

EGY ELOSZTÁSI PROBLÉMÁRA VONATKOZÓ HATÁRELOSZLÁSTÉTEL ÚJ BIZONYÍTÁSA

Írta: BÉKÉSSY ANDRÁS

IRWING WEISS bebizonyította a következő tételt [1]: Ha n számú urnában véletlenszerűen elosztunk N golyót, olyan módon, hogy bármelyik golyó a többitől függetlenül $1/n$ valószínűséggel eshet bármelyik urnába, továbbá, ha n és N a végtelenhez tart, és $\lim_{n \rightarrow \infty} N/n = \alpha > 0$ létezik, akkor az üresen maradó urnák $\xi_{n,N}$ száma aszimptotikusan normális eloszlású $ne^{-\alpha}$ középpértékkel és $\sqrt{ne^{-\alpha/2}} \sqrt{1 - (1 + \alpha)e^{-\alpha}}$ szórással.

RÉNYI ALFRÉD lényegesen enyhített feltételek mellett újból kimutatta a határeloszlás normális voltát [2]; tétele szerint a határeloszlás még akkor is normális, ha

$$(1) \quad N^2/n \rightarrow \infty, \quad N/n - \log n \rightarrow -\infty,$$

vagyis, ha $\lim_{n \rightarrow \infty} N/n$ nem létezik, ill. nullához vagy végtelenhez tart, azonban N/n változását az (1) feltételek alulról, ill. felülről korlátozzák. A teljes tétel bizonyítása három, a felhasznált eszközök szempontjából lényegesen eltérő módszerrel történt: a) az úgynevezett nyeregponthoz módszerrel arra az esetre, amikor N/n alulról és felülről is korlátos; b) a [3] cikkben kidolgozott módszerrel arra az esetre, amikor N/n alulról korlátos és $N/n - \log n \rightarrow -\infty$, végül c) speciális valószínűség-számítási megfontolásoknak és a centrális határeloszlástételnek felhasználásával arra az esetre, amikor N/n felülről korlátos és $N^2/n \rightarrow \infty$. Mind a b), mind a c) bizonyítás fedi az a) esetet is, — amely lényegében I. WEISS tétele, — tehát éppen a „legerősebb” eszközzel, a nyeregponthoz módszerrel végrehajtott vizsgálat hozta a leggyengébb a) eredményt. Ezért RÉNYI A. azt sejtette, hogy az enyhített feltételekre vonatkozó teljes tétel bebizonyítható egységes módszerrel is, és az alábbiakban ezt a sejtést kívánom igazolni.

A karakterisztikus függvények helyett az alábbiakban momentum-generáló függvényeket vezetünk be. Ennek előnye, hogy a tárgyalás áttekinthetőbb és kényelmesebb lesz. A momentum-generáló függvényeknek is — ha léteznek — I. H. CURTISS tételei szerint megvan az a folytonossági tulajdonságuk, ami a karakterisztikus függvényeknek, nevezetesen, ha a ξ_n valószínűségi változók sorozatának momentum-generáló függvényei egy $\Psi(t)$ korlátos határfüggvényhez konvergálnak egy a 0 pontot belsejében tartalmazó és $n \rightarrow \infty$ -re alulról korlátos hosszúságú — valós — intervallumon, akkor $\Psi(t)$ egy $F(x)$ eloszlásfüggvény momentum-generáló függvénye és ξ_n eloszlásfüggvényeinek sorozata a folytonossági pontokban egyenletesen $F(x)$ -hez tart.

