

PROOF. Writing again $\delta_x = \sup_{j\ge x}\rho_j$ we have

$$S_{k,m} \le \delta_{\theta_n}\sum_{\substack{|i-j|<\gamma_{n,m}\\ i\in J_k,\, j\in J_m,\, i<j}} \exp\left(-\frac{\frac{1}{2}(u_{ni}^2+u_{nj}^2)}{1+\delta_{\theta_n}}\right).$$

Now the exponential term does not exceed

$$\exp(-\tfrac{1}{2}(u_{ni}^2+u_{nj}^2)(1-\delta_{\theta_n})) \le \exp(-\tfrac{1}{2}u_{ni}^2)\exp(\tfrac{1}{2}d_k^2\delta_{\theta_n} - \tfrac{1}{2}c_m^2(1-\delta_{\theta_n}))$$
$$\le \exp(-\tfrac{1}{2}u_{ni}^2)\exp(-c_m^2(\tfrac{1}{2}-\tfrac{5}{2}\delta_{\theta_n}))$$

since $d_k \le 2c_k \le 2c_m$. Multiplying this last bound by d_k/u_{ni} (≥ 1) and summing over the indicated range for (i, j) we obtain

$$S_{k,m} \le K\gamma_{n,m}\, d_k P_k \exp(-c_m^2(\tfrac{1}{2} - \tfrac{5}{2}\delta_{\theta_n}))$$

which gives, since $P_m \ge \exp(-c_m^2/4)$ and $\gamma_{n,m} = \exp(c_m^2/8)$, $d_k \le 2c_k \le 2c_m$,

$$S_{k,m} \le Kc_m \exp(-c_m^2(\tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{8} - \tfrac{5}{2}\delta_{\theta_n}))P_k P_m \le K \exp\left(-\frac{\lambda_n^2}{16}\right)P_k P_m,$$

for some K, since $c_m \ge \lambda_n$ and $\delta_{\theta_n} \to 0$. □

Lemma 6.3.3. *Suppose that* $|r_{ij}| \le \rho_{|i-j|}$ *where* $\rho_n \log n \to 0$ *as* $n \to \infty$, $\lambda_n = \min_{1 \le i \le n} u_{ni} \to \infty$ *and* $\sum_{i=1}^n (1 - \Phi(u_{ni}))$ *is bounded. Then, for* $k \le m$,

$$S_{k,m}^1 = \sum_{\substack{\gamma_{n,m} < |i-j| \le n \\ i \in J_k,\, j \in J_m}} |r_{ij}| \exp\left(-\frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + |r_{ij}|}\right) \le \beta_n P_k P_m,$$

where $\gamma_{n,m} = e^{c_m^2/8}$ *and* $\beta_n \to 0$, β_n *depending only on n.*

PROOF. With $\delta_x = \sup_{j \ge x} \rho_j$ we have by similar arguments to those above, and writing $\gamma_{n,m} = \gamma_n$,

$$\begin{aligned} S_{k,m}^1 &\le \delta_{\gamma_n}\left\{\sum_{i \in J_k} \exp(-\tfrac{1}{2}u_{ni}^2(1 - \delta_{\gamma_n}))\right\}\left\{\sum_{j \in J_m} \exp(-\tfrac{1}{2}u_{nj}^2(1 - \delta_{\gamma_n}))\right\} \\ &\le K\delta_{\gamma_n} d_k\, d_m P_k P_m \exp(\tfrac{1}{2}(d_k^2 + d_m^2)\delta_{\gamma_n}) \\ &\le K\delta_{\gamma_n} d_m^2 \exp(d_m^2\, \delta_{\gamma_n})P_k P_m. \end{aligned}$$

Now

$$\delta_{\gamma_n} d_m^2 \le 4\delta_{\gamma_n} c_m^2 = 32\delta_{\gamma_n} \log \gamma_n,$$

where $\gamma_n = \gamma_{n,m} = \exp(c_m^2/8) \ge \exp(\lambda_n^2/8) \to \infty$, from which the result follows. □

By collecting these lemmas we can now show that Theorem 6.1.3 remains valid if the requirement that $\lambda_n = \min_{1 \le i \le n} u_{ni}$ be bounded below by $c(\log n)^{1/2}$ is replaced by the simple condition $\lambda_n \to \infty$.

Theorem 6.3.4. *Let the correlations* r_{ij} *of the normal sequence* $\{\xi_n\}$ *satisfy* $|r_{ij}| \le \rho_{|i-j|}$ *for* $i \ne j$ *where* $\rho_n < 1$ *for all* $n \ge 1$ *and* $\rho_n \log n \to 0$ *as* $n \to \infty$. *Let the constants* $\{u_{ni}\}$ *be such that* $\sum_{i=1}^n (1 - \Phi(u_{ni}))$ *is bounded and* $\lambda_n = \min_{1 \le i \le n} u_{ni} \to \infty$. *Then* (6.1.2) *holds. If also* (6.1.3) *holds for some* $\tau > 0$, *so does* (6.1.1), *i.e.* $P(\bigcap_{i=1}^n \{\xi_i \le u_{ni}\}) \to e^{-\tau}$.

PROOF. Letting A be as defined in Lemma 6.3.1 we see from Lemmas 6.1.2, 6.3.2, and 6.3.3 that

$$\sum_{\substack{1 \le i < j \le n \\ i, j \in \bigcup_{k \in A} J_k}} |r_{ij}| \exp\left(-\frac{\frac{1}{2}(u_{ni}^2 + u_{nj}^2)}{1 + |r_{ij}|}\right) \le \varepsilon_n \sum_{k, m \in A} P_k P_m + o(1),$$

where $\varepsilon_n \to 0$ as $n \to \infty$. But $\sum_{k,m \in A} P_k P_m \leq \{\sum_{i=1}^{n} (1 - \Phi(u_{ni}))\}^2$ which is bounded so that the left-hand side of (6.3.5) tends to zero as $n \to \infty$. It thus follows from Theorem 4.2.1 that (6.3.4) holds and hence so does (6.1.2) by Lemma 6.3.1.

Finally if also (6.1.3) holds then so does (6.1.1) by Lemma 6.1.1. □

6.4. Stationary Normal Sequences with Strong Dependence

We turn, in this and the remaining sections, to the second topic of the chapter, to see the effect on extremes of a more persistent dependence structure in a stationary normal sequence $\{\xi_n\}$. It was shown in Chapters 4 and 5 that $M_n = \max\{\xi_1, \ldots, \xi_n\}$ has a Type I limiting distribution, and the numbers of exceedances in disjoint intervals are asymptotically independent, provided the covariances r_n decay to zero at a rate which is not too slow. Specifically, the crucial conditions needed for these results concern the behavior of $r_n \log n$; they hold if $r_n \log n \to 0$ or, in somewhat more general circumstances, if $r_n \log n$ is not too large, too often. Results of Mittal and Ylvisaker (1975), which will be given below, show that these conditions are almost the best possible ones. For example, if $r_n \log n \to \gamma > 0$ then $a_n(M_n - b_n)$ does not tend in distribution to $\exp(-e^{-x})$ but to a convolution of $\exp(-e^{-x})$ and a normal distribution function, and further if $r_n \log n \to \infty$ in a sufficiently smooth manner (but $r_n \to 0$ still) then a different normalization is needed, and the limiting distribution is normal.

In this and the next section we will consider the case $r_n \log n \to \gamma > 0$, and use ideas from Mittal and Ylvisaker (1975) to show that then the point process of exceedances of the (usual) level u_n converges weakly to a Cox process (i.e. a mixture of Poisson processes with different intensities). The slow decay of the correlations not only changes the limiting distribution of extremes, but also destroys the asymptotic independence between extreme values in disjoint intervals. The reason for this is explained in an instructive way in the proof of Theorem 6.5.1 below, where the limiting distribution of the exceedances of u_n is obtained as the limiting distribution of the exceedances by an independent normal sequence of a random level $(x + \gamma - \sqrt{2\gamma}\zeta)/a_n + b_n$, where ζ is a standard normal variable representing "the common part" of the first n dependent variables.

The main tool for the proof will, as in Chapter 4, be the Normal Comparison Lemma (Theorem 4.2.1) which relates the distributions of the maxima of two normal sequences with different correlations. Now it is, of course, no longer sufficient to compare with an independent sequence. Instead it will be convenient to compare the distribution of M_n with that of the maximum $M_n(\delta)$ of n standard normal variables which have constant covariance

$\delta \geq 0$ between any two variables. The usefulness of this comparison stems from the fact that if $\zeta, \zeta_1, \zeta_2, \ldots$ are independent standard normal variables then $(1 - \delta)^{1/2}\zeta_1 + \delta^{1/2}\zeta, \ldots, (1 - \delta)^{1/2}\zeta_n + \delta^{1/2}\zeta$ have constant covariance δ between any two, and thus $M_n(\delta)$ has the same distribution as $(1 - \delta)^{1/2}M_n(0) + \delta^{1/2}\zeta$. The following lemma, akin to Lemma 4.3.2, will enable the use of the desired comparison to be given in the next section.

Lemma 6.4.1. *Let $d > 0$ and $\gamma \geq 0$ be constants, put $\rho_n = \gamma/\log n$ and suppose that*

$$r_n \log n \to \gamma \quad \text{as } n \to \infty. \tag{6.4.1}$$

Then, for any sequence $\{u_n\}$ such that $n(1 - \Phi(u_n))$ is bounded

$$nd\sum_{k=1}^{[nd]} |r_k - \rho_n| \exp\left(-\frac{u_n^2}{1 + w_k}\right) \to 0 \quad \text{as } n \to \infty, \tag{6.4.2}$$

where $w_k = \max(\rho_n, |r_k|)$.

Proof. As in the proof of Lemma 4.3.2 we may assume that $n(1 - \Phi(u_n))$ converges (so that (4.3.4)(i) and (ii) hold). Put $\delta(k) = \sup_{k<m\leq[nd]} w_m$. Clearly w_k, and hence $\delta(k)$, also depend on n, but we do not make this dependence explicit in the notation. Further, let α be such that $0 < \alpha < (1 - \delta(0))/(1 + \delta(0))$ for all sufficiently large n (which clearly is possible since $\sup_{k>0} r_k < 1$), and let $p = [n^\alpha]$.

As in the proof of Lemma 4.3.2, the contribution to the sum in (6.4.2) from terms up to p tends to zero, so we have only to prove that the remaining part of the sum also tends to zero. Now

$$\begin{aligned} n \sum_{k=p+1}^{[nd]} |r_k - \rho_n| \exp\left(-\frac{u_n^2}{1 + w_k}\right) &\leq n \exp\left(-\frac{u_n^2}{1 + \delta(p)}\right) \sum_{k=p+1}^{[nd]} |r_k - \rho_n| \\ &= \frac{n^2}{\log n} \exp\left(-\frac{u_n^2}{1 + \delta(p)}\right) \frac{\log n}{n} \sum_{k=p+1}^{[nd]} |r_k - \rho_n|. \end{aligned} \tag{6.4.3}$$

Since $r_n \log n \to \gamma$ there is a constant C such that $r_n \log n \leq C$, $n \geq 1$. Hence also $\delta(p) \log p \leq C$ so by (4.3.4) we have (letting K be a constant whose value may change from line to line)

$$\begin{aligned} \frac{n^2}{\log n} \exp\left(-\frac{u_n^2}{1 + \delta(p)}\right) &\leq \frac{n^2}{\log n} \exp\left(-\frac{u_n^2}{1 + C/\log n^\alpha}\right) \\ &\sim K \frac{n^2}{\log n} \left(\frac{u_n}{n}\right)^{2/(1 + C/\log n^\alpha)} \\ &\leq K n^{(2C/\log n^\alpha)/(1 + C/\log n^\alpha)} = O(1) \end{aligned} \tag{6.4.4}$$

as $n \to \infty$. Moreover, adding and subtracting $\rho_n \log n/\log k = \gamma/\log k$ and using the fact that $\log k \geq n^\alpha$ for $k > p$, gives

$$\frac{\log n}{n} \sum_{k=p+1}^{[nd]} |r_k - \rho_n| \leq \frac{1}{\alpha n} \sum_{k=p+1}^{[nd]} |r_k \log k - \gamma| + \gamma \frac{1}{n} \sum_{k=p+1}^{[nd]} \left|1 - \frac{\log n}{\log k}\right|. \tag{6.4.5}$$

Here the first term to the right tends to zero by (6.4.1). Furthermore, estimating the second sum by an integral, we obtain

$$\begin{aligned}\frac{1}{n} \sum_{k=p+1}^{[nd]} \left|1 - \frac{\log n}{\log k}\right| &\leq \frac{1}{\alpha \log n} \sum_{k=p+1}^{[nd]} \left|\log \frac{k}{n}\right| \frac{1}{n} \\ &= O\left(\frac{1}{\alpha \log n} \int_0^d |\log x| \, dx\right),\end{aligned}$$

and hence the left-hand side of (6.4.5) tends to zero. Since by (6.4.4), the first factor on the right of (6.4.3) is bounded, this concludes the proof of (6.4.2). □

It is possible to weaken the hypothesis of Lemma 6.4.1, and thus of Theorem 6.5.1 below, in the same way as (4.1.1) is weakened to (4.5.3). However, this is quite straightforward and is left to the reader.

6.5. Limits for Exceedances and Maxima when $r_n \log n \to \gamma < \infty$

In this section we investigate the limiting behavior of the point process of exceedances of the level u_n when (6.4.1) holds with $\gamma < \infty$, obtaining the asymptotic distribution of the maximum as a corollary. First recall from Chapter 5 the notation N_n for the point process of exceedances of the level u_n by the process η_n, where η_n is defined from the stationary sequence $\{\xi_j\}$ by $\eta_n(j/n) = \xi_j, j = 1, 2, \ldots; n = 1, 2, \ldots$. Further, let N be a Cox process (cf. Appendix) with (stochastic) intensity $\exp(-x - \gamma + \sqrt{2\gamma}\zeta)$, where ζ is a standard normal random variable, i.e. let N have the distribution determined by

$$\begin{aligned}P\left(\bigcap_{i=1}^{k} \{N(B_i)) = k_i\}\right) &= \int_{-\infty}^{\infty} \prod_{i=1}^{k} \left\{\frac{(m(B_i)\exp(-x - \gamma + \sqrt{2\gamma z}))^{k_i})}{k_i!}\right. \\ &\quad \left. \times \exp\{-m(B_i)e^{-x-\gamma+\sqrt{2\gamma z}}\}\right\}\phi(z)\, dz \end{aligned} \tag{6.5.1}$$

for $B_1, \dots, B_k$ disjoint positive Borel sets (and $m(B)$ denoting Lebesgue measure).

Theorem 6.5.1. *Suppose that $\{\xi_n\}$ is a stationary normal sequence with covariances $\{r_n\}$ and that $u_n = x/a_n + b_n$, with $a_n = (2 \log n)^{1/2}$ and $b_n = a_n - (2a_n)^{-1}(\log\log n + \log 4\pi)$. If $r_n \log n \to \gamma > 0$, then the point process N_n of time-normalized exceedances of the level u_n converges in distribution to N on $(0, \infty)$, where N is the Cox process defined by* (6.5.1).

PROOF. Again we have to verify (a) and (b) of Theorem A.1. As in the proof of Theorem 5.2.1, $E(N_n((c, d])) \sim (d - c)e^{-x}$, so since

$$E(N((c, d])) = E((d - c)\exp(-x - \gamma + \sqrt{2\gamma}\zeta)) = (d - c)e^{-x-\gamma}\, e^{(\sqrt{2\gamma})^2/2} = (d - c)e^{-x},$$

the first condition follows immediately.

We use the notation $M_n(c, d) = \max\{\xi_k; [cn] < k \le [dn]\}$ and write $M_n(c, d; \rho)$ for the maximum of the variables with index k, $[cn] < k \le [dn]$, in a normal sequence with constant covariance ρ between any two variables. Letting $c = c_1 < d_1 < \cdots < c_k < d_k = d$ it then follows, as noted in Section 6.4, that $M_n(c_1, d_1; \rho), \dots, M_n(c_k, d_k; \rho)$ have the same distribution as $(1 - \rho)^{1/2}M_n(c_1, d_1; 0) + \rho^{1/2}\zeta, \dots, (1 - \rho)^{1/2}M_n(c_k, d_k; 0) + \rho^{1/2}\zeta$ where $\{M_n(c_i, d_i; 0)\}_{i=1}^k$ and ζ all are independent and ζ is standard normal. Now

$$P\left(\bigcap_{i=1}^k \{N_n((c_i, d_i]) = 0\}\right) = P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i) \le u_n\}\right)$$

and with $\rho_n = \gamma/\log n$ it follows readily from Corollary 4.2.2 and Lemma 6.4.1 that

$$P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i) \le u_n\}\right) - P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i; \rho_n) \le u_n\}\right) \to 0$$

as $n \to \infty$. Thus, to prove (b) of Theorem A.1 it is enough to check that

$$P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i; \rho_n) \le u_n\}\right) \to P\left(\bigcap_{i=1}^k \{N((c_i, d_i]) = 0\}\right). \tag{6.5.2}$$

However,

$$\begin{aligned} &P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i; \rho_n) \le u_n\}\right) \\ &= P\left(\bigcap_{i=1}^k \{(1 - \rho_n)^{1/2}M_n(c_i, d_i; 0) + \rho_n^{1/2}\zeta \le u_n\}\right) \\ &= \int_{-\infty}^{\infty} P\left(\bigcap_{i=1}^k \{M_n(c_i, d_i; 0) \le (1 - \rho_n)^{-1/2}(u_n - \rho_n^{1/2}z)\}\right)\phi(z)\, dz, \end{aligned}$$

and using the expressions $a_n = (2 \log n)^{1/2}$, $b_n = a_n + O(a_n^{-1} \log \log n)$, and $\rho_n = \gamma/\log n$ we obtain

$$\begin{aligned}
(1 - \rho_n)^{-1/2}(u_n - \rho_n^{1/2} z) &= \left(1 + \frac{\rho_n}{2} + o(\rho_n)\right)\left(\frac{x}{a_n} + b_n - \rho_n^{1/2} z\right) \\
&= \frac{x}{a_n} + b_n - \left(\frac{\gamma}{\log n}\right)^{1/2} z + \frac{(\gamma/\log n)(2 \log n)^{1/2}}{2} \\
&\quad + o(a_n^{-1}) \\
&= \frac{x + \gamma - \sqrt{2\gamma} z}{a_n} + b_n + o(a_n^{-1}).
\end{aligned}$$

Hence it follows from Corollary 5.2.2 that, for fixed z,

$$\begin{aligned}
&P\left(\bigcap_{i=1}^{k} \{M_n(c_i, d_i; 0) \le (1 - \rho_n)^{-1/2}(u_n - \rho_n^{1/2} z)\}\right) \\
&\qquad \to \prod_{i=1}^{k} \exp\{-(d_i - c_i) e^{-x-\gamma+\sqrt{2\gamma} z}\},
\end{aligned}$$

and by dominated convergence this proves that

$$\begin{aligned}
&\int_{-\infty}^{\infty} P\left(\bigcap_{i=1}^{k} \{M_n(c_i, d_i; 0) \le (1 - \rho_n)^{-1/2}(u_n - \rho_n^{1/2} z)\}\right)\phi(z)\, dz \\
&\qquad \to \int_{-\infty}^{\infty} \prod_{i=1}^{k} \exp\{-(d_i - c_i) e^{-x-\gamma+\sqrt{2\gamma} z}\} \phi(z)\, dz \\
&\qquad = P\left(\bigcap_{i=1}^{k} \{N((c_i, d_i]) = 0\}\right),
\end{aligned}$$

i.e. that (6.5.2) holds. □

Corollary 6.5.2. *Suppose that the conditions of Theorem* 6.5.1 *are satisfied and that* $B_1, \ldots, B_k$ *are disjoint positive Borel sets whose boundaries have Lebesgue measure zero. Then* $P(\bigcap_{i=1}^{k} \{N_n(B_i) = k_i\})$ *tends to the expression in the right-hand side of* (6.5.1). *In particular,*

$$\begin{aligned}
P\{a_n(M_n - b_n) \le x\} &= P\{N_n((0, 1]) = 0\} \\
&\to \int_{-\infty}^{\infty} \exp(-e^{-x-\gamma+\sqrt{2\gamma} z})\phi(z)\, dz
\end{aligned}$$

as $n \to \infty$.

From this result we see (as noted in Section 6.4), that the asymptotic distribution of the maximum is now the convolution of a Type I with a normal distribution—but still with the "classical" normalizing constants. It may also be noted that it is quite straightforward to extend the above result to deal with crossings of two or more adjacent levels. However, to avoid repetition we will omit the details.

6.6. Distribution of the Maximum when $r_n \log n \to \infty$

In the case when $r_n \log n \to \infty$, the problem of exceedances of a fixed level by the dependent sequence can again be reduced to considering the exceedances of a random level by an independent sequence. But in this case the random part of the level is "too large"; in the limit the independent sequence will have either infinitely many or no exceedances of the random level. Thus it is not possible to find a normalization that makes the point process of exceedances converge weakly to a nontrivial limit, and accordingly we will only treat the one-dimensional distribution of the maximum. Since the derivation of the general result is complicated we shall consider a rather special case, which brings out the main idea of the result of Mittal and Ylvisaker (1975), while avoiding some of the technicalities of proof.

First we note that for $\gamma > 0$, $0 < q < 1$, there is a convex sequence $\{r_n\}_{n=0}^{\infty}$ with $r_0 = 1$, $r_n = \gamma/(\log n)^q$, $n \geq n_0$, for some $n_0 \geq 2$. (This is easy to see since $\gamma/(\log n)^q$ is convex for $n \geq 2$, and decreasing to zero.) By Polya's criterion $\{r_n\}$ is a covariance sequence, and we shall now consider a stationary zero mean normal sequence $\{\xi_n\}$ with this particular type of covariance. Further, since $\{r_n\}_{n=0}^{\infty}$ is convex, for each $n \geq 1$ also $(r_0 - r_n)/(1 - r_n), \ldots, (r_{n-1} - r_n)/(1 - r_n)$, 0, 0, ... is convex, so again according to Polya's criterion there is a zero mean normal sequence $\{\zeta_1^{(n)}, \zeta_2^{(n)}, \ldots\}$ with these covariances. Clearly $\xi_1, \ldots, \xi_n$ have the same distribution as $(1 - r_n)^{1/2}\zeta_1^{(n)} + r_n^{1/2}\zeta, \ldots, (1 - r_n)^{1/2}\zeta_n^{(n)} + r_n^{1/2}\zeta$, where ζ is standard normal and independent of $\{\zeta_k^{(n)}\}$.

Putting $M_n' = \max_{1 \leq k \leq n} \zeta_k^{(n)}$ the distribution of $M_n = \max_{1 \leq k \leq n} \xi_k$ therefore is the same as that of $(1 - r_n)^{1/2}M_n' + r_n^{1/2}\zeta$. This representation is the key to the proof of Theorem 6.6.3 below, but before proceeding to use it we shall prove two lemmas. The first one is a "technical" lemma of a type used several times already. In this a_n, b_n (e.g. as in Theorem 6.5.1) denote the standard normalizing constants.

Lemma 6.6.1. *Let $n' = [n \exp(-(\log n)^{1/2})]$ and, suppressing the dependence on n, let $\rho_k = (r_k - r_n)/(1 - r_n)$, $k = 1, \ldots, n$. Then, for each $\varepsilon > 0$,*

$$n \sum_{k=1}^{n'} |\rho_k - \rho_{n'}| \exp\left\{-\frac{(b_n - \varepsilon r_n^{1/2})^2}{1 + \rho_k}\right\} \to 0 \quad \text{as } n \to \infty. \tag{6.6.1}$$

PROOF. Set $p = [n^\alpha]$ where $0 < \alpha < (1 - r_1)/(1 + r_1) < (1 - \rho_1)/(1 + \rho_1)$. Since r_n is decreasing we have as in the proof of Lemma 4.3.2 that the sum up to p tends to zero, and it only remains to prove that

$$n \sum_{k=p+1}^{n'} |\rho_k - \rho_{n'}| \exp\left\{-\frac{(b_n - \varepsilon r_n^{1/2})^2}{1 + \rho_k}\right\} \to 0 \quad \text{as } n \to \infty. \tag{6.6.2}$$

Now, using (4.3.4) (with $u_n = b_n$),

$$\begin{aligned}
n \sum_{k=p+1}^{n'} \exp\left\{-\frac{(b_n - \varepsilon r_n^{1/2})^2}{1+\rho_k}\right\}
&= n \sum_{k=p+1}^{n'} \exp\left\{-\frac{b_n^2}{2} \cdot 2\,\frac{(1 - \varepsilon r_n^{1/2}/b_n)^2}{1+\rho_k}\right\} \\
&\sim Kn \sum_{k=p+1}^{n'} \left(\frac{b_n}{n}\right)^{2(1-\varepsilon r_n^{1/2}/b_n)^2/(1+\rho_k)} \\
&\le K \log n \sum_{k=p+1}^{n'} \frac{1}{n} \exp\left\{2\left(1 - \frac{(1 - \varepsilon r_n^{1/2}/b_n)^2}{1+\rho_k}\right)\log n\right\}. \qquad (6.6.3)
\end{aligned}$$

Here $1 - (1 - \varepsilon r_n^{1/2}/b_n)^2/(1+\rho_k) \le 1 - (1 - 2\varepsilon r_n^{1/2}/b_n)/(1+\rho_k) \le \rho_k + 2\varepsilon r_n^{1/2}/b_n$, and since

$$r_n^{1/2} b_n^{-1} \log n \le K'(\log n)^{(1-q)/2},$$

we obtain the bound

$$K \log n \exp\{K'(\log n)^{(1-q)/2}\} \sum_{k=p+1}^{n'} \frac{1}{n} \exp(2\rho_k \log n). \qquad (6.6.4)$$

Furthermore, for $k \ge p + 1$ we have $\log k \ge \alpha \log n$, and then

$$\begin{aligned}
\rho_k \log n &= \frac{\gamma/(\log k)^q - \gamma/(\log n)^q}{1 - \gamma/(\log n)^q} \log n \\
&\le K \frac{\log n}{(\alpha \log n)^q}\left\{1 - \left(\frac{\log k}{\log n}\right)^q\right\} \\
&\le K(\log n)^{1-q}\left\{1 - \left(\frac{\log k/n}{\log n} + 1\right)^q\right\} \\
&\le K(\log n)^{-q}\left\{-\log \frac{k}{n}\right\},
\end{aligned}$$

so that

$$\begin{aligned}
\sum_{k=p+1}^{n'} \frac{1}{n} \exp(2\rho_k \log n) &\le \sum_{k=p+1}^{n'} \frac{1}{n}\left\{\exp\left(-\log \frac{k}{n}\right)\right\}^{K(\log n)^{-q}} \\
&\le \int_0^{\exp\{-(\log n)^{1/2}\}} \{\exp(-\log x)\}^{K(\log n)^{-q}}\, dx \\
&= \int_0^{\exp\{-(\log n)^{1/2}\}} x^{-K(\log n)^{-q}}\, dx \\
&= \frac{\exp\{-(\log n)^{1/2}(1 - K(\log n)^{-q})\}}{1 - K(\log n)^{-q}}.
\end{aligned}$$

Together with (6.6.3) and (6.6.4) this implies that (6.6.2) is bounded by

$$\frac{K \log n \exp\{-(\log n)^{1/2}(1 - K(\log n)^{-q}) + K'(\log n)^{(1-q)/2}\}}{1 - K(\log n)^{-q}} \to 0,$$

as $n \to \infty$, which proves (6.6.1). □

Lemma 6.6.2. *For all* $\varepsilon > 0$

$$P(|M'_n - b_n| > \varepsilon r_n^{1/2}) \to 0 \quad \textit{as } n \to \infty,$$

for $M'_n = \max_{1 \le k \le n} \zeta_k^{(n)}$, *with* $\zeta_1^{(n)}, \ldots, \zeta_n^{(n)}$ *standard normal and with covariances* $\{\rho_k\}$ *as defined in Lemma* 6.6.1.

Proof. As above write $M_n(\rho)$ for the maximum of n standard normal variables with constant correlation ρ between any two. By definition, $\rho_k \ge 0$, and hence, by Corollary 4.2.3,

$$P\{M'_n > b_n + \varepsilon r_n^{1/2}\} \le P\{M_n(0) > b_n + \varepsilon r_n^{1/2} a_n/a_n\}.$$

Further, by the definitions, $a_n r_n^{1/2} \to \infty$ as $n \to \infty$, and by Theorem 1.5.3 it follows readily that

$$P\{M'_n > b_n + \varepsilon r_n^{1/2}\} \to 0.$$

To show that

$$P\{M'_n < b_n - \varepsilon r_n^{1/2}\} \to 0 \tag{6.6.5}$$

we place an upper bound on the difference

$$P\{M'_n < b_n - \varepsilon r_n^{1/2}\} - P\{M_n(\rho_{n'}) < b_n - \varepsilon r_n^{1/2}\}$$

by means of (4.2.5) in Corollary 4.2.2. Since $\{\rho_k\}$ is convex and thus decreasing, $\rho_k \le \rho_{n'}$ for $k \ge n'$ (and hence $(\rho_k - \rho_{n'})^+ = 0$ for $n' < k \le n$), and we obtain the upper bound

$$Kn \sum_{k=1}^{n'} (\rho_k - \rho_{n'}) \exp\left\{-\frac{(b_n - \varepsilon r_n^{1/2})^2}{1 + \rho_k}\right\}$$

which tends to zero by Lemma 6.6.1, so that

$$\limsup_{n \to \infty}(P\{M'_n < b_n - \varepsilon r_n^{1/2}\} - P\{M_n(\rho'_n) < b_n - \varepsilon r_n^{1/2}\}) \le 0. \tag{6.6.6}$$

Moreover, $M_n(\rho_{n'})$ has the same distribution as $(1 - \rho_{n'})^{1/2} M_n(0) + \rho_{n'}^{1/2}\zeta$ where $M_n(0)$ and ζ are independent, so that

$$\begin{aligned} P\{M_n(\rho_{n'}) < b_n - \varepsilon r_n^{1/2}\} &= P\{(1 - \rho_{n'})^{1/2} M_n(0) + \rho_{n'}^{1/2}\zeta < b_n - \varepsilon r_n^{1/2}\} \\ &= P\{(1 - \rho_{n'})^{1/2} r_n^{-1/2} a_n^{-1} a_n(M_n(0) - b_n) \\ &\quad + ((1 - \rho_{n'})^{1/2} - 1) b_n r_n^{-1/2} \\ &\quad + \rho_{n'}^{1/2} r_n^{-1/2}\zeta < -\varepsilon\}. \end{aligned} \tag{6.6.7}$$

For n large, using the definition of n',

$$\rho_{n'} = \frac{\gamma/(\log n')^q - \gamma/(\log n)^q}{1 - \gamma/(\log n)^q}$$

$$\sim \frac{\gamma}{(\log n)^q}\left\{\frac{1}{(1 - (\log n)^{-1/2})^q} - 1\right\} \sim \gamma q (\log n)^{-1/2-q}.$$

Thus

$$((1 - \rho_{n'})^{1/2} - 1)b_n r_n^{-1/2} \sim \tfrac{1}{2}\rho_{n'} b_n r_n^{-1/2} \sim q\sqrt{\frac{\gamma}{2}}(\log n)^{-q/2} \to 0$$

as $n \to \infty$ and also $\rho_{n'}^{1/2} r_n^{-1/2} \sim q^{1/2}(\log n)^{-1/4} \to 0$. Moreover,

$$(1 - \rho_{n'})^{1/2} r_n^{-1/2} a_n^{-1} \sim (2\gamma)^{-1/2}(\log n)^{-(1-q)/2} \to 0,$$

and since $a_n(M_n(0) - b_n)$ converges in distribution by Theorem 1.5.3 it follows that the probability in (6.6.7) tends to zero, so that from (6.6.6), $\limsup_{n\to\infty} P\{M_n' < b_n - \varepsilon r_n^{1/2}\} \le 0$, yielding (6.6.5), to complete the proof. $\square$

Theorem 6.6.3. *Suppose that the stationary, standard normal sequence $\{\xi_n\}$ has covariances $\{r_n\}$ with $\{r_n\}_{n=0}^{\infty}$ convex and $r_n = \gamma/(\log n)^q$, with $\gamma \ge 0$, $0 < q < 1$, for $n \ge n_0$, some n_0. Then*

$$P\{r_n^{-1/2}(M_n - (1 - r_n)^{1/2} b_n) \le x\} \to \Phi(x) \quad \textit{as } n \to \infty.$$

PROOF. As was noted just before Lemma 6.6.1, M_n has the same distribution as $(1 - r_n)^{1/2} M_n' + r_n^{1/2}\zeta$, where ζ is standard normal. It now follows simply from Lemma 6.6.2 that

$$P\{r_n^{-1/2}(M_n - (1 - r_n)^{1/2} b_n) \le x\} = P\{(1 - r_n)^{1/2} r_n^{-1/2}(M_n' - b_n) + \zeta \le x\}$$
$$\to P\{\zeta \le x\} = \Phi(x) \quad \text{as } n \to \infty. \qquad \square$$

Of course the hypothesis of Theorem 6.6.3 is very restrictive. The following more general result was proved by McCormick and Mittal (1976). Their proof follows similar lines as the proof of Theorem 6.6.3 above, but the arguments are much more complicated.

Theorem 6.6.4. *Suppose that the stationary standard normal sequences $\{\xi_n\}$ has covariances $\{r_n\}$ such that $r_n \to 0$ monotonically and $r_n \log n \to \infty$, monotonically for large n. Then*

$$P\{r_n^{-1/2}(M_n - (1 - r_n)^{1/2} b_n) \le x\} \to \Phi(x) \quad \textit{as } n \to \infty. \qquad \square$$

In the paper by Mittal and Ylvisaker (1975), where the above result was first proved under the extra assumption that $\{r_n\}_{n=0}^{\infty}$ is convex, it is also shown that the limit distributions in Theorems 6.5.1 and 6.6.4 are by no means the only possible ones; they exhibit a further class of limit distributions which occur when the covariance decreases irregularly. Further interesting related results are given by McCormick (1980b); in particular he obtains a double exponential limit for the "Studentized maximum", i.e. for M_n normalized by the observed mean and standard deviation.

PART III

EXTREME VALUES IN CONTINUOUS TIME

In this part of the work we shall explore extremal and related theory for continuous parameter stationary processes. As we shall see (in Chapter 13) it is possible to obtain a satisfying general theory extending that for the sequence case, described in Chapter 3 of Part II, and based on dependence conditions closely related to those used there for sequences. In particular, a general form of the Extremal Types Theorem will be obtained for the maximum

$$M(T) = \sup\{\xi(t); 0 \leq t \leq T\},$$

where $\xi(t)$ is a stationary stochastic process satisfying appropriate regularity and dependence conditions.

Before presenting this general theory, however, we shall give a detailed development for the case of stationary normal processes, for which very many explicit extremal and related results are known. For mean-square differentiable normal processes, it is illuminating and profitable to approach extremal theory through a consideration of the properties of *upcrossings* of a high level (which are analogous to the exceedances used in the discrete case). The basic framework and resulting extremal results are described in Chapters 7 and 8, respectively.

As a result of this limit theory it is possible to show that the point process of upcrossings of a level takes on an increasingly Poisson character as the level becomes higher. This and more sophisticated Poisson properties are discussed in Chapter 9, and are analogous to the corresponding results for exceedances by stationary normal sequences, given in Chapter 5.

The Poisson results provide asymptotic joint distributions for the locations and heights of any given number of the largest local maxima.

The local behaviour of a stationary normal process near a high-level upcrossing is discussed in Chapter 10, using, in particular, a simple process (the "Slepian model process") to describe the sample paths at such an upcrossing. As an interesting corollary it is possible to obtain the limiting distribution for the heights of excursions by stationary normal processes above a high level, under appropriate conditions.

In Chapter 11 we consider the joint asymptotic behaviour of the maximum and minimum of a stationary normal process, and of maxima of two or more dependent processes. In particular it is shown that—short of perfect correlation between the processes—such maxima are asymptotically independent.

While the mean square differentiable stationary normal processes form a substantial class, there are important stationary normal processes (such as the Ornstein–Uhlenbeck process) which do not possess this property. Many of these have covariance functions of the form $r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha)$ as $\tau \to 0$ for some α, $0 < \alpha < 2$ (the case $\alpha = 2$ corresponds to the mean-square differentiable processes). The extremal theory for these processes is developed in Chapter 12, using more sophisticated methods than those of Chapter 8, for which simple considerations involving upcrossings sufficed.

Finally, Chapter 13 contains the promised general extremal theory (including the Extremal Types Theorem) for stationary continuous-time processes which are not necessarily normal. This theory essentially relies on the discrete parameter results of Part II, by means of the simple device of expressing the maximum of a continuous parameter process in say time $T = n$, an integer, as the maximum of n "submaxima", over fixed intervals, viz.

$$M(n) = \max(\zeta_1, \zeta_2, \ldots, \zeta_n),$$

where $\zeta_i = \sup\{\xi(t); i - 1 \leq t \leq i\}$. It should be noted (as shown in Chapter 13) that the results for stationary normal processes given in Chapters 8 and 12 can be obtained from those in Chapter 13 by specialization. However, since most of the effort required in Chapters 8 and 12 is still needed to verify the general conditions of Chapter 13, and the normal case is particularly important, we have felt it desirable and helpful to first treat normal cases separately.

CHAPTER 7

Basic Properties of Extremes and Level Crossings

We turn our attention now to *continuous parameter* stationary processes. We shall be especially concerned with stationary *normal* processes in this and most of the subsequent chapters but begin with a discussion of some basic properties which are relevant, whether or not the process is normal, and which will be useful in the discussion of extremal behaviour in later chapters.

Our main concern is with upcrossings of a level by stationary processes, and the expected number of such upcrossings—in the particular case of a stationary normal process leading to a celebrated formula, due to S. O. Rice, for the mean number of upcrossings per unit time. The results are also extended, in various ways, to "marked crossings" and to the expected number of local maxima, obtained by considering the downcrossings of the zero level by the derivative process.

7.1. Framework

Consider a stationary process $\{\xi(t); t \geq 0\}$ having a continuous ("time") parameter $t \geq 0$. Stationarity is to be taken in the *strict* sense, i.e. to mean that any group $\xi(t_1), \dots, \xi(t_n)$ has the same distribution as $\xi(t_1 + \tau), \dots, \xi(t_n + \tau)$ for all τ. Equivalently this means that the *finite-dimensional distributions* $F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P\{\xi(t_1) \leq x_1, \dots, \xi(t_n) \leq x_n\}$ are such that $F_{t_1+\tau, \dots, t_n+\tau} = F_{t_1, \dots, t_n}$ for all choices of τ, n, and $t_1, t_2, \dots, t_n$.

It will be assumed throughout, without comment, that for each t, the d.f. $F_t(x)$ of $\xi(t)$ is continuous. It will further be assumed that, with probability one, $\xi(t)$ has continuous sample functions—that is, the functions $\{\xi(t)\}$

are a.s. continuous as functions of $t \geq 0$. Simple sufficient conditions for sample function continuity will be stated in Section 7.3 for normal processes. For the general case the reader is referred to Cramér and Leadbetter (1967, Chapter 4), and Dudley (1973).

Finally, it will be assumed that the basic underlying probability measure space has been *completed*, if not already complete. This means, in particular, that probability-one limits of r.v.'s will themselves be r.v.'s—a fact which will be useful below.

A principal aim in later chapters will be to discuss the behaviour of the maximum

$$M(T) = \sup\{\xi(t); 0 \leq t \leq T\}$$

(which is well defined and attained, since $\xi(t)$ is continuous) especially when T becomes large. It is often convenient to approximate the process $\xi(t)$ by a sequence $\{\xi_n(t)\}$ of processes taking the value $\xi(t)$ at all points of the form jq_n, $j = 0, 1, 2, \ldots$, and being linear between such points, where $q_n \downarrow 0$ as $n \to \infty$. In particular, this is useful in showing that $M(T)$ is a r.v., as the following small result demonstrates.

Lemma 7.1.1. *With the above notation, suppose that $q_n \downarrow 0$ and write $M_n(T) = \max\{\xi(jq_n);\ 0 \leq jq_n \leq T\}$. Then $M_n(T) \to M(T)$ a.s. as $n \to \infty$, and $M(T)$ is a r.v.*

Proof. $M_n(T)$ is the maximum of a finite number of r.v.'s and hence is a r.v. for each n. It is clear from a.s. continuity of $\xi(t)$ that $M_n(T) \to M(T)$ a.s. and hence by completeness $M(T)$ is a r.v. □

We shall also use the notation $M(I)$ to denote the supremum of $\xi(t)$ in any given interval I—of course, it may be similarly shown that $M(I)$ is a r.v.

7.2. Level Crossings and Their Basic Properties

In the discussion of maxima of sequences, exceedances of a level played an important role. In the continuous case a corresponding role is played by the *upcrossings* of a level for which analogous results (such as Poisson limits) may be obtained. To discuss upcrossings, it will be convenient to introduce—for any real u—a class G_u of all functions f which are continuous on the positive real line, and not identically equal to u in any subinterval. It is easy to see that the sample paths of our stationary process $\xi(t)$ are, with probability one, members of G_u. In fact, every interval contains at least one rational point, and hence

$$P\{\xi(\cdot) \notin G_u\} \leq \sum_{j=1}^{\infty} P\{\xi(t_j) = u\},$$

where $\{t_j\}$ is an enumeration of the rational points. Since $\xi(t_j)$ has a continuous distribution by assumption, $P\{\xi(t_j) = u\}$ is zero for every j.

We shall say that the function $f \in G_u$ has a *strict upcrossing* of u at the point $t_0 > 0$ if for some $\varepsilon > 0$, $f(t) \leq u$ in the interval $(t_0 - \varepsilon, t_0)$ and $f(t) \geq u$ in $(t_0, t_0 + \varepsilon)$. The continuity of f requires, of course, that $f(t_0) = u$, and the definition of G_u that $f(t) < u$ at some points $t \in (t_0 - \eta, t_0)$ and $f(t) > u$ at some points $t \in (t_0, t_0 + \eta)$ for each $\eta > 0$.

It will be convenient to enlarge this notion slightly to include also some points as upcrossings where the behaviour of f is less regular. As we shall see, these further points will not appear in practice for the processes considered in the next two chapters, but are useful in the calculations and will often actually occur for the less regular processes of Chapter 12. Specifically we shall say that the function $f \in G_u$ has an *upcrossing* of u at $t_0 > 0$ if for some $\varepsilon > 0$ and all $\eta > 0$, $f(t) \leq u$ for all t in $(t_0 - \varepsilon, t_0)$ and $f(t) > u$ for *some* t (and hence infinitely many t) in $(t_0, t_0 + \eta)$. An example of a nonstrict upcrossing of zero at t_0 is provided by the function $f(t) = t - t_0$ for $t \leq t_0$ and $f(t) = (t - t_0)\sin((t - t_0)^{-1})$ for $t > t_0$.

The following result contains basic simple facts which we shall need in counting upcrossings.

Lemma 7.2.1. *Let $f \in G_u$ for some fixed u. Then,*

(i) *if for fixed $t_1, t_2, 0 < t_1 < t_2$, we have $f(t_1) < u < f(t_2)$, then f has an upcrossing (not necessarily strict) of u somewhere in (t_1, t_2),*
(ii) *if f has an upcrossing of u at t_0 which is not strict, it has infinitely many upcrossings of u in $(t_0, t_0 + \varepsilon)$, for any $\varepsilon > 0$.*

Proof. (i) If $f(t_1) < u < f(t_2)$ with $t_1 < t_2$ write

$$t_0 = \sup\{t > t_1; f(s) \leq u \text{ for all } t_1 \leq s \leq t\}.$$

Clearly $t_1 < t_0 < t_2$ and t_0 is an upcrossing point of u by f.

(ii) If t_0 is an upcrossing point of u by f and $\varepsilon > 0$, there is certainly a point t_2 in the interval $(t_0, t_0 + \varepsilon)$ with $f(t_2) > u$. If t_0 is not a strict upcrossing there must be a point t_1 in (t_0, t_2) such that $f(t_1) < u$. By (i) there is an upcrossing between t_1 and t_2, so that (ii) follows, since $\varepsilon > 0$ is arbitrary. □

Downcrossings (strict or otherwise) may be defined by making the obvious changes, and *crossings* as points which are either up- or downcrossings. Clearly, at any crossing t_0 of u we have $f(t_0) = u$. On the other hand there may be "u-values" (i.e. points t_0 where $f(t_0) = u$) which are not crossings—such as points where f is tangential to u or points t_0 such that $f(t) - u$ is both positive and negative in every right and left neighbourhood of t_0—as for the function $u + (t - t_0)\sin((t - t_0)^{-1})$.

The above discussion applies to the sample functions of the process $\xi(t)$ satisfying the general conditions stated since, as noted, the sample functions belong to G_u with probability one. Write, now, $N_u(I)$ to denote the number of upcrossings of the level u by $\xi(t)$ in a bounded interval I, and $N_u(t) = N_u((0, t])$. We shall also sometimes write $N(t)$ for $N_u(t)$ when no confusion can arise.

In a similar way to that used for maxima, it is convenient to use the "piecewise linear" approximating processes $\{\xi_n(t)\}$ to show that $N_u(I)$ is a r.v. and, indeed, in subsequent calculations as, for example, in obtaining $E(N_u(I))$. This will be seen in the following lemma, where it will be convenient to introduce the notation

$$J_q(u) = \frac{1}{q} P\{\xi(0) < u < \xi(q)\}, q > 0. \tag{7.2.1}$$

Lemma 7.2.2. *Let I be a fixed, bounded interval. With the above general assumptions concerning the stationary process $\{\xi(t)\}$, let $\{q_n\}$ be any sequence such that $q_n \downarrow 0$ and let N_n denote the number of points jq_n, $j = 1, 2, \ldots$ such that both $(j-1)q_n$ and jq_n belong to I, and $\xi((j-1)q_n) < u < \xi(jq_n)$. Then*

(i) $N_n \le N_u(I)$,
(ii) $N_n \to N_u(I)$ *a.s. as* $n \to \infty$ *and hence* $N_u(I)$ *is a (possibly infinite-valued) r.v.,*
(iii) $E(N_n) \to E(N_u(I))$ *and hence* $E(N_u(t)) = t \lim_{q \downarrow 0} J_q(u)$, *(and hence* $E(N_u(t)) = tE(N_u(1))$*).*

Proof. (i) If for some j, $\xi((j-1)q_n) < u < \xi(jq_n)$ it follows from Lemma 7.2.1(i), that $\xi(t)$ has an upcrossing between $(j-1)q_n$ and jq_n so that (i) follows at once.

(ii) Since the distribution of $\xi(jq_n)$ is continuous and the set $\{kq_n; k = 0, 1, 2, \ldots; n = 1, 2, \ldots\}$ is countable, we see that $P\{\xi(kq_n) = u$ for any $k = 0, 1, 2, \ldots; n = 1, 2, \ldots\} = 0$, and hence we may assume that $\xi(kq_n) \ne u$ for any k and n. We may likewise assume that $\xi(t)$ does not take the value u at either endpoint of I and hence that no upcrossings occur at the endpoints.

Now, if for an integer m, we have $N_u(I) \ge m$, we may choose m distinct upcrossings $t_1, \ldots, t_m$ of u by $\xi(t)$ in the interior of I which may, by choice of $\varepsilon > 0$, be surrounded by disjoint subintervals $(t_i - \varepsilon, t_i + \varepsilon)$, $i = 1, 2, \ldots, m$, of I, such that $\xi(t) \le u$ in $(t_i - \varepsilon, t_i)$ and $\xi(\tau) > u$ for some $\tau \in (t_i, t_i + \varepsilon)$. By continuity, τ is contained in an interval—which may be taken as a subinterval of $(t_i, t_i + \varepsilon)$—in which $\xi(t) > u$. For all sufficiently large n this interval must contain a point kq_n.

Thus there are points $lq_n \in (t_i - \varepsilon, t_i)$, $kq_n \in (t_i, t_i + \varepsilon)$ such that $\xi(lq_n) < u < \xi(kq_n)$. For some j with $l < j \le k$ we must thus have $\xi((j-1)q_n) < u < \xi(jq_n)$. Since eventually each interval $(t_i - \varepsilon, t_i + \varepsilon)$ contains such a point jq_n we conclude that $N_n \ge m$ when n is sufficiently large, from which it follows at once that $\liminf_{n\to\infty} N_n \ge N_u(I)$ (finite or not). Since by (i), $\limsup_{n\to\infty} N_n \le N_u(I)$ we see that $\lim_{n\to\infty} N_n = N_u(I)$ as required.

Finally, it is easily seen that N_n is a r.v. for each n (N_n is a finite sum of r.v.'s χ_k, $\chi_k = 1$ if $\xi((k-1)q_n) < u < \xi(kq_n)$, and zero otherwise) so that, by completeness, its a.s. limit $N_u(I)$ is also a r.v., though possibly taking infinite values.

(iii) Since $N_n \to N_u(I)$ a.s., Fatou's lemma shows that $\liminf_{n\to\infty} E(N_n) \geq E(N_u(I))$. If $E(N_u(I)) = \infty$ this shows at once that $E(N_n) \to E(N_u(I))$. But the same result holds, by dominated convergence, if $E(N_u(I)) < \infty$, since $N_n \leq N_u(I)$ and $N_n \to N_u(I)$ a.s.

Finally, if $I = (0, t]$, then I contains $\nu_n \sim tq_n^{-1}$ points jq_n so that, using stationarity,

$$E(N_n) = (\nu_n - 1)P\{\xi(0) < u < \xi(q_n)\} \sim tJ_{q_n}(u).$$

Hence $tJ_{q_n}(u) \to E(N_u(t))$ from which the final conclusion of (iii) follows since the sequence $\{q_n\}$ is arbitrary. □

Corollary 7.2.3. *If $E(N_u(I)) < \infty$, or equivalently if $\liminf_{n\to\infty} J_{q_n}(u) < \infty$ for some sequence $q_n \downarrow 0$, then the upcrossings of u are a.s. strict.*

Proof. If $E(N_u(I)) < \infty$ then $N_u(I) < \infty$ a.s. and the assertion follows from (ii) of Lemma 7.2.1. □

Under mild conditions one can express $E(N_u(1)) = \lim_{q\to 0} J_q(u)$ in a simple and useful integral form. This will in the next section be adapted to normal processes and it will play a major role in the subsequent development. In Section 7.5 it will be further extended to deal with more complicated situations concerning marked crossings.

The first results of this kind were obtained by S. O. Rice (1939, 1944, 1945) for normal processes, by intuitive methods related to those used in this work. The first rigorous proofs made use of a zero-counting device developed by Kac (1943). Under successively weaker conditions the formula was verified for normal processes by Ivanov (1960), Bulinskaya (1961), Ito (1964), and Ylvisaker (1965). The general formulation we shall use here is due to Leadbetter (1966c); see also Marcus (1977).

Theorem 7.2.4. *Suppose $\xi(0)$ and $\zeta_q = q^{-1}(\xi(q) - \xi(0))$ have a joint density $g_q(u, z)$ continuous in u for all z and all sufficiently small $q > 0$, and that there exists a $p(u, z)$ such that $g_q(u, z) \to p(u, z)$ uniformly in u for fixed z as $q \to 0$. Assume furthermore that there is a function $h(z)$ with $\int_0^\infty zh(z)\,dz < \infty$ and $g_q(u, z) \leq h(z)$ for all u, q. Then*

$$E(N_u(1)) = \lim_{q\to 0} J_q(u) = \int_0^\infty zp(u, z)\,dz. \tag{7.2.2}$$

Proof. By writing the event $\{\xi(0) < u < \xi(q)\}$ as $\{\xi(0) < u < \xi(0) + q\zeta_q\} = \{\xi(0) < u\} \cap \{\zeta_q > q^{-1}(u - \xi(0))\}$ we have

$$\begin{aligned} J_q(u) &= q^{-1}P\{\xi(0) < u, \zeta_q > q^{-1}(u - \xi(0))\} \\ &= q^{-1}\int_{x=-\infty}^{u}\int_{y=q^{-1}(u-x)}^{\infty} g_q(x, y)\,dy\,dx. \end{aligned}$$

By change of variables, $x = u - qzv$, $y = z$, $(v = (u - x)/(qy))$, this is equal to

$$\int_{z=0}^{\infty} z \int_{v=0}^{1} g_q(u - qzv, z)\, dv\, dz,$$

where $g_q(u - qzv, z)$ tends pointwise to $p(u, z)$ as $q \to 0$ by the assumptions of uniform convergence and continuity. Since g_q is dominated by $h(z)$ it follows at once that the double integral tends to $\int_0^\infty zp(u, z)\, dz$. □

In many cases the limit $p(u, z)$ in (7.2.2) is simply the joint density of $\xi(0)$ and the derivative $\xi'(0)$. (This holds, for example, for normal processes, as will be seen in the next section.) If we write $p(u)$ and $p(z \mid u)$ for the density of $\xi(0)$ and the conditional density of $\xi'(0)$ given $\xi(0) = u$, respectively, (7.2.2) can then be written as

$$E(N_u(1)) = p(u) \int_0^\infty zp(z \mid u)\, dz = p(u)E(\xi'(0)^+ \mid \xi(0) = u) \qquad (7.2.3)$$

(where $\xi'(0)^+ = \max(0, \xi'(0))$), so that the mean number of upcrossings is given by the density of $\xi(0)$ multiplied by the average positive slope of the sample functions at u.

Finally, in this section we derive two small results concerning the maximum $M(T)$ and the nature of solutions to the equation $\xi(t) = u_0$, which rely only on the assumption that $E(N_u(1))$ is a continuous function of u.

Theorem 7.2.5. *Suppose that $E(N_u(1))$ is continuous at the point u_0 and, as usual that $P\{\xi(t) = u\} = 0$ for all t, so that $\xi(\cdot) \in G_u$ with probability one. Then*

(i) *with probability one, all points t such that $\xi(t) = u_0$ are either (strict) upcrossings or downcrossings,*
(ii) *the distribution of $M(T)$ is continuous at u_0, i.e. $P\{M(T) = u_0\} = 0$.*

Proof. (i) Clearly it suffices to consider just the unit interval. If $\xi(t) = u_0$ but t is neither a (strict) upcrossing nor a (strict) downcrossing it is either a tangency from below or above, i.e. for some $\varepsilon > 0$, $\xi(t) \le u_0$ $(\ge u_0)$ for all $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ or else there are infinitely many upcrossings in $(t_0 - \varepsilon, t_0)$, (and this is precluded by the finiteness of $E(N_u(1))$). Further, for each fixed u the probability of tangencies of u from below is zero. To see this, let B_u be the number of such tangencies of the level u in $(0, 1]$, write $N_u = N_u(1)$, and suppose $N_u + B_u \ge m$, so that there are at least m points $t_1, \dots, t_m$ which are either u-upcrossings or tangencies from below. Since $\xi(\cdot) \in G_u$ with probability one, there is at least one upcrossing of the level $u - 1/n$ just to the left of any t_j, for all sufficiently large n. This implies that $N_{u-1/n} \ge m$ for sufficiently large n, and hence that

$$N_u + B_u \le \liminf_{n \to \infty} N_{u-1/n},$$

and applying Fatou's lemma,

$$E(N_u) + E(B_u) \leq \liminf_{n\to\infty} E(N_{u-1/n}) = E(N_u)$$

if $E(N_u(1))$ is continuous. Since $B_u \geq 0$, we conclude that $B_{u_0} = 0$ (with probability one). A similar argument excludes tangencies from above, and hence all u-values are either up- or downcrossings (and strict).

(ii) Without loss of generality take $T = 1$. Since

$$P\{M(1) = u\} \leq P\{\xi(0) = u\} + P\{\xi(1) = u\} + P\{B_u \geq 1\}$$

the result follows from $P\{B_{u_0} = 0\} = 1$. □

7.3. Crossings by Normal Processes

Up to this point we have been considering a quite general stationary process $\{\xi(t); t \geq 0\}$. We specialize now to the case of a (stationary) *normal* or *Gaussian* process, by which we mean that the joint distribution of $\xi(t_1), \ldots, \xi(t_n)$ is multivariate normal for each choice of $n = 1, 2, \ldots$ and $t_1, t_2, \ldots, t_n$. It will be assumed without comment that $\xi(t)$ has been standardized to have zero mean and unit variance. The covariance function $r(\tau)$ will then be equal to $E(\xi(t)\xi(t+\tau))$.

Obviously $r(\tau)$ is an even function of τ, with $r(0) = E(\xi^2(t)) = 1$. Thus if r is differentiable at $\tau = 0$, its derivative must be zero there. It is of particular interest to us whether r has *two* derivatives at $\tau = 0$. If $r''(0)$ does exist (finite), it must be negative and we write $\lambda_2 = -r''(0)$. The quantity λ_2 is the *second spectral moment*, so called since we have $\lambda_2 = \int_{-\infty}^{\infty} \lambda^2\, dF(\lambda)$, where $F(\lambda)$ is the spectral d.f., i.e. $r(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau}\, dF(\lambda)$. If r is not twice differentiable at zero then $\int_{-\infty}^{\infty} \lambda^2\, dF(\lambda) = \infty$, i.e. $\lambda_2 = \infty$. When $\lambda_2 < \infty$ we have the expansion

$$r(\tau) = 1 - \frac{\lambda_2 \tau^2}{2} + o(\tau^2) \quad \text{as } \tau \to 0. \tag{7.3.1}$$

Furthermore, it may be shown that $\lambda_2 = -r''(0) < \infty$ if and only if $\xi(t)$ is differentiable in quadratic mean, i.e., if and only if there is a process $\{\xi'(t)\}$ such that $h^{-1}(\xi(t+h) - \xi(t)) \to \xi'(t)$ in quadratic mean as $h \to 0$, and that then

$$E(\xi'(t)) = 0, \qquad \operatorname{Var}(\xi'(t)) = -r''(0),$$

$\xi(t)$, $\xi'(t)$ being jointly normal and independent for each t. Furthermore

$$\operatorname{Cov}(\xi'(t), \xi'(t+\tau)) = -r''(\tau).$$

For future use we introduce also

$$\lambda_0 = \int_{-\infty}^{\infty} dF(\lambda) = r(0) = 1 \quad \text{and} \quad \lambda_4 = \int_{-\infty}^{\infty} \lambda^4 \, dF(\lambda),$$

where also $\lambda_4 = r^{(4)}(0)$, when finite. An account of these and related properties may be found in Cramér and Leadbetter (1967, Chapter 9).

To apply the general results concerning upcrossings to the normal case we require that $\xi(t)$ should have a.s. continuous sample paths. It is known (cf. Cramér and Leadbetter (1967, Section 9.5)), that if, as $\tau \to 0$

$$1 - r(\tau) = O(|\log|\tau||^{-a}) \quad \text{for some } a > 1, \tag{7.3.2}$$

it is possible to define the process $\xi(t)$ as a continuous process. This is a very weak condition which will always hold under assumptions to be used here and subsequently—for example, it is certainly guaranteed if r is differentiable at the origin, or even if $1 - r(\tau) \le C|\tau|^\alpha$ for some $\alpha > 0$, $C > 0$.

In the remainder of this and in the next chapters we shall consider a stationary normal process $\xi(t)$, standardized as above, and such that $\lambda_2 < \infty$. Then $\xi(t)$ and $\xi'(t)$ are independent normal with $\mathrm{Var}(\xi'(t)) = \lambda_2$. Their joint density

$$p(u, z) = \frac{1}{2\pi} \lambda_2^{-1/2} \exp\left(-\frac{1}{2}\left(u^2 + \frac{z^2}{\lambda_2}\right)\right)$$

is the limit of the density $g_q(u, z)$ of $\xi(0)$ and $\zeta_q = q^{-1}(\xi(q) - \xi(0))$ appearing in Theorem 7.2.4. In fact $\xi(0)$ and ζ_q are bivariate normal with mean zero and covariance matrix

$$\begin{pmatrix} 1 & q^{-1}(r(q) - r(0)) \\ q^{-1}(r(q) - r(0)) & 2q^{-2}(r(0) - r(q)) \end{pmatrix},$$

and, by (7.3.1), $q^{-1}(r(q) - r(0)) \to 0$, $2q^{-2}(r(0) - r(q)) \to \lambda_2$ as $q \to 0$, which implies the convergence of their density to $p(u, z)$. The dominated convergence required in Theorem 7.2.4 can also be checked, and hence

$$E(N_u(1)) = \int_0^\infty z p(u, z) \, dz = \frac{1}{2\pi} \lambda_2^{1/2} \exp\left(-\frac{u^2}{2}\right).$$

This is Rice's original formula for the mean number of upcrossings in normal processes. Similar reasoning shows that if $\xi(t)$ is not mean square differentiable (i.e. if $\lambda_2 = \infty$) then $E(N_u(1)) = \infty$, as suggested by Rice's formula.

However, we need a slightly more general result about $J_q(u)$, allowing for $u \to \infty$ as $q \to 0$, and we shall prove this directly along similar lines as in the proof of Theorem 7.2.4 and obtain Rice's formula via this route. To simplify notation we use the normal structure explicitly, making a slightly different transformation of variables.

Lemma 7.3.1. *Let $\{\xi(t)\}$ be a (standardized) stationary normal process with $\lambda_2 < \infty$ and write μ $(=\mu(u)) = (1/2\pi)\lambda_2^{1/2}\exp(-u^2/2)$. Let $q \to 0$ and u either be fixed or tend to infinity as $q \to 0$ in such a way that $uq \to 0$. Then*

$$J_q(u) = q^{-1}P\{\xi(0) < u < \xi(q)\} \sim \mu \quad \text{as } q \to 0.$$

Proof. By rewriting the event $\{\xi(0) < u < \xi(q)\}$ as $\{|\xi(0) + \xi(q) - 2u| < \xi(q) - \xi(0)\}$, i.e. as $\{|\zeta_1 - u| < (q/2)\zeta_2\}$ where $\zeta_1 = (\xi(0) + \xi(q))/2$, $\zeta_2 = (\xi(q) - \xi(0))/q$, are uncorrelated, and hence independent, being normal, with respective variances $\sigma_1^2 = (1 + r(q))/2$, $\sigma_2^2 = 2(1 - r(q))/q^2$, we obtain

$$\begin{aligned}
\mu^{-1}J_q(u) &= (\mu q\sigma_2)^{-1}\int_{y=0}^{\infty}\phi\left(\frac{y}{\sigma_2}\right)P\left\{|\zeta_1 - u| < \frac{qy}{2}\right\}dy \\
&= (\mu q\sigma_1\sigma_2)^{-1}\int_{y=0}^{\infty}\int_{x=u-qy/2}^{u+qy/2}\phi\left(\frac{x}{\sigma_1}\right)\phi\left(\frac{y}{\sigma_2}\right)dx\,dy \\
&= \int_{y=0}^{\infty}\frac{y}{\sigma_2^2}\exp\left(-\frac{y^2}{2\sigma_2^2}\right)\left\{\frac{\sigma_2}{2\sigma_1\mu\sqrt{2\pi}}\int_{x=-1}^{1}\phi\left(\frac{u + qxy/2}{\sigma_1}\right)dx\right\}dy.
\end{aligned} \tag{7.3.3}$$

Now the factor in braces in the integrand may be written as

$$\frac{\sigma_2}{2\sigma_1\sqrt{\lambda_2}}\int_{x=-1}^{1}\exp\left\{-\frac{u^2}{2\sigma_1^2}(1 - \sigma_1^2) - \frac{uqxy}{2\sigma_1^2} - \frac{q^2x^2y^2}{8\sigma_1^2}\right\}dx,$$

which by bounded convergence ($\sigma_1 \to 1$, $1 - \sigma_1^2 = \lambda_2 q^2/4 + o(q^2)$, $\sigma_2 \to \sqrt{\lambda_2}$) tends to 1. By replacement of $uqxy$ by $-uqy$ it is seen that the integrand of (7.3.3) is dominated by the integrable function Aye^{-cy^2} (for some constants $A, c > 0$) so that an application of dominated convergence gives

$$\lim_{q\to 0}\mu^{-1}J_q(u) = \int_0^{\infty}\frac{y}{\lambda_2}\exp\left(-\frac{y^2}{2\lambda_2}\right)dy = 1. \qquad \square$$

Rice's result is now an immediate consequence of this lemma.

Theorem 7.3.2 (Rice's Formula). *If $\{\xi(t)\}$ is a (standardized) stationary normal process with finite second spectral moment $\lambda_2(= -r''(0))$ then the mean number of upcrossings of any fixed level u per unit time is finite and given by*

$$E(N_u(1)) = \frac{1}{2\pi}\lambda_2^{1/2}\exp\left(-\frac{u^2}{2}\right). \tag{7.3.4}$$

(Hence also all upcrossings are strict.)

Proof. This follows from the case u fixed, in the above lemma, together with (iii) of Lemma 7.2.2. $\square$

The above discussion has been in terms of upcrossings. Clearly, similar results hold for downcrossings. In particular, the mean number of downcrossings is also given by (7.3.4).

7.4. Maxima of Normal Processes

In discussing the maximum of a stationary normal process $\xi(t)$ we shall find it useful to compare it with a very simple normal process $\xi^*(t)$ whose maximum is easily calculated using properties of upcrossings. Specifically let η, ζ be independent standard normal r.v.'s and define

$$\xi^*(t) = \eta \cos \omega t + \zeta \sin \omega t, \tag{7.4.1}$$

where ω is a fixed positive constant.

It is clear that $\xi^*(t)$ is normal and that $\xi^*(t_1), \dots, \xi^*(t_n)$ are jointly normal for any choice of t_i. (This follows most simply from the observation that $\sum_1^n c_i \xi^*(t_i)$ is normal for any choice of t_i and c_i.) Thus $\xi^*(t)$ is a normal process and $E(\xi^*(t)) = 0$. Its covariance function is calculated at once to be

$$\begin{aligned} r(\tau) &= E\{(\eta \cos \omega t + \zeta \sin \omega t)(\eta \cos \omega(t + \tau) + \zeta \sin \omega(t + \tau))\} \\ &= \cos \omega t \cos \omega(t + \tau) + \sin \omega t \sin \omega(t + \tau) \\ &= \cos \omega\tau. \end{aligned} \tag{7.4.2}$$

Thus $\xi^*(t)$ is strictly stationary, being normal.

Write now $\eta = A \cos \phi$ and $\zeta = A \sin \phi$, with $0 \leq \phi < 2\pi$. Then $\xi^*(t)$ may be written in its standard cosine form,

$$\xi^*(t) = A \cos(\omega t - \phi). \tag{7.4.3}$$

The Jacobian $\partial(\eta, \zeta)/\partial(A, \phi) = A$, and it follows simply that A, ϕ have joint density

$$f_{A,\phi}(x, y) = \frac{1}{2\pi} x \exp\left(-\frac{x^2}{2}\right), \qquad x \geq 0, 0 \leq y < 2\pi,$$

showing that A, ϕ are independent, A having the Rayleigh distribution $x \exp(-x^2/2)$ $(x \geq 0)$ and ϕ being uniform over $[0, 2\pi)$. The sample paths of ξ^* are thus cosine functions with angular frequency ω, and having independent random amplitude A and phase ϕ.

The distribution of the maximum $M^*(T)$ for this process can be obtained geometrically. However, it is more instructive (and simpler) to use properties of upcrossings.

Lemma 7.4.1. *For the cosine process $\xi^*(t)$ given by* (7.4.3)

$$P\{M^*(T) \leq u\} = \Phi(u) - \frac{\omega T}{2\pi} \exp\left(-\frac{u^2}{2}\right) \tag{7.4.4}$$

for $0 < T < \pi/\omega$ and $u > 0$.

PROOF. Clearly $\lambda_2 = \omega^2$ for the process $\xi^*(t)$ and, writing $N = N_u^*(T)$ for the number of upcrossings of u in $(0, T)$, we have

$$E(N) = \frac{\omega T}{2\pi} \exp\left(-\frac{u^2}{2}\right) \tag{7.4.5}$$

and

$$P\{M^*(T) > u\} = P\{\xi^*(0) > u\} + P\{\xi^*(0) \leq u, N \geq 1\}.$$

Now take $\omega T < \pi$. Then if $\xi^*(0) > u > 0$, the first upcrossing of u occurs after $t = \pi/\omega$ (see Figure 7.4.1), and hence $\{N \geq 1, \xi^*(0) > u\}$ is empty, so that

$$P\{\xi^*(0) \leq u, N \geq 1\} = P\{N \geq 1\}.$$

Thus, since $N = 0$ or 1,

$$\begin{aligned} P\{M^*(T) > u\} &= 1 - \Phi(u) + P\{N \geq 1\} = 1 - \Phi(u) + E(N) \\ &= 1 - \Phi(u) + \frac{\omega T}{2\pi} \exp\left(-\frac{u^2}{2}\right), \end{aligned} \tag{7.4.6}$$

which is equivalent to (7.4.4). □

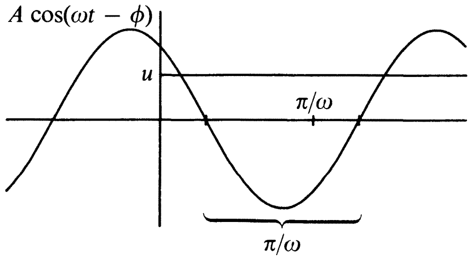


Figure 7.4.1. Upcrossings for the cosine process $\xi^*(t) = A\cos(\omega t - \phi)$.

As a matter of interest and for later use, it follows for the cosine process $\xi^*(t)$ that for fixed h, $0 < \omega h < \pi$,

$$\frac{P\{M^*(h) > u\}}{h\phi(u)} \to \left(\frac{\lambda_2}{2\pi}\right)^{1/2} \quad \text{as } u \to \infty \tag{7.4.7}$$

(since $1 - \Phi(u) \sim \phi(u)/u$ and $\lambda_2 = \omega^2$). This limit in fact holds under much more general conditions, as we shall see.

As noted above we will want to compare in the next chapter a general stationary normal process with this special process. This comparison will be made by an application of the following easy consequence of the Normal Comparison Lemma (cf. Slepian (1962)).

Theorem 7.4.2 ("Slepian's Lemma"). *Let $\{\xi_1(t)\}$ and $\{\xi_2(t)\}$ be normal processes (possessing continuous sample functions but not necessarily being stationary). Suppose that these are standardized so that $E(\xi_1(t)) = E(\xi_2(t)) = 0$, $E(\xi_1^2(t)) = E(\xi_2^2(t)) = 1$, and write $\rho_1(t, s)$ and $\rho_2(t, s)$ for their covariance functions. Suppose that for some $\delta > 0$ we have $\rho_1(t, s) \geq \rho_2(t, s)$ when $0 \leq t, s \leq \delta$. Then the respective maxima $M_1(t)$ and $M_2(t)$ satisfy*

$$P\{M_1(T) \leq u\} \geq P\{M_2(T) \leq u\}$$

when $0 \leq T \leq \delta$.

