Estimation methods for Value at Risk

1 Introduction

1.1 History of VaR

In the last few decades, risk managers have truly experienced a revolution. The rapid increase in the usage of risk management techniques has spread well beyond derivatives and is totally changing the way institutions approach their financial risk. In response to the financial disasters of the early 1990s a new method called VaR (Value at Risk) was developed as a simple method to quantify market risk (In recent years, VaR has been used in many other areas of risk including credit risk and operational risk). Some of the financial disasters of the early 1990s are:

- Figure 1 (taken from http://www.brighthub.com / money / investing / articles / 126337.aspx) shows the effect of Black Monday, which occurred on 19 October 1987. In a single day, the Dow Jones stock index (DJIA) crashed down by 22.6 percent (by 508 points), causing a negative knock on effect on other stock markets worldwide. Overall the stock market lost $0.5 trillion;
- the Japanese stock price bubble, creating a $2.7 trillion loss in capital, see Figure 2 taken from http://chovanec.wordpress.com/. According to this website, "the Nikkei Index after the Japanese bubble burst in the final days of 1989. Again, the market showed a substantial recovery for several months in mid-1990 before sliding to new lows";
- Figure 3 (taken from http: // steadfastfinances.com / 2009 / 11 / 14 / the-psychology-of-bubbles-using-hindsight-to-examine-why-we-bought - into - the - hype /) describes the dot.com bubble. During 1999 and 2000, the NASDAQ rose at a dramatic rate with all technology stocks booming. However, on 10 March 2000, the bubble finally burst, because of a sudden simultaneous sell orders in big technology companies (Dell, IBM, Cisco) on the NASDAQ. After a peak at $5048.62 on that day, the NASDAQ fell back down and has never since recovered;
- Figure 4 describes the 1997 Asian financial crisis. It first occurred at the beginning in July 1997. During that period a lot of Asia got affected by this financial crisis, leading to a pandemic spread of fear to a worldwide economic meltdown. The crisis was first triggered when the Thai baht (Thailand currency) was cut from being pegged to the US dollars and the government floated the baht. In addition, at the time Thailand was effectively bankrupt from the burden of foreign debt it acquired. Later period saw a contagious spread of the crisis to Japan and to South Asia, causing a slump in asset prices, stock market and currencies;
- the Black Wednesday, resulting in £800 million losses, see Figure 5 taken from http: // www.telegraph.co.uk / news / uknews / 1483186 / Major-was-ready-to-quit-over-Black-Wednesday.html; According to http: // en.wikipedia.org / wiki / Black_Wednesday, Black Wednesday "refers to the events of 16 September 1992 when the British Conservative government was forced to

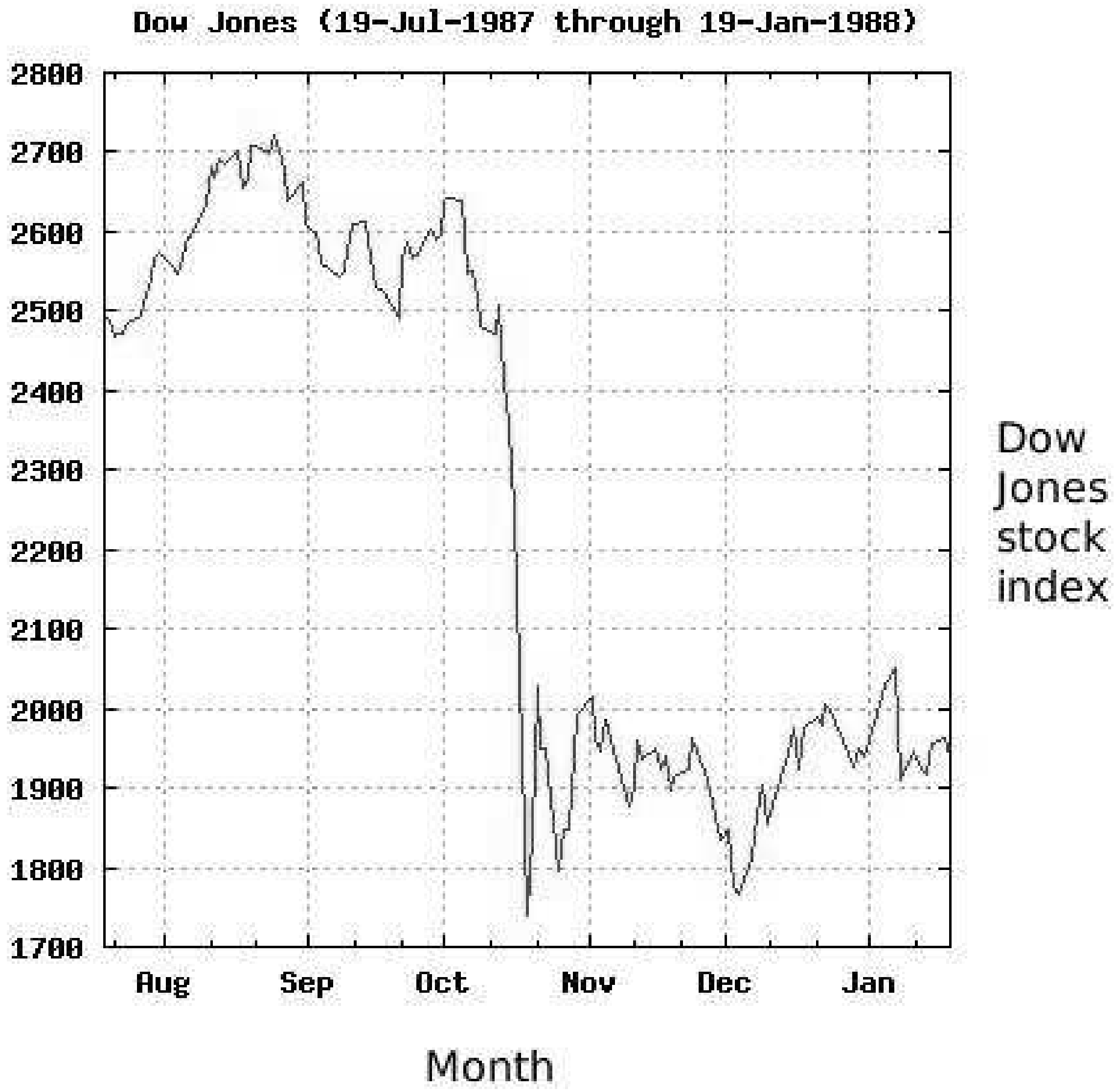

Figure 1: Black Monday crash on 19 October 1987. The Dow Jones stock index crashed down by 22.6 percent (by 508 points). Overall the stock market lost $0.5 trillion.

withdraw the pound sterling from the European Exchange Rate Mechanism (ERM) after they were unable to keep it above its agreed lower limit";

- and the infamous financial disasters of Orange County, Barings, Metallgesellschaft, Daiwa and so many more.

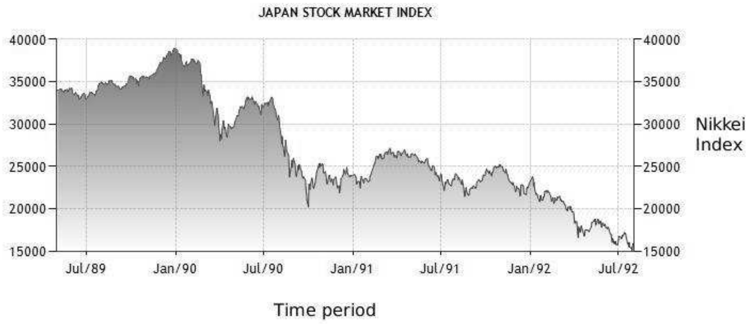

Figure 2: Japan stock price bubble near the end of 1989. A loss of \$2.7 trillion in capital. A recovery happened after mid-1990.

1.2 Definition of VaR

Till Guldimann is widely credited as the creator of value at risk (VaR) in the late 1980s. He was then the head of global research at J.P. Morgan. VaR is a method that uses standard statistical techniques to assess risk. The VaR "measures the worst average loss over a given horizon under normal market conditions at a given confidence level" (Jorion, 2011, page xxii). The value of VaR can provide users with information in two ways: as a summary measure of market risk, or an aggregate view of a portfolio's risk. Overall VaR is a forward looking risk measure and used by financial institutions, regulators, non financial corporations and asset management exposed to financial risk. The most important use of VaR has been for capital adequacy regulation under Basel II and later revisions.

Let $\{X_t, t = 1, 2, \ldots, n\}$ denote a stationary financial series with marginal cumulative distribution function (cdf) F and marginal probability density function (pdf) f. The Value at Risk for a given probability p is defined mathematically as

$$\mathrm{VaR}_p = \inf \{u : F(u) \geq p\}. \tag{1}$$

That is, VaR is the quantile of F exceeded with probability $1-p$. Figure 6 illustrates the definition given by (1).

Sometimes, VaR is defined for log-returns of the original time series. That is, if $R_t = \ln(X_{t+h}/X_t)$, $t = 1, 2, \ldots, n$ are the log-returns for some h with marginal cdf F then VaR is defined by (1). If α_h and σ_h denote the mean and standard deviation of the log-returns then one can write

$$\mathrm{VaR}_p = \alpha_h + \sigma_h \psi^{-1}(p), \tag{2}$$

where $\psi(\cdot)$ denotes the quantile function of the standardized log-returns $Z_t = (R_t - \alpha_h)/\sigma_h$.

1.3 Applications of VaR

Applications of VaR can be classified as:

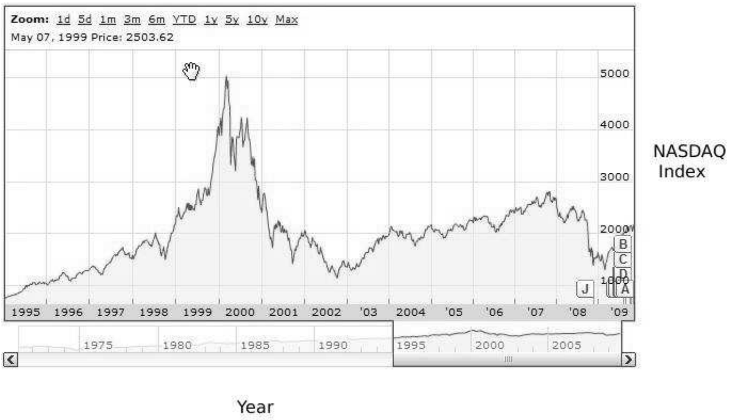

Figure 3: Dot com bubble (the NASDAQ index) during 1999 and 2000. The bubble burst on 10 March 2000. The peak on that day was \$5048.62. There is a recovery after 2002. Never recovered to attain the peak.

- Information reporting - it measures aggregate risk and corporation risk in a non technical way for easy understanding;
- Controlling risk - setting position limits for traders and business units, so they can compare diverse market risky activities;
- Managing risk - reallocating of capital across traders, products, business units and whole institutions.

Applications of value at risk have been extensive. Some recent applications and application areas have included: estimation of highly parallel architectures (Dixon *et al.*, 2012), estimation for crude oil markets (He *et al.*, 2012), multi resolution analysis based methodology in metals markets (He *et al.*, 2012), estimation of optimal hedging strategy under bivariate regime switching ARCH framework (Chang, 2011), energy markets (Cheong, 2011), Malaysian sectoral markets (Cheong and Isa, 2011), downside residential market risk (Jin and Ziobrowski, 2011), hazardous materials transportation (Kwon, 2011), operational risk in Chinese commercial banks (Lu, 2011), longevity and mortality (Plat, 2011), analysis of credit default swaps (Raunig *et al.*, 2011), exploring oil-exporting country portfolio (Sun *et al.*, 2011), Asia-focused hedge funds (Weng and Trueck, 2011), measure for waiting time in simulations of hospital units (Dehlendorff *et al.*, 2010), financial risk in pension funds (Fedor, 2010), catastrophic event modeling in the Gulf of Mexico (Kaiser *et al.*, 2010), estimating the South African equity market (Milwidsky and Mare, 2010), estimating natural disaster risks (Mondlane, 2010), wholesale price for supply chain coordination (Wang, 2010), U.S. movie box office earnings (Bi and Giles, 2009), stock market index portfolio in South Africa (Bonga-Bonga and Mutema, 2009), multi-period supply inventory coordination (Cai *et al.*, 2009), Toronto stock exchange (Dionne *et al.*, 2009), modeling volatility clustering in electricity price return series (Karandikar *et al.*, 2009), calculation for heterogeneous loan portfolios (Puzanova *et al.*, 2009), measurement of HIS stock index futures market risk (Yan and Gong, 2009), stock index futures market risk (Gong and Li, 2008), estimation of real estate values (He *et al.*, 2008), foreign exchange rates (Ku and Wang, 2008), artificial neural network (Lin and Chen, 2008), criterion for

Figure 4: Asian financial crisis (Asian dollar index) in July 1997. Not fully recovered even in 2011.

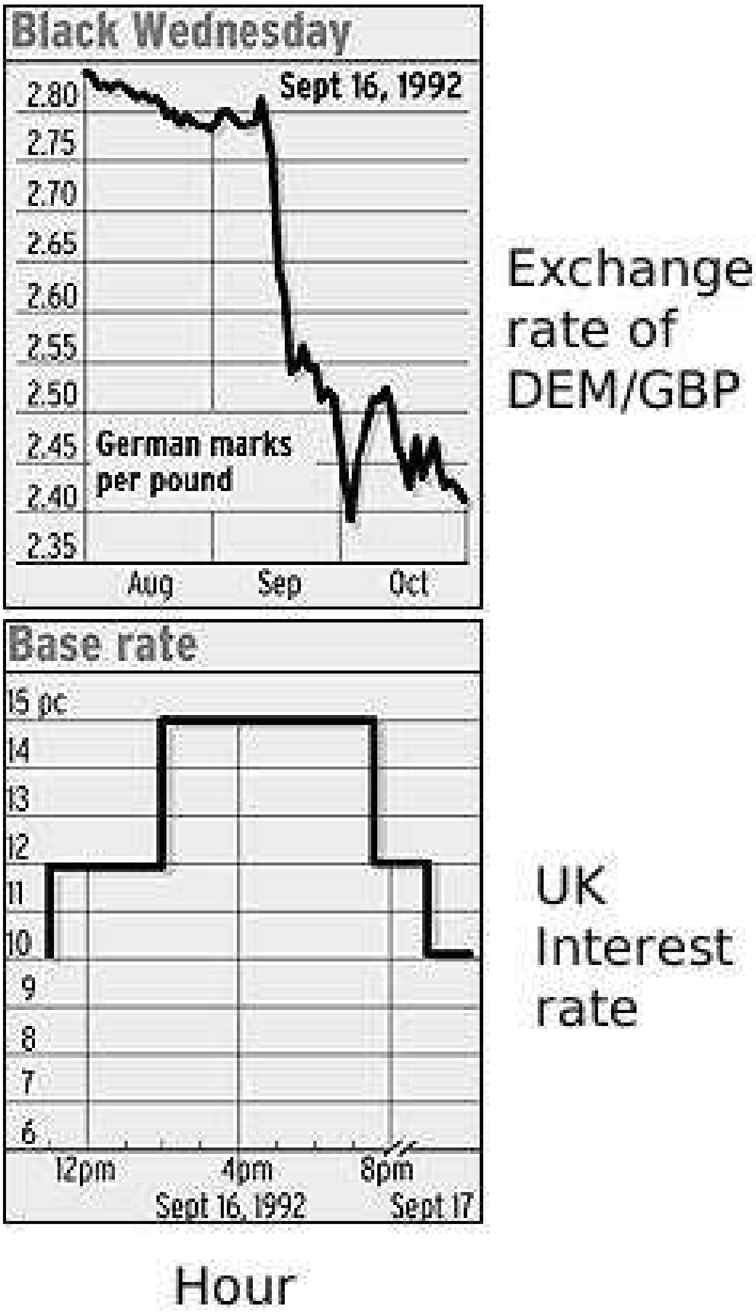

Figure 5: Black Wednesday crash of 16 September 1992. Top image shows the exchange rate of Deutsche mark to British pounds. Bottom image shows the UK interest rate on the day.

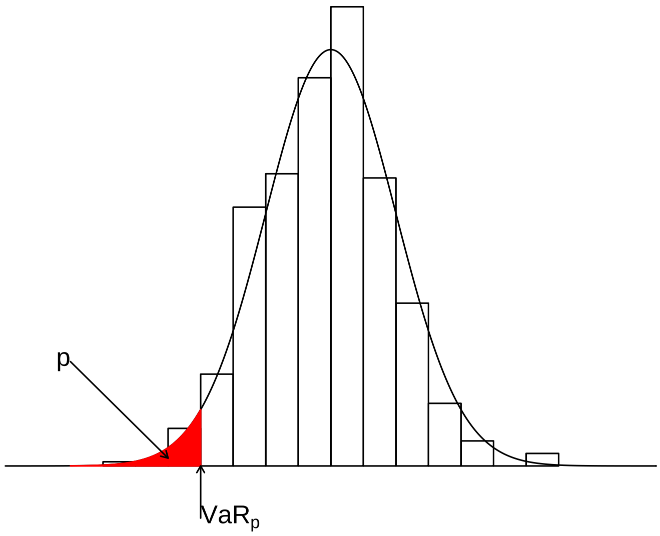

Figure 6: Value at risk illustrated.

management of storm-water (Piantadosi *et al.*, 2008), inventory control in supply chains (Yiu *et al.*, 2008), layers of protection analysis (Fang *et al.*, 2007), project finance transactions (Gatti *et al.*, 2007), storms in the Gulf of Mexico (Kaiser *et al.*, 2007), mid-term generation operation planning in electricity market environment (Lu *et al.*, 2007), Hong Kong's fiscal policy (Porter, 2007), bakery procurement (Wilson *et al.*, 2007), newsvendor models (Xu and Chen, 2007), optimal allocation of uncertain water supplies (Yamout *et al.*, 2007), futures floor trading (Lee and Locke, 2006), estimating a listed firm in China (Liu *et al.*, 2006), Asian pacific stock market (Su and Knowles, 2006), Polish power exchange (Trzpiot and Ganczarek, 2006), single loss approximation to value at risk (Böcker and Klüppelberg, 2005), real options in complex engineered systems (Hassan *et al.*, 2005), effects of bank technical sophistication and learning over time (Liu *et al.*, 2004), risk analysis of the aerospace sector (Mattedi *et al.*, 2004), Chinese securities market (Li *et al.*, 2002), risk management of investment-linked household property insurance (Zhu and Gao, 2002), project risk measurement (Feng and Chen, 2001), long-term capital management for property/casualty insurers (Panning, 1999), structure-dependent securities and FX derivatives (Singh, 1997), and mortgage backed securities (Jakobsen, 1996).

1.4 Aims

The aim of this lecture notes is to review known methods for estimating VaR given by (1). The review of methods is divided as follows: general properties (Section 2), parametric methods (Section 3), nonparametric methods (Section 4), semiparametric methods (Section 5), and computer software (Section 6). For each estimation method, we give the main formulas for computing value at risk. We have avoided giving full details for each estimation method (for example, interpretation, asymptotic properties, finite sample properties, finite sample bias, sensitivity to outliers, quality of approximations, comparison with competing estimators, advantages, disadvantages and application areas) because of space concerns. These details can be read from the cited references.

1.5 Further material

The review of value of risk presented here is not complete, but we believe we have covered most of the developments in recent years. For a fuller account of the theory and applications of value risk, we refer the readers to the following books: Bouchaud and Potters (2000, Chapter 3), Delbaen (2000, Chapter 3), Moix (2001, Chapter 6), Voit (2001, Chapter 7), Dupacova *et al.* (2002, Part 2), Dash (2004, Part IV), Franke *et al.* (2004), Tapiero (2004, Chapter 10), Meucci (2005), Pflug and Romisch (2007, Chapter 12), Resnick (2007), Ardia (2008, Chapter 6), Franke *et al.* (2008), Klugman *et al.* (2008), Lai and Xing (2008, Chapter 12), Taniguchi *et al.* (2008), Janssen *et al.* (2009, Chapter 18), Sriboonchitta *et al.* (2010, Chapter 4), Tsay (2010), Capinski and Zastawniak (2011), Jorion (2011), and Ruppert (2011, Chapter 19).

2 General properties

This section describes general properties of value at risk. The properties discussed are: ordering properties (Section 2.1), upper comonotonicity (Section 2.2), multivariate extension (Section 2.3), risk concentration (Section 2.4), Hürlimann's inequalities (Section 2.5), Ibragimov and Walden's inequalities (Section 2.6), Denis *et al.*'s inequalities (Section 2.7), Jaworski's inequalities (Section 2.8), Mesfioui and Quessy's inequalities (Section 2.9) and Slim *et al.*'s inequalities (Section 2.10).

2.1 Ordering properties

Pflug (2000) and Jadhav and Ramanathan (2009) establish several ordering properties of VaR_p. Given random variables X, Y, Y_1, Y_2 and a constant c, some of the properties given by Pflug (2000) and Jadhav and Ramanathan (2009) are:

(i) VaR_p is translation equivariant, that is $\mathrm{VaR}_p(Y+c) = \mathrm{VaR}_p(Y)+c$;

(ii) VaR_p is positively homogeneous, that is $\mathrm{VaR}_p(cY) = c\mathrm{VaR}_p(Y)$ for $c > 0$;

(iii) $\mathrm{VaR}_p(Y) = -\mathrm{VaR}_{1-p}(-Y)$;

(iv) VaR_p is monotonic with respect to stochastic dominance of order 1 (a random variable Y_1 is less than a random variable Y_2 with respect to stochastic dominance of order 1 if $E\left[\psi\left(Y_1\right)\right] \leq E\left[\psi\left(Y_2\right)\right]$ for all monotonic integrable functions ψ); that is, Y_1 is less than a random variable Y_2 with respect to stochastic dominance of order 1 then $\mathrm{VaR}_p\left(Y_1\right) \leq \mathrm{VaR}_p\left(Y_2\right)$;

(v) VaR_p is comonotone additive, that is if Y_1 and Y_2 are comonotone then $\mathrm{VaR}_p\left(Y_1+Y_2\right) = \mathrm{VaR}_p\left(Y_1\right)+\mathrm{VaR}_p\left(Y_2\right)$. Two random variables Y_1 and Y_2 defined on the same probability space $(\Omega, \mathcal{A}, P)$ are said to be comonotone if for all $w, w^{'} \in \Omega$, $\left[Y_1(w) - Y_2(w)\right]\left[Y_1\left(w^{'}\right) - Y_2\left(w^{'}\right)\right] \geq 0$ almost surely;

(vi) if $X \geq 0$ then $\mathrm{VaR}_p(X) \geq 0$;

(vii) VaR_p is monotonic, that is if $X \geq Y$ then $\mathrm{VaR}_p(X) \geq \mathrm{VaR}_p(Y)$.

Let F denote the joint cdf of (X_1, X_2) with marginal cdfs F_1 and F_2. Write $F \equiv (F_1, F_2, C)$ to mean $F\left(X_1, X_2\right) \equiv C\left(F_1\left(X_1\right), F_2\left(X_2\right)\right)$, where C is known as the copula (Nelsen, 1999), a joint cdf of uniform marginals. Let (X_1, X_2) have the joint cdf $F \equiv (F_1, F_2, C)$, $\left(X_1^{'}, X_2^{'}\right)$ have the joint cdf $F^{'} \equiv \left(F_1, F_2, C^{'}\right)$, $X = wX_1 + (1-w)X_2$, and $X^{'} = wX_1^{'} + (1-w)X_2^{'}$. Then, Tsafack (2009) shows that if $C^{'}$ is stochastically less than C then $\mathrm{VaR}_p\left(X^{'}\right) \geq \mathrm{VaR}_p(X)$ for $p \in (0,1)$.

2.2 Upper comonotonicity

If two or more assets are comonotonic then their values (whether they be small, medium, large, etc) move in the same direction simultaneously. In the real world, this may be too strong of a relation. A more realistic relation is to say that the assets move in the same direction if their values are extremely large. This weaker relation is known as *upper comonotonicity* (Cheung, 2009).

Let X_i denote the loss of the ith asset. Let $\mathbf{X} = (X_1, \ldots, X_n)$ with joint cdf $F\left(x_1, \ldots, x_n\right)$. Let $T = X_1 + \cdots + X_n$. Suppose all random variables are defined on the probability space $(\Omega, \mathcal{F}, \Pr)$. Then, a simple formula for the value at risk of T in terms of values at risk of X_i can be established if $\mathbf{X}$ is upper comonotonic.

We now define what is meant by upper comonotonicity. A subset $C \subset \mathbb{R}^n$ is said to be comonotonic if $\left(t_i - s_i\right)\left(t_j - s_j\right) \geq 0$ for all i and j whenever $(t_1, \ldots, t_n)$ and $(s_1, \ldots, s_n)$ belong to C. The random vector is said to be comonotonic if it has a comonotonic support.

Let $\mathcal{N}$ denote the collection of all zero probability sets in the probability space. Let $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup (-\infty, \ldots, -\infty)$. For a given $(a_1, \ldots, a_n) \in \mathbb{R}^n$, let $U(\mathbf{a})$ denote the upper quadrant of $(a_1, \infty) \times$

$\cdots \times (a_n, \infty)$ and let $L(\mathbf{a})$ denote the lower quadrant of $(-\infty, a_1] \times \cdots \times (-\infty, a_n]$. Let $R(\mathbf{a}) = \mathbb{R}^n \backslash (U(\mathbf{a}) \cup L(\mathbf{a}))$.

Then, the random vector $\mathbf{X}$ is said to be upper comonotonic if there exist $\mathbf{a} \in \overline{\mathbb{R}^n}$ and a zero probability set $N(\mathbf{a}) \in \mathcal{N}$ such that

(a) $\{\mathbf{X}(w) \mid w \in \Omega \backslash N(\mathbf{a})\} \cap U(\mathbf{a})$ is a comonotonic subset of $\mathbb{R}^n$;

(b) $\Pr(\mathbf{X} \in U(\mathbf{a})) > 0$;

(c) $\{\mathbf{X}(w) \mid w \in \Omega \backslash N(\mathbf{a})\} \cap R(\mathbf{a})$ is an empty set.

