



MITISZK  
Miskolc-Térségi Integrált  
Szakképző Központ

Magyarország célba ér



## VALÓSZÍNŰSÉG-SZÁMÍTÁS ÉS MATEMATIKAI STATISZTIKA

FEGYVERNEKI SÁNDOR

Miskolci Egyetem  
Gépészmérnöki és Informatikai Kar

Készült a HEFOP-3.2.2-P.-2004-10-0011-/1.0 sz.  
projekt támogatásával

Miskolc, 2007.



---

Fegyverneki Sándor

**VALÓSZÍNŰSÉG–SZÁMÍTÁS  
ÉS  
MATEMATIKAI STATISZTIKA**

Miskolci Egyetem, Alkalmazott Matematikai Tanszék  
Miskolc, 2007

# TARTALOMJEGYZÉK

|                                              |     |
|----------------------------------------------|-----|
| 1. BEVEZETÉS . . . . .                       | 1   |
| 2. A VALÓSZÍNŰSÉG FOGALMA . . . . .          | 5   |
| 3. A VALÓSZÍNŰSÉGI VÁLTOZÓ . . . . .         | 31  |
| 4. VÉLETLEN VEKTOROK . . . . .               | 51  |
| 5. MINTAVÉTEL, ALAPSTATISZTIKÁK . . . . .    | 69  |
| 6. BECSLÉSELMÉLET . . . . .                  | 73  |
| 7. HIPOTÉZISVIZSGÁLAT . . . . .              | 87  |
| 8. VÉLETLEN SZÁMOK, SZIMULÁCIÓ . . . . .     | 111 |
| 9. FÜGGELÉK . . . . .                        | 119 |
| IRODALOMJEGYZÉK . . . . .                    | 141 |
| VÉLETLEN ÉRTÉKEK I-III. BEMUTATÁSA . . . . . | 143 |

## 1. BEVEZETÉS

A mindennapi nyelvben a "véletlen" általában előreláthatatlant vagy ismeretlent jelent. Így, ha valaki az autóbuszok beérkezését figyeli anélkül, hogy a menetrendet ismerné, azt mondhatja, hogy a buszok véletlenszerűen érkeznek. Miután áttekinti az egy napi tapasztalatait, mindössze annyit mondhat, hogy a menetredhez képest bizonyos váratlan eltérések mutatkoznak. A véletlenszerűség ezen vonására, az előreláthatatlanságra, általában nem szokás hivatkozni a "véletlen" pontos matematikai definíciójában. Ha a szükséges szigorral akarjuk a véletlent definiálni, meg kell mondanunk, hogy kinek, milyen körülmények között és milyen eszközök felhasználása mellett volt az esemény előreláthatatlan. Egy ilyen definíciót miundannyiszor felül kellene vizsgálni, valahányszor a tudomány előrehaladása lehetővé teszi egy korábban előreláthatatlan esemény jóslását.

Tekintsük a legegyszerűbbnek gondolt, véletlen kísérletet, azaz egy szabályos pénzdarab feldobását (zárjuk ki a zavaró eseteket: megáll az élén, nem esik le, elgurul). Ez azt jelenti, hogy kétféle eredmény lehet: fej (jelölje 0) és írás (jelölje 1). Dobjuk fel a pénzdarab 200-szor és jegyezzük fel a sorozatot. Az ilyen lehetséges sorozatok száma

$$\begin{aligned} 2^{200} &= 1606938044258990275541962092341162602522202993782792835301376 \\ &\approx 1.6 \cdot 10^{60}. \end{aligned} \tag{1.1}$$

Ha látunk egy ilyen sorozatot, akkor ez alapján tudunk-e valamit mondani a pénzdarabról. Mikor mondhatjuk, hogy véletlen kísérleteket végeztünk? Egy véletlennek tekintett kísérlet esetén hogyan készíthetünk mérőszámokat, törvényszerűségeket és hogyan ellenőrizzük ezeket?

Léteznek-e "tökéletes" végtelen véletlen sorozatok? A válasz igenlő, sőt némely nagyon szabályosnak látszó sorozat is kielégítheti a feltételeket.

## BEVEZETÉS

---

Az egymás utáni természetes számok kettes számrendszerbeli alakját írjuk egymás mellé nagyság szerint:

$$1, 10, 11, 100, 101, 110, 111, 1000, \dots \quad (1.2)$$

Megmutatható, hogy  $k$  minden értékére és minden  $k$  hosszúságú  $B_k = b_1, b_2, \dots, b_k$  blokkra:

$$\lim_{n \rightarrow \infty} \frac{N(n, B_k)}{n - k + 1} = \lim_{n \rightarrow \infty} \frac{N(n, B_k)}{n} = \frac{1}{2^k}. \quad (1.3)$$

Alapfeladatunk a véletlen kísérletek leírása, azaz következtetés tapasztalati adatokból az ismeretlenre. Tekintsük a következő számhalmazokat:

### 1.1. táblázat: Véletlen értékek I.

3.88, 7.70, 1.65, 6.42, 4.29, 3.01, 1.41, 6.54, 3.58, 9.05, 1.12,  
1.74, 6.17, 3.38, 2.50, 4.01, 2.96, 8.48, 3.76, 5.75, 6.45, 7.89,  
1.77, 2.32, 8.83, 4.10, 9.18, 4.99, 1.30, 7.17, 2.00, 9.24, 2.25,  
7.80, 8.60, 5.84, 5.05, 8.56, 8.49, 6.95.

### 1.2. táblázat: Véletlen értékek II.

-0.340, 28.53, -12.38, -24.89, -6.636, -19.16, -4.626, 1.275,  
837.7, -11.83, 21.33, -54.07, 33.24, 1.238, -123.4, 13.61, -65.22,  
16.04, 129.9, 54.14, -16.92, -3.643, 88.06, -5.665, -89.20, 8.75,  
-5.429, -6.107, 411.4, 8.04, -5.791, 18.23, 1.357, -21.18, 51.81,  
105.8, -14.51, 10.78, 7.38, -21.65, -12.96, -3.379, 17.48, 4.310,  
15.23, -7.661, 1.853, -16.84, -19.26, -16.47.

### 1.3. táblázat: Véletlen értékek III.

## 1. FEJEZET

---

0.082, 0.668, 1.872, -1.494, 1.336, -2.547, -0.575, -3.587,  
-2.121, -0.030, -1.927, 2.916, 4.542, -2.921, 2.640, -1.718,  
2.772, 1.398, -2.021, 2.310, 0.373, 0.153, 0.118, -2.551, 1.691,  
2.121, 0.499, 0.711, -2.982, 2.799, -0.569, -3.025, 1.649, -4.036,  
0.314, 3.044, 3.482, 0.732, -2.225, -0.469, -1.882, -0.481,  
-2.581, 0.747, -2.886, -3.271, 1.317, 0.780, 1.803, -2.012,  
-4.149, -3.657, -0.214, -2.372, 0.997, -1.300, -2.129, -2.567,  
-0.223, 1.452.

Már ezeket a nem túl nagy számhalmazokat is nehezen tudjuk áttekinteni. A körülmények ismerete nélkül arról sem tudunk nyilatkozni, hogy véletlennek tekinthetők-e. Véletlen értékek I-III. bemutatása a tananyag legvégén.

A tananyag keretében megpróbálunk felépíteni olyan matematikai modelleket, amelyek segítik a véletlen megértését, a törvényszerűségeinek megismerését és a következtetést tapasztalati adatokból. Olyan módszerek kidolgozása, amelyek lehetővé teszik tapasztalati adatok információinak feldolgozását, kísérletek tervezését, számuk optimalizálását és szimulációját.

A tananyag feldolgozása során fontos a tapasztalatszerzés, a valószínűség szemléletes és absztrakt fogalmának helyes kialakítása. Változatos kombinatorikai feladatok megoldása, a kezdeti tapasztalatszerzések során a megszámlálhatóság, felsorolhatóság; később a jól leírható alapelvek(összegzés, szorzás), képletek adhatják a kezdeti háttérrel. A sejtések megfogalmazása a valószínűségekre végtelen halmazok esetén. A véletlen, a valószínűség nehéz, absztrakt fogalom, ezért fontos a tapasztalatok és a tényleges érték összevetése. A matematikai modellek segítségével, a valóság szimulálásával meg kell ismerni a módszereket, alkalmazhatóságukat és korlátaikat. Mit szabad és mit nem? Tekintsük a következő problémát: a táblázat alapján jól látható, hogy Anglia egy bizonyos térségében, a felsorolt hivatások mindegyikében, a nők sikeresebbek. Összességében mégis a férfiak tűnnek a sikeresebbnek. Lehetséges ez?

## BEVEZETÉS

---

### 1.4. táblázat: Simpson-féle paradoxon.

|              | Nők         |         |    | Férfiak     |         |    |
|--------------|-------------|---------|----|-------------|---------|----|
|              | Alkalmazott | Sikeres | %  | Alkalmazott | Sikeres | %  |
| Gazdász      | 240         | 63      | 26 | 512         | 112     | 22 |
| Informatikus | 26          | 7       | 27 | 228         | 58      | 25 |
| Mérnök       | 164         | 52      | 32 | 972         | 252     | 26 |
| Orvos        | 416         | 99      | 24 | 578         | 140     | 24 |
| Állatorvos   | 338         | 53      | 16 | 180         | 22      | 12 |
| Összesen     | 1184        | 274     | 23 | 2470        | 584     | 24 |

Röviden a tananyag felépítéséről: A 2. fejezetben elkészítjük a valószínűségi mezőt, azt a matematikai(absztrakt) modellt, amely alkalmas a véletlen kísérlet leírására, mérésére. Definiáljuk a klasszikus és a geometriai modellt és ezek néhány speciális esetét. A 3. fejezetben valós számot rendelünk a véletlen kísérlet lehetséges eredményeihez, majd ehhez a függvényeket jellemezzük valószínűségi törvényszerűségekkel illetve numerikus jellemzőkkel. A 4. fejezetben tovább visszük arra, amikor a lehetséges eredményeket több szám együttes megadásával (vektor) írhatjuk le. Az 5. fejezetben megadjuk a mintavételezés alapelveit, és a mintából számított néhány alapstatisztikát. A 6. és a 7. fejezet a minta feldolgozásának két nagy területéről szól. Először a minta alapján valószínűségeket, illetve valószínűségi törvényszerűségek ismeretlen paraméterei készitünk becsléseket. Ezután a 7. fejezetben a feltevéseink (hipotéziseink) vizsgálatát végezzük el. A 8. fejezet tartalmazza a szimulációs technikák elsajátításához szükséges minimális alapokat. A 9. fejezet a Függelék, amelyben szerepelnek a tananyag elsajátításához a jelölések, néhány közvetlenül kapcsolódó matematikai terület alapfogalmai, állításai. Továbbá néhány táblázat, hogy számítógépes háttér nélkül is oktathatóak legyenek a módszerek.

## 2. A VALÓSZÍNŰSÉG FOGALMA

### 2.1. Eseménytér, műveletek eseményekkel.

**Definíció:** Egy véletlen kísérlet lehetséges eredményeinek összességét *eseménytér*-nek (mintatér) nevezzük. Jele:  $\Omega$ . Az  $\Omega$  elemeit *elemi eseményeknek* nevezzük.

**Definíció:** Az  $\Omega$  részhalmazainak egy  $\mathcal{F}$  rendszerét  $\sigma$ -*algebrának* nevezzük, ha

- (1)  $\Omega \in \mathcal{F}$ ,
- (2)  $A \in \mathcal{F}$ , akkor  $\overline{A} \in \mathcal{F}$ ,
- (3)  $A, B \in \mathcal{F}$ , akkor  $A \cup B \in \mathcal{F}$ ,
- (4)  $A_1, A_2, \dots \in \mathcal{F}$ , akkor  $A_1 \cup A_2 \cup \dots \in \mathcal{F}$ .

Az  $\mathcal{F}$  elemeit pedig *eseményeknek* nevezzük.

**Megjegyzés:** 1. Ha csak (1), (2), (3) teljesül, akkor az  $\mathcal{F}$  halmazrendszert *algebrának* nevezzük.

2. Ha  $A, B \in \mathcal{F}$ , akkor  $A \cap B \in \mathcal{F}$ .

**2.1. Példa:** (a)  $\mathcal{F} = \{\Omega, \emptyset\}$  mindig  $\sigma$ -algebra.

(b) Legyen  $\Omega = \{1, 2, 3, 4, 5, 6, \}$ , ekkor  $\mathcal{F} = \{\Omega, \emptyset, \{5\}, \{1, 2, 3, 4, 6\}\}$   $\sigma$ -algebra.

(c) Ha  $\Omega = \{1, 2, 3, 4, 5, 6, \}$  és  $\{i\} \in \mathcal{F}$ , ( $i \in \Omega$ ) azaz minden egy elemű halmaz esemény, akkor a  $\sigma$ -algebra tulajdonságai szerint minden részhalmaz esemény. Tehát

$$\mathcal{F} = \mathcal{P}(\Omega) = 2^\Omega. \quad (2.1)$$

Ez véges esetben általában így van.  $\square$

**Definíció:** Az  $\Omega$  halmazt szokás *biztos eseménynek*, az  $\emptyset$  halmazt pedig *lehetetlen eseménynek* nevezni. Továbbá, az  $A$  esemény *bekövetkezik*, ha a kísérlet eredménye eleme az  $A$  halmaznak.

**Megjegyzés:** Az  $A \cup B$  esemény bekövetkezik, ha legalább az egyik közülük bekövetkezik, míg az  $A \cap B$  esemény akkor következik be, ha mind a kettő bekövetkezik.

**Definíció:** A  $P : \mathcal{F} \rightarrow \mathbf{R}$  nemnegatív leképezést valószínűségnek nevezzük, ha

$$(1) P(\Omega) = 1,$$

$$(2) A \cap B = \emptyset, \text{ akkor } P(A \cup B) = P(A) + P(B),$$

(3)  $A_1, A_2, \dots$  egymást kölcsönösen kizáró események  
(azaz  $A_i \cap A_j = \emptyset$ , ha  $i < j$  és  $i, j = 1, 2, \dots$ ), akkor

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \quad (2.2)$$

**Megjegyzés:** 1. Az (1)-(3) tulajdonságokat szokás a valószínűség axiómáinak nevezni.

2. Véges  $\Omega$  esetén a (3), az ún.  $\sigma$ -additivitás, nem szükséges.

**Következmény:** (1)  $P(\overline{A}) = 1 - P(A)$ .

$$(2) P(\emptyset) = 0.$$

$$(3) P(B \setminus A) = P(B) - P(A \cap B).$$

$$(4) P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

$$(5) \text{ Ha } A \subset B, \text{ akkor } P(A) \leq P(B).$$

$$(6) \text{ Ha } B_{n+1} \subset B_n \text{ és } \bigcap_{i=1}^{\infty} B_n = \emptyset, \text{ akkor } \lim_{n \rightarrow \infty} P(B_n) = 0.$$

**Bizonyítás:** (1)  $1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$ , hiszen  $A \cap \overline{A} = \emptyset$ , és így alkalmazható a 2. axióma.

(2) Alkalmazzuk (1)-et.  $\overline{\Omega} = \emptyset$ .

(3) Mivel  $B \setminus A = B \cap \overline{A}$ , és  $B = (B \cap \overline{A}) \cup (B \cap A)$ , ezért

$$P(B) = P((B \cap \overline{A}) \cup (B \cap A)) = P(B \cap \overline{A}) + P(B \cap A). \quad (2.3)$$

(4) Mivel  $A \cup B = A \cup (B \setminus A)$ , ezért a 2. axióma és a 3. következmény alapján adódik az állítás.

(5) Ha  $A \subset B$ , akkor  $A \cap B = A$ , így a 3. következmény alapján

$$0 \leq P(B \setminus A) = P(B) - P(A). \quad (2.4)$$

Itt kihasználtuk, hogy a valószínűség nemnegatív.

(6) Legyen  $A_n = B_n \setminus B_{n+1}$  ( $n \in \mathbf{N}$ ), ekkor  $A_1, A_2, \dots$  egymást kölcsönösen kizáró események. Tehát alkalmazható a 3. axióma.

$$\begin{aligned} P(B_1) &= P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \\ &= \sum_{i=1}^{n-1} P(A_i) + \sum_{i=n}^{\infty} P(A_i) \\ &= \sum_{i=1}^{n-1} P(A_i) + P(B_n) \end{aligned} \quad (2.5)$$

A sor konvergenciájának definíciójából

$$\lim_{n \rightarrow \infty} \sum_{i=n}^{\infty} P(A_i) = 0, \quad \text{azaz} \quad \lim_{n \rightarrow \infty} P(B_n) = 0. \quad (2.6)$$

**Q.E.D.**

**Megjegyzés:** 1. Az 5. következményt szokás a valószínűség monotonitásának is nevezni. Ennek fontos következménye, hogy ha  $A \in \mathcal{F}$ , akkor

$$0 \leq P(A) \leq 1, \quad (2.7)$$

mert  $\emptyset \subset A \subset \Omega$ .

2. Hasonlóan a 6. következmény a valószínűség folytonossága.

**Definíció:** Az  $(\Omega, \mathcal{F}, P)$  hármast *valószínűségi mezőnek* nevezzük.

*2.1. TÉTEL:* Adott az  $\Omega$  alaphalmaz, és az  $\mathcal{F}$   $\sigma$ -algebra. Továbbá, teljesül, hogy  $P : \mathcal{F} \rightarrow [0, 1]$ ,  $P(\Omega) = 1$ , és ha  $A \cap B = \emptyset$ , akkor  $P(A \cup B) = P(A) + P(B)$ , ekkor a következő állítások ekvivalensek:

(1)  $P$   $\sigma$ -additív, azaz teljesül a 3. axióma.

(2)  $P$  felülről folytonos, azaz ha  $A_n \subset A_{n+1}$ , akkor

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcup_{i=1}^{\infty} A_i\right); \quad (2.8)$$

(3)  $P$  alulról folytonos, azaz ha  $A_{n+1} \subset A_n$ , akkor

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{i=1}^{\infty} A_i\right); \quad (2.9)$$

(4)  $P$  folytonos a 0-ban, azaz teljesül a 6. következmény.

**Bizonyítás:** *(vázlat)* (1)  $\Rightarrow$  (2) :

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup \bigcup_{i=1}^{\infty} (A_{i+1} \setminus A_i). \quad (2.10)$$

(2)  $\Rightarrow$  (3) :

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - P\left(\bigcup_{i=1}^{\infty} \overline{A_i}\right) = 1 - \lim_{n \rightarrow \infty} P(\overline{A_n}) = \lim_{n \rightarrow \infty} P(A_n). \quad (2.11)$$

(3)  $\Rightarrow$  (4) :

$$\bigcap_{i=1}^{\infty} A_i = \emptyset. \quad (2.12)$$

(4)  $\Rightarrow$  (1) : Legyenek  $A_1, A_2, \dots$  egymást kölcsönösen kizáró események és  $B_n = \bigcup_{i=n+1}^{\infty} A_i$ , ekkor  $B_{n+1} \subset B_n$  és  $\bigcap_{i=1}^{\infty} B_n = \emptyset$ , azaz  $\lim_{n \rightarrow \infty} P(B_n) = 0$ . Viszont

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^n A_i\right) + P\left(\bigcup_{i=n+1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^n A_i\right) + P(B_n). \quad (2.13)$$

**Q.E.D.**

*2.2. TÉTEL: (Poincaré)* Az  $A_1, A_2, \dots, A_n$  eseményekre

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{i_1 < i_2 < \dots < i_k} P\left(\bigcap_{j=1}^k A_{i_j}\right), \quad (2.14)$$

ahol az összegzést az összes lehetséges  $\{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots, n\}$  esetre tekintjük.

**Megjegyzés:** A formula a 4. következmény általánosítása. Teljes indukcióval könnyen bizonyítható.

## 2.2. Klasszikus valószínűségi mező.

**Definíció:** Ha az elemi események száma véges és valószínűségük megegyezik, akkor a valószínűségi mezőt *klasszikusnak* nevezzük.

**Megjegyzés:** 1. A definíció nagyon rövidnek tűnik, ha arra gondolunk, hogy egy speciális helyzetben megadja a teljes matematikai modellt (a valószínűségi mezőt). Felmerül a kérdés, hogy a modell minden része szerepel-e benne. A válasz igen. Ha az elemi eseményeknek van valószínűsége azt úgy kell értelmezni, hogy az alaphalmaz minden egy elemű részhalmaza esemény. Ekkor viszont

$$\mathcal{F} = 2^{\Omega}, \quad (2.15)$$

azaz  $\mathcal{F}$  a hatványhalmaz.

2. Legyen  $|\Omega| = n$  és jelölje az elemi eseményeket  $\omega_i$  ( $i = 1, 2, \dots, n$ ). Ekkor

$$1 = P(\Omega) = P\left(\bigcup_{i=1}^n \{\omega_i\}\right) = \sum_{i=1}^n P(\{\omega_i\}) = nP(\{\omega_i\}). \quad (2.16)$$

Tehát  $P(\{\omega_i\}) = \frac{1}{n}$  ( $i = 1, 2, \dots, n$ ).

3. Legyen  $A \subset \Omega$  tetszőleges, ekkor felírható

$$A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\} \quad (2.17)$$

alakban. Ekkor

$$P(A) = P\left(\bigcup_{j=1}^k \{\omega_{i_j}\}\right) = \sum_{j=1}^k P(\{\omega_{i_j}\}) = kP(\{\omega_i\}) = \frac{|A|}{|\Omega|}. \quad (2.18)$$

Ezzel minden részhalmaznak meghatároztuk a valószínűségét. Tehát az ún. klasszikus képlet:

$$\text{valószínűség} = \frac{\text{kedvező esetek száma}}{\text{összes esetek száma}}. \square \quad (2.19)$$

**2.2. Példa:** Dobjunk fel egy dobókockát kétszer. Mennyi a valószínűsége, hogy legalább egy hatost dobunk!

**Megoldás:** A két dobás eredménye egy számpár, azaz

$$\Omega = \{(i, j) | 1 \leq i, j \leq 6\}, \quad |\Omega| = 36. \quad (2.20)$$

Jelölje  $A$  a legalább egy hatost, ekkor

$$A = \{(i, 6) | 1 \leq i \leq 6\} \cup \{(6, j) | 1 \leq j \leq 6\}, \quad |A| = 11. \quad (2.21)$$

Tehát a keresett valószínűség

$$P(A) = \frac{|A|}{|\Omega|} = \frac{11}{36}. \square \quad (2.22)$$

**VISSZATEVÉSES MINTAVÉTEL:** Adott  $N$  darab különböző objektum, amelyek közül  $s$  darab rendelkezik egy bizonyos tulajdonsággal, például selejt. Visszatevéssel kiveszünk  $n$  darabot. Legyen a kivett selejtek száma  $\xi$ . Mennyi a valószínűsége, hogy  $\xi = k$ , ahol  $0 \leq k \leq n$ .

$$p_k = P(\xi = k) = \frac{\binom{n}{k} s^k (N-s)^{n-k}}{N^n}. \quad (2.23)$$

Legyen  $p = \frac{s}{N}$ , akkor

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}. \quad (2.24)$$

Tehát csak a selejtaránytól függ a valószínűség.  $\square$

**2.3. Példa:** Legyen egy dobozban 3 piros és 7 kék golyó. Visszatevéssel húzunk 100-szor. Mennyi a valószínűsége, hogy a piros húzások száma( $\xi$ ) (a)  $28 \leq \xi \leq 33$ , (b)  $20 \leq \xi \leq 40$ .

**Megoldás:** (a)

$$P(28 \leq \xi \leq 33) = P(\xi = 28) + P(\xi = 29) + \dots + P(\xi = 33). \quad (2.25)$$

$$P(28 \leq \xi \leq 33) = \sum_{k=28}^{33} P(\xi = k) = \sum_{k=28}^{33} \binom{100}{k} 0.3^k (1-0.3)^{100-k}. \quad (2.26)$$

$$\sum_{k=28}^{33} \binom{100}{k} 0.3^k (1-0.3)^{100-k} \approx 0.4828916007. \quad (2.27)$$

(b)

$$P(20 \leq \xi \leq 40) = \sum_{k=20}^{40} \binom{100}{k} 0.3^k (1-0.3)^{100-k} \approx 0.9786143845. \quad (2.28)$$

**Megjegyzés:** Legyen a húzások száma 10000.

$$P(2850 \leq \xi \leq 3200) = \sum_{k=2850}^{3200} \binom{10000}{k} 0.3^k (1-0.3)^{10000-k} \approx 0.9995078261. \quad (2.29)$$

**VISSZATEVÉS NÉLKÜLI MINTAVÉTEL:** Adott  $N$  darab különböző objektum, amelyek közül  $s$  darab rendelkezik egy bizonyos tulajdonsággal, például selejt. Visszatevés nélkül kivesszünk  $n$  darabot. Legyen a kivett selejtek száma  $\xi$ . Mennyi a valószínűsége, hogy  $\xi = k$ , ahol  $0 \leq k \leq \min\{n, s\}$ .

$$p_k = P(\xi = k) = \frac{\binom{s}{k} \binom{N-s}{n-k}}{\binom{N}{n}}. \quad (2.30)$$

**Megjegyzés:** 1. Az  $n$  elemű sokaságból

$$n^k \quad (2.31)$$

számú visszatevéses és

$$n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!} \quad (2.32)$$

visszatevés nélküli  $k$  elemű minta vehető.

2. A  $p_k$  valószínűségek definíciójából következik, hogy

$$p_0 + p_1 + \dots + p_n = 1, \quad (2.33)$$

amelyből

$$\sum_{k=0}^n \binom{n}{k} s^k (N-s)^{n-k} = N^n, \quad (2.34)$$

illetve

$$\binom{s}{0} \binom{N-s}{n} + \binom{s}{1} \binom{N-s}{n-1} + \dots + \binom{s}{n} \binom{N-s}{0} = \binom{N}{n}. \quad (2.35)$$

Legyen a selejtek, a nem selejtek és a kiválasztottak száma is  $n$ , ekkor az előző azonosság

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \dots + \binom{n}{n}^2 = \binom{2n}{n}. \quad (2.36)$$

## 2. FEJEZET

---

**2.1. FELADAT:**  $n$  elemű sokaságból visszatevéssel  $k$  elemű mintát veszünk. Keressük meg annak az eseménynek a valószínűségét, hogy a mintában egyetlen elem sem fordul elő kétszer!

### Megoldás:

Kedvező esetek: visszatevés nélküli mintavétellel is megkaphatnánk. A keresett valószínűség:

$$p = \frac{n(n-1) \dots (n-k+1)}{n^k}. \quad \square \quad (2.37)$$

A következő két példa és a közelítések azt mutatják, hogy a valószínűségekre vonatkozó érzéseink nagyon megbízhatatlanok. Sőt nagy számok esetében az érzéseink gyakorlatilag használhatatlanok.

