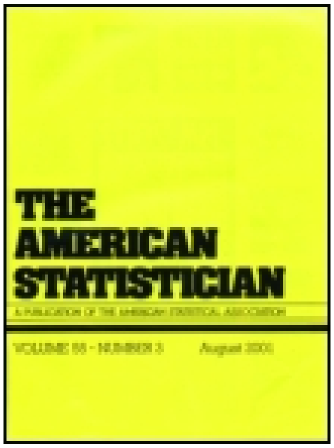

The American Statistician
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/utas20

A Few Counter Examples Useful in Teaching Central Limit Theorems

Subhash C. Bagui [a], Dulal K. Bhaumik [b] & K. L. Mehra [c]

[a] Department of Mathematics and Statistics, University of West Florida, Pensacola, FL, 32514

[b] Department of Biostatistics and Psychiatry, University of Illinois at Chicago, Chicago, IL, 60612

[c] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada

Accepted author version posted online: 21 Dec 2012.Published online: 20 Feb 2013.

To cite this article: Subhash C. Bagui, Dulal K. Bhaumik & K. L. Mehra (2013) A Few Counter Examples Useful in Teaching Central Limit Theorems, The American Statistician, 67:1, 49-56, DOI: 10.1080/00031305.2012.755361

To link to this article: http://dx.doi.org/10.1080/00031305.2012.755361

A Few Counter Examples Useful in Teaching Central Limit Theorems

Subhash C. BAGUI, Dulal K. BHAUMIK, and K. L. MEHRA

In probability theory, central limit theorems (CLTs), broadly speaking, state that the distribution of the sum of a sequence of random variables (r.v.'s), suitably normalized, converges to a normal distribution as their number n increases indefinitely. However, the preceding convergence in distribution holds only under certain conditions, depending on the underlying probabilistic nature of this sequence of r.v.'s. If some of the assumed conditions are violated, the convergence may or may not hold, or if it does, this convergence may be to a nonnormal distribution. We shall illustrate this via a few counter examples. While teaching CLTs at an advanced level, counter examples can serve as useful tools for explaining the true nature of these CLTs and the consequences when some of the assumptions made are violated.

KEY WORDS: Cauchy distribution; Laplace distribution; Logistic distribution; Sample mean; Sum of random variables; Uniform distribution.

1. INTRODUCTION

There is not a single central limit theorem (CLT), but rather an array of results concerned with sums of large numbers of random variables that, properly normalized, have limiting normal distributions. These CLTs constitute an important class of theorems in theory of probability and statistics. Their use in statistical applications is virtually endless: in devising large sample confidence intervals, tests of hypotheses, and a variety of other statistical techniques. At the beginning instructional level, CLTs often get introduced in a simplistic manner with a broad assertion that, for large n, the sample mean $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, of observations $X_1, X_2, \ldots, X_n$ in a random sample, behaves like a normal r.v., without much elaboration as to the assumptions under which this assertion holds. In applications also, CLTs are sometimes used without checking the validity of the assumptions on which they stand, leading in many situations to possibly erroneous statistical conclusions. It is usually at advanced teaching levels that any serious attention is paid toward explaining the assumptions fully. Although counter examples are used as direct evidence that a statement or a conjecture is false, they also provide insight into the roles played by different assumptions and which among these need to be strengthened in order to achieve a desired result. When discussing assumptions under which a CLT may hold in an observational or experimental setup, one may ask what might be the result if any of these assumptions are violated. Would $\bar{X}_n$ still converge in distribution, possibly to a nonnormal r.v.? If so, what is the distribution of the limiting r.v.? We examine these two questions with the help of a few counter examples. There are a wide variety of CLTs available in the literature; however, in this article we discuss mainly those for sums and weighted sums of independent and identically distributed (iid) r.v.'s.

Subhash C. Bagui, Department of Mathematics and Statistics, University of West Florida, Pensacola, FL 32514 (E-mail: *sbagui@uwf.edu*). Dulal K. Bhaumik, Department of Biostatistics and Psychiatry, University of Illinois at Chicago, Chicago, IL 60612 (E-mail: *dbhaumik@psych.uic.edu*). K. L. Mehra, Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada (E-mail: *kmehra@ualberta.ca*). The authors express their sincere thanks to the *TAS* Editor and the Associate Editor for their constructive criticism and suggestions, which helped considerably in improving the presentation of this article.

2. PRELIMINARIES

First we state a CLT that is the simplest theorem of its type dealing with sums of iid r.v.'s.

Theorem 2.1. Let $\{X_n : n \geq 1\}$ be a sequence of iid r.v.'s with mean μ, $-\infty < \mu < \infty$, and variance σ^2, $0 < \sigma^2 < \infty$, and set $S_n = \sum_{i=1}^{n} X_i$, $\bar{X}_n = [S_n/n]$ and

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}. \tag{2.1}$$

Then, $Z_n \xrightarrow{d} Z \sim N(0, 1)$, as $n \to \infty$.

The notation $\xrightarrow{d}$ stands for "convergence in distribution," $\sim$ stands for "distributed as" and $N(0, 1)$ for a normal r.v. with mean 0 and variance 1. In practice, the given theorem entails that, for large n, the distribution of $\bar{X}_n$ is approximately normal with mean μ and variance σ^2/n. For Theorem 2.1, see *Elements of Large-Sample Theory* (Lehmann 1999, p. 73). The French mathematician De Moivre was the first to prove a CLT. His work (De Moivre 1738) was extended by Laplace (1810) to sums of independent bounded r.v.'s. Later Lindeberg (1922) strengthened this CLT to unbounded r.v's, assuming among others simply finiteness of their variances. The following Theorem 2.2 is a particular version of it.

Theorem 2.2. Let $\{X_{ni} : 1 \leq i \leq n; n \geq 1\}$ be a triangular array of r.v.'s that are iid within each row with a common mean $E(X_{ni}) = \mu_n$, $-\infty < \mu_n < \infty$, and a common variance $\mathrm{var}(X_{ni}) = \sigma_n^2$, $0 < \sigma_n^2 < \infty$, for each $n = 1, 2, \ldots$ and $\{c_{ni} : 1 \leq i \leq n; n \geq 1\}$ a triangular array of finite constants, not all zeros within each row, $n = 1, 2, \ldots$. Define

$$Z_n = \frac{1}{B_n}\left(\sum_{i=1}^{n} c_{ni} X_{ni} - \mu_n \sum_{i=1}^{n} c_{ni}\right) = \frac{1}{B_n}\sum_{i=1}^{n} c_{ni}(X_{ni} - \mu_n), \tag{2.2}$$

where $B_n^2 = \text{var}(\sum_{i=1}^n c_{ni} X_{ni}) = \sigma_n^2 \sum_{i=1}^n c_{ni}^2$. Then, $Z_n \xrightarrow{d} Z \sim N(0, 1)$, as $n \to \infty$, provided that

$$\max_{1 \le i \le n} \left[c_{ni}^2 \Big/ \sum_{i=1}^n c_{ni}^2 \right] \to 0, \quad \text{as} \quad n \to \infty. \tag{2.3}$$

The condition (2.3) shall be referred to in the sequel as the "negligibility" condition. Also, note here that

$$E\left(\sum_{i=1}^n c_{ni} X_{ni}\right) = \mu_n \sum_{i=1}^n c_{ni},$$
$$\text{so that } E(Z_n) = 0 \text{ and } \text{var}(Z_n) = 1.$$

