EGY MEGFIGYELÉSSOROZAT KIEMELKEDŐ ELEMEIRŐL

Írta: RÉNYI ALFRÉD

Bevezetés

Végezzünk független megfigyeléseket egy véletlentől függő mennyiségre vonatkozólag. Legyen ξ_k $(k=1, 2, \ldots)$ a k-adik megfigyelés. Más szóval tegyük fel, hogy a $\xi_1, \xi_2, \ldots$ valószínűségi változók függetlenek és egyforma eloszlásúak. Közös eloszlásfüggvényüket jelölje $F(x)$. Tegyük fel, hogy $F(x)$ x minden értékére folytonos (vagyis, hogy a ξ_k változók minden x valós értéket 0 valószínűséggel vesznek fel). Más szóval: vegyünk egy végtelen mintát egy folytonos eloszlású sokaságból és legyen ξ_k a minta k-adik eleme. Nevezzük a ξ_k megfigyelést a megfigyeléssorozat *kiemelkedő elemének*, ha $\xi_k > \xi_j$ midőn $j < k$. Más szóval kiemelkedőnek nevezünk egy megfigyelést, ha az az összes megelőző megfigyeléseknél nagyobb. E definíció szerint ξ_1 mindig (triviálisan) kiemelkedő. Legyen $\nu_0 = 1$; jelölje ν_1 az első 1-nél nagyobb sorszámú (tehát az első nem triviális) kiemelkedő megfigyelés sorszámát, ν_2 az első ν_1-nél nagyobb (tehát a második nemtriviális) kiemelkedő elem sorszáma és általában jelölje ν_n az n-edik nemtriviális kiemelkedő elem sorszámát $(n=1, 2, 3, \ldots)$. E dolgozat tárgyát a ν_n valószínűségi változókra vonatkozó határérték- és határeloszlástételek vizsgálata képezi.

Mielőtt a ν_n $(n=1, 2, 3, \ldots)$ valószínűségi változó-sorozatra vonatkozó tételeket kimondanók, néhány általános jellegű megállapítást kívánunk tenni. Először is nyilvánvaló, hogy a ν_n változók viselkedése nem függhet az $F(x)$ eloszlásfüggvénytől. Ugyanis, ha ξ_{ν_n} a ξ_k sorozatnak n-edik kiemelkedő eleme, akkor $F(\xi_{\nu_n})$ az $F(\xi_k)$ sorozatnak ugyancsak n-edik kiemelkedő eleme. Mivel azonban az $F(\xi_k)$ $(k=1, 2, \ldots)$ változók egyenletes eloszlásúak a $(0, 1)$ intervallumban, tehát a ξ_k változókról az általánosság megszorítása nélkül eleve feltehetjük, hogy azok egyenletes eloszlásúak a $(0, 1)$ intervallumban. Az ezen feltevés mellett a ν_n változókra bebizonyított határeloszlástételek e feltevés nélkül is érvényesek lesznek az $F(x)$ folytonos eloszlásfüggvény bármely választása mellett. A ν_n változó sorozat tulajdonságai tehát „eloszlásmentesek". Ebből következik, hogy minden olyan statisztikai próba, amely a ν_n változók megfigyelésén alapszik, szintén „eloszlásmentes" lesz. Az egész problémakör rokon a rendezett minták elméletével; a rendezett minták elméletének szokásos problematikájától azonban elsősorban abban térünk el, hogy egy *végtelen* megfigyeléssorozatból indulunk ki.

Az 1. §-ban a ν_n változók közötti függőséget vizsgáljuk, és kimutatjuk, hogy a ν_n változók homogén Markov-láncot alkotnak. A 2. §-ban meghatározzuk annak valószínűségét, hogy a ξ_k sorozat első N eleme között megadott számú kiemelkedő elem legyen. A 3. §-ban bebizonyítjuk azt a meglepő tényt, hogy ha A_k jelöli azt az eseményt, hogy a ξ_k elem kiemelkedő, akkor az A_k $(k=1, 2, 3, \ldots)$ események teljesen függetlenek (annak ellenére, hogy A_k a $\xi_1, \xi_2, \ldots, \xi_k$ változók mindegyikétől függ). Ez a tény alkotja az alapját az összes további vizsgálatainknak. E tényből az 1. és 2. §. eredményeire is adódik egy újabb bizonyítás. Bebizonyítjuk, hogy 1 valószínűséggel $\lim_{n\to\infty} \sqrt[n]{v_n} = e$; ez felfogható, mint a nagy számok erős törvénye a $\log \frac{v_n}{v_{n-1}}$ gyengén függő változókra vonatkozólag.

Az 5. §-ban bebizonyítjuk, hogy $\log v_n$ határértékben normális eloszlású n várható értékkel és $\sqrt{n}$ szórással; ez felfogható mint a $\log \frac{v_n}{v_{n-1}}$ gyengén függő változókra vonatkozó centrális határeloszlástétel. A 6. § a v_n változókra vonatkozó iterált logaritmus tétellel foglalkozik, míg a 7. §-ban bizonyos, a v_n változók és permutációk ciklusok szorzataként előállítása közötti érdekes összefüggésekre mutatunk rá. Végül a 8. §-ban a v_n változók és a valós számok módosított Engel-féle sorral történő előállítása közötti meglepő összefüggést vizsgáljuk.

1. §. A kiemelkedő elemek sorszámának együttes eloszlása

Bebizonyítjuk, hogy a v_n változók homogén Markov-láncot alkotnak és egyben meghatározzuk ennek a Markov-láncnak az átmenet-valószínűségeit. Ehhez azonban szükségünk van a $v_1, v_2, \ldots, v_n$ változók együttes eloszlásának a meghatározására. Mint már a bevezetésben megmutattuk, az általánosság megszorítása nélkül feltehetjük, hogy a ξ_k változók egyenletes eloszlásúak a (0, 1) intervallumban. Legyen $1 < k_1 < k_2 < \ldots < k_n$, ahol $k_1, k_2, \ldots, k_n$ egész számok. Könnyen belátható, hogy ha[1]

$$G(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n) = \mathbf{P}(v_1 = k_1, \ldots, v_n = k_n, \xi_1 < x_0, \xi_{k_1} < x_1, \ldots \xi_{k_n} < x_n) \tag{1.1}$$

és

$$g(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n) = \frac{\partial^{n+1} G(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n)}{\partial x_0 \partial x_1 \ldots \partial x_n}, \tag{1.2}$$

akkor

$$g(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n) = x_0^{k_1-2} x_1^{k_2-k_1-1} \ldots x_{n-1}^{k_n-k_{n-1}-1}, \tag{1.3}$$

hacsak $0 < x_0 < x_1 < \ldots < x_n$, és így

$$\mathbf{P}(v_1 = k_1, \ldots, v_n = k_r) = \iint\limits_{0 < x_0 < x_1 < \ldots < x_n < 1} \ldots \int x_0^{k_1-2} x_1^{k_2-k_1-1} \ldots x_{n-1}^{k_n-k_{n-1}-1} \, dx_0 dx_1 \ldots dx_n. \tag{1.4}$$

Ennélfogva

$$\mathbf{P}(v_1 = k_1, v_2 = k_2, \ldots, v_n = k_n) = \frac{1}{(k_1 - 1)(k_2 - 1) \ldots (k_{n-1} - 1)(k_n - 1) k_n}. \tag{1.5}$$

Ilyen módon

$$\mathbf{P}(v_n = k_n | v_1 = k_1, \ldots, v_{n-1} = k_{n-1}) = \frac{k_{n-1}}{(k_n - 1) k_n}. \tag{1.6}$$

[1] Itt és a következőkben P(...) mindig a zárójelben álló esemény valószínűségét jelöli.