**2.4. Példa: (születésnap)** Egy teremben  $k$  személy tartózkodik. Mennyi a valószínűsége, hogy közülük legalább kettő születésnapja megegyezik?

$$p = 1 - \frac{365 \cdot 364 \cdot \dots \cdot (365 - k + 1)}{365^k}. \quad (2.38)$$

### 2.1. táblázat: Születésnap.

| $k$ | $1 - p$ |
|-----|---------|
| 5   | 0.97286 |
| 10  | 0.88305 |
| 15  | 0.74710 |
| 20  | 0.58856 |
| 23  | 0.49270 |
| 30  | 0.29368 |
| 40  | 0.10877 |
| 88  | 0.00001 |

**2.5. Példa: (Lottó 90-ből 5)** A lottón  $k$  alkalommal húztak ki számötöst. Mennyi a valószínűsége, hogy közülük legalább kettő számötös megegyezik?

**Megoldás:**

$$p = 1 - \frac{n(n-1) \dots (n-k+1)}{n^k}, \quad (2.39)$$

ahol

$$n = \binom{90}{5} = 43949268.$$

## 2.2. táblázat: Lottó 90-ből 5.

| $k$   | $1 - p$ |
|-------|---------|
| 500   | 0.99717 |
| 1000  | 0.98870 |
| 2000  | 0.95553 |
| 3000  | 0.90271 |
| 4000  | 0.83361 |
| 5000  | 0.75249 |
| 7806  | 0.49998 |
| 10000 | 0.32057 |
| 20000 | 0.01055 |
| 31800 | 0.00001 |

**Megjegyzés:** Jól látható, hogy  $k \ll n$ . 1957-től kb. 3000 húzás történt számunkra mégis furcsa, hogy eddig nem fordultak elő ugyanazok a számok.  $\square$

## 2.3. Geometriai valószínűségi mező.

A geometriai valószínűségi mező bevezetése, a valószínűség definíciója a klasszikus valószínűségi mező analógiájára történik. Bevezetése, alkalmazása során kiderül, hogy a szükséges elméleti alapokat majd csak a valószínűségi változóknál illetve a véletlen vektoroknál definiáljuk.

## 2. FEJEZET

---

A következő definíciót fogadjuk el a szemlélet alapján a klasszikus valószínűségi mező mintájára.

**Definíció:** Legyen  $\Omega \subset \mathbf{R}^n$ , amelynek létezik és véges a nagysága (jelölje  $m(\Omega)$ ). Továbbá legyen  $\Omega$  minden eleme (pontja) azonos "esélyű" és  $A \subset \Omega$ , amelynek szintén létezik az  $m(A)$  nagysága. A

$$P(A) = \frac{m(A)}{m(\Omega)} \quad (2.40)$$

menyiséget az  $A$  valószínűségének nevezzük.

**Megjegyzés:** 1.

$$P(A) = \frac{m(\text{kedvező esetek})}{m(\text{összes eset})}. \quad (2.41)$$

2. Egy halmaz nagyságán a hosszát, területét, térfogatát (mértékét) értjük.

3. Legyen  $\Omega = [0, 1]$  és  $m$  pedig a hosszúság, ekkor minden  $Q \in [0, 1]$  pontra csak az  $m(\{Q\}) = 0$  lehetséges. Ebből rögtön következik, hogy minden legfeljebb megszámlálhatóan végtelen halmaz nagysága (hossza) 0.

4. Létezik halmaz, amelynek nincs Lebesgue-mértéke.

*Nem mérhető halmaz konstrukciója:* Legyen  $\Omega = [0, 1]$  és  $m$  pedig a hosszúság. Az  $a, b \in \Omega$  relációban van, ha  $a - b \in \mathbf{Q}$ , azaz racionális. Ez a reláció reflexív, szimmetrikus, tranzitív. Tehát ekvivalenciareláció, amely osztályozást hoz létre. Defináljuk az  $E$  halmazt oly módon, hogy minden osztályból kiveszünk egy elemet. Ez lehetséges a halmazelmélet kiválasztási axiómája szerint. Legyen

$$\Omega \cap \mathbf{Q} = \{r_1, r_2, \dots\}, \quad (2.42)$$

$$E_n := \{x + r_n - [x + r_n] | x \in E\}, \quad (2.43)$$

ekkor az  $E_n$  halmazok páronént diszjunktak és  $\bigcup_{n=1}^{\infty} E_n = \Omega$ . Ha  $E$  mérhető, akkor  $E_n$  is és nagyságuk megegyezik. Továbbá

$$\sum_{n=1}^{\infty} m(E_n) = 1, \quad (2.44)$$

ami lehetetlen, mert a sor tagjai mind egyenlőek. Ez azt jelenti, hogy  $E$  nem mérhető.

5. Létezik kontinuum számosságú halmaz, amelynek 0 a Lebesgue-mértéke.

A *Cantor-féle triadikus halmaz*: Legyen  $E_1$  a középső része a  $[0, 1]$  intervallumnak, azaz  $E_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$ . Tehát  $x \in [0, 1] \setminus E_1$  akkor és csak akkor, ha hármas számrendszerben az első jegy (a 0 után) a 0 vagy a 2. Legyen  $E_2$  a középső részek uniója a  $[0, 1] \setminus E_1$  halmazból, azaz  $E_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$ . Tehát  $x \in [0, 1] \setminus (E_1 \cup E_2)$  akkor és csak akkor, ha hármas számrendszerben az első két jegy (a 0 után) a 0 vagy a 2. Folytassuk a konstrukciót: legyen  $E_n$  a középső részek uniója a  $[0, 1] \setminus (E_1 \cup E_2 \cup \dots \cup E_{n-1})$  halmazból. Cantor-féle triadikus halmaznak nevezzük a

$$C = [0, 1] \setminus \bigcup_{n=1}^{\infty} E_n \quad (2.45)$$

halmazt. Tehát  $x \in C$  akkor és csak akkor, ha hármas számrendszerben a számjegyei csupán a 0 vagy a 2.

A  $C$  halmaz nemmegszámlálható. A konstrukció alapján

$$m(E_n) = \left(\frac{1}{3}\right) \left(\frac{2}{3}\right)^{n-1}, \quad m(C) = 1 - \sum_{n=1}^{\infty} \left(\frac{1}{3}\right) \left(\frac{2}{3}\right)^{n-1} = 0. \quad \square \quad (2.46)$$

**2.6. Példa:** Válasszunk két számot "véletlenül" a  $[0, 5]$  intervallumból. Legyenek ezek  $b$  és  $c$ . Mennyi a valószínűsége, hogy az

$$x^2 + bx + c = 0 \quad (2.47)$$

másodfokú egyenletnek nincs valós gyöke?

**Megoldás:** Az egyenletnek nincs valós gyöke, ha a diszkriminánsa negatív, azaz  $b^2 - 4c < 0$ . Az alaphalmazunk

$$\Omega = \{(b, c) | 0 \leq b, c \leq 5\}, \quad (2.48)$$

## 2. FEJEZET

---

amely egy 5 egységnyi oldalú négyzet a  $(b, c)$  koordinátarendszerben. A kedvező eseteket a négyzet azon része adja, amikor

$$0 \leq c < \frac{b^2}{4}, \quad (2.49)$$

azaz a parabola alatti rész. Jelölje ezt  $A$ .

$$m(\Omega) = 25, \quad m(A) = \int_0^5 \frac{b^2}{4} db = \left[ \frac{b^3}{12} \right]_0^5 = \frac{125}{12}. \quad (2.50)$$

Tehát

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{\frac{125}{12}}{25} = \frac{5}{12}. \quad \square \quad (2.51)$$

**2.2. FELADAT: (Bertrand paradoxon)** Határozzuk meg annak a valószínűségét, hogy a kör egy tetszőlegesen vett húrja nagyobb legyen, mint a körbe írt egyenlő oldalú háromszög oldala ( $C$  esemény)!

**1. megoldás:** A húr helyzete meghatározott, ha megadjuk a közepét. Tudjuk, hogy az  $R$  sugarú körbe írt egyenlő oldalú háromszög oldalának közepe  $\frac{R}{2}$  távolságra van a kör középpontjától. Tehát ha a húr közepe az adott körrel koncentrikus  $\frac{R}{2}$  sugarú kör belsejében van, akkor a körbe írt egyenlő oldalú háromszög oldalánál nagyobb húrunk lesz és így

$$P(C) = \frac{\pi \left( \frac{R}{2} \right)^2}{\pi R^2} = \frac{1}{4}. \quad (2.52)$$

**2. megoldás:** Ha  $A$  és  $B$  a húr végpontjai, akkor ahhoz, hogy az  $AB$  szakasz nagyobb legyen, mint a körbe írt egyenlő oldalú háromszög oldala, az szükséges, hogy ha felvesszünk még két más  $A_1, A_2$  pontot, melyek az  $A$  ponttal együtt a kör területét három egyenlő részre osztják, a  $B$  pont az  $A_1 A_2$  íven helyezkedjék el. A keresett valószínűség tehát

$$P(C) = \frac{1}{3}. \quad (2.53)$$

**3. megoldás:** Rögzítjük a húr irányát, és merőlegesen rá egy átmérőt húzunk, melyet négy egyenlő részre osztunk (a negyedelő pontok legyenek  $E, F$ ). A húr nagyobb, mint a körbe írt egyenlő oldalú háromszög oldala, ha a közepe az  $E$  és  $F$  pontok között helyezkedik el. Tehát

$$P(C) = \frac{1}{2}. \square \quad (2.54)$$

**Megjegyzés:** Három különböző eredményt kaptunk, mert mindegyik esetben más mértéket használtunk. Az utolsó esetben invariáns a mérték az ortogonális transzformációcsoportra.

## 2.4. Feltételes valószínűség.

**2.7. Példa:** Dobjunk fel egy dobókockát kétszer. Jelölje  $A, B$  azt az eseményt, hogy a dobások összpontszáma 6 illetve 7. Legyen  $C$  az az esemény, hogy az első dobás páros szám. Ekkor

$$\Omega = \{(i, j) | 1 \leq i, j \leq 6\}, \quad |\Omega| = 36. \quad (2.55)$$

$$P(A) = \frac{5}{36}, \quad \text{mert} \quad A = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}.$$

$$P(B) = \frac{6}{36}, \quad \text{mert} \quad B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}.$$

$$P(C) = \frac{18}{36}, \quad \text{mert} \quad C = \{(i, j) | i = 2, 4, 6, 1 \leq j \leq 6\}. \quad (2.56)$$

$$P(A \cap B) = 0, \quad \text{mert} \quad A \cap B = \emptyset.$$

$$P(A \cap C) = \frac{2}{36}, \quad \text{mert} \quad A \cap C = \{(2, 4), (4, 2)\}.$$

$$P(B \cap C) = \frac{3}{36}, \quad \text{mert} \quad B \cap C = \{(2, 5), (4, 3), (6, 1)\}.$$

A továbbiaknál tételezzük fel, hogy  $C$  bekövetkezett, azaz az első dobás páros szám. Ekkor az összes eset csak a  $C$  halmazra korlátozódik.  $C$  az alaphalmaz. Ekkor

$$P(A, \text{ ha } C \text{ bekövetkezett}) = \frac{2}{18}, \quad \text{mert} \quad A \cap C = \{(2, 4), (4, 2)\}. \quad (2.57)$$

$$P(B, \text{ ha } C \text{ bekövetkezett}) = \frac{3}{18}, \quad \text{mert} \quad B \cap C = \{(2, 5), (4, 3), (6, 1)\}. \square$$

**Definíció:** Az  $A$  esemény  $B$  feltétel melletti *feltételes valószínűségének* nevezzük a

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.58)$$

mennyiséget, ha  $P(B) > 0$ .

**Megjegyzés:** A  $P(\cdot|B) : \mathcal{F} \rightarrow \mathbf{R}$  leképezés tényleg valószínűség.

**2.3. TÉTEL: (szorzási szabály)** Ha  $P(A) > 0$ ,  $P(B) > 0$ , akkor

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B). \quad (2.59)$$

**2.4. TÉTEL:** Ha az  $A_1, A_2, \dots, A_n$  eseményrendszerre  $P(\bigcap_{i=1}^{n-1} A_i) > 0$ , akkor

$$P(\bigcap_{i=1}^n A_i) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}). \quad (2.60)$$

**Definíció:** Az  $A_1, A_2, \dots$  eseményrendszert *teljes eseményrendszernek* nevezzük, ha  $A_i \cap A_j = \emptyset$ , ha  $i < j$  és  $i, j = 1, 2, \dots$ , és  $\bigcup_{i=1}^{\infty} A_i = \Omega$ .

**2.5. TÉTEL: (teljes valószínűség)** Ha  $A_1, A_2, \dots$  teljes eseményrendszer és  $P(A_i) > 0$ , ha  $i = 1, 2, \dots$ , akkor tetszőleges  $B$  esemény esetén

$$P(B) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i). \quad (2.61)$$

**Bizonyítás:**

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P(B \cap \bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{i=1}^{\infty} (B \cap A_i)) = \\ &= \sum_{i=1}^{\infty} P(B \cap A_i) = \sum_{i=1}^{\infty} P(A_i)P(B|A_i) \end{aligned} \quad (2.62)$$

Felhasználva a teljes eseményrendszer tulajdonságait, a valószínűség 3. axiómáját és a szorzási szabályt. **Q.E.D.**

**Megjegyzés:** 1.  $A$  és  $\bar{A}$  teljes eseményrendszert alkot.

2.  $A \cap B$ ,  $A \cap \bar{B}$ ,  $\bar{A} \cap B$ , és  $\bar{A} \cap \bar{B}$  teljes eseményrendszert alkot.

2.3. *FELADAT:* Bizonyítsuk be, hogy

$$\sum_{k=0}^n \binom{n+k}{k} \frac{1}{2^{n+k}} = 1. \quad (2.63)$$

**I. megoldás** (*kombinatorikus*):

Pascal-háromszög tulajdonság:

$$\binom{m+1}{l} = \binom{m}{l-1} + \binom{m}{l}. \quad (2.64)$$

Legyen

$$S_n = \sum_{k=0}^n \binom{n+k}{k} \frac{1}{2^k},$$

azaz az állítás, hogy

$$S_n = 2^n. \quad (2.65)$$

Teljes indukció szerint bizonyítunk.

$$n = 1, \quad S_1 = \binom{1}{0} \frac{1}{2^0} + \binom{2}{1} \frac{1}{2^1} = 2^1. \quad (2.66)$$

Tegyük fel, hogy  $n$ -re teljesül, hogy

$$S_n = 2^n. \quad (2.67)$$

$$S_{n+1} = \sum_{k=0}^{n+1} \binom{n+k+1}{k} \frac{1}{2^k} = 1 + \sum_{k=1}^{n+1} \left[ \binom{n+k}{k-1} + \binom{n+k}{k} \right] \frac{1}{2^k} \quad (2.68)$$

$$\begin{aligned}
 S_{n+1} &= \sum_{j=0}^n \binom{n+j+1}{j} \frac{1}{2^{j+1}} + \sum_{k=0}^{n+1} \binom{n+k}{k} \frac{1}{2^k} = \\
 &= \frac{1}{2} \sum_{j=0}^n \binom{n+j+1}{j} \frac{1}{2^j} + \binom{2n+1}{n+1} \frac{1}{2^{n+1}} + \sum_{k=0}^{n+1} \binom{n+k}{k} \frac{1}{2^k} = \\
 &= \frac{1}{2} \sum_{j=0}^{n+1} \binom{n+j+1}{j} \frac{1}{2^j} + \sum_{k=0}^{n+1} \binom{n+k}{k} \frac{1}{2^k} = \\
 &= \frac{1}{2} S_{n+1} + S_n.
 \end{aligned} \tag{2.69}$$

Hiszen

$$\binom{2n+1}{n+1} \frac{1}{2^{n+1}} = \frac{1}{2} \binom{2n+2}{n+1} \frac{1}{2^{n+1}}. \tag{2.70}$$

Tehát

$$S_{n+1} = 2S_n. \tag{2.71}$$

## II. megoldás (valószínűségi):

Fogalmazzuk át a feladatot! Egy érmével fej-írást játszunk. Addig dobáljuk, amíg a fejek vagy az írák száma eléri az  $(n+1)$ -et. Ehhez nyilván legalább  $(n+1)$  dobásra van szükségünk és  $(2n+1)$  biztosan elég. Határozzuk meg annak a valószínűségét, hogy a szükséges dobások száma  $(n+1+k)$ , ahol  $k = 0, 1, \dots, n$ !

Jelöljük ezt az eseményt  $A_k$ -val. Az  $A_k$  eseményt két részre bontjuk aszerint, hogy az utolsó dobás fej, illetve írás. Jelölje ezeket  $A_k^F$  illetve  $A_k^I$ .

$A_k^F$  akkor következik be, ha az  $(n+k+1)$ -edik dobás fej, előtte pedig  $n$  fej és  $k$  írás.

A sorrendek száma:

$$\binom{n+k}{k}. \tag{2.72}$$

Egy sorrend valószínűsége:

$$\frac{1}{2^{n+k+1}}, \tag{2.73}$$

így

$$P(A_k^F) = \binom{n+k}{k} \frac{1}{2^{n+k+1}} = P(A_k^F). \quad (2.74)$$

Tehát

$$P(A_k) = \binom{n+k}{k} \frac{1}{2^{n+k}}. \quad (2.75)$$

Mivel az  $A_k$  események teljes eseményrendszert alkotnak, valószínűségeik összege 1.  $\square$

**2.6. TÉTEL: (Bayes)** Ha  $A_1, A_2, \dots$  teljes eseményrendszer és  $P(A_i) > 0$ , ha  $i = 1, 2, \dots$ , akkor tetszőleges pozitív valószínűségű  $B$  esemény esetén

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{\infty} P(B|A_i)P(A_i)}. \quad (2.76)$$

**Bizonyítás:**

$$P(A_k|B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{\infty} P(B|A_i)P(A_i)} \quad (2.77)$$

Felhasználva a teljes valószínűség tételét és a szorzási szabályt. **Q.E.D.**

**Megjegyzés:** A Bayes-tételhez kapcsolódóan bevezethetjük a következő elnevezéseket:  $P(A_i)$  az ún. *a-priori* valószínűség és  $P(A_i|A)$  az ún. *a-posteriori* valószínűség.

**2.8. Példa:** Adott két doboz. Az elsőben 7 piros és 5 kék golyó van, míg a másodikban 4 piros és 6 kék. Valaki átrak, véletlenszerűen kiválasztott, két golyót az első dobozból a másodikba. Az eddigi információk birtokában véletlenszerűen kiveszünk egy golyót a második dobozból. Megállapítjuk, hogy ez piros. Mennyi a valószínűsége, hogy az elsőből két kéket raktak át?

**Megoldás:** A piros golyónkat egy két lépcsős kísérlet során kaptuk, amelynél az első lépésben nem tudjuk, hogy pontosan mi történt. Továbbá az eredmény birtokában ennek egy részére kell visszakövetkeztetni. Ha megvan a teljes eseményrendszer, akkor ez egy a Bayes-tétel segítségével könnyen megoldható feladat.

## 2. FEJEZET

---

Vezessük be a következő jelöléseket:  $A_i$  ( $i = 0, 1, 2$ ) jelentse azt, hogy az első lépésben  $i$  darab kék került át a második dobozba. Ekkor  $A_0, A_1, A_2$  teljes eseményrendszert alkot.  $B$  pedig jelölje azt, hogy a második dobozból pirosat húzunk. Tehát a keresett valószínűség  $P(A_2|B)$ . Viszont

$$\begin{aligned} P(A_0) &= \frac{\binom{7}{2} \binom{5}{0}}{\binom{12}{2}}, & P(B|A_0) &= \frac{6}{12}, \\ P(A_1) &= \frac{\binom{7}{1} \binom{5}{1}}{\binom{12}{2}}, & P(B|A_1) &= \frac{5}{12}, \\ P(A_2) &= \frac{\binom{7}{0} \binom{5}{2}}{\binom{12}{2}}, & P(B|A_2) &= \frac{4}{12}. \end{aligned} \tag{2.78}$$

$$P(B) = P(A_0)P(B|A_0) + P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = \frac{31}{72}. \tag{2.79}$$

$$P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(B)} = \frac{40}{341}. \square \tag{2.80}$$

**Megjegyzés:** Természetesen  $P(A_0|B)$ ,  $P(A_1|B)$  hasonlóan határozható meg.

A feltételes valószínűség definíciója előtti példában láttuk, hogy

$$P(B) = \frac{1}{6} = P(B|C), \tag{2.81}$$

azaz a valószínűség nem függ a feltételtől. Ekkor

$$P(B) = \frac{P(B \cap C)}{P(C)}, \tag{2.82}$$

vagyis

$$P(B)P(C) = P(B \cap C). \tag{2.83}$$

**Definíció:** Az  $A$  és  $B$  eseményt *sztochasztikusan függetlennek* nevezzük, ha

$$P(A \cap B) = P(A)P(B). \quad (2.84)$$

**Megjegyzés:** 1. Ha az  $A$  és  $B$  események függetlenek, akkor  $\overline{A}$  és  $B$ ,  $A$  és  $\overline{B}$  és  $\overline{A}$  és  $\overline{B}$  is függetlenek.

2. Ha  $0 < P(A) < 1$ , akkor  $A$  és  $\overline{A}$  nem függetlenek.

**2.7. TÉTEL:** Ha  $A \cap B = \emptyset$ , és  $P(A)P(B) > 0$ , akkor az  $A$  és a  $B$  esemény nem lehetnek függetlenek.

**Bizonyítás:**

$$P(A \cap B) = 0, \quad P(A)P(B) > 0. \quad (2.85)$$

Tehát nem lehetnek egyenlőek. **Q.E.D.**

**Definíció:** Az  $A_1, A_2, \dots, A_n$  eseményeket *páronként sztochasztikusan függetlennek* nevezzük, ha

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad (1 \leq i < j \leq n). \quad (2.86)$$

**Definíció:** Az  $A_1, A_2, \dots, A_n$  eseményeket *teljesen sztochasztikusan függetlennek* nevezzük, ha

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}), \quad (2.87)$$

ahol  $1 \leq i_1 < \dots < i_k \leq n$ ,  $2 \leq k \leq n$ .

**Megjegyzés:** Ha megvizsgáljuk a feltételrendszert, akkor látható, hogy a teljes függetlenség feltételeinek a száma

$$\binom{n}{2} + \binom{n}{3} + \dots + \binom{n}{n} = 2^n - \binom{n}{0} - \binom{n}{1} = 2^n - 1 - n, \quad (2.88)$$

amely nagyon gyorsan nő. Már  $n = 3$  esetén megadható példa, amely azt mutatja, hogy egyik feltétel sem elhagyható.

**2.9. Példa:** Legyen  $\Omega = \{1, 2, 3, 4\}$  és minden elemi esemény valószínűsége megegyezik. Legyen  $A = \{1, 2\}$ ,  $B = \{2, 3\}$ ,  $C = \{1, 3\}$ , ekkor

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

$$P(A \cap B) = \frac{1}{4} = P(A)P(B),$$

$$P(A \cap C) = \frac{1}{4} = P(A)P(C),$$

$$P(B \cap C) = \frac{1}{4} = P(B)P(C),$$

(2.89)

$$P(A \cap B \cap C) = 0, \quad P(A)P(B)P(C) = \frac{1}{8}.$$

Tehát páronként függetlenek, de teljesen nem.  $\square$

**2.10. Példa:** Legyen  $\Omega = \{1, 2, 3, 4, 5, 6\}$  és minden elemi esemény valószínűsége megegyezik. Legyen  $A = \{1, 2, 3, 4\}$ ,  $B = \{1, 2, 5\}$ ,  $C = \{1, 3, 6\}$ , ekkor

$$P(A) = \frac{2}{3}, \quad P(B) = P(C) = \frac{1}{2},$$

$$P(A \cap B) = \frac{1}{3} = P(A)P(B),$$

$$P(A \cap C) = \frac{1}{3} = P(A)P(C),$$

(2.90)

$$P(B \cap C) = \frac{1}{6}, \quad P(B)P(C) = \frac{1}{4},$$

$$P(A \cap B \cap C) = \frac{1}{6} = P(A)P(B)P(C).$$

Tehát sem páronként sem teljesen nem függetlenek.  $\square$

**Definíció:** Az  $\{A_1, A_2, \dots, A_n\}$  és  $\{B_1, B_2, \dots, B_m\}$  eseményrendszereket *sztochasztikusan függetlennek* nevezzük, ha  $\forall i, j$  esetén

$$P(A_i \cap B_j) = P(A_i)P(B_j) \quad (1 \leq i \leq n, \quad 1 \leq j \leq m). \quad (2.91)$$

**2.11. Példa:** Ha az  $A$  és  $B$  események függetlenek, akkor  $\overline{A}$  és  $B$ ,  $A$  és  $\overline{B}$  és  $\overline{A}$  és  $\overline{B}$  is függetlenek, azaz az  $\{A, \overline{A}\}$  és  $\{B, \overline{B}\}$  eseményrendszerek is függetlenek.

2.8. *TÉTEL:* Ha  $A_1, A_1, \dots, A_n$  független események és

$$P(A_i) < 1, \quad (i = 1, 2, \dots, n), \quad (2.92)$$

akkor

$$P\left(\bigcup_{i=1}^n A_i\right) < 1. \quad (2.93)$$

**Bizonyítás:**

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= P\left(\overline{\bigcap_{i=1}^n \overline{A_i}}\right) = 1 - P\left(\bigcap_{i=1}^n \overline{A_i}\right) = \\ &= 1 - P\left(\bigcap_{i=1}^n \overline{A_i}\right) = 1 - \prod_{i=1}^n P(\overline{A_i}). \end{aligned} \quad (2.94)$$

**Q.E.D.**

**Definíció:** *Bernoulli kísérletsorozatnak* nevezzük azt, ha adott  $A \in \mathcal{F}$  és egymástól függetlenül, azonos körülmények között elvégezzük ugyanazt a kísérletet, s "csak" azt figyeljük, hogy az  $A$  esemény bekövetkezett-e vagy sem.

**Megjegyzés:** A visszatevéses mintavétel egy ilyen kísérletsorozatot valósít meg.

2.4. *FELADAT:* Szinbádnak jogában áll  $N$  háremhölgy közül kiválasztani egyet oly módon, hogy az előtte egyenként elvonuló hölgyek valamelyikére rámutat. Tegyük fel, hogy egyértelmű (szigorúan monoton) szépségi sorrendet tud megállapítani az előtte elvonuló háremhölgyek között, és hogy a háremhölgyek bármely elvonulási sorrendje egyenlően valószínű. Szinbád  $k$  számú hölgyet elenged, majd a továbbiak közül kiválasztja az elsőt, aki szebb az összes elvonultaknál (nevezzük ezt  $k$ -stratégiának). Mennyi a valószínűsége, hogy a legszebb hölgyet választja ki? Milyen  $k$  esetén lesz a legnagyobb a valószínűsége annak, hogy a legszebb hölgyet választja ki, ha  $N$  elég nagy?