If these three conditions are satisfied then the value at risk of T can be expressed as

$$\mathrm{VaR}_p(T) = \sum_{i=1}^{n} \mathrm{VaR}_p(X_i) \tag{3}$$

for $p \in (F(a_1^*, \ldots, a_n^*), 1)$ and $\mathbf{a}^* = (a_1^*, \ldots, a_n^*)$, a comonotonic threshold as constructed in Lemma 2 of Cheung (2009).

2.3 Multivariate extension

Multivariate VaR is a much more recent topic.

Let $\mathbf{X}$ be a random vector in $\mathbb{R}^r$ with joint cdf F. Prékopa (2012) gives the following definition of multivariate VaR:

$$\mathrm{MVaR}_p = \left\{ \mathbf{u} \in \mathbb{R}^r \middle| F(\mathbf{u}) = p \right\}. \tag{4}$$

Note that MVaR may not be a single vector. It will often take the form of a set of vectors.

Prékopa (2012) gives the following motivation for multivariate VaR: "A finance company generally faces the problem of constructing different portfolios that they can sell to customers. Each portfolio produces a random total return and it is the objective of the company to have them above given levels, simultaneously, with large probability. Equivalently, the losses should be below given levels, with large probability. In order to ensure it we look at the total losses as components of a random vector and find a multivariate p-quantile or MVaR to know what are those points in the r-dimensional space (r being the number of portfolios), that should surpass the vector of total losses, to guarantee the given reliability".

Cousin and Bernardinoy (2011) provide another definition of multivariate VaR:

$$\mathrm{MVaR}_p = E[\mathbf{X} \mid \mathbf{X} \in \partial L(p)] = \begin{pmatrix} E[X_1 \mid \mathbf{X} \in \partial L(p)] \\ E[X_2 \mid \mathbf{X} \in \partial L(p)] \\ \vdots \\ E[X_r \mid \mathbf{X} \in \partial L(p)] \end{pmatrix}$$

or equivalently

$$\mathrm{MVaR}_p = E[\mathbf{X} \mid F(\mathbf{X}) = p] = \begin{pmatrix} E[X_1 \mid F(\mathbf{X}) = p] \\ E[X_2 \mid F(\mathbf{X}) = p] \\ \vdots \\ E[X_r \mid F(\mathbf{X}) = p] \end{pmatrix},$$

where $\partial L(p)$ is the boundary of the set $\left\{\mathbf{x} \in \mathbb{R}_+^r : F(\mathbf{x}) \geq p\right\}$.

Cousin and Bernardinoy (2011) establish various properties of MVaR similar to those in the univariate case. For instance,

(i) the translation equivariant property holds, that is

$$\mathrm{MVaR}_p\left(\mathbf{c}+\mathbf{X}\right) = \mathbf{c} + \mathrm{MVaR}_p\left(\mathbf{X}\right) = \begin{pmatrix} c_1 + E\left[X_1 \mid F(\mathbf{X}) = p\right] \\ c_2 + E\left[X_2 \mid F(\mathbf{X}) = p\right] \\ \vdots \\ c_r + E\left[X_r \mid F(\mathbf{X}) = p\right] \end{pmatrix};$$

(ii) the positively homogeneous property holds, that is

$$\mathrm{MVaR}_p\left(\mathbf{c}\mathbf{X}\right) = \mathbf{c}\mathrm{MVaR}_p\left(\mathbf{X}\right) = \begin{pmatrix} c_1 E\left[X_1 \mid F(\mathbf{X}) = p\right] \\ c_2 E\left[X_2 \mid F(\mathbf{X}) = p\right] \\ \vdots \\ c_r E\left[X_r \mid F(\mathbf{X}) = p\right] \end{pmatrix};$$

(iii) if F is quasi-concave (Nelson, 1999) then

$$\mathrm{MVaR}_p^i(\mathbf{X}) \geq \mathrm{VaR}_p\left(X_i\right)$$

for $i = 1, 2, \ldots, r$, where $\mathrm{MVaR}_p^i(\mathbf{X})$ denotes the ith component of $\mathrm{MVaR}_p(\mathbf{X})$;

(iv) if $\mathbf{X}$ is a comonotone non-negative random vector and if F is quasi-concave (Nelson, 1999) then

$$\mathrm{MVaR}_p^i(\mathbf{X}) = \mathrm{VaR}_p\left(X_i\right)$$

for $i = 1, 2, \ldots, r$;

(v) if $X_i = Y_i$ in distribution for every $i = 1, 2, \ldots, s$ then

$$\mathrm{MVaR}_p(\mathbf{X}) = \mathrm{MVaR}_p\left(\mathbf{Y}\right)$$

for all $p \in (0, 1)$;

(vi) if X_i is stochastically less than Y_i for every $i = 1, 2, \ldots, s$ then

$$\mathrm{MVaR}_p(\mathbf{X}) \leq \mathrm{MVaR}_p\left(\mathbf{Y}\right)$$

for all $p \in (0, 1)$.

Bivariate value at risk in the context of a bivariate normal distribution has been considered much earlier by Arbia (2002).

A matric variate extension of VaR and its application for power supply networks are discussed in Chang (2011).

2.4 Risk concentration

Let $X_1, X_2, \ldots, X_n$ denote future losses, assumed to be non-negative independent random variables with common cdf F and survival function $\overline{F}$. Degen *et al.* (2010) define *risk concentration* as

$$C(\alpha) = \frac{\mathrm{VaR}_\alpha \left[\sum_{i=1}^{n} X_i \right]}{\sum_{i=1}^{n} \mathrm{VaR}_\alpha \left(X_i \right)}.$$

If $\overline{F}$ is regularly varying with index $-1/\xi$, $\xi > 0$ (Bingham *et al.*, 1989), meaning that $\overline{F}(tx)/\overline{F}(t) \to x^{-1/\xi}$ as $t \to \infty$, then it is shown that

$$C(\alpha) \to n^{\xi - 1} \tag{5}$$

as $\alpha \to 1$. Degen *et al.* (2010) also study the rate of convergence in (5).

Suppose X_i, $i = 1, 2, \ldots, n$ are regularly varying with index $-\beta$, $\beta > 0$. According to Jang and Jho (2007), for $\beta > 1$,

$$C(\alpha) < 1$$

for all $\alpha \in [\alpha_0, 1]$ for some $\alpha_0 \in (0, 1)$. This property is referred to as subadditivity. If $C(\alpha) < 1$ holds as $\alpha \to 1$ then the property is referred to as asymptotic subadditivity. For $\beta = 1$,

$$C(\alpha) \to 1$$

as $\alpha \to 1$. This property is referred to as asymptotic comonotonicity. For $0 < \beta < 1$,

$$C(\alpha) > 1$$

for all $\alpha \in [\alpha_0, 1]$ for some $\alpha_0 \in (0, 1)$. If $C(\alpha) > 1$ holds as $\alpha \to 1$ then the property is referred to as asymptotic superadditivity.

Let $N(t)$ denote a counting process independent of $\{X_i\}$ with $E\left[N(t)\right] < \infty$ for $t > 0$. According to Jang and Jho (2007), in the case of subadditivity,

$$\mathrm{VaR}_\alpha \left[\sum_{i=1}^{N(t)} X_i \right] \leq E\left[N(t)\right] \sum_{i=1}^{N(t)} \mathrm{VaR}_\alpha \left(X_i \right)$$

for all $\alpha \in [\alpha_0, 1]$ for some $\alpha_0 \in (0, 1)$. In the case of asymptotic comonotonicity,

$$\mathrm{VaR}_\alpha \left[\sum_{i=1}^{N(t)} X_i \right] \sim E\left[N(t)\right] \sum_{i=1}^{N(t)} \mathrm{VaR}_\alpha \left(X_i \right)$$

as $\alpha \to 1$. In the case of superadditivity,

$$\mathrm{VaR}_\alpha \left[\sum_{i=1}^{N(t)} X_i \right] \geq E\left[N(t)\right] \sum_{i=1}^{N(t)} \mathrm{VaR}_\alpha \left(X_i \right)$$

for all $\alpha \in [\alpha_0, 1]$ for some $\alpha_0 \in (0, 1)$.

Suppose $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ is multivariate regularly varying with index β according to Definition 2.2 in Embrechts *et al.* (2009). If $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}$ is a measureable function such that

$$\lim_{x \to \infty} \frac{\Pr\left(\Psi(\mathbf{X}) > x\right)}{\Pr\left(X_1 > x\right)} \to q \in (0, \infty)$$

then it is shown

$$\lim_{\alpha \to 1} \frac{\mathrm{VaR}_\alpha\left(\Psi(\mathbf{X})\right)}{\mathrm{VaR}_\alpha\left(X_1\right)} \to q^{1/\beta},$$

see Lemma 2.3 in Embrechts *et al.* (2009).

2.5 Hürlimann's inequalities

Let X denote a random variable defined over $[A, B]$, $-\infty \leq A < B \leq \infty$ with mean μ, and variance σ. Hürlimann (2002) provides various upper bounds for $\mathrm{VaR}_p(X)$: for $p \leq \sigma^2 / \left\{\sigma^2 + (B - \mu)^2\right\}$,

$$\mathrm{VaR}_p(X) \leq B;$$

for $\sigma^2 / \left\{\sigma^2 + (B - \mu)^2\right\} \leq p \leq (\mu - A)^2 / \left\{\sigma^2 + (\mu - A)^2\right\}$,

$$\mathrm{VaR}_p(X) \leq \mu + \sqrt{\frac{1 - p}{p}}\sigma;$$

for $p \geq (\mu - A)^2 / \left\{\sigma^2 + (\mu - A)^2\right\}$,

$$\mathrm{VaR}_p(X) \leq \mu + \frac{(\mu - A)(B - A)(1 - p) - \sigma^2}{(B - A)p - (\mu - A)}. \tag{6}$$

The equality in (6) holds if and only if $B \to \infty$.

Now suppose X is a random variable defined over $[A, B]$, $-\infty \leq A < B \leq \infty$ with mean μ, variance σ, skewness γ and kurtosis γ_2. In this case, Hürlimann (2002) provides the following upper bound for $\mathrm{VaR}_p(X)$:

$$\mathrm{VaR}_p(X) \leq \mu + x_p \sigma,$$

where x_p is the $100(1 - p)$ percentile of the standardized Chebyshev-Markov maximal distribution. The latter is defined as the root of

$$p\left(x_p\right) = p$$

if $p \leq (1/2)\left\{1 - \gamma / \sqrt{4 + \gamma^2}\right\}$ and as the root of

$$p\left(\psi\left(x_p\right)\right) = 1 - p$$

if $p > (1/2)\left\{1 - \gamma / \sqrt{4 + \gamma^2}\right\}$, where

$$p(u) = \frac{\Delta}{q^2(u) + \Delta\left(1 + u^2\right)},$$

$$\psi(u) = \frac{1}{2}\left[\frac{A(u) - \sqrt{A^2(u) + 4q(u)B(u)}}{q(u)}\right],$$

where $\Delta = \gamma_2 - \gamma^2 + 2$, $A(u) = \gamma q(u) + \Delta u$, $B(u) = q(u) + \Delta$ and $q(u) = 1 + \gamma u - u^2$.

2.6 Ibragimov and Walden's inequalities

Let $R(\mathbf{w}) = \sum_{i=1}^{N} w_i R_i$ denote a portfolio return made up of N asset returns, R_i, and the non-negative weights w_i. Ibragimov (2009) provides various inequalities for the VaR of $R(\mathbf{w})$. They suppose that R_i are independent and identically distributed and belong to either $\underline{CS}$, the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (0, 1]$ and $\sigma > 0$ or $\overline{CSLC}$, convolutions of distributions from the class of symmetric log-concave distributions and the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in [1, 2]$ and $\sigma > 0$.

Here, $S_\alpha(\beta, \gamma, \mu)$ denotes a stable distribution specified by its characteristic function

$$\phi(t) = \begin{cases} \exp\left\{ \mathrm{i}\mu t - \gamma^\alpha |t|^\alpha \left[1 - \mathrm{i}\beta \tan\left(\pi \dfrac{\alpha}{2}\right) \mathrm{sign}(t) \right] \right\}, & \alpha \neq 1, \\ \exp\left\{ \mathrm{i}\mu t - \gamma |t| \left(1 + \mathrm{i}\beta \mathrm{sign}(t) \dfrac{2}{\pi} \ln t \right) \right\}, & \alpha = 1, \end{cases}$$

where $\mathrm{i} = \sqrt{-1}$, $\alpha \in (0, 2]$, $|\beta| \leq 1$, $\gamma > 0$ and $\mu \in \mathbb{R}$. The stable distribution contains as particular cases: the Gaussian distribution for $\alpha = 2$; the Cauchy distribution for $\alpha = 1$, and $\beta = 0$; the Lévy distribution for $\alpha = 1/2$ and $\beta = 1$; the Landau distribution for $\alpha = 1$ and $\beta = 1$; the dirac delta distribution for $\alpha \downarrow 0$ and $\gamma \downarrow 0$.

Furthermore, let $\mathcal{I}_N = \left\{ (w_1, \ldots, w_N) \in \mathbb{R}_+^N : w_1 + \cdots + w_N = 1 \right\}$. Write $\mathbf{a} \prec \mathbf{b}$ to mean that $\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]}$ for $k = 1, \ldots, N-1$ and $\sum_{i=1}^N a_{[i]} = \sum_{i=1}^N b_{[i]}$, where $a_{[1]} \geq \cdots \geq a_{[N]}$ and $b_{[1]} \geq \cdots \geq b_{[N]}$ denote the components of $\mathbf{a}$ and $\mathbf{b}$ in descending order. Let $\underline{\mathbf{w}}_N = (1/N, 1/N, \ldots, 1/N)$ and $\overline{\mathbf{w}}_N = (1, 0, \ldots, 0)$.

With these notation, Ibragimov (2009) provides the following inequalities for $\mathrm{VaR}_q\left(R(\mathbf{w})\right)$. Suppose first that $q \in (0, 1/2)$ and R_i belong to $\overline{CSLC}$. Then,

(i) $\mathrm{VaR}_{1-q}\left[R(\mathbf{v})\right] \leq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right]$ if $\mathbf{v} \prec \mathbf{w}$;

(i) $\mathrm{VaR}_{1-q}\left[R\left(\underline{\mathbf{w}}_N\right)\right] \leq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right] \leq \mathrm{VaR}_{1-q}\left[R\left(\overline{\mathbf{w}}_N\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_N$.

Suppose now that $q \in (0, 1/2)$ and R_i belong to $\underline{CS}$. Then,

(i) $\mathrm{VaR}_{1-q}\left[R(\mathbf{v})\right] \geq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right]$ if $\mathbf{v} \prec \mathbf{w}$;

(i) $\mathrm{VaR}_{1-q}\left[R\left(\overline{\mathbf{w}}_N\right)\right] \leq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right] \leq \mathrm{VaR}_{1-q}\left[R\left(\underline{\mathbf{w}}_N\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_N$.

Further inequalities for VaR are provided in Ibragimov and Walden (2011) when a portfolio return, say R, is made up of a two dimensional array of asset returns say R_{ij}. That is,

$$\begin{aligned} R(\mathbf{w}) &= \sum_{i=1}^{r} \sum_{j=1}^{c} w_{ij} R_{ij} \\ &= \sum_{i=1}^{r} w_{i0} \mathcal{R}_i + \sum_{i=1}^{r} w_{0j} C_j + \sum_{i=1}^{r} \sum_{j=1}^{c} w_{ij} U_{ij} \\ &= \mathcal{R}\left(\mathbf{w}_0^{(\mathrm{row})}\right) + C\left(\mathbf{w}_0^{(\mathrm{col})}\right) + U(\mathbf{w}), \end{aligned}$$

where $\mathcal{R}_i, i = 1, \ldots, r$ are referred to as "row effects", $C_j, j = 1, \ldots, c$ are referred to as "column effects", and $U_{ij}, i = 1, \ldots, r, j = 1, \ldots, c$ are referred to as "idiosyncratic components".

Let $\underline{\mathbf{w}}_{rc} = (1/(rc), 1/(rc), \ldots, 1/(rc))$, $\overline{\mathbf{w}}_{rc} = (1, 0, \ldots, 0)$, $\underline{\mathbf{w}}_0^{(\mathrm{row})} = (1/r, 1/r, \ldots, 1/r)$, $\overline{\mathbf{w}}_0^{(\mathrm{row})} = (1, 0, \ldots, 0)$, $\underline{\mathbf{w}}_0^{(\mathrm{col})} = (1/c, 1/c, \ldots, 1/c)$, and $\overline{\mathbf{w}}_0^{(\mathrm{col})} = (1, 0, \ldots, 0)$.

With these notation, Ibragimov and Walden (2011) provide the following inequalities for $q \in (0, 1/2)$:

(i) if $\mathcal{R}_i, C_j, U_{ij}$ belong to $\overline{CSLC}$ then $\mathrm{VaR}_{1-q}\left[R\left(\underline{\mathbf{w}}_{rc}\right)\right] \leq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right] \leq \mathrm{VaR}_{1-q}\left[R\left(\overline{\mathbf{w}}_{rc}\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(ii) if $\mathcal{R}_i, C_j, U_{ij}$ belong to $\overline{CS}$ then $\mathrm{VaR}_{1-q}\left[R\left(\underline{\mathbf{w}}_{rc}\right)\right] \geq \mathrm{VaR}_{1-q}\left[R(\mathbf{w})\right] \geq \mathrm{VaR}_{1-q}\left[R\left(\overline{\mathbf{w}}_{rc}\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(iii) if U_{ij} belong to $\overline{CSLC}$ then $\mathrm{VaR}_{1-q}\left[U\left(\underline{\mathbf{w}}_{rc}\right)\right] \leq \mathrm{VaR}_{1-q}\left[U(\mathbf{w})\right] \leq \mathrm{VaR}_{1-q}\left[U\left(\overline{\mathbf{w}}_{rc}\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(iv) if U_{ij} belong to $\overline{CS}$ then $\mathrm{VaR}_{1-q}\left[U\left(\underline{\mathbf{w}}_{rc}\right)\right] \geq \mathrm{VaR}_{1-q}\left[U(\mathbf{w})\right] \geq \mathrm{VaR}_{1-q}\left[U\left(\overline{\mathbf{w}}_{rc}\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(v) if $\mathcal{R}_i$ belong to $\overline{CSLC}$ then $\mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\underline{\mathbf{w}}_r\right)\right] \leq \mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\mathbf{w}_0^{(\mathrm{row})}\right)\right] \leq \mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\overline{\mathbf{w}}_r\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(vi) if $\mathcal{R}_i$ belong to $\overline{CS}$ then $\mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\underline{\mathbf{w}}_r\right)\right] \geq \mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\mathbf{w}_0^{(\mathrm{row})}\right)\right] \geq \mathrm{VaR}_{1-q}\left[\mathcal{R}\left(\overline{\mathbf{w}}_r\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(vii) if C_j belong to $\overline{CSLC}$ then $\mathrm{VaR}_{1-q}\left[C\left(\underline{\mathbf{w}}_c\right)\right] \leq \mathrm{VaR}_{1-q}\left[C\left(\mathbf{w}_0^{(\mathrm{col})}\right)\right] \leq \mathrm{VaR}_{1-q}\left[C\left(\overline{\mathbf{w}}_c\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$;

(viii) if C_j belong to $\overline{CS}$ then $\mathrm{VaR}_{1-q}\left[C\left(\underline{\mathbf{w}}_c\right)\right] \geq \mathrm{VaR}_{1-q}\left[C\left(\mathbf{w}_0^{(\mathrm{col})}\right)\right] \geq \mathrm{VaR}_{1-q}\left[C\left(\overline{\mathbf{w}}_c\right)\right]$ for all $\mathbf{w} \in \mathcal{I}_{rc}$.

Ibragimov and Walden (2011, Section 4) discuss an application of these inequalities to portfolio component value at risk analysis.

2.7 Denis *et al.*'s inequalities

Let $\{P_t\}$ denote prices of financial assets. The process could be modeled by

$$P_t = m + \int_0^t \sigma_s dB_s + \int_0^t b_s ds + \sum_{i=1}^{N_t} \gamma_{T_i^-} Y_i,$$

where B is a Brownian motion, $\widetilde{N}$ is a compound Poisson process independent of B, $T_1, T_2, \ldots$ are jump times for $\widetilde{N}$, b is an adapted integrable process, and σ, γ are certain random variables.

Denis *et al.* (2009) derive various bounds for the VaR of the process

$$P_t^* = \sup_{0 \leq u \leq t} P_u.$$

The following assumptions are made:

(i) for all $t > 0$, $E\left(\int_0^t \sigma_s^2 ds\right) < \infty$;

(ii) jumps of the compound Poisson process are non-negative and Y_1 is not identically equal to zero;

(iii) the process $\sum_{i=1}^{N_t} \gamma_{T_i^-} Y_i$ for $t > 0$ is well defined and integrable;

(iv) the jumps have a Laplace transform, $L(x) = E\left[\exp\left(xY_1\right)\right]$, $x < c$ for c a positive constant;

(v) there exists $\gamma^* > 0$ such that $\gamma_s \leq \gamma^*$ almost surely for all $s \in [0, t]$;

(vi) there exists $b^*(t) \geq 0$ and $a^*(t) \geq 0$ such that

$$\int_0^t \sigma_u^2 du \leq a^*(t), \ \int_0^s b_u du \leq b^*(t)$$

almost everywhere for all $s \in [0, t]$. In this case, let

$$K_t(\delta) = \delta b^*(t) + \delta^2 \frac{a^*(t)}{2} + \lambda t \left[L\left(\delta\gamma^*\right) - 1\right]$$

for $0 < \delta < c/\gamma^*$.

With these assumptions, Denis *et al.* (2009) show that

$$\mathrm{VaR}_{1-\alpha}\left(P_t^*\right) \leq \inf_{\delta < c/\gamma^*} \left\{ m + \frac{K_t(\delta) - \ln\alpha}{\delta} \right\},$$

$$\mathrm{VaR}_{1-\alpha}\left(P_t^*\right) \leq \inf_{0 < \delta < c/\gamma^*} \left\{ m + b^*(t) + \frac{a^*(t)\delta}{2} + \lambda t \frac{L\left(\delta\gamma^*\right) - 1}{\delta} - \frac{\ln\alpha}{\delta} \right\}.$$

For $\gamma \leq 0$, Denis *et al.* (2009) show that

$$\mathrm{VaR}_{1-\alpha}\left(P_t^*\right) \leq m + b^*(t) + \sqrt{-2a^*(t)\ln\alpha}.$$

If the jumps follow a simple Poisson process, Denis *et al.* (2009) show that

$$\mathrm{VaR}_{1-\alpha}\left(P_t^*\right) \leq \inf_{0 < \delta < \infty} \left\{ m + b^*(t) + \frac{a^*(t)\delta}{2} + \lambda t \frac{\exp\left(\delta\gamma^*\right) - 1}{\delta} - \frac{\ln\alpha}{\delta} \right\}.$$

If the jumps follow an exponential distribution with parameter $\nu > 0$, Denis *et al.* (2009) show that

$$\mathrm{VaR}_{1-\alpha}\left(P_t^*\right) \leq \inf_{0 < \delta < \nu/\gamma^*} \left\{ m + b^*(t) + \frac{a^*(t)\delta}{2} + \frac{\lambda t}{\nu/\gamma^* - \delta} - \frac{\ln\alpha}{\delta} \right\}.$$

About the issue of continuity / discontinuity of the market with jumps, see Walter (2015).

2.8 Jaworski's inequalities

Jaworski (2007, 2008) considers the following situation: suppose s_i, $i = 1, \ldots, n$ are the quotients of the currency rates at the end and at the beginning of an investment; suppose that the joint cdf of $(s_1, \ldots, s_n)$ is $C\left(F_1\left(s_1\right), \ldots, F_n\left(s_n\right)\right)$, where C is a copula (Nelsen, 1999) and F_i is the marginal cdf of s_i; suppose w_i is the part of the capital invested in the ith currency, where w_i are non-negative and sum to one. Then, the final investment value is

$$W_1(\mathbf{w}) = \left(w_1 s_1 + \cdots + w_n s_n\right) W_0,$$

where $\mathbf{w} = (w_1, \ldots, w_n)$. Jaworski (2007, 2008) defines the value of risk for a given $\mathbf{w}$ and a probability α as

$$\mathrm{VaR}_\alpha(\mathbf{w}) = \sup\left\{V : \Pr\left(W_0 - W_1(\mathbf{w}) \leq V\right) \leq \alpha\right\}.$$

Jaworski (2007) shows this VaR can be bounded as

$$\sum_{i=1}^{n} \mathrm{VaR}_{\alpha'}\left(\mathbf{e}_i\right) \leq \mathrm{VaR}_\alpha \leq \sum_{i=1}^{n} \mathrm{VaR}_\alpha\left(\mathbf{e}_i\right)$$

for portfolios consisting of only one currency, where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ and $\alpha^{'} = \alpha^2 / C(\alpha, \ldots, \alpha)$.

2.9 Mesfioui and Quessy's inequalities

Suppose a portfolio is made up of n assets and let $X_1, X_2, \ldots, X_n$ denote the losses for the n assets. Suppose also that the joint cdf of $(X_1, \ldots, X_n)$ is $C\left(F_1\left(x_1\right), \ldots, F_n\left(x_n\right)\right)$, where C is a copula (Nelsen, 1999), and F_i is the marginal cdf of X_i. Furthermore, define the dual of a given copula C (Definition 2.4, Mesfioui and Quessy, 2005) as

$$C^d\left(u_1, \ldots, u_n\right) = \Pr\left(U(0,1) \leq u_1 \text{ or } \cdots \text{ or } U(0,1) \leq u_n\right).$$

With these notation, Mesfioui and Quessy (2005) derive various inequalities for the value at risk of $S = X_1 + \cdots + X_n$. If C is such that $C \geq qC_L$ and $C \leq C_U^d$ for some copulas C_L and C_U then

$$\underline{\mathrm{VaR}}_\alpha \leq \mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}_\alpha,$$

where

$$\underline{\mathrm{VaR}}_\alpha = \sup_{C_U^d(u_1, \ldots, u_n) = \alpha} \sum_{i=1}^{n} F_i^{-1}\left(u_i\right)$$

and

$$\overline{\mathrm{VaR}}_\alpha = \inf_{C_L(u_1, \ldots, u_n) = \alpha} \sum_{i=1}^{n} F_i^{-1}\left(u_i\right).$$

If $X_1, X_2, \ldots, X_n$ are identical random variables with common cdf F and if $x^* \in \mathbb{R}$ is such that $f(x) = dF(x)/dx$ is non-increasing for $x \geq x^*$ then it is shown under certain conditions that

$$\mathrm{VaR}_\alpha(S) \leq nF^{-1}\left(\delta_{C_L}^{-1}(\alpha)\right),$$

where $\delta_{C_L}(t) = C_L(t, \ldots, t)$ is the diagonal section of C_L.

