TRANSFER THEOREM CONCERNING ASYMPTOTIC EXPANSIONS FOR THE DISTRIBUTION FUNCTIONS OF STATISTICS BASED ON SAMPLES WITH RANDOM SIZES

Vladimir Bening
Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
Institute for Informatics Problems, Russian Academy of Sciences

V. A. Savushkin
Dubna State University

E. I. Shunkov
Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University

Alexander Zeifman
Vologda State University, Russia
Institute of Informatics Problems, Russian Academy of Sciences
Institute of Territories Socio-Economic Development, Russian Academy of Sciences

Victor Korolev
Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
Institute for Informatics Problems, Russian Academy of Sciences

KEYWORDS

Random size; asymptotic expansion; transfer theorem; Student distribution; Laplace distribution

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. 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.

INTRODUCTION

In 1774 P. S. Laplace in his paper "Sur la probabilité des causes par les événements" (see (Kotz et al. 2001) and references in the book) introduced 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 (Kotz et al. 2001) and references in the book). 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 (Kotz et al. 2001) and references in the book). Fifty years later in scientific papers (see (Kotz et al. 2001) and references in the book) 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 (Bening and Korolev 2008) and the references therein)

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

or by its density

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

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 (Bening and Korolev 2008) which validates the use of Laplace distribution in problems of probability theory and mathe-

Proceedings 28th European Conference on Modelling and Simulation ©ECMS Flaminio Squazzoni, Fabio Baronio, Claudia Archetti, Marco Castellani (Editors)
ISBN: 978-0-9564944-8-1 / ISBN: 978-0-9564944-9-8 (CD)

matical 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 $n \to \infty$,

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

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

The asymptotically normal statistics are abundant. Paper (Bening and Korolev 2008) 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 (Bening and Korolev 2008), 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. (Korolev 1995) *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. (Bening and Korolev 2008)) 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 $x \in \mathbb{R}^1$,

$$L(x) \;=\; \int_0^\infty \Phi(x\sqrt{y})dQ(y),$$

where $Q(x)$ is the distribution function of the inverse exponential distribution

$$Q(x) \;=\; e^{-\delta/x}, \quad \delta > 0, \quad x > 0,$$

and $L(x)$ is the distribution function of the Laplace distribution corresponding to the density (1.2) with $\nu^2 = 1/\delta$.

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. (Bening and Korolev 2008) *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*

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

if and only if

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

Consider an example from (Bening and Korolev 2008) in which the random size of sample has the limiting inverse exponential distribution $Q(x)$. Let $Y_1, Y_2, \ldots$ be the independent and identically distributed random variables with some continuous distribution function. Let m be a positive integer and

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

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 (Wilks 1959). So, the distribution of $N(m)$ is the discrete Pareto distribution

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

Now, let $N^{(1)}(m), N^{(2)}(m), \ldots$ be the independent random variables with the same distribution (1.4). Then the following statement was proved in (Bening and Korolev 2008): 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 $\delta = m$. And if

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

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

Theorem 1.3. (Bening and Korolev 2008) *Let m be any positive integer. Suppose that $N^{(1)}(m), N^{(2)}(m), \ldots$ 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*

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

where $L(x)$ is the distribution function of the Laplace distribution with density (1.2) with $\nu^2 = 1/m$.

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. (Bening and Korolev 2008) and the references therein). 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 (Bening and Korolev 2008)).

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 (Bening and Korolev 2008)).

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 (Astrabadi 1985, Dadi and Marks 1987, Marks et al. 1978, Miller and Thomas 1972). In (Duttweiler and Messerschmitt 1976) the Laplace distribution is referred to as a model for speech signal in problems of encoding and decoding of analog signals. In (Epstein 1948) an application of the Laplace distribution is discussed in relation to the fracturing of materials under applied forces. In (Jones and McLachlan 1990, Kanji 1985) 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 (Hsu 1979).

This increased interest in Laplace distribution from applied sciences motivates the Laplace distribution to be investigated in mathematical statistics and theory of probability. The non-regularity 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 (Kotz et al. 2001) and references in the work).

ASYMPTOTIC EXPANSIONS

Consider random variables (r.v.'s) $N_1, N_2, \ldots$ and $X_1, X_2, \ldots$, defined on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$. By $X_1, X_2, \ldots X_n$ 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 $X_1, X_2, \ldots$ Everywhere in what follows the r.v.'s $X_1, X_2, \ldots$ are assumed independent and identically distributed.