**Megoldás:** 1. Számozzuk meg a hölgyeket 1-től  $N$ -ig szépségi sorrendben (a legszebb az egyes). Jelölje  $A_k^{(N)}$  azt az eseményt, hogy a  $k$ -stratégia mellett a legszebbet választja Szinbád. Jelölje  $B_k^{(N)}(i)$  azt az eseményt, hogy az elengedett  $k$  hölgy

## 2. FEJEZET

---

közötti legszebb éppen az  $i$ -edik a szépségi sorrendben ( $i = 1, 2, \dots, N-k+1$ ). Nyilván

$$P(A_k^{(N)} | B_k^{(N)}(1)) = 0. \quad (2.95)$$

A teljes valószínűség tétele alapján

$$P(A_k^{(N)}) = \sum_{i=2}^{N-k+1} P(A_k^{(N)} | B_k^{(N)}(i)) P(B_k^{(N)}(i)). \quad (2.96)$$

Az  $1, 2, \dots, N$  számoknak tekintsük egy véletlen permutációját. Mennyi a valószínűsége, hogy az első  $k$  szám közötti legkisebb szám éppen  $i$ -vel egyenlő?

$$P(B_k^{(N)}(i)) = \frac{\binom{N-i}{k-1} k! (N-k)!}{N!} = \frac{\binom{N-i}{k-1}}{\binom{N}{k}}, \quad (2.97)$$

és

$$\begin{aligned} P(A_k^{(N)} | B_k^{(N)}(i)) &= \sum_{j=1}^{N-k-i+2} \frac{\binom{N-k-j}{i-2} (i-2)! (N-k-i+1)!}{(N-k)!} = \\ &= \frac{1}{i-1}. \end{aligned} \quad (2.98)$$

Tehát

$$P(A_k^{(N)}) = \sum_{i=2}^{N-k+1} \frac{\binom{N-i}{k-1}}{\binom{N}{k}} \frac{1}{i-1}. \quad (2.99)$$

2. Jelölje  $C_j^{(N)}$  azt az eseményt, hogy a legszebb hölgy a  $j$ -ediknek érkezik. Ezen feltevés mellett nyilván akkor és csak akkor következik be az  $A_k^{(N)}$  esemény, ha a legszebb előtt érkezők legszebbike az első  $k$  elvonuló között van.  $P(A_k^{(N)} | C_j^{(N)}) = 0$ , ha  $j \leq k$ . Azon sorrendek közül, amelyben a legszebb hölgy a  $j$ -ediknek érkezik, nevezzük jó sorrendnek azokat, amelyeknél a legszebb előtt érkezők legszebbike az első  $k$  között van. Vegyük észre, hogy annak a valószínűsége, hogy egy sorrend

jó legyen (azon feltevés mellett, hogy alegszebb  $j$ -ediknek érkezik) nem függ attól, hogy a többieket hogyan osztjuk szét – előtte és utána érkező – csoportra. A teljes valószínűség tétele alapján

$$P(A_k^{(N)}) = \sum_{j=k+1}^N P(A_k^{(N)} | C_j^{(N)}) P(C_j^{(N)}) = \sum_{j=k+1}^N \frac{k}{j-1} \frac{1}{N}. \quad (2.100)$$

**Megjegyzés:** A két megoldás egyezősége abból adódik, hogy a

$$g_y(x) = \frac{\ln \frac{1}{1-x}}{(1-x)^y} = \frac{d}{dy} \left( \frac{1}{(1-x)^y} \right) \quad (2.101)$$

mindkét oldalát sorbafejtjük, majd vesszük  $g_k(x)$  esetén az  $x^{N-k}$  együtthatóját.

### 2.3. táblázat: Szinbád ( $N = 100$ ).

| $k$ | $P(A_k^{(N)})$ |
|-----|----------------|
| 34  | 0.3701169592   |
| 35  | 0.3707086345   |
| 36  | 0.3710145955   |
| 37  | 0.3710427787   |
| 38  | 0.3708006916   |
| 39  | 0.3702954467   |

*Fogalmazzuk át a feladatot:* Valaki kiválaszt egy  $N$  hosszúságú, tetszőleges természetes számokból álló sorozatot. Ezeket véletlenszerűen megmutatja (egyenként) egy játékosnak. A játékos feladata a legnagyobb szám eltalálása. A stratégia:  $p$ -ed részét elengedi, majd ezután kiválasztja a legnagyobbat. Milyen  $p$  esetén a legnagyobb az esélye?

A számok sorszámai 1 és  $N$  között van. A véges sok egész számra vonatkozó problémát alakítsuk át egy kontinuum számosságú valós számokra vonatkozó fel-

dattá. Az  $[1, N]$  intervallumot képezzük le a  $[0, 1]$  intervallumra. A játékos stratégiájának sikeréhez a sorozatnak a következő tulajdonságokkal kell rendelkeznie: Mindenekelőtt a legnagyobb számnak a  $(p, 1)$  intervallumban kell lennie. Ha a maximum helyét  $x$ -szel jelöljük, akkor az alábbi esetek valamelyikének be kell következnie: A második, a harmadik,  $\dots$ , az  $i$ -edik legnagyobb szám a  $(x, 1)$  intervallumba esik az  $(i + 1)$ -edik legnagyobb szám a  $(0, p)$  intervallumba esik ( $i = 0, 1, 2, \dots$ ). Ennek a valószínűsége

$$p(1 - x)^i. \quad (2.102)$$

Tehát, ha a maximum  $x$ -nél van, akkor a siker valószínűsége

$$\sum_{i=0}^{\infty} p(1 - x)^i = \frac{p}{x}. \quad (2.103)$$

Ez minden  $(p, 1)$  intervallumbeli  $x$ -re fennáll, így a siker valószínűsége

$$p \int_p^1 \frac{dx}{x} = p \ln \frac{1}{p}. \quad (2.104)$$

Ennek a maximuma pedig  $p = \frac{1}{e} \approx 0.3678794412$  esetén van.  $\square$

## 2.5. Relatív gyakoriság.

**Definíció:** Adott egy valószínűségi mező. Vizsgáljuk az  $A$  esemény bekövetkezését. Végezzünk el egy Bernoulli-kísérletsorozatot, amelynek a hossza  $n$ . Jelölje az  $A$  esemény bekövetkezéseinek a számát  $k_A$ . Ezt az  $A$  esemény *gyakoriságának* nevezzük. Míg az

$$r_A = \frac{k_A}{n} \quad (2.105)$$

mennyiséget pedig *relatív gyakoriságnak* nevezzük.

Mivel  $0 \leq k_A \leq n$ , ezért  $0 \leq r_A \leq 1$ .

$k_{\Omega} = n$ , tehát  $r_{\Omega} = 1$ .

## A VALÓSZÍNŰSÉG FOGALMA

---

Ha  $A \cap B = \emptyset$ , akkor  $k_{A \cup B} = k_A + k_B$ , ezért  $r_{A \cup B} = r_A + r_B$ .

**Megjegyzés:** Jól látható, hogy a relatív gyakoriság tulajdonságai megegyeznek a valószínűségével és mégsem jó igazi mérőszámnak.

### 3. A VALÓSZÍNŰSÉGI VÁLTOZÓ

#### 3.1. Valószínűségi változó.

**Definíció:** A  $\xi : \Omega \rightarrow \mathbf{R}$  leképezést *valószínűségi változónak* nevezzük, ha

$$\{\xi < x\} = \{\omega | \omega \in \Omega, \quad \xi(\omega) < x\} \in \mathcal{F} \quad \forall x \in \mathbf{R}. \quad (3.1)$$

**Definíció:** Az  $F(x) = P(\xi < x)$  formulával meghatározott valós függvényt a  $\xi$  valószínűségi változó *eloszlásfüggvényének* nevezzük.

**3.1. TÉTEL:** Az  $F$  valós függvény akkor és csak akkor lehet eloszlásfüggvény, ha

1.  $\lim_{x \rightarrow -\infty} F(x) = 0$ ,
2.  $\lim_{x \rightarrow \infty} F(x) = 1$ ,
3.  $F(a) \leq F(b)$ , ha  $(a < b)$ , azaz monoton növekvő,
4.  $\lim_{x \rightarrow x_0-0} F(x) = F(x_0)$ ,  $\forall x_0 \in \mathbf{R}$ , azaz balról folytonos.

**Bizonyítás:** Legyen  $A = \{\xi < a\}$  és  $B = \{\xi < b\}$ , ekkor  $A \subset B$ . Tehát a valószínűség definíciójának 5. következménye szerint  $F(a) = P(A) \leq P(B) = F(b)$ .

Legyen az  $x_1, x_2, \dots$  sorozat monoton csökkenő és  $\lim_{n \rightarrow \infty} x_n = -\infty$ , azaz nem korlátos. Továbbá  $A_n = \{\xi < x_n\}$ , ekkor  $F(x_n) = P(A_n)$ ,  $A_{n+1} \subset A_n$ , és  $\bigcap_{n=1}^{\infty} A_n = \emptyset$ . Tehát

$$\lim_{n \rightarrow \infty} P(A_n) = 0. \quad (3.2)$$

Hasonlóan legyen az  $x_1, x_2, \dots$  sorozat monoton növekvő és  $\lim_{n \rightarrow \infty} x_n = \infty$ ,  $A_n =$

$\{\xi < x_n\}$ , ekkor  $F(x_n) = P(A_n)$ ,  $A_n \subset A_{n+1}$ , és  $\bigcup_{n=1}^{\infty} A_n = \Omega$ . Tehát

$$\lim_{n \rightarrow \infty} P(A_n) = 1. \quad (3.3)$$

A balról folytonosság hasonlóan adódik. Tehát egy eloszlásfüggvény mindig teljesíti a négy tulajdonságot.

A megfordítás: Ha teljesül az 1-4. tulajdonság, akkor létezik valószínűségi mező, és azon olyan valószínűségi változó, hogy ennek eloszlásfüggvénye  $F$ .

Erre csak részleges bizonyítást adunk, amikor  $F$  szigorúan nő és folytonos ( $\exists F^{-1}$ ).

Legyen  $\Omega = [0, 1]$ ,  $\mathcal{F}$  a nyílt intervallumok által generált  $\sigma$ -algebra és  $P$  pedig egy halmaz hossza. Legyen minden  $\omega \in \Omega$  esetén

$$\xi(\omega) = F^{-1}(\omega), \quad (3.4)$$

ami folytonos és szigorúan monoton növekvő.

$$P(\xi < x) = m(\{\omega | F^{-1}(\omega) < x\}) = m(\{\omega | \omega < F(x)\}) = F(x). \quad \mathbf{Q.E.D.} \quad (3.5)$$

**3.2. TÉTEL:** Legyen  $F$  a  $\xi$  valószínűségi változó eloszlásfüggvénye és  $a, b \in \mathbf{R}$ , ekkor

1.  $P(a \leq \xi < b) = F(b) - F(a)$ ,
2.  $P(\xi = a) = F(a + 0) - F(a)$ .

**Bizonyítás:** Legyen  $A = \{\xi < a\}$  és  $B = \{\xi < b\}$ , ekkor  $A \subset B$ . Ekkor  $\{a \leq \xi < b\} = B \setminus A$ ,  $P(B \setminus A) = P(B) - P(A) = F(b) - F(a)$ .

$$P(\xi = a) = \lim_{b \rightarrow a+0} P(a \leq \xi < b) = \lim_{b \rightarrow a+0} F(b) - F(a) = F(a + 0) - F(a). \quad \mathbf{Q.E.D.} \quad (3.6)$$

**Definíció:** A  $\xi$  valószínűségi változót *diszkrétnek* nevezzük, ha a lehetséges értékek  $\xi(\Omega)$  halmazának számossága legfeljebb megszámlálhatóan végtelen.

### 3. FEJEZET

---

**Megjegyzés:** Diszkrét valószínűségi változó esetén a lehetséges értékek felírhatók egy sorozatként.

**Definíció:** Legyen a  $\xi$  valószínűségi változó lehetséges értékeinek sorozata  $x_1, x_2, \dots$ . A  $p_i = P(\xi = x_i)$  ( $i = 1, 2, \dots$ ) valószínűségek sorozatát *eloszlásnak* nevezzük.

3.3. *TÉTEL:* Ha  $p_1, p_2, \dots$  eloszlás, akkor

$$p_i \geq 0 \quad (i = 1, 2, \dots) \quad \text{és} \quad \sum_{i=1}^{\infty} p_i = 1. \quad (3.7)$$

**Definíció:** Ha létezik  $f$  nemnegatív valós függvény, melyre

$$F(x) = \int_{-\infty}^x f(t)dt, \quad \forall x \in \mathbf{R}, \quad (3.8)$$

akkor  $f$  az  $F$  eloszlásfüggvényhez tartozó *sűrűségfüggvény*.

**Megjegyzés:** 1. A sűrűségfüggvény nem egyértelmű.

2. A sűrűségfüggvény létezése azt jelenti, hogy az  $F$  eloszlásfüggvény abszolút folytonos.

3.4. *TÉTEL:* Az  $f$  valós függvény akkor és csak akkor lehet sűrűségfüggvény, ha nemnegatív és

$$\int_{-\infty}^{+\infty} f(t)dt = 1. \quad (3.9)$$

**Definíció:** A valószínűségi változót *folytonosnak* nevezzük, ha létezik a sűrűségfüggvénye.

3.5. *TÉTEL:* Legyen a  $\xi$  folytonos valószínűségi változó  $f$  sűrűségfüggvénnyel és  $a, b \in \mathbf{R}$ , ekkor  $P(\xi = a) = 0$ , és  $P(a \leq \xi < b) = \int_a^b f(x)dx$ .

**Megjegyzés:** Tetszőleges eloszlásfüggvény előállítható

$$p_1 F_1 + p_2 F_2 + p_3 F_3 \quad (3.10)$$

alakban, ahol  $p_1 + p_2 + p_3 = 1$ ,  $p_1 \geq 0$ ,  $p_2 \geq 0$ ,  $p_3 \geq 0$ ,  $F_1$  diszkrét,  $F_2$  abszolút folytonos és  $F_3$  folytonos és szinguláris eloszlásfüggvény a Lebesgue-mértékre nézve.

A  $P$  és a  $P^*$  valószínűségek szingulárisak egymásra, ha  $\exists A \in \mathcal{F}$  úgy, hogy  $P(A) = 0$  és  $P^*(\bar{A}) = 0$ . Általában egy diszkrét és egy abszolút folytonos szinguláris egymásra nézve.

**3.1. Példa:** Folytonos és szinguláris eloszlásfüggvény a Lebesgue-mértékre nézve az ún. Cantor-függvény: A Cantor-féle triadikus halmaz elkészítésekor az  $n$ -edik lépésben éppen  $2^n - 1$  intervallumot vettünk ki a  $[0, 1]$  intervallumból. Jelölje ezeket sorban  $A_1, A_2, \dots, A_{2^n-1}$ . Ekkor legyen

$$F_n(x) = \begin{cases} 0, & \text{ha } x = 0, \\ \frac{k}{2^n}, & \text{ha } x \in A_k, \quad k = 1, 2, \dots, 2^n - 1, \\ 1, & \text{ha } x = 1. \end{cases} \quad (3.11)$$

Az

$$F(x) = \lim_{n \rightarrow \infty} F_n(x), \quad x \in \mathbf{R} \quad (3.12)$$

függvényt Cantor-függvénynek nevezzük.  $F$  monoton növekvő,  $F' = 0$  majdnem mindenütt és nem abszolút folytonos.  $\square$

### 3.2. Várható érték, transzformáció.

**Definíció:** 1. Ha a  $\xi$  diszkrét valószínűségi változó lehetséges értékeinek a száma véges, azaz a lehetséges értékek

$$x_1, x_2, \dots, x_n \quad \text{és} \quad p_i = P(\xi = x_i) \quad (i = 1, 2, \dots, n), \quad (3.13)$$

akkor a

$$\sum_{i=1}^n x_i p_i \quad (3.14)$$

mennyiséget *várható érték*nek nevezzük.

### 3. FEJEZET

---

2. Ha a  $\xi$  diszkrét valószínűségi változó lehetséges értékeinek számossága megszámlálhatóan végtelen, azaz a lehetséges értékek

$$x_1, x_2, \dots, \quad \text{és} \quad p_i = P(\xi = x_i) \quad (i = 1, 2, \dots), \quad (3.15)$$

akkor a

$$\sum_{i=1}^{\infty} x_i p_i \quad (3.16)$$

mennyiséget *várható értéknek* nevezzük, ha  $\sum_{i=1}^{\infty} |x_i| p_i < +\infty$ .

3. Ha  $\xi$  folytonos valószínűségi változó  $f$  sűrűségfüggvénnyel, akkor a  $\int_{-\infty}^{+\infty} x f(x) dx$  mennyiséget *várható értéknek* nevezzük, ha

$$\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty. \quad (3.17)$$

A  $\xi$  valószínűségi változó várható értékének a jelölése:  $E(\xi)$ .  $\square$

**Megjegyzés:** A definícióban az abszolút konvergenciát, azért követeljük meg, hogy a várható érték egyértelmű legyen. A várható érték röviden  $\int_{\Omega} \xi dP$ .

**3.2. Példa:** Legyen a  $\xi$  valószínűségi változó egy dobókockával dobott pontszám. Határozzuk meg a várható értékét!

**Megoldás:**  $P(\xi = 1) = \dots = P(\xi = 6) = \frac{1}{6}$ , tehát

$$E(\xi) = \sum_{k=1}^6 k P(\xi = k) = \sum_{k=1}^6 k \frac{1}{6} = \frac{21}{6} = 3.5. \quad \square \quad (3.18)$$

**3.3. Példa:** Legyenek a  $\xi$  valószínűségi változó lehetséges értékei  $x_n = \frac{(-2)^n}{n}$ ,  $(n \in \mathbf{N})$ , és  $P(\xi = x_n) = \frac{1}{2^n}$ . Határozzuk meg a várható értékét!

**Megoldás:**

$$\sum_{n=1}^{\infty} x_n P(\xi = x_n) = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n}. \quad (3.19)$$

Ez viszont nem abszolút konvergens, így nem létezik a várható érték.  $\square$

**3.4. Példa:** Legyen a  $\xi$  valószínűségi változó eloszlásfüggvénye

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \frac{x^2}{25}, & \text{ha } 0 < x \leq 5, \\ 1, & \text{ha } x > 5, \end{cases} \quad (3.20)$$

Határozzuk meg a várható értékét!

**Megoldás:** Létezik a sűrűségfüggvénye:

$$f(x) = \begin{cases} \frac{2x}{25}, & \text{ha } 0 < x \leq 5, \\ 0, & \text{egyébként.} \end{cases} \quad (3.21)$$

$$E(\xi) = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^0 x \cdot 0 dx + \int_0^5 x \frac{2x}{25} dx + \int_5^{\infty} x \cdot 0 dx = 0 + \frac{10}{3} + 0 = \frac{10}{3}. \square \quad (3.22)$$

3.6. *TÉTEL:* 1.  $E(a\xi + b) = aE(\xi) + b, \forall a, b \in \mathbf{R}.$

2. Ha  $m \leq \xi \leq M$ , akkor  $m \leq E(\xi) \leq M$ .

**Definíció:** Legyen  $\xi$  valószínűségi változó és  $g$  valós függvény. Ha az  $\eta = g(\xi)$  függvény valószínűségi változó, akkor a  $\xi$  *transzformáltjának* nevezzük.

**Megjegyzés:** A transzformált eloszlásfüggvénye

$$F_{\eta}(y) = P(\{\omega | g(\xi(\omega)) < y\}). \quad (3.23)$$

3.7. *TÉTEL:* Ha  $g$  differenciálható és  $g'(x) \neq 0$ , akkor  $\xi$  folytonos valószínűségi változó esetén  $\eta = g(\xi)$  folytonos valószínűségi változó, melynek sűrűségfüggvénye

$$f_{\eta}(y) = \begin{cases} f_{\xi}(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, & \text{ha } a < y < b, \\ 0, & \text{egyébként,} \end{cases} \quad (3.24)$$

ahol

$$a = \min\left(\lim_{x \rightarrow -\infty} g(x), \lim_{x \rightarrow +\infty} g(x)\right), \quad b = \max\left(\lim_{x \rightarrow -\infty} g(x), \lim_{x \rightarrow +\infty} g(x)\right). \quad (3.25)$$

3.8. *TÉTEL:* Ha  $\eta = g(\xi)$  a  $\xi$  valószínűségi változó transzformáltja, akkor

$$E(\eta) = \begin{cases} \sum_{i=1}^{\infty} g(x_i)P(\xi = x_i), & \text{ha } \xi \text{ diszkrét,} \\ \int_{-\infty}^{+\infty} g(x)f_{\xi}(x)dx, & \text{ha } \xi \text{ és } \eta \text{ folytonos.} \end{cases} \quad (3.26)$$

**Definíció:** Az  $E((\xi - E(\xi))^2)$  mennyiséget a  $\xi$  valószínűségi változó *szórásnégyzetének* nevezzük. Jele:  $D^2(\xi)$ .

**Definíció:** A  $\sqrt{E((\xi - E(\xi))^2)}$  mennyiséget a  $\xi$  valószínűségi változó *szórásának* nevezzük. Jele:  $D(\xi)$ .

**Definíció:** Az  $E(\xi^k)$  mennyiséget a  $\xi$  valószínűségi változó *k-adik momentumának* nevezzük.

**Definíció:** Az  $E((\xi - E(\xi))^k)$  mennyiséget a  $\xi$  valószínűségi változó *k-adik centrális momentumának* nevezzük.

3.9. *TÉTEL:* 1.  $D(a\xi + b) = |a| D(\xi), \forall a, b \in \mathbf{R}$ .

$$2. D^2(\xi) = E(\xi^2) - E^2(\xi).$$

$$3. D^2(\xi) = E((\xi - a)^2) + (a - E(\xi))^2.$$

$$4. \min_{a \in \mathbf{R}} E((\xi - a)^2) = D^2(\xi), \text{ és ekkor } a = E(\xi).$$

**Megjegyzés:** Az utóbbi két állítás hasonló (sőt formailag azonos) a tehetetlenségi nyomatékra vonatkozó közismert Steiner-tétellel, amely azt mondja ki, hogy egy egyenesen lévő tömegeloszlás tehetetlenségi nyomatéka valamely az egyenesre merőleges forgástengelyre vonatkozólag egyenlő a súlyponton áthaladó tengelyre vonatkozó tehetetlenségi nyomatéknak és a tengely súlyponttól mért távolsága négyzetösszegével,

ha az össztömeg egységnyi; következésképpen a tehetetlenségi nyomaték akkor minimális, ha a forgástengely a súlyponton megy át.

### 3.3. Néhány diszkrét eloszlás és jellemzői.

#### 1. BINOMIÁLIS ELOSZLÁS

Legyen  $n \in \mathbf{N}$ ,  $A \in \mathcal{F}$ , és végezzünk el egy  $n$  hosszúságú Bernoulli-kísérletsorozatot. Továbbá, legyen  $\xi$  az  $A$  esemény bekövetkezéseinek a száma. Ekkor  $\xi$  eloszlása

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k}, \quad (k = 0, 1, \dots, n), \quad (3.27)$$

ahol  $P(A) = p$  és  $q = 1 - p$ , és a  $\xi$  valószínűségi változót binomiális eloszlásúnak nevezzük. Jelölés:  $\xi \sim B(n, p)$ .

*3.10. TÉTEL:*  $E(\xi) = np$ ,  $D^2(\xi) = npq$ .

**Bizonyítás:** A várható érték definíciója szerint (diszkrét véges eset)

$$\begin{aligned} E(\xi) &= \sum_{k=0}^n k P(\xi = k) = \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} = \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} q^{n-1-(k-1)} \\ &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} q^{n-1-(k-1)} = np, \end{aligned} \quad (3.28)$$

a binomiális tétel szerint.

A szórásnégyzet meghatározása hasonló felhasználva, hogy  $D^2(\xi) = E(\xi^2) - E^2(\xi)$ . **Q.E.D.**

**Megjegyzés:** A visszatevéses mintavétel binomiális eloszláshoz vezet.

#### 2. POISSON-ELOSZLÁS

### 3. FEJEZET

---

Legyen  $\lambda > 0$  rögzített konstans és  $\lambda = np_n$ , ekkor

$$\lim_{n \rightarrow \infty, \lambda = np_n} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{ahol } k = 0, 1, \dots \quad (3.29)$$

A  $\xi$  valószínűségi változót Poisson-eloszlásúnak nevezzük  $\lambda > 0$  paraméterrel, ha eloszlása

$$P(\xi = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{ahol } k = 0, 1, \dots \quad (3.30)$$

Jelölés:  $\xi \sim \text{Poisson}(\lambda)$ .

*3.11. TÉTEL:*  $E(\xi) = \lambda$ ,  $D^2(\xi) = \lambda$ .

**Bizonyítás:** A várható érték definíciója szerint (diszkrét végtelen eset)

$$\begin{aligned} E(\xi) &= \sum_{k=0}^{\infty} k P(\xi = k) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda, \end{aligned} \quad (3.31)$$

mert az exponenciális függvény sorfejtése szerint

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}. \quad (3.32)$$

A szórásnégyzet meghatározása hasonló felhasználva, hogy  $D^2(\xi) = E(\xi^2) - E^2(\xi)$ . **Q.E.D.**

### 3. GEOMETRIAI ELOSZLÁS

A binomiális eloszlás bevezetésekor használt jelölések mellett a  $\xi$  valószínűségi változó jelentse az  $A$  esemény első bekövetkezéséhez szükséges kísérletek számát. A  $\xi$  eloszlása

$$P(\xi = k) = pq^{k-1}, \quad \text{ahol } k = 1, 2, \dots \quad (3.33)$$

*3.12. TÉTEL:*  $E(\xi) = \frac{1}{p}$ ,  $D^2(\xi) = \frac{q}{p^2}$ .

**Megjegyzés:** A  $\eta = \xi - 1$  valószínűségi változót is szokás geometriai eloszlásúnak nevezni. Az  $\eta$  eloszlása

$$P(\eta = k) = pq^k, \quad \text{ahol } k = 0, 1, 2, \dots \quad (3.34)$$

3.13. *TÉTEL:*  $E(\eta) = \frac{q}{p}, D^2(\eta) = \frac{q}{p^2}.$

**Megjegyzés:**

$$P(\eta = k + m | \eta \geq m) = \frac{P(\{\eta = k + m\} \cap \{\eta \geq m\})}{P(\eta \geq m)}. \quad (3.35)$$

Viszont

$$\{\eta = k + m\} \cap \{\eta \geq m\} = \{\eta = k + m\} \quad (3.36)$$

és

$$P(\eta \geq m) = pq^m (1 + q + q^2 + \dots) = \frac{pq^m}{1 - q} = q^m. \quad (3.37)$$

Tehát

$$P(\eta = k + m | \eta \geq m) = \frac{pq^{m+k}}{q^m} = pq^k = P(\eta = k). \quad (3.38)$$

Ezzel beláttuk a geometriai eloszlás emlékezet nélküli tulajdonságát.  $\square$

### 3.4. Néhány folytonos eloszlás és jellemzői.

#### 1. EGYENLETES ELOSZLÁS

Legyen  $a, b \in \mathbf{R}$  és  $a < b$ . A  $\xi$  egyenletes eloszlású az  $(a, b)$  intervallumon, ha a sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{ha } a < x < b, \\ 0, & \text{egyébként.} \end{cases} \quad (3.39)$$

Jelölés:  $\xi \sim U(a, b)$ . Az eloszlásfüggvény

$$F(x) = \begin{cases} 0, & \text{ha } x \leq a, \\ \frac{x-a}{b-a}, & \text{ha } a < x \leq b, \\ 1, & \text{ha } x > b. \end{cases}$$

### 3. FEJEZET

---

3.14. *TÉTEL:*  $E(\xi) = \frac{a+b}{2}$ ,  $D^2(\xi) = \frac{(b-a)^2}{12}$ .