(1. 6)-ból leolvasható, hogy a ν_n $(n=0, 1, \ldots)$ változók homogén Markov-láncot alkotnak, melynek átmenetvalószínűségei

(1. 7) $$\mathbf{P}(\nu_n=k|\nu_{n-1}=l)=\frac{l}{(k-1)k}, \quad \text{ha} \quad k \geqq l+1 \geqq n+1, \quad \text{és } n=1, 2, 3, \ldots .$$

Speciálisan, ha $n=1$ (mivel $\nu_0 \equiv 1$) (1. 7)-ből következik, hogy

(1. 8) $$\mathbf{P}(\nu_1=k)=\frac{1}{(k-1)k}, \quad \text{ha} \quad k \geqq 2.$$

(1. 8)-ból leolvasható, hogy ν_1 *várható értéke végtelen nagy.* Erre a paradox tényre először S. S. WILKS [12] mutatott rá, részletesebben pedig E. J. GUMBEL [13] elemezte.

(1. 5) és (1. 7) alapján megválaszolható számos, a ν_n sorozatra vonatkozó kérdés. Meghatározhatjuk például ν_n eloszlását. Nyilvánvalóan, ha $k \geqq n+1$, akkor

(1. 9) $$\mathbf{P}(\nu_n=k)=\sum_{1<k_1<k_2<\ldots<k_{n-1}<k} \frac{1}{(k_1-1)(k_2-1)\ldots(k_{n-1}-1)(k-1)k}.$$

Határozzuk most meg rögzített $k \geqq 2$ mellett a $\mathbf{P}(\nu_n=k)$ $(n=1, 2, \ldots, k-1)$ számsorozat generátorfüggvényét, vagyis a

(1. 10) $$G_k(z)=\sum_{n=1}^{k-1} \mathbf{P}(\nu_n=k)z^n \qquad (k=2, 3, \ldots)$$

polinomot.

Könnyen belátható, hogy

(1. 11) $$G_k(z)=\frac{z}{k(k-1)}\prod_{j=1}^{k-2}\left(1+\frac{z}{j}\right)=(-1)^k\frac{\binom{1-z}{k}}{z-1}.$$

Nyilvánvaló, hogy

(1. 12) $$G_k(1)=\sum_{n=1}^{k-1} \mathbf{P}(\nu_n=k)$$

annak a valószínűsége, hogy a k szám előforduljon a ν_n számsorozatban. (1. 11)-ből

(1. 13) $$G_k(1)=\frac{1}{k(k-1)}\prod_{j=1}^{k-2}\left(\frac{j+1}{j}\right)=\frac{1}{k}.$$

Fennáll tehát a következő

1. LEMMA. *Annak valószínűsége, hogy egy folytonos eloszlású sokaságból vett végtelen minta k-adik eleme kiemelkedő legyen,* $\frac{1}{k}$*-val egyenlő.*

Az (1. 11) összefüggésből kiindulva meghatározhatjuk bármely rögzített n-re a $\mathbf{P}(\nu_n=k)$ $(k=n+1, n+2, \ldots)$ számsorozat generátorfüggvényét is.

Legyen

(1. 14) $$H(z, w)=\sum_{n=0}^{\infty}\sum_{k=n+1}^{\infty} \mathbf{P}(\nu_n=k)z^n w^k,$$

akkor

$$H(z, w) = w + \sum_{k=2}^{\infty} G_k(z) w^k = \frac{(1-w)^{1-z} - 1}{z-1}. \tag{1.15}$$

Ennélfogva

$$H(z, w) = \left(\frac{1-w}{z-1}\right) \sum_{n=0}^{\infty} \frac{z^n \left(\log \frac{1}{1-w}\right)^n}{n!} - \frac{1}{z-1},$$

tehát, ha $n \geqq 1$,

$$\sum_{k=n+1}^{\infty} \mathbf{P}(\nu_n = k) w^k = (w-1) \sum_{l=0}^{n} \frac{\left(\log \frac{1}{1-w}\right)^l}{l!} + 1. \tag{1.16}$$

2. §. A kiemelkedő elemek számának eloszlása megadott elemszámú megfigyeléssorozatban

E §-ban a következő kérdéssel foglalkozunk: Jelölje μ_N a $\xi_1, \xi_2, \ldots, \xi_N$ sorozatban a kiemelkedő elemek számát; meghatározandó μ_N eloszlása. (1. 1)-hez hasonlóan belátható, hogy ha $1 < k_1 < k_2 < \ldots < k_n \leqq N$, midőn $k_1, k_2, \ldots, k_n$ egész számok, és

$$G_N(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n) = \tag{2.1}$$
$$= \mathbf{P}(\nu_1 = k_1, \ldots, \nu_n = k_n, \nu_{n+1} > N, \xi_1 < x_0 \xi_{\nu_1}, < x_1, \ldots, \xi_{\nu_n} < x_n)$$

és

$$g_N(k_1, \ldots, k_n; x_0, \ldots, x_n) = \frac{\partial^{n+1} G_N(k_1, \ldots, k_n; x_0, x_1, \ldots, x_n)}{\partial x_0 \partial x_1 \ldots \partial x_n}, \tag{2.2}$$

akkor

$$g_N(k_1, \ldots, k_n; x_0, \ldots, x_n) = x_0^{k_1-2} x_1^{k_2-k_1-1} \ldots x_{n-1}^{k_n-k_{n-1}-1} x_n^{N-k_n}, \tag{2.3}$$

ha $0 < x_0 < x_1 < \ldots < x_n$. Ebből következik, hogy ha $1 < k_1 < k_2 < \ldots < k_n \leqq N$, akkor

$$\mathbf{P}(\nu_1 = k_1, \ldots, \nu_n = k_n, \nu_{n+1} > N) = \frac{1}{(k_1-1)(k_2-1)\ldots(k_n-1)N}. \tag{2.4}$$

Ennélfogva, ha

$$Q_N(n) = \mathbf{P}(\mu_N = n), \tag{2.5}$$

akkor $Q_1(1) = 1$ és ha $N \geqq 2$,

$$Q_N(n) = \frac{1}{N} \sum_{1 < k_1 < k_2 < \ldots < k_n \leqq N} \frac{1}{(k_1-1)(k_2-1)\ldots(k_n-1)} \qquad (n = 1, 2, 3, \ldots) \tag{2.6}$$

és így

$$H_N(x) = \sum_{n=1}^{N} Q_N(n) x^n = \frac{x}{N} \prod_{r=1}^{N-1} \left(1 + \frac{x}{r}\right). \tag{2.7}$$

Ezzel meghatároztuk μ_N generátorfüggvényét. Nyilván fennáll a következő azonosság:

$$H_N(x) = \prod_{l=1}^{N} \left(1 + \frac{x-1}{l}\right). \tag{2.8}$$

Ennek alapján kiszámíthatjuk μ_N momentumait. (2. 7)-ből, $\mathbf{M}(\zeta)$-val jelölve egy ζ valószínűségi változó várható értékét és $\mathbf{D}(\zeta)$-val a szórását,

$$\mathbf{M}(\mu_N) = \sum_{s=1}^{N} \frac{1}{s} \tag{2.9}$$

és

$$\mathbf{D}^2(\mu_N) = \sum_{s=1}^{N} \frac{1}{s} - \sum_{s=1}^{N} \frac{1}{s^2}. \tag{2.10}$$

Ilyen módon

$$\mathbf{M}(\mu_N) = \log N + O(1) \tag{2.11}$$

és

$$\mathbf{D}^2(\mu_N) = \log N + O(1). \tag{2.12}$$

Hasonlóképpen számíthatók ki μ_N magasabb momentumai is.