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

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

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

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

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 $h_i(x)$, $i = 1, \ldots, m$ 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

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

$$+ \sum_{j=1}^{l} g^{-j/2}(n) \int_{1/g(n)}^{\infty} y^{-j/2} f_j(xy^\gamma) dH(y) + \qquad (2.1)$$

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

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*

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

where

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

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

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.

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

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

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

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

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

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

where $\Gamma(\cdot)$ is the Euler's gamma-function and $\nu > 0$ is the shape parameter (if $\nu \in \mathbb{N}$, then ν is called *the number of degrees of freedom*). In practice, it can be arbitrarily small determining the typical heavy-tailed distribution. If $\nu = 2$, then the d.f. $G_2(x)$ is expressed explicitly as

$$G_2(x) = \frac{1}{2}\Big(1 + \frac{x}{\sqrt{2 + x^2}}\Big), \ \ x \in \mathbb{R}.$$

for $\nu = 1$ we have the Cauchy distribution.

For $r > 0$ let

$$H_r(x) = \frac{r^r}{\Gamma(r)} \int\limits_0^x e^{-ry} y^{r-1} dy, \quad x \geqslant 0,$$

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

$$g_r(x) = \int\limits_0^\infty \varphi(x\sqrt{y}) \frac{1 - x^2 y}{\sqrt{y}} dH_r(y), \quad x \geqslant 0. \quad (2.3)$$

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*

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

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

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

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

Laplace distribution

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

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

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

$$N(s) = \min\{i \ge 1 : \max_{1 \le j \le s} Y_j < \max_{s+1 \le k \le s+i} Y_k\}.$$

It is known that

$$\mathsf{P}(N(s) \ge k) = \frac{s}{s + k - 1}, \ \ k \ge 1 \qquad (2.4)$$

(see, e.g., (Wilks 1959 or Nevzorov 2000)). Now let $N^{(1)}(s), N^{(2)}(s), ...$ be i.i.d. r.v.'s distributed in accordance with (2.4). Define the r.v.

$$N_n(s) = \max_{1 \le j \le n} N^{(j)}(s),$$

then, as it was shown in (Bening and Korolev 2008),

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

and for an asymptotically normal statistic T_n we have

$$\mathsf{P}\big(\sigma\sqrt{n}(T_{N_n(s)}-\mu) < x\big) \longrightarrow \Lambda_{1/s}(x), \quad n \to \infty,\ x \in \mathbb{R},$$

where $\Lambda_{1/s}(x)$ is the Laplace d.f. with parameter $\theta = 1/s$.

Denote

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

Theorem 2.3. *Let the statistic T_n have the form* (2.2)*, where $X_1, X_2, \ldots$ 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 $s \in \mathbb{N}$ the r.v. $N_n(s)$ has the distribution*

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

Then

$$\sup_x \Big|\mathsf{P}\big(\sigma\sqrt{n}(T_{N_n(s)}-\mu) < x\big) - \Lambda_{1/s}(x) - \frac{\mu_3\sigma^3 l_s(x)}{6\sqrt{n}}\Big| =$$

$$= O\Big(\frac{1}{n^{1/2+\delta}}\Big), \quad n \to \infty,$$

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

APPLICATION OF STUDENT DISTRIBUTION IN INSURANCE

In the 1960s, F. Bichsel suggested a risk rating system, called the Bonus-Malus system, which was better adjusted to the individual driver risk profiles. In the 1960s, car insurers requested approval for the increase of premium rates, claiming that the current level was insufficient to cover their risks. The supervision authority was prepared to give approval only if the rates took into account individual claims experience. It was no longer acceptable that "good" drivers, who had never made a claim, should continue to pay premiums which were at the same level as "bad" drivers who had made numerous claims.

Bichsel's Problem

Let N be the number of claims made by a particular driver in a year. The model used by Bichsel for the claim number is based on the following:

1. Conditionally, given $\Theta = \theta$, the N is Poisson distributed with Poisson parameter θ, i.e.

$$\mathsf{P}(N = k \mid \Theta = \theta) = e^{-\theta}\frac{\theta^k}{k!}, \quad k = 0, 1, \ldots.$$

2. Θ has a Gamma distribution with shape parameter r and a scale parameter β with the density

$$u(\theta) = \frac{\beta^r}{\Gamma(r)}\theta^{r-1}e^{-\beta\theta}, \quad \theta \geq 0.$$

The distribution function of Θ is called the structural function of the collective and describes the personal beliefs, a priori knowledge, and experience of the actuary.