Mesfioui and Quessy (2005) also show that if X is a random variable with mean μ and variance σ^2 then

$$g_{\mu,\sigma}(\alpha) \leq \mathrm{VaR}_\alpha(X) \leq h_{\mu,\sigma}(\alpha),$$

where

$$g_{a,b}(u) = \{a - bq(1-u)\}\, I\left(u \geq \frac{b^2}{a^2+b^2}\right)$$

and

$$gh_{a,b}(u) = a + aq^2(u) I\left(u \leq \frac{b^2}{a^2+b^2}\right) + bq(u) I\left(u > \frac{b^2}{a^2+b^2}\right),$$

where $q(u) = \sqrt{u/(1-u)}$. If X_i, $i = 1, \ldots, n$ have means μ_i, $i = 1, \ldots, n$ and variances σ_i^2, $i = 1, \ldots, n$ then it is shown that

$$g_{\mu,\sigma}(\alpha) \leq \mathrm{VaR}_\alpha(S) \leq h_{\mu,\sigma}(\alpha),$$

where $\mu = \mu_1 + \cdots + \mu_n$ and $\sigma = \sigma_1 + \cdots + \sigma_n$.

2.10 Slim *et al.*'s inequalities

Suppose a portfolio is made up of d assets. Let $X_1, X_2, \ldots, X_n$ denote the losses for the n assets. Let F_i and f_i denote the cdf and the pdf of X_i. Let x_i^* denote the value for which $f_i(x)$ is non-increasing for all $x \leq x_i^*$. Given this notation, the total portfolio loss can be expressed as $S = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$ for some non-negative weights w_i summing to one. Slim *et al.* (2010) show that the VaR of S can be bounded as follows:

$$\underline{\mathrm{VaR}}_p \leq \mathrm{VaR}_p(S) \leq \overline{\mathrm{VaR}}_p,$$

where

$$\overline{\mathrm{VaR}}_p = \inf_{u_1+\cdots+u_n=\alpha+n-1} \sum_{i=1}^{n} F_i^{-1}\left(u_i\right)$$

and

$$\underline{\mathrm{VaR}}_p = \max_{1 \leq i \leq n} \left\{ F_i^{-1}(\alpha) + \sum_{1 \leq j \neq i \leq n} F_j^{-1}(n) \right\}$$

for $\alpha \leq \min\left\{F_1\left(x_1^*\right), \ldots, F_n\left(x_n^*\right)\right\}$. The use of the above results allows easy computation for explicit VaR bounds for possibly dependent risks.

3 Parametric methods

This section concentrates on estimation of value at risk when data comes from a parametric distribution and we want to make use of the parameters. The parametric methods summarized are based on: Gaussian distribution (Section 3.1), Student's t distribution (Section 3.2), Pareto positive stable distribution (Section 3.3), log folded t distribution (Section 3.4), variance covariance method (Section 3.5), Gaussian mixture distribution (Section 3.6), generalized hyperbolic distribution (Section 3.7), fourier transformation method (Section 3.8), principal components method (Section 3.9), quadratic forms (Section 3.10), elliptical distribution (Section 3.11), copula method (Section 3.12), Gram-Charlier approximation (Section 3.13), delta gamma approximation (Section 3.14), Cornish-Fisher approximation (Section 3.15), Johnson family method (Section 3.16), Tukey method (Section 3.17), asymmetric Laplace distribution (Section 3.18), asymmetric power distribution (Section 3.19), Weibull distribution (Section 3.20), ARCH models (Section 3.21), GARCH models (Section 3.22), GARCH model with heavy tails (Section 3.23), ARMA-GARCH model (Section 3.24), Markov switching ARCH model (Section 3.25), fractionally integrated GARCH model (Section 3.26), RiskMetrics model (Section 3.27), capital asset pricing model (Section 3.28), Dagum distribution (Section 3.29), location-scale distributions (Section 3.30), discrete distributions (Section 3.31), quantile regression method (Section 3.32), Brownian motion method (Section 3.33), Bayesian method (Section 3.34), and Rachev *et al.*'s method (Section 3.35).

3.1 Gaussian distribution

If $X_1, X_2, \ldots, X_n$ are observations from a Gaussian distribution with mean μ and variance σ^2 then VaR can be estimated by

$$\widehat{\mathrm{VaR}}_\alpha = \overline{X} + \Phi^{-1}(\alpha) s, \tag{7}$$

where $\overline{X}$ is the sample mean and s^2 is the sample variance

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2. \tag{8}$$

The estimator in (7) is biased and consistent. If the n in (8) is replaced by $n-1$ then (7) becomes unbiased and consistent.

3.2 Student's t distribution

If $X_1, X_2, \ldots, X_n$ are observations from a Student's t distribution with ν degrees of freedom then VaR can be estimated by (Arneric *et al.*, 2008)

$$\widehat{\mathrm{VaR}}_\alpha = \overline{X} + t_{\nu,\alpha} s \sqrt{\frac{3+\kappa}{3+2\kappa}},$$

where κ is the excess sample kurtosis and $t_{\nu,\alpha}$ is the 100α percentile of a Student's t random variable with ν degrees of freedom.

3.3 Pareto positive stable distribution

Sarabia and Prieto (2009) and Guillen *et al.* (2011) introduce the Pareto positive stable distribution specified by the cdf

$$F(x) = 1 - \exp\left\{-\lambda\left[\ln(x/\sigma)\right]^{\nu}\right\} \tag{9}$$

for $x \geq \sigma$, $\lambda > 0$ and $\nu > 0$. Here, λ and ν are shape parameters and σ is a scale parameter. The Pareto distribution is the particular case of (9) for $\nu = 1$.

The Pareto positive stable distribution has been applied to risk management, see, for example, Guillen *et al.* (2011). If X is a random variable having the cdf (9) then it is easy to see that

$$\mathrm{VaR}_{\alpha} = \sigma \exp\left\{\left[-\frac{1}{\lambda}\ln(1-\alpha)\right]^{1/\nu}\right\}$$

for $0 < \alpha < 1$. So, if $\left(\widehat{\sigma}, \widehat{\lambda}, \widehat{\nu}\right)$ are maximum likelihood estimators of (σ, λ, ν) then

$$\widehat{\mathrm{VaR}}_{\alpha} = \widehat{\sigma} \exp\left\{\left[-\frac{1}{\widehat{\lambda}}\ln(1-\alpha)\right]^{1/\widehat{\nu}}\right\}$$

for $0 < \alpha < 1$.

3.4 Log folded t distribution

Brazauskas and Kleefeld (2011) introduce the log folded t distribution specified by the quantile function

$$F^{-1}(u) = \exp\left\{\sigma Q_{T(\nu)}((u+1)/2)\right\}$$

for $0 < u < 1$, where $\sigma > 0$ is a scale parameter, $\nu > 0$ is a shape parameter, and $Q_{T(\nu)}(\cdot)$ denotes the quantile function of a Student's t random variable with ν degrees of freedom. Brazauskas and Kleefeld (2011) also provide an application of this distribution to risk management.

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the log folded t distribution with order statistics $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$. Brazauskas and Kleefeld (2011) show that the value at risk can be estimated by

$$\widehat{\mathrm{VaR}}_{1-\alpha} = \exp\left\{\widehat{\sigma} Q_{T(\nu)}(1-\alpha/2)\right\},$$

where

$$\widehat{\sigma} = \left[\frac{1}{n}\sum_{i=1}^{n} \ln^2 X_i\right]^{1/2}$$

or

$$\widehat{\sigma} = \frac{1}{c(a,b)\left(n - m_n - m_n^*\right)} \sum_{i=m_n+1}^{n-m_n^*} \ln X_{i:n},$$

where

$$c(a,b) = \frac{1}{1-a-b}\int_a^{1-b} Q_{T(\infty)}((u+1)/2)du,$$

where m_n and m_n^* are integers $0 \leq m_n < n - m_n^* \leq n$ such that $m_n/n \to a$ and $m_n^*/n \to b$ as $n \to \infty$, where a and b are trimming proportions with $0 \leq a + b < 1$.

3.5 Variance covariance method

Suppose the portfolio return, say R, is made up of N asset returns, R_i, $i = 1, 2, \ldots, N$, as

$$R = \sum_{i=1}^{N} w_i R_i,$$

where w_i are non-negative weights summing to one. Suppose also $\mathrm{E}\left(R_i\right) = \mu_i$, $\mathrm{Var}\left(R_i\right) = \sigma_i^2$ and $\mathrm{Cov}\left(R_i, R_j\right) = \sigma_i \sigma_j \rho_{ij}$. The variance covariance method suggests that the value at risk of R can be approximated by

$$\mathrm{VaR}_\alpha(R) = \sum_{i=1}^{N} w_i \mu_i + \Phi^{-1}(\alpha) \sqrt{\sum_{i=1}^{N} w_i \sigma_i^2 + \sum_{i,j=1, i \neq j}^{N} w_i w_j \sigma_i \sigma_j \rho_{ij}}.$$

An estimator can be obtained by replacing the parameters μ_i, σ_i and ρ_{ij} by their maximum likelihood estimators.

3.6 Gaussian mixture distribution

Let $\{P_t\}$ denote the financial asset prices and let $R_t = \ln P_t - \ln P_{t-1}$ denote the log-return corresponding to the original financial series. Zhang and Cheng (2005) consider the model that R_t have a Gaussian mixture distribution specified by the pdf

$$f(r) = \sum_{k=1}^{K} p_k \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ -\frac{\left(r - \mu_k\right)^2}{2\sigma_k^2} \right\}$$

for $K \geq 1$, where the mixing coefficients p_k sum to one. Let VaR_α^k denote the VaR corresponding to the kth component, that is

$$\int_{-\infty}^{\mathrm{VaR}_\alpha^k} \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ -\frac{\left(r - \mu_k\right)^2}{2\sigma_k^2} \right\} dr = \alpha.$$

Let VaR_α denote the VaR corresponding to the mixture model, that is

$$\int_{-\infty}^{\mathrm{VaR}_\alpha} \sum_{k=1}^{K} p_k \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ -\frac{\left(r - \mu_k\right)^2}{2\sigma_k^2} \right\} dr = \alpha.$$

Then, Theorem 1 in Zhang and Cheng (2005) shows that

$$\min_{1 \leq k \leq K} \mathrm{VaR}_\alpha^k \leq \mathrm{VaR}_\alpha \leq \max_{1 \leq k \leq K} \mathrm{VaR}_\alpha^k$$

always holds.

Furthermore, let α^k denote the significance level of VaR corresponding to the kth component, that is

$$\alpha^{(k)} = \int_{-\infty}^{\mathrm{VaR}} \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ -\frac{\left(r - \mu_k\right)^2}{2\sigma_k^2} \right\} dr.$$

Let α denote the significance level of VaR corresponding to the mixture model, that is

$$\alpha = \int_{-\infty}^{\mathrm{VaR}} \sum_{k=1}^{K} p_k \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{ -\frac{(r-\mu_k)^2}{2\sigma_k^2} \right\} dr = \alpha.$$

Then, Theorem 2 in Zhang and Cheng (2005) shows that

$$\min_{1 \le k \le K} \alpha^{(k)} \le \alpha = \sum_{k=1}^{K} p_k \alpha^{(k)} \le \max_{1 \le k \le K} \alpha^{(k)}$$

always holds.

3.7 Generalized hyperbolic distribution

Suppose the log-returns, $R_t = \ln X_t - \ln X_{t-1}$, follow the model

$$R_t = \sigma_t \epsilon_t,$$

where σ_t is the volatility process and ϵ_t are independent and identical random variables with zero mean and unit variance. Let $\mathrm{VaR}_{\alpha,t}$ denote the corresponding value at risk. Suppose ϵ_t are independent and identical and have the generalized hyperbolic distribution specified by the pdf

$$f(x) = \frac{(\eta/\delta)^\lambda}{\sqrt{2\pi} K_\lambda(\delta\eta)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left\{\sqrt{\delta^2+(x-\mu)^2}/\alpha\right\}^{1/2-\lambda}} \exp\left[\beta(x-\mu)\right],$$

where $\mu \in \mathbb{R}$ is a location parameter, $\alpha \in \mathbb{R}$ is a shape parameter, $\beta \in \mathbb{R}$ is an asymmetry parameter, $\delta \in \mathbb{R}$ is a scale parameter, $\lambda \in \mathbb{R}$, $\eta = \sqrt{\alpha^2 - \beta^2}$, and $K_\nu(\cdot)$ denotes the modified Bessel function of order ν.

Tian and Chan (2010) propose a method based on saddlepoint approximation for computing $\mathrm{VaR}_{\alpha,t}$. It can be described as follows:

1. Estimate σ_t^2 by

$$\widehat{\sigma}_t^2 = \left(\sum_{j=1}^{m} \omega_j R_{t-j} \right)^2$$

   for $m > 1$, where ω_j are some non-negative weights summing to one;

2. Compute $\widehat{t}$ as the root of $\kappa^{'}\left(\widehat{t}\right) = t$, where $\kappa^{'}(\cdot)$ is defined in step 3;

3. Compute $\widehat{q}_p$ as the root of

$$p = \begin{cases} \exp\left\{ \kappa\left(\widehat{t}\right) - \widehat{t}t + \frac{1}{2}\widehat{t}^2\kappa^{''}\left(\widehat{t}\right) \right\} \Phi\left(-\sqrt{\widehat{t}^2\kappa^{''}\left(\widehat{t}\right)} \right), & \text{if } t > E, \\ \frac{1}{2}, & \text{if } t = E, \\ 1 - \exp\left\{ \kappa\left(\widehat{t}\right) - \widehat{t}t + \frac{1}{2}\widehat{t}^2\kappa^{''}\left(\widehat{t}\right) \right\} \Phi\left(-\sqrt{\widehat{t}^2\kappa^{''}\left(\widehat{t}\right)} \right), & \text{if } t < E, \end{cases}$$

where

$$
\begin{aligned}
E &= \mu + \frac{\delta\beta K_{\lambda+1}(\delta\eta)}{\eta K_\lambda(\delta\eta)}, \\
\kappa(z) &= \mu z + \ln\eta^\lambda - \lambda\ln\eta + \ln K_\lambda(\delta\eta) - \ln K_\lambda(\delta\eta), \\
\kappa^{'}(z) &= \mu + \frac{\delta(\beta+z)K_{\lambda+1}(\delta\eta)}{\eta K_\lambda(\delta\eta)}, \\
\kappa^{''}(z) &= \frac{\delta K_{\lambda+1}(\delta\eta)}{\eta K_\lambda(\delta\eta)} + \frac{\delta^2(\beta+z)^2 K_{\lambda+2}(\delta\eta)}{\eta^2 K_\lambda(\delta\eta)} - \frac{\delta^2(\beta+z)^2 K^2_{\lambda+1}(\delta\eta)}{\eta^2 K^2_\lambda(\delta\eta)};
\end{aligned}
$$

4. Estimate $\mathrm{VaR}_{\alpha,t}$ by $\widehat{\mathrm{VaR}}_{\alpha,t} = \widehat{\sigma}_t \widehat{q}_p$.

3.8 Fourier transformation method

Siven *et al.* (2009) suggest a method for computing VaR by approximating the cdf F by a Fourier series. The approximation is given by the following result due to Hughett (1998): suppose

(a) that there exists constants A and $\alpha > 1$ such that $F(-y) \leq A|y|^{-\alpha}$ and $1 - F(y) \leq A|y|^{-\alpha}$ for all $y > 0$,

(b) that there exist constants B and $\beta > 0$ such that $|\phi(u)| \leq B\,|u/(2\pi)|^{-\beta}$ for all $u \in \mathbb{R}$, where $\phi(\cdot)$ denotes the characteristic function corresponding to $F(\cdot)$;

Then, for constants $0 < l < 2/3$, $T > 0$ and $N > 0$, the cdf F can be approximated as

$$
F(x) \approx \frac{1}{2} + 2\sum_{k=1}^{N/2-1} \mathrm{Re}\left(G[k]\exp\left(2\pi \mathrm{i} kx/T\right)\right),
$$

where $\mathrm{i} = \sqrt{-1}$, $\mathrm{Re}(\cdot)$ denotes the real part, and

$$
G(k) = \frac{1-\cos(2\pi lk)}{2\pi \mathrm{i} k}\phi\left(-2\pi k/T\right).
$$

An estimator for VaR_p is obtained by solving the equation

$$
\frac{1}{2} + 2\sum_{k=1}^{N/2-1} \mathrm{Re}\left(G[k]\exp\left(2\pi \mathrm{i} kx/T\right)\right) = p
$$

for x.

3.9 Principal components method

Brummelhuis *et al.* (2002) use an approximation based on the principal component method to compute VaR. If $\mathbf{S}(t) = (S_1(t), \ldots, S_n(t))$ is a vector of risk factors over time t and if $\Pi(t, \mathbf{S}(t))$ is a random variable they define VaR to be

$$
\Pr\left[\Pi\left(0, \mathbf{S}(0)\right) - \Pi\left(t, \mathbf{S}(t)\right) > \mathrm{VaR}\right] = \alpha. \tag{10}
$$

This equation is too general to be solved. So, Brummelhuis *et al.* (2002) consider the quadratic approximation

$$\Pi\left(t, \mathbf{S}(t)\right) - \Pi\left(0, \mathbf{S}(0)\right) \approx \Theta t + \boldsymbol{\Delta\xi} + \frac{1}{2}\boldsymbol{\xi\Gamma\xi}^T$$

and assume that $\boldsymbol{\xi}$ is normally distributed with mean $\mathbf{m}$ and covariance matrix $\mathbf{V}$. Under this approximation, we can rewrite (10) as

$$\Pr\left[\Theta + \boldsymbol{\Delta\xi} + \frac{1}{2}\boldsymbol{\xi\Lambda\xi}^T \leq -\mathrm{VaR}\right] = \alpha.$$

Let $\mathbf{V} = \mathbf{H}^T\mathbf{H}$ denote the Cholesky decomposition and let

$$\begin{aligned}
\widetilde{\Theta} &= \Theta + \mathbf{m}\boldsymbol{\Delta} + \frac{1}{2}\mathbf{m}\boldsymbol{\Gamma}\mathbf{m}^T,\\
\widetilde{\boldsymbol{\Delta}} &= \left(\boldsymbol{\Delta} + \mathbf{m}\boldsymbol{\Gamma}\right)\mathbf{H}^T,\\
\widetilde{\boldsymbol{\Gamma}} &= \mathbf{H}\boldsymbol{\Gamma}\mathbf{H}^T.
\end{aligned}$$

Also let $\widetilde{\boldsymbol{\Gamma}} = \mathbf{P}\widetilde{\mathbf{D}}\mathbf{P}^T$ denote the principal components decomposition of $\widetilde{\boldsymbol{\Gamma}}$, $\mathbf{v} = \widetilde{\boldsymbol{\Delta}}\mathbf{P}\widetilde{\mathbf{D}}^{-1}$, and $T = \widetilde{\Theta} - \frac{1}{2}\mathbf{v}\mathbf{D}\mathbf{v}^T$. With these notation, Brummelhuis *et al.* (2002) show that VaR can be approximated by

$$\mathrm{VaR} = K - T,$$

where K is the root of

$$\frac{1}{(2\pi)^{n/2}} \int_{\frac{1}{2}\mathbf{z}\widetilde{\mathbf{D}}\mathbf{z}^T \leq -\mathrm{VaR} - T} \exp\left\{-\frac{1}{2}|z - v|^2\right\} dz = \alpha.$$

3.10 Quadratic forms

Suppose the financial series are realizations of a quadratic form

$$V = \theta + \boldsymbol{\delta}^T\mathbf{Y} + \frac{1}{2}\mathbf{Y}^T\boldsymbol{\Lambda}\mathbf{Y} = \theta + \sum_{j=1}^{m}\left(\delta_j Y_j + \frac{1}{2}\lambda_j Y_j^2\right),$$

where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)^T$ is a standard normal vector, $\delta = (\delta_1, \delta_2, \ldots, \delta_m)^T$ and $\boldsymbol{\Lambda}$ = diag $(\lambda_1, \lambda_2, \ldots, \lambda_m)$. Examples include non-linear positions like options in finance or the modelling of bond prices in terms of interest rates (duration and convexity). Here, λ's are the eigenvalues sorted in ascending order. Suppose there are $n \leq m$ distinct eigenvalues. Let i_j denote the highest index of the jth distinct eigenvalue with multiplicity μ_j. For $j = 1, 2, \ldots, n$, let

$$\begin{aligned}
V_j &= \begin{cases} \dfrac{1}{2}\lambda_{i_j} \displaystyle\sum_{\ell=i_{j-1}+1}^{i_j} \left(\dfrac{\delta_\ell}{\lambda_{i_j}} + Y_\ell\right)^2, & \text{if } \lambda_{i_j} \neq 0,\\ \lambda_{i_j} \displaystyle\sum_{\ell=i_{j-1}+1}^{i_j} \delta_\ell Y_\ell, & \text{if } \lambda_{i_j} = 0, \end{cases}\\
\overline{\delta}_j^2 &= \sum_{\ell=i_{j-1}+1}^{i_j} \delta_\ell^2,\\
a_j^2 &= \overline{\delta}_j^2 / \lambda_{i_j}^2.
\end{aligned}$$

Let b_j denote the moment generating function of $V - V_j$ evaluated at $1/\lambda_{i_j}$. With this notation, Jaschke *et al.* (2004) derive various approximations for VaR. The first of these applicable for $\lambda_{i_1} < 0$ is

$$\mathrm{VaR}_\alpha \approx \lambda_{i_1} \ln b_1 + \frac{\lambda_{i_1}}{2} \chi^2_{\mu_1, 1-\alpha} \left(a_1^2 \right),$$

where $\chi^2_{\mu,\alpha}(\delta)$ denotes the 100α percentile of a non-central chisquare random variable with degrees of freedom μ and non-centrality parameter δ. The second of the approximations applicable for $\lambda_{i_1} = 0$ and $\lambda_{i_n} = 0$ is

$$\mathrm{VaR}_\alpha \approx \theta - \sum_{j=2}^{n} \frac{\overline{\delta}_j^2}{2\lambda_{i_j}} + \left(\widetilde{F}_1^t \right)^{-1} (\alpha),$$

where

$$\widetilde{F}_1^t(x) = \left[\frac{|\overline{\delta}_1|}{\sqrt{2\pi}} \exp\left(-\sum_{j=2}^{n} a_j^2/2 \right) \prod_{j=2}^{n} \left| \overline{\delta}_1^2 / \lambda_{i_j} \right|^{\mu_j/2} \right] \frac{\exp\left[-x^2 / \left(2\overline{\delta}_1^2 \right) \right]}{(-x)^{1+\sum_{j=2}^{n} \mu_j/2}}.$$

The third of the approximations applicable for $\lambda_{i_1} > 0$ and $\lambda_{i_n} < 0$ is

$$\mathrm{VaR}_\alpha \approx \theta - \sum_{j=1}^{n} \frac{\overline{\delta}_j^2}{2\lambda_{i_j}} + \left(\frac{m\alpha}{2d} \right)^{2/m},$$

where

$$d = \frac{1}{\Gamma(m/2)} \prod_{j=1}^{n} \left| \lambda_{i_j} \right|^{-\mu_j/2} \exp\left(-\sum_{j=1}^{n} a_j^2/2 \right).$$

3.11 Elliptical distribution

Suppose a portfolio return, say R, is made up of n asset returns, say R_i, $i = 1, 2, \ldots, n$, as $R = \delta_1 R_1 + \cdots + \delta_n R_n = \boldsymbol{\delta}^T \mathbf{R}$, where δ_i are non-negative weights summing to one, $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^T$ and $\mathbf{R} = (R_1, \ldots, R_n)^T$. Kamdem (2005) derives various expressions for the value at risk of R by supposing that $\mathbf{R}$ has an elliptically symmetric distribution.

If $\mathbf{R}$ has the joint pdf $f_{\mathbf{R}}(\mathbf{r}) = \mid \boldsymbol{\Sigma} \mid^{-1/2} g\left((\mathbf{r} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{r} - \boldsymbol{\mu}) \right)$, where $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{\Sigma}$ is the variance-covariance matrix, and $g(\cdot)$ is a continuous and integrable function over $\mathbb{R}$, then it is shown that

$$\mathrm{VaR}_\alpha(R) = \boldsymbol{\delta}^T \boldsymbol{\mu} + q \sqrt{\boldsymbol{\delta}^T \boldsymbol{\Sigma} \boldsymbol{\delta}},$$

where q is the root of

$$G(s) = \alpha,$$

where

$$G(s) = \frac{\pi^{(n-1)/2}}{\Gamma\left((n-1)/2\right)} \int_s^{-\infty} \int_{z_1^2}^{\infty} \left(u - z_1^2 \right)^{(n-3)/2} g(u) du dz_1. \tag{11}$$

If $\mathbf{R}$ follows a mixture of elliptical pdfs given by

$$f_{\mathbf{R}}(\mathbf{r}) = \sum_{i=1}^{m} \beta_j \left|\boldsymbol{\Sigma}_j\right|^{-1/2} g_j \left(\left(\mathbf{r} - \boldsymbol{\mu}_j\right)^T \boldsymbol{\Sigma}_j^{-1} \left(\mathbf{r} - \boldsymbol{\mu}_j\right) \right),$$

where $\boldsymbol{\mu}_j$ is the mean vector for the jth elliptical pdf, $\boldsymbol{\Sigma}_j$ is the variance-covariance matrix for the jth elliptical pdf, and β_j are non-negative weights summing to one, then it is shown that the value at risk of $\mathbf{R}$ is the root of

$$\sum_{j=1}^{m} \beta_j G_j \left(\frac{\boldsymbol{\delta}^T \boldsymbol{\mu}_j + \mathrm{VaR}_\alpha}{\sqrt{\boldsymbol{\delta}^T \boldsymbol{\Sigma}_j \boldsymbol{\delta}}} \right) = \alpha,$$

where $G_j(\cdot)$ is defined as in (11).

