

# A MATEMATIKAI STATISZTIKA ORVOSI-BIOLÓGIAI ALKALMAZÁSÁNAK NÉHÁNY PROBLÉMÁJÁRÓL

JUVANCZ IRÉNEUSZ és LIPTÁK TAMÁS

## ÖSSZEFOGLALÁS

A közlemény első része az alkalmazások elvi kérdéseivel foglalkozik. Az orvosi-biológiai kutatás természetéből adódó speciális problémákat két súlyponti kérdés körül csoportosítja. Az egyik problémakör a szignifikancia-vizsgálathoz, mégpedig magának a szignifikanciának helyes értelmezéséhez kapcsolódik, a másik problémakör pedig az orvosi-biológiai kutatásban gyakran szükségszerűen *kiselemszámú* mintavételből ered. Az első kérdéscsoportra vonatkozó vizsgálatok befejezetlensége miatt a cikk csupán a probléma felvetésével és jelentőségével foglalkozik. A másik kérdéscsoportot illetően kismintákra is kidolgozott eloszlásmentes statisztikai módszerek bevezetését javasolja, azaz olyan módszerek alkalmazását, melyek a vizsgált statisztikai sokaságra vonatkozó csaknem minden előzetes feltevéstől függetlenek.

A dolgozat második része a fent említett javaslatokhoz kapcsolódva ismerteti a rendstatisztika (más kifejezéssel: variációs sorok elmélete) főbb tételeit s közli a legújabb ezirányú szovjet eredményeket is. Ezek felhasználásával a közlemény harmadik része kismintákra is használható eloszlásmentes módszereket dolgoz ki szignifikancia-vizsgálathoz. Végül ezen új módszerek gyakorlati alkalmazását néhány orvosi vonatkozású példán illusztrálja és megadja a számításokhoz szükséges táblázatokat, melyeket az Alkalmazott Matematikai Intézet Orvosi Matematikai Statisztikai Csoportja dolgozott ki.

A közlemény az orvosi-biológiai alkalmazásokból indul ki, de az itt tárgyalt elvi kérdések közül több általánosabb jellegű, a rendstatisztikán alapuló eloszlásmentes statisztikai eljárások pedig az alkalmazások széles területén érdeklődésre és felhasználásra tarthatnak számot.

### 1.

#### *Elvi kérdések*

Mint az AMI orvos-statisztikai csoportja, egy éve foglalkozunk rendszeresen orvosi és biológiai kísérletek és megfigyelések eredményeinek értékelésével és több esetben már ezek tervezésében is résztvettünk, továbbá orvosi szaklapok matematikai-statisztikai vonatkozású cikkeit lektoráltuk. Munkánk során közeli kapcsolatba kerültünk az orvosi-biológiai kutatásokkal, a hozzánk forduló orvosokkal és kutatókkal történt rendszeres megbeszéléseink alkalmával egyre jobban megismertük a matematikai statisztika alkalmazásának itt felmerülő speciális problémáit és körülményeit. Közleményünk keretében munkánk során leszűrt tapasztalataink közül a legfontosabbakat szeretnénk ismertetni és egyes kérdésekre konkrét megoldást adni. Bár sok kérdésre vonatkozólag csak a megoldás irányát tudjuk vázolni, úgy gondoljuk, már a

problémák felismerése, pontos körvonalazása és ismertetése is fontos lépés azok megoldása felé.

Mindenekelőtt szeretnénk vázolni azt, miért van szükség az orvosi-biológiai kutatás terén *matematikai statisztika alkalmazására*. Ugyanis éppen azok az okok, melyek az alkalmazást szükségessé teszik, más tényezőkkel egybefonódva egyben meg is nehezítik azt.

Matematikai statisztika (és általában valószínűségszámítás) alkalmazására abban az esetben van szükség, ha valamely természeti jelenség törvényeinek, lefolyásának, kimenetelének vizsgálatakor valamilyen okból nem vehetjük számításba a jelenséget egyértelműen meghatározó tényezők teljes összességét. Ez bekövetkezhet ismereteink hiányossága miatt, vagy például azért, mert az összes ható tényezők figyelembevétele túlságosan bonyolulttá, megengedhetetlenül hosszadalmassá tenné a vizsgálatot. Ez utóbbi eset áll elő például akkor, ha játékkockával dobunk: a kocka összes lényeges fizikai adatainak, a feldobás módjának és bizonyára egész sor egyéb tényezőnek ismeretében nyilvánvalóan pontosan előre megmondhatnánk, hogy adott kiinduló helyzetből a kocka melyik lapjára fog esni: de nyilvánvaló, hogy ez gyakorlatilag kivihetetlen. Ha azonban ilyenkor a szóbanforgó jelenség nem egyszeri, individuális előfordulású, hanem azonos körülmények között elvben tetszőlegesen sokszor megismételhető, bizonyos tapasztalati tények és szimmetriák figyelembevételével a szóbanforgó jelenségre vonatkozólag mégis objektív törvényszerűségeket állapíthatunk meg: ezek segítségével újabb és újabb tényezők hatását vehetjük figyelembe. Az ilyen típusú, ú. n. *véletlen tömegjelenségek* vizsgálatával foglalkozik a valószínűségszámítás. Ennek egyik fejezete a matematikai statisztika, amely az általában ismeretlen eloszlású statisztikai sokaságokból vett minták segítségével igyekszik becsléseket, információkat nyerni a statisztikai sokaságról. A matematikai statisztika másik s az orvosi gyakorlatban leginkább előforduló feladata annak megítélése, hogy egy bizonyos tényező hatással van-e a vizsgált jelenségre. Erre szolgál az ú. n. szignifikancia-vizsgálat. A kísérletező azt kérdezi, van-e hatása egy bizonyos tényezőnek a vizsgált jelenségre. Ebből a célból két kísérletet végez. Az első kísérlet során a szóbanforgó tényező hatását a vizsgált jelenségre minél kisebbre igyekszik csökkenteni, a másodikban pedig éppen fordítva, ezt a hatást fokozza. A két kísérlet eredményeként nyert megfigyelések két *mintát* alkotnak (»kontroll« és »kezelt« egyedekből álló minták). A két minta alapján kell ítéletet mondanunk arról, van-e lényeges hatása a szóbanforgó tényezőnek a vizsgált jelenségre? A matematikus a következő gondolatmenet alapján mond ítéletet: »Először felteszem, hogy a beavatkozás hatástalan volt, azaz a kísérletező által adott két minta ugyanabból a statisztikai sokaságból származik (»null-hipotézis«). Alkalmos módon definiálok valamilyen »távolság«-ot két minta között (ilyen például  $\chi^2$ ,  $t$  stb.). Meghatározom e »távolság« valószínűség-eloszlását arra az esetre, ha két minta egymástól függetlenül ugyanabból a statisztikai sokaságból származik. Ennek birtokában kiszámítom annak valószínűségét, hogy egyazon sokaságból egymástól függetlenül kivett két minta így definiált »távolság«-a nagyobb vagy egyenlő legyen a két megadott mintán ténylegesen észlelt »távolság«-nál. Ha ez a valószínűség a probléma természetétől függő *szintnél* kisebb, úgy a két minta által reprezentált sokaságok közti eltérést *szignifikánsnak* ítélem.

A kísérletező azonban gyakran félreérti a »szignifikáns eltérés«-t kimondó véleményt. A »szignifikáns eltérés« ítéletéből, azaz a null-hipotézis elvetéséből

azt a következtetést vonja le, hogy a két sokaság jelentősen eltér egymástól, azaz »távolság«-uk meghaladja a gyakorlati vagy elméleti szempontból már lényegesnek tekinthető minimális fokot. Azonban a matematikai szignifikancia — azaz a null-hipotézis elvetése — csak azt jelenti, hogy az adott szigorúsági szint mellett, adott határozottsággal állíthatjuk, hogy a két sokaság nem azonos.

Feladatunk tehát null-hipotézis helyett olyan statisztikai hipotézis alkalmazása, mely jobban illeszkedik a gyakorlat által felvetett kérdésekhez, melynek segítségével meg tudjuk ítélni, hogy az észlelt eltérés *szignifikánsan* meghaladja-e azt a legkisebb eltérést, melyet már szakmailag — esetünkben tehát orvosi-biológiai szempontból is — *jelentősnek* tekintünk. Ezen a téren végzett vizsgálatainkat még nem fejeztük be, de járható utat találtunk a probléma megoldására.

A statisztikai ítélet fenti általános logikai problémája mellett foglalkoznunk kell a matematikai statisztika orvosi-biológiai alkalmazhatóságának, illetve helyes alkalmazásának speciális kérdéseivel is.

A matematikai statisztika alkalmazását orvosi-biológiai kutatásokra megnehezíti elsősorban az, hogy ezen a területen a megfigyelések és kísérletek tervezésében és berendezésében a »*ceteris paribus*« elvét általában nem könnyű, néha szinte lehetetlen megvalósítani. Az élő szervezet működésének bonyolultsága miatt ugyanis a vizsgálni kívánt jelenséget gyakran nem sikerül eléggé kiemelni a többi közül és így szándékunk ellenére, néha egyenesen tudunk nélkül olyan tényezők befolyása is érvényesülhet, melyeknek azonos vagy ellentétes irányú hatása megzavarja: meghamisítja vagy elfedi a kísérlet, megfigyelés eredményeit. Tüdőgyulladás lefolyását például nem csak a vizsgált gyógyszer hatássósága határozza meg, hanem a betegség előrehaladott volta, a beteg általános állapota, kora stb. Ez a körülmény növeli a vizsgált anyag diszperzióját, megnehezíti, határozatlanabbá teszi a statisztikai értékelést.

További nehézséget jelent az a tény, hogy az orvosi-biológiai mérések igen pontatlanok: 10% mérési hiba nem számít szokatlanul nagyoknak és kivételesen jóknak kell elfogadnunk az 1—2%-os hibával dolgozó mérési eljárásokat.

Az említett nehézségeket a kísérletek és megfigyelések helyes tervezése mellett bizonyos mértékben azzal is ellensúlyozhatnánk, hogy a megfigyelések számát növelnénk, nagy-elemszámú mintákkal dolgoznánk. Ez igen előnyös lenne egyéb szempontból is: a klasszikus matematikai statisztikának ugyanis csaknem minden módszere nagy-elemszámú mintavételre van kidolgozva, hiszen a felhasznált tételek rendszerint határeloszlásokra vonatkoznak (például:  $\chi^2$ ). A nyugati irodalomban és tudományos gyakorlatban divatos és kis-elemszámú mintákkal is dolgozó eljárások mind ú. n. *paraméteres módszerek*. Paraméteresnek nevezünk egy statisztikai eljárást akkor, ha az a vizsgált statisztikai sokaság eloszlástípusának előzetes ismeretét feltételezi s csupán az ismeretlen paraméterek becslésével vagy más hasonló, az eloszlástípus segítségével jól megközelíthető kérdéssel foglalkozik. Például a szignifikancia-vizsgálatnál gyakran alkalmazott Student-féle *t*-eljárás (»szignifikáns differencia«) és a biológiai gyakorlatban különösen sűrűn alkalmazott varianciaanalízis (»szórás-elemzés«, »Fisher-féle *z*-próba«) arra az előzetes feltevésre épülnek, hogy a vizsgált sokaságok normális (Gauss-féle) eloszlásúak. Fenti módszerek alkalmazása előtt tehát kellő lelkiismeretességgel meg kell győződ-nünk arról, hogy indokolt-e ez a feltevés, megnyugtató módon alátámasztják-e

elméleti meggondolások és tapasztalatok. Ennek a feltétlenül szükséges *ellenőrzésnek* emulasztása az elmélet és gyakorlat elszakadására, formalizmusra, helytelen következtetésekre vezethet. Éppen ebből a szempontból tartjuk jellemzőnek azt a tényt, hogy a nyugati statisztikai iskola kutatói elhanyagolják az a priori feltevések ellenőrzésének kérdését, a tudományos gyakorlatban pedig a fentemlített paraméteres módszereket az alapfeltevés ellenőrzése nélkül jóformán bármilyen statisztikai sokaság vizsgálatára alkalmazzák. A normális eloszlásra kidolgozott paraméteres módszereket kritikátlanul alkalmazó kutatók legfeljebb arra hivatkoznak, hogy »a centrális határeloszlás-tétel következményeként a valóságban előforduló folytonos eloszlások majdnem mind normális eloszlások«. Az idézett centrális határeloszlás-tétel azonban — szemléletesen kifejezve — csak azt mondja ki, hogy ha valamely mennyiség értékét igen nagyszámú véletlentől függő tényező határozza meg, melyek egymástól függetlenek és egymagában mindegyikük hatása igen kicsiny, hatásaik pedig összeadódnak, úgy e véletlentől függő mennyiség közelítőleg normális eloszlású. De a természetben előforduló esetekben ezek a hatások nem szükségképpen additív természetűek, hiszen például gyógyszerek hatásmechanizmusával kapcsolatban is beszélünk »hatványozódó« hatásokról. Így nem is csodálkozhatunk azon, hogy a biológiai és orvosi kutatás terén sűrűn fordulnak elő nem-normális eloszlású folytonos statisztikai sokaságok. Orvosi-biológiai alkalmazások területén tehát különösen el kell ítélnünk a paraméteres módszerek kritikátlan, formalista használatát. Meg kell jegyeznünk, hogy ezek használata még abban az esetben is nehézségekbe ütközhet, ha ismerjük a vizsgált sokaság eloszlástípusát, de ez nem egyezik a megfelelő paraméteres módszer alkalmazásához szükséges eloszlástípussal. Elég nagy-elemszámú minta esetén természetesen kellő biztonsággal megbecsülhetjük azon konstans paramétereket, melyek segítségével a sokaságot a megfelelő típusba transzformálhatjuk s ezekre az átszámított adatokra már alkalmazhatjuk a paraméteres eljárást; kis-minták esetén azonban erre nincs lehetőség.

