Advanced Studies in Contemporary Mathematics
28 (2018), No. 2, pp. 187 - 200

www.jangjeon.or.kr
http://dx.doi.org/10.17777/ascm2018.28.2.187

Transfer theorems concerning asymptotic expansions for the distribution functions of statistics based on samples with random sizes

Bening V. E. *

Keywods

Asymptotic normality; sample of random size; asymptotic expansions; Laplace and Student distributions; transfer theorem.

2010 Mathematical Subject Classification: Primary 60F05

Secondary 62E20; 91B30; 91B70

Abstract

In the paper, we discuss the transformation of the asymptotic expansion for the distribution of a statistic admitting Edgeworth expansion if the sample size is replaced by a random variable.We demonstrate that all those statistics that are regarded as asymptotically normal in the classical sense, become asymptotically Laplace or Student if the sample size is random. We especially separate the case where the Student distribution parameter ("the number degrees of freedom") is small. We show that the Student distribution with arbitrary "number of degrees of freedom"can be obtained as the limit when the sample size is random. We emphasize the possibility of using a family of Student distributions as a comfortable model with heavy tails since in this case many relations, in particular, a likelihood function, have the explicit form (unlike stable laws). Thus, the Laplace and Student distributions may be used as an asymptotic approximation in descriptive statistics being a convenient heavy-tailed alternative to stable laws.

1 Introduction

In classical problems of mathematical statistics, the size of the available sample, i. e., the number of available observations, is traditionally assumed to be deterministic. In the asymptotic settings it plays the role of infinitely increasing *known* parameter. At the same time, in practice very often the data to be analyzed is collected or registered during a certain period of time and the flow of informative events each of which brings a next observation forms a random point process. Therefore, the number of available observations is unknown till the end of the process

*Peoples' Friendship University of Russia (RUDN University), 6 Miklukho - Maklaya St., Moscow, 117198, Russian Federation; bening@yandex.ru

of their registration and also must be treated as a (random) observation. For example, this is so in insurance statistics where during different accounting periods different numbers of insurance events (insurance claims and/or insurance contracts) occur and in high performance information systems where due to the stochastic character of the intensities of information flows, the size of data available for the statistical analysis can be often regarded as random. Say, the statistical algorithms applied in high-frequency financial applications must take into consideration that the number of events in a limit order book during a time unit essentially depends on the intensity of order flows. Moreover, contemporary statistical procedures of insurance and financial mathematics do take this circumstance into consideration as one of possible ways of dealing with heavy tails. However, in other fields such as medical statistics or quality control this approach has not become conventional yet although the number of patients with a certain disease varies from month to month due to seasonal factors or from year to year due to some epidemic reasons and the number of failed items varies from lot to lot. In these cases the number of available observations as well as the observations themselves are unknown beforehand and should be treated as random to avoid underestimation of risks or error probabilities.

In asymptotic settings, statistics constructed from samples with random sizes are special cases of random sequences with random indices. The randomness of indices usually leads to that the limit distributions for the corresponding random sequences are heavy-tailed even in the situations where the distributions of non-randomly indexed random sequences are asymptotically normal see, e. g., [1] – [4]. For example, if a statistic which is asymptotically normal in the traditional sense, is constructed on the basis of a sample with random size having negative binomial distribution, then instead of the expected normal law, the Student distribution with power-type decreasing heavy tails appears as an asymptotic law for this statistic.

We use conventional notation: $\mathbb{R}$ is the set of real numbers, $\mathbb{N}$ is the set of natural numbers, $\Phi(x)$ and $\varphi(x)$ are the distribution function (d.f.) and the probability density of the standard normal law, respectively, the symbol $\Longrightarrow$ denotes the weak convergence.

1.1 Laplace distribution as an asymptotic approximation

In 1774 P. S. Laplace introduced in his paper «Sur la probabilité des causes par les événements»(see [5] and references in the book) a native probabilistic law for the error of measurement in the following formulation: «the logarithm of the frequency of an error (without regard to sign) is a linear function of the error». Later in 1911 the famous economist and probabilist J. M. Keynes obtained the first law error again from the assumption that the most probable value of the measured quantity is equal to the median of measurements (see [5]). Later in 1923 E. B. Wilson suggested that the frequency we actually meet in everyday work in economics, biometrics, or vital statistics often fails to conform closely to the normal distribution, and that Laplace's first law should be considered as a candidate for fitting data in economics and health sciences (see [5] and references in the book). Fifty years later in scientific papers (see [5]) one could often find appeals for using the first Laplace's law as the main hypothesis instead of the normal distribution for the economical, biometrical and demographic data.

Nowadays the first Laplace's law is called the Laplace distribution. The distribution is defined by its characteristic function (see [2] and references in the paper)

$$f(t) \;=\; \frac{1}{1+\lambda^2\,t^2}, \quad t \;\in\; \mathbb{R}^1, \tag{1.1}$$

or by its density

$$l(x) = \frac{1}{2\,\lambda}\exp\Big\{-\frac{|x|}{\lambda}\Big\}, \quad \lambda > 0, \quad x \in \mathbb{R}^1. \tag{1.2}$$

Another name – double exponential distribution – shows an opportunity to obtain it as the difference between two independent identically distributed exponential random variables which are often used for modeling of lifetime of an observable object.