Proof. Define $M_n^{(1)}$ and $M_n^{(2)}$ relative to $\xi_1(t)$, $\xi_2(t)$ as in Lemma 7.1.1 where $q_n = 2^{-n}$. Then, with probability one $M_n^{(1)} \uparrow M_1(T)$, so that $\{M_n^{(1)} \leq u\} \downarrow \{M_1(T) \leq u\}$ and hence $P\{M_n^{(1)} \leq u\} \to P\{M_1(T) \leq u\}$ as $n \to \infty$. Similarly $P\{M_n^{(2)} \leq u\} \to P\{M_2(T) \leq u\}$. But it is clear from Corollary 4.2.3 that $P\{M_n^{(2)} \leq u\} \leq P\{M_n^{(1)} \leq u\}$ so that the desired result follows. □

7.5. Marked Crossings

The material in the remainder of this chapter will not be used until Chapter 9 and subsequent chapters. The reader may wish to proceed directly to Chapter 8, and return to this section when needed.

We shall consider situations where we not only register the occurence of an upcrossing, but also the value of some other random variable connected with the upcrossing. We may, for example, be interested in the derivative $\xi'(t_i)$ at upcrossing points t_i of u by $\xi(\cdot)$ or the value $\xi(s_i)$ at downcrossing points s_i of zero by $\xi'(\cdot)$, i.e. at points where $\xi(t)$ has a local maximum. We shall refer to these as *marked crossings* and, for example, regard $\xi'(t_i)$ and $\xi(s_i)$ as *marks* attached to the crossings at t_i and s_i. We shall here develop methods for dealing with such marks, along similar lines to those leading to Rice's formula (although with some increase in complexity).

We shall let $\{\zeta(t); t \geq 0\}$ and $\{\eta(t); t \geq 0\}$ be jointly stationary processes with continuous sample paths and consider level crossings in $\zeta(t)$ marked by $\eta(t)$. In the examples above $\zeta(t)$ would be $\xi(t)$ and $-\xi'(t)$, respectively, and $\eta(t)$ would be $\xi'(t)$ and $\xi(t)$. Denote by $\{t_i\}$ the upcrossings of the level u by $\zeta(t)$, and let, for any interval A, $N_u(I; A)$ be the number of t_i's in I such that $\eta(t_i) \in A$, and write $N_u(T; A) = N_u((0, T]; A)$. The notation $N_u(I)$, $N_u(T)$ will have the same meaning as before, e.g. $N_u(I) = N_u(I; (-\infty, \infty))$.

Further define

$$J_q(u; A) = \frac{1}{q} P\{\zeta(0) < u < \zeta(q), \eta(0) \in A\}.$$

Lemma 7.5.1. *Let I be a bounded interval, $q_n \to 0$ as $n \to \infty$, and let $N_n(A)$ be the number of points $jq_n \in I$ (with $(j-1)q_n \in I$) such that*

$$\zeta((j-1)q_n) < u < \zeta(jq_n) \quad \text{and} \quad \eta((j-1)q_n) \in A.$$

Then

(i) *if A is an open interval,*

$$\liminf_{n\to\infty} N_n(A) \geq N_u(I; A), \quad a.s.$$

(ii) *if, for every v,*

$$P\{\zeta(t) = u, \eta(t) = v \text{ for some } t \in I\} = 0 \tag{7.5.1}$$

then, for any interval A,

$$N_u(I; A) = \lim_{n\to\infty} N_n(A), \quad a.s.,$$

(iii) *if A is an open interval,*

$$E(N_u(I; A)) \leq \liminf_{n\to\infty} E(N_n(A))$$

and, if (7.5.1) *holds, and $E(N_u(1)) < \infty$,*

$$E(N_n(A)) \to E(N_u(I; A))$$

and, for $I = (0, 1]$,

$$E(N_u(I; A)) = \lim_{q\downarrow 0} J_q(u; A).$$

PROOF. (i) Suppose that $N_u(I; A) \geq m$ and that $\zeta(t)$ has upcrossings of u at $t_1, \ldots, t_m$ in the interior of I, with $\eta(t_i) \in A$, $i = 1, \ldots, m$. (By the continuity of the distribution of $\zeta(t)$ no upcrossings occur at the endpoints of I.) Since $\eta(t)$ is continuous and A is open, we can surround the t_i's by disjoint subintervals $(t_i - \varepsilon, t_i + \varepsilon)$ of I in which $\eta(t) \in A$. It then follows as in the proof of Lemma 7.2.2(ii) that $\liminf_{n\to\infty} N_n(A) \geq m$.

(ii) First assume $N_u(I; A) = m < \infty$, and let $t_1, \ldots, t_m$ be as in (i). If (a, b) is the interior of A, (7.5.1) precludes $\eta(t_i) = a$ or b, so that $\eta(t_i) \in (a, b)$, and we may therefore take disjoint intervals $(t_i - \varepsilon, t_i + \varepsilon)$ in which $\eta(t) \in (a, b)$. Write J_n for the set of j's such that $(j-1)q_n$ and jq_n both belong to $(t_i - \varepsilon, t_i + \varepsilon)$ for some i, and J_n^* for the set of j's such that $(j-1)q_n$ and jq_n belong to I but $j \notin J_n$. Clearly t_i is the only upcrossing of u by $\zeta(t)$ for $t \in (t_i - \varepsilon, t_i + \varepsilon)$, and therefore by Lemma 7.2.1(i),

$$\limsup_{n\to\infty} \sum_{j\in J_n} \chi_j \leq m, \tag{7.5.2}$$

where

$$\chi_j = \begin{cases} 1, & \text{if } \zeta((j-1)q_n) < u < \zeta(jq_n) \text{ and } \eta((j-1)q_n) \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, if

$$\limsup_{n\to\infty} \sum_{j\in J_n^*} \chi_j > 0,$$

then for n arbitrarily large there are $j_n \in J_n^*$ with $\chi_{j_n} = 1$ and hence a sequence of integers $\{\tilde{n}\}$ such that $j_{\tilde{n}} q_{\tilde{n}} \to \tau$, with $\tau \notin (t_i - \varepsilon, t_i + \varepsilon)$, $i = 1, \ldots, m$, and $\chi_{j_{\tilde{n}}} = 1$. From the continuity of $\eta(t)$ it follows that $\eta(\tau) \in [a, b]$, and hence, by (7.5.1), $\eta(\tau) \in (a, b)$. Thus $\eta(t) \in (a, b) \subset A$ for $t \in (\tau - \varepsilon', \tau + \varepsilon')$ for some $\varepsilon' > 0$, which can be taken small enough to make $t_i \notin (\tau - \varepsilon', \tau + \varepsilon')$, $i = 1, \ldots, m$. Further, for $\tilde{n}$ large enough, both $(j_{\tilde{n}} - 1)q_{\tilde{n}}$ and $j_{\tilde{n}} q_{\tilde{n}}$ belong to $(\tau - \varepsilon', \tau + \varepsilon')$ and thus, by Lemma 7.2.1(i), $\zeta(t)$ has a u-upcrossing in $(\tau - \varepsilon', \tau + \varepsilon')$ which contradicts $N_u(I; A) = m$. This shows that

$$\limsup_{n \to \infty} \sum_{j \in J_n^*} \chi_j = 0,$$

which together with (7.5.2) proves that $\limsup_{n \to \infty} N_n(A) \le N_u(I; A)$, a.s. Since furthermore, (7.5.1) implies that $N_u(I; A) = N_u(I; (a, b))$, part (i) gives

$$\liminf_{n \to \infty} N_n(A) \ge \liminf_{n \to \infty} N_n((a, b)) \ge N_u(I; (a, b)) = N_u(I; A).$$

Hence $N_n(A) \to N_u(I; A) = m < \infty$ a.s. as asserted. If $N_u(I; A) = \infty$, the conclusion follows from part (i) with (a, b) replacing A, since $N_u(I; A) = N_u(I, (a, b))$ by (7.5.1).

(iii) The first conclusion follows at once from Fatou's lemma and part (i), while it follows from part (ii) that $E(N_u(I; A)) = \lim_{n \to \infty} E(N_n(A))$, since $N_n(A) \le N_u(I)$ and $E(N_u(I)) < \infty$ by assumption. Further, if $I = (0, 1)$, there are approximately q_n^{-1} points $jq_n \in I$, so that

$$E(N_n(A)) \sim q_n^{-1} E(\chi_1) = J_{q_n}(u; A).$$

The last assertion of (iii) follows, since the sequence $\{q_n\}$ is arbitrary. □

We shall now evaluate the limit of $J_q(u; A)$ for the case when $\zeta(t)$ and $\eta(t)$ are jointly normal processes with zero means. Let $r(t)$ denote the covariance function of $\zeta(t)$, and write $\lambda_0 = r(0) = \operatorname{Var}(\zeta(t))$. As was noted earlier, if $\zeta(t)$ is quadratic mean differentiable then $\lambda_2 = -r''(0) = \operatorname{Var}(\zeta'(t)) < \infty$ and $\zeta(t)$ and $\zeta'(t)$ are independent for each t and normal with joint density

$$p(u, z) = \lambda_0^{-1/2} \phi(u \lambda_0^{-1/2}) \lambda_2^{-1/2} \phi(z \lambda_2^{-1/2}). \tag{7.5.3}$$

Further it can be shown that the three processes $\{\zeta(t)\}$, $\{\zeta'(t)\}$, and $\{\eta(t)\}$ are jointly normal, and that the crosscovariances and covariances can be obtained as limits, e.g.

$$\operatorname{Cov}(\zeta'(t), \eta(t + \tau)) = \lim_{h \to 0} E(h^{-1}(\zeta(t + h) - \zeta(t))\eta(t + \tau)).$$

Conditional distributions can also be defined, using ratios of density functions when they exist, e.g. for a measurable set A, we define

$$P\{\eta(0) \in A \mid \zeta(0) = u, \zeta'(0) = z\}$$
$$= \int_{y \in A} \frac{p_{\zeta(0), \zeta'(0), \eta(0)}(u, z, y)}{p(u, z)} \, dy,$$

where $p_{\zeta(0),\zeta'(0),\eta(0)}$ is the density function of $\zeta(0)$, $\zeta'(0)$, $\eta(0)$. In the sequel, conditional probabilities will always be understood as defined in this way.

Lemma 7.5.2. *Let $\{\zeta(t)\}$ and $\{\eta(t)\}$ be jointly normal zero mean processes such that $\zeta(0)$, $\zeta'(0)$, $\eta(0)$ have a nonsingular distribution. Assume further that $\{\eta(t)\}$ has continuous sample paths and that $\zeta(t)$ is differentiable in quadratic mean. Then, for any measurable set A and any u,*

$$\lim_{q\downarrow 0} J_q(u; A) = \int_{z=0}^{\infty} zp(u, z)P\{\eta(0) \in A \mid \zeta(0) = u, \zeta'(0) = z\}\, dz.$$

Proof. Write $\eta = \eta(0)$, and as in the proof of Lemma 7.3.1 introduce the independent normal r.v.'s $\zeta_1 = (\zeta(0) + \zeta(q))/2$, $\zeta_2 = (\zeta(q) - \zeta(0))/q$ with variances $\sigma_1^2 = (r(0) + r(q))/2$, $\sigma_2^2 = 2(r(0) - r(q))/q^2$, and note that

$$\begin{aligned} J_q(u; A) &= q^{-1}P\left\{|\zeta_1 - u| < \frac{q\zeta_2}{2}, \eta \in A\right\} \\ &= (q\sigma_1\sigma_2)^{-1} \int_{z=0}^{\infty} \int_{x=u-qz/2}^{u+qz/2} \phi\left(\frac{x}{\sigma_1}\right)\phi\left(\frac{z}{\sigma_2}\right) \\ &\qquad \times P\{\eta \in A \mid \zeta_1 = x, \zeta_2 = z\}\, dx\, dz \\ &= \int_{z=0}^{\infty} \frac{z}{\sigma_2}\phi\left(\frac{z}{\sigma_2}\right) \int_{x=-1}^{1} \frac{1}{2\sigma_1}\phi\left(\frac{u + xqz/2}{\sigma_1}\right) \\ &\qquad \times P\left\{\eta \in A \mid \zeta_1 = u + \frac{xqz}{2}, \zeta_2 = z\right\} dx\, dz. \end{aligned} \tag{7.5.4}$$

To obtain the limit of the conditional normal probability

$$P\{\eta \in A \mid \zeta_1 = v, \zeta_2 = z\}$$

as $q \to 0$ (and $v \to u$) we note that since $\{\zeta(t)\}$ and $\{\eta(t)\}$ are jointly normal processes, the conditional distribution of $\eta = \eta(0)$ given $\zeta_1 = (\zeta(0) + \zeta(q))/2 = v$, $\zeta_2 = (\zeta(q) - \zeta(0))/q = z$, is also normal with mean

$$m_q(v, z) = E(\eta) + v\sigma_1^{-2}\operatorname{Cov}(\eta, \zeta_1) + z\sigma_2^{-2}\operatorname{Cov}(\eta, \zeta_2)$$

and variance

$$V_q = \operatorname{Var}(\eta) - \sigma_1^{-2}\operatorname{Cov}^2(\eta, \zeta_1) - \sigma_2^{-2}\operatorname{Cov}^2(\eta, \zeta_2),$$

see e.g. Rao (1972) p. 522.

Since $\zeta_1 \to \zeta(0)$, $\zeta_2 \to \zeta'(0)$ in quadratic mean as $q \to 0$ it follows that $\operatorname{Cov}(\eta, \zeta_1) \to \operatorname{Cov}(\eta(0), \zeta(0))$, $\operatorname{Cov}(\eta, \zeta_2) \to \operatorname{Cov}(\eta(0), \zeta'(0))$ as $q \to 0$. Since furthermore, $\sigma_1^2 \to \lambda_0 = r(0) = \operatorname{Var}(\zeta(0))$, $\sigma_2^2 \to \lambda_2 = -r''(0) = \operatorname{Var}(\zeta'(0))$, and $\zeta(0), \zeta'(0), \eta(0)$ are nonsingular by assumption we have $\operatorname{Var}(\eta \mid \zeta(0), \zeta'(0)) = V_0 = \lim_{q\to 0} V_q > 0$. With $m_q = m_q(u + kqz/2, z)$, $E(\eta \mid \zeta(0) = u, \zeta'(0) = z) = m_0 = \lim_{q\to 0} m_q$, and thus dominated convergence gives that for all x and z,

$$\begin{aligned} P\left\{\eta \in A \mid \zeta_1 = u + \frac{xqz}{2}, \zeta_2 = z\right\} &= \int_A \frac{1}{\sqrt{V_q}} \phi\left(\frac{y - m_q}{\sqrt{V_q}}\right) dy \\ &\to \int_A \frac{1}{\sqrt{V_0}} \phi\left(\frac{y - m_0}{\sqrt{V_0}}\right) dy \\ &= P\{\eta(0) \in A \mid \zeta(0) = u, \zeta'(0) = z\} \end{aligned}$$

as $q \to 0$. Again by dominated convergence it follows that

$$\begin{aligned} J_q(u; A) \to \int_{z=0}^{\infty} &\frac{z}{\sqrt{\lambda_2}} \phi\left(\frac{z}{\sqrt{\lambda_2}}\right) \frac{1}{\sqrt{\lambda_0}} \phi\left(\frac{u}{\sqrt{\lambda_0}}\right) \\ &\times P\{\eta(0) \in A \mid \zeta(0) = u, \quad \zeta'(0) = z\}\, dz \end{aligned}$$

which by (7.5.3) is the conclusion of the lemma. □

7.6. Local Maxima

As an application of the marked crossings theory we end this chapter with some comments concerning *local maxima*. To avoid technicalities we assume that $\{\xi(t)\}$ is stationary normal and, in addition, has sample functions which are, with probability one, everywhere continuously differentiable. Sufficient conditions for differentiability can be found in Cramér and Leadbetter (1967, Chap. 9), and they require slightly more than finiteness of the second spectral moment λ_2; cf. the condition (7.3.2) for sample function continuity.

Clearly then $\xi(t)$ has a local maximum at t_0 if and only if $\xi'(t)$ has a downcrossing of zero at t_0, and a number of results for local maxima can therefore trivially be obtained from corresponding results for downcrossings.

In particular, to ensure that $\xi(t)$ has only finitely many local maxima in a finite time, it suffices that $\lambda_4 = r^{(4)}(0) < \infty$, where λ_4 is the fourth spectral moment $\int_{-\infty}^{\infty} \lambda^4 \, dF(\lambda)$.

If $\lambda_4 < \infty$, then $\xi(t)$ has also a second derivative $\xi''(t)$, defined in quadratic mean, and $\xi(t)$, $\xi'(t)$, $\xi''(t)$ are jointly normal with mean zero and the covariance matrix

$$\begin{bmatrix} \lambda_0 & 0 & -\lambda_2 \\ 0 & \lambda_2 & 0 \\ -\lambda_2 & 0 & \lambda_4 \end{bmatrix},$$

where we assume $\lambda_0 = 1$. Further $\xi(t)$, $\xi'(t)$, $\xi''(t)$ have a nonsingular distribution provided $\xi(t)$ is not of the form $\xi(t) = A\cos(\omega t - \phi)$. (In fact, the determinant of the covariance matrix is

$$\lambda_2(\lambda_0\lambda_4 - \lambda_2^2) = \lambda_2\left\{\int dF(\lambda)\int \lambda^4\, dF(\lambda) - \left(\int \lambda^2\, dF(\lambda)\right)^2\right\},$$

which is zero only if F is concentrated at two symmetric points.) If $\lambda_4 < \infty$ we also have the analogue of (7.3.1),

$$\operatorname{Cov}(\xi'(t), \xi'(t+\tau)) = -r''(\tau) = \lambda_2 - \tfrac{1}{2}\lambda_4\tau^2 + o(\tau^2) \quad \text{as } \tau \to 0,$$

and, normalizing to variance one, we obtain

$$\operatorname{Cov}(\lambda_2^{-1/2}\xi'(t), \lambda_2^{-1/2}\xi'(t+\tau)) = 1 - \frac{1}{2}\frac{\lambda_4}{\lambda_2}\tau^2 + o(\tau^2) \quad \text{as } \tau \to 0. \tag{7.6.1}$$

We will temporarily use the notation $N'(T)$ for the number of local maxima of $\xi(t)$, $0 < t \le T$. Since $N'(T)$ is just the number of downcrossing zeros for $\xi'(t)$, we obtain from (7.6.1) and Rice's formula (7.3.4) that the expected number of local maxima in $(0, T]$ is

$$E(N'(T)) = \frac{T}{2\pi}\left(\frac{\lambda_4}{\lambda_2}\right)^{1/2}.$$

In Chapter 9 we shall study heights and locations of *high local maxima*. Write $N'_u(T)$ for the number of local maxima of $\xi(t)$, $0 < t < T$, whose height exceeds u, i.e. with the previous notation, if $\xi(t)$ has local maxima at the time points $\{s_i\}$, then $N'_u(T)$ is the number of $s_i \in (0, T)$ such that $\xi(s_i) > u$.

Lemma 7.6.1. *If $\{\xi(t)\}$ is stationary normal, with continuously differentiable sample paths, and with a quadratic mean second derivative $\xi''(t)$ with $\operatorname{Var}(\xi''(t)) = \lambda_4 < \infty$ such that $\xi(t)$, $\xi'(t)$, $\xi''(t)$ have a nonsingular distribution, then*

(i) $$E(N'_u(T)) = T\int_{x=u}^{\infty}\int_{z=-\infty}^{0} |z|\, p(x, 0, z)\, dz\, dx, \tag{7.6.2}$$

where $p(x, y, z)$ is the joint density of $\xi(t)$, $\xi'(t)$, $\xi''(t)$, and

(ii) *assuming $\xi(t)$ to be standardized with mean zero and unit variance*

$$E(N'_u(T)) = \frac{T}{2\pi}\left\{\left(\frac{\lambda_4}{\lambda_2}\right)^{1/2}\left(1 - \Phi\left(u\left(\frac{\lambda_4}{D}\right)^{1/2}\right)\right) + (2\pi\lambda_2)^{1/2}\phi(u)\Phi\left(\frac{u\lambda_2}{D^{1/2}}\right)\right\}, \tag{7.6.3}$$

where $D = \lambda_4 - \lambda_2^2$.

PROOF. We shall use Lemmas 7.5.1 and 7.5.2, identifying $\zeta(t) = -\xi'(t)$ and $\eta(t) = \xi(t)$. By assumption, $\{\zeta(t)\}$ and $\{\eta(t)\}$ satisfy the hypotheses of Lemma 7.5.2, with $\mathrm{Var}(\zeta'(t)) = \lambda_4$, so that for any open interval A,

$$\lim_{q \downarrow 0} J_q(0; A) = \int_{z=0}^{\infty} z f_{\zeta(0),\zeta'(0)}(0, z) P\{\eta(0) \in A \mid \zeta(0) = 0, \zeta'(0) = z\}\, dz$$

$$= \int_{z=-\infty}^{0} |z| p(0, z) P\{\xi(0) \in A \mid \xi'(0) = 0, \xi''(0) = z\}\, dz, \quad (7.6.4)$$

where $f_{\zeta(0),\zeta'(0)}(x, z) = p(-x, -z)$ is the density of $\zeta(0), \zeta'(0)$ $(= -\xi'(0), -\xi''(0))$. By Theorems 7.2.5 and 7.3.2, all t such that $\zeta(t) = \xi'(t) = 0$ are either (strict) upcrossing or downcrossing points. Lemma 7.5.1(iii) implies that, writing $N_0(T; V_\varepsilon)$ for the number of maxima in $(0, T]$ with height in $V_\varepsilon = (v - \varepsilon, v + \varepsilon)$,

$$P\{\xi'(t) = 0, \xi(t) = v \text{ for some } t \in (0, T]\} \le 2E(N_0(T; V_\varepsilon))$$
$$\le 2T \liminf_{q \downarrow 0} J_q(0; V_\varepsilon),$$

since $E(N_n(V_\varepsilon)) \sim J_{q_n}(0; V_\varepsilon)$. By (7.6.4) the right-hand side can be made arbitrarily small by choosing ε small. Thus $\zeta(t), \eta(t)$ satisfy condition (7.5.1) and by Lemma 7.5.1(iii) and stationarity

$$E(N_0(T; (u, \infty))) = TE(N_0(1; (u, \infty))) = T \lim_{q \to 0} J_q(0; (u, \infty)).$$

Inserting

$$P\{\xi(0) \in (u, \infty) \mid \xi'(0) = 0, \xi''(0) = z\} = \int_u^{\infty} \frac{p(x, 0, z)}{p(0, z)}\, dx$$

into (7.6.4), part (i) follows.

Part (ii) follows after some calculation by inserting the normal density

$$p(x, 0, z) = (2\pi)^{-3/2}(\lambda_2 D)^{-1/2} \exp\left\{-\frac{(\lambda_4 x^2 + 2\lambda_2 xz + z^2)}{2D}\right\},$$

into (7.6.2). □

CHAPTER 8
Maxima of Mean Square Differentiable Normal Processes

In this chapter the theory of maxima of mean square differentiable stationary normal processes will be developed under simple conditions—giving analogous results to those of Chapter 4. This will be approached using the properties of upcrossings developed in the previous chapter and will result in the limiting double exponential distribution for the maximum, with the appropriate scale and location normalization similar to that in Chapter 4.

There are many important normal processes which are not differentiable (such as the Ornstein–Uhlenbeck process) and in Chapter 12 we shall develop a general theory for extremes of normal processes, including differentiable and many nondifferentiable processes as special cases.

However, we think it is illuminating to treat the regular case separately, since it allows for much simpler proofs, due to the possibility of a comparison, via Slepian's Lemma, with the cosine process of Section 7.4.

8.1. Conditions

We shall assume throughout the chapter that $\{\xi(t); t \geq 0\}$ is a stationary, normal process with $E(\xi(t)) = 0$, $E(\xi^2(t)) = 1$, $E(\xi(t)\xi(t + \tau)) = r(\tau)$ where the spectral moment $\lambda_2 = -r''(0)$ exists, finite. Equivalently, this requires that the mean number of upcrossings of any level per time unit is finite (Theorem 7.3.2), and also equivalently that the covariance function has the following representation,

$$r(\tau) = 1 - \frac{\lambda_2 \tau^2}{2} + o(\tau^2) \quad \text{as } \tau \to 0. \tag{8.1.1}$$

A more general class of processes with covariance function of the type

$$r(\tau) = 1 - C|\tau|^{\alpha} + o(|\tau|^{\alpha}) \quad \text{as } \tau \to 0,$$

where $0 < \alpha \leq 2$ is considered in Chapter 12. This includes the regular processes as the special case $\alpha = 2$, while $\alpha < 2$ implies $\lambda_2 = \infty$ and thus that the process is nondifferentiable and has an infinite mean number of upcrossings.

As for normal sequences, the double exponential limit

$$P\{a_T(M(T) - b_T) \leq x\} \to \exp(-e^{-x}) \quad \text{as } T \to \infty$$

(for $M(T) = \sup\{\xi(t); 0 \leq t \leq T\}$ as in Chapter 7) will be derived under the weak condition

$$r(t)\log t \to 0 \quad \text{as } t \to \infty. \tag{8.1.2}$$

This is the continuous time analogue of (4.1.1), and it will be used to derive a version of Lemma 4.3.2 before starting the main development. Still weaker conditions corresponding to (4.5.4) will be studied in Chapter 12. In particular, these conditions will include the case $\int_0^\infty r^2(t)\,dt < \infty$, sometimes used in the literature.

In the following lemma we shall consider a level u which increases with the time period T in such way that $E(N_u(T))$ remains constant, i.e. $T\mu$ remains constant where $\mu = E(N_u(1)) = (1/2\pi)\,\lambda_2^{1/2}\exp(-u^2/2)$.

To obtain the asymptotic distribution of $M(T)$, the maximum of the continuous process, it will again be convenient to approximate by the sequence $\{\xi(kq); k = 1, 2, \ldots\}$ obtained by sampling the process at the points $\{kq; k = 1, 2, \ldots\}$, where we shall let $q \to 0$ as $u \to \infty$ (or equivalently $T \to \infty$). The rate of decrease of q (as specified below in Lemma 8.1.1) will be a compromise between two requirements. On the one hand, the sampled process shall approximate the continuous process sufficiently well, and on the other hand, the sampling points should be sufficiently far apart to avoid too high a dependence between consecutive values $\xi(kq)$, $\xi((k+1)q)$.

The statement that "a property holds provided $\psi = \psi(u) \downarrow 0$ sufficiently slowly" is to be taken to have the meaning that there exists some $\psi_0(u) \downarrow 0$ for which the property holds, and it holds for any $\psi(u)$ such that $\psi(u) \to 0$ but $\psi_0(u) \leq \psi(u)$ as $u \to \infty$. The following is the promised continuous analogue of Lemma 4.3.2.

Lemma 8.1.1. *Let $\varepsilon > 0$ be given.*

(i) *If $r(t) \to 0$ as $t \to \infty$, then* $\sup\{|r(t)|; |t| \geq \varepsilon\} = \delta < 1$.

(ii) *Suppose that* (8.1.1) *and* (8.1.2) *both hold. Let $T \sim \tau/\mu$, where τ is fixed and* $\mu = E(N_u(1)) = (1/2\pi)\,\lambda_2^{1/2}\exp(-u^2/2)$, *so that $u \sim (2\log T)^{1/2}$ as $T \to \infty$ (as is easily checked). If $qu = q(u)u \downarrow 0$ sufficiently slowly as $u \to \infty$ then*

$$\frac{T}{q}\sum_{\varepsilon \leq kq \leq T} |r(kq)|\exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0 \quad \textit{as } T \to \infty.$$

PROOF. (i) As in the discrete case (cf. remarks preceding Lemma 4.3.2) if $r(t) = 1$ for any $t > 0$, then $r(t) = 1$ for arbitrarily large values of t which contradicts $r(t) \to 0$. Hence $|r(t)| < 1$ for $|t| \geq \varepsilon$, and since $r(t)$ is continuous and tends to zero as $t \to \infty$, we must have $|r(t)|$ bounded away from 1 in $|t| \geq \varepsilon$, and (i) follows.

(ii) As in the discrete case, choose a constant β such that $0 < \beta < (1 - \delta)/(1 + \delta)$. Letting K be a generic constant,

$$\begin{aligned}\frac{T}{q} \sum_{\varepsilon \leq kq \leq T^\beta} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} &\leq \frac{T^{\beta+1}}{q^2} \exp\left\{-\frac{u^2}{1+\delta}\right\} \\ &= K \frac{T^{\beta+1}}{q^2} \mu^{2/(1+\delta)} \\ &\leq \frac{K}{q^2} T^{\beta+1-2/(1+\delta)} \\ &\leq \frac{K}{q^2 u^2} (\log T) T^{\beta+1-2/(1+\delta)}\end{aligned}$$

since $u^2 \sim 2 \log T$, as noted. If γ is chosen so that $0 < \gamma < (1 - \delta)/(1 + \delta) - \beta$, the last expression is dominated by $K(qu)^{-2} T^{-\gamma}$ which tends to zero provided $uq \to 0$ more slowly than $T^{-\gamma/2}$ $(\leq K \exp(-\gamma u^2/4))$. Hence this sum tends to zero.

By writing

$$\exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} = \exp(-u^2) \exp\left\{\frac{u^2 |r(kq)|}{1 + |r(kq)|}\right\}$$

we see that the remaining sum does not exceed

$$\frac{T}{q} \exp(-u^2) \sum_{T^\beta < kq \leq T} |r(kq)| \exp\{u^2 |r(kq)|\}.$$

Again as in the discrete case, if $\delta(t) = \sup_{s \geq t} |r(s) \log s|$ then $\delta(t) \to 0$ as $t \to \infty$ and for $s \geq t > 1$ we have $|r(s)| \leq \delta(t)/\log s \leq \delta(t)/\log t$. Thus for $kq \geq T^\beta$, $u^2 |r(kq)| \leq K \log T \delta(T^\beta)/\log T^\beta = (K/\beta)\, \delta(T^\beta)$ which tends to zero, uniformly in k. Hence the exponential term $\exp\{u^2 |r(kq)|\}$ is certainly bounded in (k, u). It is thus sufficient to show that

$$\frac{T}{q} \exp(-u^2) \sum_{T^\beta < kq \leq T} |r(kq)| \to 0 \quad \text{as } T \to \infty.$$

But this does not exceed

$$K \frac{T}{q} T^{-2} \frac{T}{q} \frac{\delta(T^\beta)}{\log T^\beta} \leq \frac{K\, \delta(T^\beta)}{q^2 u^2}$$

which again tends to zero provided $qu \to 0$ sufficiently slowly (i.e. slower than $\delta(T^\beta)^{1/2}$). □

8.2. Double Exponential Distribution of the Maximum

Having proved the technical Lemma 8.1.1, we now proceed to the main derivation of the extremal results under the assumption that $r''(0)$ exists (i.e. $\lambda_2 < \infty$) and that (8.1.2) holds. The condition $\lambda_2 < \infty$ guarantees that the point process of upcrossings of a level u will have a finite intensity. The case $\lambda_2 = \infty$ is also of interest, and, as noted, will be treated in Chapter 12, but requires the use of more complex methods.

Our basic technique here is to divide the interval $(0, T)$ (where T becomes large) into n pieces of fixed length h ($n = [T/h]$). Then $M(T)$ will clearly be close to $M(nh)$ which is the maximum of n r.v.'s $\zeta_j = M((j-1)h, jh)$, $j = 1, 2, \ldots, n$, (the $\{\zeta_j\}$ forming a stationary sequence). Thus we might expect that the methods used for sequences would apply here and this is the case (although we shall organize our arguments slightly differently to better suit the present purposes).

It is therefore not surprising that the tail of the distribution of the ζ_i, i.e. $P\{M(h) > u\}$ (for fixed h) plays a central role. In fact the same asymptotic form (7.4.7) holds for this tail probability here, as did for the special process $\xi^*(t) = \eta \cos \omega t + \zeta \sin \omega t$. In this present chapter it will be sufficient to obtain the following somewhat weaker result. In this we shall use Slepian's Lemma (Theorem 7.4.2) to compare maxima of $\xi(t)$ and $\xi^*(t)$ along the lines of a procedure originally used by S. M. Berman (1971a).

Lemma 8.2.1. *Suppose that the (standardized) stationary normal process* $\{\xi(t)\}$ *satisfies* (8.1.1). *Then, with the above notation,*

(i) *for all* $h > 0$, $P\{M(h) > u\} \le 1 - \Phi(u) + \mu h$ *so that*

$$\limsup_{u \to \infty} P\{M(h) > u\}/(\mu h) \le 1,$$

(ii) *given* $\theta < 1$ *there exists* $h_0 = h_0(\theta)$ *such that for* $0 \le h \le h_0$

$$P\{M(h) > u\} \ge 1 - \Phi(u) + \theta\mu h \tag{8.2.1}$$

so that $\liminf_{u\to\infty} P\{M(h) > u\}/(\mu h) \ge \theta$ *for* $0 \le h \le h_0 = h_0(\theta)$.

Proof. (i) follows since

$$\begin{aligned} P\{M(h) > u\} &\le P\{\xi(0) > u\} + P\{N_u(h) \ge 1\} \\ &\le 1 - \Phi(u) + E(N_u(h)). \end{aligned}$$

The second result (ii) follows simply from Slepian's Lemma (Theorem 7.4.2) by comparison with the simple process $\xi^*(t)$ given by (7.4.1). For if $\omega = \theta\lambda_2^{1/2}$ we have, by (8.1.1), $r(t) \le \cos \omega t$ for $0 \le t \le h_0 < \pi/\omega$, ($h_0 = h_0(\theta) > 0$). But this shows that the covariance function of $\xi(t)$ is dominated by that of $\xi^*(t)$ in $[0, h_0]$ and hence $P\{M(h) > u\} \ge P\{M^*(h) > u\}$ for $h \le h_0$, (with M^* as in (7.4.4)), which then gives (ii). □

Our remaining task is to approximate the maximum $M(T)$ (for increasing T) by the maxima over suitable, separated, fixed length subintervals, and show asymptotic independence of the maxima over these intervals. First we give a simple but useful lemma. In this, for $q > 0$, N_u and $N_u^{(q)}$ will denote the number of upcrossings of u in a fixed interval I of length h, by the process $\{\xi(t)\}$, and the sequence $\{\xi(kq)\}$, respectively. More precisely, $N_u^{(q)}$ is the number of $kq \in I$ such that $(k-1)q \in I$ and $\xi((k-1)q) < u < \xi(kq)$ (cf. Lemma 7.2.2 with q for q_n and $N_n = N_u^{(q_n)}$).

Lemma 8.2.2. *If* (8.1.1) *holds, with the above notation and $qu \to 0$ as $u \to \infty$ then, as $u \to \infty$*

(i) $E(N_u^{(q)}) = h\mu + o(\mu)$,

(ii) $P\{M(I) \le u\} = P\{\xi(kq) \le u, kq \in I\} + o(\mu)$,

where each $o(\mu)$-term is uniform in all such intervals I of length $h \le h_0$ for any fixed $h_0 > 0$.

Proof. The number of points $kq \in I$ with $(k-1)q \in I$ is clearly $(h/q) - \beta$ where $0 \le \beta \le 2$. Hence with $J_q(u)$ defined by (7.2.1), Lemma 7.3.1 implies that

$$\begin{aligned}E(N_u^{(q)}) &= \left(\frac{h}{q} + \beta\right) P\{\xi(0) < u < \xi(q)\}\\ &= (h + \beta q)J_q(u)\\ &= \mu h(1 + o(1)) + O(\mu q),\end{aligned}$$

where the o- and O-terms are uniform in h so that (i) clearly holds with $o(\mu)$ uniform in $0 < h \le h_0$.

To prove (ii), note that if a is the left-hand endpoint of I,

$$\begin{aligned}0 &\le P\{\xi(kq) \le u, kq \in I\} - P\{M(h) \le u\}\\ &\le P\{\xi(a) > u\} + P\{\xi(a) < u, N_u \ge 1, N_u^{(q)} = 0\}\\ &\le 1 - \Phi(u) + P\{N_u - N_u^{(q)} \ge 1\}.\end{aligned}$$

The first term is $o(\phi(u)) = o(\mu)$, independent of h. Since $N_u - N_u^{(q)}$ is a non-negative integer-valued random variable (cf. Lemma 7.2.2(i)), the second term does not exceed $E(N_u - N_u^{(q)})$ which by (i) is $o(\mu)$, uniformly in $(0, h_0]$. Hence (ii) follows. □

Now let $u, T \to \infty$ in such a way that $T\mu \to \tau > 0$. Fix $h > 0$ and write $n = [T/h]$. Divide the interval $[0, nh]$ into n pieces each of length h. Fix ε, $0 < \varepsilon < h$ and divide each piece into two—of length $h - \varepsilon$ and ε, respectively. By doing so we obtain n pairs of intervals $I_1, I_1^*, \ldots, I_n, I_n^*$, alternately of length $h - \varepsilon$ and ε, making up the whole interval $[0, T]$ apart from one further piece which is contained in the next pair, I_{n+1}, I_{n+1}^*.

Lemma 8.2.3. *As $T \to \infty$ let $q \to 0$, $u \to \infty$ in such a way that $qu \to 0$ and $T\mu \to \tau > 0$. Then*

(i)
$$\limsup_{T\to\infty} \left| P\left\{M\left(\bigcup_1^n I_j\right) \le u\right\} - P\{M(nh) \le u\}\right| \le \frac{\tau}{h}\varepsilon,$$

(ii)
$$P\left\{\xi(kq) \le u,\, kq \in \bigcup_1^n I_j\right\} - P\left\{M\left(\bigcup_1^n I_j\right) \le u\right\} \to 0.$$

Proof. For (i) note that

$$\begin{aligned} 0 &\le P\left\{M\left(\bigcup_1^n I_j\right) \le u\right\} - P\{M(nh) \le u\} \\ &\le nP\{M(I_1^*) > u\} \\ &\sim \frac{\tau\varepsilon}{h}\frac{P\{M(I_1^*) > u\}}{\mu\varepsilon} \end{aligned}$$

since $n = [T/h] \sim \tau/(\mu h)$. Since I_1^* has length ε, (i) follows from Lemma 8.2.1(i).

To prove (ii) we note that the expression on the left is non-negative and dominated by

$$\sum_{j=1}^n \left(P\{\xi(kq) \le u,\, kq \in I_j\} - P\{M(I_j) \le u\}\right)$$

which by Lemma 8.2.2(ii) does not exceed $no(\mu) = [T/h]o(\mu) = o(1)$, (the $o(\mu)$-term being uniform in the I_j's), as required. □

The next lemma, implying the asymptotic independence of maxima, is formulated in terms of the condition (8.2.2), also appearing in Lemma 8.1.1.

Lemma 8.2.4. *Suppose $r(t) \to 0$ as $t \to \infty$ and that, as $T \to \infty$, and $u \to \infty$,*

$$\frac{T}{q}\sum_{\varepsilon \le kq \le T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0 \tag{8.2.2}$$

for each $\varepsilon > 0$ and some q such that $qu \to 0$. Then if $T\mu \to \tau$,

(i)
$$P\left\{\xi(kq) \le u,\, kq \in \bigcup_1^n I_j\right\} - \prod_{j=1}^n P\{\xi(kq) \le u,\, kq \in I_j\} \to 0,$$

(ii)
$$\limsup_{T\to\infty} \left|\prod_{j=1}^n P\{\xi(kq) \le u,\, kq \in I_j\} - P^n\{M(h) \le u\}\right| \le \frac{2\tau}{h}\varepsilon$$

for each ε, $0 < \varepsilon < h$.

Proof. To show (i) we use Corollary 4.2.2 and compare the maximum of $\xi(kq)$, $kq \in \bigcup_1^n I_j$ under the full covariance structure, with the maximum of $\xi(kq)$, assuming variables arising from different I_j-intervals are independent.

To formalize this, let $\Lambda^1 = (\lambda_{ij}^1)$ be the covariance matrix of $\xi(kq)$, $kq \in \bigcup I_j$ and let $\Lambda^0 = (\lambda_{ij}^0)$ be the modification obtained by writing zeros in the off-diagonal blocks (which would occur if the groups were independent of each other); e.g. with $n = 3$,

$$\Lambda^1 = \left[\begin{array}{c:c:c} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \hdashline \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \hdashline \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{array}\right], \qquad \Lambda^0 = \left[\begin{array}{c:c:c} \Lambda_{11} & 0 & 0 \\ \hdashline 0 & \Lambda_{22} & 0 \\ \hdashline 0 & 0 & \Lambda_{33} \end{array}\right].$$

From (4.2.3) we obtain

$$\left| P\left\{\xi(kq) \le u, kq \in \bigcup_1^n I_j\right\} - \prod_{j=1}^n P\{\xi(kq) \le u, kq \in I_j\} \right|$$

$$\le \frac{1}{2\pi} \sum_{1 \le i < j \le L} |\lambda_{ij}^1 - \lambda_{ij}^0| (1 - \rho_{ij}^2)^{-1/2} \exp\left(-\frac{u^2}{1 + \rho_{ij}}\right), \quad (8.2.3)$$

where L is the total number of kq-points in $\bigcup_1^n I_j$, and $\rho_{ij} = |\lambda_{ij}^1|$. Since all terms with i, j in the same diagonal block vanish, while otherwise $\sup \rho_{ij} = \delta < 1$ by Lemma 8.1.1(i), we see that the double sum does not exceed

$$K \sum{}^* \rho_{ij} \exp\left(-\frac{u^2}{1 + \rho_{ij}}\right),$$

where $\sum^*$ indicates that the summation is carried out over $i < j$ with (i, j) in the off-diagonal blocks only. But ρ_{ij} is of the form $|r(kq)|$ where there are not more than T/q terms with the same k-value. Thus, since the minimum value of kq is at least ε, we obtain the bound

$$\left| P\left\{\xi(kq) \le u, kq \in \bigcup_1^n I_j\right\} - \prod_{j=1}^n P\{\xi(kq) \le u, kq \in I_j\} \right|$$

$$\le K \frac{T}{q} \sum_{\varepsilon \le kq \le T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\}$$

which tends to zero by assumption (8.2.2) so that (i) follows.

To prove (ii), note that by Lemma 8.2.2(ii),

$$0 \le P\{\xi(kq) \le u, kq \in I_j\} - P\{M(I_j) \le u\} = o(\mu)$$

(uniformly in j) and

$$0 \le P\{M(I_j) \le u\} - P\{M(h) \le u\} \le P\{M(I_j^*) > u\}$$

so that by Lemma 8.2.1(i), for sufficiently large n (uniformly in j)

$$0 \le P_j - P \le 2\mu\varepsilon,$$

where $P_j = P\{\xi(kq) \le u, kq \in I_j\}$, $P = P\{M(h) \le u\}$. Hence

$$0 \le \prod_{j=1}^{n} P_j - P^n \le (\max P_j)^n - P^n \le 2n\mu\varepsilon$$

(using the fact that $y^n - x^n \le n(y - x)$ for $0 < x < y < 1$). Part (ii) now follows since $n\mu \sim T\mu/h \to \tau/h$. □

The basic extremal theorem now follows readily.

Theorem 8.2.5. *Let $u, T \to \infty$ in such a way that $T\mu = (T/2\pi)\lambda_2^{1/2}\exp(-u^2/2) \to \tau \ge 0$. Suppose that $r(t)$ satisfies* (8.1.1) *and either* (8.1.2) *or the weaker condition* (8.2.2) *for some q such that $qu \to 0$ as $T \to \infty$ (cf. Lemma* 8.1.1). *Then*

$$P\{M(T) \le u\} \to e^{-\tau} \quad \text{as } T \to \infty. \tag{8.2.4}$$

PROOF. If $T\mu(u) \to 0$ then $P\{M(T) > u\} \le 1 - \Phi(u) + T\mu(u) \to 0$ and it only remains to prove the result for $\tau > 0$. By Lemma 8.1.1 the assumption (8.2.2) of Lemma 8.2.4 holds. From Lemmas 8.2.3 and 8.2.4 we obtain

$$\limsup_{T\to\infty} |P\{M(nh) \le u\} - P^n\{M(h) \le u\}| \le \frac{3\tau}{h}\varepsilon$$

and since $\varepsilon > 0$ is arbitrary it follows that

$$P\{M(nh) \le u\} - P^n\{M(h) \le u\} \to 0.$$

Further, since $nh \le T < (n + 1)h$, it follows along now familiar lines that

$$0 \le P\{M(nh) \le u\} - P\{M(T) \le u\} \le P\{N_u(h) \ge 1\} \le \mu h$$

which tends to zero, so that

$$P\{M(T) \le u\} = P^n\{M(h) \le u\} + o(1).$$

This holds for any fixed $h > 0$. Suppose now that θ is fixed, $0 < \theta < 1$, and h chosen with $0 < h < h_0$ where $h_0 = h_0(\theta)$ is as in Lemma 8.2.1(ii), from whence it follows that

$$P\{M(h) > u\} \ge \theta\mu h(1 + o(1)) = \frac{\theta\tau}{n}(1 + o(1))$$

and hence

$$\begin{aligned} P\{M(T) \le u\} &= (1 - P\{M(h) > u\})^n + o(1) \\ &\le \left(1 - \frac{\theta\tau}{n} + o\left(\frac{1}{n}\right)\right)^n + o(1) \end{aligned}$$

so that

$$\limsup_{T\to\infty} P\{M(T) \le u\} \le e^{-\theta\tau}.$$

By letting $\theta \uparrow 1$ we see that lim sup $P\{M(T) \le u\} \le e^{-\tau}$. That the opposite inequality holds for the lim inf is seen in a similar way, but even more simply, from Lemma 8.2.1(i) (no θ being involved) so that the entire result follows. □

Corollary 8.2.6. *Suppose $r(t)$ satisfies* (8.1.1) *and* (8.1.2), *and let $E = E_T$ be any interval of length γT for a constant $\gamma > 0$. Then $P\{M(E) \le u\} \to e^{-\gamma\tau}$ as $T \to \infty$.*

Proof. By stationarity we may take E to be an interval with left endpoint at zero, so that $P\{M(E) \le u\} = P\{M(\gamma T) \le u\}$. It is simply checked that the process $\eta(t) = \xi(\gamma t)$ satisfies the conditions of the theorem, and has mean number of upcrossings per unit time given by $\mu_\eta = \gamma\mu$, so that $\mu_\eta T \to \gamma\tau$. Writing M_η for the maximum of η the result follows at once since $P\{M(\gamma T) \le u\} = P\{M_\eta(T) \le u\} \to e^{-\gamma\tau}$. □

It is now a simple matter to obtain the double exponential limiting law for $M(T)$ under a linear normalization. This is similar to the result of Theorem 4.3.3 for normal sequences.

Theorem 8.2.7. *Suppose that the* (*standardized*) *stationary normal process* $\{\xi(t)\}$ *satisfies* (8.1.1) *and* (8.1.2) (*or* (8.2.2)). *Then*

$$P\{a_T(M(T) - b_T) \le x\} \to \exp(-e^{-x}) \quad \text{as } T \to \infty, \tag{8.2.5}$$

where

$$\begin{aligned} a_T &= (2 \log T)^{1/2}, \\ b_T &= (2 \log T)^{1/2} + \left(\log \frac{\lambda_2^{1/2}}{2\pi}\right) \Big/ (2 \log T)^{1/2}. \end{aligned} \tag{8.2.6}$$

Proof. Write $\tau = e^{-x}$ and define

$$u^2 = 2\left(\log T + x + \log \frac{\lambda_2^{1/2}}{2\pi}\right) \tag{8.2.7}$$

so that

$$T\mu = T\frac{\lambda_2^{1/2}}{2\pi}\exp\left(-\frac{u^2}{2}\right) = e^{-x} = \tau.$$

Hence (8.2.4) holds. But it follows from (8.2.7) that

$$\begin{aligned} u &= (2 \log T)^{1/2}\left[1 + \frac{x + \log(\lambda_2^{1/2}/2\pi)}{2 \log T} + o\left(\frac{1}{\log T}\right)\right] \\ &= \frac{x}{a_T} + b_T + o(a_T^{-1}) \end{aligned}$$

so that (8.2.4) gives $P\{a_T(M(T) - b_T) + o(1) \le x\} \to e^{-\tau}$ from which (8.2.5) follows at once. □

It is of interest to note in passing that this calculation is somewhat simpler—due to the absence of a log u-term in (8.2.7), than the corresponding calculation in the discrete case (cf. Theorem 1.5.3).

In the discrete case we obtained Poisson limiting behaviour for the exceedances of a high level. Corresponding results hold for the point processes of high-level upcrossings under the conditions of this chapter. These are readily obtained from the present extremal theory by means of our familiar point process convergence theorem, as in the discrete case, resulting in a number of interesting consequences concerning local maxima, height of excursions, etc. We will defer such a discussion to Chapters 9 and 10. However, it is worth noting here that historically the asymptotic Poisson distribution of the number of high-level upcrossings was proved first (under more restrictive conditions) by Volkonski and Rozanov (1961). Cramér (1965) noted the connection with the maximum given, e.g. by

$$\{N_u(T) = 0\} = \{M(T) \leq u\} \cup \{N_u(T) = 0, \xi(0) > u\},$$

which led to the determination of the asymptotic distribution of $M(T)$, and subsequent extremal development.

CHAPTER 9
Point Processes of Upcrossings and Local Maxima

In the limit theory for the maximum of a stationary normal process $\xi(t)$, as developed in Chapter 8, substantial use was made of upcrossings, and of the obvious fact that the maximum exceeds u if there is at least one upcrossing of the level u. However, the upcrossings have an interest in their own right, and as we shall see here, they also contain considerable information about the local structure of the process. This chapter is devoted to the asymptotic Poisson character of the point process of upcrossings of increasingly high levels, and of the point process formed by the local maxima of the process.

Indeed, only a little more effort is needed to prove Poisson convergence of the upcrossings, once the limiting theory for the maximum is available. The main step in the proofs is that maxima over disjoint intervals are asymptotically independent, from which convergence of the point processes of upcrossings of several levels follows by means of the basic point process convergence theorem. This then easily yields complete Poisson convergence of the point process formed by the local maxima, and as a consequence, the joint asymptotic distribution of heights and locations of the highest local maxima.

In our derivation of the results we shall make substantial use of the regularity condition $\lambda_2 < \infty$, which ensures that the upcrossings do not appear in clusters, and remain separated as the level increases. Similar results will be established in Chapter 12 for the case $\lambda_2 = \infty$, with regular upcrossings replaced by so-called ε-upcrossings. For the results about local maxima we shall require in addition that $\lambda_4 < \infty$, which in a similar way can be weakened by replacing maxima by "ε-maxima".

9.1. Poisson Convergence of Upcrossings

Corresponding to each level u we have defined $\mu = \mu(u) = (1/2\pi)\lambda_2^{1/2} \exp(-u^2/2)$ to be the mean number of u-upcrossings per time unit by the stationary normal process $\xi(t)$, and, as in Chapter 8, we consider $T = T(u)$ such that $T\mu \to \tau$ as $u \to \infty$, where $\tau > 0$ is a fixed number. Let N_T^* be the time-normalized point process of u-upcrossings, defined by

$$N_T^*(B) = N_u(TB) = \#\{u\text{-upcrossings by } \xi(t);\, t/T \in B\},$$

for any real Borel set B, i.e. N_T^* has a point at t if ξ has a u-upcrossing at tT. Note that we define N_T^* as a point process on the entire real line, and that the only significance of the time T is that of an appropriate scaling factor. This is a slight shift in emphasis from Chapter 8, where we considered $u_T = x/a_T + b_T$ as a height normalization for the maximum over the increasing time interval $(0, T]$.

Let N be a Poisson process on the real line with intensity τ. To prove point process convergence under suitable conditions, we need to prove different forms of asymptotic independence of maxima over disjoint intervals. For the one-level result, that N_T^* converges in distribution to N, we need only the following partial independence, (given in Qualls (1968)).

Lemma 9.1.1. *Let* $0 < c = c_1 < d_1 \le c_2 < \cdots \le c_r < d_r = d$ *be fixed numbers* $E_i = (Tc_i, Td_i]$, *and* $M(E_i) = \sup\{\xi(t); Tc_i < t \le Td_i\}$. *Then, if* $r(t)$ *satisfies* (8.1.1) *and* (8.1.2),

$$P\left(\bigcap_{i=1}^r \{M(E_i) \le u\}\right) - \prod_{i=1}^r P\{M(E_i) \le u\} \to 0$$

as $u \to \infty$, $T\mu \to \tau \ge 0$.

Proof. The proof is similar to that of Lemmas 8.2.3 and 8.2.4. Recall the construction in Lemma 8.2.3, and divide the positive real line into intervals I_1, $I_1^*, I_2, \ldots$ of lengths $h - \varepsilon$ and ε, alternately. We can then approximate $M(E_i)$ by the maximum on the parts of the separated intervals I_k which are contained in E_i. Write n for the number of I_k^*'s which have nonempty intersection with $\bigcup_{i=1}^r E_i$. We at once obtain

$$\begin{aligned} 0 &\le P\left(\bigcap_{i=1}^r \left\{M\left(\bigcup_k I_k \cap E_i\right) \le u\right\}\right) - P\left(\bigcap_{i=1}^r \{M(E_i) \le u\}\right) \\ &\le nP\{M(I_1^*) > u\}, \end{aligned}$$

where (writing $|E|$ for the length of an interval E),

$$n \sim h^{-1} \sum_{i=1}^r |E_i| = \frac{T}{h} \sum_{i=1}^r (d_i - c_i) \le \frac{T(d-c)}{h}.$$

Since Lemma 8.2.1(i) implies that

$$\limsup_{u\to\infty} \mu^{-1}P\{M(I_1^*) > u\} \le \varepsilon,$$

we therefore have

$$\limsup_{u\to\infty}\left|P\left(\bigcap_{i=1}^{r}\left\{M\left(\bigcup_k I_k\cap E_i\right)\le u\right\}\right) - P\left(\bigcap_{i=1}^{r}\{M(E_i)\le u\}\right)\right| \le \frac{\tau\varepsilon(d-c)}{h}. \tag{9.1.1}$$

Now, let $q\to 0$ as $u\to\infty$ so that $qu\to 0$. The discrete approximation of maxima in terms of $\xi(jq)$, $jq\in\bigcup_k I_k\cap E_i$ is then obtained as in Lemma 8.2.3(ii). In fact, since there are $n+\delta$ intervals I_k which intersect $\bigcup E_i$ (where $|\delta|\le r$), we have

$$\begin{aligned}0 &\le P\left(\bigcap_{i=1}^{r}\left\{\xi(jq)\le u, jq\in\bigcup_k I_k\cap E_i\right\}\right) - P\left(\bigcap_{i=1}^{r}\left\{M\left(\bigcup_k I_k\cap E_i\right)\le u\right\}\right)\\ &\le \sum_i\sum_k (P\{\xi(jq)\le u, jq\in I_k\cap E_i\} - P\{M(I_k\cap E_i)\le u\})\\ &= (n+\delta)o(\mu) = o(1) \quad \text{as } u\to\infty,\end{aligned} \tag{9.1.2}$$

by Lemma 8.2.2(ii).

Furthermore,

$$P\left(\bigcap_{i=1}^{r}\left\{\xi(jq)\le u, jq\in\bigcup_k I_k\cap E_i\right\}\right) - \prod_{i=1}^{r} y_i \to 0, \tag{9.1.3}$$

where $y_i = \prod_k P\{\xi(jq)\le u,\ jq\in I_k\cap E_i\}$, the proof this time being a rephrasing of the proof of Lemma 8.2.4(i), ((8.1.2) implies (8.2.2) with Td replacing T).

By combining (9.1.1), (9.1.2), and (9.1.3) we obtain

$$\limsup_{u\to\infty}\left|P\left(\bigcap_{i=1}^{r}\{M(E_i)\le u\}\right) - \prod_{i=1}^{r} y_i\right| \le \frac{\tau\varepsilon(d-c)}{h}$$

and in particular, for $i = 1,\dots,r$,

$$\limsup_{u\to\infty} |P\{M(E_i)\le u\} - y_i| \le \frac{\tau\varepsilon(d-c)}{h}.$$

Hence, writing $x_i = P\{M(E_i)\le u\}$, we have

$$\limsup_{u\to\infty}\left|P\left(\bigcap_{i=1}^{r}\{M(E_i)\le u\}\right) - \prod_{i=1}^{r} P\{M(E_i)\le u\}\right| \le \frac{\tau\varepsilon(d-c)}{h} + \limsup_{u\to\infty}\left|\prod_{i=1}^{r} y_i - \prod_{i=1}^{r} x_i\right|.$$

But, with $z = \max_i |y_i - x_i|$ (so that $\limsup z \le \tau\varepsilon(d - c)/h$),

$$\prod_{i=1}^{r} y_i - \prod_{i=1}^{r} x_i \le \prod_{i=1}^{r} (x_i + z) - \prod_{i=1}^{r} x_i \le 2^r z,$$

with a similar relation holding with y_i and x_i interchanged, and hence

$$\limsup_{u\to\infty} \left| \prod_{i=1}^{r} y_i - \prod_{i=1}^{r} x_i \right| \le \frac{2^r \tau\varepsilon(d - c)}{h}.$$

Since ε is arbitrary, this proves the lemma. □

Theorem 9.1.2. *Let $u \to \infty$ and $T \sim \tau/\mu$, where $\mu = (1/2\pi)\, \lambda_2^{1/2} \exp(-u^2/2)$, and suppose the stationary normal process $\xi(t)$ satisfies* (8.1.1) *and* (8.1.2). *Then the time-normalized point process N_T^* of u-upcrossings converges in distribution to a Poisson process with intensity τ on the positive real line.*

Proof. By the basic convergence theorem for simple point processes, Theorem A.1, it is sufficient to show that, as $u \to \infty$,

(a) $E(N_T^*((c, d])) \to E(N((c, d])) = \tau(d - c)$ for all $0 < c < d$, and
(b) $P\{N_T^*(B) = 0\} \to P\{N(B) = 0\} = \exp\{-\tau \sum_{i=1}^{r} (d_i - c_i)\}$ for all sets B of the form $\bigcup_{i=1}^{r} (c_i, d_i]$, $0 < c_1 < d_1 < \cdots < c_r < d_r$.

Here, part (a) is trivially satisfied, since

$$E(N_T^*((c, d])) = E(N_u((Tc, Td])) = T\mu\,(d - c) \to \tau(d - c).$$

For part (b), we have for the u-upcrossings,

$$P\{N_T^*(B) = 0\} = P\left(\bigcap_{i=1}^{r} \{N_T^*((c_i, d_i]) = 0\}\right) = P\left(\bigcap_{i=1}^{r} \{N_u(E_i) = 0\}\right),$$

where $E_i = (Tc_i, Td_i]$. Now it is easy to see that we can work with maxima instead of crossings, since

$$\begin{aligned} 0 &\le P\left(\bigcap_{i=1}^{r} \{N_u(E_i) = 0\}\right) - P\left(\bigcap_{i=1}^{r} \{M(E_i) \le u\}\right) \\ &= P\left(\bigcap_{i=1}^{r} \{N_u(E_i) = 0\} \cap \bigcup_{i=1}^{r} \{M(E_i) > u\}\right) \\ &\le \sum_{i=1}^{r} P\{\xi(Tc_i) > u\} \to 0 \quad \text{as } u \to \infty, \end{aligned}$$

and since furthermore Lemma 9.1.1 and Corollary 8.2.6 imply that

$$\lim_{u\to\infty} P\left(\bigcap_{i=1}^{r} \{M(E_i) \le u\}\right) = \lim_{u\to\infty} \prod_{i=1}^{r} P\{M(E_i) \le u\} = \prod_{i=1}^{r} \exp\{-\tau(d_i - c_i)\},$$

we have proved part (b). □

One immediate consequence of the distributional convergence of N_T^*, is the asymptotic Poisson distribution of the number of u-upcrossings in increasing Borel sets $T \cdot B$.

Corollary 9.1.3. *Under the conditions of Theorem* 9.1.2 *if B is any positive Borel set whose boundary has Lebesgue measure zero, then*

$$P\{N_T^*(B) = r\} \to e^{-\tau m(B)}(\tau m(B))^r/r!, \qquad r = 0, 1, \ldots, \tag{9.1.4}$$

as $u \to \infty$, where $m(B)$ is the Lebesgue measure of B. The joint distribution of $N_T^*(B_1), \ldots, N_T^*(B_n)$ corresponding to disjoint B_j (with boundaries which have Lebesgue measure zero) converges to the product of the corresponding Poisson probabilities.

9.2. Full Independence of Maxima in Disjoint Intervals

In this section we shall prove that, without any extra assumptions, the maxima in disjoint intervals are actually asymptotically independent, and not only "independent on the diagonal", as in Lemma 9.1.1. To prove this we must allow for different levels in different intervals, with correspondingly different crossing intensities.

To this end, and also for use in the next section, we shall examine the relationship between the intensity τ and the height u of a level for which $T\mu \to \tau$. If $T = \tau/\mu = \tau 2\pi\lambda_2^{-1/2}\exp(u^2/2)$ we have

$$u^2 = 2 \log T \left(1 - \frac{\log \tau + \log(2\pi/\lambda_2^{1/2})}{\log T}\right)$$

or

$$u = (2 \log T)^{1/2} - \frac{\log \tau + \log(2\pi/\lambda_2^{1/2})}{(2 \log T)^{1/2}} + o((\log T)^{-1/2}). \tag{9.2.1}$$

However, any level which differs from u by $o((\log T)^{-1/2})$ will do equally well in Theorem 9.1.2, and it is often convenient as in Theorem 8.2.7 to use the level obtained by deleting the last term in (9.2.1) entirely. (The reader should check that for this choice the relation $T\mu \to \tau$ also holds.) If we write

$$u_\tau = (2 \log T)^{1/2} - \frac{\log \tau + \log(2\pi/\lambda_2^{1/2})}{(2 \log T)^{1/2}} \tag{9.2.2}$$

we have, for $\tau > \tau^* > 0$,

$$u_{\tau^*} - u_\tau = \frac{\log \tau/\tau^*}{(2 \log T)^{1/2}} \sim \frac{\log \tau/\tau^*}{u_\tau} > 0, \tag{9.2.3}$$

so that levels corresponding to different intensities τ, τ^* (under the same time-normalization T) become increasingly close to each other, the difference being of the order $1/u_\tau$. Note that (9.2.3) holds for any u_{τ^*}, u_τ which satisfy $T\mu(u_{\tau^*}) \to \tau^*$, $T\mu(u_\tau) \to \tau$, and not only for the particular choice (9.2.2).

We shall prove the full asymptotic independence of maxima in disjoint increasing intervals under the conditions (8.1.1) and (8.1.2), or condition (8.2.2) for some u satisfying $T\mu(u) \to \tau$ with $0 < \tau < \infty$, i.e.

$$\frac{T}{q} \sum_{\varepsilon \le kq \le T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0, \tag{9.2.4}$$

for each $\varepsilon > 0$, and some $q \to 0$ such that $uq \to 0$.

Note that if $\{u_*\}$ is another family of levels such that

$$0 < \liminf_{T\to\infty} T\mu(u_*) \le \limsup_{T\to\infty} T\mu(u_*) < \infty$$

then it follows simply that $u_*^2 - u^2$ is bounded. Hence (9.2.4) is satisfied also with u replaced by u_*. Further (9.2.4) holds with $T' = Td$ replacing T since then $T'\mu \to \tau' = \tau d$.

Theorem 9.2.1. *Suppose that $r(t)$ satisfies* (8.1.1) *and either* (8.1.2), (*or the weaker condition* (9.2.4), *for some family of levels $u \to \infty$ as $T \to \infty$ such that $T\mu(u) \to \tau > 0$*). *Let $0 \le c = c_1 < d_1 \le c_2 < \cdots \le c_s < d_s = d$ be fixed, and write $E_i = (Tc_i, Td_i]$. Then for any s levels $u_{T,1}, \ldots, u_{T,s}$,*

$$\left| P\left(\bigcap_{i=1}^{s} \{M(E_i) \le u_{T,i}\}\right) - \prod_{i=1}^{s} P\{M(E_i) \le u_{T,i}\}\right| \to 0, \tag{9.2.5}$$

as $T \to \infty$.

Proof. We shall first assume that there is a constant $\delta > 0$ such that

$$\delta \le T\mu(u_{T,i}) \le \frac{1}{\delta} \tag{9.2.6}$$

for all sufficiently large T and all i so that (9.2.4) holds with u replaced by any of the $u_{T,i}$ and T replaced by Td. However, under this hypothesis, the proof of (9.2.5) goes step by step as the proof of Lemma 9.1.1, with the appropriate changes for u and T in (9.2.4).

Next, to remove assumption (9.2.6) we shall introduce truncated levels $v_{T,1}, \ldots, v_{T,s}$ as follows. Let $\delta > 0$ be given and let u_T^δ and $u_T^{1/\delta}$ be the positive solutions of $T\mu(u) = \delta$ and $T\mu(u) = 1/\delta$, respectively (these solutions exist for sufficiently large T). Define

$$v_{T,i} = \begin{cases} u_T^\delta, & \text{if } u_{T,i} > u_T^\delta, \\ u_{T,i}, & \text{if } u_T^{1/\delta} < u_{T,i} \le u_T^\delta, \\ u_T^{1/\delta}, & \text{if } u_{T,i} \le u_T^{1/\delta}. \end{cases}$$

Then clearly

$$\begin{aligned}
&\left|P\left(\bigcap_{i=1}^{s}\{M(E_i)\le u_{T,i}\}\right)-P\left(\bigcap_{i=1}^{s}\{M(E_i)\le v_{T,i}\}\right)\right|\\
&\qquad\le\sum_{i=1}^{s}P\{M(E_i)\le u_T^{1/\delta}\}+\sum_{i=1}^{s}P\{M(E_i)>u_T^{\delta}\}\\
&\qquad\to\sum_{i=1}^{s}e^{-(d_i-c_i)/\delta}+\sum_{i=1}^{s}(1-e^{-(d_i-c_i)/\delta})
\end{aligned}\tag{9.2.7}$$

Obviously, the same bounds hold for

$$\left|\prod_{i=1}^{s}P\{M(E_i)\le u_{T,i}\}-\prod_{i=1}^{s}P\{M(E_i)\le v_{T,i}\}\right|,$$

(since the proof of (9.2.7) does not use dependence or independence of the $M(E_i)$'s).

Since the $v_{T,i}$'s satisfy (9.2.6), and since

$$\begin{aligned}
&\left|P\left(\bigcap_{i=1}^{s}\{M(E_i)\le u_{T,i}\}\right)-\prod_{i=1}^{s}P\{M(E_i)\le u_{T,i}\}\right|\\
&\qquad\le\left|P\left(\bigcap_{i=1}^{s}\{M(E_i)\le u_{T,i}\}\right)-P\left(\bigcap_{i=1}^{s}\{M(E_i)\le v_{T,i}\}\right)\right|\\
&\qquad\quad+\left|P\left(\bigcap_{i=1}^{s}\{M(E_i)\le v_{T,i}\}\right)-\prod_{i=1}^{s}P\{M(E_i)\le v_{T,i}\}\right|\\
&\qquad\quad+\left|\prod_{i=1}^{s}P\{M(E_i)\le v_{T,i}\}-\prod_{i=1}^{s}P\{M(E_i)\le u_{T,i}\}\right|
\end{aligned}$$

it follows that

$$\begin{aligned}
&\limsup_{T\to\infty}\left|P\left(\bigcap_{i=1}^{s}\{M(E_i)\le u_{T,i}\}\right)-\prod_{i=1}^{s}P\{M(E_i)\le u_{T,i}\}\right|\\
&\qquad\le\sum_{i=1}^{s}e^{-(d_i-c_i)/\delta}+\sum_{i=1}^{s}(1-e^{-(d_i-c_i)\delta}).
\end{aligned}$$

Letting $\delta\to 0$ now concludes the proof of (9.2.5) in the general case. □

From the theorem and Corollary 8.2.6 we immediately obtain the following result.

Theorem 9.2.2. *Let $u^{(1)},\dots,u^{(r)}$ be levels such that*

$$T\mu(u^{(i)})=\frac{T}{2\pi}\lambda_2^{1/2}\exp\left\{-\frac{(u^{(i)})^2}{2}\right\}\to\tau_i$$

as $T \to \infty$, and suppose $r(t)$ satisfies (8.1.1), *and either* (8.1.2) *or the weaker condition* (9.2.4) *with some family $\{u\}$ such that $T\mu(u) \to \tau \to 0$. Then, for any $0 \le c = c_1 < d_1 \le c_2 < \cdots \le c_s < d_s = d\ (E_i = (Tc_i, Td_i])$,*

$$P\left(\bigcap_{i=1}^{s} \{M(E_i) \le u_{T,i}\}\right) \to \exp\left\{-\sum_{i=1}^{s} \tau_i'(d_i - c_i)\right\}, \tag{9.2.8}$$

where each $u_{T,i}$ is one of $u^{(1)}, \ldots, u^{(r)}$ and $T\mu(u_{T,i}) \to \tau_i'$.

9.3. Upcrossings of Several Adjacent Levels

The Poisson convergence theorem, Theorem 9.1.2, implies that any of the time-normalized point processes of upcrossings of levels $u^{(1)} \ge \cdots \ge u^{(r)}$ are asymptotically Poisson if $T\mu(u^{(i)}) \to \tau_i > 0$ as $T, u^{(i)} \to \infty$. We shall now investigate the *dependence* between these point processes, following similar lines to those in Chapter 5. This dependence was first described by Qualls (1969); a point process formulation being given in Lindgren et al. (1975).

To describe their dependence we shall represent the upcrossings as points in the plane, rather than on the line, letting the upcrossings of the level $u^{(i)}$ define a point process on a fixed line L_i as was done in Chapter 5. However, for the normal process treated in this chapter, the lines $L_1, \ldots, L_r$ can be chosen to have a very simple relation to the process itself by utilizing the process

$$\xi_T(t) = a_T(\xi(tT) - b_T),$$

where time has been normalized by a factor T and height by

$$a_T = (2 \log T)^{1/2},$$
$$b_T = (2 \log T)^{1/2} + \log(\lambda_2^{1/2}/2\pi)/(2 \log T)^{1/2}$$

as usual.