3.12 Copula method

Suppose a portfolio return, say R, is made up of two asset returns, R_1 and R_2, as $R = wR_1 + (1-w)R_2$, where w is the portfolio weight for asset 1 and $1-w$ is the portfolio weight for asset 2. Huang *et al.* (2009) consider computation of VaR for this situation by supposing that the joint cdf of (R_1, R_2) is $C\left(F_1\left(R_1\right), F_2\left(R_2\right)\right)$, where C is a copula (Nelsen, 1999), F_i is the marginal cdf of R_i and f_i is the marginal pdf of R_i. Then, the cdf of R is

$$\begin{aligned}
\Pr\left(R \leq r\right) &= \Pr\left(wR_1 + (1-w)R_2 \leq r\right) \\
&= \Pr\left(R_1 \leq \frac{r}{w} - \frac{(1-w)R_2}{w}\right) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{r/w-(1-w)R_2/w} c\left(F_1\left(r_1\right), F_2\left(r_2\right)\right) f_1\left(r_1\right) f_2\left(r_2\right) dr_1 dr_2,
\end{aligned}$$

where c is the copula pdf. So, $\mathrm{VaR}_p(R)$ can be computed by solving the equation

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\mathrm{VaR}_p(R)/w-(1-w)R_2/w} C\left(F_1\left(r_1\right), F_2\left(r_2\right)\right) dr_1 dr_2 = p.$$

In general, this equation will have to solved numerically or by simulation.

Franke *et al.* (2011) consider the more general case that the portfolio return R is made up of n asset returns, R_i, $i = 1, 2, \ldots, n$; that is

$$R = \sum_{i=1}^{n} w_i R_i$$

for some non-negative weights summing to one. Suppose as above that the joint cdf of $(R_1, \ldots, R_n)$ is $C\left(F_1\left(R_1\right), \ldots, F_n\left(R_n\right)\right)$, where F_i is the marginal cdf of R_i and f_i is the marginal pdf of R_i. Then, the cdf of R is

$$\Pr\left(R \leq r\right) = \int_{\mathcal{U}} c\left(u_1, \ldots, u_n\right) du_1 \cdots du_n,$$

where

$$\mathcal{U} = \left\{ [0,1]^{n-1} \times [0, u_n(r)] \right\},$$

and

$$u_n(r) = F_n\left(r/w_n - \sum_{i=1}^{n-1} w_i F_i^{-1}\left(u_i\right)/w_n\right).$$

So, $\mathrm{VaR}_p(R)$ can be computed by solving the equation

$$\int_{\mathcal{U}} c\left(u_1, \ldots, u_n\right) du_1 \cdots du_n = p.$$

Again, this equation will have to be computed by numerical integration or simulation.

3.13 Gram-Charlier approximation

Simonato (2011) suggests a number of approximations for computing (2). The first of these is based on Gram Charlier expansion.

Let $\kappa_3 = E\left[Z_t^3\right]$ denote the skewness coefficient and $\kappa_4 = E\left[Z_t^4\right]$ the kurtosis coefficient of the standardized log-returns. Simonato (2011) suggests the approximation

$$\mathrm{VaR}_\alpha = \alpha_h + \sigma_h \psi_{GC}^{-1}(p),$$

where $\psi_{GC}^{-1}(\cdot)$ is the inverse function of

$$\psi_{GC}(k) = \Phi(k) - \frac{\kappa_3}{6}\left(k^2-1\right)\phi(k) - \frac{\kappa_4-3}{24} k\left(k^2-3\right)\phi(k),$$

where $\Phi(\cdot)$ denotes the standard normal cdf and $\phi(\cdot)$ denotes the standard normal pdf.

3.14 Delta gamma approximation

Let $\mathbf{R} = \left(R_1, \ldots, R_n\right)^T$ denote a vector of returns normally distributed with zero means and covariate matrix $\boldsymbol{\Sigma}$. Suppose the return of an associated portfolio takes the general form $Y = g(\mathbf{R})$. It will be difficult to find the value of risk of Y for general $g(\cdot)$. Some approximations are desirable. The delta gamma approximation is a commonly used approximation (Feuerverger and Wong, 2000).

Suppose we can approximate $\mathbf{Y} = \mathbf{a}_1^T\mathbf{R} + \mathbf{R}^T\mathbf{B}_1\mathbf{R}$ for $\mathbf{a}_1$ a $n \times 1$ vector and $\mathbf{B}_1$ a $n \times n$ matrix. Let $\boldsymbol{\Sigma} = \mathbf{H}\mathbf{H}^T$ denote the Cholesky decomposition. Let $\lambda_1, \ldots, \lambda_n$ and $\mathbf{P}_1, \ldots, \mathbf{P}_n$ denote the eigenvalues and eigenvectors of $\mathbf{H}^T\mathbf{B}_1\mathbf{H}$. Let $\mathbf{a}_j$ denote the entries of $\mathbf{P}^T\mathbf{H}^T\mathbf{a}_1$, where $\mathbf{P} = \left(\mathbf{P}_1, \ldots, \mathbf{P}_n\right)$. Then, the delta gamma approximation is that

$$Y \stackrel{d}{=} \sum_{j=1}^{n}\left(a_j Z_j + \lambda_j Z_j^2\right), \tag{12}$$

where $Z_1, \ldots, Z_n$ are independent standard normal random variables. The value of risk can be obtained by inverting the distribution of the right hand side of (12).

3.15 Cornish-Fisher approximation

Another approximation suggested by Simonato (2011) is based on Cornish-Fisher expansion. With the notation as in Section 3.13, the approximation is

$$\mathrm{VaR}_\alpha = \alpha_h + \sigma_h \psi_{CF}^{-1}(p),$$

where $\psi_{CF}^{-1}(\cdot)$ is the inverse function of

$$\begin{aligned} \psi_{CF}^{-1}(p) &= \Phi^{-1}(p) + \frac{\kappa_3}{6}\left[\left(\Phi^{-1}(p)\right)^2 - 1\right] + \frac{\kappa_4 - 3}{24}\left[\left(\Phi^{-1}(p)\right)^3 - 3\Phi^{-1}(p)\right] \\ & \quad - \frac{\kappa_3^2}{36}\left[2\left(\Phi^{-1}(p)\right)^3 - 5\Phi^{-1}(p)\right], \end{aligned}$$

where $\Phi^{-1}(\cdot)$ denotes the standard normal quantile function.

3.16 Johnson family method

A third approximation suggested by Simonato (2011) is based on the Johnson family of distributions due to Johnson (1949).

Let Y denote a standard normal random variable. A Johnson random variable can be expressed as

$$Z = c + d g^{-1}\left(\frac{Y - a}{b}\right),$$

where

$$g^{-1}(u) = \begin{cases} \exp(u), & \text{for the lognormal family,} \\ \left[\exp(u) - \exp(-u)\right]/2, & \text{for the unbounded family,} \\ 1/\left[1 + \exp(-u)\right], & \text{for the bounded family,} \\ u, & \text{for the normal family.} \end{cases}$$

Here, a, b, c and d are unknown parameters determined, for example, by the method of moments, see Hill *et al.* (1976).

With the notation as above, the approximation is

$$\mathrm{VaR}_\alpha = \alpha_h + \sigma_h \psi_J^{-1}(p; a, b, c, d),$$

where

$$\psi_J^{-1}(p; a, b, c, d) = c + d g^{-1}\left(\frac{\Phi^{-1}(p) - a}{b}\right),$$

where $\Phi^{-1}(\cdot)$ denotes the standard normal quantile function.

3.17 Tukey method

Jiménez and Arunachalam (2011) present a method for approximating VaR based on Tukey's g and h family of distributions.

Let Y denote a standard normal random variable. A Tukey's g and h random variable can be expressed as

$$Z = g^{-1}\left[\exp(gY) - 1\right]\exp\left(hY^2/2\right)$$

for $g \neq 0$ and $h \in \mathbb{R}$. The family of lognormal distributions is contained as the particular case for $h = 0$. The family of Tukey's h distribution is contained as the limiting case for $g \to 0$.

With the notation as in Section 3.13, the approximation suggested by Jiménez and Arunachalam (2011) is

$$\mathrm{VaR}_p = A + BT_{g,h}\left(\Phi^{-1}(p)\right),$$

where A and B are location and scale parameters. For $g = 0$ and $h = 1$, Z is a normal random variable with mean μ and standard deviation σ, so $A = \mu$ and $B = \sigma$. For $g = 0.773$ and $h = -0.09445$, Z is an exponential random variable with parameter λ, so $A = (1/\lambda)\ln 2$ and $B = g/\lambda$. For $g = 0$ and $h = 0.057624$, Z is a Student's t random variable with ten degrees of freedom, so $A = 0$ and $B = 1$.

3.18 Asymmetric Laplace distribution

Trindade and Zhu (2007) consider the case that the log-returns of $X_1, X_2, \ldots, X_n$ is a random sample from the asymmetric Laplace distribution given by the pdf

$$f(x) = \frac{\kappa\sqrt{2}}{\tau\left(1+\kappa^2\right)} \left\{ \begin{array}{ll} \exp\left(-\dfrac{\kappa\sqrt{2}}{\tau}\left|x-\theta\right|\right), & \text{if } x \geq \theta, \\ \exp\left(-\dfrac{\sqrt{2}}{\kappa\tau}\left|x-\theta\right|\right), & \text{if } x < \theta \end{array} \right.$$

for $x \in \mathbb{R}$, $\tau > 0$ and $\kappa > 0$. The maximum likelihood estimator of VaR_α is derived as

$$\widehat{\mathrm{VaR}}_\alpha = -\frac{\widehat{\tau}\ln\left[\left(1+\widehat{\kappa}^2\right)(1-\alpha)\right]}{\widehat{\kappa}\sqrt{2}},$$

where $(\widehat{\tau}, \widehat{\kappa})$ are the maximum likelihood estimators of (τ, κ). Trindade and Zhu (2007) show further that

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_\alpha - \mathrm{VaR}_\alpha\right) \to N\left(0, \sigma^2\right)$$

in distribution as $n \to \infty$, where $\sigma^2 = \tau^2\left[(\omega-1)^2\kappa^2 + 2\omega^2\right] / \left(4\kappa^2\right)$ and $\omega = \ln\left[\left(1+\kappa^2\right)(1-\alpha)\right]$.

3.19 Asymmetric power distribution

Komunjer (2007) introduces the asymmetric power distribution as a model for risk management. A random variable, say X, is said to have this distribution if its pdf is

$$f(x) = \left\{ \begin{array}{ll} \dfrac{\delta^{1/\lambda}}{\Gamma(1+1/\lambda)}\exp\left[-\dfrac{\delta}{\alpha^\lambda}|x|^\lambda\right], & \text{if } x \leq 0, \\ \dfrac{\delta^{1/\lambda}}{\Gamma(1+1/\lambda)}\exp\left[-\dfrac{\delta}{(1-\alpha)^\lambda}|x|^\lambda\right], & \text{if } x > 0 \end{array} \right. \tag{13}$$

for $x \in \mathbb{R}$, where $0 < \alpha < 1$, $\lambda > 0$ and $\delta = 2\alpha^\lambda (1-\alpha)^\lambda / \left\{\alpha^\lambda + (1-\alpha)^\lambda\right\}$. Note that λ is a shape parameter and α is a scale parameter. The cdf corresponding to (13) is shown to be (Lemma 1, Komunjer, 2007)

$$F(x) = \begin{cases} \alpha \left[1 - \mathcal{I}\left(\dfrac{\delta}{\alpha^\lambda}\sqrt{\lambda}|x|^\lambda, 1/\lambda\right)\right], & \text{if } x \leq 0, \\ 1 - (1-\alpha)\left[1 - \mathcal{I}\left(\dfrac{\delta}{(1-\alpha)^\lambda}\sqrt{\lambda}|x|^\lambda, 1/\lambda\right)\right], & \text{if } x > 0, \end{cases} \tag{14}$$

where $\mathcal{I}(x, \gamma) = \int_0^{x\sqrt{\gamma}} t^{\gamma-1} \exp(-t) dt / \Gamma(\gamma)$. Inverting (14) as in Lemma 2 of Komunjer (2007), we can express $\mathrm{VaR}_p(X)$ as

$$\mathrm{VaR}_p(X) = \begin{cases} -\left[\dfrac{\alpha^\lambda}{\delta\sqrt{\lambda}}\right]^{1/\lambda} \left[\mathcal{I}^{-1}\left(1 - \dfrac{p}{\alpha}, \dfrac{1}{\lambda}\right)\right]^{1/\lambda}, & \text{if } p \leq \alpha, \\ -\left[\dfrac{(1-\alpha)^\lambda}{\delta\sqrt{\lambda}}\right]^{1/\lambda} \left[\mathcal{I}^{-1}\left(1 - \dfrac{1-p}{1-\alpha}, \dfrac{1}{\lambda}\right)\right]^{1/\lambda}, & \text{if } p > \alpha, \end{cases} \tag{15}$$

where $\mathcal{I}^{-1}(\cdot, \cdot)$ denotes the inverse function of $\mathcal{I}(\cdot, \cdot)$. An estimator of $\mathrm{VaR}_p(X)$ can be obtained by replacing the parameters in (15) by their maximum likelihood estimators, see Proposition 2 in Komunjer (2007).

3.20 Weibull distribution

Gebizlioglu *et al.* (2011) consider estimation of VaR based on the Weibull distribution. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a Weibull distribution with the cdf specified by $F(x) = 1 - \exp\left\{-(x/\theta)^\beta\right\}$ for $x > 0$, $\theta > 0$ and $\beta > 0$. Then, the estimator for VaR is

$$\widehat{\mathrm{VaR}}_\alpha = \{-\ln(1-\alpha)\}^{1/\widehat{\beta}}\, \widehat{\theta}.$$

Gebizlioglu *et al.* (2011) consider various methods for obtaining the estimators $\widehat{\theta}$ and $\widehat{\beta}$. By the method of maximum likelihood, $\widehat{\theta}$ and $\widehat{\beta}$ are the simultaneous solutions of

$$\frac{\overline{x}^2}{s^2} = \frac{\{\Gamma(1+1/\beta)\}^2}{\Gamma(1+2/\beta) - \Gamma^2(1+1/\beta)}$$

and

$$\widehat{\theta} = \frac{\overline{x}}{\Gamma\left(1 + 1/\widehat{\beta}\right)},$$

where $\overline{x}$ is the sample mean and s^2 is the sample variance. By Cohen and Whitten (1982)'s modified method of maximum likelihood, $\widehat{\theta}$ and $\widehat{\alpha}$ are the simultaneous solutions of

$$-\frac{nX_{(1)}^\beta}{\ln[n/(n+1)]} = \sum_{i=1}^n X_i^\beta$$

and

$$\widehat{\theta} = \left(\frac{1}{n}\sum_{i=1}^n X_i^{\widehat{\beta}}\right)^{1/\widehat{\beta}},$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics in ascending order. By Tiku (1967, 1968) and Tiku and Akkaya (2004)'s modified method of maximum likelihood,

$$\widehat{\theta} = \exp\left(\widehat{\delta}\right), \ \widehat{\beta} = 1/\widehat{\eta},$$

where

$$\begin{aligned}
&\widehat{\delta} = K + D\widehat{\eta}, \ \widehat{\eta} = \left\{B + \sqrt{B^2 + 4nC}\right\}/(2n), \\
&K = \sum_{i=1}^{n} \beta_i X_{(i)}/m, \ D = \sum_{i=1}^{n} (\alpha_i - 1)/m, \\
&B = \sum_{i=1}^{n} (\alpha_i - 1)\left(X_{(i)} - K\right), \ C = \sum_{i=1}^{n} \beta_i \left(X_{(i)} - K\right)^2, \\
&m = \sum_{i=1}^{n} \beta_i, \ \alpha_i = \left[1 - t_{(i)}\right] \exp\left(t_{(i)}\right), \\
&\beta_i = \exp\left(t_{(i)}\right), \ t_{(i)} = \ln\left(-\ln\left(1 - i/(n+1)\right)\right).
\end{aligned}$$

By the least squares method, $\widehat{\theta}$ and $\widehat{\alpha}$ are those minimizing

$$\sum_{i=1}^{n} \left(1 - \exp\left\{-\left[X_{(i)}/\theta\right]^{\beta}\right\} - \frac{i}{n+1}\right)^2$$

with respect to θ and α. By the weighted least squares method, $\widehat{\theta}$ and $\widehat{\alpha}$ are those minimizing

$$\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left(1 - \exp\left\{-\left[X_{(i)}/\theta\right]^{\beta}\right\} - \frac{i}{n+1}\right)^2$$

with respect to θ and α. By the percentile method, $\widehat{\theta}$ and $\widehat{\alpha}$ are those minimizing

$$\sum_{i=1}^{n} \left\{X_{(i)} - \theta \left[-\ln\left(1 - \frac{i}{n+1}\right)\right]^{1/\beta}\right\}^2$$

with respect to θ and α.

3.21 ARCH models

ARCH models are popular in finance. Suppose the log-returns, say R_t, of $\{X_1, X_2, \ldots, X_n\}$ follow the ARCH model specified by

$$R_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \beta_0 + \sum_{j=1}^{k} \beta_j R_{t-j}^2,$$

where ϵ_i are independent and identical random variables with zero mean, unit variance, pdf $f(\cdot)$ and cdf $F(\cdot)$, and $\boldsymbol{\beta} = (\beta_0, \beta_1 \ldots, \beta_k)^T$ is an unknown parameter vector satisfying $\beta_0 > 0$ and $\beta_j \geq 0$, $j = 1, 2, \ldots, k$. If $\widehat{\boldsymbol{\beta}} = \left(\widehat{\beta}_0, \widehat{\beta}_1 \ldots, \widehat{\beta}_k\right)^T$ are the maximum likelihood estimators then the residuals are

$$\widehat{\epsilon}_t = R_t/\widehat{\sigma}_t,$$

where

$$\widehat{\sigma}_t^2 = \widehat{\beta}_0 + \sum_{j=1}^{k} \widehat{\beta}_j R_{t-j}^2.$$

Taniai and Taniguchi (2008) show that VaR for this ARCH model can be approximated by

$$\widehat{\mathrm{VaR}}_p \approx \widehat{\sigma}_{n+1} \left[F^{-1}(p) + \widehat{\sigma} \Phi^{-1}(\alpha)/\sqrt{n} \right],$$

where

$$\begin{aligned} \sigma^2 \quad = \quad & \frac{1}{f^2\left(F^{-1}(p)\right)} \left[p(1-p) \right. \\ & + F^{-1}(p) f\left(F^{-1}(p)\right) \left\{ \int_{-\infty}^{F^{-1}(p)} u^2 f(u) du - p \right\} \boldsymbol{\tau}^T \mathbf{U}^{-1} \mathbf{V} \\ & \left. + \frac{1}{4} \left(F^{-1}(p)\right)^2 f^2\left(F^{-1}(p)\right) \boldsymbol{\tau}^T \mathbf{U}^{-1} \mathbf{S} \mathbf{U}^{-1} \boldsymbol{\tau} \right], \end{aligned}$$

where $\mathbf{V} = E\left[\sigma_t^2 \mathbf{W}_{t-1}\right]$, $\mathbf{S} = 2E\left[\sigma_t^4 \mathbf{W}_{t-1} \mathbf{W}_{t-1}^T\right]$, $\mathbf{W} = \left(1, R_t^2, \ldots, R_{t-k+1}^2\right)^T$, $\mathbf{U} = E\left[\mathbf{W}_{t-1} \mathbf{W}_{t-1}^T\right]$, $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_k)^T$, $\tau_0 = E\left[1/\sigma_t^2\right]$, and $\tau_j = E\left[R_{t-j}^2/\sigma_t^2\right]$, $j = 1, 2, \ldots, k$.

3.22 GARCH models

Suppose the financial returns, say R_t, satisfy the model

$$\left[1 - \phi(L)\right] R_t = \left[1 - \theta(L)\right] \epsilon_t, \ \epsilon_t = \eta_t \sqrt{h_t}, \tag{16}$$

where η_t are independent and identical standard normal random variables, R_t is the return at time t, L denotes the lag operator satisfying $L R_t = R_{t-1}$, $\phi(L)$ is the polynomial $\phi(L) = 1 - \sum_{i=1}^{r} \phi_i L^i$, $\theta(L)$ is the polynomial $\theta(L) = 1 + \sum_{i=1}^{s} \theta_i L^i$, h_t is the conditional variance, and η_t are independent and identical residuals with zero means and unit variances. One popular specification for h_t is

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}. \tag{17}$$

This corresponds to the GARCH (p, q) model.

For the model given by (16) and (17), Chan (2009b) proposes the following algorithm for computing VaR:

1. Estimate the maximum likelihood estimates of the parameters in (16) and (17);
2. Using the parameter estimates, compute the standardized residuals $\widehat{\eta}_t = \left(R_t - \widehat{r}_t\right)/\widehat{h}_t$;
3. Compute the first k sample moments for $\widehat{\eta}_t$;

4. Compute

$$\widehat{p}(\eta_t) = \exp\left(\sum_{i=1}^{k} \lambda_i \eta_t^i\right) \Big/ \int \exp\left(\sum_{i=1}^{k} \lambda_i \eta_t^i\right) d\eta_t.$$

The parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$ are determined from the sample moments of step 3 in a way explained in Chan (2009a) and Rockinger and Jondeau (2002);

5. Compute $\widehat{\mathrm{VaR}}_p$ as the root of the equation

$$\int_{-\infty}^{K} \widehat{p}(\eta_t)\, d\eta_t = p.$$

3.23 GARCH model with heavy tails

Chan *et al.* (2007) consider the case that financial returns, say R_t, come from a GARCH (p, q) specified by

$$R_t = \sigma_t \epsilon_t, \ \sigma_t^2 = c + \sum_{i=1}^{p} b_i R_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2,$$

where R_t is strictly stationary with $ER_t^2 < \infty$, and ϵ_t are zero mean, unit variance, independent and identical random variables independent of $\{R_{t-k}, k \geq 1\}$. Further, Chan *et al.* (2007) assume that ϵ_t have heavy tails, that is their cdf, say G, satisfies

$$\lim_{x \to \infty} \frac{1 - G(xy)}{1 - G(x)} = y^{-\gamma}, \ \lim_{x \to \infty} \frac{G(-x)}{1 - G(x)} = d$$

for all $y > 0$, where $\gamma > 0$ and $d \geq 0$. Chan *et al.* (2007) show that the VaR for this model given by

$$\mathrm{VaR}_\alpha = \inf\left\{x : \Pr\left(R_{n+1} \leq x | R_{n+1-k}, k \geq 1\right) \geq \alpha\right\}$$

can be estimated by

$$\widehat{\mathrm{VaR}}_\alpha = \widetilde{\sigma}_{n+1}\left(\widehat{a}, \widehat{b}, \widehat{c}\right)(1 - \alpha)^{-1/\widehat{\gamma}} \left(\frac{k}{m}\right)^{1/\widehat{\gamma}} \widehat{\epsilon}_{m,m-k},$$

where

$$
\begin{aligned}
\widetilde{\sigma}_t^2(a,b,c) &= \frac{c}{1-\sum_{j=1}^{q} a_j} + \sum_{i=1}^{p} b_i R_{t-i}^2 \\
&+ \sum_{i=1}^{p} b_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \cdots \sum_{j_k=1}^{q} a_{j_1} \cdots a_{j_k} R_{t-i-j_1-\cdots-j_k}^2 I\left\{t-i-j_1-\cdots-j_k \geq 1\right\}, \\
L_\nu(a,b,c) &= \sum_{t=\nu}^{n} \left\{R_t^2/\widetilde{\sigma}_t^2(a,b,c) + \ln \widetilde{\sigma}_t^2(a,b,c)\right\}, \\
\left(\widehat{a},\widehat{b},\widehat{c}\right) &= \operatorname{argmin}_{(a,b,c)} L_\nu(a,b,c), \\
\widehat{\epsilon}_t &= R_t/\widetilde{\sigma}_t^2\left(\widehat{a},\widehat{b},\widehat{c}\right), \\
\widehat{\gamma} &= \left\{\frac{1}{k}\sum_{i=1}^{k} \ln \frac{\widehat{\epsilon}_{m,m-i+1}}{\widehat{\epsilon}_{m,m-k}}\right\}^{-1},
\end{aligned}
$$

where $\nu = \nu(n) \to \infty$ and $\nu/n \to 0$ as $n \to \infty$, $m = n - \nu + 1$, $\widehat{\epsilon}_{m,1} \leq \widehat{\epsilon}_{m,2} \leq \cdots \leq \widehat{\epsilon}_{m,m}$ are the order statistics of $\widehat{\epsilon}_\nu, \widehat{\epsilon}_{\nu+1}, \ldots, \widehat{\epsilon}_n$, and $k = k(m) \to \infty$ and $k/m \to 0$ as $n \to \infty$. Chan *et al.* (2007) also establish asymptotic normality of $\widehat{\mathrm{VaR}}_\alpha$.

3.24 ARMA-GARCH model

Suppose the financial returns, say R_t, $t = 1, 2, \ldots, T$, satisfy the ARMA(p,q)-GARCH(r,s) model specified by

$$
\begin{aligned}
R_t &= a_0 + \sum_{i=1}^{p} a_i R_{t-i} + \epsilon_t + \sum_{j=1}^{q} b_j \epsilon_{t-j}, \\
\sigma_t^2 &= c_0 + \sum_{i=1}^{r} c_i \epsilon_{t-i}^2 + \sum_{j=1}^{s} d_j \sigma_{t-j}^2, \\
\epsilon_t &= z_t \sigma_t,
\end{aligned}
$$

where z_t are independent standard normal random variables. For this model, Hartz *et al.* (2006) show that the h-step ahead forecast of value at risk can be estimated by

$$
\widehat{\mu}_{T+h} + \widehat{\sigma}_{T+h} \Phi^{-1}(\alpha),
$$

where

$$\begin{aligned}
\widehat{\epsilon}_t &= R_t - \widehat{a}_0 - \sum_{i=1}^{p} \widehat{a}_i R_{t-i} - \sum_{j=1}^{q} \widehat{b}_j \widehat{\epsilon}_{t-j}, \\
\widehat{\sigma}_t^2 &= \widehat{c}_0 + \sum_{i=1}^{r} \widehat{c}_i \widehat{\epsilon}_{t-i}^2 + \sum_{j=1}^{s} \widehat{d}_j \widehat{\sigma}_{t-j}^2, \\
\widehat{\mu}_{T+h} &= \widehat{a}_0 + \sum_{i=1}^{p} \widehat{a}_i R_{T+h-i} + \sum_{j=1}^{q} \widehat{b}_j \widehat{\epsilon}_{T+h-j}, \\
\widehat{\sigma}_{T+h}^2 &= \widehat{c}_0 + \sum_{i=1}^{r} \widehat{c}_i \widehat{\epsilon}_{T+h-i}^2 + \sum_{j=1}^{s} \widehat{d}_j \widehat{\sigma}_{T+h-j}^2.
\end{aligned}$$

The parameter estimators required can be obtained, for example, by the method of maximum likelihood.