A $\xi_{n,N}$ változóknak, az üres urnák számának különböző N értékekre vonatkozó $\Phi_{n,N}(t)$ karakterisztikus függvényeit generáló függvény [2], (10) szerint

$$G(t, z) = \sum_{N=0}^{\infty} \Phi_{n,N}(t) \frac{(nz)^N}{N!} = (e^z + e^{\mu} - 1)^n,$$

ahol

$$\Phi_{n,N}(t) = E\{e^{i\xi_{n,N}t}\},$$

ha tehát karakterisztikus függvények helyett a

$$\psi_{n,N}(t) = \Phi_{n,N}(-it) = E\{e^{\xi_{n,N}t}\}$$

momentum-generáló függvényeket vesszük kiindulási alapul, akkor az ezeket generáló függvény $(e^z + e^t - 1)^n$ és

$$(2) \quad \psi_{n,N}(t) = \frac{N!}{n^N 2\pi i} \oint \frac{(e^z + p)^n}{z^{N+1}} dz,$$

ahol $p = e^t - 1$.

TÉTEL. Ha $N \rightarrow \infty$ és $n \rightarrow \infty$,

$$(3) \quad N^2/n \rightarrow \infty$$

és

$$(4) \quad \frac{e^{N/n}}{n} \rightarrow 0,$$

akkor

$$(5) \quad \psi_{N,n} \left(\frac{s}{D_n \sqrt{n}} \right) e^{-\frac{s \sqrt{ne^{-\alpha_n}}}{D_n}} \rightarrow s^2/2,$$

$$\text{ahol } \alpha_n = \frac{N+1}{n} \text{ és } D_n^2 = e^{-\alpha_n} [1 - (1 + \alpha_n)e^{-\alpha_n}].$$

Az (5) relációból CURTISS említett tétele szerint következik, hogy

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\xi_{n,N} - ne^{-\alpha_n}}{D_n \sqrt{n}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

A bizonyításhoz a (2) integrált átalakítjuk. Legyen b olyan pont, amelyben (2) logaritmikus deriváltja eltűnik, tehát amely az

$$\frac{ne^b}{e^b + p} - \frac{N+1}{b} = 0,$$

vagy

$$(6) \quad b = \alpha_n(1 + pe^{-b})$$

egyenletnek tesz eleget. Az (5) egyenletnek mind pozitív, mind negatív p esetében egyetlen olyan valós pozitív gyöke van, amelyik p -vel nullához tart, ezt válasszuk kitétetett nyeregpontnak és ezt jelöljük a következőkben b -vel. Legyen $c = b - \alpha_n$, akkor

$$(7) \quad ce^c = p\alpha_n e^{-\alpha_n}.$$

Legyen a (2) egyenlet integráljának integrációs útja a 0 középpontú b sugarú kör. Legyen tehát $z = be^{i\varphi}$ és emeljük ki az integrál alól az integrandusnak a b nyereg-pontban felvett értékét. A (6) egyenlet felhasználásával a következő alakra jutunk:

$$(8) \quad \psi_{n,N}(t) = F \cdot \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left[1 + \frac{\alpha_n}{b} (e^{-b+be^{i\varphi}} - 1) \right]^n e^{-iN\varphi} d\varphi,$$

ahol

$$(9) \quad F = \frac{N!}{n^N} \cdot \frac{(e^b - p)^n}{b^N} \cdot \frac{1}{\sqrt{2\pi N}}.$$

Az integrál (8) alakja kényelmes lesz annak az esetnek tárgyalásánál, amelynél α_n felülről korlátos, ha azonban α_n felülről nem korlátos, — de korlátos alulról —, akkor kényelmesebb integrációs útnak a b ponton átmenő, az imaginárius tengellyel párhuzamos, mindkét irányban a végtelenbe húzódó egyenest választani. Könnyen belátható, hogy ez az út megengedett, ekvivalens a (2) egyenletben meghatározott, a 0 pontot körülvevő zárt görbével. Legyen tehát $z = b + bi\varphi$ és

$$(10) \quad \psi_{n,N}(t) = F \cdot \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 + \frac{\alpha_n}{b} (e^{-b+bi\varphi} - 1) \right]^n \frac{d\varphi}{(1+i\varphi)^{N+1}}.$$