Now, $\xi_T(t) = x$ if and only if $\xi(tT) = x/a_T + b_T$, and clearly the mean number of upcrossings of the level x by $\xi_T(t)$ in an interval of length h is equal to $(Th/2\pi)\lambda_2^{1/2} \exp\{-(x/a_T + b_T)^2/2\}$, which by comments following (9.2.2) equals $h\tau(1 + o(1))$ as $T \to \infty$, with $\tau = e^{-x}$. Therefore, let $x_1 \ge x_2 \ge \cdots \ge x_r$ be a set of fixed numbers, defining horizontal lines $L_1, L_2, \ldots, L_r$, (see Fig. 9.3.1(b)) and consider the point process in the plane formed by the upcrossings of any of these lines by the process $\xi_T(t)$. Here the dependence between points on different lines is not quite as simple as it was in Chapter 5 since, unlike an exceedance, an upcrossing of a high level is not an upcrossing of a lower level and there may in fact even be more upcrossings of the higher than of the lower level; see Figure 9.3.1, which shows the relation between the upcrossings of levels $u^{(i)} = x_i/a_T + b_T$ by $\xi(t)$, and of levels x_i by $\xi_T(t)$. As is seen, local irregularities in the process $\xi(t)$ can cause the appearance of extra upcrossings of a high level not present in the lower levels.

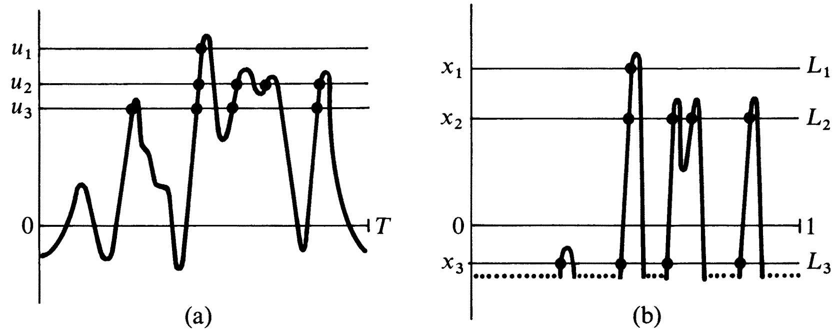


Figure 9.3.1. Point processes of upcrossings of several high levels, (a) $\xi(t)$, $0 \le t \le T$, (b) $\xi_T(t)$, $0 \le t \le 1$.

Let N_T^* denote the point process in the plane formed by the upcrossings of the fixed levels $x_1 \ge x_2 \ge \cdots \ge x_r$ by the process $\xi_T(t) = a_T(\xi(tT) - b_T)$, and let its components on the separate lines be $N_T^{(1)}, \ldots, N_T^{(r)}$, so that

$$N_T^*(B) = \# \{\text{upcrossings in } B \text{ of } L_1, \ldots, L_r \text{ by } \xi_T(t)\}$$
$$= \sum_{i=1}^{r} N_T^{(i)}(B \cap L_i),$$

for Borel sets $B \subseteq R^2$.

We shall now prove that N_T^* converges in distribution to a point process N in the plane, which is of a type already encountered in connection with exceedances in Chapter 5. The points of N are concentrated on the lines $L_1, \ldots, L_r$ and its distribution is determined by the joint distribution of its components $N^{(1)}, \ldots, N^{(r)}$ on the separate lines $L_1, \ldots, L_r$.

As in Chapter 5, let $\{\sigma_{1j}; j = 1, 2, \ldots\}$ be the points of a Poisson process $N^{(r)}$ with parameter $\tau_r = e^{-x_r}$ on L_r. Let β_j, $j = 1, 2, \ldots$ be i.i.d. random variables, independent of $N^{(r)}$, with distribution defined by

$$P\{\beta_j = s\} = \begin{cases} (\tau_{r-s+1} - \tau_{r-s})/\tau_r, & s = 1, \ldots, r-1, \\ \tau_1/\tau_r, & s = r, \end{cases}$$

so that $P\{\beta_j \ge s\} = \tau_{r-s+1}/\tau_r$ for $s = 1, 2, \ldots, r$.

Construct the processes $N^{(r-1)}, \ldots, N^{(1)}$ by placing points $\sigma_{2j}, \sigma_{3j}, \ldots, \sigma_{\beta_j j}$ on the $\beta_j - 1$ lines $L_{r-1}, \ldots, L_{r-\beta_j+1}$, vertically above σ_{1j}, $j = 1, 2, \ldots$, and finally define N to be the sum of the r processes $N^{(1)}, \ldots, N^{(r)}$.

As before, each $N^{(k)}$ is Poisson on L_k, since it is obtained from the Poisson process $N^{(r)}$ by independent deletion of points, with deletion probability $1 - P\{\beta_j \ge r - k + 1\} = 1 - \tau_k/\tau_r$, and it has intensity $\tau_r(\tau_k/\tau_r) = \tau_k$. Furthermore, $N^{(k)}$ is obtained from $N^{(k+1)}$ by a binomial thinning, with deletion probability $1 - \tau_k/\tau_{k+1}$. Of course, N itself is not Poisson in the plane.

When proving the main result, that N_T^* tends in distribution to N, we need to show that asymptotically, there are not more upcrossings of a higher level than of a lower level. With a convenient abuse of notation, write $N_T^{(i)}(I)$ for the number of points of $N_T^{(i)}$ with time-coordinate in I.

Lemma 9.3.1. *Suppose $x_i > x_j$, and consider the point processes $N_T^{(i)}$ and $N_T^{(j)}$ of upcrossings by $\xi_T(t)$ of the levels x_i and x_j, respectively. Under the conditions of Theorem* 9.1.2,

$$P\{N_T^{(i)}(I) > N_T^{(j)}(I)\} \to 0$$

as $T \to \infty$, for any bounded interval I.

PROOF. By stationarity it is sufficient to prove the lemma for $I \subset (0, 1]$. Let $I_k = ((k-1)/n, k/n]$, $k = 1, \ldots, n$, for fixed n and recall the notation (9.2.2), $u_{\tau_j} = x_j/a_T + b_T$, $\tau_j = e^{-x_j}$. Since by Theorem 7.3.2 all crossings are strict, the event $\{N_T^{(i)}(I) > N_T^{(j)}(I)\}$ implies that one of the events

$$\bigcup_{k=0}^{n} \left\{\xi\left(\frac{kT}{n}\right) > u_{\tau_j}\right\}$$

or

$$\bigcup_{k=1}^{n} \{N_T^{(i)}(I_k) \geq 2\}$$

occurs, so that Boole's inequality and stationarity give

$$\begin{aligned} P\{N_T^{(i)}(I) > N_T^{(j)}(I)\} &\leq \sum_{k=0}^{n} P\left\{\xi\left(\frac{kT}{n}\right) > u_{\tau_j}\right\} + \sum_{k=1}^{n} P\{N_T^{(i)}(I_k) \geq 2\} \\ &= (n+1)P\{\xi(0) > u_{\tau_j}\} + nP\{N_T^{(i)}(I_1) \geq 2\}. \end{aligned}$$

Obviously, $(n+1)P\{\xi(0) > u_{\tau_j}\} \to 0$, while by Corollary 9.1.3

$$P\{N_T^{(i)}(I_1) \geq 2\} \to 1 - \exp\left(-\frac{\tau_i}{n}\right) - \frac{\tau_i}{n}\exp\left(-\frac{\tau_i}{n}\right),$$

which implies

$$\limsup_{T\to\infty} P\{N_T^{(i)}(I) > N_T^{(j)}(I)\} \leq n\left(1 - \exp\left(-\frac{\tau_i}{n}\right) - \frac{\tau_i}{n}\exp\left(-\frac{\tau_i}{n}\right)\right).$$

Since n is arbitrary and

$$n\left(1 - \exp\left(-\frac{\tau}{n}\right) - \frac{\tau}{n}\exp\left(-\frac{\tau}{n}\right)\right) \to 0$$

as $n \to \infty$, this proves the lemma. □

Theorem 9.3.2. *Suppose that $r(t)$ satisfies* (8.1.1) *and* (8.1.2) (*or, more generally* (9.2.4) *for some $\{u\}$ such that $T\mu(u) \to \tau > 0$*), *let $\tau_1 < \tau_2 < \cdots < \tau_r$ be real positive numbers, and let N_T^* be the point process of upcrossings of the levels $x_1 > x_2 > \cdots > x_r$ ($\tau_i = e^{-x_i}$) by the normalized process $\xi_T(t) = a_T(\xi(tT) - b_T)$ represented on the lines $L_1, \ldots, L_r$. Then as $T \to \infty$, N_T^* tends in* *distribution to the point process N on $(0, \infty) \times R$, described above, with points on the horizontal lines L_i, $i = 1, \ldots, r$, generated by a Poisson process $N^{(r)}$ on L_r with intensity τ_r, and a sequence of successive binomial thinnings $N^{(k)}$ with deletion probabilities* $1 - \tau_k/\tau_{k+1}$, $k = 1, \ldots, r - 1$.

PROOF. This is similar to the proof of Theorem 5.5.1, in that one has to show that

(a) $E(N_T^*(B)) \to E(N(B))$ for all sets B of the form $(c, d] \times (\gamma, \delta]$, $0 < c < d$, $\gamma < \delta$, and
(b) $P\{N_T^*(B) = 0\} \to P\{N(B) = 0\}$ for all sets B which are finite unions of disjoint sets of this form.

Here, if $B = (c, d] \times (\gamma, \delta]$ and $(\gamma, \delta]$ contains exactly the lines $L_s, \ldots, L_t$,

$$E(N_T^*(B)) = E\left(\sum_{k=s}^{t} N_T^{(k)}((c, d])\right) = \sum_{k=s}^{t} T(d - c)\mu(u^{(k)}) \to (d - c)\sum_{k=s}^{t} \tau_k$$
$$= E(N(B))$$

so that (a) is satisfied.

To prove (b), as in the proof of Theorem 5.5.1 write B in the form

$$B = \bigcup_{k=1}^{m} C_k = \bigcup_{k=1}^{m}\left((c_k, d_k] \times \bigcup_{j=1}^{j_k} (\gamma_{kj}, \delta_{kj}]\right),$$

where $(c_k, d_k]$ and $(c_l, d_l]$ are disjoint for $k \neq l$. For each k, let m_k be the index of the lowest L_j that intersects C_k, i.e. $L_{m_k} \cap C_k \neq \emptyset$, $L_j \cap C_k = \emptyset$ for $j > m_k$.

Clearly, if $N_T^{(m_k)}((c_k, d_k]) = 0$ then either $N_T^*(C_k) = 0$ or there is an index $i < m_k$ such that $N_T^{(i)}((c_k, d_k]) > 0$, i.e. in $(c_k, d_k]$ there are more upcrossings of a higher level than of a lower level. Since obviously $N_T^*(C_k) = 0$ implies $N_T^{(m_k)}((c_k, d_k]) = 0$,

$$0 \leq P\left(\bigcap_{k=1}^{m} \{N_T^{(m_k)}((c_k, d_k]) = 0\}\right) - P\{N_T^*(B) = 0\} \to 0,$$

since the difference is bounded by the probability that some higher level has more upcrossings than a lower level, which tends to zero by Lemma 9.3.1.

But

$$\{N_T^{(m_k)}((c_k, d_k]) = 0\} \cap \{\xi_T(c_k) \leq x_{m_k}\} = \{M((Tc_k, Td_k]) \leq u_{T,k}\},$$

where $u_{T,k} = x_{m_k}/a_T + b_T$, so that Theorem 9.2.2 implies that

$$\lim_{T\to\infty} P\left(\bigcap_{k=1}^{m} \{N_T^{(m_k)}((c_k, d_k]) = 0\}\right) = \lim_{T\to\infty} P\left(\bigcap_{k=1}^{m} \{M((Tc_k, Td_k]) \leq u_{T,k}\}\right)$$
$$= \prod_{k=1}^{m} \exp\{-\tau_k'(d_k - c_k)\},$$

where $\tau_k' = \tau_{m_k} = e^{-x_{m_k}}$. Clearly this is just $P\{N(B) = 0\}$, and thus the proof of (b) is complete. □

Corollary 9.3.3. *Let $\{\xi(t)\}$ satisfy the conditions of Theorem 9.3.2, and let $B_1, \ldots, B_r$ be positive Borel sets, whose boundaries have Lebesgue measure zero. Then, for integers $m_j^{(k)}$,*

$$P\{N_T^{(k)}(B_j) = m_j^{(k)}, j = 1, \ldots, s, k = 1, \ldots, r\}$$
$$\to P\{N^{(k)}(B_j) = m_j^{(k)}, j = 1, \ldots, s, k = 1, \ldots, r\}.$$

For example, for disjoint B_1 and B_2, (with $|B| = m(B)$ for Lebesgue measure)

$$P\{N_T^{(1)}(B_1) = m_1^{(1)}, N_T^{(1)}(B_2) = m_2^{(1)}, N_T^{(2)}(B_2) = m_2^{(2)}\}$$
$$\to \exp(-\tau_1|B_1|)\frac{(\tau_1|B_1|)^{m_1^{(1)}}}{m_1^{(1)}!} \cdot \exp(-\tau_2|B_2|)\frac{(\tau_2|B_2|)^{m_2^{(2)}}}{m_2^{(2)}!}$$
$$\cdot \binom{m_2^{(2)}}{m_2^{(1)}}\left(\frac{\tau_1}{\tau_2}\right)^{m_2^{(1)}}\left(1 - \frac{\tau_1}{\tau_2}\right)^{m_2^{(2)} - m_2^{(1)}}$$

9.4. Location of Maxima

So far, we have examined the extremal properties of a normal process $\xi(t)$ by sections at certain (increasing) levels. Even if this gives perfect information about the height of the global maximum of the process, it does not directly tell us where this maximum occurs or how it is related to possible lower *local maxima*.

The maximum of any continuous process $\xi(t)$, $0 \le t \le T$, is attained in $[0, T]$. However, the maximum level may be reached many times, or even infinitely often. But there will—by continuity of $\xi(t)$—be a first occasion on which $\xi(t)$ attains its maximum in $[0, T]$, and we denote this by $L(T)$.

We state the first result concerning $L(T)$ as a lemma, though it is rather obvious.

Lemma 9.4.1. *$L(T)$ is a r.v. For $0 \le t \le T$, $P\{L(T) \le t\} = P\{M((0, t]) \ge M((t, T])\}$.*

PROOF. Both statements follow from the equivalence of the events $\{L(T) \le t\}$ and $\{M((0, t]) \ge M((t, T])\}$, the latter being measurable since $M((0, t])$ and $M((t, T])$ are r.v.'s.

The distribution of $L(T)$ can have a jump at 0 and at T as simple examples (such as the process $\xi(t) = A\cos(t - \phi)$ with $T < 2\pi$) show. However, for general continuous processes a simple condition precludes the possibility of any other jumps in the distribution of $L(T)$.

Specifically we will say that $\xi(t)$ has a *derivative in probability* at t_0 if there exists an r.v. η such that

$$\frac{\xi(t_0 + h) - \xi(t_0)}{h} \to \eta \quad \text{in probability as } h \to 0.$$

Clearly, if ξ has a q.m. or probability-one derivative, it has a derivative in probability (with the same value).

Theorem 9.4.2. *Suppose that $\xi(t)$ has a derivative in probability at t (where $0 < t < T$), and that the distribution of this derivative is continuous at zero. Then $P\{L(T) = t\} = 0$.*

PROOF. Let η denote the derivative in probability at t. Clearly

$$\{L(T) = t\} \subset \left\{\frac{\xi(t) - \xi(t-h)}{-h} \le 0\right\} \cap \left\{\frac{\xi(t) - \xi(t+h)}{h} \ge 0\right\}$$

for all $h > 0$ such that $0 \le t - h$ and $t + h \le T$.

Now $(\xi(t+h) - \xi(t))/h \to \eta$ in probability as $h \downarrow 0$ and there exists a sequence $\{h_n\}$ such that $(\xi(t + h_n) - \xi(t))/h_n \to \eta$ with probability one. By considering a subsequence of $\{h_n\}$ we may also arrange that $(\xi(t - h_n) - \xi(t))/(-h_n) \to \eta$ with probability one, i.e. on a set B with $P(B) = 1$. We see that $\eta = 0$ on $\{L(T) = t\} \cap B$, i.e. $\{L(T) = t\} \cap B \subset \{\eta = 0\}$ and hence

$$P\{L(T) = t\} \le P(\{L(T) = t\} \cap B) + P(B^c) \le P\{\eta = 0\} + P(B^c) = 0,$$

when η has a continuous distribution at zero. □

Turning to stationary processes, one may be tempted to conjecture that if $\xi(t)$ is stationary, then $L(T)$ is uniformly distributed over $(0, T)$. For example, this is so if $\xi(t) = A\cos(t - \phi)$, with ϕ uniformly distributed over $(0, 2\pi]$, for $T = 2\pi$. (If A has a Rayleigh distribution and is independent of ϕ, $\xi(t)$, of course, is normal.)

If $T < 2\pi$, there is a positive probability of L being 0 or T, and $L(T)$ is not strictly uniform. However, its distribution is still uniform *between* 0 and T as a simple calculation shows.

In general, however, L need *not* be uniform in the open interval $(0, T)$, even if $\xi(t)$ is normal and stationary. As an example of this, let ϕ_1, ϕ_2, A_1, and A_2 be independent, with ϕ_1 and ϕ_2 uniform over $(0, 2\pi]$, and with A_1 and A_2 having Rayleigh distributions, and put $\xi(t) = A_1\cos(t - \phi) + A_2\cos(100t - \phi_2)$. Then $\xi(t)$ is a stationary normal process, and (e.g. by drawing a picture) it can be seen that if $A_1 < A_2$, $\phi_1 \in (3\pi/2, 2\pi]$, and $\phi_2 \in (\pi/4, 3\pi/4)$ then the maximum of $\xi(t)$ over $[0, \pi/2]$ is attained in the interval $(0, \pi/100]$. Hence $P\{L(\pi/2) \le \pi/100\} \le P\{A_1 < A_2, 3\pi/2 < \phi_1 \le 2\pi, \pi/4 < \phi_2 \le 3\pi/4\} = (\frac{1}{2})(\frac{1}{4})(\frac{1}{4}) = \frac{1}{32} > \frac{1}{50} = (\pi/100)/(\pi/2)$, and $L(\pi/2)$ cannot be uniform over $(0, \pi/2)$.

However, for a stationary *normal* process the distribution of L is always *symmetric* in the entire interval $[0, T]$, and possible jumps at 0 and T are equal in magnitude. This follows from the reversibility of a stationary normal process in the sense that $\{\xi(-t)\}$ has the same distribution as $\{\xi(t)\}$.

One method of removing boundaries, like 0 and T, is to let them disappear to infinity, and one may ask whether $L = L(T)$ might be asymptotically uniform as $T \to \infty$. For normal processes, this follows simply from the asymptotic independence of maxima over disjoint intervals, as was previously mentioned. We state these results here, as simple consequences of Theorem 9.2.2.

Theorem 9.4.3. *Let $\{\xi(t)\}$ be a stationary normal process (standardized as usual) with $\lambda_2 < \infty$, and suppose that $r(t)\log t \to 0$ as $t \to \infty$. Then*

$$P\{L(T) \le lT\} \to l \quad \text{as } T \to \infty \qquad (0 \le l \le 1).$$

Proof. With the usual notation, if $0 \le l \le 1$, $l^* = 1 - l$, and

$$X_T = a_{lT}(M((0, lT]) - b_{lT}),$$

$$Y_T = a_{l^*T}(M((lT, T]) - b_{l^*T}),$$

where a's and b's are given by (8.2.6), then

$$P\{X_T \le x, Y_T \le y\} \to \exp\{-e^{-x} - e^{-y}\}$$

as $T \to \infty$. Furthermore,

$$\begin{aligned} P\{L(T) \le lT\} &= P\{M((0, lT]) \ge M((lT, T])\} \\ &= P\left\{X_T - \frac{a_{lT}}{a_{l^*T}} Y_T \ge a_{lT}(b_{l^*T} - b_{lT})\right\}, \end{aligned}$$

where $a_{l^*T}/a_{lT} \to 1$ and $a_{lT}(b_{l^*T} - b_{lT}) \to \log(l^*/l)$. As $T \to \infty$ the above probability tends to

$$P\{X - Y \ge \log l^*/l\},$$

where X and Y are independent r.v.'s with common d.f. $\exp(-e^{-x})$, and an evaluation of this probability yields the desired value l. $\square$

9.5. Height and Location of Local Maxima

One consequence of Theorem 9.4.3 is that asymptotically the global maximum is attained in the interior $(0, T)$ and thus also is a local maximum. For sufficiently regular processes one might consider also smaller, but still high, *local maxima*, which are separated from $L(T)$.

We first turn our attention to continuously differentiable normal processes which are twice differentiable in quadratic mean.

In analogy to the development in Chapter 5, we shall consider the point process in the plane, which is formed by the suitably transformed local maxima of $\xi(t)$. (Note that since the process $\xi(t)$ is continuous, the path of $a_T(\xi(tT) - b_T)$ is also continuous, and although its visits to any bounded rectangle $B \subset R^2$ are approximately Poisson in number, they are certainly not *points*.)

Suppose $\xi(t)$, $0 \le t \le T$ has local maxima at the points s_i with height $\xi(s_i)$. Let a_T and b_T be the normalizing constants defined by (8.2.6), and define a point process in the plane by putting points at $(T^{-1}s_i, a_T(\xi(s_i) - b_T))$. We recall from Section 9.1 that asymptotically the upcrossings of the fixed level x by $a_T(\xi(t) - b_T)$ form a Poisson process with intensity $\tau = e^{-x}$ when time is normalized to t/T, and that an upcrossing of a level x is accompanied by an upcrossing of the higher level y with a probability $e^{-y}/e^{-x} = e^{-(y-x)}$.

When investigating the Poisson character of local maxima, a question of some interest is to what extent high-level upcrossings and high local maxima can replace each other. Obviously there must be at least one local maximum between an upcrossing of a certain level u and the next downcrossing of the same level, so that, loosely speaking, there are at least as many high maxima as there are high upcrossings. As will now be seen there are, with high probability, no more. In fact we shall see that this is true even when $T \to \infty$ in such a way that $T\mu = T(1/2\pi)\lambda_2^{1/2}\exp(-u^2/2)$ converges.

First recall the notation from Section 7.6

$$N'_u((a, b]) = \#\{s_i \in (a, b];\ \xi(s_i) > u\},$$
$$N'_u(T) = N'_u((0, T]).$$

Lemma 9.5.1. *If $\lambda_4 < \infty$ and $T \sim \tau/\mu = \tau 2\pi\lambda_2^{-1/2}\exp(u^2/2)$, then*

(i) $$E(N'_u(T)) \to \tau$$

and, with $N_u(T) = \#\{$upcrossings of u by $\xi(t), 0 \le t \le T\}$

(ii) $$P\{|N'_u(T) - N_u(T)| \ge 1\} \to 0$$

as $u \to \infty$.

Proof. First note that at least one of the following events occur,

$$\{N'_u(T) \ge N_u(T)\} \quad \text{or} \quad \{\xi(T) \ge u\}$$

and that in the latter case, $N'_u(T) \ge N_u(T) - 1$. Therefore

$$\begin{aligned} P\{|N'_u(T) - N_u(T)| \ge 1\} &\le E(|N'_u(T) - N_u(T)|) \\ &\le E(N'_u(T) - N_u(T)) + 2P\{\xi(T) \ge u\}, \end{aligned}$$

and since $E(N_u(T)) \sim \tau$ and $P\{\xi(T) \ge u\} \to 0$, (ii) is a direct consequence of (i). But $E(N'_u(T))$ is given by (7.6.3) and since $1 - \Phi(x) \sim \phi(x)/x$ as $x \to \infty$, it follows that, for some constant K, with $D = \lambda_4 - \lambda_2^2$,

$$\frac{T}{2\pi}\left(\frac{\lambda_4}{\lambda_2}\right)^{1/2}\left(1 - \Phi\left(u\left(\frac{\lambda_4}{D}\right)^{1/2}\right)\right) \le K\frac{T}{u}\phi\left(u\left(\frac{\lambda_4}{D}\right)^{1/2}\right) \le K\frac{T}{u}\phi(u) \to 0$$

while

$$\frac{T}{2\pi}(2\pi\lambda_2)^{1/2}\phi(u)\Phi\left(\frac{u\lambda_2}{D^{1/2}}\right) = \frac{T}{2\pi}\lambda_2^{1/2}\exp\left(-\frac{u^2}{2}\right)(1 + o(1)) \to \tau$$

as $u \to \infty$, which proves (i). $\square$

Theorem 9.5.2. *Suppose the standardized stationary normal process $\{\xi(t)\}$ has continuously differentiable sample paths and is twice quadratic mean differentiable (i.e. $\lambda_4 < \infty$), and suppose that $r(t) \log t \to 0$ as $t \to \infty$. Then the point process $N_T'^*$ of normalized local maxima $(s_i/T, a_T(\xi(s_i) - b_T))$ converges in distribution to a Poisson process N' on $(0, \infty) \times R$ with intensity measure equal to the product of Lebesgue measure and that defined by the increasing function $-e^{-x}$.*

Proof. By now familiar reasoning from Theorem A.1, it is sufficient to show that

(a) $E(N_T'^*(B)) \to E(N'(B)) = (d - c)(e^{-\gamma} - e^{-\delta})$ *for any set B of the form* $(c, d] \times (\gamma, \delta]$, $0 < c < d$, $\gamma < \delta$, and
(b) $P\{N_T'^*(B) = 0\} \to P\{N'(B) = 0\}$ for sets B which are finite unions of sets of this form.

To prove (a), we use Lemma 9.5.1(i). Then, with $u^{(1)} = \delta/a_T + b_T$, $u^{(2)} = \gamma/a_T + b_T$,

$$\begin{aligned} E(N_T'^*(B)) &= E(N'_{u^{(2)}}((Tc, Td])) - E(N'_{u^{(1)}}((Tc, Td])) \\ &\to (d - c)e^{-\gamma} - (d - c)e^{-\delta} = (d - c)(e^{-\gamma} - e^{-\delta}), \end{aligned}$$

since $T\mu(u^{(i)}) \to e^{-\delta}, e^{-\gamma}$ for $i = 1, 2$.

Part (b) is a consequence of Lemma 9.5.1(ii) and the multilevel upcrossing theorem, Theorem 9.3.2. Let $N_u(I)$, as before, denote the number of u-upcrossings by $\xi(t)$, $t \in I$, and write the set B in the form $\bigcup_j E_j \times F_j$, where $E_j = (c_j, d_j]$ are disjoint and each F_j is a finite union of disjoint intervals. Suppose first that there is only one set E_j, i.e. $B = E \times \bigcup_k G_k$, where $G_k = (\gamma_k, \delta_k]$, and write $u^{(2k-1)} = \delta_k/a_T + b_T$, $u^{(2k)} = \gamma_k/a_T + b_T$. According to Lemma 9.5.1(ii) asymptotically every upcrossing of the high level u is accompanied by one (and no more) local maximum above that level, and hence

$$\begin{aligned} P\{N_T'^*(E \times G_k) = 0\} &= P\{N'_{u^{(2k)}}((Tc, Td]) = N'_{u^{(2k-1)}}((Tc, Td])\} \\ &= P\{N_{u^{(2k)}}((Tc, Td]) = N_{u^{(2k-1)}}((Tc, Td])\} + o(1). \end{aligned}$$

By Theorem 9.3.2, with $\tau_{2k} = e^{-\gamma_k}$, $\tau_{2k-1} = e^{-\delta_k}$,

$$\begin{aligned} P\{N_{u^{(2k)}}&((Tc, Td]) = N_{u^{(2k-1)}}((Tc, Td])\} \\ &\to \sum_{j=0}^{\infty} \exp\{-\tau_{2k}(d - c)\} \frac{(\tau_{2k}(d - c))^j}{j!} \left(\frac{\tau_{2k-1}}{\tau_{2k}}\right)^j \\ &= \exp\{-(\tau_{2k} - \tau_{2k-1})(d - c)\} \\ &= \exp\{-(d - c)(-e^{-\delta_k} - (-e^{-\gamma_k}))\} \\ &= P\{N'(E \times G_k) = 0\}. \end{aligned}$$

By slightly extending the argument we obtain

$$\begin{aligned}P\left\{N_T'^*\left(E \times \bigcup_k G_k\right) = 0\right\} &= P\left(\bigcap_k \{N_{u^{(2k)}}((Tc, Td]) = N_{u^{(2k-1)}}((Tc, Td])\}\right) \\ &\quad + o(1) \\ &\to \exp\left\{-(d-c)\sum_k (e^{-\gamma_k} - e^{-\delta_k})\right\} \\ &= P\left\{N'\left(E \times \bigcup_k G_k\right) = 0\right\},\end{aligned}$$

and we have proved part (b) for sets B of the simple form $B = E \times \bigcup_k G_k$. The general proof of part (b) is only notationally more complex. □

The limiting Poisson process in Theorem 9.5.2 has exactly the same distribution as that in Theorem 5.7.2 for $\xi(t)$ normal, since $\log G(s) = -e^{-s}$ in this case. This means that all conclusions that can be drawn from that theorem about asymptotic properties of the normalized point process $a_n(\xi_i - b_n)$ also carry over to the normalized point process of local maxima $a_T(\xi(s_i) - b_T)$.

As an example we shall use Theorem 9.5.2 to give the simultaneous distribution of location and height of the two largest local maxima of $\xi(t), t \in (0, T]$. Let $M_1(T)$ be the highest and $M_2(T)$ the second highest local maximum, and $L_1(T)$, $L_2(T)$ their locations.

Theorem 9.5.3. *Suppose $\{\xi(t)\}$ satisfies the hypotheses of Theorem 9.5.2. Then*

$$\begin{aligned}P\{a_T(M_1(T) - b_T) \le x_1, L_1(T) \le l_1 T, a_T(M_2(T) - b_T) \le x_2, L_2(T) \le l_2 T\} \\ \to l_1 l_2 \exp(-e^{-x_2})(1 + e^{-x_2} - e^{-x_1}) \qquad (9.5.1)\end{aligned}$$

as $T \to \infty$, for $0 \le l_1, l_2 \le 1, x_2 \le x_1$.

Proof. The asymptotic distribution of the *heights* of the two highest local maxima,

$$\begin{aligned}P\{a_T(M_1(T) - b_T) \le x_1, a_T(M_2(T) - b_T) \le x_2\} \\ \to \exp(-e^{-x_2})(1 + e^{-x_2} - e^{-x_1}),\end{aligned}$$

follows in the same way as Theorem 5.6.2 formula (5.6.4) from the observation above that the limiting point process of normalized local maxima $(s_i/T, a_T(\xi(s_i) - b_T))$, $0 \le s_i \le T$, is the same as that of a normalized sequence of independent normal r.v.'s $(i/n, a_n(\xi_i - b_n))$, $i = 1, \ldots, n$.

But also the location of the local maxima can be obtained in this way. Suppose, e.g., $l_1 < l_2$, and write I, J, K for the intervals $(0, l_1 T]$, $(l_1 T, l_2 T]$, $(l_2 T, T]$, respectively. With $u^{(1)} = x_1/a_T + b_T$, $u^{(2)} = x_2/a_T + b_T$ the event in (9.5.1) can be expressed in terms of the highest and second highest local maxima over I, J, K as (with obvious notation)

$$\begin{aligned}\{M_1(I) \le u^{(1)}, M_2(I) \le u^{(2)}, M_1(J) \le u^{(2)}, M_1(J) \le M_1(I), M_1(K) \\ \le M_2(I \cup J)\}\end{aligned}$$

and the limit of the probability of this event, when expressed in terms of the point process $N_T'^*$ of local maxima, is again the same as it would be for the point process of normalized independent r.v.'s. For such a process obviously $L_1(T)/T$ and $L_2(T)/T$ are independent and uniformly distributed over (0, 1) and independent of the heights of the maxima, which proves the theorem. $\square$

9.6. Maxima Under More General Conditions

We have investigated the local maxima under the rather restrictive assumption that $\xi(t)$ is twice differentiable (in quadratic mean), i.e. $\lambda_4 < \infty$. If $\lambda_4 = \infty$ the mean number of zeros of $\xi'(t)$ is infinite, by Rice's formula, and in fact infinitely close to every local maximum there may be infinitely many more, which precludes the possibility of a Poisson type limit theorem for the locations of local maxima.

One way of circumventing this difficulty is to exclude a small interval around each high maximum from further considerations, starting with the highest. To be more precise, let

$$M_1(T) = \sup\{\xi(t);\ t \in (0, T)\}$$

be the global maximum, and

$$L_1(T) = \inf\{t > 0;\ \xi(t) = M_1(T)\}$$

its location. For $\varepsilon > 0$ an arbitrary but fixed constant, let $I_1 = (0, L_1(T) - \varepsilon) \cup (L_1(T) + \varepsilon, T)$, and define

$$M_{2,\varepsilon}(T) = \sup\{\xi(t);\ t \in I_1\},$$
$$L_{2,\varepsilon}(T) = \inf\{t \in I_1;\ \xi(t) = M_{2,\varepsilon}(T)\}.$$

Proceeding recursively, with

$$I_k = I_{k-1} \cap [L_{k-1,\varepsilon}(T) - \varepsilon, L_{k-1,\varepsilon}(T) + \varepsilon]^c,$$

we get a sequence $M_{k,\varepsilon}(T)$, $L_{k,\varepsilon}(T)$, $(M_{1,\varepsilon}(T) = M_1(T), L_{1,\varepsilon}(T) = L_1(T))$ of heights and locations of *ε-maxima*, and there are certainly only a finite number of these in any finite interval. In fact, it is not difficult to relate these variables to the point processes of upcrossings (in the same way as regular local maxima can be replaced by upcrossings of high levels if $\lambda_4 < \infty$) and thereby obtain the following Poisson limit theorem, the proof of which is omitted.

Theorem 9.6.1. *Suppose $\{\xi(t)\}$ is a standardized normal process with $\lambda_2 < \infty$ and with $r(t)\log t \to 0$ as $t \to \infty$. Then the point process $N_T^{(\varepsilon)}$ of normalized ε-maxima $(L_{i,\varepsilon}(T)/T, a_T(M_{i,\varepsilon}(T) - b_T))$ converges in distribution to the same Poisson process N' in the plane as in Theorem* 9.5.2.

Note that the limiting properties are independent of the ε chosen. We shall return to this topic for more irregular processes in Chapter 12.

CHAPTER 10

Sample Path Properties at Upcrossings

Our main concern in the previous chapter has been the numbers and locations of upcrossings of high levels, and the relations between the upcrossings of several adjacent levels. For instance, we know from Theorem 9.3.2 and relation (9.2.3) that for a standard normal process each upcrossing of the high level $u = u_\tau$ with a probability $p = \tau^*/\tau$ is accompanied by an upcrossing also of the level

$$u_{\tau^*} = u - \frac{\log p}{u},$$

asymptotically independently of all other upcrossings of u_τ and u_{τ^*}.

We shall in this chapter show that the empirical distributions of the values of ξ(t) after a u-upcrossing converge, and shall represent the limiting distributions as the distribution of a certain model process. By studying in more detail the behaviour of this model process, we will then attempt to throw some further light on the structure of the sample paths of $\xi(t)$ near an upcrossing of a high level u.

10.1. Marked Upcrossings

We assume that $\{\xi(t)\}$ is a stationary normal process on the entire real line with $E(\xi(t)) = 0$, $E(\xi^2(t)) = 1$ and covariance function $r(\tau)$ satisfying

$$r(\tau) = 1 - \tfrac{1}{2}\lambda_2 \tau^2 + o(\tau^2) \quad \text{as } \tau \to 0. \tag{10.1.1}$$

With a slightly more restrictive assumption,

$$-r''(\tau) = \lambda_2 + O(|\log|\tau||^{-a}) \quad \text{as } \tau \to 0 \tag{10.1.2}$$

for some $a > 1$, we can assume that $\{\xi(t)\}$ has continuously differentiable sample paths, (cf. condition (7.3.2) for sample function continuity) and we will do so since it serves our purposes of illustration. We also assume throughout this chapter that for each choice of distinct nonzero points $s_1, \ldots, s_n$, the distribution of $\xi(0)$, $\xi'(0)$, $\xi(s_1), \ldots, \xi(s_n)$ is nonsingular. (A sufficient condition for this is that the spectral distribution function $F(\lambda)$ has a continuous component; see Cramér and Leadbetter (1967, Section 10.6).)

Since $\lambda_2 = -r''(0) < \infty$ the number of upcrossings of the level u in any bounded interval has a finite expectation and so will be finite with probability one. Let

$$\cdots < t_{-1} < t_0 < t_1 < t_2 < \cdots$$

with $t_0 \leq 0 < t_1$, be the locations of the upcrossings of u by $\{\xi(t)\}$, and note that $|t_k| \to \infty$ as $|k| \to \infty$. As before, we denote by N_u the point process of upcrossings with events at $\{t_k\}$.

In order to retain information about $\{\xi(t)\}$ near its upcrossings, we now attach to each t_k a *mark* η_k. In Chapter 7 each mark was simply a real number (e.g. in Section 7.6 where the marks were the values of $\xi(t)$ at the downcrossing zeros of $\xi'(t)$). Here the useful marks are more abstract, and in fact we take η_k to be the function defined by

$$\eta_k(t) = \xi(t_k + t).$$

Thus, the mark $\{\eta_k(t)\}$ is the entire sample function of $\{\xi(t)\}$ translated back by the distance t_k. By assumption ξ is continuously differentiable and has finite number of upcrossings in finite intervals. Since further $\xi'(t_k) > 0$ at any upcrossing it is easily seen that each $\eta_k(t)$ is a r.v. For, with $\tau_1^{(n)} = \min\{i/n, i = 1, 2, \ldots; \xi((i-1)/n) < u < \xi(i/n)\}$, clearly $\tau_1^{(n)} \to t_1$ and $\xi(\tau_1^{(n)} + t) \to \xi(t_1 + t)$, a.s., so that $\eta_1(t)$, being a limit of r.v.'s is a r.v. In general, let $\tau_k^{(n)} = \min\{i/n,\ i = 1, 2, \ldots;\ \xi(t_{k-1} + (i-1)/n) < u < \xi(t_{k-1} + i/n)\}$, leading to $\xi(t_{k-1} + \tau_k^{(n)} + t) \to \xi(t_k + t)$, and hence also $\eta_k(t)$ is a r.v.

In particular, $\eta_k(0) = u$, while for small t-values $\eta_k(t)$ describes the behaviour of the ξ-process in the immediate vicinity of its kth u-upcrossing

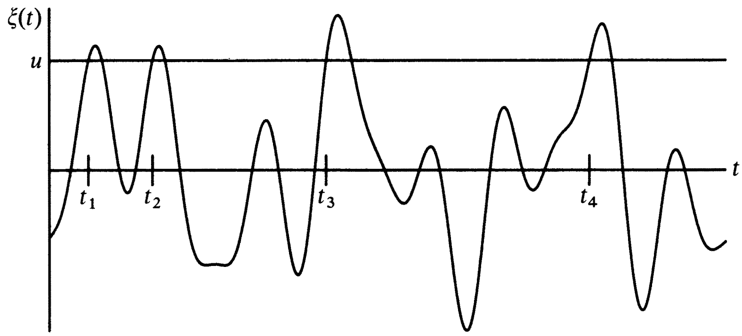


Figure 10.1.1. Point process $\{t_k\}$ of upcrossings for $\xi(t)$, $t \geq 0$.

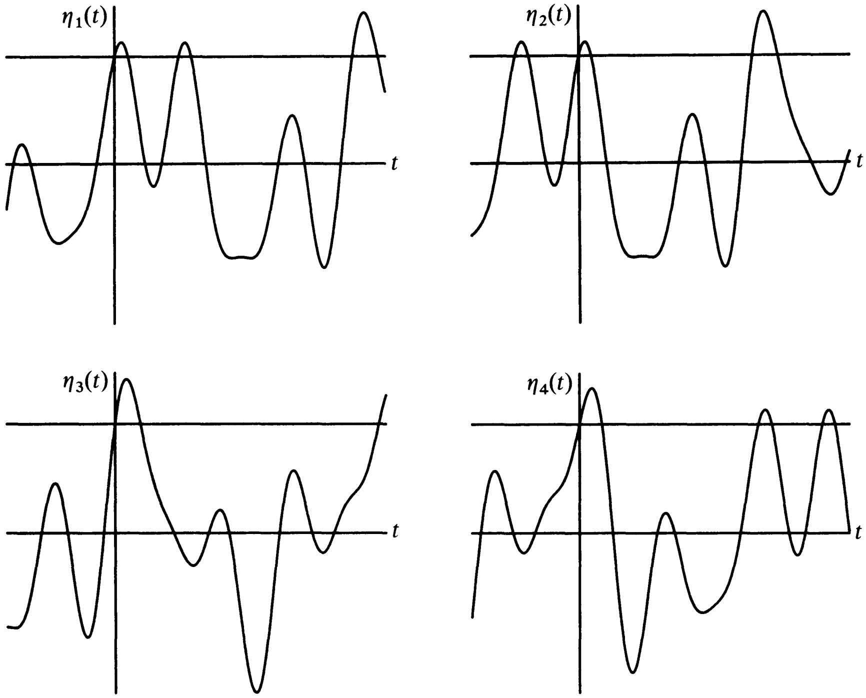


Figure 10.1.2. Realization of marks $\eta_k(t) = \xi(t_k + t)$ describing $\xi(\cdot)$ after upcrossings.

t_k. Of course, any of the marks, say $\{\eta_0(t)\}$, contains perfect information about all the upcrossings and all other marks, so different $\{\eta_k(t)\}$ are totally dependent.

The marks are furthermore constrained by the requirement that t_1 is the first upcrossing of u after zero, t_2 the second and so on, which suggests that the different marks ..., $\{\eta_1(t)\}$, $\{\eta_2(t)\}$, ... are not identically distributed.

A realization of $\{\xi(t)\}$ generates a realization of the sequence of marks $\{\eta_1(t)\}$, $\{\eta_2(t)\}$, ..., and this is what will actually be observed over a long period of time. For each such realization one can form the *empirical distribution* of the observed values $\eta_k(s)$, $k = 1, 2, \ldots$, i.e. of the values $\xi(t_k + s)$ of the process a fixed time s after the u-upcrossings. We shall see below that this empirical distribution converges, for an increasing observation interval, and we shall consider the limit as the marginal distribution of an "arbitrary" mark at time s. Similarly, one can obtain joint distributions of an "arbitrary" mark at times $s_1, \ldots, s_n$ and of several consecutive marks. Such distributions are the main topic of this chapter. In point process theory similar objects are called *Palm distributions* (or *Palm measures*), and they formalize the notion of conditional distribution given that the point process has a point at a specified time τ.

10.2. Empirical Distributions of the Marks at Upcrossings

The point process N_u on the real line formed by the upcrossings of the level u by $\{\xi(t)\}$ is stationary and without multiple events, i.e. the joint distribution of $N_u(t + I_j), j = 1, \ldots, n$ does not depend on t, and $N_u(\{t\})$ is either 0 or 1; (here $t + I_j$ is the set I_j translated by an amount t, i.e. $t + I_j = \{t + s; s \in I_j\}$). Further, N_u is jointly stationary with $\xi(t)$, i.e. the distribution of $N_u(t + I_j), \xi(t + s_j), j = 1, \ldots, n$ does not depend on t.

Let $s_j, j = 1, \ldots, n$, and $y_j, j = 1, \ldots, n$, be fixed numbers. For each upcrossing, t_k, it is then possible to check whether $\xi(t_k + s_j) \le y_j, j = 1, \ldots, n$, or not. Now, starting from the point process N_u, form a new process $\tilde{N}_u$ by deleting all points t_k in N_u that do not satisfy $\xi(t_k + s_j) \le y_j, j = 1, \ldots, n$. The relative number of points in this new point process tells us how likely it is (for the particular sample function) that a point t_k in N_u is accompanied by a mark $\eta_k(t)$, satisfying $\eta_k(s_j) \le y_j, j = 1, \ldots, n$.

More explicitly, we shall assume that $\xi(t)$ has been observed in the time interval $(0, T]$ and, writing $s = (s_1, \ldots, s_n)$, $y = (y_1, \ldots, y_n)$, define the empirical distribution functions F_s^T of a mark as the relative frequency of upcrossings satisfying $\xi(t_k + s_j) \le y_j, j = 1, \ldots, n$, i.e.

$$F_s^T(y) = \frac{\tilde{N}_u(T)}{N_u(T)} = \frac{\#\{t_k \in (0, T]; \xi(t_k + s_j) \le y_j, j = 1, \ldots, n\}}{\#\{t_k \in (0, T]\}}.$$

It is a characteristic property of *ergodic* processes that empirical distributions such as F_s^T, as $T \to \infty$ converge a.s. to a specific, nonstochastic limit, in this case (as will be seen) to

$$F_s(y) = \frac{E(\tilde{N}_u(1))}{E(N_u(1))}.$$

We recall here that a (strictly) stationary sequence $\{\xi_n\}$ is ergodic if, for every (measurable) function $h(\xi.)$ of the entire sequence, the time average of h over successively translated copies of the sequence

$$\frac{1}{n} \sum_{k=1}^{n} h(\xi_{.+k})$$

as $n \to \infty$ converges, with probability one, to $E(h(\xi.))$ provided this is finite. Here $\xi_{.+k}$ denotes the original sequence $\xi.$ translated k time units to the left. For a continuous time process the sum has to be replaced by an integral.

Furthermore, if $\{\xi(t)\}$ is a continuous time ergodic process and the r.v.'s ζ_i are defined as a function of $\xi(t), i - 1 < t \le i$, i.e. $\zeta_i = g(\xi(t), i - 1 < t \le i)$, then $\{\zeta_i\}$ is a stationary ergodic sequence; for more details about ergodicity, see, for example, Breiman (1968).

For normal processes and sequences there exists a simple criterion for ergodicity, viz. that the spectral distribution should contain no discrete part, i.e. the spectral d.f. should be a continuous function; this was proved by Maruyama (1949) and Grenander (1950).

Theorem 10.2.1. *If the process $\{\xi(t)\}$ is ergodic and $E(N_u(1)) < \infty$, then, with probability one,*

$$F_s^T(y) \to F_s(y) = \frac{E(\tilde{N}_u(1))}{E(N_u(1))} \quad \textit{as } T \to \infty. \tag{10.2.1}$$

PROOF. If $\{\xi(t)\}$ is ergodic, then $\zeta_i = \#\{t_k \in (i-1, i]\}$ and $\tilde{\zeta}_i = \#\{t_k \in (i-1, i];\ \xi(t_k + s_j) \le y_j, j = 1, \ldots, n\}$ are also ergodic, and thus

$$\frac{1}{T}\sum_{i=1}^{T} \zeta_i \to E(\zeta_1), \qquad \frac{1}{T}\sum_{i=1}^{T} \tilde{\zeta}_i \to E(\tilde{\zeta}_1) \tag{10.2.2}$$

when $T \to \infty$ through integer values. Now, e.g.

$$\frac{1}{T} N_u(T) = \frac{[T]}{T} \cdot \frac{1}{[T]} \sum_{i=1}^{[T]} \zeta_i + \frac{1}{T} \zeta_T', \quad \text{say},$$

where $0 \le \zeta_T' \le \zeta_{[T]+1}$. Since $([T] + 1)^{-1}\zeta_{[T]+1} \to 0$ by (10.2.2) it follows that

$$\frac{1}{T} N_u(T) \to E(\zeta_1) = E(N_u(1)),$$

and similarly

$$\frac{1}{T} \tilde{N}_u(T) \to E(\tilde{\zeta}_1) = E(\tilde{N}_u(1)),$$

with probability one, as $T \to \infty$, which proves (10.2.1). □

Up to this point we have, as customarily, not been specific about the basic underlying probability space. Now, let C_1 be the space of continuously differentiable functions defined on the real line. We shall denote a typical element in C_1 by ξ and let $\xi(t)$ be the value of ξ at t, in complete agreement with previous usage. We can then think of the probability measure P for our original process $\{\xi(t)\}$ as defined on C_1 or, more precisely, on the smallest σ-field $\mathscr{F}$ of subsets of C_1 which makes all the projections $\xi \curvearrowright \xi(t)$ measurable.

As a generalization of the point process $\tilde{N}_u$ studied above, define, for each $E \in \mathscr{F}$, a new point process $\tilde{N}_{u,E}$ by deleting from N_u all points t_k for which $\xi(t_k + \cdot) \notin E$, i.e. for which the function $\eta_k(\cdot) = \xi(t_k + \cdot)$ does not belong to E. Then, motivated by (10.2.1), we define a second probability measure P^u on C_1 by

$$P^u(E) = \frac{E(\tilde{N}_{u,E}(1))}{E(N_u(1))} = \frac{E(\#\{t_k \in (0, 1];\ \xi(t_k + \cdot) \in E\})}{E(\#\{t_k \in (0, 1]\})}.$$

We shall call P^u the *Palm distribution* or the *ergodic distribution* of ξ after a u-upcrossing. By additivity of expectations, P^u is in fact a probability measure, and the finite-dimensional distribution of $\xi(s_j)$, $j = 1, \dots, n$ are clearly exactly the functions $F_{\mathbf{s}}$ appearing on (10.2.1). Thus

$$P^u\{\xi(s_j) \le y_j, j = 1, \dots, n\} = F_{\mathbf{s}}(\mathbf{y})$$
$$= \lim_{T\to\infty} \frac{\#\{t_k \in (0, T]; \xi(t_k + s_j) \le y_j, j = 1, \dots, n\}}{\#\{t_k \in (0, T]\}}.$$

Obviously, the result of Theorem 10.2.1 about the empirical distributions of $\eta_k(t) = \xi(t_k + t)$ holds also with $\tilde{N}_u$ replaced by $\tilde{N}_{u,E}$, for arbitrary $E \in \mathscr{F}$, i.e. with P-probability one,

$$\frac{\tilde{N}_{u,E}(T)}{N_u(T)} \to \frac{E(\tilde{N}_{u,E}(1))}{E(N_u(1))} = P^u(E).$$

This leads to the convergence of the empirical distributions of many other interesting functionals, such as the excursion time, i.e. that time from an upcrossing to the next downcrossing of the same level, and the maximum in intervals of fixed length following an upcrossing, i.e.

$$P^u\left\{\sup_{t\in I} \xi(t) \le x\right\} = \lim_{T\to\infty} \frac{\#\left\{t_k \in (0, T]; \sup_{t\in I} \xi(t_k + t) \le x\right\}}{\#\{t_k \in (0, T]\}};$$

for more examples, see Lingren (1977).

The measure P^u formalizes the notion of a conditional distribution of the process given that it has a u-upcrossing at time 0, and by Theorem 10.2.1 it describes the long run properties of $\xi(t_k + t)$ as a function of t, when t_k runs through the set of all positive u-upcrossings. One can therefore interpret the P^u-distribution of $\xi(t)$ as "the conditional distribution of the original process at time t after an arbitrary u-upcrossing". In particular

$$P^u\{\xi(t) \text{ has an upcrossing of } u \text{ at } t = 0\} = 1,$$

so that $\eta_0 = \xi$ with P^u-probability one.

If $\xi(t)$ is stationary normal, with continuously differentiable sample paths, it can be shown (although this involves crossing theory for nonstationary processes) that further, with obvious notation, $E^u(N_u(t)) < \infty$ for all $t > 0$, so that, under P^u, there are only a finite number of u-upcrossings in any finite interval, and thus the marks $\eta_k(t)$, $k = 1, 2, \dots$ are well defined. The following result gives further motivation for thinking of P^u as the distribution of an arbitrary mark in the original process, and makes precise the intuitive notion that under P^u the marks $\eta_k(t) = \xi(t_k + t)$ have the same distribution, for $k = 0, 1, \dots$, all being equally "arbitrary".

Theorem 10.2.2. *Suppose $\{\xi(t)\}$ is stationary normal with continuously differentiable sample paths (e.g. satisfies the general hypothesis of the previous* *section). Then the sequence of marks $\{\eta_0(t)\}, \{\eta_1(t)\}, \ldots$ is stationary under the Palm distribution P^u, in the sense that the finite-dimensional P^u-distribution of*

$$\eta_{k+k_1}(s_j), \ldots, \eta_{k+k_r}(s_j), \qquad j = 1, \ldots, n$$

is independent of k. In particular, for a fixed t, all the $\eta_j(t), j = 0, 1, \ldots,$ form a stationary real sequence, and hence have the same P^u-distribution, whereas they are nonidentically distributed under P.

Proof. We only show that, under P^u, the distribution of $\eta_j(t)$ is the same as that of $\eta_{j-1}(t)$. A full proof is only notationally more complicated.

Put $\mu = E(N_u(0, 1]) = E(\#\{t_k \in (0, 1]\})$. Then

$$\begin{aligned} P^u\{\eta_j(t) \le y\} &= \mu^{-1}E(\#\{t_k \in (0, 1]; \xi(t_{k+j} + t) \le y\}) \\ &= (\mu T)^{-1}E(\#\{t_k \in (0, T]; \xi(t_{k+j} + t) \le y\}). \quad (10.2.3) \end{aligned}$$

Similarly,

$$P^u\{\eta_{j-1}(t) \le y\} = (\mu T)^{-1}E(\#\{t_k \in (0, T]; \xi(t_{k+j-1} + t) \le y\}). \quad (10.2.4)$$

Now take a pair of adjacent points t_k and t_{k+1}. We see that t_k is counted in (10.2.3) if and only if

$$t_k \in (0, T] \quad \text{and} \quad \xi(t_{k+j} + t) \le y$$

while t_{k+1} contributes to (10.2.4) if and only if

$$t_{k+1} \in (0, T] \quad \text{and} \quad \xi(t_{k+1+j-1} + t) \le y,$$

i.e. if and only if

$$t_{k+1} \in (0, T] \quad \text{and} \quad \xi(t_{k+j} + t) \le y.$$

Hence the numbers in (10.2.3) and (10.2.4) differ at most by $+1$ or -1 so that

$$|P^u\{\eta_j(t) \le y\} - P^u\{\eta_{j-1}(t) \le y\}| \le \frac{1}{\mu T} \to 0 \quad \text{as} \quad T \to \infty. \qquad \square$$

Palm probabilities can also be obtained as limits of ordinary conditional probabilities given a point, i.e. an upcrossing, not exactly at 0, but somewhere nearby. Let t_0 be the last u-upcrossing for $\xi(t)$ prior to 0. Then

$$\begin{aligned} &P^u\{\xi(s_j) \le y_j, j = 1, \ldots, n\} \\ &\quad = \lim_{h \downarrow 0} P\{\xi(t_0 + s_j) \le y_j, j = 1, \ldots, n \mid -h < t_0 \le 0\}, \quad (10.2.5) \end{aligned}$$

where it may be shown that the limit (10.2.5) exists, and equals the ratio (10.2.1). In fact, (10.2.5) can be taken as a definition of the Palm distributions, an approach which was taken by Kac and Slepian (1959), who also termed it the *horizontal window* conditioning of crossings, indicating that the sample path $\xi(t)$ has to pass through a horizontal window $\{(t, y); -h \le t \le 0, y = u\}$. This is in contrast to *vertical window* conditioning which requires that $u \le \xi(0) \le u + h$, $\xi'(0) > 0$, so that the process has to pass through a vertical window $\{(t, x); t = 0, u \le x \le u + h\}$ with positive slope. The horizontal and vertical conditioned distributions will differ in one particular respect, namely, in the marginal distribution of the derivative $\eta_k'(0) = \xi'(t_k)$ of the mark at the upcrossing; cf. the end of the next section

10.3. The Slepian Model Process

We will devote the rest of this chapter to a more detailed study of the properties of the marks under the Palm distribution, in particular, as the level u gets high. In view of Theorem 10.2.2 all $\{\eta_k(t)\}$ have the same P^u distribution and we pick $\eta_0(t) = \xi(t)$ as a typical representative.

Our tool will be an explicit representation of the P^u-distribution of $\xi(t)$ in terms of a simple process, originally introduced by D. Slepian (1963) and therefore in this work termed a *Slepian model process.* The following theorem uses the definition of Palm distributions and forms the basis for the Slepian representation.

Theorem 10.3.1. *Under the hypothesis of Theorem* 10.2.2 *let* $\mu = E(N_u(1)) = (1/2\pi)\lambda_2^{1/2} \exp(-u^2/2)$. *Then for* $t \neq 0$,

$$P^u\{\xi(t) \le y\} = \int_{x=-\infty}^{y} \left\{\mu^{-1} \int_{z=0}^{\infty} zp(u, z)p(x \mid u, z)\, dz\right\} dx, \quad (10.3.1)$$

where $p(u, z)$ *is the joint density of* $\xi(0)$ *and its derivative* $\xi'(0)$, *and* $p(x \mid u, z)$ *is the conditional density of* $\xi(t)$ *given* $\xi(0) = u$, $\xi'(0) = z$. *Thus the* P^u*-distribution of* $\xi(t) = \eta_0(t)$ *is absolutely continuous, with density*

$$\mu^{-1} \int_{z=0}^{\infty} zp(u, z)p(x \mid u, z)\, dz.$$

The n-dimensional P^u*-distribution of* $\xi(s_1), \dots, \xi(s_n)$ *is obtained by replacing* $p(x \mid u, z)$ *by* $p(x_1, \dots, x_n \mid u, z)$, *the conditional P-density of* $\xi(s_1), \dots, \xi(s_n)$ *given* $\xi(0) = u$, $\xi'(0) = z$.

Proof. The one-dimensional form (10.3.1) is a direct consequence of Lemmas 7.5.1(iii) and 7.5.2 since we have assumed that $\xi(0)$ and $\xi(t)$ have a non-singular distribution. We can take $\zeta(s) = \xi(s)$, $\zeta'(s) = \xi'(s)$, $\eta(s) = \xi(s + t)$ and, in the same way as in the proof of Lemma 7.6.1, check that

$$P\{\zeta(t) = u, \eta(t) = v \quad \text{for some} \quad t \in (0, 1]\} = 0$$

so that

$$E(\tilde{N}_u(1)) = \int_{z=0}^{\infty} zp(u, z)P\{\xi(t) \le y \mid \xi(0) = u, \xi'(0) = z\}\, dz$$

$$= \int_{x=-\infty}^{y} \int_{z=0}^{\infty} zp(u, z)p(x \mid u, z)\, dz\, dx.$$

The multivariate version is proved in an analgous way. □

Theorem 10.3.1 states that the joint density of $\xi(s_1), \dots, \xi(s_n)$ under P^u is given by

$$\mu^{-1} \int_{z=0}^{\infty} z p(u, z) p(x_1, \dots, x_n | u, z)\, dz, \tag{10.3.2}$$

where $p(x_1, \dots, x_n | u, z)$ is the conditional P-density of $\xi(s_1), \dots, \xi(s_n)$ given $\xi(0) = u$, $\xi'(0) = z$. We shall now evaluate (10.3.2) in order to obtain the Slepian model process.

With $\mu = (1/2\pi)\lambda_2^{1/2} \exp(-u^2/2)$, and using the fact that $\xi(0)$ and $\xi'(0)$ are independent and normal with $E(\xi'(0)) = 0$, $E(\xi'(0)^2) = \lambda_2$, we have that

$$p(u, z) = \frac{\mu}{\lambda_2} \exp\left(-\frac{z^2}{2\lambda_2}\right)$$

and we can write (10.3.2) in the form

$$\int_{z=0}^{\infty} \frac{z}{\lambda_2} \exp\left(-\frac{z^2}{2\lambda_2}\right) p(x_1, \dots, x_n | u, z)\, dz. \tag{10.3.3}$$

The covariance matrix of $\xi(0)$, $\xi'(0)$, $\xi(s_1), \dots, \xi(s_n)$ is

$$\begin{bmatrix} 1 & 0 & r(s_1) & \dots & r(s_n) \\ 0 & \lambda_2 & -r'(s_1) & \dots & -r'(s_n) \\ r(s_1) & -r'(s_1) & 1 & \dots & r(s_n - s_1) \\ \vdots & \vdots & \vdots & & \vdots \\ r(s_n) & -r'(s_n) & r(s_1 - s_n) & \dots & 1 \end{bmatrix}$$

From standard properties of conditional normal densities—see Rao (1973, p. 522)—it follows that $p(x_1, \dots, x_n | u, z)$ is an n-variate normal density and that

$$E(\xi(s_i) | \xi(0) = u, \xi'(0) = z) = u r(s_i) - \frac{z r'(s_i)}{\lambda_2} \tag{10.3.4}$$

and

$$\begin{aligned} &\operatorname{Cov}(\xi(s_i), \xi(s_j) | \xi(0) = u, \xi'(0) = z) \\ &\qquad = r(s_i - s_j) - r(s_i) r(s_j) - \frac{r'(s_i) r'(s_j)}{\lambda_2}. \end{aligned} \tag{10.3.5}$$

The density (10.3.3) is therefore a mixture of n-variate normal densities, all with the same covariances (10.3.5), but with different means (10.3.4), and mixed in proportion to the Rayleigh density

$$\frac{z}{\lambda_2} \exp\left(-\frac{z^2}{2\lambda_2}\right), \qquad z > 0. \tag{10.3.6}$$

Now we are ready to introduce the Slepian model process. Let ζ have a Rayleigh distribution with density (10.3.6) and let $\{\kappa(t), t \in R\}$ be a non-stationary normal process, independent of ζ with zero mean, and with the covariance function

$$r_\kappa(s, t) = \mathrm{Cov}(\kappa(s), \kappa(t)) = r(s - t) - r(s)r(t) - \frac{r'(s)r'(t)}{\lambda_2}.$$

That this actually is a covariance function follows from (10.3.5). Hence on some probability space which need not be specified further, there exist κ and ζ with these properties. The process

$$\xi_u(t) = ur(t) - \frac{\zeta r'(t)}{\lambda_2} + \kappa(t) \tag{10.3.7}$$

is called a *Slepian model process* for $\xi(t)$ after u-upcrossings. Obviously, conditional on $\zeta = z$, the process (10.3.7) is normal with mean and covariances given by the right-hand side of (10.3.4) and (10.3.5), respectively, and so its finite-dimensional distributions are given by the densities (10.3.3).

Theorem 10.3.2. *Under the hypothesis of Theorem* 10.2.2, *the finite-dimensional Palm distributions of the mark* $\{\eta_0(t)\}$ *and thus, by Theorem* 10.2.2, *of all marks* $\{\eta_k(t)\}$, *are equal to the finite-dimensional distributions of the Slepian model process*

$$\xi_u(t) = ur(t) - \frac{\zeta r'(t)}{\lambda_2} + \kappa(t)$$

i.e.

$$P^u\{\xi(s_j) \in B_j, j = 1, \ldots, n\} = P\{\xi_u(s_j) \in B_j, j = 1, \ldots, n\}$$

for any Borel sets $B_1, \ldots, B_n$.

One should note that the height of the level u enters in $\xi_u(t)$ only via the function $ur(t)$, while ζ and $\kappa(t)$ are the same for all u. This makes it possible to obtain the Palm distributions for the marks at crossings of any level u by introducing just one random variable ζ and one stochastic process $\{\kappa(t)\}$. In the sequel we will use the fact that u enters only through the term $ur(t)$ to derive convergence theorems for $\xi_u(t)$ as $u \to \infty$. These are then translated into distributional convergence under the Palm distribution P^u, by Theorem 10.3.2, and thus, for ergodic processes, to the limiting empirical distributions by Theorem 10.2.1.

As noted in the previous section the same reasoning applies to the limiting empirical distributions of certain other functionals. In particular, this includes maxima, and therefore it is of interest to examine the asymptotic properties of maxima in the Slepian model process.

Some simple facts about the model process $\{\xi_u(t)\}$ should be mentioned here. It may be shown that $\{\kappa(t)\}$ is continuously differentiable, and clearly $E(\kappa(t)) = E(\kappa'(t)) = 0, E(\kappa(0)^2) = E(\kappa'(0)^2) = 0$ so that $P\{\kappa(0) = \kappa'(0) = 0\} = 1$. Since $\lambda_2 = -r''(0)$ one has

$$\xi_u'(0) = ur'(0) - \frac{\zeta r''(0)}{\lambda_2} + \kappa'(0) = \zeta$$

so that ζ is simply the derivative of $\xi_u(t)$ at zero. From Theorem 10.3.2 this immediately translates into a distributional result for the derivative at upcrossings.

Corollary 10.3.3. *The Palm distribution of the derivative $\eta_k'(0)$ of a mark at a u-upcrossing does not depend on u, and it has the Rayleigh density* (10.3.6).

The value of ζ determines the slope of $\xi_u(t)$ at 0. For large t-values the dominant term in $\xi_u(t)$ will be $\kappa(t)$, if $r(t)$ and $r'(t) \to 0$ as $t \to \infty$. (A sufficient condition for this is that the process $\{\xi(t)\}$ has a spectral density, in which case it is also ergodic.) Then $r_\kappa(\tau + s, \tau + t) \to r(s - t)$ as $\tau \to \infty$ so that $\xi_u(t)$ for large t has asymptotically the same covariance structure as the unconditioned process $\xi(t)$, simply reflecting the fact that the influence of the upcrossing vanishes.

With the vertical window conditioning mentioned in Section 10.2, an explicit model representation similar to (10.3.7) can be defined. The only difference is that the derivative ζ of the slope at 0 has a (positive) truncated normal distribution. Note that such a model does not describe the empirical behaviour of $\xi(t)$ near upcrossings of a fixed level u.

10.4. Excursions Above a High Level

We now turn to the asymptotic form of the marks at high-level crossings under the limiting empirical distribution P^u. For this we will simply investigate the model process $\xi_u(t)$ since, as shown in Theorem 10.3.2, its distribution is equal to the P^u-distribution of a mark.

The length and height of an excursion over the high level u will turn out to be of the order u^{-1}, so we normalize $\xi_u(t)$ by expanding the scale by a factor u. Before proceeding to details we give a heuristic argument motivating the precise results to be obtained, by introducing the expansions

$$\begin{aligned} r\left(\frac{t}{u}\right) &= 1 - \frac{\lambda_2 t^2}{2u^2}(1 + o(1)), \\ r'\left(\frac{t}{u}\right) &= -\frac{\lambda_2 t}{u}(1 + o(1)) \end{aligned} \tag{10.4.1}$$

as $t/u \to 0$, which follow from (10.1.1), and by noting that hence $\kappa(t/u = o(t/u)$ as $t/u \to 0$. Inserting this into $\xi_u(t)$ and omitting all o-terms we obtain

$$\xi_u(t/u) \approx u\left(1 - \frac{\lambda_2 t^2}{2u^2}\right) + \frac{\zeta t}{u} = u + \frac{1}{u}\left\{\zeta t - \frac{\lambda_2 t^2}{2}\right\} \tag{10.4.2}$$

as $u \to \infty$ and t is fixed.

The polynomial $\zeta t - \lambda_2 t^2/2$ in (10.4.2) has its maximum for $t = \zeta/\lambda_2$ with a maximum value of $\zeta^2/2\lambda_2$ and therefore we might expect that $\xi_u(t)$ has a maximum of the order $u + (1/u)\,\zeta^2/2\lambda_2$. Hence, the probability that the maximum exceeds $u + v/u$ should be approximately

$$P\left\{\frac{\zeta^2}{2\lambda_2} > v\right\} = \int_{\sqrt{2\lambda_2 v}}^{\infty} \frac{z}{\lambda_2} \exp\left(-\frac{z^2}{2\lambda_2}\right) dz = e^{-v}.$$

The following theorem justifies the approximations made above.

Theorem 10.4.1. *Suppose r satisfies* (10.1.2) *and $r(t) \to 0$ as $t \to \infty$. Then for each $\tau > 0$, $v > 0$*

$$P\left\{\sup_{0 \le t \le \tau} \xi_u(t) > u + \frac{v}{u}\right\} \to e^{-v} \quad \textit{as } u \to \infty,$$

i.e the normalized height of the excursion of $\xi_u(t)$ over u is asymptotically exponential.

Proof. We first prove that the maximum of $\xi_u(t)$ occurs near zero. Choose a function $\delta(u) \to \infty$ as $u \to \infty$ such that $\delta(u)/u \to 0$ and $\delta^2(u)/u \to \infty$. Then

$$P\left\{\sup_{\delta(u)/u \le t \le \tau} \xi_u(t) > u\right\} \to 0, \tag{10.4.3}$$

since the probability is at most

$$P\left\{\sup_{\delta(u)/u \le t \le \tau} (\xi_u(t) - ur(t)) + \sup_{\delta(u)/u \le t \le \tau} ur(t) > u\right\}$$

$$\le P\left\{\sup_{0 \le t \le \tau}\left(-\frac{\zeta r'(t)}{\lambda_2} + \kappa(t)\right) > u\left(1 - \sup_{\delta(u)/u \le t \le \tau} r(t)\right)\right\}.$$

Here

$$\sup_{0 \le t \le \tau}\left(-\frac{\zeta r'(t)}{\lambda_2} + \kappa(t)\right)$$

is a proper (i.e. finite-valued) random variable, and (since $1 - r(t) = \lambda_2 t^2/2 + o(t^2)$ as $t \to 0$, and the joint distribution of $\xi(0)$ and $\xi(t)$ is nonsingular for all t, so that $r(t) < 1$ for $t \neq 0$) $1 - r(t) \geq Kt^2$ for $0 \leq t \leq \tau$, some K depending on τ, so that

$$u\left(1 - \sup_{\delta(u)/u \leq t \leq \tau} r(t)\right) \geq Ku\frac{\delta^2(u)}{u^2} \to \infty$$

which implies (10.4.3).