3.25 Markov switching ARCH model

Suppose the financial returns, say R_t, $t = 1, 2, \ldots, T$, satisfy the Markov switching ARCH model specified by

$$\begin{aligned}
R_t &= u_{s_t} + \epsilon_t, \\
\epsilon_t &= \left(g_{s_t} w_t\right)^{1/2}, \\
w_t &= \left(h_t e_t\right)^{1/2}, \\
h_t &= a_0 + a_1 w_{t-1}^2 + \cdots + a_q w_{t-q}^2,
\end{aligned}$$

where e_t are standard normal random variables, s_t is an unobservable random variable assumed to follow a first-order Markov process, and w_t is a typical ARCH (q) process. This model is due to Bollerslev (1986). An estimator of the value at risk at time t can be obtained by inverting the cdf of R_t with its parameters replaced by their maximum likelihood estimators.

3.26 Fractionally integrated GARCH model

Suppose the financial returns, say R_t, $t = 1, 2, \ldots, T$, satisfy the fractionally integrated GARCH model specified by

$$\begin{aligned}
R_t &= \sigma_t \epsilon_t, \\
\sigma_t^2 &= w + \sum_{i=1}^{p} \beta_i \left(\sigma_{t-i}^2 - R_{t-i}^2\right) - \sum_{i=1}^{\infty} \lambda_i R_{t-i}^2,
\end{aligned}$$

where ϵ_t are random variables with zero means and unit variances. This model is due to Baillie *et al.* (1996). An estimator of the value at risk at time t can be obtained by inverting the cdf of R_t with its parameters replaced by their maximum likelihood estimators. This of course depends on the distribution of ϵ_t. If, for example, ϵ_t are normally distributed then $\widehat{\mathrm{VaR}}_{t,\alpha} = \widehat{\sigma}_{t+1} \Phi^{-1}(\alpha)$, where $\widehat{\sigma}_{t+1}$ may be the maximum likelihood estimator of σ_{t+1}.

3.27 RiskMetrics model

Suppose $\{R_t\}$ are the log-returns of $\{X_1, X_2, \ldots, X_n\}$ and let Ω_t denote the information up to time t. The RiskMetrics model (RiskMetrics, 1996) is specified by

$$\begin{aligned}
&R_t = \epsilon_t, \\
&\epsilon_t \left| \Omega_{t-1} \sim N\left(0, \sigma_t^2\right), \right. \\
&\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)\epsilon_{t-1}^2, \ 0 < \lambda < 1.
\end{aligned}$$

The value at risk for this model can be computed by inverting

$$\Pr\left(R_t < \mathrm{VaR}_{t,\alpha}\right) = \alpha$$

with the parameters, σ_t^2 and λ, replaced by their maximum likelihood estimators.

3.28 Capital asset pricing model

Let R_i denote the return on asset i, let R_f denote the "risk-free rate", and let R_m denote the "return on the market portfolio". With this notation, Fernandez (2006) considers the capital asset pricing model given by

$$R_i - R_f = \alpha_i + \beta_i \left(R_m - R_f\right) \epsilon_i$$

for $i = 1, 2, \ldots, k$, where ϵ_i are independent random variables with $\mathrm{Var}\left(\epsilon_i\right) = \sigma_{\epsilon_i}^2$, and $\mathrm{Var}\left(R_m\right) = \sigma_m^2$. It is easy to see that

$$\begin{aligned}
\mathrm{Var}\left(R_i\right) &= \beta_i^2 \sigma_m^2 + \sigma_e^2, \\
\mathrm{Cov}\left(R_i, R_j\right) &= \beta_i \beta_j \sigma_m^2.
\end{aligned}$$

Fernandez (2006) shows that the value at risk of the portfolio of k assets can be expressed as

$$\mathrm{VaR}_\alpha = V_0 \Phi^{-1}(\alpha) \sqrt{\mathbf{w}^T \left(\boldsymbol{\beta}\boldsymbol{\beta}^T \sigma_m^2 + \mathbf{E}\right) \mathbf{w}}, \tag{18}$$

where $\mathbf{w}$ is a $k \times 1$ vector of portfolio weights, V_0 is the initial value of the portfolio, $\boldsymbol{\beta} = \left(\beta_1, \ldots, \beta_k\right)^T$ and $\mathbf{E} = \mathrm{diag}\left(\sigma_{\epsilon_1}^2, \ldots, \sigma_{\epsilon_k}^2\right)^T$. An estimator of (18) can be obtained by replacing the parameters by their maximum likelihood estimators.

3.29 Dagum distribution

The Dagnum distribution is due to Dagum (1977, 1980). It has the pdf and cdf specified by

$$f(x) = \beta\lambda\delta \exp(-\delta x) \left[1 + \lambda \exp(-\delta x)\right]^{-\beta-1}$$

and

$$F(x) = \left[1 + \lambda \exp(-\delta x)\right]^{-\beta},$$

respectively, for $x > 0$, $\lambda > 0$, $\beta > 0$ and $\delta > 0$. Domma and Perri (2009) discuss an application of this distribution for VaR estimation. They show that

$$\widehat{\mathrm{VaR}}_p = \frac{1}{\widehat{\delta}} \ln\left(\frac{\widehat{\lambda}}{p^{-1/\widehat{\beta}} - 1}\right),$$

where $\left(\widehat{\lambda}, \widehat{\beta}, \widehat{\delta}\right)$ are maximum likelihood estimators of (λ, β, δ) based on $\{X_1, X_2, \ldots, X_n\}$ being a random sample coming from the Dagnum distribution. Domma and Perri (2009) show further that

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_p - \mathrm{VaR}_p\right) \to N\left(0, \sigma^2\right)$$

in distribution as $n \to \infty$, where $\sigma = \mathbf{g}\mathbf{I}^{-1}\mathbf{g}^T$ and

$$\mathbf{g} = \left[-\frac{p^{-1/\beta}\ln p}{\delta\beta^2\left(p^{-1/\beta}-1\right)}, \frac{1}{\lambda\delta}, -\frac{1}{\delta^2}\ln\left(\frac{\lambda}{p^{-1/\beta}-1}\right)\right].$$

Here, $\mathbf{I}$ is the expected information matrix of $\left(\widehat{\lambda}, \widehat{\beta}, \widehat{\delta}\right)$. An explicit expression for the matrix is given in the appendix of Domma and Perri (2009).

3.30 Location-scale distributions

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a location-scale family with cdf $F_{\mu,\sigma}(x) = F_0\left((x-\mu)/\sigma\right)$ and pdf $f_{\mu,\sigma}(x)$. Then,

$$\mathrm{VaR}_p = \mu + z_p\sigma, \tag{19}$$

where $z_p = F_0^{-1}(p)$. The point estimator for VaR is

$$\widehat{\mathrm{VaR}}_p = \widehat{\mu}_n + z_p c_n \widehat{\sigma}_n,$$

where

$$\widehat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\widehat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \widehat{\mu}_n\right)^2,$$

and

$$c_n = \left(E\left[\widehat{\sigma}_n/\sigma\right]\right)^{-1}.$$

Bae and Iscoe (2012) propose various confidence intervals for VaR. Based on $c_n = 1 + O\left(n^{-1}\right)$ and asymptotic normality, Bae and Iscoe (2012) propose the interval

$$\widehat{\mu}_n + z_p\widehat{\sigma}_n \pm \frac{\widehat{\sigma}_n}{\sqrt{n}} z_{(1+\alpha)/2}\sqrt{1 + \frac{z_p^2}{4}(\kappa - 1) + z_p\omega}, \tag{20}$$

where α is the confidence level, κ is the kurtosis of $F_0(x)$, and ω is the skewness of $F_0(x)$. Based on Bahadur (1966)'s almost sure representation of the sample quantile of a sequence of independent random variables, Bae and Iscoe (2012) propose the interval

$$\widehat{\xi}_p \pm \frac{1}{\sqrt{n}} z_{(1+\alpha)/2}\frac{\sqrt{p(1-p)}}{f_{\mu,\sigma}\left(\widehat{\xi}_p\right)},$$

where ξ_p is the pth quantile and $\widehat{\xi}_p$ is its sample counterpart.

Sometimes the financial series of interest is strictly positive. In this case, if $X_1, X_2, \ldots, X_n$ is a random sample from a log location-scale family with cdf $G_{\mu,\sigma}(x) = \ln F_0\left((x-\mu)/\sigma\right)$, then (19) and (20) generalize to

$$\mathrm{VaR}_p = \exp\left(\mu + z_p \sigma\right)$$

and

$$\exp\left(\widehat{\mu}_n + z_p \widehat{\sigma}_n \pm \frac{\widehat{\sigma}_n}{n} z_{(1+\alpha)/2} \sqrt{1 + \frac{z_p^2}{4}(\kappa - 1) + z_p \omega}\right),$$

respectively, as noted by Bae and Iscoe (2012).

3.31 Discrete distributions

Göb (2011) considers VaR estimation for the three most common discrete distributions: Poisson, binomial and negative binomial. Let

$$L_c(\lambda) = \sum_{y=0}^{c} \frac{\lambda^y \exp(-\lambda)}{y!}.$$

Then, the VaR for the Poisson distribution is

$$\mathrm{VaR}_p(\lambda) = \inf\left\{c = 0, 1, \ldots \mid L_c(\lambda) \geq p\right\}.$$

Letting

$$L_{n,c}(r) = \sum_{y=0}^{c} \binom{n}{y} r^y (1-r)^{n-y},$$

the VaR for the binomial distribution is

$$\mathrm{VaR}_p(r) = \inf\left\{c = 0, 1, \ldots \mid L_{n,c}(r) \geq p\right\}.$$

Letting

$$H_{n,c}(r) = \sum_{y=0}^{c} \binom{n+y-1}{y} (1-r)^y r^n,$$

the VaR for the negative binomial distribution is

$$\mathrm{VaR}_p(r) = \inf\left\{c = 0, 1, \ldots \mid H_{n,c}(r) \geq p\right\}.$$

Göb (2011) derives various properties of these VaR measures in terms of their parameters. For the Poisson distribution, the following properties were derived:

(a) for fixed $p \in (0,1)$, $\mathrm{VaR}_p(\lambda)$ is increasing in $\lambda \in [0, \infty)$ with $\lim_{\lambda \to \infty} \mathrm{VaR}_p(\lambda) = \infty$. There are values $0 = \lambda_{-1} < \lambda_0 < \lambda_1 < \lambda_2 < \cdots$, $\lim_{c \to \infty} \lambda_c = \infty$, such that, for $c \in \mathbb{N}_0$, $\mathrm{VaR}_p(\lambda) = c$ on the interval $(\lambda_{c-1}, \lambda_c]$ and $L_c(\lambda) > L_c\left(\lambda_c\right) = p$ for $\lambda \in (\lambda_{c-1}, \lambda_c)$. In particular, $\lambda_0 = -\ln(p)$.

(b) For fixed $\lambda > 0$, $c = -1, 0, 1, \ldots$, let $p_c = L_c(\lambda)$. Then, for $c = 0, 1, 2, \ldots$, $\mathrm{VaR}_p(\lambda) = c$ for $p \in (p_{c-1}, p_c]$.

For the binomial distribution, the following properties were derived:

(a) for fixed $p \in (0,1)$, $\mathrm{VaR}_p(r)$ is increasing in $r \in [0,1]$. There are values $0 = r_{-1} < r_0 < r_1 < r_2 < \cdots < r_n = 1$ such that, for $c \in \{0, \ldots, n\}$, $\mathrm{VaR}_p(r) = c$ on the interval $(r_{c-1}, r_c]$ and $L_{n,c}(r) > L_{n,c}(r_c) = p$ for $r \in (r_{c-1}, r_c)$. In particular, $r_0 = 1 - p^{1/n}$ and $r_{n-1} = (1-p)^{1/n}$.

(b) For fixed $0 < r < 1$, $c = -1, 0, 1, \ldots$, let $p_c = L_{n,c}(r)$. Then, for $c = 0, 1, 2, \ldots, n$, $\mathrm{VaR}_p(r) = c$ for $p \in (p_{c-1}, p_c]$.

For the negative binomial distribution, the following properties were derived:

(a) for fixed $p \in (0,1)$, $\mathrm{VaR}_p(r)$ is decreasing in $r \in [0,1]$. There are values $1 = r_{-1} > r_0 > r_1 > r_2 > \cdots$, $\lim_{c \to \infty} r_c = 0$, such that, for $c \in \mathbb{N}_0$, $\mathrm{VaR}_p(r) = c$ on the interval $[r_c, r_{c-1})$ and $H_{n,c}(r) > H_{n,c}(r_c) = p$ for $r \in (r_c, r_{c-1})$. In particular, $r_0 = p^{1/n}$.

(b) For fixed $0 < r < 1$, $c = -1, 0, 1, \ldots$, let $p_c = H_{n,c}(r)$. Then, for $c = 0, 1, 2, \ldots, n$, $\mathrm{VaR}_p(r) = c$ for $p \in (p_{c-1}, p_c]$.

Empirical estimation of the three VaR measures can be based on asymptotic normality.

3.32 Quantile regression method

Quantile regressions have been used to estimate value at risk, see Koenker and Basset (1978), Koenker and Portnoy (1997), Chernozhukov and Umantsev (2001) and Engle and Manganelli (2004). The idea is to regress the value at risk on some known covariates. Let X_t at time t denote the financial variable, let $\mathbf{z}_t$ denote a $k \times 1$ vector of covariates at time t, let $\boldsymbol{\beta}_\alpha$ denote a $k \times 1$ vector of regression coefficients, and let $\mathrm{VaR}_{t,\alpha}$ denote the corresponding value at risk. Then, the quantile regression model can be rewritten as

$$\mathrm{VaR}_{t,\alpha} = g\left(\mathbf{z}_t; \boldsymbol{\beta}_\alpha\right). \tag{21}$$

In the linear case, (21) could take the form

$$\mathrm{VaR}_{t,\alpha} = \mathbf{z}_t^T \boldsymbol{\beta}_\alpha.$$

The parameters in (21) can be estimated by least squares as in standard regression.

3.33 Brownian motion method

Cakir and Raei (2007) describe simulation schemes for computing value at risk for single asset and multiple asset portfolios. Let P_t denote the price at time t, let T denote a holding period divided into small intervals of equal length Δt, let ΔP_t denote the change in P_t over Δt, let Z_t denote a standard normal shock, let μ denote the mean of returns over the holding period T, and let σ denote the standard deviation of returns over the holding period T. With these notation, Cakir and Raei (2007) suggest the model

$$\frac{\Delta P_t}{P_t} = \mu \Delta t + \sigma \sqrt{\Delta t} Z_t. \tag{22}$$

Under this model, the VaR for single asset portfolios can be computed as follows:

(i) starting with P_t, simulate $P_t, P_{t+1}, \ldots, P_T$ using (22);

(ii) repeat step (i) ten thousand times;

(iii) compute the empirical cdf over the holding period;

(iii) compute $\widehat{\mathrm{VaR}}_\alpha$ as 100α percentile of the empirical cdf.

The VaR for multiple asset portfolios can be computed as follows:

(i) suppose the price at time t for the ith asset follows

$$\frac{\Delta P_t^i}{P_t^i} = \mu^i \Delta t + \sigma^i \sqrt{\Delta t} Z_t^i \tag{23}$$

for $i = 1, 2, \ldots, N$, where N is the number of assets, and the notation is the same as that for single asset portfolios. The standard normal shocks, $Z_t^i, i = 1, 2, \ldots, N$, need not be correlated;

(ii) starting with P_t^i, $i = 1, 2, \ldots, N$, simulate $P_t^i, P_{t+1}^i, \ldots, P_T^i$, $i = 1, 2, \ldots, N$ using (23);

(iii) compute the portfolio price for the holding period as the weighted sum of the individual asset prices;

(iv) repeat steps (ii) and (iii) ten thousand times;

(v) compute the empirical cdf of the portfolio price over the holding period;

(vi) compute $\widehat{\mathrm{VaR}}_\alpha$ as 100α percentile of the empirical cdf.

3.34 Bayesian method

Pollard (2007) defines a Bayesian value at risk. Let X_t denote the financial variable of interest at time t. Let $p\left(X_t \mid \mathbf{\Theta}, Z_t\right)$ denote the posterior pdf of X_t given some parameters $\mathbf{\Theta}$ and "state" variables Z_t. Pollard (2007) defines the Bayesian value at risk at time t as

$$\mathrm{VaR}_\alpha = \left\{ x : \int_{-\infty}^{x} p\left(y \mid \mathbf{\Theta}, Z_{t+1}\right) dy = \alpha \right\}. \tag{24}$$

The "state" variables Z_t are assumed to follow a transition pdf $f\left(Z_t, Z_{t+1}\right)$.

Pollard (2007) also proposes several methods for estimating (24). One of them is the following:

(i) Use Markov Chain Monte Carlo to simulate N samples, $\left\{ \left(Z_t^{(n)}, \mathbf{\Theta}^{(n)} \right), n = 1, 2, \ldots, N \right\}$, from the joint conditional posterior pdf of $\left(Z_t, \mathbf{\Theta}\right)$ given $Y_t = \left\{X_\tau, \tau = 1, 2, \ldots, t\right\}$;

(ii) For n from 1 to N, simulate $Z_{t+1}^{(n)}$ from the conditional posterior pdf of Z_{t+1} given $\mathbf{\Theta}^{(n)}$ and $Z_t^{(n)}$;

(iii) For n from 1 to N, simulate $X_{t+1}^{(n)}$ from the conditional posterior pdf of X_{t+1} given $\mathbf{\Theta}^{(n)}$ and $Z_{t+1}^{(n)}$;

(iv) Compute the empirical cdf

$$\widehat{G}(x) = \frac{1}{N} \sum_{n=1}^{N} I\left\{ X_{t+1}^{(n)} \leq x \right\}; \tag{25}$$

(v) Estimate VaR as $\widehat{G}^{-1}(\alpha)$.

3.35 Rachev *et al.*'s method

Let $R = \sum_{i=1}^{n} w_i R_i$ denote a portfolio return made up of n asset returns, R_i, and the non-negative weights w_i summing to one. Suppose R_i are independent $S_\alpha \left(\alpha_i, \beta_i, 0\right)$ random variables. Then, it can be shown that (Rachev *et al.*, 2003) $R \sim S_\alpha \left(\alpha_p, \beta_p, 0\right)$, where

$$\alpha_p = \left[\sum_{i=1}^{n} \left(|w_i| \, \sigma_i \right)^\alpha \right]^{1/\alpha}$$

and

$$\beta_p = \frac{\sum_{i=1}^{n} \operatorname{sign}(w_i) \, \beta_i \left(|w_i| \, \sigma_i \right)^\alpha}{\sum_{i=1}^{n} \left(|w_i| \, \sigma_i \right)^\alpha}.$$

Hence, the value of risk of R can be estimated by the following algorithm due to Rachev *et al.* (2003):

- estimate α_i and β_i (to obtain say $\widehat{\alpha}_i$ and $\widehat{\beta}_i$) using possible data on the ith asset return;
- estimate α_p and β_p by

  $$\widehat{\alpha}_p = \left[\sum_{i=1}^{n} \left(|w_i| \, \widehat{\sigma}_i \right)^\alpha \right]^{1/\alpha}$$

  and

  $$\widehat{\beta}_p = \frac{\sum_{i=1}^{n} \operatorname{sign}(w_i) \, \widehat{\beta}_i \left(|w_i| \, \widehat{\sigma}_i \right)^\alpha}{\sum_{i=1}^{n} \left(|w_i| \, \widehat{\sigma}_i \right)^\alpha},$$

  respectively;
- estimate $\mathrm{VaR}_p(R)$ as the pth quantile of $S_\alpha \left(\widehat{\alpha}_p, \widehat{\beta}_p, 0 \right)$.

4 Nonparametric methods

This section concentrates on estimation methods for value at risk when the data are assumed to come from no particular distribution. The nonparametric methods summarized are based on: historical method (Section 4.1), filtered historical method (Section 4.2), importance sampling method (Section 4.3), bootstrap method (Section 4.4), kernel method (Section 4.5), Chang *et al.*'s estimators (Section 4.6), Jadhav and Ramanathan's method (Section 4.7), and Jeong and Kang's method (Section 4.8).

4.1 Historical method

Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics in ascending order corresponding to the original financial series $X_1, X_2, \ldots, X_n$. The historical method suggests to estimate value at risk by

$$\widehat{\mathrm{VaR}}_\alpha(X) = X_{(i)}$$

for $\alpha \in ((i-1)/n, i/n]$.

4.2 Filtered historical method

Suppose the log-returns, $R_t = \ln X_t - \ln X_{t-1}$, follow the model, $R_t = \sigma_t \epsilon_t$, discussed before, where σ_t is the volatility process and ϵ_t are independent and identical random variables with zero means. Let $\epsilon_{(1)} \leq \epsilon_{(2)} \leq \cdots \leq \epsilon_{(n)}$ denote the order statistics of $\{\epsilon_t\}$. The filtered historical method suggests to estimate value at risk by

$$\widehat{\mathrm{VaR}}_\alpha = \epsilon_{(i)} \widehat{\sigma}_t$$

for $\alpha \in ((i-1)/n, i/n]$, where $\widehat{\sigma}_t$ denotes an estimator of σ_t at time t. This method is due to Hull and White (1998) and Barone-Adesi *et al.* (1999).

4.3 Importance sampling method

Suppose $\widehat{F}(\cdot)$ is the empirical cdf of $X_1, X_2, \ldots, X_n$. As seen in Section 4.1, an estimator for VaR is $\widehat{F}^{-1}(\alpha)$. This estimator is asymptotically normal with variance equal to

$$\frac{\alpha(1-\alpha)}{n f^2\left(\mathrm{VaR}_\alpha\right)}.$$

This can be large if α is closer to zero or one. There are several methods for variance reduction. One popular method is importance sampling. Suppose $G(\cdot)$ is another cdf and let $S(x) = \widehat{F}(dx)/G(dx)$ and

$$\widehat{S}(x) = \frac{1}{n} \sum_{i=1}^{n} I\left\{X_i \leq x\right\} S\left(X_i\right).$$

Hong (2011) shows that $\widehat{S}^{-1}(p)$ under certain conditions can provide estimators for VaR with smaller variance.

4.4 Bootstrap method

Suppose $\widehat{F}(\cdot)$ is the empirical cdf of $X_1, X_2, \ldots, X_n$. The bootstrap method can be described as follows:

1. simulate B independent sample from $\widehat{F}(\cdot)$;
2. for each sample estimate VaR_α, say $\widehat{\mathrm{VaR}}_\alpha^{(i)}$ for $i = 1, 2, \ldots, B$, using the historical method;
3. take the estimate of VaR as the mean or the median of $\widehat{\mathrm{VaR}}_\alpha^{(i)}$ for $i = 1, 2, \ldots, B$.

One can also construct confidence intervals for VaR based on the bootstrapped estimates $\widehat{\mathrm{VaR}}_\alpha^{(i)}$, $i = 1, 2, \ldots, B$.

4.5 Kernel method

Kernels are commonly used to estimate pdfs. Let $K(\cdot)$ denote a symmetric kernel, i.e., a symmetric pdf. The kernel estimator of F can be given by

$$\widehat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} G\left(\frac{x - X_i}{h}\right), \tag{26}$$

where h is a smoothing bandwidth and

$$G(x) = \int_{-\infty}^{x} K(u) du.$$

A variable width version of (26) is

$$\widehat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{hT_i} G\left(\frac{x - X_i}{hT_i}\right), \tag{27}$$

where $T_i = d_k(x_i)$ is the distance of X_i from its kth nearest neighbor among the remaining $(n-1)$ data points and $k = n^{-1/2}$. The kernel estimator of VaR, say $\widehat{\mathrm{VaR}}_p$, is then the root of the equation

$$\widehat{F}(x) = p \tag{28}$$

for x, where $\widehat{F}(\cdot)$ is given by (26) or (27). According to Sheather and Marron (1990), $\widehat{\mathrm{VaR}}_p$ could also be estimated by

$$\widehat{\mathrm{VaR}}_p = \frac{\displaystyle\sum_{i=1}^{n} \widehat{F}\left((i - 1/2)/n - p\right) X_{(i)}}{\displaystyle\sum_{i=1}^{n} \widehat{F}\left((i - 1/2)/n - p\right)},$$

where $\widehat{F}(\cdot)$ is given by (26) or (27) and $\left\{X_{(i)}\right\}$ are the ascending order statistics of X_i.

The estimator in (28) is due to Gourieroux *et al.* (2000). Its properties have been studied many authors. For instance, Chen and Tang (2005) show under certain regularity conditions that

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_p - \mathrm{VaR}_p\right) \to N\left(0, \sigma^2(p) f^{-2}\left(\mathrm{VaR}_p\right)\right)$$

in distribution as $n \to \infty$, where

$$\sigma^2(p) = \lim_{n\to\infty} \sigma^2(p;n),$$

$$\sigma^2(p;n) = \left\{ p(1-p) + 2\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \gamma(k) \right\},$$

$$\gamma(k) = \mathrm{Cov}\left\{ I\left(X_1 < \mathrm{VaR}_p\right), I\left(X_{k+1} < \mathrm{VaR}_p\right)\right\}.$$

Here, $I\{\cdot\}$ denotes the indicator function.

