Liet. matem. rink., T. 43, spec. nr., 538–542

Imties pločio ir centro perkėlimo teorema

Darius PETRONAITIS (MII, KTU)
el. paštas: dariuspet@centras.lt

1. Įvadas

Sakykime, kad $\{X_1, X_2, \ldots, X_{N_n}\}$ – paprastoji atsitiktinė imtis iš generalinės aibės su pasiskirstymo funkcija $F(x)$, o$\{N_n, n \geqslant 1\}$ imties tūrius nusakanti, teigiamas reikšmes įgyjančių atsitiktinių didžių seka. Be to, $\frac{N_n}{n}$ pagal tikimybę konverguoja į atsitiktinį dydį τ $(P\{\tau < q\} = A(q))$.

Pažymėkime šiuos žymenis:

$$R_{N_n} = Z_{N_n} - W_{N_n}, \quad M_{N_n} = \frac{Z_{N_n} + W_{N_n}}{2},$$

čia $Z_{N_n} = \max(X_j, j = \overline{1, N_n})$, o $W_{N_n} = \min(X_j,\ j = \overline{1, N_n})$.

Šiame darbe tirsime tiesiškai normuotų statistikų R_{N_n} ir M_{N_n} greitųjų statistinių procedūrų skirstinių asimptotiką:

$$\lim_{n\to\infty} P\Big\{\frac{R_{N_n} - A_{1,n}}{B_{1,n}} < x\Big\}, \quad \lim_{n\to\infty} P\Big\{\frac{M_{N_n} - A_{2,n}}{B_{2,n}} < x\Big\}.$$

Tiriant tiesiškai normuotų greitųjų statistinių procedūrų asimptotiką, nagrinėsime tik netrivialų konvergavimą, kurio terminą pateikė L. de Hanas [1]. Silpnas ekstremalių statistikų darinių konvergavimas, kai vienas iš ekstremumų nusveria kitą, vadinamas trivialiuoju konvergavimu. Taigi, mus domina situacija, kai į galutinę išraišką savo indėlį įneša abu ekstremumai.

APIBRĖŽIMAS. Sakome, kad funkcija $F(x)$ priklauso maksimumo skirstinio $H_{i,\alpha}(x)$ (arba minimumo skirstinio $L_{i,\beta}(x)$), $i \in \{1, 2, 3\}$ traukos sričiai, jei egzistuoja tokios konstantų sekos $a_n, b_n > 0$, (arba $c_n, d_n > 0$), su kuriomis

$$\lim_{n\to\infty} F^n(a_n + b_n x) = H_{i,\alpha}(x), \tag{1}$$

$$(\text{arba } \lim_{n\to\infty}(1 - F(c_n + d_n x))^n = 1 - L_{i,\beta}(x)); \tag{2}$$

čia $H_{i,\alpha}(x)$(arba $L_{i,\beta}(x)$) – neišsigimusi pasiskirstymo funkcija. Simboliškai tai žymėsime taip: $F \in D_Z(H_{i,\alpha})$ (arba $F \in D_W(L_{i,\beta})$), $i \in \{1, 2, 3\}$.

$H_{i,\alpha}(x)$ ir $L_{i,\beta}(x)$, $i \in \{1, 2, 3\}$ pavidalas pateiktas [2] (2.4.1 ir 2.4.2 teoremos), čia α, β yra teigiamos konstantos. Vartosime šiuos žymenis:

$$x_* = \inf\{x\colon F(x) > 0\} \geqslant -\infty, \quad x^* = \sup\{x\colon F(x) < 1\} \leqslant \infty,$$

$$Z_{n,m} = \frac{Z_n - a_m}{b_m}, \quad W_{n,m} = \frac{W_n - c_m}{d_m},$$
$$R_{n,m} = \frac{R_n - A_{1,m}}{B_{1,m}}, \quad M_{n,m} = \frac{M_n - A_{2,m}}{B_{2,m}},$$

čia $b_n > 0$, $d_n > 0$, $B_{i,n} > 0$, $i = \overline{1,2}$ ir $a_n, c_n\ A_{i,n}$, $i = \overline{1,2}$ normalizavimo konstantos, o $n, m \geqslant 1$.

2. Rezultatai ir jų įrodymas

Teikiame teoremą, kuri apibendrina rezultatus gautus [3] straipsnyje (atsisakome N_n ir X_j, $j = \overline{1, N_n}$ nepriklausomumo).