We now present the reasoning from [2] which validates the use of Laplace distribution in problems of probability theory and mathematical statistics as the limiting distribution for samples of random size. Consider random variables $N_1, N_2, \ldots, X_1, X_2, \ldots$ defined on a common measurable space $(\Omega, \mathcal{A})$. Let P be a probability measure over $(\Omega, \mathcal{A})$. Suppose that the random variables N_n take on positive integers for any $n \geq 1$ and do not depend on $X_1, X_2, \ldots$. Define the random variable T_{N_n} for some statistic $T_n = T_n(X_1, \ldots, X_n)$ and any $n \geq 1$ by

$$T_{N_n}(\omega) = T_{N_n(\omega)}(X_1(\omega), \ldots, X_{N_n(\omega)}(\omega)),$$

for every outcome $\omega \in \Omega$. The statistic T_n is called asymptotically normal if there exist real numbers $\sigma > 0$ and $\mu \in \mathbb{R}^1$ such that, as $n \to \infty$,

$$\mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) \Longrightarrow \Phi(x), \tag{1.3}$$

where $\Phi(x)$ is the standard normal distribution function.

The asymptotically normal statistics are abundant. Paper [2] contains some examples of these statistics: the sample mean (assuming nonzero variances), the central order statistics or the maximum likelihood estimators (under weak regularity conditions) and many others. The following lemma, proved in [2], gives the necessary and sufficient conditions under which the distributions of asymptotically normal statistics based on samples of random size converge to a predetermined distribution $F(x)$.

Lemma 1.1. ([2]) *Let $\{d_n\}_{n\geq 1}$ be an increasing and unbounded sequence of positive numbers. Suppose that $N_n \to \infty$ in probability as $n \to \infty$. Let T_n be an asymptotically normal statistic as in (1.3). Then a necessary and sufficient condition for a distribution function $F(x)$ to satisfy*

$$\mathsf{P}(\sigma\sqrt{d_n}(T_{N_n} - \mu) < x) \Longrightarrow F(x), \quad n \to \infty$$

is that there exists a distribution function $H(x)$ satisfying

$$H(x) = 0, \quad x < 0;$$

$$F(x) = \int_0^\infty \Phi(x\sqrt{y})\,dH(y), \quad x \in \mathbb{R}^1;$$

$$\mathsf{P}(N_n < d_n x) \Longrightarrow H(x), \quad n \to \infty.$$

It is well known (see e.g. [2]) that the Laplace distribution can be expressed in terms of a scale mixture of normal distributions (with zero mean) with an inverse exponential mixing distribution, i.e., for any $x \in \mathbb{R}^1$,

$$L_\lambda(x) = \int_0^\infty \Phi(\,x\sqrt{y})\,dQ_\lambda(y),$$

where $Q_\lambda(x)$ is the distribution function of the inverse exponential distribution

$$Q_\lambda(x) = e^{-2\lambda^2/x}, \quad x > 0,$$

and $L_\lambda(x)$ is the distribution function of the Laplace distribution corresponding to the density (1.2).

Recall that the inverse exponential distribution is the distribution of the random variable

$$V = \frac{1}{U},$$

where the random variable U has the exponential distribution, and the inverse exponential distribution is a special case of the Fréchet distribution which is well known in asymptotic theory of order statistics as the type II extreme value distribution.

Lemma 1.1 can be applied to derive the following theorem which gives the necessary and sufficient conditions for the Laplace distribution to be the limiting distribution of the asymptotically normal statistics based on samples of random size.

Theorem 1.2. ([2]) *Let $\sigma > 0$ and $\{d_n\}_{n\geq 1}$ be an increasing and unbounded sequence of positive numbers. Suppose that $N_n \to \infty$ in probability as $n \to \infty$. Let T_n be an asymptotically normal statistic as in (1.3). Then*

$$\mathsf{P}(\sigma\sqrt{d_n}(T_{N_n} - \mu) < x) \Longrightarrow L_\lambda(x), \quad n \to \infty$$

if and only if

$$\mathsf{P}(N_n < d_n x) \Longrightarrow Q_\lambda(x), \quad n \to \infty.$$

Consider an example in which the random size of sample has the limiting inverse exponential distribution $Q_\lambda(x)$. Let $Y_1, Y_2, \dots$ be the independent and identically distributed random variables with some continuous distribution function. Let m be a positive integer and

$$N(m) = \min\{\, n \geq 1 : \max_{1\leq\, j \leq\, m} Y_j < \max_{m+1 \leq\, k\, \leq\, m+n} Y_k \,\}, \quad m \in \mathbb{N}.$$

The random variable $N(m)$ denotes the number of additional observations needed to exceed the current maximum obtained with m observations. The distribution of the random variable $N(m)$ was obtained by S.S. Wilks ([6]). So, the distribution of $N(m)$ is the discrete Pareto distribution

$$\mathsf{P}(N(m) \geq k) = \frac{m}{m+k-1}, \quad k \geq 1. \tag{1.4}$$

Now, let $N^{(1)}(m), N^{(2)}(m), \dots$ be the independent random variables with the same distribution (1.4). Then the following statement was proved in [2]: for any $x > 0$,

$$\lim_{n\to\infty} \mathsf{P}\Big(\frac{1}{n} \max_{1\leq\, j\leq\, n} N^{(j)}(m) < x \Big) = e^{-m/x}.$$

Therefore, the limit is the distribution function of the inverse exponential distribution with $2\lambda^2 = m$. And if

$$N_n = \max_{1\,\leq\, j \leq\, n} N^{(j)}(m), \tag{1.5}$$

then Theorem 1.2 (with $d_n = n$) gives the Laplace distribution as the limiting distribution of regular statistics.