The unconditional distribution of the number of claims is

$$\mathsf{P}(N = k) = \int_0^\infty \mathsf{P}(N = k \mid \Theta = \theta)u(\theta)d\theta =$$

$$= \int_0^\infty e^{-\theta}\frac{\theta^k}{k!}\frac{\beta^r}{\Gamma(r)}\theta^{r-1}e^{-\beta\theta}d\theta =$$

$$= C^k_{r+k-1}\, p^r(1-p)^k, \quad k = 0, 1, \ldots,$$

where $p = \frac{\beta}{\beta+1}$, and $N \equiv N_{p,r}$ is the negative binomial random variable with parameters p and r.

Approximation of the Aggregate Claim Amount

Consider the statistic which is the average of claim amounts

$$T_n = \frac{1}{n}\sum_{i=1}^n X_i,$$

where X_i is a claim size of each claim. Suppose that $X_1, \ldots, X_n$ are iid random variables, and $\mathsf{E}X_i = \mu$, $\mathsf{D}X_i = v^2$, $\sigma^2 = 1/v^2$. By CLT, we have

$$\mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) \longrightarrow \Phi(x), \quad n \to \infty.$$

From our results we have an approximate formula for the aggregate claim amount for small β

$$\sum_{i=1}^{N_{p,r}} X_i \approx \frac{1}{\sigma}\sqrt{\frac{p}{r}}\, N_{p,r}\, S_{2r} + \mu,$$

where $p = \frac{\beta}{\beta+1} \approx 0$, and S_{2r} is the Student distrubuted random variable with parameter $2r$.

Research supported by the Russian Foundation for Basic Research (projects 12-07-00115a, 12-07-00109a, 14-07-00041a).

REFERENCES

Asrabadi, B.R. 1985. "The exact confidence interval for the scale parameter and the MVUE of the Laplace distribution." *Communications in Statistics. Theory and Methods.* 14, 713–733.

Bening, V.E. and V.Yu. Korolev. 2008. "Some statistical problems related to the Laplace distribution." *Informatics and its Applications.* 2(2), 19–34.

Dadi, M.I. and R.J. Marks. 1987. "Detector relative efficiencies in the presence of Laplace noise." *IEEE Trans. Aerospace Electron. Systems.* 23(4), 568–582.

Duttweiler, D.L. and D.G. Messerschmitt. 1976. "Nearly instantaneous companding for nonuniformly quantizied PCM." *IEEE Trans. Comm.* 24(8), 864–873.

Epstein, B. 1948. "Application of the theory of extreme values in fracture problems." *J. Amer. Statist. Assoc.* 43(243), 403–412.

Hsu, D.A. 1979. "Long-tailed distributions for position errors in navigation." *Appl. Statist.* 28(1), 62–72.

Jones, P.N. and G.J. McLachlan. 1990. "Laplace-normal mixtures fitted to wind shear data." *J. Appl. Statistics.* 17(2), 271–276.

Kanji, G.K. 1985. "A mixture model for wind shear data." *J. Appl. Statistics.* 12(1), 49–58.

Korolev, V.Yu. 1995. "Convergence of random sequences with independent random indices. II." *Theory of Probability and Its Applications.* 40(4), 770–772.

Kotz, S., T.J. Kozubowski and K. Podgórski. 2001. *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance.* Birkhäuser, Boston.

Marks, R.J., G.L. Wise, D.G. Haldeman and J.L. Whited. 1978. "Detection in Laplace noise." *IEEE Trans. Aerospace Electron. Systems.* 14(6), 866–871.

Miller, J.H. and J.B. Thomas. 1972. "Detectors for discrete-time signals in non-Gaussian noise." *IEEE Trans. Inform. Theory.* 18(2), 241–250.

Nevzorov, V. B. 2000. *Records. Mathematical Theory.* Fazis, Moscow (in Russian).

Wilks, S.S. 1959. "Recurrence of extreme observations." *Journal of American Mathematical Society.*, 1(1), 106–112.

AUTHOR BIOGRAPHIES

VLADIMIR BENING is Doctor of Science in physics and mathematics; professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University; senior scientist, Institute of Informatics Problems, Russian Academy of Sciences. His email is `bening@yandex.ru`.

VLADISLAV SAVUSHKIN is PhD student, Dubna State University. His email is `savushkinva@mail.ru`

EGOR SHUNKOV is PhD student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Heard of Department of Applied Mathematics, Vologda State University; senior scientist, Institute of Informatics Problems, Russian Academy of Sciences; principal scientist, Institute of Territories Socio-Economic Development, Russian Academy of Sciences. His email is `a_zeifman@mail.ru` and his personal webpage at `http://uni-vologda.ac.ru/ zai/eng.html`.

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Russian Academy of Sciences. His email is `victoryukorolev@yandex.ru`.