Fejtegetéseinkből tehát az következik, hogy az orvosi-biológiai kutatás vonalán fokozott jelentősége van a nagy-elemszámú mintavételnek. Csökkenti a diszperziót, kiegyensúlyozza a mérési hibákat, lehetővé teszi a határeloszlás-tételek és (esetleges transzformációk végrehajtása után) a különben igen értékes paraméteres módszerek alkalmazását. Ugyanakkor, mikor hangsúlyozzuk a nagy-elemszámú mintavétel előnyeit, a kísérletezőnek is igazat kell adnunk abban, hogy az orvosi-biológiai kutatások terén a nagy-elemszámú mintavételnek, azaz nagyszámú megfigyelés végrehajtásának gyakorlati, olykor elvi akadályai lehetnek. Így egyes kísérleteket technikai okokból nem lehet megismételni, ritka betegség megfigyelésekor pedig néha kénytelenek vagyunk 5—6 esetre szorítkozni. Ha a megfigyelés »tárgya« az ember, nem is szabad »kísérleteznünk«; a behatás módját és a kontrollok kiválasztását úgy kell elvégeznünk, hogy mindkét csoportba jutó embereknek a lehető legtöbbet használjunk.

Egyik fő problémánk tehát az, hogy a nagy-elemszámú mintavétel minden előnye ellenére általában mégis kis-elemszámú mintákkal vagyunk kénytelenek dolgozni. Ezen a téren két utat látunk a megoldásra. Az első lehetőség annak vizsgálata, hogy a standard klasszikus eloszlásokra (főleg a normális eloszlásra) kidolgozott paraméteres eljárások milyen *enyhébb feltételek* mellett és milyen *közelítéssel* alkalmazhatók ezeket többé-kevésbé jól megközelítő eloszlású sokaságok vizsgálatára. Megjegyezzük, hogy ezen a téren már történtek kez-

deményező lépések a nyugati irodalomban is; de gyakorlatilag hatás nélkül. A másik út olyan módszerek kidolgozása és alkalmazása, melyek *függetlenek a vizsgált sokaság eloszlástípusától*, legfeljebb egészen enyhe feltételek teljesülését, például folytonos eloszlást követelnek meg a sokaságtól. Az ú. n. rendstatisztika (variációs sorok elmélete) kifejlesztésével ezen a téren főleg a szovjet valószínűségszámítási és matematikai-statisztikai iskolának vannak komoly eredményei. Közleményünk most következő második részében összefoglaljuk a rendstatisztikának az alkalmazások szempontjából legfontosabb tételeit, a harmadik részben pedig kis-elemszámú mintavétellel dolgozó szignifikanciavizsgálatra használható eloszlásmentes módszereket közlünk, melyek a rendstatisztika legújabb szovjet eredményein alapulnak.

2.

*A rendstatisztika néhány alapvető tétele*

Adott  $n$  számú  $x_1, x_2, \dots, x_n$  valós számok közül a legkisebbet jelölje  $R_1(x_1, x_2, \dots, x_n)$  [ $= \min(x_1, x_2, \dots, x_n)$ ], a nagyság szerint másodikat  $R_2(x_1, x_2, \dots, x_n), \dots$ , végül pedig a legnagyobbikat  $R_n(x_1, x_2, \dots, x_n)$  [ $= \max(x_1, x_2, \dots, x_n)$ ]. Ekkor tetszőleges  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változókhoz hozzárendelhetjük ezek »rendezett sorozatát« a következő relációkkal:

$$\begin{aligned}
 \xi_1^* &= R_1(\xi_1, \xi_2, \dots, \xi_n) = \min(\xi_1, \xi_2, \dots, \xi_n) \\
 \xi_2^* &= R_2(\xi_1, \xi_2, \dots, \xi_n) \\
 &\dots\dots\dots \\
 \xi_n^* &= R_n(\xi_1, \xi_2, \dots, \xi_n) = \max(\xi_1, \xi_2, \dots, \xi_n).
 \end{aligned}
 \tag{1}$$

A  $\xi_1^*, \xi_2^*, \dots, \xi_n^*$  valószínűségi változók ú. n. »variációs sor«-t alkotnak. A rendstatisztika ezen variációs sorokkal foglalkozik.

Szorítkozzunk most arra az esetre, mikor a  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változók egyazon folytonos  $F(x)$  eloszlásfüggvénnyel rendelkeznek. Minthogy folytonos eloszlású statisztikai sokaságból vett minta egyes elemei éppen ilyen valószínűségi változók, ezért a továbbiakban a megfelelő matematikai-statisztikai terminológiát használjuk.

A rendstatisztika alábbi tételei két csoportba oszthatók. Az első típusú, ú. n. »egy-mintás« tételekben egy folytonos eloszlású statisztikai sokaságból vett mintának és a sokaság elméleti eloszlásfüggvényének bizonyos »távolság«-ait értelmezzük a minta empirikus eloszlásfüggvényének segítségével és meghatározzuk a nyert valószínűségi változók eloszlásfüggvényeit, illetve ezek határértékét arra az esetre, amikor a minta elemszáma minden határon túl nő. A második-típusú, ú. n. »két-mintás« tételek két minta empirikus eloszlásfüggvényeinek hasonlóan definiált távolságaira vonatkozó eloszlásokat és határeloszlásokat adják meg.

Először is definiáljuk az *empirikus eloszlásfüggvény* fogalmát. Egy  $n$ -elemű minta empirikus eloszlásfüggvénye:

$$S_n(x) = \frac{k_n(x)}{n},
 \tag{2}$$

ahol  $k_n(x)$  a minta azon elemeinek számát jelenti, melyek kisebbek, mint  $x$ . Másszóval, ha a mintavételi változók variációs sora :

$$(3) \quad \xi_1^* \equiv \xi_2^* \equiv \dots \equiv \xi_n^*,$$

akkor

$$(4) \quad S_n(x) = \begin{cases} 0, & \text{ha } x \equiv \xi_1^* \\ \frac{1}{n}, & \text{ha } \xi_1^* < x \equiv \xi_2^* \\ \frac{2}{n}, & \text{ha } \xi_2^* < x \equiv \xi_3^* \\ \dots & \dots \\ \frac{\nu}{n}, & \text{ha } \xi_\nu^* < x \equiv \xi_{\nu+1}^* \\ \dots & \dots \\ \frac{n-1}{n}, & \text{ha } \xi_{n-1}^* < x \equiv \xi_n^* \\ 1, & \text{ha } \xi_n^* < x. \end{cases}$$

A minta empirikus eloszlásfüggvénye tehát a minta  $x$ -nél kisebb elemeinek viszonylagos gyakori ága, mint  $x$  függvénye. Grafikonja egy lépcső-vonal, mely a minta két szomszédos eleme között az  $x$ -tengellyel párhuzamos egyenes szakasz, a minta egyes elemeinek megfelelő helyen pedig szakadáspontja van, ahol a görbe » $\frac{1}{n}$ -nel feljebb folytatódik«. Az ugráspontokban az alsó egyenes szakasz jobb végpontja a görbepont. A mintavétel során természetesen a minta elemei valószínűségi változók, a minta empirikus eloszlásfüggvényének értéke tehát minden rögzített  $x$  helyen valószínűségi változó.

Rátérünk a tételek ismertetésére.

### Egymintás tételek

Egy folytonos  $F(x)$  eloszlásfüggvényű statisztikai sokaságból  $n$ -elemű mintát veszünk, ennek empirikus eloszlásfüggvényét  $S_n(x)$ -szel jelöljük. Definíáljuk a  $D_n^+$ ,  $D_n^-$ ,  $D_n$ ,  $Q'_n(a)$ , és  $Q''_n(a,b)$  valószínűségi változókat a következőképpen :

$$(5) \quad D_n^+ = \sup_{-\infty < x < +\infty} (S_n(x) - F(x))$$

$$(6) \quad D_n^- = \sup_{-\infty < x < +\infty} (F(x) - S_n(x))$$

$$(7) \quad D_n = \sup_{-\infty < x < +\infty} |S_n(x) - F(x)| = \max(D_n^+, D_n^-)$$

$$(8) \quad Q'_n(a) = \sup_{a \leq F(x)} \left( \frac{S_n(x)}{F(x)} - 1 \right)$$

$$(9) \quad Q''_n(a,b) = \sup_{a \leq F(x) \leq b} \left( \frac{S_n(x)}{F(x)} - 1 \right),$$

ahol  $0 < a < b < 1$  tetszőleges rögzített értékek. [A (8) és (9) definíciókkal

kapcsolatban megjegyezzük, hogy a  $\sup_{-\infty < x < +\infty} \left( \frac{S_n(x)}{P(x)} - 1 \right)$  változó pozitív valószínűséggel felvehet egy tetszőleges nagy fix értéknél nagyobb értékeket.]

A) *GLIVENKO tétele (a statisztika alaptétele)* [8]

$$(10) \quad \mathbf{V} \left\{ \lim_{n \rightarrow \infty} D_n = 0 \right\} = 1.$$

Ez a tétel tehát azt mondja ki, hogy a minta empirikus eloszlásfüggvénye 1 valószínűséggel egyenletesen konvergál az elméleti eloszlásfüggvényhez. Mintavétel útján tehát kimerítő képet szerezhethünk a vizsgált statisztikai sokaságról: ebben az értelemben valóban a statisztika alaptételének nevezhetjük.

B) *KOLMOGOROV tételei* [16]: tetszőleges valós  $z$ -re:

$$(11) \quad \lim_{n \rightarrow \infty} \mathbf{V} \{ \sqrt{n} D_n^+ < z \} = \lim_{n \rightarrow \infty} \mathbf{V} \{ \sqrt{n} D_n^- < z \} = \\ = K^*(z) = \begin{cases} 0, & \text{ha } z \leq 0 \\ 1 - e^{-2z^2}, & \text{ha } z > 0. \end{cases}$$

$$(12) \quad \lim_{n \rightarrow \infty} \mathbf{V} \{ \sqrt{n} D_n < z \} = K(z) = \begin{cases} 0, & \text{ha } z \leq 0 \\ \sum_{\nu=-\infty}^{\infty} (-1)^\nu e^{-2\nu^2 z^2}, & \text{ha } z > 0. \end{cases}$$

C) *SZMIRNOV tétele* [29.]: tetszőleges valós  $x, y$  esetén:

$$(13) \quad \lim_{n \rightarrow \infty} \mathbf{V} \{ \sqrt{n} D_n^- < x; \sqrt{n} D_n^+ < y \} = S(x, y),$$

ahol  $S(x, y)$  az alábbi függvény:

$$(14) \quad S(x, y) = \begin{cases} 0, & \text{ha } \min(x, y) \leq 0 \\ \sum_{\nu=-\infty}^{\infty} e^{-2\nu^2(x+y)^2} - \sum_{\nu=1}^{\infty} (e^{-2(\nu x + \nu - 1)y^2} + e^{-2(\nu - 1)x + \nu y^2}), & \text{ha } \min(x, y) > 0 \end{cases}$$

D) *RÉNYI tételei\**:

$$(15) \quad \lim_{n \rightarrow \infty} \mathbf{V} \{ \sqrt{n} Q_n'(a) < z \} = H_1(z; a) = \begin{cases} 0, & \text{ha } z \leq 0 \\ \sqrt{\frac{2}{\pi}} \int_0^z \sqrt{\frac{a}{1-a}} e^{-\frac{t^2}{2}} dt, & \text{ha } z > 0 \end{cases}$$

\* A szerző ezirányú dolgozatai sajtó alatt vannak.

$$(16) \quad \lim_{n \rightarrow \infty} \mathbf{V} \left\{ \sqrt{n} Q_n''(a, b) < z \right\} = H_0(z; a, b) = \\ = \frac{1}{\pi} \sqrt{\frac{b}{1-b}} \int_{-\infty}^z e^{-\frac{b}{1-b} \cdot x^2} \left( \frac{(z-x) \sqrt{\frac{ab}{b-a}}}{\int_0^{\frac{(z-x) \sqrt{\frac{ab}{b-a}}}{2}} e^{-\frac{t^2}{2}} dt} \right) dx; \quad (-\infty < z < +\infty).$$

### Kétmintás tételek

Egy folytonos eloszlásfüggvényű statisztikai sokaságból egymástól függetlenül két mintát veszünk, az egyik  $m$ -elemű, a másik  $n$ -elemű minta. Empirikus eloszlásfüggvényeiket  $r$ -ndre  $S_m(x)$ -szel ill.  $T_n(x)$ -szel jelöljük. Az (5), (6) és (7) egyenletek analógiájára definiáljuk a  $D_{m,n}^+$ ,  $D_{m,n}^-$  és  $D_{m,n}$  valószínűségi változókat:

$$(17) \quad D_{m,n}^+ = \sup_{-\infty < x < +\infty} (S_m(x) - T_n(x))$$

$$(18) \quad D_{m,n}^- = \sup_{-\infty < x < +\infty} (T_n(x) - S_m(x))$$

$$(19) \quad D_{m,n} = \sup_{-\infty < x < +\infty} |S_m(x) - T_n(x)| = \max(D_{m,n}^+, D_{m,n}^-).$$

Definiáljuk továbbá a  $C_{m,n}$  valószínűségi változót a következőképpen: jelentse a  $C_{m,n}$  az első,  $S_m(x)$  empirikus eloszlásfüggvényű minta elemei közül azon  $\xi_k^*$  elemek számát, melyekre fennáll:

$$(20) \quad S_m(\xi_k^*) \cong T_n(\xi_k^*).$$

E) *SZMIRNOV tételei* [29]: tetszőleges valós  $x, y$  és  $z$  számok esetén  $m$  és  $n$  tartsanak a végtelenhez úgy, hogy közben  $\frac{m}{n} = \text{konstans}$  legyen; ekkor

$$(21) \quad \lim \mathbf{V} \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n}^+ < z \right\} = \lim \mathbf{V} \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n}^- < z \right\} = K^*(z)$$

$$(22) \quad \lim \mathbf{V} \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n} < z \right\} = K(z)$$

$$(23) \quad \lim \mathbf{V} \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n}^- < x; \sqrt{\frac{mn}{m+n}} D_{m,n}^+ < y \right\} = S(x, y),$$

ahol  $K^*(z)$ ,  $K(z)$  és  $S(x, y)$  a (11), (12) és (13) relációkban szereplő függvények.

