A RENDEZETT MINTÁK ELMÉLETÉRŐL

RÉNYI ALFRÉD lev. tag

Bevezetés

A matematikai statisztikában egyre nagyobb jelentőségre tesz szert a rendezett minták elmélete. A rendezett minták elméletének jelentőségét az adja meg, hogy eredményei bármely (folytonos) eloszlású statisztikai sokaságra alkalmazhatók; módszerei tehát ú. n. „nem-paraméteres" módszerek, amelyeknek alkalmazási köre sokkal tágabb, mint a „paraméteres" módszereké, amelyek csak bizonyos megadott típusú eloszlások esetében alkalmazhatók. A módszer másik fő előnye a gyakorlati alkalmazások szempontjából, — különösen a minőségellenőrzés matematikai statisztikai módszereit illetőleg — abban rejlik, hogy sokkal kevesebb számolást igényel, mint más ismert statisztikai módszerek.

A század eleje óta többen, így *K. Pearson* [1], *L. v. Bortkiewicz* [2], *E. L. Dodd* [3], *L. H. C. Tippet* [4] és *M. Fréchet* [5] foglalkoztak olyan részletkérdésekkel, amelyek a rendezett minták elmélete körébe sorolhatók, azonban szovjet matematikusok, *A. N. Kolmogorov* [6], *V. I. Glivenko* [7], *N. V. Szmirnov* [8], *B. V. Gnyegyenko* [9] és tanítványaik voltak azok, akik, felismerve a problémakör nagy elméleti és gyakorlati jelentőségét, azt összefüggő elméletté fejlesztették.

Különösen az elmult három évben számos cikk foglalkozott ezzel a problémakörrel; ezek közül kiemeljük *B. V. Gnyegyenko* és *V. S. Koroljuk* [10], *B. V. Gnyegyenko* és *E. L. Rvacseva* [11], *B. V. Gnyegyenko* és *V. S. Mihalevics* [12], *V. S. Mihalevics* [13], *J. D. Kvit* [14], *G. M. Mania* [15] és *I. I. Gihman* [16] munkáit. Eredményeik továbbfejlesztésébe számos más matematikus is bekapcsolódott: kiemeljük ezek közül *W. Feller* [17], *J. L. Doob* [18], *F. J. Massey* [19], *M. D. Donsker* [20], *T. W. Anderson* és *D. A. Darling* [21] munkáit. A problémakör 1947-ig terjedőleg viszonylag teljes bibliográfiája megtalálható *S. S. Wilks* [30] tanulmányában, amely 90 dolgozat felsorolását tartalmazza.

Jelen dolgozat célja egy új módszer bemutatása, amelynek segítségével a rendezett minták elméletének ma már klasszikus eredményei meglepő egyszerűséggel nyerhetők és amely lehetővé teszi számos új tétel bebizonyítását is. A módszer lényege abban áll, hogy a rendezett mintákra vonatkozó problémákat független valószínűségi változók összegeinek vizsgálatára vezeti vissza. Az 1. § tartalmazza a módszer ismertetését, a 2. § néhány ismert tétel egyszerű bizonyítását az új módszer segítségével; a 3. § a módszerrel elérhető új eredmények megfogalmazását tartalmazza az empírikus és elméleti eloszlásfüggvény összehasonlítására vonatkozólag, amelyek *A. N. Kolmogorov* és *N. V. Szmirnov* alapvető eredményeihez csatlakoznak. Jelentse $F_n(x)$ egy $F(x)$ folytonos eloszlású statisztikai sokaságból vett n elemű minta (empírikus) eloszlásfüggvényét (vagyis $F_n(x)$ az x-nél kisebb elemek relatív gyakoriságát jelenti a mintában); *Kolmogorov* meghatározta $|F_n(x)-F(x)|$ felső határának, *Szmirnov* pedig az $(F_n(x)-F(x))$ felső határának határeloszlását; a 4. §-ban az $\frac{F_n(x)-F(x)}{F(x)}$, ill. $\left|\frac{F_n(x)-F(x)}{F(x)}\right|$ viszonylagos eltérések felső határának határeloszlását határozzuk meg. Ehhez módszerünkön kívül a 4. §-ban közölt segédtételekre van szükség, amelyek *Erdős Pál* és *M. Kac* [22] eredményeit általánosítják. Ezek a segédtételek önmagukban is érdekességgel bírnak. Az 5. § a 3. §-ban megfogalmazott eredmények bizonyítását, a 6. § pedig a tételeinkben szereplő határeloszlásfüggvények numerikus kiszámítására vonatkozó megjegyzéseket és ezek kiszámítását szolgáló táblázatokat tartalmaz. Az 1. §-ban ismertetett módszerre *S. Malmquist* [23] egy tételének elemzése útján jutottam; az 1. § tartalmazza többek között *Malmquist* tételének egy új és egyszerű bizonyítását is. Egy másik egyszerű bizonyítást ugyanerre a tételre *Hajós Györggyel* együtt találtunk és egy közös dolgozatunkban [24] közöljük. Mindezek a vizsgálatok az Alkalmazott Matematikai Intézet Valószínűségszámítási és Matematikai Statisztikai Osztályainak közös szemináriumában folytatott megbeszélésekből indultak ki. A dolgozatban foglalt eredmények egy részét előadtam 1952 szeptemberében Krakowban és Wroclawban, továbbá 1953 januárjában a berlini Humboldt-egyetem kongresszusán*, valamint 1953 szeptemberében Vársóban a VIII. Lengyel Matematikai Kongresszuson. Ez utóbbi alkalommal *A. N. Kolmogorov* akadémikus néhány igen értékes megjegyzést tett, melyekért ezúton mondok neki köszönetet. Az Intézet munkatársai közül köszönettel tartozom *Lipták Tamás*nak, aki néhány részletszámítás kidolgozásával, továbbá *Palásti Ilonának* és *Várnai Péternének*, akik a numerikus számítások elvégzésével vettek részt a dolgozat elkészítésében.

1. §. Egy új módszer a rendezett minták elméletében

Kiindulunk a következő speciális esetből: legyen adva egy exponenciális eloszlású ζ valószínűségi változó értékére vonatkozó n-elemű minta, vagyis n független megfigyelés eredménye, amelyeket jelöljenek $\zeta_1, \zeta_2, \ldots, \zeta_n$; másszóval $\zeta_1, \zeta_2, \ldots, \zeta_n$ független és ugyanazon exponenciális eloszlásfüggvénnyel bíró valószínűségi változók ($n = 1, 2, 3, \ldots$). Szükségünk lesz az exponenciális

* Ez az előadás a kongresszus közleményeiben „Eine neue Methode in der Theorie der geordneten Stichproben" címen jelenik meg.

eloszlás következő jólismert tulajdonságára: ha ζ exponenciális eloszlású, úgy

$$P(\zeta < x+y \mid \zeta \geqq y) = P(\zeta < x) \tag{1.1}$$

(ha $x > 0$ és $y \geqq 0$)*.

Ez a tulajdonság egyébként egyértelműen jellemzi az exponenciális eloszlást. Legyen ugyanis ζ eloszlásfüggvénye $F(x)$, úgy

$$P(\zeta < x+y \mid \zeta \geqq y) = \frac{F(x+y)-F(y)}{1-F(y)},$$

tehát (1.1) ekvivalens a következő relációval:

$$\Phi(x+y) = \Phi(x)\Phi(y), \tag{1.2}$$

ahol $\Phi(x) = 1-F(x)$; márpedig ismeretes, hogy a $0 \leqq \Phi(x) \leqq 1$ feltételnek eleget tevő függvények közül a $\Phi(x) \equiv 0$ és $\Phi(x) \equiv 1$ triviális esetektől eltekintve a $\Phi(x) = e^{-\lambda x}$ $(\lambda > 0)$ függvényekre és csak ezekre áll fenn az (1.2) függvényegyenlet.

Különösen szemléletessé válik (1.1) jelentése, ha a ζ valószínűségi változót mint egy véletlentől függő időtartamú esemény időtartamát interpretáljuk. Ezen interpretáció mellett az (1.1) állítás úgy fogalmazható meg, hogy egy exponenciális eloszlású időtartamú esemény esetében, amely egy y időpontban még folyamatban van, az, hogy még meddig fog folyni, független y-tól, vagyis attól, hogy mióta folyik.

Ha a $\zeta_1, \zeta_2, \ldots, \zeta_n$ számokat nagyság szerint elrendezzük és bevezetjük a

$$\zeta_k^* = R_k(\zeta_1, \zeta_2, \ldots, \zeta_n) \quad (1 \leqq k \leqq n) \tag{1.3}$$

jelölést, ahol az $R_k(x_1, x_2, \ldots, x_n)$ n-változós függvény az $x_1, x_2, \ldots, x_n$ számok közül a nagyság szerint k-adikat jelenti $(k = 1, 2, \ldots, n)$, tehát például $\zeta_1^* = \min\limits_{1 \leqq k \leqq n} \zeta_k$ és $\zeta_n^* = \max\limits_{1 \leqq k \leqq n} \zeta_k$, úgy a $\zeta_1^* \leqq \zeta_2^* \leqq \ldots \leqq \zeta_n^*$ *rendezett minta* elemeinek eloszlása és együttes eloszlása igen egyszerűen meghatározható. Ebből a célból interpretáljuk a ζ_k változókat mint egymástól független események időtartamait; akkor ζ_k^* annak az eseménynek az időtartamát jelenti, amelyik k-adikként ér véget.

Számítsuk ki mindenekelőtt a $\zeta_{k+1}^* - \zeta_k^*$ különbségek eloszlásait. Ha $\zeta_k^* = y$, úgy

$$P(\zeta_{k+1}^* - \zeta_k^* > x \mid \zeta_k^* = y) = P(\zeta_{k+1}^* > x+y \mid \zeta_k^* = y) \tag{1.4}$$

és az (1.4) jobboldalán annak valószínűsége áll, hogy az y időpontban még folyamatban lévő $n-k$ esemény közül egyik sem ér véget $x+y$ időpontig; ennek értéke viszont (1.1) szerint

$$[P(\zeta > x)]^{n-k} = e^{(n-k)\lambda x}$$

és így $\zeta_{k+1}^* - \zeta_k^*$ feltételes eloszlásfüggvénye $\zeta_k^* = y$ feltétel mellett

$$P(\zeta_{k+1}^* - \zeta_k^* < x \mid \zeta_k^* = y) = 1 - e^{(n-k)\lambda x}. \tag{1.5}$$

* $P(A)$ az A esemény valószínűségét, $P(A \mid B)$ az A eseménynek a B feltételre vonatkozó feltételes valószínűségét jelöli.

Mivel az (1.5) alatti eloszlás nem függ y-tól, (1.5) egyben megadja $\zeta_{k+1}^*-\zeta_k^*$ feltétel nélküli eloszlásfüggvényét is; ugyanis a teljes valószínűség tétele szerint

$$P(\zeta_{k+1}^*-\zeta_k^*<x)=\int\limits_0^\infty P(\zeta_{k+1}^*-\zeta_k^*<x\,|\,\zeta_k^*=y)\,dG_k(y)=1-e^{-(n-k)\lambda x}, \quad (1.6)$$

ahol $G_k(y)$-nal jelöltük ζ_k^* eloszlásfüggvényét. Tehát a $\zeta_{k+1}^*-\zeta_k^*$ különbségek maguk is exponenciális eloszlásúak $\frac{1}{(n-k)\lambda}$ várható értékkel és így a

$$\delta_{k+1}=(n-k)(\zeta_{k+1}^*-\zeta_k^*) \qquad (k=0,1,2,\ldots,n-1) \qquad (1.7)$$

változók mind $\frac{1}{\lambda}$ várható értékű exponenciális eloszlással rendelkeznek. ((1.7)-ben definíciószerűen $\zeta_0^*\equiv 0$).

Az elmondottakból az is következik, hogy a $\delta_1,\delta_2,\ldots,\delta_n$ változók összességükben függetlenek. Ugyanis egyszerűen belátható, hogy a

$$P(\zeta_{k+1}^*-\zeta_k^*<x\,|\,\zeta_1^*=y_1,\zeta_2^*-\zeta_1^*=y_2,\ldots,\zeta_k^*-\zeta_{k-1}^*=y_k);(k=1,2,\ldots,n-1) \quad (1.8)$$

valószínűség nem függ az $y_1, y_2,\ldots,y_k$ változóktól; ez nyilvánvaló, hiszen a feltételek azt jelentik, hogy $\zeta_1^*=y_1,\zeta_2^*=y_1+y_2,\ldots,\zeta_k^*=y_1+y_2+\cdots+y_k$, vagyis megadják, hogy a $t=0$ időpontban egyidejűleg kezdődő n esemény közül elsőnek befejeződő k esemény mely időpontokban fejeződött be; a feltételek azt jelentik, hogy a $t=y_1+y_2+\cdots+y_k$ időpontban még $n-k$ esemény van folyamatban, és annak a valószínűsége, hogy ezek közül legalább egy véget ér $t+x$ időpont előtt, $1-e^{(n-k)\lambda x}$. Így tehát az (1.8) baloldalán álló valószínűség $1-e^{-(n-k)\lambda x}$-szel egyenlő, vagyis nem függ az $y_1,y_2,\ldots,y_n$ változóktól és ez equivalens azzal, hogy a $\zeta_{k+1}^*-\zeta_k^*$ változók és így a δ_k változók is $(k=1,2,\ldots,n)$ összességükben függetlenek. Ilyen módon tehát a ζ_k^* változók előállíthatók

$$\zeta_k^*=\frac{\delta_1}{n}+\frac{\delta_2}{n-1}+\cdots+\frac{\delta_k}{n-k+1}, \qquad k=1,2,\ldots,n \qquad (1.9)$$

alakban független és egyforma eloszlású változókból képezett lineáris kifejezések formájában. (1.9) úgy is kifejezhető, hogy *a ζ_k^* változók Markov-láncot (mégpedig ú. n. additív láncot) alkotnak.* (1.9) segítségével ζ_k^* eloszlása, ill. a ζ_k^* változók közül tetszőleges számú változó együttes eloszlása explicit alakban meghatározható.

Nézzük most meg, hogy az elmondottak hogyan alkalmazhatók általában rendezett minták vizsgálatára. Legyen ξ egy tetszőleges, folytonos és monoton *növekvő** $F(x)$ eloszlásfüggvénnyel rendelkező valószínűségi változó, legyen $(\xi_1,\xi_2,\ldots,\xi_n)$ a ξ változó értékére vonatkozó független megfigyelésekből álló n-elemű minta, vagyis legyenek $\xi_1,\xi_2,\ldots,\xi_n$ összességükben független és ugyanazon

* Azon, hogy $F(x)$ monoton növekvő, azt értjük, hogy szigorúan növekvő abban a legkisebb (a, b) intervallumban, melyre $F(a)=0$ és $F(b)=1$, ahol $a=-\infty$ és $b=+\infty$ is lehet.

(folytonos) $F(x)$ eloszlásfüggvénnyel rendelkező valószínűségi változók. Rendezzük el a ξ_k mintaelemeket nagyság szerint, vagyis képezzük a $\xi_k^* = R_k(\xi_1, \xi_2, \ldots, \xi_n)$ új változókat $(k = 1, 2, \ldots, n)$.

A rendstatisztika feladata a ξ_k^* változók vizsgálatában áll; ez azonban visszavezethető arra a speciális esetre, amikor a ζ_k változók exponenciális eloszlásúak (és így — az (1.9) képlet szerint — független valószínűségi változók összegeinek vizsgálatára), mégpedig a következőképpen: legyen

$$\eta_k = F(\xi_k) \quad \text{és} \quad \zeta_k = \log \frac{1}{\eta_k} \qquad (k = 1, 2, \ldots, n), \tag{1.10}$$

jelentse $\eta_k^* = F(\xi_k^*)$ az $\eta_1, \eta_2, \ldots, \eta_n$ változók közül a nagyság szerinti k-adikat, vagyis legyen $\eta_k^* = R_k(\eta_1, \eta_2, \ldots, \eta_n)$, továbbá legyen

$$\zeta_k^* = \log \frac{1}{\eta_{n+1-k}^*}. \qquad (k = 1, 2, \ldots, n). \tag{1.11}$$

Mivel $\log \frac{1}{x}$ monoton *csökkenő* függvény, tehát

$$\zeta_k^* = R_k(\zeta_1, \zeta_2, \ldots, \zeta_n), \qquad (k = 1, 2, \ldots, n) \tag{1.12}$$

vagyis ζ_k^* a $\zeta_1, \zeta_2, \ldots, \zeta_n$ változók közül a nagyság szerinti k-adik. Mivel a ξ_k változókról feltettük, hogy függetlenek, következik, hogy a ζ_k változók is függetlenek.