**Bizonyítás:**

$$E(\xi) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{a+b}{2}. \quad (3.40)$$

$$E(\xi^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^3}{3} \right]_a^b = \frac{1}{b-a} \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}. \quad (3.41)$$

Tehát a szórásnégyzet

$$D^2(\xi) = E(\xi^2) - E^2(\xi) = \frac{a^2 + ab + b^2}{3} - \left( \frac{a+b}{2} \right)^2 = \frac{(b-a)^2}{12}. \quad \mathbf{Q.E.D.} \quad (3.42)$$

**Megjegyzés:** Az egyenletes eloszlás adja a geometriai valószínűségi mező elméleti alapját.

3.15. *TÉTEL:* Ha  $F$  szigorúan monoton növekvő eloszlásfüggvény és  $\xi$   $F$  eloszlású, akkor  $\eta = F(\xi)$  egyenletes eloszlású a  $[0, 1]$  intervallumon. Fordítva, ha  $\xi \sim U(0, 1)$ , akkor  $\eta = F^{-1}(\xi)$  éppen  $F$  eloszlású.

**Bizonyítás:** Legyen  $x \in [0, 1]$ .

$$F_\eta(x) = P(\eta < x) = P(F(\xi) < x) = P(\xi < F^{-1}(x)) = F(F^{-1}(x)) = x. \quad \square \quad (3.43)$$

## 2. EXPONENCIÁLIS ELOSZLÁS

A  $\xi$  exponenciális eloszlású  $\lambda > 0$  paraméterrel, ha a sűrűségfüggvénye

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{ha } x \geq 0, \\ 0, & \text{egyébként.} \end{cases} \quad (3.44)$$

Jelölés:  $\xi \sim \text{Exp}(\lambda)$ . Az eloszlásfüggvény

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{ha } x > 0. \end{cases} \quad (3.45)$$

3.16. *TÉTEL:*  $E(\xi) = \frac{1}{\lambda}$ ,  $D^2(\xi) = \frac{1}{\lambda^2}$ .

Örökifjú tulajdonság:  $P(\xi \geq a + b | \xi \geq a) = P(\xi \geq b)$ , ahol  $a > 0$ ,  $b > 0$ .

### 3. NORMÁLIS ELOSZLÁS

Legyen  $m \in \mathbf{R}$ ,  $\sigma > 0$ . Az  $\eta$  normális eloszlású, ha a sűrűségfüggvénye

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad x \in \mathbf{R}. \quad (3.46)$$

Jelölés:  $\eta \sim N(m, \sigma^2)$ . Ha  $m = 0$  és  $\sigma = 1$ , akkor a valószínűségi változót standard normális eloszlásúnak nevezzük. Jelölje a sűrűségfüggvényét  $\varphi$  és az eloszlásfüggvényét  $\Phi$ . Ha  $\xi$  standard normális eloszlású, akkor az  $\eta = \sigma\xi + m$  valószínűségi változó  $F$  eloszlásfüggvényére jellemző, hogy

$$F(x) = \Phi\left(\frac{x-m}{\sigma}\right). \quad (3.47)$$

3.17. *TÉTEL:*  $E(\xi) = m$ ,  $D^2(\xi) = \sigma^2$ .

**Megjegyzés:** 1. A  $\varphi$  függvény írja le a Gauss-görbét(harang görbét).

2.  $\Phi(0) = 0.5$  és  $\Phi(-x) = 1 - \Phi(x)$ .

3.18. *TÉTEL: (Moivre-Laplace)* Legyen a  $\xi$  valószínűségi változó binomiális eloszlású  $n$  és  $p$  paraméterrel és  $0 \leq a < b \leq n$  egész, akkor

$$\begin{aligned} P(a \leq \xi \leq b) &= \sum_{k=a}^b \binom{n}{k} p^k q^{n-k} \approx \\ &\approx \Phi\left(\frac{b - np + \frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np - \frac{1}{2}}{\sqrt{npq}}\right). \end{aligned} \quad (3.48)$$

**Bizonyítás:** Legyen  $p_k = P(\xi = k)$  ( $k = 0, 1, 2, \dots, n$ ), és

$$x = \frac{k - np}{\sqrt{npq}} \quad (3.49)$$

### 3. FEJEZET

---

A  $p_k$  és  $p_{k+1}$  valószínűségeket átalakítjuk  $\varphi(x)$  és  $\varphi(x + \delta x)$  értékévé, azaz

$$\varphi(x) = \sqrt{npq}p_k \quad \text{és} \quad \delta x = \frac{1}{\sqrt{npq}} \quad (3.50)$$

$$\varphi(x + \delta x) = \sqrt{npq}p_{k+1}. \quad (3.51)$$

$$\begin{aligned} \delta\varphi &= \varphi(x + \delta x) - \varphi(x) \\ &\approx \sqrt{npq}(p_{k+1} - p_k) \end{aligned} \quad (3.52)$$

$$\frac{\delta\varphi}{\varphi} \approx \frac{p_{k+1} - p_k}{p_k} = \frac{p_{k+1}}{p_k} - 1. \quad (3.53)$$

Azonban

$$p_k = \binom{n}{k} p^k q^{n-k} \quad \text{és} \quad p_{k+1} = \binom{n}{k+1} p^{k+1} q^{n-k-1}. \quad (3.54)$$

Tehát

$$\begin{aligned} \frac{\delta\varphi}{\varphi} &\approx \frac{\binom{n}{k+1} p^{k+1} q^{n-k-1}}{\binom{n}{k} p^k q^{n-k}} - 1 \\ &= \frac{(n-k)p}{(k+1)q} - 1 \\ &= \frac{np - kp - kq - q}{(k+1)q} = \frac{np - k - q}{(k+1)q}. \end{aligned} \quad (3.55)$$

Azonban

$$k = np + \sqrt{npq}x \quad \text{és} \quad \delta x = \frac{1}{\sqrt{npq}}. \quad (3.56)$$

Így

$$\begin{aligned} \frac{1}{\varphi} \frac{\delta\varphi}{\varphi} &\approx \frac{[-\sqrt{npq}x - q]\sqrt{npq}}{[np + \sqrt{npq}x + 1]q} \\ &= \frac{-np x - \sqrt{npq}}{np + \sqrt{npq}x + 1} \\ &= \frac{-x - \sqrt{\frac{q}{np}}}{1 + \sqrt{\frac{q}{np}}x + \frac{1}{np}} \end{aligned} \quad (3.57)$$

Ha  $n \rightarrow \infty$  (ekkor  $\delta x \rightarrow 0$ ), akkor

$$\begin{aligned}\frac{1}{\varphi} \frac{d\varphi}{dx} &= -x, \\ \ln \varphi(x) &= -\frac{1}{2}x^2 + \ln c, \\ \varphi(x) &= ce^{-\frac{1}{2}x^2}.\end{aligned}\tag{3.58}$$

Mivel  $\varphi$  sűrűségfüggvény, ezért

$$1 = \int_{-\infty}^{\infty} \varphi(x) dx = \int_{-\infty}^{\infty} ce^{-\frac{1}{2}x^2} dx = c \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = c\sqrt{2\pi}.\tag{3.59}$$

Tehát

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \quad \mathbf{Q.E.D.}\tag{3.60}$$

#### 4. CAUCHY ELOSZLÁS

Legyen  $c \in \mathbf{R}$ ,  $s > 0$ . Az  $\eta$  Cauchy eloszlású, ha a sűrűségfüggvénye

$$f(x) = \frac{1}{\pi s \left[ 1 + \left( \frac{x-c}{s} \right)^2 \right]}, \quad x \in \mathbf{R}.\tag{3.61}$$

Nem létezik a várható érték. Az eloszlásfüggvény

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan \left( \frac{x-c}{s} \right).\tag{3.62}$$

**Megjegyzés:** Szokás csak a  $c = 0$ ,  $s = 1$  esetet (standard) Cauchy-eloszlásnak nevezni.

#### 5. WEIBULL ELOSZLÁS

A Weibull-eloszlás paramétereire többféle elterjedt jelölésrendszer van. Az eltérő jelölések használatát egyértelműen magyarázza, hogy a Weibull-eloszlás igen széles

körben, a legkülönbébb tudományterületeken alkalmazták, valamint a paramétereknek sokféle meghatározási módja is ismeretes és az egyes megoldásoknál a változók átírása jelentős egyszerűsítéseket eredményez. Mi a következőkben az

$$F_c(x) = \begin{cases} 1 - \exp(-x^c), & \text{ha } x \geq 0 \\ 0, & \text{ha } x < 0 \end{cases} \quad (3.63)$$

jelölést alkalmazzuk a standard Weibull-eloszlás jelölésére. Ebből a lineáris transzformáltak eloszlása

$$F_c\left(\frac{x-a}{b}\right). \quad (3.64)$$

Tehát ez az eloszláscsalád háromparaméteres, amelyből a  $c$  az ún. alakparaméter (típusparaméter). Viszont lényeges, hogy aszimmetrikus eloszlás.

**Megjegyzés:** Az eloszlás  $c = 1$  esetén az exponenciális eloszlást,  $c = 2$  a Rayleigh eloszlás adja, míg  $c = 3.57$  közelében az eloszlás közel szimmetrikussá válik és jól közelíti a normális eloszlást. Megfelelő paraméter választással az is elérhető, hogy a Weibull- eloszlás jól közelítse a lognormális és Gamma-eloszlásokat.

### 3.5. A Poisson-eloszlás származtatása.

Tekintsük a következő tulajdonságokkal rendelkező időben lejátszódó folyamatot: legyenek a bekövetkezések egymástól függetlenek, hosszú intervallumon az intenzitás egyenletes (a bekövetkezések várható száma  $\lambda h$ ,  $h$  az intervallum hossza), míg egységnyi intervallumon a várható számuk  $\lambda > 0$ . Továbbá, nagyon kicsi  $\delta t$  intervallumon legfeljebb majdnem csak egyszer következik be.

Az  $o(g(x))$  (kis ordó) jelölést alkalmazva folyamatunk matematikai modellje a következő:

**Megjegyzés:**  $o(g(x)) = h(x)$ , ha  $x \rightarrow a$ , azt jelenti, hogy

$$\lim_{x \rightarrow a} \frac{h(x)}{g(x)} = 0. \quad (3.65)$$

Legyen  $p_k(t) = P(\text{ pontosan } k \text{ darab bekövetkezés } t \text{ hosszúságú intervallumon})$ .  
Ekkor

$$\begin{aligned} p_k(t + \delta t) &= p_k(t)p_0(\delta t) + p_{k-1}(t)p_1(\delta t) + o(\delta t), \quad \text{ha } k \geq 1, \\ p_0(t + \delta t) &= p_0(t)p_0(\delta t) + o(\delta t). \end{aligned} \quad (3.66)$$

A kezdeti feltételek pedig

$$p_0(0) = 1, \quad p_k(0) = 0, \quad \text{ha } k \geq 1. \quad (3.67)$$

Továbbá, felírva nagyon kicsi intervallumra a folyamat tulajdonságait, a bekövetkezések várható száma alapján, a következő összefüggéseket kapjuk

$$\begin{aligned} p_1(\delta t) &= \lambda \delta t + o(\delta t), \\ p_0(\delta t) &= 1 - \lambda \delta t + o(\delta t). \end{aligned} \quad (3.68)$$

Ezeket behelyettesítve

$$\begin{aligned} p_k(t + \delta t) &= p_k(t)(1 - \lambda \delta t) + p_{k-1}(t)\lambda \delta t + o(\delta t), \quad \text{ha } k \geq 1, \\ p_0(t + \delta t) &= p_0(t)(1 - \lambda \delta t) + o(\delta t). \end{aligned} \quad (3.69)$$

Rendezzük át az egyenleteket:

$$\begin{aligned} \frac{p_k(t + \delta t) - p_k(t)}{\delta t} + \lambda p_k(t) &= \lambda p_{k-1}(t) + o(1), \quad \text{ha } k \geq 1, \\ \frac{p_0(t + \delta t) - p_0(t)}{\delta t} + \lambda p_0(t) &= o(1). \end{aligned} \quad (3.70)$$

Ha  $\delta t \rightarrow 0$ , akkor

$$\begin{aligned} p'_k(t) + \lambda p_k(t) &= \lambda p_{k-1}(t), \quad \text{ha } k \geq 1, \\ p'_0(t) + \lambda p_0(t) &= 0. \end{aligned} \quad (3.71)$$

A második sorból rögtön adódik, hogy

$$\begin{aligned} \frac{d}{dt} [e^{\lambda t} p_0(t)] &= 0, \\ e^{\lambda t} p_0(t) &= p_0(0). \end{aligned} \quad (3.72)$$

### 3. FEJEZET

---

A  $p_0(0) = 1$  kezdeti feltétel szerint pedig  $e^{\lambda t} p_0(t) = 1$ .

Hasonlóan a  $k \geq 1$  esetekre

$$\begin{aligned} \frac{d}{dt} [e^{\lambda t} p_k(t)] &= \lambda e^{\lambda t} p_{k-1}(t), \\ [e^{\lambda t} p_k(t)]_0^x &= \int_0^x \lambda e^{\lambda t} p_{k-1}(t) dt. \end{aligned} \quad (3.73)$$

A  $p_k(0) = 0$  ( $k \geq 1$ ) kezdeti feltétel alapján

$$e^{\lambda x} p_k(x) = \int_0^x \lambda e^{\lambda t} p_{k-1}(t) dt. \quad (3.74)$$

Vezessük be a  $Q_k(x) = e^{\lambda x} p_k(x)$  jelölést, ekkor

$$\begin{aligned} Q_k(x) &= \lambda \int_0^x Q_{k-1}(t) dt, \quad \text{ha } k \geq 1, \\ Q_0(x) &= 1. \end{aligned} \quad (3.75)$$

Ebből rögtön adódik, hogy

$$\begin{aligned} Q_1(x) &= \lambda \int_0^x dt = \lambda x, \quad \text{azaz } p_1(x) = \lambda x e^{-\lambda x}; \\ Q_2(x) &= \lambda \int_0^x \lambda t dt = \frac{\lambda^2 x^2}{2}, \quad \text{azaz } p_2(x) = \frac{\lambda^2 x^2}{2} e^{-\lambda x}; \\ Q_3(x) &= \lambda \int_0^x \frac{\lambda^2 t^2}{2} dt = \frac{\lambda^3 x^3}{6}, \quad \text{azaz } p_3(x) = \frac{\lambda^3 x^3}{6} e^{-\lambda x}; \dots \end{aligned} \quad (3.76)$$

Általában pedig

$$Q_k(x) = \frac{\lambda^k x^k}{k!}, \quad \text{azaz } p_k(x) = \frac{\lambda^k x^k}{k!} e^{-\lambda x}. \quad (3.77)$$

Végül egy rögzített  $t$  hosszúságú intervallumon, mivel a bekövetkezések várható száma  $\mu = \lambda t$ , megkapjuk a Poisson-eloszlást.

$$P(\xi = k) = p_k(t) = \frac{\lambda^k t^k}{k!} e^{-\lambda t} = \frac{\mu^k}{k!} e^{-\mu}. \quad \mathbf{Q.E.D.} \quad (3.78)$$

### 3.6. Eloszlástípusok.

Mint látni fogjuk nagyon sok becslés a hely- (lokációs, centrum) és a skálaparaméterre vonatkozik vagy arra vezeti vissza a problémákat. Mit is jelentenek ezek a paraméterek? Tudjuk, hogy a normális eloszlások esetén van az ún. standard normális eloszlás, amelynél az egyik paraméter az origóba esik, a másik pedig az egységgel egyezik meg. Ha egy ilyen eloszlású  $\xi$  valószínűségi változóhoz tekintjük azokat az  $\eta$  valószínűségi változókat, amelyek vele lineárisan függenek össze:

$$\eta = a\xi + b \quad (a > 0, b \in \mathbf{R}), \quad (3.79)$$

akkor ismeretes, hogy ez a transzformáció geometriailag a mértékegység és a kezdőpont megváltoztatását jelenti, amelyek meghatározása a gyakorlati problémákban nagyon fontos. Mint könnyen látható a  $\xi$  és az  $\eta$  valószínűségi változók eloszlásfüggvényei között az

$$F_\xi(x) = F_\eta(ax + b) \quad (3.80)$$

összefüggés áll fenn.

**Definíció:** Az  $F(x)$  és  $G(x)$  eloszlásfüggvények típusa megegyezik, ha valamely  $a > 0$  és  $b$  konstansok mellett érvényes a

$$G(x) = F(ax + b) \quad (3.81)$$

azonosság.

**Megjegyzés:** 1. A típushoz való tartozás relációja reflexív, szimmetrikus és tranzitív, így ekvivalenciareláció, amelyik osztályozást határoz meg az eloszlásfüggvények halmazán.

### 3. FEJEZET

---

2. Azokat az eloszlásokat, amelyeknél a kezdőpont (a helyparaméter) az origóba esik, a mértékegység (a skálaparaméter) pedig a valós egység standard eloszlásoknak nevezzük, amelyek természetesen reprezentálják osztályukat. Ezek persze függenek a méréseknél használt "fizikai" mértékegységektől.

3. Valódinak nevezzük egy eloszlást, ha nem egy pontra koncentrált.

*3.19. TÉTEL:* Ha az eloszlásfüggvények valamely  $\{F_n(x)\}$  sorozata gyengén konvergál az  $F(x)$  valódi eloszlásfüggvényhez, akkor az  $\{a_n > 0\}$ ,  $\{b_n\}$  számsorozatok tetszőleges választása esetén az  $\{F_n(a_n x + b_n)\}$  sorozat csupán egy az  $F(x)$  típushoz tartozó valódi eloszláshoz konvergálhat.

*3.20. TÉTEL:* Az  $\{F_n(x)\}$  eloszlásfüggvények valamely sorozatára az

$$F_n(a_n x + b_n) \rightarrow F(x), \quad (3.82)$$

$$F_n(\alpha_n x + \beta_n) \rightarrow F(x) \quad (3.83)$$

relációk akkor és csak akkor érvényesek egyidejűleg, ha

$$\frac{\alpha_n}{a_n} \rightarrow 1, \quad \frac{b_n - \beta_n}{b_n} \rightarrow 0, \quad (3.84)$$

ahol  $\{a_n > 0\}$ ,  $\{\alpha_n\}$ ,  $\{b_n\}$ ,  $\{\beta_n\}$  valós számsorozatok,  $F(x)$  pedig egy valódi eloszlásfüggvény.

### 3.7. Generátor-, karakterisztikus függvény.

**Definíció:** Legyen  $\xi$  egy nemnegatív egész értékű valószínűségi változó és legyen  $p_j = P(\xi = j)$ ,  $(j = 0, 1, 2, \dots)$ . A

$$G_\xi(z) = \sum_{j=0}^{\infty} p_j z^j = E(z^\xi) \quad (3.85)$$

függvényt a  $\xi$  generátorfüggvényének nevezzük.

3.21. *TÉTEL:* Legyen  $\xi$  és  $\eta$  nemnegatív egész értékű valószínűségi változó, ekkor

- (a)  $G_\xi(z)$  konvergens, ha  $|z| \leq 1$ .
- (b)  $\xi$  és  $\eta$  eloszlása akkor és csak akkor egyezik meg, ha  $G_\xi(z) = G_\eta(z)$ .
- (c)  $p_n = \frac{1}{n!} \frac{d^n G_\xi(z)}{dz^n} \Big|_{z=0}$ ,  $n = 0, 1, 2, \dots$
- (d)  $E(\xi) = G'_\xi(1)$  és  $D^2(\xi) = G''_\xi(1) + G'_\xi(1) - (G'_\xi(1))^2$ .

**Definíció:** Legyen  $\xi$  valószínűségi változó a

$$\varphi_\xi(t) = E(e^{it\xi}), \quad t \in \mathbf{R} \quad (3.86)$$

függvényt a  $\xi$  *karakterisztikus függvényének* nevezzük.

3.22. *TÉTEL:* Legyen  $\xi$  és  $\eta$  valószínűségi változó, ekkor

- (a)  $F_\xi = F_\eta$  akkor és csak akkor, ha  $\varphi_\xi = \varphi_\eta$ .
- (b)  $|\varphi_\xi(t)| \leq \varphi_\xi(0) = 1$ ,  $\forall t \in \mathbf{R}$ .
- (c)  $\varphi_\xi^{(k)} = i^k E(\xi^k)$ , ha  $E(\xi^k)$  létezik.

3.23. *TÉTEL:* Ha a  $\varphi$  karakterisztikus függvény abszolút integrálható, akkor a  $\xi$  valószínűségi változónak létezik a sűrűségfüggvénye, és

$$f_\xi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \varphi_\xi(u) du. \quad (3.87)$$

## 4. VÉLETLEN VEKTOROK

### 4.1. A kétdimenziós véletlen vektor.

**Definíció:** A  $(\xi, \eta) : \Omega \rightarrow \mathbf{R}^2$  leképezést *(kétdimenziós) véletlen vektornak* nevezzük, ha

$$\{\xi < x, \eta < y\} = \{\omega | \omega \in \Omega, \quad \xi(\omega) < x, \eta(\omega) < y\} \in \mathcal{F} \quad \forall x, y \in \mathbf{R}. \quad (4.1)$$

**Definíció:** Az  $F(x, y) = P(\xi < x, \eta < y)$  formulával meghatározott valós értékű függvényt a  $(\xi, \eta)$  véletlen vektor *együttes eloszlásfüggvényének* nevezzük. Az

$$F_\xi(x) = \lim_{y \rightarrow +\infty} F(x, y), \quad F_\eta(y) = \lim_{x \rightarrow +\infty} F(x, y) \quad (4.2)$$

függvényeket *peremeloszlásfüggvénynek* nevezzük

**4.1. TÉTEL:** Az  $F$  függvény akkor és csak akkor lehet együttes eloszlásfüggvény, ha

$$1. \quad \lim_{x \rightarrow -\infty} F(x, y) = 0, \quad \lim_{y \rightarrow -\infty} F(x, y) = 0,$$

$$2. \quad \lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F(x, y) = 1,$$

3.  $F$  mindkét változójában balról folytonos,

4.  $F(b, d) - F(b, c) - F(a, d) + F(a, c) \geq 0$ ,  $\forall a < b, c < d$  esetén, azaz teljesül az ún. "téglalap" tulajdonság.

**Megjegyzés:** A téglalap tulajdonságból következik, hogy mindkét változójában monoton növekvő.

**Definíció:** A  $(\xi, \eta)$  véletlen vektort *diszkrétnek* nevezzük, ha a lehetséges értékek számossága legfeljebb megszámlálhatóan végtelen.

**Definíció:** Legyen a  $\xi$ , illetve  $\eta$  valószínűségi változó lehetséges értékeinek sorozata  $x_1, x_2, \dots$ , illetve  $y_1, y_2, \dots$ . A  $P(\xi = x_i, \eta = y_j) = p_{ij}$  ( $i, j = 1, 2, \dots$ ) valószínűségek sorozatát *együttes eloszlás*nak nevezzük. A

$$q_i = \sum_{j=1}^{\infty} p_{ij}, \quad (i = 1, 2, \dots), \quad (4.3)$$

$$r_j = \sum_{i=1}^{\infty} p_{ij}, \quad (j = 1, 2, \dots) \quad (4.4)$$

valószínűség sorozatokat *peremeloszlás*nak nevezzük. Minden  $r_j > 0$  esetén a  $\xi$  *feltételes eloszlása* adott  $\eta = y_j$  mellett

$$P(\xi = x_i | \eta = y_j) = \frac{p_{ij}}{r_j}. \quad (4.5)$$

Az

$$E(\xi | \eta = y_j) = \sum_{i=1}^{\infty} x_i \frac{p_{ij}}{r_j} \quad (4.6)$$

mennyiséget *feltételes várható érték*nek nevezzük. Az

$$E(\xi | \eta = y_j) = m_2(y_j) \quad (4.7)$$

függvényt a  $\xi$ -nek az  $\eta$ -ra vonatkozó *regressziós függvény*ének nevezzük.

4.2. **TÉTEL:** Ha  $p_{ij}$  ( $i, j = 1, 2, \dots$ ) együttes eloszlás, akkor

$$p_{ij} \geq 0 \quad (i, j = 1, 2, \dots) \quad \text{és} \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1. \quad (4.8)$$

**Definíció:** Ha létezik  $f$  nemnegatív valós értékű függvény, melyre

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du, \quad \forall x, y \in \mathbf{R}, \quad (4.9)$$

akkor  $f$  az  $F$  eloszlásfüggvényhez tartozó *együttes sűrűségfüggvény*. Az

$$f_{\xi}(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_{\eta}(y) = \int_{-\infty}^{+\infty} f(x, y) dx \quad (4.10)$$

függvényeket *peremsűrűségfüggvény*nek nevezzük.

**4.3. TÉTEL:** Az  $f$  függvény akkor és csak akkor lehet együttes sűrűségfüggvény, ha nemnegatív és

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dy dx = 1. \quad (4.11)$$

**Definíció:** A  $(\xi, \eta)$  véletlen vektort *folytonosnak* nevezzük, ha létezik az együttes sűrűségfüggvénye.

**Definíció:** A  $\xi$  és  $\eta$  valószínűségi változót *függetlennek* nevezzük, ha

$$F(x, y) = F_{\xi}(x)F_{\eta}(y), \quad \forall x, y \in \mathbf{R}. \quad (4.12)$$

**Megjegyzés:** A függetlenség megfelelői diszkrét illetve folytonos esetben:

$$p_{ij} = q_i r_j, \quad (i, j = 1, 2, \dots), \quad (4.13)$$

$$f(x, y) = f_{\xi}(x)f_{\eta}(y) \quad \forall x, y \in \mathbf{R}. \quad (4.14)$$

**Definíció:** Legyen  $(\xi, \eta)$  véletlen vektor. Az  $F(x|y)$  az *feltételes eloszlásfüggvénye* a  $\xi$ -nek  $\eta = y$  esetén, ha

$$F(x|y) = P(\xi < x | \eta = y) = \lim_{h \rightarrow 0+0} P(\xi < x | y \leq \eta < y + h). \quad (4.15)$$

**Megjegyzés:** Ha léteznek a feltételes valószínűségek.

**Definíció:** Ha létezik  $f_{\xi|\eta}$  nemnegatív valós értékű függvény, melyre

$$F(x|y) = \int_{-\infty}^x f_{\xi|\eta}(u|y) du, \quad \forall x, y \in \mathbf{R}, \quad (4.16)$$

akkor  $f_{\xi|\eta}$  a  $\xi$ -nek az  $\eta$ -ra vonatkozó *feltételes sűrűségfüggvénye*.