4.6 Chang *et al.*'s estimators

Chang *et al.* (2003) propose several non-parametric estimators for the VaR of log-returns, say R_t with pdf $f(\cdot)$. The first of these is $\widehat{\mathrm{VaR}}_\alpha = (1-w)R_{(m)} + wR_{(m+1)}$, where $m = [n\alpha + 0.5]$ and $w = n\alpha - m + 0.5$, where $[x]$ denotes the greatest integer less than or equal to x. This estimator is shown to have the asymptotic distribution

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_\alpha - \mathrm{VaR}_\alpha\right) \to N\left(0, \alpha(1-\alpha)(p) f^{-2}\left(\mathrm{VaR}_\alpha\right)\right)$$

in distribution as $n \to \infty$. It is sometimes referred to as the historical simulation estimator. The second of the proposed estimators is

$$\begin{aligned} \widehat{\mathrm{VaR}}_\alpha &= \sum_{i=1}^{n} R_{(i)} \Big[B_{i/n}\left((n+1)\alpha, (n+1)(1-\alpha)\right) \\ &\qquad - B_{(i-1)/n}\left((n+1)\alpha, (n+1)(1-\alpha)\right) \Big]. \end{aligned}$$

This estimator is shown to have the asymptotic distribution

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_\alpha - \mathrm{VaR}_\alpha\right) \to N\left(0, \alpha(1-\alpha)(p) f^{-2}\left(\mathrm{VaR}_\alpha\right)\right)$$

in distribution as $n \to \infty$. The third of the proposed estimators is

$$\widehat{\mathrm{VaR}}_\alpha = \sum_{i=1}^{n} k_{i,n} R_{(i)},$$

where

$$k_{i,n} = B_{q_{i,n}}\left((n+1)\alpha, (n+1)(1-\alpha)\right) - B_{q_{i-1,n}}\left((n+1)\alpha, (n+1)(1-\alpha)\right),$$

$$q_{0,n} = 0, \; q_{i,n} = \sum_{j=1}^{i} w_{j,n}, j = 1, 2, \ldots, n,$$

$$w_{i,n} = \begin{cases} \dfrac{1}{2}\left[1 - \dfrac{n-2}{\sqrt{n(n-1)}}\right], & \text{if } i = 1, n, \\ \dfrac{1}{\sqrt{n(n-1)}}, & \text{if } i = 2, 3, \ldots, n-1. \end{cases}$$

This estimator is shown to have the asymptotic distribution

$$\sqrt{n}\left(\widehat{\mathrm{VaR}}_\alpha - \mathrm{VaR}_\alpha\right) \to N\left(0, \alpha(1-\alpha)(p) f^{-2}\left(\mathrm{VaR}_\alpha\right)\right)$$

in distribution as $n \to \infty$.

4.7 Jadhav and Ramanathan's method

Jadhav and Ramanathan (2009) provide a collection of non-parametric estimators for VaR_α. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics in ascending order corresponding to $X_1, X_2, \ldots, X_n$. For given α, define $i = [n\alpha + 0.5]$, $j = [n\alpha]$, $k = [(n+1)\alpha]$, $g = n\alpha - j$, $h = (n+1)\alpha - k$ and $r = [(p+1)\alpha]$. The collection provided is

$$
\begin{aligned}
\widehat{\mathrm{VaR}}_\alpha &= (1-g)X_{(j)} + gX_{(j+1)}, \\
\widehat{\mathrm{VaR}}_\alpha &= \begin{cases} X_{(j)}, & \text{if } g < 0.5, \\ X_{(j+1)}, & \text{if } g \geq 0.5, \end{cases} \\
\widehat{\mathrm{VaR}}_\alpha &= \begin{cases} X_{(j)}, & \text{if } g = 0, \\ X_{(j+1)}, & \text{if } g > 0, \end{cases} \\
\widehat{\mathrm{VaR}}_\alpha &= (1-h)X_{(k)} + hX_{(k+1)}, \\
\widehat{\mathrm{VaR}}_\alpha &= \begin{cases} \dfrac{X_{(j)} + X_{(j+1)}}{2}, & \text{if } g = 0, \\ X_{(j+1)}, & \text{if } g > 0, \end{cases} \\
\widehat{\mathrm{VaR}}_\alpha &= X_{(j+1)}, \\
\widehat{\mathrm{VaR}}_\alpha &= (0.5 + i - np)X_{(i)} + (0.5 - i + np)X_{(i+1)}, \ 0.5 \leq n\alpha \leq n - 0.5, \\
\widehat{\mathrm{VaR}}_\alpha &= \sum_{m=1}^{n} W_{n,m} X_{(m)}, \\
\widehat{\mathrm{VaR}}_\alpha &= \sum_{m=r}^{r+n-p} \frac{\binom{m-1}{r-1}\binom{n-m}{p-r}}{\binom{n}{p}} X_{(m)},
\end{aligned}
$$

where

$$
W_{n,m} = I_{m/n}\left(\alpha(n+1), (1-\alpha)(n+1)\right) - I_{(m-1)/n}\left(\alpha(n+1), (1-\alpha)(n+1)\right),
$$

where $I_x(a,b)$ denote the incomplete beta function ratio defined by

$$
I_x(a,b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1} dt}{B(a,b)} = \frac{\int_0^x t^{a-1}(1-t)^{b-1} dt}{\int_0^1 t^{a-1}(1-t)^{b-1} dt}.
$$

The last of the estimators in the collection is due to Kaigh and Lachenbruch (1982). The second last is due to Harrell and Davis (1982).

4.8 Jeong and Kang's method

Suppose the log-returns, $R_t = \ln X_t - \ln X_{t-1}$, follow the model, $R_t = \sigma_t \epsilon_t$, discussed before. Let $\mathrm{VaR}_{\alpha,t}$ denote the corresponding VaR. Jeong and Kang (2009) propose a fully non-parametric estimator for the $\mathrm{VaR}_{\alpha,t}$ defined by

$$
\Pr\left(R_t < \mathrm{VaR}_{\alpha,t} \middle| \mathcal{F}_{t-1}\right) = \alpha,
$$

where $\mathcal{F}_t$ is the σ-field generated by $(\sigma_s)_{s \leq t}$. Let

$$Q_n(\alpha) = \begin{cases} X_{(s)}, & \text{if } (s-1)/n < \alpha \leq s/n, \\ X_{(1)}, & \text{if } \alpha = 0, \end{cases}$$

$$a_i = \int_0^1 (\alpha - s)^i K\left(\frac{\alpha - s}{h}\right) ds,$$

and

$$A_i(\alpha) = \int_0^1 (\alpha - s)^i K\left(\frac{\alpha - s}{h}\right) Q_n(s) ds$$

for some kernel function $K(\cdot)$ with bandwidth h. With this notation, Jeong and Kang (2009) propose the estimator

$$\widehat{\text{VaR}}_{\alpha,t} = \widehat{\sigma}_t \widehat{q}_2,$$

where

$$\widehat{\sigma}_t^2 = \frac{1}{\widehat{m}} \sum_{p=t-\widehat{m}}^{t-1} R_p^2$$

and

$$\widehat{q}_2 = \frac{A_0(\alpha)\left(a_2 a_4 - a_3^2\right) - A_1(\alpha)\left(a_1 a_4 - a_2 a_3\right) + A_2(\alpha)\left(a_1 a_3 - a_2^2\right)}{a_0\left(a_2 a_4 - a_3^2\right) - a_1\left(a_1 a_4 - a_2 a_3\right) + a_2\left(a_1 a_3 - a_2^2\right)}.$$

Here, $\widehat{m}$ can be determined using a recursive algorithm presented in Section 2.1 of Jeong and Kang (2009).

5 Semiparametric methods

This section concentrates on estimation methods for value at risk which have both parametric and nonparametric elements. The semiparametric methods summarized are based on: extreme value theory method (Section 5.1), generalized Pareto distribution (Section 5.2), Matthys *et al.*'s method (Section 5.3), Araújo Santos *et al.*'s method (Section 5.4), Gomes and Pestana's method (Section 5.5), Beirlant *et al.*'s method (Section 5.6), Caeiro and Gomes's method (Section 5.7), Figueiredo *et al.*'s method (Section 5.8), Li *et al.*'s method (Section 5.9), Gomes *et al.*'s method (Section 5.10), Wang's method (Section 5.11), M-estimation method (Section 5.12), and the generalized Champernowne distribution (Section 5.13).

5.1 Extreme value theory method

Let $M_n = \max\{R_1, R_2, \ldots, R_n\}$ denote the maximum of financial returns. Extreme value theory says that under suitable conditions there exist norming constants $a_n > 0$ and b_n such that

$$\Pr\left\{a_n\left(M_n - b_n\right) \leq x\right\} \to \exp\left\{-(1+\xi x)^{-1/\xi}\right\}$$

in distribution as $n \to \infty$. The parameter ξ is known as the extreme value index. It controls the tail behavior of the extremes.

There are several estimators proposed for ξ. One of the earliest estimators due to Hill (1975) is

$$\widehat{\xi} = \frac{1}{k} \sum_{i=1}^{k} \ln \frac{R_{(i)}}{R_{(k+1)}}, \tag{29}$$

where $R_{(1)} > R_{(2)} > \cdots R_{(k)} > \cdots > R_{(n)}$ are the order statistics in descending order. Another earliest estimator due to Pickands (1975) is

$$\widehat{\xi} = \frac{1}{\ln 2} \ln \frac{R_{(k+1)} - R_{(2k+1)}}{R_{(2k+1)} - R_{(4k+1)}}. \tag{30}$$

The tails of F for most situations in finance take the Pareto form, that is

$$1 - F(x) = C x^{-1/\xi} \tag{31}$$

for some constant C. Embrechts *et al.* (1997, page 334) propose estimating C by $\widehat{C} = (k/n) R_{k+1}^{1/\widehat{\xi}}$.

Combining (29) and (31), Odening and Hinrichs (2003) propose estimating VaR by

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(k+1)} \left(\frac{k}{np} \right)^{\widehat{\xi}}. \tag{32}$$

This estimator is actually due to Weissman (1978).

An alternative approach is to suppose that the maximum of financial returns follows the generalized extreme value cdf (Fisher and Tippett, 1928) given by

$$G(x) = \exp \left\{ - \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-1/\xi} \right\} \tag{33}$$

for $1 + \xi (x - \mu)/\sigma > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$. In this case, the value at risk can be estimated by

$$\widehat{\mathrm{VaR}}_p = \widehat{\mu} - \frac{\widehat{\sigma}}{\widehat{\xi}} \left[1 - \{ - \ln p \}^{-\widehat{\xi}} \right],$$

where $\left(\widehat{\mu}, \widehat{\sigma}, \widehat{\xi} \right)$ are the maximum likelihood estimators of (μ, σ, ξ). Prescott and Walden (1990) provide details of maximum likelihood estimation for the generalized extreme value distribution.

The Gumbel distribution is the particular case of (33) for $\xi = 0$. It has the cdf specified by

$$G(x) = \exp \left\{ - \exp \left(- \frac{x - \mu}{\sigma} \right) \right\}$$

for $\mu \in \mathbb{R}$ and $\sigma > 0$. If the maximum of financial returns follows this cdf then the value at risk can be estimated by

$$\widehat{\mathrm{VaR}}_p = \widehat{\mu} - \widehat{\sigma} \ln \{ - \ln p \},$$

where $(\widehat{\mu}, \widehat{\sigma})$ are the maximum likelihood estimators of (μ, σ).

For more on extreme value theory, estimation of the tail index and applications, we refer the readers to Longin (1996, 2000), Beirlant *et al.* (2015), Fraga Alves and Neves (2015) and Gomes *et al.* (2015).

5.2 Generalized Pareto distribution

The Pareto distribution is a popular model in finance. Suppose the log-return, say R_t, of $X_1, X_2, \ldots, X_n$ comes from the generalized Pareto distribution with cdf specified by

$$F(y) = \frac{N_u}{n} \left(1 + \gamma \frac{y - u}{\sigma}\right)^{-1/\gamma}$$

for $u < y < \infty$, $\sigma > 0$ and $\gamma \in \mathbb{R}$, where u is some threshold and N_u is the number of observed exceedances above u.

For this model, several estimators are available for the VaR. Let $R_{(1)} \leq R_{(2)} \leq \cdots \leq R_{(n)}$ denote the order statistics in ascending order. The first estimator due to Pickands (1975) is

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k+1)} + \frac{1}{1 - 2^{-\widehat{\gamma}}} \left[\left(\frac{k}{(n+1)p}\right)^{\widehat{\gamma}} - 1\right] \left(R_{(n-k+1)} - R_{(n-2k+1)}\right),$$

where

$$\widehat{\gamma} = \frac{1}{\ln 2} \ln \frac{R_{(n-k+1)} - R_{(n-2k+1)}}{R_{(n-2k+1)} - R_{(n-4k+1)}}$$

for $k \neq n/4$. The second estimator due to Dekkers *et al.* (1989) is

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k)} + \frac{\widehat{a}}{\widehat{\gamma}} \left[\left(\frac{k}{np}\right)^{\widehat{\gamma}} - 1\right],$$

where

$$\begin{aligned}
\widehat{\gamma} &= M_{k+1}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_{k+1}^{(1)}\right)^2}{M_{k+1}^{(2)}}\right)^{-1}, \\
M_{(k+1)}^{\ell} &= \frac{1}{k} \sum_{i=1}^{k} \left[\ln R_{(n-i+1)} - \ln R_{(n-k)}\right]^{\ell}, \ \ell = 1, 2, \\
\widehat{a} &= \frac{R_{(n-k)} M_{(k+1)}^{(1)}}{\rho_1}, \\
\rho_1 &= \begin{cases} 1, & \text{if } \gamma \geq 0, \\ \dfrac{1}{1-\gamma}, & \text{if } \gamma < 0. \end{cases}
\end{aligned}$$

Suppose now that the returns are from the alternative generalized Pareto distribution with cdf specified by

$$F(x) = 1 - \left(1 + \xi \frac{x - u}{\sigma}\right)^{-1/\xi}$$

for $1 + \xi(x - u)/\sigma > 0$. Then, the VaR is

$$\mathrm{VaR}_p = u + \frac{\sigma}{\xi}\left[(1-p)^{-\xi} - 1\right]. \tag{34}$$

If $\widehat{\sigma}$ and $\widehat{\xi}$ are the maximum likelihood estimators of σ and ξ, respectively, then the maximum likelihood estimator of VaR is

$$\widehat{\mathrm{VaR}}_p = \widehat{u} + \frac{\widehat{\sigma}}{\widehat{\xi}} \left[(1-p)^{-\widehat{\xi}} - 1 \right].$$

There are several methods for constructing confidence intervals for (34). One popular method is the bias-corrected method due to Efron and Tibshirani (1993). This method based on bootstrapping can be described as follows:

1. Given a random sample $\mathbf{r} = (r_1, r_2, \ldots, r_n)$, calculate the maximum likelihood estimate $\widehat{\boldsymbol{\theta}} = \left(\widehat{\sigma}, \widehat{\xi}\right)$ and $\widehat{\boldsymbol{\theta}}_{(i)}$, the maximum likelihood estimate with the ith data point, r_i, removed;
2. Simulate $\mathbf{r}^{*i} = \{r_1^*, r_2^*, \ldots, r_n^*\}$ from the generalized Pareto distribution with parameters $\widehat{\boldsymbol{\theta}}$;
3. Compute the maximum likelihood estimate, say $\widehat{\boldsymbol{\theta}}^{*i}$, for the sample simulated in step 2;
4. Repeat steps 2 and 3, B times;
5. Compute

$$\alpha_1 = \Phi \left(\widehat{z}_0 + \frac{\widehat{z}_0 + \Phi^{-1}(\alpha)}{1 - \widehat{a} \left(\widehat{z}_0 + \Phi^{-1}(\alpha) \right)} \right)$$

and

$$\alpha_2 = \Phi \left(\widehat{z}_0 + \frac{\widehat{z}_0 + \Phi^{-1}(1-\alpha)}{1 - \widehat{a} \left(\widehat{z}_0 + \Phi^{-1}(1-\alpha) \right)} \right),$$

where

$$\widehat{z}_0 = \Phi^{-1} \left(\frac{\displaystyle\sum_{i=1}^{B} I \left\{ \widehat{\mathrm{VaR}}^{*i} < \widehat{\mathrm{VaR}} \right\}}{B} \right)$$

and

$$\widehat{a} = \frac{\displaystyle\sum_{i=1}^{n} I \left(\overline{\widehat{\mathrm{VaR}}} - \widehat{\mathrm{VaR}}_{(i)} \right)^3}{6 \left\{ \displaystyle\sum_{i=1}^{n} I \left(\overline{\widehat{\mathrm{VaR}}} - \widehat{\mathrm{VaR}}_{(i)} \right)^2 \right\}^{3/2}},$$

where $\overline{\widehat{\mathrm{VaR}}}$ is the mean of $\widehat{\mathrm{VaR}}^{*i}$;

6. Compute the bias-correct confidence interval for VaR as

$$\left(\widehat{\mathrm{VaR}}^{*(\alpha_1)}, \widehat{\mathrm{VaR}}^{*(\alpha_2)} \right),$$

where $\widehat{\mathrm{VaR}}^{*(\alpha)}$ is the 100α percentile of $\widehat{\mathrm{VaR}}^{*i}$.

Note that $\widehat{\boldsymbol{\theta}}^{*i}$ and $\widehat{\mathrm{VaR}}^{*i}$ are the bootstrap replicates of $\boldsymbol{\theta}$ and VaR, respectively.

5.3 Matthys *et al.*'s method

Several improvements have been proposed on (32). The one due to Matthys *et al.* (2004) takes account of censoring. Suppose only N of the n are actually observed, the remaining are considered to be censored or missing. In this case, Matthys *et al.* (2004) show that VaR can be estimated by

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k)} \left[\frac{k+1}{(n+1)p} \right]^{\widehat{\gamma}} \exp \left\{ -\frac{\widehat{b}}{\widehat{\rho}} \left[1 - \left(\frac{(n+1)p}{k+1} \right)^{-\widehat{\rho}} \right] \right\},$$

where

$$\begin{aligned}
H_{k,n}^{(c)} &= \frac{1}{k-n+N} \left[\sum_{j=n-N+1}^{k} \ln \frac{R_{(n-j+1)}}{R_{(n-k)}} + (n-N) \ln \frac{R_{(N)}}{R_{(n-k)}} \right], \\
C &= (n-N)/k, \ Z_{j,k} = j \ln \frac{R_{(n-j+1)}}{R_{(n-j)}}, \\
\widehat{\rho} &= -\frac{1}{\ln \lambda} \ln \frac{H_{[\lambda^2 k],n}^{(c)} - H_{[\lambda k],n}^{(c)}}{H_{[\lambda k],n}^{(c)} - H_{k,n}^{(c)}}, \\
\widehat{\gamma} &= \frac{1}{k-n+N} \sum_{j=n-N+1}^{k} Z_j - \widehat{b} \frac{1 - C^{1-\widehat{\rho}}}{(1-C)(1-\widehat{\rho})}, \\
\widehat{b} &= \frac{\displaystyle \frac{1}{k-n+N} \sum_{j=n-N+1}^{k} \left[\left(\frac{j}{k+1} \right)^{-\widehat{\rho}} - \frac{1 - C^{1-\widehat{\rho}}}{(1-C)(1-\widehat{\rho})} \right] Z_j}{\left[\displaystyle \frac{1 - C^{1-2\widehat{\rho}}}{(1-C)(1-2\widehat{\rho})} - \frac{1 - C^{1-\widehat{\rho}}}{(1-C)(1-\widehat{\rho})} \right]^2}.
\end{aligned}$$

Here, λ is a tuning parameter and takes values in the unit interval. Among other properties, Matthys *et al.* (2004) establish asymptotic normality of VaR_{1-p}.

5.4 Araújo Santos *et al.*'s method

The improvement of (32) due to Araújo Santos *et al.* (2006) takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n_q)} + \left(R_{(n-k)} - R_{(n_q)} \right) \left(\frac{k}{np} \right)^{H_n},$$

where $n_q = [nq] + 1$ and

$$H_n = \frac{1}{k} \sum_{i=1}^{k} \ln \frac{R_{(n-i+1)} - R_{(n_q)}}{R_{(n-k)} - R_{(n_q)}}.$$

5.5 Gomes and Pestana's method

The improvement of (32) due to Gomes and Pestana (2007) takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k+1)} \exp \left[\overline{H}(k) \ln \left(\frac{k}{np} \right) \right],$$

where

$$\overline{H}(k) = H(k)\left[1 - \frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right],$$

$$H(k) = \frac{1}{k}\sum_{i=1}^{k} U_i,\ U_i = i\ln\left(\frac{R_{(n-i+1)}}{R_{(n-i)}}\right),$$

$$\widehat{\rho} = \min\left[0, \frac{3\left(T_n^{(\tau)}(k)-1\right)}{T_n^{(\tau)}(k)-3}\right],$$

$$T_n^{(\tau)}(k) = \begin{cases} \dfrac{\left(M_n^{(1)}(k)\right)^{\tau} - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}}, & \text{if } \tau \neq 0, \\ \dfrac{\ln\left(M_n^{(1)}(k)\right) - \frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right)}{\frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right) - \frac{1}{3}\ln\left(M_n^{(3)}(k)/6\right)}, & \text{if } \tau = 0, \end{cases}$$

$$M_n^{(j)}(k) = \frac{1}{k}\sum_{i=1}^{k}\left[\ln R_{(n-i+1)} - \ln R_{(n-k)}\right]^j,$$

$$\widehat{\beta} = \left(\frac{k}{n}\right)^{\widehat{\rho}}\frac{d_{\widehat{\rho}}(k)D_0(k) - D_{\widehat{\rho}}(k)}{d_{\widehat{\rho}}(k)D_{\widehat{\rho}}(k) - D_{2\widehat{\rho}}(k)},$$

$$d_\alpha(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha},\ D_\alpha(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha}U_i.$$

Here, τ is a tuning parameter. Under suitable conditions, Gomes and Pestana (2007) show further that

$$\frac{\sqrt{k}}{\ln k - \ln(np)}\left(\widehat{\mathrm{VaR}}_{1-p} - \mathrm{VaR}_{1-p}\right) \to N\left(0, \xi^2\right)$$

in distribution as $n \to \infty$.

5.6 Beirlant *et al.*'s method

The improvement of (32) due to Beirlant *et al.* (2004) takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k)}\left[\frac{k+1}{(n+1)p}\right]^{\widehat{\gamma}}\exp\left\{-\frac{\widehat{\gamma}\widehat{\beta}}{\widehat{\rho}}\left(\frac{n+1}{k+1}\right)^{\widehat{\rho}}\left[1 - \left(\frac{(n+1)p}{k+1}\right)^{-\widehat{\rho}}\right]\right\},$$

where $\widehat{\rho}$ is as given by Section 5.5, and

$$\widehat{\gamma} = \frac{1}{k}\sum_{i=1}^{k} i\ln\frac{R_{(n-i+1)}}{R_{(n-i)}},$$

$$\widehat{\beta} = \left(\frac{k}{n}\right)^{\widehat{\rho}}\frac{d_{\widehat{\rho}}(k)D_0(k) - D_{\widehat{\rho}}(k)}{d_{\widehat{\rho}}(k)D_{\widehat{\rho}}(k) - D_{2\widehat{\rho}}(k)},$$

$$d_\alpha(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha},\ D_\alpha(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha}U_i.$$

This estimator is shown to be consistent.

5.7 Caeiro and Gomes's method

Caeiro and Gomes (2008, 2009) propose several improvements on (32). The first of these takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k)} \left(\frac{k}{np} \right)^{\widehat{\gamma}} \left\{ 1 - \frac{\widehat{\gamma}\widehat{\beta}}{\widehat{\rho}} \left(\frac{n}{k} \right)^{\widehat{\gamma}} \left[1 - \left(\frac{(n+1)p}{k+1} \right)^{-\widehat{\rho}} \right] \right\},$$

where $\widehat{\rho}$ and $\widehat{\beta}$ are as given by Section 5.5, and $\widehat{\gamma}$ is as given by Section 5.6. The second of these takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n-k)} \left(\frac{k}{np} \right)^{\widehat{\gamma}} \exp \left\{ -\frac{\widehat{\gamma}\widehat{\beta}}{\widehat{\rho}} \left(\frac{n}{k} \right)^{\widehat{\gamma}} \left[1 - \left(\frac{(n+1)p}{k+1} \right)^{-\widehat{\rho}} \right] \right\},$$

where $\widehat{\rho}$ and $\widehat{\beta}$ are as given by Section 5.5, and $\widehat{\gamma}$ is as given by Section 5.6. The third of these takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = \frac{R_{(n-[k/2])} - R_{(n-k)}}{2^{\widehat{\gamma}} - 1} \left(\frac{k}{np} \right)^{\widehat{\gamma}} \left[1 - B_{1/2} \left(\widehat{\gamma}; \widehat{\rho}, \widehat{\beta} \right) \right],$$

where $\widehat{\rho}$ and $\widehat{\beta}$ are as given by Section 5.5, $\widehat{\gamma}$ is as given by Section 5.6, and $B_x(a,b)$ denotes the incomplete beta function defined by

$$B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1} dt.$$

The fourth of these takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = \frac{R_{(n-[k/2])} - R_{(n-k)}}{2^{\overline{H}(k)} - 1} \left(\frac{k}{np} \right)^{\overline{H}(k)} \left[1 - B_{1/2} \left(\overline{H}(k); \widehat{\rho}, \widehat{\beta} \right) \right],$$

where $\widehat{\rho}$, $\widehat{\beta}$ and $\overline{H}(k)$ are as given in Section 5.5. All of these estimators are shown to be consistent and asymptotically normal.

5.8 Figueiredo *et al.*'s method

The latest improvement of (32) is due to Figueiredo *et al.* (2012). It takes the expression

$$\widehat{\mathrm{VaR}}_{1-p} = R_{(n_q)} + \left(R_{(n-k)} - R_{(n_q)} \right) \left(\frac{k}{np} \right)^{\overline{H}_n},$$

where $n_q = [nq] + 1$ and

$$\overline{H}_n = H_n \left[1 - \frac{\widehat{\beta}(n/k)^{\widehat{\rho}}}{1 - \widehat{\rho}} \right]$$

with $\left(\widehat{\beta}, \widehat{\rho} \right)$ as defined in Section 5.5 and H_n as defined in Section 5.4.

