LMD konf. darbai, 380–383
1997, Vilnius

Bootstrap approximation for probabilities of large deviations

L. Saulis (VGTU, MII)

Let $X, X_1, \ldots, X_n, \ldots$ be a sequence of independent identically distributed random variables (r.v) with $\mathbf{E}X = 0$ and $\mathbf{E}X^2 = 1$. Denote

$$F_n(x) = \mathbf{P}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j > x\right).$$

Let

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X})^2,$$

denote the mean and variance of the sample $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$, and $\mathcal{X}^* = \{X_1^*, X_2^*, \ldots, X_n^*\}$ – resample drawn at random with replacement from $\mathcal{X}$, so that each X_i^* has probability $\frac{1}{n}$ of being equal to any given one of the X_j's:

$$\mathbf{P}(X_i^* = X_j \mid \mathcal{X}) = \frac{1}{n}, \quad 1 \leqslant i, j \leqslant n.$$

The bootstrap approximation to F_n is $\widehat{F}_n$, defined by

$$\widehat{F}_n(x) := \mathbf{P}\left(\frac{1}{\hat{\sigma}\sqrt{n}}\sum_{j=1}^{n}(X_j^* - \overline{X}) > x \,\middle|\, \mathcal{X}\right). \tag{1}$$

1. EDGEWORTH APPROXIMATION

If X is non-singular and if $\mathbf{E}|X|^{r+2} < \infty$ $(r \geqslant 1)$, then $F_n(x)$ an r-term Edgeworth expansion

$$F_n(x) = G_n(x) + o\left(n^{-r/2}\right), \quad n \to \infty,$$

uniformly in x, where

$$G_n(x) = 1 - \Phi(x) - \sum_{j=1}^{r} n^{-j/2} p_j(x)\varphi(x), \tag{2}$$

φ and Φ are standart normal density and d.f., and $p_j(x)$ is a polynomial of degree $2j-1$. The coefficients in $p_j(x)$ depend on moments $\mu_l = \mathbf{E}(X^l)$ up to the $(j+2)$-th.

An estimate $\hat{p}_j$ of the polynomial is obtainable by replacing the moment $\mu_l = \mathbf{E}(X^l)$ in coefficients by its standardised estimate

$$\hat{\mu}_l := \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X})^l/\hat{\sigma}^l.$$

Thus we are led to the empiric Edgeworth approximation:

$$\widehat{G}_n(x) = 1 - \Phi(x) - \sum_{j=1}^{r} n^{-j/2}\hat{p}_j(x)\varphi(x). \tag{3}$$

Both $\widehat{F}_n(x)$ and $\widehat{G}_n(x)$ are approximations to $F_n(x)$. Their relative perfomance it easy to deduce in the case of fixed x. Let

$$\hat{\rho}_n = \hat{\rho}_n(x) := \frac{\widehat{F}_n(x) - F_n(x)}{\widehat{G}_n(x) - G_n(x)}, \tag{4}$$

where $x = x(n)$ diverges as n increases.

In interpreting our next result the reader should bear in mind that „the smaller the value of $|\hat{\rho}_n|$, the better is the performance of the bootstrap". In particular, the bootstrap outperforms r-term Edgeworth approximation if $\hat{\rho} \to \infty$.

2. FORMULAE FOR LARGE DEVIATIONS

We say that the r.v. X satisfies condition (L), if $\exists : \gamma \geqslant 0$ and $\theta_0 > 0$ such that

$$\mathbf{E}\exp\{\theta_0|X|^{1/(1+\gamma)}\} < \infty \tag{L}$$

and Crammer's condition

$$\limsup_{|t|\to\infty} |\mathbf{E}e^{itX}| < 1. \tag{C}$$

THEOREM 1. *Let a r.v. X with $\mathbf{E}X = 0$ and $\mathbf{E}X^2 = 1$ satisfies the conditions (L) and (C), then for arbitrary integer r, $1 \leqslant r < \infty$, in the interval*

$$0 \leqslant x \leqslant (\sqrt{n})^{1/(1+2\gamma)}$$

as $n \to \infty$ the ralation of large deviations

$$\frac{F_n(x)}{1-\Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda^{[s+r]}\left(\frac{x}{\sqrt{n}}\right)\right\}\left\{1 + \sum_{j=0}^{r-1} n^{-j/2}L_j(x) + O\left(\left(\frac{x+1}{\sqrt{n}}\right)^r\right)\right\}$$

holds. Here $s = [1/r]$, $[a]$ is the integer part a,

$$\lambda^{[m]}(t) = \sum_{k=0}^{m-1} \lambda_k t^k.$$

In particular,

$$\lambda_0 = \frac{1}{6}\Gamma_3(X),$$
$$\lambda_1 = \frac{1}{24}\big(\Gamma_4(X) - 3\,\Gamma_3^2(X)\big),$$
$$\lambda_2 = \frac{1}{120}\big(\Gamma_5(X) - 10\,\Gamma_3(X)\Gamma_4(X) + 15\,\Gamma_3^3(X)\big), \ldots$$

Formulas for $L_j(x)$ are presented in (L. Saulis, V. Statulevičius. *Limit Theorems for large deviations.* Dordrecht, Boston, London: Kluwer Academic Publishers 1991, 232 p.). In particular,

$$L_0(x) \equiv 0,$$
$$L_1(x) = \frac{1}{6}\mathbf{E}(X^3)\{(x^2-1)\psi(x) - x^3\},$$
$$\psi(x) = \frac{\varphi(x)}{1-\Phi(x)} = x + \frac{1}{x} - \frac{2}{x^3} + \frac{10}{x^5} - \ldots$$

Note, that

$$L_1(x) = -\frac{1}{2}\mathbf{E}(X^3)\left\{\frac{1}{x} - \frac{4}{x^3} + O\left(\frac{1}{x^5}\right)\right\}$$

as $x \to \infty$.