Teorema. *Tarkime, kad* $F \in D_Z(H_{i,\alpha})$, $F \in D_W(L_{i,\alpha})$, $i \in \{1,2,3\}$ *ir kiekvienam* $\varepsilon > 0$ $\lim_{n\to\infty} P\Big\{\big|\frac{N_n}{n} - \tau\big| \geqslant \varepsilon\Big\} = 0$. Tada

$$\lim_{n\to\infty} P\{R_{N_n,n} < x\} = \Psi_{i,\alpha}(x), \qquad \lim_{n\to\infty} P\{M_{N_n} < x\} = \Phi_{i,\alpha}(x),$$

čia $\Psi_{i,\alpha}(x) = \int_0^\infty F_{i,\alpha}(x,q)\,\mathrm{d}A(q)$, $\Phi_{i,\alpha}(x) = \int_0^\infty T_{i,\alpha}(x,q)\,\mathrm{d}A(q)$, *o* $F_{i,\alpha}(x,q)$ *ir* $T_{i,\alpha}(x,q)$ *nusakomos taip*:

1. *Jei* $i = 1$ *ir* $0 < \rho < \infty$, *kai* $\rho = \lim_{x\to\infty} \frac{F(-x)}{1-F(x)}$, *tada*

$$F_{1,\alpha}(x,q) = 1 - \int_{-\infty}^{\infty} L_{1,\alpha}(q^{-\frac{1}{\alpha}}\rho^{-\frac{1}{\alpha}}(y-x))\,\mathrm{d}H_{1,\alpha}(q^{-\frac{1}{\alpha}}y), \tag{3}$$
$$T_{1,\alpha}(x,q) = \int_{-\infty}^{\infty} L_{1,\alpha}(q^{-\frac{1}{\alpha}}\rho^{-\frac{1}{\alpha}}(x-y))\,\mathrm{d}H_{1,\alpha}(q^{-\frac{1}{\alpha}}y). \tag{4}$$

Normalizavimo konstantos parenkamos taip:

$$A_{1,n} = a_n - c_n, \quad B_{1,n} = b_n, \quad A_{2,n} = \frac{a_n + c_n}{2}, \quad B_{2,n} = \frac{b_n}{2}. \tag{5}$$

2. *Jei* $i = 2$ *ir* $0 < \rho < \infty$, *kai* $\rho = \lim_{x\to+0} \frac{F(x_*+x)}{1-F(x^*-x)}$, *tada*

$$F_{2,\alpha}(x,q) = 1 - \int_{-\infty}^{\infty} L_{2,\alpha}(q^{\frac{1}{\alpha}}\rho^{\frac{1}{\alpha}}(y-x))\,\mathrm{d}H_{2,\alpha}(q^{\frac{1}{\alpha}}y), \tag{6}$$
$$T_{2,\alpha}(x,q) = \int_{-\infty}^{\infty} L_{2,\alpha}(q^{\frac{1}{\alpha}}\rho^{\frac{1}{\alpha}}(x-y))\,\mathrm{d}H_{2,\alpha}(q^{\frac{1}{\alpha}}y). \tag{7}$$

Normalizavimo konstantos parenkamos tokiu pat būdu, kaip pirmame punkte.

Jei $i = 3$ *ir egzistuoja konkreti diferencijuojama funkcija* P, *kuri taškui* x_* *gretimų iš dešinės taškų aibę atvaizduoja į taškui* x^* *gretimų iš kairės taškų aibę taip, kad*

$1 - F(P(x)) \sim F(x)$, $x \to x_* + 0$ *ir* $-\infty < \rho < 0$, *kai* $\rho = \lim_{x \to x_*+0} P'(x)$, *tada*

$$F_{3,0}(x, q) = 1 - \int_{-\infty}^{\infty} L_{3,0}(-\rho(y - x) + ln(q))\, \mathrm{d}H_{3,0}(y - ln(q)), \tag{8}$$

$$T_{3,0}(x, q) = \int_{-\infty}^{\infty} L_{3,0}(-\rho(x - y) + ln(q))\, \mathrm{d}H_{3,0}(y - ln(q)). \tag{9}$$

Normalizavimo konstantos parenkamo tokiu pat būdu, kaip pirmame punkte.

1 lema. *Tarkime* $(X_1, X_2, \ldots, X_n)$ *– paprastoji atsitiktinė imtis iš generalinės aibės su pasiskirstymo funkcija* $F(x)$ ($F \in D_Z(H_{ia}),\ F \in D_W(L_{ia}), ß \in \{1, 2, 3\}$), tuomet $\lim_{n\to\infty} P\{R_{n,n} < x\} = F_{i,\alpha}(x, 1)$, $i \in \{1, 2, 3\}$, *kurios pavidalai pateikti teoremoje. Tada atsitiktinio tūrio imčiai (*N_n *tenkina teoremos sąlygas)*:

$$\lim_{n\to\infty} P\{(R_{N_n,N_n} < x) \cap E\} = F_{i,\alpha}(x, 1)P(E), i \in \{1, 2, 3\}, \tag{10}$$

čia E *– bet koks diskretus atsitiktinis įvykis (nepriklausantis nuo* n.