In view of (10.4.3) we now need only show that

$$P\left\{\sup_{0 \leq t \leq \delta(u)/u} \xi_u(t) > u + \frac{v}{u}\right\} = P\left\{\sup_{0 \leq t \leq \delta(u)} u\left(\xi_u\left(\frac{t}{u}\right) - u\right) > v\right\} \to e^{-v}$$

for $v > 0$. By (10.4.1)

$$\begin{aligned} u\left(\xi_u\left(\frac{t}{u}\right) - u\right) &= -u^2\left(1 - r\left(\frac{t}{u}\right)\right) - \frac{\zeta u r'(t/u)}{\lambda_2} + u\kappa\left(\frac{t}{u}\right) \\ &= -\frac{\lambda_2 t^2}{2}\cdot(1 + o(1)) + \zeta t(1 + o(1)) + t\frac{\kappa(t/u)}{t/u} \end{aligned} \tag{10.4.4}$$

uniformly for $0 \leq t \leq \delta(u)$ as $u \to \infty$. Since $\{\kappa(t)\}$ has a.s. continuously differentiable sample paths with $\kappa(0) = \kappa'(0) = 0$, and $\delta(u)/u \to 0$,

$$\sup_{0 \leq t \leq \delta(u)} \frac{\kappa(t/u)}{t/u} \to 0 \quad \text{(a.s.) as } u \to \infty.$$

This implies that the maximum of $u(\xi_u(t/u) - u)$ is asymptotically determined by the maximum of $-\lambda_2 t^2/2 + \zeta t$ and that

$$\begin{aligned} \lim_{u \to \infty} P\left\{\sup_{0 \leq t \leq \tau} \xi_u(t) > u + \frac{v}{u}\right\} &= P\left\{\sup_t\left(-\frac{\lambda_2 t^2}{2} + \zeta t\right) > v\right\} = P\left\{\frac{\zeta^2}{2\lambda_2} > v\right\} \\ &= e^{-v} \end{aligned}$$

as was to be shown. □

As mentioned above, distributional results and limits for the model process $\{\xi_u(t)\}$ carry over to similar results and limits for marks $\eta_k(t) = \xi(t_k + t)$, i.e. for the ergodic behaviour of the original process $\{\xi(t)\}$ after t_k.

In particular, Theorem 10.4.1 has the corollary that the limiting empirical distribution of the normalized maxima after upcrossings of a level u, for ergodic processes, is approximately exponential for large values of u, i.e.

$$\lim_{T \to \infty} \frac{\#\left\{t_k \in (0, \mathrm{T}];\ \sup_{0 \leq t \leq \tau} \xi(t_k + t) > u + v/u\right\}}{\#\{t_k \in (0, T]\}} \to e^{-v}$$

as $u \to \infty$, a.s. This clarifies the observation at the beginning of this chapter, that an excursion over the high level u exceeds the level $u - (\log p)/u$ with probability $e^{\log p} = p$.

It should be noted here, even if not formally proved, that the excursions emerging from different upcrossings are asymptotically independent. This explains the asymptotic independence of extinctions of crossings with increasing levels.

The following theorem follows from (10.4.4).

Theorem 10.4.2. *Suppose r satisfies* (10.1.2) *and* $r(t) \to 0$ *as* $t \to \infty$. *Then with probability one the normalized model process* $\tilde{\xi}_u(t) = u(\xi_u(t/u) - u)$ *tends uniformly for* $|t| \le \tau$ *to a parabola*

$$\tilde{\xi}_\infty(t) = -\frac{\lambda_2 t^2}{2} + \zeta t$$

in the sense that, with probability one,

$$\sup_{|t| \le \tau} |\tilde{\xi}_u(t) - \tilde{\xi}_\infty(t)| \to 0$$

as $u \to \infty$.

This theorem throws some light upon the discrete approximation used in the proof of the maximum and Poisson theorems in previous chapters. The choice of spacing in the discrete grid, q, appeared to be chosen there for purely technical reasons. Theorem 10.4.2 shows why it works. By the theorem the natural time scale for excursions over the high level u is u^{-1}, so the spacing $q = o(u^{-1})$ of the q-grid catches high maxima with an increasing number of grid points.

CHAPTER 11

Maxima and Minima and Extremal Theory for Dependent Processes

Trivially, extremes in two or more mutually independent processes are independent. In this chapter we shall establish the perhaps somewhat surprising fact that, asymptotically, independence of extremes holds for normal processes even when they are highly correlated. However, we shall first consider the asymptotic independence of maxima and minima in one normal process. Since minima of $\xi(t)$ are maxima for $-\xi(t)$, this can in fact be regarded as a special case of independence between extremes in two processes, namely between the maxima in the completely dependent processes $\xi(t)$ and $-\xi(t)$.

11.1. Maxima and Minima

For a standardized stationary normal process, $\{\xi(t)\}$ and $\{-\xi(t)\}$ have the same distribution. Writing

$$m(T) = \inf\{\xi(s); 0 \leq s \leq T\},$$

clearly $m(T) = -\sup\{-\xi(s);\ 0 \leq s \leq T\}$, and hence $m(T)$ has the same asymptotic behaviours as $-M(T)$. If $\{\xi(t)\}$ satisfies the hypotheses of Theorem 8.2.5, then $P\{m(T) \geq -v\} \to e^{-\eta}$ as $T, v \to \infty$ and $T\nu \to \eta > 0$ for $\nu = (1/2\pi)\lambda_2^{1/2} \exp(-v^2/2)$. It follows, under the hypotheses of Theorem 8.2.7, that

$$P\{a_T(m(T) + b_T) \leq x\} \to 1 - \exp(-e^x)$$

with the same normalizations as for maxima, i.e.

$$a_T = (2 \log T)^{1/2},$$

$$b_T = (2 \log T)^{1/2} + \frac{\log(\lambda_2^{1/2}/2\pi)}{(2 \log T)^{1/2}}.$$

It was shown by Berman (1971b) that, without any additional assumptions, the minimum $m(T)$ and the maximum $M(T)$ are asymptotically independent in analogy with the asymptotic independence of minima and maxima of independent sequences, established in Theorem 1.8.3. We shall now obtain Berman's result (Theorem 11.1.5).

With the same notation and technique as in Chapter 8, we let N_u and $N_u^{(q)}$ be the number of u-upcrossings by the process $\{\xi(t); 0 \leq t \leq h\}$ and the sequence $\{\xi(jq); 0 \leq jq \leq h\}$, and define similarly D_{-v} and $D_{-v}^{(q)}$ to be the number of *downcrossings* of the level $-v$. We first observe that if, for $h = 1$,

$$\mu = E(N_u) = \frac{1}{2\pi} \lambda_2^{1/2} \exp\left(-\frac{u^2}{2}\right),$$

$$\nu = E(D_{-v}) = \frac{1}{2\pi} \lambda_2^{1/2} \exp\left(-\frac{v^2}{2}\right),$$

and if $u, v \to \infty$ so that $T\mu \to \tau > 0$ and $T\nu \to \eta > 0$, then we have $u \sim v$ and

$$u - v \sim \frac{\log(\eta/\tau)}{u}, \tag{11.1.1}$$

cf. (9.2.3). In particular, this implies that if $q \to 0$ so that $uq \to 0$ then also $vq \to 0$.

The following lemma contains the necessary discrete approximation and separation of maxima. As in Chapter 8, we split the increasing interval $(0, T]$ into $n = [T/h]$ pieces, each divided into two, I_k and I_k^*, of length $h - \varepsilon$ and ε, respectively. Write $m(I)$ for the minimum of $\xi(t)$ over the interval I.

Lemma 11.1.1. *If* (8.1.1) *holds then, as $u, v \to \infty$ and $uq, vq \to 0$,*

(i) $P\{m(I) < -v, M(I) > u\}$

$$= P\left\{\min_{jq \in I} \xi(jq) < -v, \max_{jq \in I} \xi(jq) > u\right\} + o(\nu + \mu),$$

where $o(v + \mu)$ is uniform in all intervals I of length $h \le h_0$, for fixed $h_0 > 0$,

(ii)
$$\limsup_{T\to\infty} \left| P\left\{ -v \le m\left(\bigcup_{k=1}^{n} I_k\right) \le M\left(\bigcup_{k=1}^{n} I_k\right) \le u \right\} - P\{-v \le m(nh) \le M(nh) \le u\} \right| \le \frac{\eta + \tau}{h}\varepsilon,$$

(iii)
$$P\left\{ -v \le \xi(jq) \le u, jq \in \bigcup_{k=1}^{n} I_k \right\} - P\left\{ -v \le m\left(\bigcup_{k=1}^{n} I_k\right) \le M\left(\bigcup_{k=1}^{n} I_k\right) \le u \right\} \to 0.$$

Proof. (i) By Lemma 8.2.2(i) applied to $\{\xi(t)\}$ and $\{-\xi(t)\}$,

$$\begin{aligned} 0 &\le P\{m(I) < -v, M(I) > u\} - P\left\{\min_{jq\in I} \xi(jq) < -v, \max_{jq\in I} \xi(jq) > u\right\} \\ &\le E(D_{-v} - D_{-v}^{(q)}) + E(N_u - N_u^{(q)}) + P\{\xi(0) < -v\} + P\{\xi(0) > u\} \\ &= o(v + \mu). \end{aligned}$$

Parts (ii) and (iii) follow as in Lemma 8.2.3; we shall not repeat the details here. □

The following two lemmas give the asymptotic independence of both maxima and minima over the separate I_k-intervals.

Lemma 11.1.2. *Let $\xi_1, \ldots, \xi_n$ be standard normal variables with covariance matrix $\Lambda^1 = (\Lambda_{ij}^1)$, and $\eta_1, \ldots, \eta_n$ similarly with covariance matrix $\Lambda^0 = (\Lambda_{ij}^0)$, and let $\rho_{ij} = \max(|\Lambda_{ij}^1|, |\Lambda_{ij}^0|)$. Further, let $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ be vectors of real numbers and write*

$$w = \min(|u_1|, \ldots, |u_n|, |v_1|, \ldots, |v_n|).$$

Then

$$\begin{aligned} P\{-v_j < \xi_j \le u_j \text{ for } j = 1, \ldots, n\} &- P\{-v_j < \eta_j \le u_j \text{ for } j = 1, \ldots, n\} \\ &\le \frac{4}{2\pi} \sum_{1\le i<j\le n} |\Lambda_{ij}^1 - \Lambda_{ij}^0|(1 - \rho_{ij}^2)^{-1/2} \exp\left(-\frac{w^2}{1 + \rho_{ij}}\right). \qquad (11.1.2) \end{aligned}$$

Proof. This requires only a minor variation of the proof of Theorem 4.2.1. It follows as in that proof (and using the same notation) that the left-hand side of (11.1.2) equals

$$\int_0^1 \left\{ \sum_{i<j} (\Lambda_{ij}^1 - \Lambda_{ij}^0) \int_{-\mathbf{v}}^{\mathbf{u}} \cdots \int \frac{\partial^2 f_h}{\partial y_i \, \partial y_j} \, d\mathbf{y} \right\} dh.$$

By performing the integrations on y_i and y_j we obtain four terms, containing the integrands

$$f_h(y_i = u_i, y_j = u_j),\ -f_h(y_i = u_i, y_j = -v_j),\ -f_h(y_i = -v_i, y_j = u_j),$$

and

$$f_h(y_i = -v_i, y_j = -v_j),$$

respectively. Each of these terms can be estimated as in the proof of Theorem 4.2.1 and the lemma follows. □

Lemma 11.1.3. *Suppose $r(t) \to 0$ as $t \to \infty$, and that $T\mu \to \tau$, $Tv \to \eta$ as $T \to \infty$, and that, if $qu, qv \to 0$ sufficiently slowly*

$$\frac{T}{q} \sum_{\varepsilon \le kq \le T} |r(kq)| \exp\left(-\frac{w^2}{1 + |r(kq)|}\right) \to 0 \tag{11.1.3}$$

for every $\varepsilon > 0$, where $w = \min(u, v) > 0$. (This holds in particular if $r(t) \log t \to 0$ as $t \to \infty$.) Then

(i) $P\left\{-v \le \xi(kq) \le u, kq \in \bigcup_{j=1}^{n} I_j\right\} - \prod_{j=1}^{n} P\{-v \le \xi(kq) \le u, kq \in I_j\} \to 0$ *as $T \to \infty$,*

(ii) $\limsup_{T\to\infty} \left| \prod_{j=1}^{n} P\{-v \le \xi(kq) \le u, kq \in I_j\} - P^n\{-v \le m(h) \le M(h) \le u\}\right| \le \frac{(\eta + \tau)}{h}\varepsilon.$

Proof. Part (i) follows from Lemma 11.1.2 in the same way as Lemma 8.2.4(i) follows from Theorem 4.2.1. As for part (ii), note that

$$\begin{aligned} 0 &\le P\{-v < m(I_j) \le M(I_j) \le u\} - P\{-v < m(h) \le M(h) \le u\} \\ &\le P\{m(I_j^*) \le -v\} + P\{M(I_j^*) > u\} \le (v + \mu)(\varepsilon + o(1)) \end{aligned}$$

for T sufficiently large, by Lemma 8.2.1(i). It now follows from Lemma 11.1.1(i) and Lemma 8.2.2(ii) that for T sufficiently large,

$$0 \le P_j - P \le (v + \mu)(\varepsilon + o(1))$$

where $P_j = P\{-v < \xi(kq) \le u,\ kq \in I_j\}$, $P = P\{-v < m(h) \le M(h) \le u\}$. Hence

$$0 \le \prod_{j=1}^{n} P_j - P^n \le n(v + \mu)(\varepsilon + o(1)) \sim \frac{(\eta + \tau)}{h}(\varepsilon + o(1))$$

as in Lemma 8.2.4(ii). □

The final essential step in the proof of asymptotic independence of $m(T)$ and $M(T)$ is to show that an interval of fixed length h does not contain both large positive and large negative values of $\xi(t)$.

Lemma 11.1.4. *If* (8.1.1) *holds, there is an* $h_0 > 0$ *such that as* $u, v \to \infty$,

$$P\{M(h) > u, m(h) < -v\} = o(v + \mu)$$

for all $h < h_0$.

PROOF. By Lemma 11.1.1(i) it is enough to prove that

$$P\left\{\max_{0 \le jq \le h} \xi(jq) > u, \min_{0 \le jq \le h} \xi(jq) < -v\right\} = o(\mu + v)$$

for some q satisfying $uq, vq \to 0$. By stationarity, this probability is bounded by

$$\frac{h}{q} P\left\{\xi(0) > u, \min_{-h \le jq \le h} \xi(jq) < -v\right\}$$

$$\le \frac{h}{q} \sum_{-h \le jq \le h} P\{\xi(0) > u, \xi(jq) < -v\}. \tag{11.1.4}$$

Here, for $jq \neq 0$,

$$P\{\xi(0) > u, \xi(jq) < -v\} = \int_u^\infty \phi(x) P\{\xi(jq) < -v \mid \xi(0) = x\}\, dx$$

$$= \int_u^\infty \phi(x)\Phi\left(\frac{-v - xr(jq)}{\sqrt{1 - r^2(jq)}}\right) dx, \tag{11.1.5}$$

since, conditional on $\xi(0) = x$, $\xi(jq)$ is normal with mean $xr(jq)$ and variance $1 - r^2(jq)$. Now, choose $h_0 > 0$ such that $0 < r(t) < 1$ for $0 < |t| \le h_0$, which is possible by (8.1.1). If $0 < |jq| \le h < h_0$, $x \ge 0$, then

$$(-v - xr(jq))/\sqrt{1 - r^2(jq)} \le -v,$$

for $u, v > 0$, so that (11.1.5) is bounded by

$$\int_u^\infty \phi(x)\Phi(-v)\, dx = (1 - \Phi(u))(1 - \Phi(v)) \sim \frac{\phi(u)\phi(v)}{uv}.$$

Together with (11.1.4) and (11.1.5) this shows that

$$0 \le P\left\{\max_{0 \le jq \le h} \xi(jq) > u, \min_{0 \le jq \le h} \xi(jq) < -v\right\} \le \frac{h^2}{q^2} \cdot \frac{\phi(u)\phi(v)}{uv}$$

$$= h^2 \cdot \frac{\phi(u)\phi(v)}{(\mu + v)} \cdot \frac{1}{quqv} (\mu + v) = o(1)(\mu + v),$$

since $\phi(u)\phi(v)/(\mu + v) = O(\mu v/(\mu + v)) \to 0$, and q can be chosen to make $quqv \to 0$ arbitrarily slowly. $\square$

Theorem 11.1.5. *Let $u = u_T \to \infty$ and $v = v_T \to \infty$ as $T \to \infty$, in such a way that*

$$T\mu = \frac{T}{2\pi}\lambda_2^{1/2}\exp\left(-\frac{u^2}{2}\right) \to \tau \geq 0, \qquad Tv = \frac{T}{2\pi}\lambda_2^{1/2}\exp\left(-\frac{v^2}{2}\right) \to \eta \geq 0.$$

Suppose the stationary normal process $\xi(t)$ satisfies (8.1.1) *and either* (8.1.2) *or the weaker condition* (11.1.3). *Then*

$$P\{-v < m(T) \leq M(T) \leq u\} \to e^{-\eta-\tau} \quad \text{as } T \to \infty,$$

and hence

$$P\{a_T(m(T) + b_T) \leq x,\, a_T(M(T) - b_T) \leq y\} \to (1 - \exp(-e^x))\exp(-e^{-y})$$

with $a_T = (2\log T)^{1/2}$, $b_T = (2\log T)^{1/2} + \log(\lambda_2^{1/2}/2\pi)/(2\log T)^{1/2}$. Thus the normalized minimum and maximum are asymptotically independent.

PROOF. By Lemma 11.1.1(ii) and (iii), and Lemma 11.1.3(i) and (ii) we have

$$\limsup_{T\to\infty} |P\{-v < m(nh) \leq M(nh) \leq u\} - P^n\{-v < m(h) \leq M(h) \leq u\}| \leq \frac{2(\eta + \tau)}{h}\varepsilon,$$

for arbitrary $\varepsilon > 0$, and hence

$$P\{-v < m(nh) \leq M(nh) \leq u\} - P^n\{-v < m(h) \leq M(h) \leq u\} \to 0 \quad (11.1.6)$$

as $n \to \infty$. Furthermore, by Lemma 11.1.4

$$P^n\{-v < m(h) \leq M(h) \leq u\} = (1 - P\{m(h) \leq -v\} - P\{M(h) > u\} + o(v + \mu))^n.$$

Arguing as in the proof of Theorem 8.2.5, the result now follows from Lemma 8.2.1, using the fact that $n(v + \mu) \sim (T/h)(v + \mu) \to (\eta + \tau)/h$. □

This theorem, of course, has ramifications similar to those for maxima from Theorem 8.2.5, but before discussing them we give a simple corollary about the absolute maximum of $\xi(t)$.

Corollary 11.1.6. *If $u \to \infty$, $T\mu \to \tau$, then*

$$P\left\{\sup_{0\leq t\leq T} |\xi(t)| \leq u\right\} \to e^{-2\tau},$$

and furthermore

$$P\left\{a_T\left(\sup_{0\leq t\leq T} |\xi(t)| - b_T\right) \leq x + \log 2\right\} \to \exp(-e^{-x}).$$

As for the maximum alone, it is now easy to prove asymptotic independence of maxima and minima over several disjoint intervals with lengths proportional to T. As a consequence one has a Poisson convergence theorem for the two point processes of upcrossings of u and downcrossings of $-v$, the limiting Poisson processes being independent. Furthermore, the point process of downcrossings of several low levels converges to a point process with Poisson components obtained by successive binomial thinning, as in Theorem 9.3.2, and these downcrossings processes are asymptotically independent of the upcrossings processes. Of course, it then follows that the entire point process of local maxima, considered in Theorem 9.5.2, is also asymptotically independent of the point process of normalized local minima.

11.2. Extreme Values and Crossings for Dependent Processes

One remarkable feature of dependent *normal* processes is that, regardless of how high the correlation—short of perfect correlation—the number of high-level crossings in the different processes are asymptotically independent, as shown in Lindgren (1974). This will now be proved, again by means of the Normal Comparison Lemma.

Let $\{\xi_1(t)\}, \ldots, \{\xi_p(t)\}$ be jointly normal processes with zero means, variances one and covariance functions $r_k(\tau) = \operatorname{Cov}(\xi_k(t), \xi_k(t + \tau))$. We shall assume that they are jointly stationary, i.e. $\operatorname{Cov}(\xi_k(t), \xi_l(t + \tau))$ does not depend on t, and we write

$$r_{kl}(\tau) = \operatorname{Cov}(\xi_k(t), \xi_l(t + \tau))$$

for the *cross-covariance* function. Suppose further that each r_k satisfies (8.1.1), possibly with different λ_2's, i.e.

$$r_k(t) = 1 - \frac{\lambda_{2,k} t^2}{2} + o(t^2), \qquad t \to 0, \tag{11.2.1}$$

and that

$$\begin{gathered} r_k(t) \log t \to 0, \\ r_{kl}(t) \log t \to 0 \quad \text{as } t \to \infty, \end{gathered} \tag{11.2.2}$$

for $1 \leq k$, $l \leq p$. To exclude the possibility that $\xi_k(t) \equiv \pm \xi_l(t + t_0)$ for some $k \neq l$, and some choice of t_0 and $+$ or $-$, we assume that

$$\max_{k \neq l} \sup_t |r_{kl}(t)| < 1. \tag{11.2.3}$$

However, we note here that if $\inf_t r_{kl}(t) = -1$ for some $k \neq l$, there is a t_0 such that $r_{kl}(t_0) = -1$, which means that $\xi_k(t) \equiv -\xi_l(t + t_0)$. A maximum in $\xi_l(t)$ is therefore a minimum in $\xi_k(t)$, and as was shown in the first section of this chapter, maxima and minima are asymptotically independent. In fact, with some increase in the complexity of proof, condition (11.2.3) can be relaxed to $\max_{k \neq l} \sup_t r_{kl}(t) < 1$.

Define

$$M_k(T) = \sup\{\xi_k(t); 0 \le t \le T\}, \qquad k = 1, \dots, p,$$

and let $u_k = u_k(T)$ be levels such that

$$T\mu_k = \frac{T}{2\pi} \lambda_{2,k}^{1/2} \exp\left(-\frac{u_k^2}{2}\right) \to \tau_k \ge 0$$

as $T \to \infty$. Write $u = \min\{u_1, \dots, u_p\}$.

To prove asymptotic independence of the $M_k(T)$ we approximate by the maxima over separated intervals I_j, $j = 1, \dots, n$, with $n = [T/h]$ for h fixed, and then replace the continuous maxima by the maxima of the sampled processes to obtain asymptotic independence of maxima over different intervals. We will only briefly point out the changes which have to be made in previous arguments. The main new argument to be used here concerns the maxima of $\xi_k(t)$, $k = 1, \dots, p$, over one fixed interval, I, say.

We first state the asymptotic independence of maxima over disjoint intervals.

Lemma 11.2.1. *If* r_k, r_{kl} *satisfy* (11.2.1)–(11.2.3) *for* $1 \le k, l \le p$, *and if* $T\mu_k \to \tau_k \ge 0$, *then for* $h > 0$ *and* $n = [T/h]$,

$$P\{M_k(nh) \le u_k, k = 1, \dots, p\} - P^n\{M_k(h) \le u_k, k = 1, \dots, p\} \to 0.$$

PROOF. This corresponds to (11.1.6) in the proof of Theorem 11.1.5, and is proved by similar means. It is only the relation

$$P\left\{\xi_k(jq) \le u_k, jq \in \bigcup_{r=1}^n I_r, k = 1, \dots, p\right\} - \prod_{r=1}^n P\{\xi_k(jq) \le u_k, jq \in I_r, k = 1, \dots, p\} \to 0, \quad (11.2.4)$$

corresponding to Lemma 11.1.3(i), that has to be given a different proof.

Identifying $\xi_1, \dots, \xi_n$ in Corollary 4.2.2 with $\xi_1(jq), \dots, \xi_p(jq), jq \in \bigcup_{r=1}^n I_r$, and $\eta_1, \dots, \eta_n$, analogously, but with variables from different I_r-intervals independent, (4.2.6) gives, since $\sup |r_{kl}(t)| \le 1$, for $k \neq l$ and $\max_k \sup_{t \ge \varepsilon} r_k(t) < 1$,

$$\left| P\left\{\xi_k(jq) \le u_k, jq \in \bigcup_{r=1}^n I_r, k = 1, \dots, p\right\} - \prod_{r=1}^n P\{\xi_k(jq) \le u_k, jq \in I_r, k = 1, \dots, p\} \right|$$
$$\le K \sum_{k=1}^p {\sum_{i<j}}^* |r_k((i-j)q)| \exp\left\{-\frac{u^2}{1 + |r_k((i-j)q)|}\right\}$$
$$+ K \sum_{1 \le k \neq l \le p} {\sum_{i<j}}^* |r_{kl}((i-j)q)| \exp\left\{-\frac{u^2}{1 + |r_{kl}((i-j)q)|}\right\} \quad (11.2.5)$$

where $\sum^*$ as before indicates that the sum is taken over i, j such that iq and jq belong to different I_r. Since both $r_k(t)\log t \to 0$, $r_{kl}(t)\log t \to 0$ and $\sup|r_{kl}(t)| < 1$ we can, as in Lemma 8.1.1(ii), conclude that both sums in (11.2.5) tend to zero. $\square$

Lemma 11.2.2. *If* r_k, r_{kl} *satisfy* (11.2.1) *and* (11.2.3), *for* $1 \le k, l \le p$,

(i) $$P\{M_k(h) > u_k, M_l(h) > u_l\} = o(\mu_k + \mu_l) \quad \text{for } k \neq l,$$

and

(ii) $$P\{M_k(h) \le u_k, k = 1, \ldots, p\} = 1 - \sum_{k=1}^{p} P\{M_k(h) > u_k\} + o\left(\sum_{k=1}^{p} \mu_k\right).$$

Proof. (i) As in the proof of Lemma 11.1.4 it is enough to prove that, if $q \to 0$ so that $u_k q \sim u_l q \to 0$ sufficiently slowly,

$$P\left\{\max_{0 \le jq \le h} \xi_k(jq) > u_k, \max_{0 \le jq \le h} \xi_l(jq) > u_l\right\} = o(\mu_k + \mu_l).$$

Since, for $r = k, l$,

$$P\left\{\max_{0 \le jq \le h} \xi_r(jq) > u_r\right\} = O(\mu_r),$$

and therefore

$$\begin{aligned} P\left\{\max_{0 \le jq \le h} \xi_k(jq) > u_k\right\} P\left\{\max_{0 \le jq \le h} \xi_l(jq) > u_l\right\} &= O(\mu_k \mu_l) \\ &= o(\mu_k + \mu_l), \end{aligned}$$

it clearly suffices to prove that

$$\begin{aligned} P\left\{\max_{0 \le jq \le h} \xi_k(jq) > u_k, \max_{0 \le jq \le h} \xi_l(jq) > u_l\right\} \\ - P\left\{\max_{0 \le jq \le h} \xi_k(jq) > u_k\right\} P\left\{\max_{0 \le jq \le h} \xi_l(jq) > u_l\right\} \\ = o(\mu_k + \mu_l). \end{aligned} \tag{11.2.6}$$

To estimate the difference we again use Corollary 4.2.2 with Λ^1 defined by r_k, r_l, and r_{kl}, and Λ^0 obtained by taking r_{kl} identically zero for $k \neq l$. Elementary calculations show that the difference in (11.2.6) equals

$$P\left\{\max_{0 \le jq \le h} \xi_k(jq) \le u_k, \max_{0 \le jq \le h} \xi_l(jq) \le u_l\right\} - P\left\{\max_{0 \le jq \le h} \xi_k(jq) \le u_k\right\} P\left\{\max_{0 \le jq \le h} \xi_l(jq) \le u_l\right\},$$

which by Corollary 4.2.2 is bounded in modulus by

$$K \sum_{0 \le iq, jq \le h} |r_{kl}((i-j)q)|(1 - r_{kl}^2((i-j)q))^{-1/2} \cdot \exp\left\{-\frac{u^2}{1 + |r_{kl}((i-j)q)|}\right\} \tag{11.2.7}$$

with $u = \min(u_k, u_l)$.

Now, by (11.2.3), $\sup|r_{kl}(t)| = 1 - \delta$ for some $\delta > 0$, and using this, we can bound (11.2.7) by

$$\begin{aligned} Kh^2q^{-2}\exp\left(-\frac{u^2}{1+1-\delta}\right) &= Kh^2\frac{\phi(u)}{\mu_k+\mu_l}(uq)^{-2}u^2\exp\left\{-u^2\cdot\frac{\delta}{2(2-\delta)}\right\} \\ &\quad \times(\mu_k+\mu_l) = o(\mu_k+\mu_l) \end{aligned}$$

if $uq \to 0$ sufficiently slowly, since $\phi(u)/(\mu_k + \mu_l)$ is bounded.

(ii) This follows immediately from part (i) and the inequality

$$\begin{aligned} &\sum_{k=1}^{p} P\{M_k(h) > u_k\} - \sum_{1 \le k < l \le p} P\{M_k(h) > u_k, M_l(h) > u_l\} \\ &\qquad \le P\left(\bigcup_{k=1}^{p}\{M_k(h) > u_k\}\right) \le \sum_{k=1}^{p} P\{M_k(h) > u_k\}. \qquad \square \end{aligned}$$

Reasoning as in the proof of Theorem 8.2.5, and using Lemma 11.2.1 and Lemma 11.2.2(ii) we get the following result.

Theorem 11.2.3. *Let $u_k = u_k(T) \to \infty$ as $T \to \infty$, so that*

$$T\mu_k = \frac{T}{2\pi}\lambda_{2,k}^{1/2}\exp(-u_k^2/2) \to \tau_k > 0, \quad 1 \le k \le p,$$

and suppose that the jointly stationary normal processes $\xi_k(t)$ satisfy (11.2.1)–(11.2.3). *Then*

$$P\{M_k(T) \le u_k, k = 1, \ldots, p\} \to \exp\left(-\sum_{k=1}^{p}\tau_k\right)$$

as $T \to \infty$.

Under the same conditions as in Theorem 11.2.3, the time-normalized point processes of upcrossings of one or several levels tend jointly in distribution to p independent, binomially thinned, Poisson processes. Similarly, under the conditions of Theorem 9.5.2, the point processes of normalized local maxima converge to p independent Poisson processes in the plane. We shall state the latter result as a theorem, leaving its proof, and the former result, to the reader.

For $k = 1, \ldots, p$, suppose that $\xi_k(t)$ has local maxima at the points $s_0^{(k)} < 0 \le s_1^{(k)} < s_2^{(k)} < \cdots$, and let $N'_{k,T}$ be the point process $(s_i^{(k)}/T, a_T(\xi_k(s_i^{(k)}) - b_{T,k})$ of normalized local maxima.

Theorem 11.2.4. *Suppose the standardized stationary normal processes* $\{\xi_1(t)\}, \ldots, \{\xi_p(t)\}$ *have continuously differentiable sample paths, and are twice differentiable in quadratic mean, and suppose that the covariance and crosscovariance functions* $r_k(t)$ *and* $r_{kl}(t)$ *satisfy* (11.2.1)–(11.2.3). *Then the point processes* $N'_{1,T}, \ldots, N'_{p,T}$ *of normalized maxima are asymptotically independent and each converges to a Poisson process on* $(0, \infty) \times R$ *with intensity measure* $dt \times e^{-x}\,dx$ *as* $T \to \infty$, *as in Theorem* 9.5.2.

We end this chapter with an example which illustrates the extraordinary character of extremes in normal processes.

Let $\{\zeta(t)\}$ and $\{\eta(t)\}$ be independent standardized normal processes whose covariance functions r_ζ and r_η satisfy (8.1.1) and (8.1.2), let c_k, $k = 1, \ldots, p$, satisfying $|c_k| < 1$, and $c_k \neq c_l$, $k \neq l$, be constants, and define

$$\xi_k(t) = c_k \zeta(t) + (1 - c_k^2)^{1/2} \eta(t).$$

Then the processes $\xi_k(t)$, $k = 1, \ldots, p$ are jointly normal and their covariance functions $r_k(t)$ and crosscovariance functions

$$r_{kl}(t) = c_k c_l r_\zeta(t) + (1 - c_k^2)^{1/2}(1 - c_l^2)^{1/2} r_\eta(t)$$

satisfy (11.2.1)–(11.2.3). Thus, even though $\xi_1(t), \ldots, \xi_p(t)$ are linearly dependent, their maxima are asymptotically independent.

We can illustrate this geometrically by representing $(\zeta(t), \eta(t))$ by a point moving randomly in the plane. The upcrossings of a level u_k by $\xi_k(t)$ then correspond to outcrossings of the line

$$c_k x + (1 - c_k^2)^{1/2} y = u_k$$

by $(\zeta(t), \eta(t))$, as illustrated in Figure 11.2.1. The times of these outcrossings form asymptotically independent Poisson processes when suitably normalized if $T\mu_k \to \tau_k > 0$.

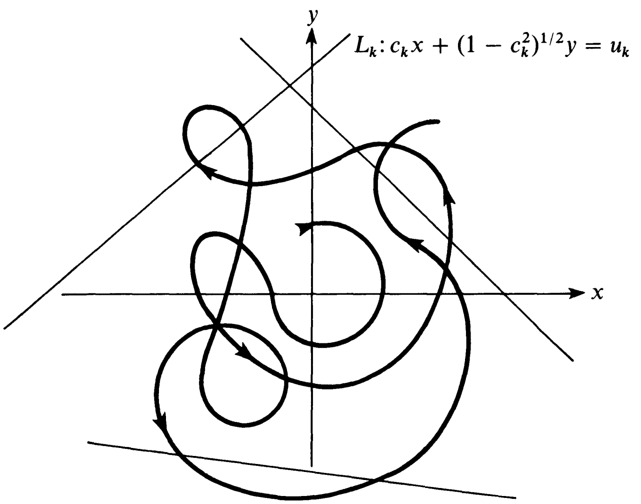


Figure 11.2.1. Outcrossings of straight lines L_k by a bivariate normal process $(\xi_1(t), \xi_2(t))$.

CHAPTER 12

Maxima and Crossings of Nondifferentiable Normal Processes

The basic assumption of the previous chapters has been that the covariance function $r(\tau)$ of the stationary normal process $\xi(t)$ has an expansion $r(\tau) = 1 - \lambda_2\tau^2/2 + o(\tau^2)$ as $\tau \to 0$. In this chapter we shall consider the more general class of covariances which have the expansion $r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha)$ as $\tau \to 0$, where the positive constant α may be less than 2. This includes covariances of the form $\exp(-|\tau|^\alpha)$, the case $\alpha = 1$ being that of the Ornstein–Uhlenbeck process. Since the mean number of upcrossings of any level per unit time is infinite when $\alpha < 2$, the methods of Chapter 8 do not apply in such cases. However, it will be shown by different methods that the double exponential limiting law for the maximum still applies with appropriately defined normalizing constants, if (8.1.2) (or a slightly weaker version) holds. This, of course, also provides an alternative derivation of the results of Chapter 8 when $\alpha = 2$. Finally, while clearly no Poisson result is possible for upcrossings when $\alpha < 2$, it will be seen that Poisson limits may be obtained for the related concept of ε-upcrossings, defined similarly to the ε-maxima of Chapter 9.

12.1. Introduction and Overview of the Main Result

Throughout the chapter it will be assumed that $\{\xi(t)\}$ is a (zero mean and unit variance) stationary normal process with covariance function $r(\tau)$ satisfying

$$r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha) \quad \text{as } \tau \to 0, \tag{12.1.1}$$

where α is a constant, $0 < \alpha \leq 2$, and C is a positive constant. As noted in Chapter 7, this assumption is, in particular, sufficient to guarantee continuity of the sample paths of the process, and thus ensure that the maximum

$M(T) = \sup\{\xi(t); 0 \le t \le T\}$ is well defined and finite for each T. Our main result (Theorem 12.3.5) is that, under (12.1.1) and the now familiar decay condition $r(t) \log t \to 0$, $M(T)$ has the Type I limiting distribution, viz.

$$P\{a_T(M(T) - b_T) \le x\} \to \exp(-e^{-x}).$$

Here the constant a_T is the same as in Chapter 8, but the constant b_T depends on α:

$$a_T = (2 \log T)^{1/2},$$

$$b_T = (2 \log T)^{1/2} + \frac{1}{(2 \log T)^{1/2}} \times \left\{\frac{2-\alpha}{2\alpha} \log \log T + \log(C^{1/\alpha} H_\alpha (2\pi)^{-1/2} 2^{(2-\alpha)/2\alpha})\right\},$$

where H_α is a certain strictly positive constant, ($H_1 = 1$, $H_2 = \pi^{-1/2}$).

This remarkable result was first obtained by Pickands (1969a, b), although his proofs were not quite complete. Complements and extensions have been given by Berman (1971c), Qualls and Watanabe (1972), and Lindgren *et al.* (1975). While we shall not follow the method of Pickands it does have some particularly interesting features in that it uses a generalized notion of upcrossings which makes it possible to obtain a Poisson type result also for $\alpha < 2$. Briefly, given an $\varepsilon > 0$, the function $f(t)$ is said to have an ε-upcrossing of the level u at t_0 if $f(t) \le u$ for all $t \in (t_0 - \varepsilon, t_0)$, and, for all $\eta > 0$, $f(t) > u$ for some $t \in (t_0, t_0 + \eta)$. Clearly, this is equivalent to requiring that it has a (nonstrict or strict) upcrossing there, and furthermore $f(t) \le u$ for all t in $(t_0 - \varepsilon, t_0)$. An ε-upcrossing is always an upcrossing, while obviously an upcrossing need not be an ε-upcrossing. Clearly the number of ε-upcrossings in, say, a unit interval is bounded (by $1/\varepsilon$) and hence certainly has a finite mean. Even if this mean cannot be calculated as easily as the mean number of ordinary upcrossings, its limiting form for large u has a simple relation to the extremal results for $M(T)$. In particular, as we shall see, it does not depend on the ε chosen. As noted we shall not use ε-upcrossings in our main result, but will show (in Section 12.4) how Poisson results may be obtained for them.

The main complication in the derivation of the main result, as compared with the case $\alpha = 2$, concerns the tail distribution of $M(h)$ for h fixed, which cannot be approximated by the tail distribution of the simple cosine process if $\alpha < 2$. Our proof is organized in a number of parts, and it may be useful to get a "bird's-eye view" from the following summary of the main steps.

1. Find the tail of the distribution of $\max\{\xi(jq); j = 0, \ldots, n-1\}$ for a *fixed n*:

$$P\left\{\max_{0 \le j < n} \xi(jq) > u\right\} \sim \frac{\phi(u)}{u} C^{1/\alpha} H_\alpha(n, a)$$

as $u \to \infty$, $q \to 0$, $q \sim a u^{-2/\alpha}$, $a > 0$ and n fixed; $H_\alpha(n, a)$ a constant (Lemma 12.2.3).

2. Find the tail of the distribution of $\max\{\xi(jq); 0 \le jq \le h\}$ for a fixed $h > 0$:

$$P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\} \sim h \frac{u^{2/\alpha}\phi(u)}{u} C^{1/\alpha} H_\alpha(a),$$

where

$$\frac{H_\alpha(n, a)}{na} \to H_\alpha(a)$$

as $n \to \infty$ (Lemma 12.2.4).

3. Approximate $\sup_{0 \le t \le h} \xi(t)$ by $\max_{0 \le jq \le h} \xi(jq)$ for fixed h:

$$\limsup_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0 \le jq \le h} \xi(jq) < u - \frac{a^\beta}{u}, \sup_{0 \le t \le h} \xi(t) > u\right\} \to 0$$

as $a \to 0$ (Lemma 12.2.5)

and

$$\limsup_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{u - \frac{a^\beta}{u} < \max_{0 \le jq \le h} \xi(jq) \le u\right\} \to 0 \quad \text{as } a \to 0$$

(Lemma 12.2.6).

4. Find the tail distribution of $\sup_{0 \le t \le h} \xi(t)$ for a fixed h:

$$P\left\{\sup_{0 \le t \le h} \xi(t) > u\right\} \sim h \frac{u^{2/\alpha}\phi(u)}{u} C^{1/\alpha} H_\alpha$$

(Theorem 12.2.9), where

$$H_\alpha = \lim_{a \to 0} H_\alpha(a) > 0$$

(Lemmas 12.2.7 and 12.2.8).

5. Once the tail distribution and its discrete approximation are obtained, continue as in Chapter 8 to prove asymptotic independence of maxima in disjoint intervals under suitable covariance conditions, e.g. $r(t) \log t \to 0$.

12.2. Maxima Over Finite Intervals

We start the derivation with a general result, of some interest in its own right, giving an estimate of the probability of a large deviation. This is a special case of a result announced by Fernique (1964), a proof having been given by Marcus (1970).

Lemma 12.2.1. *If* $\{\xi(t); 0 \le t \le 1\}$ *is a normal process with mean zero,* $\operatorname{Var}(\xi(0)) = \sigma^2 \ge 0$, *such that*

$$E((\xi(s) - \xi(t))^2) \le C|t - s|^\alpha \tag{12.2.1}$$

for some α, $0 < \alpha \le 2$, *then there exists a constant* $c_\alpha > 0$, *only depending on* α, *such that for all* x,

$$P\left\{\sup_{0 \le t \le 1} \xi(t) > x\right\} \le 4 \exp\left(-\frac{c_\alpha x^2}{C}\right) + \tfrac{1}{2}\exp\left(-\frac{x^2}{8\sigma^2}\right).$$

If $\sigma^2 = 0$ *the last term is zero.*

Proof. The idea of the proof is to express all $t \in [0, 1]$ in dyadic form, $t = a_1 2^{-1} + a_2 2^{-2} + \cdots$, with $a_k = 0, 1$, and then write

$$\xi(t) = \xi(0) + (\xi(a_1 2^{-1}) - \xi(0)) + (\xi(a_1 2^{-1} + a_2 2^{-2}) - \xi(a_1 2^{-1})) + \cdots \tag{12.2.2}$$

and construct bounds for each of the terms in this telescoping sum. Define for $p = 0, 1, \ldots; k = 0, 1, \ldots, 2^p - 1$,

$$\zeta(k, p) = |\xi(k2^{-p} + 2^{-p-1}) - \xi(k2^{-p})|,$$

and note that since $\xi(t)$ is normal with mean zero,

$$P\left\{\frac{\zeta(k, p)}{E(\zeta(k, p)^2)^{1/2}} \ge x\right\} = 2\int_x^\infty \phi(y)\, dy \le e^{-x^2/2} \quad \text{for } x \ge 0. \tag{12.2.3}$$

Now, let $\beta > 0$ and introduce the event

$$A = \bigcup_{p=0}^\infty \left\{\max_{0 \le k \le 2^p - 1} \frac{\zeta(k, p)}{E(\zeta(k, p)^2)^{1/2}} \ge (2\beta(p + 1))^{1/2}\right\}.$$

If $\beta > \log 2$, Boole's inequality, together with (12.2.3) implies

$$\begin{aligned} P(A) &\le \sum_{p=0}^\infty \sum_{k=0}^{2^p - 1} P\left\{\frac{\zeta(k, p)}{E(\zeta(k, p)^2)^{1/2}} \ge (2\beta(p + 1))^{1/2}\right\} \\ &\le \sum_{p=0}^\infty 2^p e^{-\beta(p+1)} = \frac{e^{-\beta}}{1 - e^{-(\beta - \log 2)}}, \end{aligned}$$

so that if $\beta > 2 \log 2$, then $P(A) \le 2 \exp(-\beta)$, so that

$$P(A) \le 4e^{-\beta}. \tag{12.2.4}$$

But for $\beta \le 2 \log 2$ this holds trivially (since $P(A) \le 1$) and we can therefore use (12.2.4) for all values of $\beta \ge 0$.

Next, note that on the complementary event A^c,

$$|\zeta(k, p)| \le E(\zeta(k, p)^2)^{1/2} \cdot (2\beta(p + 1))^{1/2}$$

for $p = 0, 1, \ldots; k = 0, 1, \ldots, 2^p - 1$, and that (12.2.1) implies that

$$E(\zeta(k, p)^2) \le C2^{-(p+1)\alpha}.$$

Thus, by (12.2.2) we conclude that on A^c,

$$|\xi(t)-\xi(0)| \le \sum_{p=0}^{\infty} C^{1/2} 2^{-(p+1)\alpha/2}(2\beta(p+1))^{1/2} = \left(\frac{\beta C}{4c_\alpha}\right)^{1/2}, \quad \text{say,}$$

and that consequently, by (12.2.4),

$$P\left\{\sup_{0\le t\le 1} |\xi(t)-\xi(0)| > \left(\frac{\beta C}{4c_\alpha}\right)^{1/2}\right\} \le 4e^{-\beta}.$$

The conclusion of the lemma now follows by choosing $\beta = c_\alpha x^2/C$, since

$$\begin{aligned} P\left\{\sup_{0\le t\le 1} \xi(t) > x\right\} &\le P\left\{\sup_{0\le t\le 1} |\xi(t)-\xi(0)| > \frac{x}{2}\right\} + P\left\{\xi(0) > \frac{x}{2}\right\} \\ &\le 4\exp\left(-\frac{c_\alpha x^2}{C}\right) + \tfrac{1}{2}\exp\left(-\frac{x^2}{8\sigma^2}\right), \end{aligned}$$

using the inequality in (12.2.3). □

With this result out of the way we return to the process $\{\xi(t)\}$ with mean zero, variance one, and covariance function $r(t)$. When considering the distribution of continuous or discrete maxima like $\sup_{0\le t\le h}\xi(t)$ and $\max_{0\le jq\le h}\xi(jq)$ for small values of h, it is natural to condition on the value of $\xi(0)$. In fact, the local behaviour (12.1.1) of $r(t)$ is reflected in the local variation of $\xi(t)$ around $\xi(0)$. For normal processes this involves no difficulty of definition if one considers $\xi(t)$ only at a finite number of points, say $t = t_j, j = 1, \ldots, n$, since conditional probabilities are then defined in terms of ratios of density functions (cf. Section 7.5). Thus we can write, with $t_0 = 0$,

$$P\left\{\max_{j=0,\ldots,n} \xi(t_j) \le u\right\} = \int_{-\infty}^{u} \phi(x) P\left\{\max_{j=1,\ldots,n} \xi(t_j) \le u \mid \xi(0) = x\right\} dx,$$

where the conditional probability can be expressed by means of a (conditional) normal density function (cf. Chapter 10). In particular, the conditional probability is determined by conditional means and covariances.

For maxima over an interval we have, e.g. with $t_j = hj2^{-n}, j = 0, \ldots, 2^n$,

$$P\left\{\sup_{0\le t\le h} \xi(t) \le u\right\} = \lim_{n\to\infty} P\left\{\max_{t_j} \xi(t_j) \le u\right\},$$

which by dominated convergence equals

$$\int_{-\infty}^{u} \phi(x) \lim_{n\to\infty} P\left\{\max_{t_j} \xi(t_j) \le u \mid \xi(0) = x\right\} dx.$$

Now, if the conditional means and covariances of $\xi(t)$ given $\xi(0)$ are such that the normal process they define is continuous, we define

$$P\left\{\sup_{0\le t\le h} \xi(t) \le u \mid \xi(0) = x\right\}$$

to be the probability that, in a continuous normal process with mean $E(\xi(t) \mid \xi(0) = x)$ and covariance function $\operatorname{Cov}(\xi(s), \xi(t) \mid \xi(0) = x)$, the maximum does not exceed u. Then clearly

$$\lim_{n \to \infty} P\left\{\max_{t_j} \xi(t_j) \le u \mid \xi(0) = x\right\} = P\left\{\sup_{0 \le t \le h} \xi(t) \le u \mid \xi(0) = x\right\},$$

so that

$$P\left\{\sup_{0 \le t \le h} \xi(t) \le u\right\} = \int_{-\infty}^{u} \phi(x) P\left\{\sup_{0 \le t \le h} \xi(t) \le u \mid \xi(0) = x\right\} dx.$$

In the applications below, the conditional distributions define a continuous process, and we shall without further comment use relations like this.

To obtain nontrivial limits as $u \to \infty$ we introduce the rescaled process

$$\xi_u(t) = u(\xi(tq) - u),$$

where we shall let q tend to zero as $u \to \infty$. Here we have to be a little more specific about this convergence than in Chapter 8, and shall assume that $u^{2/\alpha} q \to a > 0$, and let a tend to zero at a later stage.

Lemma 12.2.2. *Suppose $u \to \infty$, $q \to 0$ so that $u^{2/\alpha} q \to a > 0$. Then*

(i) *the conditional distributions of $\xi_u(t)$ given that $\xi_u(0) = x$, are normal with*

$$E(\xi_u(t) \mid \xi_u(0) = x) = x - Ca^\alpha |t|^\alpha (1 + o(1)),$$

$$\operatorname{Cov}(\xi_u(s), \xi_u(t) \mid \xi_u(0) = x) = Ca^\alpha(|s|^\alpha + |t|^\alpha - |t - s|^\alpha) + o(1)$$

where for fixed x the $o(1)$ are uniform for $\max(|s|, |t|) \le t_0$, *for all $t_0 > 0$,*

(ii) *for all $t_0 > 0$ there is a constant K, not depending on a or x, such that, for $|s|, |t| \le t_0$,*

$$\operatorname{Var}(\xi_u(s) - \xi_u(t) \mid \xi_u(0) = x) \le Ka^\alpha |t - s|^\alpha.$$

PROOF. (i) Since the process $\{\xi_u(t)\}$ is normal with mean $-u^2$ and covariance function

$$\operatorname{Cov}(\xi_u(s), \xi_u(t)) = u^2 r((t - s)q)$$

we obtain (see, for example, Rao (1973, p. 522)) that the conditional distributions are normal with

$$\begin{aligned} E(\xi_u(t) \mid \xi_u(0) = x) &= -u^2 + u^2 r(tq) \cdot \frac{1}{u^2}(x + u^2) \\ &= -u^2 + (1 - Cq^\alpha |t|^\alpha + |t|^\alpha o(q^\alpha))(x + u^2) \\ &= x - (x + u^2)(Cq^\alpha |t|^\alpha + |t|^\alpha o(q^\alpha)) \\ &= x - Ca^\alpha |t|^\alpha (1 + o(1)) \quad \text{as } q \to 0, \end{aligned}$$

since $u^2q^\alpha \to a^\alpha > 0$ and x is fixed. Furthermore,

$$\begin{aligned}\operatorname{Cov}(\xi_u(s), \xi_u(t) \mid \xi_u(0) = x) &= u^2(r((t-s)q) - r(sq)r(tq))\\ &= u^2(1 - Cq^\alpha|t-s|^\alpha - (1 - Cq^\alpha|s|^\alpha)\\ &\quad \times (1 - Cq^\alpha|t|^\alpha) + o(q^\alpha))\\ &= Ca^\alpha(|s|^\alpha + |t|^\alpha - |t-s|^\alpha) + o(1),\end{aligned}$$

uniformly for $\max(|s|, |t|) \le t_0$.

(ii) Since $\xi_u(s) - \xi_u(t)$ and $\xi_u(0)$ are normal with variances $2u^2(1 - r((t-s)q))$ and u^2, respectively, and covariance $u^2(r(sq) - r(tq))$, we have, for some constant K,

$$\begin{aligned}\operatorname{Var}(\xi_u(s) - \xi_u(t) \mid \xi_u(0) = x) &= 2u^2(1 - r((t-s)q)) - u^2(r(sq) - r(tq))^2\\ &\le 2Cu^2q^\alpha|t-s|^\alpha + o(u^2q^\alpha|t-s|^\alpha)\\ &\le Ka^\alpha|t-s|^\alpha,\end{aligned}$$

for $|s|, |t| \le t_0$. □

The first step in the derivation of the tail distribution of $M(h) = \sup\{\xi(t); 0 \le t \le h\}$ is to consider the maximum taken over a fixed number of points, $0, q, \ldots, (n-1)q$.

Lemma 12.2.3. *For each C there is a constant $H_\alpha(n, a) < \infty$ such that, if $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$, then*

$$\frac{1}{\phi(u)/u} P\left\{\max_{0 \le j < n} \xi(jq) > u\right\} \to C^{1/\alpha}H_\alpha(n, a).$$

Proof. We have

$$\begin{aligned}P\left\{\max_{0 \le j < n} \xi(jq) > u\right\} &= P\left\{\max_{0 \le j < n} \xi_u(j) > 0\right\}\\ &= P\{\xi_u(0) > 0\} + P\left\{\xi_u(0) \le 0, \max_{0 < j < n} \xi_u(j) > 0\right\},\end{aligned}$$

where $P\{\xi_u(0) > 0\} = P\{\xi(0) > u\} = 1 - \Phi(u) \sim \phi(u)/u$. Since, furthermore, $\xi_u(0)$ is normal with mean $-u^2$ and variance u^2, we have

$$\begin{aligned}&P\left\{\xi_u(0) \le 0, \max_{0 < j < n} \xi_u(j) > 0\right\}\\ &= \int_{-\infty}^0 \frac{1}{u}\phi\left(u + \frac{x}{u}\right) P\left\{\max_{0 < j < n} \xi_u(j) > 0 \mid \xi_u(0) = x\right\} dx\\ &= \frac{\phi(u)}{u}\int_{-\infty}^0 \exp\left(-x - \frac{x^2}{2u^2}\right) P\left\{\max_{0 < j < n} (\xi_u(j) - x) > -x \mid \xi_u(0) = x\right\} dx.\end{aligned} \tag{12.2.5}$$

By Lemma 12.2.2(i), for any fixed x,

$$E(\xi_u(j) - x \mid \xi_u(0) = x) \to -Ca^\alpha |j|^\alpha,$$

$$\mathrm{Cov}(\xi_u(i) - x, \xi_u(j) - x \mid \xi_u(0) = x) \to Ca^\alpha(|i|^\alpha + |j|^\alpha - |i - j|^\alpha)$$

as $q \to 0$. Since limits of covariances are covariances, one can define normal r.v.'s, $Y_a(j)$, $0 < j < n$, with means and covariances depending on $a = \lim qu^{2/\alpha}$,

$$E(Y_a(j)) = -Ca^\alpha |j|^\alpha,$$

$$\mathrm{Cov}(Y_a(i), Y_a(j)) = Ca^\alpha(|i|^\alpha + |j|^\alpha - |i - j|^\alpha).$$

Now convergence of moments implies convergence in distribution for jointly normal r.v.'s (as can easily be seen, e.g. using characteristic functions). If $A = (-x, \infty)$, (with boundary $\partial A = \{-x\}$), then clearly

$$P\left\{\max_{0<j<n} Y_a(j) \in \partial A\right\} \le \sum_{j=1}^{n-1} P\{Y_a(j) = -x\}$$

which is zero since the one-dimensional distributions of $Y_a(1), \ldots, Y_a(n-1)$ are all continuous.

It follows that

$$P\left\{\max_{0<j<n} (\xi_u(j) - x) > -x \mid \xi_u(0) = x\right\} \to P\left\{\max_{0<j<n} Y_a(j) > -x\right\}.$$

To be able to use the dominated convergence theorem in (12.2.5) we note that, by Lemma 12.2.2(i) for $x < 0$

$$\begin{aligned} P&\left\{\max_{0<j<n} (\xi_u(j) - x) > -x \mid \xi_u(0) = x\right\} \\ &\le \sum_{j=1}^{n} P\{\xi_u(j) - x > -x \mid \xi_u(0) = x\} \\ &\le n(1 - \Phi(c' - c''x)) \le n\frac{\phi(c' - c''x)}{c' - c''x} \end{aligned}$$

for some constants $c', c'' > 0$. This shows that the convergence in (12.2.5) is dominated, and we obtain

$$\frac{1}{\phi(u)/u} P\left\{\max_{0\le j<n} \xi(jq) > u\right\} \to 1 + \int_{-\infty}^{0} e^{-x} P\left\{\max_{0<j<n} Y_a(j) > -x\right\} dx < \infty,$$

which proves the existence and finiteness of the constant $H_\alpha(n, a)$. □

For future use we note the following expression for the constant $H_\alpha(n, a)$:

$$H_\alpha(n, a) = C^{-1/\alpha}\left(1 + \int_{-\infty}^{0} e^{-x} P\left\{\max_{0<j<n} Y_a(j) > -x\right\} dx\right). \quad (12.2.6)$$

Lemma 12.2.4. *Suppose $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$, and take h such that $\sup_{\varepsilon \le t \le h} r(t) < 1$ for all $\varepsilon > 0$. Then, for each C,*

(i) *there is a constant $H_\alpha(a) < \infty$ such that*

$$\frac{H_\alpha(n, a)}{na} \to H_\alpha(a) \quad \text{as } n \to \infty,$$

and

$$\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\} \to hC^{1/\alpha}H_\alpha(a),$$

(ii) *$H_\alpha(a_0) > 0$ for some $a_0 > 0$.*

PROOF. (i) Let n be a fixed integer, write $m = [h/nq]$, and

$$B_r = \left\{\max_{rn \le j < (r+1)n} \xi(jq) > u\right\}.$$

Then

$$P\left(\bigcup_{r=0}^{m-1} B_r\right) \le P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\} \le P\left(\bigcup_{r=0}^{m} B_r\right),$$

where, by Lemma 12.2.3,

$$\begin{aligned} P\left(\bigcup_{r=0}^{m} B_r\right) &\le (m+1)P(B_0) \sim (m+1)\frac{\phi(u)}{u} C^{1/\alpha}H_\alpha(n, a) \\ &\sim \frac{hu^{2/\alpha}}{na}\frac{\phi(u)}{u} C^{1/\alpha}H_\alpha(n, a) \end{aligned}$$

since $1/q \sim u^{2/\alpha}/a$ by assumption. Hence

$$\limsup_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\} \le hC^{1/\alpha}\frac{H_\alpha(n, a)}{na} < \infty. \tag{12.2.7}$$

Furthermore

$$\begin{aligned} P\left(\bigcup_{r=0}^{m-1} B_r\right) &\ge \sum_{r=0}^{m-1} P(B_r) - \sum_{r \ne s} P(B_r \cap B_s) \\ &\ge mP(B_0) - m\sum_{r=1}^{m-1} P(B_0 \cap B_r) \end{aligned}$$

so that

$$\begin{aligned} \liminf_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left(\bigcup_{r=0}^{m-1} B_r\right) &\ge hC^{1/\alpha}\frac{H_\alpha(n, a)}{na} \\ &\quad - \limsup_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} m\sum_{r=1}^{m-1} P(B_0 \cap B_r) \\ &= hC^{1/\alpha}\frac{H_\alpha(n, a)}{na} - \rho_n, \quad \text{say.} \end{aligned} \tag{12.2.8}$$

We shall now show that

$$\rho_n = \limsup_{u\to\infty} \frac{1}{u^{2/\alpha}\phi(u)/u}\, m \sum_{r=1}^{m-1} P(B_0 \cap B_r) \to 0 \quad \text{as } n \to \infty. \tag{12.2.9}$$

By Boole's inequality and stationarity

$$\begin{aligned} m\sum_{r=1}^{m-1} P(B_0 \cap B_r) &= m\sum_{r=1}^{m-1} P\left(\bigcup_{0\le i<n}\ \bigcup_{rn\le j<(r+1)n} \{\xi(iq) > u, \xi(jq) > u\}\right) \\ &\le m\sum_{i=0}^{n-1}\sum_{j=n}^{mn} P\{\xi(iq) > u, \xi(jq) > u\} \\ &\le m\sum_{j=1}^{n} jP\{\xi(0) > u, \xi(jq) > u\} \\ &\quad + mn\sum_{j=n+1}^{mn} P\{\xi(0) > u, \xi(jq) > u\}. \end{aligned} \tag{12.2.10}$$

To estimate these sums we use different techniques for small and large values of jq. Let $\varepsilon > 0$ be such that

$$1 \ge 1 - r(t) \ge \frac{C}{2}|t|^\alpha \quad \text{for } |t| \le \varepsilon.$$

Assume $jq \le \varepsilon$, and write $r = r(jq)$. Since the conditional distribution of $\xi(jq)$ given $\xi(0) = x$, is normal with mean rx and variance $1 - r^2$,

$$\begin{aligned} P\{\xi(0) > u, \xi(jq) > u\} &= \int_u^\infty \phi(x) P\{\xi(jq) > u \mid \xi(0) = x\}\, dx \\ &= \int_u^\infty \phi(x)\left(1 - \Phi\left(\frac{u - xr}{\sqrt{1 - r^2}}\right)\right) dx \\ &= \left[-(1 - \Phi(x))\left(1 - \Phi\left(\frac{u - xr}{\sqrt{1 - r^2}}\right)\right)\right]_{x=u}^\infty \\ &\quad + \int_u^\infty (1 - \Phi(x)) \frac{r}{\sqrt{1 - r^2}}\,\phi\left(\frac{u - xr}{\sqrt{1 - r^2}}\right) dx \\ &\le (1 - \Phi(u))\left(1 - \Phi\left(u\sqrt{\frac{1 - r}{1 + r}}\right)\right) \\ &\quad + \int_u^\infty (1 - \Phi(u)) \frac{r}{\sqrt{1 - r^2}}\,\phi\left(\frac{u - xr}{\sqrt{1 - r^2}}\right) dx \\ &= 2(1 - \Phi(u))\left(1 - \Phi\left(u\sqrt{\frac{1 - r}{1 + r}}\right)\right) \\ &\le 2\frac{\phi(u)}{u}\left(1 - \Phi\left(u\sqrt{\frac{1 - r}{1 + r}}\right)\right). \end{aligned} \tag{12.2.11}$$

Here,

$$\frac{1-r}{1+r} = \frac{1-r(jq)}{1+r(jq)} \geq \frac{C}{4}|jq|^\alpha \geq Kj^\alpha a^\alpha u^{-2}, \tag{12.2.12}$$

for some constant $K > 0$, and thus if $nq \leq \varepsilon$,

$$\begin{aligned}
m\sum_{j=1}^{n} jP\{\xi(0) > u, \xi(jq) > u\} \\
&\leq 2m\sum_{j=1}^{n} j\frac{\phi(u)}{u}(1 - \Phi(\sqrt{Kj^\alpha a^\alpha})) \\
&\sim \frac{2h}{q}\frac{\phi(u)}{u}\cdot\frac{1}{n}\sum_{j=1}^{n} j(1 - \Phi(\sqrt{Kj^\alpha a^\alpha})) \\
&\leq \frac{u^{2/\alpha}\phi(u)}{u}\cdot\frac{K'h}{a}\frac{1}{n}\sum_{j=1}^{\infty} j(1 - \Phi(\sqrt{Kj^\alpha a^\alpha}))
\end{aligned} \tag{12.2.13}$$

for some (generic) constant K'. Since $1 - \Phi(x) \leq \phi(x)/x$ the sum

$$\sum_{j=1}^{\infty} j(1 - \Phi(\sqrt{Kj^\alpha a^\alpha}))$$

is convergent, and since $nq \to 0$ as $u \to \infty$ (n fixed)

$$\limsup_{u\to\infty} \frac{1}{u^{2/\alpha}\phi(u)/u}\, m\sum_{j=1}^{n} jP\{\xi(0) > u, \xi(jq) > u\} = O(1/n). \tag{12.2.14}$$

For the second sum in (12.2.10) we get, again using (12.2.11) and (12.2.12), as $n \to \infty$,

$$mn\sum_{j=n+1}^{[\varepsilon/q]} P\{\xi(0) > u, \xi(jq) > u\} \leq \frac{K'h}{qu^{2/\alpha}}\frac{u^{2/\alpha}\phi(u)}{u}\sum_{j=n+1}^{\infty}(1 - \Phi(\sqrt{Kj^\alpha a^\alpha})), \tag{12.2.15}$$

where the sum is convergent. For terms with $jq \geq \varepsilon$ we use the estimate from Corollary 4.2.2, (using $\delta = \sup_{\varepsilon\leq t\leq h}|r(t)| < 1$),

$$P\{\xi(0) > u, \xi(jq) > u\} \leq (1 - \Phi(u))^2 + K'\exp\left(-\frac{u^2}{1 + |r(jq)|}\right),$$

which implies that

$$\begin{aligned}
mn\sum_{j=[\varepsilon/q]+1}^{mn} P\{\xi(0) > u, \xi(jq) > u\} \\
&\leq \left(\frac{h}{q}\right)^2(1 - \Phi(u))^2 + \frac{K'h}{q}\sum_{\varepsilon<jq\leq h}\exp\left(-\frac{u^2}{1 + |r(jq)|}\right).
\end{aligned}$$

Since again $\delta < 1$, this is bounded by

$$\frac{h^2}{q^2}(1-\Phi(u))^2 + \frac{K'h^2}{q^2}\exp\left(-\frac{u^2}{1+\delta}\right)$$
$$= \frac{u^{2/\alpha}\phi(u)}{u}\left(K'\frac{u}{u^{2/\alpha}q^2}\exp\left(-\frac{u^2}{2}\cdot\frac{1-\delta}{1+\delta}\right) + o(1)\right)$$
$$= \frac{u^{2/\alpha}\phi(u)}{u}\cdot o(1) \quad \text{as } u \to \infty, \tag{12.2.16}$$

since $u/(u^{2/\alpha}q^2) \sim u^{1+2/\alpha}/a^2$.

Together, (12.2.15) and (12.2.16) imply that

$$\limsup_{u\to\infty} \frac{1}{u^{2/\alpha}\phi(u)/u}\, mn \sum_{j=n+1}^{mn} P\{\xi(0) > u, \xi(jq) > u\}$$
$$= K' \sum_{j=n+1}^{\infty} (1 - \Phi(\sqrt{Kj^\alpha a^\alpha})) \to 0 \quad \text{as } n \to \infty,$$

and combining this with (12.2.14) and (12.2.10) we obtain (12.2.9).

Thus we have shown that

$$hC^{1/\alpha}\cdot\frac{H_\alpha(n,\alpha)}{na} - \rho_n \le \liminf_{u\to\infty}\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0\le jq\le h}\xi(jq) > u\right\}$$
$$\le \limsup_{u\to\infty}\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0\le jq\le h}\xi(jq) > u\right\}$$
$$\le hC^{1/\alpha}\cdot\frac{H_\alpha(n,a)}{na},$$

where $\rho_n \to 0$ as $n \to \infty$. Since $H_\alpha(n, a) < \infty$ for all n, and the lim inf and lim sup do not depend on n, this implies the existence of

$$\lim_{n\to\infty}\frac{H_\alpha(n,a)}{na} = H_\alpha(a),$$

which is then the joint value of lim inf and lim sup. Furthermore this proves that $H_\alpha(a) < \infty$.

(ii) Again take $\varepsilon > 0$ small enough to make (12.1.2) hold for $|jq| \le \varepsilon$. Applying (12.2.11) we obtain for $|h| \le \varepsilon$,

$$P\left\{\max_{0\le jq\le h}\xi(jq) > u\right\}$$
$$\ge [h/q]P\{\xi(0) > u\} - [h/q]\sum_{0<jq\le h} P\{\xi(0) > u, \xi(jq) > u\}$$
$$\ge [h/q](1-\Phi(u)) - 2[h/q]\sum_{0<jq\le h}\frac{\phi(u)}{u}\left(1 - \Phi\left(u\sqrt{\frac{1-r(jq)}{1+r(jq)}}\right)\right)$$
$$\ge [h/q]\frac{\phi(u)}{u}\left(1 + o(1) - 2\sum_{j=1}^{[h/q]+1}(1-\Phi(\sqrt{Kj^\alpha a^\alpha}))\right).$$

But $[h/q] \sim (h/a)u^{2/\alpha}$, and

$$\sum_{j=1}^{[h/q]+1} (1 - \Phi(\sqrt{Kj^\alpha a^\alpha})) \le \sum_{j=1}^{\infty} (1 - \Phi(\sqrt{Kj^\alpha a^\alpha}))$$
$$\le \sum_{j=1}^{\infty} \frac{\phi(\sqrt{Kj^\alpha a^\alpha})}{\sqrt{Kj^\alpha a^\alpha}} \to 0 \quad \text{as } a \to \infty,$$

and hence there is then certainly one a_0 that makes the sum less than 1/2, giving $P\{\max_{0 \le jq \le h} \xi(jq) > u\} \ge \eta u^{2/\alpha}\phi(u)/u$ for some $\eta > 0$, from which (ii) clearly follows. □

The following three lemmas relate the continuous maximum $\sup_{0 \le t \le h} \xi(t)$ to the discrete one $\max_{0 \le jq \le h} \xi(jq)$. We first prove that we can neglect the probability that the discrete maximum is less than $u - \gamma/u$ and the continuous is greater than u, as $\gamma = a^\beta \to 0$.

Lemma 12.2.5. *Let $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$, and let $\gamma = a^\beta$ for some positive constant $\beta < \alpha/2$. Then*

$$v_a = \limsup_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\max_{0 \le jq \le h} \xi(jq) \le u - \frac{\gamma}{u}, \sup_{0 \le t \le h} \xi(t) > u\right\} \to 0$$
as $a \to 0$.

