Dolne oszacowania wartości rekordowych

Agnieszka Goroncy
Uniwersytet Mikołaja Kopernika, Toruń

Tomasz Rychlik
Uniwersytet Mikołaja Kopernika, IM PAN, Toruń

XXXV Konferencja *Statystyka Matematyczna*, Wisła.

8 grudnia 2009

Rozważany model

$X_1, X_2, \ldots \overset{iid}{\sim} F$ - ciągła.

Momenty

$$\begin{aligned}
\mu &= \mathbb{E}X_1 = \int_0^1 F^{-1}(x)dx, \\
\sigma_p^p &= \mathbb{E}|X_1 - \mu|^p = \int_0^1 |F^{-1}(x) - \mu|^p \, dx > 0, \quad 1 \leq p < \infty, \\
\sigma_\infty &= \operatorname{ess\,sup} |X_1 - \mu| = \sup_{0<x<1} |F^{-1}(x) - \mu| \\
&= \max\{\mu - F^{-1}(0+), F^{-1}(1-) - \mu\}.
\end{aligned}$$

k-te rekordy

Czasy rekordowe:

$$
\begin{aligned}
T_0^{(k)} &= k, \\
T_n^{(k)} &= \min\left\{ j > T_{n-1}^{(k)} : \ X_j > X_{T_{n-1}^{(k)}+1-k:T_{n-1}^{(k)}} \right\}, \quad n = 1, 2, \ldots .
\end{aligned}
$$

Wartości rekordowe:

$$
\begin{aligned}
R_0^{(k)} &= X_k, \\
R_n^{(k)} &= X_{T_n^{(k)}+1-k:T_n^{(k)}}, \quad n = 1, 2 \ldots .
\end{aligned}
$$

Rozkłady n-tych wartości rekordowych

$$
\begin{aligned}
F_n^{(k)}(x) &= 1 - [1 - F(x)]^k \sum_{i=0}^{n} \frac{k^i}{i!} (-\ln[1 - F(x)])^i, \\
f_n^{(k)}(x) &= \frac{k^{n+1}}{n!} (-\ln[1 - F(x)])^n [1 - F(x)]^{k-1} f(x).
\end{aligned}
$$

Jeżeli $F \sim U(0,1)$, to

$$
\begin{aligned}
G_n^{(k)}(x) &= 1 - (1 - x)^k \sum_{i=0}^{n} \frac{k^i}{i!} [-\ln(1 - x)]^i, \quad 0 < x < 1, \\
g_n^{(k)}(x) &= \frac{k^{n+1}}{n!} [-\ln(1 - x)]^n (1 - x)^{k-1}, \quad 0 < x < 1.
\end{aligned}
$$

Sformuowanie problemu

Znaleźć optymalne dolne oszacowania dla

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_p} = \int_0^1 \frac{F^{-1}(x) - \mu}{\sigma_p} g_n^{(k)}(x)dx. \quad (1)$$

dla $1 \leq p \leq \infty$.

UWAGA: Oszacowania górne dla (1) są znane:

$\mathbb{E}\frac{R_n - \mu}{\sigma_2}$ - Nagaraja (1978),
$\mathbb{E}\frac{R_n - \mu}{\sigma_p}$ - Raqab (2000),
$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_2}$ - Raqab (1997),
$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_p}$ - Raqab, Rychlik (2002).

Klasyczne (zwykłe) rekordy ($k = 1$)

Oszacowania dolne dla

$$\mathbb{E}\frac{R_n - \mu}{\sigma_p} = \int_0^1 \frac{F^{-1}(x) - \mu}{\sigma_p} g_n(x) dx$$

są nieujemne.

Wyniki dla $1 < p \leq \infty$

Twierdzenie

$$\mathbb{E}\frac{R_n - \mu}{\sigma_p} \geq 0.$$

Gdy $1 < p < \infty$, *to* '$=$' *osiągana w granicy przez mieszaniny rozkładów jednostajnych na przedziałach* $[\alpha - \delta, \alpha + \delta]$ *oraz* $[\beta - \delta, \beta + \delta]$, *z prawdopodobieństwami* θ *oraz* $1 - \theta$, *odpowiednio, gdzie* $\alpha = \alpha(p, \theta, \delta) < \beta = \beta(p, \theta, \delta)$ *są rozwiązaniami równań:*

$$\begin{aligned}
\mu &= \theta\alpha + (1 - \theta)\beta, \\
2\delta(p + 1)\sigma_p^p &= \theta[(\mu - \alpha + \delta)^{p+1} - (\mu - \alpha - \delta)^{p+1}] \\
&+ (1 - \theta)[(\beta - \mu + \delta)^{p+1} - (\beta - \mu - \delta)^{p+1}],
\end{aligned}$$

gdy $\theta \nearrow 1$, *zaś* $\delta \searrow 0$.