A következő tételek ugyanezen valószínűségi változók határeloszlása helyett azok pontos eloszlását adják meg az  $m = n$  esetben.



F) GNYEGYENKO és KOROLJUK tételei [10]: tetszőleges valós  $z$  esetén a  $\gamma = [nz]$  jelöléssel

$$(24) \quad V\{D_{n,n}^+ < z\} - V\{D_{n,n}^- < z\} = \begin{cases} 0, & \text{ha } z \leq 0 \\ 1 - \frac{\binom{2n}{n-\gamma}}{\binom{2n}{n}}, & \text{ha } z > 0. \end{cases}$$

$$(25) \quad V\{D_{n,n} < z\} = \begin{cases} 0, & \text{ha } z \leq 0 \\ \frac{1}{\binom{2n}{n}} \sum_{\nu=-[n/\gamma]}^{[n/\gamma]} (-1)^\nu \binom{2n}{n-\nu\gamma}, & \text{ha } z > 0. \end{cases}$$

Itt és a továbbiak során is a következő jelöléseket használjuk:

1. tetszőleges valós  $t$  szám esetén  $[t]$  jelentse  $t$  »egész részét«, azaz azon legnagyobb egész számot, mely  $t$ -nél még nem nagyobb;

2. tetszőleges valós  $t$  és tetszőleges egész  $k$  esetén  $\binom{t}{k}$  (másik szokásos jelölése:  $C_n^k$ ) jelentse a következőt:

$$(26) \quad \binom{t}{k} = \begin{cases} 0, & \text{ha } k < 0 \\ 1, & \text{ha } k = 0 \\ \frac{t(t-1)\dots(t-k+1)}{1 \cdot 2 \cdot 3 \dots k}, & \text{ha } 0 < k \leq t. \\ 0, & \text{ha } k > t \end{cases}$$

G) GNYEGYENKO és RVACSEVA tételei [11]: tetszőleges valós  $x$  és  $y$  esetén

$$(27) \quad V\{D_{n,n}^- < x; D_{n,n}^+ < y\} = \Psi_n(x,y),$$

ahol  $\alpha = [nx]$ ,  $\beta = [ny]$  és

$$(28) \quad \Psi_n(x,y) = \begin{cases} 0, & \text{ha } \max(x,y) \leq \frac{1}{n} \\ 1, & \text{ha } \min(x,y) > 1 - \frac{1}{n} \\ \frac{1}{\binom{2n}{n}} \left| \sum_{\nu=-[\frac{n}{\alpha \cdot \beta}]}^{[\frac{n}{\alpha \cdot \beta}]} \binom{2n}{n-\nu(\alpha+\beta)} - \sum_{\nu=1}^{[\frac{n-\alpha}{\alpha \cdot \beta}]} \binom{2n}{n+\alpha-\nu(\alpha+\beta)} - \sum_{\nu=1}^{[\frac{n-\beta}{\alpha \cdot \beta}]} \binom{2n}{n+\beta-\nu(\alpha+\beta)} \right| \end{cases}$$

egyébként.

H) GNYEGYENKO és MIHALEVICS tétele [13], [14]:

$$(29) \quad V\{C_{n,n} - k\} = \frac{1}{n+1}; \quad k = 0, 1, 2, \dots, n:$$

tehát  $C_{n,n}$  összes lehetséges értékei egyenlő valószínűségűek. Mihalevics később [23] cikkében bebizonyítja ugyanezt az összefüggést  $C_{n,n}$  helyett általában a  $C_{n,np}$  változóra is, ahol  $p$  egész szám.

A bizonyításokkal itt nem foglalkozhatunk. Természetesen nem célunk ennek a közleménynek kereteiben a rendstatisztika összes tételeit ismertetni még felsorolás formájában sem. Megemlítjük, hogy a variációs sorok elméletének kialakítását *V. I. Glivenko* [8], *A. N. Kolmogorov* [16], *N. V. Szmirnov* [27], [28], [29], [30], *H. Cramér* [3] és *R. Mises* [24] alapvető munkáitól számíthatjuk. Azóta ezen a téren a szovjet kutatók közül *G. M. Manyia* [18], *I. D. Kvit* [17], *M. U. Pajvin* [25], *B. V. Gnyegyenko* [9], [10], [11], [12], [13], [14], *V. Sz. Koroljuk* [10], *E. L. Rv cseva* [11] és *V. Sz. Mihalevics* [13], [14], [23] érték el lényeges eredményeket. Hazánkban *Rényi Alfréd* járult hozzá a variációs sorok modern elméletének kiépítéséhez\*, amennyiben azt bizonyos mértékben a független valószínűségi változók összegeire vonatkozó elméletre vezette vissza, továbbá meghatározta a (8) és (9) változók és még más hasonló változók határeloszlásait. *H. Cramér* és *R. Mises* idézett munkái után a nyugati statisztikai irodalomban hosszú ideig csak *A. Wald* és *J. Wolfowitz* bonyolult formalizmussal dolgozó cikke képviselte az idevágó irodalmat. *W. Feller* 1948-ban megjelent és *Kolmogorov* tételeinek bizonyítását egyszerűsítő munkája után *J. L. Doob* [6] ezen tételekre vonatkozó heurisztikus bizonyítási eljárása, melyet *M. D. Donsker* [5] tett szabattossá, jelentettek lényegesen újat, majd *K. L. Chung* [2], *M. Kac* [15], *Anderson* és *Darling* [1] munkái fejlesztették tovább a variációs sorok elméletét. *F. J. Massey* [19], [20], [21], [22] négy cikkben foglalkozik *Kolmogorov* és *Szmirnov* említett tételeivel és azok alkalmazásához szükséges egyes numerikus táblázatok kidolgozásával járul hozzá az elméleti eredményekhez.

A mi szempontunkból lényeges, hogy *Gnyegyenko*, *Koroljuk*, *Rvaceva* és *Mihalevics* (24)–(29) alatt ismertetett tételei lehetővé teszik, hogy a variációs sorok elméletének alkalmazását kis-mintákkal dolgozó szignifikancia-vizsgálatra is kiterjesszük. Ezzel foglalkozunk közleményünk 3. részében.

### 3.

*Eloszlásmentes módszerek alkalmazása szignifikancia-vizsgálatra kis-elemszámú mintavételnél*

A rendstatisztikák előbb ismertetett tételei közül (10)–(23) határeloszlástételeket nagy-elemszámú mintavétel esetén alkalmazhatjuk. Erre vonatkozólag utalunk *Gnyegyenko* [9] könyvére, továbbá *Szentmártony Tibor* [26] munkájára, melyek foglalkoznak e tételek statisztikai alkalmazásával és *Szmirnov*nak az »Annals of Mathematical Statistics« c. folyóiratban is megjelent [30] közleménye nyomán megadják a  $K(z)$  függvény táblázatát is.

\* Ezirányú dolgozatai sajtó alatt vannak.

Mi csak az orvosi-biológiai alkalmazásokban leggyakrabban előforduló kis-elemszámú mintavétellel dolgozó szignifikancia-vizsgálatokra használjuk fel a rendsztatistika egyes tételeit, mégpedig a (25), (27) és (29) alatt szereplőket.

Mint már említettük, két-mintás szignifikancia-vizsgálatra szolgáló eljárást úgy nyerhetünk, hogy meghatározzuk egy olyan változó valószínűségi eloszlását, mely alkalmas módon jellemzi két, egyazon sokaságból vett, minta »egymástól való eltérését«, »távolságát«. Nyilvánvaló, hogy ilyen »távolságot« sokféleképpen értelmezhetünk, s ezek eloszlása általában függhet a szóbanforgó statisztikai sokaság eloszlástípusától, sőt a típuson belül a sokaság eloszlásfüggvényében szereplő paraméterek értékétől is. Észszerű tehát ezt a »távolságot« úgy definiálni, hogy annak eloszlása, de legalábbis határeloszlása minél kevésbé függjön a sokaság eloszlástípusától és jellemzőitől. Így a Pearson-féle  $\chi^2$ -változót ilyen »távolság«-ként használva, az eloszlás véges esetben polinomiális eloszlás, melynek paraméterei a sokaság eloszlásától, sőt az adatok csoportosításától is függhenek, határértékben azonban ez a függés eltűnik. Minthogy a nagy-elemszámú mintákra gyakorlatilag elegendő közelítéssel a határeloszlást alkalmazhatjuk, ilyen esetekben a szóbanforgó »távolság« alkalmazásával nagy-elemszámú mintavételre eloszlásmentes módszert nyerünk. A már szóvátett okoknál fogva azonban a mi alkalmazási területünkön gyakran kis-elemszámú mintavételre vagyunk kénytelenek szorítkozni s így nem elégedhetünk meg a határértékekre szóló függetlenséggel. Közleményünk első részében kifejtett okok mind amellettszólnak, hogy igen előnyös lenne olyan »távolság«-definíciók megalkotása, melyek véges esetben sem függhenek a sokaság eloszlásától.

Könnyű belátni, hogy abban az esetben, ha a sokaság eloszlásfüggvénye folytonos, a (17) — (18) — (19) egyenletekkel definiált valószínűségi változók mind rendelkeznek ezzel a tulajdonsággal. Ezen változók mindegyikének értéke adott minták esetén csak attól függ, hogy a két minta elemeinek egyesítésével nyert és nagyság szerint rendezett sorozat tagjai milyen sorrend szerint tartoznak az egyik, illetve másik mintához. De nyilvánvaló, hogy azonos — és folytonos — eloszlású változók esetén minden ilyen sorrend egyenlő valószínűségű s így e definíciónak megfelelő »távolságok« eloszlása nem függ  $F(x)$ -től. Meg kell jegyeznünk, hogy éppen ez az észrevétel könnyíti meg a (10) — (29) tételek bizonyítását.

A »távolság« definíciója során két esetet kell megkülönböztetnünk. Az egyik esetben a kísérlet végrehajtása előtt a priori feltevésünk van arra vonatkozólag, hogy a behatás milyen irányban fogja befolyásolni a vizsgált jelenséget, tehát *azt kérdezzük, hogy a behatás szignifikánsan befolyásolta-e az adott irányban a vizsgált egyedeket*. Ilyen hipotézisnek megfelelő »távolság«-definíció tehát egyirányú, nem-szimmetrikus a két mintára vonatkozólag. A másik esetben ilyen előzetes feltevésünk nincs, s így csupán *azt vizsgáljuk, okozott-e valamilyen irányú szignifikáns eltérést a behatás*. Ilyenkor a »távolság«-nak szimmetrikusnak kell lennie. Nyilvánvaló, hogy ha az első esetben a kísérleti eredmények valóban igazolják a hatás irányára vonatkozó előzetes feltevésünket, egyirányú módszer használatával sokkal határozottabb statisztikai ítéletet mondhatunk, mint szimmetrikus módszer segítségével. Természetesen hangsúlyoznunk kell, hogy egyirányú módszerek használata csak akkor jogosult, ha valóban előzetes feltevésre alapul, utólagos használatuk meghamisítja a statisztikai ítéletet.