Theorem 1.3. ([2]) *Let m be any positive integer. Suppose that $N^{(1)}(m), N^{(2)}(m), \dots$ are independent random variables having the same distribution (1.4), and a random variable N_n is defined by (1.5). Let T_n be an asymptotically normal statistic as in (1.3). Then*

$$\mathsf{P}(\sigma\sqrt{n}(T_{N_n} - \mu) < x) \Longrightarrow L_{\lambda(m)}(x), \quad n \to \infty,$$

where $L_{\lambda(m)}(x)$ *is the distribution function of the Laplace distribution with density (1.2) with* $\lambda(m) = \sqrt{m/2}$.

Further, the Laplace distribution plays the same role in the theory of geometric random sums as the normal distribution plays in the classical probability theory (see e.g. [7]). In turn, the geometric random sums play an important role in the investigation of speculative processes. The reason of increasing usage of the Laplace distribution is also its representation as a scale mixture of some well known distributions. For example, the Laplace distribution can be represented as a scale mixture of symmetrized Rayleigh-Rice distribution with the mixing χ^2-distribution with 1 degree of freedom (see Corollary 3.2 in [2]).

The Laplace distribution as a probabilistic model for applications is also attractive because of its extremal entropy property. This property often motivates a choice of Laplace distribution as a model for the error of measurements when the accuracy randomly varies from one measurement to the next (see [2]).

In applied economics and science, the popularity of Laplace distribution as a mathematical (probabilistic) model is explained by the fact that the Laplace distribution has heavier tails than the normal distribution does. So, in communication theory, the Laplace distribution is considered as a probabilistic model for some types of random noise in problems of detection of a known constant signal (see [7] – [10]). In [11] the Laplace distribution is referred to as a model for speech signal in problems of encoding and decoding of analog signals. In [12] an application of the Laplace distribution is discussed in relation to the fracturing of materials under applied forces. In [13] and [14] authors give examples of application of Laplace distribution in aerodynamics, when the gradient of airspeed change against its duration is modeled by mixtures of the Laplace distribution with the normal distribution. Modeling of the error distributions in navigation with Laplace distribution is investigated in [15].

This increased interest in Laplace distribution from applied sciences motivates the Laplace distribution to be investigated in mathematical statistics and theory of probability. The nonregularity of the Laplace distribution makes known difficulties of its use in problems of testing statistical hypotheses. But the asymptotic methods of testing statistical hypotheses developed in last decades now allow to use the Laplace distribution in mathematical statistics (see [5] and references in the work).

1.2 Student distribution as an asymptotic approximation

The Student distribution is an absolutely continuous probability distribution given by the density

$$p_\gamma(x) = \frac{\Gamma\big((\gamma+1)/2\big)}{\sqrt{\pi\gamma}\,\Gamma(\gamma/2)} \Big(1 + \frac{x^2}{\gamma}\Big)^{-\frac{\gamma+1}{2}}, \qquad -\infty < x < \infty. \tag{1.6}$$

Here $\gamma > 0$ is a parameter, $\Gamma(\,\cdot\,)$ is the Euler gamma function

$$\Gamma(z) = \int_0^\infty e^{-y}\, y^{z-1}\, dy, \qquad z > 0.$$

In particular, for $\gamma = 1$, density (1.6) is

$$p_1(x) = \frac{1}{\pi(1+x^2)}, \qquad -\infty < x < \infty,$$

which corresponds to the Cauchy distribution. It is not difficult to see that the Student distribution with the parameter γ has no moments of order $\delta \geqslant \gamma$. If $\gamma = n$ is a positive integer, and if $X, X_1, \ldots, X_n$ are independent random variables with the same standard normal distribution, then the r.v.

$$Y = \frac{\sqrt{n} \cdot X}{\sqrt{X_1^2 + \ldots + X_n^2}} \tag{1.7}$$

has the Student distribution with the parameter n, which is called *the number of degrees of freedom*. Representatiion (1.7) is the basis of testing hypotheses on a mean of normal samples proposed in 1908 by W. S. Gossett, who used the pseudonym "Student" for his paper "On the probable error of a mean" [16].

The important role played by the Student distribution in mathematical statistics in analysis of normal samples is well known. Here the parameter γ is closely related to the sample size and takes positive integer values. However, we can say that in these problems the Student distribution plays an auxiliary role as an abstract, ideal theoretical model.

The descriptive statistic does not use the Student distribution as an analytic model "adjusted" to experimental data. (Recently papers appeared in which the Student distribution is applied (without a theoretical basis) for describing some financial indices, in particular, increments of logarithms of stock exchange prices. We mention the works of Praetz [17] and Blattberg and Gonedes [18]). This can be confirmed by the fact that there is no book on the theory (or practice) of statistical estimation which considers ax estimation problem for the parameter of the Student distribution.

In the field of applied mathematics there seems to be insufficient trust in the Student distribution as a model for describing the statistical behavior of real data. This lack of trust is related to the fact that, istead of normal and Poisson distributions which are used as limits in the central limit theorem, and instead of the Poisson theorem on rare events, the Student distribution is not considered an asymptotic approximation.