Vizsgáljuk most meg a ζ_k változók eloszlását. Ha $0 \leqq x \leqq 1$, úgy $y = F^{-1}(x)$-szel jelölve az $x = F(y)$ függvény inverz függvényét (amely — $F(x)$ szigorúan monoton növekvő lévén — a $0 \leqq x \leqq 1$ intervallumban egyértelműen definiálva van), nyerjük:

$$P(\zeta_k < x) = P\left(\log \frac{1}{F(\xi_k)} < x\right) = P(\xi_k > F^{-1}(e^{-x})) = 1 - F(F^{-1}(e^{-x})) = 1 - e^{-x},$$

ha $0 \leqq x \leqq 1$, tehát a $\zeta_1, \zeta_2, \ldots, \zeta_n$ változók függetlenek és exponenciális eloszlással bírnak, mégpedig 1 várható értékkel. Ilyen módon maguk a ξ_k^* valószínűségi változók előállíthatók

$$\xi_k^* = F^{-1}(e^{-\zeta_{n+1-k}^*}) = F^{-1}\left(e^{-\left(\frac{\delta_1}{n} + \frac{\delta_2}{n-1} + \ldots + \frac{\delta_{n+1-k}}{k}\right)}\right) \qquad (k = 1, 2, \ldots, n) \tag{1.13}$$

alakban, ahol $\delta_1, \delta_2, \ldots, \delta_n$ független és ugyanazon $1 - e^{-x}$ exponenciális eloszlásfüggvénnyel bíró valószínűségi változók. Eredményünkből az is következik, hogy az

$$\frac{\eta_k^*}{\eta_{k+1}^*} = e^{-\frac{\delta_{n+1-k}}{k}} \tag{1.14}$$

hányadosok függetlenek egymástól (itt definíciószerűen $\eta_{n+1}^* = 1$), tekintve, hogy a δ_{n+1-k}^* változók, mint láttuk, függetlenek. Egy másik következménye (1.13)-nak, hogy *a* $\xi_n^*, \xi_{n-1}^*, \ldots, \xi_1^*$ *változók Markov-láncot alkotnak*; ugyanis (1.13) szerint

$$\xi_{n-k}^* = F^{-1}\left[e^{-\log \frac{1}{F(\xi_{n-k+1}^*)} - \frac{\delta_{k+1}}{n-k}}\right], \tag{1.15}$$

továbbá ξ^*_{n-k+1} és δ_{k+1} függetlenek egymástól, tekintve, hogy

$$\xi^*_{n-k+1} = F^{-1}\left[e^{-\left(\frac{\delta_1}{n}+\frac{\delta_2}{n-1}+\cdots+\frac{\delta_k}{n-k+1}\right)}\right].$$

Arra a tényre, hogy a ξ^*_k változók ($k = 1, 2, \ldots, n$), vagyis egy rendezett minta elemei, Markov-láncot alkotnak, először *A. N. Kolmogorov* mutatott rá 1933-ban [6b] dolgozatában. A jelen dolgozatban foglalt új módszer ebből az alapvető megállapításból indul ki; ebben a tényben rejlő lehetőségek teljes kiaknázása azonban csak azáltal vált lehetségessé, hogy a $\zeta^*_k = \log \dfrac{1}{F(\xi^*_{n+k-1})}$ transzformáció segítségével a $\{\xi^*_k\}$ Markov-láncot additív lánccá alakítottuk át. Ezzel kapcsolatban érdekes megvizsgálni általában a következő kérdést: mely $\{\xi_t\}$ Markov-folyamatok esetében adható meg egy $G_t(x)$ függvénysereg úgy, hogy az $\eta_t = G_t(\xi_t)$ változók additív Markov-folyamatot alkossanak. Ehhez szükséges, hogy a $P(\xi_t < x \mid \xi_s = y) = F(x, s, y, t)$ eloszlásfüggvény eleget tegyen a

$$\frac{\partial F}{\partial x}\frac{\partial F}{\partial y}\left(\frac{\partial F}{\partial y}\frac{\partial^3 F}{\partial x^2 \partial y} - \frac{\partial F}{\partial x}\frac{\partial^3 F}{\partial x \partial y^2}\right) = \frac{\partial^2 F}{\partial x \partial y}\left(\frac{\partial^2 F}{\partial x^2}\left(\frac{\partial F}{\partial y}\right)^2 - \frac{\partial^2 F}{\partial y^2}\left(\frac{\partial F}{\partial x}\right)^2\right)$$

differenciálegyenletnek. A kérdésre más alkalommal vissza kívánunk térni.

Az η_k változók nyilvánvalóan egyenletes eloszlásúak a (0, 1) intervallumban, hiszen ha $0 < x < 1$, úgy

$$P(\eta_k < x) = P(\xi_k < F^{-1}(x)) = F(F^{-1}(x)) = x \tag{1.16}$$

és így az $\eta^*_k = R_k(\eta_1, \eta_2, \ldots, \eta_n)$ változók nem mások, mint egy, a (0, 1) intervallumban egyenletes eloszlású sokaságból vett n-elemű minta nagyság szerint elrendezett elemei.

(1. 14)-ből következik, hogy

$$\left(\frac{\eta^*_k}{\eta^*_{k+1}}\right)^k = e^{-\delta_{n+1-k}}$$

és mivel $P(e^{-\delta_{n+1-k}} < x) = P\left(\delta_{n+1-k} < \log\frac{1}{x}\right) = e^{-\log\frac{1}{x}} = x$, tehát *az* $\left(\frac{\eta^*_k}{\eta^*_{k+1}}\right)^k$ *változók függetlenek és egyforma, mégpedig egyenletes eloszlásúak* a (0, 1) intervallumban. Ez *S. Malmquist* tétele, melyet a bevezetésben említettünk; ennek a tételnek egy elemi bizonyítása a [24] dolgozatban található; a fent adott bizonyítás még egy fokkal egyszerűbb.

2. §. A rendezett minták elméletének felépítése az új módszerrel

Az elmondottak alapján a rendezett minták elemeinek határeloszlására vonatkozó eredményekhez igen egyszerűen lehet eljutni. Hogy erről képet adjunk, bebizonyítjuk a következő tételeket:*

* Ebben a §-ban is az 1. § ban bevezetett jelöléseket használjuk.

1. 1. tétel: Ha $k \geqq 1$ rögzített pozitív egész szám, úgy

$$\lim_{n\to\infty} P(n\zeta_k^* < x) = \int_0^x \frac{t^{k-1}e^{-t}}{(k-1)!}\,dt, \qquad (x>0),$$

vagyis $n\zeta_k^*$ határértékben k-adrendű Γ-eloszlással rendelkezik.

Bizonyítás: Mint láttuk,

$$\zeta_k^* = \frac{\delta_1}{n} + \frac{\delta_2}{n-1} + \cdots + \frac{\delta_k}{n+1-k},$$

ahol $\delta_1, \delta_2, \ldots, \delta_n$ független, exponenciális eloszlású valószínűségi változók, $1-e^{-x}$ $(x \geq 0)$ eloszlásfüggvénnyel; tehát $\frac{\delta_j}{n+1-j}$ sűrűségfüggvénye

$$(n+1-j)e^{-(n+1-j)x} \qquad (x>0)$$

és így egyszerű számolással következik, hogy ζ_k^* sűrűségfüggvénye

$$g_k(t) = \binom{n}{k} k e^{-nt} \cdot \sum_{j=0}^{k-1} (-1)^{k-1-j} \binom{k-1}{j} e^{jt},$$

tehát $n\zeta_k^*$ sűrűségfüggvénye

$$\frac{1}{n} g_k\left(\frac{t}{n}\right) = \binom{n-1}{k-1} e^{-t}\left(1-e^{-\frac{t}{n}}\right)^{k-1} = \frac{\left[n\left(1-e^{-\frac{t}{n}}\right)\right]^{k-1} e^{-t} \prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)}{(k-1)!}. \qquad (2.\ 1)$$

Mivel

$$\lim_{n\to\infty} n\left(1-e^{-\frac{t}{n}}\right) = t,$$

tehát $n\zeta_k^*$ sűrűségfüggvénye, ha $n \to \infty$, konvergál a

$$\frac{t^{k-1}e^{-t}}{(k-1)!}$$

függvényhez, vagyis egy k-adrendű Γ-eloszlás sűrűségfüggvényéhez.

Az eredményt előre láthattuk volna a következő meggondolással:

$$n\zeta_k^* = \delta_1 + \delta_2 + \cdots + \delta_k + \sum_{j=2}^{k} \frac{(j-1)\delta_j}{n+1-j}; \qquad (2.\ 2)$$

azonban $\delta_1+\delta_2+\cdots+\delta_k$ sűrűségfüggvénye $\frac{t^{k-1}e^{-t}}{(k-1)!}$, a $\sum_{j=2}^{k} \frac{(j-1)\delta_j}{n+1-j}$ változó viszont sztochasztikusan 0-hoz konvergál, ha $n \to \infty$.

A következő segédtételre van szükség, hogy ezt a meggondolást szabatos bizonyítássá építsük ki:

*1. lemma: Legyen $\gamma_n = \alpha_n + \beta_n$, ahol α_n és β_n valószínűségi változók, α_n eloszlásfüggvénye $G_n(x)$ továbbá**

$$\lim_{n\to\infty} M(\beta_n) = 0 \quad \textit{és} \quad \lim_{n\to\infty} D(\beta_n) = 0.$$

Tegyük fel, hogy a $G_n(x)$ eloszlásfüggvényeknek létezik a határértéke, vagyis fennáll, hogy

$$\lim_{n\to\infty} G_n(x) = G(x)$$

a $G(x)$ eloszlásfüggvény minden folytonossági helyén. Jelölje $F_n(x)$ a γ_n eloszlásfüggvényét; akkor

$$\lim_{n\to\infty} F_n(x) = G(x).$$

Bizonyítás: Feltehetjük az általánosság megszorítása nélkül, hogy $M(\beta_n) = 0$. Ekkor a Csebisev-egyenlőtlenség szerint $P(|\beta_n| > \varepsilon) < \dfrac{D^2(\beta_n)}{\varepsilon^2}$, tehát tetszőleges $\varepsilon > 0$ és $\delta > 0$ számokhoz megadható olyan $n_0(\varepsilon, \delta)$ pozitív egész szám, hogy ha $n > n_0$, úgy $P(|\beta_n| > \varepsilon) < \delta$. De akkor

$$F_n(x) = P(\gamma_n < x) \leqq P(\alpha_n < x + \varepsilon) + P(\beta_n < -\varepsilon); \tag{2.3}$$

ugyanis, ha $\alpha_n + \beta_n < x$, úgy vagy $\alpha_n < x + \varepsilon$, vagy $\alpha_n \geqq x + \varepsilon$, de ugyanakkor $\beta_n < -\varepsilon$ és így teljes valószínűség tételének alkalmazásával adódik (2.3). Hasonlóképpen

$$F_n(x) = P(\gamma_n < x) \geqq P(\alpha_n < x - \varepsilon) - P(\beta_n > \varepsilon); \tag{2.4}$$

ugyanis, ha $\alpha_n < x - \varepsilon$, akkor vagy $\alpha_n + \beta_n < x$, vagy $\alpha_n + \beta_n \geqq x$, de ugyanakkor $\beta_n > \varepsilon$. Tehát

$$G_n(x-\varepsilon) - \delta \leqq F_n(x) \leqq G_n(x+\varepsilon) + \delta.$$

Elvégezve az $n \to \infty$ határátmenetet és figyelembevéve, hogy ε és δ tetszőleges kicsinynek választhatók, következik, hogy

$$G(x-0) \leqq \varliminf_{n\to\infty} F_n(x) \leqq \varlimsup_{n\to\infty} F_n(x) \leqq G(x+0),$$

vagyis, hogy $G(x)$ minden x folytonossági helyén

$$\lim_{n\to\infty} F_n(x) = G(x).$$

Ezzel az 1. lemmát bebizonyítottuk.

Az 1.1. tétel ebből a lemmából azonnal következik, mivel

$$M\left(\sum_{j=2}^{k} \frac{(j-1)\delta_j}{n+1-j}\right) = \sum_{j=2}^{k} \frac{j-1}{n+1-j} \quad \text{és} \quad D^2\left(\sum_{j=2}^{k} \frac{(j-1)\delta_j}{n+1-j}\right) = \sum_{j=2}^{k} \frac{(j-1)^2}{(n+1-j)^2}$$

és így a lemma feltételei teljesülnek.

Az 1.1. tétel segítségével egyszerűen meghatározhatjuk η_k^* határeloszlását is; ugyanis $\eta_{n+1-k}^* = e^{-\zeta_k^*}$ és így

$$P\left(n(1-\eta_{n+1-k}^*) < x\right) = P\left(\zeta_k^* \leqq \log \frac{1}{1-\frac{x}{n}}\right).$$

* Itt és a továbbiakban $M(\beta)$-val a β valószínűségi változó várható értékét, $D(\beta)$-val pedig a szórását jelöljük.

Mivel $\lim_{n\to\infty} \frac{\log \frac{1}{1-\frac{x}{n}}}{\frac{x}{n}} = 1$ és a Γ-eloszlás folytonos, következik, hogy $n(1-\eta^*_{n+1-k})$ határértékben szintén k-adrendű Γ-eloszlású, $\frac{t^{k-1}e^{-t}}{(k-1)!}$ $(t \geqq 0)$ sűrűségfüggvénnyel. Mármost az η_k valószínűségi változók egyenletes eloszlásúak a $(0, 1)$ intervallumban és függetlenek; az egyenletes eloszlás szimmetriája következtében ugyanez áll fenn az $(1-\eta_k)$ változókra is $(k = 1, 2, \ldots, n)$ és így

$$\eta^*_k = R_k(\eta_1, \eta_2, \ldots, \eta_n) \quad \text{és} \quad 1-\eta^*_{k+1-k} = R_k(1-\eta_1, 1-\eta_2, \ldots, 1-\eta_n)$$

egyforma eloszlásúak. Ezért tehát fennáll a következő

*1.2. tétel: Az $n\eta^*_k$ és $n(1-\eta^*_{n+1-k})$ változók eloszlása, ha $k \geqq 1$ rögzítve van és $n \to \infty$, konvergál a $\frac{t^{k-1}e^{-t}}{(k-1)!}$ $(t \geqq 0)$ sűrűségfüggvényű k-adrendű Γ-eloszláshoz.**

Az 1.2. tétel segítségével meghatározhatjuk ξ^*_k és ξ^*_{n+1-k} határeloszlásait is; ezek azonban — ellentétben az η^*_k és ζ^*_k változók határeloszlásaitól — függni fognak az $F(x)$ eloszlásfüggvénytől. (Lásd [8]-at.)

Most bebizonyítjuk a következő

*1.3. tételt:*** η^*_k és η^*_{n+1-k} határértékben függetlenek, ha $n \to \infty$ ugyanakkor pedig k és j rögzítve vannak, vagyis*

$$\lim_{n\to\infty} P\left(\eta^*_k < \frac{x}{n}, 1-\eta^*_{n+1-j} < \frac{y}{n}\right) = \int_0^x \int_0^y \frac{u^{k-1}v^{j-1}}{(k-1)!\,(j-1)!} e^{-(u+v)}\,du\,dv.$$

Bizonyítás: Mindenekelőtt bebizonyítunk egy segédtételt:

2. lemma: Legyen $\gamma_n = \alpha_n + \beta_n$, ahol α_n és β_n valószínűségi változók, $\lim_{n\to\infty} M(\beta_n) = 0$ és $\lim_{n\to\infty} D(\beta_n) = 0$, továbbá legyen δ_n független α_n-től. Ha $F_n(x)$ jelenti α_n eloszlásfüggvényét és $G_n(y)$ a δ_n eloszlásfüggvényét, továbbá a $\lim_{n\to\infty} F_n(x) = F(x)$ és $\lim_{n\to\infty} G_n(y) = G(y)$ határeloszlások léteznek, akkor

$$\lim_{n\to\infty} P(\gamma_n < x, \delta_n < y) = F(x)\,G(y),$$

vagyis γ_n és δ_n határértékben függetlenekké válnak egymástól.