Gyakorlatilag az esetek többségében a behatástól azt várjuk, hogy a statisztikai sokaságot valamilyen irányban »eltolja«, azaz csak a sokaság várható értékét változtatja meg, az eloszlás egyéb jellemzőit nem befolyásolja lényegesen. Így azon »távolság«-oknak van különösen lényeges szerepük, melyek bizonyos mértékben jellemzik, mérik az ilyen típusú tranzlációkat.

a) »*C*«-kritérium. Foglalkozzunk mindenképp a (29) alatti tétel alkalmazásával, mely *Gnyegyenkotól* és *Mihalevicstől* származik. Azonos elemszámú minták esetén előzetes tájékozódásra használhatjuk abban az esetben, ha a priori feltevésünknek megfelelően a behatás adott irányban »eltolta« a vizsgált sokaságot és a »lemaradó« minta empirikus eloszlásfüggvénye teljesen a másik felett fekszik, legfeljebb közös pontja van vele. Null-hipotézis mellett ugyanis annak valószínűsége, hogy ennek az esetnek megfelelően  $C_{n,n} = n$  legyen, (29) szerint

$$(30) \quad V \{ C_{n,n} = n \} = \frac{1}{n+1}.$$

5%-szint mellett ilyenkor tehát már »ránézésre« is szignifikánsnak ítéelhetjük az eltérést, ha a minták legalább 20–20 elemszámúak.

b) »Abszolút-*D*«-kritérium. A (25) alatti tételt, mely *Gnyegyenkotól* és *Koroljuktól* származik, »szimmetrikus távolság« definiálására használhatjuk. E végből írjuk a tételt a következő alakba :

$$(31) \quad v_n(c) = V \left\{ D_{n,n} \cong \frac{c}{n} \right\} = \frac{2}{\binom{2n}{n}} \sum_{\nu=1}^{\lfloor n/c \rfloor} (-1)^{\nu-1} \binom{2n}{n-\nu c} ; c = 1, 2, \dots, n.$$

$D_{n,n}$  is eloszlásmentes »távolságot« értelmmez. Null-hipotézis mellett annak valószínűségét, hogy  $D_{n,n} \cong \frac{c}{n}$ , a (31) képletből nyerjük. Ezen  $v_n(c)$  valószínűségek értékeit az I. táblázatba foglaltuk. Ha  $D_{n,n}$  olyan nagy, hogy a megfelelő  $v_n(c)$  valószínűségi érték a használt valószínűségi szintnél (orvosi-biológiai vonatkozásoknál általában 5%) kisebb, úgy a két minta egymástól való eltérését szignifikánsnak ítéelhetjük. Az »abszolút-*D*«-kritériumot abban az esetben használhatjuk, ha a behatás irányára vonatkozólag nincs előzetes hipotézisünk.

c) »*Kétoldali-D*«-kritérium. Előzetes egyirányú hipotézis esetére az előbbinél jóval hatékonyabb eljárást nyerünk a *Gnyegyenkotól* és *Rvaacsevától* származó (27) tétel alkalmazásával, amely (azonos elemszámú minták esetén) a  $D_{n,n}^-$  és  $D_{n,n}^+$  változók együttes eloszlását adja meg. Mint említettük, bennünket elsősorban a két minta egymástól való »eltoltságának«, »tranzlációjának« jellemzése érdekel. Világos, hogy az elsőnek választott minta annál inkább »lemarad« a másik mögött, annál inkább mondhatjuk róla, hogy »balra toldott« az  $x$ -tengelyen a másodiktól, minél nagyobb  $D_{n,n}^+$  és ugyanakkor minél kisebb  $D_{n,n}^-$ ; tehát minél nagyobb az első minta empirikus eloszlásfüggvényének a második feletti »maximális kiugrása« és ugyanakkor minél kisebb a

második mintának az első feletti »maximális kiugrása«. Írjuk tehát a (27) alatti tételt a következő alakba :

$$\begin{aligned}
 w_n(a,b) &= V \left\{ D_{n,n}^- \cong \frac{a}{n} ; D_{n,n}^+ \cong \frac{b}{n} \right\} = \\
 (32) \quad &= \frac{1}{\binom{2n}{n}} \left| \sum_{\nu=1,2,\dots} \binom{2n}{n+a+1-\nu(a+b+1)} + \right. \\
 &\quad \left. + \sum_{\nu=1,2,\dots} \binom{2n}{n+b-\nu(a+b+1)} - 2 \sum_{\nu=1,2,\dots} \binom{2n}{n-\nu(a+b+1)} \right|.
 \end{aligned}$$

Itt  $a$  és  $b$  nem-negatív egész számok lehetnek és természetesen  $0 \cong a + b \cong n$ . Null-hipotézis mellett  $w_n(a, b)$  a valószínűsége annak, hogy megfelelő sorrendben véve a mintákat (tehát az »elmaradónak« várt mintát elsőnek, a másikat pedig másodiknak választva),  $D_{n,n}^- \cong \frac{a}{n}$  és  $D_{n,n}^+ \cong \frac{b}{n}$  legyen. Adott esetben a II. táblázatból kiolvasott  $w_n(a, b)$  értéket a szinthez hasonlítva ítéltjük meg, szignifikánsan eltér-e egymástól a várt irányban a két minta?

Az elmondottnak orvosi vonatkozású példákkal való szemléltetése előtt igyekszünk gyakorlati szempontból is világosan meghatározni, milyen esetben, milyen feltételek betartása mellett alkalmazhatjuk két minta, megfigyelési adatsor közti eltérés szignifikanciájának megítélésére a fenti módszereket.

Három feltételnek kell teljesülnie. 1. A vizsgált statisztikai sokaságnak *folytonos eloszlásúnak* kell lennie. A kimondottan diszkrét eloszlású sokaságtól eltekintve (pl. a baktérium-szám, megbetegedettek száma stb.) az orvosi-biológiai kutatásban szereplő sokaságok majdnem mind ilyenek. Tudjuk azonban, hogy az egyes adatokra vonatkozó mérési eredmények többé-kevésbé hibás értékeket szolgáltatnak, ezért ezeket megfelelő módon kerekíteni szokták. Ilyen kerekített értékekből álló adatsorokat, mintákat azonban — túlságos kerekítés után — már nem tekinthetjük folytonos eloszlású sokaságot reprezentáló mintáknak, hiszen a kerekítés miatt az egyes adatok csak ugrás-szerű, diszkrét értéket vehetnek fel. Különösen fontos ezt hangsúlyoznunk azért, mert az orvosi-biológiai gyakorlatban előforduló mérések — mint azt már szóvátettük — általában igen nagy hibával dolgoznak, az egyes adatok tehát csak erős kerekítés után értékelhetők egyenként. Ezt a tényezőt azonban a módszerek alkalmazásakor aránylag könnyen kiküszöbölhetjük azáltal, ha tudományos kísérletekben a *mérőműszer megengedte legnagyobb pontossággal olvassuk le az eredményeket* s ezeket kerekítés nélkül használjuk fel statisztikai értékelésre. Igen hasznos, ha egyetlen adatra vonatkozólag több mérést végzünk s a pontosan leolvasott értékek ugyanannyi jegyig számított számtani középértékével dolgozunk. Természetesen hangsúlyozzuk, hogy a kerekítés nélkül nyert adatok utolsó néhány jegye az egyénekenkénti jellemzés szempontjából értéktelen lehet, a mérési hibaokozta torzítás azonban a kísérlet egészének statisztikai értékelésében kiegyenlítődik s hatása legfeljebb a statisztikai ítélet határozottságának rovására megy. A folytonossági feltétel teljesülésének azt a gyakorlatilag elegendő (bár nem szükséges) kritériumát adhatjuk meg, hogy a két vizsgált minta »egyesítésével« nyert adatsorban *minden egyes érték csak egyszer forduljon elő*, tehát nem fordulhat elő sem az, hogy az egyik mintában

egy érték többször szerepel, sem pedig az, hogy a két mintának van »közös» eleme. Megjegyezzük, hogy G. M. Manyia [18] eredményei feljogosítanak bennünket arra, hogy úgy véljük, ezeket a módszereket idővel csoportosított adatokra is alkalmazhatjuk. — 2. Az alkalmazhatóság második feltétele az, hogy a *mintáknak*, melyekre a szignifikancia-vizsgálatot alkalmazzuk, *egymástól függetleneknek* kell lenniök. Ki kell tehát zárnunk e módszerek alkalmazásánál az ú. n. önkontroll esetét, melynél ugyanazokon az egyéneken végezzük el a behatást is, akiken előzőleg a kontrolladatokat leolvastuk. Ezeket a módszereket tehát csak akkor használhatjuk, ha a »kontroll« és »kezelt« egyedek egymástól független csoportot képeznek. — 3. A harmadik megszorítás arra vonatkozik, hogy a megfelelő (25), (27) és (29) alatti tételek egyelőre csak *azonos elemszámú minták* esetén alkalmazhatók. Minthogy elméletileg is ez a legoptimálisabb elosztása a kísérleti egyedeknek a két csoport között, a kísérleteket már lehetőleg eszerint tervezzük meg. Ha a két minta már adott és különböző elemszámúak, úgy az egyik mintából megfelelő számú elem kisorsolásával már egyenlő elemszámú mintákat kaphatunk. Ebben az esetben természetesen információt veszünk, amit a kísérletek alkalmasabb tervezésével esetleg elkerülhettünk volna. Megjegyezzük, hogy a rendstatisztika további kifejlődésével módot nyerhetünk majd egyenlőtlen elemszámú minták ilyen módszerekkel történő értékelésére is.

A példákra rátérve megjegyezzük, hogy tapasztalatunk szerint az ismertett három új módszert konkrétan grafikus úton alkalmazhatjuk a legegyszerűbben. A két minta empirikus eloszlásfüggvényeit kockás vagy milliméteres papírra rajzoljuk fel és ezután könnyen leolvashatjuk a megfelelő változók értékeit. Érdemes megjegyeznünk, hogy az  $x$ -tengelyen történő monoton (sorrendtartó) transzformáció esetén — mint már említettük — ezek az értékek változatlanok maradnak és így nem szükséges a valódi léptéket felmérnünk az  $x$ -tengelyre: ehelyett alábbi egyszerűbb módon járhatunk el: A két mintában szereplő összes adatokat egyetlen adatsorrá egyesítjük és nagyság szerint növekedőleg elrendezzük. Ezután az  $x$ -tengelyen ekvidisztáns távolságokban ezen egyesített adatsor minden egyes adata helyett ugyanezen sorrendben egy-egy pontot jelölünk ki, de például kék, illetve piros színnel, aszerint, hogy az illető adat az első vagy a második mintához tartozik. A következő lépés az, hogy ezen kék, illetve piros pontokból álló »transzformált« minták empirikus eloszlásfüggvényeit rajzoljuk meg a megfelelő színekkel. Természetesen kockás papírnál az  $x$ -tengely egységét alkalmasan  $n$  kockával egyenlőnek válasszuk (ha  $n$ - $n$ -elemszámú mintákról van szó), mivel ekkor éppen 1 kockával kell feljebb ugorjunk az egyes minta-elemeknél, tehát a lépcső-vonalak »fokai« 1 kockányi magasságúak. A  $D_{n,n}^-$ ,  $D_{n,n}$  ill.  $D_{n,n}$  differenciák értékeit ezután az ábrából könnyen leolvashatjuk.

#### 4.

#### *Példák*

Az 1. példa egy konkrét minta empirikus eloszlásfüggvényének grafikonját adja meg, de itt még az  $x$ -tengelyen a szokásos skálát alkalmazzuk. — 20 emberen lemérve a glomerulus filtratum mennyiségét, például a következő eredményeket kapjuk (itt és a továbbiakban is mindig nagyság szerint növekedőleg elrendezve adjuk meg a mintákat):

*Glomerulus filtratum mennyisége ml/min egységben:*

81,8	84,1	86,0	88,2
89,7	91,5	92,3	93,9
94,6	96,2	97,3	98,7
100,5	101,0	103,5	104,4
105,3	107,6	109,2	111,3

Ezen adatok empirikus eloszlásfüggvényét az 1. ábrán láthatjuk. Az  $x$ -tengelyen ebben a példában a szokásos léptéket használjuk.

2. *példa.* Két különböző községben levő általános iskolákban megmérték 24—24 12-éves fiú mellkaskörfogatát. Az  $X$ -községben lévő iskola igen jó sporteredményeket ért el, a másik,  $Y$ -községbeli lemaradt a versenyeken. A kérdés az, szignifikánsan nagyobb-e a jobban sportoló  $X$ -községbeli tanulók mellkaskörfogata az  $Y$ -községbelieknél? A mérések eredményei például a következők voltak:

*Mellkaskörfogat cm-egységben belégzéskor:*

*$X$ -községben:*

66,7	67,4	68,5	69,2
70,3	71,0	73,1	74,9
76,6	77,5	78,4	78,9
79,0	80,6	81,0	81,7
82,1	82,5	83,7	84,0
84,6	85,4	85,9	86,3

*$Y$ -községben:*

63,5	64,3	65,5	66,3
67,8	68,2	69,4	70,5
71,4	71,8	72,2	72,3
73,7	73,9	74,3	75,0
75,3	76,4	77,1	78,2
78,6	80,3	82,8	83,0

A 2. ábráról először is látjuk, hogy az  $Y$ -minta empirikus eloszlásfüggvénye konzekvensen a másik felett halad s így a » $C$ «-kritériumot alkalmazva rögtön láthatjuk, hogy az adott irányú eltérés 5%-os szint mellett szignifikáns!

$$V \{C_{24,24} = 24\} = \frac{1}{25} = 4\% < 5\%.$$

A »kétoldali  $D$ «-kritérium szerint természetesen még inkább szignifikánsnak találjuk az eltérést: a 2. ábrából

$$D_{24,24}^+ = \frac{10}{24} \text{ és } D_{24,24}^- = 0, \text{ tehát}$$

(33)

$$w_{24}(0,10) = 0,52\% \ll 5\%.$$

3. *példa.* Azt vizsgáljuk, befolyásolja-e egy bizonyos diéta a testsúly növekedését. 20—20 egyforma súlyú és korú, azonos nemű és beltenyészetből származó patkányon kísérletet végzünk. Az első 20 állat a diétát kapja (»kezeltek«), a másik 20-at pedig a szokásos módon tápláljuk (»kontrollok«). Megfelelő idő múlva lemérjük mind a 40 patkány súlygyarapodását és például a következő eredményeket kapjuk :

*Súlygyarapodás gr.-egységben 2 hét múlva:*

1. »kezeltek«:

18,2	20,1	22,0	24,1
25,9	26,8	28,9	31,1
31,9	32,8	37,7	41,0
44,8	47,1	50,0	54,8
60,2	63,1	68,4	71,3

2. »kontrollok«:

30,2	34,3	35,1	36,2
38,6	39,4	40,0	42,3
42,9	44,2	45,9	49,1
51,1	52,0	53,1	53,9
56,2	58,4	62,0	67,3

A grafikus módszert alkalmazva láthatjuk, hogy a minták adott sorszámozása mellett (lásd a 3. ábrát)

$$D_{20;20}^+ = \frac{9}{20}, \quad D_{20;20}^- = \frac{2}{20} \quad \text{és} \quad D_{20;20} = \max(D_{20;20}^+, D_{20;20}^-) = \frac{9}{20}.$$

Az »abszolút-D«-kritérium alkalmazásánál az I. táblázatból kiolvashatjuk, hogy

$$v_{20}(9) = 3,35\% < 5\%.$$

és így 5%-os szint mellett az eltérést szignifikánsnak ítéelhetjük.