It is necessary to emphasize that in view of the relative simplicity of representation, the Student distribution could be a comfortable analytic model describing probabilistic statistical properties of large risks since it is more heavy-tailed than the normal law. For example, the Student distribution could be a comfortable alternative to stable laws, which are frequently applied to describe these properties. The advantage of the Student distribution over stable models is in the fact that the statistical analysis of Student models is simplier since, in this case, the likelihood function can be written in explicit form in terms of elementary functions while, for stable laws, this is impossible (excluding four exceptioxx). At the same time, for $0 < \gamma \leqslant 2$ as $|x| \to \infty$ tails of the Student distribution decrease exponentially, a fact which coincides with the asymptotic (as $|x| \to \infty$) behavior of tails of stable laws.

Our further arguments will be based on the two following lemmas.

Let $N_{p,r}$ be an r.v. with a negative binomial distribution

$$\mathsf{P}(N_{p,r} = k) = C_{r+k-2}^{k-1} p^r (1-p)^{k-1}, \quad k = 1, 2, \ldots \tag{1.8}$$

Here $r > 0$ and $p \in (0, 1)$ are parameters and, for noninteger r,the value C_{r+k-2}^{k-1} is defined in the following way:

$$C_{r+k-2}^{k-1} = \frac{\Gamma(r+k-1)}{(k-1)! \cdot \Gamma(r)}.$$

In particular, for $r = 1$ relation (1.8) gives a geometrical distribution. It is known that

$$\mathsf{E}\, N_{p,r} = \frac{r(1-p)+p}{p},$$

so $\mathsf{E}\, N_{p,r} \to \infty$ as $p \to 0$.

A negative binomial distribution with the integer positive r admits an illustrative interpretation in terms of Bernoulli trials. Namely, an r.v. with distribution (1.5) is the number of Bernoully trials tested before the rth failure if the probability of success in a single trial is equal to $1-p$.

Lemma 1.4. *For any fixed $r > 0$,*

$$\lim_{p\to 0} \mathsf{P}\Big(\frac{N_{p,r}}{\mathsf{E}\, N_{p,r}} < x\Big) \;=\; G_{r,r}(x),$$

uniformly in $x \in \mathbb{R}$, where $G_{r,r}(x)$ is the gamma distribution function with the shape parameter which coincides with the scale parameter and equals r,

$$G_{\alpha,\lambda}(x) \;=\; \begin{cases} 0, & x \leqslant 0, \\ \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^x e^{-\lambda y}\, y^{\alpha-1}\, dy, & x > 0. \end{cases}$$

The proof is a simple exercise on the application of characteristic functions argument.

Lemma 1.5. *Let $\gamma > 0$ be arbitrary and let $\{d_n\}_{n\geqslant 1}$ be some infinitely increasing sequence of positive numbers. Suppose that $N_n \to \infty$ in probability as $n \to \infty$. Let the statistic T_n be asymptotically normal in the sense of (1.3). In order to have*

$$\mathsf{P}\left(\sigma\sqrt{d_n}\big(T_{N_n} - \mu\big) < x\right) \Longrightarrow F_\gamma(x), \qquad n\to\infty,$$

where $F_\gamma(x)$ is the Student distribution function with the parameter γ, it is necessary and sufficient that

$$\mathsf{P}(N_n < d_n x) \Longrightarrow G_{\gamma/2,\gamma/2}(x), \quad n\to\infty.$$

Proof. It is not difficult to prove that for arbitrary $\gamma > 0$ the density $p_\gamma(x)$ of the Student distribution with parameter γ (see (1.6)) can be represented in the form

$$p_\gamma(x) \;=\; \mathsf{E}\, \sqrt{U_{\gamma/2}}\,\varphi\big(x\sqrt{U_{\gamma/2}}\big), \tag{1.9}$$

where $U_{\gamma/2}$ is a r.v. with distribution function $G_{\gamma/2,\gamma/2}(x)$. Indeed

$$\mathsf{E}\, \sqrt{U_{\gamma/2}}\,\varphi\big(x\sqrt{U_{\gamma/2}}\big) \;=$$

$$= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2}\sqrt{\pi}\,\Gamma(\gamma/2)} \int_0^\infty \exp\Big\{-u\Big(\frac{x^2+\gamma}{2}\Big)\Big\}\, u^{(\gamma-1)/2} du \;=$$

$$= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2}\sqrt{\pi}\,\Gamma(\gamma/2)} \Big(\frac{x^2+\gamma}{2}\Big)^{-(\gamma+1)/2} \int_0^\infty \exp\{-z\}\, z^{(\gamma+1)/2-1} dz \;=$$

$$= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2}\sqrt{\pi}\,\Gamma(\gamma/2)} \Big(\frac{x^2+\gamma}{2}\Big)^{-(\gamma+1)/2} \Gamma\Big(\frac{\gamma+1}{2}\Big) \;=$$

$$= \frac{\Gamma\big((\gamma+1)/2\big)}{\sqrt{\pi\gamma}\,\Gamma(\gamma/2)} \Big(1+\frac{x^2}{\gamma}\Big)^{-\frac{\gamma+1}{2}} \;=\; p_\gamma(x).$$