The next Theorem 2.2a extends Theorem 2.2 to the case when the triangular arrays $\{X_{ni} : 1 \le i \le n; n \ge 1\}$ consist of r.v.'s are independent within each row, not necessarily identically distributed, and satisfy an additional condition involving $(2 + \delta)$th $(\delta > 0)$ absolute moments of X_{ni}'s:

Theorem 2.2a. Let $\{X_{ni} : 1 \le i \le n; n \ge 1\}$ be a triangular array of r.v.'s that are independent within each row with means $E(X_{ni}) = \mu_{ni}, -\infty < \mu_{ni} < \infty$, and variances $\text{var}(X_{ni}) = \sigma_{ni}^2$, $0 \le \sigma_{ni}^2 < \infty$, not all zeros within each row, $n = 1, 2, \ldots$ and another triangular array $\{c_{ni} : 1 \le i \le n; n \ge 1\}$ of constants, not all zeros within each row, $n = 1, 2, \ldots$. Now define

$$Z_n = \frac{1}{B_n}\left(\sum_{i=1}^n c_{ni} X_{ni} - \sum_{i=1}^n c_{ni}\mu_{ni}\right)$$
$$= \frac{1}{B_n} \sum_{i=1}^n c_{ni}(X_{ni} - \mu_{ni}), \tag{2.2a}$$

where $B_n^2 = \text{var}(\sum_{i=1}^n c_{ni} X_{ni}) = \sum_{i=1}^n c_{ni}^2 \sigma_{ni}^2$, and assume that

$$\left(1/B_n^{2+\delta}\right) \sum_{i=1}^n |c_{ni}|^{2+\delta} E|X_{ni} - \mu_{ni}|^{2+\delta} = o(1), \tag{2.3a}$$

where the notation $\eta_n = o(\xi_n)$ stands for $(\eta_n/\xi_n) \to 0$ as $n \to \infty$. Then $Z_n \xrightarrow{d} Z \sim N(0, 1)$, *as* $n \to \infty$.

Remark 2.1. Theorems 2.2 and 2.2a are only special cases of a more general Lindeberg–Feller (LF) CLT for triangular arrays $\{Y_{ni} : i = 1, 2, \ldots, n; n \ge 1\}$ of r.v.'s that are independent within each row. This general LF Theorem gives a sufficient condition—referred to as the Lindeberg condition—which if satisfied by the triangular array, yields the asymptotic normality of the sum $S_n' = \sum_{i=1}^n Y_{ni}$ (see the Appendix; cf. Serfling 1980, pp. 31–32). In the setup of Theorem 2.2, it is easily seen that the "negligibility" condition (2.3) ensures for its defined triangular array $\{Y_{ni}\}$ (see the Appendix for definition) that this Lindeberg condition is satisfied. Similarly, in the setup of Theorem 2.2a, the same conclusion is ensured for its defined triangular array $\{Y_{ni}\}$ (see the Appendix for definition) by the assumed condition (2.3a) (Serfling 1980, Corollary 1.9.3, p. 32). Brief proofs of these assertions are given in the Appendix.

Theorems of the type 2.2 or 2.2a are useful to statisticians in many ways (Lehmann 1999, p. 102). For example, weighted sums of independent r.v.'s commonly occur in statistical applications as efficient estimators of unknown parameters. To check the asymptotic normality of such estimators, one can apply Theorem 2.2 or Theorem 2.2a, whichever is appropriate for the estimator under consideration. This asymptotic normality would be useful in large sample testing and derivation of large sample confidence intervals and other inferential procedures for these parameters. Ahead we discuss two such estimators: The first one (A) is that of the regression coefficient β in the standard linear regression model covered by Theorem 2.2 and the second (B) is that of the common mean μ based on n independent observations—not necessarily identically distributed—with known finite variances, which is covered by Theorem 2.2a.

(A) Application of Theorem 2.2: Estimation of a regression coefficient. Consider the simple linear regression model

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \quad (i = 1, 2, \ldots, n), \tag{2.4}$$

where $\varepsilon_1, \ldots, \varepsilon_n$ are iid r.v.'s with mean 0 and variance σ^2, $(0 < \sigma^2 < \infty)$, x_i's are (known) constants and α and β unknown parameters. The least squares estimator of β is given by

$$\hat{\beta}_n = \sum_{i=1}^n Y_i(x_i - \bar{x}_n) \Big/ \sum_{i=1}^n (x_i - \bar{x}_n)^2 = \sum_{i=1}^n c_{ni} Y_i,$$

where $c_{ni} = (x_i - \bar{x}_n)/\sum_{i=1}^n (x_i - \bar{x}_n)^2$, $1 \le i \le n$, $\bar{Y}_n = \sum_{i=1}^n Y_i/n$, and $\bar{x}_n = \sum_{i=1}^n x_i/n$. Because $E(Y_i) \equiv \mu_i = \alpha + \beta x_i$, $\sum_{i=1}^n c_{ni} = 0$ and $\sum_{i=1}^n c_{ni} x_i = 1$, it follows that $E(\hat{\beta}_n) = \sum_{i=1}^n c_{ni}(\alpha + \beta x_i) = \beta$ and $\text{var}(\hat{\beta}_n) = \sigma^2(\sum_{i=1}^n c_{ni}^2) = \sigma^2/\sum_{i=1}^n (x_i - \bar{x}_n)^2$. Since, $(Y_i - \mu_i)$, $i = 1, 2, \ldots, n$ are iid, by Theorem 2.2 we obtain

$$Z_n = (\hat{\beta}_n - \beta) \Big/ \sqrt{\text{var}(\hat{\beta}_n)} = \sum_{i=1}^n c_{ni}(Y_i - \mu_i) \Big/ \sigma \sqrt{\sum_{i=1}^n c_{ni}^2}$$
$$= \sum_{i=1}^n \left[(x_i - \bar{x}_n) \Big/ \sqrt{\sum_{i=1}^n (x_i - \bar{x}_n)^2} \right]$$
$$\times [(Y_i - \mu_i)/\sigma] \xrightarrow{d} Z \sim N(0, 1), \tag{2.5}$$

as $n \to \infty$, provided the "negligibility" condition (2.3) holds, namely, that

$$\left[\max_{1 \le i \le n} (x_i - \bar{x}_n)^2 \Big/ \sum_{i=1}^n (x_i - \bar{x}_n)^2 \right]$$
$$= \left[\max_{1 \le i \le n} c_{ni}^2 \Big/ \sum_{i=1}^n c_{ni}^2 \right] \to 0, \tag{2.6}$$

as $n \to \infty$. The condition (2.6) is satisfied, for example, if $x_j = j\lambda$ for a constant $-\infty < \lambda < \infty$. □