* Az η^*_k változók eloszlása véges n-re pontosan is kiszámítható és határátmenettel némi számolással így is bebizonyítható az 1.2. tétel (l. H. Cramér [25] 370—371). Ezt a tételt itt azért bizonyítottuk be módszerünkkel, hogy annak felhasználását először egy egyszerű eseten keresztül mutassuk meg.

** Lásd H. Cramér [25] 371. Ezt az ismert tételt azért tárgyaljuk itt, mert módszerünk jobban megvilágítja a tételben kifejezésre jutó tény mélyebb hátterét, mint a tétel ismert bizonyítása.

Bizonyítás: Válasszuk (mint az 1. lemma bizonyitásánál) n értékét olyan nagyra, hogy ha $n > n_0$, úgy $P(|\beta_n| > \varepsilon) < \delta$. — Az 1. lemma bizonyításánál alkalmazott meggondoláshoz hasonlóan belátható, hogy ha $n > n_0$, ahol n_0 az ε és δ pozitív számok megválasztásától függ, úgy

$$P(\alpha_n < x-\varepsilon, \delta_n < y)-\delta \leqq P(\gamma_n < x, \delta_n < y) < P(\alpha_n < x+\varepsilon, \delta_n < y)+\delta \quad (2.5)$$

és mivel feltevésünk szerint α_n és δ_n függetlenek, tehát

$$P(\alpha_n < x \pm \varepsilon, \delta_n < y) = F_n(x \pm \varepsilon) G_n(y)$$

és ugyanúgy, mint az 1. lemma bizonyitásánál, nyerjük, hogy olyan x értékekre, amelyek $F(x)$-nek folytonossági helyei

$$\lim_{n\to\infty} P(\gamma_n < x, \delta_n < y) = F(x)G(y).$$

Ezzel a 2. lemmát bebizonyítottuk.

Mármost

$$\zeta^*_{n+1-k} - \log n = \frac{\delta_{j+1}}{n-j} + \cdots + \frac{\delta_{n+1-k}}{k} - \log n + \zeta^*_j,$$

ahol $\zeta^*_j = \frac{\delta_1}{n} + \frac{\delta_2}{n-1} + \cdots + \frac{\delta_j}{n+1-j}$ és így

$$\lim_{n\to\infty} M(\zeta^*_j) = \lim_{n\to\infty} D(\zeta^*_j) = 0.$$

Mivel továbbá a $\frac{\delta_{j+1}}{n-j} + \cdots + \frac{\delta_{n+1-k}}{k} - \log n$ összeg független ζ^*_j-tól és $\lim_{n\to\infty} n \log \frac{1}{1-\frac{y}{n}} = y$, továbbá már tudjuk, hogy

$$\lim_{n\to\infty} P(n\zeta^*_j < y) = \int_0^y \frac{t^{j-1}}{(j-1)!} e^{-t}\, dt,$$

tehát

$$\lim_{n\to\infty} P\left(\eta^*_k < \frac{x}{n}, 1-\eta^*_{n+1-j} < \frac{y}{n}\right) = \lim_{n\to\infty} P\left(\zeta^*_{n+1-k} - \zeta^*_j \geqq \log\frac{n}{x}\right) \cdot \int_0^y \frac{t^{j-1}e^{-t}}{(j-1)!}dt. \quad (2.6)$$

De az 1. lemma szerint

$$\lim_{n\to\infty} P\left(\zeta^*_{n+1-k} - \zeta^*_j \geqq \log\frac{n}{x}\right) = \lim_{n\to\infty} P\left(\zeta^*_{n+1-k} \geqq \log\frac{n}{x}\right) \quad (2.7)$$

és az 1.2. tétel szerint

$$\lim_{n\to\infty} P\left(\zeta^*_{n+1-k} \geqq \log\frac{n}{x}\right) = \lim_{n\to\infty} P\left(\eta^*_k < \frac{x}{n}\right) = \int_0^x \frac{t^{k-1}e^{-t}}{(k-1)!}\, dt. \quad (2.8)$$

A (2.6), (2.7) és (2.8) összefüggésekből az 1.3. tétel állítása következik.

Az 1.3. tétel segítségével kiszámíthatjuk az $\eta^*_n - \eta^*_1$ különbség határeloszlását is. Ennek azért van jelentősége, mert $\eta^*_n - \eta^*_1$ nem más, mint az

$(\eta_1, \eta_2, \ldots \eta_n)$ minta *terjedelme* és egy minta terjedelmét fel lehet használni az eloszlás szórásának becslésére. Mivel nagy n-re 1-hez közeli valószínűséggel η_n^* közel van 1-hez és η_1^* közel van 0-hoz, nyilván az $n[1-(\eta_n^*-\eta_1^*)]$ változót kell vizsgálnunk, mivel $n\eta_1^*$ és $n(1-\eta_n^*)$ az 1.3. tétel szerint határértékben függetlenek, összegük határeloszlása egyenlő határeloszlásaik kompozíciójával. Mivel mind $n\eta_1^*$, mind pedig $n(1-\eta_n^*)$ határeloszlásának sűrűségfüggvénye e^{-x}, ha $x\geqq 0$, $n[1-(\eta_n^*-\eta_1^*)]$ határeloszlásának sűrűségfüggvénye $\int\limits_0^x e^{-(x-y)}e^{-y}dy=xe^{-x}$, ha $x\geqq 0$, tehát $n[1-(\eta_n^*-\eta_1^*)]$ határértékben 2-rendű Γ-eloszlású valószínűségi változó.

Az η_k^* valószínűségi változók határeloszlásai segítségével a ξ_k^* valószínűségi változók határeloszlásai meghatározhatók; ehhez azonban, mivel $\xi_k^*=F^{-1}(\eta_k^*)$, szükséges az $F^{-1}(x)$ inverz függvény ismerete. Mivel ez az inverz függvény az esetek többségében nem írható fel explicit alakban, ez bizonyos számítástechnikai nehézségekkel jár.

Eddig a rendezett minta ξ_k^* elemeinek (ill. η_k^* és ζ_k^* transzformáltjainak) határeloszlását azon feltevés mellett vizsgáltuk, hogy a k sorszám (ill. $j=n+1-k$) rögzítve van és ugyanakkor $n\to\infty$; ezt a problémakört a minta „szélső elemei" vizsgálatának nevezik. Most áttérünk a ξ_k^* (ill. η_k^*, ill. ζ_k^*) elemek határeloszlásának vizsgálatára azon feltevés mellett, hogy n-el együtt k is végtelenhez tart, mégpedig úgy, hogy $|k-nq|$ korlátos, ahol q állandó $(0<q<1)$. A ξ_{k+1}^* változóra, ahol $k=[nq]$, ez teljesül; ezt a változót a minta *q-quantilisének* nevezik; speciálisan, ha $n=2m+1$, úgy ξ_{m+1}^* a minta mediánja; nyilvánvaló, hogy a minta q-quantilise nem más, mint a minta empírikus eloszlásfüggvényének q-quantilise és így a minta mediánja nem más, mint a minta empírikus eloszlásfüggvényének mediánja; ennek megfelelően, ha n páros, $n=2m$, úgy a minta mediánjának a $\frac{\xi_m^*+\xi_{m+1}^*}{2}$ számot nevezzük.

Be fogjuk bizonyítani a következő tételt,* amely tartalmazza azt az állítást, hogy a minta q-quantilisei határértékben normális eloszlásúak, feltéve, hogy az $F(x)$ eloszlásfüggvény eleget tesz bizonyos egyszerű feltételeknek.

1.4. tétel: A $\xi_1, \xi_2, \ldots, \xi_n$ független valószínűségi változók közös $F(x)$ eloszlásfüggvényének létezzék az $f(x)=F'(x)$ sűrűségfüggvénye és tegyük fel, hogy az (a, b) intervallumban, $(a<b)$, $f(x)$ folytonos és pozitív; akkor, ha $0<F(a)<q<F(b)<1$, továbbá, ha $|k_n-nq|$ korlátos $\left(\text{tehát a fortiori } \lim\limits_{n\to\infty}\frac{k_n}{n}=q\right)$, *úgy $\xi_{k_n}^*$ határértékben normális eloszlású $Q=F^{-1}(q)$ körül, vagyis az $F(x)$*

* Az irodalomban ez a tétel a $\left|\frac{k_n}{n}\right|=q$ feltevés mellett szerepel; a $|k_n-nq|\leqq A$ feltétel ennél kevesebbet kíván meg.

eloszlásfüggvény q-quantilise, mint középérték körül, mégpedig

$$\lim_{n\to\infty} P\left(\frac{\xi^*_{k_n}-Q}{\frac{1}{f(Q)}\sqrt{\frac{q(1-q)}{n}}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt.$$

Bizonyítás: Először vizsgáljuk meg ζ^*_{n+1-k} határeloszlását:

$$\zeta^*_{n+1-k}=\sum_{j=1}^{n+1-k_n}\frac{\delta_j}{n+1-j}, \tag{2.9}$$

ahol a δ_j változók függetlenek és exponenciális eloszlásúak, $1-e^{-x}$ $(x\geqq 0)$ eloszlásfüggvénnyel, tehát $M(\delta_j)=D^2(\delta_j)=1$ és

$$M(|\delta_j-1|^3)=\int_0^\infty |x-1|^3e^{-x}dx<\int_0^1 e^{-x}dx+\int_0^\infty x^3e^{-x}dx\leqq 7.$$

De akkor

$$M_n=M(\zeta^*_{n+1-k_n})=\sum_{j=1}^{n+1-k_n}\frac{1}{n+1-j}$$

$$S_n^2=D^2(\zeta^*_{n+1-k_n})=\sum_{j=1}^{n+1-k_n}\frac{1}{(n+1-j)^2} \tag{2.10}$$

és

$$K_n^3=\sum_{j=1}^{n+1-k_n}M\left(\left|\frac{\delta_j-1}{n+1-j}\right|^3\right)\leqq 7\sum_{j=1}^{n+1-k_n}\frac{1}{(n+1-j)^3}$$

és így

$$\frac{K_n}{S_n}\leqq\frac{7}{k_n} \tag{2.11}$$

tehát, ha $n\to\infty$, úgy $\frac{K_n}{S_n}\to 0$; ez azt jelenti, hogy a (2.9) összegek sorozatára alkalmazható a központi határeloszlástétel Ljapunov-féle alakja és így

$$\lim_{n\to\infty} P\left(\frac{\zeta^*_{n+1-k_n}-M_n}{S_n}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{\infty}^{x}e^{-\frac{t^2}{2}}dt. \tag{2.12}$$

Mármost ismeretes, hogy

$$\sum_{k=1}^{m}\frac{1}{k}=\ln m+C+\Delta_m,$$

ahol C az Euler-féle állandó, továbbá megadható olyan $A>0$ állandó, hogy $|\Delta_m|<\frac{A}{m}$. Egyszerű számolással belátható továbbá, hogy ha $H\geqq h\geqq 2$, úgy

$$\sum_{k=h}^{H}\frac{1}{k^2}\leqq\sum_{k=h}^{H}\frac{1}{k(k-1)}=\frac{1}{h-1}-\frac{1}{H}=\frac{1}{h}-\frac{1}{H}+\frac{1}{h(h-1)},$$

és

$$\sum_{k=h}^{H} \frac{1}{k^2} \geqq \sum_{k=h}^{H} \frac{1}{k(k+1)} = \frac{1}{h} - \frac{1}{H+1} = \frac{1}{h} - \frac{1}{H} + \frac{1}{H(H+1)}$$

tehát

$$\sum_{k=h}^{H} \frac{1}{k^2} = \frac{1}{h} - \frac{1}{H} + \frac{\vartheta}{h(h-1)}, \text{ ahol } 0 < \vartheta < 1. \tag{2. 13}$$

Ennélfogva

$$M_n = \log \frac{n}{k_n} + \varepsilon_n' \tag{2. 14}$$

és

$$S_n^2 = \frac{1}{k_n} - \frac{1}{n} + \varepsilon_n'', \tag{2. 15}$$

ahol $|\varepsilon_n'| < \frac{2A}{k_n}$ és $|\varepsilon_n''| < \frac{B}{k_n^2}$, A és B az n-től nem függő állandók; ennélfogva

$$\frac{\zeta_{n+1-k_n}^* - M_n}{S_n} = \frac{\zeta_{n+1-k_n}^* - \ln \frac{n}{k_n}}{\sqrt{\frac{n-k_n}{nk_n}}} + \varepsilon_n''', \tag{2. 16}$$

ahol $|\varepsilon_n'''| \leqq W \sqrt{\frac{n}{k_n(n-k_n)}}$ és $W > 0$ n-től független állandó. De ha $n \to \infty$, úgy $\frac{k_n}{n} \to q$ és így $\varepsilon_n''' \to 0$ és ezért

$$\lim_{n\to\infty} P\left(\frac{\zeta_{n+1-k_n}^* - \log \frac{n}{k_n}}{\sqrt{\frac{n-k_n}{nk_n}}} < x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \tag{2. 17}$$

A rövidség kedvéért vezessük be a $q_n = \frac{k_n}{n}$ jelölést, akkor tehát (2. 17) szerint

$$\lim_{n\to\infty} P\left(\frac{\zeta_{n+1-k_n}^* - \log \frac{1}{q_n}}{\sqrt{\frac{1-q_n}{nq_n}}} < x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \tag{2. 18}$$

Mivel $\zeta_{n+1-k_n}^* = \log \frac{1}{F(\xi_{k_n}^*)}$, (2. 18)-ból következik, hogy

$$\lim_{n\to\infty} P\left(\xi_{k_n}^* > F^{-1}\left(q_n e^{-x\sqrt{\frac{1-q_n}{nq_n}}}\right)\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \tag{2. 19}$$

Viszont a differenciálszámítás középértéktétele értelmében

$$F^{-1}(y) - F^{-1}(x) = (y-x)\left(\frac{dF^{-1}(x)}{dx}\right)_{x=z}, \text{ ahol } x < z < y$$

és így

$$F^{-1}\left(q_n e^{x\sqrt{\frac{1-q_n}{nq_n}}}\right)=F^{-1}(q_n)+q_n\left(e^{x\sqrt{\frac{1-q_n}{nq_n}}}-1\right)\left(\frac{dF^{-1}(x)}{dx}\right)_{x=q_n\vartheta_n},$$

ahol $e^{x\sqrt{\frac{1-q_n}{nq_n}}}<\vartheta_n<1$; azonban $q_n\left(e^{-x\sqrt{\frac{1-q_n}{nq_n}}}-1\right)=-x\sqrt{\frac{(1-q_n)q_n}{n}}(1+\varrho_n)$, ahol $\lim\limits_{n\to\infty}\varrho_n=0$, tehát (2. 19)-ből következik, hogy

$$\lim_{n\to\infty}P\left(\frac{\xi^*_{k_n}-F^{-1}(q_n)}{\sqrt{\frac{q_n(1-q_n)}{n}}(1+\varrho_n)\left(\frac{dF^{-1}(x)}{dx}\right)_{x=q_n\vartheta_n}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt. \tag{2. 20}$$

Most szükségünk lesz a következő egyszerű lemmára:

3. lemma: Ha $\eta_n=a_n\alpha_n+b_n$, *ahol* $\lim\limits_{n\to\infty}a_n=a>0$ *és* $\lim\limits_{n\to\infty}b_n=b$, *továbbá* $F_n(x)$ *jelöli* $a\alpha_n+b$ *eloszlásfüggvényét és* $G_n(x)$ *az* η_n *eloszlásfüggvényét és létezik* $\lim\limits_{n\to\infty}G_n(x)=G(x)$, *akkor*

$$\lim_{n\to\infty}F_n(x)=G(x)$$

$G(x)$ *minden folytonossági helyén.*

Bizonyítás:

$$F_n(x)=P(a\alpha_n+b<x)=G_n\left(\frac{a_n}{a}x+\frac{b_na-ba_n}{a}\right)$$

és mivel $n\to\infty$ esetén $\lim\limits_{n\to\infty}\left(\frac{a_n}{a}x+\frac{b_na-ba_n}{a}\right)=x$, tehát

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}G_n(x)=G(x),$$

ami nem más, mint lemmánk állítása.