However, density (1.9) corresponds to the distribution function $\mathsf{E}\,\Phi\big(x\sqrt{U_{\gamma/2}}\big)$. Now the needed statement follows from Lemma 1.1 with an account of the identifiability of scale mixtures of normal laws, according to which if $\mathsf{E}\,\Phi\big(xV_1\big) \equiv \mathsf{E}\,\Phi\big(xV_2\big)$ for some positive random variables V_1 and V_2, then $V_1 \stackrel{d}{=} V_2$ (see, for example, [19]). $\square$

Corollary 1.6. *Let $r > 0$ be arbitrary. Suppose that for any $n \geqslant 1$ a r.v. N_n has the negative binomial distribution with parameters $p = 1/n$ and r. Let the statistic T_n be asymptotically normal in the sense of (2.1). Then*

$$\mathsf{P}\left(\sigma\sqrt{rn}\big(T_{N_n} - \mu\big) < x\right) \Longrightarrow F_{2r}(x), \qquad n \to \infty$$

uniformly in $x \in \mathbb{R}$, where $F_{2r}(x)$ is the Student distribution function with parameter $\gamma = 2r$.
Proof. By Lemma 1.4 we have

$$\frac{N_n}{nr} = \frac{N_n}{\mathsf{E}\,N_n} \cdot \frac{\mathsf{E}\,N_n}{nr} = \frac{N_n}{\mathsf{E}\,N_n} \cdot \frac{r(n-1)+1}{nr} = \frac{N_n}{\mathsf{E}\,N_n}\Big[1 + O\Big(\frac{1}{n}\Big)\Big] \Longrightarrow U_r$$

as $n \to \infty$, where U_r is a r.v. with gamma distribution with a shape parameter coinciding with scale parameter and equal to r. Now the needed statement directly follows from Lemma 1.5.

Remark 1.7. The Cauchy distribution ($\gamma = 1$) arises in the situation described in Corollary 1.6, when the sample size N_n has a negative binomial distribution with parameters $p = 1/n$, $r = 1$, and n is large.

Remark 1.8. When the sample size N_n has a negative binomial distribution with parameters $p = 1/n,\ r = 1$ (i.e., geometric distribution with parameter $p = 1/n$), in a limit as $n \to \infty$ we obtain a Student distribution with parameter $\gamma = 2$ to which the following distribution function corresponds:

$$F_2(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{2+x^2}}\right), \qquad x \in \mathbb{R}. \tag{1.10}$$

This distribution was first described as a limit for a sample median constructed by a sample of a random size having a geometric distribution in [20] (we note that this work does not say the distribution function in the right-hand side of (1.10) corresponds to the Student distribution).

The main conclusion from the results given above can be formulated as follows. If the number of random factors determining an observed value of a r.v. is a r.v. itself and its distribution can be approximated with the help of gamma distribution having identical parameters (for example, is a negative binomial distribution with a probability of success close to one; see Lemma 1.4), then the functions of random factor values which, in the classic situation are asymptotically normal, really are asymptotically Student. Hence, since gamma models with the identical parameters and negative binomial models are widely used, the Student distribution can be considered as quite a reasonable model in problems of applied statistics.

2 Asymptotic expansions

In this section we consider conditions under which the distribution functions of the statistics based on samples with random sizes possess Edgeworth expansions. We briefly recall the setup and some notation.

Consider random variables $N_1, N_2, ...$ and $X_1, X_2, ...$, defined on the same probability space $(\Omega,\ \mathcal{A},\ \mathsf{P})$. By $X_1, X_2, ...X_n$ we will mean statistical observations whereas the r.v. N_n will be regarded as the random sample size depending on the parameter $n \in \mathbb{N}$. Assume that for each $n \geq 1$ the r.v. N_n takes only natural values (i.e., $N_n \in \mathbb{N}$) and is independent of the sequence $X_1, X_2, ...$ Everywhere in what follows the r.v.'s $X_1, X_2, ...$ are assumed independent and identically distributed.

For every $n \geq 1$ by $T_n = T_n(X_1, ..., X_n)$ denote a statistic, i.e., a real-valued measurable function of $X_1, ..., X_n$. For each $n \geq 1$ we define a r.v. T_{N_n} by setting $T_{N_n}(\omega) \equiv T_{N_n(\omega)}(X_1(\omega), ..., X_{N_n(\omega)}(\omega))$, $\omega \in \Omega$.

The following condition determines the asymptotic expansion (a.e.) for the distribution function of T_n with a non-random sample size.

Condition 1. *There exist $l \in \mathbb{N}$, $\mu \in \mathbb{R}$, $\sigma > 0$, $\alpha > l/2$, $\nu \geqslant 0$, $C_1 > 0$, a differentiable d.f. $F(x)$ and differentiable bounded functions $f_j(x), j = 1, ..., l$ such that*

$$\sup_x \Big| \mathsf{P}\big(\sigma n^{\nu}(T_n - \mu) < x\big) - F(x) - \sum_{j=1}^{l} n^{-j/2} f_j(x) \Big| \leq \frac{C_1}{n^{\alpha}}, \quad n \in \mathbb{N}.$$

The following condition determines the a.e. for the d.f. of the normalized random index N_n.