**Megjegyzés:**

$$f_{\xi|\eta}(x|y) = \frac{f(x, y)}{f_{\eta}(y)}. \quad (4.17)$$

**Definíció:** A feltételes sűrűségfüggvény segítségével meghatározott feltételes várható értéket *regressziós függvénynek* nevezzük, azaz az

$$\int_{-\infty}^{+\infty} x f_{\xi|\eta}(x|y) dx = m_2(y) \quad (4.18)$$

függvényt a  $\xi$ -nek az  $\eta$ -ra vonatkozó regressziós függvényének nevezzük, illetve az

$$\int_{-\infty}^{+\infty} y f_{\eta|\xi}(y|x) dy = m_1(x) \quad (4.19)$$

függvényt a  $\eta$ -nak az  $\xi$ -re vonatkozó regressziós függvényének nevezzük.

**Megjegyzés:** A  $\min_g E((\xi - g(\eta))^2)$  értékét, akkor kapjuk, ha  $g$  megegyezik a regressziós függvényével.

4.4. *TÉTEL:* Ha  $(\xi, \eta)$  véletlen vektor és  $g : \mathbf{R}^2 \rightarrow \mathbf{R}$  olyan függvény, hogy  $g(\xi, \eta)$  valószínűségi változó, akkor

$$E(g(\xi, \eta)) = \begin{cases} \sum_{i,j} g(x_i, y_j) p_{ij}, & \text{ha } (\xi, \eta) \text{ diszkrét,} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dy dx, & \text{ha } (\xi, \eta) \text{ folytonos.} \end{cases} \quad (4.20)$$

**Definíció:** A

$$\text{cov}(\xi, \eta) = E((\xi - E(\xi))(\eta - E(\eta))) \quad (4.21)$$

mennyiséget *kovarianciának* nevezzük. Az

$$r(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{D(\xi)D(\eta)} \quad (4.22)$$

menntis6get pedig *korrelációs egy6tthathónak* nevezz6k.

**Megjegyzés:** A korrelációs egy6tthathó az 6sszef6ggést próbálja meg mérni. Ha  $\xi$  és  $\eta$  független, akkor  $r(\xi, \eta) = 0$ , fordítva nem igaz. Pl. ha  $\xi \sim N(0, 1)$ ,  $\eta = \xi^2$ , akkor  $r(\xi, \eta) = 0$ .

**Definíció:** Legyen  $g$  valós függvény. Az

$$I(\xi, \eta) = \sqrt{1 - \frac{D^2(\eta - g(\xi))}{D^2(\eta)}} \quad (4.23)$$

menntis6get *korrelációs indexnek* nevezz6k.

**4.5. TÉTEL:**

1.  $E(\xi + \eta) = E(\xi) + E(\eta)$ .
2.  $D^2(\xi + \eta) = D^2(\xi) + D^2(\eta) + 2cov(\xi, \eta)$ .
3.  $E(E(\xi|\eta = y)) = E(\xi)$ .
4.  $|cov(\xi, \eta)| \leq D(\xi)D(\eta)$ , azaz  $|r(\xi, \eta)| \leq 1$ .
5.  $0 \leq I(\xi, \eta) \leq 1$ .

## 4.2. Néhány folytonos eloszlás és jellemzői.

### 1. EGYENLETES ELOSZLÁS

A  $(\xi, \eta)$  véletlen vektor egyenletes eloszlású az  $A \subset \mathbf{R}^2$  tartományon, ha

$$f(x, y) = \begin{cases} \frac{1}{|A|}, & \text{ha } (x, y) \in A, \\ 0, & \text{egyébként.} \end{cases} \quad (4.24)$$

**4.1. Példa:** Legyen a  $(\xi, \eta)$  véletlen vektor egyenletes eloszlású a  $(0, 0)$ ,  $(1, 0)$ ,  $(0, 2)$  pontok által meghatározott háromszögön. Határozzuk meg az egy6ttes sűrűségfüggvényt, az egy6ttes eloszlásfüggvényt, a peremeloszlásfüggvényeket, a várható értékeket, és a  $P(\xi < 0.5|\eta = 0.5)$  valószínűséget!

**Megoldás:** A háromszög területe 1. Tehát

$$f(x, y) = \begin{cases} 1, & \text{ha } 0 \leq x \leq 1, 0 \leq y \leq 2 - 2x, \\ 0, & \text{egyébként.} \end{cases} \quad (4.25)$$

$$F(x, y) = \begin{cases} 0, & \text{ha } x \leq 0, \text{ vagy } y \leq 0, \\ xy, & \text{ha } 0 < x \leq 1, 0 < y \leq 2 - 2x, \\ xy - \frac{(2x + y - 2)^2}{4}, & \text{ha } 0 < x \leq 1, 2 - 2x < y \leq 2, \\ 2x - x^2, & \text{ha } 0 < x \leq 1, \text{ és } y > 2, \\ y - \frac{y^2}{4}, & \text{ha } 0 < y \leq 2, \text{ és } x > 1, \\ 1, & \text{ha } x > 1, \text{ és } y > 2. \end{cases} \quad (4.26)$$

$$F_\xi(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 2x - x^2, & \text{ha } 0 < x \leq 1, \\ 1, & \text{ha } x > 1, \end{cases} \quad (4.27)$$

$$F_\eta(y) = \begin{cases} 0, & \text{ha } y \leq 0, \\ y - \frac{y^2}{4}, & \text{ha } 0 < y \leq 2, \\ 1, & \text{ha } y > 2, \end{cases} \quad (4.28)$$

$$E(\xi) = \frac{1}{3}, \quad E(\eta) = \frac{2}{3}, \quad P(\xi < 0.5 | \eta = 0.5) = \frac{2}{3}. \quad \square \quad (4.29)$$

## 2. NORMÁLIS ELOSZLÁS

A  $(\xi, \eta)$  véletlen vektor normális eloszlású, ha

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp[-Q], \quad (4.30)$$

$$Q = \frac{1}{2(1-\rho^2)} \left[ \left( \frac{x-m_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x-m_1}{\sigma_1} \right) \left( \frac{y-m_2}{\sigma_2} \right) + \left( \frac{y-m_2}{\sigma_2} \right)^2 \right], \quad (4.31)$$

ahol  $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$ .

$$Q = \frac{1}{2\sigma_1^2(1-\rho^2)} \left[ x - m_1 - \rho \frac{\sigma_1}{\sigma_2} (y - m_2) \right]^2 + \frac{(y - m_2)^2}{2\sigma_2^2}. \quad (4.32)$$

$$\begin{aligned}
 f_\eta(y) &= \int_{-\infty}^{+\infty} f(x, y) dx = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-m_2)^2}{2\sigma_2^2}\right] \times \\
 &\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_1^2(1-\rho^2)} \left(x-m_1-\rho\frac{\sigma_1}{\sigma_2}(y-m_2)\right)^2\right] dx = \quad (4.33) \\
 &= \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-m_2)^2}{2\sigma_2^2}\right],
 \end{aligned}$$

mert az integrál értéke 1, hiszen egy olyan valószínűségi változó sűrűségfüggvénye, amely eloszlása

$$N\left(m_1 + \rho\frac{\sigma_1}{\sigma_2}(y-m_2), \sigma_1^2(1-\rho^2)\right). \square \quad (4.34)$$

**Megjegyzés:** 1. Rögtön látható, hogy a két perem  $N(m_1, \sigma_1^2)$  és  $N(m_2, \sigma_2^2)$  eloszlású, valamint

$$\begin{aligned}
 m_1(x) &= m_2 + \rho\frac{\sigma_2}{\sigma_1}(x-m_1), \\
 m_2(y) &= m_1 + \rho\frac{\sigma_1}{\sigma_2}(y-m_2).
 \end{aligned} \quad (4.35)$$

Tehát a regressziós függvények egyenesek.

2. Hasonló integrálással adódik, hogy  $\rho$  a korrelációs együttható.

### 4.3. Az n-dimenziós véletlen vektor.

A véletlen vektor és a hozzákapcsolódó fogalmak definícióját csak kétdimenziós esetben adtuk meg, de nagyon egyszerűen kiterjeszthetők véges sok valószínűségi változó esetére. Például, a  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változókat függetlennek nevezük, ha

$$F(x_1, x_2, \dots, x_n) = F_{\xi_1}(x_1)F_{\xi_2}(x_2) \cdots F_{\xi_n}(x_n) \quad \forall x_1, x_2, \dots, x_n \in \mathbf{R}. \quad (4.36)$$

4.6. *TÉTEL:* Az  $F(x_1, x_2, \dots, x_n)$  függvény akkor és csak akkor együttes eloszlásfüggvény, ha minden változójában balról folytonos, és

$$\lim_{x_i \rightarrow -\infty} F(x_1, x_2, \dots, x_n) = 0, \quad (i = 1, 2, \dots, n), \quad (4.37)$$

$$\lim_{x_i \rightarrow +\infty (i=1,2,\dots,n)} F(x_1, x_2, \dots, x_n) = 1, \quad (4.38)$$

$$\sum_{K=e_1+e_2+\dots+e_n} (-1)^K F(e_1 a_1 + (1-e_1)b_1, \dots, e_n a_n + (1-e_n)b_n) \geq 0 \quad (4.39)$$

$\forall a_i \leq b_i$  ( $i = 1, 2, \dots, n$ ) és az összegzést  $\forall K$  esetében vesszük, ahol az  $e_1, e_2, \dots, e_n$  értéke 0 és 1 lehet.

4.7. *TÉTEL:* Legyenek  $\xi_1, \xi_2, \dots, \xi_n$  független valószínűségi változók, melyeknek rendre  $F_{\xi_1}, F_{\xi_2}, \dots, F_{\xi_n}$  az eloszlásfüggvénye. Ekkor

(a) az  $\eta(\omega) = \max\{\xi_1(\omega), \dots, \xi_n(\omega)\}$  ( $\forall \omega \in \Omega$ ) valószínűségi változó eloszlásfüggvénye

$$F_\eta(y) = F_{\xi_1}(y) F_{\xi_2}(y) \cdots F_{\xi_n}(y). \quad (4.40)$$

(b) az  $\eta(\omega) = \min\{\xi_1(\omega), \dots, \xi_n(\omega)\}$  ( $\forall \omega \in \Omega$ ) valószínűségi változó eloszlásfüggvénye

$$F_\eta(z) = 1 - (1 - F_{\xi_1}(z))(1 - F_{\xi_2}(z)) \cdots (1 - F_{\xi_n}(z)). \quad (4.41)$$

#### 4.4. Valószínűségi változók összege.

4.8. *TÉTEL: (konvolúció)* Legyen  $(\xi, \eta)$  véletlen vektor és  $Z = \xi + \eta$ , ekkor teljesülnek a következő állítások:

(a) Ha  $\xi$  és  $\eta$  független diszkrét valószínűségi változók, amelyek mindegyikének lehetséges értékei  $0, 1, 2, \dots$ , akkor  $Z$  értékei  $k = i + j$  ( $i, j = 0, 1, 2, 3, \dots$ ) és

$$P(Z = k) = \sum_{i+j=k} P(\xi = i)P(\eta = j) = \sum_{i=0}^k P(\xi = i)P(\eta = k - i). \quad (4.42)$$

(b) Ha  $\xi$  és  $\eta$  független valószínűségi változók, akkor

$$P(Z < z) = \int_{-\infty}^{+\infty} f_{\xi}(x)F_{\eta}(z-x)dx = \int_{-\infty}^z f_Z(x)dx, \quad (4.43)$$

ahol  $Z$  sűrűségfüggvénye

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{\xi}(x)f_{\eta}(z-x)dx = \int_{-\infty}^{+\infty} f_{\xi}(z-y)f_{\eta}(y)dy. \quad (4.44)$$

4.9. *TÉTEL:* Ha  $\xi$  és  $\eta$  független nemnegatív egész értékű valószínűségi változó, akkor

$$G_{\xi+\eta}(z) = G_{\xi}(z)G_{\eta}(z). \quad (4.45)$$

4.10. *TÉTEL:* Ha  $\xi$  és  $\eta$  független valószínűségi változó, ekkor

$$\varphi_{\xi+\eta}(t) = \varphi_{\xi}(t)\varphi_{\eta}(t), \quad \forall t \in \mathbf{R}. \quad (4.46)$$

### 1. $\chi_n^2$ -ELOSZLÁS

**Definíció:** Legyen  $\xi_1, \xi_2, \dots, \xi_n \sim N(0, 1)$ , amelyek teljesen függetlenek, akkor

$$\xi = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2 \quad (4.47)$$

valószínűségi változót  $n$  szabadságfokú  $\chi_n^2$ -eloszlásúnak nevezzük. Jelölés:  $\xi \sim \chi_n^2$ .

4.11. *TÉTEL:*  $E(\xi) = n$ ,  $D^2(\xi) = 2n$ .

**Megjegyzés:** Ha  $n = 2$ , akkor  $\xi$  exponenciális eloszlású, azaz  $\xi \sim \text{Exp}(0.5)$ .

### 2. GAMMA-ELOSZLÁS

Legyen  $\alpha > 0$ ,  $\lambda > 0$ . Az  $\xi$  Gamma-eloszlású, ha a sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}, & \text{ha } x > 0, \\ 0, & \text{ha } x \leq 0. \end{cases} \quad (4.48)$$

Jelölés:  $\xi \sim \text{Gamma}(\alpha, \lambda)$ .

$$4.12. \text{ TÉTEL: } E(\xi) = \frac{\alpha}{\lambda}, D^2(\xi) = \frac{\alpha}{\lambda^2}$$

**Megjegyzés:** 1. Ha  $\alpha = 1$ , akkor éppen az exponenciális eloszlást kapjuk.

2. Független exponenciális eloszlású valószínűségi változók összege Gamma-eloszlás.

3.  $\chi_n^2$ -eloszlás:  $\alpha = \frac{n}{2}$ ,  $\lambda = \frac{1}{2}$ , azaz

$$\xi \sim \text{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right).$$

## 4.5. Egyenlőtlenségek.

4.13. *TÉTEL: (Markov-egyenlőtlenség)* Legyen a  $\eta$  nemnegatív valószínűségi változó, melynek létezik a várható értéke, ekkor  $\forall c > 0$  esetén

$$P(\eta \geq c) \leq \frac{E(\eta)}{c}. \quad (4.49)$$

**Bizonyítás:** Folytonos eset:

$$E(\eta) = \int_0^{\infty} x f(x) dx \geq \int_c^{\infty} x f(x) dx \geq \int_c^{\infty} c f(x) dx = cP(\eta \geq c). \quad (4.50)$$

Diszkrét eset:

$$\begin{aligned} E(\eta) &= \sum_{i=1}^{\infty} x_i P(\eta = x_i) \geq \sum_{x_i \geq c} x_i P(\eta = x_i) \geq \\ &\geq \sum_{x_i \geq c} c P(\eta = x_i) \geq c \sum_{x_i \geq c} P(\eta = x_i) = cP(\eta \geq c). \quad \mathbf{Q.E.D.} \end{aligned} \quad (4.51)$$

**Megjegyzés:**

$$P(\eta \leq c) \geq 1 - \frac{E(\eta)}{c}. \quad (4.52)$$

4.14. **TÉTEL: (Csebisev-egyenlőtlenség)** Ha a  $\xi$  valószínűségi változónak létezik a szórásnégyzete, akkor  $\forall \varepsilon > 0$  esetén

$$P(|\xi - E(\xi)| \geq \varepsilon) \leq \frac{D^2(\xi)}{\varepsilon^2}. \quad (4.53)$$

**Bizonyítás:** Legyen  $\eta = (\xi - E(\xi))^2$ , ekkor  $\eta \geq 0$ ,  $E(\eta) = D^2(\xi)$ . Alkalmazzuk  $\eta$ -ra a Markov-egyenlőtlenséget, ha  $c = \varepsilon^2$ . **Q.E.D.**

**Megjegyzés:**

$$P(|\xi - E(\xi)| \leq \varepsilon) \geq 1 - \frac{D^2(\xi)}{\varepsilon^2}. \quad (4.54)$$

4.15. **TÉTEL: (Jensen)** Ha  $f$  konvex függvény és  $\xi$  olyan valószínűségi változó, amelyre létezik  $E(f(\xi))$  és  $f(E(\xi))$ , akkor

$$E(f(\xi)) \geq f(E(\xi)). \quad (4.55)$$

**Bizonyítás:** Legyen  $L$  a támasztóegyenes az  $f$  függvényhez az  $(E(\xi), f(E(\xi)))$  pontban, akkor

$$E(f(\xi)) \geq E(L(\xi)) = L(E(\xi)) = f(E(\xi)). \quad \text{Q.E.D.} \quad (4.56)$$

**Megjegyzés:** 1.  $E(\xi^2) \geq E^2(\xi)$ .

2. Ha  $k \geq 1$ , akkor

$$E(|\xi|^k) \geq E^k(|\xi|). \quad (4.57)$$

4.16. **TÉTEL: (Cauchy-Bunyakovszkij-Schwarz)** Ha létezik  $E(\xi^2)$  és  $E(\eta^2)$ , és  $E(\xi) = E(\eta) = 0$ , akkor

$$|E(\xi\eta)| \leq \sqrt{E(\xi^2)E(\eta^2)}. \quad (4.58)$$

**Bizonyítás:** Legyen  $t \in \mathbf{R}$  tetszőleges, ekkor

$$0 \leq E((t\xi + \eta)^2) = t^2 E(\xi^2) + 2E(\xi\eta) + E(\eta^2). \quad (4.59)$$

Ez utóbbi a  $t$  változónak egy másodfokú kifejezése, amely sohasem negatív. Tehát, mint másodfokú egyenletnek nincs két különböző valós gyöke. Tehát a diszkrimináns nem pozitív, azaz

$$D = 4E^2(\xi\eta) - 4E(\xi^2)E(\eta^2) \leq 0. \quad (4.60)$$

Átrendezve kapjuk az állítást. **Q.E.D.**

**Megjegyzés:** 1. A bizonyításból rögtön adódik, hogy egyenlőség akkor és csak akkor, ha  $t\xi + \eta = 0$  1-valószínűséggel.

2. Ha létezik  $E(\xi)$ ,  $E(\eta)$ ,  $E(\xi^2)$  és  $E(\eta^2)$ , akkor

$$\begin{aligned} E(\xi - E(\xi)) &= 0, & D^2(\xi) &= E((\xi - E(\xi))^2), \\ E(\eta - E(\eta)) &= 0, & D^2(\eta) &= E((\eta - E(\eta))^2), \\ cov(\xi, \eta) &= E((\xi - E(\xi))(\eta - E(\eta))). \end{aligned} \quad (4.61)$$

Tehát

$$|cov(\xi, \eta)| \leq D(\xi)D(\eta), \quad (4.62)$$

azaz

$$|r(\xi, \eta)| \leq 1. \quad (4.63)$$

Egyenlőség akkor és csak akkor, ha  $\eta = a\xi + b$  ( $a, b \in \mathbf{R}$ ) 1-valószínűséggel.

## 4.6. Nagy számok gyenge törvényei.

**4.17. TÉTEL: (nagy számok gyenge törvénye)** Legyen  $\xi_1, \xi_2, \dots$  független, azonos eloszlású valószínűségi változók sorozata. Létezik a szórásnégyzet. Ekkor tetszőleges  $\varepsilon > 0$  esetén

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{\xi_1 + \dots + \xi_n}{n} - E(\xi_1)\right| \geq \varepsilon\right) = 0. \quad (4.64)$$

**Bizonyítás:**

$$E\left(\frac{\xi_1 + \dots + \xi_n}{n}\right) = E(\xi_1), \quad (4.65)$$

$$D^2 \left( \frac{\xi_1 + \dots + \xi_n}{n} \right) = \frac{D^2(\xi_1)}{n}. \quad (4.66)$$

Ekkor alkalmazzuk a Csebisev-egyenlőtlenséget

$$P \left( \left| \frac{\xi_1 + \dots + \xi_n}{n} - E(\xi_1) \right| \geq \varepsilon \right) \leq \frac{D^2(\xi_1)}{n\varepsilon^2} \rightarrow 0, \quad (4.67)$$

ha  $n \rightarrow \infty$ . **Q.E.D.**

**Megjegyzés:** Legyen  $A$  esemény,  $P(A) = p$ , és  $S_n$  az  $A$  esemény gyakorisága az első  $n$  kísérletből egy Bernoulli kísérletsorozatnál. Ekkor tetszőleges  $\varepsilon > 0$  esetén

$$\lim_{n \rightarrow +\infty} P \left( \left| \frac{S_n}{n} - p \right| \geq \varepsilon \right) = 0. \quad (4.68)$$

$S_n \sim B(n, p)$ , így

$$P \left( \left| \frac{S_n}{n} - P(A) \right| \geq \varepsilon \right) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}. \square \quad (4.69)$$

## 4.7. Polinomiális eloszlás.

Egy urnában  $n$  különböző fajtájú golyó van. Legyenek ezek a típusok

$$a_1, a_2, \dots, a_n. \quad (4.70)$$

Az  $a_i$  típus kihúzása jelentse az  $A_i$  eseményt és tudjuk, hogy

$$P(A_i) = p_i, \quad (1 \leq i \leq n). \quad (4.71)$$

Húzzunk az urnából visszatevéssel  $K$ -szor. Ekkor

$$\Omega = \{\omega | \omega = (a_{i_1}, \dots, a_{i_K})\}, \quad \text{azaz} \quad |\Omega| = n^K. \quad (4.72)$$

Legyen  $k_i$  az  $A_i$  esemény bekövetkezéseinek a száma egy adott  $\omega \in \Omega$  elemi esemény(mintarealizáció) esetén. Míg  $\xi_i$  jelentse az  $A_i$  esemény bekövetkezéseinek a számát.

Ekkor

$$P(\omega) = p_1^{k_1} p_2^{k_2} \cdot \dots \cdot p_n^{k_n}.$$

$$P(\xi_1 = k_1, \dots, \xi_n = k_n) = \frac{K!}{k_1! \dots k_n!} p_1^{k_1} p_2^{k_2} \cdot \dots \cdot p_n^{k_n}. \quad (4.73)$$

Ez utóbbi a  $(\xi_1, \dots, \xi_n)$  együttes eloszlása és polinomiális (multinomiális) eloszlásnak nevezzük.

**Megjegyzés:** 1. A polinomiális eloszlás egydimenziós peremeloszlásai binomiális eloszlások.

2.

$$P(\xi_1 = k_1, \dots, \xi_n = k_n) = p_1^{k_1} p_2^{k_2} \cdot \dots \cdot p_n^{k_n} \binom{K}{k_1} \binom{K-k_1}{k_2} \binom{K-k_1-k_2}{k_3} \dots \binom{k_n}{k_n}. \quad (4.74)$$

#### 4.8. Transzformáció n-dimenzióban.

**Definíció:** Az együttes eloszlásfüggvénye az  $(\eta_1, \dots, \eta_n)$  véletlen vektornak, amely a  $(\xi_1, \dots, \xi_n)$  véletlen vektorból áll elő úgy, hogy

$$\eta_k = g_k(\xi_1, \dots, \xi_n) \quad k = 1, 2, \dots, n \quad (4.75)$$

a következő

$$F_{\eta_1, \dots, \eta_n}(y_1, \dots, y_n) = P_{\xi_1, \dots, \xi_n}(\{\omega | g_k(\xi_1(\omega), \dots, \xi_n(\omega)) < y_k, \quad k = 1, 2, \dots, n\}). \quad (4.76)$$

Most tekintsük azt az esetet, amikor  $g_k$  ( $k = 1, 2, \dots, n$ ) függvényeknek folytonos az első parciális deriváltjuk minden  $(x_1, \dots, x_n)$  pontban úgy, hogy

$$J(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \dots & \frac{\partial g_n}{\partial x_n} \end{pmatrix} \neq 0. \quad (4.77)$$

Ha a  $(\xi_1, \dots, \xi_n)$  véletlen vektornak létezik az  $f_{\xi_1, \dots, \xi_n}$  együttes sűrűségfüggvénye és a

$$y_k = g_k(x_1, x_2, \dots, x_n), \quad k = 1, 2, \dots, n, \quad (4.78)$$

egyenletrendszernek minden  $(x_1, \dots, x_n)$  pontban pontosan egy megoldása van, akkor a  $(\eta_1, \dots, \eta_n)$  véletlen vektornak az együttes sűrűségfüggvénye

$$f_{\eta_1, \dots, \eta_n}(y_1, y_2, \dots, y_n) = f_{\xi_1, \dots, \xi_n}(x_1, x_2, \dots, x_n) |J(x_1, x_2, \dots, x_n)|^{-1}.$$

**4.2. Példa:** Legyen a  $(\xi_1, \xi_2)$  véletlen vektor sűrűségfüggvénye

$$f(x_1, x_2) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp \left\{ -\frac{(x_1 - m)^2 - 2\rho(x_1 - m)(x_2 - m) + (x_2 - m)^2}{2\sigma^2(1-\rho^2)} \right\}. \quad (4.79)$$

Tehát a  $(\xi_1, \xi_2)$  véletlen vektor normális eloszlású, ahol

$$E(\xi_1) = E(\xi_2) = m, \quad D^2(\xi_1) = D^2(\xi_2) = \sigma^2 \quad \text{és} \quad r(\xi_1, \xi_2) = \rho. \quad (4.80)$$

Készítsük el a  $\xi_1, \xi_2$  mintaelemekből (nem függetlenek) az átlagot (jelölje  $\eta_1$ ) és a tapasztalati szórásnégyzet kétszeresét (jelölje  $\eta_2$ ), azaz

$$\eta_1 = \frac{\xi_1 + \xi_2}{2}, \quad \eta_2 = \frac{(\xi_1 - \xi_2)^2}{2}. \quad (4.81)$$

Az  $(\eta_1, \eta_2)$  véletlen vektor sűrűségfüggvényének meghatározásához a megfelelő transzformációk:

$$y_1 = \frac{x_1 + x_2}{2}, \quad y_2 = \frac{(x_1 - x_2)^2}{2}. \quad (4.82)$$

Ez a transzformáció leképezi az  $A = \mathbf{R}^2$  halmazt a  $B = \{(y_1, y_2) | y_1 \in \mathbf{R}, y_2 \geq 0\}$  halmazra. De ez a transzformáció nem egyértelmű és a  $B$  halmaz minden elemének (kivéve, amikor  $y_2 = 0$ ) két elem felel meg az  $A$  halmazban. Így az inverzfüggvények két csoportja tartozik a transzformációhoz:

$$\begin{aligned} x_1 &= y_1 - \sqrt{\frac{y_2}{2}}, & x_1 &= y_1 + \sqrt{\frac{y_2}{2}}, \\ x_2 &= y_1 + \sqrt{\frac{y_2}{2}}, & x_2 &= y_1 - \sqrt{\frac{y_2}{2}}. \end{aligned} \quad (4.83)$$

## VÉLETLEN VEKTOROK

---

Továbbá az  $A$  halmaz nem írható fel két olyan diszjunkt halmaz uniójaként, amelyek mindegyike esetén a transzformációnk a  $B$  halmazra képez. A problémát az  $A$  halmaznak azok a pontjai okozzák, amelyek az  $x_1 = x_2$  egyenletű egyenesen fekszenek. Ekkor ugyanis  $y_2 = 0$ . Azonban mondhatjuk azt, hogy  $f(x_1, x_2) = 0$  minden olyan pontban, ahol  $x_1 = x_2$ . Ezt megtehetjük anélkül, hogy az eloszlás megváltozna, hiszen ezen pontok valószínűsége 0.