Alkalmazva ezt a lemmát, legyen $\alpha_n=\sqrt{n}(\xi^*_{k_n}-Q)$

$$a_n=\frac{1}{\sqrt{q_n(1-q_n)}(1+\varrho_n)\left(\frac{dF^{-1}(x)}{dx}\right)_{x=q_n\vartheta_n}} \quad \text{és } b_n=\sqrt{n}(Q-F^{-1}(q_n))a_n,$$

akkor, mivel $\left(\frac{dF^{-1}(x)}{dx}\right)_{x=q}=\frac{1}{f(F^{-1}(q))}=\frac{1}{f(Q)}$, tehát $\lim\limits_{n\to\infty}a_n=\frac{f(q)}{\sqrt{q(1-q)}}$ és mivel $|q_n-q|\leq\frac{D}{n}$, ahol D n-től független állandó, tehát

$$b_n=\sqrt{n}(Q-F^{-1}(q_n))a_n\to 0,$$

ennélfogva lemmánk értelmében

$$\lim_{n\to\infty}P\left(\frac{\xi^*_{k_n}-Q}{\frac{1}{f(Q)}\sqrt{\frac{q(1-q)}{n}}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt, \tag{2. 21}$$

amit bizonyitani akartunk.

Tételünk állítása úgy jellemezhető, hogy az n-elemű minta q-quantilise nagy n esetében közelítőleg normális eloszlású Q, vagyis a sokaság q-quantilise körül, $\frac{1}{f(Q)}\sqrt{\frac{q(1-q)}{n}}$ szórással.

Ilyen módon a minta mediánja segítségével megadhatók azok a határok, amelyek közé 1-hez tetszőlegesen közeli valószínűséggel az eloszlás mediánja esik. Speciálisan szimmetrikus eloszlás — pl. a normális eloszlás — esetében az eloszlás mediánja megegyezik a várható értékkel és így ezen az úton a várható értékre is következtethetünk.

A rendezett minták elmélete igen jól felhasználható a tömeggyártás gyártásközbeni statisztikai minőségellenőrzésénél,* ami a valószínűségszámítás egyik fontos alkalmazási területe. Tegyük fel, hogy egy automatagépen gyártott gépalkatrész valamely méretének (amely példányról példányra bizonyos apró véletlen ingadozásokat mutat, tehát valószínűségi változónak tekinthető) eloszlásfüggvénye, normális gyártási körülmények között, a (folytonos) $F(x)$ függvény; a gyártási folyamat ellenőrzése céljából bizonyos időközönként egy n-elemű — pl. 5-elemű — mintát veszünk.

A minta elemeinek vizsgált méretét lemérjük és a kapott eredményeket az ú. n. *kontrollkártyára* a mintavétel időpontjának megfelelő abszcisszán át húzott függőleges egyenesre felmérjük és egy ponttal jelöljük; a felmérés következtében a minta elemei maguktól nagyság szerint rendeződnek el. Annak megállapítása céljából, hogy a gyártási folyamatban nem lépett-e fel valamilyen rendellenesség (az automatagép beállításának eltolódása, a megmunkálást végző gép valamely részének kopása, stb.) a kontrollkártyára 5, párhuzamos egyenesek által alkotott sávot rajzolunk fel, amelyek rendre megadják azokat a határokat, amelyek közé az 5-elemű minta nagyság szerint legkisebb, második, harmadik, negyedik, ill. legnagyobb elemének egy megadott valószínűséggel — pl. $95^0/_0$-valószínűséggel — esniük kell, normális gyártási körülmények között. Ezen határok meghatározása az elmondottak alapján igen egyszerűen történhetik, ugyanis ha ξ_k^* jelenti a minta nagyság szerinti k-ik elemét ($k=1, 2, 3, 4, 5$), úgy a

$$\zeta_k^* = \log \frac{1}{F(\xi_{6-k}^*)}$$

változók eloszlását, ill. együttes eloszlását, mint láttuk, pontosan ki tudjuk számítani.

Ennek a módszernek a gyakorlati alkalmazásával a minőségellenőrzésnél a Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Mate-

* Lásd erre vonatkozólag *L. I. Braginszki* [26] munkáját; *Braginszki* számításai szabatossá tehetők a rendezett minták elméletének felhasználásával; a gyakorlati alkalmazásoknál célszerű a kontrollkártyákat a szabatos számítások alapján elkészíteni.

matikai Statisztikai Osztálya foglalkozik, a módszer felhasználásához szükséges táblázatok elkészítése is folyamatban van.

Nem folytatjuk itt a módszerrel nyerhető tételek bemutatását, csak összefoglalásul kiemeljük, hogy módszerünk lényege abban áll, hogy segítségével mindezeket a tételeket az (1.9) képlet segítségével a valószínűségszámítás legfejlettebb fejezetének, a független valószínűségi változók összegei határeloszláselméletének ismert tételeiből vezethetjük le.

3. §. Az elméleti és empirikus eloszlásfüggvény összehasonlítására vonatkozó új tételek megfogalmazása

Az eddigi elmondottakkal csak azt mutattuk meg, hogyan lehet módszerünkkel a rendstatisztika bizonyos jólismert klasszikus eredményeit egységes új eljárással egyszerűen bebizonyítani. Most rátérünk arra, hogy milyen új eredmények nyerhetők ugyanezzel a módszerrel.

A. N. Kolmogorov [6b] egy alapvető tételt bizonyított be, amely kritériumot szolgáltat arra vonatkozólag, hogy egy minta származhat-e egy megadott eloszlású statisztikai sokaságból, vagyis megmutatja, hogy hogyan lehet a minta elemeinek eloszlásából a teljes sokaság ismeretlen eloszlására következtetni. Legyen

$$\left.\begin{aligned} F_n(x) &= 0, \quad \text{ha} \quad x \leqq \xi_1^*, \\ F_n(x) &= \frac{k}{n}, \quad \text{ha} \quad \xi_k^* < x \leqq \xi_{k+1}^* \quad (k = 0, 1, 2, \ldots, n-1) \\ F_n(x) &= 1, \quad \text{ha} \quad x > \xi_n^*, \end{aligned}\right\} \tag{3.1}$$

és

vagyis jelölje $F_n(x)$ a minta empirikus eloszlásfüggvényét, másszóval a minta x-nél kisebb elemeinek relatív gyakoriságát. *Kolmogorov* tétele a következőképpen szól:

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{-\infty<x<+\infty} |F_n(x) - F(x)| < y\right) = \begin{cases} \sum\limits_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2y^2}, & \text{ha } y > 0 \\ 0 & \text{ha } y \leqq 0. \end{cases} \tag{3.2}$$

Kolmogorov tétele tehát az empirikus és elméleti eloszlásfüggvények abszolút eltérése felső határának határeloszlását adja meg, amely határeloszlás nem függ az $F(x)$ eloszlásfüggvénytől, amelyről a tétel érvényességéhez csak azt kell feltenni, hogy folytonos. *Kolmogorov* tétele az $|F_n(x) - F(x)|$ eltérést ugyanolyan súllyal veszi figyelembe, akármekkora is $F(x)$ értéke; ezáltal pl. az $|F_n(x) - F(x)| = 0{,}01$ eltérés ugyanolyan súllyal esik latba olyan x helyen, ahol $F(x) = 0{,}5$, vagyis ahol ez az eltérés 2%, mint egy olyan x pontban, ahol $F(x) = 0{,}01$, vagyis ahol az eltérés 100%-ot tesz ki. Ezen úgy segíthetünk, hogy $|F_n(x) - F(x)|$ helyett az $\frac{|F_n(x) - F(x)|}{F(x)}$ hányadost, vagyis $F_n(x)$

viszonylagos hibáját tekintjük. Ilyen módon természetesen merül fel a gondolat, hogy megvizsgáljuk az empirikus és elméleti eloszlásfüggvények közti viszonylagos eltérés abszolút értéke felső határának, vagyis $\frac{|F_n(x)-F(x)|}{F(x)}$ felső határának határeloszlását.

Kolmogorov tételéhez hasonló tételt bizonyított be N. V. *Szmirnov* az empirikus és elméleti eloszlásfüggvények egyoldalú eltérésére vonatkozólag. *Szmirnov* tétele a következőképpen szól:

$$\lim_{n\to\infty} P(\sqrt{n} \sup_{-\infty<x<+\infty} (F_n(x)-F(x))<y) = 1-e^{-2y^2}. \qquad (y \geqq 0) \qquad (3.3)$$

Ennél a tételnél is felmerült a viszonylagos eltérés vizsgálatának szükségessége.

Ezeket a problémákat sikerült a fent vázolt módszer segítségével megoldani. A probléma megoldása során egy természetes korlátozást kell bevezetnünk: mivel $F(x)$ tetszőleges kis értékeket is felvesz, nem célszerű $\frac{F_n(x)-F(x)}{F(x)}$-nek az egész $-\infty<x<+\infty$ intervallumra, hanem csak valamely $x_a \leqq x<+\infty$ intervallumra vonatkozó felső határát vizsgálni, ahol x_a az $F(x_a)=a>0$ feltétellel definiált abszcissza; a értéke azonban tetszőleges kicsiny pozitív szám lehet. Az 5. §-ban a következő eredményeket fogjuk bebizonyítani:

3. 1. tétel:

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{a\leqq F(x)\leqq 1}\left(\frac{F_n(x)-F(x)}{F(x)}\right)<y\right)=\begin{cases}\sqrt{\frac{2}{\pi}}\int\limits_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{t^2}{2}}dt, & \text{ha } y>0\\ 0 & \text{ha } y\leqq 0.\end{cases} \qquad (3.4)$$

3. 2. tétel:

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{a\leqq F(x)\leqq 1}\left|\frac{F_n(x)-F(x)}{F(x)}\right|<y\right)=\begin{cases}\frac{4}{\pi}\sum\limits_{k=0}^{\infty}(-1)^k\frac{e^{-\frac{(2k+1)^2\pi^2(1-a)}{8y^2a}}}{2k+1}, & \text{ha } y>0\\ 0, & \text{ha } y\leqq 0.\end{cases} \qquad (3.5)$$

Megvizsgálhatjuk a $\frac{F_n(x)-F(x)}{F(x)}$ változónak, ill. abszolút értékének azon $x_a \leqq x \leqq x_b$ intervallumbeli felső határának határeloszlását is, ahol $F(x_a)=a$, $F(x_b)=b$ és $0<a<b<1$; az erre vonatkozó tételek a következők:

3. 3. tétel:

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{a\leqq F(x)\leqq b}\left(\frac{F_n(x)-F(x)}{F(x)}\right)<y\right)=\frac{1}{\pi}\int\limits_{-\infty}^{y\sqrt{\frac{b}{1-b}}} e^{-\frac{u^2}{2}}\left[\int\limits_0^{\left(\sqrt{\frac{b}{1-b}}y-u\right)\sqrt{\frac{a(1-b)}{b-a}}} e^{-\frac{t^2}{2}}dt\right]du$$

$$(-\infty<y<+\infty) \qquad (3.6)$$

és

3. 4. tétel:

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{a \leq F(x) \leq b} \left| \frac{F_n(x)-F(x)}{F(x)} \right| < y\right) = \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{e^{-\frac{(2k+1)^2\pi^2(1-a)}{8y^2a}}}{2k+1} E_k, \tag{3.7}$$

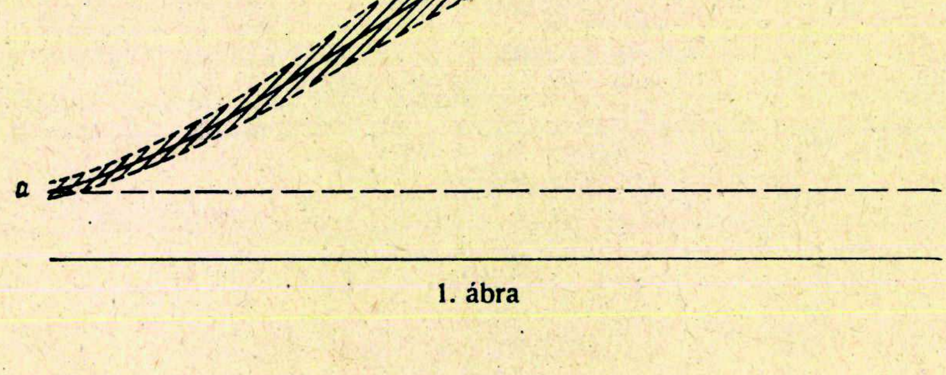
ha $y > 0$, ahol

$$E_k = 1 - \frac{2}{\sqrt{2\pi}} \int_{y\sqrt{\frac{b}{1-b}}}^{\infty} e^{-\frac{u^2}{2}} du + \varrho_k$$

és

$$\varrho_k = \frac{2e^{-\frac{by^2}{2(1-b)}}}{\sqrt{2\pi}\, y \sqrt{\frac{a}{1-a}}} \int_0^{(2k+1)\frac{\pi}{2}} e^{\frac{(1-b)u^2}{2by^2}} \sin u\, du. \tag{3.8}$$

Ezek a tételek kritériumokat szolgáltatnak annak a hipotézisnek az ellenőrzéséhez, hogy egy $\xi_1, \xi_2, \ldots, \xi_n$ minta egy $F(x)$ eloszlásfüggvényű sokaságból származik. Ezeknek a kritériumoknak a jellegzetessége abban áll, hogy egy olyan sávot adnak meg $F(x)$ körül, amelyben a hipotézis helyessége esetén az $F_n(x)$ empirikus eloszlásfüggvényeknek bizonyos megadott valószínűséggel haladnia kell, amely sávnak a szélessége minden x pontban $F(x)$-szel arányos (lásd az 1. ábrát). Ez a sáv azonban nem szimmetrikus. Ezen úgy segíthetünk, hogy a kritériumot kétszer alkalmazzuk, először a $(\xi_1, \xi_2, \ldots, \xi_n)$ mintára és az $F(x)$ eloszlásfüggvényre, azután a $(-\xi_1, -\xi_2, \ldots, -\xi_n)$ mintára és az $1-F(-x) = G(x)$ eloszlásfüggvényre.



1. ábra

Mutassunk rá a 3.3. tétel egy igen meglepő következményére: *Szmirnov* (3. 3) tételéből következik, hogy

$$\lim_{n\to\infty} P\Big(\sup_{-\infty<x<+\infty} (F_n(x)-F(x))<0\Big)=0, \tag{3. 9}$$

vagyis annak a valószínűsége, hogy az empirikus eloszlásfüggvény végig az elméleti eloszlásfüggvény alatt haladjon, 0-hoz tart, ha $n\to\infty$. A 3. 1. tételből következik, hogy

$$\lim_{n\to\infty} P\Big(\sup_{x_a\leqq x<+\infty} (F_n(x)-F(x))<0\Big)=0, \tag{3.10}$$

vagyis ugyanez áll fenn akkor is, ha csak a $(x_a, +\infty)$ intervallumot vizsgáljuk; ezzel szemben a 3.3. tétel értelmében

$$\lim_{n\to\infty} P\Big(\sup_{x_a\leqq x\leqq x_b} (F_n(x)-F(x))<0\Big)=\frac{1}{\pi}\int_0^\infty e^{-\frac{u^2}{2}}\left[\int_0^{u\sqrt{\frac{a(1-b)}{b-a}}} e^{-\frac{t^2}{2}}\,dt\right]du>0, \tag{3.11a}$$

vagyis határértékben is pozitív a valószínűsége annak, hogy azon a szakaszon, ahol $F(x)$ értéke valamely $a>0$ és $b<1$ szám közé esik, az empirikus eloszlásfüggvény végig az elméleti eloszlásfüggvény alatt haladjon, akármilyen kicsiny pozitív szám is a és akármilyen közel is van b az 1-hez. Ezen eredménynek nyilvánvalóan jelentősége van a statisztikai gyakorlat szempontjából is.