A (2. 7) képlet a következő alakban is felírható:

$$H_N(x) = \sum_{n=1}^{N} \frac{|S_N^{(n)}| x^n}{N!}, \tag{2.13}$$

ahol $S_N^{(n)}$ az elsőfajú Stirling-számokat jelölik (l. [1] 142.). Így tehát

$$Q_N(n) = \frac{|S_N^{(n)}|}{N!}. \tag{2.14}$$

Ilyen módon az elsőfajú Stirling-számok egyszerű valószínűségszámítási értelmezését nyertük. (A Stirling-számok egy (bonyolultabb) valószínűségszámítási interpretációját lásd [1] 166—167.)

3. §. A különböző megfigyelések kiemelkedő voltának függetlensége

Jelölje A_k azt az eseményt, hogy egy folytonos eloszlású sokaságból vett végtelen minta k-adik eleme kiemelkedő legyen. E §-ban a következő egyszerű, de alapvető tényt bizonyítjuk be

2. LEMMA. *Ha A_k $(k = 1, 2, \ldots)$ jelöli azt az eseményt, hogy egy folytonos eloszlású sokaságból vett minta k-adik eleme kiemelkedő, akkor az $A_1, A_2, \ldots, A_k, \ldots$ események függetlenek.*

E lemmára, annak fontosságára való tekintettel, három különböző bizonyítást adunk.

Első bizonyítás. Ha $1<j_1<j_2<\ldots<j_k$ tetszőleges egész számok, akkor az 1. és 2. §-ban elvégzett meggondolásokhoz hasonlóan adódik, hogy (események szorzatán azok együttes bekövetkezését értve)

$$(3.1)\quad \mathbf{P}(A_{j_1}A_{j_2}\ldots A_{j_k})=\underset{0<x_1<x_2<\ldots<x_k<1}{\int\int\ldots\int} x_1^{j_1-1}x_2^{j_2-j_1-1}\ldots x_k^{j_k-j_{k-1}-1}dx_1dx_2\ldots dx_k$$

és így

$$(3.2)\quad \mathbf{P}(A_{j_1}A_{j_2}\ldots A_{j_k})=\frac{1}{j_1j_2\ldots j_k},$$

mivel továbbá az 1. lemma szerint $P(A_k)=\frac{1}{k}$, tehát

$$(3.3)\quad \mathbf{P}(A_{j_1}A_{j_2}\ldots A_{j_k})=\mathbf{P}(A_{j_1})\mathbf{P}(A_{j_2})\ldots\mathbf{P}(A_{j_k}).$$

Ezzel a 2. lemmát bebizonyítottuk.

Második bizonyítás. Rendezzük el a $\xi_1,\xi_2,\ldots,\xi_n$ valószínűségi változókat nagyság szerint. Legyen közülük a legnagyobb $\xi_{l_1^{(n)}}$, a második legnagyobb $\xi_{l_2^{(n)}}$, és általában jelölje $\xi_{l_r^{(n)}}$ a nagyság szerint r-ediket ($r=1,2,\ldots,n$). Az $(l_1^{(n)},l_2^{(n)},\ldots,l_n^{(n)})$ számsorozat az $1,2,\ldots,n$ számok egy permutációja lesz, és mivel a változók feltevés szerint függetlenek és egyforma eloszlásúak, tehát a lehetséges $n!$ permutáció mindegyike ugyanazzal a valószínűséggel, vagyis $\frac{1}{n!}$ valószínűséggel lép fel. A $\mathbf{P}(A_{j_1}A_{j_2}\ldots A_{j_k})$ valószínűség $(1<j_1<j_2<\ldots<j_k)$ tehát kiszámítható oly módon, hogy meghatározzuk, hogy mely $(l_1^{(j_k)},l_2^{(j_k)},\ldots,l_{j_k}^{(j_k)})$ permutációk esetében lesznek $\xi_{j_1},\xi_{j_2},\ldots,\xi_{j_k}$ kiemelkedő tagok és ezt a számot osztjuk $j_k!$-sal. Nyilvánvaló, hogy ahhoz, hogy $\xi_{j_1},\xi_{j_2},\ldots,\xi_{j_k}$ kiemelkedő tagok legyenek, a következő feltételek teljesülése szükséges és elégséges: a) A $\xi_1,\xi_2,\ldots,\xi_{j_k}$ számok közül ξ_{j_k} legyen a legnagyobb, vagyis $l_1^{(j_k)}=j_k$. b) A $j_k-1,j_k-2,\ldots,j_{k-1}+1$ számok tetszőlegesen helyezkedhetnek el az $(l_1^{(j_k)},l_2^{(j_k)},\ldots,l_{j_k}^{(j_k)})$ sorozatban, j_{k-1} viszont az ezek elhelyezése után üresen maradó helyek közül a balról legelsőre kell hogy kerüljön. c) Általában, miután az $r\geqq j_s$ számokat $(1\leqq s\leqq k)$ elhelyeztük, a $j_s-1,j_s-2,\ldots,j_{s-1}+1$ számokat $(j_0=1)$ tetszés szerint helyezhetjük el az üres helyekre; az ezek után még üresen maradó helyek közül a balról legelső helyre kell helyezni a j_{s-1} számot.

A mondottakból következik, hogy a $j_1,j_2,\ldots,j_k$ számok elhelyezése tekintetében soha nincs lehetőség választásra, e számok egyértelműen meghatározott helyre kerülnek; ha azonban $m<j_k$ és $m\neq j_i$ $(i=1,2,\ldots,k)$, akkor az m szám elhelyezésére m lehetőség áll fenn. Azon $(l_1^{(j_k)},l_2^{(j_k)},\ldots,l_{j_k}^{(j_k)})$ permutációk száma tehát, amelyek létrejötte esetén az $A_{j_1},A_{j_2},\ldots,A_{j_k}$ események bekövetkeznek $\frac{j_k!}{j_1j_2\ldots j_k}$ és így

$$(3.4)\quad \mathbf{P}(A_{j_1}A_{j_2}\ldots A_{j_k})=\frac{1}{j_1j_2\ldots j_k}.$$

Ezzel a 2. lemmára egy újabb bizonyítást nyertünk.