Condition 2. *There exist $m \in \mathbb{N}$, $\beta > m/2$, $C_2 > 0$, a function $0 < g(n) \uparrow \infty, n \to \infty$, a d.f. $H(x), H(0+) = 0$ and functions $h_i(x)$, $i = 1, ..., m$ with bounded variation such that*

$$\sup_{x \geqslant 0} \Big| \mathsf{P}\Big(\frac{N_n}{g(n)} < x\Big) - H(x) - \sum_{i=1}^{m} n^{-i/2} h_i(x) \Big| \leq \frac{C_2}{n^{\beta}}, \quad n \in \mathbb{N}.$$

Define the function $G_n(x)$ as

$$G_n(x) = \int_{1/g(n)}^{\infty} F(xy^{\nu})\, dH(y) + \sum_{i=1}^{m} n^{-i/2} \int_{1/g(n)}^{\infty} F(xy^{\nu}) dh_i(y) +$$

$$+ \sum_{j=1}^{l} g^{-j/2}(n) \int_{1/g(n)}^{\infty} y^{-j/2} f_j(xy^{\nu})\, dH(y) +$$

$$+ \sum_{j=1}^{l} \sum_{i=1}^{m} n^{-i/2} g^{-j/2}(n) \int_{1/g(n)}^{\infty} y^{-j/2} f_j(xy^{\nu})\, dh_i(y). \tag{2.1}$$

Theorem 2.1. *Let the statistic $T_n = T_n(X_1, ..., X_n)$ satisfy Condition 1 and the r.v. N_n satisfy Condition 2. Then there exists a constant $C_3 > 0$ such that*

$$\sup_x \big| \mathsf{P}\big(\sigma g^{\nu}(n)(T_{N_n} - \mu) < x\big) - G_n(x)\big| \leq C_1 \,\mathsf{E}\, N_n^{-\alpha} + \frac{C_3 + C_2 M_n}{n^{\beta}},$$

where

$$M_n = \sup_x \int_{1/g(n)}^{\infty} \Big| \frac{\partial}{\partial y} \big(F(xy^{\nu}) + \sum_{j=1}^{l} (yg(n))^{-j/2} f_j(xy^{\nu})\big)\Big|\, dy$$

and the function $G_n(x)$ is defined by (2.1).

The proof directly follows from the formula of total probability, according to which

$$\sup_x \big|\mathsf{P}\big(\sigma\, g^{\nu}(n)(T_{N_n} - \mu) < x\big) - G_n(x)\big| \leq I_{1n} + I_{2n},$$

where

$$I_{1n} = \sup_x \Big| \int_{1/g(n)}^{\infty} \Big(F(xy^\nu) + \sum_{j=1}^{l} \big(yg(n)\big)^{-j/2} f_j(xy^\nu) \Big) \times$$

$$\times\, d\Big(\mathsf{P}\Big(\frac{N_n}{g(n)} < y \Big) - H(y) - \sum_{i=1}^{m} n^{-i/2}\, h_i(y) \Big) \Big|,$$

$$I_{2n} = \sum_{k=1}^{\infty} \sup_x \Big| \mathsf{P}\Big(\sigma k^\nu (T_k - \mu) < x\Big(\frac{k}{g(n)}\Big)^\nu \Big) -$$

$$- F\Big(x\,\Big(\frac{k}{g(n)}\Big)^\nu\Big) - \sum_{j=1}^{l} k^{-j/2} f_j\Big(x\Big(\frac{k}{g(n)}\Big)^\nu\Big) \Big| \mathsf{P}(N_n = k).$$

Now

$$I_{1n} \leqslant \frac{C_3}{n^\beta} + \sup_x \Big| \int_{1/g(n)}^{\infty} \Big(\mathsf{P}\Big(\frac{N_n}{g(n)} < y\Big) - H(y) - \sum_{i=1}^{m} n^{-i/2}\, h_i(y) \Big) \times$$

$$\times\, d\, \Big(F(x\, y^\nu) + \sum_{j=1}^{l} \big(yg(n)\big)^{-j/2} f_j(xy^\nu) \Big) \Big| \leqslant$$

$$\leqslant \frac{C_3}{n^\beta} + \sup_x \int_{1/g(n)}^{\infty} \Big| \mathsf{P}\Big(\frac{N_n}{g(n)} < y\Big) - H(y) - \sum_{i=1}^{m} n^{-i/2}\, h_i(y) \Big| \times$$

$$\times \Big| \frac{\partial}{\partial\, y} \left(F(xy^\nu) + \sum_{j=1}^{l} (yg(n))^{-j/2}\, f_j(xy^\nu) \right) \Big|\, dy \leqslant \frac{C_3}{n^\beta} + \frac{C_2 M_n}{n^\beta}$$

and

$$I_{2n} \leq C_1 \sum_{k=1}^{\infty} \frac{1}{k^\alpha}\, \mathsf{P}(N_n = k) = C_1\, \mathsf{E}\, N_n^{-\alpha}.$$

□

Let $\Phi(x)$ and $\varphi(x)$ respectively denote the d.f. of the standard normal law and its density.

Lemma 2.1. *Let $l = 1$, $0 < g(n) \uparrow \infty$, $F(x) = \Phi(x)$, $f_1(x) = \frac{1}{6}\mu_3\sigma^3(1 - x^2)\varphi(x)$. Then the quantity M_n in Theorem* 2.1 *satisfies the inequality $M_n \leqslant 2 + \widetilde{C}|\mu_3|\sigma^3$, where*

$$\widetilde{C} = \frac{1}{3}\sup_{u\geqslant 0}\big\{\varphi(u)(u^4 + 2u^2 + 1)\big\} = \frac{16}{3\sqrt{2\pi e^3}} \approx 0.47.$$

Consider some examples of application of Theorem 2.1.