Azt, hogy a (3.11a) baloldalán álló határérték pozitív, *Gihman* [16] is bebizonyította és azt az eredményt kapta, hogy

$$\lim_{n\to\infty} P\Big(\sup_{x_a\leqq x\leqq x_b} (F_n(x)-F(x))<0\Big)=\frac{1}{\pi}\arc \sin\sqrt{\frac{a(1-b)}{b(1-a)}}. \tag{3.11b}$$

Gihman közli továbbá, hogy — személyes közlése szerint — a (3.11b) eredményt már *Gnyegyenko* ismerte. A (3.11a) és (3.11b) jobboldalán álló kifejezések természetesen azonosak; ez a következő meggondolással látható be: (3.11a) jobboldala nem más, mint annak kétszeres valószínűsége, hogy az (x, y) síkon normális eloszlású és $\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}$ sűrűségfüggvényű pont a $0<x<\infty$, $0<y<x\sqrt{\frac{a(1-b)}{b-a}}$ szögtérbe essék; ez a valószínűség nem más, mint

$$\frac{2\operatorname{arc\,tg}\sqrt{\frac{a(1-b)}{b-a}}}{2\pi}=\frac{1}{\pi}\arc\sin\sqrt{\frac{a(1-b)}{b(1-a)}}. \tag{3.12}$$

Ugyanis az $\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}$ sűrűségfüggvényű normális eloszlás körszimmetriája folytán a φ nyílású szögtérnek megfelelő valószínűség $\frac{\varphi}{2\pi}$.

A 3. 1.—3. 4. tételeket az 5. §-ban fogjuk bebizonyítani. Előbb azonban a 4. §-ban néhány segédtételt bizonyítunk be, amelyek önmagukban is érdekességgel bírnak.

4. §. Néhány új határeloszlástétel

Legyen adva egy valószínűségi változók szériáiból álló sorozat:

$$\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,N_n} \qquad (n=1,2,\ldots).$$

Tegyük fel, hogy az összes $\xi_{n,k}$ valószínűségi változók várható értéke zérus és szórása véges, továbbá, hogy az ugyanazon szériához tartozó valószínűségi változók összességükben függetlenek és kielégítik a Lindeberg-féle feltételt; másszóval bevezetve az $F_{n,k}(x)=P(\xi_{n,k}<x)$ jelölést és feltéve, hogy $M(\xi_{n,k})=$ $=\int\limits_{-\infty}^{+\infty} x\,dF_{n,k}(x)=0$, továbbá bevezetve az $S_{n,k}=\sum\limits_{\nu=1}^{k}\xi_{n,\nu}$ és

$$B_n^2=D^2(S_{n,N_n})=\sum_{k=1}^{N_n} D^2(\xi_{n,k}) \qquad (n=1,2,\ldots)$$

jelöléseket, teljesül a következő reláció:

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{N_n}\int\limits_{|x|>\varepsilon B_n} x^2\,dF_{n,k}(x)=0, \quad \text{ha } \varepsilon>0. \tag{4. 1}$$

A fenti feltételeknek eleget tevő széria-sorozatokra vonatkozólag a következő három tételt fogjuk bebizonyítani:

4. 1. tétel:

$$\lim_{n\to\infty}P(\max_{1\leqq k\leqq N_n}S_{n,k}<xB_n)=\begin{cases}\sqrt{\dfrac{2}{\pi}}\displaystyle\int_0^x e^{-\frac{t^2}{2}}dt, & \text{ha } x>0\\ 0, & \text{ha } x\leqq 0.\end{cases} \tag{4. 2}$$

4. 2. tétel:

$$\lim_{n\to\infty}P(\max_{1\leqq k\leqq N_n}|S_{n,k}|<xB_n)=\begin{cases}\dfrac{4}{\pi}\displaystyle\sum_{k=0}^{\infty}\frac{(-1)^k e^{-\frac{(2k+1)^2\pi^2}{8x^2}}}{2k+1}, & \text{ha } x>0\\ 0, & \text{ha } x\leqq 0.\end{cases} \tag{4. 3}$$

4. 3. tétel:

$$\lim_{n\to\infty}P(-yB_n\leqq\min_{1\leqq k\leqq N_n}S_{n,k}\leqq\max_{1\leqq k\leqq N_n}S_{n,n}<xB_n)=$$

$$=\begin{cases}\dfrac{4}{\pi}\displaystyle\sum_{k=0}^{\infty}\frac{e^{-\frac{(2k+1)^2\pi^2}{2(x+y)^2}}\sin(2k+1)\pi\frac{x}{x+y}}{2k+1}, & \text{ha } x>0 \text{ és } y>0;\\ 0, & \text{ha } x<0 \text{ vagy } y<0.\end{cases} \tag{4. 4}$$

Megjegyzés: Ha $y=x$, úgy a 4. 3. tétel a 4. 2. tételre redukálódik. A 4. 2. tétel tehát speciális esete a 4. 3. tételnek.

4. 4. tétel: Ha $A_n^2 = D^2(S_{n,M_n})$, *ahol* $1 \leq M_n < N_n$ *és* $\lim_{n\to\infty} \frac{A_n}{B_n} = \lambda$ $(0 \leq \lambda < 1)$, *úgy*

$$\lim_{n\to\infty} P(\max_{M_n < k \leq N_n} |S_{n,k}| < yB_n) =$$

$$= \begin{cases} \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k e^{-\frac{(2k+1)^2\pi^2}{8y^2}} \left[1 - \frac{2}{\sqrt{2\pi}} \int_{\frac{y}{\lambda}}^{\infty} e^{-\frac{u^2}{2}} du + \varrho_k\right], & \text{ha } y > 0; \\ 0, & \text{ha } y \leq 0 \end{cases} \tag{4. 5}$$

ahol

$$\varrho_k = \frac{2\lambda e^{-\frac{y^2}{2\lambda^2}}}{\sqrt{2\pi} y} \int_0^{(2k+1)\frac{\pi}{2}} e^{\frac{\lambda^2 u^2}{2y^2}} \sin u \, du.$$

Megjegyzés: Ha $M_n = 1$ (és így $\lambda = 0$), úgy a 4. 2. tétel a 4. 4. tételből speciális esetként adódik.

Megjegyezzük, hogy arra a speciális esetre, amelyben az összes szóbanforgó $\xi_{n,k}$ valószínűségi változók azonos eloszlásúak, a 4. 1. és 4. 2. tételek állítását *Erdős Pál* és *M. Kac* [22] bizonyították be.* A fenti általánosabb tételek bebizonyítása céljából az ő bizonyításukat annyiban módosítjuk, hogy az ott alkalmazott többdimenziós határeloszlástétel helyett közönséges (egydimenziós) határeloszlástételt alkalmazunk. Ez lehetővé teszi, hogy eredményeiket messzemenően általánosítjuk. A 4. 1.—4. 4. tételeket az 5 §-ban segédtételként használjuk a

$$\sup_{a \leq F(x) \leq b} \left(\frac{F_n(x) - F(x)}{F(x)}\right) \quad \text{és} \quad \sup_{a \leq F(x) \leq b} \left|\frac{F_n(x) - F(x)}{F(x)}\right|$$

változók határeloszlásának meghatározására, ahol $F_n(x)$ a folytonos $F(x)$ eloszlásfüggvényű ξ valószínűségi változóra vonatkozó n független megfigyelés eredményeinek empirikus eloszlásfüggvénye és $0 < a < b < 1$. Térjünk rá a 4. 1. tétel bizonyítására. Legyen

$$P_n(x) = P(\max_{1 \leq k \leq N_n} S_{n,k} < xB_n).$$

Legyenek az $\eta_1, \eta_2, \ldots, \eta_k, \ldots$ valószínűségi változók független, 0 várható értékű és 1 szórású normális eloszlású változók és vezessük be a

$$\zeta_k = \sum_{\nu=1}^{k} \eta_\nu \qquad k = 1, 2, \ldots$$

* Gondolatmenetüket *M. D. Donsker* fogalmazta meg általánosságban (l. [20b]). Lásd továbbá *A. Wald* dolgozatát, [27], amelyben egyforma eloszlású változók esetében, továbbá *A. Wald* [28] és *K. L. Chung* [29] munkáit, amelyekben korlátos harmadik momentumú változók esetében vizsgálják az első n részletösszegek maximumának határeloszlását; utóbbi erre a speciális esetre maradéktag-becslést is ad.

valószínűségi változókat. Először is azt bizonyítjuk be, hogy tetszőleges $\varepsilon > 0$ és pozitív egész k esetén

$$\varliminf_{n\to\infty} P_n(x) \geqq P(\max(\zeta_1, \zeta_2, \ldots, \zeta_k) < (x-\varepsilon)\sqrt{k} - \frac{1}{\varepsilon^2 k} \tag{4.6}$$

és

$$\varlimsup_{n\to\infty} P_n(x) \leqq P(\max(\zeta_1, \zeta_2, \ldots, \zeta_k) < x\sqrt{k}). \tag{4.7}$$

Legyen ugyanis m_j $(j=1, 2, \ldots, k)$ azon legkisebb pozitív egész szám, melyre

$$\sum_{r=1}^{m_j} D^2(\xi_{n,r}) \geqq \frac{j}{k} B_n^2.$$

Nyilvánvaló, hogy $1 \leqq m_1 \leqq m_2 \leqq \ldots \leqq m_k = N_n$. Definiáljuk most a következő valószínűségi változókat:

$$\Delta_{n,1} = S_{n,m_1}, \Delta_{n,j} = S_{n,m_j} - S_{n,m_{j-1}} \qquad (j = 2, 3, \ldots, k). \tag{4.8}$$

Könnyen beláthatjuk, hogy minden rögzített j esetén $(j=1, 2, \ldots, k)$ a

$$\xi_{n,m_{j-1}}, \xi_{n,m_{j-1}+2}, \ldots, \xi_{n,m_j} \qquad (n = 1, 2, \ldots) \tag{4.9}$$

széria-sorozatra szintén fennáll a Lindeberg-féle feltétel. Bevezetve ugyanis a

$$B_{n,j}^2 = D^2(\Delta_{n,j})$$

jelölést és felhasználva a (4. 1)-ből triviálisan következő

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sup_{1\leqq k\leqq N_n} D^2(\xi_{n,k}) = 0$$

összefüggést, nyerjük, hogy tetszőleges $\delta > 0$ esetén

$$\frac{1-\delta}{k} B_n^2 \leqq B_{n,j}^2 \leqq \frac{1+\delta}{k} B_n^2, \qquad \text{ha} \quad n > n_0(\delta). \tag{4.10}$$

Így tetszőleges $\varepsilon > 0$ mellett

$$\frac{1}{B_{n,j}^2} \sum_{r=m_{j-1}+1}^{m_j} \int_{|x|>\varepsilon B_{n,j}} x^2 dF_{n,r}(x) \leqq \frac{k}{1-\delta} \frac{1}{B_n^2} \sum_{r=1}^{N_n} \int_{|x|>\varepsilon' B_n} x^2 dF_{n,r}(x),$$

ha $n > n_0(\delta)$, ahol $\varepsilon' = \varepsilon \sqrt{\frac{1-\delta}{k}}$. A (4. 9) alatti sorozatra tehát valóban teljesül a Lindeberg-feltétel. A központi határeloszlástétel szerint tehát

$$\lim_{n\to\infty} P(\Delta_{n,j} < x B_{n,j}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \qquad (j = 1, 2, \ldots, k) \tag{4.11}$$

De a $\Delta_{n,1}, \Delta_{n,2}, \ldots, \Delta_{n,k}$ valószínűségi változók függetlenek és

$$\lim_{n\to\infty} \frac{B_{n,j}}{B_n} = \frac{1}{\sqrt{k}} \qquad (j = 1, 2, \ldots, k),$$

tehát

$$\lim_{n\to\infty} P\left(\frac{\Delta_{n,j}}{B_n} < \frac{x_j}{\sqrt{k}};\ j=1,2,\ldots,k\right) = \frac{1}{(\sqrt{2\pi})^k}\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_k} e^{-\frac{1}{2}\left(\sum_{j=k}^{k} t_j^2\right)}dt_1\ldots dt_k, \tag{4.12}$$

vagyis

$$\lim_{n\to\infty} P\left(\frac{\Delta_{n,1}+\Delta_{n,2}+\cdots+\Delta_{n,j}}{B_n} < x,\ j=1,2,\ldots,k\right) = \frac{1}{(\sqrt{2\pi})^k}\int\int_{(T_x)}\cdots\int e^{-\frac{1}{2}\left(\sum_{j=1}^{k} t_j^2\right)}dt_1\ldots dt_k, \tag{4.13}$$

ahol az integrációt az alábbi egyenlőtlenségekkel definiált T_x tartományra kell kiterjeszteni:

$$T_x: \{-\infty < t_1+t_2+\cdots+t_j < x\sqrt{k},\quad j=1,2,\ldots,k\}.$$

Így a következőt olvashatjuk le

$$\lim_{n\to\infty} P(\max_{1\leqq j\leqq k} S_{n,m_j} < xB_n) = P(\max_{1\leqq j\leqq k} \delta_j < x\sqrt{k}). \tag{4.14}$$

Legyen most

$$P_{n,k}(x) = P(\max_{1\leqq j\leqq k} S_{n,m_j} < xB_n)$$

és jelentse $\Pi_{n,r}(x)$ annak a valószínűségét, hogy $S_{n,r}$ az első az S_n, $(j=1,2,\ldots,r)$ összegek közül, amely $\geqq xB_n$, vagyis legyen

$$\Pi_{n,r}(x) = P(S_{n,r} \geqq xB_n;\ \max_{1\leqq j\leqq r-1} S_{n,j} < xB_n).$$

Nyilvánvalóan $\sum_{r=1}^{N_n} \Pi_{n,r}(x) = 1 - P_n(x) \leqq 1$.

Legyen $m_{j-1} < r \leqq m_j$; bevezetve a

$$\Pi_{n,r}^{(1)}(x) = P(S_{n,r} \geqq xB_n;\ \max_{1\leqq j\leqq r-1} S_{n,j} < xB_n;\ |S_{n,m_j} - S_{n,r}| \geqq \varepsilon B_n) \tag{4.15}$$

és a

$$\Pi_{n,r}^{(2)}(x) = P(S_{n,r} \geqq xB_n;\ \max_{1\leqq j\leqq r-1} S_{n,j} < xB_n;\ |S_{n,m_j} - S_{n,r}| < \varepsilon B_n) \tag{4.16}$$

jelöléseket, nyilvánvalóan

$$\Pi_{n,r}(x) = \Pi_{n,r}^{(1)}(x) + \Pi_{n,r}^{(2)}(x).$$

Alkalmazzuk a Csebisev-féle egyenlőtlenséget:

$$\Pi_{n,r}^{(1)}(x) = \Pi_{n,r}(x) P(|S_{n,m_j} - S_{n,r}| \geqq \varepsilon B_n) \leqq \Pi_{n,r}(x)\frac{B_{n,j}^2}{\pi^2 B_n^2}$$

és vegyük figyelembe a (4. 10) alatti relációt; így kapjuk, hogy

$$\Pi_{n,r}^{(1)}(x) \leqq \Pi_{n,r}(x)\frac{1+\delta}{\varepsilon^2 k},$$

tehát

$$1-P_n(x)=\sum_{r=1}^{N_n} \Pi_{n,r}(x) \leqq \frac{1+\delta}{\varepsilon^2 k}+\sum_{r=1}^{N_n} \Pi_{n,r}^{(2)}(x).$$

Viszont

$$\sum_{r=1}^{n} \Pi_{n,r}^{(2)}(x) \leqq 1-P_{n,k}(x-\varepsilon), \text{ tekintve, hogy ha}$$

$$S_{n,r} \geqq xB_n \text{ és } |S_{n,m_j}-S_{n,r}|<\varepsilon B_n, \text{ úgy } S_{n,m_j}>(x-\varepsilon)B_n;$$

így

$$1-P_n(x) \leqq \frac{1+\delta}{\varepsilon^2 k}+1-P_{n,k}(x-\varepsilon).$$

Ha még figyelembevesszük a triviális $P_n(x) \leqq P_{n,k}(x) \quad (k=1,2,\ldots)$ egyenlőtlenségeket is, nyerjük, hogy

$$P_{n,k}(x-\varepsilon)-\frac{1+\delta}{\varepsilon^2 k} \leqq P_n(x) \leqq P_{n,k}(x) \tag{4. 17}$$

A fenti (4. 17) alatti relációt (4. 14)-vel összevetve éppen a kívánt (4. 6) és (4. 7) alatti egyenlőtlenségeket kapjuk.