PROOF. By Boole's inequality and stationarity

$$P\left\{\max_{0 \le jq \le h} \xi(jq) \le u - \frac{\gamma}{u}, \sup_{0 \le t \le h} \xi(t) > u\right\}$$
$$\le ([h/q] + 1) P\left\{\xi(0) \le u - \frac{\gamma}{u}, \sup_{0 \le t \le q} \xi(t) > u\right\},$$

and with $\xi_u(s) = u(\xi(sq) - u)$, we can write

$$P\left\{\xi(0) \le u - \frac{\gamma}{u}, \sup_{0 \le t \le q} \xi(t) > u\right\}$$
$$= P\left\{\xi_u(0) \le -\gamma, \sup_{0 \le s \le 1} \xi_u(s) > 0\right\}$$
$$= \int_{y=-\infty}^{-\gamma} \frac{1}{u}\phi\left(u + \frac{y}{u}\right) P\left\{\sup_{0 \le s \le 1} \xi_u(s) > 0 \mid \xi_u(0) = y\right\} dy.$$

By Lemma 12.2.2(i), the conditional distributions of $\xi_u(s)$ given $\xi_u(0) = y$ are normal with mean

$$\mu(s) = y - Ca^\alpha |s|^\alpha (1 + o(1)) \quad \text{as } q \to 0$$

with the $o(1)$ uniform in $|s| \le 1$. Here $\mu(s) < y$ for small q, and

$$P\left\{\sup_{0 \le s \le 1} \xi_u(s) > 0 \mid \xi_u(0) = y\right\} \le P\left\{\sup_{0 \le s \le 1} (\xi_u(s) - \mu(s)) > -y \mid \xi_u(0) = y\right\},$$

where, conditional on $\xi_u(0)$, $\xi_u(s) - \mu(s)$ is a nonstationary normal process with mean zero and, by Lemma 12.2.2(ii), incremental variance

$$\operatorname{Var}(\xi_u(s) - \xi_u(t) \mid \xi_u(0) = y) \le Ka^\alpha |t - s|^\alpha,$$

for some constant K which is independent of a and y. Fernique's lemma (Lemma 12.2.1) implies that, with $c = c_\alpha/K$,

$$P\left\{\sup_{0 \le s \le 1} (\xi_u(s) - \mu(s)) > -y \mid \xi_u(0) = y\right\} \le 4\exp(-ca^{-\alpha}y^2),$$

and thus, using K and c as generic constants,

$$\begin{aligned}
\frac{1}{u^{2/\alpha}\phi(u)/u} & P\left\{\max_{0 \le jq \le h} \xi(jq) \le u - \frac{\gamma}{u}, \sup_{0 \le t \le h} \xi(t) > u\right\} \\
&\le \frac{Kh}{qu^{2/\alpha}\phi(u)} \int_{-\infty}^{-\gamma} \phi\left(u + \frac{y}{u}\right) \cdot \exp(-ca^{-\alpha}y^2)\, dy \\
&\le \frac{K}{qu^{2/\alpha}} \int_{-\infty}^{-\gamma} \exp(-y - ca^{-\alpha}y^2)\, dy \\
&\le \frac{K}{qu^{2/\alpha}} \int_{-\infty}^{-\gamma} \exp(-ca^{-\alpha}y^2)\, dy \\
&\sim Ka^{\alpha/2-1}\Phi(-ca^{\beta-\alpha/2}).
\end{aligned}$$

Clearly, this tends to zero as $a \to 0$, since $\beta < \alpha/2$, which proves the lemma. □

Lemma 12.2.6. *If $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$, and $\gamma = a^\beta$ for some constant $\beta > 0$, then, with h as in Lemma 12.2.4,*

$$\lim_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{u - \frac{\gamma}{u} < \max_{0 \le jq \le h} \xi(jq) \le u\right\} = h(e^\gamma - 1)C^{1/\alpha}H_\alpha(a).$$

Proof. Since $u^{2/\alpha}q \to a > 0$ implies $(u - \gamma/u)^{2/\alpha}q \to a$, and since furthermore

$$\left(u - \frac{\gamma}{u}\right)^{2/\alpha} \frac{\phi(u - \gamma/u)}{u - \gamma/u} \sim e^\gamma \frac{u^{2/\alpha}\phi(u)}{u}$$

as $u \to \infty$, it follows from Lemma 12.2.4(i) that

$$\begin{aligned}
\frac{1}{u^{2/\alpha}\phi(u)/u} & P\left\{u - \frac{\gamma}{u} < \max_{0 \le jq \le h} \xi(jq) \le u\right\} \\
&= \frac{1}{u^{2/\alpha}\phi(u)/u}\left(P\left\{\max_{0 \le jq \le h} \xi(jq) > u - \frac{\gamma}{u}\right\} - P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\}\right) \\
&\to he^\gamma C^{1/\alpha}H_\alpha(a) - hC^{1/\alpha}H_\alpha(a).
\end{aligned}$$

□

Lemma 12.2.7. *Under the conditions of Lemma* 12.2.4,

(i) $$hC^{1/\alpha}H_\alpha(a) \le \liminf_{u\to\infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\sup_{0\le t\le h} \xi(t) > u\right\}$$
$$\le \limsup_{u\to\infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\sup_{0\le t\le h} \xi(t) > u\right\}$$
$$\le v_a + h(e^\gamma - 1)C^{1/\alpha}H_\alpha(a) + hC^{1/\alpha}H_\alpha(a) < \infty, \tag{12.2.17}$$

for $\gamma = a^\beta$, *where, by Lemma* 12.2.5, $v_a \to 0$ *as* $a \to 0$,

(ii) $$\lim_{a\to 0} H_\alpha(a) = H_\alpha, \quad say$$

exists finite, and

$$\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\sup_{0\le t\le h} \xi(t) > u\right\} \to hC^{1/\alpha}H_\alpha, \tag{12.2.18}$$

(iii) H_α *is independent of* C.

PROOF. Since

$$P\left\{\max_{0\le jq\le h} \xi(jq) > u\right\} \le P\left\{\sup_{0\le t\le h} \xi(t) > u\right\}$$
$$\le P\left\{\max_{0\le jq\le h} \xi(jq) \le u - \frac{\gamma}{u}, \sup_{0\le t\le h} \xi(t) > u\right\}$$
$$+ P\left\{u - \frac{\gamma}{u} < \max_{0\le jq\le h} \xi(jq) \le u\right\}$$
$$+ P\left\{\max_{0\le jq\le h} \xi(jq) > u\right\},$$

part (i) follows directly from Lemmas 12.2.4, 12.2.5, and 12.2.6.

Further, the middle limits in (12.2.17) are independent of a, and it follows that $\limsup_{a\to 0} H_\alpha(a) < \infty$. Therefore

$$h(e^\gamma - 1)C^{1/\alpha}H_\alpha(a) \to 0$$

as $a \to 0$, and since $v_a \to 0$ it follows as in the proof of Lemma 12.2.4(i) that $\lim_{a\to 0} H_\alpha(a)$ exists, finite and (12.2.18) holds.

For part (iii), note that if $\xi(t)$ satisfies (12.1.1) then the covariance function $\tilde{r}(\tau)$ of $\tilde{\xi}(t) = \xi(t/C^{1/\alpha})$ satisfies

$$\tilde{r}(\tau) = 1 - |\tau|^\alpha + o(|\tau|^\alpha) \quad \text{as } \tau \to 0.$$

Furthermore,

$$\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\sup_{0\le t\le h} \xi(t) > u\right\} = \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{\sup_{0\le t\le hC^{1/\alpha}} \tilde{\xi}(t) > u\right\},$$

which by (ii) shows that H_α does not depend on C. □

One immediate consequence of (12.2.18) and Lemma 12.2.4(i) is that

$$
\begin{aligned}
\limsup_{u\to\infty} & \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{ \max_{0\le jq\le h} \xi(jq) \le u,\ \sup_{0\le t\le h} \xi(t) > u \right\} \\
&= \limsup_{u\to\infty} \left(\frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{ \sup_{0\le t\le h} \xi(t) > u \right\} \right. \\
&\qquad \left. - \frac{1}{u^{2/\alpha}\phi(u)/u} P\left\{ \max_{0\le jq\le h} \xi(jq) > u \right\} \right) \\
&= hC^{1/\alpha}(H_\alpha - H_\alpha(a)) \to 0 \quad \text{as } a \to 0. \qquad (12.2.19)
\end{aligned}
$$

Of course, (12.2.18) has its main interest if $H_\alpha > 0$, but to prove this requires a little further work as follows.

Lemma 12.2.8. $H_\alpha > 0$.

Proof. We have from Lemma 12.2.4(i) and (ii) that there is an $a_0 > 0$ such that

$$
H_\alpha(a_0) = \lim_{n\to\infty} \frac{H_\alpha(n, a_0)}{na_0} > 0.
$$

Let the $Y_a(j)$ be as in the proof of Lemma 12.2.3, i.e. normal with mean $-Ca^\alpha|j|^\alpha$ and covariances $Ca^\alpha(|i|^\alpha + |j|^\alpha - |j - i|^\alpha)$. Then we have from (12.2.6),

$$
C^{1/\alpha}H_\alpha(n, a) = 1 + \int_{-\infty}^{0} e^{-x} P\left\{ \max_{0<j<n} Y_a(j) > -x \right\} dx,
$$

$$
C^{1/\alpha}H_\alpha(nk, a) = 1 + \int_{-\infty}^{0} e^{-x} P\left\{ \max_{0<j<nk} Y_a(j) > -x \right\} dx,
$$

$$
C^{1/\alpha}H_\alpha(n, ak) = 1 + \int_{-\infty}^{0} e^{-x} P\left\{ \max_{0<j<n} Y_{ak}(j) > -x \right\} dx.
$$

Here $Y_a(jk), j = 1, \ldots, n$ have the same distributions as $Y_{ak}(j), j = 1, \ldots, n$, which implies

$$
H_\alpha(n, ak) \le H_\alpha(nk, a)
$$

for $k = 1, 2, \ldots$, the r.v.'s in $H_\alpha(n, ak)$ forming a subset of those appearing in $H_\alpha(nk, a)$. Thus

$$
0 < H_\alpha(a_0) = \lim_{n\to\infty} \frac{H_\alpha(n, a_0)}{na_0} \le \lim_{n\to\infty} \frac{H_\alpha(nk, a_0/k)}{nk \cdot (a_0/k)} = H_\alpha(a_0/k),
$$

and since $H_\alpha(a_0/k) \to H_\alpha$ as $k \to \infty$, the lemma follows. □

By combining Lemmas 12.2.7 and 12.2.8 we obtain the tail of the distribution of the maximum $M(h) = \sup\{\xi(t); 0 \le t \le h\}$ over a fixed interval.

Theorem 12.2.9. *If $r(t)$ satisfies* (12.1.1), *then for each fixed $h > 0$ such that* $\sup_{\varepsilon \le t \le h} r(t) = \delta_\varepsilon < 1$ *for all* $\varepsilon > 0$,

$$\lim_{u \to \infty} \frac{1}{u^{2/\alpha}\phi(u)/u} P\{M(h) > u\} = hC^{1/\alpha}H_\alpha,$$

where $H_\alpha > 0$ is a finite constant depending only on α.

Remark 12.2.10. In the proof of Theorem 12.2.9 we obtained the *existence* of the constant H_α by rather tricky estimates, starting with

$$H_\alpha(n, a) = C^{-1/\alpha}\left(1 + \int_{-\infty}^{0} e^{-x} P\left\{\max_{0 \le j \le n} Y_a(j) > -x\right\} dx\right).$$

By pursuing these estimates further one can obtain a related expression for H_α,

$$H_\alpha = \lim_{T \to \infty} T^{-1} \int_{-\infty}^{0} e^{-x} P\left\{\sup_{0 \le t \le T} Y_0(t) > -x\right\} dx,$$

where $\{Y_0(t)\}$ is a nonstationary normal process with mean $-|t|^\alpha$ and covariances $|s|^\alpha + |t|^\alpha - |t - s|^\alpha$. However, this does not seem to be very instructive, nor of much help in computing H_α.

It should be noted, though, that the proper time-normalization of the distribution of $M(h)$ only depends on the covariance function through the time-scale $C^{1/\alpha}$ and on the constant H_α. Therefore, if one can find the limiting form of the tail of the distribution of $M(h)$ (for some h) for one single process satisfying (12.1.1) one also knows the value of H_α for that particular α. For $\alpha = 2$ this is easily done, by considering the simple cosine-process (7.4.3). By comparing (7.4.7) and Theorem 12.2.9, we find $H_2 = 1/\sqrt{\pi}$.

The only other value of α for which the tail of the distribution of $M(h)$ has been found is $\alpha = 1$. In fact, explicit expressions for the entire distribution of $M(h)$ are known for the normal process with triangular covariance function $r(t) = 1 - |t|$, $|t| \le 1$, see Slepian (1961), and as a result one has $H_1 = 1$. In particular, this shows that for the Ornstein–Uhlenbeck process, with $r(t) = \exp(-|t|)$, $P\{M(h) > u\} \sim hu\phi(u)$. □

Before proceeding to the maxima over increasing intervals we formulate the following lemma for later reference.

Lemma 12.2.11. *Suppose $\{\xi(t)\}$ satisfies* (12.1.1), *let $h > 0$ be fixed such that* $\sup_{\varepsilon \le t \le h} r(t) < 1$ *for all* $\varepsilon > 0$, *and let* $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$. *Then for every interval I of length h,*

$$0 \le P\{\xi(jq) \le u, jq \in I\} - P\{M(I) \le u\} \le \mu h \rho_a + o(\mu),$$

where $\mu = C^{1/\alpha}H_\alpha u^{2/\alpha}\phi(u)/u$, $\rho_a = 1 - H_\alpha(a)/H_\alpha \to 0$ *as* $a \to 0$, *and the* $o(\mu)$*-term is the same for all intervals of length h.*

Proof. By stationarity

$$\begin{aligned}0 &\le P\{\xi(jq) \le u, jq \in I\} - P\{M(I) \le u\} \\ &\le P\{\xi(0) > u\} + P\{\xi(jq) \le u, jq \in [0, h]\} - P\{M(h) \le u\},\end{aligned}$$

where $P\{\xi(0) > u\} \le \phi(u)/u = o(\mu)$. Therefore the result is immediate from Lemma 12.2.4(i),

$$\mu^{-1} P\left\{\max_{0 \le jq \le h} \xi(jq) > u\right\} = \frac{hH_\alpha(a)}{H_\alpha} + o(1),$$

and (12.2.18),

$$\mu^{-1} P\{M(h) > u\} = h + o(1). \qquad \square$$

12.3. Maxima Over Increasing Intervals

The covariance condition (8.1.2), i.e. $r(t) \log t \to 0$ as $t \to \infty$, is also sufficient to establish the double exponential limit for the maximum $M(T) = \sup\{\xi(t); 0 \le t \le T\}$ in this general case. We then let $T \to \infty$, $u \to \infty$ so that

$$T\mu = TC^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u \to \tau > 0,$$

i.e. $TP\{M(h) > u\} \to \tau h$. Taking logarithms we get

$$\log T + \log(C^{1/\alpha} H_\alpha (2\pi)^{-1/2}) + \frac{2 - \alpha}{\alpha} \log u - \frac{u^2}{2} \to \log \tau,$$

implying

$$u^2 \sim 2 \log T,$$

or $\log u = \frac{1}{2} \log 2 + \frac{1}{2} \log \log T + o(1)$, which gives

$$\begin{aligned}u^2 = 2 \log T &+ \frac{2 - \alpha}{\alpha} \log \log T - 2 \log \tau \\ &+ 2 \log(C^{1/\alpha} H_\alpha (2\pi)^{-1/2} 2^{(2-\alpha)/2\alpha}) + o(1).\end{aligned} \qquad (12.3.1)$$

Lemma 12.3.1. *Let $\varepsilon > 0$ be given, and suppose* (12.1.1) *and* (8.1.2) *both hold. Let $T \sim \tau/\mu$ for $\tau > 0$ fixed and with $\mu = C^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u$, so that $u \sim (2 \log T)^{1/2}$ as $T \to \infty$, and let $q \to 0$ as $u \to \infty$ in such a way that $u^{2/\alpha} q \to a > 0$. Then*

$$\frac{T}{q} \sum_{\varepsilon \le kq \le T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0 \quad \text{as } T \to \infty. \qquad (12.3.2)$$

Proof. This lemma corresponds to Lemma 8.1.1. First, we split the sum in (12.3.2) at T^β, where β is a constant such that $0 < \beta < (1 - \delta)/(1 + \delta)$,

$\delta = \sup\{|r(t)|; t \geq \varepsilon\} < 1$. Then, with the generic constant K, since $\exp(-u^2/2) \leq K/T$, $u^2 \sim 2 \log T$,

$$\begin{aligned}
\frac{T}{q} \sum_{\varepsilon \leq kq \leq T^\beta} |r(kq)| \exp\left\{-\frac{u^2}{1+|r(kq)|}\right\} &\leq \frac{T^{\beta+1}}{q^2} \exp\left\{-\frac{u^2}{1+\delta}\right\} \\
&\leq \frac{K}{q^2} T^{\beta+1-2/(1+\delta)} \sim \frac{K}{(u^{2/\alpha}q)^2} (\log T)^{2/\alpha} T^{\beta+1-2/(1+\delta)} \\
&\to 0 \quad \text{as } T \to \infty,
\end{aligned}$$

and $u^{2/\alpha}q \to a > 0$.

With $\delta(t) = \sup\{|r(s) \log s|; s \geq t\}$, we have $|r(t)| \leq \delta(t)/\log t$ as $t \to \infty$, and hence for $kq \geq T^\beta$,

$$\exp\left\{-\frac{u^2}{1+|r(kq)|}\right\} \leq \exp\left\{-u^2\left(1 - \frac{\delta(T^\beta)}{\log T^\beta}\right)\right\},$$

so that the remaining sum is bounded by

$$\begin{aligned}
&\frac{T}{q} \sum_{T^\beta < kq \leq T} |r(kq)| \exp\left\{-u^2\left(1 - \frac{\delta(T^\beta)}{\log T^\beta}\right)\right\} \\
&\leq \left(\frac{T}{q}\right)^2 \exp\left\{-u^2\left(1 - \frac{\delta(T^\beta)}{\log T^\beta}\right)\right\} \frac{1}{\log T^\beta} \cdot \frac{q}{T} \sum_{T^\beta < kq \leq T} |r(kq)| \log kq.
\end{aligned} \tag{12.3.3}$$

Since $r(t) \log t \to 0$, we also have

$$\frac{q}{T} \sum_{T^\beta < kq \leq T} |r(kq)| \log kq \to 0$$

as $T \to \infty$, while for the remaining factor in (12.3.3) we have to use the more precise estimate from (12.3.1),

$$u^2 = 2 \log T + \frac{2-\alpha}{\alpha} \log \log T + O(1).$$

Since $\delta(t) \to 0$ as $t \to \infty$, we see that for some constant $K > 0$, since $\beta < 1$,

$$\exp\left\{-u^2\left(1 - \frac{\delta(T^\beta)}{\log T^\beta}\right)\right\} \leq K \exp(-u^2) \leq K T^{-2} (\log T)^{-(2-\alpha)/\alpha}.$$

Thus, since $u^2 \sim 2 \log T$ and $u^{2/\alpha}q \to a$,

$$\begin{aligned}
&\frac{T}{q} \sum_{T^\beta < kq \leq T} |r(kq)| \exp\left\{-u^2\left(1 - \frac{\delta(T^\beta)}{\log T^\beta}\right)\right\} \\
&\leq \left(\frac{T}{q}\right)^2 T^{-2} (\log T)^{-(2-\alpha)/\alpha} \frac{1}{\log T^\beta} o(1) \\
&= \frac{1}{(u^{2/\alpha}q)^2} (\log T)^{2/\alpha} (\log T)^{-(2-\alpha)/\alpha} \frac{1}{\log T^\beta} o(1) = o(1),
\end{aligned}$$

and this concludes the proof of the lemma. □

We can now proceed along similar lines to the proof of Theorem 8.2.5. First, take a fixed $h > 0$, write $n = [T/h]$, and divide $[0, nh]$ into h intervals of length h, and then split each interval into subintervals I_k, I_k^* of length $h - \varepsilon$ and ε, respectively. We then show asymptotic independence of maxima, first giving the following lemmas, corresponding to Lemmas 8.2.3 and 8.2.4, respectively.

Lemma 12.3.2. *Suppose $u \to \infty$, $q \to 0$, $u^{2/\alpha}q \to a > 0$, (12.1.1) holds, and $T\mu \to \tau > 0$. Then*

(i) $$\limsup_{u\to\infty}\left|P\left\{M\left(\bigcup_1^n I_k\right) \le u\right\} - P\{M(nh) \le u\}\right| \le \tau\cdot\frac{\varepsilon}{h},$$

(ii) $$\limsup_{u\to\infty}\left|P\left\{\xi(jq) \le u, jq \in \bigcup_1^n I_k\right\} - P\left\{M\left(\bigcup_1^n I_k\right) \le u\right\}\right| \le \tau\rho_a,$$

where $\rho_a \to 0$ as $a \to 0$.

Proof. Part (i) follows at once from Boole's inequality and Theorem 12.2.9, since

$$0 \le P\left\{M\left(\bigcup_1^n I_k\right) \le u\right\} - P\{M(nh) \le u\} \le nP\{M(I_1^*) > u\} \sim n\mu\varepsilon \to \tau\cdot\frac{\varepsilon}{h},$$

since $n\mu \sim T\mu/h \to \tau/h$.

Part (ii) follows similarly from Lemma 12.2.11, which implies

$$\begin{aligned}0 &\le P\left\{\xi(jq) \le u, jq \in \bigcup_1^n I_k\right\} - P\left\{M\left(\bigcup_1^n I_k\right) \le u\right\}\\ &\le n\max_k(P\{\xi(jq) \le u, jq \in I_k\} - P\{M(I_k) \le u\})\\ &\le n\mu(h-\varepsilon)\rho_a + no(\mu) \to \tau\left(1 - \frac{\varepsilon}{h}\right)\rho_a \le \tau\rho_a,\end{aligned}$$

where $\rho_a = 1 - H_\alpha(a)/H_\alpha \to 0$ as $a \to 0$. □

Lemma 12.3.3. *Let $r(t) \to 0$ as $t \to \infty$, and suppose that, as $u^{2/\alpha}q \to a > 0$, (12.3.2) holds for each $\varepsilon > 0$. Then, as $T \to \infty$, $u^{2/\alpha}q \to a$,*

(i) $$P\left\{\xi(jq) \le u, jq \in \bigcup_1^n I_k\right\} - \prod_{k=1}^n P\{\xi(jq) \le u, jq \in I_k\} \to 0,$$

(ii) $$\limsup_{u\to\infty}\left|\prod_{k=1}^n P\{\xi(jq) \le u, jq \in I_k\} - P^n\{M(h) \le u\}\right| \le \tau\left(\rho_a + \frac{\varepsilon}{h}\right),$$

where $\rho_a \to 0$ as $a \to 0$.

PROOF. The proof of part (i) is identical to that of Lemma 8.2.4(i). As for part (ii),

$$0 \le \prod_{k=1}^{n} P\{\xi(jq) \le u, jq \in I_k\} - \prod_{k=1}^{n} P\{M(I_k) \le u\}$$
$$\to \tau\left(1 - \frac{\varepsilon}{h}\right)\rho_a \le \tau\rho_a$$

as in the proof of Lemma 12.3.2(ii), which does not use dependence or independence between variables in different intervals. Furthermore, by stationarity

$$\begin{aligned} 0 &\le \prod_{k=1}^{n} P\{M(I_k) \le u\} - P^n\{M(h) \le u\} \\ &= P^n\{M(I_1) \le u\} - P^n\{M(h) \le u\} \\ &\le n(P\{M(I_1) \le u\} - P\{M(h) \le u\}) \\ &\le nP\{M(I_1) > u\} \sim n\mu\varepsilon \to \tau \cdot \frac{\varepsilon}{h}. \end{aligned}$$ □

Theorem 12.3.4. *Let $\{\xi(t)\}$ be a stationary normal process with zero mean and suppose $r(t)$ satisfies* (12.1.1) *and* (8.1.2), *i.e.*

$$r(t) = 1 - C|t|^\alpha + o(|t|^\alpha) \quad \text{as } t \to 0$$

and

$$r(t) \log t \to 0 \quad \text{as } t \to \infty.$$

If $u = u_T \to \infty$ so that $T\mu = TC^{1/\alpha}H_\alpha u^{2/\alpha}\phi(u)/u \to \tau > 0$, then

$$P\{M(T) \le u\} \to e^{-\tau} \quad \text{as } T \to \infty.$$

PROOF. By Lemma 12.3.1, condition (12.3.2) of Lemma 12.3.3 is satisfied, and by Lemmas 12.3.2 and 12.3.3 we then have

$$\limsup_{u \to \infty} |P\{M(nh) \le u\} - P^n\{M(h) \le u\}| \le 2\tau\left(\rho_a + \frac{\varepsilon}{h}\right),$$

where $\rho_a \to 0$ as $a \to 0$. Letting $\varepsilon \to 0$ and $a \to 0$ this shows that

$$\lim_{u \to \infty} (P\{M(nh) \le u\} - P^n\{M(h) \le u\}) = 0.$$

By Theorem 12.2.9, $P\{M(h) \le u\} = 1 - \mu h + o(\mu)$ and hence, as $\mu \sim \tau/T$, $n \sim T/h$,

$$P^n\{M(h) \le u\} = (1 - \mu h + o(\mu))^n \to e^{-\tau}.$$

Since furthermore

$$M(nh) \le M(T) \le M((n+1)h),$$

this proves the theorem. □

As is easily checked, the choice $u_T = x/a_T + b_T$, with a_T and b_T given by (12.1.2), satisfies $T\mu \to \tau = e^{-x}$, cf. (12.3.1), and we immediately have the following theorem.

Theorem 12.3.5. *Suppose $\{\xi(t)\}$ satisfies the conditions of Theorem 12.3.4, and that, with H_α as in Remark 12.2.10,*

$$a_T = (2\log T)^{1/2},$$

$$b_T = (2\log T)^{1/2} + \frac{1}{(2\log T)^{1/2}} \times \left\{\frac{2-\alpha}{2\alpha}\log\log T + \log(C^{1/\alpha}H_\alpha(2\pi)^{-1/2}2^{(2-\alpha)/2\alpha})\right\}.$$

Then $P\{a_T(M(T) - b_T) \le x\} \to \exp(-e^{-x})$ as $T \to \infty$.

12.4. Asymptotic Properties of ε-upcrossings

As mentioned in Section 12.1, the asymptotic Poisson character of upcrossings applies also to nondifferentiable normal processes, if one considers ε-upcrossings instead of ordinary upcrossings. To prove this, we need to evaluate the expectation of $N_{\varepsilon,u}(T)$, the number of ε-upcrossings of u by $\xi(t)$, $0 \le t \le T$.

Lemma 12.4.1. *Suppose $r(t)$ satisfies (12.1.1). Then, with h as in Theorem 12.2.9, with $\varepsilon = h/2$,*

$$\lim_{u\to\infty} \frac{E(N_{\varepsilon,u}(1))}{u^{2/\alpha}\phi(u)/u} = C^{1/\alpha}H_\alpha.$$

Proof. Write

$$A = \{\xi(t) > u \text{ for some } t \in (-\varepsilon, 0]\},$$

$$B = \{\xi(t) > u \text{ for some } t \in (0, \varepsilon)\}.$$

From Theorem 12.2.9 we have, for $2\varepsilon \le h$, as $u \to \infty$,

$$P(A \cup B) \sim 2\varepsilon C^{1/\alpha}H_\alpha u^{2/\alpha-1}\phi(u),$$

$$P(A) \sim \varepsilon C^{1/\alpha}H_\alpha u^{2/\alpha-1}\phi(u),$$

$$P(B) \sim \varepsilon C^{1/\alpha}H_\alpha u^{2/\alpha-1}\phi(u).$$

Hence

$$P(B \cap A^c) = P(A \cup B) - P(A) \sim \varepsilon C^{1/\alpha} H_\alpha u^{2/\alpha - 1} \phi(u),$$

and

$$P(B \cap A^c) \leq P\{N_{\varepsilon,u}(\varepsilon) = 1\} = E(N_{\varepsilon,u}(\varepsilon)) \leq P(B),$$

since there is at most one ε-upcrossing in $[0, \varepsilon]$. Hence

$$E(N_{\varepsilon,u}(\varepsilon)) \sim \varepsilon C^{1/\alpha} H_\alpha u^{2/\alpha - 1} \phi(u)$$

and thus

$$E(N_{\varepsilon,u}(1)) = \frac{1}{\varepsilon} E(N_{\varepsilon,u}(\varepsilon)) \sim C^{1/\alpha} H_\alpha u^{2/\alpha - 1} \phi(u)$$

as required. □

In particular, the lemma implies that asymptotically the mean number of ε-upcrossings of a suitably increasing level is independent of the choice of $\varepsilon > 0$, ($\varepsilon \leq h/2$), and this leads us directly to the following Poisson result obtained in Lindgren *et al.* (1975). Let N_T^* be the point process on $(0, \infty)$ defined by

$$N_T^*(B) = N_{\varepsilon,u}(T \cdot B),$$

where the level u is chosen so that $T\mu = TC^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u \sim \tau > 0$, and let N be a Poisson process with intensity τ.

Theorem 12.4.2. *Suppose that the assumptions of Theorem* 12.3.4 *are satisfied. Then the time-normalized point process N_T^* of ε-upcrossings of the level u converges in distribution to N as $u \to \infty$, where N is a Poisson process with intensity τ on the positive real line.*

Proof. As in the proof of Theorem 9.1.2 we only have to check that for $0 < c < d$

(a) $\lim_{T \to \infty} E(N_T^*((c, d])) = E(N((c, d])) = \tau(d - c)$, and if $R_i = (c_i, d_i]$ (disjoint), $U = \bigcup_{i=1}^m R_i$, then

(b) $$P\{N_T^*(U) = 0\} \to P\{N(U) = 0\} = \prod_{i=1}^m e^{-\tau m(R_i)}.$$

By Lemma 12.4.1, $E(N_T^*((c, d])) = E(N_{\varepsilon,u}((Tc, Td])) = TE(N_{\varepsilon,u}((c, d])) \sim T(d - c)\mu \sim \tau(d - c)$, which proves (a). For part (b) the same steps as in the proof of Theorem 9.1.2 go through, with only obvious changes. □

In previous chapters we have encountered a variety of results, related to the Poisson convergence of upcrossings of an increasing level. There are no further difficulties in extending these results to cover ε-upcrossings. However, we do not want to lengthen an already long journey over an ocean of lemmas. We mention that a little further generality may be obtained throughout by including a function of slow growth (or perhaps slow decrease) instead of C in (12.1.1). This has been considered by Berman (1971b), and also by Qualls and Watanabe (1972).

12.5. Weaker Conditions at Infinity

As already noted, the above extremal results may also be generalized by weakening the condition (8.1.2), which describes the behaviour of the correlation at distant points. We shall proceed as in the discrete case (Section 4.5), following Leadbetter *et al.* (1978) and Mittal (1979). Of course, we cannot expect a substantial weakening of (8.1.2) since it is clearly close to being a necessary condition.

Let $h(t)$ be any function and define

$$\begin{aligned} \theta_T(h) &= \{t \in (0, T]; |r(t)| \log t > h(t)\}, \\ l_T(h) &= \text{Lebesgue measure of } \theta_T(h). \end{aligned} \tag{12.5.1}$$

By analogy with the conditions for discrete time we will place restrictions on the amount of time that $|r(t)| \log t$ is large by requiring that there is some nonincreasing function h with $h(t) \downarrow 0$ as $t \uparrow \infty$ such that

$$l_T(h) = O(T/(\log T)^\gamma) \quad \text{for some } \gamma > \max(0, 1 - 1/\alpha) \tag{12.5.2}$$

and some constant $K > 0$ such that

$$l_T(K) = O(T^\eta) \quad \text{for some } \eta < 1. \tag{12.5.3}$$

Obviously, the condition $r(t) \log t \to 0$ as $t \to \infty$ implies that $\theta_T(h)$ is empty if, e.g. $h(t) = \sup_{s \geq t} |r(s)| \log s$, so that (12.5.2) is actually weaker than (8.1.2); see Mittal (1979) for examples. In fact, (12.5.2) is also weaker than some other conditions which have been used on occasions. For example, since $\int_0^T |r(t)|^p \, dt \geq l_T(h)(h(T)/\log T)^p$ if h is decreasing, $\int_0^\infty r^2(t) \, dt < \infty$ implies that $l_T(h) = O((\log T/h(T))^2)$ for all h, so that (12.5.2) is indeed weaker than the condition $\int_0^\infty r^2(t) \, dt < \infty$, sometimes used in the literature.

Theorem 12.5.1. *Let $\xi(t)$ be a (zero mean) stationary normal process with covariance function $r(t) \to 0$ as $t \to \infty$, and satisfying* (12.1.1), (12.5.2), *and* (12.5.3). *Let $u = u_T \to \infty$ so that $T\mu \to \tau > 0$. Then*

$$P\{M(T) \leq u\} \to e^{-\tau} \quad \text{as } T \to \infty.$$

Proof. In the proof of Theorem 12.3.4 the condition (8.1.2) was used exclusively to prove (12.3.2). Thus, to prove the theorem, let $u \to \infty$ as $T \to \infty$ so that $T \sim \tau/\mu$ with $\tau > 0$ fixed and $\mu = C^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u$, and take $q \to 0$ so that $u^{2/\alpha} q \to a > 0$. Then, as we shall see, (12.1.1) and (12.5.1)–(12.5.3) imply that for $\varepsilon > 0$,

$$\frac{T}{q} \sum_{\varepsilon \leq kq \leq T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0.$$

as $T \to \infty$, proving the theorem.

Let $\rho(t) = \sup_{s \geq t} |r(s)|$, let β satisfy $0 < \beta < (1 - \delta)/(1 + \delta)$ for $\delta = \rho(\varepsilon)$, and split the sum in (12.3.2) into two parts at T^β, i.e., let $\sum'$ be the sum over $\varepsilon \leq kq \leq T^\beta$ and $\sum''$ the sum over $T^\beta < kq \leq T$. As in the proof of Lemma 12.3.1

$$\frac{T}{q}\sum{}' = \frac{T}{q}\sum_{\varepsilon \leq kq \leq T^\beta} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\} \to 0$$

if $qu^{2/\alpha} \to a > 0$.

For the remaining sum $\sum''$ we need a bound on the number of terms for which $|r(kq)| \log kq$ is not bounded by a small function. Define, for a function h,

$$n_T(h) = \#\{k;\, T^\beta < kq \leq T, |r(kq)| \log kq > h(kq)\}$$

in analogy with $l_T(h)$ in (12.5.1).

Since $r(t)$ satisfies a Lipschitz condition at 0 it does so uniformly for all t. In fact, if $\alpha' < \min(1, \alpha)$ then

$$|r(t + s) - r(t)| \leq C|s|^{\alpha'},$$

for some constant C, see Boas (1967, Theorem 1). We will use this to give a bound for $n_T(h)$ in terms of $l_T(h/2)$. Let γ be as in condition (12.5.2) and take α' such that $\alpha/(1 + \gamma\alpha) < \alpha' < \min(1, \alpha)$. Note that we can always find such an α' and that $1/\alpha' - 1/\alpha - \gamma < 0$. We will show that for all nonincreasing functions h,

$$n_T(h) \leq C'\left(\frac{\log T}{h(T)}\right)^{1/\alpha'} l_T\left(\frac{h}{2}\right), \tag{12.5.4}$$

if T is large enough. Since, for $t \geq kq > 1$, $|r(t)| \log t > (|r(kq)| - C|t - kq|^{\alpha'}) \log kq$ we see that if

$$|r(kq)| \log kq > h(kq)$$

and $t \leq T$ is such that

$$kq < t < kq + \left(\frac{h(T)}{2C \log T}\right)^{1/\alpha'}$$

then

$$|r(t)| \log t > \frac{h(t)}{2}.$$

We have $q \sim a/u^{2/\alpha}$, $u \sim (2 \log T)^{1/2}$ and thus

$$(h(T)/\log T)^{1/\alpha'}/q \sim \frac{h(T)^{1/\alpha'}}{a} (\log T)^{-1/\alpha' + 1/\alpha} \to 0$$

since $\alpha > \alpha'$. This implies that for T large enough the kq which contribute to $n_T(h)$ also contribute disjoint intervals of length $(h(T)/(2C \log T))^{1/\alpha'}$ to $l_T(h/2)$, and we get (12.5.4) with $C' = (2C)^{1/\alpha'}$.

We can now proceed by splitting the sum $\sum''$ according to whether $kq \in \theta_T(2K)$ or not. Recalling the notation $\rho(t) = \sup_{s \geq t} |r(s)|$, we have

$$\begin{aligned}
\frac{T}{q}\sum{}'' &= \frac{T}{q}\sum_{T^\beta < kq \leq T} |r(kq)| \exp\left\{-\frac{u^2}{1+|r(kq)|}\right\} \\
&\leq \frac{T}{q} n_T(2K) \exp\left\{-\frac{u^2}{1+\rho(T^\beta)}\right\} \\
&\quad + \frac{T}{q}\sum_{T^\beta < kq \leq T,\, kq \in \theta_T(2K)^c} |r(kq)| \exp\left\{-u^2\left(1 - \frac{2K}{\log T^\beta}\right)\right\} \qquad (12.5.5)
\end{aligned}$$

(where the c denotes complement). For large T, the bound (12.5.4) applies to $n_T(2K)$, and therefore the first term in (12.5.5) is

$$C'\frac{T}{q}\left(\frac{\log T}{2K}\right)^{1/\alpha'} l_T(K) T^{-2/(1+\rho(T^\beta))} = \frac{C'}{u^{2/\alpha}q}(\log T)^{1/\alpha'}$$

as $T \to \infty$, since $\eta < 1$ by (12.5.3) and since $\rho(T^\beta) \to 0$.

The second term in (12.5.5) is bounded by

$$\left(\frac{T}{q}\right)^2 \exp\left\{-u^2\left(1 - \frac{2K}{\beta \log T}\right)\right\} \frac{1}{\beta \log T} \cdot \frac{q}{T}\sum |r(kq)| \log kq = F_1 \cdot F_2, \qquad (12.5.6)$$

say, where the sum is extended over all kq such that $T^\beta < kq \leq T$ and $kq \in \theta_T(2K)^c$. We will see that F_1 is bounded and that $F_2 \to 0$ as $T \to \infty$ so that $F_1 \cdot F_2 \to 0$. We start with F_2, introducing the function h that appears in (12.5.2) and split the sum according to whether $kq \in \theta_T(2h)$ or not, giving

$$\begin{aligned}
F_2 &= \frac{q}{T}\sum |r(kq)| \log kq \\
&\leq \frac{q}{T}\sum_{\substack{kq \in \theta_T(2h)^c \\ kq \leq T}} + \frac{q}{T}\sum_{kq \in \theta_T(2h) \cap \theta_T(2K)^c} \\
&\leq \frac{q}{T}\cdot\frac{T}{q} 2h(T^\beta) + \frac{q}{T}\cdot 2K n_T(2h) \\
&\leq 2h(T^\beta) + 2KC'\frac{q}{T}\left(\frac{\log T}{h(T)}\right)^{1/\alpha'} l_T(h) \\
&= 2h(T^\beta) + h(T)^{-1/\alpha'}(\log T)^{1/\alpha' - 1/\alpha - \gamma}(u^{2/\alpha}q)O(1)
\end{aligned}$$

say, by condition (12.5.2). Since $1/\alpha' - 1/\alpha - \gamma < 0$, we can deduce that $F_2 \to 0$ as $T \to \infty$, provided $h(t)$ decreases sufficiently slowly. Note that if (12.5.2) is satisfied for some function h, then it is satisfied for all functions which decrease more slowly. The remaining factor F_1 in (12.5.6) is given by

$$F_1 = \left(\frac{T}{q}\right)^2 \exp\left\{-u^2\left(1 - \frac{2K}{\beta \log T}\right)\right\} \frac{1}{\beta \log T}.$$

Using the fact that $u^2 = 2\log T + 2(1/\alpha - \frac{1}{2})\log\log T + O(1)$ we obtain

$$\frac{O(1)}{q^2 \log T} \exp\left\{-2\left(\frac{1}{\alpha} - \frac{1}{2}\right)\log\log T\right\} = \frac{O(1)}{q^2(\log T)^{2/\alpha}} = O(1).$$

Thus, $F_1 \cdot F_2 \to 0$, and we have proved that also the second term in (12.5.5) tends to zero, which completes the proof of the theorem. □

Remark 12.5.2. As in discrete time, one would be inclined to consider a condition like

$$\frac{q}{T} \sum_{T^\beta < kq \le T} |r(kq)| \log kq \exp\{\gamma |r(kq)| \log kq\} \to 0 \tag{12.5.7}$$

as $T \to \infty$, for some $\beta < 1, \gamma > 2$ which in fact can replace (12.5.2). However, (12.5.7) contains the somewhat arbitrary spacing q, and a more natural condition for a continuous time process would restrict the size of

$$\int_1^T |r(t)| \log t \exp\{\gamma |r(t)| \log t\}\, dt.$$

However, it is not clear how this might be done, in relation to (12.5.7). □

CHAPTER 13

Extremes of Continuous Parameter Stationary Processes

Our primary task in this chapter will be to discuss continuous parameter analogues of the sequence results of Chapter 3, and, in particular, to obtain a corresponding version of the Extremal Types Theorem which applies in the continuous parameter case. This will be taken up in the first section, using a continuous parameter analogue of the dependence restriction $D(u_n)$. Limits for probabilities $P\{M(T) \leq u_T\}$ are then considered for arbitrary families of constants $\{u_T\}$, leading, in particular, to a determination of domains of attraction.

The theory is applied in two cases—first to normal processes, providing an alternative approach to the derivation of the results of Chapter 12, and then to stationary processes with finite upcrossing intensities. Finally, for this latter class, general Poisson results are obtained for upcrossings of high levels, giving, as applications, asymptotic distributions (and joint distributions) of kth largest local maxima.

13.1. The Extremal Types Theorem

Throughout this chapter we consider a (strictly) stationary process $\{\xi(t); t \geq 0\}$ satisfying the general conditions stated at the start of Chapter 7. In particular, it will be assumed that $\xi(t)$ has a.s. continuous sample functions, continuous one-dimensional distributions, and that the underlying probability space is complete. As shown in Lemma 7.1.1, it then follows that $M(I) = \sup\{\xi(t); t \in I\}$ is a r.v. for any interval I and, in particular, so is $M(T) = M((0, T])$.

Our main interest in this section concerns asymptotic distributional properties of $M(T)$, and especially what forms of nondegenerate limiting distribution are possible, in the sense that $P\{a_T(M(T) - b_T) \le x\}$ converges to a nondegenerate d.f. G for some constants $a_T > 0, b_T$, as $T \to \infty$. Following Leadbetter and Rootzen (1982), we shall find that interesting forms of the Extremal Types Theorem hold under natural continuous parameter analogues of the dependence restriction used in the discrete case. In fact our approach to the continuous situation is to relate it to the discrete case as in Chapter 8 by considering a sequence of "submaxima". Specifically, for some $h > 0$ (to be conveniently chosen) let

$$\zeta_i = \sup\{\xi(t); (i-1)h \le t \le ih\}, \tag{13.1.1}$$

so that for any $n = 1, 2, \ldots$ we have

$$M(nh) = \max(\zeta_1, \zeta_2, \ldots, \zeta_n). \tag{13.1.2}$$

It is apparent that the properties of $M(T)$ as $T \to \infty$ may be obtained from those of $M(nh)$ by writing $n = [T/h]$ and thus approximating T by nh.

As noted above, we shall consider a continuous parameter analogue (to be called $C(u_T)$) of the condition $D(u_n)$, used for sequences. The condition $C(u_T)$ will be used in ensuring that the stationary sequence $\{\zeta_n\}$ defined by (13.1.1) satisfies $D(u_n)$. However, before introducing this condition we note a preliminary form of the Extremal Types Theorem which simply *assumes* that the sequence $\{\zeta_n\}$ satisfies $D(u_n)$. This result follows immediately from the sequence case and clearly illustrates the central ideas required in the continuous parameter context. The more complete version (Theorem 13.1.5) to be given later, of course simply requires finding appropriate conditions, of which the main one will be $C(u_T)$, on $\xi(t)$, to guarantee that $\{\zeta_n\}$ will satisfy $D(u_n)$.

Theorem 13.1.1. *Suppose that for some families of constants $\{a_T > 0\}$, $\{b_T\}$,*

$$P\{a_T(M(T) - b_T) \le x\} \xrightarrow{w} G(x) \quad \text{as } T \to \infty \tag{13.1.3}$$

for some nondegenerate G, and that the $\{\zeta_n\}$ sequence defined by (13.1.1) *satisfies $D(u_n)$ whenever $u_n = x/a_{nh} + b_{nh}$ for some fixed $h > 0$ and all real x. Then G is one of the three extreme value types.*

PROOF. Since (13.1.3) holds, in particular, as $T \to \infty$ through values nh and the ζ_n-sequence is clearly stationary, the result follows by replacing ξ_n by ζ_n in Theorem 3.3.3 and using (13.1.2). □

Although we shall not make further use of the fact, it is interesting to note that this at once implies that the Extremal Types Theorem holds under "strong mixing" assumptions as the following corollary shows.

Corollary 13.1.2. *Theorem* 13.1.1 *holds, in particular, if the $D(u_n)$ condition is replaced by the assumption that $\{\xi(t)\}$ is strongly mixing. For then the sequence $\{\zeta_n\}$ is strongly mixing and hence satisfies $D(u_n)$.*

We now introduce the continuous analogue of the condition $D(u_n)$, stated in terms of the finite-dimensional distribution functions $F_{t_1 \cdots t_n}$ of $\xi(t)$, again writing $F_{t_1 \cdots t_n}(u)$ for $F_{t_1 \cdots t_n}(u, \ldots, u)$. The points t_i will be members of a discrete set $\{jq_T; j = 1, 2, 3, \ldots\}$ where $\{q_T\}$ is a family of constants tending to zero as $T \to \infty$ at a rate to be specified later.

The condition *$C(u_T)$ will be said to hold for the process $\xi(t)$ and the family of constants $\{u_T; T > 0\}$, with respect to the constants $q_T \to 0$, if for any points $s_1 < s_2 < \cdots < s_p < t_1 < \cdots < t_{p'}$ belonging to $\{kq_T; 0 \le kq_T \le T\}$ and satisfying $t_1 - s_p \ge \gamma$, we have*

$$|F_{s_1 \cdots s_p, t_1 \cdots t_{p'}}(u_T) - F_{s_1 \cdots s_p}(u_T) F_{t_1 \cdots t_{p'}}(u_T)| \le \alpha_{T,\gamma} \tag{13.1.4}$$

where $\alpha_{T,\gamma_T} \to 0$ for some family $\gamma_T = o(T)$, as $T \to \infty$.

As in the discrete case we may (and do) take $\alpha_{T,\gamma}$ to be nonincreasing as γ increases and also note that the condition $\alpha_{T,\gamma_T} \to 0$ for some $\gamma_T = o(T)$ may be replaced by

$$\alpha_{T,\lambda T} \to 0 \quad \text{as } T \to \infty \tag{13.1.5}$$

for each fixed $\lambda > 0$.

The $D(u_n)$ condition for $\{\zeta_n\}$ required in Theorem 13.1.1 will now be related to $C(u_T)$ by approximating crossings and extremes of the continuous parameter process, by corresponding quantities for a "sampled version". To achieve the approximation we require two conditions involving the maximum of $\xi(t)$ in fixed (small) time intervals. These conditions are given here in a form which applies very generally—readily verifiable sufficient conditions for important cases are given later in this chapter.

It will be convenient to introduce a function $\psi(u)$ which will generally describe the form of the tail of the distribution of the maximum $M(h)$ in a fixed interval $(0, h]$ as u becomes large. Specifically as needed we shall make one or more of the following successively stronger assumptions:

$$P\{\xi(0) > u\} = o(\psi(u)), \tag{13.1.6}$$

$$P\{M(q) > u\} = o(\psi(u)) \quad \text{for any } q = q(u) \to 0 \quad \text{as } u \to \infty, \tag{13.1.7}$$

there exists $h_0 > 0$ such that

$$\limsup_{u \to \infty} \frac{P\{M(h) > u\}}{h\psi(u)} \le 1 \quad \text{for } 0 < h \le h_0, \tag{13.1.8}$$

$$P\{M(h) > u\} \sim h\psi(u) \quad \text{as } u \to \infty \quad \text{for } 0 < h \le h_0. \tag{13.1.9}$$

Note that equation (13.1.9) commonly holds and specifies that the tail of the distribution of $M(h)$ is asymptotically proportional to $\psi(u)$, whereas

(13.1.8) is a weaker assumption which is sometimes convenient as a sufficient condition for the yet weaker (13.1.7) and (13.1.6). As we shall see later, $\psi(u)$ can also be identified with the mean number of upcrossings of the level u per unit time, $\mu(u)$, in important cases when this is finite. In any case it is, of course, possible to define $\psi(u)$ to be $P\{M(h_0) > u\}/h_0$ for some fixed $h_0 > 0$, or some asymptotically equivalent function and then attempt to verify any of the above conditions which may be needed.

We shall also require an assumption relating "continuous and discrete" maxima in fixed intervals. Specifically we assume, as required, that for each $a > 0$ there is a family of constants $\{q\} = \{q_a(u)\}$ tending to zero as $u \to \infty$ for each $a > 0$, such that for any fixed $h > 0$,

$$\limsup_{u\to\infty} \frac{P\{M(h) > u, \xi(jq) \le u, 0 \le jq \le h\}}{\psi(u)} \to 0 \quad \text{as } a \to 0. \tag{13.1.10}$$

Finally a condition which is sometimes helpful in verifying (13.1.10) is

$$\limsup_{u\to\infty} \frac{P\{\xi(0) \le u, \xi(q) \le u, M(q) > u\}}{q\psi(u)} \to 0 \quad \text{as } a \to 0. \tag{13.1.11}$$

Here the constant a specifies the rate of convergence to zero of $q_a(u)$—as a decreases, the grid of points $\{q_a(u)\}$ tends to become (asymptotically) finer, and for small a the maximum of $\xi(t)$ on the discrete grid approximates the continuous maximum well, as will be seen below. (Simpler versions of (13.1.10) and (13.1.11) would be to assume the existence of *one* family $q = q(u)$ of constants such that the upper limits in (13.1.10) and (13.1.11) are zero for this family. It can be seen that one can do this without loss of generality in the theorems below, but it seems that, as was the case in Chapter 12, the conditions involving the parameter a may often be easier to check.)

The following lemma contains some simple but useful relationships.

Lemma 13.1.3. (i) *If* (13.1.8) *holds, so does* (13.1.7) *which in turn implies* (13.1.6). *Hence* (13.1.9) *clearly implies* (13.1.8), (13.1.7), *and* (13.1.6).

(ii) *If I is any interval of length h and* (13.1.6) *and* (13.1.10) *both hold, then there are constants λ_a such that*

$$0 \le \limsup_{u\to\infty} \frac{P\{\xi(jq) \le u, jq \in I\} - P\{M(I) \le u\}}{\psi(u)} \ge \lambda_a \to 0 \tag{13.1.12}$$

as $a \to 0$, where $q = q_a(u)$ is as in (13.1.10), *the convergence being uniform in all intervals of this fixed length h.*

(iii) *If* (13.1.7) *and* (13.1.11) *hold, so does* (13.1.10) *and hence, by* (ii) *so does* (13.1.12).

(iv) *If* (13.1.9) *holds and $I_1 = (0, h]$, $I_2 = (h, 2h]$ with $0 < h \le h_0/2$, then $P\{M(I_1) > u, M(I_2) > u\} = o(\psi(u))$ as $u \to \infty$.*

PROOF. (i) If (13.1.8) holds and $q \to 0$ as $u \to \infty$, then for any fixed $h > 0$, q is eventually smaller than h and $P\{M(q) > u\} \le P\{M(h) > u\}$, so that

$$\limsup_{u\to\infty} \frac{P\{M(q) > u\}}{\psi(u)} \le \limsup_{u\to\infty} \frac{P\{M(h) > u\}}{\psi(u)} \le h$$

by (13.1.8). Since h is arbitrary it follows that $P\{M(q) > u\}/\psi(u) \to 0$, giving (13.1.7). It is clear that (13.1.7) implies (13.1.6) since

$$P\{\xi(0) > u\} \le P\{M(q) > u\},$$

which proves (i).

To prove (ii) we assume that (13.1.6) and (13.1.10) hold and let I be an interval of fixed length h. Since the numbers of points jq in I and in $(0, h]$ differ by at most 2, it is readily seen from stationarity that

$$P\{\xi(jq) \le u, jq \in I\} \le P\{\xi(jq) \le u, 0 \le jq \le h\} + P\{\xi(0) > u\} + P\{\xi(h) > u\}$$

so that

$$\begin{aligned} 0 &\le P\{\xi(jq) \le u, jq \in I\} - P\{M(I) \le u\} \\ &\le P\{\xi(jq) \le u, 0 \le jq \le h\} - P\{M(h) \le u\} + 2P\{\xi(0) > u\} \\ &= P\{M(h) > u, \xi(jq) \le u, 0 \le jq \le h\} + 2P\{\xi(0) > u\} \end{aligned}$$

from which (13.1.12) follows at once by (13.1.6) and (13.1.10), so that (ii) follows.

To prove (iii) we note that there are at most $[h/q]$ complete intervals $((j-1)q, jq]$ in $(0, h]$ with perhaps a smaller interval remaining so that

$$P\{M(h) > u, \xi(jq) \le u, 0 \le jq \le h\} \le \frac{h}{q} P\{\xi(0) \le u, \xi(q) \le u, M(q) > u\} + P\{M(q) > u\}$$

so that (13.1.10) easily follows from (13.1.11) and (13.1.7).

Finally if (13.1.9) holds and $I_1 = (0, h]$, $I_2 = (h, 2h]$ with $0 < h \le h_0/2$, then

$$P\{M(I_2) > u\} = P\{M(I_1) > u\} = h\psi(u)(1 + o(1))$$

and

$$P(\{M(I_1) > u\} \cup \{M(I_2) > u\}) = P\{M(I_1 \cup I_2) > u\} = 2h\psi(u)(1 + o(1))$$

so that

$$\begin{aligned} P\{M(I_1) > u, M(I_2) > u\} &= P\{M(I_1) > u\} + P\{M(I_2) > u\} \\ &\quad - P(\{M(I_1) > u\} \cup \{M(I_2) > u\}) \\ &= o(\psi(u)) \end{aligned}$$

as required. □

For $h > 0$, let $\{T_n\}$ be a sequence of time points such that $T_n \in (nh, (n+1)h]$ and write $v_n = u_{T_n}$. It is then relatively easy to relate $D(v_n)$ for the sequence $\{\zeta_n\}$ to the condition $C(u_T)$ for the process $\xi(t)$, as the following lemma shows.

Lemma 13.1.4. *Suppose that* (13.1.6) *holds for some function* $\psi(u)$ *and let* $\{q_a(u)\}$ *be a family of constants for each* $a > 0$ *with* $q_a(u) > 0$, $q_a(u) \to 0$ *as* $u \to \infty$, *and such that* (13.1.10) *holds. If* $C(u_T)$ *is satisfied with respect to the family* $q_T = q_a(u_T)$ *for each* $a > 0$, *and* $T\psi(u_T)$ *is bounded, then the sequence* $\{\zeta_n\}$ *defined by* (13.1.1) *satisfies* $D(v_n)$, *where* $v_n = u_{T_n}$ *is as above.*

PROOF. For a given n, let $i_1 < i_2 < \cdots < i_p < j_1 < \cdots < j_{p'} < n$, $j_1 - i_p \geq l$. Write $I_r = ((i_r - 1)h, i_r h]$, $J_s = ((j_s - 1)h, j_s h]$. For brevity write q for the elements in one of the families $\{q_a(\cdot)\}$ and let

$$A_q = \bigcap_{r=1}^{p} \{\xi(jq) \leq v_n, jq \in I_r\}, \quad A = \bigcap_{r=1}^{p} \{\zeta_{i_r} \leq v_n\},$$

$$B_q = \bigcap_{s=1}^{p'} \{\xi(jq) \leq v_n, jq \in J_s\}, \quad B = \bigcap_{s=1}^{p'} \{\zeta_{j_s} \leq v_n\}.$$

It follows in an obvious way from Lemma 13.1.3(ii) that

$$\begin{aligned} 0 &\leq \limsup_{n\to\infty} \{P(A_q B_q) - P(AB)\} \leq \limsup_{n\to\infty} (p + p')\psi(v_n)\lambda_a \\ &\leq \limsup_{n\to\infty} n\psi(v_n)\lambda_a \leq K\lambda_a \end{aligned}$$

for some constant K (since $nh \sim T_n$ and $T_n\psi(v_n)$ is bounded) and where $\lambda_a \to 0$ as $a \to 0$. Similarly,

$$\limsup_{n\to\infty} |P(A_q) - P(A)| \leq K\lambda_a, \qquad \limsup_{n\to\infty} |P(B_q) - P(B)| \leq K\lambda_a.$$

Now

$$\begin{aligned} |P(A \cap B) - P(A)P(B)| &\leq |P(A \cap B) - P(A_q \cap B_q)| \\ &\quad + |P(A_q \cap B_q) - P(A_q)P(B_q)| + P(A_q)|P(B_q) - P(B)| \\ &\quad + P(B)|P(A_q) - P(A)| \\ &= R_{n,a} + |P(A_q \cap B_q) - P(A_q)P(B_q)|, \end{aligned} \tag{13.1.13}$$

where $\limsup_{n\to\infty} R_{n,a} \leq 3K\lambda_a$.

Since the largest jq in any I_r is at most $i_p h$, and the smallest in any J_s is at least $(j_1 - 1)h$, their difference is at least $(l-1)h$. Also the largest jq in $J_{p'}$ does not exceed $j_{p'}h \leq nh \leq T_n$ so that from (13.1.4) and (13.1.13),

$$|P(A \cap B) - P(A)P(B)| \leq R_{n,a} + \alpha^{(a)}_{T_n,(l-1)h}, \tag{13.1.14}$$

in which the dependence of $\alpha_{T,l}$ on a is explicitly indicated. Write now $\alpha^*_{n,l} = \inf_{a>0} \{R_{n,a} + \alpha^{(a)}_{T_n,(l-1)h}\}$. Since the left-hand side of (13.1.14) does not depend on a we have

$$|P(A \cap B) - P(A)P(B)| \leq \alpha^*_{n,l},$$

which is precisely the desired conclusion of the lemma, provided we can show that $\lim_{n\to\infty} \alpha^*_{n,[\lambda n]} = 0$ for any $\lambda > 0$ (cf. (3.2.3)). But, for any $a > 0$,

$$\alpha^*_{n,[\lambda n]} \le R_{n,a} + \alpha^{(a)}_{T_n,(\lambda n-1)h} \le R_{n,a} + \alpha^{(a)}_{T_n,\lambda T_n/2}$$

when n is sufficiently large (since $\alpha^{(a)}_{T,l}$ decreases in l), and hence by (13.1.5)

$$\limsup_{n\to\infty} \alpha^*_{n,[\lambda n]} \le 3K\lambda_a,$$

and since a is arbitrary and $\lambda_a \to 0$ as $a \to 0$, it follows that $\alpha^*_{n,[\lambda n]} \to 0$ as desired. □

The general continuous version of the Extremal Types Theorem is now readily restated in terms of conditions on $\xi(t)$ itself.

Theorem 13.1.5. *With the above notation for the stationary process* $\{\xi(t)\}$ *satisfying* (13.1.6) *for some function* ψ, *suppose that there are constants* $a_T > 0$, b_T *such that*

$$P\{a_T(M(T) - b_T) \le x\} \xrightarrow{w} G(x)$$

for a nondegenerate G. *Suppose that* $T\psi(u_T)$ *is bounded and* $C(u_T)$ *holds for* $u_T = x/a_T + b_T$ *for each real* x, *with respect to families of constants* $\{q_a(u)\}$ *satisfying* (13.1.10). *Then* G *is one of the three extreme value distributional types.*

Proof. This follows at once from Theorem 13.1.1 and Lemma 13.1.4, by choosing $T_n = nh$. □

As noted the conditions of this theorem are of a general kind, and more specific sufficient conditions will be given in the applications later in this chapter.

13.2. Convergence of $P\{M(T) \le u_T\}$

The Extremal Types Theorem involved consideration of

$$P\{a_T(M(T) - b_T) \le x\},$$

which may be rewritten as $P\{M(T) \le u_T\}$ with $u_T = x/a_T + b_T$. We turn now to the question of convergence of $P\{M(T) \le u_T\}$ as $T \to \infty$ for families u_T which are not necessarily linear functions of a parameter x. (This is analogous to the convergence of $P\{M_n \le u_n\}$ for sequences, of course.) These results are of interest in their own right, but also since they make it possible to simply modify the classical criteria for domains of attraction to the three limiting distributions, to apply in this continuous parameter context.

Our main purpose is to demonstrate the equivalence of the relations $P\{M(h) > u_T\} \sim \tau/T$ and $P\{M(T) \le u_T\} \to e^{-\tau}$ under appropriate conditions. The following condition will be referred to as $C'(u_T)$ and is analogous to $D'(u_n)$ defined in Chapter 3, for sequences.

The condition *$C'(u_T)$ will be said to hold for the process $\{\xi(t)\}$ and the family of constants $\{u_T; T > 0\}$, with respect to the constants $q = q_T \to 0$ if*

$$\limsup_{T\to\infty} \frac{T}{q} \sum_{h<jq<\varepsilon T} P\{\xi(0) > u_T, \xi(jq) > u_T\} \to 0$$

as $\varepsilon \to 0$, for some $h > 0$.

The following lemma will be useful in obtaining the desired equivalence.

Lemma 13.2.1. *Suppose that* (13.1.9) *holds for some function ψ, and let $\{u_T\}$ be a family of levels such that $C'(u_T)$ holds with respect to families $\{q_a(u)\}$ satisfying* (13.1.10), *for each $a > 0$, with h in $C'(u_T)$ not exceeding $h_0/2$ in* (13.1.9). *Then $T\psi(u_T)$ is bounded, and writing $n' = [n/k]$, for n and k integers,*

$$0 \le \limsup_{n\to\infty}(n'P\{M(h) > v_n\} - P\{M(n'h) > v_n\}) = o(k^{-1}) \quad \text{as } k \to \infty, \tag{13.2.1}$$

with $v_n = u_{T_n}$, for any sequence $\{T_n\}$ with $T_n \in (nh, (n+1)h]$.

Proof. We shall use the extra assumption

$$\liminf_{T\to\infty} T\psi(u_T) > 0, \tag{13.2.2}$$

in proving $T\psi(u_T)$ bounded and (13.2.1). It is then easily checked (e.g. by replacing $T\psi(u_T)$ by $\max(1, T\psi(u_T))$ in the proof) that the result also holds without the extra assumption.

Now, write $I_j = ((j-1)h, jh]$, $j = 1, 2, \ldots$ and $M_k(I) = \max\{\xi(jq; jq \in I\}$, for any interval I. We shall first show that (assuming (13.2.2) holds)

$$0 \le \limsup_{n\to\infty} \frac{1}{T_n\psi(v_n)}(n'P\{M(h) > v_n\} - P\{M(n'h) > v_n\}) = o(k^{-1}) \tag{13.2.3}$$

as $k \to \infty$. The expression in (13.2.3) is clearly non-negative, and by stationarity and the fact that $M \ge M_q$, does not exceed

$$\limsup_{n\to\infty} \frac{1}{T_n\psi(v_n)} \sum_{j=1}^{n'} (P\{M(I_j) > v_n\} - P\{M_q(I_j) > v_n\})$$

$$+ \limsup_{n\to\infty} \frac{1}{T_n\psi(v_n)}\left[\sum_{j=1}^{n'} P\{M_q(I_j) > v_n\} - P\{M_q(n'h) > v_n\}\right].$$

By Lemma 13.1.3(ii), the first of the upper limits does not exceed $\lambda_a \limsup_{n\to\infty} n'/T_n = \lambda_a/(hk)$, where $\lambda_a \to 0$ as $a \to 0$. The expression in the second upper limit may be written as

$$\begin{aligned}
&\frac{1}{T_n\psi(v_n)}\left[\sum_{j=1}^{n'} P\{M_q(I_j) > v_n\} - \sum_{j=1}^{n'} P\left\{M_q(I_j) > v_n, M_q\left(\bigcup_{l=j+1}^{n'} I_l\right) \le v_n\right\}\right] \\
&\quad\le \frac{1}{T_n\psi(v_n)} \sum_{j=1}^{n'} P\{M_q(I_j) > v_n, M_q(I_{j+1}) > v_n\} \\
&\quad\quad + \frac{1}{T_n\psi(v_n)} \sum_{j=1}^{n'} P\left\{M_q(I_j) > v_n, M_q\left(\bigcup_{l=j+2}^{n'} I_l\right) > v_n\right\} \\
&\quad\le \frac{n'}{T_n} o(1) + \frac{n'h}{qT_n\psi(v_n)} \sum_{h \le jq \le n'h} P\{\xi(0) > v_n, \xi(jq) > v_n\},
\end{aligned}$$

by Lemma 13.1.3(iv) and some obvious estimation using stationarity. By $C'(u_T)$, using (13.2.2), the upper limit (over n) of the last term is readily seen to be $o(k^{-1})$ for each $a > 0$, and (13.2.3) now follows by gathering these facts.

Further, by (13.2.3) and (13.1.9)

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{T_n\psi(v_n)} &\ge \liminf_{n\to\infty} \frac{1}{T_n\psi(v_n)} P\{M(n'h) > v_n\} \\
&\ge \liminf_{n\to\infty} \frac{1}{T_n\psi(v_n)} n'P\{M(h) > v_n\} \\
&\quad - \limsup_{n\to\infty} \frac{1}{T_n\psi(v_n)} [n'P\{M(h) > v_n\} - P\{M(n'h) > v_n\}] \\
&= \frac{1}{k} - o\left(\frac{1}{k}\right),
\end{aligned}$$

and hence $\liminf_{n\to\infty}(T_n\psi(v_n))^{-1} > 0$. Thus $T_n\psi(u_{T_n})$ is bounded for any sequence $\{T_n\}$ satisfying $nh < T_n \le (n+1)h$, which readily implies that $T\psi(u_T)$ is bounded. Finally, (13.2.1) then follows at once from (13.2.3). □

Corollary 13.2.2. *Under the conditions of the lemma, if*

$$\lambda_{n,k} = |n'h\psi(v_n) - P\{M(n'h) > v_n\}|,$$

then $\limsup_{n\to\infty} \lambda_{n,k} = o(k^{-1})$ *as* $k \to \infty$.

PROOF. Noting that $n'\psi(v_n)$ is bounded, this follows at once from the lemma, by (13.1.9). □

Our main result now follows readily.

Theorem 13.2.3. *Suppose that* (13.1.9) *holds for some function* ψ, *and let* $\{u_T\}$ *be a family of constants such that for each* $a > 0$, $C(u_T)$ *and* $C'(u_T)$ *hold with* *respect to the family* $\{q_a(u)\}$ *of constants satisfying* (13.1.10), *with* h *in* $C'(u_T)$ *not exceeding* $h_0/2$ *in* (13.1.9). *Then*

$$T\psi(u_T) \to \tau > 0 \tag{13.2.4}$$

if and only if

$$P\{M(T) \le u_T\} \to e^{-\tau}. \tag{13.2.5}$$

PROOF. If (13.1.9), (13.1.10), and $C'(u_T)$ hold as stated, then $T\psi(u_T)$ is bounded according to Lemma 13.2.1 and by Lemma 13.1.4 the sequence of "submaxima" $\{\zeta_n\}$ defined by (13.1.1) satisfies $D(v_n)$, with $v_n = u_{T_n}$, for any sequence $\{T_n\}$ with $T_n \in (nh, (n+1)h]$. Hence from Lemma 3.3.2 writing $n' = [n/k]$,

$$P\{M(nh) \le v_n\} - P^k\{M(n'h) \le v_n\} \to 0 \quad \text{as } n \to \infty. \tag{13.2.6}$$

Clearly it is enough to prove that

$$T_n\psi(v_n) \to \tau > 0 \tag{13.2.7}$$

if and only if

$$P\{M(T_n) \le v_n\} \to e^{-\tau}, \tag{13.2.8}$$

for any sequence $\{T_n\}$ with $T_n \in (nh, (n+1)h]$. Further, $T\psi(u_T)$ bounded implies that $\psi(u_T) \to 0$ as $T \to \infty$ so that

$$\begin{aligned} 0 &\le P\{M(nh) \le v_n\} - P\{M(T_n) \le v_n\} \\ &\le P\{M(h) > v_n\} \sim h\psi(v_n) \to 0, \end{aligned}$$

and thus (13.2.8) holds if and only if

$$P\{M(nh) \le v_n\} \to e^{-\tau}. \tag{13.2.9}$$

Hence it is sufficient to prove that (13.2.7) and (13.2.9) are equivalent under the hypothesis of the theorem.

Suppose now that (13.2.7) holds so that, in particular,

$$n'h\psi(v_n) \to \frac{\tau}{k} \quad \text{as } n \to \infty. \tag{13.2.10}$$

With the notation of Corollary 13.2.2 we have

$$1 - n'h\psi(v_n) - \lambda_{n,k} \le P\{M(n'h) \le v_n\} \le 1 - n'h\psi(v_n) + \lambda_{n,k} \tag{13.2.11}$$

so that, letting $n \to \infty$,

$$\begin{aligned} 1 - \frac{\tau}{k} - o(k^{-1}) &\le \liminf_{n\to\infty} P\{M(n'h) \le v_n\} \\ &\le \limsup_{n\to\infty} P\{M(n'h) \le v_n\} \\ &\le 1 - \frac{\tau}{k} + o(k^{-1}). \end{aligned}$$

By taking kth powers throughout and using (13.2.6) we obtain

$$\begin{aligned}\left(1 - \frac{\tau}{k} - o(k^{-1})\right)^k &\le \liminf_{n\to\infty} P\{M(nh) \le v_n\}\\ &\le \limsup_{n\to\infty} P\{M(nh) \le v_n\}\\ &\le \left(1 - \frac{\tau}{k} + o(k^{-1})\right)^k,\end{aligned}$$

and letting k tend to infinity proves (13.2.9).

Hence (13.2.7) implies (13.2.9) under the stated conditions. We shall now show that conversely (13.2.9) implies (13.2.7). The first part of the above proof still applies so that (13.2.6) and the conclusion of Corollary 13.2.2, and hence (13.2.11), hold. A rearrangement of (13.2.11) gives

$$\begin{aligned}1 - P\{M(n'h) \le v_n\} - \lambda_{n,k} &\le n'h\psi(v_n)\\ &\le 1 - P\{M(n'h) \le v_n\} + \lambda_{n,k}.\end{aligned}$$

But it follows from (13.2.6) and (13.2.9) that $P\{M(n'h) \le v_n\} \to e^{-\tau/k}$ and hence, using Corollary 13.2.2, that

$$\begin{aligned}1 - e^{-\tau/k} - o(k^{-1}) &\le \liminf_{n\to\infty} n'h\psi(v_n)\\ &\le \limsup_{n\to\infty} n'h\psi(v_n)\\ &\le 1 - e^{-\tau/k} + o(k^{-1}).\end{aligned}$$

Multiplying through by k and letting $k \to \infty$ shows that $T_n\psi(v_n) \sim nh\psi(v_n) \to \tau$, and concludes the proof that (13.2.9) implies (13.2.7). □

13.3. Associated Sequence of Independent Variables

With a slight change of emphasis from Chapter 3 we say that any i.i.d. sequence $\hat{\zeta}_1, \hat{\zeta}_2, \ldots$ whose marginal d.f. F satisfies

$$1 - F(u) \sim P\{M(h) > u\}$$

for some $h > 0$, is an *independent sequence associated with* $\{\xi(t)\}$. If (13.1.9) holds this is clearly equivalent to the requirement

$$1 - F(u) \sim h\psi(u) \quad \text{as } u \to \infty. \tag{13.3.1}$$

Theorem 13.2.3 may then be related to the corresponding result for i.i.d. sequences in the following way.

