On the Laplace and Student distributions as an alternative to the normal laws in some asymptotic problems of mathematical statistics

Vladimir Bening, Roman Korolev

Lomonosov Moscow State University

The 33rd Finnish Summer School on Probability Theory and Statistics

Päivölä

June 6th-10th, 2011

Introduction

In the communication, we discuss the transformation of the asymptotic distribution of asymptotically normal statistics if the sample size is replaced by a random variable.

Consider random variables $N_1, N_2, \ldots$ and $X_1, X_2, \ldots$ defined on the same probability space $(\Omega, \mathcal{A}, \mathrm{P})$. Assume that for each $n \geqslant 1$ the random variable N_n takes only natural values and is independent of iid sequence $X_1, X_2, \ldots$.

Introduction

Let $T_n = T_n(X_1, \ldots, X_n)$ be a statistic, i.e. a measurable function of the sample $X_1, \ldots, X_n$. We say that a statistic T_n is asymptotically normal with parameters $(\mu, 1/\sigma^2)$, $\sigma > 0$, if

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

where $\Phi(x)$ is the standard normal distribution function.
For each $n \geqslant 1$ define the random variables T_{N_n} by setting

$$T_{N_n}(\omega) \equiv T_{N_n(\omega)}(X_1(\omega), \ldots, X_{N_n(\omega)}(\omega)), \ \ \omega \in \Omega.$$

Introduction

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.

Main theorem

We can prove the following general result.
Theorem 1. *Let $\{d_n\}$ be an infinitely increasing sequence of positive numbers. Assume that $N_n \to \infty$ in probability as $n \to \infty$. Let a statistic T_n be asymptotically normal, i.e.*

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

Then

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

if and only if

$$\mathrm{P}(N_n < d_n x) \longrightarrow H(x), \quad n \to \infty, \quad x > 0,$$

and

$$F(x) = \int_0^\infty \Phi(x\sqrt{y})dH(y),$$

that is $F(x)$ is a scale mixture of normal law.

Sketch of the proof (only «if» part). By the formula of total probability we have

$$\mathrm{P}(\sigma\sqrt{d_n}(T_{N_n}-\mu)<x)-F(x)=$$

$$=\sum_{k=1}^{\infty}\mathrm{P}(N_n=k)\mathrm{P}(\sigma\sqrt{k}(T_k-\mu)<\sqrt{k/d_n}x)-F(x)=$$

$$=\sum_{k=1}^{\infty}\mathrm{P}(N_n=k)\Big(\Phi\big(\sqrt{k/d_n}x\big)-F(x)\Big)+$$

$$+\sum_{k=1}^{\infty}\mathrm{P}(N_n=k)\Big(\mathrm{P}(\sigma\sqrt{k}(T_k-\mu)<\sqrt{k/d_n}x)-\Phi\big(\sqrt{k/d_n}x\big)\Big)\equiv$$

$$\equiv J_1+J_2.$$

Sketch of the proof

Using the definition of $F(x)$, we obtain

$$J_1 = \int_0^\infty \Phi(x\sqrt{y})d\mathrm{P}(N_n < d_n y) - \int_0^\infty \Phi(x\sqrt{y})dH(y) =$$

$$= \int_0^\infty \Phi(x\sqrt{y})d\big(\mathrm{P}(N_n < d_n y) - H(y)\big) =$$

$$= -\int_0^\infty \big(\mathrm{P}(N_n < d_n y) - H(y)\big)d\Phi(x\sqrt{y}) \longrightarrow 0, \quad n \to \infty,$$

since

$$\mathrm{P}(N_n < d_n y) \longrightarrow H(y), \quad n \to \infty, \quad y > 0.$$

Sketch of the proof

Assuming for simplicity that the following rate of convergence holds

$$\sup_x \left| \mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) - \Phi(x) \right| \leqslant \frac{C}{n^\alpha}, \quad C > 0, \quad \alpha > 0.$$

Using this inequality, we have

$$|J_2| = \left| \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\Big(\mathsf{P}(\sigma\sqrt{k}(T_k - \mu) < \sqrt{k/d_n}x) - \Phi\big(\sqrt{k/d_n}x\big)\Big) \right| \leqslant$$

$$\leqslant C \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\frac{1}{k^\alpha} = C\ \mathsf{E}\frac{1}{N_n^\alpha} =$$

$$= \frac{C}{d_n^\alpha}\mathsf{E}(N_n/d_n)^{-\alpha} \longrightarrow 0, \quad n \to \infty,$$

since

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

Student distribution

Let $P_\gamma(x)$ be the Student distribution function corresponding to the density

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

where $\gamma > 0$ is the parameter (if γ is natural, then it is called *the number of degrees of freedom*). In our situation parameter γ may be arbitrary small, and we have typical heavy-tailed distribution.