Tekintsük most azon speciális esetet, amikor $\xi_{n,k}$ csak a $+1$ és -1 értékeket veszi fel és

$$P(\xi_{n,k}=+1)=P(\xi_{n,k}=-1)=\frac{1}{2} \qquad (k=1,2,\ldots,n;\, n=1,2,\ldots)$$

Ekkor

$$P_n(x)=\frac{1}{2^n}\sum_{\substack{-n<\nu<[x\sqrt{n}]\\ \nu\equiv n \bmod 2}}\left\{\binom{n}{\frac{n+\nu}{2}}-\binom{n}{\frac{n-\nu}{2}+[x\sqrt{n}]+1}\right\}, \tag{4. 18}$$

ha $x\sqrt{n}$ nem egész szám, és így a Moivre—Laplace tételből következik, hogy

$$\lim_{n\to\infty} P_n(x)=\sqrt{\frac{2}{\pi}}\int_0^x e^{-\frac{t^2}{2}}dt. \tag{4. 19}$$

Ezért, mivel (4. 6) szerint

$$\overline{\lim_{k\to\infty}}\, P(\max(\zeta_1,\zeta_2,\ldots,\zeta_k)<x\sqrt{k}) \leqq \lim_{n\to\infty} P_n(x+\varepsilon) \tag{4. 20}$$

és (4. 7)-ből

$$\underline{\lim_{k\to\infty}}\, P(\max(\zeta_1,\zeta_2,\ldots,\zeta_k)<x\sqrt{k} \geqq \lim_{n\to\infty} P_n(x) \tag{4. 21}$$

következik, hogy

$$\lim_{k\to\infty} P(\max(\zeta_1,\zeta_2,\ldots,\zeta_k)<x\sqrt{k})=\sqrt{\frac{2}{\pi}}\int_0^x e^{-\frac{t^2}{2}}dt, \tag{4. 22}$$

és így a (4. 6) és (4. 7) relációkat még egyszer alkalmazva, (4. 2) következik.

A most közölt bizonyítás alapgondolata a következőkben foglalható össze: kimutattuk, hogy a $\xi_{n,k}$ változók speciális választása mellett fennáll (4. 2); ebből (4. 6) és (4. 7) segítségével következtettünk arra, hogy (4. 2) akkor is fennáll, ha $\xi_{n,k}=\eta_k$ és az η_k változók normális eloszlásúak; ebből újból (4. 6) és (4. 7) segítségével következett, hogy (4. 2) fennáll tetszőleges, a tétel feltételeinek eleget tevő $\xi_{n,k}$ változók esetében is.

A 4. 2. és 4. 3. tétel bizonyítása ugyanezen a gondolaton alapszik és a megfelelő módosításokkal lépésről-lépésre megegyezik a 4. 1. tétel bizonyításával.

Elegendő a 4. 3. tétel bizonyításával foglalkozni, mivel az, mint arra rámutattunk, a 4. 2. tételt speciális esetként tartalmazza. A bizonyítás első részét felesleges részletezni, csak a második részével foglalkozunk.

Legyen újból $P(\xi_{n,k}=1)=P(\xi_{n,k}=-1)=\frac{1}{2}$ $(k=1,2,\ldots,N_n;\ n=1,2,\ldots,)$ és tegyük fel, hogy a $\xi_{n,k}$ $(k=1,2,\ldots,N_n)$ változók függetlenek, úgy a síkbeli bolyongás elméletéből jólismert meggondolással következik, hogy ha $B=[y\sqrt{n}]+1$ és $A=[x\sqrt{n}]+1$, úgy

$$P(-y\sqrt{n}<S_{n,k}<x\sqrt{n};\ k=1,2,\ldots,n)= \tag{4. 23}$$
$$=\sum_{-B<k<A}\left[v_k+\sum_{\nu=1}^{\infty}\{(v_{2\nu(A+B)+k}-v_{2\nu(A+B)-2B-k})+(v_{-2\nu(A+B)+k}-v_{-2\nu(A+B)+2A-k})\}\right],$$

ahol

$$v_k=\begin{cases}\dfrac{1}{2^n}\dbinom{n}{\frac{n+k}{2}}, & \text{ha}\quad k\equiv n \mod 2,\\ 0 & \text{egyébként.}\end{cases} \tag{4. 24}$$

Mivel a Moivre—Laplace-tétel szerint

$$\sum_{\frac{n}{2}-b\frac{\sqrt{n}}{2}<k<\frac{n}{2}+a\frac{\sqrt{n}}{2}}\binom{n}{k}\frac{1}{2^n}\to\frac{1}{\sqrt{2\pi}}\int_{-b}^{a}e^{-\frac{u^2}{2}}du, \tag{4. 25}$$

ebből egyszerű számolással következik, hogy

$$\lim_{n\to\infty}P(-y\sqrt{n}<S_{n,k}<x\sqrt{n};\ k=1,2,\ldots,n)=$$
$$=\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{e^{-\frac{(2k+1)^2\pi^2}{2(x+y)^2}}\sin(2k+1)\dfrac{\pi x}{x+y}}{2k+1}. \tag{4. 26}$$

Ugyanúgy, mint a 3. 1. tétel esetében, következik, hogy ugyanez a határértéke a (4. 4) baloldalán álló valószínűségnek az általános esetben is.

A 3. 4. tétel a 3. 3. tételből a következőképpen vezethető le: a teljes valószínűség tétele szerint, tekintetbe véve, hogy S_{n,M_n} és $S_{n,k}-S_{n,M_n}$ függetlenek, ha $k>M_n$, továbbá hogy

$$P(\max_{M_n<k\leq N_n}|S_{n,k}|<yB_n)=P(-yB_n<(S_{n,M_n}+(S_{n,k}-S_{n,M_n}))<yB_n),$$

következik, hogy

$$P(\max_{M_n < k \leqq N_n} |S_{n,k}| < yB_n) =$$

$$= \int_{-\infty}^{+\infty} P\Big(-(x+y)B_n < S_{n,k} - S_{n,M_n} < (y-x)B_n\Big) dF_n(x), \tag{4.27}$$

ahol $F_n(x)$ az $\frac{S_{n,M_n}}{B_n}$ változó eloszlásfüggvényét jelenti. Mivel feltételeink értelmében az $S_{n,M_n} = \sum_{k=1}^{M_n} \xi_{n,k}$ összegekre érvényes a Lindeberg-féle feltétel és $D\left(\frac{S_{n,M_n}}{B_a}\right) = \frac{A_n}{B_n} \to \lambda$, tehát (minden véges intervallumon x-ben egyenletesen)

$$\lim_{n \to \infty} F_n(x) = \frac{1}{\sqrt{2\pi\lambda}} \int_{-\infty}^{x} e^{-\frac{u^2}{2\lambda^2}} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{\lambda}} e^{-\frac{u^2}{2}} du. \tag{4.28}$$

Mivel továbbá a 4. 3. tétel szerint — figyelembevéve, hogy

$$D(S_{n,N_n} - S_{n,M_n}) = \sqrt{B_n^2 - A_n^2},$$

következik, hogy

$$\lim_{n \to \infty} P(-(x+y)B_n < S_{n,k} - S_{n,M_n} < (y-x)B_n;\ k = 1, 2, \ldots, n) =$$

$$= \begin{cases} \dfrac{4}{\pi} \sum\limits_{k=0}^{\infty} \dfrac{e^{-\frac{(2k+1)^2 n^2 (1-\lambda^2)}{8y^2}}}{2k+1} \sin (k+1) \dfrac{\pi(y-x)}{2y}, & \text{ha } -y \leqq x \leqq y \\ 0, & \text{ha } |x| > y, \end{cases} \tag{4.29}$$

tehát

$$\lim_{n \to \infty} P(\max_{M_n < k \leqq N_n} |S_{n,k}| < yB_n) =$$

$$= \begin{cases} \dfrac{4}{\pi} \sum\limits_{k=0}^{\infty} \dfrac{e^{-\frac{(2k+1)^2 \pi^2 (1-\lambda^2)}{8y^2}}}{2k+1} \displaystyle\int_{-y}^{+y} \dfrac{e^{-\frac{x^2}{2\lambda^2}}}{\sqrt{2\pi\lambda}} \sin (2k+1)\pi \dfrac{y-x}{2y} dx, & \text{ha } y > 0 \\ 0, & \text{ha } y \leqq 0. \end{cases} \tag{4.30}$$

Ebből egyszerű számolással következik a 3. 4. tétel, ugyanis

$$\int_{-y}^{+y} \frac{e^{-\frac{x^2}{2\lambda^2}}}{\sqrt{2\pi\lambda}} \sin (2k+1)\pi \frac{y-x}{2y} dx = (-1)^k e^{\frac{(2k+1)^2 \pi^2 \lambda^2}{8y^2}} \cdot \left(\int_{-\frac{y}{\lambda}}^{+\frac{y}{\lambda}} \frac{e^{-\frac{1}{2}\left(t - \frac{(2k+1)i\lambda}{2y}\pi\right)^2}}{\sqrt{2\pi}} dt \right). \tag{4.31}$$

Mármost, mivel $e^{-\frac{t^2}{2}}$ integrálja egy zárt görbén eltűnik, tehát ha $a > 0$ és b

valós, úgy

$$\int_{-a}^{+a} \frac{e^{-\frac{(t-ib)^2}{2}}}{\sqrt{2\pi}} dt = \int_{-a-ib}^{a-ib} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt = \int_{-a}^{+a} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt +$$

$$+ \int_{-a-ib}^{-a} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt + \int_{a}^{a-ib} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt = \int_{-a}^{+a} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt + \frac{2e^{-\frac{a^2}{2}}}{\sqrt{2\pi}} \int_{a}^{b} e^{\frac{b^2}{2}} \sin av\, dv.$$

Ennélfogva

$$\int_{-y}^{+y} \frac{e^{-\frac{x^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda} \sin(2k+1)\pi \frac{y-x}{2y} dx =$$

$$= (-1)^k e^{-\frac{(2k+1)^2\pi^2\lambda^2}{8y^2}} \left[1 - \frac{2}{\sqrt{2\pi}} \int_{\frac{y}{\lambda}}^{\infty} e^{-\frac{t^2}{2}} dt + \frac{2e^{-\frac{y^2}{2\lambda^2}}\lambda}{\sqrt{2\pi y}} \int_{0}^{(2k+1)\frac{\pi}{2}} e^{\frac{\lambda^2 v^2}{2y^2}} \sin v\, dv\right].$$

Ezzel a 4. 4. tételt bebizonyítottuk.

5. §. Kolmogorov és Szmirnov kritériumaival analóg új kritériumok bizonyítása

Legyen ξ folytonos $F(x)$ eloszlásfüggvényű valószínűségi változó, jelentsék $\xi_1, \xi_2, \ldots, \xi_n$ a ξ-re vonatkozó n számú független megfigyelés eredményeit, azaz legyenek $\xi_1, \xi_2, \ldots, \xi_n$ összességükben független valószínűségi változók, ugyanazon $F(x)$ folytonos eloszlásfüggvénnyel. Jelöljük $F_n(x)$-szel ezen változók empírikus eloszlásfüggvényét.

Bebizonyítjuk a 3. §.-ban megfogalmazott tételeket, az 1. §-ban kifejtett módszerrel és a 4. §. tételeinek felhasználásával.

Legyen $\eta_k = F(\xi_k)$ és $\zeta_k = \log \frac{1}{\eta_k}$, továbbá $\eta_k^* = F(\xi_k^*)$ és $\zeta_k^* = R_k(\zeta_1, \zeta_2, \ldots, \zeta_n)$, úgy az η_k változók a (0, 1) intervallumban egyenletes eloszlásúak és empírikus eloszlásfüggvényük

$$G_n(x) = F_n(F^{-1}(x))$$

ahol $y = F^{-1}(x)$ az $x = F(y)$ függvény inverz függvénye. De könnyen látható, hogy

$$\sup_{a \leq F(x) \leq 1} \left(\frac{F_n(x) - F(x)}{F(x)}\right) = \sup_{a \leq x \leq 1} \left(\frac{G_n(x) - x}{x}\right), \qquad (5.\ 1)$$

tehát az (5. 1) baloldalán álló változó helyett vizsgálhatjuk a vele azonos

$$\sup_{a \leq x \leq 1} \left(\frac{G_n(x) - x}{x}\right)$$

változót. Az η_k^* változók — mint láttuk — Markov-láncot alkotnak. Láttuk továbbá, hogy a $\delta_{k+1}=(n-k)(\zeta_{k+1}^*-\zeta_k^*)$ változók összességükben függetlenek és 1 várható értékű exponenciális eloszlásúak, azaz

$$P(\delta_k < x) = 1 - e^{-x} \qquad (x \geqq 0).$$

Láttuk, hogy a δ_k változók segítségével a $\zeta_k^* = \ln \frac{1}{\eta_{n+1-k}^*}$ változókat független valószínűségi változók összegeként állíthatjuk elő a következőképpen:

$$\zeta_k^* = \sum_{j=1}^{k} \frac{\delta_j}{n+1-j}.$$

Térjünk rá most a 3. 1. tétel bizonyítására. Mindenekelőtt könnyen belátható, hogy a (3. 4) reláció helyett elegendő bebizonyítani, hogy

$$\lim_{n\to\infty} P\left(\sqrt{n} \sup_{a \leqq G_n(x)} \left(\frac{G_n(x)-x}{x}\right) < y\right) = \sqrt{\frac{2}{\pi}} \int_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{t^2}{2}} dt \qquad (y > 0). \quad (5.2)$$

Ugyanis ha $|G_n(x)-x| \leqq \varepsilon$, úgy abból, hogy $G_n(x) \geqq a+\delta$, következik, hogy $x \geqq G_n(x)-\varepsilon \geqq a$ és így ezen feltevés mellett

$$\sup_{a \leqq x} \frac{G_n(x)-x}{x} \geqq \sup_{G_n(x) \geqq a+\varepsilon} \frac{G_n(x)-x}{x},$$

vagyis abból, hogy $\sup_{a \leqq x} \frac{G_n(x)-x}{x} < \frac{y}{\sqrt{n}}$, következik, hogy $\sup_{G_n(x) \geqq a+\varepsilon} \frac{G_n(x)-x}{x} \leqq \leqq \frac{y}{\sqrt{n}}$. De ha A és B tetszőleges események és $AB \subset A'B$, úgy

$$P(A) = P(A\bar{B}) + P(AB) \leqq P(\bar{B}) + P(A'B) \leqq P(\bar{B}) + P(A')$$

Alkalmazva ezt az egyenlőtlenséget arra az esetre, amikor B a $|G_n(x)-x| \leqq \varepsilon$ eseményt, A a $\sup_{a \leqq x}\left(\frac{G_n(x)-x}{x}\right) < \frac{y}{\sqrt{n}}$ eseményt és A' a $\sup_{G_n(x) \geqq a+\varepsilon} \frac{G_n(x)-x}{x} < \frac{y}{\sqrt{n}}$ eseményt jelenti, következik, hogy

$$P\left(\sqrt{n} \sup_{a \leqq x} \frac{G_n(x)-x}{x} < y\right) \leqq P(|G_n(x)-x| > \varepsilon) + P\left(\sqrt{n} \sup_{a+\varepsilon \leqq G_n(x)} \frac{G_n(x)-x}{x} < y\right).$$

Hasonlóképpen látható be, hogy

$$P\left(\sqrt{n} \sup_{a-\varepsilon \leqq G_n(x)} \frac{G_n(x)-x}{x} < y\right) \leqq P(|G_n(x)-x| > \varepsilon) + P\left(\sqrt{n} \sup_{a \leqq x} \frac{G_n(x)-x}{x} < y\right).$$

Mivel $\lim_{n\to\infty} P(|G_n(x)-x| > \varepsilon) = 0$, hacsak $\varepsilon > 0$, következik, hogy ha (5. 2) teljesül, úgy

$$\overline{\lim_{n\to\infty}} P\left(\sqrt{n} \sup_{a \leqq x} \frac{G_n(x)-x}{x} < y\right) \leqq \sqrt{\frac{2}{\pi}} \int_0^{y\sqrt{\frac{a+\varepsilon}{1-a-\varepsilon}}} e^{-\frac{t^2}{2}} dt$$

és

$$\lim_{n\to\infty} P\left(\sqrt{n}\sup_{a\leq x}\frac{G_n(x)-x}{x}<y\right)\geqq\sqrt{\frac{2}{\pi}}\int\limits_0^{y\sqrt{\frac{a-\varepsilon}{1-a+\varepsilon}}} e^{-\frac{t^2}{2}}dt.$$

Mivel ε tetszőleges kicsinynek választható és az integrál a felső határának folytonos függvénye, következik, hogy

$$\lim_{n\to\infty} P\left(\sqrt{n}\sup_{a\leq x}\frac{G_n(x)-x}{x}<y\right)=\sqrt{\frac{2}{\pi}}\int\limits_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{t^2}{2}}dt.$$

Tehát (3. 4) valóban következik (5. 2)-ből és így a 3. 1. tétel bizonyításához elegendő kimutatni (5. 2) érvényességét.