(B) *Application of Theorem* 2.2a: *Estimation of a common mean*. Suppose that $X_1, X_2, \ldots, X_n$ are independent with common mean $E(X_i) = \mu$, $-\infty < \mu < \infty$, and (known) variances var $(X_i) = \sigma_i^2$, $0 < \sigma_i^2 < \infty$, $i = 1, 2, \ldots, n$. Consider the (unbiased) weighted linear estimator $\delta_n = \sum_{i=1}^n w_{ni} X_i$ of μ, where $w_{ni} = (1/\sigma_i^2)/\sum_{i=1}^n (1/\sigma_i^2)$, $1 \le i \le n$. If X_i's are normally distributed, then δ_n is the maximum likelihood estimator (MLE) of the common mean μ. In general, the estimator δ_n is a (generalized) least squares estimator (LSE) of μ, for it minimizes $\Lambda = \sum_{i=1}^{n} (1/\sigma_i^2)(X_i - \mu)^2$, a weighted sum of squares of distances of observations X_i, $i = 1, 2, \ldots, n$, from their common mean μ. The variance B_n^2 of δ_n is $B_n^2 = \sum_{i=1}^{n} w_{ni}^2 \sigma_i^2 = 1/\sum_{i=1}^{n} (1/\sigma_i^2)$. Therefore, δ_n is consistent for μ if $\sum_{i=1}^{n} (1/\sigma_i^2) \to \infty$, as $n \to \infty$. For the asymptotic normality of the estimator δ_n, we shall assume that

$$\sum_{i=1}^{n} \left(1/\sigma_i^{2+\delta}\right) E\,|Z_i|^{2+\delta} = o\left(\left[\sum_{i=1}^{n} \sigma_i^{-2}\right]^{(1+\delta/2)}\right), \quad (2.7)$$

as $n \to \infty$, where $Z_i = [(X_i - \mu)/\sigma_i]$, $i = 1, 2, \ldots, n$. Setting $c_{ni} = w_{ni}\sigma_i = [\sigma_i^{-1}/(\sum_{i=1}^{n} \sigma_i^{-2})]$, $i = 1, 2, \ldots, n$, $\delta_n - \mu = \sum_{i=1}^{n} w_{ni}(X_i - \mu) = \sum_{i=1}^{n} c_{ni} Z_i$, where Z_i's are independent r.v.'s with means 0 and variances 1 but possibly not identically distributed. We shall now apply Theorem 2.2a to deduce the asymptotic normality of δ_n: First note that, since, $c_{ni} = [\sigma_i^{-1}/(\sum_{i=1}^{n} \sigma_i^{-2})]$, $i = 1, 2, \ldots, n$, the LHS of the assumed condition (2.3a) in the present application equals

$$\left(\sum_{i=1}^{n} \sigma_i^{-2}\right)^{(1+\delta/2)} \sum_{i=1}^{n} c_{ni}^{2+\delta} E|Z_i|^{2+\delta}$$
$$= \left(\sum_{i=1}^{n} \sigma_i^{-2}\right)^{-(1+\delta/2)} \sum_{i=1}^{n} \sigma_i^{-(2+\delta)} E|Z_i|^{2+\delta} = o(1),$$

as $n \to \infty$, on account of assumption (2.7). Accordingly, by Theorem 2.2a

$$Z_n = (\delta_n - \mu)/\sqrt{\mathrm{var}(\delta_n)}$$
$$= \sqrt{\sum_{i=1}^{n} \left(\sigma_i^{-2}\right)} \sum_{i=1}^{n} w_{ni}\,(X_i - \mu) \xrightarrow{d} Z \sim N(0, 1),$$

as $n \to \infty$. This completes the proof of asymptotic normality of δ_n. Also, note here that the assumption (2.7) clearly implies (in the present setup) that $(\sum_{i=1}^{n} \sigma_i^{-2})$ must $\to \infty$, as $n \to \infty$. Accordingly, (2.7) also ensures that $B_n^2 \to 0$, as $n \to \infty$, and thereby the consistency of the of the estimator δ_n. □

Remark 2.2. The conditions (2.6) and (2.7) in applications (A) and (B) are realistic conditions and can hold in practice, at least approximately, in a variety of situations. We demonstrate this in the Appendix with three illustrations I_1, I_2, and I_3, the first one pertaining to condition (2.6) and the next two to condition (2.7).

3. COUNTER EXAMPLES

In this section, we consider five counter examples. The first two are with respect to Theorem 2.1, a CLT for sums of iid r.v.'s, the next two with respect to Theorem 2.2, a CLT for weighted sums of triangular arrays of r.v.'s that are iid within each row, and the last counter example is with respect to Theorem 2.2a dealing with weighted sums of triangular arrays of r.v.'s that are independent within each row but not necessarily identically distributed.

Theorem 2.1 asserts that the proximity of $\bar{X}_n'$s distribution to that of a normal r.v. for large samples requires additional assumptions beyond the observations being simply iid, namely, the requirement of existence and finiteness of their means and variances. One may ask the question as to what happens to the large sample distribution of $\bar{X}_n$ when these conditions do not hold. We shall investigate this in the following two examples:

Example 3.1. Let $X_i, i = 1, 2, \ldots, n$ be iid observations from a Cauchy $C(\mu, \delta)$ distribution with density

$$f_X(x) = \frac{\delta}{\pi} \frac{1}{\delta^2 + (x - \mu)^2}, \; -\infty < x < \infty, \quad (3.1)$$

where μ, $-\infty < \mu < \infty$, and δ, $\delta > 0$, are the location and scale parameters, respectively. The Cauchy family (3.1) is an important class of densities arising in many applied contexts in Physics, Statistics, and other related disciplines. It is well known that the mean of a Cauchy distribution does not exist and its second moment is infinite. Therefore, Theorem 2.1 for the asymptotic normality of $\bar{X}_n$ does not apply. It turns out that the distribution of $\bar{X}_n$ remains the same as that of a single observation X_1 in this case, regardless of the value of n (see the Appendix for a proof of this assertion). Hence, trivially, $\bar{X}_n$ converges in distribution to a $C(\mu, \delta)$ r.v., instead of being asymptotically normal.

Even more extreme examples exist in which means and variances do not exist, with tails of distributions so heavy that $\bar{X}_n$'s are more variable than their respective single observations. An example of such a distribution follows.

Example 3.2. Consider a random variable X with density

$$f_X(x) = (1/\sqrt{2\pi x^3})e^{-1/(2x)}, \quad x > 0. \quad (3.2)$$

The mean and variance do not exist for this density. Here, in fact, distributionally $X = 1/Z^2$, where $Z \sim N(0, 1)$. The given density is a special case of an inverse-gamma family of densities (Johnson, Kotz, and Balakrishnan 1995, p. 401; Casella and Berger 2002, p. 51) given, for parameters $\alpha > 0$, $\beta > 0$, by

$$f_X(x; \alpha, \beta) = (\beta^\alpha/\Gamma(\alpha))x^{-\alpha-1}e^{-(\beta/x)}, \quad x > 0, \quad (3.3)$$

with the shape and scale parameters α and β, respectively, each set equal to $\frac{1}{2}$. (It is easy to see that if $Y \sim$ the gamma density $g_Y(y; \alpha, \beta) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}e^{-\beta x}$, $x > 0$, then the density (3.3), as the name inverse gamma suggests, is that of the r.v. $X = (1/Y)$.) The mean and variance of density (3.3) are, respectively, $[\beta/(\alpha - 1)]$ and $[\beta^2/(\alpha - 1)^2(\alpha - 2)]$, which exist only if $\alpha > 1$ and $\alpha > 2$, respectively. As for the density (3.2), its mean and variance do not exist, as can be easily checked, so that the conditions of Theorem 2.1 are not satisfied for this example. It turns out that $\bar{X}_n$ in this example has the same distribution as that of nX_1 for each n. (A sketch of the proof is given in the Appendix.) Clearly then, $\bar{X}_n$ is much more variable than a single observation X_1 and increases by an order of n, instead of converging in distribution to a limiting r.v., that is, certainly not to a normal r.v., as $n \to \infty$.

It should be mentioned that the density (3.2) corresponds to an important class of distributions in applications. It is the distribution of first passage times in a one-dimensional Brownian motion. It is also the limiting distribution of normalized average $[\bar{X}_n/n]$ of waiting times $X_1, X_2, \ldots, X_n$ of successive returns to the origin in a symmetric random walk and is typical of limiting distributions, without expectation, of such waiting time averages of recurrence of events in many physical and economic processes (Feller 1968, p. 90, 246).