Student distribution

Let $N_{p,r}$ be a random variables with the negative binomial distribution

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

Here $r > 0$ and $p \in (0,1)$ are parameters and for non-integer r the quantity C_{r+k-1}^{k-1} is defined as

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

In particular, for $r = 1$ random variable $N_{p,1}$ has the geometric distribution. It is well known that

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

so that

$$\mathsf{E}N_{p,r} \longrightarrow \infty, \quad p \to 0.$$

Student distribution

Theorem 2. *Let $r > 0$ be arbitrary. Assume that for any $n \geqslant 1$ the random variable $N_{p,r}$ has negative binomial distribution with parameters $p = 1/n$ and r. Let a statistic T_n is asymptotically normal.*
Then

$$\mathrm{P}(\sigma\sqrt{r\ n}(T_{N_{1/n,r}} - \mu) < x) \longrightarrow P_{2r}(x), \quad n \to \infty,$$

uniformly in $x \in \mathbb{R}^1$, where $P_{2r}(x)$ is the Student distribution with parameter $\gamma = 2r$.

Remark 1. *The Cauchy distribution* ($\gamma = 1$) *appears as a limit in Theorem 2 if the sample size* $N_{p,r}$ *has the negative binomial distribution with the parameters* $p = 1/n$ *and* $r = 1/2$.

Student distribution

Remark 2. *If the sample size $N_{p,r}$ has the negative binomial distribution with the parameters $p = 1/n$ and $r = 1$ (i.e. the geometric distribution with parameter $p = 1/n$), then, as $n \to \infty$, the limit distribution is Student with $\gamma = 2$ and the distribution function which can be found explicitly*

$$P_2(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{2 + x^2}}\right), \quad x \in \mathbb{R}^1.$$

Laplace distribution

Let $\Lambda_\gamma(x)$ be the Laplace distribution function corresponding to the density

$$\lambda_\gamma(x) \;=\; \frac{1}{2\gamma} e^{-|x|/\gamma}, \quad x \in \mathbb{R}^1,$$

where $\gamma > 0$ is the scale parameter.

Laplace distribution

Let $Y_1, Y_2, \ldots$ be iid random variables with continuous distribution function. Define random variables

$$N(m) = \min\{i \geqslant 1 : \max_{1\leqslant j\leqslant m} Y_j < \max_{m+1\leqslant k\leqslant m+i} Y_k\}.$$

It is well known that

$$\mathrm{P}(N(m) \geqslant k) = \frac{m}{m+k-1}, \quad k \geqslant 1.$$

Let $N^{(1)}(m), N^{(2)}(m), \ldots$ be iid random variables distributed as random variable $N(m)$ and define the random variables

$$N_n = \max_{1\leqslant j\leqslant n} N^{(j)}(m).$$

Laplace distribution

Theorem 3. *Let* $m \in \{1, 2, \ldots\}$ be arbitrary. Then

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

and for an asymptotically normal statistic T_n

$$\mathrm{P}(\sigma\sqrt{n}(T_{N_n} - \mu) < x) \longrightarrow \Lambda_{1/m}(x) \quad n \to \infty,$$

where $\Lambda_{1/m}(x)$ is the Laplace distribution function with the parameter $\gamma = 1/m$.

Laplace distribution

We also consider the accuracy of the Student and Laplace approximation. For example, the following result holds.
Theorem 4. *Assume that, as $n \to \infty$,*

$$\sup_x \Big| \mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) - \Phi(x) \Big| = \mathcal{O}(n^{-\alpha}), \quad n \to \infty, \quad \alpha > 0.$$

Then for $m \in \{1, 2, \dots\}$

$$\sup_x \Big| \mathsf{P}\Big(\frac{N_n}{n} < x\Big) - e^{-m/x} \Big| \leqslant \frac{C_m}{n}, \quad C_m > 0,$$

and

$$\sup_x \Big| \mathsf{P}(\sigma\sqrt{n}(T_{N_n} - \mu) < x) - \Lambda_{1/m}(x) \Big| = \mathcal{O}(n^{-\min\{1,\alpha\}}), \quad n \to \infty.$$

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.

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

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

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

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.

Bichsel's Problem

The unconditional distribution of the number of claims is

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

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

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

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 $2r$.