Az F mennyiség a $p = e^{s/D_n \sqrt{N}} - 1$ és a (6) egyenlettel definiált b mennyiségen keresztül függ, s , n és N értékétől, az aszimptotikus értékének meghatározására szolgáló számolás tehát teljesen elemi; először is a tétel (3), ill. (4) feltétele szerint $p \rightarrow 0$, tehát $c/\alpha_n \rightarrow 0$, $\alpha_n p e^{-\alpha_n} = 0 \left(\frac{1}{\sqrt{N}} \right)$ és ennél fogva a (7) egyenlet ismételt felhasználásával

$$(11) \quad c = p\alpha_n e^{-\alpha_n} - p^2 \alpha_n^2 e^{-2\alpha_n} + 0 \left(\frac{1}{n\sqrt{N}} \right).$$

Alkalmazzuk továbbá (9)-re a *Stirling*-formulát, b helyett vezessük be c -t és akkor

$$(12) \quad \log F = nc + (n - N) \log \left(1 + \frac{c}{\alpha_n} \right) + o(1).$$

Ha α_n alulról korlátos, akkor (11) szerint $c/\alpha_n = 0 \left(\frac{1}{\sqrt{N}} \right)$ tehát (12)-ből F aszimptotikus kifejezése:

$$(13) \quad F \sim e^{s^2/2 + s\sqrt{N}e^{-\alpha_n/D_n}}$$

minden nehézség nélkül adódik. Ha azonban α_n alulról nem korlátos, hanem csökkenését csak a (3) szerinti $n\alpha_n^2 \rightarrow \infty$ feltétel korlátozza, akkor a (12)-ben szereplő $n \log \left(1 + \frac{c}{\alpha_n} \right)$ kifejezésnek $o(1)$ -ig való kifejtése némi nehézséget okoz, mert c/α_n

csak $O(1/\alpha_n\sqrt{n})$ nagyságrendű. Ebben az esetben például az

$$\begin{aligned} n \log \left(1 + \frac{c}{\alpha_n} \right) &= \\ &= n \log (1 + pe^{-\alpha_n} - p^2 \alpha_n e^{-2\alpha_n}) + 0 \left(\frac{1}{\alpha_n \sqrt{n}} \right) = \\ &= n \log (1 + p) + n \log \left(1 - \frac{p(1 - e^{-\alpha_n}) + \alpha_n p^2 e^{-2\alpha_n}}{1 + p} \right) + 0 \left(\frac{1}{\alpha_n \sqrt{n}} \right) \end{aligned}$$

átalakítás vezet célhoz, mert $p(1 - e^{-\alpha_n})$ már $O(1/\sqrt{n})$ nagyságrendű, és ennél fogva a logaritmikus kifejezések $O(1/n\sqrt{n})$ -ig könnyen kifejezhetők. Összevonások és egyszerűsítések után ugyancsak (13) adódik.

(8) szerint tehát

$$(14) \quad \psi_{n,N} \left(\frac{s}{D_n \sqrt{n}} \right) \sim e^{s^2/2 + s\sqrt{n}e^{-\alpha_n/D_n}} \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left[1 + \frac{\alpha_n}{b} (e^{-b+be^{i\varphi}} - 1) \right]^n e^{-iN\varphi} d\varphi$$

illetőleg (10) szerint

$$(15) \quad \psi_{n,N} \left(\frac{s}{D_n \sqrt{n}} \right) \sim e^{s^2/2 + s\sqrt{n}e^{-\alpha_n/D_n}} \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 + \frac{\alpha_n}{b} (e^{-b+bi\varphi} - 1) \right]^n \frac{d\varphi}{(1+i\varphi)^{N+1}}.$$

ennél fogva még csak annyit kell megmutatni, hogy a tétel feltételei mellett

$$(16) \quad \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left[1 + \frac{\alpha_n}{b} (e^{-b+be^{i\varphi}} - 1) \right]^n e^{-iN\varphi} d\varphi \rightarrow 1,$$

illetőleg

$$(17) \quad \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 + \frac{\alpha_n}{b} (e^{-b+bi\varphi} - 1) \right]^n \frac{d\varphi}{(1+i\varphi)^{N+1}} \rightarrow 1.$$

Az utóbbi állítások bizonyítását a nyeregponthoz (ill. Laplace-módszer) szokásos technikájával végezzük el, és pedíg a (16) integrálra, ha α_n felülről, a (17) integrálra, ha alulról korlátos.

