

O PEWNEJ NIERÓWNOŚCI DLA MEDIAN L-STATYSTYK

Ryszard Zieliński

IMPAN Warszawa
e-mail: R.Zielinski@impan.gov.pl

TWIERDZENIE.

Jeżeli $T = \sum_{j=k}^m \lambda_j X_{j:n}$ jest L -statystyką taką, że $\lambda_k > 0$, $\lambda_m > 0$, $k \leq m$,
 $\lambda_k + \lambda_{k+1} + \dots + \lambda_m = 1$, to

$$(*) \quad m(U_{k:n}) \leq \text{Med}(F, F(T)) \leq m(U_{m:n})$$

gdzie

$\text{Med}(F, S)$ jest medianą rozkładu $S = S(X_1, \dots, X_n)$, gdy próba X_1, \dots, X_n pochodzi z rozkładu F

oraz

$m(U_{j:n})$ jest medianą j -tej statystyki pozycyjnej $U_{j:n}$ z próby o liczności n z rozkładu jednostajnego $U(0, 1)$.

Granice w nierówności (*) są dokładne w tym sensie, że

$$(\forall \varepsilon > 0)(\exists F \in \mathcal{F}) \text{Med}(F, F(T)) > m(U_{m:n}) - \varepsilon,$$

$$(\forall \eta > 0)(\exists G \in \mathcal{F}) \text{Med}(G, G(T)) < m(U_{k:n}) + \eta.$$

\mathcal{F} - rodzina wszystkich rozkładów o ciągłych i ściśle rosnących dystrybuantach

DOWÓD:

Z nierówności $X_{k:n} \leq T \leq X_{m:n}$ mamy

$$U_{k:n} = F(X_{k:n}) \leq F(T) \leq F(X_{m:n}) = U_{m:n}$$

Osiągalność prawego ograniczenia:

Bierzemy rodzinę rozkładów $F_\alpha(x) = x^\alpha$, $0 < x < 1$, $\alpha > 0$.

$$X_{j:n} = F_\alpha^{-1}(U_{j:n}) = U_{j:n}^{1/\alpha}$$

$$\begin{aligned} F_\alpha(T) &= \left(\lambda_k U_{k:n}^{1/\alpha} + \lambda_{k+1} U_{k+1:n}^{1/\alpha} + \dots + \lambda_{m-1} U_{m-1:n}^{1/\alpha} + \lambda_m U_{m:n}^{1/\alpha} \right)^\alpha \\ &= U_{m:n} \left[\lambda_k \left(\frac{U_{k:n}}{U_{m:n}} \right)^{1/\alpha} + \lambda_{k+1} \left(\frac{U_{k+1:n}}{U_{m:n}} \right)^{1/\alpha} + \dots \right. \\ &\quad \left. + \lambda_{m-1} \left(\frac{U_{m-1:n}}{U_{m:n}} \right)^{1/\alpha} + \lambda_m \right]^\alpha \end{aligned}$$

Gdy $\alpha \rightarrow 0$ to $F_\alpha(T) \rightarrow U_{m:n}$, więc $Med(F_\alpha, F_\alpha(T)) \rightarrow m(U_{m:n})$.

Osiągalność prawego ograniczenia: $G_\alpha(x) = 1 - (1 - x)^\alpha$

□

WNIOSEK.

Jeżeli L-statystyka

$$T = \sum_{j=k}^m \lambda_j X_{j:n}$$

$$k < m, \lambda_k > 0, \lambda_m > 0, \lambda_k + \lambda_{k+1} + \dots + \lambda_m = 1$$

$$\lambda_j = \lambda_j(q), j = k, \dots, m$$

jest estymatorem q -ego kwantyla $x_q(F)$ pewnego nieznanego rozkładu $F \in \mathcal{F}$, to błąd estymacji może być dowolnie duży w tym sensie, że

$$(\forall C > 0)(\exists F \in \mathcal{F})$$

$$|Med(F, T) - x_q(F)| > C$$

DOWÓD

Mamy

$$\frac{1}{2} = P_F\{T \leq \text{Med}(F, T)\} = P_F\{F(T) \leq F(\text{Med}(F, T))\}$$

czyli

$$\text{Med}(F, F(T)) = F(\text{Med}(F, T)).$$

Z Twierdzenia

$$(*) \quad m(U_{k:n}) \leq \text{Med}(F, F(T)) \leq m(U_{m:n})$$

czyli

$$F^{-1}(m(U_{k:n})) \leq \text{Med}(F, T) \leq F^{-1}(m(U_{m:n}))$$

oraz

$$F^{-1}(m(U_{k:n}))(-x_q(F)) \leq \text{Med}(F, T) - x_q(F) \leq F^{-1}(m(U_{m:n}))(-x_q(F)).$$

Dla $k < m$ mamy $m(U_{k:n}) < m(U_{m:n})$ więc $F^{-1}(m(U_{k:n}))(-x_q(F))$ może być dowolnie małe oraz $F^{-1}(m(U_{m:n}))(-x_q(F))$ może być dowolnie duże. Ograniczenia są osiągalne, więc $|\text{Med}(F, T) - x_q(F)|$ może być dowolnie duże. \square

ILUSTRACJA NUMERYCZNA (symulacje)

ESTYMATORY:

Davis and Steinberg (1986)

$$X_{(n+1)/2:n}, \quad \text{gdy } n \text{ jest nieparzyste;} \\ (X_{n/2:n} + X_{n/2+1:n})/2, \quad \text{gdy } n \text{ jest parzyste}$$

Harrell and Davis (1982)

$$HD = \frac{n!}{[(\frac{n-1}{2})!]^2} \sum_{j=1}^n \left[\int_{(j-1)/2}^{j/n} [u(1-u)]^{(n-1)/2} du \right] X_{j:n},$$

Kaigh and Cheng (1991) dla n nieparzystych

$$KC = \frac{1}{\binom{2n-1}{n}} \sum_{j=1}^n \binom{\frac{n-3}{2} + j}{\frac{n-1}{2}} \binom{\frac{3n-1}{2} - j}{\frac{n-1}{2}} X_{j:n}.$$

ROZKŁADY:

Pareto

$$1 - \frac{1}{x^\alpha}, \quad x > 1, \quad \text{tłuste ogony, nie ma momentów rzędu } k \geq \alpha,$$

Potęgowy (Beta)

$$x^\alpha, \quad x \in (0, 1), \quad \text{bez ogonów, wszystkie momenty,}$$

Wykładniczy

$$1 - \text{Exp}\{-\alpha x\}, \quad x > 0, \quad \text{bardzo regularny,}$$

$n = 9, sym = 9,999$

Distribution	Mediana	HD	KC	$\frac{X_{5:10}+X_{6:10}}{2}$
Pareto				
$\alpha = 1/2$	4	7.72	13.71	4.13
$\alpha = 1/4$	16	255	1107	18.45
$\alpha = 1/8$	256	3.3×10^6	2.8×10^7	383
Power				
$\alpha = 1/2$	0.25	0.2780	0.2919	0.2535
$\alpha = 1/4$	0.0625	0.1055	0.1286	0.0692
$\alpha = 1/8$	0.0039	0.0241	0.0432	0.0053
Exponential				
$\alpha = 1/2$	1.3863	1.5138	1.6235	1.4079
$\alpha = 1/4$	2.7726	3.0571	3.2731	2.8036
$\alpha = 1/8$	5.5452	6.0595	6.4897	5.6143