2.1 Student distribution

Let $X_1, X_2, ...$ be i.i.d. r.v.'s with $\mathsf{E}X_1 = \mu$, $0 < \mathsf{D}X_1 = \sigma^{-2}$, $\mathsf{E}|X_1|^{3+2\delta} < \infty$, $\delta \in (0, \frac{1}{2})$ and $\mathsf{E}(X_1 - \mu)^3 = \mu_3$. For each n let

$$T_n = \tfrac{1}{n}(X_1 + ... + X_n). \tag{2.2}$$

Assume that the r.v. X_1 satisfies the Cramér Condition (C)

$$\limsup_{|t|\to\infty} |\mathsf{E}\ \exp\{itX_1\}| < 1.$$

Let $G_\gamma(x)$ be the Student d.f. with parameter $\gamma > 0$ corresponding to the density (see (1.6))

$$p_\gamma(x) = \frac{\Gamma(\gamma + 1/2)}{\sqrt{\pi\ \gamma}\Gamma(\gamma/2)}\Big(1 + \frac{x^2}{\gamma}\Big)^{-(\gamma+1)/2}, \quad x \in \mathbb{R},$$

where $\Gamma(\cdot)$ is the Euler's gamma-function and $\gamma > 0$ is the shape parameter.

For $r > 0$ let

$$H_r(x) = \frac{r^r}{\Gamma(r)}\int_0^x e^{-ry}\, y^{r-1}\, dy, \quad x \geqslant 0,$$

be the gamma-d.f. with parameter $r > 0$. Denote

$$g_r(x) = \int_0^\infty \varphi(x\sqrt{y})\,\frac{1 - x^2 y}{\sqrt{y}}\, dH_r(y), \quad x \geqslant 0. \tag{2.3}$$

Applying Theorem 2.1, we obtain the following result.

Theorem 2.2. *Let the statistic T_n have the form (2.2), where $X_1, X_2, ...$ are i.i.d. r.v.'s with $\mathsf{E}X_1 = \mu$, $0 < \mathsf{D}X_1 = \sigma^{-2}$, $\mathsf{E}|X_1|^{3+2\delta} < \infty$, $\delta \in (0, \frac{1}{2})$ and $\mathsf{E}(X_1 - \mu)^3 = \mu_3$. Moreover, assume that the r.v. X_1 satisfies the Cramér Condition (C). Assume that for some $r > 0$ the r.v. N_n has the negative binomial distribution*

$$\mathsf{P}(N_n = k) = \frac{(k + r - 2)\cdots r}{(k-1)!}\,\frac{1}{n^r}\,\Big(1 - \frac{1}{n}\Big)^{k-1}, \quad k \in \mathbb{N}.$$

Let $G_{2r}(x)$ be the Student d.f. with parameter $\gamma = 2r$ and $g_r(x)$ be defined by (2.3). Then for $r > 1/(1+2\delta)$, as $n \to \infty$, we have

$$\sup_x\Big|\mathsf{P}\Big(\sigma\sqrt{r(n-1)+1}\,(T_{N_n} - \mu) < x\Big) - G_{2r}(x) - \frac{\mu_3\sigma^3 g_r(x)}{6\sqrt{r(n-1)+1}}\Big| =$$

$$= \begin{cases} O\Big(\Big(\dfrac{\log n}{n}\Big)^{1/2+\delta}\Big), & r = 1,\\ O\big(n^{-\min(1, r(1/2+\delta))}\big), & r > 1,\\ O\big(n^{-r(1/2+\delta)}\big), & (1+2\delta)^{-1} < r < 1. \end{cases}$$

2.2 Laplace distribution

Consider the Laplace d.f. $\Lambda_\lambda(x)$ corresponding to the density

$$l_\lambda(x) = \frac{1}{2\lambda}\exp\Big\{-\frac{|x|}{\lambda}\Big\}, \quad \lambda > 0, \quad x \in \mathbb{R}.$$

Let $Y_1, Y_2, \dots$ be i.i.d. r.v.'s with a continuous d.f. Set

$$N(m) = \min\{\, i \geq 1 : \max_{1 \leq j \leq m} Y_j < \max_{m+1 \leq k \leq m+i} Y_k \,\}, \qquad m \in \mathbb{N}.$$

It is known that (see (1.4))

$$\mathsf{P}(N(m) \geq k) = \frac{m}{m+k-1}, \quad k \geq 1. \tag{2.4}$$

Now let $N^{(1)}(m), N^{(2)}(m), \dots$ be i.i.d. r.v.'s distributed in accordance with (2.4). Define the r.v.

$$N_n(m) = \max_{1 \leq j \leq n} N^{(j)}(m),$$

then, as it was shown in [2],

$$\lim_{n\to\infty} \mathsf{P}\Big(\frac{N_n(m)}{n} < x\Big) = e^{-m/x}, \quad x > 0,$$

and for an asymptotically normal statistic T_n (see (1.3)) we have

$$\mathsf{P}\big(\sigma\sqrt{n}(T_{N_n(m)} - \mu) < x\big) \longrightarrow \Lambda_{\lambda(m)}(x), \quad n \to \infty, \; x \in \mathbb{R},$$

where $\Lambda_{\lambda(m)}(x)$ is the Laplace d.f. with parameter $\lambda(m) = \sqrt{m/2}$.