Legyen α_n felülről korlátos. Osszuk a (16) integrált három részre:

$$\frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} + \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{\delta}^{\pi} + \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{-\delta} = I_1 + I_2 + I_3,$$

ahol δ tetszőlegesen kicsiny, de rögzített pozitív szám, értékéről később rendelkezünk. Az I_1 integrál, amely a nyeregponthoz (a $\varphi=0$ pont) közvetlen környezetét tartalmazza, a „lényeges” rész, az I_2 és I_3 az I_1 -hez képest kicsiny:

$$I_1 \rightarrow 1, \quad I_2 \text{ és } I_3 = o(1),$$

ugyanis

$$\begin{aligned} |I_2| &\leq \max_{\delta \leq \varphi \leq \pi} \frac{\sqrt{N\pi}}{\sqrt{2}} \left| 1 + \frac{\alpha_n}{b} (e^{-b+be^{i\varphi}} - 1) \right|^n \leq \\ &\leq \max_{\delta \leq \varphi \leq \pi} \frac{\sqrt{N\pi}}{\sqrt{2}} e^{-nb(1-\cos\varphi)} \left| 1 + \frac{p\alpha_n}{b} (e^{-be^{i\varphi}} - e^{-b}) \right|^n \leq \\ &\leq \max_{\delta \leq \varphi \leq \pi} \frac{\sqrt{N\pi}}{\sqrt{2}} e^{-nb(1-\cos\varphi) + n|p|\alpha_n e^{-b} C}, \end{aligned}$$

ahol C valamilyen állandó, amely csak b felső korlátjától és δ -tól függ. Ha α_n felülről korlátos, akkor b is és

$$|I_2| \leq \max_{\delta \leq \varphi \leq \pi} \frac{\sqrt{N\pi}}{\sqrt{2}} e^{-nb[(1-\cos\varphi) - |p| \cdot C \cdot \frac{\alpha_n}{b} e^{-\alpha_n \cdot \varphi}]} = o(1),$$

hiszen $\alpha_n/b \rightarrow 1$ és $|p| \rightarrow 0$, $1 - \cos \varphi$ pedig alulról korlátos, ha $\varphi \geq \delta$, míg $bn \sim N$.

Az I_1 integrál integrandusának logaritmusát fejtsük ki φ hatványai szerint φ^3 -ig, az eredmény

$$(18) \quad \left[1 + \frac{\alpha_n}{b} (e^{-b+be^{i\varphi}} - 1) \right]^n e^{-iN\varphi} = e^{-n\alpha_n \frac{\varphi^2}{2} (1+b-\alpha) + O(n\alpha\varphi^3)}$$

és ennél fogva (α_n felülről korlátos lévén)

$$\begin{aligned} I_1 &= \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} e^{-\frac{N\varphi^2}{2} [1+o(1)] + O(N\varphi^3)} d\varphi = \\ &= \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} e^{-\frac{N\varphi^2}{2} [1+o(1)]} d\varphi + O\left(\frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} N\varphi^3 e^{-\frac{N\varphi^2}{2} [1+O(\varphi)]} d\varphi\right) \sim 1 + M, \end{aligned}$$

ahol

$$\begin{aligned} M &= O\left(N^{3/2} \int_{-\delta}^{\delta} \varphi^3 e^{-\frac{N\varphi^2}{2} [1+o(1)+O(\varphi)]} d\varphi\right) < \\ &< O\left(N^{3/2} \int_{-\delta}^{\delta} \varphi^3 e^{-\frac{N\varphi^2}{2} [1-\eta(\delta)]} d\varphi\right) = O\left(\frac{1}{\sqrt{N}}\right), \end{aligned}$$

hiszen δ tetszőlegesen kicsiny lehet, válasszuk tehát olyan kicsinyre, hogy $1 - \eta(\delta) > 0$ legyen.