4. *példa.* Vizsgáljuk két diureticum hatását. Előzetes állatkísérletek alapján az a feltevésünk, hogy az »A«-készítmény diureticus hatása nagyobb a »B«-készítményénél. Feltevésünk alátámasztására 16 betegen lemérjük az »A« készítmény, másik 16 betegen pedig a »B«-készítmény hatására fellépő diuresist. Kérdés, szignifikánsan nagyobb-e az »A«-készítmény diureticus hatása a »B«-készítményénél?

*Napi diuresis 100 ml.-egységben:*

»A«-készítmény:

8,3	11,4	20,1	26,2
27,2	36,0	36,6	37,3
38,4	41,9	42,7	44,4
45,0	47,5	48,6	51,1



»B«-készítmény:

8,7	13,8	18,9	21,7
23,9	25,5	26,8	28,9
30,0	31,2	32,3	33,9
35,5	40,6	43,7	46,6

Mint hogy itt egyirányú, előzetes hipotézisen alapuló vizsgálatról van szó, felhasználhatjuk a »kétoldali-D«-kritériumot. A 4. ábráról leolvashatjuk, hogy

$$D_{16;16}^+ = \frac{8}{16} \text{ és } D_{16;16}^- = \frac{1}{16}, \text{ tehát a II. táblázatból}$$

$$w_{16}(1,8) = 1,45\% < 5\%.$$

Tehát mondhatjuk, hogy az »A«-készítmény szignifikánsan hatásosabb a »B«-készítménynél 5%-os szint mellett.

Úgy véljük, hogy a bemutatott példák jól szemléltetik azt, hogy tágabb alkalmazhatóságuk mellett ezek a módszerek technikai egyszerűségükkel is kitűnnek.

IRODALOM

- [1.] T. W. Anderson és D. A. Darling: *Annals of Math. Statistics* 23. (1952) 193–212. old.
- [2.] K. L. Chung: *Transactions of the Am. Math. Society* 67. (1949) 36–50. old.
- [3.] H. Cramér: *Skandinavisk Aktuarietidskrift*, 11. (1928) 13–74. old. és 141–180. old.
- [4.] M. D. Donsker: *Memoirs of the Am. Math. Society* 6. (1951) 12. old.
- [5.] M. D. Donsker: *Annals of Math. Statistics* 23. (1952) 277–281. old.
- [6.] J. L. Doob: *Annals of Math. Statistics* 20. (1949) 393–403. old.
- [7.] W. Feller: *Annals of Math. Statistics* 19. (1948) 177–189. old.
- [8.] В. И. Гливенко: *Giornale dell'Istituto Italiano degli Attuari* 4 (1933). 92–99. old.
- [9.] Б. В. Гнеденко: *Курс теории вероятностей*, Гиттл. Москва Ленинград (1950) (Ft. 14. ) főleg: 278–292. old.
- [10.] Б. В. Гнеденко és В. С. Корольюк: *Доклады А. Н.* 80. (1951) 525–528. old.
- [11.] Б. В. Гнеденко és Е. Л. Рвачева: *Доклады А. Н.* 82. (1952) 513–516. old.
- [12.] Б. В. Гнеденко: *Доклады А. Н.* 82. (1952). 661–663. o.d.
- [13.] Б. В. Гнеденко és В. С. Михалевич: *Доклады А. Н.* 82. (1952) 841–843. old.
- [14.] Б. В. Гнеденко és В. С. Михалевич: *Доклады А. Н.* 85. (1952) 25–27. old.
- [15.] M. Kac: *Proceedings of the National Academy of Sciences of USA* 35. (1949) 252–257. old.
- [16.] А. Н. Колмогоров: *Известия Ак. Наук. Сер. физ.-мат.* (1933) 363–372. old. és *Giornale dell'Istituto Italiano degli Attuari* 4. (1933) 1–11. old. és 83–91. old.
- [17.] И. Д. Квнт: *Доклады А. Н.* 71. (1950) 229–231. old.
- [18.] Г. М. Манна: *Доклады А. Н.* 69 (1949) 495–497. old.
- [19.] F. J. Massey: *Annals of Math. Statistics* 21. (1950) 116–119. old.
- [20.] F. J. Massey: *Annals of Math. Statistics* 21. (1950) 440–443. old.
- [21.] F. J. Massey: *Annals of Math. Statistics* 22. (1951) 125–128. old.
- [22.] F. J. Massey: *Annals of Math. Statistics* 23. (1952) 435–441. old.
- [23.] В. С. Михалевич: *Доклады А. Н.* 85 (1952). 485–488. old.
- [24.] R. Mises: *»Wahrscheinlichkeitsrechnung«*, Deuticke Wien. 1931.
- [25.] А. У. Паявин: *Автореферат Моск. пед. инст. им. В. П. Потемкина* (1951).
- [26.] Szentmártony Tibor: *Matematikai statisztika a műszaki gyakorlatban*, Mérnöki Továbbképző Int. Kiadv. Bpest, (1950.) (Ft. 27.50) főleg: 98–99. old.
- [27.] Н. В. Смирнов: *Comptes Rendus Ac. Sc. Paris.* 202. (1936) 449. old.

- [28.] Н. В. Смирнов: Матем. Сборник. 2. (44) (1937) 973 - 993. old.  
 [29.] Н. В. Смирнов: Матем. Сборник. 6 (44) (1939) 3 26. old. és Бюллетен Мат. 2. (1939).  
 [30.] Н. В. Смирнов: Бюллетен (1939). мат. 2. és Annals of Math. Statistics 19. (1948) 279—281. old.  
 [31.] A. Wald és J. Wolfowitz: Annals of Math. Statistics 10. (1939) 105—118. old.

## О НЕКОТОРЫХ ВОПРОСАХ МЕДИЦИНСКО-БИОЛОГИЧЕСКОГО ПРИМЕНЕНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

И. ЮВАНЦ — Т. ЛИПТАК

### Резюме

Первая часть работы трактует о принципиальных вопросах применений. Она группирует специальные проблемы, проистекающие из характера медицинско-биологического исследования, вокруг двух центральных вопросов. Один круг проблем связан с исследованием «сигнификации» именно с правильным толкованием самой «сигнификации». Другой круг проблем относится к выборкам медицинско-биологического исследования, имеющим часто, в силу необходимости, лишь малое число элементов. Что касается другой группы вопросов, работа предлагает введение статистических методов, независимых от распределения и вырабатываемые также и для малых выборок, т. е. она предлагает применение методов, независимых от почти всех предварительных предположений, относящихся к исследуемой статистической совокупности.

Вторая часть работы, связанная с вышеуказанным предложением, познакомит с главными теоремами теории вариационных рядов и сообщает новейшие советские результаты в этой области. С помощью этих результатов, третья часть работы вырабатывает для исследования сигнификационные методы которые не зависят от распределения совокупности, и являются применимы и при малых выборках. Наконец, иллюстрируется практическое применение этих новых методов несколькими примерами медицинской научно-исследовательской работы, и сообщаются необходимые для вычислений таблицы, составленные Группой Медицинской Математической Статистики Института Прикладной Математики.

Работа исходит из медицинско-биологических применений, а большинство трактующих принципиальных вопросов имеет более общий характер. Статистические методы, изложенные в этой работе, могут широко применяться.

## SUR QUELQUES PROBLÈMES DE L'APPLICATION MÉDICO-BIOLOGIQUE DE LA STATISTIQUE MATHÉMATIQUE

I. JUVAN CZ — T. LIPTÁK

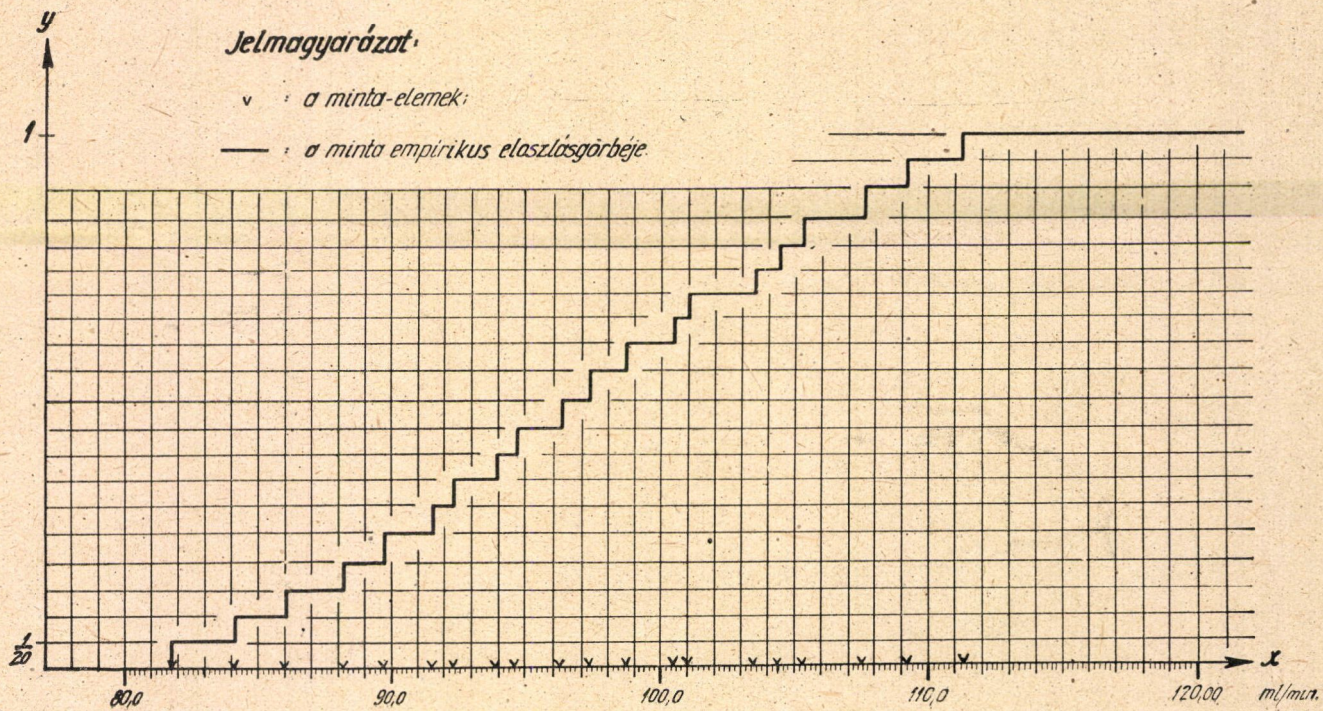
### RÉSUMÉ

La première partie de la publication s'occupe des questions de principe des applications. Il groupe les problèmes spéciaux, qui résultent de la nature des recherches médico-biologiques, autour de deux questions: l'une des sphères de problèmes s'attache aux recherches de signification, notamment à l'interprétation correcte de la signification elle-même, tandis que l'autre sphère de problème dérive de la prise d'échantillon qui, dans les recherches médico-biologiques, souvent ne contient, par nécessité, qu'un nombre réduit d'éléments. Tandis qu'en ce qui regarde le premier groupe de problèmes l'article, vu le fait que les recherches concernant ce groupe de problèmes ne sont pas encore achevées, ne s'occupe que de la seule formulation du problème et de son importance, il propose, en ce qui regarde le deuxième groupe de problèmes, l'introduction de méthodes statistiques «exemptes de distribution» élaborées pour les petits échantillons aussi, c'est-à-dire l'application de méthodes indépendantes de presque toute supposition préalable.

La seconde partie de l'article s'attachant à la proposition mentionnée décrit les théorèmes principaux de la statistique d'ordre, (en d'autres termes: de la théorie des séries de variations) et renseigne aussi sur les résultats soviétiques les plus récents dans

cette direction. En utilisant ces résultats, la publication, dans sa troisième partie, élabore des méthodes exemptes de distribution qui peuvent être employées, dans le cas des petits échantillons aussi, pour l'examen de la signification. Enfin l'article illustre l'application pratique de ces méthodes nouvelles par quelques exemples de rapports médicaux et publie les tables nécessaires pour les calculs qui ont été élaborés par le groupe de statistique mathématique médicale de l'Institut des Mathématiques Appliquées.

L'article prend comme point de départ les problèmes qui se présentent dans les applications médico-biologiques, mais un certain nombre des questions de principe discutées ici sont de caractère plus général, et l'on peut s'attendre que les méthodes statistiques exemptes de distribution, basées sur la statistique d'ordre vont soulever un intérêt général et seront utilisées dans un large domaine d'application.