Legyen tehát  $A = \mathbf{R}^2 \setminus \{(x_1, x_2) | x_1 = x_2\}$ . Ez a mintatér az  $A_1 = \{(x_1, x_2) | x_2 < x_1\}$  és  $A_2 = \{(x_1, x_2) | x_2 > x_1\}$  diszjunkt halmazok uniója. Ekkor a transzformációnk kölcsönösen egyértelmű az  $A_i$  ( $i = 1, 2$ ) halmazok mindegyikéről az új

$$B = \{(y_1, y_2) | y_1 \in \mathbf{R}, y_2 > 0\} \quad (4.84)$$

halmazra. Az  $(\eta_1, \eta_2)$  véletlen vektor sűrűségfüggvénye most már meghatározható. A transzformációk Jacobi-determinánsaira teljesül, hogy

$$|J_1| = |J_2| = \frac{1}{\sqrt{2y_2}}. \quad (4.85)$$

Tehát az  $(\eta_1, \eta_2)$  véletlen vektor sűrűségfüggvénye

$$\begin{aligned} g(y_1, y_2) &= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \frac{2}{\sqrt{2y_2}} \exp \left\{ -\frac{(2-2\rho)(y_1-m)^2 + (1+\rho)y_2}{2\sigma^2(1-\rho^2)} \right\} \frac{1}{\sqrt{2y_2}} = \\ &= \frac{\sqrt{2}}{\sqrt{2\pi}\sigma\sqrt{1+\rho}} \exp \left\{ -\frac{(y_1-m)^2}{\sigma^2(1+\rho)} \right\} \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho}} \frac{1}{\sqrt{y_2}} \exp \left\{ -\frac{y_2}{2\sigma^2(1-\rho)} \right\}. \end{aligned} \quad (4.86)$$

Ebből jól látható, hogy  $\eta_1$  és  $\eta_2$  sztochasztikusan függetlenek (ez ismert független valószínűségi változók esetére). Továbbá,  $\eta_1$  eloszlása normális, amelyre

$$E(\eta_1) = m \quad \text{és} \quad D^2(\eta_1) = \frac{\sigma^2(1+\rho)}{2}. \quad (4.87)$$

Míg

$$\frac{\eta_2}{\sigma^2(1-\rho)} \quad (4.88)$$

eloszlása 1-szabadságfokú  $\chi^2$ .

Legyen  $\eta = a\xi$  ( $a > 0$ ), ekkor az eloszlásfüggvényekre illetve a sűrűségfüggvényekre teljesül, hogy

$$F_\eta(x) = F_\xi\left(\frac{x}{a}\right), \quad f_\eta(x) = \frac{1}{a}f_\xi\left(\frac{x}{a}\right), \quad (4.89)$$

azaz ha

$$f_\xi(x) = \frac{1}{\Gamma(\frac{1}{2})\sqrt{2}} \frac{1}{\sqrt{x}} e^{-\frac{x}{2}}, \quad (4.90)$$

ahol  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ , és ha  $a = \sigma^2(1 - \rho)$ , akkor

$$f_\eta(x) = \frac{1}{\sqrt{2\pi\sigma^2(1 - \rho)x}} \exp\left\{-\frac{x}{2\sigma^2(1 - \rho)}\right\}. \quad (4.91)$$

Továbbá ez azt jelenti, hogy

$$\frac{\eta_2}{2} = \frac{(\xi_1 - \xi_2)^2}{4} \quad (4.92)$$

sűrűségfüggvénye

$$\frac{1}{\Gamma(\frac{1}{2})\sigma\sqrt{1 - \rho}\sqrt{x}} \exp\left\{-\frac{y_2}{\sigma^2(1 - \rho)}\right\}. \quad (4.93)$$

A szakasz eredményeit a következőképpen összegezhetjük: Ha adott két korrelált, normális eloszlású valószínűségi változó, akkor a belőlük képzett számtani átlag és tapasztalati szórásnégyzet független. Továbbá, a tapasztalati szórásnégyzet eloszlása egy olyan 1-szabadságfokú  $\chi^2$  eloszlás, melynek a várható értéke

$$\frac{\sigma^2(1 - \rho)}{2}. \quad \square \quad (4.94)$$

## 1. STUDENT-ELOSZLÁS

**Definíció:** Legyen  $\xi_0, \xi_1, \xi_2, \dots, \xi_n \sim N(0, 1)$ , amelyek teljesen függetlenek, akkor

$$\xi = \frac{\xi_0}{\sqrt{\frac{\xi_1^2 + \xi_2^2 + \dots + \xi_n^2}{n}}} \quad (4.95)$$

valószínűségi változót  $n$  szabadságfokú Student-eloszlásúnak nevezzük.

Jelölés:  $\xi \sim t_n$ .

4.18. *TÉTEL:*  $E(\xi^\alpha)$  csak akkor létezik, ha  $\alpha < n$ . A sűrűségfüggvénye

$$f_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}. \quad (4.96)$$

**Megjegyzés:** Szokás  $t$ -eloszlásnak is nevezni.

2. Ha  $n = 1$ , akkor éppen a Cauchy-eloszlást kapjuk.

2. Ha  $n \rightarrow \infty$ , akkor a standard normális eloszlást kapjuk.

## 2. $F_{n,m}$ -ELOSZLÁS

**Definíció:** Legyen  $\xi \sim \chi_n^2$ ,  $\eta \sim \chi_m^2$  és teljesen függetlenek, ekkor

$$\zeta = \frac{m\xi}{n\eta} \quad (4.97)$$

valószínűségi változót  $n, m$  szabadságfokú  $F_{n,m}$ -eloszlásúnak nevezzük.

Jelölés:  $\zeta \sim F_{n,m}$ .

## 4.9. Centrális határeloszlás-tétel.

4.19. *TÉTEL: (centrális határeloszlás-tétel)* Legyen  $\xi_1, \xi_2, \dots$  független, azonos eloszlású valószínűségi változók sorozata és létezik az  $E(\xi_i) = \mu$  és  $D^2(\xi_i) = \sigma^2 > 0$ .

Ha  $S_n = \sum_{k=1}^n \xi_k$ , akkor

$$\lim_{n \rightarrow +\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} < x\right) = \Phi(x), \quad x \in \mathbf{R}, \quad (4.98)$$

ahol  $\Phi$  a standard normális eloszlásfüggvény.

**Megjegyzés:** Speciális esete a Moivre-Laplace tétel.

## 5. MINTAVÉTEL, ALAPSTATISZTIKÁK

A matematikai statisztika alapfeladatát nagy általánosságban a következőképpen fogalmazhatjuk meg: *”Következtetés tapasztalati adatokból események ismeretlen valószínűségeire vagy valószínűségi változók ismeretlen eloszlásfüggvényeire és ezek paramétereire.”* (Vincze, 1975). Továbbá a matematikai statisztika feladata olyan módszerek kidolgozása, amelyek segítségével tapasztalati adatokból a keresett elméleti értékekre a lehető legtöbb információt nyerhetjük. De feladata maguknak a kísérleteknek a tervezése és számuk optimalizálása is.

A statisztikai következtetés: a bekövetkezés esetlegessége. Csak a valószínűség ismert (átlagosság, az esetek 100%-a, relatív gyakoriság). Nem tudjuk megmondani, hogy bekövetkezik vagy nem.

A matematikai statisztika főbb fejezetei: becslélmélet (pont, intervallum), hipotézisvizsgálat, a mintavétel elmélete.

### 5.1. Minta, mintavétel.

Minthogy mind a hipotézisvizsgálat mind a becslélmélet következtetései tapasztalati megfigyelések alapján történik, ezért a mintavétel elmélete a matematikai statisztika alapvető és egyben bevezető fejezetének tekinthető, amelynek egyes részei csak az elmélet különböző részei során tárgyalhatók. Pl. egy kísérlet tervezése már attól függ, hogy a kísérlet kimenetele alapján milyen becslési vagy hipotézisvizsgálati módszert alkalmazunk.

**Definíció:** Az  $(\Omega, \mathcal{F}, \mathcal{P})$  hármast *statisztikai mezőnek* nevezzük, ahol

$$\mathcal{P} = \{P_\vartheta | \vartheta \in \Theta \subset \mathbf{R}^k\}. \quad (5.1)$$

**Megjegyzés:** Feladat az igazi  $\vartheta$  paraméterre való következtetés. Egy  $\xi$  valószínűségi változó kerül megfigyelésre, amelynek lehetséges értékei az  $\mathcal{X}$  mintateret alkotják és ennek bizonyos részhalmazai a  $\mathcal{B}$   $\sigma$ -algebrát. A statisztikai mező generálja hozzá a valószínűségeket. Legyenek ezek  $\mathcal{P}^*$ , ahol ha  $B \in \mathcal{B}$ , akkor

$$\{\omega | \xi(\omega) \in B\} \in \mathcal{F}, \quad P_\vartheta^*(B) = P_\vartheta(\{\omega | \xi(\omega) \in B\}). \quad (5.2)$$

Tekinthetjük ezt is statisztikai mezőnek. Valójában a következtetés  $\xi$ -ről történik  $\vartheta$ -ra.

**Definíció:** A

$$\xi_1, \xi_2, \dots, \xi_n \quad (5.3)$$

valószínűségi változók összeséget mintának nevezzük, ha azonos eloszlásúak.

**Megjegyzés:** 1. Ha a valószínűségi változók függetlenek, azonos eloszlásúak, akkor független mintának nevezzük (a legfontosabb esetekben a minta ilyen lesz).

2. Gyakorlati követelmény: jellemezze az összeséget, ahonnan származik, továbbá minnél több információ az ismeretlen eloszlásra. Hogyan biztosítható, hogy teljesüljön az azonos eloszlás, függetlenség, véletlenszerűség.

3. Megkövetelt nagyságrendek: (a) "nagy" minta (százaz nagyságrend): elméleti érték becslése, (b) "kis" minta (4-30): statisztikai hipotézis ellenőrzése (a kísérlet költséges, sokszor kell elvégezni).

4. A mintavétel módszerei: (a) egyszerű véletlen; (b) kétfokozatú, többfokozatú, szekvenciális (részsokaságok monotonitása, csomagolás, költség); (c) rétegezett, csoportos (egylépéses, kétlépéses).

**Definíció:** Az

$$x_1, x_2, \dots, x_n \quad (5.4)$$

tényleges mérési adatok összeséget mintarealizációnak nevezzük.

## 5.2. A statisztikai minta jellemzői.

**Definíció:** Legyen  $\xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathcal{X}$ , ekkor  $t(\xi)$  statisztika, ha  $t$  mérhető függvény.

**Megjegyzés:** Statisztika a mintaelemek mérhető függvénye.

A következőkben megadunk néhány használatos statisztikát:

Átlag (mintaközép):

$$\bar{\xi} = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} \quad (5.5)$$

A minta elemeit sorba rendezzük.  $\xi_1^*$  jelölje a legkisebbet. A rendezett minta:

$$\xi_1^* \leq \xi_2^* \leq \dots \leq \xi_n^*. \quad (5.6)$$

**Megjegyzés:** Ne felejtsük el, hogy függvények esetében pontonként kell alkalmaznunk a rendezést.

**Definíció:** Adott az  $F$  eloszlásfüggvény és a  $p$  valószínűség. Az  $x_p$   $p$ -kvantilis, ha  $p = F(x_p)$ . Ha  $p = 0.5$  mediánnak, míg  $0.25$  és  $0.75$  esetén alsó illetve felső kvartilisnek nevezzük.

**Megjegyzés:** Jelölje  $\xi_p$  a  $p$ -kvantilis tapasztalati megfelelőjét, azaz  $\xi_p = \xi_{[np]+1}^*$ , ekkor aszimptotikusan

$$\xi_p \sim N \left( x_p, \frac{1}{f^2(x_p)} \frac{p(1-p)}{n} \right), \quad (5.7)$$

ahol  $f$  az  $F$ -hez tartozó sűrűségfüggvény.

A medián tapasztalati megfelelője  $\xi_1, \xi_2, \dots, \xi_n$  esetén.

$$\text{med}\{\xi_i\} = \begin{cases} \xi_{m+1}^*, & \text{ha } n = 2m + 1, \\ \frac{\xi_m^* + \xi_{m+1}^*}{2}, & \text{ha } n = 2m. \end{cases} \quad (5.8)$$

Medián abszolút eltérés:

$$MAD\{\xi_i\} = \text{med}\{|\xi_i - \text{med}\{\xi_i\}|\}. \quad (5.9)$$

Mintaterjedelem:  $\xi_n^* - \xi_1^*$ .

Tapasztalati momentumok:  $\frac{\sum_{i=1}^n \xi_i^k}{n}$ .

Tapasztalati szórásnégyzet és korrigáltja:

$$s_n^2 = \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n}, \quad s_n^{*2} = \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n-1}. \quad (5.10)$$

Szórási együttható:  $\frac{s_n}{\bar{\xi}}$

(szórás nagysága az értékekhez képest,  $\xi > 0$ ).

Tapasztalati eloszlásfüggvény:

$$F_n^*(x) = \begin{cases} 0 & , \text{ ha } x \leq \xi_1^*, \\ \frac{k}{n} & , \text{ ha } \xi_k^* < x \leq \xi_{k+1}^* \quad (k = 1, 2, \dots, n-1), \\ 1 & , \text{ ha } \xi_n^* < x. \end{cases} \quad (5.11)$$

**5.1. TÉTEL: (Glivenko – a matematikai statisztika alaptétele)** Ha a  $\xi_1, \xi_2, \dots, \xi_n$  független minta, akkor

$$P\left(\lim_{n \rightarrow \infty} \sup_{-\infty < x < +\infty} |F_n^*(x) - F(x)| = 0\right) = 1.$$

**Megjegyzés:** Még sokféle statisztika használatos. Ezek közül ki kell emelni a hisztogramokat, amelyekkel majd a hipotézisvizsgálatoknál foglalkozunk.

## 6. BECSLÉSELMÉLET

A becsléselelméletben gyakran feltesszük, hogy a megfigyelt mennyiségek független valószínűségi változók, közös  $F_{\vartheta_0}$  eloszlással, amely egy meghatározott  $\{F_{\vartheta} | \vartheta \in \Theta\}$  eloszláshalmazba tartozik. A paraméterter  $\Theta$  általában az  $\mathbf{R}^k$  egy részhalmaza. Megpróbáljuk  $\vartheta_0$  értékét a megfigyelések alapján megbecsülni, azaz keressük azt a leképezést, amely az összes megfigyelések halmazát  $\Theta$ -ba képezi le és  $\vartheta_0$ -hoz közeli értéket vesz fel nagy valószínűséggel, ha  $F_{\vartheta_0}$  a valódi eloszlás.

**6.1. FELADAT:** Hány kockával dobhatunk a legnagyobb valószínűséggel egy hatost?

**Megoldás:**  $n$  kocka esetén a valószínűség

$$p_n = \frac{n5^{n-1}}{6^n}. \quad (6.1)$$

$$p_n - p_{n+1} = \frac{n5^{n-1}}{6^n} - \frac{(n+1)5^n}{6^{n+1}} = \frac{(n-5)5^{n-1}}{6^{n+1}}. \quad (6.2)$$

$n < 5$ , akkor  $p_{n+1} > p_n$ ;

$n \geq 6$ , akkor  $p_{n+1} < p_n$ ;

$n = 5$ , akkor  $p_5 = p_6 = \frac{5^5}{6^5} \cdot \square$

**6.2. FELADAT:** Állatok számának becslése az újra elfogott állatok számából.

Egy tóból kifognak 1000 halat, mindegyiket megjelölik piros ponttal, majd visszadobják a tóba. Bizonyos idő elteltével ismét kifognak 1000 halat, és közülük százon piros pontot találnak. Milyen következtetés vonható le ennek alapján a tóban lévő halak számáról?

**Megoldás:** A halak számának meghatározása a statisztikai becslések egy tipikus

példája. A lehetséges módszerek közül a visszatevés nélküli mintavétel esetére visszavezethető módszert mutatjuk meg.

Természetesen feltesszük, hogy a két halászat eredménye egy-egy véletlen minta, melyet a tó összes halaiból álló sokaságból vettünk. Feltesszük továbbá, hogy a tó halainak száma a két halászat között nem változott.

**Általános eset:**

Legyen  $n$  a tó halainak száma;  $m$  az első, míg  $r$  a második alkalommal kifogott halak száma;  $k$  a második fogás megjelölt halainak a száma. Jelölje  $p_k(n)$  annak a valószínűségét, hogy a második fogásban pontosan  $k$  megjelölt hal lesz. Ekkor

$$p_k(n) = \frac{\binom{m}{k} \binom{n-m}{r-k}}{\binom{n}{r}}. \quad (6.3)$$

A gyakorlatban  $m$ ,  $r$  és  $k$  megfigyelhető, de az  $n$  ismeretlen rögzített szám.

Biztosan tudjuk, hogy  $n \geq m + r - k$ . Példánkban  $n \geq 1900$ .

Ha feltesszük, hogy  $n = 1900$ , akkor

$$p_{100}(1900) \approx 10^{-430}. \quad (6.4)$$

Hasonlóan, ha  $n$  nagy, pl.  $n = 100000$ , akkor

$$p_{100}(100000) \approx 10^{-66}. \quad (6.5)$$

Keressük meg azt az  $n$  értéket, amelyre  $p_k(n)$  a legnagyobb értéket veszi fel, hiszen erre az  $n$ -re lesz kapott megfigyelésünk a legvalószínűbb. Jelölje ezt a számot  $\hat{n}$ . Az ilyen becslést szokás maximum likelihood becslésnek nevezni.

Tekintsük a következő hányadost:

$$\frac{p_k(n)}{p_k(n-1)} = \frac{(n-m)(n-r)}{(n-m-r+k)n}. \quad (6.6)$$

Egyszerű számolással adódik, hogy

$$\frac{p_k(n)}{p_k(n-1)} < 1, \quad \text{ha} \quad nk > mr, \quad (6.7)$$

$$\frac{p_k(n)}{p_k(n-1)} > 1, \quad \text{ha} \quad nk < mr. \quad (6.8)$$

Tehát ha  $n$  növekszik, akkor kezdetben  $p_k(n)$  nő, majd csökkenni kezd. Maximumát arra a legnagyobb  $n$  egész számra éri el, mely még kisebb, mint

$$\hat{n} = \frac{mr}{k} = 10000. \square \quad (6.9)$$

**Megjegyzés:** A valóságban  $n$  értéke esetleg ennél nagyobb vagy kisebb, így felvetődik az a kérdés, hogy az  $n$  nagy valószínűséggel milyen határok közé esik.

| $n$   | $p_{100}(n)$ |        |
|-------|--------------|--------|
| 7920  | 0.00101      |        |
| 9000  | 0.02153      |        |
| 10000 | 0.04429      | (6.10) |
| 11000 | 0.02590      |        |
| 13000 | 0.00102.     |        |

Jó lenne tudni, hogy melyek azok az  $n$  értékek, amelyek nagy eséllyel számításba jönnek, azaz mely tartományba esik az  $n$  nagy valószínűséggel (intervallumbecslés).

**Definíció:** Legyen adott a

$$\xi_1, \xi_2, \dots, \xi_n \quad (6.11)$$

minta, melynek sűrűségfüggvénye  $f$  (vagy diszkrét eloszlása  $p$ ), és ez a  $\vartheta$  paramétertől függ. Tehát adott a

$$\{f(., \vartheta) | \vartheta \in \Theta \subset \mathbf{R}^k\}. \quad (6.12)$$

*Pontbecslésnek* nevezzük a mintaelemek mérhető függvényét (statisztika), ahol a becslés és a paraméter koordinátáinak a száma megegyezik, azaz

$$\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n) \in \Theta. \quad (6.13)$$

*Intervallumbecslésnek* nevezzük a  $\Gamma$  tartományt  $1 - \alpha$  megbízhatósági szinttel, ha

$$\Gamma \subset \Theta \text{ és } P(\vartheta \in \Gamma) = 1 - \alpha. \quad (6.14)$$

**Megjegyzés:** 1. Az absztraktabb elmélet szerint legyen adott az  $(\mathcal{X}, \mathcal{F}, P_\vartheta)$ ,  $\vartheta \in \Theta \subset \mathbf{R}^k$  statisztikai tér, a  $P_\vartheta$  eloszlássereg dominált a  $\mu$  mértékkel, azaz léteznek a sűrűségfüggvények, s ekkor nem kell megkülönböztetni a diszkrét és folytonos esetek jelöléseit.

2. Probléma: Hogyan készítsünk becsléseket? Hogyan válasszunk a becslések közül? Mikor elfogadható a kiválasztott becslés?

3. Sok esetben nem a paramétert becsüljük, hanem valamilyen függvényét.

## 6.1. Pontbecslés.

**Definíció:** Legyen adott a  $\xi_1, \xi_2, \dots, \xi_n$  minta  $f(x_1, x_2, \dots, x_n; \vartheta)$  sűrűségfüggvénnyel. A  $\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n)$  (röviden  $\hat{\vartheta}_n$ ) a  $\vartheta$  paraméter *torzítatlan* becslése, ha

$$E(\hat{\vartheta}_n) = \vartheta. \quad (6.15)$$

$\hat{\vartheta}_n$  a  $\vartheta$  paraméter *aszimptotikusan torzítatlan* becslése, ha

$$\lim_{n \rightarrow \infty} E(\hat{\vartheta}_n) = \vartheta. \quad (6.16)$$

**Megjegyzés:** 1. Az átlag a várható érték torzítatlan becslése.

2. A tapasztalati szórásnégyzet nem torzítatlan becslése a szórásnégyzetnek, viszont aszimptotikusan torzítatlan. A korrigált tapasztalati szórásnégyzet torzítatlan becslése a szórásnégyzetnek.

3. Ha  $\hat{\vartheta}_n$  és  $\tilde{\vartheta}_n$  torzítatlan, akkor  $a\hat{\vartheta}_n + (1-a)\tilde{\vartheta}_n$  szintén torzítatlan.

**Definíció:** Adott a  $\hat{\vartheta}_n$  és a  $\tilde{\vartheta}_n$  torzítatlan becslés. A  $\hat{\vartheta}_n$  *hatásosabb*, ha

$$D^2(\hat{\vartheta}_n) \leq D^2(\tilde{\vartheta}_n). \quad (6.17)$$

**Definíció:** Az

$$a_1\xi_1 + a_2\xi_2 + \dots + a_n\xi_n, \quad (6.18)$$

statisztikát, ahol  $(a_i \in \mathbf{R}, i = 1, 2, \dots, n)$ , és  $a_1 + \dots + a_n = 1$ , a várható érték *lineáris becslésének* nevezzük.

**Megjegyzés:** Az átlag a leghatásosabb lineáris becslés.

**Definíció:** Legyen adott a  $\xi_1, \xi_2, \dots, \xi_n$  minta  $f(x_1, x_2, \dots, x_n; \vartheta)$  sűrűségfüggvénnyel. A minta *likelihood függvénye*:

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = f(\xi_1, \xi_2, \dots, \xi_n; \vartheta). \quad (6.19)$$

A minta *loglikelihood függvénye*:

$$l(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = -\ln(L(\vartheta; \xi_1, \xi_2, \dots, \xi_n)). \quad (6.20)$$

**Definíció:** A  $\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n)$  *elégseges* statisztika, ha

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = g(\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n; \vartheta))h(\xi_1, \xi_2, \dots, \xi_n), \quad (6.21)$$

ahol  $g$  és  $h$  megfelelő függvény.

**Megjegyzés:** A szokásos definíció a  $\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n)$  elégseges statisztika, ha a  $\forall A \in \mathcal{B}$  esetén a  $P_\vartheta(A|\hat{\vartheta}_n = x)$  feltételes valószínűségek megadhatók úgy, hogy ne fűggenek  $\vartheta$ -tól. Jelentése lényegében az, hogy  $\hat{\vartheta}_n$  a paraméterösszességre vonatkozó minden információt tartalmaz. Az általunk megadott definíciót a Neyman faktorizációs tétel alapján kapjuk.

**6.1. Példa:** Legyen  $p(x; \vartheta) = \vartheta^x(1-\vartheta)^{1-x}$ ,  $x = 0, 1$  és  $0 < \vartheta < 1$ . Adott a  $\xi_1, \xi_2, \dots, \xi_n$  független minta ezzel az eloszlással. Legyen  $\hat{\vartheta}_n = \xi_1 + \xi_2 + \dots + \xi_n$ .

**Megoldás:**

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = p(\xi_1; \vartheta) \dots p(\xi_n; \vartheta) = \binom{n}{\hat{\vartheta}_n} \vartheta^{\hat{\vartheta}_n} (1 - \vartheta)^{n - \hat{\vartheta}_n} \frac{1}{\binom{n}{\hat{\vartheta}_n}} \quad (6.22)$$

Tehát  $\hat{\vartheta}_n$  elégséges becslés  $\vartheta$ -ra.  $\square$

**Definíció:** A  $\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n)$  sorozat *konzisztens becsléssorozat* a  $\vartheta$  paraméterre, ha

$$\lim_{n \rightarrow \infty} P(|\hat{\vartheta}_n - \vartheta| > \varepsilon) = 0, \quad (6.23)$$

minden  $\varepsilon > 0$  esetén. A  $\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n)$  sorozat *erősen konzisztens becsléssorozat* a  $\vartheta$  paraméterre, ha  $\forall n$  esetén

$$E(\hat{\vartheta}_n) = \vartheta, \quad \text{és} \quad \lim_{n \rightarrow \infty} D^2(\hat{\vartheta}_n) = 0. \quad (6.24)$$

**Megjegyzés:** 1. Az erősen konzisztens becsléssorozat konzisztens.

2. Az átlagok sorozata erősen konzisztens becsléssorozat a várható értékre.

## 6.2. Pontbecslési módszerek.

Számos tulajdonság, távolság, eltérés alapján készíthetünk becsléseket. Pl. minimum  $\chi^2$  módszer, maximum entrópia módszer.

**Definíció:** Legyen adott a  $\xi_1, \xi_2, \dots, \xi_n$  minta  $f$  sűrűségfüggvénnyel. A  $\hat{\vartheta}_n$  *maximum likelihood becslés* a  $\vartheta$  paraméterre, ha

$$\max_{\vartheta} L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = L(\hat{\vartheta}_n; \xi_1, \xi_2, \dots, \xi_n). \quad (6.25)$$

**Megjegyzés:** 1. Ezzel ekvivalens, hogy

$$\min_{\vartheta} l(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = l(\hat{\vartheta}_n; \xi_1, \xi_2, \dots, \xi_n). \quad (6.26)$$

2. Ha a minta független, akkor

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = f(\xi_1, \xi_2, \dots, \xi_n; \vartheta) = f(\xi_1; \vartheta) \cdots f(\xi_n; \vartheta), \quad (6.27)$$

$$\max_{\vartheta} \prod_{i=1}^n f(\xi_i; \vartheta) = \prod_{i=1}^n f(\xi_i; \hat{\vartheta}_n), \quad (6.28)$$

$$\min_{\vartheta} - \sum_{i=1}^n \ln(f(\xi_i; \vartheta)) = - \sum_{i=1}^n \ln(f(\xi_i; \hat{\vartheta}_n)). \quad (6.29)$$

3. Általában maximum likelihood becslés nem egyezik meg a legvalószínűbb esettel.