THEOREM 2. *Let a r.v. X with $\mathbf{E}X = 0$ and $\mathbf{E}X^2 = 1$ satisfies the conditions (L) and (C), then for arbitrary integer r, $r \geqslant 1 < \infty$, in the interval*

$$0 \leqslant x \leqslant (\sqrt{n})^{1/(1+2\gamma)}$$

as $n \to \infty$ the ralation of large deviations

$$\frac{\widehat{F}_n(x)}{1-\Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\hat{\lambda}^{[s+r]}\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + \sum_{j=0}^{r-1} n^{-j/2}\widehat{L}_j(x) + \widehat{R}_r(x)\right)$$

holds. Here $\widehat{R}_r(x) = O\left(\left(\frac{x+1}{\sqrt{n}}\right)^r\right)$ with probability 1,

$$\hat{\lambda}^{[m]}(t) = \sum_{k=0}^{m-1}\hat{\lambda}_k t^k,$$

where $\hat{\lambda}_k$ are Crammer coefficients expressed in sample moments

$$\hat{\mu}_l := \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X})^l/\hat{\sigma}^l.$$

In particular,

$$\hat{\lambda}_0 = \frac{1}{6}\hat{\mu}_3, \qquad \hat{\lambda}_1 = \frac{1}{24}(\hat{\mu}_4 - 3\,\hat{\mu}_3^2 - 3), \ldots$$
$$\widehat{L}_0(x) \equiv 0,$$
$$\widehat{L}_1(x) = \frac{1}{6}\hat{\mu}_3\{(x^2-1)\psi(x) - x^3\} = -\frac{1}{2}\hat{\mu}_3\left\{\frac{1}{x} - \frac{4}{x^3} + O\left(\frac{1}{x^5}\right)\right\} \quad \text{if } x \to \infty.$$

3. COMPARISON OF BOOTSTRAP AND EDGEWORTH APPROXIMATION

$$\hat{\rho}_n = \hat{\rho}_n(x) := \frac{\widehat{F}_n(x) - F_n(x)}{\widehat{G}_n(x) - G_n(x)}.$$

(i) If $x \to \infty$ and $x = o(n^{(r-1)/(6r)})$ $(\gamma > \frac{2r+1}{2(r-1)}, r \geqslant 2)$, then $\hat{\rho}_n \to 1$ in probability.

In this case the bootstrap and r-term Edgeworth approximation to $F_n(x)$ are *asymptotically equivalent.*

(ii) If $\frac{x}{n^{(r-1)/(6r)}} \to \infty$ and $\frac{x}{n^{1/3}} \to 0$ $(\frac{1}{4} < \gamma < \frac{2r+1}{2(r-1)})$, then $\hat{\rho} \to 0$ in probability. In this case the bootstrap approximation to $F_n(x)$ is asymptotically much better than the r-th Edgeworth approximation.

(iii) The asymptotical behaviour of $\hat{\rho}_n$ is investigated as $0 < \gamma < 1/4$ as well. In this case the sign of moments $\mathbf{E}(X^3)$ is essential.

If $E(X^3) < 0$, then the bootstrap is asymptotically better than any r-term Edgeworth approximation. The situation is a little different if $E(X^3) > 0$. In that case the ratio of the absolute values of bootstrap and Edgeworth approximation errors diverges to $+\infty$ with probability $\frac{1}{2}$ and converges to unity with probability $\frac{1}{2}$. In this case the two methods are asymptotically equivalent with probability $\frac{1}{2}$, and the Edgeworth method is better with probability $\frac{1}{2}$.

REFERENCES

[1] P. Hall, On the relative performance of bootstrap and Edgeworth approximations of a distribution function, *J. Multivariate Anal.*, **35** (1990), 108–129.

[2] L. Saulis, V. Statulevičius, *Limit Theorems for Large Deviations*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1991.

„Bootstrap" aproksimacija didžiųjų nuokrypių zonose

L. Saulis

Darbe nagrinėjama nepriklausomų vienodai pasiskirsčiusių atsitiktinių dydžių $X_j, j = 1, 2, \ldots, n$, funkcijos $F_n(x) = \mathbf{P}(\sum_{j=1}^{n} X_j > x)$ aproksimavimas funkcija $\widehat{F}_n(x)$, apibrėžta lygybe (1) („bootstrap" aproksimacija) didžiųjų nuokrypių zonose. Įrodytas pastarosios ir Edgeworth aproksimacijų ekvivalentumas.