1. ábra

y

*Jelmagyarázat*

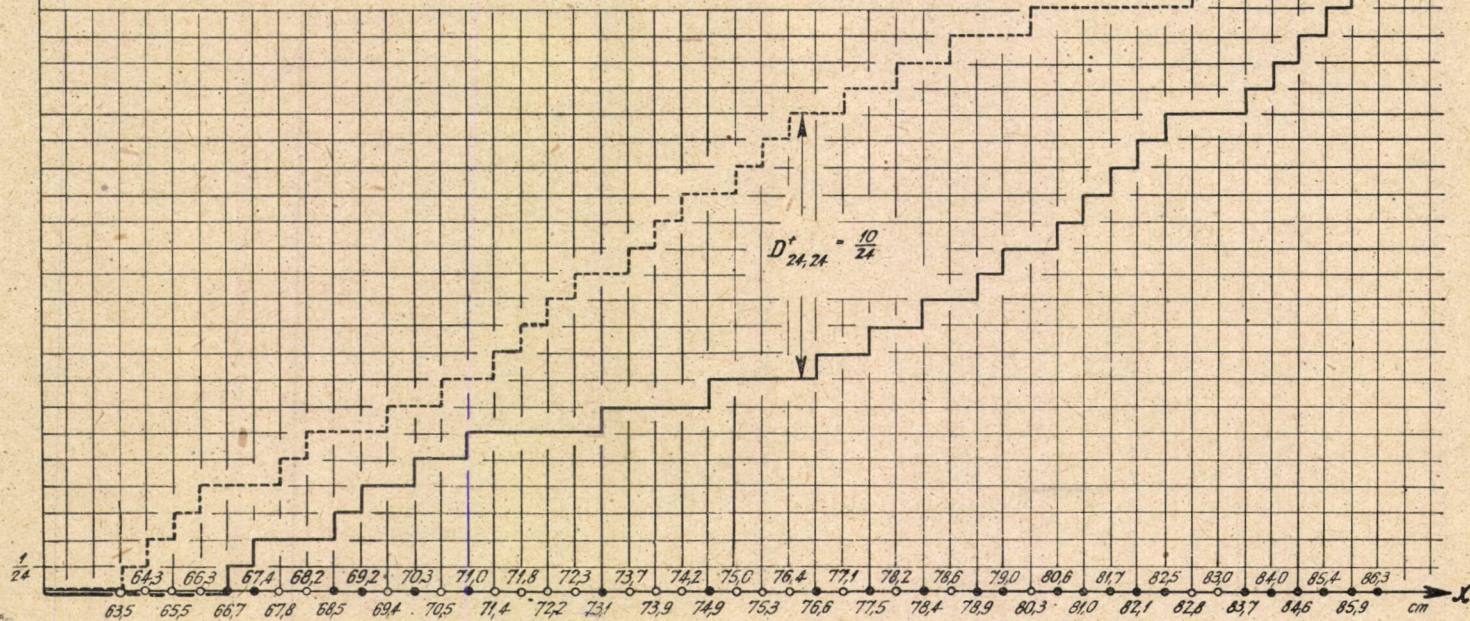
• • Y - minta-elem.

--- • Y - minta empirikus eloszlásgörbéje ( $S_n(x)$ )

• • X - minta-elem.

— • X - minta empirikus eloszlásgörbéje ( $T_n(x)$ )

f



2. ábra

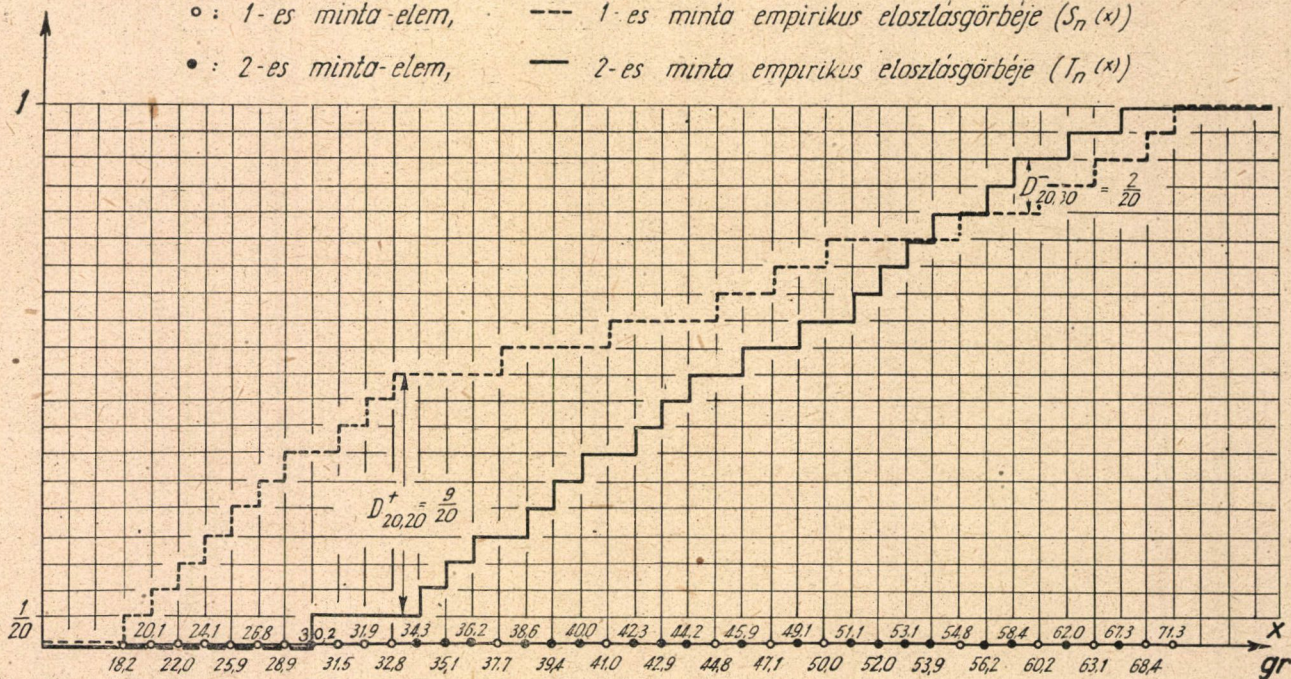
### Jelmagyarázat:

○ : 1-es minta-elem,

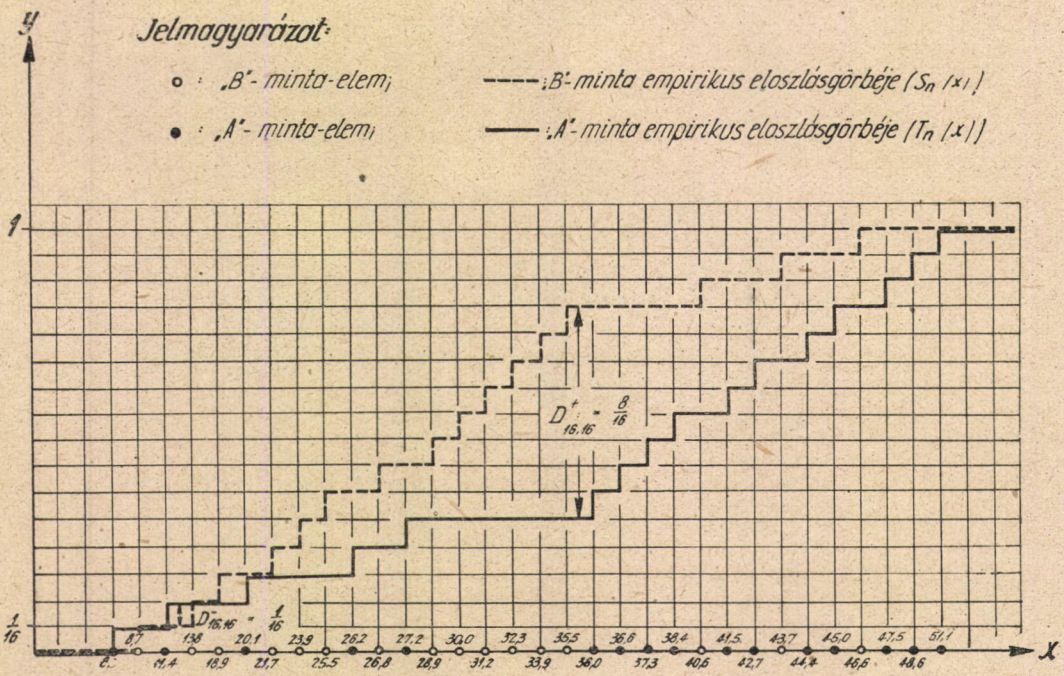
--- 1-es minta empirikus eloszlásgörbéje ( $S_n(x)$ )

● : 2-es minta-elem,

— 2-es minta empirikus eloszlásgörbéje ( $T_n(x)$ )



3. ábra



4. ábra

## I. Táblázat.

Az «abszolút-D»-kritérium alkalmazásához szükséges  $v_n(c)$  valószínűségek értékei kismintákra

$$v_n(c) = V \left\{ D_{n,n} \geq \frac{c}{n} \right\} = \frac{2}{\binom{2n}{n}} \sum_{v=1}^{\left[ \frac{n}{c} \right]} (-1)^{v-1} \binom{2n}{n-vc}$$

100  $v_n(c)$  %

n \ c	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
5	100,00	87,30	35,71	7,94	0,79	—	—	—	—	—	—	—	—	—	—	—	—	—
6	100,00	93,07	47,40	14,29	2,60	0,22	—	—	—	—	—	—	—	—	—	—	—	—
7	100,00	96,27	57,52	21,21	5,30	0,82	0,06	—	—	—	—	—	—	—	—	—	—	—
8	100,00	98,01	66,01	28,27	8,70	1,87	0,25	0,02	—	—	—	—	—	—	—	—	—	—
9	100,00	98,95	73,01	35,17	12,59	3,36	0,63	0,07	0,00	—	—	—	—	—	—	—	—	—
10	100,00	99,45	78,69	41,75	16,78	5,24	1,23	0,21	0,02	0,00	—	—	—	—	—	—	—	—
11	100,00	99,71	83,26	47,92	21,15	5,77	1,70	0,38	0,06	0,01	0,00	—	—	—	—	—	—	—
12	100,00	99,85	86,90	53,61	25,58	9,95	3,14	0,79	0,15	0,02	0,00	—	—	—	—	—	—	—
13	100,00	99,92	89,78	58,82	29,99	12,65	4,43	1,26	0,29	0,05	0,01	0,00	—	—	—	—	—	—
14	100,00	99,96	92,06	63,55	34,33	15,49	5,90	1,88	0,49	0,10	0,02	0,00	—	—	—	—	—	—
15	100,00	99,98	93,83	67,81	38,55	18,44	7,55	2,62	0,77	0,18	0,04	0,01	0,00	—	—	—	—	—
16	100,00	99,99	95,23	71,64	42,63	21,45	9,33	3,50	1,12	0,30	0,07	0,01	0,00	—	—	—	—	—
17	100,00	99,99	96,31	75,06	46,54	24,50	11,24	4,50	1,56	0,46	0,12	0,02	0,00	—	—	—	—	—
18	100,00	100,00	97,15	78,10	50,26	27,54	13,24	5,60	2,07	0,67	0,18	0,04	0,01	0,00	—	—	—	—
19	100,00	100,00	97,81	80,81	53,79	30,57	15,32	6,81	2,67	0,92	0,28	0,07	0,02	0,00	—	—	—	—
20	100,00	100,00	98,31	83,20	57,13	33,56	17,45	8,11	3,35	1,23	0,40	0,11	0,03	0,01	0,00	—	—	—
21	100,00	100,00	98,70	85,31	60,28	36,50	19,63	9,48	4,11	1,59	0,55	0,17	0,04	0,01	0,00	—	—	—
22	100,00	100,00	99,01	87,17	63,24	39,37	21,84	10,93	4,93	2,00	0,73	0,24	0,07	0,02	0,00	—	—	—
23	100,00	100,00	99,24	88,80	66,01	42,18	24,06	12,43	5,83	2,47	0,95	0,32	0,10	0,03	0,01	0,00	—	—
24	100,00	100,00	99,42	90,24	68,60	44,90	26,28	13,98	6,78	2,99	1,20	0,43	0,14	0,04	0,01	0,00	—	—
25	100,00	100,00	99,55	91,50	71,02	47,55	28,50	15,58	7,79	3,56	1,48	0,56	0,19	0,06	0,02	0,00	—	—
26	100,00	100,00	99,66	92,60	73,27	50,10	30,71	17,20	8,85	4,18	1,81	0,71	0,26	0,08	0,02	0,01	0,00	—
27	100,00	100,00	99,74	93,57	75,37	52,56	32,90	18,86	9,96	4,84	2,17	0,98	0,33	0,11	0,03	0,01	0,00	—
28	100,00	100,00	99,80	94,41	77,32	54,94	35,06	20,53	11,10	5,55	2,56	1,09	0,42	0,15	0,05	0,01	0,00	—
29	100,00	100,00	99,85	95,14	79,12	57,22	37,20	22,21	12,29	6,30	2,99	1,31	0,53	0,20	0,07	0,02	0,01	0,00
30	100,00	100,00	99,88	95,78	80,80	59,41	39,29	23,91	13,50	7,09	3,46	1,56	0,65	0,25	0,09	0,03	0,01	0,00