Wyniki dla $1 < p \leq \infty$, c.d.

Gdy $p = \infty$, to '=' osiągana w granicy przez mieszaniny rozkładów jednostajnych na przedziałach $[\mu - \sigma_\infty, \mu - \sigma_\infty + \delta]$ oraz $[\mu + \frac{\theta}{1-\theta}(\sigma_\infty - \delta), \mu + \frac{\theta}{1-\theta}\sigma_\infty]$ z prawdopodobieństwami θ oraz $1 - \theta$, odpowiednio, gdy δ i θ dążą do 0, i przez mieszaniny rozkładów jednostajnych na przedziałach $[\mu - \frac{1-\theta}{\theta}\sigma_\infty, \mu - \frac{1-\theta}{\theta}(\sigma_\infty - \delta)]$ i $[\mu + \sigma_\infty - \delta, \mu + \sigma_\infty]$ z prawdopodobieństwami θ oraz $1 - \theta$, odpowiednio, gdy $\delta \searrow 0$ zaś $\theta \nearrow 1$.

Wyniki dla $p = 1$

Twierdzenie

$$\mathbb{E}\frac{R_n - \mu}{\sigma_1} \geq \frac{1}{2},$$

'=' osiągana w granicy przez mieszaniny rozkładów jednostajnych na przedziałach $[\mu - \frac{\sigma_1}{2\theta} - \delta, \mu - \frac{\sigma_1}{2\theta} + \delta,]$ *oraz* $[\mu + \frac{\sigma_1}{2(1-\theta)} - \delta, \mu + \frac{\sigma_1}{2(1-\theta)} + \delta,]$ *z prawdopodobieństwami* θ *oraz* $1 - \theta$*, odpowiednio, gdy* θ *i* δ *dążą jednocześnie do* 0.

UWAGA

Wszystkie rozkłady, dla których oszacowania są osiągane nie zależą od n.

Przyrosty zwykłych rekordów

Dla $1 \leq m < n$ z $1 \leq p \leq \infty$ oraz $m = 0 < n$ z $1 < p \leq \infty$ mamy

$$\mathbb{E}\frac{R_n - R_m}{\sigma_p} = \mathbb{E}\frac{R_n - \mu}{\sigma_p} - \mathbb{E}\frac{R_m - \mu}{\sigma_p} \geq 0,$$

Jedyny wyjątek dla $m = 0$, $p = 1$:

$$\mathbb{E}\frac{R_n - R_0}{\sigma_1} = \mathbb{E}\frac{R_n - \mu}{\sigma_1} \geq \frac{1}{2}, \qquad n \geq 1.$$

n-te wartości k-tych rekordów, $k \geq 2$

Oszacowania dla

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_p} = \int_0^1 \frac{F^{-1}(x) - \mu}{\sigma_p} g_n^{(k)}(x)dx$$

są niedodatnie.

Wyniki dla $1 < p < \infty$

Twierdzenie

Niech $k \geq 2$, zaś $1 < p < \infty$. Mamy

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_p} \geq -B_q = -B_q(k, n),$$

gdzie

$$\begin{aligned} B_q^q &= \alpha[g_n^{(k)}(\alpha) - g_n^{(k)}(d^*)]^q + \int_\alpha^{d^*} (g_n^{(k)}(x) - g_n^{(k)}(d^*))^q dx \\ &+ \int_{d^*}^1 (g_n^{(k)}(d^*) - g_n^{(k)}(x))^q dx, \end{aligned}$$

dla $d^* \in (\alpha, 1)$ będącego jedynym rozwiązaniem równania

$$\begin{aligned} \int_{d^*}^1 (g_n^{(k)}(d^*) - g_n^{(k)}(x))^{q-1} dx &= \int_\alpha^{d^*} (g_n^{(k)}(x) - g_n^{(k)}(d^*))^{q-1} dx \\ &+\alpha(g_n^{(k)}(\alpha) - g_n^{(k)}(d^*))^{q-1}, \end{aligned}$$