Harmadik bizonyítás. Ismeretes a rendezett minták elméletéből, hogy ha $\xi_1,\xi_2,\ldots,\xi_n$ független és egyforma eloszlású valószínűségi változók, akkor ξ_n értéke ugyanakkora, tehát $\frac{1}{n}$ valószínűséggel esik azon n intervallum mindegyikébe, amelyekre a $\xi_1, \xi_2, \ldots, \xi_{n-1}$ változók értékei a valós tengelyt felbontják. Ugyanis annak valószínűsége, hogy ξ_n a $\xi_1, \xi_2, \ldots, \xi_{n-1}$ változók által meghatározott intervallumok közül felülről az l-edikbe essék ($l=1, 2, \ldots, n$, ahol tehát elsőnek nevezzük az $x \geqq \max(\xi_1, \xi_2, \ldots, \xi_{n-1})$ intervallumot és l-ediknek az $x < \min(\xi_1, \xi_2, \ldots, \xi_{n-1})$ intervallumot) azt jelenti, hogy ξ_n a $\xi_1, \xi_2, \ldots, \xi_n$ sorozatnak nagyság szerint felülről az l-edik eleme, és mivel az $1, 2, \ldots, n$ számoknak $(n-1)!$ olyan permutációja van, amelyben n az l-edik helyen áll, tehát e valószínűség $\frac{(n-1)!}{n!} = \frac{1}{n}$. Speciálisan tehát $\frac{1}{n}$ annak a valószínűsége, hogy ξ_n kiemelkedő legyen, vagyis $\mathbf{P}(A_n) = \frac{1}{n}$. Ez egyben az 1. lemma egy újabb bizonyítása. Másrészt az említett kombinatorikai meggondolásból az is következik, hogy az A_n esemény feltételes valószínűsége is $\frac{1}{n}$ minden olyan feltétel mellett, amely a $\xi_1, \xi_2, \ldots, \xi_{n-1}$ változóknak pusztán az egymás közötti nagyság szerinti sorrendjére vonatkozik, de e változók értékére nézve megszorítást nem jelent. Ha ugyanis az említett feltételt B-vel jelölve a B esemény a $\xi_1, \xi_2, \ldots, \xi_{n-1}$ változók elvileg lehetséges $(n-1)!$ nagyság szerinti elrendeződése közül $N \leqq (n-1)!$ elrendeződés esetén teljesül, akkor

$$\mathbf{P}(A_n|B) = \frac{N}{Nn} = \frac{1}{n}. \tag{3.5}$$

Ha most B-nek választjuk az $A_{j_1}A_{j_2} \ldots A_{j_{k-1}}$ eseményt és $n=j_k$, ahol $1 < j_1 < j_2 < \ldots \ldots < j_k$, akkor azonnal adódik a 2. lemma állítása.

A 2. lemmára adott harmadik bizonyítás segítségével bebizonyíthatjuk a 2. lemma következő általánosítását is.

3. LEMMA. *Legyenek $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ egyforma, folytonos eloszlású független valószínűségi változók és jelölje α_k azt, hogy ξ_k a $\xi_1, \xi_2, \ldots, \xi_k$ számsorozatot nagyság szerint csökkenőleg rendezve el hányadik helyen áll $(k=1, 2, \ldots)$. Akkor az $\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots$ valószínűségi változók teljesen függetlenek, $\alpha_1 \equiv 1$, és ha $n=2, 3, \ldots$*

$$\mathbf{P}(\alpha_n = l) = \frac{1}{n} \quad \textit{midőn} \quad l = 1, 2, \ldots, n. \tag{3.6}$$

Bizonyítás. A 3. lemma összes állításának bebizonyításához nyilván elegendő megmutatni, hogy ha $n \geqq 2$ tetszőleges, és továbbá $a_1, a_2, \ldots, a_n$ tetszőleges olyan egész számok, melyekre $1 \leqq a_k \leqq k$ $(k=2, 3, \ldots, n)$, akkor

$$\mathbf{P}(\alpha_2 = a_2, \alpha_3 = a_3, \ldots \alpha_n = a_n) = \frac{1}{n!}. \tag{3.7}$$

Az $a_2, a_3, \ldots, a_n$ számok minden megengedett választása $(1 \leqq a_k \leqq k;$ $k=2, 3, \ldots, n)$ egyértelműen meghatározza a $\xi_1, \xi_2, \ldots, \xi_n$ számsorozat nagyság szerinti elrendeződését, tehát a fentiekben $(l_1^{(n)}, l_2^{(n)}, \ldots, l_n^{(n)})$-nel jelölt permutációk közül egy meghatározott permutációnak felel meg; ugyanis a_k pontosan azt adja meg, hogy a $\xi_1, \xi_2, \ldots, \xi_k$ számokat nagyság szerint csökkenőleg elrendezve, ξ_k-t hány olyan ξ_j előzi meg, amelyre $j < k$. Ebből már a 3. lemma állítása következik.

A 3. lemma nyilvánvaló módon speciális esetként tartalmazza a 2. lemmát, hiszen (1)-ből azonnal adódik, figyelembe véve, hogy n független valószínűségi változó közül bizonyosokat kiválasztva, azok is függetlenek lesznek, hogy

$$(3.8)\qquad \mathbf{P}(A_{j_1}A_{j_2}\dots A_{j_k})=\mathbf{P}(\alpha_{j_1}=j_1,\alpha_{j_2}=j_2,\dots,\alpha_{j_k}=j_k)=\frac{1}{j_1 j_2\dots j_k}.$$

Ez a 2. lemma egy újabb (negyedik) bizonyításának tekintendő.

Valójában azonban a 2. lemmára adott második, harmadik és negyedik bizonyítás ugyananannak a kombinatorikai gondolatnak a különböző variánsai. Tulajdonképpen tehát a 2. lemmához csak két, lényegében különböző út vezet: beláthatjuk a 2. lemmát analitikusan (l. az első bizonyítást), vagy kombinatorikai meggondolás segítségével.

A 2. lemma segítségével az 1. és 2. § eredményeire új bizonyítások adódnak. Így pl. (1. 5) a következőképpen látható be: ha $l_1, l_2, \dots, l_{k_n-n}$ jelölik az $1, 2, \dots, k_n-1$ sorozat $k_1, k_2, \dots, k_{n-1}$-től különböző tagjait, akkor ($\bar{A}$-sal jelölve az A eseménnyel ellentétes eseményt)

$$(3.9)\qquad \mathbf{P}(\nu_1=k_1,\dots,\nu_n=k_n)=\mathbf{P}(A_{k_1}A_{k_2}\dots A_{k_n}\bar{A}_{l_1}\bar{A}_{l_2}\dots\bar{A}_{l_{k_n-n}})$$

és így

$$(3.10)\qquad \mathbf{P}(\nu_1=k_1,\dots,\nu_n=k_n)=\frac{1}{k_1k_2\dots k_n}\prod_{j=1}^{k_n-n}\left(1-\frac{1}{l_j}\right)=\frac{1}{(k_1-1)\dots(k_n-1)k_n}.$$

Az pedig, hogy a ν_n változók Markov-láncot alkotnak, egyenesen magától értetődővé válik, ugyanis általában igaz, hogy ha az η_k változók függetlenek és egész értékűek, és ν_n jelöli az n-edik olyan k indexet, amelyre $\eta_k=a$, ahol a egy rögzített egész szám, akkor a ν_n változók Markov-láncot alkotnak.