5.9 Li *et al.*'s method

Let p_n be such that $p_n \to 0$ and $np_n \to q > 0$ as $n \to \infty$. Li *et al.* (2010) derive estimators for VaR_{1-p_n} for large n. They give the estimator

$$\widehat{\mathrm{VaR}}_{1-p_n} = \widehat{c}^{1/\widehat{\alpha}} p_n^{-1/\widehat{\alpha}} \left[1 + \widehat{\alpha}^{-1} \widehat{c}^{-\widehat{\beta}/\widehat{\alpha}} \widehat{d} p_n^{\widehat{\beta}/\widehat{\alpha}-1}\right],$$

where

$$\widehat{c} = \frac{\widehat{\alpha}\widehat{\beta}}{\widehat{\alpha}-\widehat{\beta}} R_{(n-k)}^{\widehat{\alpha}} \left[\frac{1}{\widehat{\beta}} - \frac{1}{k}\sum_{i=1}^{k} \ln \frac{R_{(n-i+1)}}{R_{(n-k)}}\right]$$

and

$$\widehat{d} = \frac{\widehat{\alpha}\widehat{\beta}}{\widehat{\beta}-\widehat{\alpha}} R_{(n-k)}^{\widehat{\beta}} \left[\frac{1}{\widehat{\alpha}} - \frac{1}{k}\sum_{i=1}^{k} \ln \frac{R_{(n-i+1)}}{R_{(n-k)}}\right],$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are the simultaneous solutions of the equations

$$\frac{1}{k}\sum_{i=1}^{k} Q_i^{-1}(\alpha, \beta) = 1$$

and

$$\frac{1}{k}\sum_{i=1}^{k} Q_i^{-1}(\alpha, \beta) \ln \frac{R_{(n-i+1)}}{R_{(n-k)}} = \frac{1}{\beta},$$

where

$$Q_i(\alpha, \beta) = \frac{\alpha}{\beta}\left[1 + \frac{\alpha\beta}{\alpha-\beta} H(\alpha)\right] \left(\frac{R_{(n-i+1)}}{R_{(n-k)}}\right)^{\beta-\alpha} - \frac{\alpha\beta}{\alpha-\beta} H(\alpha)$$

and

$$H(\alpha) = \frac{1}{\alpha} - \frac{1}{k}\sum_{i=1}^{k} \ln \frac{R_{(n-i+1)}}{R_{(n-k)}}.$$

Li *et al.* (2010) show under suitable conditions that

$$\frac{\sqrt{k}}{\ln k - \ln (np_n)} \left[\frac{\widehat{\mathrm{VaR}}_{1-p_n}}{F^{-1}(1-p_n)} - 1\right] \to N\left(0, \frac{\beta^4}{\alpha^2(\beta-\alpha)^4}\right)$$

in distribution as $n \to \infty$.

5.10 Gomes *et al.*'s method

Gomes *et al.* (2011) propose a bootstrap based method for computing VaR. The method can be described as follows:

1. For an observed sample, $r_1, r_2, \ldots, r_n$, compute $\widehat{\rho}$ as in Section 5.5 for $\tau = 0$ and $\tau = 1$;
2. Compute the median of $\widehat{\rho} = \widehat{\rho}(k)$, say M, for $k \in \left(\left[n^{0.995}\right], \left[n^{0.999}\right]\right)$. Also compute
$$I_\tau = \sum_{k \in ([n^{0.995}],[n^{0.999}])} \left(\widehat{\rho}(k) - M\right)^2$$
for $\tau = 0, 1$. Choose the tuning parameter, τ, as zero if $I_0 \leq I_1$ and as one otherwise;
3. Compute $\widehat{\rho} = \widehat{\rho}\left(\left[n^{0.999}\right]\right)$ and $\widehat{\beta} = \widehat{\beta}\left(\left[n^{0.999}\right]\right)$ using the formulas in Section 5.5 and the chosen tuning parameter;
4. Compute $\overline{H}(k)$, $k = 1, 2, \ldots, n-1$ in Section 5.5 with the estimates $\widehat{\rho}$ and $\widehat{\beta}$ in step 3;
5. Set $n_1 = \left[n^{0.95}\right]$ and $n_2 = \left[n_1^2/n\right] + 1$;
6. Generate B bootstrap samples $\left(r_1^*, r_2^*, \ldots, r_{n_2}^*\right)$ and $\left(r_1^*, r_2^*, \ldots, r_{n_2}^*, r_{n_2+1}^*, \ldots, r_{n_1}^*\right)$ from the empirical cdf of $r_1, r_2, \ldots, r_n$;
7. Compute $\overline{H}([k/2]) - \overline{H}(k)$ for the bootstrap samples in step 6. Let $t_{1,\ell}(k)$, $\ell = 1, 2, \ldots, B$ denote the estimates for the bootstrap samples of size n_1. Let $t_{2,\ell}(k)$, $\ell = 1, 2, \ldots, B$ denote the estimates for the bootstrap samples of size n_2;
8. Compute
$$\mathrm{MSE}_1(j, k) = \frac{1}{B} \sum_{i=1}^{B} t_{j,\ell}^2(k)$$
and
$$\mathrm{MSE}_2(j, k) = \ln^2\left(\frac{k}{np}\right) \mathrm{MSE}_1$$
for $j = 1, 2$ and $k = 1, 2, \ldots, n_j - 1$;
9. Compute
$$\widehat{P}(j) = \arg \min_{1 \leq k \leq n_j - 1} \mathrm{MSE}_1(j, k), \ \ \widehat{Q}(j) = \arg \min_{1 \leq k \leq n_j - 1} \mathrm{MSE}_2(j, k)$$
for $j = 1, 2$;
10. Compute
$$\widehat{k}_0 = \min\left\{n - 1, \left[\frac{\left(1 - 4^{\widehat{\rho}}\right)^{2/(1-\widehat{\rho})} \widehat{P}^2(1)}{\widehat{P}\left(\left[n_1^2/n\right] + 1\right)}\right] + 1\right\};$$
11. Compute $\overline{H}\left(\widehat{k}_0\right)$ with the estimates $\widehat{\rho}$ and $\widehat{\beta}$ in step 3;
12. Compute
$$\widehat{\ell}_0 = \min\left\{n - 1, \left[\frac{\left(1 - 4^{\widehat{\rho}}\right)^{2/(1-\widehat{\rho})} \widehat{Q}^2(1)}{\widehat{Q}\left(\left[n_1^2/n\right] + 1\right)}\right] + 1\right\};$$

13. Finally, estimate VaR_{1-p} as

$$\widehat{\mathrm{VaR}}_{1-p} = r_{\left(n-\widehat{\ell}_0+1\right)} \left(\frac{\widehat{\ell}_0}{np} \right)^{\overline{H}\left(\widehat{\ell}_0\right)}.$$

5.11 Wang's method

Wang (2010) combine the historical method in Section 4.1 with the generalized Pareto model in Section 5.2 to suggest the following estimator for VaR:

$$\widehat{\mathrm{VaR}}_p = \begin{cases} R_{(i)}, \ p \in ((i-1)/n, i/n], & \text{if } p < p_0, \\ u + \dfrac{\widehat{\sigma}}{\widehat{\xi}} \left[(1-p)^{-\widehat{\xi}} - 1 \right], & \text{if } p \geq p_0, \end{cases}$$

where $\widehat{\sigma}$ and $\widehat{\xi}$ are the maximum likelihood estimators of σ and ξ, respectively, and p_0 is an appropriately chosen threshold.

5.12 M-estimation method

Iqbal and Mukherjee (2012) provide an M-estimator for VaR. They consider a GARCH (1, 1) model for returns $R_1, \ldots, R_n$ specified by

$$R_t = \sigma_t \epsilon_t,$$

where

$$\sigma_t^2 = \omega_0 + \alpha_0 R_{t-1}^2 + \beta_0 \sigma_{t-1}^2 + \gamma_0 I \left(R_{t-1} < 0 \right) R_{t-1}^2$$

and ϵ_t are independent and identical random variables symmetric about zero. The unknown parameters are $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \gamma_0, \beta_0)^T$ and they belong to the parameter space, the set of all of all $\boldsymbol{\theta} = (\omega, \alpha, \gamma, \beta)^T$ with $\omega, \alpha, \beta > 0$, $\alpha + \gamma \geq 0$ and $\alpha + \beta + \gamma/2 < 1$. The M-estimator, say $\widehat{\boldsymbol{\theta}}_T$, is obtained by solving the equation

$$\sum_{t=1}^n \widehat{m}_t(\boldsymbol{\theta}) = 0,$$

where

$$\widehat{m}_t(\boldsymbol{\theta}) = (1/2) \left\{ 1 - H \left(R_t / \widehat{v}_t^{1/2}(\boldsymbol{\theta}) \right) \right\} \left[\dot{\widehat{v}}_t(\boldsymbol{\theta}) / \widehat{v}_t(\boldsymbol{\theta}) \right]$$

and

$$\widehat{v}_t(\boldsymbol{\theta}) = \frac{\omega}{1-\beta} + I(t \geq 2) \left\{ \alpha \sum_{j=1}^{t-1} \beta^{j-1} R_{t-j}^2 + \gamma \sum_{j=1}^{t-1} I \left(R_{t-j} < 0 \right) \beta^{j-1} R_{t-j}^2 \right\},$$

where $H(x) = x\psi(x)$ for some skew-symmetric function $\psi : \mathbb{R} \to \mathbb{R}$ and $\dot{\widehat{v}}_t(\boldsymbol{\theta})$ denotes the derivative of $\widehat{v}_t(\boldsymbol{\theta})$. Iqbal and Mukherjee (2012) propose that VaR_p can be estimated by $\widehat{v}_t^{1/2}\left(\widehat{\boldsymbol{\theta}}_T\right)$ multiplied by the $([np]+1)$th order statistic of $\left\{ R_t / \left\{ \widehat{v}_t \left(\widehat{\boldsymbol{\theta}}_T \right) \right\}^{1/2}, t = 2, 3, \ldots, n \right\}$.

5.13 Generalized Champernowne distribution

Champernowne generalized distribution was introduced by Buch-Larsen *et al.* (2005) as a model for insurance claims. A random variable, say X, is said to have this distribution if its cdf is

$$F(x) = \frac{(x+c)^a - c^a}{(x+c)^a + (M+c)^a - 2c^a} \tag{35}$$

for $x > 0$, where $\alpha > 0$, $c > 0$ and $M > 0$ is the median. Charpentier and Oulidi (2010) provide estimators of $\mathrm{VaR}_p(X)$ based on beta kernel quantile estimators. They suggest the following algorithm for estimating $\mathrm{VaR}_p(X)$:

- suppose $X_1, X_2, \ldots, X_n$ is a random sample from (35);
- let $\left(\widehat{M}, \widehat{\alpha}, \widehat{c}\right)$ denote the estimators of the parameters (M, α, c); if the method of maximum likelihood is used then the estimators can be obtained by maximizing the log-likelihood given by

$$\begin{aligned}\ln L(\alpha, M, c) &= n\left\{\ln a + \ln\left[(M+c)^a - c^a\right]\right\} + (a-1)\sum_{i=1}^{n} \ln\left(X_i + c\right) \\ &\quad -2\sum_{i=1}^{n} \ln\left[(X_i+c)^a + (M+c)^a - 2c^a\right];\end{aligned}$$

- transform $Y_i = F\left(X_i\right)$, where $F(\cdot)$ is given by (35) with (M, α, c) replaced by $\left(\widehat{M}, \widehat{\alpha}, \widehat{c}\right)$;
- estimate the cdf of $Y_1, Y_2, \ldots, Y_n$ as

$$\widehat{F}_{n,Y}(y) = \frac{\displaystyle\sum_{i=1}^{n} \int_0^y K_\beta\left(Y_i; b, t\right) dt}{\displaystyle\sum_{i=1}^{n} \int_0^1 K_\beta\left(Y_i; b, t\right) dt},$$

  where $K_\beta\left(\cdot; b, t\right)$ is given by either

$$K_\beta(u; b, t) = k_{t/b+1,(1-t)/b+1}(u) = \frac{u^{t/b}(1-u)^{(1-t)/b}}{B\left(t/b+1, (1-t)/b+1\right)}$$

  or

$$K_\beta(u; b, t) = \begin{cases} k_{t/b,(1-t)/b}(u), & \text{if } t \in [2b, 1-2b], \\ k_{\rho_b(t),(1-t)/b}(u), & \text{if } t \in [0, 2b), \\ k_{t/b,\rho_b(1-t)}(u), & \text{if } t \in (1-2b, 1], \end{cases}$$

  where $\rho_b(t) = 2b^2 + 2.5 - \sqrt{4b^4 + 6b^2 + 2.25 - t^2 - t/b}$;
- solve $\widehat{F}_{n,Y}(q) = p$ for q by using some Newton algorithm;
- estimate $\mathrm{VaR}_p(X)$ by $\widehat{\mathrm{VaR}}_p(X) = F^{-1}_{\widehat{M},\widehat{\alpha},\widehat{c}}(q)$.

6 Computer software

Software for computing value at risk and related quantities are widely available. Some software available from the R package (R Development Core Team, 2015) are:

- the package actuar due to Vincent Goulet, Sébastien Auclair, Christophe Dutang, Xavier Milhaud, Tommy Ouellet, Louis-Philippe Pouliot and Mathieu Pigeon. According to the authors, this package provides "additional actuarial science functionality, mostly in the fields of loss distributions, risk theory (including ruin theory), simulation of compound hierarchical models and credibility theory. The package also features 17 new probability laws commonly used in insurance, most notably heavy tailed distributions";
- the package ghyp due to David Luethi and Wolfgang Breymann. According to the authors, this package "provides detailed functionality for working with the univariate and multivariate Generalized Hyperbolic distribution and its special cases (Hyperbolic (hyp), Normal Inverse Gaussian (NIG), Variance Gamma (VG), skewed Student-t and Gaussian distribution). Especially, it contains fitting procedures, an AIC-based model selection routine, and functions for the computation of density, quantile, probability, random variates, expected shortfall and some portfolio optimization and plotting routines as well as the likelihood ratio test. In addition, it contains the Generalized Inverse Gaussian distribution";
- the package PerformanceAnalytics due to Peter Carl, Brian G. Peterson, Kris Boudt, and Eric Zivot. According to the authors, this package "aims to aid practitioners and researchers in utilizing the latest research in analysis of non-normal return streams. In general, it is most tested on return (rather than price) data on a regular scale, but most functions will work with irregular return data as well, and increasing numbers of functions will work with P & L or price data where possible";
- the package crp.CSFP due to Matthias Fischer, Kevin Jakob and Stefan Kolb. According to the authors, this package models "credit risks based on the concept of "CreditRisk+", First Boston Financial Products, 1997 and "CreditRisk+ in the Banking Industry", Gundlach & Lehrbass, Springer, 2003";
- the package fAssets due to Diethelm Wuertz and many others;
- the package fPortfolio due to the Rmetrics Core Team and Diethelm Wuertz;
- the package CreditMetrics due to Andreas Wittmann;
- the package fExtremes due to Diethelm Wuertz and many others;
- the package rugarch due to Alexios Ghalanos.

Some other software available for computing value at risk and related quantities are:

- the package EC - VaR due to Rho - Works Advanced Analytical Systems, http: // www . rhoworks . com / ecvar.php. According to the authors, this package implements "Conditional Value - at Risk, BetaVaR, Component VaR, traditional VaR and back testing measures for portfolios composed of stocks, currencies and indexes. An integrated optimizer can solve for the minimum CVaR portfolio based on market data, while a module capable of doing Stochastic Simulation allows to graph all feasible portfolios on CVaR - Return space. EC -

VaR employs a full - valuation historical - simulation approach to estimate Value - at - Risk and other risk indicators";

- the package VaR calculator and simulator due to Lapides Software Development Inc, http: // members.shaw.ca / lapides / var.html. According to the authors, this package implements "simple, robust, down to earth implementation of JP Morgan's RiskMetrics. Build to answer day to day needs of medium size organisations. Ideal for managers with focus on performance, end result and value. Allows one to calculate the value at risk of any portfolio. Calculates correlations, volatilities, valuates complex financial instruments and employs 2 methods: Analytical VaR calculation and Monte Carlo simulation";

- the package NtInsight for asset liability management due to Numerical Technologies, http: // www.numtech.com / financial-risk-management-software / . According to the producers, this package is used by "banks and insurance companies that handles massive and complicated financial simulation without oversimplified approximations. It provides asset/liability management professionals an integrated balance sheet management environment to monitor, analyze and manage liquidity risks, interest - rate risks, and earnings - at - risk";

- the package Protecht.ALM due to David Tattam and David Bergmark from the company Protecht, http: // www.protecht.com.au / risk-management-software / asset-liability-risk. According to the authors, this package provides "a full analysis and measurement of interest rate risk using variety of complimentary best practice measures such as VaR, PVBP and gap reporting. Also offers web based scenario and risk reporting for in - house reporting of exposures";

- the package ProFintm Risk due to the company Entrion, http:// www.entrion.com / software /. According to the authors, this package provides "a multi commodity Energy risk application that calculates VaR. The result is a system that minimizes the resource needed for daily risk calculator; which in turn, changes the focus from calculating risk to managing risk. VaR is calculated using the Delta - Normal method and this method calculates VaR using commodity prices and positions, volatilities, correlations and risk statistics. This application calculates volatilities and correlations using exponentially weighted historical prices";

- the package ALM Optimizer for asset allocation software due to Bob Korkie from the company RMKorkie & Associates, http://assetallocationsoftware.org/. According to the author, this package provides "risk and expected return of Markowitz efficient portfolios but extended to include recent technical advances on the definition of risk, adjustments for input bias, non normal distributions, and enhancements that allow for overlays, risk budgets, and investment horizon adjustments". Also the package "is a true Portfolio Optimizer with lognormal asset returns and user specified return or surplus optimization; optimization, risk, and rebalancing horizons; volatility, expected shortfall, and two value at risk (VaR) risk variables tailored to the risk horizon; and user specified portfolio constraints including risk budget constraints";

- the package QuantLib due to StatPro, http://www.statpro.com/portfolio-analytics-products/risk-management-software/. According to the authors, this package provides "access to a complete universe of pricing functions for risk assessment covering every asset class from equity, interest rate-linked products to mortgage-backed securities". The package has key features including "Multiple ex-ante risk measures including Value-at-Risk and CVaR (expected shortfall) at a variety of confidence levels, potential gain, volatility, tracking error and diversification grade. These measures are available in both absolute and relative basis";

- the package FinAnalytica's Cognity risk management due to FinAnalytica, http: // www . finanalytica . com / daily-risk-statistics / . According to the authors, this package provides "more accurate fat-tailed VaR estimates that do not suffer from the over-optimism of normal distributions. But Cognity goes beyond VaR and also provides the downside Expected Tail Loss (ETL) measure - the average or expected loss beyond VaR. As compared with volatility and VaR, ETL, also known as Conditional Value at Risk (CVaR) and Expected Shortfall (ES), is a highly informative and intuitive measure of extreme downside losses. By combining ETL with fat-tailed distributions, risk managers have access to the most accurate estimate of downside risk available today";

- the package CVaR Expert due to CVaR Expert Rho - Works Advanced Analytical Systems, http://www.rhoworks.com/software/detail/cvarxpert.htm. According to the authors, this package implements "total solution for measuring, analyzing and managing portfolio risk using historical VaR and CVaR methodologies. Traditional Value-at-Risk, Beta VaR, Component VaR, Conditional VaR and backtesting modules are incorporated on the current version, which lets you work with individual assets, portfolios, asset groups and multi currency investments (Enterprise Edition). An integrated optimizer can solve for the minimum CVaR portfolio based on market data and investor preferences, offering the best risk benchmark that can be produced. A module capable of doing Stochastic Simulation allows you to graph the CVaR-Return space for all feasible portfolios";

- the Kamakura Risk Manager software (KRM) due to ZSL Inc, http: // www.zsl.com / solutions / banking-finance / enterprise-risk-management-krm. According to the authors, KRM "completely integrates credit portfolio management, market risk management, asset and liability management, Basel II and other capital allocation technologies, transfer pricing, and performance measurement. KRM is also directly applicable to operational risk, total risk, and accounting and regulatory requirements using the same analytical engine, GUI and reporting, and its vision is that completely integrated risk solution based on common assumptions and methodologies. KRM offers, dynamic value at risk and expected shortfall, historical value at risk measurement, Monte Carlo value at risk measurement, etc";

- the package G@RCH 6, OxMetrics due to Timberlake Consultants Limited, http: // www . timberlake . co . uk / ?id=64#garch. According to the authors, the package is "dedicated to the estimation and forecasting of univariate ARCH-type models. G@RCH provides a user-friendly interface (with rolling menus) as well as some graphical features (through the OxMetrics graphical interface). G@RCH helps the financial analysis: value-at-risk, expected shortfall, backtesting (Kupiec LRT, dynamic quantile test); forecasting, realized volatility".

7 Conclusions

We have reviewed the current state of the most popular risk measure, value at risk, with emphasis on recent developments. We have reviewed ten of its general properties, including upper comonotonicity and multivariate extensions; thirty five of its parametric estimation methods, including time series, quantile regression and Bayesian methods; eight of its nonparametric estimation methods, including historical methods and bootstrapping; thirteen of its semiparametric estimation methods, including extreme value theory and M-estimation methods; twenty known computer software, including those based on the R platform.

This review could encourage further research with respect to measures of financial risk. Some open problems to address are: further multivariate extensions of risk measures and corresponding estimation methods; development of a comprehensive R package implementing a wide range of parametric, nonparametric and semiparametric estimation methods, no such packages are available to date; estimation based on nonparametric Bayesian methods; estimation methods suitable for big data; and so on.

References

[1] Araújo Santos, A., Fraga Alves, M. I. and Gomes, M. I. (2006). Peaks over random threshold methodology for tail index and quantile estimation. *Revstat*, **4**, 227-247.

[2] Arbia, G. (2002). Bivariate value-at-risk. *Statistica*, **62**, 231-247.

[3] Ardia, D. (2008). *Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications*. Springer Verlag, Berlin.

[4] Arneric, J., Jurun, E. and Pivac, S. (2008). Parametric forecasting of value at risk using heavy tailed distribution. In: *Proceedings of the 11th International Conference on Operational Research*, pp. 65-75.

[5] Bae, T. and Iscoe, I. (2012). Large-sample confidence intervals for risk measures of locationscale families. *Journal of Statistical Planning and Inference*, **142**, 2032-2046.

[6] Bahadur, R. (1966). A note on quantiles in large samples. *Annals of Mathematical Statistics*, **37**, 577-580.

[7] Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **74**, 3-30.

[8] Barone-Adesi, G., Giannopoulos, K. and Vosper, L. (1999). VaR without correlations for nonlinear portfolios. *Journal of Futures Markets*, **19**, 583-602.

[9] Beirlant, J., Figueiredo, F., Gomes, M. I. and Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators. *Journal of Statistical Planning and Inference*, **138**, 1851-1870.

[10] Beirlant, J., Herrmann, K. and Teugels, J. L. (2015). The estimation of the tail index. To appear in *Extreme Events in Finance*, edited by F. Longin, Wiley.

[11] Bi, G. and Giles, D. E. (2009). Modelling the financial risk associated with U.S. movie box office earnings. *Mathematics and Computers in Simulation*, **79**, 2759-2766.

[12] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press, Cambridge.

[13] Böcker, K. and Klüppelberg, C. (2005). Operational VaR: A closed-form approximation. *Risk*, **18**, 90-93.

[14] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **28**, 307-327.

[15] Bonga-Bonga, L. and Mutema, G. (2009). Volatility forecasting and value-at-risk estimation in emerging markets: The case of the stock market index portfolio in South Africa. *South African Journal of Economic and Management Sciences*, **12**, 401-411.

[16] Bouchaud, J. -P. and Potters, M. (2000). *Theory of Financial Risks: From Statistical Physics to Risk Management*. Cambridge University Press, Cambridge.

[17] Brazauskas, V. and Kleefeld, A. (2011). Folded and log-folded-t distributions as models for insurance loss data. *Scandinavian Actuarial Journal*, 59-74.

[18] Brummelhuis, R., Cordoba, A., Quintanilla, M. and Seco, L. (2002). Principal component value at risk. *Mathematical Finance*, **12**, 23-43.

[19] Buch-Larsen, T., Nielsen, J. P., Guillen, M. and Bolance, C. (2005). Kernel density estimation for heavy-tailed distribution using the Champernowne transformation. *Statistics*, **6**, 503-518.

[20] Caeiro, F. and Gomes, M. I. (2008). Minimum-variance reduced-bias tail index and high quantile estimation. *Revstat*, **6**, 1-20.

[21] Caeiro, F. and Gomes, M. I. (2009). Semi-parametric second-order reduced-bias high quantile estimation. *Test*, **18**, 392-413.

[22] Cai, Z. -Y., Xin, R. and Xiao, R. (2009). Value at risk management in multi-period supply inventory coordination. In: *Proceedings of the 2009 IEEE International Conference on e-Business Engineering*, pp. 335-339.

[23] Cakir, S. and Raei, F. (2007). Sukuk versus Eurobonds: Is there a difference in value-at-risk? IMF Working Paper WP/07/237.

[24] Capinski, M. and Zastawniak, T. (2011). *Mathematics for Finance*. Springer Verlag, London.

[25] Chan, F. (2009a). Modelling time-varying higher moments with maximum entropy density. *Mathematics and Computers in Simulation*, **79**, 2767-2778.

[26] Chan, F. (2009b). Forecasting value-at-risk using maximum entropy density. In: *Proceedings of the 18th World IMACS / MODSIM Congress*, pp. 1377-1383.

[27] Chan, N. H., Deng, S. -J., Peng, L. and Xia, Z. (2007). Interval estimation of value-at-risk based on GARCH models with heavy tailed innovations. *Journal of Econometrics*, **137**, 556-576.

[28] Chang, C. -S. (2011). A matrix-based VaR model for risk identification in power supply networks. *Applied Mathematical Modelling*, **35**, 4567-4574.