Wyniki dla $1 < p < \infty$, c.d.

oraz α, będącego jedynym rozwiązaniem równania

$$\alpha g_n^{(k)}(\alpha) = G_n^{(k)}(\alpha).$$

'=' zachodzi dla granicznego rozkładu postaci

$$\frac{F^{-1}(x)-\mu}{\sigma_p} = \begin{cases} -\frac{(\theta+g_n^{(k)}(\alpha))^{1/(p-1)}}{B_q}, & x \in [0,\alpha), \\ -\frac{(\theta+g_n^{(k)}(x))^{1/(p-1)}}{B_q}, & x \in [\alpha, d^*), \\ \frac{(-g_n^{(k)}(x)-\theta)^{1/(p-1)}}{B_q}, & x \in [d^*, 1). \end{cases}$$

Wyniki dla $p = 1$

Twierdzenie

Niech $k \geq 2$, zaś $p = 1$. Mamy

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_1} \geq -\frac{1}{2}g_n^{(k)}(\alpha) = -\frac{1}{2}\frac{k^{n+1}}{n!}[-\ln(1-\alpha)]^n(1-\alpha)^{k-1},$$

gdzie α jest jedynym rozwiązaniem równania

$$\alpha g_n^{(k)}(\alpha) = G_n^{(k)}(\alpha).$$

'$=$' zachodzi dla granicznego rozkładu trzypunktowego postaci

$$\begin{aligned} \mathbb{P}\left(X = \mu - \frac{\sigma_1}{2\alpha}\right) &= \alpha, \\ \mathbb{P}(X = \mu) &= 1 - \epsilon - \alpha, \\ \mathbb{P}\left(X = \mu + \frac{\sigma_1}{2\epsilon}\right) &= \epsilon, \end{aligned}$$

przy $\epsilon \longrightarrow 0$.

Wyniki dla $p = \infty$

Twierdzenie

Niech $k \geq 2$, $p = \infty$ oraz niech α będzie jedynym rozwiązaniem równania

$$\alpha g_n^{(k)}(\alpha) = G_n^{(k)}(\alpha).$$

Jeżeli $\alpha < \frac{1}{2}$, to

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_\infty} \geq 1 - 2G_n^{(k)}(\alpha) = \frac{1}{2^{k-1}}\sum_{i=1}^{n}\frac{(k\ln(2))^i}{i!} - 1,$$

zaś '$=$' zachodzi dla granicznego rozkładu dwupunktowego postaci

$$\begin{aligned}\mathbb{P}(X = \mu + \sigma_\infty) &= \frac{1}{2},\\ \mathbb{P}(X = \mu - \sigma_\infty) &= \frac{1}{2}.\end{aligned}$$

Wyniki dla $p = \infty$, c.d.

Jeżeli $\alpha > \frac{1}{2}$, *to*

$$\mathbb{E}\frac{R_n^{(k)} - \mu}{\sigma_\infty} \geq 1 - g_n^{(k)}(\alpha) = 1 - \frac{k^{n+1}}{n!}[-\ln(1-\alpha)]^n(1-\alpha)^{k-1},$$

zaś '$=$' zachodzi dla granicznego rozkładu dwupunktowego postaci

$$\begin{aligned} \mathbb{P}(X = \mu - \sigma_\infty \frac{1-\alpha}{\alpha}) &= \alpha, \\ \mathbb{P}(X = \mu + \sigma_\infty) &= 1 - \alpha. \end{aligned}$$

Przyrosty k-tych rekordów

Dla $p > 1$ lub $p = 1$, $1 \leq m < n$, mamy

$$\mathbb{E}\frac{R_n^{(k)} - R_m^{(k)}}{\sigma_p} \geq 0,$$

z jedynym wyjątkiem dla $m = 0$, $p = 1$:

$$\mathbb{E}\frac{R_n^{(k)} - R_0^{(k)}}{\sigma_1} \geq \frac{1}{2}.$$