Ha α_n felülről nem, de alulról korlátos, akkor tekintsük (16) helyett (17) integrál-

ját. Mint előbb, bontsuk az integrált három részre

$$\frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} + \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{\delta}^{\infty} + \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{-\delta} = I_1 + I_2 + I_3.$$

Az I_1 lényeges rész integrandusát az előbbi módon sorbafejtve ugyancsak a (18) kifejezésre jutunk, most azonban, mivel felülről nem korlátos, a maradéktag $O(\alpha^3 \varphi^3)$ és az $I_1 \rightarrow 1$ igaz voltát csak abban az esetben tudjuk kimutatni az előbbi módszerrel, ha $\alpha^2 \delta \rightarrow 0$, vagyis ha $\delta = O(\alpha^{-2})$. Legyen tehát pl. $\delta = O(n^{-1/5})$ — ez elegendő a (4) feltétel szerint, — és akkor $I_1 \rightarrow 1$, továbbá

$$|I_2| \leq \frac{\sqrt{N}}{\sqrt{2\pi}} = \frac{(1+\delta^2)}{(1+\delta^2)^{N/2}} \max_{\delta \leq \varphi < \infty} \left| 1 + \frac{\alpha_n}{b} (e^{i\varphi} - 1) \right|^n \int_0^{\infty} \frac{d\varphi}{1+\varphi^2},$$

de

$$\left| 1 + \frac{\alpha_n}{b} (e^{i\varphi} - 1) \right|^n \leq \left| 1 + \frac{2\alpha_n}{b} \right|^n$$

és

$$\alpha_n/b = O(pe^{-\alpha_n}) = O(n^{-1/2}e^{-\alpha_n/2}) \text{ miatt}$$

továbbá

$$|I_2| = O(\sqrt{N}e^{-N\delta^2/2 + O(\sqrt{N})});$$

míg

$$\frac{\sqrt{N}}{N\delta^2} = \frac{1}{\alpha_n^{9/10} N^{1/10}} \rightarrow 0$$

és így $|I_2| = o(1)$.

A tétel ezzel bizonyítást nyert. Megjegyzendő, azt, hogy az $\alpha_n \rightarrow \infty$ esetben a (16) integrálja 1-hez tart, a (4)-nél enyhébb feltétel mellett ($\alpha_n^2 n^{-1/5} \rightarrow 0$) bizonyítottuk be, tehát még az $e^{\alpha_n}/n \rightarrow \gamma > 0$ esetben is igaz, hogy

$$\psi_{n,N}(t) \sim F(t).$$

$F(t)$ aszimptotikus értékét erre az esetre kiszámítva

$$\psi_{n,N}(t) \sim F(t) \sim e^{p/\gamma} = e^{1/\gamma(e^t - 1)}$$

adódik, és mivel ez a Poisson-eloszlás momentum-generáló függvénye, $\xi_{n,N}(t)$ határ-eloszlása az $1/\gamma$ középtértékű Poisson-eloszlás, ami az irodalomból jólismert eredmény (lásd pl. [5]).

IRODALOM

- [1] I. WEISS, Limiting distribution in some occupancy problems, *Applied Math. and Stat. Laboratory, Stanford University Technical Report*, No. 28, 1955.
- [2] A. RÉNYI: Three new proofs and a generalisation of a theorem of Irwing Weiss, *MTA Mat. Kut. Int. Közl., A. sorozat*, sajtó alatt.
- [3] P. ERDŐS—A. RÉNYI, On the central limit theorem for samples from a finite population, *MTA Mat. Kut. Int. Közleményei* 4 (1959) 49—61.
- [4] I. H. CURTISS, A note on the theory of moment generating functions, *Ann. Math. Stat.* 13 (1942) 430—433.
- [5] W. FELLER, *An introduction to probability theory and its applications*, Vol. I. 2. ed., 1957, 93—94.

(Beérkezett: 1962. IX. 17)