Theorem 13.3.1. *Let* $\{u_T\}$ *be a family of constants such that the conditions of Theorem* 13.2.3 *hold, and let* $\hat{\zeta}_1, \hat{\zeta}_2, \ldots$ *be an associated independent sequence.*

Let $0 < \rho < 1$. *If*

$$P\{M(T) \le u_T\} \to \rho \quad \text{as } T \to \infty \tag{13.3.2}$$

then

$$P\{\hat{M}_n \le v_n\} \to \rho \quad \text{as } n \to \infty \tag{13.3.3}$$

with $v_n = u_{nh}$. *Conversely, if* (13.3.3) *holds for some sequence* $\{v_n\}$ *then* (13.3.2) *holds for any* $\{u_T\}$ *such that* $\psi(u_T) \sim \psi(v_{[T/h]})$, *provided the conditions of Theorem* 13.2.3 *hold.*

PROOF. If (13.3.2) holds, and $\rho = e^{-\tau}$, Theorem 13.2.3 and (13.3.1) give

$$1 - F(u_{nh}) \sim h\psi(u_{nh}) \sim \frac{\tau}{n},$$

so that $P\{\hat{M}_n \le u_{nh}\} \to e^{-\tau}$, giving (13.3.3). Conversely, (13.3.3) and (13.3.1) imply that $h\psi(v_n) \sim 1 - F(v_n) \sim \tau/n$ and hence

$$T\psi(u_T) \sim T\psi(v_{[T/h]}) \sim \frac{T\tau}{h[T/h]} \to \tau$$

so that (13.3.2) holds by Theorem 13.2.3. □

These results show how the function ψ may be used in the classical criteria for domains of attraction to determine the asymptotic distribution of $M(T)$. We write $D(G)$ for the domain of attraction to the (extreme value) d.f. G, i.e. the set of all d.f.'s F such that $F^n(x/a_n + b_n) \to G(x)$ for some sequences $\{a_n > 0\}$, $\{b_n\}$.

Theorem 13.3.2. *Suppose that the conditions of Theorem* 13.2.3 *hold for all families* $\{u_T\}$ *of the form* $u_T = x/a_T + b_T$, *where* $a_T > 0$, b_T *are given constants, and that*

$$P\{a_T(M(T) - b_T) \le x\} \to G(x). \tag{13.3.4}$$

Then (13.3.1) *holds for some* $F \in D(G)$. *Conversely, suppose* (13.1.9) *holds and that* (13.3.1) *is satisfied for some* $F \in D(G)$, *let* $a_n' > 0$, b_n' *be constants such that* $F^n(x/a_n' + b_n') \to G(x)$, *and define* $a_T = a'_{[T/h]}$, $b_T = b'_{[T/h]}$. *Then* (13.3.4) *holds, provided the conditions of Theorem* 13.2.3 *are satisfied for each* $u_T = x/a_T + b_T$, $-\infty < x < \infty$.

PROOF. If (13.3.4) holds, together with the conditions stated, Theorem 13.3.1 applies, so that, in particular,

$$P\{a_{nh}(\hat{M}_n - b_{nh}) \le x\} \to G(x),$$

where $\hat{M}_n$ is the maximum of the associated sequence of independent variables $\hat{\zeta}_1, \ldots, \hat{\zeta}_n$. It follows at once that their marginal d.f. F belongs to $D(G)$, and (13.3.1) is immediate by definition.

Conversely, suppose (13.3.1) holds for some d.f. $F \in D(G)$, and let $\hat{\zeta}_1, \hat{\zeta}_2, \ldots$ be an i.i.d. sequence with marginal d.f. F, and suppose that for $v_n = x/a_n' + b_n'$,

$$P\{\hat{M}_n \le v_n\} \to G(x) \quad \text{as } n \to \infty.$$

Then clearly, for $a_T = a'_{[T/h]}$, $b_T = b'_{[T/h]}$, and $u_T = x/a_T + b_T$,

$$\psi(u_T) = \psi(v_{[T/h]})$$

so that Theorem 13.3.1 applies, giving (13.3.4). □

13.4. Stationary Normal Processes

Although we have obtained the asymptotic distributional properties of the maximum of stationary normal processes directly, it is of interest to see how these may be obtained as applications of the general theory of this chapter. This does not lessen the work involved, of course, since the same calculations in the "direct route" are used to verify the conditions in the general theory. However, the use of the general theory does also give insight and perspective regarding the principles involved. We deal here with the more general normal processes considered in Chapter 12. This will include the normal processes with finite second spectral moments considered in Chapter 8, of course. The latter processes may also be treated as particular cases of general processes with finite upcrossing intensities—a class dealt with in the next section.

Suppose then that $\xi(t)$ is a stationary normal process with zero mean and covariance function (12.1.1), viz.

$$r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha) \quad \text{as } \tau \to 0, \tag{13.4.1}$$

where $0 < \alpha \le 2$. The major result to be obtained is Theorem 12.3.4 restated here.

Theorem 12.3.4. *Let $\{\xi(t)\}$ be a zero-mean stationary normal process, with covariance function $r(t)$ satisfying* (13.4.1) *and*

$$r(t) \log t \to 0 \quad \text{as } t \to \infty. \tag{13.4.2}$$

If $u = u_T \to \infty$ and $\mu = \mu(u) = C^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u$ (with H_α defined in Theorem 12.2.9), *and if $T\mu(u_T) \to \tau$, then $P\{M(T) \le u\} \to e^{-\tau}$ as $T \to \infty$.*

Proof from the General Theory. Write $\psi(u) = \mu(u)$ so that $T\psi(u_T) \to \tau$. Theorem 12.2.9 shows at once that (13.1.9) holds (for all $h > 0$). Define $q_a(u) = au^{-2/\alpha}$, and note that (13.1.10) holds, by (12.2.19). Hence the result will follow at once if $C(u_T)$, $C'(u_T)$ are both shown to hold with respect to $\{q_a(u)\}$ for each $a > 0$.

It is easily seen in a familiar way that $C(u_T)$ holds. For by Corollary 4.2.2 the left-hand side of (13.1.4) (with $q_a(u)$ for $q(u)$, $u = u_T$) does not exceed

$$K \sum_{i=1}^{p} \sum_{j=1}^{p'} |r(t_j - s_i)| \exp\left\{-\frac{u^2}{1 + |r(t_j - s_i)|}\right\}$$

which is dominated by

$$K \frac{T}{q} \sum_{\gamma \le kq \le T} |r(kq)| \exp\left\{-\frac{u^2}{1 + |r(kq)|}\right\},$$

and this tends to zero for each $\{q\} = \{q_a\}$, a fixed, by Lemma 12.3.1. If we identify this expression with $\alpha_{T,\gamma}$ then (13.1.5) holds almost trivially since $\alpha_{T,\lambda T} \le \alpha_{T,\gamma}$ for any fixed γ when $\lambda T > \gamma$.

$C'(u_T)$ follows equally simply by Corollary 4.2.4, which gives

$$|P\{\xi(0) > u, \xi(jq) > u\} - (1 - \Phi(u))^2| \le K |r(jq)| \exp\left\{-\frac{u^2}{1 + |r(jq)|}\right\}$$

so that

$$\begin{aligned} &\frac{T}{q} \sum_{h < jq \le \varepsilon T} P\{\xi(0) > u, \xi(jq) > u\} \\ &\qquad \le \frac{\varepsilon T^2}{q^2}(1 - \Phi(u))^2 + K \frac{T}{q} \sum_{h < jq \le \varepsilon T} |r(jq)| \exp\left\{-\frac{u^2}{1 + |r(jq)|}\right\}. \end{aligned}$$

The second term tends to zero as $T \to \infty$ again by Lemma 12.3.1. The first term is asymptotically equivalent to

$$\frac{\varepsilon T^2}{q^2} \frac{(\phi(u))^2}{u^2} \to \frac{\varepsilon \tau^2}{a^2 C^{2/\alpha} H_\alpha^2}$$

by the definitions of q and $\psi(u)$, and the fact that $T\psi(u) \to \tau$. Since $\varepsilon\tau^2/a^2 \to 0$ for each fixed a as $\varepsilon \to 0$, $C'(u_T)$ follows. □

Finally, we note that the "double exponential limiting distribution" for the maximum $M(T)$ (Theorem 12.3.5) follows exactly as before from Theorem 12.3.4.

13.5. Processes with Finite Upcrossing Intensities

We show now how some of the conditions required for the general theory may be simplified when the mean number $\mu(u)$ of upcrossings of each level u per unit time is finite. This includes the particular normal cases with finite second spectral moments already covered in Chapter 8 and in the preceding section but, of course, not the "nondifferentiable" processes with $\alpha < 2$.

We use the notation of Chapter 7 in addition to that of the present chapter and assume that $\mu = \mu(u) = E(N_u(1)) < \infty$ for each value of u. Writing as in (7.2.1) for $q > 0$,

$$J_q(u) = \frac{P\{\xi(0) < u < \xi(q)\}}{q} \tag{13.5.1}$$

it is clear that

$$J_q(u) \leq \frac{P\{N_u(q) \geq 1\}}{q} \leq \frac{E(N_u(q))}{q} = \mu \tag{13.5.2}$$

and it follows from Lemma 7.2.2(iii) that

$$J_q(u) \to \mu \quad \text{as } q \to 0 \tag{13.5.3}$$

for each fixed u.

In the *normal* case we saw (Lemma 7.3.1) that $J_q(u) \sim \mu(u)$ as $q \to 0$ in such a way that $uq \to 0$. Here we shall use a variant of this property assuming as needed that, for each $a > 0$, there are constants $q_a(u) \to 0$ as $u \to \infty$ such that, with $q_a = q_a(u)$, $\mu = \mu(u)$,

$$\liminf_{u\to\infty} \frac{J_{q_a}(u)}{\mu} \geq \nu_a, \tag{13.5.4}$$

where $\nu_a \to 1$ as $a \to 0$. (As indicated below this is readily verified when $\xi(t)$ is normal when we may take $q_a(u) = a/u$.)

We shall assume as needed that

$$P\{\xi(0) > u\} = o(\mu(u)) \quad \text{as } u \to \infty, \tag{13.5.5}$$

which clearly holds for the normal case but more generally is readily verified if, for example, for some $q = q(u) \to 0$ as $u \to \infty$,

$$\limsup_{u\to\infty} \frac{P\{\xi(0) > u, \xi(q) > u\}}{P\{\xi(0) > u\}} < 1, \tag{13.5.6}$$

since (13.5.6) implies that $\liminf_{u\to\infty} qJ_q(u)/P\{\xi(0) > u\} > 0$, from which it follows that $P\{\xi(0) > u\}/J_q(u) \to 0$, and hence (13.5.5) holds since $J_q(u) \leq \mu(u)$.

We may now recast the conditions (13.1.8) and (13.1.9) in terms of the function $\mu(u)$, identifying this function with $\psi(u)$.

Lemma 13.5.1. (i) *Suppose* $\mu(u) < \infty$ *for each* u *and that* (13.5.5) (*or the sufficient condition* (13.5.6)) *holds. Then* (13.1.8) *holds with* $\psi(u) = \mu(u)$.
(ii) *If* (13.5.4) *holds for some family* $\{q_a(u)\}$ *then* (13.1.11) *holds with* $\psi(u) = \mu(u)$.

PROOF. Since clearly

$$P\{M(h) > u\} \leq P\{N_u(h) \geq 1\} + P\{\xi(0) > u\} \leq \mu h + P\{\xi(0) > u\},$$

(13.1.8) follows at once from (13.5.5), which proves (i).

Now, if (13.5.4) holds, then with $q = q_a(u)$, $\mu = \mu(u)$,

$$\begin{aligned} P\{\xi(0) &\le u, \xi(q) \le u, M(q) > u\} \\ &= P\{\xi(0) \le u, M(q) > u\} - P\{\xi(0) \le u < \xi(q)\} \\ &\le P\{N_u(q) \ge 1\} - qJ_q(u) \\ &\le \mu q - \mu q \nu_a(1 + o(1)) \end{aligned}$$

so that

$$\limsup_{u \to \infty} \frac{P\{\xi(0) \le u, \xi(q) \le u, M(q) > u\}}{q\mu} \le 1 - \nu_a,$$

which tends to zero as $a \to 0$, giving (13.1.11). □

In view of this lemma, the Extremal Types Theorem now applies to processes of this kind using the more readily verifiable conditions (13.5.4) and (13.5.5), as follows.

Theorem 13.5.2. *Theorem* 13.1.5 *holds for a stationary process* $\{\xi(t)\}$ *with* $\psi(u) = \mu(u) < \infty$ *for each* u *if the conditions* (13.1.6) *and* (13.1.10) *are replaced by* (13.5.4) *and* (13.5.5) (*or by* (13.5.4) *and* (13.5.6)).

PROOF. By (i) of the previous lemma the condition (13.5.5) (or its sufficient condition (13.5.6)) implies (13.1.8) and hence both (13.1.6) and (13.1.7). On the other hand (ii) of the lemma shows that (13.5.4) implies (13.1.11) which together with (13.1.7) implies (13.1.10) by Lemma 13.1.3(iii). □

The condition (13.1.10) also occurs in Theorem 13.2.3 and may, of course, be replaced by (13.5.4) there, since (13.1.7) is implied by (13.1.9) which is assumed in that theorem.

Finally, we note that while (13.5.5) and (13.5.6) are especially convenient to give (13.1.8) (Lemma 13.5.1(i)), the verification of (13.1.9) still requires obtaining

$$\liminf_{u \to \infty} \frac{P\{M(h) > u\}}{h\mu(u)} \ge 1 \quad \text{for } 0 \le h \le h_0.$$

This, of course, follows for all normal processes considered by Theorem 12.2.9 with $\alpha = 2$. There are a number of independent simpler derivations of this when $\alpha = 2$, one of these being along the lines of the "cosine-process" comparison in Chapter 7. The actual comparison used there gave a slightly weaker result, which was, however, sufficient to yield the desired limit theory by the particular methods employed.

13.6. Poisson Convergence of Upcrossings

It was shown in Chapter 9 that the upcrossings of one or more high levels by a normal process $\xi(t)$ take on a specific Poisson character under appropriate conditions. It was assumed, in particular, that the covariance function

$r(t)$ of $\xi(t)$ satisfied (8.1.1) so that the expected number of upcrossings per unit time, $\mu = E(N_u(1))$, is finite.

Corresponding results are obtainable for ε-upcrossings by normal processes when $r(t)$ satisfies (13.4.1) with $\alpha < 2$ and indeed the proof is indicated in Chapter 12 for the single level result (Theorem 12.4.2).

For general stationary processes the same results may be proved under conditions used in this present chapter, including C, C'. Again when $\mu = E(N_u(1)) < \infty$ the results apply to actual upcrossings, while if $\mu = \infty$ they apply to ε-upcrossings. We shall state and briefly indicate the proof of the specific theorem for a single level in the case when $\mu < \infty$.

As in previous discussions, we consider a time period T and a level u_T both increasing in such a way that $T\mu \to \tau > 0$ ($\mu = \mu(u_T)$), and define a normalized point process of upcrossings by

$$N_T^*(B) = N_{u_T}(TB), \qquad (N_T^*(t) = N_{u_T}(tT))$$

for each interval (or more general Borel set) B, so that, in particular,

$$E(N_T^*(1)) = E(N_{u_T}(T)) = \mu T \to \tau.$$

This shows that the "intensity" (i.e. mean number of events per unit time) of the normalized upcrossing point process converges to τ. Our task is to show that the upcrossing point process actually converges in distribution to a Poisson process with mean τ.

The derivation of this result is based on the following two extensions of Theorem 13.2.3, which are proved by similar arguments to those used in obtaining Theorem 13.2.3, and in Chapter 9.

Theorem 13.6.1. *Under the conditions of Theorem* 13.2.3, *if* $0 < \theta < 1$ *and* $\mu T \to \tau$, *then*

$$P\{M(\theta T) \le u_T\} \to e^{-\theta\tau} \quad \text{as } T \to \infty. \tag{13.6.1}$$

Theorem 13.6.2. *If* $I_1, I_2, \ldots, I_r$ *are disjoint subintervals of* $[0, 1]$ *and* $I_j^* = TI_j = \{t; t/T \in I_j\}$ *then under the conditions of Theorem* 13.2.3 *if* $\mu T \to \tau$,

$$P\left(\bigcap_{j=1}^{r} \{M(I_j^*) \le u_T\}\right) - \prod_{j=1}^{r} P\{M(I_j^*) \le u_T\} \to 0, \tag{13.6.2}$$

so that by Theorem 13.6.1

$$P\left(\bigcap_{j=1}^{r} \{M(I_j^*) \le u_T\}\right) \to \exp\left(-\tau \sum_{j=1}^{r} \theta_j\right), \tag{13.6.3}$$

where θ_j *is the length of* I_j, $1 \le j \le r$.

It is now a relatively straightforward matter to show that the point processes N_T^* converges (in the full sense of weak convergence) to a Poisson process N with intensity τ.

Theorem 13.6.3. *Under the conditions of Theorem* 13.2.3, *if* $T\mu \to \tau$ *where* $\mu = \mu(u_T)$, *then the family* N_T^* *of normalized point processes of upcrossings of* u_T *on the unit interval converges in distribution to a Poisson process* N *with intensity* τ *on that interval as* $T \to \infty$.

PROOF. Again by Theorem A.1 it is sufficient to prove that

(i) $E(N_T^*((c, d])) \to E(N((c, d])) = \tau(d - c)$ as $T \to \infty$ for all c, d, $0 < c < d \le 1$, and
(ii) $P\{N_T^*(B) = 0\} \to P\{N(B) = 0\}$ as $T \to \infty$ for all sets B of the form $\bigcup_1^r B_j$ where r is any integer and B_j are disjoint intervals $(c_j, d_j] \subset (0, 1]$.

Now (i) follows trivially since

$$E(N_T^*((c, d])) = \mu T(d - c) \to \tau(d - c).$$

To obtain (ii) we note that

$$\begin{aligned} 0 &\le P\{N_T^*(B) = 0\} - P\{M(TB) \le u_T\} \\ &= P\{N_u(TB) = 0, M(TB) > u_T\} \\ &\le \sum_{j=1}^{r} P\{\xi(Tc_j) > u_T\} \end{aligned}$$

since if the maximum in $TB = \bigcup_{j=1}^r (Tc_j, Td_j]$ exceeds u_T, but there are no upcrossings of u_T in these intervals, then $\xi(t)$ must exceed u_T at the initial point of at least one such interval. But the last expression is just

$$rP\{\xi(0) > u_T\} \to 0$$

as $T \to \infty$. Hence

$$P\{N_T^*(B) = 0\} - P\{M(TB) \le u_T\} \to 0.$$

But $P\{M(TB) \le u_T\} = P(\bigcap_{j=1}^r \{M(TB_j) \le u_T\}) \to \exp\{-\tau\Sigma(d_j - c_j)\}$ by Theorem 13.6.2 so that (ii) follows since $P\{N(B) = 0\} = \exp\{-\tau\Sigma(d_j - c_j)\}$. □

Corollary 13.6.4. *If* B_j *are disjoint (Borel) subsets of the unit interval and if the boundary of each* B_j *has zero Lebesgue measure then*

$$P\{N_T^*(B_j) = r_j, 1 \le j \le r\} \to \prod_{j=1}^{r} \exp\{-\tau m(B_j)\} \frac{(\tau m(B_j))^{r_j}}{r_j!},$$

where $m(B_j)$ *denotes the Lebesgue measure of* B_j.

PROOF. This is an immediate consequence of the full distributional convergence proved (cf. Appendix). □

The above results concern convergence of the point processes of upcrossings of u_T to a Poisson process in the unit interval. A slight modification, requiring C and C' to hold for all families $u_{\theta T}$ in place of u_T for all $\theta > 0$, enables a corresponding result to be shown for the upcrossings on the whole positive real line, but we do not pursue this here. Instead we show how Theorem 13.6.3 yields the asymptotic distribution of the kth largest local maximum in $(0, T]$.

Suppose, then, that $\xi(t)$ has a continuous derivative a.s. and (cf. Chapters 7 and 9) define $N_u'(T)$ to be the number of local maxima in the interval $(0, T]$ for which the process value exceeds u, i.e. the number of downcrossing points t of zero by ξ' in $(0, T]$ such that $\xi(t) > u$. Clearly $N_u'(T) \geq N_u(T) - 1$ since at least one local maximum occurs between two upcrossings. It is also reasonable to expect that if the sample function behaviour is not too irregular, there will tend to be just one local maximum above u between most successive upcrossings of u when u is large, so that $N_u'(T)$ and $N_u(T)$ will tend to be approximately equal. The following result makes this precise.

Theorem 13.6.5. *With the above notation let $\{u_T\}$ be constants such that $P\{\xi(0) > u_T\} \to 0$, and that $T\mu$ $(= T\mu(u_T)) \to \tau > 0$ as $T \to \infty$. Suppose that $E(N_u'(1))$ is finite for each u and that $E(N_u'(1)) \sim \mu(u)$ as $u \to \infty$. Then, writing $u_T = u$, $E(|N_u'(T) - N_u(T)|) \to 0$. If also the conditions of Theorem 13.6.3 hold (so that $P\{N_u(T) = r\} \to e^{-\tau}\tau^r/r!$) it follows that $P\{N_u'(T) = r\} \to e^{-\tau}\tau^r/r!$.*

Proof. As noted above, $N_u'(T) \geq N_u(T) - 1$, and it is clear, moreover, that if $N_u'(T) = N_u(T) - 1$, then $\xi(T) > u$. Hence

$$\begin{aligned} E(|N_u'(T) - N_u(T)|) &= E(N_u'(T) - N_u(T)) + 2P\{N_u'(T) = N_u(T) - 1\} \\ &\leq TE(N_u'(1)) - \mu T + 2P\{\xi(T) > u\}, \end{aligned}$$

which tends to zero as $T \to \infty$ since $P\{\xi(T) > u_T\} = P\{\xi(0) > u_T\} \to 0$ and $TE(N_{u_T}'(1)) - \mu T = \mu T((1 + o(1)) - 1) \to 0$, so that the first part of the theorem follows. The second part now follows immediately since the integer-valued r.v. $N_u'(T) - N_u(T)$ tends to zero in probability, giving $P\{N_u'(T) \neq N_u(T)\} \to 0$ and hence $P\{N_u'(T) = r\} - P\{N_u(T) = r\} \to 0$ for each r. □

Now write $M^{(k)}(T)$ for the kth largest local maximum in the interval $(0, T)$. Since the events $\{M^{(k)}(T) \leq u\}$, $\{N_u'(T) < k\}$ are identical we obtain the following corollary.

Corollary 13.6.6. *Under the conditions of the theorem*

$$P\{M^{(k)}(T) \leq u_T\} \to e^{-\tau} \sum_{s=0}^{k-1} \frac{\tau^s}{s!}.$$

As a further corollary we obtain the limiting distribution of $M^{(k)}(T)$ in terms of that for $M(T)$.

Corollary 13.6.7. *Suppose that $P\{a_T(M(T) - b_T) \leq x\} \to G(x)$ and that the conditions of Theorem 13.2.3 hold with $u_T = x/a_T + b_T$ for each real x (and $\psi = \mu$). Suppose also that $E(N'_u(1)) \sim E(N_u(1))$ as $u \to \infty$. Then*

$$P\{a_T(M^{(k)}(T) - b_T) \leq x\} \to G(x) \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!}, \tag{13.6.4}$$

where $G(x) > 0$ (and zero if $G(x) = 0$).

PROOF. This follows from Corollary 13.6.6 by writing $G(x) = e^{-\tau}$ since Theorem 13.2.3 implies that $T\mu \to \tau$. □

Note that, by Lemma 9.5.1(i), for a stationary *normal* process with finite second and fourth spectral moments $E(N'_u(1)) \sim \mu$ so that Theorem 13.6.5 and its corollaries apply.

The relation (13.6.4) gives the asymptotic distribution of the kth largest local maximum $M^{(k)}(T)$ as a corollary of Theorem 13.6.5. Further, it is clearly possible to generalize Theorem 13.6.5 to give "full Poisson convergence" for the point process of local maxima of height above u and indeed to generalize Theorem 9.5.2 and obtain joint distributions of heights and positions of local maxima in this general situation.

13.7. Interpretation of the Function $\psi(u)$

The function $\psi(u)$ used throughout this chapter describes the tail of the distribution of the maximum $M(h)$ in a fixed interval h, in the sense of (13.1.9), viz.

$$P\{M(h) > u\} \sim h\psi(u) \quad \text{for } 0 < h \leq h_0.$$

We have seen how ψ may be calculated for particular cases—as $\psi(u) = \mu(u)$ for processes with a finite upcrossing intensity $\mu(u)$ and as

$$\psi(u) = K\phi(u)u^{(2/\alpha)-1}$$

for normal processes satisfying (13.4.1). Berman (1982a) has recently considered another general method for obtaining ψ (or at least many of its properties) based on the asymptotic distribution of the amount of time spent above a high level.

Specifically Berman considers the time $L_t(u)$ which a stationary process spends above the level u in the interval $(0, t)$ and proves the basic result

$$\lim_{u\to\infty} \frac{P\{vL_t(u) > x\}}{E(vL_t(u))} \to -\Gamma'(x),$$

at all continuity points $x > 0$ of Γ' (under given conditions). Here $v = v(u)$ is a certain function of u and $\Gamma(x)$ is an absolutely continuous nonincreasing function with density Γ', and t is fixed.

While this result does not initially apply at $x = 0$, it is extended to so apply giving, since the events $\{M(t) > u\}$, $\{L_t(u) > 0\}$ are equivalent,

$$\begin{aligned} P\{M(t) > u\} &\sim -\Gamma'(0)E(vL_t(u)) \\ &= -\Gamma'(0)vt(1 - F(u)), \end{aligned}$$

where F is the marginal d.f. of the process, since it is very easily shown that $E(L_t(u)) = t(1 - F(u))$. Hence we may—under the stated condition—obtain ψ as

$$\psi(u) = -\Gamma'(0)v(u)(1 - F(u)).$$

It is required in this approach that F have such a form that it belongs to the domain of attraction of the Type I extreme value distribution and it follows (though not immediately) that $M(h)$ has a Type I limit so that (e.g. from the theory of this chapter) a limiting distribution for $M(T)$ as $T \to \infty$ must (under appropriate conditions) also be of Type I. However a number of important cases are covered in this approach including stationary normal processes, certain Markov processes, and so-called χ^2-processes. Further, the approach gives considerable insight into the central ideas governing extremal properties.

PART IV

APPLICATIONS OF EXTREME VALUE THEORY

A substantial section of this volume has been directed towards showing that the classical theory of extremes still applies, under specified general assumptions, to a wide variety of dependent sequences and continuous parameter processes. It is tempting, by way of applications, to give examples which simply demonstrate how the classical extremal distributions do apply to such dependent situations. We have done this only to a limited extent, for two reasons. First, the literature abounds with applications of the classical theory, and many of these are really *dependent* cases although assumed independent. More importantly, however, we feel that each potential application should be understood as well as possible in terms of its underlying physical principles so that extremal theory may be thoughtfully applied in the light of such principles, rather than by routine "trial and error" fitting.

Our approach in this part is therefore to primarily include applications which we feel do profit from a discussion of such underlying principles, occasionally involving modest extensions to the general theory, and to point out where difficulties may occur. For a more extensive compilation of the fitting of extremal distributions, under classical assumptions, we refer to the literature, e.g. Gumbel (1958); Harter (1978) contains a comprehensive listing of references.

CHAPTER 14
Extreme Value Theory and Strength of Materials

Extreme value distributions have found widespread use for the description of strength of materials and mechanical structures, often in combination with stochastic models for the loads and forces acting on the material. Thus it is often assumed that the maximum of several loads follows one of the extreme value distributions for maxima. More important, and also less obvious, is that the strength of a piece of material, such as a strip of paper or glass fibre, is sometimes determined by the strength of its weakest part, and then perhaps follows one of the extreme value distributions for minima. Based on this so-called *weakest link principle* much of the work has been directed towards a study of *size effects* in the testing of materials. By this we mean the empirical fact that the strength of a piece of material varies with its dimensions in a way which is typical for the type of material and the geometrical form of the object. An early attempt towards a statistical theory for this was made more than a century ago by Chaplin (1880, 1882); see also Lieblein (1954) and Harter (1977).

14.1. Characterizations of the Extreme Value Distributions

Here we shall state and discuss some precise conditions under which the extreme value distributions would appear as distributions for material strength, and we also present some illustrating data.

Suppose that a piece of material, such as a glass fibre or an iron bar, is subject to tension, and that it breaks if the tension exceeds the inherent strength of the material. It has been found by experience that the breaking tension varies stochastically from piece to piece but that its distribution depends on the size of the material in quite a regular manner.

Let ξ_L be the random strength of a piece of material L with length l, and suppose the material can be divided, at least hypothetically, into smaller pieces, $L_1, \ldots, L_n$ of arbitrary lengths $l_1, \ldots, l_n$, and with (random) strength $\xi_{L_1}, \ldots, \xi_{L_n}$, respectively. We say that the material is stochastically

(i) *brittle* if $\xi_L = \min(\xi_1, \ldots, \xi_{L_n})$,
(ii) *homogeneous* if the marginal distributions of $\xi_{L_1}, \ldots, \xi_{L_n}$ depend only on $l_1, \ldots, l_n$,
(iii) *disconnected* if $\xi_{L_1}, \ldots, \xi_{L_n}$ are independent for all disjoint subdivisions $L_1, \ldots, L_n$ of L.

Of these properties, (ii) and (iii) are of purely statistical character, while (i) depends on the mechanism involved in a failure. All properties have definite physical meaning, and any specific material can, at least approximately, have one or more of the three properties. Later we will discuss what happens when some of the requirements are relaxed, and see how far the present theory can describe properties of nonbrittle, inhomogeneous, and weakly connected materials.

Suppose a material satisfies (i)–(iii), and let $F_l(x)$ be the (non-degenerate) d.f. of the strength of a piece with length l. Then

$$1 - F_1(x) = (1 - F_{1/n}(x))^n. \tag{14.1.1}$$

Now, any d.f. $F_1(x)$ satisfies (14.1.1) for some d.f. $F_{1/n}(x)$, for example, taking $F_{1/n}(x) = 1 - (1 - F_1(x))^{1/n}$, so that in order to obtain a simple structure, we therefore have to introduce one extra restriction on the material.

A material is called stochastically

(iv) *size-stable* if the distribution of ξ_L is of the same type regardless of the length l, i.e. there are constants $a_l > 0$, b_l and a d.f. F, such that $F_l(x) = F(a_l(x - b_l))$.

This is an *ad hoc* notion, but is frequently used in this context, since it leads to considerable mathematical simplification. Unlike properties (i)–(iii) it has no strict physical meaning, but is sometimes defended from the intuitively appealing principle that the simplest model is also the physically most realistic one.

For a size-stable material, writing $\alpha_n = a_{1/n}a_1^{-1}$, $\beta_n = b_1 - \alpha_n b_{1/n}$,

$$F_{1/n}(x) = F_1(\alpha_n x + \beta_n)$$

so that, by (14.1.1),

$$1 - F_1(x) = (1 - F_1(\alpha_n x + \beta_n))^n. \tag{14.1.2}$$

Thus, F_1 is *min-stable*, and hence, according to Theorem 1.8.4, F can be taken to have one of the following forms.

$$\text{Type I:}\quad F(x) = 1 - \exp(-e^x), \qquad -\infty < x < \infty,$$

$$\text{Type II:}\quad F(x) = \begin{cases} 1 - \exp(-(-x)^{-\alpha}), & \text{for some } \alpha > 0, \quad x < 0 \\ 1, & x \geq 0, \end{cases}$$

$$\text{Type III:}\quad F(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp(-x^\alpha), \quad \text{for some } \alpha > 0, & x \geq 0. \end{cases}$$

Thus strength distributions for materials satisfying (i)–(iv) are of one of the three minimum extremal types. If, furthermore, the scale of measurement is such that the measured strength ξ_L is bounded from below, $\xi_L \geq x_0$ for some $x_0 \in (-\infty, \infty)$, then the only possibility is the Type III (or Weibull) distribution, with location parameter x_0, and general scale parameter, i.e.

$$F(x) = 1 - \exp\{-(a(x - x_0))^\alpha\}, \qquad x \geq x_0, \tag{14.1.3}$$

for some constant $\alpha > 0$. Often $x_0 = 0$ is a natural choice, expressing the fact that strength can never be negative. However, even in cases when the strength must be positive, one may prefer to use one of the other types, e.g., the double exponential,

$$F(x) = 1 - \exp\{-e^{a(x-b)}\}, \qquad -\infty < x < \infty, \tag{14.1.4}$$

since if the location parameter b is large enough, the probability of a value below the *a priori* lower bound is negligible. If (14.1.4) gives a better fit to data than (14.1.3) in the important region of variation it is therefore perhaps not necessary to favour the latter. However, great care must be taken when the aim of the analysis is prediction of very low values, as the following example shows.

Example 14.1.1. (Yield strength of high-tensile steel). Test specimens of high-tensile steel of a special quality were tested for yield strength. The samples were selected from the entire production of a specific manufacturer during two years (1975 and 1976). The empirical distribution of 100 observed values of yield strength is plotted on a nonstandard Weibull probability paper in Figure 14.1.1 together with a fitted Weibull distribution,

$$(-\log(1 - F(x)))^{1/\alpha} = a(x - x_0), \qquad x \geq x_0$$

with $\alpha = 2.7$, $x_0 = 463$ MPa; (MPa = MegaPascal = 10^6 Newton per m^2). There is a good fit to the Weibull distribution, quite in accordance with theory. One could even be tempted to use $x_0 = 463$ MPa as a true lower strength limit, and to predict the frequency of low strength values from this distribution function.

For example, since the 0.1 % fractile of the fitted distribution is 470 MPa it appears that only 1 out of 1000 test probes should have a strength less than 470 MPa.

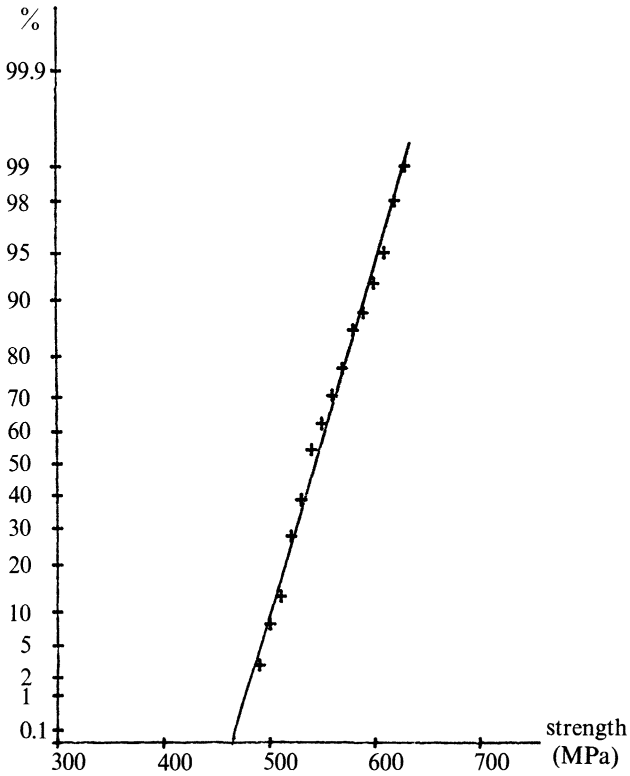


Figure 14.1.1. Empirical distribution of 100 yield strength values with fitted Weibull d.f. $F(x) = 1 - \exp(-a(x - x_0)^\alpha)$ with $a = 0.011$, $x_0 = 463$, $\alpha = 2.7$.

However, Figure 14.1.2 shows the observed distribution of strength in a sample of 6262 items, in which as many as 25 are less than 470 MPa, and 12 even less than 460 MPa. Hence, clearly the excellent Weibull fit obtained from the 100 values in the centre of the range does, for some reason, not extend to the tails. One explanation for the noted discrepancy between the behaviour in the tail and in the central region of the distribution could be that Figure 14.1.1 describes the variation in the material under some normal production conditions, while the extremely low strength values in Figure 14.1.2 are due to gross effects, such as misclassification of the product or external disturbances of the production.

Regardless of the explanation, the example shows that great care has to be used when extrapolating from the normal range of variation to very extreme ranges, even in cases where theory and observations seem at first sight to agree.

An account of the data presented above can be found in Öfverbeck and Östberg (1977), together with a discussion of gross error effects on construction steel. □

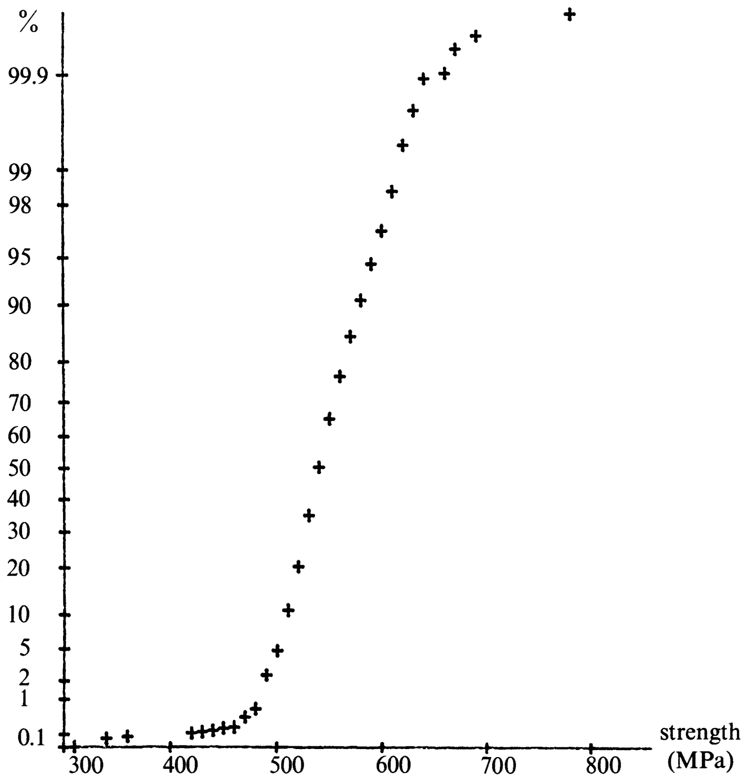


Figure 14.1.2. Empirical distribution of 6262 yield strength values; scales chosen to give a straight line for the Type III (Weibull) distribution.

14.2. Size Effects in Extreme Value Distributions

The conditions (i)–(iii) introduced in Section 14.1 state that the strength of a piece of material is determined by its weak points, of stochastic strength and location, distributed completely at random over the material. Together with the stability condition (iv) this led to a characterization of the possible strength distributions as the three types of extreme value distribution for minima. A simple consequence of the conditions is that the strength decreases in a specific way as the size of tested material increases. Here we shall give an example to illustrate this, and also show how an observed deviation from the simple law of decrease can suggest possible violations of one or both of the conditions (ii) and (iii). Harter (1977) gives an extensive discussion of the literature of size effects from the earliest days; the bibliographic part of Harter's work is published separately as Harter (1978).

If conditions (i)–(iv) hold, then the d.f. $F_l(x)$ of the strength of a piece with length l is

$$F_l(x) = 1 - (1 - F_1(x))^l = F(a_l(x - b_l)). \tag{14.2.1}$$

Thus, F_l being min-stable, it is simple to obtain the relation between size l and distributional parameters such as mean strength and standard deviation.

For example, if $F(x) = 1 - \exp(-e^x)$, writing $a = a_1$, $b = b_1$, $F_1(x) = 1 - \exp(-e^{a(x-b)})$ is a double exponential distribution with scale and location parameters a^{-1} and b, and then

$$F_l(x) = 1 - \exp(-e^{a(x-b+a^{-1}\log l)}),$$

so that $F_l(x)$ is double exponential with parameters

$$a_l = a,$$
$$b_l = b - a^{-1}\log l.$$

From this we can obtain the effect of size on other parameters, such as the mean strength m_l, and the standard deviation σ_l. Since, in the standardized case with $a = 1$, $b = 0$, $m_1 = -\gamma$, (where γ is Euler's constant with the approximate value 0.577), and $\sigma_1 = \pi/\sqrt{6}$ we have in the general case

$$m_l = b - a^{-1}(\gamma + \log l),$$
$$\sigma_l = a^{-1}\pi/\sqrt{6}.$$

Similar calculations may be done for the Type II and III distributions. Table 14.2.1 gives location and scale parameters and the mean and standard deviation of strength as functions of the tested length l (expressed in terms of $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$). For completeness, all three types of distributions are included in the table, even though the Type II d.f. is not commonly used as strength distribution.

Example 14.2.1. (Size effect on strength of paper strips). A strip of paper will burst when it is subjected to a tension which exceeds the strength at its weakest point. We shall discuss here the mean strength of paper strips (of constant width) with length varying from 8 cm up to 10 m.

If an experiment is carried out several times with strips of various lengths one can check the fit of the extreme value distributions and the predicted dependence of the mean strength m_l on the length l. Figures 14.2.1 and 14.2.2 show observed mean values (measured in the unit kN, kiloNewton, per meter of paper width) obtained from experiments with paper strips of width 5 cm, plotted on two different scales, chosen so that a Type I (Figure 14.2.1)

Table 14.2.1

Type I (double exponential, $F(x) = 1 - \exp(-e^x)$)
$F_l(x) = F(a_l(x - b_l)) = 1 - \exp(-e^{a(x-b+a^{-1}\log l)})$
$a_l = a$
$b_l = b - a^{-1} \log l$
$m_l = b - a^{-1}(\gamma + \log l)$
$\sigma_l = a^{-1}\pi/\sqrt{6}$
Type II ($F(x) = 1 - \exp(-(-x)^{-\alpha}), x < 0$)
$F_l(x) = F(a_l(x - b_l)) = 1 - \exp(-l(-a(x - x_0))^{-\alpha}), \quad x < x_0$
$a_l = al^{-1/\alpha}$
$b_l = b = x_0$
$m_l = x_0 - l^{1/\alpha}a^{-1}\Gamma(1 - 1/\alpha)$ (if $\alpha > 1$)
$\sigma_l = l^{1/\alpha}a^{-1}\{\Gamma(1 - 2/\alpha) - \Gamma^2(1 - 1/\alpha)\}^{1/2}$ (if $\alpha > 2$)
Type III (Weibull distribution, $F(x) = 1 - \exp(-x^\alpha), x > 0$)
$F_l(x) = F(a_l(x - b_l)) = 1 - \exp(-l(a(x - x_0))^\alpha), \quad x > x_0$
$a_l = al^{1/\alpha}$
$b_l = b = x_0$
$m_l = x_0 + l^{-1/\alpha}a^{-1}\Gamma(1 + 1/\alpha)$
$\sigma_l = l^{-1/\alpha}a^{-1}\{\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)\}^{1/2}$

and a Type III (Figure 14.2.2) distribution would give a straight line for m_l.

Two qualities of paper were tested and for one of the qualities three series of experiments were carried out. As is seen from the diagrams, mean strength clearly decreases with increasing length, but no decisive conclusion can be drawn from these data regarding the best fitting distribution. We are grateful to Dr Bengt Hållberg and Svenska Cellulosa Aktiebolaget, SCA, for making the data available. □

There is no difficulty in generalizing the theory above to apply in more than one dimension and to describe area and volume effects on the strength of materials.

We shall now briefly discuss the possibility of handling nonbrittle, inhomogeneous and/or weakly connected materials for which one or more of (i)–(iii) does not hold. As mentioned previously, the size-stability condition (iv) cannot be removed completely, since then any strength distribution would apply.

If a material is (stochastically) *nonbrittle* the strength of the whole is not equal to the strength of its weakest part.

A number of different strength models would then be possible. If, for example, in a bundle of parallel fibres, the total load were distributed on the separate fibres in proportion to their individual strength, then the total

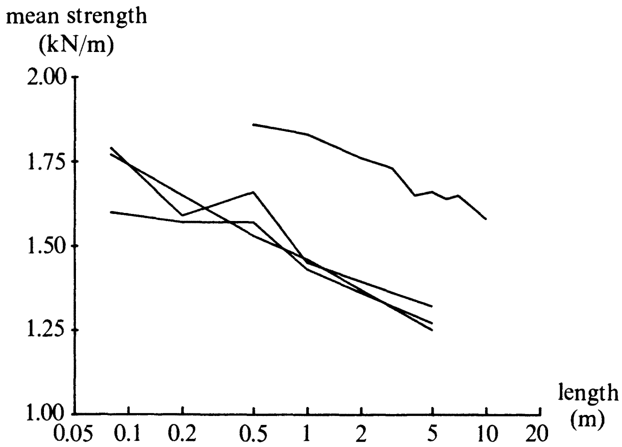


Figure 14.2.1. Mean bursting strength of paper strips of varying length; scales chosen to give straight lines for the Type I distribution.

strength would be the sum of the strengths of the fibres, leading to a normal distribution for total strength, with mean proportional to the number of fibres. Also the case where the total load is distributed equally over all remaining, nonfailed parts of the material, those parts failing whose strength is less than its share of the total load, will lead to an asymptotic normal distribution for the total strength, see, for example, Smith (1980) and references therein.

In a stochastically *inhomogeneous* material the strength distribution for a small piece of the material varies with its location. However, if the conditions (i), (iii), and (iv) are satisfied, one can still obtain one of the extreme value distributions under mild conditions on the inhomogeneity. One simple

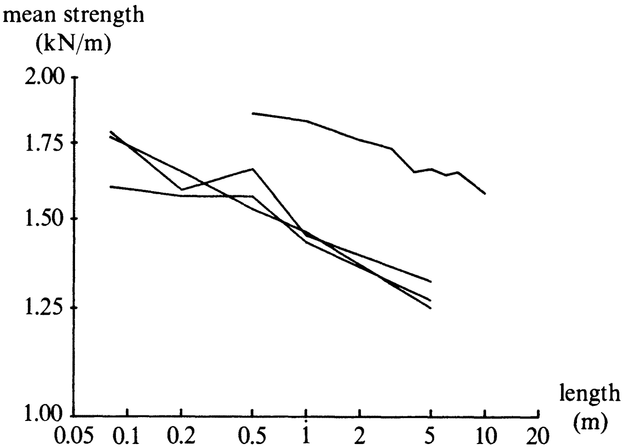


Figure 14.2.2. Mean bursting strength of paper strips of varying length; scales chosen to give straight lines for the Type III distribution.

kind of inhomogeneity is obtained by measuring the size of the material not as its physical dimension but in terms of an integral of a positive local size function, as will now be described.

Let $\lambda(x)$, $x \in L$ be a non-negative, integrable function, and define, for $L_i \subset L$,

$$l_i = \int_{L_i} \lambda(x)\, dx. \tag{14.2.2}$$

Let L_i, $i = 1, \ldots, n$ be disjoint parts of L, $\bigcup_{i=1}^n L_i = L$, and let, as before, $\xi_{L_1}, \ldots, \xi_{L_n}$ be the strengths of the separate pieces.

The material is called stochastically

(ii′) *inhomogeneous with size function* λ, if the marginal distributions of $\xi_{L_1}, \ldots, \xi_{L_n}$ depend only on $l_1, \ldots, l_n$ as defined by (14.2.2).

Now suppose a material satisfies (i), (ii′), (iii), (iv), and let $F_l(x)$ be the d.f. of the strength of a piece L with

$$l = \int_L \lambda(x)\, dx.$$

Then it is readily seen that (14.1.1) and (14.1.2) still hold, so that $F_1(x)$ is min-stable and one of the three extremal types for minima. Using a, b as the parameters in $F_1(x)$, we have in the Weibull and double exponential case

$$F_1(x) = 1 - \exp\{-(a(x - x_0))^\alpha\}, \qquad x > x_0,$$
$$F_1(x) = 1 - \exp\{-e^{a(x-b)}\},$$

respectively.

Starting from the local size function $\lambda(x)$ one can simply derive explicit expressions for the location- and scale-parameters in the strength distribution for a piece L with size $l = \int_L \lambda(x)\, dx$, which is $F^{(L)}(x) = F_l(x) = Fa_l(x - b_l)$, with a_l and b_l given in Table 14.2.1. To obtain these expressions, define the functions $a(x)$, $b(x)$ as

$$\begin{aligned} a(x) &= a\lambda^{1/\alpha}(x) \quad \text{(Weibull)}, \\ b(x) &= b - a^{-1} \log \lambda(x) \quad \text{(double exponential)} \end{aligned} \tag{14.2.3}$$

and

$$\begin{aligned} A(L) &= \left(\int_L a(x)^\alpha\, dx\right)^{1/\alpha}, \\ B(L) &= -a^{-1} \log \int_L e^{-ab(x)}\, dx. \end{aligned} \tag{14.2.4}$$

We can then write the d.f. for minimum strength ξ_L of a piece L as

$$\begin{aligned} F^{(L)}(x) &= 1 - \exp\{-(A(L)(x - x_0))^\alpha\}, \\ F^{(L)}(x) &= 1 - \exp\{-e^{a(x - B(L))}\}, \end{aligned} \tag{14.2.5}$$

respectively. These formulae motivate the use of $a(x)^{-1}$ and $b(x)$ as "local" scale and location functions in standard extremal models.

The average strength is given by

$$E(\xi_L) = \begin{cases} x_0 + A(L)^{-1}\Gamma(1 + 1/\alpha), \\ -a^{-1}\gamma + B(L), \end{cases}$$

in the two cases.

The most interesting generalization of the properties (i)–(iv), from the point of view taken in this book, is to (weakly) *connected* materials, in which there is a dependence of the strength over separate parts of the material. However, this will lead to strength distributions which are not necessarily any of the extreme value types for any finite test size, although they may be asymptotically so as size increases, under natural conditions as in Chapter 13.

Let the local strength parameters (14.2.3), $a(x)$, $x \in L$ or $b(x)$, $x \in L$, be, not deterministic functions but, stochastic processes with some distribution that depends on the irregularity and connectedness of the material. Then $A(L)$ and $B(L)$ in (14.2.4) are r.v.'s with d.f.'s $F_{A(L)}$ and $F_{B(L)}$, say, and if we let (14.2.5) define the conditional d.f. of ξ_L given $a(x)$ or $b(x)$, then

$$F^{(L)}(x) = 1 - \int_{s=0}^{\infty} \exp\{-(s(x - x_0))^\alpha\}\, dF_{A(L)}(s)$$

or

$$F^{(L)}(x) = 1 - \int_{s=-\infty}^{\infty} \exp\{-e^{a(x-s)}\}\, dF_{B(L)}(s), \tag{14.2.6}$$

respectively. Furthermore, the strengths of disjoint pieces would be dependent, through the outcome of the processes $a(x)$ and $b(x)$. The average strength will be

$$E(\xi_L) = \begin{cases} x_0 + E(A(L)^{-1})\Gamma(1 + 1/\alpha), \\ a^{-1}\gamma + E(B(L)). \end{cases}$$

The concept of a random local strength, averaging to a random location parameter $B(L)$ in the d.f. $F^{(L)}(x)$ given by (14.2.6), is formally somewhat analogous to the "strong dependence case" for the maxima of normal sequences treated in Sections 6.4 and 6.5 in which one has to subtract a slowly varying mean level before obtaining the double exponential limit; compare formula (14.2.6) with the mixed double exponential limit for maxima in Corollary 6.5.2.

Example 14.2.2. It has been noted in experiments that the average strength does not always decrease with increasing length as in any one case in Table 14.1, but can fall off with varying rates for different lengths. For example, the strength of glass fibre seems to decrease more rapidly when the length exceeds a certain limit. Some authors, for example, Metcalfe and Smitz (1964) explain this by postulating that weak points of the fibres appear in certain semi-stochastic patterns, and that they tend to cluster at regular intervals. This can obviously be modelled by means of the random strength parameters $A(L)$ and $B(L)$ as described above. □

CHAPTER 15

Application of Extremes and Crossings under Dependence

In this chapter we shall present some examples of continuous parameter processes and sequences with dependence where extreme value theory may be applied for descriptive or predictive purposes.

First, the important distinction between continuous time and discrete time extremes is illustrated, and the Poisson character of exceedances and crossings discussed in some examples. Descriptive models for some physical phenomena are discussed in connection with domains of attraction, particularly relative to mixtures of distributions with random scale or location parameter. An important problem, lying behind many extremal problems, is that of extrapolating extreme value distributions over expanding intervals (cf. the discussion of size effect in Chapter 14). This gives rise to several statistical problems concerning the choice of proper normalizing constants and the effect of nonstationarity. A discussion of these and some examples of local extremes for the description of random waves concludes the chapter.

15.1. Extremes in Discrete and Continuous Time

Many physical phenomena are (seemingly) continuous by nature, and the mathematical models which describe them are most naturally phrased in terms of continuous time processes. For extreme values, it is, in many cases, the continuous time extremes which are of primary interest. However, most statistical observations are obtained from some sampling procedure, and this will have the obvious effect that extremes in many cases become less accentuated. Since there are fairly complete (asymptotic) theories for extremes in both the continuous and the discrete case (at least for normal processes) a natural question concerns the relation between the two. The following example illustrates some practical consequences of the difference between discrete and continuous extremes.

Example 15.1.1 (Extreme temperatures). Experiments with thermometers and temperature measurements were started in Uppsala, Sweden, around 1712–1713, shortly after Fahrenheit had constructed his temperature scale

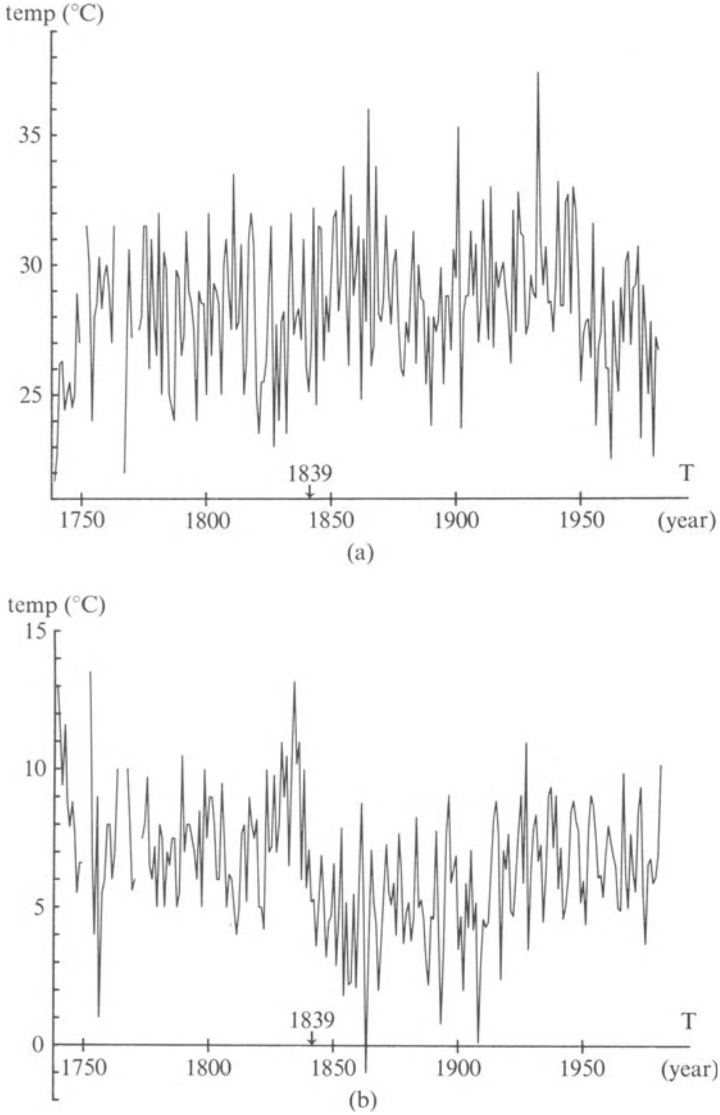
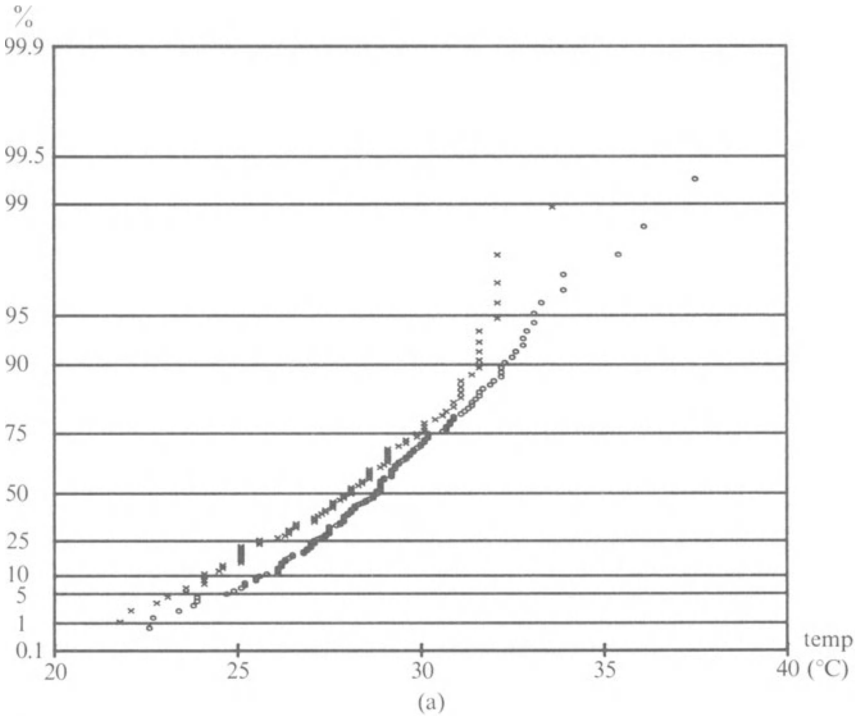


Figure 15.1.1. (a) Maximum and (b) minimum recorded temperature in Uppsala during the month of July; note the introduction of a max- and min-thermometer in 1839.



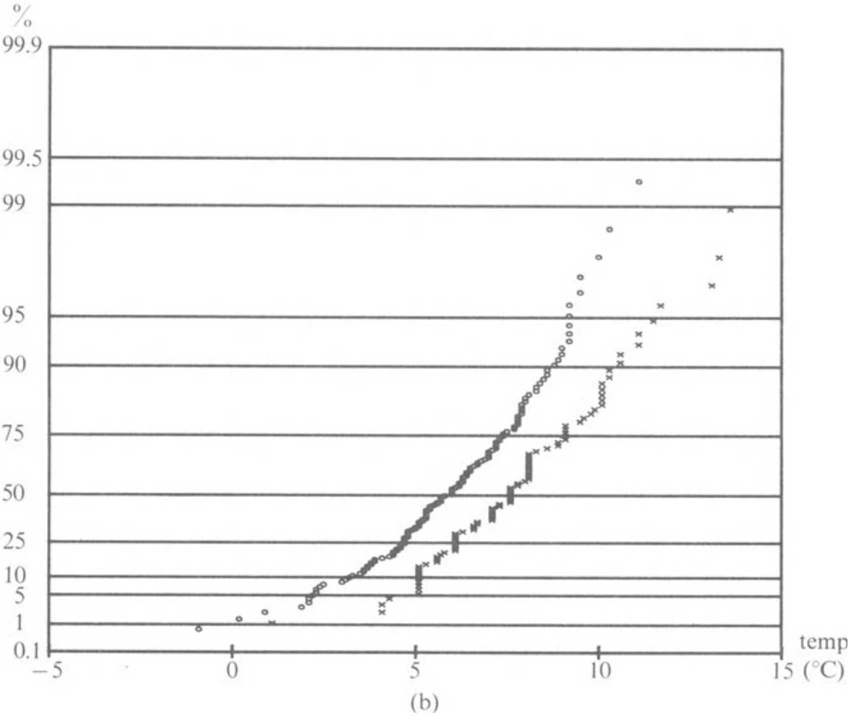


Figure 15.1.2. Observed d.f. of (a) maximum and (b) minimum temperature in Uppsala during July; X = discrete recordings, 1739–1838, O = continuous recordings, 1839–1981.

in 1709. With Anders Celsius, these measurements became more regular and there is an almost complete series of daily temperature data from January 1739.

During the first century of this period measurements were made only three or four times a day, and sometimes at irregular intervals. In 1839 observations of the true maxima and minima were made possible by the installation of a max- and min-thermometer. The observed daily maximum or minimum for the period 1739–1838 should therefore be expected to be less extreme than the recordings from 1839 onwards. This effect can be expected to be particularly large for the minimum temperature during the summer months, since the lowest temperatures then occur very early in the morning at that latitude, while the discrete recordings were normally made when the sun had been up for several hours. This is illustrated in Figure 15.1.1 which shows the observed *monthly* maxima and minima for the month of July for each year (except for a few missing years during the 1770s). The data consists of $M_{1739}, \ldots, M_{1838}$ (and $m_{1739}, \ldots, m_{1838}$), where, e.g. M_i is the maximum of the *discrete* set of July temperatures in year i, and of $M_{1839}, \ldots, M_{1981}$ (and $m_{1839}, \ldots, m_{1981}$), with M_j equal to the absolute (continuous) July maximum in year j. The d.f.'s of these data are plotted in Figure 15.1.2 with O for continuous and X for discrete recordings. There is a clear shift towards more extreme values with the introduction of the max- and min-thermometer, as is also seen, at least for minima, in the complete time series in Figure 15.1.1. The data are plotted on double exponential probability paper, but the fit is not particularly good, and there is no strong reason why it should be; the maximum temperature over a month being determined more by the prevailing weather conditions during that month than by any extreme weather that might occur in particular months.

The temperature data in this example have been extracted from the original manuscripts by Sverker Hellström, Uppsala, to whom we are grateful for permission to use them. □

15.2. Poisson Exceedances and Exponential Waiting Times

The Poisson character of exceedances or crossings of a high level, and the exponential waiting times between successive exceedances, is coupled to the weak dependence of maxima in adjacent intervals, as has been observed empirically in connection with rare events emerging from unlikely combinations of innocent causes, such as extreme floods or storm winds and extreme loads in mechanical structures.

Example 15.2.1 (High river flows). Todorovic (1979) has obtained the observed frequencies of $N(T)$, the number of times in T days, that the water flow in the Greenbrier River, West Virginia, exceeds the level 17,000 cubic ft.

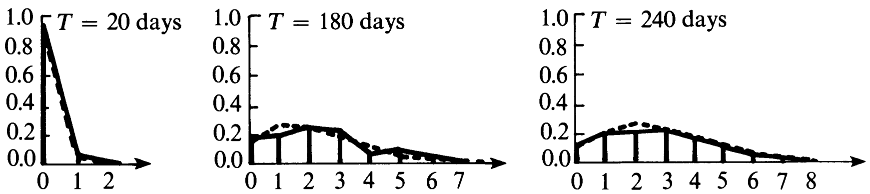


Figure 15.2.1. Observed (———) and fitted Poisson distribution (---) for the number of exceedances in time T; flood data from Todorovic (1979).

The total observation period was 72 years between 1896 and 1967. The observed frequences, together with theoretical Poisson probabilities are shown in Figure 15.2.1 for a few values of T.

In this type of river and climate, the Poisson distribution gives a reasonably good fit to the observed distribution. The time between exceedances is also large and the independence therefore plausible.

In other areas, with a climate that changes between wet and dry periods, the exceedances may occur in clumps and, of course, then deviate from the Poisson distribution. □

When no clumps of exceedances occur, one can use the approximation

$$P\{N(T) = 0\} \approx \exp\{-T\mu(u)\}, \tag{15.2.1}$$

where $\mu(u)$ is the mean number of exceedances of u per time unit ($\mu(u) = P\{\xi_i > u\}$ if time is discrete). In other cases, better approximations to the probability of no exceedances may be obtained from the crossing rate for a smooth, enveloping process.

However, even if (15.2.1) provides a good approximation to the probability of no exceedance, there remains the more difficult problem of assessing the form of the crossing intensity $\mu(u)$ as a function of u. In discrete time, $\mu(u) = P\{\xi_i > u\}$ is given by the marginal d.f., while in continuous time it is, by (7.2.3),

$$\mu(u) = f_{\xi(0)}(u)E(\xi'(0)^+ \mid \xi(0) = u) = \int_{z=0}^{\infty} z f_{\xi(0),\,\xi'(0)}(u, z)\, dz.$$

What is needed then, apart from the density $f_{\xi(0)}(u)$, is a measure of the average steepness of the sample functions at various levels. For normal processes, $\mu(u)$ is given by (7.3.4) and it can also be calculated exactly for a few non-normal processes, some of them functions of multivariate normal processes, such as the χ^2-process; see Belayev (1968), Belayev and Nosko (1969), Veneziano (1979), Hasofer (1976), and Lindgren (1980b,c).

Even if the mean crossing rate has a known functional form, some parameters have to be estimated from observations before the formula can be used in practice. For example, if $\xi(t)$ is a stationary normal process with mean m and variance λ_0,

$$\mu(u) = \frac{1}{2\pi}\left(\frac{\lambda_2}{\lambda_0}\right)^{1/2} \exp\left\{-\frac{1}{2\lambda_0}(u-m)^2\right\}$$

so there are three parameters, m, λ_0, and $\nu = (1/2\pi)(\lambda_2/\lambda_0)^{1/2}$ = mean number of upcrossings of the level m in unit time, to be estimated. Of course, for large values of $u - m$, predictions by means of this formula can be highly unreliable.

Example 15.2.2 (Break frequency of paper; cf. Example 14.2.1). A paper web which runs through a printing press is subject to a tension which can cause a web break. Since a break usually starts at one of the side edges of the web we can model this phenomenon by considering the *local strength* of the paper (at any of the side edges) as a stationary stochastic process $\xi(t)$ with one-dimensional parameter t running along the length of the paper.

In Example 14.2.1 the object of study was $\min\{\xi(t); 0 \leq t \leq T\}$ for small values of T (between 0.08 m and 10 m), and from that example we can draw the conclusion that the mean local strength is certainly not less than 1.5 kN/m for those particular qualities of paper.

In the model a web break starts when the local strength $\xi(t)$ has a downcrossing of the tension level u, which in a printing press takes values somewhere around 0.3–0.5 kN/m. Suppose now that for small values of u the downcrossing rate is the same as for a normal process,

$$\mu(u) = \nu \exp\left\{-\frac{1}{2\lambda_0}(u-m)^2\right\}.$$

(This is consistent with Example 14.2.1 which did not exclude a Type I decrease of mean strength with increasing length.) Further assume that we can run the press at a series of different tension levels. An experiment carried out on tension level u_i will result in an observed distance to the first web break, which we can denote by y_i. The tension can be varied between each experiment. We thus obtain a sequence of tension levels u_i, and observed corresponding distances to web breaks, y_i, $i = 1, \ldots, n$.

By the Poisson character of extreme crossings, for small values of u_i the y_i's are approximately independent and exponentially distributed with mean $1/\mu(u_i)$, so one can write

$$\log y_i = -\log \nu + \frac{1}{2\lambda_0}(u_i - m)^2 + \log e_i, \tag{15.2.2}$$

where $\log e_i$, $i = 1, \ldots, n$, are approximately independent r.v.'s with d.f. $P\{\log e_i \leq x\} = 1 - \exp(-e^x)$, $-\infty < x < \infty$. Thus (15.2.2) is a regression equation from which m, λ_0, and $\log \nu$ can be estimated. However, the levels

u_i need to be spread out over a wide range of values in order to give good estimates of the parameters and this is not always possible under experimental conditions; for a discussion of efficiency, see Hållberg and de Maré (1976). □

15.3. Domains of Attraction and Extremes from Mixed Distributions

In this section we shall discuss two important questions related to the domains of attraction, namely, the effect on the extreme value distribution of a randomly varying parameter in a parent distribution, and the influence of deterministic components in nonstationary cases. We start with a discussion of some models for wind variation.

Example 15.3.1. The variation in wind speed and direction has long been studied and documented in statistical terms and has an interesting history. Notable studies were made by Gustav Eiffel on top of the Eiffel tower during the early years after its construction; see Eiffel (1900). As an example, Figure 15.3.1(a) shows a recording of speed and direction during a day at the top of the tower (upper curve) and on the ground (lower curve).

The speed was measured as the time it takes the wind to run 5000 m, which means that there is an averaging over a time interval with a length depending on the current wind speed. This averaging was partly due to the construction of the measuring and recording devices and, as Eiffel noted, has a considerable influence on the observed high speeds. To partly compensate for this, Eiffel installed an extra, more sensitive, anemometer which was engaged only when the slow anemometer showed values exceeding a certain level. Figure 15.3.1(a) and (b) shows examples of the slow and sensitive meter recordings.

In Figure 15.3.2 empirical d.f.'s for wind speed on top of the tower are shown for two different months, (a) January and July, based on measurements from 1890 to 1895, and (b) for combined data for all twelve months in a year. Since there are some theoretical arguments for using a Rayleigh distribution in connection with speed measurements and random directions, the scales are chosen so that the Rayleigh distribution, $F(v) = 1 - \exp(-v^2/2\sigma^2)$, is represented by a straight line:

$$\log(-\log(1 - F(v))) = 2 \log v - \log(2\sigma^2).$$

As is seen, the curves for the individual months show some definite curvature, while the combined data show a good fit to the Rayleigh distribution. This has sometimes been taken as an indication that wind speed should be described by the Rayleigh distribution. As we shall see, there is at least some further rationale for this. □

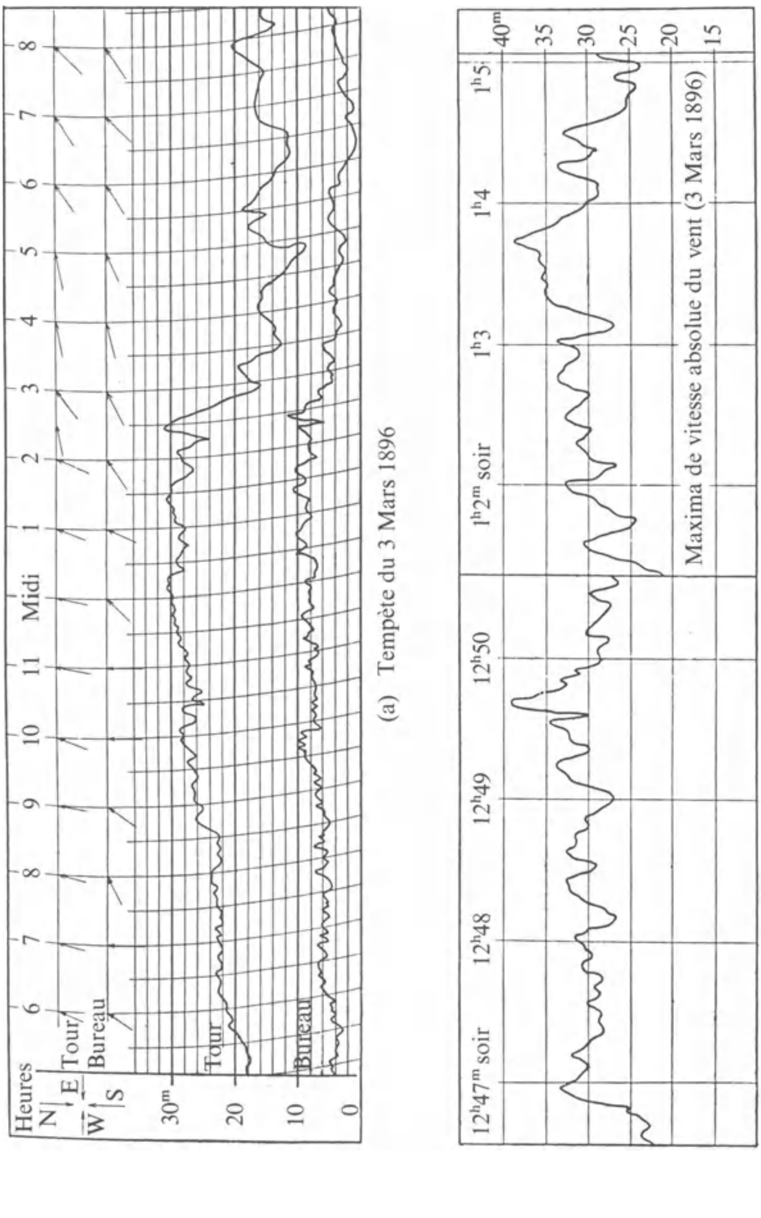


Figure 15.3.1. G. Eiffel's registration of wind speed, (a) with slow anemometer and (b) with sensitive anemometer showing extreme winds.