Theorem 2.2 for weighted sums requires the "negligibility" condition (2.3) to be satisfied in order for Z_n to converge in distribution to a normal r.v., as $n \to \infty$. The "negligibility" condition implies that no summand $[c_{ni}(X_{ni} - \mu_n)/B_n]$ in the sum Z_n in (2.2) contributes excessively to the variance of Z_n. It is apparent that the "negligibility" condition (2.3) rules out situations where Z_n depends only on a negligible proportion of the summands. This occurs, for example, (in an extreme case, say) when $c_{n1} = c_{n2} = \cdots = c_{nk} = 1$, for some fixed k, and $c_{n,k+1} = \cdots = c_{nn} = 0$. Under this setup, $\max_{1\le i\le n} \{c_{ni}^2/\sum_{i=1}^n c_{ni}^2\} = 1/k$, which does not go to 0 as $n \to \infty$. So, the "negligibility" condition does not hold. On the other hand, if $c_{ni} = i$, then $\sum_{i=1}^n c_{ni}^2 = \frac{n(n+1)(2n+1)}{6}$, so that

$$\max_{1\le i\le n} \left[c_{ni}^2 \Big/ \sum_{i=1}^n c_{ni}^2 \right] = \frac{6n^2}{n(n+1)(2n+1)} \to 0, \quad \text{as } n \to \infty. \tag{3.4}$$

Thus, in view of (3.4), the "negligibility" condition holds here. Consequently, by Theorem 2.2 it follows that

$$Z_n = \frac{1}{B_n} \sum_{i=1}^n i(X_{ni} - \mu_n) \xrightarrow{d} Z \sim N(0,1), \quad \text{as} \quad n \to \infty.$$

In case the "negligibility" condition of Theorem 2.2 does not hold, it may still occur that Z_n converges to a limiting r.v., but the distribution of the limiting r.v. will typically be nonnormal. We shall illustrate this via two counter examples namely, Examples 3.3 and 3.4 where the "negligibility" condition is not satisfied. In Examples 3.3 and 3.4 ahead, we have only simple sequences $\{X_i : 1 \le i \le n; n \ge 1\}$ of iid r.v.'s with corresponding simple sequences $\{c_i : 1 \le i \le n; n \ge 1\}$ of weights, instead of triangular arrays of iid (within each row) r.v.'s and their corresponding constant weights. Theorem 2.2, indeed, does cover this simpler case.

In Example 3.3 ahead, the weighted sum U_n equals the decimal representation, up to n places, of a number chosen in the unit interval (0, 1) randomly, that is, $U_n = 0.X_1X_2X_3\cdots X_n = \sum_{i=1}^n (X_i/10^i)$, with weights $c_i = 1/10^i$, $1 \le i \le n$, which do not satisfy the "negligibility" condition (2.3) of Theorem 2.2. Its asymptotic normality, thus, cannot be concluded using this theorem.

Example 3.3. Let $X_1, X_2, \ldots, X_n$ be n iid r.v.'s with

$$P(X_i = j) = \frac{1}{10} \quad \text{for} \quad j = 0, 1, 2, \ldots, 9.$$

Define $U_n = \sum_{i=1}^n (X_i/10^i) = \sum_{i=1}^n c_i X_i$ with $c_i = 1/10^i$. Note that $\sum_{i=1}^n c_i^2 = \sum_{i=1}^n (1/10^{2i}) = \sum_{i=1}^n a^i = a(1-a^n)/(1-a)$, where $a = 1/10^2$, which yields $\max_{1\le i\le n} [c_i^2/\sum_{i=1}^n c_i^2] = \frac{(0.01)(99)}{(1-(0.01)^n)} \to 0.99$, as $n \to \infty$. Thus, the "negligibility" condition (2.3) of Theorem 2.2 stands violated and the theorem does not apply. The sequence $\{U_n\}$, however, does converge in distribution to a random variable. Intuitively, the limiting distribution of U_n should be $U(0, 1)$ and, indeed, it turns out to be so. To see this, note that the moment generating function (m.g.f.) of X_1 is given by

$$M_{X_1}(t) = E(e^{tX_1}) = \sum_{x=0}^{9} e^{tx}\left(\frac{1}{10}\right) = \frac{e^{10t}-1}{10(e^t-1)},$$

so that the m.g.f. of U_n is

$$\begin{aligned} M_{U_n}(t) = E(e^{tU_n}) &= \prod_{i=1}^n M_{X_1}\left(\frac{t}{10^i}\right) \\ &= \prod_{i=1}^n \frac{(e^{t/10^{i-1}}-1)}{10(e^{t/10^i}-1)} = \frac{(e^t-1)}{10^n(e^{t/10^n}-1)}. \end{aligned} \tag{3.5}$$

As $n \to \infty$, $10^n(e^{t/10^n} - 1) = 10^n(1 + \frac{t}{10^n} + \frac{t^2}{2(10)^{2n}} + \cdots - 1) = t + \frac{t^2}{2(10)^n} + \cdots \to t$. Therefore, from (3.5) we obtain

$$\lim_{n\to\infty} M_{U_n}(t) = \frac{(e^t-1)}{t} = \int_0^1 e^{tx}dx = M_U(t), \tag{3.6}$$

the m.g.f. of $U \sim$ Uniform(0, 1) r.v. The Equation (3.6) is a criterion for convergence in distribution (Rao 1966, p. 83; Casella and Berger 2002, p. 66, 235) and implies that $U_n \xrightarrow{d} U \sim U(0,1)$, as $n \to \infty$. In this example, the weights do not satisfy the "negligibility" condition of Theorem 2.2, however, U_n does converge in distribution to a limiting r.v. that has a nonnormal distribution.

It would be instructive in the preceding context to add that, since, $E(U_n) = \frac{1}{2}[1-(0.1)^n] \to \frac{1}{2}$ and $\text{var}(U_n) = \sigma_n^2(U_n) = \frac{1}{12}[1-(0.1)^{2n}] \to \frac{1}{12}$, as $n \to \infty$, $\frac{1}{2}$ and $\frac{1}{12}$ being, respectively, the mean $E(U)$ and $\text{var}(U) = \sigma^2(U)$ of a $U(0,1)$ r.v., the normalized U_n and U r.v.'s, namely, $U_n^* = (U_n - E(U_n))/\sigma_n(U_n)$ and $U^* = (U - E(U))/\sigma(U)$, respectively, also satisfy $U_n^* \xrightarrow{d} U^*$, as $n \to \infty$. Thus, the normalized U_n converges in distribution to a normalized $U(0, 1)$ r.v.