Hasonlóképpen látható be (2. 8) is (és ily módon (2. 6), mert az (2. 8)-ból következik).

Legyen ugyanis

$$(3.11)\qquad \varepsilon_k=\begin{cases}1, & \text{ha az } A_k \text{ esemény bekövetkezik,}\\ 0 & \text{egyébként.}\end{cases}$$

Akkor

$$(3.12)\qquad \mu_N=\varepsilon_1+\varepsilon_2+\dots+\varepsilon_N.$$

Mivel pedig az ε_k változók a 2. lemma szerint függetlenek, tehát μ_N generátorfüggvénye egyenlő az ε_k változók $(k=1, 2, \dots, N)$ generátorfüggvényeinek szorzatával. Tekintve, hogy

$$\mathbf{P}(\varepsilon_k=1)=\frac{1}{k}\quad \text{és}\quad \mathbf{P}(\varepsilon_k=0)=1-\frac{1}{k},$$

tehát ε_k generátorfüggvénye $1+\frac{x-1}{k}$ $(k=1, 2, \dots, N)$, ennélfogva μ_N generátorfüggvénye

$$(3.13)\qquad H_N(x)=\prod_{k=1}^{N}\left(1+\frac{x-1}{k}\right).$$

4. §. A nagy számok törvénye a kiemelkedő megfigyelésekre vonatkozólag

E §-ban a következő tételt bizonyítjuk be:

1. TÉTEL. *Ha ν_n jelöli egy folytonos sokaságból vett végtelen minta n-edik kiemelkedő elemének sorszámát, akkor* 1 *valószínűséggel*

(4. 1) $$\lim_{n\to\infty} \sqrt[n]{\nu_n} = e,$$

(ahol e a természetes logaritmus alapja).

Bizonyítás. Kiindulunk abból a nyilvánvaló összefüggésből, hogy

(4. 2) $$\mathbf{P}(\nu_n \leqq k) = \mathbf{P}(\mu_k \geqq n)$$

és így

(4. 3) $$\mathbf{P}(\nu_n > k) = \mathbf{P}(\mu_k < n).$$

Az 1. tétel állításának bebizonyítása céljából válasszunk egy tetszőleges kis ε pozitív számot. Akkor, ha $\{x\}$ jelöli a legkisebb x-nél nem kisebb egész számot,

(4. 4) $$\mathbf{P}(\nu_n > e^{n(1+\varepsilon)}) = \mathbf{P}(\mu_{\{e^{n(1+\varepsilon)}\}} < n)$$

és ha $[x]$ jelöli x egész részét, akkor

(4. 5) $$\mathbf{P}(\nu_n \leqq e^{n(1-\varepsilon)}) = \mathbf{P}(\mu_{[e^{n(1-\varepsilon)}]} \geqq n).$$

Mivel a Csebisev-egyenlőtlenség szerint

(4. 6) $$\mathbf{P}(|\mu_N - M(\mu_N)| \geqq \lambda \mathbf{D}(\mu_N)) \leqq \frac{1}{\lambda^2},$$

tehát (2. 12) és (2. 13) miatt

(4. 7) $$\mathbf{P}(|\mu_N - \log N| \geqq \varepsilon \log N) = O\left(\frac{1}{\log N}\right).$$

(4. 4)-ből, (4. 5)-ből és (4. 7)-ből következik, hogy

(4. 8) $$\mathbf{P}(\sqrt[n]{\nu_n} > e^{1+\varepsilon}) + \mathbf{P}(\sqrt[n]{\nu_n} < e^{1-\varepsilon}) = O\left(\frac{1}{n}\right).$$

Ilyen módon a

(4. 9) $$\sum_{k=1}^{\infty} \left(\mathbf{P}(\sqrt[k^2]{\nu_{k^2}} > e^{1+\varepsilon}) + \mathbf{P}(\sqrt[k^2]{\nu_{k^2}} < e^{1-\varepsilon})\right)$$

sor ε minden pozitív értékére konvergens, amiből a Borel—Cantelli lemma (l. [2]) segítségével szokásos meggondolással következik, hogy 1 valószínűséggel

(4. 10) $$\lim_{k\to+\infty} \sqrt[k^2]{\nu_{k^2}} = e.$$

Bevezetve a $\beta_n = \sqrt[n]{\nu_n}$ jelölést, ha $k^2 < n < (k+1)^2$, akkor ν_n monotonitása miatt

$$(\beta_{k^2})^{\left(\frac{k}{k+1}\right)^2} < \beta_n < (\beta_{(k+1)^2})^{\left(\frac{k+1}{k}\right)^2}, \tag{4.11}$$

tehát (4. 10)-ből azonnal következik az 1. tétel állítása. Az 1. tétel állítása nyilván kimondható a következő ekvivalens alakban is.

1′. TÉTEL. *Ha μ_N jelenti egy folytonos sokaságból vett végtelen minta első N eleme között a kiemelkedő elemek számát, akkor* 1 *valószínűséggel fennáll a*

$$\lim_{N\to\infty} \frac{\mu_N}{\log N} = 1 \tag{4.12}$$

reláció.

Az 1′. tétel speciális esete a nagy számok erős törvénye egy általános alakjának (l. [3], [4], [5]), mely szerint, ha az η_k $(k = 1, 2, \ldots)$ nemnegatív változók függetlenek, $\mathbf{M}(\eta_k) = M_k$, és $\mathbf{D}(\eta_k) = D_k$, bevezetve továbbá az $A_n = \sum_{k=1}^{n} M_k$ jelölést, teljesülnek a következő feltételek:

a) $$\lim_{n\to\infty} A_n = +\infty$$

és

b) a $\sum_{n=1}^{\infty} \frac{D_n^2}{A_n^2}$ sor konvergens,

akkor 1 valószínűséggel fennál, hogy

$$\lim_{N\to+\infty} \frac{\sum_{k=1}^{N} \eta_k}{A_N} = 1. \tag{4.13}$$

E tételt az $\eta_k = \varepsilon_k$ változókra alkalmazva, azonnal adódik az 1′ tétel, figyelembe véve, hogy ez esetben $A_N \sim \log N$ és $D_N^2 = \frac{1}{N} - \frac{1}{N^2}$, tehát

$$\frac{D_N^2}{A_N^2} \sim \frac{1}{N \log^2 N},$$

vagyis az említett tétel a) és b) feltételei teljesülnek. Megjegyzendő, hogy az 1. tétel fent adott bizonyításának módszere az említett általános tétel bizonyítására is felhasználható.