**6.2. Példa:**  $p(x; \vartheta) = \vartheta^x (1 - \vartheta)^{1-x}$ ,  $x = 0, 1$  és  $0 < \vartheta < 1$ .

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = \vartheta^{\xi_1 + \xi_2 + \dots + \xi_n} (1 - \vartheta)^{n - (\xi_1 + \xi_2 + \dots + \xi_n)}, \quad (6.30)$$

$$l = -(\xi_1 + \xi_2 + \dots + \xi_n) \ln \vartheta - (n - (\xi_1 + \xi_2 + \dots + \xi_n)) \ln(1 - \vartheta), \quad (6.31)$$

$$\frac{dl}{d\vartheta} = -\frac{\xi_1 + \xi_2 + \dots + \xi_n}{\vartheta} + \frac{n - (\xi_1 + \xi_2 + \dots + \xi_n)}{1 - \vartheta}, \quad (6.32)$$

$$\hat{\vartheta}_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}. \square \quad (6.33)$$

**6.3. Példa:** Adott a  $\xi_1, \xi_2, \dots, \xi_n$  független minta és  $\xi_i \sim N(\mu, \sigma^2)$ . Határozzuk meg a  $\vartheta = (\mu, \sigma^2)$  maximum likelihood becslését!

**Megoldás:** Tudjuk, hogy

$$f_{\xi_i}(x; \vartheta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \quad (6.34)$$

$$L(\vartheta; \xi_1, \xi_2, \dots, \xi_n) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi_i - \mu)^2}{2\sigma^2}\right), \quad (6.35)$$

$$\begin{aligned} l &= - \sum_{i=1}^n \ln\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi_i - \mu)^2}{2\sigma^2}\right)\right) = \\ &= \frac{n}{2} \ln(2\pi) + \frac{n}{2} \ln(\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (\xi_i - \mu)^2. \end{aligned} \quad (6.36)$$

$$\begin{aligned}\frac{\partial l}{\partial \mu} &= -\frac{1}{\sigma^2} \sum_{i=1}^n (\xi_i - \mu) = 0, \\ \frac{\partial l}{\partial (\sigma^2)} &= -\frac{1}{2\sigma^4} \sum_{i=1}^n (\xi_i - \mu)^2 = 0,\end{aligned}\tag{6.37}$$

amely egyenletrendszerből a megoldás

$$\hat{\vartheta}_n = (\hat{\mu}, \hat{\sigma}^2), \text{ ahol } \hat{\mu} = \bar{\xi} = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}, \hat{\sigma}^2 = s_n^2 = \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n}. \square \tag{6.38}$$

**Definíció:** Jelölje

$$I(\vartheta) = E_{\vartheta} \left( \left( \frac{\partial f(\xi; \vartheta)}{\partial \vartheta} \right)^2 \right). \tag{6.39}$$

a Fisher-féle információmennyiséget.

**6.1. TÉTEL: (Cramer-Rao)** A  $\hat{\vartheta}_n$  torzítatlan becslés  $\vartheta$ -ra a  $\xi_1, \xi_2, \dots, \xi_n$  független minta alapján, akkor

$$D^2(\hat{\vartheta}_n) \geq \frac{1}{nI(\vartheta)}. \tag{6.40}$$

**Megjegyzés:** 1. Ha van elégséges statisztika, akkor a maximum likelihood módszer ennek valamely függvényéhez vezet.

2. Ha van minimális szórású becslés (Cramer-Rao egyenlőtlenségben egyenlőség van), akkor a maximum likelihood becslés ilyen.

3. Számos esetben megoldási és torzítási problémák vannak. Kis mintás esetben szükséges a becslések korrigálása.

4. Előfordul, hogy a maximum likelihood becslés nem egyértelmű, de kiválasztható konzisztens becsléssorozat, amely aszimptotikusan minimális szórású és normális, azaz

$$\lim_{n \rightarrow \infty} D^2(\hat{\vartheta}_n) n I(\vartheta) = 1, \tag{6.41}$$

$$\lim_{n \rightarrow \infty} P(\sqrt{nI(\vartheta)} \hat{\vartheta}_n < x) = \Phi(x), \quad \forall x \in \mathbf{R}. \quad \square \quad (6.42)$$

### 6.3. A momentumok módszere.

Ha egy valószínűségi változónak létezik a várható értéke, akkor a nagy számok törvénye alapján ha  $\xi_1, \xi_2, \dots$  független, vele azonos eloszlású valószínűségi változók sorozata, akkor a részletösszegek átlaga tart a várható értékhez. Továbbá tudjuk, hogy ha általában nem is, de elég gyenge feltételek mellett a momentumok meghatározzák az eloszlást. Ez adta az ötletet, hogy becsljük meg az elméleti momentumokat a tapasztalati momentumokkal, azaz legyen

$$\mu_k = E(\xi^k) \quad \text{és} \quad m_k = \frac{\xi_1^k + \xi_2^k + \dots + \xi_n^k}{n}, \quad (6.43)$$

$$\mu_k \approx m_k. \quad (6.44)$$

**Megjegyzés:** 1. Nyilván a megoldáshoz szükséges, hogy az egyenletek száma megegyezzen a paraméterek számával. A felhasznált momentumok  $k$  rendje legyen kicsi, mert ha  $k \geq 1$ , akkor

$$E(|\xi|^k) \geq E^k(|\xi|), \quad (6.45)$$

és a közelítés hibája szintén a momentumokkal mérhető.

2. Egyszerűen és gyorsan felírhatóak az egyenletek, teljesül a torzítatlanság. Viszont már egyszerű esetekben is gond van a momentumok létezésével. Pl. Cauchy-eloszlás,  $\xi \sim \chi_1^2$  reciproka. További gondot okozhat az egyenletrendszer megoldása. Pl. Weibull-eloszlás.  $\square$

### 6.4. A minimum $\chi^2$ módszere.

Ez a módszer, akkor alkalmazható, ha a megfigyeléseink csoportosíthatóak ( $n$  megfigyelés  $k$  csoportba). Jelölje a csoportonkénti gyakoriságokat  $n_1, n_2, \dots, n_k$ , mi alatt az elméleti valószínűségek értéke  $p_i(\vartheta)$  ( $i = 1, 2, \dots, k$ ). A gyakorlatban a  $\vartheta \in$

$\mathbf{R}^m$  értéke ismeretlen, így becsülnünk kell a mintából. Ez úgy történik, hogy minimalizálunk – egy megfelelő eltérés, távolság segítségével – a gyakoriságok és az  $np_i(\vartheta)$  ( $i = 1, 2, \dots, k$ ) várható értékek között. Most csak a következő kettőt adjuk (Pearson szerint)

$$\chi^2(\vartheta) = \sum_{i=1}^k \frac{(n_i - np_i(\vartheta))^2}{np_i(\vartheta)}, \quad (6.46)$$

$$\chi_M^2(\vartheta) = \sum_{i=1}^k \frac{(n_i - np_i(\vartheta))^2}{n_i}. \quad (6.47)$$

A becsléseket megkaphatjuk a

$$-\frac{1}{2} \frac{\partial \chi^2(\vartheta)}{\partial \vartheta_j} = \sum_{i=1}^k \left[ \frac{n_i - np_i(\vartheta)}{p_i(\vartheta)} + \frac{(n_i - np_i(\vartheta))^2}{2np_i^2(\vartheta)} \right] \frac{\partial p_i(\vartheta)}{\partial \vartheta_j} = 0 \quad (6.48)$$

egyenletrendszerből ( $j = 1, 2, \dots, m$ ). Nagy  $n$  esetén az egyenlet megoldása könnyebbé tehető a szögletes zárójelben a második tag elhagyásával.

Míg a módosított esetben a becsléseket megkaphatjuk a

$$-\frac{1}{2} \frac{\partial \chi_M^2(\vartheta)}{\partial \vartheta_j} = \sum_{i=1}^k \frac{n_i - np_i(\vartheta)}{n_i n} \frac{\partial p_i(\vartheta)}{\partial \vartheta_j} = 0 \quad (6.49)$$

egyenletrendszerből ( $j = 1, 2, \dots, m$ ).  $\square$

## 6.5. Intervallumbecslések.

Az eddigiek során arra törekedtünk, hogy megfigyeléseink alapján egyetlen értékkel becsüljük az ismeretlen paramétert. Ebben a szakaszban a feladat megadni egy  $\Gamma \subset \Theta$  tartományt, amelyre  $P(\vartheta \in \Gamma) = 1 - \alpha$ .

**Definíció:** Legyen a  $\vartheta \in \mathbf{R}$ , a  $\xi_1, \xi_2, \dots, \xi_n$  minta. A

$$\hat{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n) \leq \vartheta \leq \tilde{\vartheta}_n(\xi_1, \xi_2, \dots, \xi_n) \quad (6.50)$$

$1 - \alpha$  megbízhatóságú konfidenciaintervallum  $\vartheta$  paraméterre, ha

$$P_{\vartheta}(\hat{\vartheta}_n \leq \vartheta \leq \tilde{\vartheta}_n) = 1 - \alpha. \quad (6.51)$$

**Megjegyzés:** 1. Véletlen egy intervallum, ha legalább az egyik végpontja valószínűségi változó.

2. A konfidenciaintervallum annál jobb minnél rövidebb. Ennek mérése például várható értékkel.

3. Alkalmazások: kontrollkártyák (átlag-szórás, medián-MAD, medián-minta-terjedelem).

**6.4. Példa:**  $1 < \xi < 2 \iff 2 < 2\xi$  és  $\xi < 2$ , s így

$$P(1 < \xi < 2) = P(\xi < 2 < 2\xi). \quad (6.52)$$

**6.5. Példa:**  $\xi \sim \chi_{16}^2$ . Mennyi a valószínűsége, hogy  $26.3 \in [\xi, 3.3\xi]$ ?

**Megoldás:**

$$P(\xi < 26.3 < 3.3\xi) = P(7.97 < \xi < 26.3) \approx 0.90. \quad (6.53)$$

Az intervallum hossza:  $2.3\xi$ . Várható értéke:  $E(2.3\xi) = 2.3 \cdot 16 = 36.8$ .  $\square$

**6.6. Példa:** Konfidenciaintervallum készítése a Csebisev-egyenlőtlenség alapján. Fontos, mert ha viszonylag keveset tudunk az eloszlásról, azaz nem ismerjük "csak" azt, hogy létezik a  $E(\xi) = \mu$  és a  $D(\xi) = \sigma$ , ekkor

$$P(|\xi - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}. \quad (6.54)$$

Továbbá

$$E(\bar{\xi}) = \mu, \quad D(\bar{\xi}) = \frac{\sigma}{\sqrt{n}}. \quad (6.55)$$

Tehát

$$P(\bar{\xi} - k\sigma\sqrt{n} \leq \mu \leq \bar{\xi} + k\sigma\sqrt{n}) \geq 1 - \frac{1}{k^2}. \quad (6.56)$$

**Megjegyzés:** 1. A Csebisev-egyenlőtlenség alapján nem tudjuk biztosítani az egyenlőséget, de ha  $\alpha$  kicsi, azaz kisebb, mint 0.05, akkor már megfelelő lehet az információ tartalom.

2. A  $k = 3, k = 4$  adja a minőségellenőrzésben szokásos  $6\sigma, 8\sigma$  szabályt.

6.3. *FELADAT:* Legyen

$$\xi_1, \xi_2, \dots, \xi_n \sim N(\mu, \sigma_0^2), \quad (6.57)$$

független minta. Készítsünk  $P(x_a < \mu < x_b) = 1 - \alpha$  konfidenciaintervallumot a várható értékre, ha a  $\sigma_0$  szórás ismert!

**Megoldás:** Tudjuk, hogy

$$\frac{\bar{\xi} - \mu}{\sigma_0} \sqrt{n} \sim N(0, 1). \quad (6.58)$$

Legyen  $0 < a < b$ , ekkor

$$\Phi(a + x) - \Phi(a) > \Phi(b + x) - \Phi(b), \quad x > 0, \quad (6.59)$$

azaz

$$\Phi(b) - \Phi(a) > \Phi(b + x) - \Phi(a + x). \quad (6.60)$$

Ez alapján (standard esetben) a legrövidebb intervallum, akkor adódik, ha a 0-ra szimmetrikus.

$$P\left(\bar{\xi} - u_{\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}} \leq \mu \leq \bar{\xi} + u_{\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}}\right) = 1 - \alpha, \quad (6.61)$$

ahol a  $\sigma_0$  ismert és  $\Phi(u_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$ .  $\square$

**Megjegyzés:** A centrális határeloszlás-tétel szerint független mintára

$$\lim_{n \rightarrow \infty} P\left(\frac{\bar{\xi} - \mu}{\sigma_0} \sqrt{n} < x\right) = \Phi(x), \quad \forall x \in \mathbf{R}. \quad (6.62)$$

Tehát ha a mintaelemszám elég nagy, akkor általában alkalmazható a normális eloszlás.

6.4. *FELADAT:* Legyen

$$\xi_1, \xi_2, \dots, \xi_n \sim N(\mu, \sigma^2), \quad (6.63)$$

független minta. Készítsünk  $P(x_a < \mu < x_b) = 1 - \alpha$  konfidenciaintervallumot a várható értékre, ha a  $\sigma$  szórás nem ismert!

**Megoldás:** Tudjuk, hogy

$$\frac{\bar{\xi} - \mu}{s_n^*} \sqrt{n} \sim t_{n-1}. \quad (6.64)$$

Hasonólan az előzőhöz

$$P\left(\bar{\xi} - t_{\frac{\alpha}{2}} \frac{s_n^*}{\sqrt{n}} \leq \mu \leq \bar{\xi} + t_{\frac{\alpha}{2}} \frac{s_n^*}{\sqrt{n}}\right) = 1 - \alpha, \quad (6.65)$$

ahol az  $F_{n-1}(t_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$ .  $\square$

6.5. *FELADAT:* Legyen

$$\xi_1, \xi_2, \dots, \xi_n \sim N(\mu, \sigma^2), \quad (6.66)$$

független minta. Készítsünk  $P(x_a < \sigma^2 < x_b) = 1 - \alpha$  konfidenciaintervallumot a szórásnégyzetre!

**Megoldás:** Tudjuk, hogy

$$\frac{ns_n^2}{\sigma^2} \sim \chi_{n-1}^2. \quad (6.67)$$

Ez alapján meghatározunk egy intervallumot, de ez általában nem a legrövidebb, hiszen a  $\chi^2$ -eloszlások nem szimmetrikusak (a legrövidebb nehéz feladat).

$$P\left(\frac{ns_n^2}{b} \leq \sigma^2 \leq \frac{ns_n^2}{a}\right) = 1 - \alpha,$$

ahol a  $\chi_{n-1}^2(a) = \frac{\alpha}{2}$  és  $\chi_{n-1}^2(b) = 1 - \frac{\alpha}{2}$ .  $\square$

6.6. *FELADAT:* Legyen

$$\xi_1, \xi_2, \dots, \xi_n \sim \text{Exp}(\lambda), \quad (6.68)$$

független minta. Készítsünk  $P(x_a < \lambda < x_b) = 1 - \alpha$  konfidenciaintervallumot a paraméterére!

**Megoldás:** Független exponenciális eloszlású valószínűségi változók összege Gamma-eloszlású, azaz

$$n\bar{\xi}\lambda \sim \text{Gamma}(n, 1), \quad P\left(\frac{a}{n\bar{\xi}} \leq \lambda \leq \frac{b}{n\bar{\xi}}\right) = 1 - \alpha, \quad (6.69)$$

ahol a  $\text{Gamma}(n, 1)(a) = \frac{\alpha}{2}$  és  $\text{Gamma}(n, 1)(b) = 1 - \frac{\alpha}{2}$ .  $\square$

**Megjegyzés:** Az exponenciális eloszlású valószínűségi változó esetén elkészített konfidenciaintervallum alapján adható egy általános módszer intervallumbecslés meghatározására. Legyen

$$\xi_1, \xi_2, \dots, \xi_n \sim F_\theta, \quad (6.70)$$

független minta ( $F_\theta$  abszolút folytonos), ekkor

$$F_\theta(\xi_1), F_\theta(\xi_2), \dots, F_\theta(\xi_n) \sim U(0, 1), \quad (6.71)$$

független minta. Továbbá

$$-\ln(F_\theta(\xi_1)), -\ln(F_\theta(\xi_2)), \dots, -\ln(F_\theta(\xi_n)) \sim \text{Exp}(1). \quad (6.72)$$

Tehát

$$-\sum_{i=1}^n \ln F_\theta(\xi_i) \sim \text{Gamma}(n, 1), \quad (6.73)$$

$$P(x_a < -\sum_{i=1}^n \ln F_\theta(\xi_i) < x_b) = 1 - \alpha. \quad (6.74)$$

A megfelelő egyenletek megoldása általában nem könnyű.  $\square$

## 7. HIPOTÉZISVIZSGÁLAT

### 7.1. A hipotézisvizsgálat alapfogalmai.

Tekintsünk egy véletlen jelenséget, amelyet jellemez az  $F(., \vartheta)$  eloszlásfüggvény, ahol  $\vartheta$  (skalár vagy vektor) a paramétertérhez tartozik. A hipotézis egy egyszerű állítás, hogy  $F(., \vartheta)$  jellemzi-e a véletlen jelenséget. Az állításnak a következő kétféle típusa van:

1. Ismert az  $F(., \vartheta)$  alakja a  $\vartheta$  kivételével.
2. Az  $F$  alakja ismeretlen.

Mielőtt definiálnánk az alapfogalmakat vizsgáljuk meg a következő példát.

**7.1. Példa:** Legyen  $\xi$  egy véletlen kísérlet eredménye. Ismert, hogy  $\xi \sim N(\vartheta, 100)$ . Az eddigi tapasztalat  $\vartheta = 75$ .

*SEJTÉS:*  $\vartheta > 75$ , de lehetséges  $\vartheta \leq 75$ .

$$\begin{aligned} H_0 : \quad \vartheta \leq 75 & \quad \text{alaphipotézis,} \\ H_1 : \quad \vartheta > 75 & \quad \text{ellenhipotézis (alternatív).} \end{aligned} \tag{7.1}$$

Ez a két feltevés felosztotta a paraméterteret. Melyiket fogadjuk el, ha adott a

$$\xi_1, \xi_2, \dots, \xi_n \sim N(\vartheta, 100) \tag{7.2}$$

független minta.

**Döntés:** megadunk egy szabályt, amely minden

$$x_1, x_2, \dots, x_n \tag{7.3}$$

mintarealizáció esetén eldönti melyiket válasszuk.

Az  $\mathcal{X}$  mintatér felosztása:

$$\mathcal{X} = C \cup \overline{C} \quad (7.4)$$

Ha  $(x_1, x_2, \dots, x_n) \in C$ , akkor elutasítjuk a  $H_0$  hipotézist (elfogadjuk a  $H_1$  hipotézist).

Ha  $(x_1, x_2, \dots, x_n) \in \overline{C}$ , akkor elfogadjuk a  $H_0$  hipotézist (elutasítjuk a  $H_1$  hipotézist).

### 1. Teszt:

$$n = 25, \Omega = \{(x_1, x_2, \dots, x_{25}) | x_i \in \mathbf{R}, i = 1, 2, \dots, 25\}. \quad (7.5)$$

$$C = \{(x_1, x_2, \dots, x_{25}) | x_1 + \dots + x_{25} > 25 \cdot 75\}. \quad (7.6)$$

Az elutasítás valószínűsége:

$$P((\xi_1, \xi_2, \dots, \xi_{25}) \in C) = P(\bar{\xi} > 75) = K_1(\vartheta), \quad (7.7)$$

azaz függ a paramétertől (erőfüggvény).

$$\begin{aligned} K_1(\vartheta) &= 1 - \Phi\left(\frac{75 - \vartheta}{2}\right), \quad \bar{\xi} \sim N(\vartheta, 4). \\ K_1(73) &= 0.159, \quad K_1(75) = 0.5, \\ K_1(77) &= 0.841, \quad K_1(79) = 0.977. \end{aligned} \quad (7.8)$$

### 2. Teszt:

$$C = \{(x_1, x_2, \dots, x_{25}) | x_1 + \dots + x_{25} > 25 \cdot 78\}. \quad (7.9)$$

Az elutasítás valószínűsége:

$$\begin{aligned} P((\xi_1, \xi_2, \dots, \xi_{25}) \in C) &= P(\bar{\xi} > 78) = K_2(\vartheta). \\ K_2(\vartheta) &= 1 - \Phi\left(\frac{78 - \vartheta}{2}\right), \quad \bar{\xi} \sim N(\vartheta, 4). \\ K_2(73) &= 0.006, \quad K_2(75) = 0.067, \\ K_2(77) &= 0.309, \quad K_2(79) = 0.691. \end{aligned} \quad (7.10)$$

**3. Teszt:**  $K_3(\vartheta)$  legyen olyan, hogy

$$K_3(75) = 0.159, \quad K_3(77) = 0.841. \quad (7.11)$$

Adjunk meg ilyen tesztet!

$$C = \{(x_1, x_2, \dots, x_n) | x_1 + \dots + x_{25} > nc\}. \quad (7.12)$$

$$K_3(\vartheta) = P(\bar{\xi} > c) = 1 - \Phi\left(\frac{c - \vartheta}{10} \sqrt{n}\right), \quad \bar{\xi} \sim N(\vartheta, \frac{100}{n}). \quad (7.13)$$

$$1 - \Phi\left(\frac{c - 75}{10} \sqrt{n}\right) = 0.159, \quad 1 - \Phi\left(\frac{c - 77}{10} \sqrt{n}\right) = 0.841. \quad (7.14)$$

$$\frac{c - 75}{10} \sqrt{n} = 1, \quad \frac{c - 77}{10} \sqrt{n} = -1. \quad (7.15)$$

$$n = 100, \quad c = 76, \quad K_3(73) = 0.001, \quad K_3(79) = 0.999. \quad (7.16)$$

A megadott helyeken történt vizsgálatok alapján a 3. teszt,  $K_3$  tulajdonságai (erőfüggvénye) szerint, "kívánatosabb."  $\square$

**Definíció:** Az  $F(., \vartheta)$  egy eloszlásösszesség a mintatéren, ahol a  $\vartheta \in \Theta$  paraméter. A

$$\begin{aligned} H_0 : \quad \vartheta &\in \Theta_0, & \text{alaphipotézis,} \\ H_1 : \quad \vartheta &\in \Theta_1, & \text{ellenhipotézis (alternatív).} \end{aligned} \quad (7.17)$$

ahol  $\Theta_0 \cup \Theta_1 = \Theta$ .

**Megjegyzés:** A hipotézis aszerint egyszerű vagy összetett, hogy  $\Theta_i$  ( $i = 1, 2$ ) egy vagy több elemű.

**Definíció:** Legyen  $\xi \in \mathcal{X}$ , ahol  $\mathcal{X}$  a mintatér és  $C \subset \mathcal{X}$ . A  $C$  halmazt *kritikus tartománynak* nevezzük, ha teljesül rá, hogy

1. Ha  $\xi \in C$ , akkor elvetjük a  $H_0$  alaphipotézist és elfogadjuk a  $H_1$  ellenhipotézist.
2. Ha  $\xi \notin C$ , akkor elfogadjuk a  $H_0$  alaphipotézist.

**Definíció:** Ha adott  $C$  kritikus tartomány esetén elfogadjuk vagy elvetjük a hipotézist a paraméterre vagy az  $F$  alakjára azt *statisztikai próbának* nevezzük.

**Definíció:** Egy próbát  $\alpha$ -szintűnek (*szignifikancia szint*) nevezzük, ha

$$P(\xi \in C|H_0) \leq \alpha. \quad (7.18)$$

A  $P(\xi \in C|H_0)$  valószínűséget *elsőfajú hibának* nevezzük, a  $P(\xi \notin C|H_1)$  valószínűséget pedig *másodfajú hibának*.

Paraméteres esetben

$$\sup_{\vartheta \in \Theta_0} P(\xi \in C|\vartheta) = \alpha \quad (7.19)$$

a próba szintje.

**Definíció:** A

$$W(\vartheta) = P(\xi \in C|\vartheta) \quad (7.20)$$

függvényt a  $C$  kritikus tartományhoz tartozó *erőfüggvénynek* nevezzük.

**Megjegyzés:**  $1 - W(\vartheta) = P(\xi \in C|\vartheta)$ , ahol  $\vartheta \in \Theta_1$ , és

$$\alpha = \sup_{\vartheta \in \Theta_0} W(\vartheta). \quad (7.21)$$

**Definíció:** A próba *konzisztens*, ha az erőfüggvény egyhez tart (ha a mintaelemszám tart a végtelenhez) minden  $\vartheta \in \Theta_1$  esetén (az ellenhipotézis esetén).

**Definíció:** Egy próbát *torzítatlannak* nevezzük  $\alpha$ -szinten, ha

$$\begin{aligned} P(\xi \in C|\vartheta) &\leq \alpha, & \text{ha } \vartheta \in \Theta_0, \\ P(\xi \in C|\vartheta) &\geq \alpha, & \text{ha } \vartheta \in \Theta_1. \end{aligned} \quad (7.22)$$

**Definíció:** Az  $\alpha$ -szintű  $C_1$  és  $C_2$  kritikus tartománnyal definiált próbák közül  $C_1$ -et *jobbnek* (*erősebbnek*) nevezzük a  $C_2$ -nél, ha

$$P(\xi \in C_1|\vartheta) \geq P(\xi \in C_2|\vartheta), \quad \text{ha } \vartheta \in \Theta_1. \quad (7.23)$$

A  $C$  kritikus tartománnyal definiált  $\alpha$ -szintű próba *egyenletesen legjobb (legerősebb)*, ha jobb minden  $\alpha$ -szintű próbánál.

## 7.2. A likelihood hányados próba.

**Definíció:** A

$$\phi_C(x) = \begin{cases} 1, & \text{ha } x \in C, \\ 0, & \text{ha } x \notin C \end{cases} \quad (7.24)$$

függvényt determinisztikus próbafüggvénynek nevezzük.

**Definíció:** Legyen  $0 \leq \phi \leq 1$  mérhető függvény. Ha adott  $x \in \mathcal{X}$  esetén  $\phi(x)$  valószínűséggel utasítjuk el a  $H_0$  alaphipotézist és  $1 - \phi(x)$  valószínűséggel fogadjuk el, akkor véletlenített próbafüggvénynek nevezzük. A vele végzett próbát pedig véletlenített próbának.

**Megjegyzés:** 1. A  $T$  statisztika segítségével hozunk létre próbát, ha pl.

$$\phi(x) = \begin{cases} 1, & \text{ha } T(x) \geq a, \\ 0, & \text{ha } T(x) < a. \end{cases} \quad (7.25)$$

2. Ha  $T$  nem folytonos, akkor  $\alpha$ -terjedelmű próba konstruálásánál előfordulhat, hogy

$$\begin{aligned} (1) \quad & P(T(x) > a | H_0) < \alpha \\ (2) \quad & P(T(x) \geq a | H_0) > \alpha. \end{aligned} \quad (7.26)$$

Ekkor pontosan  $\alpha$ -szintű próbát nem tudunk determinisztikusan meghatározni. Ha azonban a próbafüggvényt a következő alakban definiáljuk:

$$\phi(x) = \begin{cases} 1, & \text{ha } T(x) > a, \\ p, & \text{ha } T(x) = a, \\ 0, & \text{ha } T(x) < a, \end{cases} \quad (7.27)$$

akkor  $p$  és  $a$  egyértelműen megválasztható úgy, hogy a próba terjedelme  $\alpha$  legyen. (1) és (2) alapján megválasztjuk  $a$  értékét. Ha valamelyikben egyenlőség áll fenn, akkor  $p = 0$ . Ellenkező esetben  $\exists 0 < p < 1$ , hogy

$$P(T(x) \geq a | H_0) + pP(T(x) = a | H_0) = \alpha. \quad (7.28)$$

Ekkor a próba terjedelme  $E(\phi(\xi)|H_0)$ .