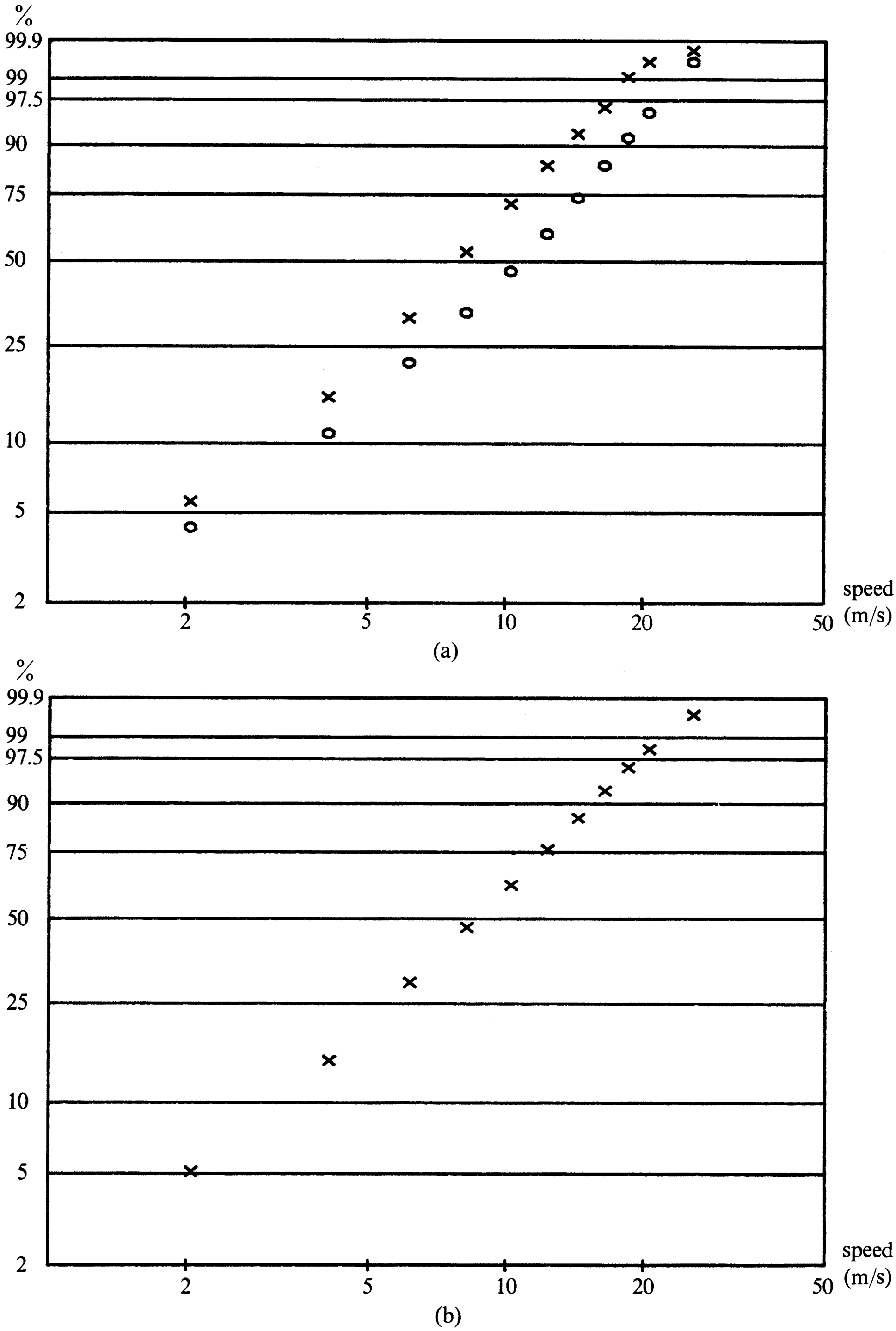


Figure 15.3.2. Observed d.f. of wind speed on the Eiffel tower (scales chosen to give a straight line for the Rayleigh distribution); (a) ○ = January, × = July, (b) × = Yearly data.

The horizontal wind blowing at a specific point at time t can be represented as a vector $W(t) = (\xi_1(t), \xi_2(t))$ with components $\xi_1(t)$ and $\xi_2(t)$ which are the wind velocities along, say, the N–S and E–W directions, respectively. The total wind speed is then

$$\xi(t) = |W(t)| = (\xi_1^2(t) + \xi_2^2(t))^{1/2}$$

(with wind direction $\theta(t)$, i.e. $\xi_1(t) = \xi(t)\cos\theta(t)$, $\xi_2(t) = \xi(t)\sin\theta(t)$).

Suppose $\xi_1(t)$ and $\xi_2(t)$ are independent normal processes with mean m_1, m_2 and the same covariance function $r(t) = \operatorname{Cov}(\xi_i(s), \xi_i(s+t))$, $i = 1, 2$, and variance $\sigma^2 = r(0) = \operatorname{Var}(\xi_i(t))$. If both components have mean zero, $m_1 = m_2 = 0$, then

$$P\{\xi(t) > v\} = P\{\xi_1^2(t) + \xi_2^2(t) > v^2\} = \exp\left(-\frac{v^2}{2\sigma^2}\right), \qquad v > 0,$$

so that $\xi(t)$ has a Rayleigh distribution with density

$$f_{\xi(t)}(v) = v\sigma^{-2}\exp\left(-\frac{v^2}{2\sigma^2}\right), \qquad v > 0. \tag{15.3.1}$$

If m_1 or m_2 are not zero, i.e. if there is a prevailing wind direction, the wind speed has a noncentral Rayleigh distribution with parameter $\lambda = (m_1^2 + m_2^2)^{1/2}$ and probability density function

$$\begin{aligned} f_{\xi(t)}(v) &= v\sigma^{-2}\exp\left(-\frac{\lambda^2 + v^2}{2\sigma^2}\right)\sum_{j=0}^{\infty}\frac{(\lambda^2 v^2/4\sigma^4)^j}{(j!)^2} \\ &= v\sigma^{-2}\exp\left(-\frac{\lambda^2 + v^2}{2\sigma^2}\right) I_0\left(\frac{\lambda v}{\sigma^2}\right), \qquad v > 0, \end{aligned} \tag{15.3.2}$$

where $I_0(x)$ is the modified Bessel function of the first kind of order 0; see Johnson and Kotz (1970, Section 28.3).

Due to these relations between the normal and the Rayleigh distribution it is not surprising that the latter has been used as a parent distribution for wind speed measurements by many authors. However, in the example above (Example 15.3.1) the Rayleigh distribution did not fit the monthly data well.

Note also that if $\xi(t) = (\xi_1^2(t) + \xi_2^2(t))^{1/2}$ is a Rayleigh process with normal components, then the averages

$$\bar{\xi}_i(T) = \frac{1}{T}\int_0^T \xi_i(t)\,dt$$

are still normal, so that

$$\bar{\xi}(t) = (\bar{\xi}_1^2(t) + \bar{\xi}_2^2(t))^{1/2}$$

is also Rayleigh, but the observed mean of $\xi(t)$,

$$\bar{\xi}(t) = \frac{1}{T}\int_0^T (\xi_1^2(t) + \xi_2^2(t))^{1/2}\,dt$$

is not Rayleigh. The distributional properties of filtered wind speed data has been studied by Sharpe (1974).

We turn now to a discussion of extreme wind data and the effect of nonstationarity and time-varying parameters. The Rayleigh distributions (both central and noncentral) belong to the domain of attraction for the Type I double exponential extreme value distribution. One would therefore expect that the maximum wind $M(T)$, taken over an interval of length T where stationary conditions hold, would follow this distribution, i.e.

$$P\{M(T) \le u\} \simeq \exp(-\,e^{-a_T(u-b_T)})$$

for some constants $a_T > 0$, b_T depending on the length of measurement and the correlation structure of the process.

Example 15.3.2 (Yearly maximum of wind speed). Table 15.3.1 and Figure 15.3.3 show the observed yearly maximum of the 1-hour mean wind in London, Ontario for the years 1939–1961 ($T = 1$ year, unit of speed = m/s). (The data have been compiled from Davenport (1978).) The straight line in the figure is the d.f. for a double exponential distribution as fitted by Davenport: $F(v) = 1 - \exp(-e^{-(v-17)/3})$. □

As was seen in the previous example a Type I extreme value distribution can be reasonably well fitted to the yearly maxima of wind at a specific point. On the other hand, in Example 15.3.1 different short-term distributions had to be fitted for each month, and there is no reason to believe that

Table 15.3.1

Speed	Year	Speed	Year
14.3	1958	18.3	1947
14.8	1944	18.3	1949
15.2	1960	18.8	1954
15.6	1946	19.2	1959
16.1	1939	20.1	1941
16.1	1956	20.1	1955
16.5	1940	20.6	1952
16.5	1945	22.4	1942
17.4	1943	23.2	1948
17.4	1951	24.6	1957
17.4	1953	25.9	1950
17.4	1961		

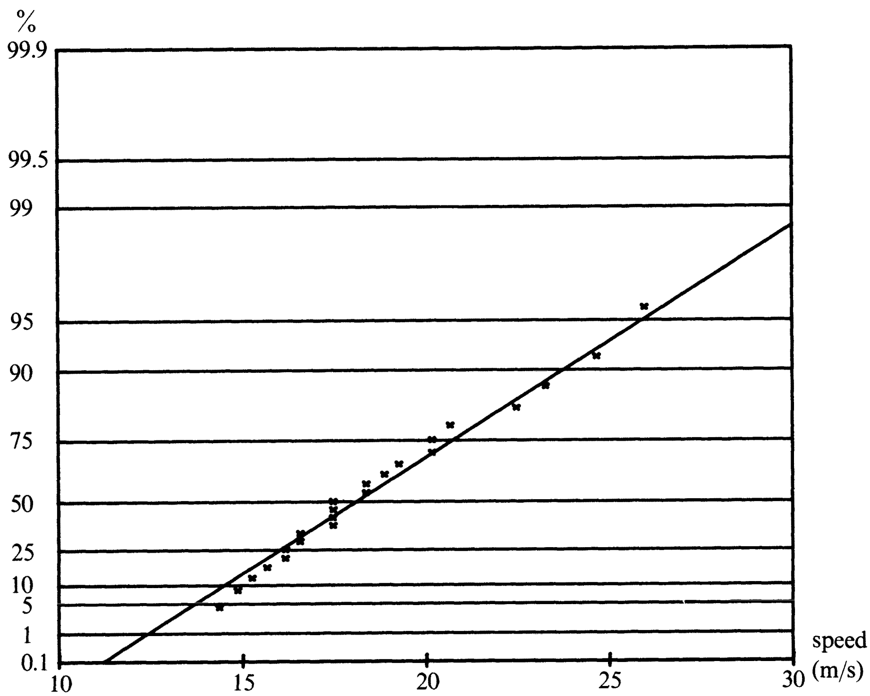


Figure 15.3.3. Observed d.f. of maximum yearly wind speed 1939–61, plotted on double exponential probability paper (from Davenport (1977)).

it is not also the case in Example 15.3.2. There are many other situations, e.g. in modelling rainfall, waveheight, etc., when one has to use different models for different periods of time, and let the parameters of the model vary with time. This variation can then be considered either as deterministic, repeating a certain pattern from one period of time to the next, or as a random function with its own distributional properties. The former case is parallel to the considerations leading to formula (14.2.5) for the strength of materials, and it involves taking the maximum of variables with nonidentical distributions. A theory for normal sequences with varying means was presented in Chapter 6, and an example will be given at the end of the next section.

In the latter case, when the variation is random, one is actually facing a situation in which the individual observations are identically distributed, and follow some *mixed distribution*, in analogy with formula (14.2.6). An important question then is to what extent extreme observations are due to extreme parameter values or to extreme experimental outcomes.

Simiu and Filliben (1976) have demonstrated that the double exponential distribution does not describe extreme winds in climates characterized by the occurrence of special types of winds, such as hurricanes, which are considerably stronger than the usual winds. They conclude that in such cases knowledge of the frequency and characteristics of hurricanes is vital for a reliable prediction of strong winds. Similar problems occur in ocean engineering where the frequency of different weather types is used to mix the effects calculated from stationary short-term wind and wave models.

As an example we will discuss in more detail a scale mixture F_ξ of Rayleigh distributions,

$$\begin{aligned} 1 - F_\xi(v) = P\{\xi > v\} &= \int_0^\infty P\{\xi > v \mid \sigma = s\}\, dF_\sigma(s) \\ &= \int_0^\infty \exp\left(-\frac{v^2}{2s^2}\right) dF_\sigma(s), \end{aligned} \tag{15.3.3}$$

where F_σ is the d.f. of a random scale parameter in (15.3.1) (assuming $m_1 = m_2 = 0$ to be constant).

As follows simply from Theorem 1.6.1, each conditional distribution $1 - \exp(-v^2/2s^2)$ belongs to the domain of attraction to the Type I extreme value distribution, which means that the maxima, suitably normalized have approximately a double exponential distribution. When s varies as the r.v. σ, extreme values of ξ may occur due to large values of σ, and the question we shall deal with is the domain of attraction for the mixed distribution (15.3.3) for different mixing distributions F_σ. In the Type II case there is the following simple and satisfying answer: A scale mixture F_ξ of Rayleigh distributions belongs to the domain of attraction for the Type II extreme value distribution with parameter α if and only if the same is true for the scale distribution F_σ.

Furthermore, if $\{a_n > 0\}$ is a sequence of scale parameters for F_σ such that, as $n \to \infty$, $F_\sigma(x/a_n)^n \to \exp(-x^{-\alpha})$, $x > 0$, and $c_\alpha = 2^{\alpha/2}\Gamma(\alpha/2 + 1)$, then $a'_n = a_{[c_\alpha n]}$ is an appropriate sequence of scale parameters for F_ξ, and

$$F_\xi\left(\frac{x}{a'_n}\right)^n \to \exp(-x^{-\alpha}) \quad \text{as} \quad n \to \infty, x > 0.$$

We shall prove this here making use of the explicit form of the Rayleigh distribution, even though it seems likely that a direct estimate of the tail of the distribution of the mixture would work just as well. Let

$$F_{1/(2\sigma^2)}(x) = 1 - F_\sigma\left(\frac{1}{\sqrt{2x}}\right)$$

be the d.f. of $1/(2\sigma^2)$ and let

$$L(t) = \int_0^\infty e^{-tx}\, dF_{1/(2\sigma^2)}(x)$$

be its Laplace transform, so that

$$1 - F_\xi(v) = \int_0^\infty e^{-v^2 x}\, dF_{1/(2\sigma^2)}(x) = L(v^2).$$

Now, by Theorem 1.6.2, F_σ belongs to the domain of attraction for the Type II extreme value distribution if and only if it is regularly varying with exponent α, i.e. if

$$\frac{1 - F_\sigma(tx)}{1 - F_\sigma(t)} \to x^{-\alpha} \quad \text{as } t \to \infty,$$

and hence with $s = 1/(2t^2)$,

$$\frac{F_{1/(2\sigma^2)}(s/x^2)}{F_{1/(2\sigma^2)}(s)} \to (x^2)^{-\alpha/2} \quad \text{as } s \to 0,$$

so that $F_{1/(2\sigma^2)}(s)$ is regularly varying at 0 with exponent $\alpha/2$. By a Tauberian theorem for Laplace transforms (see Feller (1971), Section XIII.5, Theorem 3, formula (5.7)), this is equivalent to $L(t)$ being regularly varying at ∞ and

$$L(t) \sim \Gamma\left(\frac{\alpha}{2} + 1\right) F_{1/(2\sigma^2)}(1/t) \quad \text{as } t \to \infty.$$

Thus

$$\begin{aligned} 1 - F_\xi(v) = L(v^2) &\sim \Gamma\left(\frac{\alpha}{2} + 1\right) F_{1/(2\sigma^2)}\left(\frac{1}{v^2}\right) \\ &= \Gamma\left(\frac{\alpha}{2} + 1\right)\left(1 - F_\sigma\left(\frac{v}{\sqrt{2}}\right)\right) \\ &\sim 2^{\alpha/2}\,\Gamma\left(\frac{\alpha}{2} + 1\right)(1 - F_\sigma(v)) \quad \text{as } v \to \infty, \end{aligned} \tag{15.3.4}$$

proving the equivalence of domains of attraction.

Further, $F_\sigma(x/a_n)^n \to \exp(-x^{-\alpha})$ if and only if

$$n(1 - F_\sigma(x/a_n)) \to x^{-\alpha}.$$

With $a_n' = a_{[c_\alpha n]}$, (15.3.4) then shows that

$$\begin{aligned} n(1 - F_\xi(x/a_n')) &\sim [c_\alpha n] c_\alpha^{-1}\left(1 - F_\xi\left(\frac{x}{a_{[c_\alpha n]}}\right)\right) \\ &\sim [c_\alpha n]\left(1 - F_\sigma\left(\frac{x}{a_{[c_\alpha n]}}\right)\right) \to x^{-\alpha}, \end{aligned}$$

so that a_n' is a proper choice of scale parameters.

Thus in the Type II case with F_σ falling off regularly at infinity i.e. $1 - F_\sigma(x) \sim x^{-\alpha}L(x)$ and $L(x)$ varying slowly, the maximum $M_n^{(\xi)}$ of the independent variables $\xi_1, \ldots, \xi_n$, each with d.f.

$$1 - \int_0^\infty \exp\left(-\frac{v^2}{2s^2}\right) dF_\sigma(s),$$

has an asymptotic Type II distribution

$$P\{a_n' M_n^{(\xi)} \le x\} \to \exp(-x^{-\alpha}).$$

In this case the type of the maximum is therefore determined by the domain of attraction of the d.f. F_σ of the parameter values.

15.4. Extrapolation of Extremes Over an Extended Period of Time

Suppose $\{\xi(t)\}$ is a stationary continuous time process, for which the maximum is attracted to the double exponential law, so that with some $a_T > 0, b_T$,

$$P\{M(T) \le u\} = P\left\{\sup_{0 \le t \le T} \xi(t) \le u\right\} \approx \exp(-e^{-a_T(u-b_T)}) \quad (15.4.1)$$

and we want to use knowledge of a_T, b_T, for some T and "extrapolate" to the distribution of $M(nT)$ for some fixed integer n. If T is large we can expect that the maxima over the intervals $(0, T]$, $(T, 2T], \ldots, ((n-1)T, nT]$ are approximately independent, so that

$$\begin{aligned} P\{M(nT) \le u\} &\approx P^n\{M(T) \le u\} \approx \exp(-ne^{-a_T(u-b_T)}) \\ &= \exp(-e^{-a_T(u-b_T-a_T^{-1}\log n)}), \end{aligned}$$

and hence we would have

$$\begin{aligned} a_{nT} &= a_T, \\ b_{nT} &= b_T + a_T^{-1} \log n \end{aligned} \quad (15.4.2)$$

as a possible choice for the location and scale parameter for the maximum over an extended period of time; cf. the discussion of the size effect on strength in Section 14.2.

These relations hold as soon as $\{\xi(t)\}$ belongs to the domain of attraction of the double exponential law. For a differentiable normal process

$$\begin{aligned} a_T &= (2 \log T)^{1/2}, \\ b_T &= a_T + a_T^{-1} \log(\sqrt{\lambda_2}/2\pi) \end{aligned} \quad (15.4.3)$$

by Theorem 8.2.7, where $\lambda_2 = \operatorname{Var}(\xi'(t))$, which gives, as $T \to \infty$,

$$\begin{aligned} a_{nT} &= (2 \log nT)^{1/2} = (2 \log T)^{1/2}\left(1 + \frac{1}{2}\frac{\log n}{\log T} + o\left(\frac{1}{\log T}\right)\right) \\ &= a_T + a_T^{-1} \log n + o(a_T^{-1}), \\ b_{nT} &= a_{nT} + a_{nT}^{-1} \log(\sqrt{\lambda_2}/2\pi) \\ &= a_T + a_T^{-1} \log n + a_T^{-1} \log(\sqrt{\lambda_2}/2\pi) + o(a_T^{-1}) \\ &= b_T + a_T^{-1} \log n + o(a_T^{-1}), \end{aligned}$$

agreeing with (15.4.2) up to terms of low order, as it should.

However, these scale and location parameters are dependent on the time-scale of the process, in a way which sometimes can be a source of confusion. A change of scale units in (15.4.1) and (15.4.3), so that time T is replaced by $T' = cT$, will change $(2 \log T)^{1/2}$ into $(2 \log T + 2 \log c)^{1/2}$ and $\lambda_2^{1/2}$ into $c^{-1}\lambda_2^{1/2}$, so the approximation (15.4.1) will be different. Of course, whatever the time scale, there is a precise limiting distribution for the maximum, the problem is that the error in using (15.4.1) for finite time intervals depends on the time scale.

One natural way of counting time, which is often used, is in terms of zero crossing distances or mean period length. Let (assuming $\lambda_0 = 1$),

$$v = \frac{1}{2\pi}\lambda_2^{1/2}$$

be the mean number of zero upcrossings per (old) time unit, and introduce the new time scale $T' = vT$, which counts the time in terms of the expected number of zero upcrossings. Then $\log(v^{-1}\lambda_2^{1/2}/2\pi) = \log 1 = 0$, so the (standardized) approximation would be

$$P\{M(T) \le u\} \approx \exp(-e^{-(2 \log vT)^{1/2}(u-(2 \log vT)^{1/2})})$$

with equal scale and location parameter,

$$a_T = b_T = (2 \log vT)^{1/2}. \tag{15.4.4}$$

The approximating distribution has mean and standard deviation

$$m_T = (2 \log vT)^{1/2} + \frac{\gamma}{(2 \log vT)^{1/2}},$$

$$\sigma_T = \frac{\pi}{\sqrt{6}}(2 \log vT)^{1/2},$$

which can be used as standard approximations of the limiting mean and standard deviations.

Example 15.4.1 (Extremes in air pollution data). The United States Federal short-term standard for sulfur dioxide (SO_2) requires that the 3-*hour mean concentration* of SO_2 should not exceed 50 pphm (parts per hundred million) more than once a year. To see how this can be complied with we shall discuss data from 19 years observations of 1-*hour* mean concentrations from Long Beach, California (taken from Roberts (1979b)). Clearly the 3-hour means have less accentuated extremes than the 1-hour means so this procedure will certainly not underestimate the frequency of exceedances.

Figure 15.4.1 shows the daily maxima of the hourly averages for each of the 365 days of 1979.

As is seen in the diagram the data are clearly correlated. Furthermore, from a complete set of data from 1956–1974, shown in Table 15.4.1, it is seen

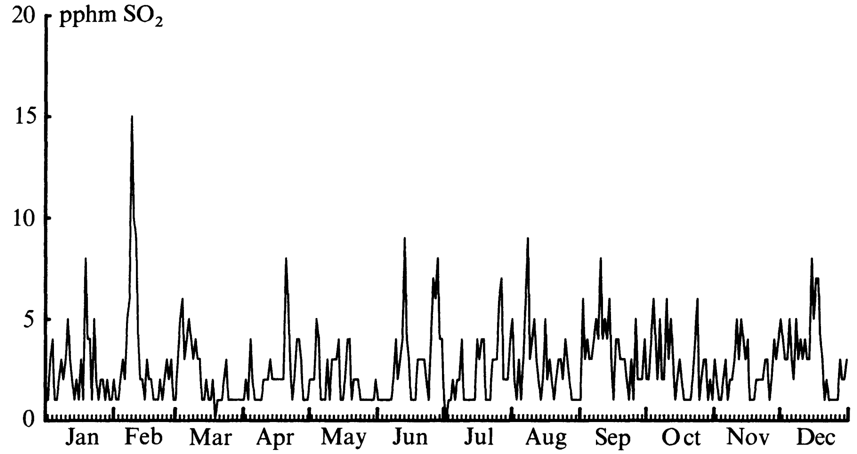


Figure 15.4.1. Observed values of daily maximum of 1-hour mean SO_2 concentration at Long Beach, California, 1979.

that the highest values are concentrated to the winter months, so that what we actually have is an example of a nonstationary correlated sequence.

The d.f. of mean concentration of air pollutants has been studied and described by many authors, and it appears that a lognormal distribution would

Table 15.4.1 Sulfur dioxide, 1-hour average concentrations (pphm); monthly and annual maxima and annual averages.

Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Ann. Max.	Ann. Ave.
1956	47	31	44	12	13	3	14	21	33	26	40	32	47	4.0
1957	22	19	20	32	20	23	18	16	13	14	41	25	41	3.0
1958	15	13	20	12	24	13	37	20	32	27	27	68	68	3.4
1959	20	32	20	15	3	6	8	15	17	15	20	20	32	2.1
1960	22	18	23	20	8	13	14	9	13	16	27	20	27	1.9
1961	25	20	20	16	10	10	8	10	12	16	14	43	43	1.9
1962	20	13	15	18	10	12	10	10	11	11	14	7	20	1.5
1963	12	18	27	21	2	7	4	4	15	10	18	18	27	1.3
1964	16	10	3	3	19	9	16	25	4	14	18	21	25	1.4
1965	16	18	9	14	8	10	18	18	14	12	17	14	18	2.6
1966	27	33	25	10	17	30	13	18	22	15	25	23	33	3.0
1967	30	40	32	10	8	7	8	26	10	40	18	17	40	2.5
1968	51	30	18	22	10	19	22	25	26	29	50	40	51	3.1
1969	37	13	55	14	9	10	13	17	33	13	15	44	55	2.5
1970	23	19	10	11	15	12	25	40	25	20	12	8	40	2.4
1971	22	36	20	28	10	15	20	55	38	41	26	25	55	2.5
1972	30	32	18	27	37	13	23	19	21	31	25	13	37	2.5
1973	10	8	8	12	11	16	25	16	11	28	10	23	28	1.9
1974	8	9	9	13	8	14	9	9	25	11	19	15	34	1.7
Average	23.8	21.6	20.8	16.3	12.7	13.7	16.0	19.6	19.7	20.4	22.9	25.0		2.4

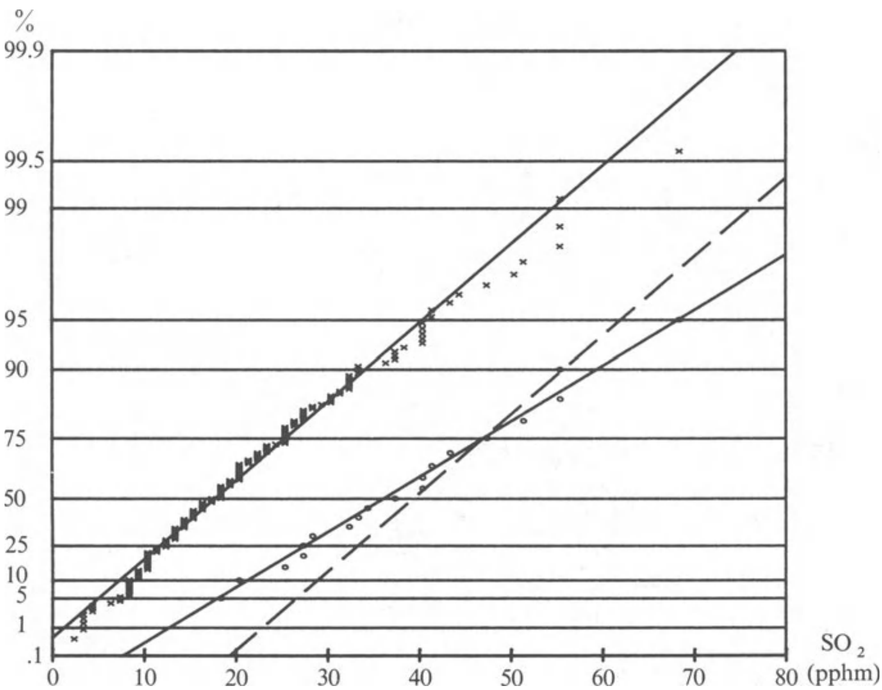


Figure 15.4.2. Observed d.f. of monthly maxima (×), yearly maxima (O) from 1956 to 1974 at Long Beach; —— fitted double exponential distributions, - - - d.f. for yearly maximum extrapolated from monthly maxima.

give a reasonable fit, at least in the center of the distribution. The lognormal distribution belongs to the domain of attraction to the Type I distribution (see Example 1.7.4), and we shall try to fit a double exponential distribution to the extremal data in Table 15.4.1. Of course, a good fit to lognormality in the centre of the distribution does not ensure that this will work well.

Figure 15.4.2 shows the empirical d.f.'s of the 19-yearly, and $19 \times 12 = 228$ monthly maxima, plotted on double exponential probability paper. It is worth noting that the theory in Chapters 3 and 13 might be applicable to the correlated data here and that then the data should give an acceptable plot on the appropriate extreme value paper. However, the non-stationarity, and perhaps also the dependence from month to month, affects the choice of scale and location parameters in the approximation formula (15.4.1) as will now be demonstrated.

Assuming a double exponential distribution for M(month) and M(year), one can estimate parameters, e.g. by the maximum-likelihood method. Roberts (1979a,b), from which the complete set of data has been taken, uses a variant of the least-squares method and obtains the estimated distributions

$$P\{M(\text{month}) \le u\} = \exp(-e^{-0.115(u-14.5)}),$$

$$P\{M(\text{year}) \le u\} = \exp(-e^{-0.081(u-31.5)}).$$

The estimated probability that the 1-hour averages exceed 50 pphm at least once in one year is

$$P\{M(\text{year}) > 50\} = 1 - \exp(-e^{-0.081(50-31.5)}) = 0.2,$$

while the observed frequency is $4/19 \approx 0.21$. This should be compared to (15.4.2) from which one would expect

$$a_{\text{year}} = a_{\text{month}} = 0.115,$$

$$b_{\text{year}} = b_{\text{month}} + a_{\text{month}}^{-1} \log 12 = 14.5 + \frac{2.48}{0.115} = 36.1$$

Using a_{month} and b_{month} to estimate the probability of exceeding 50 pphm during one year, would then give

$$P\{M(\text{year}) > 50\} \approx 1 - \exp(-e^{-0.115(50-36.1)}) = 0.18,$$

and the correspondingly estimated d.f. of M(year) is represented by the dashed line in Figure 15.4.2.

As seen from Figure 15.4.2, the location of the extrapolated distribution agrees reasonably well with that of F_{year} but at the same time it is more concentrated, so that the probability of very high or very low yearly maxima will be underestimated. It is not clear whether this is caused by the nonstationarity or by the strong correlation (or by both), but it is evident from the data in Table 15.4.1 that there are several runs of very low or very high values within individual years, and that high yearly maxima tend to encourage several high monthly maxima that same year. □

As was seen in Chapter 6, for normal sequences the "stationary" theory still applies to sequences with nonstationary mean or correlation structure, provided the location parameter is adequately adjusted and the correlations show the standard $\log^{-1}$-decay with time.

Example 15.4.2 (Nonstationary ozone data). Horowitz (1980) has applied extreme-value theory to correlated, nonstationary ozone data. Assuming daily maximal 1-hour mean concentrations η_i to be lognormal, he allows for a time-dependent mean value function over the year, with residuals which are correlated normal variables with zero means and common variance σ^2. Normalizing the mean in terms of σ we obtain the model

$$\log \eta_i = \sigma(m_i + \xi_i), \qquad i = 1, \ldots, n\,(=365), \tag{15.4.5}$$

where a second degree polynomial is fitted for m_i and ξ_i are assumed to be correlated standard normal variables. As Figure 15.4.3 shows there is clearly a need for a time-dependent model, and the model (15.4.5) at least gives a good fit for the marginal distribution of η_i.

By Theorem 6.2.1, under mild conditions on the function m_i we then have that

$$P\left\{\max_{1 \le i \le n} \log \eta_i \le \sigma\left(\frac{x}{a_n} + b_n + m_n^*\right)\right\} \to \exp(-e^{-x}),$$

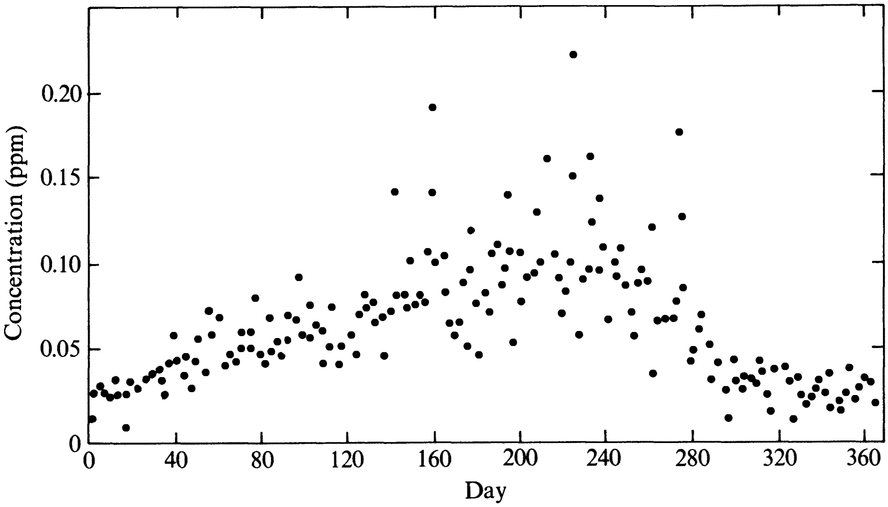


Figure 15.4.3. Observed daily maximum 1-hour average ozone concentrations.

where m_n^* is defined by (6.2.2). Writing $d_n = \exp(\sigma(b_n + m_n^*))$, $c_n = a_n/(\sigma d_n)$,

$$\exp\left(\sigma\left(\frac{x}{a_n} + b_n + m_n^*\right)\right) = d_n\left(1 + \frac{\sigma x}{a_n} + o(a_n^{-1})\right) = d_n + \frac{x}{c_n} + o(c_n^{-1}),$$

and hence

$$P\left\{\max_{1 \le i \le n} \eta_i \le \frac{x}{c_n} + d_n\right\} \approx \exp(-e^{-x})$$

for large n.

This is in sharp contrast to the extremal distribution that would appear for the yearly maximum of a sequence of (dependent) *identically* distributed variables, each with a marginal distribution equal to the observed "pooled" dis ribution of the 365 daily values, not accounting for a *nonstationary mean.*

It is worth noting that the application of the normal extreme value theory to this nonstationary situation is not restricted by an assumption of stationary correlation structure, as was shown by the results of Sections 6.2–6.3. □

15.5. Local Extremes—Application to Random Waves

In this last section we shall discuss some applications of continuous time extremal theory which relates to the behaviour of the process as a sequence of *random waves*, in particular, the sequence of successive local maxima and minima.

Let $\xi(t)$ be a continuously differentiable normal process, twice differentiable in quadratic mean, with local maxima at s_i, $0 \le s_1 < s_2 < \cdots < s_{N'} \le T$, where $N' = N'(T)$ is the (random) number of local maxima in the interval $[0, T]$ as in Section 7.6. Further, let s_i' be the locations of the local minima, indexed to make $s_i < s_i' < s_{i+1}$. For example, thinking of $\xi(t)$ as describing the height of the sea level above the mean level at a specific point at time t, a natural terminology is to call $\xi(s_1), \xi(s_1'), \xi(s_2), \ldots$ a sequence of random wave characteristics. The values $\xi(s_i), \xi(s_i')$ are the *apparent amplitudes*, while the drops in height, $\xi(s_i) - \xi(s_i')$, are the *apparent waveheights*. The extreme values of $\xi(s_i)$ and $\xi(s_i) - \xi(s_i')$ are of importance in many fields, marine sciences providing one prominent example.

It is sometimes easier to obtain reliable measurements of the waveheights and amplitudes than to get continuous measurements of $\xi(t)$ itself. Since

$$\max(\xi(s_1), \ldots, \xi(s_{N'})) = \sup\{\xi(t); s_1 \le t \le s_{N'}\},$$

which equals $M(T) = \sup\{\xi(t); 0 \le t \le T\}$ if $M(T)$ is not attained at 0 or T, one can nevertheless use the continuous time extremal results to obtain the asymptotic distribution of the maximum of $\xi(s_i)$, $i = 1, 2, \ldots$

We now specialize to a stationary normal process with $\lambda_4 = \mathrm{Var}(\xi''(t)) < \infty$ and general variance $\lambda_0 = \mathrm{Var}(\xi(t))$. This requires only obvious changes in previous standardized ($\lambda_0 = 1$) formulae since $\tilde{\xi}(t) = \lambda_0^{-1/2}\xi(t)$ satisfies $\mathrm{Var}(\tilde{\xi}(t)) = 1$, $\mathrm{Var}(\tilde{\xi}'(t)) = \lambda_2/\lambda_0$, $\mathrm{Var}(\tilde{\xi}''(t)) = \lambda_4/\lambda_0$. If, as in Section 7.6, $N_u'(T)$ denotes the number of local maxima s_i in $[0, T]$ such that $\xi(s_i) > u$, then by (7.6.3) with $\nu = (1/2\pi)(\lambda_2/\lambda_0)^{1/2}$, $\nu' = (1/2\pi)(\lambda_4/\lambda_2)^{1/2}$, $\varepsilon = (1 - (\nu/\nu')^2)^{1/2} = (1 - \lambda_2^2/\lambda_0\lambda_4)^{1/2}$,

$$E(N_u'(T)) = T\left\{\nu'\left(1 - \Phi\left(\frac{u}{\varepsilon\sqrt{\lambda_0}}\right)\right) + \nu \exp\left(-\frac{u^2}{2\lambda_0}\right)\Phi\left(\frac{u}{\varepsilon\sqrt{\lambda_0}}\frac{\nu}{\nu'}\right)\right\}.$$

Since the total expected number of maxima is $T\nu'$, the ratio $E(N_u'(T))/T\nu'$ may be regarded as a measure of how likely it is that a local maximum is greater than u, and we define the function $F_{\max}$ by

$$F_{\max}(u) = 1 - \frac{E(N_u'(1))}{\nu'}.$$

Evidently $F_{\max}(u)$ is nondecreasing, continuous, and $F_{\max}(u) \to 0$ (or 1) as $u \to -\infty$ (or $+\infty$) so that $F_{\max}$ is a d.f. Write

$$\begin{aligned} f_{\max}(u) &= \frac{d}{du} F_{\max}(u) \\ &= \frac{\varepsilon}{\sqrt{\lambda_0}} \phi\left(\frac{u}{\varepsilon\sqrt{\lambda_0}}\right) \\ &\quad + \left(\frac{1 - \varepsilon^2}{\lambda_0}\right)^{1/2} \frac{u}{\sqrt{\lambda_0}} \exp\left(-\frac{u^2}{2\lambda_0}\right)\Phi\left((1 - \varepsilon^2)^{1/2} \frac{u}{\varepsilon\sqrt{\lambda_0}}\right) \end{aligned}$$

for the corresponding density function. If the process $\xi(t)$ is ergodic $N_u'(T)/N'(T) \to 1 - F_{\max}(u)$, with probability one (cf. Section 10.2), so that

$F_{\max}(u)$ is in that case the limit of the empirical distribution of the apparent amplitudes. The form of this distribution depends only on the so-called *spectral width parameter* $\varepsilon = (1 - \lambda_2^2/\lambda_0\lambda_4)^{1/2}$. Values of ε near 0 give a narrow banded spectrum, with one dominating frequency, while values ε near 1 give a broad-band spectrum.

Example 15.5.1. The apparent amplitudes of sea level were observed by a floating buoy off South Uist in the Hebrides. During several storms the empirical distribution of the local wave maxima and minima were recorded and the spectral parameters estimated. Figure 15.5.1(a)–(c) shows the result of one such recording with moderate spectral width. The figures show in histogram form the distribution of the sea level $\xi(t)$ sampled twice per second, the local maxima $\xi(s_i)$, and the local minima $\xi(s_i')$. The density function $f_{\max}(u)$ (and the corresponding $f_{\min}(u) = f_{\max}(-u)$ for minima) based on normal process theory is also shown, with $\varepsilon = 0.617$ estimated from data. The observation time was $T = 17.4$ min.

As is seen, the current sea level $\xi(t)$ is reasonably normal in this example, and it is striking how well the theoretical densities $f_{\max}(u)$ and $f_{\min}(u)$ describe the observed amplitudes, although they are fitted only via indirect estimation of the spectral moments $\lambda_0, \lambda_2, \lambda_4$. □

The apparent waveheight $\xi(s_i) - \xi(s_i')$ is frequently studied in oceanographic surveys, but no closed form expression for its density is known. Some simple, and for many processes, reliable approximations have been given by Lindgren and Rychlik (1982), and by Cavanie *et al.* (1976).

In oceanography the severity of the seas is often described in terms of the *significant waveheight* or the *significant amplitude*. Let $u_{1/3}$ be such that

$$1 - F_{\max}(u_{1/3}) = 1/3,$$

so that one-third of the amplitudes exceed $u_{1/3}$, and define

$$A_s = 3\int_{u_{1/3}}^{\infty} u f_{\max}(u)\,du$$

to be the mean of the one-third highest waves. Then A_s is called the significant amplitude. The significant waveheight H_s is defined similarly, and one often assumes $H_s = 2A_s$, an approximation which seems to be good if ε is small, i.e. $v \approx v'$ so that there is only one local maximum between every zero upcrossing. Furthermore, for small values of ε,

$$A_s \approx \sqrt{\lambda_0}\left(\sqrt{2\log 3} + 3\sqrt{2\pi}\left(1 - \Phi\left(\sqrt{2\log 3}\right)\right)\right) \approx 2.00\sqrt{\lambda_0}. \quad (15.5.1)$$

For any given significant amplitude one can ask for the maximum amplitude one is likely to encounter over a period of time T. Since, as noted above, $M_{N'} = \max\{\xi(s_1), \ldots, \xi(s_{N'})\} = \sup\{\xi(t); 0 \le t \le T\} = M(T)$, except when $M(T)$ is attained in $[0, s_1)$ or in $(s_{N'}, T]$, this can be solved within the present theory.

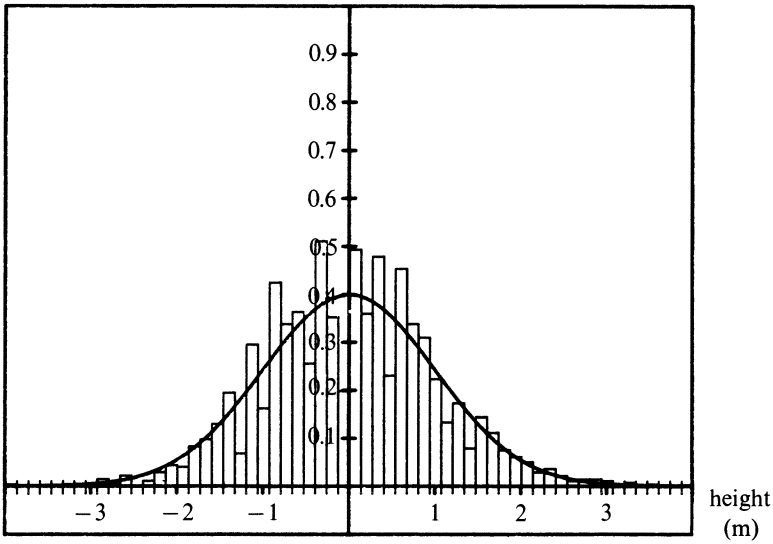


Figure 15.5.1(a). Histogram of observed water level; estimated spectral width parameter $\varepsilon = 0.617$.

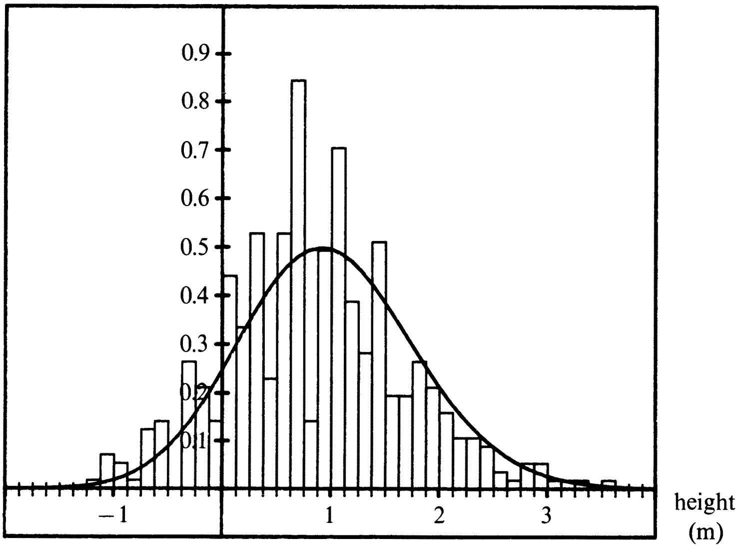


Figure 15.5.1(b). Height of local maxima; estimated spectral width parameter $\varepsilon = 0.617$.

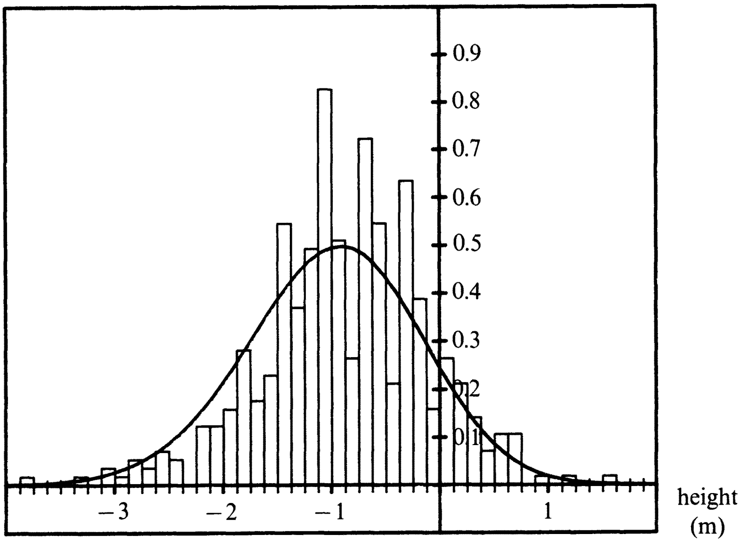


Figure 15.5.1(c). Height of local minima; estimated spectral width parameter $\varepsilon = 0.617$.

Let a_T and b_T be the standardized normalizing constants given by (15.4.4), again with general variance $\lambda_0 = \mathrm{Var}(\xi(t))$,

$$a_T = \lambda_0^{-1/2}(2 \log \nu T)^{1/2}, \qquad b_T = \lambda_0^{1/2}(2 \log \nu T)^{1/2}. \tag{15.5.2}$$

Then, by Theorem 8.2.7,

$$P\{a_T(M(T) - b_T) \le x\} \to \exp(-e^{-x})$$

so that

$$P\{M_{N'} \le u\} \approx \exp(-e^{-a_T(u-b_T)}),$$

where $N' = N'(T)$ is a random variable. All that is needed is therefore estimates of $\sqrt{\lambda_0}$ and ν, or equivalently of the *significant mean period* $T_s = 1/\nu$. (For small values of ε one can estimate $\sqrt{\lambda_0}$ by $A_s/1.66$, according to (15.5.1.)

□

Example 15.5.2. The highest encountered amplitude, together with the estimated variance λ_0 and mean period $T = 1/\nu$, were recorded during 12-minute periods eight times per day in the winter months January–March, 1973 at the Seven Stones Light Vessel between Cornwall and the Isles of Scilly.

This results in a total of $8 \times 90 = 720$ triples of observations,

$$M_{N'}^{(j)}, \lambda_0^{(j)}, T_s^{(j)}, \qquad j = 1, \ldots, 720,$$

where the $M_{N'}^{(j)}$ are the maximal wave amplitudes over separate intervals of length $T = 12$ min. Taken over the entire 3-month period the $\lambda_0^{(j)}$ (and $M_{N'}^{(j)}$) vary considerably as is seen in Figure 15.5.2 (a) and (b).

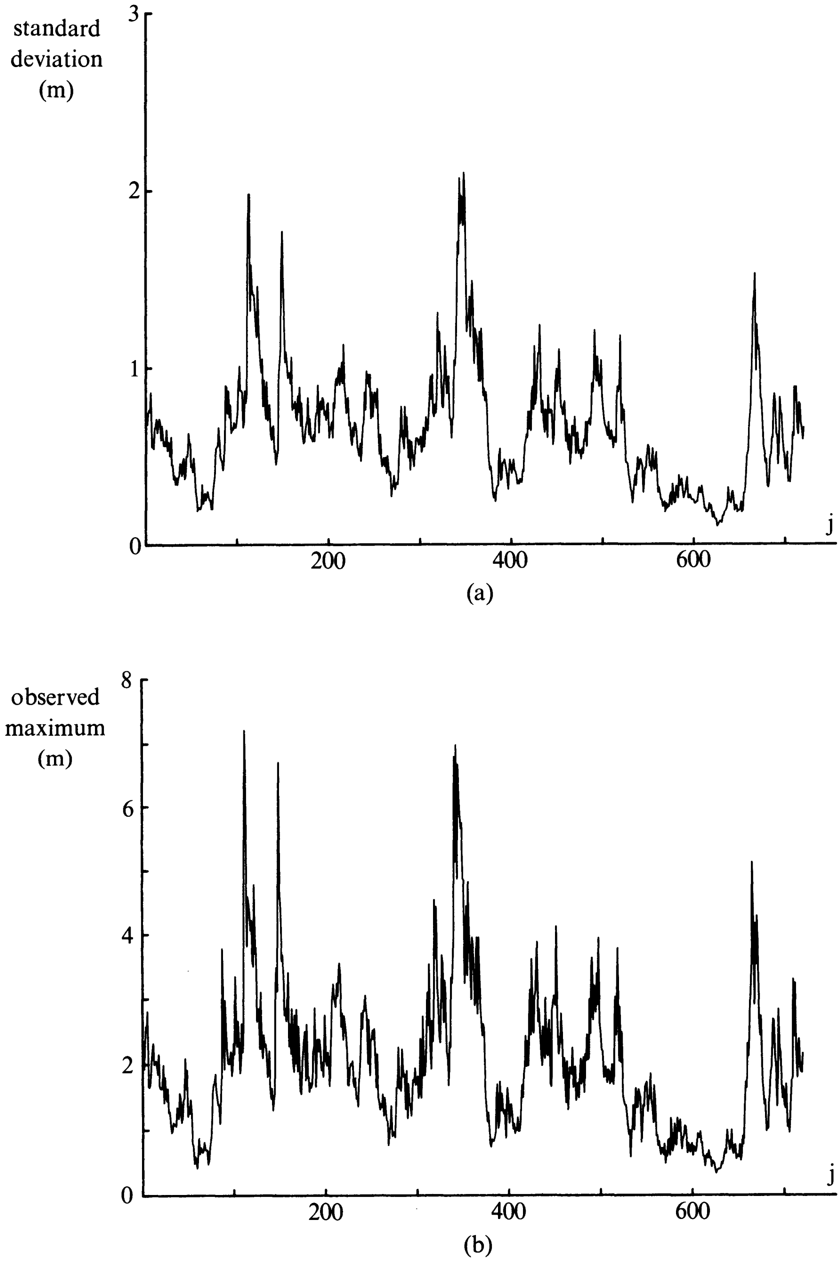


Figure 15.5.2. Plot of temporal variation in standard deviation (a) of observed water level over 12-minute periods, and (b) of observed maximum water level over the same periods, January–March.

In Figure 15.5.3 the normalized amplitudes $a_T^{(j)}(M_{N'}^{(j)} - b_T^{(j)})$ are plotted on double exponential probability paper with

$$a_T^{(j)} = (\lambda_0^{(j)})^{-1/2}\left(2\log\frac{T}{T_s^{(j)}}\right)^{1/2},$$

$$b_T^{(j)} = (\lambda_0^{(j)})^{1/2}\left(2\log\frac{T}{T_s^{(j)}}\right)^{1/2}.$$

This choice of normalizing constants is appropriate if the $M_{N'}^{(j)}$ are observations of the maximal wave amplitudes of a (nonstationary) normal process whose correlation structure is so slowly varying that it can be regarded as stationary over each individual interval.

The theoretical (asymptotic, as $T \to \infty$) d.f. of these standardized variables is $\exp(-e^{-x})$, and in Figure 15.5.3 this d.f. is illustrated by the straight dashed line. As seen in Figure 15.5.3 there is a considerable difference between the observed and the theoretical d.f. of standardized wave maxima. There may be physical reasons for this (such as breaking waves) but there are also two plausible statistical explanations. The observation interval $T = 12$ min may be too short for the asymptotic theory to be effective, and there may be long-ranging effects that shift the values of $a_T^{(j)}(M_{N'}^{(j)} - b_T^{(j)})$ up and down at a slow rate. To see this phenomenon we have plotted the entire

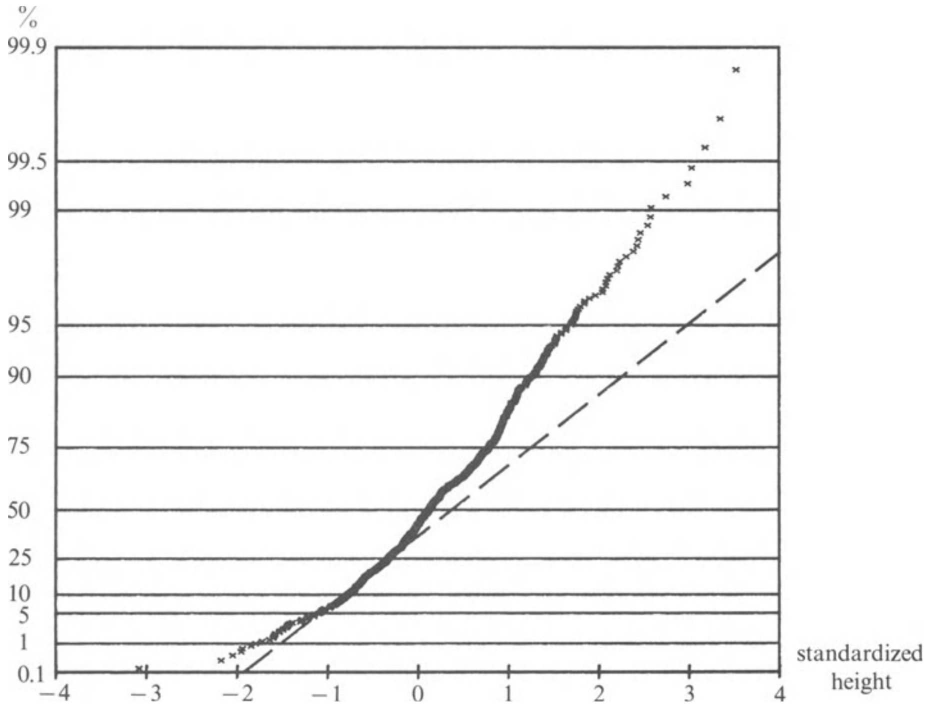


Figure 15.5.3. Observed d.f. of standardized observed maximum over 12 minutes, January–March plotted on double exponential probability paper.

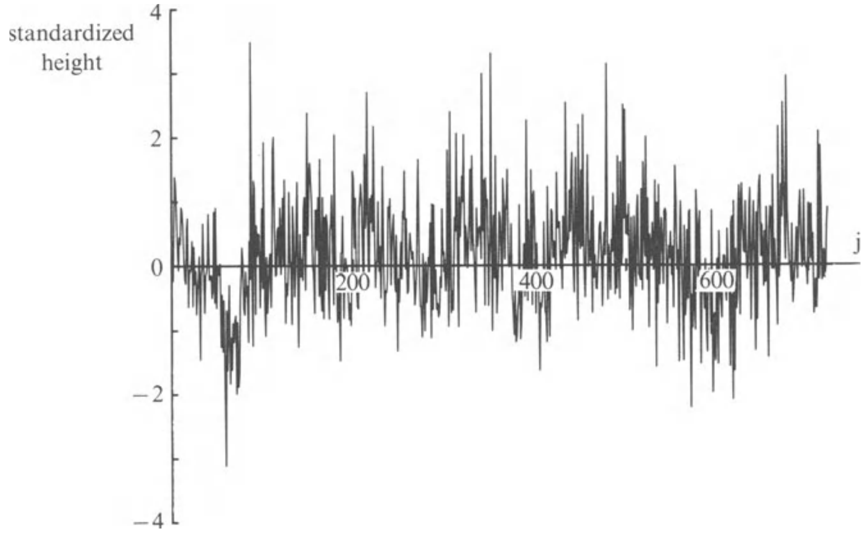


Figure 15.5.4. Plot of temporal variation in standardized maxima over 12-minute periods.

series of normalized values in Figure 15.5.4. As can be seen there, such a slow variation seems to be present, and this might be the main reason for the deviation shown in Figure 15.5.3.

The data behind Examples 15.5.1 and 15.5.2 were made available by the courtesy of P. Challenor at the Institute of Oceanographic Sciences, Wormley, England. □

APPENDIX

Some Basic Concepts of Point Process Theory

Intuitively, by a *point process*, we generally mean a series of events occurring in time (or space, or both) according to some statistical law. For example, the events may be radioactive disintegrations or telephone calls, occurring in time, or the positions of a certain variety of plant in a field (two-dimensional space). The cases of particular interest to us are when the events are the instants of occurrence of exceedances of a level u by a stochastic sequence $\{\xi_n\}$ (Chapter 5) or of the upcrossings of a level u by a continuous parameter process $\{\xi(t)\}$ (Chapter 9).

These point processes occur in one dimension (which we may regard as "time" if we wish). We may simultaneously consider exceedances or upcrossings of more than one level, and obtain a point process in the plane (cf. Chapters 5 and 9).

Point process theory may be discussed in a quite abstract setting, leading to a very satisfying general theory, and we refer the interested reader to the books by Kallenberg (1976) and Matthes *et al.* (1978) for this. Here we shall just indicate some of the main concepts regarding point processes on the real line, and in the plane.

If I is any finite interval on the real line, the number of events, $N(I)$ say, of a point process occurring in I must be a random variable. More generally, for any bounded Borel set B, $N(B)$ should be an r.v. Further, the number of events in the union of finitely or countably many disjoint sets, is the sum of the numbers in each set, i.e. $N(B) = \sum_1^\infty N(B_i)$ if B_i are disjoint (Borel) sets whose union is B. That is, $N(\cdot)$ is a *measure* on the Borel sets. Again the value of $N(B)$ must be an integer as long as it is finite. Hence the following formal definition naturally suggests itself.

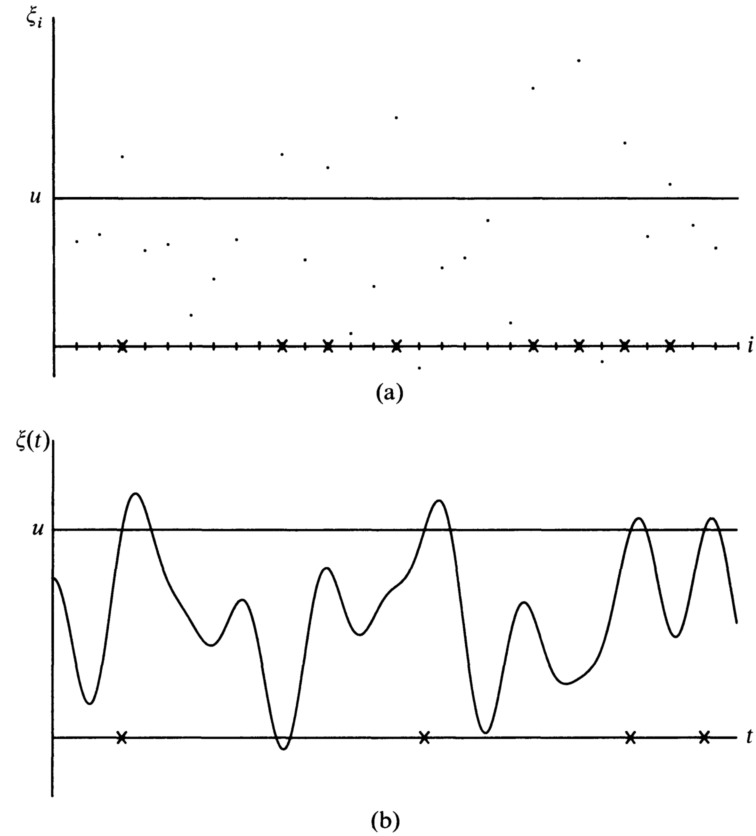


Figure A.1. (a) Point process of exceedances. (b) Point process of upcrossings.

A point process in a rectangle $S \subset R^n$ is a family of non-negative integer (or $+\infty$)-valued r.v.'s $N_\omega(B)$ defined for each Borel set $B \subset S$ and such that for each ω, $N_\omega(\cdot)$ is a measure on the Borel sets, being finite valued on bounded sets.

In the definition some useful extra generality has been gained by considering point processes defined on rectangles $S \subset R^n$ (where S may be finite or infinite, and open or closed at each of its boundaries) in addition to point processes on all of R^n. The requirement that a Borel set B is bounded has, in this context, the precise technical meaning that the diameter of B is finite and that the closure of B in R^n is contained in S. Thus, for instance, if $S = R^n$, then any set with finite diameter is bounded, while if $S = (0, 1] \subset R^1$, say, then, e.g. the set $(0, a]$, $a > 0$, has finite diameter but is not bounded.

Note that the same definition may be used for more abstract topological spaces than $S \subset R^n$ by using the Borel sets of that space in lieu of those in R^n. Here, as noted, we shall mainly consider just the line and the plane.

If τ is a random variable we may consider the trivial point process consisting of just one event occurring at the value which τ takes, i.e. the point process δ_τ where $\delta_\tau(B) = 1$ if $\tau \in B$ and $\delta_\tau(B) = 0$ otherwise. Thus δ_τ represents unit mass at the randomly chosen point τ. More generally, if τ_j are random variables (for $j = 1, 2, 3, \ldots$ or $j = 0, \pm 1 \pm 2, \ldots$) we may define a point process $N = \sum_j \delta_{\tau_j}$ provided that $N(B) = \sum_j \delta_{\tau_j}(B)$ is finite a.s. for bounded sets B. For this point process, the events occur at the random points $\{\tau_j\}$. Going a little further still, we may conveniently include possible *multiple events* by writing $N = \sum_j \beta_j \delta_{\tau_j}$ where β_j are non-negative integer-valued r.v.'s and the τ_j taken distinct.

In fact for a space with sufficient structure (such as the real line or plane) it may be shown that any point process N may be represented in terms of its *atoms* in this way, i.e. $N = \sum_j \beta_j \delta_{\tau_j}$, where the τ_j are a.s. distinct random elements of the space, and β_j are non-negative, integer-valued random variables. In the case where the β_j are each unity a.s., we say that the point process *has no multiple events* or is *simple*.

If $B_1, \ldots, B_k$ are bounded Borel sets, $N(B_1), \ldots, N(B_k)$ are random variables and have a joint distribution—termed a *finite-dimensional distribution* of the point process. In fact, the probabilistic properties of interest concerning the point process are specified uniquely by the collection of all such finite-dimensional distributions, i.e. for all choices of k and the sets $B_1, \ldots, B_k$. Of course, to define a point process starting from finite-dimensional distributions, we must choose these distributions in an appropriately consistent manner, so that the r.v.'s $N(B)$ will not only be well defined but will be non-negative, countably additive in B, etc. We refer the interested reader to Kallenberg (1976) for details.

If N is a point process the measure λ defined on the Borel sets (of the space involved) by

$$\lambda(B) = E(N(B))$$

is termed the *intensity measure* of the point process. Note that, unlike $N(B)$ itself, $\lambda(B)$ may be infinite even when B is bounded (since while a r.v. is finite valued, its mean need not be).

Although we shall not need them here, it is of interest to note that the probabilistic properties of a point process N may also be summarized by various generating functionals. In our judgement the most natural and useful of these is the Laplace transform $L_N(f)$ defined for non-negative measurable functions f by

$$L_N(f) = E\left(\exp\left(-\int f\,dN\right)\right) = E(\exp(-\textstyle\sum \beta_j f(\tau_j)))$$

when N is represented as $\sum \beta_j \delta_{\tau_j}$. Such generating functionals have properties and uses analogous to those of characteristic functions, moment generating functions, and Laplace transforms of random variables. In particular, if $f(x) = t\chi_B(x)$ (where $\chi_B(x) = 1$ or 0 according as $x \in B$ or $x \notin B$) we have $L_N(f) = E(e^{-tN(B)})$. This is simply the Laplace transform (or moment generating function evaluated at $-t$) of the r.v. $N(B)$, and it uniquely specifies the distribution of $N(B)$. Similarly joint Laplace transforms for r.v.'s $N(B_1), \ldots, N(B_k)$ may be specified by taking $f = \sum_{i=1}^{k} t_i \chi_{B_i}$.

Probably the most useful point process—both in its own right, and also as a "building block" for other types—is the *Poisson process*. This may be specified by its intensity measure $\lambda(B)$ which may be taken to be any measure which is finite on bounded sets. The point process N is said to be Poisson with this intensity if for each (bounded) B, $N(B)$ is a Poisson r.v. with mean $\lambda(B)$, and $N(B_1), \ldots, N(B_k)$ are independent for any choice of k, and disjoint $B_1, \ldots, B_k$. The existence of such a process N is easily shown under very general circumstances though we do not do so here. Further, N has the Laplace transform $L_N(f) = \exp\{-\int (1 - e^{-f(u)})\, d\lambda(u)\}$, and N is simple if the intensity λ is absolutely continuous.

It is readily checked that the choice $f(x) = t\chi_B(x)$ yields the Laplace transform of a Poisson r.v. with mean $\lambda(B)$. Incidentally this process may be called the "general Poisson process". The usual (stationary) Poisson process on the real line arises when $\lambda(B)$ is a constant multiple of Lebesgue measure $m(B)$, i.e. $\lambda(B) = \tau m(B)$, in which case we say that the Poisson process has intensity τ.

As noted above, the Poisson process may be used as a building block for the construction of other point processes. In particular, a most useful case arises if the intensity measure λ is itself allowed to be stochastic. Such a point process is no longer Poisson, but may be profitably thought of as "Poisson with a (stochastically) varying mean rate". We refer to such a process as a *doubly stochastic Poisson* or, more commonly, a *Cox process*. Specific use is made of such processes in Chapter 6, where the distribution is explicity given.

A notion which will be useful to us is that of *thinning* of a point process—and, in particular, of a Poisson process. Thinning refers to the removal of some of the events of the point process by a (usually) probabilistic mechanism which can be quite complicated. In its simplest form—with which we shall be concerned here—each event is removed or retained *independently*, with probabilities $1 - p$, p, say. For example, if N is a Poisson process with intensity measure λ, and N^* a point process obtained from N by such independent thinning, we have, for a Borel set B,

$$\begin{aligned} P\{N^*(B) = r\} &= \sum_{s=r}^{\infty} P\{N(B) = s\}\, P\{N^*(B) = r | N(B) = s\} \\ &= \sum_{s=r}^{\infty} \frac{e^{-\lambda(B)}(\lambda(B))^s}{s!} \binom{s}{r} p^r (1 - p)^{s-r}, \end{aligned}$$

since given that $N(B) = s$, $N^*(B)$ is binomial with parameters (s, p). This expression reduces simply to yield

$$P\{N^*(B) = r\} = \frac{e^{-p\lambda(B)}(p\lambda(B))^r}{r!}$$

so that $N^*(B)$ is a Poisson r.v. with mean $p\lambda(B)$. Similarly, it may be seen that $N^*(B_1), \ldots, N^*(B_k)$ are independent whenever $B_1, \ldots, B_k$ are disjoint, so that N^* is clearly a Poisson process with intensity measure $p\lambda$—an intuitively appealing result of which use is made, e.g. in Chapter 5.

There are numerous structural properties of point process theory—such as existence, uniqueness, simplicity, infinite divisibility and so on—which we do not go into here. It will, however, be of interest to mention *convergence* of a sequence of point processes, and to state a useful theorem in this connection.

Suppose that $\{N_n\}$ is a sequence of point processes on a rectangle $S \subset R^n$ and that N is a point process. Then we may say that N_n *converges in distribution* to N (written $N_n \xrightarrow{d} N$) if the sequence of vector r.v.'s $(N_n(B_1), \ldots, N_n(B_k))$ converges in distribution to $(N(B_1), \ldots, N(B_k))$ for each choice of k, and all bounded Borel sets $B_i \subset S$ such that $N(\partial B_i) = 0$ a.s., $i = 1, \ldots, k$, (writing ∂B for the boundary of the set B).

A point process may be viewed as a random element of a certain metric space (whose points are measures) and convergence in distribution of N_n to N becomes weak convergence of the distributions of N_n to that of N. Here we do not need this general viewpoint since the above definition is equivalent to it. However, at the end of this appendix we shall make a few more comments on the general approach, and exemplify how proofs can be simplified once results from the general theory are available.

The main result which we shall need is the following simple sufficient condition for convergence in distribution. This is a special case of a theorem of Kallenberg (1976), stated here for semiclosed (finite or infinite) intervals and rectangles—obvious modifications apply to other types of set S.

Theorem A.1. (i) *Let N_n, $n = 1, 2, \ldots$, and N be point processes on the semi-closed interval S in the real line, N being simple. Suppose that*

(a) $E(N_n((c, d])) \to E(N((c, d]))$ *for all* $-\infty < c < d < \infty$ *such that* $[c, d] \subset S$, *and*

(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ *for all B of the form* $\bigcup_1^k (c_i, d_i]$, *with* $[c_i, d_i] \subset S$, *for* $i = 1, \ldots, k$; $k = 1, 2, \ldots$.