The following example is also quite interesting. In this example, we consider a weighted sum of iid Laplace r.v.'s with weights not satisfying the "negligibility" condition (2.3), so that Theorem 2.2 is not applicable. However, this weighted sum does converge in distribution, as $n \to \infty$, but to a logistic r.v.:

Example 3.4. (Knight 2000). Let $X_1, X_2, \ldots, X_n$ be n iid r.v.'s with common density

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < \infty. \tag{3.7}$$

This is known as the "standard" Laplace density, a member of the Laplace family of symmetric densities $f_{\mu,\lambda}(x) = (1/2\lambda)e^{-|x-\mu|/\lambda}$, $-\infty < x < \infty$, with the location parameter μ, $-\infty < \mu < \infty$, and the scale parameter $\lambda > 0$, set equal to 0 and 1, respectively. It can be clearly seen that $E(X_1) = 0$, $\text{var}(X_1) = 2$, and its m.g.f. evaluates to

$$M_{X_1}(t) = E[e^{tX_1}] = \frac{1}{2}\int_{-\infty}^{\infty} e^{tx}e^{-|x|}dx$$
$$= \frac{1}{2}\frac{1}{(1-t)} + \frac{1}{2}\frac{1}{(1+t)} = \frac{1}{1-t^2} \quad \text{for } |t| < 1. \quad (3.8)$$

Now define

$$L_n = \sum_{k=1}^{n}(X_k/k) = \sum_{k=1}^{n} c_k X_k, \text{ with } c_k = 1/k. \quad (3.9)$$

Since, $\sum_{k=1}^{n} c_k^2 = \sum_{k=1}^{n}(1/k^2)$, by Lemma A.1 in the Appendix, $\max_{1\le i\le n}[c_i^2/\sum_{i=1}^{n} c_i^2] \to [1/\sum_{i=1}^{\infty}(1/k^2)] = \frac{6}{\pi^2}$, as $n \to \infty$, so that the "negligibility" condition (2.3) for the weighted sum L_n in (3.9) does not hold. Nevertheless, as $n \to \infty$, L_n still converges in distribution to a r.v. as we shall see ahead. Since, X_i's are iid, the m.g.f. of L_n in view of (3.8), is given by

$$M_{L_n}(t) = E(e^{tL_n}) = \prod_{k=1}^{n} M_{X_1}(t/k) = \prod_{k=1}^{n}\left(\frac{k^2}{k^2-t^2}\right),$$
$$\to \prod_{k=1}^{\infty}\left(\frac{k^2}{k^2-t^2}\right) = \Gamma(1+t)\Gamma(1-t) \text{ for } |t| < 1, \quad (3.10)$$

as $n \to \infty$, the last equality in (3.10) holding by virtue of Lemma A.3 in the Appendix. In fact, the RHS expression of (3.10) is the m.g.f. of the Logistic $L(0, 1)$ density $f_L(x) = e^x/(1+e^x)^2$, $-\infty < x < \infty$. To see this note that in (3.11) ahead, after making the substitution $u = [e^x/(1+e^x)]$ for the variable of integration and then using the Beta integral, we have

$$M_L(t) = \int_{-\infty}^{\infty} e^{tx}\frac{e^x}{(1+e^x)^2}dx = \int_0^1 u^t(1-u)^{-t}du$$
$$= \frac{\Gamma(1+t)\Gamma(1-t)}{\Gamma(2)} = \Gamma(1+t)\cdot\Gamma(1-t), \quad (3.11)$$

so that from (3.10) and (3.11), we obtain

$$M_{L_n}(t) \to \Gamma(1+t)\cdot\Gamma(1-t) = M_L(t), \ |t| < 1. \quad (3.12)$$

Thus, we have $L_n \xrightarrow{d} L$, as $n \to \infty$ where the r.v. $L \sim L(0, 1)$ with density f_L defined already. This density is that of the "standard" Logistic distribution, a member of the logistic family of symmetric densities $f_{\mu,\lambda}(x) = \frac{1}{\lambda}\cdot e^{(\frac{x-\mu}{\lambda})}/(1+e^{(\frac{x-\mu}{\lambda})})^2$, $-\infty < x < \infty$, with location and scale parameters μ, $-\infty < \mu < \infty$, and $\lambda(> 0)$ set equal to 0 and 1, respectively. The mean and variance of the "standard" logistic density are 0 and $\pi^2/3$, respectively (Hájek and Šidák 1967, p. 125). In this example, the "negligibility" condition is again not satisfied, but the sum L_n of (3.9) converges in distribution to a limiting r.v. with the Logistic $L(0, 1)$ distribution.

Here again, it would be instructive to add that, since, $E(L_n) = E(L) = 0$ and $\text{var}(L_n) = \sigma_n^2(L_n) = 2\sum_{k=1}^{n}(1/k^2) \to \pi^2/3$ as $n \to \infty$, $\pi^2/3$ being the $\text{var}(L) = \sigma^2(L)$ of the logistic $L(0, 1)$ distribution, the normalized L_n and L r.v.'s, namely, $L_n^* = L_n/\sigma_n(L_n)$ and $L^* = L/\sigma(L)$, respectively, also satisfy $L_n^* \xrightarrow{d} L^*$, as $n \to \infty$.

Remark 3.1. For an iid sequence $\{X_n : n \ge 1\}$ of Laplace r.v.'s with a nonzero finite mean $E(X_1) = \mu$, the same variance $\text{var}(X_1) = 2$ and density $f_X(x) = \frac{1}{2}e^{-\frac{1}{2}|x-\mu|}$, $-\infty < x < \infty$, L_n without normalization does not converge in distribution to a r.v., as $n \to \infty$. To see this, first note that, since, $\sum_{k=1}^{n}(1/k) \to \infty$ as $n \to \infty$, the mean $E(L_n) = \mu\sum_{k=1}^{n}(1/k) \to -\infty$ or $+\infty$, depending on whether $\mu < 0$ or > 0, and further that the m.g.f. of L_n (see (3.9)), namely,

$$M_{L_n}(t) = E(e^{tL_n}) = \prod_{k=1}^{n} E\left(e^{(t/k)X_1}\right)$$
$$= \frac{1}{2}\prod_{k=1}^{n}\int e^{(t/k)x}e^{-|x-\mu|}dx = \frac{1}{2}\prod_{k=1}^{n}\int e^{(t/k)(y+\mu)}e^{-|y|}dy$$
$$= e^{t\mu\sum_{k=1}^{n}(1/k)}\prod_{k=1}^{n}\left(\frac{k^2}{k^2-t^2}\right) \to 0 \text{ or } +\infty, \quad (3.13)$$

by (3.10), depending on whether $\mu < 0$ or > 0. Thus, on account of (3.13), L_n does not converge in distribution to a proper r.v.

However, if we set $L_n' = [L_n - E(L_n)] = \sum_{k=1}^{n}(X_k'/k)$, where $X_k' = [X_k - E(X_k)]$, which $\sim \text{Lap}(0, 1)$, the standard Laplace r.v. with density (3.7), then as in Example 3.4, L_n' converges in distribution to a standard logistic $L(0, 1)$ r.v., as $n \to \infty$.

The following Example 3.5 provides a counter example when the conditions of existence and finiteness of means and variances in Theorem 2.2a are violated.

Example 3.5. Let $\{X_{ni} : 1 \le i \le n, n \ge 1\}$ be a triangular array of Cauchy r.v.'s that are independent within each row with location and scale parameters, respectively, given by μ_{ni}, $-\infty \le \mu_{ni} \le \infty$, and δ_{ni}, $0 < \delta_{ni} < \infty$, $1 \le i \le n$, $n = 1, 2, \ldots$ and let $\bar{X}_n = [\sum_{i=1}^{n} X_{ni}/n]$ denote the mean of their nth row. Since, the moments of Cauchy r.v.'s either do not exist or are infinite, it is clear that Theorem 2.2a is not applicable to conclude a limiting distribution for $\bar{X}_n$, since, the conditions of this theorem stand violated. The convergence in distribution, however, of $\bar{X}_n$—normalized or not—does take place as we shall see. It is shown in the Appendix that $\bar{X}_n \xrightarrow{d} X$, a Cauchy $C(\mu, \delta)$ r.v., provided $\bar{\mu}_n = [\sum_{i=1}^{n}\mu_{ni}/n] \to \mu$, $-\infty < \mu < \infty$, and $\bar{\delta}_n = [\sum_{i=1}^{n}\delta_{ni}/n] \to \delta$, $0 < \delta < \infty$, as $n \to \infty$. Thus, in this case, $\bar{X}_n \xrightarrow{d}$ to a nonnormal r.v. as $n \to \infty$.