[29] Chang, K. L. (2011). The optimal value at risk hedging strategy under bivariate regime switching ARCH framework. *Applied Economics*, **43**, 2627-2640.

[30] Chang, Y. P., Hung, M. C. and Wu, Y. F. (2003). Nonparametric estimation for risk in value-at-risk estimator. *Communications in Statistics—Simulation and Computation*, **32**, 1041-1064.

[31] Charpentier, A. and Oulidi, A. (2010). Beta kernel quantile estimators of heavy-tailed loss distributions. *Statistics and Computing*, **20**, 35-55.

[32] Chen, S. X. and Tang, C. Y. (2005). Nonparametric inference of value-at-risk for dependent financial returns. *Journal of Financial Econometrics*, **3**, 227-255.

[33] Cheong, C. W. (2011). Univariate and multivariate value-at-risk: Application and implication in energy markets. *Communications in Statistics—Simulation and Computation*, **40**, 957-977.

[34] Cheong, C. W. and Isa, Z. (2011). Bivariate value-at-risk in the emerging Malaysian sectoral markets. *Journal of Interdisciplinary Mathematics*, **14**, 67-94.

[35] Chernozhukoy, V. and Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics*, **26**, 271-292.

[36] Cheung, K. C. (2009). Upper comonotonicity. *Insurance: Mathematics and Economics*, **45**, 35-40.

[37] Cohen, A. C. and Whitten, B. (1982). Modified maximum likelihood and modified moment estimators fort the three-parameter Weibull distribution. *Communications in Statistics—Theory and Methods*, **11**, 2631-2656.

[38] Cousin, A. and Bernardinoy, E. D. (2011). A multivariate extension of value-at-risk and conditional-tail-expectation. ArXiv: 1111.1349v1

[39] Dagum, C. (1977). A new model for personal income distribution: Specification and estimation. *Economie Appliquée*, **30**, 413-437.

[40] Dagum, C. (1980). The generation and distribution of income, the Lorenz curve and the Gini ratio. *Economie Appliquée*, **33**, 327-367.

[41] Dash, J. W. (2004). *Quantitative Finance and Risk Management: A Physicist's Approach*. World Scientific Publishing Company, River Edge, New Jersey.

[42] Degen, M., Lambrigger, D. D. and Segers, J. (2010). Risk concentration and diversification: Second-order properties. *Insurance: Mathematics and Economics*, **46**, 541-546.

[43] Dehlendorff, C., Kulahci, M., Merser, S. and Andersen, K. K. (2010). Conditional value at risk as a measure for waiting time in simulations of hospital units. *Quality Technology and Quantitative Management*, **7**, 321-336.

[44] Dekkers, A. L. M., Einmahl, J. H. J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, **17**, 1833-1855.

[45] Delbaen, F. (2000). *Coherent Risk Measures*. Scuola Normale Superiore, Classe di Scienze.

[46] Denis, L., Fernandez, B. and Meda, A. (2009). Estimation of value at risk and ruin probability for diffusion processes with jumps. *Mathematical Finance*, **19**, 281-302.

[47] Dionne, G., Duchesne, P. and Pacurar, M. (2009). Intraday value at risk (IVaR) using tick-by-tick data with application to the Toronto stock exchange. *Journal of Empirical Finance*, **16**, 777-792.

[48] Dixon, M. F., Chong, J. and Keutzer, K. (2012). Accelerating value-at-risk estimation on highly parallel architectures. *Concurrency and Computation—Practice and Experience*, **24**, 895-907.

[49] Domma, F. and Perri, P. F. (2009). Some developments on the log-Dagum distribution. *Statistical Methods and Applications*, **18**, 205-220.

[50] Dupacova, J., Hurt, J. and Stepan, J. (2002). *Stochastic Modeling in Economics and Finance*. Kluwer Academic Publishers, Dordrecht.

[51] Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

[52] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modeling Extremal Events for Insurance and Finance*. Springer Verlag, Berlin.

[53] Embrechts, P., Lambrigger, D. D. and Wuthrich, M. V. (2009). Multivariate extremes and the aggregation of dependent risks: Examples and counter-examples. *Extremes*, **12**, 107-127.

[54] Embrechts, P., Neslehova, J. and Wuthrich, M. V. (2009). Additivity properties for value-at-risk under Archimedean dependence and heavy-tailedness. *Insurance: Mathematics and Economics*, **44**, 164-169.

[55] Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, **22**, 367-381.

[56] Fang, J., Mannan, M., Ford, D., Logan, J. and Summers, A. (2007). Value at risk perspective on layers of protection analysis. *Process Safety and Environmental Protection*, **85**, 81-87.

[57] Fedor, M. (2010). Financial risk in pension funds: Application of value at risk methodology. Chapter 9 of *Pension Fund Risk Management: Financial and Actuarial Modelling*, pp. 185-209.

[58] Feng, Y. J. and Chen, A. D. (2001). The application of value-at-risk in project risk measurement. In: *Proceedings of the 2001 International Conference on Management Science and Engineering*, pp. 1747-1750.

[59] Fernandez, V. (2006). The CAPM and value at risk at different time scales. *International Review of Financial Analysis*, **15**, 203-219.

[60] Feuerverger, A. and Wong, A. C. M. (2000). Computation of value at risk for non-linear portfolios. *Journal of Risk*, **3**, 37-55.

[61] Figueiredo, F., Gomes, M. I., Henriques-Rodrigues, L. and Miranda, M. C. (2012). A computational study of a quasi-PORT methodology for VaR based on second-order reduced-bias estimation. *Journal of Statistical Computation and Simulation*, **82**, 587-602.

[62] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, **24**, 180-290.

[63] Fraga Alves, I. and Neves, C. (2015). Extreme value theory: An introductory overview. To appear in *Extreme Events in Finance*, edited by F. Longin, Wiley.

[64] Franke, J., Hardle, W. K. and Hafner, C. M. (2004). *Statistics of Financial Markets: An Introduction.* Springer Verlag, Berlin.

[65] Franke, J., Hardle, W. K. and Hafner, C. M. (2008). *Statistics of Financial Markets.* Springer Verlag, Berlin.

[66] Franke, J., Hardle, W. K. and Hafner, C. M. (2011). Copulae and value at risk. In: *Statistics of Financial Markets*, pp. 405-446.

[67] Gatti, S., Rigamonti, A. and Saita, F. (2007). Measuring value at risk in project finance transactions. *European Financial Management*, **13**, 135-158.

[68] Gebizlioglu, O. L., Senoglu, B. and Kantar, Y. M. (2011). Comparison of certain value-at-risk estimation methods for the two-parameter Weibull loss distribution. *Journal of Computational and Applied Mathematics*, **235**, 3304-3314.

[69] Göb, R. (2011). Estimating value at risk and conditional value at risk for count variables. *Quality and Reliability Engineering International*, **27**, 659-672.

[70] Gomes, M. I., Caeiro, F., Henriques-Rodrigues, L. and Manjunath, B. G. (2015). Bootstrap methods in statistics of extremes. To appear in *Extreme Events in Finance*, edited by F. Longin, Wiley.

[71] Gomes, M. I., Mendonca, S. and Pestana, D. (2011). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology. *Communications in Statistics—Theory and Methods*, **40**, 2946-2968.

[72] Gomes, M. I. and Pestana, D. (2007). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, **102**, 280-292.

[73] Gong, Z. and Li, D. (2008). Measurement of HIS stock index futures market risk based on value-at-risk. In: *Proceedings of the 15th International Conference on Industrial Engineering and Engineering Management*, pp. 1906-1911.

[74] Gouriéroux, C., Laurent, J. -P. and Scaillet, O. (2000). Sensitivity analysis of values at risk. *Journal of Empirical Finance*, **7**, 225-245.

[75] Guillen, M., Prieto, F. and Sarabia, J. M. (2011). Modelling losses and locating the tail with the Pareto positive stable distribution. *Insurance: Mathematics and Economics*, **49**, 454-461.

[76] Harrell, F. E. and Davis, C. E. (1982). A new distribution free quantile estimator. *Biometrika*, **69**, 635-640.

[77] Hartz, C., Mittnik, S. and Paolella, M. (2006). Accurate value-at-risk forecasting based on the (good old) normal-GARCH model. Center for Financial Studies (CFS), Working Paper Number 2006/23.

[78] Hassan, R., de Neufville, R. and McKinnon, D. (2005). Value-at-risk analysis for real options in complex engineered systems. In: *Proceedings of the International Conference on Systems, Man and Cybernetics*, pp. 3697-3704.

[79] He, K., Lai, K. K. and Xiang, G. (2012). Portfolio value at risk estimate for crude oil markets: A multivariate wavelet denoising approach. *Energies*, **5**, 1018-1043.

[80] He, K., Lai, K. K. and Yen, J. (2012). Ensemble forecasting of value at risk via multi resolution analysis based methodology in metals markets. *Expert Systems with Applications*, **39**, 4258-4267.

[81] He, K., Xie, C. and Lai, K. K. (2008). Estimating real estate value at risk using wavelet denoising and time series model. In: *Proceedings of the 8th International Conference on Computational Science*, part II, pp. 494-503.

[82] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, **13**, 331-341.

[83] Hill, I., Hill, R. and Holder, R. (1976). Fitting Johnson curves by moments. *Applied Statistics*, **25**, 180-192.

[84] Hong, L. J. (2011). Monte Carlo estimation of value-at-risk, conditional value-at-risk and their sensitivities. In: *Proceedings of the 2011 Winter Simulation Conference*, pp. 95-107.

[85] Huang, J. J., Lee, K. J., Liang, H. M. and Lin, W. F. (2009). Estimating value at risk of portfolio by conditional copula-GARCH method. *Insurance: Mathematics and Economics*, **45**, 315-324.

[86] Hughett, P. (1998). Error bounds for numerical inversion of a probability characteristic function. *SIAM Journal on Numerical Analysis*, **35**, 1368-1392.

[87] Hull, J. and White, A. (1998). Incorporating volatility updating into the historical simulation method for VaR. *Journal of Risk*, **1**, 5-19.

[88] Hürlimann, W. (2002). Analytical bounds for two value-at-risk functionals. *Astin Bulletin*, **32**, 235-265.

[89] Ibragimov, R. (2009). Portfolio diversification and value at risk under thick-tailedness. *Quantitative Finance*, **9**, 565-580.

[90] Ibragimov, R. and Walden, J. (2011). Value at risk and efficiency under dependence and heavy-tailedness: Models with common shocks. *Annals of Finance*, **7**, 285-318.

[91] Iqbal, F. and Mukherjee, K. (2012). A study of value-at-risk based on M-estimators of the conditional heteroscedastic models. *Journal of Forecasting*, **31**, 377-390.

[92] Jadhav, D. and Ramanathan, T. V. (2009). Parametric and non-parametric estimation of value-at-risk. *Journal of Risk Model Validation*, **3**, 51-71.

[93] Jakobsen, S. (1996). Measuring value-at-risk for mortgage backed securities. In: *Risk Management in Volatile Financial Markets*, volume 32, pp. 184-206.

[94] Jang, J. and Jho, J. H. (2007). Asymptotic super(sub)additivity of value-at-risk of regularly varying dependent variables. Preprint, MacQuarie University, Sydney, Australia.

[95] Janssen, J., Manca, R. and Volpe di Prignano, E. (2009). *Mathematical Finance*. John Wiley and Sons, Hoboken, New Jersey.

[96] Jaschke, S., Klüppelberg, C. and Lindner, A. (2004). Asymptotic behavior of tails and quantiles of quadratic forms of Gaussian vectors. *Journal of Multivariate Analysis*, **88**, 252-273.

[97] Jaworski, P. (2007). Bounds for value at risk for asymptotically dependent assets - the copula approach. In: *Proceedings of the 5th EUSFLAT Conference*, Ostrava, Czech Republic.

[98] Jaworski, P. (2008). Bounds for value at risk for multiasset portfolios. *Acta Physica Polonica*, A, **114**, 619-627.

[99] Jeong, S. -O. and Kang, K. -H. (2009). Nonparametric estimation of value-at-risk. *Journal of Applied Statistics*, **36**, 1225-1238.

[100] Jiménez, J. A. and Arunachalam, V. (2011). Using Tukey's g and h family of distributions to calculate value-at-risk and conditional value-at-risk. *Journal of Risk*, **13**, 95-116.

[101] Jin, C. and Ziobrowski, A. J. (2011). Using value-at-risk to estimate downside residential market risk. *Journal of Real Estate Research*, **33**, 389-413.

[102] Johnson, N. L. (1949). System of frequency curves generated by methods of translation. *Biometrika*, **36**, 149-176.

[103] Jorion, P. (2001). *Value at Risk: The New Benchmark for Managing Financial Risk*, second edition. McGraw-Hill, New York.

[104] Kaigh, W. D. and Lachenbruch, P. A. (1982). A generalized quantile estimator. *Communications in Statistics—Theory and Methods*, **11**, 2217-2238.

[105] Kaiser, M. J., Pulsipher, A. G., Darr, J., Singhal, A., Foster, T. and Vojjala, R. (2010). Catastrophic event modeling in the Gulf of Mexico II: Industry exposure and value at risk. *Energy Sources Part B—Economics Planning and Policy*, **5**, 147-154.

[106] Kaiser, M. J., Pulsipher, A. G., Singhal, A., Foster, T. and Vojjala, R. (2007). Industry exposure and value at risk storms in the Gulf of Mexico. *Oil and Gas Journal*, **105**, 36-42.

[107] Kamdem, J. S. (2005). Value-at-risk and expected shortfall for linear portfolios with elliptically distributed risk factors. *International Journal of Theoretical and Applied Finance*, **8**, doi: 10.1142/S0219024905003104

[108] Karandikar, R. G., Deshpande, N. R. and Khaparde, S. A. (2009). Modelling volatility clustering in electricity price return series for forecasting value at risk. *European Transactions on Electrical Power*, **19**, 15-38.

[109] Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2008). *Loss Models*. John Wiley and Sons, Hoboken, New Jersey.

[110] Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, **46**, 33-50.

[111] Koenker, R. and Portnoy, S. (1997). Quantile regression. Working Paper 97-0100, University of Illinois at Urbana-Champaign.

[112] Komunjer, I. (2007). Asymmetric power distribution: Theory and applications to risk measurement. *Journal of Applied Econometrics*, **22**, 891-921.

[113] Ku, Y. -H. H. and Wang, J. J. (2008). Estimating portfolio value-at-risk via dynamic conditional correlation MGARCH model - an empirical study on foreign exchange rates. *Applied Economics Letters*, **15**, 533-538.

[114] Kwon, C. (2011). Conditional value-at-risk model for hazardous materials transportation. In: *Proceedings of the 2011 Winter Simulation Conference*, pp. 1703-1709.

[115] Lai, T. L. and Xing, H. (2008). *Statistical Models and Methods for Financial Markets*. Springer Verlag, New York.

[116] Lee, J. and Locke, P. (2006). Dynamic trading value at risk: Futures floor trading. *Journal of Futures Markets*, **26**, 1217-1234.

[117] Li, D. Y., Peng, L. and Yang, J. P. (2010). Bias reduction for high quantiles. *Journal of Statistical Planning and Inference*, **140**, 2433-2441.

[118] Li, K., Yu, X. Y. and Gao, F. (2002). The validity analysis of value-at-risk technique in Chinese securities market. In: *Proceedings of the 2002 International Conference on Management Science and Engineering*, pp. 1518-1521.

[119] Lin, H. -Y. and Chen, A. -P. (2008). Application of dynamic financial time-series prediction on the interval artificial neural network approach with value-at-risk model. In: *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks*, pp. 3918-3925.

[120] Liu, C. C., Ryan, S. G. and Tan, H. (2004). How banks' value-at-risk disclosures predict their total and priced risk: Effects of bank technical sophistication and learning over time. *Review of Accounting Studies*, **9**, 265-294.

[121] Liu, R., Zhan, Y. R. and Lui, J. P. (2006). Estimating value at risk of a listed firm in China. In: *Proceedings of the 2006 International Conference on Machine Learning and Cybernetics*, pp. 2137-2141.

[122] Longin, F. (1996). The asymptotic distribution of extreme stock market returns. *Journal of Business*, **69**, 383-408.

[123] Longin, F. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance*, **24**, 1097-1130.

[124] Lu, G., Wen, F., Chung, C. Y. and Wong, K. P. (2007). Conditional value-at-risk based mid-term generation operation planning in electricity market environment. In: *Proceedings of the 2007 IEEE Congress on Evolutionary Computation*, pp. 2745-2750.

[125] Lu, Z. (2011). Modeling the yearly Value-at-Risk for operational risk in Chinese commercial banks. *Mathematics and Computers in Simulation*, **82**, 604-616.

[126] Mattedi, A. P., Ramos, F. M., Rosa, R. R. and Mantegna, R. N. (2004). Value-at-risk and Tsallis statistics: Risk analysis of the aerospace sector. *Physica A—Statistical Mechanics and Its Applications*, **344**, 554-561.

[127] Matthys, G., Delafosse, E., Guillou, A. and Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance: Mathematics and Economics*, **34**, 517-537.

[128] Mesfioui, M. and Quessy, J. F. (2005). Bounds on the value-at-risk for the sum of possibly dependent risks. *Insurance: Mathematics and Economics*, **37**, 135-151.

[129] Meucci, A. (2005). *Risk and Asset Allocation*. Springer Verlag, Berlin.

[130] Milwidsky, C. and Mare, E. (2010). Value at risk in the South African equity market: A view from the tails. *South African Journal of Economic and Management Sciences*, **13**, 345-361.

[131] Moix, P. -Y. (2001). *The Measurement of Market Risk: Modelling of Risk Factors, Asset Pricing, and Approximation of Portfolio Distributions*. Springer Verlag, Berlin.

[132] Mondlane, A. I. (2010). Value at risk in a volatile context of natural disaster risk. In: *Proceedings of the 10th International Multidisciplinary Scientific Geo-Conference*, pp. 277-284.

[133] Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer Verlag, New York.

[134] Odening, M. and Hinrichs, J. (2003). Using extreme value theory to estimate value-at-risk. *Agricultural Finance Review*, **63**, 55-73.

[135] Panning, W. H. (1999). The strategic uses of value at risk: Long-term capital management for property/casualty insurers. *North American Actuarial Journal*, **3**, 84-105.

[136] Pflug, G. Ch. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In: *Probabilistic Constrained Optimization: Methodology and Applications*, editor S. Uryasev, Kluwer, pp. 272-281.

[137] Pflug, G. Ch. and Romisch, W. (2007). *Modeling, Measuring and Managing Risk*. World Scientific Publishing Company, Hackensack, New Jersey.

[138] Piantadosi, J., Metcalfe, A. V. and Howlett, P. G. (2008). Stochastic dynamic programming (SDP) with a conditional value-at-risk (CVaR) criterion for management of storm-water. *Journal of Hydrology*, **348**, 320-329.

[139] Pickands III, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, **3**, 119-131.

[140] Plat, R. (2011). One-year value-at-risk for longevity and mortality. *Insurance: Mathematics and Economics*, **49**, 462-470.

[141] Pollard, M. (2007). Bayesian value-at-risk and the capital charge puzzle. http://www.apra.gov.au/AboutAPRA/WorkingAtAPRA/Documents/Pollard-M_Paper-for-APRA.pdf

[142] Porter, N. (2007). Revenue volatility and fiscal risks - An application of value-at-risk techniques to Hong Kong's fiscal policy. *Emerging Markets Finance and Trade*, **43**, 6-24.

[143] Prékopa, A. (2012). Multivariate value at risk and related topics. *Annals of Operations Research*, **193**, 49-69.

[144] Prescott, P. and Walden, A. T. (1990). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika*, **67**, 723-724.

[145] Puzanova, N., Siddiqui, S. and Trede, M. (2009). Approximate value at risk calculation for heterogeneous loan portfolios: Possible enhancements of the Basel II methodology. *Journal of Financial Stability*, **5**, 374-392.

[146] R Development Core Team (2015). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria.

[147] Rachev, S. T., Schwartz, E. and Khindanova, I. (2003). Stable modeling of market and credit value at risk. Chapter 7 of *Handbook of Heavy Tailed Distributions in Finance*, editor S. T. Rachev, pp. 249-328.

[148] Raunig, B. and Scheicher, M. (2011). A value-at-risk analysis of credit default swaps. *Journal of Risk*, **13**, 3-29.

[149] Resnick, S. I. (2007). *Heavy-Tail Phenomena*. Springer Verlag, New York.

[150] RiskMetrics Group (1996). *RiskMetrics-Technical Document*. Morgan J.P.

[151] Rockinger, M. and Jondeau, E. (2002). Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics*, **106**, 119-142.

[152] Ruppert, D. (2011). *Statistics and Data Analysis for Financial Engineering*. Springer Verlag, New York.

[153] Sarabia, J. M. and Prieto, F. (2009). The Pareto-positive stable distribution: A new descriptive method for city size data. *Physica A—Statistical Mechanics and Its Applications*, **388**, 4179-4191.

[154] Sheather, S. J. and Marron, J. S. (1990). Kernel quantile estimators. *Journal of the American Statistical Association*, **85**, 410-416.

[155] Simonato, J. -G. (2011). The performance of Johnson distributions for computing value at risk and expected shortfall. *Journal of Derivatives*, **19**, 7-24.

[156] Singh, M. K. (1997). Value at risk using principal components analysis. *Journal of Portfolio Management*, **24**, 101-112.

[157] Siven, J. V., Lins, J. T. and Szymkowiak-Have, A. (2009). Value-at-risk computation by Fourier inversion with explicit error bounds. *Finance Research Letters*, **6**, 95-105.

[158] Slim, S., Gammoudi, I. and Belkacem, L. (2012). Portfolio value at risk bounds using extreme value theory. *International Journal of Economics and Finance*, **4**, 204-215.

[159] Sriboonchitta, S., Wong, W. -K., Dhompongsa, S. and Nguyen, H. T. (2010). *Stochastic Dominance and Applications to Finance, Risk and Economics*. CRC Press, Boca Raton, Florida.

[160] Su, E. and Knowles, T. W. (2006). Asian Pacific stock market volatility modelling and value at risk analysis. *Emerging Markets Finance and Trade*, **42**, 18-62.

[161] Sun, X., Tang, L. and He, W. (2011). Exploring the value at risk of oil exporting country portfolio: An empirical analysis from the FSU region. *Procedia Computer Science*, **4**, 1675-1680.

[162] Taniai, H. and Taniguchi, M. (2008). Statistical estimation errors of VaR under ARCH returns. *Journal of Statistical Planning and Inference*, **138**, 3568-3577.

[163] Taniguchi, M., Hirukawa, J. and Tamaki, K. (2008). *Optimal Statistical Inference in Financial Engineering*. Chapman and Hall/CRC, Boca Raton, Florida.

[164] Tapiero, C. (2004). *Risk and Financial Management*. John Wiley and Sons, Hoboken, New Jersey.

[165] Tian, M. Z. and Chan, N. H. (2010). Saddle point approximation and volatility estimation of value-at-risk. *Statistica Sinica*, **20**, 1239-1256.

[166] Tiku, M. L. (1967). Estimating the mean and standard deviation from censored normal samples. *Biometrika*, **54**, 155-165.

[167] Tiku, M. L. (1968). Estimating the parameters of lognormal distribution from censored samples. *Journal of the American Statistical Association*, **63**, 134-140.

[168] Tiku, M. L. and Akkaya, A. D. (2004). *Robust Estimation and Hypothesis Testing*. New Age International, New Delhi.

[169] Trindade, A. A. and Zhu, Y. (2007). Approximating the distributions of estimators of financial risk under an asymmetric Laplace law. *Computational Statistics and Data Analysis*, **51**, 3433-3447.

[170] Trzpiot, G. and Ganczarek, A. (2006). Value at risk using the principal components analysis on the Polish power exchange. In: *From Data and Information Analysis to Knowledge Engineering*, pp. 550-557.

[171] Tsafack, G. (2009). Asymmetric dependence implications for extreme risk management. *Journal of Derivatives*, **17**, 7-20.

[172] Tsay, R. S. (2010). *Analysis of Financial Time Series*, third edition. John Wiley and Sons, Hoboken, New Jersey.

[173] Voit, J. (2001). *The Statistical Mechanics of Financial Markets*. Springer Verlag, Berlin.

[174] Walter, Ch. (2015). Jumps in financial modelling: Pitting the Black-Scholes model refinement programme against the Mandelbrot programme. To appear in *Extreme Events in Finance*, edited by F. Longin, Wiley.

[175] Wang, C. (2010). Wholesale price for supply chain coordination via conditional value-at-risk minimization. *Applied Mechanics and Materials*, **20-23**, 88-93.

[176] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, **73**, 812-815.

[177] Weng, H. and Trueck, S. (2011). Style analysis and value-at-risk of Asia-focused hedge funds. *Pacific—Basin Finance Journal*, **19**, 491-510.

[178] Wilson, W. W., Nganje, W. E. and Hawes, C. R. (2007). Value-at-risk in bakery procurement. *Review of Agricultural Economics*, **29**, 581-595.

[179] Xu, M. and Chen, F. Y. (2007). Tradeoff between expected reward and conditional value-at-risk criterion in newsvendor models. In: *Proceedings of the 2007 IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 1553-1557.

[180] Yamout, G. M., Hatfield, K. and Romeijn, H. E. (2007). Comparison of new conditional value-at-risk-based management models for optimal allocation of uncertain water supplies. *Water Resources Research*, **43**, W07430.

[181] Yan, D. and Gong, Z. (2009). Measurement of HIS stock index futures market risk based on value at risk. In: *Proceedings of the 2009 International Conference on Information Management, Innovation Management and Industrial Engineering*, volume 3, pp. 78-81.

[182] Yiu, K. F. C., Wang, S. Y. and Mak, K. L. (2008). Optimal portfolios under a value-at-risk constraint with applications to inventory control in supply chains. *Journal of Industrial and Management Optimization*, **4**, 81-94.

[183] Zhang, M. H. and Cheng, Q. S. (2005). An approach to VaR for capital markets with Gaussian mixture. *Applied Mathematics and Computation*, **168**, 1079-1085.

[184] Zhu, B. and Gao, Y. (2002). Value at risk and its use in the risk management of investment-linked household property insurance. In: *Proceedings of the 2002 International Conference on Management Science and Engineering*, pp. 1680-1683.