Atlikus nežymius pakeitimus ji įrodoma tai pat, kaip [2] (6.2.3 lema). Dėl apimties ribotumo jos nepateiksiu.

2 lema. *Tarkime tenkinamos* 1 *lemos sąlygos. Tuomet* $\forall \varepsilon > 0$

$$\begin{aligned} &\lim_{n\to\infty} P\Big\{\Big|\frac{B_{1,n}}{B_{1,N_n}} - B^*_{1,\tau}\Big| \geqslant \varepsilon\Big\} = 0, \\ &\lim_{n\to\infty} P\Big\{\Big|\frac{A_{1,N_n} - A_{1,n}}{B_{1,N_n}} + A^*_{1,\tau}\Big| \geqslant \varepsilon\Big\} = 0, \end{aligned} \tag{11}$$

čia $B^*_{1,q}$ *ir* $A^*_{1,q}$, *apibrėžiamos sąryšiu*:

$$F_{i,\alpha}(x, q) = F_{i,\alpha}(A^*_{1,q} + B^*_{1,q}x, 1), \quad i \in \{1, 2, 3\}, \tag{12}$$

funkcijos $F_{i,\alpha}(x, q),\ i \in \{1, 2, 3\}$ *pavidalai pateikti teoremoje.*

Sprendžiame (12) lygtį. Pavyzdžiui, kai $i = 1$ atlikę integravimo kintamojo pakeitimą gauname $F_{1,\alpha}(x, q) = F_{1,\alpha}(xq^{-\frac{1}{\alpha},1})$, tuomet (12) lygtį pakeičiame lygtimi $F_{1,\alpha}(xq^{-\frac{1}{\alpha},1}) = F_{1,\alpha}(A^*_{1,q} + B^*_{1,q}x, 1)$. Be to įrodoma, kad $F_{1,\alpha}(x, 1)$ tolydi, nemažėjanti, tapatingai nelygi konstantai funkcija. Tuomet paskutinės lygties sprendiniai $B^*_{1,q} = q^{-\frac{1}{\alpha}},\ A^*_{1,q} = 0$. Išnagrinėję analogiškai kitus atvejus nustatome, kad $A^*_{1,q},\ B^*_{1,q}$ yra tolydžios ir monotoninės funkcijos q atžvilgiu ($B^*_{1,t}$ lygus $q^{\pm\frac{1}{\alpha}}$ arba 1, $A^*_{1,q}$ lygi nuliui arba $-(1 - \frac{1}{\rho}) \ln(q)$). Remiantis šiais faktais įrodome šią lemą naudodamiesi [2] (6.2.4 lemos) įrodymo metodika.

Teoremos įrodymas. Teoremą įrodysime statistikai R_{N_n}, o statistikai M_{N_n} įrodymas analogiškas, bet dėl straipsnio apimties ribojimo jo nepateiksime. Fiksuokime taškus $\tau_0 < \tau_1 < \ldots < \tau_m$ ir nagrinėkime įvykius: $E_0 = \{\tau < \tau_0\}$, $E_k = \{\tau_{k-1} \leqslant \tau < \tau_k\}$, $k = \overline{1, m}$ ir $E_{m+1} = \{\tau \geqslant \tau_m\}$. Tuomet remiantis 1 lema gauname:

$\lim_{n\to\infty} P\{(R_{N_n,N_n} < x) \cap E_k\} = F_{i,\alpha}(x,1)P(E_k)$, $k = \overline{0, m+1}$, $i \in \{1,2,3\}$. Pertvarkykime reiškinį R_{N_n,N_n} taip:

$$\begin{aligned} R_{N_n,N_n} &= R_{N_n,n}\frac{B_{1,n}}{B_{1,N_n}} + \frac{A_{1,n} - A_{1,N_n}}{B_{1,N_n}} \\ &= R_{N_n,n}B^*_{1,\tau} + A^*_{1,\tau} + (\frac{B_{1,n}}{B_{1,N_n}} - B^*_{1,\tau})R_{N_n,n} + \frac{A_{1,n} - A_{1,N_n}}{B_{1,N_n}} - A^*_{1,\tau}. \end{aligned}$$