5. A kiemelkedő elemekre vonatkozó centrális határeloszlástétel

E §-ban a következő tételt bizonyítjuk be:

2. TÉTEL. *Jelölje ν_n egy folytonos eloszlású sokaságból vett végtelen minta n-edik kiemelkedő elemét, akkor x minden valós értékére*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\log \nu_n - n}{\sqrt{n}} < x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du. \tag{5.1}$$

Bizonyítás. (4. 2) miatt a 2. tétel állítása ekvivalens a következővel:

2'. TÉTEL. *Ha μ_N jelöli egy folytonos eloszlású sokaságból vett végtelen minta első N eleme között a kiemelkedő elemek számát, akkor x minden valós értékére*

$$\lim_{N\to+\infty} \mathbf{P}\left(\frac{\mu_N - \log N}{\sqrt{\log N}} < x\right) = \Phi(x). \tag{5. 2}$$

(5. 2) azonban azonnal következik a centrális határeloszlás-tétel Ljapunoff-féle alakjából, (l. [2]) ugyanis

$$\mu_N = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_N,$$

ahol

$$\mathbf{M}(\varepsilon_k) = \frac{1}{k},$$

$$\mathbf{D}^2(\varepsilon_k) = \frac{1}{k} - \frac{1}{k^2}$$

$$\mathbf{M}\left(\left|\varepsilon_k - \frac{1}{k}\right|^3\right) = \left(\frac{1}{k} - \frac{1}{k^2}\right)\left(1 - \frac{2}{k} + \frac{2}{k^2}\right) \leqq \frac{1}{k},$$

tehát

$$\frac{\sqrt[3]{\sum_{k=1}^{N} \mathbf{M}\left(\left|\varepsilon_k - \frac{1}{k}\right|^3\right)}}{\sqrt{\sum_{k=1}^{N} \mathbf{D}^2(\varepsilon_k)}} = O\left(\frac{1}{\sqrt[6]{\log N}}\right).$$

6. §. A kiemelkedő elemekre vonatkozó iterált-logaritmus tétel

Fennáll a következő

3. TÉTEL. *Ha ν_n jelöli egy folytonos sokaságból vett végtelen minta n-edik kiemelkedő elemének a sorszámát, akkor 1 valószínűséggel*

$$\limsup_{n\to+\infty} \frac{\log \nu_n - n}{\sqrt{2n \log\log n}} = +1. \tag{6. 1}$$

és

$$\liminf_{n\to-\infty} \frac{\log \nu_n - n}{\sqrt{2n \log\log n}} = -1. \tag{6. 2}$$

A 3. tétel ugyanis (4. 2) miatt ekvivalens a következővel:

3'. TÉTEL. *Ha μ_N jelöli egy folytonos sokaságból vett végtelen minta első N eleme között a kiemelkedőek számát, akkor 1 valószínűséggel*

$$\limsup_{N\to+\infty} \frac{\mu_N - \log N}{\sqrt{2 \log N \cdot \log\log\log N}} = +1 \tag{6. 3}$$

és

$$\liminf_{N\to+\infty} \frac{\mu_N - \log N}{\sqrt{2N\log N\cdot \log\log\log N}} = -1. \tag{6. 4}$$

Mármost a 3' tétel egyszerűen annyit jelent, hogy az ε_k ($k=1, 2, \ldots$) független változókra alkalmazható az iterált logaritmus-tétel, ami abból következik, hogy az ε_k változók összességükben korlátosak és μ_N szórásnégyzete $+\infty$-hez tart (l. [6]).

7. §. Egy véletlen permutáció ciklusai számának eloszlásáról

A μ_N valószínűségi változó felfogható oly módon is, hogy találomra kiválasztjuk az $1, 2, \ldots, N$ számok egy permutációját, (úgy, hogy minden egyes permutáció kiválasztásának valószínűsége ugyanakkora, vagyis $\frac{1}{N!}$ legyen) és megszámoljuk, hogy hány olyan szám van e permutációban, amelyet nem előz meg a permutációban nála nagyobb elem (tehát amely nála nagyobb elemmel nem alkot inverziót) és μ_N-nel jelöljük az ilyen — a továbbiakban kiemelkedőnek nevezett — elemek számát. A μ_N valószínűségi változó eloszlására vonatkozó eredményeink (azaz a (2. 15) explicit formula, a μ_N generátorfüggvényének (2. 8) alatti kifejezése, ennek a képletnek a 3 §-ban adott interpretációja és a 2' tétel) egyben egy másik, véletlen permutációkra vonatkozó problémára is választ adnak, mégpedig arra, hogy egy véletlen permutáció cikluselőállításában szereplő ciklusok számának mi az eloszlása. Fennáll ugyanis a következő összefüggés:

4. LEMMA. *Jelölje γ_N az $1, 2, \ldots, N$ számok egy találomra kiválasztott permutációja cikluselőállításában szereplő ciklusok számát, és μ_N az $1, 2, \ldots, N$ számok egy találomra kiválasztott permutációja kiemelkedő elemeinek a számát, akkor γ_N és μ_N egyforma eloszlásúak, tehát*

$$\mathbf{P}(\gamma_N = k) = \mathbf{P}(\mu_N = k) \qquad (k=1, 2, \ldots, N) \tag{7. 1}$$

Bizonyítás. Ahhoz, hogy a 4. lemmát bebizonyítsuk, elegendő megadni N elem permutációi halmazának egy olyan önmagára való kölcsönösen egyértelmű leképezését, amelynél egy k ciklusból álló Π permutációnak egy k kiemelkedő elemmel bíró Π' permutáció felel meg. E leképezést a következőképpen nyerjük: állítsuk elő a szóban forgó permutációt ciklusok szorzataként; rendezzük el a ciklusokat oly módon, hogy minden ciklusnak megkeressük a legnagyobb elemét és a ciklusokat úgy írjuk egymás mellé, hogy a ciklusok legnagyobb elemei növekvő sorozatot alkossanak. Ezek után egy-egy cikluson belül rendezzük el az elemeket úgy, hogy a ciklus felírása az abban szereplő legnagyobb számmal kezdődjék. Most már csak annyit kell tennünk, hogy a ciklusokat elhatároló zárójeleket elhagyjuk, és az így kapott számsorozat lesz a Π permutációhoz hozzárendelt Π' permutáció. Például, ha $N=9$ és Π az

$$5\ 1\ 3\ 2\ 7\ 9\ 8\ 4\ 6$$

permutáció, akkor Π ciklusok szorzataként való előállítása a fenti előírások szerint elrendezve

$$\Pi = (3)\,(842157)\,(96)$$

és így Π' a

$$3\ 8\ 4\ 2\ 1\ 5\ 7\ 9\ 6$$

permutáció.

Nyilvánvaló, hogy Π' kiemelkedő elemeinek száma egyenlő lesz Π ciklusai számával, hiszen Π ciklusai legnagyobb elemei és csak azok lesznek a kiemelkedő elemek Π'-ben. A leképezés nyilván kölcsönösen egyértelmű, hiszen Π-nek ciklusok szorzataként való felírása egyértelműen meghatározza Π-t és a cikluselőállítás felírására vonatkozó utasításaink is teljesen egyértelműek.

Ezzel tehát bebizonyítottuk, hogy γ_N generátorfüggvénye is $H_N(x)$-szel egyenlő (l. (2. 8)) továbbá, hogy γ_N határértékben szintén normális eloszlású $\log N$ várható értékkel és $\sqrt{\log N}$ szórással. A permutációk ciklusszámának eloszlására vonatkozó ezen eredmények már régebben is ismertek voltak (l. pl. [7], [8]); a fentiekből egy újabb bizonyítás adódik ezen eredményekre.