Denote

$$l_m(x) = \int_0^\infty \varphi(x\sqrt{y})\,\frac{1-x^2y}{\sqrt{y}}\, de^{-m/y}, \quad x \in \mathbb{R}. \tag{2.5}$$

Theorem 2.3. *Let the statistic T_n have the form (2.2), where $X_1, X_2, \dots$ are i.i.d. r.v.'s with $\mathsf{E}X_1 = \mu$, $0 < \mathsf{D}X_1 = \sigma^{-2}$, $\mathsf{E}|X_1|^{3+2\delta} < \infty$, $\delta \in (0, \frac{1}{2})$ and $\mathsf{E}(X_1 - \mu)^3 = \mu_3$. Moreover, assume that the r.v. X_1 satisfies the Cramér Condition (C). Assume that for some $m \in \mathbb{N}$ the r.v. $N_n(m)$ has the distribution*

$$\mathsf{P}(N_n(m) = k) = \Big(\frac{k}{m+k}\Big)^n - \Big(\frac{k-1}{m+k-1}\Big)^n, \quad k \in \mathbb{N}.$$

Then

$$\sup_x \Big|\mathsf{P}\big(\sigma\sqrt{n}\,(T_{N_n(m)} - \mu) < x\big) - \Lambda_{\lambda(m)}(x) - \frac{\mu_3\sigma^3 l_m(x)}{6\sqrt{n}}\Big| = O\Big(\frac{1}{n^{1/2+\delta}}\Big), \quad n \to \infty,$$

where $l_m(x)$ is defined in (2.5).

Acknowledgement

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008).

E-mail: bening@yandex.ru

Address: Russia, 119992, Moscow, GSP-1, Vorob'evy Gory, Lomonosov Moscow State University, faculty of Computational Mathematics and Cybernetics.

References

1. *Bening V. E., Korolev V. Yu. Generalized Poisson Models and their Applications in Insurance and Finance*, *W. de Gruyter*, (Berlin, Germany) ISBN 978 - 3 - 11 - 093601 - 8, 2012, 433 p.
2. *Bening V. E., Korolev V. Yu.* Some statistical problems related to the Laplace distribution, Informatics and Its Applications, 2008, Vol. 2, No. 2, pp. 19 – 34.
3. *Gnedenko B. V., Korolev V. Yu. Random Summation. Limit Theorems and Applications*, *Boca Raton: CRC Press*, 1996, 267 p.
4. *Bening V. E., Korolev V. Yu.* On an application of the Student distribution in the theory of probability and mathematical statistics, Theory of Probability and Its Applications, 2005, Vol. 49, No. 3, pp. 377–391.
5. *Kotz S., Kozubowski T. J., Podgórski K. The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*,— Birkhäuser Boston, USA, 2001, 349 p.
6. *Wilks S. S.* Recurrence of extreme observations, *Journal of American Mathematical Society*, 1959. vol. 1, No. 1, p. 106 – 112.
7. *Asrabadi B. R.* The exact confidence interval for the scale parameter and the MVUE of the Laplace distribution. *Communications in Statistics. Theory and Methods*, 1985, vol. 14, No. 3, p. 713 – 733.
8. *Dadi M. I., Marks R. J. II.* Detector relative efficiencies in the presence of Laplace noise. *IEEE Trans. Aerospace Electron. Systems*, 1987, AES - 23(4), p. 568 – 582.
9. *Marks R. J., Wise G. L., Haldeman D. G., Whited J. L.* Detection in Laplace noise, *IEEE Trans. Aerospace Electron. Systems*, 1978, AES-14(6), p. 866 – 871.
10. *Miller J. H., Thomas J. B.* Detectors for discrete-time signals in non-Gaussian noise, *IEEE Trans. Inform. Theory*, 1972, IT-18(2), p. 241 – 250.
11. *Duttweiler D. L., Messerschmitt D. G.* Nearly instantaneous companding for nonuniformly quantizied PCM. *IEEE Trans. Comm.*, 1976, COM-24(8), p. 864 – 873.
12. *Epstein B.* Application of the theory of extreme values in fracture problems. *J. Amer. Statist. Assoc.*, 1948, vol. 43, No. 243, p. 403 – 412.

13. *Jones P. N. McLachlan G. J.* Laplace-normal mixtures fitted to wind shear data. // *J. Appl. Statistics*, 1990, vol. 17, No. 2, p. 271 – 276.

14. *Kanji G. K.* A mixture model for wind shear data. // *J. Appl. Statistics*, 1985, vol. 12, No. 1, p. 49 – 58.

15. *Hsu D. A.* Long-tailed distributions for position errors in navigation. // *Appl. Statist.*, 1979, vol. 28, No. 1, p. 62 – 72.

16. *Student.* On the probable error of a mean. // *Biometrika*, 1908, vol. 6, p. 1 – 25.

17. *Praetz P. D.*, The distribution of share-price changes. // *J. Business*, 1972, vol. 45, p. 49 – 55.

18. *Blattberg R, Gonedes N.*, A comparison of the stable and Student distributions as statistical models of stock prices. // *J. Business*, 1974, vol. 47, p. 244 – 250.

19. *Teicher H.*, Identifiability of mixtures. // *Ann. Math. Statist.*, 1961, vol. 32, p. 244 – 248.

20. *Gnedenko B. V.*, Estimating the unknown parameters of a distribution with a random number of independent observations. // *Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR*, // 1989, vol. 92, p. 146 – 150 (in Russian).