II. Táblázat (1. rész)

A »kétoldali-D«-kritérium alkalmazásához szükséges  $w_n(a,b)$  valószínűségek értékei kismintákra

$$w_n(a,b) = V \left\{ D_{n,n}^- \cong \frac{a}{n}; D_{n,n}^+ \cong \frac{b}{n} \right\} = \frac{1}{\binom{2n}{n}} \left\{ \sum_{v=1}^n \binom{2n}{n+a+1-v(a+b+1)} + \sum_{v=1}^n \binom{2n}{n+b-v(a+b+1)} - 2 \sum_{v=1}^n \binom{2n}{n-v(a+b+1)} \right\}$$

100  $w_n(a,b)$  ‰

n	a \ b	100 $w_n(a,b)$ ‰									
		1	2	3	4	5	6	7	8	9	10
4	0	22,86	18,57	8,57	1,43	—	—	—	—	—	—
	1	58,57	37,14	11,43	—	—	—	—	—	—	—
5	0	16,67	16,27	10,32	3,17	0,40	—	—	—	—	—
	1	51,98	39,68	17,06	3,97	—	—	—	—	—	—
	2	75,79	46,83	17,86	—	—	—	—	—	—	—
6	0	14,29	14,18	10,82	4,65	1,08	0,11	—	—	—	—
	1	46,32	39,50	21,21	6,93	1,30	—	—	—	—	—
	2	72,73	50,97	23,59	7,14	—	—	—	—	—	—
7	0	12,50	12,47	10,64	5,71	1,86	0,35	0,03	—	—	—
	1	41,64	37,94	23,89	9,79	2,59	0,41	—	—	—	—
	2	68,97	53,06	28,35	10,55	2,65	—	—	—	—	—
8	0	11,11	11,10	10,12	6,37	2,61	0,69	0,11	0,01	—	—
	1	37,76	35,79	25,37	12,28	4,23	0,92	0,12	—	—	—
	2	65,07	53,65	32,08	13,89	4,34	0,93	—	—	—	—
	3	81,12	60,37	33,69	14,13	4,35	—	—	—	—	—
9	0	10,00	10,00	9,47	6,72	3,25	1,69	0,24	0,03	0,00	—
	1	34,54	33,49	25,95	14,30	5,67	1,60	0,31	0,04	—	—
	2	61,29	53,22	34,83	16,99	6,22	1,67	0,32	—	—	—
	3	79,09	62,14	37,55	17,55	6,29	1,68	—	—	—	—

n	a \ b													
		1	2	3	4	5	6	7	8	9	10	11	12	13
10	0	9,09	9,09	8,81	6,83	3,76	1,49	0,42	0,08	0,01	0,00	—	—	—
	1	31,82	31,26	25,89	15,84	7,17	2,42	0,60	0,10	0,01	—	—	—	—
	2	57,76	52,12	36,74	19,75	8,19	2,60	0,62	0,10	—	—	—	—	—
	3	76,76	63,04	40,72	20,77	8,37	2,62	0,62	—	—	—	—	—	—
11	0	8,33	8,33	8,19	6,78	4,15	1,88	0,63	0,16	0,03	0,00	—	—	—
	1	29,49	29,20	25,42	16,93	8,54	3,30	0,97	0,21	0,03	0,00	—	—	—
	2	54,53	50,61	37,93	22,11	10,14	3,67	1,03	0,22	0,03	—	—	—	—
	3	74,27	63,27	43,26	23,74	10,51	3,73	1,04	0,22	—	—	—	—	—
	4	85,24	68,47	44,89	24,11	10,57	3,73	1,04	—	—	—	—	—	—
12	0	7,69	7,69	7,62	6,63	4,42	2,23	0,86	0,25	0,06	0,01	0,00	—	—
	1	27,47	27,32	24,70	17,65	9,73	4,20	1,42	0,37	0,07	0,01	0,00	—	—
	2	51,57	48,87	38,55	24,07	12,01	4,83	1,55	0,39	0,07	0,01	—	—	—
	3	71,74	62,98	45,22	26,41	12,65	4,96	1,57	0,39	0,08	—	—	—	—
	4	83,93	69,46	47,57	27,05	12,78	4,98	1,57	0,39	—	—	—	—	—
13	0	7,14	7,14	7,10	6,42	4,59	2,53	1,09	0,37	0,10	0,02	0,00	—	—
	1	25,71	25,64	23,83	18,05	10,74	5,08	1,93	0,58	0,14	0,02	0,00	—	—
	2	48,89	47,04	38,71	25,64	13,76	6,04	2,16	0,63	0,14	0,02	0,00	—	—
	3	69,24	62,29	46,67	28,78	14,73	6,28	2,21	0,63	0,14	0,02	—	—	—
	4	82,42	70,00	49,81	29,75	14,97	6,32	2,21	0,63	0,14	—	—	—	—
14	0	6,67	6,67	6,65	6,18	4,68	2,78	1,32	0,50	0,15	0,04	0,01	0,00	—
	1	24,17	24,13	22,88	18,21	11,56	5,92	2,47	0,84	0,23	0,05	0,01	0,00	—
	2	46,45	45,19	38,52	26,87	15,35	7,26	2,85	0,92	0,24	0,05	0,01	0,00	—
	3	66,80	61,33	47,69	30,84	16,73	7,65	2,94	0,94	0,24	0,05	0,01	—	—
	4	80,80	70,18	51,66	32,22	17,11	7,73	2,95	0,94	0,24	0,05	—	—	—
	5	88,37	74,01	53,04	32,61	17,20	7,75	2,95	0,94	0,24	—	—	—	—

n	a \ b	1	2	3	4	5	5	7	8	9	10	11	12	13	14
15	0	6,25	6,25	6,24	5,92	4,71	2,99	0,64	0,22	0,06	0,01	0,00	—	—	—
	1	22,79	22,77	21,93	18,17	12,21	6,69	3,03	1,13	0,35	0,09	0,02	0,00	—	—
	2	44,23	43,37	38,07	27,78	16,76	8,46	3,59	1,28	0,38	0,09	0,02	0,00	—	—
	3	64,45	60,16	48,32	32,60	18,60	9,04	3,74	1,31	0,38	0,09	0,02	0,00	—	—
	4	79,09	70,04	48,22	34,45	19,19	9,19	3,77	1,31	0,38	0,09	0,02	—	—	—
	5	87,52	74,67	55,00	35,03	19,33	9,22	3,77	1,31	0,38	0,09	—	—	—	—
16	0	5,88	5,88	5,88	5,66	4,69	3,15	1,73	0,78	0,29	0,09	0,02	0,00	—	—
	1	21,57	21,56	20,98	17,99	12,70	7,38	3,58	1,45	0,49	0,14	0,03	0,01	0,00	—
	2	42,20	41,62	37,43	28,42	18,00	9,62	4,37	1,68	0,55	0,15	0,03	0,01	0,00	—
	3	62,21	58,86	48,64	34,08	20,35	10,43	4,60	1,74	0,56	0,15	0,03	0,01	0,00	—
	4	77,34	69,66	54,33	36,45	21,16	10,67	4,65	1,75	0,56	0,15	0,03	0,01	—	—
	5	86,54	75,08	56,68	37,26	21,40	10,72	4,66	1,75	0,56	0,15	0,03	—	—	—
17	0	5,56	5,56	5,55	5,40	4,63	3,27	1,90	0,92	0,37	0,13	0,04	0,01	0,00	—
	1	20,47	20,46	20,07	17,70	13,05	8,00	4,12	1,80	0,66	0,21	0,05	0,01	0,00	—
	2	40,35	39,96	36,66	28,82	19,06	10,72	5,16	2,13	0,75	0,23	0,06	0,01	0,00	—
	3	60,08	57,47	48,70	35,30	21,94	11,80	5,50	2,22	0,77	0,23	0,06	0,01	0,00	—
	4	75,58	69,08	55,21	38,21	23,04	12,14	5,60	2,24	0,78	0,23	0,06	0,01	0,00	—
	5	85,46	75,27	58,11	39,31	23,38	12,24	5,62	2,25	0,78	0,23	0,06	0,01	—	—
18	0	5,26	5,26	5,26	5,16	4,55	3,36	2,06	1,06	0,46	0,17	0,05	0,01	0,00	—
	1	19,47	19,47	19,21	17,34	13,28	8,52	4,64	2,16	0,86	0,29	0,08	0,02	0,00	—
	2	38,65	38,38	35,80	29,02	19,96	11,74	5,96	2,62	1,00	0,33	0,09	0,02	0,00	—
	3	58,07	56,03	48,54	36,27	23,39	13,13	6,44	2,76	1,03	0,33	0,09	0,02	0,00	—
	4	73,83	68,34	55,85	39,76	24,80	13,61	6,58	2,79	1,04	0,33	0,09	0,02	0,00	—
	5	84,30	75,26	59,30	41,17	25,28	13,75	6,61	2,80	1,04	0,33	0,09	0,02	0,00	—
	6	90,17	78,48	60,69	41,65	25,42	13,78	6,62	2,80	1,04	0,33	0,09	0,02	—	—

n	a \ b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
19	0	5,00	5,00	5,00	4,93	4,45	3,41	2,19	1,19	0,55	0,22	0,07	0,02	0,01	0,00		
	1	18,57	18,57	18,39	16,93	13,41	8,97	5,12	2,52	1,07	0,39	0,12	0,03	0,01	0,00		
	2	37,08	36,90	34,89	29,05	20,70	12,68	6,75	3,13	1,27	0,45	0,14	0,04	0,01	0,00		
	3	56,16	54,58	48,20	37,03	24,69	14,40	7,38	3,33	1,32	0,46	0,14	0,04	0,01	0,00		
	4	72,10	67,48	56,26	41,11	26,44	15,07	7,59	3,39	1,33	0,46	0,14	0,04	0,01	0,00		
	5	83,09	75,08	60,28	42,85	27,08	15,25	7,64	3,40	1,34	0,46	0,14	0,04	0,01	0,00	—	—
	6	89,53	78,61	62,01	43,50	27,29	15,30	7,66	3,40	1,34	0,46	0,14	0,04	0,01	—	—	—
20	0	4,76	4,76	4,76	4,72	4,34	3,44	2,30	1,31	0,65	0,27	0,10	0,03	0,01	0,00		
	1	17,75	17,75	17,63	16,48	13,45	9,35	5,57	2,88	1,59	0,51	0,17	0,05	0,01	0,00		
	2	35,63	35,51	33,94	28,94	21,29	13,54	7,51	3,66	1,57	0,59	0,19	0,06	0,01	0,00		
	3	54,36	53,14	47,72	37,60	25,84	15,61	8,33	3,94	1,65	0,61	0,20	0,06	0,01	0,00		
	4	70,40	66,52	56,48	42,26	27,95	16,44	8,62	4,03	1,67	0,61	0,20	0,06	0,01	0,00		
	5	81,84	74,76	61,07	44,37	28,78	16,72	8,70	4,05	1,68	0,61	0,20	0,06	0,01	0,00		—
	6	88,81	79,10	63,15	45,20	29,07	16,81	8,72	4,05	1,68	0,61	0,20	0,06	0,01	0,00	—	—
	7	92,50	81,08	63,98	45,48	29,15	16,83	8,73	4,05	1,68	0,62	0,20	0,06	0,01	—	—	—
21	0	4,55	4,55	4,55	4,51	4,22	3,44	2,39	1,43	0,74	0,33	0,13	0,04	0,01	0,00		
	1	17,00	17,00	16,92	16,02	13,02	9,65	5,99	3,23	1,53	0,64	0,23	0,07	0,02	0,00		
	2	34,29	34,21	32,99	28,73	21,75	14,31	8,25	4,20	1,89	0,75	0,26	0,08	0,02	0,00		
	3	52,66	51,72	47,13	38,00	26,85	16,75	9,28	4,58	2,01	0,79	0,27	0,08	0,02	0,00		
	4	68,74	65,49	56,52	43,23	29,35	17,79	9,65	4,70	2,04	0,79	0,27	0,08	0,02	0,00		
	5	80,57	74,32	61,69	45,73	30,39	18,17	9,78	4,73	2,05	0,80	0,27	0,08	0,02	0,01	0,00	
	6	88,03	79,17	64,14	46,77	30,77	18,29	9,81	4,74	2,05	0,80	0,27	0,08	0,02	0,01	0,00	—
	7	92,14	81,50	65,17	47,15	30,89	18,32	9,82	4,74	2,05	0,80	0,27	0,08	0,02	0,02	—	—
22	0	4,35	4,35	4,35	4,33	4,09	3,43	2,47	1,53	0,83	0,39	0,16	0,06	0,02	0,01	0,00	
	1	16,30	16,30	16,25	15,56	13,34	9,89	6,36	3,58	1,77	0,78	0,30	0,10	0,03	0,01	0,00	
	2	33,04	32,99	32,05	28,43	22,09	14,99	8,96	4,74	2,23	0,94	0,35	0,11	0,03	0,01	0,00	
	3	51,05	50,32	46,45	38,24	27,73	17,81	10,20	5,23	2,40	0,99	0,36	0,12	0,03	0,01	0,00	
	4	67,13	64,41	56,43	44,04	30,62	19,08	10,69	5,40	2,45	1,00	0,36	0,12	0,03	0,01	0,00	
	5	79,28	73,78	62,14	46,93	31,89	19,57	10,86	5,45	2,46	1,00	0,36	0,12	0,03	0,01	0,00	
	6	87,19	79,13	64,99	48,20	32,38	19,74	10,91	5,46	2,47	1,00	0,36	0,12	0,03	0,01	0,00	
	7	91,72	81,81	66,24	48,69	32,55	19,79	10,92	5,46	2,47	1,00	0,36	0,12	0,03	0,01	0,00	—