For further study of interesting counter examples in probability theory and statistics, see *Counter Examples in Probability and Statistics* (Romano and Siegel 1986) and *Elements of Large-Sample Theory* (Lehmann 1999).

4. CONCLUDING REMARKS

In this article, we have presented a few interesting examples that illustrate what may happen when the assumptions of CLTs are violated. We showed via these examples that even with the violation of assumptions—either that of the existence of moments or of the "negligibility" condition—the convergence in distribution of the (normalized) sum Z_n in the given theorems can still happen but typically not to a normal r.v. This answers partially the question as to the consequences of violation of a CLTs assumptions.

Examples of this article can be used in graduate classes when teaching CLTs. These examples will help students appreciate the use of CLTs and also enhance their understanding of these theorems. The students and teachers will find these counter examples instructive and challenging as they are based on important results, lemmas, and theorems. Finally, this article may stimulate some pedagogical interest in the area of CLTs.

APPENDIX

The following results were used at various stages in the presentation of counter examples in Section 3.

Lemma A.1.

$$\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}\,.$$

The result may be found in the book *Calculus A New Horizon* (Anton 1999, p. 643, 646).

Lemma A.2. The Euler limiting form of Gamma function is given by

$$\Gamma(z)=\lim_{n\to\infty}\frac{n!n^z}{z(z+1)\cdots(z+n)}. \tag{A.1}$$

The given Euler form of Gamma function can be found in the book Handbook of Complex Variables (Krantz 1999, P. 156).

Lemma A.3. For any real $|t|<1$, $\prod_{k=1}^{\infty}\frac{k^2}{(k^2-t^2)}=\Gamma(1-t)\Gamma(1+t)$.

Proof. We can express that

$$\prod_{k=1}^{\infty}\frac{k^2}{(k^2-t^2)}=\lim_{n\to\infty}\prod_{k=1}^{n}\frac{k^2}{(k^2-t^2)}=\lim_{n\to\infty}n^t\prod_{k=1}^{n}\frac{k}{(k+t)} \times\lim_{n\to\infty}n^{-t}\prod_{k=1}^{n}\frac{k}{(k-t)}. \tag{A.2}$$

Also, in view of Lemma A.2, we have for the two products on the right of (A.2) that

$$\begin{aligned}\lim_{n\to\infty}n^t\prod_{k=1}^{n}\frac{k}{(k+t)}&=t\lim_{n\to\infty}\frac{n!n^t}{t(t+1)\cdots(t+n)}\\&=t\Gamma(t)=\Gamma(1+t),\\\lim_{n\to\infty}n^{-t}\prod_{k=1}^{n}\frac{k}{(k-t)}&=(-t)\lim_{n\to\infty}\frac{n!n^{-t}}{(-t)(-t+1)\cdots(-t+n)}\\&=(-t)\Gamma(-t)=\Gamma(1-t).\end{aligned} \tag{A.3}$$

The proof of the lemma now follows from Equations (A.2) and (A.3). □

Ahead, we give the three illustrations I_1, I_2, and I_3, referred to in Remark 2.2. The first one pertains to the satisfaction of condition (2.6) and the latter two to that of condition (2.7):

I_1. Suppose that in Application (A), $x_j=\lambda j^\kappa+\gamma$ for $j=1,2,\ldots,n$ and some constants λ, γ $(-\infty<\lambda,\ \gamma<\infty)$ and $\kappa>0$. Then, using the result that $\sum_{j=1}^{n} j^\kappa\approx[n^{\kappa+1}/(\kappa+1)]$ for large n (Feller 1968, p. 255), it can be easily verified that $[\max_{1\le j\le n}(x_j-\bar{x})^2/\sum_{j=1}^{n}(x_j-\bar{x})^2]=O(1/n)$, where the notation $\eta_n=O(\xi_n)$ stands for $[|\eta_n/\xi_n|]\le M$ for some positive real number M, as $n\to\infty$. Thus, the required condition (2.6) is satisfied.

I_2. Suppose for the sequence $\{X_j:1\le j\le n;n\ge1\}$ in Application (B) that (i) $E|Z_j|^{2+\delta}\le\Delta$ for some finite $\Delta>0$ and that (ii) $\sigma_j^2=\lambda j^\xi$ for some constants λ and ξ satisfying $\lambda>0$ and $(2/(2+\delta))<\xi\le1$, with δ the same as in (2.7). Then, since, for $[(2+\delta)\xi/2]>1$ and $\xi\le1$, $[\sum_{i=1}^{n} j^{-(\xi(2+\delta)/2)}]\to a$ constant and $[\sum_{j=1}^{n} j^{-\xi}]\to\infty$ as $n\to\infty$, it follows that

$$\left[\sum_{j=1}^{n}\sigma_j^{-(2+\delta)}E\left|Z_j\right|^{2+\delta}\Big/\left(\sum_{j=1}^{n}\sigma_j^{-2}\right)^{1+\delta/2}\right]\le\Delta\left[\sum_{j=1}^{n}j^{-[\xi(2+\delta)/2]}\Big/\left(\sum_{j=1}^{n}j^{-\xi}\right)^{1+\delta/2}\right]=o(1), \tag{2.8}$$

as $n\to\infty$. The inequality (2.8) establishes (2.7) for the sequence $\{X_j:1\le j\le n;n\ge1\}$ in this illustration.

I_3. Suppose now instead that in Application (B), (i) the sequence $\{X_j:1\le j\le n;n\ge1\}$ of observations consists of sample means $\{\bar{Y}_j:1\le j\le n;n\ge1\}$ based on independent iid samples of varying sizes k_j, $j=1,2,\ldots,n$, from the same population with mean μ, $-\infty<\mu<\infty$, variance σ^2, $0<\sigma^2<\infty$ and a finite third absolute moment μ_3. Then the common mean of X_j's $(=\bar{Y}_j'\text{s})$ is μ and the variances $\sigma_j^2=\sigma^2/k_j$, $j=1,2,\ldots,n$. (ii) Suppose that for some constants $\nu>0$ and $\eta>2$, $k_j=(\nu\vee j^{[\delta/(\eta(2+\delta))]})$, where $(a\vee b)=\max(a,b)$, $j=1,2,\ldots,n$. Under the preceding two assumptions, the condition (2.7) holds. To see this, first observe that under (i), $E|X_j-\mu|^{2+\delta}\le[E|X_j-\mu|^3]^{(2+\delta)/3}\le(\mu_3+\mu_2\mu_1+\mu_1^3)^{(2+\delta)/3}\le\Delta$ for some constant $\Delta>0$, where μ_k is kth order absolute moment about μ, $k=1,2,3$. Thus, in view of assumptions (i) and (ii), it follows that

$$\begin{aligned}&\left[\sum_{j=1}^{n}\left(\sigma_j^{-2}\right)^{2+\delta}E\left|X_j-\mu\right|^{2+\delta}\Big/\left(\sum_{j=1}^{n}\sigma_j^{-2}\right)^{1+\delta/2}\right]\\&\le\Delta\left[\sum_{j=1}^{n}k_j^{2+\delta}\Big/\sigma^{2+\delta}\left(\sum_{j=1}^{n}k_j\right)^{1+\delta/2}\right]\\&\le\Delta[(\nu^{2+\delta}\vee n^{\delta/\eta})/n^{\delta/2}]\le[\Delta/n^{\delta[(\eta-2)/2\eta]}]=o(1),\end{aligned} \tag{2.9}$$

as $n\to\infty$, where we have used in (2.9) the inequalities $1\le k_j\le k_j^{2+\delta}\le[\eta^{2+\delta}\vee n^{\delta/\eta}]$ and $\eta>2$, the last inequality in (2.9) holding for sufficiently large n. Thus, (2.9) establishes (2.7) in this case.