Iš 2 lemos išplaukia, kad nagrinėjamos lygybės paskutinis narys artėja į nulį pagal tikimybę. Parodysime, kad ir priešpaskutinis narys artėja į nulį pagal tikimybę. Bet kokiam r teisinga nelygybė

$$P\left\{\left|\Big(\frac{B_{1,n}}{B_{1,N_n}} - B^*_{1,\tau}\Big)R_{N_n,n}\right| \geqslant \varepsilon\right\} \leqslant P\left\{|R_{N_n,n}| \geqslant r\right\} + P\left\{\left|\frac{B_{1,n}}{B_{1,N_n}} - B^*_{1,\tau}\right| \geqslant \frac{\varepsilon}{r}\right\},$$

o iš 2 lemos išplaukia, kad šios nelygybės paskutinis narys artėja į nulį. Kadangi $\lim_{n\to\infty} P\{|R_{N_n,n}| \geqslant r\} = 0$, $r \to \infty$, tai pertvarkymo antras narys artėja į nulį pagal tikimybę. Tuomet $\lim_{n\to\infty} P\{R_{N_n,N_n} < x\} = \lim_{n\to\infty} P\{R_{N_n,n}B^*_{1,\tau} + A^*_{1,\tau} < x\}$. Remiantis gautais faktais ir 1 lema gauname:

$$\lim_{n\to\infty} P\left\{(R_{N_n,n}B^*_{1,\tau} + A^*_{1,\tau} < x) \cap E_k\right\} = F_{i,\alpha}(x,1)P(E_k). \tag{13}$$

Kiekviename intervale $[\tau_{k-1}, \tau_k)$, $k = \overline{1, m}$ fiksuokime tašką $\tau(k)$. Tuomet, remiantis 2 lemos (12) formule pažymėtu sąryšiu ir (13) iš formulės išplaukia, kad

$$\begin{aligned} &\lim_{n\to\infty} P\left\{(R_{N_n,n}B^*_{1,\tau} + A^*_{1,\tau} < A_{\tau(k)} + B_{\tau(k)}x) \cap E_k\right\} \\ &\quad = F_{i,\alpha}(x,\tau(k))P(E_k). \end{aligned}$$

Kadangi $F_{i,\alpha}(x,1), A^*_{1,t}, B^*_{1,t} > 0$ – tolydžios funkcijos, o τ_{k-1} ir τ_k parinkti pakankamai artimi, tai $|P\{(R_{N_n,n} < x) \cap E_k\} - F_{i,\alpha}(x,\tau(k))P(E_k)| < \frac{\varepsilon}{m}$, $k = \overline{1, m}$, $n > n^*$. Jei τ_0, τ_m parenkamos taip, kad būtu tenkinama sąlyga $P(E_0) + P(E_{m+1}) < \varepsilon$, tai

$$\begin{aligned} &|P\{R_{N_n,n} < x\} - \int_0^\infty F_{i,\alpha}(x,q)\,\mathrm{d}P(\tau < q)| \\ &\quad = \Bigg|\sum_{k=0}^{m+1} P\{(R_{N_n,n} < x) \cap E_k\} - \sum_{k=1}^{m} F_{i,\alpha}(x,\tau(k))P(E_k) \\ &\qquad + \sum_{k=1}^{m} F_{i,\alpha}(x,\tau(k))P(E_k) - \int_0^\infty F_{i,\alpha}(x,q)\,\mathrm{d}P(\tau < q)\Bigg| \\ &\quad < \left|\sum_{k=1}^{m} P\{(R_{N_n,n} < x) \cap E_k\} - \sum_{k=1}^{m} F_{i,\alpha}(x,\tau(k))P(E_k)\right| \\ &\qquad + P(E_0) + P(E_{m+1}) \end{aligned}$$

$$+\left|\sum_{k=1}^{m} F_{i,\alpha}(x,\tau(k))P(E_k) - \int_{\tau_0}^{\tau_m} F_{i,\alpha}(x,q)\,\mathrm{d}P(\tau<q)\right|$$
$$+P(E_0)+P(E_{m+1}) < 4\varepsilon,$$

kai $n > n^*$. Kadangi $\varepsilon > 0$ laisvai parenkamas, tai

$$\lim_{n\to\infty} P\{R_{N_n,n} < x\} = \int_0^\infty F_{i,\alpha}(x,q)\,\mathrm{d}P(\tau<q), \quad i\in\{1,2,3\}.$$

Teorema įrodyta.

Literatūra

[1] L. de Haan, Weak limits of sample range, *Journal Applied Probability*, **11**, 836–841 (1974).
[2] Я. Галамбош, *Асимптотическая теория экстремальных порядковых статистик*, Наука, Москва (1984).
[3] A. Aksomaitis, D. Petronaitis, Greitųjų statistinių procedūrų perkėlimo teorema, *Liet. Matem. Rink.*, **42**(spec.nr.), 710–713 (2002).

Move theorem of the sample range and centre

D. Petronaitis

The limit distribution functions are obtained for the sample range and sample centre with random indices under nonrandom normalization.