n	b		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	a																	
23	0		4,17	4,17	4,17	4,15	3,98	3,41	2,53	1,62	0,91	0,45	0,20	0,08	0,03	0,01	0,00	
	1		15,67	15,67	15,63	15,10	13,20	10,07	6,26	3,91	2,02	0,93	0,38	0,14	0,04	0,01	0,00	
	2		31,88	31,85	31,12	28,05	22,33	15,59	9,62	5,28	2,59	1,14	0,45	0,16	0,05	0,01	0,00	
	3		49,53	48,97	45,71	38,35	28,47	18,79	11,10	5,89	2,82	1,21	0,47	0,16	0,05	0,01	0,00	
	4		65,57	62,45	56,21	44,63	31,77	20,31	11,72	6,12	2,89	1,23	0,47	0,16	0,05	0,01	0,00	
	5		77,99	73,15	62,45	48,00	33,29	20,93	11,94	6,19	2,91	1,24	0,47	0,16	0,05	0,01	0,00	
	6		86,32	78,98	65,69	49,51	33,91	21,15	12,02	6,21	2,91	1,24	0,47	0,16	0,05	0,01	0,00	
	7		91,24	82,02	67,18	50,13	34,14	21,24	12,04	6,21	2,91	1,24	0,47	0,16	0,05	0,01	0,00	
	8		93,83	83,44	67,79	50,36	34,21	21,25	12,04	6,21	2,91	1,24	0,47	0,16	0,05	0,01	0,00	—
24	0		4,00	4,00	4,00	3,99	3,85	3,37	2,57	1,71	1,00	0,52	0,24	0,10	0,03	0,01	0,00	
	1		15,08	15,08	15,05	14,64	13,03	10,20	6,98	4,22	2,26	1,08	0,46	0,18	0,06	0,02	0,00	
	2		30,80	30,78	30,22	27,63	22,47	16,11	10,24	5,82	2,96	1,36	0,56	0,21	0,07	0,02	0,01	0,00
	3		48,09	47,66	44,93	38,34	29,10	19,70	11,97	6,56	3,25	1,46	0,59	0,21	0,07	0,02	0,01	0,00
	4		64,06	62,17	55,88	45,21	32,81	21,47	12,73	6,85	3,35	1,49	0,60	0,22	0,07	0,02	0,01	0,00
	5		76,70	72,45	62,64	48,93	34,60	22,24	13,02	6,95	3,38	1,49	0,60	0,22	0,07	0,02	0,01	0,00
	6		85,41	78,74	66,28	50,71	35,36	22,53	13,12	6,98	3,39	1,50	0,60	0,22	0,07	0,02	0,01	0,00
	7		90,72	82,15	68,02	51,47	35,65	22,63	13,15	6,99	3,39	1,50	0,60	0,22	0,07	0,02	0,01	0,00
	8		93,61	83,80	68,77	51,76	35,75	22,66	13,16	6,99	3,39	1,50	0,60	0,22	0,07	0,02	0,01	0,00
25	0		3,85	3,85	3,85	3,84	3,73	3,33	2,60	1,78	1,08	0,58	0,28	0,12	0,05	0,02	0,00	
	1		14,53	14,53	14,51	14,20	12,84	10,28	7,24	4,51	2,51	1,25	0,56	0,22	0,08	0,03	0,01	0,00
	2		29,79	29,78	29,35	27,16	22,54	16,55	10,82	6,34	3,34	1,59	0,68	0,26	0,09	0,03	0,01	0,00
	3		46,73	46,40	44,11	38,24	29,62	20,52	12,81	7,23	3,70	1,72	0,73	0,28	0,10	0,03	0,01	0,00
	4		62,61	62,03	55,46	45,60	33,74	22,57	13,72	7,60	3,84	1,76	0,74	0,28	0,10	0,03	0,01	0,00
	5		75,43	71,70	62,71	49,73	35,80	23,49	14,09	7,73	3,88	1,78	0,74	0,28	0,10	0,03	0,01	0,00
	6		84,47	78,42	66,74	51,79	36,72	23,86	14,22	7,77	3,89	1,78	0,74	0,28	0,10	0,03	0,01	0,00
	7		90,14	82,19	68,75	52,70	37,09	23,99	14,26	7,79	3,89	1,78	0,74	0,28	0,17	0,03	0,01	0,00
	8		93,34	84,08	69,65	53,07	37,22	24,04	14,28	7,80	3,90	1,78	0,74	0,28	0,10	0,03	0,01	0,00

n	a \ b	b																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
26	0	3,70	3,70	3,70	3,70	3,62	3,27	2,62	1,84	1,15	0,64	0,32	0,14	0,06	0,02	0,01	0,00	
	1	14,02	14,02	14,01	13,76	12,62	10,32	7,45	4,78	2,75	1,42	0,66	0,28	0,10	0,04	0,01	0,00	
	2	28,85	28,83	28,50	26,67	22,53	16,92	11,35	6,84	3,72	1,84	0,82	0,33	0,12	0,04	0,01	0,00	
	3	45,44	45,19	43,27	38,05	30,03	21,27	13,60	7,89	4,17	2,01	0,88	0,35	0,13	0,04	0,01	0,00	
	4	61,20	59,89	54,97	45,87	34,56	23,60	14,68	8,35	4,34	2,07	0,89	0,36	0,13	0,04	0,01	0,00	
	5	74,16	70,90	62,67	50,42	36,91	24,69	15,14	8,52	4,40	2,08	0,90	0,36	0,13	0,04	0,01	0,00	
	6	83,52	78,03	67,10	52,76	38,00	25,15	15,31	8,58	4,42	2,09	0,90	0,36	0,13	0,04	0,01	0,00	
	7	89,54	82,16	69,39	53,84	38,46	25,32	15,37	8,60	4,42	2,09	0,90	0,36	0,13	0,04	0,01	0,00	
	8	93,02	84,30	70,45	54,30	38,63	25,38	15,39	8,60	4,42	2,09	0,90	0,36	0,13	0,04	0,01	0,00	
	9	94,85	85,31	70,90	54,47	38,69	25,40	15,40	8,60	4,42	2,09	0,90	0,36	0,13	0,04	0,01	0,00	
27	0	3,57	3,57	3,57	3,57	3,51	3,21	2,63	1,90	1,22	0,70	0,36	0,17	0,07	0,03	0,01	0,00	
	1	13,55	13,55	13,54	13,35	12,39	10,33	7,64	5,03	2,98	1,59	0,77	0,34	0,13	0,05	0,02	0,00	
	2	27,96	27,95	27,70	26,16	22,46	17,23	11,83	7,32	4,11	2,09	0,97	0,41	0,16	0,05	0,02	0,00	
	3	44,22	44,03	42,42	37,79	30,36	21,94	14,36	8,55	4,65	2,31	1,05	0,44	0,16	0,06	0,02	0,00	
	4	59,85	58,75	54,42	46,04	35,28	24,57	15,62	9,10	4,87	2,39	1,07	0,44	0,17	0,06	0,02	0,01	0,00
	5	72,92	70,06	62,55	50,99	37,93	25,84	16,17	9,32	4,94	2,41	1,08	0,44	0,17	0,06	0,02	0,01	0,00
	6	82,55	77,58	67,37	53,63	39,20	26,39	16,39	9,40	4,97	2,42	1,08	0,44	0,17	0,06	0,02	0,01	0,00
	7	88,89	82,05	69,93	54,89	39,76	26,61	16,47	9,42	4,98	2,42	1,08	0,44	0,17	0,06	0,02	0,01	0,00
	8	92,67	84,45	71,17	55,44	39,98	26,69	16,50	9,43	4,98	2,42	1,08	0,44	0,17	0,06	0,02	0,01	0,00
	9	94,71	85,62	71,72	55,66	40,05	26,72	16,50	9,43	4,98	2,42	1,08	0,44	0,17	0,06	0,02	0,01	0,00
28	0	3,45	3,45	3,45	3,45	3,40	3,15	2,63	1,94	1,28	0,76	0,40	0,20	0,09	0,03	0,01	0,00	
	1	13,10	13,10	13,10	12,95	12,15	10,31	8,15	5,26	3,20	1,76	0,88	0,40	0,17	0,06	0,02	0,01	0,00
	2	27,12	27,11	26,92	25,63	22,35	17,48	12,27	7,78	4,49	2,36	1,13	0,50	0,20	0,07	0,02	0,01	0,00
	3	43,06	42,91	41,57	37,47	30,59	22,54	15,07	9,19	5,13	2,62	1,23	0,53	0,21	0,08	0,02	0,01	0,00
	4	58,54	57,63	53,82	46,12	35,91	25,46	16,52	9,85	5,40	2,73	1,27	0,54	0,21	0,08	0,02	0,01	0,00
	5	71,69	69,19	62,35	51,47	38,86	26,92	17,18	10,12	5,50	2,76	1,28	0,54	0,21	0,08	0,02	0,01	0,00
	6	81,57	77,08	67,54	54,40	40,32	27,59	17,46	10,22	5,54	2,77	1,28	0,54	0,21	0,08	0,02	0,01	0,00
	7	88,22	81,89	70,39	55,86	40,98	27,86	17,56	10,26	5,55	2,78	1,28	0,54	0,21	0,08	0,02	0,01	0,00
	8	92,28	84,56	71,82	56,52	41,26	27,96	17,59	10,27	5,55	2,78	1,28	0,54	0,21	0,08	0,02	0,01	0,00
	9	94,53	85,88	72,46	56,79	41,36	28,00	17,60	10,27	5,55	2,78	1,28	0,54	0,21	0,08	0,02	0,01	0,00

n \ a \ b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
29	0	3,33	3,33	3,33	3,33	3,09	2,62	1,98	1,34	0,81	0,45	0,22	0,10	0,04	0,02	0,01	0,00	
	1	12,69	12,69	12,57	12,57	11,90	10,26	7,92	5,47	3,42	1,94	1,00	0,47	0,20	0,08	0,03	0,01	0,00
	2	26,33	26,33	26,18	25,10	22,19	17,67	12,66	8,22	4,86	2,63	1,30	0,59	0,25	0,09	0,03	0,01	0,00
	3	41,96	41,84	40,73	37,10	30,75	23,07	15,74	9,81	5,62	2,95	1,43	0,64	0,26	0,10	0,03	0,01	0,00
	4	57,28	56,53	53,18	46,12	36,44	26,28	17,39	10,59	5,95	3,08	1,48	0,65	0,26	0,10	0,03	0,01	0,00
	5	70,49	68,31	62,07	51,85	39,70	27,95	18,17	10,92	6,08	3,13	1,49	0,66	0,27	0,10	0,03	0,01	0,00
	6	80,59	76,53	67,64	55,10	41,37	28,73	18,50	11,05	6,12	3,14	1,50	0,66	0,27	0,10	0,03	0,01	0,00
	7	87,53	81,67	70,78	56,74	42,15	29,07	18,63	11,10	6,14	3,15	1,50	0,66	0,27	0,10	0,03	0,01	0,00
	8	91,86	84,59	72,39	57,51	42,48	29,20	18,68	11,11	6,14	3,15	1,50	0,66	0,27	0,10	0,03	0,01	0,00
	9	94,33	86,10	73,15	57,84	42,61	29,24	18,69	11,12	6,14	3,15	1,50	0,66	0,27	0,10	0,03	0,01	0,00
	10	95,62	86,81	73,47	57,98	42,66	29,26	18,70	11,12	6,14	3,15	1,50	0,66	0,27	0,10	0,03	0,01	0,00
30	0	3,23	3,23	3,23	3,22	3,20	3,02	2,61	2,01	1,39	0,87	0,49	0,25	0,12	0,05	0,02	0,01	0,00
	1	12,30	12,30	12,30	12,21	11,65	10,19	8,01	5,66	3,62	2,11	1,12	0,54	0,24	0,10	0,04	0,01	0,00
	2	25,59	25,58	25,47	24,57	21,99	17,81	13,02	8,64	5,23	2,90	1,48	0,69	0,30	0,12	0,04	0,01	0,00
	3	40,91	40,82	39,89	36,69	30,85	23,54	16,36	10,42	6,10	3,29	1,64	0,75	0,32	0,12	0,04	0,02	0,00
	4	56,07	55,45	52,50	46,04	36,90	27,05	18,21	11,32	6,50	3,46	1,70	0,77	0,33	0,13	0,04	0,02	0,00
	5	69,31	67,41	61,74	52,15	40,46	28,92	19,12	11,72	6,66	3,52	1,72	0,78	0,33	0,13	0,04	0,02	0,00
	6	79,61	75,94	67,66	55,68	42,34	29,83	19,52	11,88	6,72	3,54	1,73	0,78	0,33	0,13	0,04	0,02	0,00
	7	86,82	81,60	71,08	57,54	43,24	30,23	19,68	11,94	6,74	3,54	1,73	0,78	0,33	0,13	0,04	0,02	0,00
	8	91,41	84,57	72,89	58,44	43,65	30,40	19,74	11,96	6,75	3,54	1,73	0,78	0,33	0,13	0,04	0,02	0,00
	9	94,10	86,26	73,77	58,84	43,81	30,46	19,76	11,97	6,75	3,54	1,73	0,78	0,33	0,13	0,04	0,02	0,00
	10	95,54	87,09	74,16	59,00	43,87	30,48	19,77	11,97	6,75	3,54	1,73	0,78	0,33	0,13	0,04	0,02	0,00