Proof of Example 3.1. Let $X_1, X_2, \ldots, X_n$ be an iid sample from a Cauchy $C(\mu, \delta)$ distribution defined by the density (3.1). We need to establish that $\bar{X}_n =^d X_1$ for each n, which we shall show via the characteristic function route. As a first step, we show that the characteristic function of a $C(0, 1)$ r.v. Z evaluates to

$$\varphi_Z(t) = E(e^{itZ}) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx = e^{-|t|}, \quad -\infty < t < \infty. \tag{A.4}$$

To see this evaluation, consider a r.v. Y with the "standard" Laplace density $f_Y(y) = \frac{1}{2}e^{-|y|}$, $-\infty < y < \infty$. Using the same steps as for calculating the m.g.f. (3.8), we obtain $\varphi_Y(t) = \frac{1}{1+t^2}$, $-\infty < t < \infty$, which on account of the Inversion Theorem (Rao 1966, p. 86; Feller 1971, p. 509) yields $\frac{1}{2\pi}\int_{-\infty}^{\infty}[e^{-ity}\varphi_Y(t)]dt = f_Y(y)$, or equivalently that $\frac{1}{\pi}\int_{-\infty}^{\infty}[e^{-ity}/(1+t^2)]dt = e^{-|y|}, -\infty < y < \infty$. By changing the sign of the variable of integration t and then interchanging the variables t and y, we arrive at the Equation (A.4). Since a $C(\mu, \delta)$ r.v. X can be expressed (distributionally) as $X =^d \delta Z + \mu$, where Z is a $C(0, 1)$ r.v., the characteristic function of X, in view of (A.4), calculates to

$$\begin{aligned}\varphi_X(t) &= E(e^{itX}) = E\left[e^{it(\delta Z+\mu)}\right] = e^{i\mu t}\varphi_Z(\delta t)\\ &= e^{i\mu t-\delta|t|}, \quad -\infty < t < \infty.\end{aligned} \tag{A.5}$$

We can now, using (A.5), derive the characteristic function of $\bar{X}_n$ as

$$\begin{aligned}\varphi_{\bar{X}_n}(t) &= E\left(e^{it\bar{X}_n}\right) = \prod_{k=1}^{n} E\left(e^{i\frac{t}{n}X_k}\right) = \left(\varphi_X\left(\frac{t}{n}\right)\right)^n = (e^{i\mu\frac{t}{n}-\delta|\frac{t}{n}|})^n\\ &= e^{i\mu t-\delta|t|} = \varphi_{X_1}(t), \quad -\infty < t < \infty.\end{aligned} \tag{A.6}$$

The Equation (A.6) and the Inversion Theorem together imply the desired result $\bar{X}_n =^d X_1$. The proof is complete. □

Proof of Example 3.2. Let the density function (3.2) and the corresponding distribution function be denoted by $f_{(1)}(x)$ and $F_{(1)}(x)$, respectively. Now consider a more general family of densities by taking $X = \gamma^2/Z^2$, where $Z \sim N(0, 1)$ with distribution function Φ and $0 < \gamma < \infty$. The corresponding distribution and the density functions are given by $F_{(\gamma)}(x) = 2[1 - \Phi(\gamma/\sqrt{x})], x > 0$ and $f_{(\gamma)}(x) = (1/\sqrt{2\pi x^3})\gamma e^{-\gamma^2/2x}$, $x > 0$, respectively. (Note from the inverse-gamma density definition (3.3) that the density $f_{(\gamma)}$ in this example is, in fact, a special case of (3.3) with $\alpha = 1/2$ and $\beta = (\gamma^2/2)$.) It can be shown that the preceding family is closed under convolutions, so that $f_{(\gamma)} * f_{(\lambda)} = f_{(\gamma+\lambda)}$. Accordingly, if $X_1, X_2, \ldots, X_n$ are independent r.v.'s with density $f_{(1)}(x)$, $x > 0$, then the preceding equation coupled with an induction argument yields that the density of $S_n = \sum_{i=1}^n X_i$ is $f_{(n)}(x)$, $x > 0$; or equivalently that $P[(\bar{X}_n/n) \le x] = P[S_n \le n^2 x] = F_{(n)}(n^2 x) = 2[1 - \Phi(n/\sqrt{n^2 x})] = 2[1 - \Phi(1/\sqrt{x})] = F_{(1)}(x)$. The last equation implies that $\bar{X}_n =^d nX_1$, so that as $n \to \infty$, $\bar{X}_n$ tends to increase by an order of n (Feller 1971, p. 52; Romano and Siegel 1986, pp. 59–60). □

Proof of Example 3.5. In distributional sense, each X_{nj} in the rectangular array in this example can be expressed as $(\delta_{nj}X_j + \mu_{nj})$, $1 \le j \le n$, $n = 1, 2, \ldots,$ with $X_1, X_2, \ldots, X_n$ as iid $C(0, 1)$ r.v.'s. Since, as shown in the proof of Example 3.1, the characteristic function of X_j a $C(0, 1)$ r.v. is given by $\varphi_{X_j}(t) = e^{-|t|}$, $-\infty < t < \infty$, the characteristic function of each X_{nj} evaluates to

$$\begin{aligned}\varphi_{X_{nj}}(t) &= E\{e^{itX_{nj}}\} = e^{it\mu_{nj}} \cdot \varphi_{X_j}(\delta_{nj}t)\\ &= e^{it\mu_{nj}-\delta_{nj}|t|}, -\infty < t < \infty,\end{aligned} \tag{A.7}$$

so that from (A.7) we obtain the characteristic function of $\bar{X}_n$ as

$$\varphi_{\bar{X}_n}(t) = \prod_{j=1}^{n} e^{it(\mu_{nj}/n)-|t|(\delta_{nj}/n)} = e^{it\bar{\mu}_n-\bar{\delta}_n|t|} \to e^{it\mu-\delta|t|}, \quad -\infty < t < \infty, \tag{A.8}$$

as $n \to \infty$. The expression on the RHS of (A.8) is the characteristic function of a $C(\mu, \delta)$ r.v. The convergence (A.8), being a criterion for convergence in distribution (Serfling 1980, p. 16; Casella and Berger 2002, p. 84, 235), implies that $\bar{X}_n \xrightarrow{d} X$, where X is a $C(\mu, \delta)$. The proof is complete. □

Deduction of Theorems 2.2 and 2.2a from the LF Theorem. First we state the *LF Theorem*. Let $\{X_{ni} : 1 \le i \le n; n \ge 1\}$ be a triangular array of r.v.'s that are independent within each row, $n = 1, 2, \ldots,$ with finite means $E(X_{ni}) = \mu_{ni}$ and variances $\mathrm{var}(X_{ni}) = \sigma_{ni}^2$, $i = 1, 2, \ldots, n$. Define $Y_{ni} = (X_{ni} - \mu_{ni})/B_n$, where $B_n^2 = \mathrm{var}(\sum_{i=1}^n X_{ni}) = \sum_{i=1}^n \sigma_{ni}^2$, assuming that $0 < B_n < \infty$, $n = 1, 2, \ldots$. Then, as $n \to \infty$, the "uniform asymptotic negligibility" (UAN) condition for the array $\{Y_{ni} : 1 \le i \le n; n \ge 1\}$, namely, that $Y_{ni} \xrightarrow{p} 0$, uniformly in $1 \le i \le n$ (i.e., $\max_{1\le i\le n} P[|Y_{ni}| \ge \varepsilon] \to 0$ as $n \to \infty$, for any $\varepsilon > 0$, however small), and that

$$Z_n = \sum_{i=1}^{n} Y_{ni} \xrightarrow{d} Z \sim N(0, 1), \tag{A.9}$$

both together hold if and only if, for every $\varepsilon > 0$,

$$\sum_{i=1}^{n} E\left\{Y_{ni}^2 \cdot I_{[|Y_{ni}|\ge\varepsilon]}\right\} \to 0, \tag{A.10}$$

as $n \to \infty$ (Lindeberg condition).