Szükségünk lesz még a következő összefüggésre:

$$\sqrt{n}\sup_{a\leq G_n(x)}\left(\frac{G_n(x)-x}{x}\right)=\sqrt{n}\max_{na\leq k\leq n}\left(\frac{\frac{k}{n}}{\eta_k^*}-1\right). \tag{5. 3}$$

Ez abból következik, hogy $G_n(x)$ állandó η_k^* és η_{k+1}^* között és így egy $\eta_k^*<x<\eta_{k+1}^*$ szakaszon $\frac{G_n(x)-x}{x}=\frac{G_n(x)}{x}-1$ felső határa

$$\frac{G_n(\eta_k^*+0)}{\eta_k^*}-1=\frac{\frac{k}{n}}{\eta_k^*}-1.$$

A bizonyításhoz szükségünk van még a 4. 1. tételre.

Alkalmazzuk ezt a tételt a

$$\frac{\delta_j-1}{n+1-j}\qquad (j=1,2,\ldots,[n(1-a)]+1;\ n=1,2,\ldots)$$

változó-szériákból álló sorozatra, amely eleget tesz a Lindeberg-féle feltételnek (sőt, még a Ljapunov-féle feltételnek is), ha $0<a<1$. Ekkor (1. 11)-re való tekintettel tetszőleges $z>0$ esetén nyerjük, hogy

$$\lim_{n\to\infty} P\left(\max_{an\leq k\leq n}\left(\ln\frac{1}{\eta_k^*}-\sum_{\nu=k}^{n}\frac{1}{\nu}\right)<z\sqrt{\sum_{an\leq k\leq n}\frac{1}{k^2}}\right)=\sqrt{\frac{2}{\pi}}\int\limits_0^{z} e^{-\frac{t^2}{2}}dt. \tag{5. 4}$$

Mivel $k\geq an$ és $0<a<1$ esetén

$$\sum_{\nu=k}^{n}\frac{1}{\nu}=\ln\frac{n}{k}+O\left(\frac{1}{n}\right)\text{ és }\sqrt{\sum_{an\leq k\leq n}\frac{1}{k^2}}=\sqrt{\frac{1-a}{an}}+O\left(\frac{1}{n}\right),$$

az (5. 4) alatti relációból:

$$\lim_{n\to\infty} P\left(\max_{an\leq k\leq n}\ln\frac{\frac{k}{n}}{\eta_k^*}<z\sqrt{\frac{1-a}{an}}\right)=\sqrt{\frac{2}{\pi}}\int\limits_0^{z} e^{-\frac{t^2}{2}}dt, \tag{5. 5}$$

tehát végül, bevezetve a $z\sqrt{\frac{1-a}{a}}=y$ jelölést,

$$\lim_{n\to\infty} P\left(\sqrt{n}\max_{an\leq k\leq n}\left(\frac{\frac{k}{n}}{\eta_k^*}-1\right)<y\right)=\sqrt{\frac{2}{\pi}}\int\limits_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{t^2}{2}}\,dt, \tag{5.6}$$

ha $y>0$. (5.6)-ból (5.2) és (5.3) figyelembevételével a 3.1. tétel következik.

Térjünk át most a 3.3. tétel bizonyítására. A

$$\tau_1=\sqrt{n}\max_{na\leq k\leq nb}\left(\log\frac{1}{\eta_k^*}-\sum_{\nu=k}^{n}\frac{1}{\nu}\right)=\sqrt{n}\max_{na\leq k\leq nb}\sum_{j=1}^{n+1-k}\frac{\delta_j-1}{n+1-j} \tag{5.7}$$

valószínűségi változót két változó összegeként állíthatjuk elő, $\tau_1=\tau_2+\tau_3$ alakban, ahol

$$\tau_2=\sqrt{n}\sum_{1\leq j\leq n+1-nb}\frac{\delta_j-1}{n+1-j} \tag{5.8}$$

és

$$\tau_3=\sqrt{n}\max_{na\leq n+1-k\leq nb}\sum_{j=1}^{k}\frac{\delta_j-1}{n+1-j}. \tag{5.9}$$

Nyilvánvaló, hogy τ_2 határértékben $\sqrt{\frac{1-b}{b}}$ szórású normális eloszlású valószínűségi változó, a 3.1. tétel bizonyításából pedig láthatjuk, hogy

$$\lim_{n\to\infty} P(\tau_3\sqrt{b}<z)=\sqrt{\frac{2}{\pi}}\int\limits_0^{z\sqrt{\frac{a}{b+a}}} e^{-\frac{t^2}{2}}\,dt \qquad (z>0). \tag{5.10}$$

Figyelembevéve továbbá, hogy τ_2 és τ_3 függetlenek, (5.8)-ból és (5.10)-ből következik, hogy

$$\lim_{n\to\infty} P(\tau_1<y)=\frac{1}{\pi}\sqrt{\frac{b}{1-b}}\int\limits_{-\infty}^{y} e^{-\frac{bu^2}{2(1-b)}}\int\limits_0^{(y-u)\sqrt{\frac{ab}{b-a}}} e^{-\frac{t^2}{2}}\,dt\,du. \tag{5.11}$$

Ezzel a (3.3) tételt bebizonyítottuk.

Ugyanezen módszerrel bizonyíthatjuk be a 3.2. tételt is:

Az (5.3) alatti reláció helyét itt

$$\sqrt{n}\sup_{a\leq G_n(x)}\left|\frac{G_n(x)-x}{x}\right|=\sqrt{n}\max_{na\leq k\leq n}\left(\left|\frac{\frac{k}{n}}{\eta_k^*}-1\right|,\left|\frac{\frac{k}{n}}{\eta_{k+1}^*}-1\right|\right) \tag{5.12}$$

veszi át, melyből következik, hogy

$$\sqrt{n}\max_{na\leq k\leq n}\left|\frac{\frac{k}{n}}{\eta_k^*}-1\right|\leq\sqrt{n}\sup_{a\leq G_n(x)}\left|\frac{G_n(x)-x}{x}\right|\leq$$
$$\leq\sqrt{n}\max_{na\leq k\leq n}\left|\frac{\frac{k}{n}}{\eta_k^*}-1\right|+\frac{1}{a\sqrt{n}}. \tag{5.13}$$

Ugyanis, ha $\eta^*_{k+1} < \frac{k}{n}$, úgy $\left|\dfrac{\frac{k}{n}}{\eta^*_{k+1}}-1\right| = \dfrac{\frac{k}{n}}{\eta^*_{k+1}}-1 < \dfrac{\frac{k+1}{n}}{\eta^*_{k+1}}-1 = \left|\dfrac{\frac{k+1}{n}}{\eta^*_{k+1}}-1\right|$,

ha viszont $\eta^*_{k+1} \geqq \frac{k}{n}$, úgy $k \geqq na$ esetén $\eta^*_{k+1} \geqq a$ és így

$$\left|\frac{\frac{k}{n}}{\eta^*_{k+1}}-1\right| = 1-\frac{\frac{k}{n}}{\eta^*_{k-1}} \leqq 1-\frac{\frac{k+1}{n}}{\eta^*_{k-1}}+\frac{1}{n\eta^*_{k+1}} \leqq \left|\frac{\frac{k+1}{n}}{\eta^*_{k+1}}-1\right|+\frac{1}{na}.$$

Tehát minden körülmények között

$$\left|\frac{\frac{k}{n}}{\eta^*_{k+1}}-1\right| \leqq \left|\frac{\frac{k+1}{n}}{\eta^*_{k+1}}-1\right|+\frac{1}{na} \qquad (k \geqq na)$$

és így, mivel $\max\left|\dfrac{\frac{k+1}{n}}{\eta^*_{k-1}}-1\right| = \max\left|\dfrac{\frac{k}{n}}{\eta^*_{k}}-1\right|$, tehát

$$\max_{na\leqq k\leqq n}\left(\left|\frac{\frac{k}{n}}{\eta^*_{k}}-1\right|, \left|\frac{\frac{k}{n}}{\eta^*_{k+1}}-1\right|\right) \leqq \max_{na\leqq k\leqq n}\left|\frac{\frac{k}{n}}{\eta^*_{k}}-1\right|+\frac{1}{na}.$$

Az itt szereplő

$$\sqrt{n}\max_{na\leqq k\leqq n}\left|\frac{\frac{k}{n}}{\eta^*_{k}}-1\right|$$

változó határeloszlása megegyezik a

$$\sqrt{n}\max_{na\leqq k\leqq n}\left|\log\frac{1}{\eta^*_k}-\sum_{\nu=k}^{n}\frac{1}{\nu}\right| = \sqrt{n}\max_{na\leqq k\leqq n}\left|\sum_{j=1}^{n+1-k}\frac{\delta_j-1}{n+1-j}\right|$$

változó határeloszlásával, melyet a (4. 2) tétel segítségével határozhatunk meg.

A 3. 4. tétel bizonyításához egyidejűleg kell alkalmaznunk azokat az eljárásokat, amelyeket a 3. 2. és a 3. 3. tétel bizonyításánál alkalmaztunk; a 3. 4. tétel bizonyításához a 4. 2. tétel helyett a 4. 4. tételt használjuk fel.

A bizonyítás alapgondolata a következő:

$$\sqrt{n}\sup_{a\leqq F(x)\leqq b}\left|\frac{F_n(x)-F(x)}{F(x)}\right|$$

határeloszlása megegyezik a $\varkappa_1 = \sqrt{n}\max\limits_{na\leqq k\leqq nb}\left(\left|\dfrac{\frac{k}{n}}{\eta^*_k}-1\right|\right)$ határeloszlásával, ennélfogva megegyezik

$$\varkappa_2 = \sqrt{n}\max_{na\leqq k\leqq nb}\left|\ln\frac{\frac{k}{n}}{\eta^*_k}\right| = \sqrt{n}\max_{na\leqq k\leqq nb}\left|\sum_{j=1}^{n+1-k}\frac{\delta_j-1}{n+1-j}\right|$$

határeloszlásával és így a 4. 4. tétel akalmazható, mégpedig, mivel az abban szereplő A_n ill. B_n konstansok értékei itt: $B_n = \sqrt{\frac{1-a}{an}+O\left(\frac{1}{n}\right)}$ és $A_n =$

$$=\sqrt{\frac{1-b}{bn}}+O\left(\frac{1}{n}\right), \text{ tehát}$$

$$\lambda=\lim_{n\to\infty}\frac{A_n}{B_n}=\sqrt{\frac{a(1-b)}{b(1-a)}},$$

és így bevezetve az $S_{n,r}=\sum_{j=1}^{r}\frac{\delta_j-1}{n+1-j}$ jelölést,

$$\lim_{n\to\infty}P\left(\sqrt{n}\sup_{a\leq F(x)\leq b}\left|\frac{F_n(x)-F(x)}{F(x)}\right|<y\right)=$$

$$=\lim_{n\to\infty}P\left(\max_{n+1-nb\leq r\leq n+1-na}|S_{n,r}|\leq y\sqrt{\frac{a}{1-a}}\,B_n\right)=$$

$$=\frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{e^{-\frac{(2k+1)^2\pi^2(1-a)}{8y^2a}}}{2k+1}\left[1-\frac{2}{\sqrt{2\pi}}\int_{y\sqrt{\frac{b}{1+b}}}^{\infty}e^{-\frac{u^2}{2}}du+\varrho_k\right],$$

ahol

$$\varrho_k=\frac{2\sqrt{\frac{1-b}{b}}\,e^{\frac{y^2b}{2(1-b)}}}{\sqrt{2\pi}y}\int_0^{(2k+1)\frac{\pi}{2}}e^{-\frac{u^2(1-b)}{2by^2}}\sin u\,du.$$

Ezzel a 3. 4. tételt be is bizonyítottuk.

6. §. Megjegyzések a 3. 1—3. 4 tételekben szereplő határeloszlásfüggvényekről

A 3. 1. tételben szereplő határeloszlásfüggvény értékeit a normális eloszlás táblázatából olvashatjuk le. A 3. 2. tételben szereplő határeloszlásfüggvény értékeit úgy határozhatjuk meg, hogy az

$$L(z)=\frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{e^{-\frac{(2k+1)^2\pi^2}{8z^2}}}{2k+1}\qquad(z>0)\qquad(6.1)$$

függvénybe a $z=y\sqrt{\frac{a}{1-a}}$ értéket helyettesítjük be. A fejezet végén táblázatot adunk az $L\left(y\sqrt{\frac{a}{1-a}}\right)$ függvényre, a különböző értékei mellett. Az $L\left(y\sqrt{\frac{a}{1-a}}\right)$ eloszlásfüggvény menete a különböző értékeire a 2. ábrán látható.