Kézenfekvő feltenni azt a kérdést, hogy annak megfelelően, hogy μ_N-et előállítottuk mint N független valószínűségi változó összegét, lehet-e ennek megfelelően kombinatorikailag interpretálható módon előállítani γ_N-et is mint független valószínűségi változó összegét úgy, hogy a k-adik az 1 értéket $\frac{1}{k}$ valószínűséggel, a 0 értéket $1-\frac{1}{k}$ valószínűséggel vegye fel. Azt, hogy ez valóban lehetséges, már W. FELLER megmutatta (l. [7]).

8. §. Valós számok módosított és közönséges Engel-féle sorral való előállításáról

E §-ban először a következő tételt bizonyítjuk be.

4. TÉTEL. *Vegyünk egy végtelen mintát egy folytonos eloszlású sokaságból és legyenek $\nu_0=1, \nu_1, \nu_2, \dots, \nu_n, \dots$ a minta kiemelkedő elemeinek sorszámai. Legyen*

$$\alpha=\frac{1}{\nu_1}+\sum_{n=2}^{\infty}\frac{1}{(\nu_1-1)(\nu_2-1)\dots(\nu_{n-1}-1)\nu_n}. \tag{8.1}$$

Akkor az α valószínűségi változó egyenletes eloszlású a $(0, 1)$ intervallumban.

A 4. tétel bizonyításához szükségünk lesz a következő segédtételre

5. LEMMA. *Legyen x tetszőleges valós szám a $(0, 1]$ balról nyílt, jobbról zárt intervallumban, akkor x egyértelműen előállítható az*

$$x=\frac{1}{k_1}+\sum_{n=2}^{\infty}\frac{1}{(k_1-1)(k_2-1)\dots(k_{n-1}-1)k_n} \tag{8.2}$$

alakban, ahol $k_1, k_2, \dots, k_n$ 1-nél nagyobb egész számok egy szigorúan növekvő végtelen sorozata, amely a következőképpen határozható meg: k_1-nek a legkisebb, $\frac{1}{x}$-nél nagyobb számot választjuk, és ha már $k_1, k_2, \dots, k_{N-1}$ meg vannak választva, $(N=2, 3, \dots)$ k_N-nek a legkisebb olyan egész számot választjuk, amelyre $\frac{1}{k_1}+\sum_{n=2}^{N}\frac{1}{(k_1-1)(k_2-1)\dots(k_{n-1}-1)k_n}<x.$

1. m e g j e g y z é s. A (8. 2) előállítás rokon a valós számok *Engel-féle sor* alakjában való, vagyis

$$x=\sum_{n=1}^{\infty}\frac{1}{q_1 q_2\dots q_n} \tag{8. 3}$$

alakú előállításával, ahol $q_1, q_2, \dots, q_n, \dots$ 1-nél nagyobb egész számok egy *nemcsökkenő* végtelen sorozata: a q_n számokat a következő algoritmus szolgáltatja: q_1 a legkisebb egész szám, amely $\frac{1}{x}$-nél nagyobb, és ha már $q_1, q_2, \dots, q_{N-1}$ meg vannak választva ($N=2, 3, \dots$), q_N-nek a legkisebb olyan egész számot választjuk, amelyre $\sum_{n=2}^{N}\frac{1}{q_1 q_2\dots q_n}<x$. A két eljárás hasonlósága miatt a (8. 2) előállítást az x szám *módosított Engel-féle sor*ának nevezzük.

2. m e g j e g y z é s. Ha x racionális szám, akkor x előállítható (8. 2)-höz hasonló *véges* sor alakjában is.

Az 5. lemma bizonyítása. Ha a lemma szerinti algoritmust alkalmazzuk, akkor először is $k_N>k_{N-1}$, ugyanis $\frac{1}{k}+\frac{1}{(k-1)k}=\frac{1}{k-1}$ miatt, ha $k_N\leqq k_{N-1}$ volna, akkor k_{N-1} helyett eggyel kisebb szám is választható lett volna, ami ellentmond az algoritmus szabályainak; másrészt

$$\left|x-\frac{1}{k_1}-\sum_{n=1}^{N}\frac{1}{(k_1-1)(k_2-1)\dots(k_{n-1}-1)k_n}\right|\leqq\frac{1}{(k_1-1)\dots(k_N-1)k_N} \tag{8. 4}$$

és mivel $k_N>N$, tehát a (8. 4) jobb oldala 0-hoz tart, ha $N\to\infty$, miért is (8. 2) fennáll.

A 4. tétel bizonyítása. Legyen $1<k_1<k_2<\dots<k_N$,

$$a_N=\frac{1}{k_1}+\sum_{n=2}^{N}\frac{1}{(k_1-1)\dots(k_{n-1}-1)k_n} \tag{8. 5}$$

és

$$b_N=a_N+\frac{1}{(k_1-1)\dots(k_N-1)k_N}. \tag{8. 6}$$

Annak valószínűsége, hogy α beleessék az $(a_N, b_N]$ intervallumba, nyilván egyenlő annak valószínűségével, hogy $v_1=k_1, \dots, v_N=k_N$ legyen, és így (1. 5) szerint e valószínűség

$$\frac{1}{(k_1-1)\dots(k_N-1)k_N}=b_N-a_N, \tag{8. 7}$$

vagyis a keresett valószínűség egyenlő az (a_N, b_N) intervallum hosszával. Ebből következik, hogy tetszőleges $(a, b]$ intervallumra $(0\leqq a<b\leqq 1)$

$$\mathbf{P}(a<\alpha\leqq b)=b-a, \tag{8. 8}$$

tekintve, hogy minden intervallum előállítható véges vagy megszámlálhatóan végtelen sok idegen $(a_N, b_N]$ alakú intervallum egyesítéseként. Ezzel a 4. tételt bebizonyítottuk.

A 4. tétel úgy is interpretálható, hogy ha az x számot a $(0, 1]$ intervallumban találomra választjuk, egyenletes eloszlás szerint, és a k_n számsorozatot x (8. 2) alakú előállításával definiáljuk, akkor a k_n $(n=1, 2, \ldots)$ valószínűségi változókra ugyanazok a statisztikus törvényszerűségek érvényesek, mint amelyeket a v_n változókra az 1., 4., 5. és 6. §-ban bebizonyítottunk, tehát például igaz az, hogy annak a valószínűsége, hogy a k pozitív egész szám $(k \geqq 2)$ előforduljon a k_n sorozatban, $\frac{1}{k}$-val egyenlő, hogy majdnem minden x számra $\lim_{n\to\infty} \sqrt[n]{k_n} = e$, továbbá, hogy ha $E_n(y)$ jelöli azon x számok halmazát $(0 \leqq x \leqq 1)$, amelyekre $\frac{\log k_n - n}{\sqrt{n}} < y$, akkor $E_n(y)$ Lebesgue-mértéke $\Phi(y)$-hoz konvergál, ha $n \to +\infty$, s í. t.