Deduction of Theorem 2.2. To see that under the setup of Theorem 2.2, the Lindeberg condition (A.10) is satisfied, note that in this case, upon setting $Z_{ni} = (X_{ni} - \mu_n)/\sigma_n$ with mean 0 and variance 1, $B_n^2 = \mathrm{var}(\sum_{i=1}^n c_{ni}X_{ni}) = \sigma_n^2\sum_{i=1}^n c_{ni}^2$ and $Y_{ni} = c_{ni}Z_{ni}/\sqrt{\sum_{i=1}^n c_{ni}^2}$, $1 \le i \le n$, so that for any $\varepsilon > 0$ and $\eta_n = \max_{1\le i\le n}|c_{ni}|/\sqrt{\sum_{i=1}^n c_{ni}^2}$,

$$\begin{aligned}\sum_{i=1}^{n} E\left\{Y_{ni}^2 I_{[|Y_{ni}|\ge\varepsilon]}\right\} &= \sum_{i=1}^{n}\left[c_{ni}^2 \Big/ \sum_{i=1}^{n} c_{ni}^2\right]\\ &\quad\times E\left\{Z_{ni}^2 I_{[(|c_{ni}Z_{ni}|/\sum_{i=1}^n c_{ni}^2)\ge\varepsilon]}\right\}\\ &\le E\left\{Z_{n1}^2 I_{[\eta_n|Z_{n1}|\ge\varepsilon]}\right\} \to 0,\end{aligned} \tag{A.11}$$

since, $\eta_n \to 0$, as $n \to \infty$, on the account of the "negligibility" condition (2.3) of Theorem 2.2. This establishes the required Lindeberg condition (A.10) of the preceding LF Theorem and, consequently, Theorem 2.2 follows.

It is worth mentioning that the "UAN" condition for the triangular array $\{Y_{ni} : 1 \le i \le n, n \ge 1\}$ in Theorem 2.2—a consequence of (A.11) and the "if" part of the LF Theorem—follows directly also from the "negligibility" condition (2.3), since

$$\max_{1\le i\le n} P[|Y_{ni}| \ge \varepsilon] \le \max_{1\le i\le n} \{\mathrm{var}(Y_{ni})/\varepsilon^2\} = \varepsilon^{-2} \max_{1\le i\le n} \left\{ c_{ni}^2 \Big/ \sum_{i=1}^n c_{ni}^2 \right\} = \varepsilon^{-2}\eta_n^2 \to 0,$$

as $n \to \infty$. Thus, $Y_{ni} \xrightarrow{p} 0$, as $n \to \infty$, uniformly in $1 \le i \le n$.

Remark A.1. In the deduction of Theorem 2.2, we have shown that the "negligibility" condition (2.3) implies the Lindeberg condition (A.10), from which Theorem 2.2 follows in view of the preceding LF Theorem. In fact, in Theorem 2.2 the "negligibility" condition is equivalent to the Lindeberg condition. To see the reverse implication that (A.10) also implies (2.3), note that for given $\varepsilon > 0$, however small, $[c_{ni}^2/\sum_{i=1}^n c_{ni}^2] = \mathrm{var}(Y_{ni}) \le E[|Y_{ni}|^2 I_{[|Y_{ni}|\ge\varepsilon]}] + \varepsilon^2$, so that

$$\max_{1\le i\le n} \left[c_{ni}^2 \Big/ \sum_{i=1}^n c_{ni}^2 \right] \le \sum_{i=1}^n E[|Y_{ni}|^2 I_{[|Y_{ni}|\ge\varepsilon]}] + \varepsilon^2, \quad \text{(A.12)}$$

with the Lindeberg condition (A.10) implying the convergence to zero, as $n \to \infty$, of the first term on the RHS of (A.12). Since, $\varepsilon > 0$ may be taken arbitrarily small, it follows from (A.12) that $\max_{1\le i\le n}[c_{ni}^2/\sum_{i=1}^n c_{ni}^2] \to 0$, as $n \to \infty$. This establishes (2.3) and, consequently, the equivalence of the "negligibility" and Lindeberg conditions in Theorem 2.2.

Deduction of Theorem 2.2a. To the notation of Theorem 2.2a add the notation $Y_{ni} = [c_{ni}(X_{ni} - \mu_{ni})/B_n]$, where $B_n^2 = \sum_{i=1}^n c_{ni}^2\sigma_{ni}^2$. Then the r.v. Z_n of Equation (2.2a) and condition (2.3a) of Theorem 2.2a reduce, respectively, to $Z_n = \sum_{i=1}^n Y_{ni}$ and $\sum_{i=1}^n |Y_{ni}|^{2+\delta} \to 0$, as $n \to \infty$ for a $\delta > 0$. Theorem 2.2a now follows by the LF Theorem stated, on account of the last convergence, since, $\sum_{i=1}^n E\{Y_{ni}^2 I_{[|Y_{ni}|\ge\varepsilon]}\} \le \varepsilon^{-\delta}\sum_{i=1}^n E|Y_{ni}|^{2+\delta} \to 0$, as $n \to \infty$.

[Received July 2011. Revised October 2012.]

REFERENCES

Anton, H. (1999), *Calculus A New Horizon*, New York: Wiley. [54]

Casella, G., and Berger, R. (2002), *Statistical Inference* (2nd ed.), Pacific Grove, CA: Duxbury. [51,52,55]

De Moivre, A. (1738), *The Doctrine of Chances* (2nd ed.), London: W. Pearson. [49]

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications* (Vol. I), New York: Wiley. [52,54]

——— (1971), *An Introduction to Probability Theory and Its Applications* (Vol. II), New York: Wiley. [55]

Johnson, N., Kotz, S., and Balakrishnan, N. (1995), *Continuous Univariate Distributions* (2nd ed., Vol. 2), NewYork: Wiley. [51]

Hájek, J., and Šidák, Z. (1967), *Theory of Rank Tests*, New York: Academic Press. [53]

Knight, K. (2000), *Mathematical Statistics*, Boca Raton, FL: Chapman and Hall/CRC. [52]

Krantz, S. G. (1999), "The Gamma and Beta Functions, §13.1," in *Handbook of Complex Variables*, pp. 155–158, Boston, MA: Birkhauser. [54]

Laplace, P. S. (1810), Mémoire sur les Intégrales Définies et leur Application aux Probabilitiés, Reproduced in *Oeuvres de Laplace*, 12, 357–412. [49]

Lehmann, E. L. (1999), *Elements of Large-Sample Theory*, New York: Springer-Verlag. [49,50,53]

Lindeberg, J. W. (1922), "Eine neue Herleitung des Exponentialgesetz in der Wahrscheinlichkeit-srechnung," *Mathematische Zeitschrift*, 15, 211–225. [49]

Rao, C. R. (1966), *Linear Statistical Inference and Its Applications*, New York: Wiley. [52,55]

Romano, J. P., and Siegel, A. F. (1986), *Counter Examples in Probability and Statistics*, Monterey, CA: Wadsworth and Brooks/Cole. [53,55]

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: Wiley. [50,55]