A 3. 3. tételben szereplő határeloszlásfüggvényt közelítőleg a következőképpen számíthatjuk ki:

$$F(y,a,b)=\frac{1}{\pi}\int_{\infty}^{y\sqrt{\frac{b}{1-b}}}e^{-\frac{u^2}{2}}\int_0^{\left(y-u\sqrt{\frac{1-b}{b}}\right)\sqrt{\frac{ab}{b-a}}}e^{-\frac{v^2}{2}}dv\,du=\frac{1}{\pi}\iint_T e^{-\frac{u^2+v^2}{2}}du\,dv\qquad(6.2)$$

ahol T egy végtelen háromszögalakú tartomány az (u, v) síkon, amelyet a következő egyenlőtlenségek definiálnak:

$$-\infty < u < y\sqrt{\frac{b}{1-b}}, \quad 0 \leqq v \leqq \sqrt{\frac{a(1-b)}{b-a}}\left(y\sqrt{\frac{b}{1-b}} - u\right) \tag{6. 3}$$

Az $r = \sqrt{u^2+v^2}$, $\varphi = \operatorname{arc\,tg}\frac{v}{u}$ polárkoordinátákat bevezetve, kapjuk, hogy

$$F(y, a, b) = \frac{1}{\pi}\int_0^{\pi-\alpha}\left(1-e^{-\frac{ay^2}{2(1-a)\sin^2(\varphi+\alpha)}}\right)d\varphi, \tag{6. 4}$$

ahol $\operatorname{tg}\alpha = \sqrt{\frac{a(1-b)}{b-a}}$ és $0 < \alpha < \frac{\pi}{2}$, tehát

$$F(y, a, b) = \frac{1}{\pi}\int_\alpha^{\pi}\left(1-e^{-\frac{ay^2}{2(1-a)\sin^2\beta}}\right)d\beta. \tag{6. 5}$$

Mivel

$$\frac{1}{\pi}\int_0^{\pi}\left(1-e^{-\frac{ay^2}{2(1-a)\sin^2\beta}}\right)d\beta = \sqrt{\frac{2}{\pi}}\int_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{u^2}{2}}du, \tag{6. 6}$$

tehát végül nyerjük a következő közelítő kifejezést:

$$F(y, a, b) = \sqrt{\frac{2}{\pi}}\int_0^{y\sqrt{\frac{a}{1-a}}} e^{-\frac{u^2}{2}}du - \frac{\operatorname{arc\,tg}\sqrt{\frac{a(1-b)}{b-a}}}{\pi}(1-R), \tag{6. 7}$$

ahol

$$R = \frac{1}{\alpha}\int_0^{\alpha} e^{-\frac{ay^2}{2(1-a)\sin^2\beta}}d\beta \text{ és } \alpha = \operatorname{arc\,tg}\sqrt{\frac{a(1-b)}{b-a}}. \tag{6. 8}$$

Ha $1-b=\varepsilon$ kicsiny, úgy R általában elhanyagolható, kivéve y igen kis értékeit, tekintve, hogy

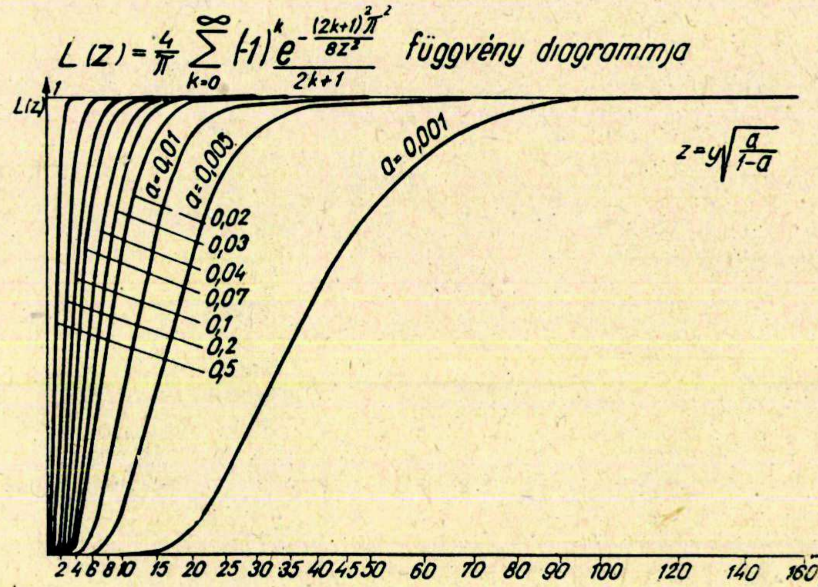
$$R \leqq e^{-\frac{ay^2}{2(1-a)\sin^2\alpha}} = e^{-\frac{by^2}{2(1-b)}}. \tag{6. 9}$$

A 3. 4. tételben szereplő határeloszlásfüggvény a következőképpen számítható ki közelítőleg: az integrálszámítás második középértéktételéből következik, hogy

$$\begin{aligned} |\varrho_k| &= \frac{2e^{-\frac{by^2}{(2(1-b)}}\sqrt{\frac{1-b}{b}}}{\sqrt{2\pi}\,y}\left|\int_0^{(2k+1)\frac{\pi}{2}} e^{-\frac{(1-b)u^2}{2by^2}}\sin u\,du\right| \leqq \\ &\leqq \frac{2e^{-\frac{by^2}{2(1-b)}}\sqrt{\frac{1-b}{b}}\,e^{-\frac{(2k+1)^2\pi^2(1-b)}{8y^2b}}}{\sqrt{2\pi}\,y}. \end{aligned} \tag{6. 10}$$

Ilyen módon a 3. 4. tételben szereplő határeloszlásfüggvény a következő alakra hozható, a (6. 1) jelölés felhasználásával:

$$\frac{4}{\pi} L\left(y\sqrt{\frac{a}{1-a}}\right)\left(1-\frac{2}{\sqrt{2\pi}}\int\limits_{y\sqrt{\frac{b}{1-b}}}^{\infty} e^{-\frac{u^2}{2}}du\right)+\Delta, \tag{6. 11}$$



2. ábra

ahol, amint egyszerű számolással belátható

$$\Delta < \frac{2\lambda e^{-\frac{by^2}{2(1-b)}}\sqrt{\frac{1-b}{b}}}{\pi\sqrt{2\pi}\,y}\log\frac{1+e^{-\frac{\pi^2(b-a)}{8aby^2}}}{1-e^{-\frac{\pi^2(b-a)}{8aby^2}}}, \tag{6. 12}$$

amelyből láthatjuk, hogy ha b igen közel van 1-hez, úgy Δ elhanyagolható. Figyeljük meg, hogy a főtag első tényezője csak a-tól, második tényezője csak b-től függ: ez igen leegyszerűsíti a kiszámítást; ugyanis az első tényezőt az $L(z)$ függvény táblázatából, a másodikat a normális eloszlás táblázatából meghatározhatjuk.

Magyar Tudományos Akadémia
Alkalmazott Matematikai Intézete.

IRODALOM

[1] *K. Pearson,* Note on Francis Galton's problem. Biometrika *1* (1902) 390—399.

[2] *L. v. Bortkiewicz,* Variationsbreite und mittlere Fehler. Sitzungsber. Berl. Math. Ges. *21* (1922) 3—11.

[3] *E. L. Dodd,* The greatest and the least variate under general laws of error. Trans. Amer. Math. Soc. *25* (1923) 525—539.

[4] *L. H. C. Tippet,* On the extreme individuals and the range of samples taken from a normal population. Biometrika *17* (1925) 264—387.

[5] *M. Fréchet,* Sur la loi de probabilité de l'écart maximum. Ann. Soc. Polon. Math. *6* (1928) 92—116.

[6] *a)* А. Н. Колмогоров: Известия Акад. Наук СССР Сер. физ-мат.. (1933) 363—372.

b) A. N. Kolmogoroff, Giornale dell'Istitutio Italiano degli Attuari *4* (1933) 1—11 és 83—91.

[7] *V. I. Glivenko,* Sulla determinazione empirica delle leggi di probabilita. Giorn. Ist. Ital. Attuari *4* (1933) 92—99.

[8] *a) N. V. Smirnoff,* Über die Verteilung des allgemeinen Gliedes in der Variationsreihe, Metron *12* (1935) 59—81.

b) Н. В. Смирнов: Comptes Rendus Ac. Sci. Paris. *202* (1936) 449.

c) Н. В. Смирнов: Матем. Сборник *2* (44) (1937) 973 993.

d) Н. В. Смирнов: Матем. Сборник *6* (44) (1939) 3—26 és Вюллемен Мат. *2* (1939).

e) Н. В. Смирнов: Бюллетен Мат. *2* (1939) és Annals of Math. Statistics *19* (1948) 279—281.

[9] *B. V. Gnedenko,* Sur la distribution limite du terme maximum d'une série aléatoire, Ann. of Math. *44* (1943) 423—453.

[10] Б. В. Гнеденко и В. С. Королюк: О максимальном расхождении двух эмпирических распределений, Д. А. Н, *80* (1951), 525—528.

[11] Б. В. Гнеденко и Е. Л. Рвачева: Об одной задаче сравнения двух эмпирических распределений, Д. А. Н *82* (1952) 513—516.

[12] Б. В. Гнеденко и В. С. Михалевич: Две теорема о поведений эмпирических функций распределения, Д. А. Н. *85* (1952) 841—845 és 25—27.

[13] В. С. Михалевич: О взаимнои расположении двух эмпирических функций распределения, Д. А. Н. *85* (1952) 485—488.

[14] И. Д. Квит, Д. А. Н. *71* (1950) 229—231.

[15] Г. М. Мания, Обобщение критерия А. Н. Колмогорова для оценки закона распределения по эмпирическим данным, Д. А. Н. (1949(495—497.

[16] И. Н. Гихман: Об емпирической функций распределения в случае группировки данных. Д. А. Н. *82* (1952) 837—840.

[17] *W. Feller,* On the Kolmogorov—Smirnov limit theorems for empirical distributions. Ann. Math. Statistics *19* (1948) 177—189.

[18] *J. L. Doob,* Heuristic approach to the Kolmogorov—Smirnov theorems. Ann. Math. Statistics *20* (1949) 393—403.

[19] *a) F. J. Massey,* A note on the estimation of a distribution function by confidence limits. Ann. Math. Statistics *21* (1950) 116—119.

b) A note on the power of a non-parametric test. Ann. Math. Statistics *21* (1950), 440—443.

c) Distribution table for the deviation between two sample cumulatives. Ann. Math Statistics *23* (1952) 435—441.

[20] *a) M. D. Donsker,* Justification and extension of Doob's heuristic approach to the Kolmogorov—Smirnov theorems. Ann. Math. Statistics *23* (1952) 277—281.

b) An invariance principle for certain probability limit theorems, Memoirs Amer. Math. Soc. *6* (1951) 1—12.

[21] *T. W. Anderson—D. A. Darling,* Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes Ann. Math. Statistics *23* (1952), 193—212.

[22] *P. Erdős—M. Kac,* Bull. Amer. Math. Soc. *52* (1946) 292.

[23] *S. Malmquist,* Skand. Actuarietidskrift. *33* (1950) 214.

[24] *Hajós György—Rényi Alfréd,* A rendezett minták elméletének néhány alapvető kérdéséről, sajtó alatt a III. Oszt. Közleményeiben.

[25] *H. Cramér,* Mathematical methods of statistics. Princeton, 1946.

[26] Л. И. Брагински: Оперативый статистический контроль качества в машиностроении, Москва, Машгиз, 1951.

[27] *A. Wald,* Limit distribution of the maximum and minimum of successive cumulative sums of random variables. Bull. Amer. Math. Soc. *53* (1947) 142—153.

[28] *A. Wald,* On the distribution of the maximum of successive cumulative sums of independently but not identically distributed chance variables, Bull. Amer. Math. Soc. *54* (1948) 422—430.

[29] *K. L. Chung,* Asymptotic distribution of the maximum cumulative sum of independent random variables, Bull. Amer. Math. Soc. *54* (1948) 1162—1170.

[30] *S. S. Wilks,* Order Statistics. Bull. Amer. Math. Soc. *54* (1948) 6—50.

y \ a	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,1	0,2	0,3	0,4	0,5
0,1													0,0000	0,0000
0,5											0,0000	0,0000	0,0008	0,0092
1,0						0,0000	0,0000	0,0000	0,0000	0,0000	0,0092	0,0716	0,2001	0,3708
1,5			0,0000	0,0000	0,0000	0,0002	0,0009	0,0023	0,0050	0,0092	0,1420	0,3543	0,5591	0,7328
2,0		0,0000	0,0001	0,0008	0,0036	0,0101	0,0212	0,0367	0,0563	0,0791	0,3708	0,6193	0,7951	0,9082
2,5		0,0001	0,0022	0,0112	0,0299	0,0578	0,0925	0,1320	0,1730	0,2155	0,5778	0,7966	0,9714	0,9751
3,0	0,0000	0,0015	0,0157	0,0474	0,0941	0,1487	0,2061	0,2632	0,3184	0,3708	0,7328	0,9009	0,9915	0,9954
3,5	0,0001	0,0092	0,0491	0,1135	0,1879	0,2629	0,3341	0,3994	0,4598	0,5140	0,8398	0,9561	0,9978	0,9991
4,0	0,0006	0,0291	0,1052	0,2001	0,2942	0,3804	0,4570	0,5244	0,5835	0,6353	0,9082	0,9823	0,9995	0,9999
4,5	0,0031	0,0643	0,1776	0,2950	0,4001	0,4902	0,5665	0,6311	0,6860	0,7328	0,9511	0,9936	0,9999	1,0000
5,0	0,0096	0,1135	0,2582	0,3895	0,4985	0,5873	0,6594	0,7193	0,7683	0,8088	0,9751	0,9979	1,0000	
5,5	0,0225	0,1726	0,3511	0,4784	0,5863	0,6723	0,7374	0,7903	0,8326	0,8665	0,9887	0,9994		
6,0	0,0428	0,2375	0,4204	0,5591	0,6627	0,7409	0,8006	0,8463	0,8817	0,9081	0,9954	0,9999		
6,5	0,0707	0,3045	0,4952	0,6310	0,7282	0,7989	0,8509	0,8895	0,9181	0,9395	0,9977	1,0000		
7,0	0.1053	0,3708	0,5639	0,6939	0,7834	0,8461	0,8904	0,9220	0,9446	0,9607	0,9991			
7,5	0,1452	0,4347	0,6193	0,7484	0,8294	0,8839	0,9207	0,9460	0,9633	0,9752	0,9996			
8,0	0,1889	0,4950	0,6811	0,7951	0,8671	0,9135	0,9436	0,9634	0,9763	0,9847	0,9999			
8,5	0,2348	0,5513	0,7301	0,8345	0,8977	0,9365	0,9606	0,9756	0,9850	0,9908	1,0000			
9,0	0,2819	0,6032	0.7731	0,8696	0,9221	0,9540	0,9729	0,9849	0,9907	0,9946				
9,5	0,3290	0,6510	0,8104	0,8950	0,9410	0,9713	0,9817	0,9898	0,9944	0,9969				
10,0	0,3754	0,6938	0,8427	0,9175	0,9564	0,9770	0,9878	0,9936	0,9967	0,9983				
10,5	0,4205	0,7328	0,8704	0,9358	0,9680	0,9840	0,9921	0,9961	0,9981	0,9991				
11,0	0,4640	0,7678	0,8939	0,9505	0,9768	0,9891	0,9949	0,9976	0,9989	0,9995				
11,5	0,5055	0,7992	0,9137	0,9622	0,9833	0,9927	0,9968	0,9986	0,9994	0,9997				
12,0	0,5450	0,8271	0,9303	0,9713	0,9882	0,9951	0,9980	0,9992	0,9997	0,9999				
12,5	0,5824	0.8517	0,9441	0,9784	0,9917	0,9968	0,9988	0,9995	0,9998	0,9999				
13,0	0,6174	0,8734	0,9555	0,9841	0,9943	0,9980	0,9993	0,9997	0,9999	1,0000				
13,5	0,6509	0,8924	0,9648	0,9883	0,9961	0,9987	0,9996	0,9999	1,0000					
14,0	0,6812	0,9090	0,9724	0,9915	0,9973	0,9992	0,9998	0,9999						
14,5	0,7099	0,9234	0,9780	0,9938	0,9982	0,9995	0,9999	1,0000						
15,0	0,7367	0,9358	0,9833	0,9956	0,9988	0,9997	0,9999							
15,5	0,7615	0,9464	0,9872	0,9969	0,9992	0,9998	1,0000							
16,5	0,7844	0,9555	0,9902	0,9978	0,9995	0,9999								
16,5	0,8055	0,9631	0,9927	0,9985	0,9997	0,9999								
17,0	0,8249	0,9697	0,9944	0,9990	0,9998	1,0000								

17,5	0,8428	0,9752	0,9958	0,9993	0,9999
18,0	0,8591	0,9797	0,9969	0,9995	0,9999
18,5	0,8740	0,9836	0,9977	0,9997	1,0000
19,0	0,8876	0,9867	0.9983	0,9998	
19,5	0,9000	0,9893	0,9988	0,9999	
20,0	0,9112	0,9915	0,9991	0,9999	
20,5	0,9213	0,9932	0,9994	0,9999	
21,0	0,9304	0,9946	0,9996	1,0000	
21,5	0,9386	0,9957	0,9997		
22,0	0,9460	0,9967	0,9998		
22,5	0,9526	0,9974	0,9998		
23,0	0,9590	0,9980	0,9999		
23,5	0,9636	0,9984	0,9999		
24,0	0,9696	0,9983	1,0000		
24,5	0,9724	0,9991			
25,0	0,9760	0,9993			
26	0,9821	0,9996			
27	0,9867	0,9998			
28	0,9902	0,9999			
29	0,9929	0,9999			
30	0,9949	1,0000			
35	0,9991				
40	0,9999				
43	1,0000				

Az $L\left(y\sqrt{\frac{a}{1-a}}\right)$ függvény értékeinek táblázata.