E tételek az Engel-féle sorokra vonatkozó megfelelői ismertek; pl. az, hogy majdnem minden x-re $\lim \sqrt[n]{q_n} = e$, E. BORELtől [9], míg az hogy a $\frac{\log q_n - n}{\sqrt{n}} < y$ egyenlőtlenségnek eleget tevő x számok halmazának Lebesgue-mértéke $\Phi(y)$-hoz konvergál, ha $n \to \infty$, P. LÉVYtől származik [10]. E tételek és ezek bizonyos élesítéseinek részletes bizonyítása a [10] dolgozatban található meg. Jelen dolgozat módszere csekély módosítása útján az említett közönséges Engel-féle sorokra vonatkozó, BORELtől és LÉVYtől származó, továbbá ERDŐS PÁL, SZÜSZ PÉTER és a szerző által talált tételekre új és az eddig ismerteknél lényegesen egyszerűbb bizonyításokat találhatunk.

Jelölje ugyanis $\varepsilon_k(x)$ azt, hogy a k szám *hányszor* fordul elő az x valós szám közönséges Engel-féle sorának nevezői között, tehát a (8. 3) által definiált $q_n = q_n(x)$ $(n=1, 2, \ldots)$ számsorozatban. Könnyen belátható, hogy ha x a $(0, 1]$ intervallumban egyenletes eloszlású valószínűségi változó, akkor

$$\mathbf{P}(\varepsilon_k(x) = r) = \frac{k-1}{k^{r+1}} \qquad (r=0, 1, 2, \ldots) \tag{8.9}$$

(l. [10] 1. tétel). Azt is egyszerűen beláthatjuk továbbá, hogy az $\varepsilon_k(x)$ valószínűségi változók $(k=1, 2, 3, \ldots)$ *függetlenek*. Ugyanis annak a valószínűsége, hogy $q_1(x) = q_1, q_2(x) = q_2, \ldots, q_n(x) = q_n$ legyen (ahol most $q_1, q_2, \ldots, q_n$ pozitív egész számok egy rögzített nemcsökkenő sorozata), nyilvánvalóan

$$\frac{1}{q_1 q_2 \cdots q_n (q_n - 1)} \tag{8.10}$$

és így annak a valószínűsége, hogy a $q_n(x)$ $(n=1, 2, \ldots)$ sorozatban a 2 szám r_2-ször, a 3 szám r_3-szor, ..., a k szám (pontosan) r_k-szor forduljon elő,

$$\mathbf{P}(\varepsilon_2 = r_2, \ldots, \varepsilon_k = r_k) = \frac{1}{2^{r_2} 3^{r_3} \ldots k^{r_k}(k-1)} - \frac{1}{2^{r_2} 3^{r_3} \ldots (k-1)^{r_{k-1}} k^{r_k+1}(k-1)}, \tag{8.11}$$

és így

$$\mathbf{P}(\varepsilon_2 = r_2, \ldots, \varepsilon_k = r_k) = \prod_{l=2}^{k} \frac{1}{l^{r_l}} = \prod_{l=2}^{k} \frac{l-1}{l^{r_l+1}}, \tag{8.12}$$

tehát (8. 9)-re való tekintettel

$$\mathbf{P}(\varepsilon_2 = r_2, \ldots, \varepsilon_k = r_k) = \prod_{l=2}^{k} \mathbf{P}(\varepsilon_l = r_l), \tag{8.13}$$

amiből az $\varepsilon_2, \ldots, \varepsilon_k$ változók függetlensége már leolvasható.

Legyen

$$\mu_N(x) = \sum_{k=1}^{N} \varepsilon_k(x), \tag{8.14}$$

vagyis jelölje $\mu_N(x)$ azt, hogy a $q_1(x), q_2(x), \ldots$ sorozatnak hány N-nél nem nagyobb tagja van. Akkor nyilván fennáll a

$$\mathbf{P}(\mu_N(x) \geqq n) = \mathbf{P}(q_n(x) \leqq N) \tag{8.15}$$

azonosság és ennek segítségével belátható, hogy

$$\lim_{n \to +\infty} \sqrt[n]{q_n(x)} = e \tag{8.16}$$

ekvivalens azzal, hogy

$$\lim_{N \to +\infty} \frac{\mu_N(x)}{\log N} = 1. \tag{8.17}$$

Az, hogy (8. 17) majdnem minden x-re fennáll, ugyanúgy következik a nagy számok törvényének a 4. §-ban említett általános alakjából, mint az 1′ tétel. Hasonlóképpen látható be a $q_n(x)$-re vonatkozó centrális határeloszlástétel és iterált logaritmustétel (l. [11]) is. Így tehát a megfigyeléssorozatok kiemelkedő elemeire vonatkozó vizsgálataink meglepő módon az Engel-féle sorok elméletének lényeges egyszerűsítéséhez is elvezettek.

Befejezésül még csak egy megjegyzést kívánunk tenni. Jelölje $q_n(x)$ $(n = 1, 2, \ldots)$ az x szám közönséges Engel-féle sorának nevezőit és jelölje k_n^* a nagyság szerint n-edik pozitív egész számot a $q_n(x)$ számsorozatban, vagyis a $q_n(x)$ számsorozatból töröljük az ismétléseket, tehát minden előforduló számot csak egyszer tartsunk meg; akkor a k_n^* számsorozat pontosan ugyanazokkal a statisztikai tulajdonságokkal rendelkezik, mint a módosított Engel-féle sorok k_n nevezői, illetve mint a folytonos sokaságból vett végtelen minta kiemelkedő elemeinek v_n indexei. Így tehát a kiemelkedő mintaelemek elmélete a közönséges Engel-féle sorokkal is közvetlen kapcsolatban áll.

IRODALOM

[1] CH. JORDAN: *Calculus of finite differences*, Budapest, 1939.
[2] RÉNYI A.: *Valószínűségszámítás*, Tankönyvkiadó, Budapest, 1954.
[3] M. LOÈVE: *Probability theory*, Van Nostrand, New York, 1955., 238 o.
[4] RÉNYI A.: A számjegyek eloszlása valós számok Cantor-féle előállításaiban, *Matematikai Lapok* **7** (1956) 77—100.
[5] RÉNYI A.: A new axiomatic theory of probability, *Acta Math. Acad. Sci. Hung.* **6** (1955) 185—336.
[6] A. KOLMOGOROFF: Über das Gesetz des iterierten Logarithmus, *Math. Ann.* 1927.
[7] W. FELLER: *An introduction to probability theory and its applications*, New York, 1950. Ch. 10. § 7.
[8] J. RIORDAN: *An introduction to combinatorial analysis*, Wiley, New York, 1958., 70—71 o.
[9] É. BOREL: Sur les développements unitaires normaux, *Comptes Rendus* **225** (1947) 51.
[10] P. LÉVY, Remarques sur un théoreme de M. Émile Borel, *Comptes Rendus* **225** (1947) 918—919.
[11] P. ERDŐS, A. RÉNYI, P. SZÜSZ: On Engel's and Sylvester's series, *Annales Univ. Sci. Budapest, Sect. Math.* **1** (1958) 7—32.
[12] S. S. WILKS: Recurrence of extreme observations, *Journal of the Australian Math. Soc.* **1** (1959) 106—122.
[13] E. J. GUMBEL: The return period of order statistics, *Annals of the Institute of Statistical Mathematics* **12** (1961) 249—256.

(Beérkezett: 1962. I. 21.)