Then $N_n \xrightarrow{d} N$.

(ii) *The same is true for point processes on a semi-closed rectangle S in the plane, if the semi-closed intervals $(c, d]$, $(c_i, d_i]$ are replaced by semi-closed rectangles $(c, d] \times (\gamma, \delta]$, $(c_i, d_i] \times (\gamma_i, \delta_i]$.*

The remarkable feature of this result is that convergence of the probability of occurrence of no events in certain given sets is essentially sufficient to guarantee convergences of quantities like $P\{N_n(B) = r\}$ and corresponding joint probabilities. The simple conditions (a) and (b) are often readily verified.

The following useful proposition is a main step in Kallenberg's proof of Theorem A.1. (The idea of the proof is that (a) ensures that to each sequence of integers there is a subsequence such that N_n converges to some simple point process along the subsequence. Proposition A.2 and (b) of Theorem A.1 then shows that the limit has the same distribution as N, which completes the proof.)

Proposition A.2. (i) *Suppose that N and N' are simple point processes on a semiclosed interval S in the real line and that $P\{N(B) = 0\} = P\{N'(B) = 0\}$ for all B of the form $\bigcup_1^k (c_i, d_i]$, with $[c_i, d_i] \subset S$ for $i = 1, \ldots, k$; $k = 1, 2, \ldots$. Then N and N' have the same distribution.*

(ii) *The same is true for point processes on a semi-closed rectangle in the plane, if the semi-closed intervals $(c_i, d_i]$ are replaced by semi-closed rectangles $(c_i, d_i] \times (\gamma_i, \delta_i]$.*

Our next result concerns convergence of a sequence of point processes in the plane, to a Poisson process in the plane, and shows how this property is preserved under suitable transformations of the points of each member of the sequence, and of the limit. Obviously this result could be stated in much greater generality but the form given here is sufficient for our applications.

Theorem A.3. *Let N_n, $n = 1, 2, \ldots$, and N be point processes on a semi-closed rectangle S in the plane, N being simple, and $\tau(x)$ a strictly decreasing continuous real function. Define new point processes $\{N_n'\}$, N' such that if N_n (N) has an atom at (s, t) then N_n' (N') has an atom at $(s, \tau^{-1}(t))$, where τ^{-1} is the inverse function of τ.*

(i) *If $N_n \xrightarrow{d} N$ then $N_n' \xrightarrow{d} N'$.*

(ii) *If N is Poisson, with intensity measure λ, then N' is Poisson with intensity λT^{-1}, where T denotes the transformation of the plane given by $T(s, t) = (s, \tau^{-1}(t))$. If λ is Lebesgue measure on the plane, the intensity λT^{-1} is the product of linear Lebesgue measure and the measure defined by the monotone function τ.*

PROOF. (i) It is readily checked that for any rectangle $B = (c, d] \times (\gamma, \delta]$

$$N'(B) = N((c, d] \times [\tau(\delta), \tau(\gamma))) = N(T^{-1}(B))$$

and hence (by uniqueness of extensions of measures) $N'(B) = N(T^{-1}(B))$ for all Borel sets B. This holds also with N_n', N_n replacing N', N.

Suppose now that B is a Borel set such that $N'(\partial B) = 0$ a.s. (again ∂B denotes the boundary of B). Now it may be seen (using the continuity of τ) that $\partial T^{-1}(B) \subset T^{-1}(\partial B)$ so that $N(\partial T^{-1}(B)) \leq N(T^{-1}(\partial B)) = N'(\partial B) = 0$ a.s. Since $N_n \xrightarrow{d} N$ we thus have $N_n(T^{-1}(B)) \xrightarrow{d} N(T^{-1}(B))$ or $N_n'(B) \xrightarrow{d} N'(B)$. This result extends simply to show that $(N_n'(B_1), \ldots, N_n'(B_k)) \xrightarrow{d} (N'(B_1), \ldots, N'(B_k))$ whenever $N'(\partial B_i) = 0$ a.s. for each $i = 1, \ldots, k$ and hence $N_n' \xrightarrow{d} N'$ as required.

(ii) If N is Poisson with intensity λ, and B is any Borel set in S

$$P\{N'(B) = r\} = P\{N(T^{-1}(B)) = r\} = \frac{e^{-\lambda T^{-1}(B)}(\lambda T^{-1}(B))^r}{r!}$$

for each $r = 0, 1, 2, \ldots$, so that $N'(B)$ is Poisson with mean $\lambda T^{-1}(B)$. Independence of $N'(B_1), \ldots, N'(B_k)$ for disjoint $B_1, \ldots, B_k$ follows from the fact that $T^{-1}(B_1), \ldots, T^{-1}(B_k)$ are also disjoint, and hence $N(T^{-1}(B_1)), \ldots, N(T^{-1}(B_k))$ are independent.

The last statement of (ii) follows simply if λ is Lebesgue measure,

$$\begin{aligned}\lambda T^{-1}((c, d] \times (\gamma, \delta]) &= \lambda\{(c, d] \times [\tau(\delta), \tau(\gamma))\} \\ &= (d - c)(\tau(\delta) - \tau(\gamma)),\end{aligned}$$

noting also that τ is continuous. □

As promised, we shall make some comments (without proofs) on the general approach to point processes. The reader should perhaps be warned that this presupposes some knowledge of weak convergence of probability measures on metric spaces, which is not needed elsewhere in the book.

Let, as before, S be a rectangle in R^n and let M denote the set of positive integer-valued measures of S which are finite on bounded sets. A sequence $\{v_n\} \subset M$ is said to converge vaguely to a measure $v \in M$ (notation: $v_n \xrightarrow{v} v$) if $\int f dv_n \to \int f dv$ for all functions f which are continuous and vanish outside some bounded set. The notion of vague convergence is particularly straightforward if v is simple, as can be seen from the following easily proven proposition.

Proposition A.4. *Suppose v is simple with atoms at the distinct points $t_1, t_2, \ldots \in S$, i.e. $v = \sum_k \delta_{t_k}$. Then $v_n \xrightarrow{v} v$ if and only if there are bounded rectangles $S_j \uparrow S$ with $v(\partial S_j \cap S) = 0$ such that if $t_{k_1}, \ldots, t_{k_l}$ are the atoms of v which are contained in S_j then for n large, v_n has precisely l atoms $t_{n,1}, \ldots, t_{n,l}$ in S_j and they can be ordered so that $t_{n,i} \to t_{k_i}$, as $n \to \infty$, $i = 1, \ldots, l$.*

Vague convergence induces a topology on M, and we let $\mathscr{M}$ be the σ-algebra generated by the open sets of this topology. The space $(M, \mathscr{M})$ is Polish, i.e. there exists some metric on M which generates the topology of M and which makes M complete. A point process N can be defined as a random element in $(M, \mathscr{M})$ and convergence in distribution of point processes is just ordinary convergence of random elements in a metric space, as set forth, e.g. in Billingsley's (1968) monograph. Thus by definition $N_n \xrightarrow{d} N$ if $E(h(N_n)) \to E(h(N))$ for all continuous bounded functions $h: M \to R^1$. As noted above, this definition of convergence can be shown to be equivalent to the more elementary definition given on p. 309.

An important result then is that $N_n \xrightarrow{d} N$ implies convergence in distribution of a wide class of functions of N_n. Let h be a function from $(M, \mathscr{M})$

to some metric space R (e.g. R may be M itself) and let D_h be the set of discontinuity points of h, i.e. $v \in D_h$ if there is a sequence $\{v_n\}$ such that $v_n \xrightarrow{v} v$ but $h(v_n) \not\to h(v)$. Then, if $P\{N \in D_h\} = 0$ (h is then said to be a.s. N-continuous) and $N_n \xrightarrow{d} N$ it follows that $h(N_n) \xrightarrow{d} h(N)$.

This can be used to give an alternative easy proof of Theorem A.3. In fact, if $v \in M$ is simple and $v' = h(v)$ is such that if v has an atom at (s, t) then v' has an atom at $(s, \tau^{-1}(t))$, then it is immediate from Proposition A.4 that h is continuous, and hence Theorem A.3(i) follows. A further example of the usefulness of the general theory is given in connection with record times in Section 5.8.

As a final note, it is apparent that the concept of a point process may be generalized to include measures N for which $N(B)$ is not necessarily integer valued. This generalization leads to a natural setting for point processes within the framework of the theory of random measures—a viewpoint developed in detail by Kallenberg (1976).

Bibliography

(For journal abbreviations see Mathematical Reviews, Index, 1981)

Adler, R. J. (1978). Weak convergence results for extremal processes generated by dependent random variables. *Ann. Probab.* **6**, 660–667.

Adler, R. J. (1981). *The Geometry of Random Fields.* New York: Wiley.

Afonja, B. (1972). The moments of the maximum of correlated normal and t-variates. *J. Roy. Statist. Soc., Ser. B*, **34**, 251–262.

Anderson, C. W. (1970). Extreme value theory for a class of discrete distributions with applications to some stochastic processes. *J. Appl. Probab.* **7**, 99–113.

Anderson, C. W. (1980). Local limit theorems for the maxima of discrete random variables. *Math. Proc. Cambridge Philos. Soc.* **88**, 161–165.

Antle, C. E. and Rademaker, F. (1972). An upper confidence limit on the maximum of m future observations from a type I extreme value distribution. *Biometrika* **59**, 475–477.

Arov, D. Z. and Bobrov, A. A. (1960). The extreme terms of a sample and their role in the sum of independent variables. *Theory Probab. Appl.* **5**, 377–396.

Balkema, A. A. (1973). *Monotone Transformations and Limit Laws.* Amsterdam: Mathematical Centre Tracts.

Balkema, A. A. and Haan, L. de (1972). On R. von Mises' condition for the domain of attraction of $\exp(-\exp(-x))$. *Ann. Math. Statist.* **43**, 1352–1354.

Balkema, A. A. and Haan, L. de (1978a). Limit distributions for order statistics, I. *Theory Probab. Appl.* **23**, 77–92.

Balkema, A. A. and Haan, L. de (1978b). Limit distributions for order statistics, II. *Theory Probab. Appl.* **23**, 341–358.

Balkema, A. A. and Resnick, S. I. (1977). Max-infinite divisibility. *J. Appl. Probab.* **14**, 309–319.

Bar-David, I. (1975a). Radon–Nikodym derivatives, passages and maxima for a Gaussian process with particular covariance and mean. *J. Appl. Probab.* **12**, 724–733.

Bar-David, I. (1975b). Sample functions of a Gaussian process cannot be recovered from their zero crossings. *IEEE Trans. Inform. Theory* **21**, 86–90.

Bar-David, I. and Ein-Gal, M. (1975). Passages and maxima for a particular Gaussian process. *Ann. Probab.* **3**, 549–556.

Bar-David, I. and Nemirovsky, A. (1972). Level crossings of nondifferentiable shot processes. *IEEE Trans. Inform. Theory* **18**, 27–34.

Barbé, A. (1976). A measure of the mean level-crossing activity of stationary normal processes. *IEEE Trans. Inform. Theory* **22**, 96–102.

Barlow, R. E. (1971). Averaging time and maxima for air pollution concentrations. Proc. 38th Session ISI, Washington D.C. pp. 663–676.

Barlow, R. E., Gupta, S. S., and Panchapakesan, S. (1969). On the distribution of the maximum and minimum of ratios of order statistics. *Ann. Math. Statist.* **40**, 918–934.

Barndorff-Nielsen, O. (1961). On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. *Math. Scand.* **9**, 383–394.

Barndorff-Nielsen, O. (1963). On the limit behaviour of extreme order statistics. *Ann. Math. Statist.* **34**, 992–1002.

Barndorff-Nielsen, O. (1964). On the limit distribution of the maximum of a random number of random variables. *Acta Math. Acad. Sci. Hungar.* **15**, 399–403.

Belayev, Yu. K. (1966). On the number of intersections of a level by a Gaussian stochastic process, I. *Theory Probab. Appl.* **11**, 106–113.

Belayev, Yu. K. (1967a). On the number of level crossings by a Gaussian stochastic process, II. *Theory Probab. Appl.* **12**, 392–404.

Belayev, Yu. K. (1967b). Bursts and shines of random fields. *Soviet Math. Dokl.* **8**, 1107–1109.

Belayev, Yu. K. (1967c). Limit theorem for the number of high level crossings by a stationary Gaussian process. *Soviet Math. Dokl.* **8**, 449–450.

Belayev, Yu. K. (1968). On the number of exits across the boundary of a region by a vector stochastic process. *Theory Probab. Appl.* **13**, 320–324.

Belayev, Yu. K. (1970). Random flows generated by random processes. *Math. Notes* **8**, 691–698.

Belayev, Yu. K. (1972). Point processes and first passage problems. *Proc. 6th Berkeley Symposium on Math. Statist. and Probability*, Ed. L. M. LeCam, Vol. 3, pp. 1–17.

Belayev, Yu. K. and Nosko, V. P. (1969). Characteristics of excursions above a high level for a Gaussian process and its envelope. *Theory Probab. Appl.* **14**, 296–309.

Belayev, Yu. K., Nosko, V. P., and Svesnikov, A. A. (1975). Probabilistic characteristics of excursions of a two-dimensional random field. *Theory Probab. Appl.* **6**, 25–30.

Bergström, H. (1953). Eine Theorie der stabilen Verteilungsfunktionen. *Arch. Math. (Basel)* **4**, 380–391.

Berman, S. M. (1961). Convergence to bivariate limiting extreme value distributions. *Ann. Inst. Statist. Math.* **13**, 217–223.

Berman, S. M. (1962a). A law of large numbers for the maximum in a stationary Gaussian sequence. *Ann. Math. Statist.* **33**, 93–97.

Berman, S. M. (1962b). Limiting distribution of the maximum term in sequences of dependent random variables. *Ann. Math. Statist.* **33**, 894–908.

Berman, S. M. (1962c). Equally correlated random variables. *Sankhyā*, *Ser. A*, **24**, 155–156.

Berman, S. M. (1964a). Limiting distribution of the maximum of a diffusion process. *Ann. Math. Statist.* **35**, 319–329.

Berman, S. M. (1964b). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35**, 502–516.

Berman, S. M. (1971a). Excursions above high levels for stationary Gaussian processes. *Pacific J. Math.* **36**, 63–79.

Berman, S. M. (1971b). Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes. *Ann. Math. Statist.* **42**, 927–945.

Berman, S. M. (1971c). Maxima and high level excursions of stationary Gaussian processes. *Trans. Amer. Math. Soc.* **160**, 65–85.

Berman, S. M. (1972a). Maximum and high level excursion of a Gaussian process with stationary increments. *Ann. Math. Statist.* **43**, 1247–1266.

Berman, S. M. (1972b). A class of limiting distributions of high level excursions of Gaussian processes. *Z. Wahrsch. verw. Gebiete* **21**, 121–134.
Berman, S. M. (1973). Excursions of stationary Gaussian processes above high moving barriers. *Ann. Probab.* **1**, 365–387.
Berman, S. M. (1974). Sojourns and extremes of Gaussian processes. *Ann. Probab.* **2**, 999–1026 (Correction (1980), **8**, 999).
Berman, S. M. (1979). High level sojourns for strongly dependent Gaussian processes. *Z. Wahrsch. verw. Gebiete* **50**, 223–236.
Berman, S. M. (1980). A compound Poisson limit for stationary sums, and sojourns of Gaussian processes. *Ann. Probab.* **8**, 511–538.
Berman, S. M. (1982a). Sojourns and extremes of stationary processes. *Ann. Probab.* **10**, 1–46.
Berman, S. M. (1982b). Sojourns and extremes of a diffusion process on a fixed interval. *Techn. Report*, Courant Institute of Math. Sci., New York Univ.
Billingsley, P. (1968). *Convergence of Probability Measures.* New York: Wiley.
Blake, I. F. and Lindsey, W. C. (1973). Level-crossing problems for random processes. *IEEE Trans. Inform. Theory* **19**, 295–315.
Boas, R. P. (1967). Lipschitz behaviour and integrability of characteristic functions. *Ann. Math. Statist.* **38**, 32–36.
Boatwright, P. A. and Riter, S. (1973). Level selection for a level crossing receiver. *IEEE Trans. Comm.* **21**, 953–954.
Bobrov, A. A. (1954). The growth of the maximal summand in the sums of independent random variables. *Mat. Sb. Kievskogo Gosuniv.* **5**, 15–38.
Bofinger, E. and Bofinger, V. J. (1965). The correlation of maxima in samples drawn from a bivariate normal distribution. *Austral. J. Statist.* **7**, 57–61.
Bofinger, V. J. (1970). The correlation of maxima in several bivariate non-normal distributions. *Austral. J. Statist.* **12**, 1–7.
Bortkievicz, L. von (1922). Variationsbreite und mittlere Fehler. *Sitzungsber. Berliner Math. Ges.* **21**.
Breiman, L. (1968). *Probability.* Reading, Massachusetts: Addison-Wesley.
Brown, B. M. and Resnick, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14**, 732–739.
Buhl, S. (1969). The joint distribution of the number of crossings of several levels by a normal stochastic process. *Techn. Report*, Dept. of Statistics, Univ. of Aarhus.
Bulinskaya, E. V. (1961). On the mean number of crossings of a level by a stationary Gaussian process. *Theory Probab. Appl.* **6**, 435–438.
Bury, K. V. (1974). Distribution of smallest lognormal and gamma extremes. *Statist. Hefte* **15**, 105–114.
Cabaña, E. M. (1970). The vibrating string forced by white noise. *Z. Wahrsch. verw. Gebiete* **15**, 111–130.
Cabaña, E. M. (1972a). On barrier problems for the vibrating string. *Z. Wahrsch. verw. Gebiete* **22**, 13–24.
Cabaña, E. M. (1972b). On the barrier problem for sine series with independent Gaussian coefficients. *Z. Wahrsch. verw. Gebiete* **24**, 215–219.
Cabaña, E. M. and Wschebor, M. (1981). An estimate for the tails of the distribution of the supremum for a class of stationary multiparameter Gaussian processes. *J. Appl. Probab.* **18**, 536–541.
Cacoullos, T. and DeCicco, H. (1967). On the distribution of the bivariate range. *Technometrics* **9**, 476–480.
Campbell, J. W. and Tsokos, C. P. (1973). The asymptotic distribution of maxima in bivariate samples. *J. Amer. Statist. Assoc.* **68**, 734–739.
Cartwright, D. E. and Longuet-Higgins, M. S. (1956). The statistical distribution of the maxima of a random function. *Proc. Roy. Soc. London, Ser. A*, **237**, 212–232.

Cavanie, A., Arhan, M. and Ezraty, R. (1976). A statistical relationship between individual heights and periods of storm waves. *Proceedings BOSS'76*, pp. 354–360. Trondheim: Tapir.

Chakrabarti, S. K. and Cooley, R. P. (1977). Statistical distribution of periods and heights of ocean waves. *J. Geophys. Res.* **82**, 1363–1368.

Chan, L. K. and Jarvis, G. A. (1970). Convergence of moments of some extreme order statistics. *J. Statist. Res.* **4**, 37–42.

Chandler, K. N. (1952). The distribution and frequency of record values. *J. Roy. Statist. Soc., Ser. B*, **14**, 220–228.

Chaplin, W. S. (1880). The relation between the tensile strengths of long and short bars. *Van Nostrand's Engineering Magazine* **23**, 441–444.

Chaplin, W. S. (1882). On the relative tensile strengths of long and short bars. *Proc. Engineer's Club, Philadelphia* **3**, 15–28.

Chen, Pei-De (1966). The limiting distribution of the variation of a sequence of independent random variables. *Acta Math. Sinica* **16**, 87–103.

Chernick, M. R. (1980). A limit theorem for the maximum term in a particular EARMA (1, 1) sequence. *J. Appl. Probab.* **17**, 869–873. (Correction (1981) **18**, 569.)

Chernick, M. R. (1981a). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *Ann. Probab.* **9**, 145–149.

Chernick, M. R. (1981b). On strong mixing and Leadbetter's *D* condition. *J. Appl. Probab.* **18**, 764–769.

Chibisov, D. M. (1964). On limit distributions for order statistics. *Theory Probab. Appl.* **9**, 142–148.

Chuberkis, V. P. (1971). Crossings of fixed levels by two stationary random processes. *Automat. Remote Control* **32**, 1581–1587.

Cleveland, W. S. (1980). Discussion of Horowitz (1980). *Technometrics* **22**, 479–482.

Clough, D. J. (1969). An asymptotic extreme value sampling theory for estimation of a global maximum. *Canad. Op. Res. Soc. J.* **7**, 102–115.

Clough, D. J. and Kotz, S. (1965). Extreme-value distributions with a special queuing model application. *Canad. Op. Res. Soc. J.* **3**, 96–109.

Cohen, J. W. (1967). The distribution of the maximum number of customers present simultaneously during a busy period for the queueing systems M/G/1 and G/M/1. *J. Appl. Probab.* **4**, 162–179.

Cramér, H. (1962). On the maximum of a normal stationary stochastic process. *Bull. Amer. Math. Soc.* **68**, 512–516.

Cramér, H. (1965). A limit theorem for the maximum values of certain stochastic processes. *Theory Probab. Appl.* **10**, 126–128.

Cramér, H. (1966). On the intersections between the trajectories of a normal stationary process and a high level. *Ark. Mat.* **6**, 337–349.

Cramér, H. and Leadbetter, M. R. (1965). The moments of the number of crossings of a level by a stationary normal process. *Ann. Math. Statist.* **36**, 1656–1663.

Cramér, H. and Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*. New York: Wiley.

Cramér, H., Leadbetter, M. R., and Serfling, R. J. (1971). On distribution function–moment relationships in a stationary point process. *Z. Wahrsch. verw. Gebiete* **18**, 1–8.

Cuzick, J. M. (1975). Conditions for finite moments of the number of zero crossings for Gaussian processes. *Ann. Probab.* **3**, 849–858.

Cuzick, J. M. (1981). Boundary crossing probabilities for stationary Gaussian processes and Brownian motion. *Trans. Amer. Math. Soc.* **263**, 469–492.

Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. *Proc. Roy. Soc. London, Ser. A*, **183**, 405–435.

Darling, D. A. (1952). The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.* **73**, 95–107.

Davenport, A. G. (1978). Wind structure and wind climate. *Int. Res. Seminar on Safety of Structures Under Dynamic Loading*, Eds. I. Holand, D. Kavlie, G. Moe, R. Sigbjörnsson, pp. 238–256. Trondheim: Tapir.

David, H. A. (1981). *Order Statistics*, 2nd ed. New York: Wiley.

David, H. A. and Galambos, J. (1974). The asymptotic theory of concomitants of order statistics. *J. Appl. Probab.* **11**, 762–770.

Davis, R. A. (1979). Maxima and minima of stationary sequences. *Ann. Probab.* **7**, 453–460.

Davis, R. A. (1981). Limit laws for upper and lower extremes from stationary mixing sequences. *Techn. Report*, Massachusetts Institute of Technology.

Davis, R. A. (1982). Limit laws for the maximum and minimum of stationary sequences. *Z. Wahrsch. verw. Gebiete* **61**, 31–42.

Deheuvels, P. (1974). Valeur extremales d'échantillons croissants d'une variable aléatoire réelle. *Ann. Inst. H. Poincaré* **10**, 89–114.

Deheuvels, P. (1978). Caracterisation complete des lois extremes multivariées et de la convergence aux types extremes. *Publ. Inst. Statist. Univ. Paris* **23**.

Denzel, G. E. and O'Brien, G. L. (1975). Limit theorems for extreme values of chain-dependent processes. *Ann. Probab.* **3**, 773–779.

Deo, C. M. (1973a). A note on strong-mixing Gaussian sequences. *Ann. Probab.* **1**, 186–187.

Deo, C. M. (1973b). An iterated logarithm law for maxima of nonstationary Gaussian processes. *J. Appl. Probab.* **10**, 402–408.

Diebolt, J. (1979). Sur la loi de maximum de certains processus stochastiques. *Publ. Inst. Stat. Univ. Paris* **24**, 31–68.

Dodd, E. L. (1923). The greatest and the least variate under general laws of error. *Trans. Amer. Math. Soc.* **25**, 525–539.

Dudley, R. M. (1973). Sample functions of the Gaussian process. *Ann. Probab.* **1**, 66–103.

Durrett, R. and Resnick, S. I. (1978). Functional limit theorems for dependent variables. *Ann. Probab.* **6**, 829–846.

Dwass, M. (1964). Extremal processes. *Ann. Math. Statist.* **35**, 1718–1725.

Dwass, M. (1966). Extremal processes, II. *Illinois J. Math.* **10**, 381–391.

Dziubdziela, W. (1978). On convergence rates in the limit laws for extreme order statistics. *Trans. 7th Prague Conf. and 1974 European Meeting of Statisticians*, Vol. B, pp. 119–127.

Dziubdziela, W. and Kopocinski, B. (1976). Limiting properties of the kth record values. *Zastos. Mat.* **15**, 187–190.

Dziubdziela, W. and Szczotka, W. (1979). On weak convergence of order statistics. *Bull. Acad. Polon. Sci. Sér. Sci. Math.* **27**, 499–501.

Earle, M. D., Effermeyer, C. C., and Evans, D. J. (1974). Height-period joint probabilities in hurricane Camille. *J. Waterways, Harbors and Coastal Engineering Div. ASCE* **3**, 257–264.

Eddy, W. F. (1980). The distribution of the convex hull of a Gaussian sample. *J. Appl. Probab.* **17**, 686–695.

Eiffel, G. (1900). *Travaux Scientifiques exécutés a la Tour de Trois Cent Metres, de 1889 a 1900*. Paris: L. Maretheux.

Epstein, B. (1948). Applications of the theory of extreme values in fracture problems. *J. Amer. Statist. Assoc.* **43**, 403–412.

Epstein, B. (1960). Elements of the theory of extreme values. *Technometrics* **2**, 27–41.

Esary, J. D. and Marshall, A. W. (1974). Multivariate distributions with exponential minimums. *Ann. Statist.* **2**, 84–98.

Ewing, J. A. (1973). Mean length of runs of high waves. *J. Geophys. Res.* **12**, 1933–1936.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. New York: Wiley.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed. New York: Wiley.
Fernique, X. (1964). Continuité des processus Gaussiens. *Comptes Rendus* **258**, 6058–6060.
Fertig, K. W. and Mann, N. R. (1980). An accurate approximation to the sampling distribution of the studentized extreme value statistic. *Technometrics* **22**, 83–97.
Fieger, W. (1965). Eine für beliebige Call-Prozesse geltende Verallgemeinerung der Palmschen Formeln. *Math. Scand.* **16**, 121–147.
Fieger, W. (1967). Zero points of stochastic processes. *Techn. Report*, Univ. of Aarhus.
Fieger, W. (1969a). Die Anzahl der lokalen Maximalstellen eines stationären Gauszchen Prozesses. *Math. Scand.* **25**, 218–226.
Fieger, W. (1969b). Die Transformation von Burkill-Unterteilungsintegralen. *Math. Ann.* **183**, 115–129.
Fieger, W. (1971). Die Anzahl der Gamma-Niveau-Kreuzungspunkte von stochastishen Prozessen. *Z. Wahrsch. verw. Gebiete* **18**, 227–260.
Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24**, 180–190.
Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.* **6**, 93–116.
Fréchet, M. (1943). *Les Probabilités Associées a un Systeme d'Évenements Compatibles et Dépendents*, 2 ed. Paris: Hermann.
Freudenthal, A. M. (1968). Statistical approach to brittle fracture. In: *Fracture: An Advanced Treatise*, Vol. II, Ed. H. Liebowitz, pp. 592–619.
Galambos, J. (1972). On the distribution of the maximum of random variables. *Ann. Math. Statist.* **43**, 516–521.
Galambos, J. (1973a). A general Poisson limit theorem of probability theory. *Duke Math. J.* **40**, 581–586.
Galambos, J. (1973b). The distribution of the maximum of a random number of random variables with applications. *J. Appl. Probab.* **10**, 122–129.
Galambos, J. (1974). A limit theorem with applications in order statistics. *J. Appl. Probab.* **11**, 219–222.
Galambos, J. (1975a). Order statistics of samples from multivariate distributions. *J. Amer. Statist. Assoc.* **70**, 674–680.
Galambos, J. (1975b). Limit laws for mixtures with applications to asymptotic theory of extremes. *Z. Wahrsch. verw. Gebiete* **32**, 197–207.
Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. New York: Wiley.
Galambos, J. and Seneta, E. (1975). Record times. *Proc. Amer. Math. Soc.* **50**, 383–387.
Gallot, S. A. (1966). A bound for the maximum of a number of random variables. *J. Appl. Probab.* **3**, 556–558.
Gaver, D. P. (1976). Random record models. *J. Appl. Probab.* **13**, 538–547.
Geffroy, J. (1958/1959). Contributions a la théorie des valeurs extrémes. *Publ. Inst. Stat. Univ. Paris* **7/8**, 37–185.
Geman, D. (1970). Horizontal-window conditioning and the zeros of stationary Gaussian processes. Ph.D. dissertation, Northwestern Univ., Evanston, Illinois.
Geman, D. (1972). On the variance of the number of zeros of a stationary Gaussian process. *Ann. Math. Statist.* **43**, 977–982.
Geman, D. (1973). A note on the distribution of hitting times. *Ann. Probab.* **1**, 854–856.
Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* **8**, 1–67.
Glick, N. (1978). Breaking records and breaking boards. *Amer. Math. Monthly* **85**, 2–26.
Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423–453.

Green, R. F. (1976a). Outlier-prone and outlier-resistant distributions. *J. Amer. Statist. Assoc.* **71**, 502–505.
Green, R. F. (1976b). Partial attraction of maxima. *J. Appl. Probab.* **13**, 159–163.
Greig, M. (1967). Extremes in a random assembly. *Biometrika* **54**, 273–282.
Grenander, U. (1950). Stochastic processes and statistical inference. *Ark. Mat.* **1**, 195–277.
Grenander, U. (1965). A limit theorem for sums of minima of stochastic variables. *Ann. Math. Statist.* **36**, 1041–1042.
Gumbel, E. J. (1958). *Statistics of Extremes*. New York: Columbia Univ. Press.
Gumbel, E. J. and Goldstein, N. (1964). Analysis of empirical bivariate extremal distributions. *J. Amer. Statist. Assoc.* **59**, 794–816.
Gumbel, E. J. and Mustafi, C. K. (1967). Some analytical properties of bivariate extremal distributions. *J. Amer. Statist. Assoc.* **62**, 569–588.
Haan, L. de (1970). On regular variation and its application to the weak convergence of sample extremes. *Amsterdam Math. Centre Tracts* **32**, 1–124.
Haan, L. de (1971). A form of regular variation and its application to the domain of attraction of the double exponential distribution. *Z. Wahrsch. verw. Gebiete* **17**, 241–258.
Haan, L. de (1974). Weak limits of sample range. *J. Appl. Probab.* **11**, 836–841.
Haan, L. de (1976). Sample extremes: an elementary introduction. *Statist. Neerlandica* **30**, 161–172.
Haan, L. de and Hordijk, A. (1972). The rate of growth of sample maxima. *Ann. Math. Statist.* **43**, 1185–1196.
Haan, L. de and Resnick, S. I. (1973). Almost sure limit points of record values. *J. Appl. Probab.* **10**, 528–542.
Haan, L. de and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. verw. Gebiete* **40**, 317–337.
Haan, L. de and Resnick, S. I. (1982a). Local limit theorems for sample extremes. *Ann. Probab.* **10**, 396–413.
Haan, L. de and Resnick, S. I. (1982b). Asymptotically balanced functions and stochastic compactness of extremes. *Techn. Report 8206/S*, Econ. Inst., Erasmus Univ.
Haan, L. de and Ridder, G. (1979). Stochastic compactness of sample extremes. *Ann. Probab.* **7**, 290–303.
Haghighi-Talab, D. and Wright, C. (1973). On the distribution of records in a finite sequence of observations, with an application to a road traffic problem. *J. Appl. Probab.* **10**, 556–571.
Hall, P. (1978a). On the extreme terms of a sample from the domain of attraction of a stable law. *J. London Math. Soc.* **18**, 181–191.
Hall, P. (1978b). Representations and limit theorems for extreme value distributions. *J. Appl. Probab.* **15**, 639–644.
Hall, P. (1978c). Some asymptotic expansions of moments of order statistics. *Stochastic Process. Appl.* **7**, 265–275.
Hall, P. (1979). The rate of convergence of normal extremes. *J. Appl. Probab.* **16**, 433–439.
Hall, P. (1980). Estimating probabilities for normal extremes. *Adv. Appl. Probab.* **12**, 491–500.
Hall, W. J. and Wellner, J. A. (1979). The rate of convergence in law of the maximum of an exponential sample. *Statist. Neerlandica* **33**, 151–154.
Hållberg, B. and Maré, J. de (1976). Extreme value theory and break frequency of paper. *Bull Inst. Statist. Inst.* **46**, Book 3, 357–361.
Harris, R. (1970). An application of extreme value theory to reliability theory. *Ann. Math. Statist.* **41**, 1456–1465.

Harter, H. L. (1977). A survey of the literature on the size effect on material strength. *Techn. Report AFFDL-TR-77-11*, Wright-Patterson Air Force Base, pp. 1–450.

Harter, H. L. (1978). A bibliography of extreme-value theory. *Intern. Statist. Rev.* **46**, 279–306.

Hasofer, A. M. (1968). A statistical theory of the brittle fracture. *Int. J. Fracture Mechanics* **4**, 439–452.

Hasofer, A. M. (1970). On the derivative and the upcrossings of the Rayleigh process. *Austral. J. Statist.* **12**, 150–151.

Hasofer, A. M. (1972). Wind-load design based on statistics. *Proc. Inst. Civil Eng.* **51**, 69–82.

Hasofer, A. M. (1976). The upcrossing rate of a class of stochastic processes. In: *Studies in Probability and Statistics*, Ed. Williams, pp. 153–159. Amsterdam: North Holland.

Hasofer, A. M. and Sharpe, K. (1969). The analysis of wind gusts. *Austral. Meteorol. Mag.* **17**, 198–214.

Hebbar, H. V. (1980). Almost sure limit points of maxima of stationary Gaussian sequences. *Ann. Probab.* **8**, 393–399.

Helstrom, C. W. (1957). The distribution of the number of crossings of a Gaussian stochastic process. *IEEE Trans. Inform. Theory* **3**, 232–237.

Heyde, C. C. (1971). On the growth of the maximum queue length in a stable queue. *Oper. Res.* **19**, 447–452.

Höglund, T. (1972). Asymptotic normality of sums of minima of random variables. *Ann. Math. Statist.* **43**, 351–353.

Holmes, P. T. and Strawderman, W. E. (1969). A note on the waiting times between record observations. *J. Appl. Probab.* **6**, 711–714.

Holst, L. (1981). On convergence of the coverage by random arcs on a circle and the largest spacing. *Ann. Probab.* **9**, 648–655.

Horowitz, J. (1980). Extreme values from a nonstationary stochastic process: An application to air quality analysis. *Technometrics* **22**, 469–478.

Hüsler, J. (1977). On limit distributions of first crossing points of Gaussian sequences. *Stochastic Process. Appl.* **6**, 65–75.

Hüsler, J. (1979). The limiting behaviour of the last exit time for sequences of independent, identically distributed random variables. *Z. Wahrsch. verw. Gebiete* **50**, 159–164.

Hüsler, J. (1980). Limit distribution of the last exit time for stationary random sequences. *Z. Wahrsch. verw. Gebiete* **52**, 301–308.

Hüsler, J. (1981). A note on the functional law of iterated logarithm for maxima of Gaussian sequences. *J. Multivariate Anal.* **11**, 273–279.

Hüsler, J. (1982). Asymptotic approximation of crossing probabilities of stationary random sequences. To appear in *Z. Wahrsch. verw. Gebiete.*

Ibragimov, I. A. and Maslova, N. B. (1971). On the expected number of real zeros of random polynomials, I. Coefficients with zero means. *Theory Probab. Appl.* **16**, 228–248.

Iglehart, D. L. (1972). Extreme values in the GI/G/1 queue. *Ann. Math. Statist.* **43**, 627–635.

Ito, K. (1964). The expected number of zeros of continuous stationary Gaussian processes. *J. Math. Kyoto Univ.* **3**, 207–216.

Ivanov, V. A. (1960). On the average number of crossings of a level by sample functions of a stochastic process. *Theory Probab. Appl.* **5**, 319–323.

Jeruchim, M. C. (1976). On the estimation of error probability using generalized extreme-value theory. *IEEE Trans. Inform. Theory* **22**, 108–110.

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*, Vol. 2. New York: Wiley.

Judickaja, P. I. (1974a). On the maximum of a Gaussian sequence. *Theory Probab. Math. Statist.* **2**, 259–267.

Judickaja, P. I. (1974b). The asymptotic behaviour of the maximum of a Gaussian field. *Theory Probab. Math. Statist.* **3**, 243–250.

Judickaja, P. I. (1975). Asymptotic inequalities for maxima of nondifferentiable normal fields. *Theory Probab. Math. Statist.* **6**, 135–144.

Juncosa, M. L. (1949). The asymptotic behavior of the minimum in a sequence of random variables. *Duke Math. J.* **16**, 609–618.

Kac, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* **49**, 314–320.

Kac, M. and Slepian, D. (1959). Large excursions of Gaussian processes. *Ann. Math. Statist.* **30**, 1215–1228.

Kalinauskaite, N. (1973). The effect of the maximum modulus of a term on the sums of independent random vectors I. *Lietuv. Mat. Rink.* **13**, 117–123.

Kalinauskaite, N. (1974). The effect of the maximum modulus of a term on the sums of independent random vectors II. *Lietuv. Mat. Rink.* **14**, 41–48.

Kallenberg, O. (1976). *Random Measures.* Berlin: Akademie-Verlag, London–New York: Academic Press.

Karr, A. F. (1976). Two extreme value processes arising in hydrology. *J. Appl. Probab.* **13**, 190–194.

Kedem, B. (1978). Sufficiency and the number of level crossings by a stationary process. *Biometrika* **65**, 207–210.

Kolchin, V. F. (1969). On the limiting behaviour of extreme order statistics in a polynomial scheme. *Theory Probab. Appl.* **14**, 458–469.

Kostyukov, A. I. (1978). Curve-crossing problems for Gaussian stochastic processes and its application to neural modeling. *Biol. Cybernetics* **29**, 187–191.

Kudô, A. (1958). On the distribution of the maximum value of an equally correlated sample from a normal population. *Sankhyā, Ser. A*, **20**, 309–316.

Lai, T. L. and Robbins, H. (1976). Maximally dependent random variables. *Proc. Nat. Acad. Sci. U.S.A.* **73**, 286–288.

Lai, T. L. and Robbins, H. (1978). A class of dependent random variables and their maxima. *Z. Wahrsch. verw. Gebiete* **42**, 89–111.

Lamperti, J. (1964). On extreme order statistics. *Ann. Math. Statist.* **35**, 1726–1737.

Leadbetter, M. R. (1963). On the normal stationary process—areas outside given levels. *J. Roy. Statist. Soc., Ser. B*, **25**, 189–194.

Leadbetter, M. R. (1965). On crossings of arbitrary curves by certain Gaussian processes. *Proc. Amer. Math. Soc.* **16**, 60–68.

Leadbetter, M. R. (1966a). On streams of events and mixtures of streams. *J. Roy. Statist. Soc., Ser. B*, **28**, 218–227.

Leadbetter, M. R. (1966b). Local maxima of stationary processes. *Proc. Cambridge Phil. Soc.* **62**, 263–268.

Leadbetter, M. R. (1966c). On crossings of levels and curves by a wide class of stochastic processes. *Ann. Math. Statist.* **37**, 260–267.

Leadbetter, M. R. (1967). A note on the moments of the number of axis-crossings by a stochastic process. *Bull. Amer. Math. Soc.* **73**, 129–132.

Leadbetter, M. R. (1971). Extreme value theory for stochastic processes. *Proc. 5th Princeton Systems Symposium.*

Leadbetter, M. R. (1973). Point processes generated by level crossings. In: *Stochastic point processes*, Ed. P. A. W. Lewis. New York: Wiley.

Leadbetter, M. R. (1974). On extreme values in stationary sequences. *Z. Wahrsch. verw. Gebiete* **28**, 289–303.

Leadbetter, M. R. (1976). Weak convergence of high level exceedances by a stationary sequence. *Z. Wahrsch. verw. Gebiete* **34**, 11–15.

Leadbetter, M. R. (1982). Extremes and local dependence in stationary sequences. *Center for Stoch. Proc.* Report 12, Statistics Dept. Univ. of N.C.

Leadbetter, M. R. and Cryer, J. D. (1965a). On the mean number of curve crossings by non-stationary normal processes. *Ann. Math. Statist.* **36**, 509–516.

Leadbetter, M. R. and Cryer, J. D. (1965b). The variance of the number of zeros of a stationary normal process. *Bull. Amer. Math. Soc.* **71**, 561–563.

Leadbetter, M. R. and Rootzén, H. (1982). Extreme value theory for continuous parameter stationary processes. *Z. Wahrsch. verw. Gebiete* **60**, 1–20.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1978). Conditions for the convergence in distribution of maxima of stationary normal processes. *Stochastic Process. Appl.* **8**, 131–139.

Lévy, P. (1939). Sur la division d'un segment par des points choisis au hasard. *C.R. Acad. Sci. Paris* **208**, 147–149.

Lieblein, J. (1954). Two early papers on the relation between extreme values and tensile strength. *Biometrika* **41**, 559–560.

Lindgren, G. (1970). Some properties of a normal process near a local maximum. *Ann. Math. Statist.* **41**, 1870–1883.

Lindgren, G. (1971). Extreme values of stationary normal processes. *Z. Wahrsch. verw. Gebiete* **17**, 39–47.

Lindgren, G. (1972a). Wave-length and amplitude in Gaussian noise. *Adv. Appl. Probab.* **4**, 81–108.

Lindgren, G. (1972b). Wave-length and amplitude for a stationary Gaussian process after a high maximum. *Z. Wahrsch. verw. Gebiete* **23**, 293–326.

Lindgren, G. (1973). Discrete wave-analysis of continuous stochastic processes. *Stochastic Process. Appl.* **1**, 83–105.

Lindgren, G. (1974). A note on the asymptotic independence of high level crossings for dependent Gaussian processes. *Ann. Probab.* **2**, 535–539.

Lindgren, G. (1975a). Prediction from a random time point. *Ann. Probab.* **3**, 412–423.

Lindgren, G. (1975b). Prediction of catastrophes and high level crossings. *Bull. Int. Statist. Inst.* **46**, Book 2, 225–240.

Lindgren, G. (1977). Functional limits of empirical distributions in crossing theory. *Stochastic Process. Appl.* **5**, 143–149.

Lindgren, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Stochastic Process. Appl.* **11**, 93–117.

Lindgren, G. (1980a). Model processes in nonlinear prediction with application to detection and alarm. *Ann. Probab.* **8**, 775–792.

Lindgren, G. (1980b). Point processes of exits by bivariate Gaussian processes and extremal theory for the χ^2-process and its concomitants. *J. Multivariate Anal.* **10**, 181–206.

Lindgren, G. (1980c). Extreme values and crossings for the χ^2-process and other functions of multidimensional Gaussian processes, with reliability applications. *Adv. Appl. Probab.* **12**, 746–774.

Lindgren, G. (1981). Jumps and bumps on random roads. *J. Sound and Vibration* **78**, 383–395.

Lindgren, G. (1982a). On the shape and duration of FM-clicks. *Techn. Report 1982:1*, Dept. of Math. Statist., Univ. of Lund.

Lindgren, G. (1982b). Extremal rank and transformation of variables for extremes of functions of multivariate Gaussian processes. *Techn. Report 1982:4*, Dept. of Math. Statist., Univ. of Lund.

Lindgren, G. and Rychlik, I. (1982). Wave characteristic distributions for Gaussian waves—wave-length, amplitude and steepness. *Ocean Engineering* **9**, 411–432.

Lindgren, G., Maré, J. de., and Rootzén, H. (1975). Weak convergence of high level crossings and maxima for one or more Gaussian processes. *Ann. Probab.* **3**, 961–978.

Logan, B. F. and Shepp, L. A. (1968a). Real zeros of random polynomials. *Proc. London Math. Soc.* **18**, 29–35.
Logan, B. F. and Shepp, L. A. (1968b). Real zeros of random polynomials II. *Proc. London Math. Soc.* **18**, 308–314.
Loynes, R. M. (1965). Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36**, 993–999.
Macak, I. K. (1979a). On the number of crossings of a curve by a random process. *Theory Probab. Math. Statist.* **17**, 101–110.
Macak, I. K. (1979b). On the tangencies and intersections of a random process with a continuous curve. *Theory Probab. Math. Statist.* **18**, 103–107.
Malevich, T. L. (1974). Some bounds for the probabilities of events generated by Gaussian processes and fields, and applications to problems of level crossings. *Theory Probab. Appl.* **19**, 142–153.
Malevich, T. L. (1979). On conditions for the moments of the number of zeros of Gaussian stationary processes to be finite. *Theory Probab. Appl.* **24**, 741–753.
Malevich, T. L. and Toisov, N. (1979). On bounds for the maximum of a stationary Gaussian process. *Theory Probab. Math. Statist.* **18**, 93–101.
Marcus, M. (1970). A bound for the distribution of the maximum of continuous Gaussian processes. *Ann. Math. Statist.* **41**, 305–309.
Marcus, M. B. (1972). Upper bounds for the asymptotic maxima of continuous Gaussian processes. *Ann. Math. Statist.* **43**, 522–533.
Marcus, M. B. (1974). Asymptotic maxima of continuous Gaussian processes. *Ann. Probab.* **2**, 702–713.
Marcus, M. B. (1977). Level crossings of a stochastic process with absolutely continuous sample paths. *Ann. Probab.* **5**, 52–71.
Marcus, M. B. and Pinsky, M. (1969). On the domain of attraction of $\exp(-\exp(-x))$. *J. Math. Anal. Appl.* **28**, 440–449.
Mardia, K. V. (1964). Asymptotic independence of bivariate extremes. *Calcutta Statist. Assoc. Bull.* **13**, 172–178.
Maré, J. de (1977a). Reconstruction of a stationary Gaussian process from its sign-changes. *J. Appl. Probab.* **14**, 38–57.
Maré, J. de (1977b). The behaviour of a non-differentiable stationary Gaussian process after a level crossing. *Stochastic Process. Appl.* **6**, 77–86.
Maré, J. de (1980). Optimal prediction of catastrophes with applications to Gaussian processes. *Ann. Probab.* **8**, 841–850.
Maruyama, G. (1949). The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyusyu Univ.* A4, 45–106.
Matthes, K., Kerstan, J., and Mecke, J. (1978). *Infinitely Divisible Point Processes.* New York: Wiley.
McCord, J. R. (1964). On asymptotic moments of extreme statistics. *Ann. Math. Statist.* **35**, 1738–1745.
McCormick, W. P. (1980a). An extension to a strong law result of Mittal and Ylvisaker for the maxima of stationary Gaussian processes. *Ann. Probab.* **8**, 498–510.
McCormick, W. P. (1980b). Weak convergence for the maxima of stationary Gaussian processes using random normalization. *Ann. Probab.* **8**, 483–497.
McCormick, W. P. and Mittal, Y. (1976). On weak convergence of the maximum. *Techn. Report No. 81*, Dept. of Statist., Stanford Univ.
Mejzler, D. (1978). Limit distributions for the extreme order statistics. *Canad. Math. Bull.* **21**, 447–459.
Metcalfe, A. G. and Smitz, G. K. (1964). Effect of length on the strength of glass fibres. *Proc. Amer. Soc. Testing and Materials* **64**, 1075–1093.
Mikhailov, V. G. (1974). Asymptotic independence of vector components of multivariate extreme order statistics. *Theory Probab. Appl.* **19**, 817–821.

Miroshin, R. N. (1973). A sufficient condition for finiteness of moments of the number of zeros of a differentiable stationary Gaussian process. *Theory Probab. Appl.* **18**, 454–463.

Miroshin, R. N. (1974). A necessary condition for finiteness of the moments of the number of zeros of a differentiable Gaussian stationary process. *Theory Probab. Appl.* **19**, 570–577.

Mises, R. von (1923). Über die Variationsbreite einer Beobachtungsreihe. In: *Selected Papers*, Vol. II, pp. 129–134, Providence, R.I.: Amer. Math. Soc.

Mises, R. von (1954). La distribution de la plus grande de n valeurs. In: *Selected papers*, Vol. II, pp. 271–294. Providence, R.I.: Amer. Math. Soc.

Mittal, Y. (1974). Limiting behaviour of maxima in stationary Gaussian sequences. *Ann. Probab.* **2**, 231–242.

Mittal, Y. (1978). Maxima of partial samples in Gaussian sequences. *Ann. Probab.* **6**, 421–432.

Mittal, Y. (1979). A new mixing condition for stationary Gaussian processes. *Ann. Probab.* **7**, 724–730.

Mittal, Y. and Ylvisaker, D. (1975). Limit distributions for the maxima of stationary Gaussian processes. *Stochastic Process. Appl.* **3**, 1–18.

Mittal, Y. and Ylvisaker, D. (1976). Strong laws for the maxima of stationary Gaussian processes. *Ann. Probab.* **4**, 357–371.

Mollison, D., Buneman, O. P., and Salter, S. H. (1976). Wave power availability in the NE Atlantic. *Nature* **263**, 223–226.

Mori, T. (1976). Limit laws for maxima and second maxima for strong-mixing processes. *Ann. Probab.* **4**, 122–126.

Mori, T. (1981). The relation of sums and extremes of random variables. *Proc. 43rd ISI Meeting*, Buenos Aires.

Nagaev, A. V. (1970). On the role of the largest of the order statistics in the formation of a large deviation of a sum of independent random variables. *Soviet Math. Dokl.* **11**, 972–974.

Nagaev, A. V. (1971). Limit distribution of the extreme terms of a variational series under conditions of large deviations. *Theory Probab. Appl.* **16**, 126–140.

Nair, K. A. (1976). Bivariate extreme value distributions. *Comm. Statist.*, A **5**, 575–581.

Nair, K. A. (1981). Asymptotic distributions and moments of normal extremes. *Ann. Probab.* **9**, 150–153.

Nakagawa, T. (1978). Discrete extreme distributions. *IEEE Trans. Reliab.* **27**, 367–368.

National Bureau of Standards (1953). Probability tables for the analysis of extreme value data. *Appl. Math. Ser.* **22**.

Neuts, M. F. (1967). Waitingtimes between record observations. *J. Appl. Probab.* **4**, 206–208.

Newell, G. F. (1962). Asymptotic extreme value distribution for one-dimensional diffusion processes. *J. Math. Mech.* **11**, 481–496.

Newell, G. F. (1964). Asymptotic extremes for m-dependent random variables. *Ann. Math. Statist.* **35**, 1322–1325.

Nisio, M. (1967). On the extreme values of Gaussian processes. *Osaka J. Math.* **4**, 313–326.

O'Brien, G. L. (1974a). Limit theorems for the maximum term of a stationary process. *Ann. Probab.* **2**, 540–545.

O'Brien, G. L. (1974b). The limiting distribution of maxima of random variables defined on a denumerable Markov chain. *Ann. Probab.* **2**, 103–111.

O'Brien, G. L. (1974c). The maximum term of uniformly mixing stationary processes. *Z. Wahrsch. verw. Gebiete* **30**, 57–63.

Ochi, M. K. (1973). On prediction of extreme values. *J. Ship Res.* **17**, 29–37.

Öfverbeck, P. and Östberg, G. (1977). A study of outliers in statistical distributions of mechanical properties of structural steel. *Techn. Report*, Dept. of Structural Engineering, Univ. of Lund., pp. 1–37.

Osborn, P. W. and Schilling, D. L. (1971). Threshold analysis of phase-locked-loop demodulators using most likely noise. *IEEE Trans. Comm.* **19**, 31–41.

Pakshirajan, R. P. and Hebbar, H. V. (1977). Limit theorems for extremes of absolute values of Gaussian sequences. *Sankhyā, Ser. A*, **39**, 191–195.

Palm, C. (1943). Intensitätsschwankungen in Fernsprechverkehr. *Ericson Technics*, pp. 1–189.

Phoenix, S. L. and Taylor, H. M. (1973). The asymptotic strength distribution of a general fiber bundle. *Adv. Appl. Probab.* **5**, 200–216.

Pickands, J. III (1967a). Sample sequences of maxima. *Ann. Math. Statist.* **38**, 1570–1574.

Pickands, J. III (1967b). Maxima of stationary Gaussian processes. *Z. Wahrsch. verw. Gebiete* **7**, 190–233.

Pickands, J. III (1968). Moment convergence of sample extremes. *Ann. Math. Statist.* **39**, 881–889.

Pickands, J. III (1969a). Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145**, 51–73.

Pickands, J. III (1969b). Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.* **145**, 75–86.

Pickands, J. III (1971). The two-dimensional Poisson process and extremal processes. *J. Appl. Probab.* **8**, 745–756.

Pickands, J. III (1981). Multivariate extreme value distributions. *Proc. 43rd ISI Meeting*, Buenos Aires.

Piterbarg, V. I. (1978). Asymptotic expansions for the probability of large excursions of Gaussian processes. *Soviet Math. Dokl.* **19**, 1279–1283.

Piterbarg, V. I. (1981). Comparison of distribution functions of maxima of Gaussian processes. *Teor. Ver. Prim.* **26**, 702–719 (in Russian).

Piterbarg, V. I. and Prisjaznjuk, V. P. (1979). Asymptotics of the probability of large excursions for a nonstationary Gaussian process. *Theory Probab. Math. Statist.* **18**, 131–144.

Piterbarg, V. I. and Prisjaznjuk, V. P. (1981). Exact asymptotic behavior of the probability of a large span of a stationary Gaussian process. *Theory Probab. Appl.* **26**, 468–484.

Posner, E. C. (1965). The application of extreme-value theory to error-free communication. *Technometrics* **7**, 517–529.

Poznjakov, V. V. (1972). Approximate formulas for the probability of noncrossing of level by stationary process. In: *Col. Janos Bolyai, 9th European Meeting of Statisticians*, Vol. II, Ed. J. Gani, pp. 637–640.

Puri, M. L. and Tran, L. T. (1979). Local maxima of the sample functions of the N-parameter Bessel process. *Stochastic Process. Appl.* **9**, 137–145.

Qualls, C. (1968). On a limit distribution of high level crossings of a stationary Gaussian process. *Ann. Math. Statist.* **39**, 2108–2113.

Qualls, C. (1969). On the joint distribution of crossings of high multiple levels by a stationary Gaussian process. *Ark. Mat.* **8**, 129–137.

Qualls, C. (1970). On the number of zeros of a stationary Gaussian random trigonometric polynomial. *J. London Math. Soc.* **2**, 216–220.

Qualls, C. and Watanabe, H. (1972). Asymptotic properties of Gaussian processes. *Ann. Math. Statist.* **43**, 580–596.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley.

Reiss, R.-D. (1981). Uniform approximation to distributions of extreme order statistics. *Adv. Appl. Probab.* **13**, 533–547.

Rényi, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hungar.* **4**, 191–229.

Resnick, S. I. (1971). Tail equivalence and its applications. *J. Appl. Probab.* **8**, 136–156.

Resnick, S. I. (1972). Products of distribution functions attracted to extreme value laws. *J. Appl. Probab.* **8**, 781–793.
Resnick, S. I. (1973a). Limit laws for record values. *Stochastic Process. Appl.* **1**, 67–82.
Resnick, S. I. (1973b). Record values and maxima. *Ann. Probab.* **1**, 650–662.
Resnick, S. I. (1973c). Extremal processes and record value times. *J. Appl. Probab.* **10**, 864–868.
Resnick, S. I. (1974). Inverses of extremal processes. *Adv. Appl. Probab.* **6**, 392–406.
Resnick, S. I. (1975). Weak convergence to extremal processes. *Ann. Probab.* **3**, 951–960.
Resnick, S. I. and Rubinovitch, M. (1973). The structure of extremal processes. *Adv. Appl. Probab.* **5**, 287–307.
Rice, S. O. (1939). The distribution of the maxima of a random curve. *Amer. J. Math.* **61**, 409–416.
Rice, S. O. (1944, 1945). Mathematical analysis of random noise. *Bell Syst. Tech. J.* **23**, 282–332, and **24**, 46–156.
Roberts, E. M. (1979a). Review of statistics of extreme values with applications to air quality data, Part I: review. *J. Air Pollut. Control Assoc.* **29**, 632–637.
Roberts, E. M. (1979b). Review of statistics of extreme values with applications to air quality data, Part II: applications. *J. Air Pollut. Control Assoc.* **29**, 733–740.
Roes, P. B. M. (1970). On the expected number of crossings of a level in certain stochastic processes. *J. Appl. Probab.* **7**, 766–770.
Rootzén, H. (1974). Some properties of convergence in distribution of sums and maxima of dependent random variables. *Z. Wahrsch. verw. Gebiete* **29**, 295–307.
Rootzén, H. (1978). Extremes of moving averages of stable processes. *Ann. Probab.* **6**, 847–869.
Rootzén, H. (1982). The rate of convergence of extremes of stationary normal sequences. To appear in *Adv. Appl. Probab.*
Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.* **42**, 43–47.
Rossberg, H. J. (1964a). Über die stochastische Unabhängigkeit gevisser Funktionen von Ranggrössen. *Math. Nachr.* **28**, 157–167.
Rossberg, H. J. (1964b). Die asymptotische Unabhängigkeit der kleinsten und grössten Werte einer Stichprobe vom Stichprobenmittel. *Math. Nachr.* **28**, 305–318.
Sahran, J. (1977). Joint distributions based on zeros, horizontal crossings and maximum deviations. *Z. Angew. Math. Mech.* **57**, 610–613.
Sarhan, A. E. and Greenberg, B. G. (1962). *Contributions to Order Statistics.* New York: Wiley.
Sato, H. and Tanabe, M. (1979). Level dependence of the variance of the number of level-crossing points of a stationary Gaussian process. *IEEE Trans. Inform. Theory* **25**, 127–129.
Sellars, F. (1975). Maximum heights of ocean waves. *J. Geophys. Res.* **80**, 398–404.
Sen, P. K. (1961). A note on the large sample behaviour of extreme sample values from distributions with finite endpoints. *Bull. Calcutta Statist. Assoc.* **10**, 106–115.
Sen, P. K. (1972). On weak convergence of extremal processes for random sample sizes. *Ann. Math. Statist.* **43**, 1355–1362.
Serfling, R. J. (1975). A general Poisson approximation theorem. *Ann. Probab.* **3**, 726–731.
Serfling, R. J. (1978). Some elementary results on Poisson approximation in a sequence of Bernoulli trials. *SIAM Rev.* **20**, 567–579.
Serfozo, R. (1980). High-level exceedances of regenerative and semi-stationary processes. *J. Appl. Probab.* **17**, 423–431.
Sethuraman, J. (1965). On a characterization of the three limiting types of the extreme. *Sankhyā, Ser. A*, **27**, 357–364.

Sharpe, K. (1974). On the effect of the anemometer on the stochastic properties of wind-speed data. *Techn. Report 34*, School of Math. Sci., Univ. of Melbourne.

Sharpe, K. (1978). Some properties of the crossings process generated by a stationary χ^2-process. *Adv. Appl. Probab.* **10**, 373–391.

Sharpe, K. (1979). Level crossings of a constructed process. *J. Appl. Probab.* **16**, 206–212.

Shenker, M. (1981). The mean number of real roots for one class of random polynomials. *Ann. Probab.* **9**, 510–512.

Shepp, L. A. (1971). First passage time for a particular Gaussian process. *Ann. Math. Statist.* **42**, 946–951.

Shepp, L. A. and Slepian, D. (1976). First-passage time for a particular stationary periodic Gaussian process. *J. Appl. Probab.* **13**, 27–38.

Shorrock, R. W. (1972a). A limit theorem for inter-record times. *J. Appl. Probab.* **9**, 219–223 (Correction (1972) **9**, 877).

Shorrock, R. W. (1972b). On record values and record times. *J. Appl. Probab.* **9**, 316–326.

Shorrock, R. W. (1973). Record values and inter-record times. *J. Appl. Probab.* **10**, 543–555.

Shorrock, R. W. (1974). On discrete time extremal processes. *Adv. Appl. Probab.* **6**, 580–592.

Shorrock, R. W. (1975). Extremal processes and random measures. *J. Appl. Probab.* **12**, 316–323.

Shur, M. G. (1965). On the maximum of a Gaussian stationary process. *Theory Probab. Appl.* **10**, 354–357.

Sibuya, M. (1960). Bivariate extremal distributions I. *Ann. Inst. Statist. Math.* **11**, 195–210.

Siddiqui, M. M. and Biodini, R. W. (1975). The joint distribution of record values and inter-record times. *Ann. Probab.* **3**, 1012–1013.

Simiu, E. and Filliben, J. J. (1976). Probability distributions of extreme wind speeds. *J. Structural Div. ASCE* **102**, 1861–1877.

Singh, C. (1967). On the extreme values and range of samples from non-normal populations. *Biometrika* **54**, 541–550.

Singpurwalla, N. D. (1971). Statistical fatigue models, a survey. *IEEE Trans. Reliab.* **20**, 185–189.

Singpurwalla, N. D. (1972). Extreme values from a lognormal law with applications to air pollution problems. *Technometrics* **14**, 703–711.

Sirao, T. and Watanabe, H. (1970). On the upper and lower class for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **147**, 301–331.

Slepian, D. (1961). First passage time for a particular Gaussian process. *Ann. Math. Statist.* **32**, 610–612.

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Techn. J.* **41**, 463–501.

Slepian, D. (1963). On the zeros of Gaussian noise. In: *Time Series Analysis*, Ed. M. Rosenblatt, pp. 104–115. New York: Wiley.

Smirnov, N. V. (1952). Limit distributions for the terms of a variational series. *Amer. Math. Soc. Transl. Ser. I*, **11**, 82–143.

Smirnov, N. V. (1967). Some remarks on limit laws for order statistics. *Theory Probab. Appl.* **12**, 336–337.

Smith, R. L. (1980). A probability model for fibrous composites with local load sharing. *Proc. Roy. Soc. London, Ser. A*, **372**, 539–553.

Srivastava, O. P. (1967). Asymptotic independence of certain statistics connected with the extreme order statistics in a bivariate distribution. *Sankhyā, Ser. A*, **29**, 175–182.

Suthasinekul, S. (1976). Level crossing patterns of time series: application to signal processing. *Proc. 3rd Int. Conf. Pattern Recognition*, pp. 524–528.

Tata, M. N. (1969). On outstanding values in a sequence of random variables. *Z. Wahrsch. verw. Gebiete* **12**, 9–20.

Teugels, J. L. (1981). Limit theorems on order statistics. *Ann. Probab.* **9**, 868–880.

Thom, H. C. S. (1971). Asymptotic extreme value distributions of wave heights in the open ocean. *J. Marine Res.* **29**, 19–27.

Tiago de Oliveira, J. (1958). Extremal distributions. *Rev. Fac. Sci. Lisboa, Ser. A*, **7**, 215–227.

Tiago de Oliveira, J. (1961). The asymptotical independence of the sample mean and the extremes. *Rev. Fac. Sci., Lisboa, Ser. A*, **8**, 299–310.

Tiago de Oliveira, J. (1962/63). Structure theory of bivariate extremes: extensions. *Estudos Math. Estat. Econom.* **7**, 165–195.

Tiago de Oliveira, J. (1968). Extremal processes: definition and properties. *Publ. Inst. Statist. Univ. Paris* **17**, 25–36.

Tiago de Oliveira, J. (1970). Biextremal distributions: statistical decision. *Trab. Estat. Invest. Oper.* **21**, 107–117.

Tiago de Oliveira, J. (1975a). Bivariate extremes: extensions. *Bull. Int. Stat. Inst.* **46**, Book 2, 241–252.

Tiago de Oliveira, J. (1975b). Bivariate and multivariate extremal distributions. In: *Statistical Distributions in Scientific Work*, Ed. G. P. Patil *et al.*, Vol. 1, pp. 355–361. Dordrecht-Holland: Reidel.

Tiago de Oliveira, J. (1976). Asymptotic behaviour of maxima with periodic disturbances. *Ann. Inst. Statist. Math.* **28**, 19–23.

Tiago de Oliveira, J. (1980). Bivariate extremes: foundations and statistics. In: *Proc. Fifth Int. Symp. Multivar. Anal.*, Ed. P. R. Krishnaia, pp. 349–366. Amsterdam: North Holland.

Tippett, L. H. C. (1925). On the extreme individuals and the range of samples taken from a normal population. *Biometrika* **17**, 364–387.

Todorovic, P. (1979). A probabilistic approach to analysis and prediction of floods. *Proc. 43rd Session ISI*, Buenos Aires.

Tsirel'son, V. S. (1975). The density of the distribution of the maximum of a Gaussian process. *Theory Probab. Appl.* **20**, 847–856.

Umbach, D. (1978). On the approximate behaviour of the posterior distribution for an extreme multivariate observation. *J. Multivariate Anal.* **8**, 518–531.

Veneziano, D. (1979). Envelopes of vector random processes and their crossing rates. *Ann. Probab.* **7**, 62–74.

Veneziano, D., Grigoriou, M., and Cornell, C. A. (1977). Vector process models for system reliability. *J. Engineering Mechanics Div. ASCE* **103**, 441–460.

Vervaat, W. (1973). Limit theorems for records from discrete distributions. *Stochastic Process. Appl.* **1**, 317–334.

Villasenor, J. (1978). Attraction of convolutions to extreme distributions. Techn. Rep. Colegio de Postgraduados, México.

Villasenor, J. A. (1981). Norming constants for maxima attracted to $\exp(-\exp(-x))$. *Proc. 43rd Session ISI*, Buenos Aires.

Volkonskii, V. A. (1960). An ergodic theorem on the distribution of the duration of fades. *Theory Probab. Appl.* **5**, 323–326.

Volkonskii, V. A. and Rozanov, Yu. A. (1959). Some limit theorems for random functions, I. *Theory Probab. Appl.* **4**, 178–197.

Volkonskii, V. A. and Rozanov, Yu. A. (1961). Some limit theorems for random functions, II. *Theory Probab. Appl.* **6**, 186–198.

Walsh, J. E. (1969a). Asymptotic independence between largest and smallest of a set of independent observations. *Ann. Inst. Statist. Math.* **21**, 287–289.

Walsh, J. E. (1969b). Approximate distributions for largest and for smallest of a set of independent observations. *South African Statist. J.* **3**, 83–89.

Walsh, J. E. (1970). Sample sizes for approximate independence of largest and smallest order statistics. *J. Amer. Statist. Assoc.* **65**, 860–863.

Watson, G. S. (1954). Extreme values in samples from m-dependent stationary stochastic processes. *Ann. Math. Statist.* **25**, 798–800.

Watts, J. H. V. (1977). Limit theorems and representations for order statistics from dependent sequences. Ph.D. thesis, Statistics Dept., Univ. of N.C.

Watts, V. (1980). The almost sure representation of intermediate order statistics. *Z. Wahrsch. verw. Gebiete* **54**, 281–285.

Watts, V., Rootzén, H., and Leadbetter, M. R. (1982). Convergence in distribution of intermediate order statistics from stationary sequences. *Ann. Probab.* **10**, 653–662.

Webber, M. (1979). Analyse asymptotique des processus Gaussienns stationaires. *Inst. Res. Math. Avan. Strasbourg.*

Weinstein, S. B. (1973). Theory and application of some classical and generalized asymptotic distributions of extreme values. *IEEE Trans. Inform. Theory* **19**, 148–154.

Weiss, L. (1969). The joint asymptotic distribution of the k-smallest sample spacings. *J. Appl. Probab.* **6**, 442–448.

Weissman, I. (1975a). Extremal processes generated by independent nonidentically distributed random variables. *Ann. Probab.* **3**, 172–177.

Weissman, I. (1975b). On location and scale functions of a class of limiting processes with applications to extreme value theory. *Ann. Probab.* **3**, 178–181.

Weissman, I. (1975c). Multivariate extremal processes generated by independent non-identically distributed random variables. *J. Appl. Probab.* **12**, 477–487.

Welsch, R. E. (1971). A weak convergence theorem for order statistics from strong-mixing processes. *Ann. Math. Statist.* **42**, 1637–1646.

Welsch, R. E. (1972). Limit laws for extreme order statistics from strong-mixing processes. *Ann. Math. Statist.* **43**, 439–446.

Welsch, R. E. (1973). A convergence theorem for extreme values from Gaussian sequences. *Ann. Probab.* **1**, 398–404.

Williams, D. (1973). On Rényi's "record" problem and Engel's series. *Bull. London Math. Soc.* **5**, 235–237.

Wong, E. (1966). Some results concerning the zero-crossings of Gaussian noise. *SIAM J. Appl. Math.* **14**, 1246–1254.

Woods, R. E. and Gonzales, R. C. (1982). Sampling considerations for multilevel-crossing analysis. *IEEE Trans. Pattern Anal. Machine Intelligence* **4**, 117–123.

Wu, Chuan-Yi (1966). The types of limit distributions for some terms of a variational series. *Sci. Sinica* **15**, 749–762.

Yang, M. C. K. (1975). On the distribution of the inter-record times in an increasing population. *J. Appl. Probab.* **12**, 148–154.

Ylvisaker, N. D. (1965). The expected number of zeros of a stationary Gaussian process. *Ann. Math. Statist.* **36**, 1043–1046.

Ylvisaker, N. D. (1966). On a theorem of Cramér and Leadbetter. *Ann. Math. Statist.* **37**, 682–685.

Ylvisaker, N. D. (1968). A note on the absence of tangencies in Gaussian sample paths. *Ann. Math. Statist.* **39**, 261–262.

Zähle, U. (1980). A Poisson limit theorem for rare events of a Gaussian sequence. *Theory Probab. Appl.* **25**, 90–104.

Zidek, J. V., Navin, F. P. D., and Lockhart, R. (1979). Statistics of extremes: an alternate method with application to bridge design codes. *Technometrics* **21**, 185–191.

List of Special Symbols

$\#\{\ \}$	number of elements
$\Gamma(\alpha)$	gamma function, $\int_0^\infty x^{\alpha-1}e^{-x}\,dx$
$[\ ]$	integer part
$m(B)$	Lebesgue measure

$f(t) \sim g(t)$, $o(\psi(t))$, $O(\psi(t))$ have their standard meaning, i.e., $f(t)/g(t) \to 1$, $o(\psi(t))/\psi(t) \to 0$, and $O(\psi(t))/\psi(t)$ is bounded, respectively.

Index