**7.1. TÉTEL: (Neyman-Pearson)** Legyen adott egy  $\xi$  ( $\mathcal{X} = \mathbf{R}$ ) megfigyelés, amely  $f_0$  vagy  $f_1$  sűrűségfüggvényű eloszlásból származhat.

$$\begin{aligned} H_0 : f &= f_0 && \text{alaphipotézis,} \\ H_1 : f &= f_1 && \text{ellenhipotézis.} \end{aligned} \tag{7.29}$$

Ekkor létezik egy pontosan  $\alpha$ -szintű legerősebb próba, amelynek próbafüggvénye

$$\phi(x) = \begin{cases} 1, & \text{ha } \frac{f_1}{f_0} > a, \\ p, & \text{ha } \frac{f_1}{f_0} = a, \\ 0, & \text{ha } \frac{f_1}{f_0} < a, \end{cases} \tag{7.30}$$

ahol  $p$  és  $a$  úgy választhatók, hogy a próba terjedelme  $\alpha$  legyen.

**Megjegyzés:** Tegyük fel, hogy a sűrűségfüggvények folytonosak és ugyanazon a tartományon pozitívak. Ekkor a bizonyítás sokkal egyszerűbb. Legyen

$$C = \{x \mid \frac{f_1(x)}{f_0(x)} > a\}, \tag{7.31}$$

ahol  $a$  értékét az határozza meg, hogy

$$\alpha = P(\xi \in C|H_0) = \int_C f_0(x) dx. \tag{7.32}$$

Legyen  $D$  tetszőleges kritikus tartomány, amelyre

$$P(\xi \in D|H_0) \leq \alpha. \tag{7.33}$$

Megjegyezzük, hogy

$$0 \leq (\phi_C(x) - \phi_D(x)) (f_1(x) - a f_0(x)), \quad \forall x \in \mathcal{X}. \tag{7.34}$$

$$\begin{aligned} 0 &\leq \int_{\mathcal{X}} (\phi_C(x) - \phi_D(x)) (f_1(x) - a f_0(x)) dx = \\ &= P(\xi \in C|H_1) - P(\xi \in D|H_1) - a (P(\xi \in C|H_0) - P(\xi \in D|H_0)) = \\ &= P(\xi \in C|H_1) - P(\xi \in D|H_1) - a (\alpha - P(\xi \in D|H_0)) \leq \\ &\leq P(\xi \in C|H_1) - P(\xi \in D|H_1). \end{aligned} \tag{7.35}$$

Tehát  $P(\xi \notin C|H_1) \leq P(\xi \notin D|H_1)$ .  $\square$

**Definíció:** Legyen

$$L_\xi(H) = \sup_{\vartheta \in \Theta} L(\vartheta; \xi_1, \xi_2, \dots, \xi_n). \quad (7.36)$$

A  $H_0, H_1$  hipotézisek alapján definiáljuk a likelihood hányadost

$$L_\xi(H_0, H_1) = \frac{L_\xi(H_1)}{L_\xi(H_0)}. \quad (7.37)$$

**Megjegyzés:** Ha  $T$  elégséges statisztika a  $\vartheta$  paraméterre, akkor a faktorizáció szerint  $L_\xi(H_0, H_1)$  egyszerűen a  $T$  függvénye.

**Definíció:** Egy próbát, amelynek a kritikus tartománya

$$C = \{x | L_\xi(H_0, H_1) > a\} \quad (7.38)$$

alakú, valamely  $a$  konstansra, likelihood hányados próbának nevezzük.

**Megjegyzés:** 1. Az  $a$  értékét úgy határozzuk meg, hogy rögzítjük a próba  $\alpha$ -szintjét és  $\alpha = P(\xi \in C|H_0)$ .

2. A likelihood hányados próba optimális egyszerű hipotézisek esetén. A legtöbb szokásos próba likelihood hányados próba, habár más statisztikából is felépíthető.

3. Az általánosított likelihood hányados próba. Eddig diszjunkt hipotézis paramétertartományokat tekintettünk. Most legyen

$$\begin{aligned} H_0 : \quad \vartheta &\in \Theta_0, & \text{alaphipotézis,} \\ H_1 : \quad \vartheta &\in \Theta, & \text{ellenhipotézis.} \end{aligned} \quad (7.39)$$

Tegyük fel, hogy  $\Theta_1$  alapján  $k$  darab szabad paraméter van, míg  $\Theta_0$  alapján  $k - m$  darab szabad paraméter van, Jelölés:  $|\Theta_1| = k, |\Theta_0| = k - m$ .

**7.2. TÉTEL:** Tegyük fel, hogy  $\Theta_0 \subset \Theta_1$  és  $|\Theta_1| - |\Theta_0| = m$ . Ekkor megfelelő feltételek esetén, a  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$  független mintára

$$\lim_{n \rightarrow \infty} 2 \ln L_\xi(H_0, H_1) \sim \chi_m^2, \quad (7.40)$$

ha  $H_0$  igaz. Ha  $H_0$  nem igaz, akkor  $2 \ln L_\xi$  egyre nagyobb. Elutasítjuk  $H_0$ -t, ha  $2 \ln L_\xi > a$ , ahol  $\alpha = P(\chi_m^2 > a)$  egy közel  $\alpha$ -szintű próbát ad.

**Megjegyzés:** Azt mondjuk, hogy  $2 \ln L_\xi$  aszimptotikusan  $\chi_m^2$ -eloszlású.

### 7.3. Néhány általánosított likelihood hányados próba.

7.3. *TÉTEL:* Ha a

$$\xi_1, \xi_2, \dots, \xi_n, \quad \xi_i \sim N(\mu, \sigma^2), \quad (7.41)$$

független minta, akkor

$$\begin{aligned} \max_{\mu} L(\mu, \sigma^2; \xi_1, \xi_2, \dots, \xi_n) &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{2\sigma^2}\right), \\ \max_{\sigma^2} L(\mu, \sigma^2; \xi_1, \xi_2, \dots, \xi_n) &= \left(2\pi \frac{\sum_{i=1}^n (\xi_i - \mu)^2}{n}\right)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\right), \\ \max_{\mu, \sigma^2} L(\mu, \sigma^2; \xi_1, \xi_2, \dots, \xi_n) &= \left(2\pi \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n}\right)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\right). \square \end{aligned} \quad (7.42)$$

#### 1. Egymintás $u$ -próba.

Adott a  $\xi_1, \xi_2, \dots, \xi_n$ , ( $\xi_i \sim N(\mu, \sigma^2)$ ), független minta és  $\sigma^2$  ismert.

$$\begin{aligned} H_0 : \quad & \mu = \mu_0, \\ H_1 : \quad & \mu \neq \mu_0. \end{aligned} \quad (7.43)$$

$$L_\xi(H_0, H_1) = \exp\left(\frac{1}{2\sigma^2} n(\bar{\xi} - \mu_0)^2\right), \quad (7.44)$$

azaz elutasítjuk  $H_0$ -t, ha  $(\bar{\xi} - \mu_0)^2$  nagy. Ez nem meglepetés, hiszen

$$u = \frac{\bar{\xi} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1). \quad (7.45)$$

A kritikus tartomány ugyanúgy készíthető, mint a konfidenciaintervallum.

**Megjegyzés:**  $2 \ln L_\xi(H_0, H_1) = u^2 \sim \chi_1^2$ .

### 2. Kétmintás $u$ -próba.

Adott a  $\xi_1, \xi_2, \dots, \xi_m$ , ( $\xi_i \sim N(\mu_1, \sigma^2)$ ), és  $\eta_1, \eta_2, \dots, \eta_n$ , ( $\eta_i \sim N(\mu_2, \sigma^2)$ ), független minta és  $\sigma^2$  ismert.

$$\begin{aligned} H_0 : \quad & \mu_1 = \mu_2, \\ H_1 : \quad & \mu_1 \neq \mu_2. \end{aligned} \tag{7.46}$$

$$L_\xi(H_0, H_1) = \exp \left( \frac{1}{2\sigma^2} \frac{mn}{m+n} (\bar{\xi} - \bar{\eta})^2 \right), \tag{7.47}$$

azaz elutasítjuk  $H_0$ -t, ha  $(\bar{\xi} - \bar{\eta})^2$  nagy. Ezt felhasználva kapjuk, hogy

$$u = \frac{\bar{\xi} - \bar{\eta}}{\sigma} \sqrt{\frac{mn}{m+n}} \sim N(0, 1). \tag{7.48}$$

A kritikus tartomány ugyanúgy készíthető, mint az egymintás esetben.

**Megjegyzés:**  $2 \ln L_\xi(H_0, H_1) = u^2 \sim \chi_1^2$ .

### 3. Egymintás $t$ -próba.

Adott a  $\xi_1, \xi_2, \dots, \xi_n$ , ( $\xi_i \sim N(\mu, \sigma^2)$ ), független minta és  $\sigma^2$  ismeretlen.

$$\begin{aligned} H_0 : \quad & \mu = \mu_0, \\ H_1 : \quad & \mu \neq \mu_0. \end{aligned} \tag{7.49}$$

$$L_\xi(H_0, H_1) = \left( 1 + \frac{t^2}{n-1} \right)^{-\frac{n}{2}}, \tag{7.50}$$

ahol

$$t = \frac{\bar{\xi} - \mu_0}{s_n^*} \sqrt{n}, \tag{7.51}$$

azaz elutasítjuk  $H_0$ -t, ha  $t^2$  nagy. Ekkor

$$t = \frac{\bar{\xi} - \mu_0}{s_n^*} \sqrt{n} \sim t_{n-1}. \tag{7.52}$$

A kritikus tartomány ugyanúgy készíthető, mint a konfidenciaintervallum.

#### 4. Kétmintás $t$ -próba.

Adott a  $\xi_1, \xi_2, \dots, \xi_m$ , ( $\xi_i \sim N(\mu_1, \sigma^2)$ ), és  $\eta_1, \eta_2, \dots, \eta_n$ , ( $\eta_i \sim N(\mu_2, \sigma^2)$ ), független minta és  $\sigma^2$  ismeretlen.

$$\begin{aligned} H_0 : \quad & \mu_1 = \mu_2, \\ H_1 : \quad & \mu_1 \neq \mu_2. \end{aligned} \tag{7.53}$$

$$t = \frac{\bar{\xi} - \bar{\eta}}{\sqrt{\frac{(m-1)s_m^{*2} + (n-1)s_n^{*2}}{m+n-2}}} \sqrt{\frac{mn}{m+n}} \sim t_{m+n-2}. \tag{7.54}$$

A kritikus tartomány ugyanúgy készíthető, mint az egymintás esetben.

**Megjegyzés:** 1. Ha  $\sigma_1 \neq \sigma_2$ , de  $m = n$ , akkor alkalmazható az egyesített  $t$ -próba, hiszen

$$\xi_i - \eta_i \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2). \tag{7.55}$$

2. **(Scheffé)** Ha  $\sigma_1 \neq \sigma_2$ , de  $m \neq n$ . Tegyük fel, hogy  $m < n$ , és

$$Z_i = \xi_i - \sqrt{\frac{m}{n}}\eta_i + \frac{1}{\sqrt{mn}} \sum_{j=1}^m \eta_j - \bar{\eta}, \tag{7.56}$$

akkor

$$E(Z_i) = \mu_1 - \mu_2, \quad D^2(Z_i) = \sigma_1^2 + \frac{m}{n}\sigma_2^2, \quad cov(Z_i, Z_j) = 0 (i \neq j), \tag{7.57}$$

így

$$\begin{aligned} H_0 : \quad & \mu_1 - \mu_2 = 0, \\ H_1 : \quad & \mu_1 - \mu_2 \neq 0 \end{aligned} \tag{7.58}$$

esetre készíthetünk egymintás  $t$ -próbát.

#### 5. $\chi^2$ -próba.

## 7. FEJEZET

---

Adott a  $\xi_1, \xi_2, \dots, \xi_n$ , ( $\xi_i \sim N(\mu, \sigma^2)$ ), független minta és  $\sigma^2$  ismeretlen.

$$\begin{aligned} H_0 : \quad & \sigma = \sigma_0, \\ H_1 : \quad & \sigma \neq \sigma_0. \end{aligned} \tag{7.59}$$

$$L_\xi(H_0, H_1) = \left(\frac{Y}{n}\right)^{-\frac{n}{2}} \exp\left(\frac{Y-n}{2}\right), \tag{7.60}$$

ahol

$$Y = \frac{ns_n^2}{\sigma_0^2} \sim \chi_{n-1}^2. \tag{7.61}$$

A kritikus tartomány ugyanúgy készíthető, mint a konfidenciaintervallum.

### 6. $F$ -próba.

Adott a  $\xi_1, \xi_2, \dots, \xi_m$ , ( $\xi_i \sim N(\mu_1, \sigma^2)$ ), és  $\eta_1, \eta_2, \dots, \eta_n$ , ( $\eta_i \sim N(\mu_2, \sigma^2)$ ), független minták (egyástól is).

$$\begin{aligned} H_0 : \quad & \sigma_1 = \sigma_2, \\ H_1 : \quad & \sigma_1 \neq \sigma_2. \end{aligned} \tag{7.62}$$

$$L_\xi(H_0, H_1) = \frac{m^{m/2} n^{n/2}}{(m+n)^{(m+n)/2}} \frac{\left(1 + \frac{m-1}{n-1} F\right)^{(m+n)/2}}{\left(\frac{m-1}{n-1} F\right)^{m/2}}, \tag{7.63}$$

ahol

$$F = \frac{s_m^{*2}}{s_n^{*2}} \sim F_{m-1, n-1}, \tag{7.64}$$

azaz elutasítjuk  $H_0$ -t, ha  $F$  nagyon kicsi vagy nagyon nagy. Ekkor a kritikus tartomány készítése

$$\int_0^{x_a} f_F(x) dx + \int_{x_b}^{\infty} f_F(x) dx = \alpha. \tag{7.65}$$

**Megjegyzés:** Ha

$$F \sim F_{m,n}, \text{ akkor } \frac{1}{F} \sim F_{n,m}. \tag{7.66}$$

#### 7.4. A Pearson-féle $\chi^2$ statisztika és alkalmazásai.

Legyen az  $A_1, A_2, \dots, A_k$  teljes eseményrendszer. Végezzünk el  $n$  Bernoulli-kísérletet a megfigyelésükre és jelölje  $\xi_i$  az  $A_i$  gyakoriságát, ekkor  $(\xi_1, \xi_2, \dots, \xi_k)$  polinomiális eloszlású. Írjuk fel a következő hipotéziseket.

$$\begin{aligned} H_0 &: p_i = p_i(\vartheta), \quad \text{ha } \vartheta \in \Theta_0, \\ H_1 &: p_i \text{ tetszőleges.} \end{aligned} \quad (7.67)$$

Ekkor

$$2 \ln L(H_0, H_1) = 2 \sum_{i=1}^k \eta_i \ln \hat{p}_i - 2 \sum_{i=1}^k \eta_i \ln p_i(\hat{\vartheta}) = 2 \sum_{i=1}^k \eta_i \ln \left( \frac{\hat{p}_i}{p_i(\hat{\vartheta})} \right), \quad (7.68)$$

ahol  $\hat{p}_i = \frac{\eta_i}{n}$  és  $\hat{\vartheta}$  a  $\vartheta$  maximum likelihood becslése a  $H_0$  teljesülése esetén.

Vezessük be a következő jelöléseket:

$$o_i = \eta_i, \quad e_i = np_i(\hat{\vartheta}), \quad \delta_i = o_i - e_i. \quad (7.69)$$

Ekkor

$$\begin{aligned} 2 \ln L(H_0, H_1) &= 2 \sum_{i=1}^k \eta_i \ln \left( \frac{\hat{p}_i}{p_i(\hat{\vartheta})} \right) \\ &= 2 \sum_{i=1}^k o_i \ln \left( \frac{o_i}{e_i} \right) \\ &= 2 \sum_{i=1}^k (\delta_i + e_i) \ln \left( 1 + \frac{\delta_i}{e_i} \right) \\ &= 2 \sum_{i=1}^k (\delta_i + e_i) \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\delta_i^m}{m e_i^m} \\ &\approx \sum_{i=1}^k \frac{\delta_i^2}{e_i} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \end{aligned} \quad (7.70)$$

Ez utóbbit szokás Pearson-féle  $\chi^2$  statisztikának nevezni, mert ha  $H_0$  esetén  $\vartheta \in \mathbf{R}^m$  és becsljük, akkor

$$2 \ln L(H_0, H_1) \sim \chi_{k-m-1}^2 \quad (7.71)$$

aszimptotikusan.

**7.2. Példa:** Egy dobókocka dobálása során a következő gyakoriságokat kaptuk:

1 – 7db              2 – 6db              3 – 10db

4 – 6db              5 – 8db              6 – 3 db

Ekkor

$$\overline{X} = 3.275, \quad \chi^2 = 4.1 < 11.071 \approx \chi_{5,0.05}^2. \square$$

**Megjegyzés:** 1. A Pearson-féle  $\chi^2$ -próba alkalmazható tetszőleges eloszlás vizsgálatára, azaz illeszkedésvizsgálatra. Adott az  $F$  eloszlásfüggvény. Osszuk fel a  $(-\infty, +\infty)$  intervallumot. Legyen

$$-\infty = x_0 < x_1 < \dots < x_k = +\infty, \quad (7.72)$$

$$p_i = F(x_{i+1}) - F(x_i), \quad (i = 1, 2, \dots, k). \quad (7.73)$$

A felosztás módjára nincs általános szabály.

3. A szokásos alkalmazások: illeszkedésvizsgálat, függetlenség- és homogenitás-vizsgálat.

2. A hisztogram az alapstatisztikák közé tartozik, de csak most jutottunk el odáig, hogy a Pearson-féle  $\chi^2$ -próba kívánalmait szerint készítsük el.

Az  $[a, b]$  intervallum tartalmazza az adatokat.

$$a = d_0 < d_1 < \dots < d_k = b. \quad (7.74)$$

A felosztáskor figyeljük a darabszámot, kiugró értékeket és általában legyenek egyenlő hosszúak az intervallumok (kivéve a széleken). Adjuk meg a  $[d_{i-1}, d_i)$  intervallumba eső adatok számát ( $o_i$ ) minden  $i$ -re. Az  $o_i$  gyakorisággal arányos oszlopot rajzolunk a  $[d_{i-1}, d_i)$  intervallumra.

Gyakorisághisztogram:

$$\sum_{i=1}^k \frac{o_i}{d_i - d_{i-1}} (d_i - d_{i-1}) = n. \quad (7.75)$$

Sűrűséghisztogram:

$$\sum_{i=1}^k \frac{\frac{o_i}{n}}{d_i - d_{i-1}} (d_i - d_{i-1}) = 1. \square \quad (7.76)$$

### 7.5. Az $\omega^2$ -próba.

Legyen  $\xi, \eta$  két független valószínűségi változó, amelyek eloszlásfüggvénye  $F$  és  $G$ . Ekkor

$$P(\xi < \eta) = \int F dG. \quad (7.77)$$

Ha mind a  $\xi$  mind a  $\eta$  valószínűségi változó vizsgálatára adott lenne egy minta, akkor a

$$\delta_1 = \left| \int F dG - 0.5 \right| = |P(\xi < \eta) - 0.5| = 0 \quad (7.78)$$

feltevés nem más, mint a Wilcoxon-próba nullhipotézise. Tehát a  $\delta_1 = 0$  teljesülése azt fejezi ki, hogy a  $G$  eloszlásfüggvényű mintaelemekhez meghatározunk az  $F$  eloszlástípusból egy olyan hely- és skálaparaméterrel rendelkezőt, hogy a Wilcoxon-próba nullhipotézise teljesüljön.

Sajnos a  $\delta_1$ , nem viselkedik távolságként, hiszen a

$$|P(\xi < \eta) - 0.5| = 0 \quad (7.79)$$

nem csak akkor teljesül, ha  $F = G$ . Viszont észrevehetjük, hogy a  $\delta_1$  kapcsolódik az ún. Cramer-Mises

$$d(F, G) = \int_{-\infty}^{+\infty} (F(x) - G(x))^2 dG(x) \quad (7.80)$$

eltéréshez. Ehhez az eltéréshez kapcsolódik az  $\omega^2$ -illeszkedésvizsgálat, amelynek a  $\chi^2$ -próbával szemben óriási előnye, hogy nincs szükség az értékek csoportosítására, viszont a mintát növekvő sorrendbe kell rendezni.

Annak ellenőrzésére, hogy az eloszlástípust jól választottuk-e meg a Cramer-Mises eltérése alapján  $\omega^2$ -próbát is szokás használni.

Az  $F_n^*(x)$ -nek a feltételezett  $F(x)$ -től vett eltérése mértékéül az

$$\omega_n^2 = n \int_{-\infty}^{+\infty} (F(x) - F_n^*(x))^2 dF(x) \quad (7.81)$$

mennyiséget használják a következő Mises-Szmirnov tétel alapján.

**7.4. TÉTEL:** Tetszőleges  $F(x)$  folytonos eloszlásfüggvényű  $\xi$  valószínűségi változóra, minden  $x > 0$  esetén igaz, hogy

$$\lim_{n \rightarrow \infty} P(\omega_n^2 < x) = a_1(x), \quad (7.82)$$

ahol az  $a_1(x)$  függvény nem függ  $\xi$ -től.

**Megjegyzés:** 1. Az  $\omega_n^2$  statisztikát a következőképpen lehet meghatározni az  $F$  eloszlásfüggvény és a minta segítségével:

$$\omega_n^2 = \frac{1}{12n} + \sum_{k=1}^n \left( F(\xi_k^*) - \frac{k-0.5}{n} \right)^2. \quad (7.83)$$

2. Az  $a_1(x)$  eloszlásfüggvénynek csak a karakterisztikus függvénye adható meg közvetlenül használható formában, ezért a próba alkalmazásához szükséges a következő táblázat:

|          |   |      |      |      |      |      |      |      |       |       |        |
|----------|---|------|------|------|------|------|------|------|-------|-------|--------|
| $x$      | 0 | 0.12 | 0.15 | 0.18 | 0.24 | 0.35 | 0.46 | 0.74 | 0.87  | 1.17  | 1.49   |
| $a_1(x)$ | 0 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | 0.95 | 0.99 | 0.995 | 0.999 | 0.9998 |

3. Az  $\omega^2$ -próba végrehajtása: Rögzítünk egy  $p$  megbízhatósági szintet és az

$$a_1(x_p) = p$$

egyenletből kiszámítjuk a megfelelő  $x_p$  kvantilis értékét illetve a táblázat segítségével ellenőrizzük, hogy a feltételezett  $F$  eloszlásfüggvény és az  $F_n^*$  empirikus eloszlásfüggvényből kiszámított  $\omega_n^2$  milyen  $x_p$ -hez viszonyítva. Ha  $\omega_n^2 > x_p$ , akkor a feltevésünk nem fogadható el. Természetesen feltesszük, hogy az  $n$  érték elég nagy, ami a szakirodalom szerint azt jelenti, hogy  $n > 50$ .

## 7.6. Regressziós egyenesek összehasonlítása.

Adott két mérési adatsor, amelyek alapján elkészíthető két regressziós egyenes. Feladatunk annak vizsgálata mikor egyezik meg a két egyenes statisztikailag.

A regressziós egyenes meghatározása: Adott  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . Keressük azt az egyenest,

$$y = bx + a, \quad (7.84)$$

amelyre

$$Q(a, b) = \sum_{i=1}^n (y_i - bx_i - a)^2 \quad (7.85)$$

minimális.

Vezessük be a következő jelöléseket:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n}, \quad (7.86)$$

$$Q_x = \sum_{i=1}^n x_i^2 - n\bar{x}^2, \quad Q_y = \sum_{i=1}^n y_i^2 - n\bar{y}^2, \quad Q_{xy} = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}. \quad (7.87)$$

Ekkor a regressziós egyenes együtthatói

$$b_{yx} = \frac{Q_{xy}}{Q_x}, \quad a_{yx} = \bar{y} - b_{yx}\bar{x}. \quad (7.88)$$

A hibák, a konfidencia-intervallumok és a hipotézisvizsgálatok leírásához bevezetett jelölésekben az  $s$  a korrigált tapasztalati szórásra utal.

$$Q_{y \cdot x} = Q_y - b_{yx}Q_{xy} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = Q(a_{yx}, b_{yx}), \quad (7.89)$$

$$r = \frac{Q_{xy}}{\sqrt{Q_x Q_y}}, \quad s_x^2 = \frac{Q_x}{n-1}, \quad s_y^2 = \frac{Q_y}{n-1}, \quad s_{xy}^2 = \frac{Q_{xy}}{n-1}, \quad (7.90)$$

$$s_{y.x}^2 = \frac{Q_{y.x}}{n-2}, \quad s_{b_{yx}}^2 = \frac{s_{y.x}^2}{Q_x}, \quad s_{a_{yx}}^2 = s_{y.x}^2 \left( \frac{1}{n} + \frac{\bar{x}^2}{Q_x} \right). \quad (7.91)$$

Két egyenes összehasonlítása: Adott két minta. A mintaelemszámok  $n$  és  $m$ . A formulákban az 1 illetve 2 indexek az aktuális mintára utalnak. A  $b_1$  és  $b_2$  regressziós együtthatók összehasonlításához először végezzük el a következő  $F$ -próbát:

$$\frac{s_{y_1.x_1}^2}{s_{y_2.x_2}^2} > F_{(n-2, m-2; 0.05)}. \quad (7.92)$$

A további lépéseinket az  $F$ -próba alapján két irányba folytathatjuk:

1. Ha az  $F$ -próba elfogad, azaz feltehetjük, hogy a szórásnégyzetek megegyeznek, akkor a

$$\hat{t} = \frac{b_1 - b_2}{\sqrt{\frac{Q_{y_1.x_1} + Q_{y_2.x_2}}{n+m-4} \left[ \frac{1}{Q_{x_1}} + \frac{1}{Q_{x_2}} \right]}} \quad (7.93)$$

statisztika közelítőleg Student-eloszlású  $n+m-4$  szabadságfokkal.

2. Ha az  $F$ -próba elutasít, azaz feltehetjük, hogy a szórásnégyzetek különböznek, akkor elég nagy minták esetén ( $n > 20, m > 20$ ) a

$$\hat{t} = \frac{b_1 - b_2}{\sqrt{\frac{s_{y_1.x_1}^2}{Q_{x_1}} + \frac{s_{y_2.x_2}^2}{Q_{x_2}}}} \quad (7.94)$$

statisztika közelítőleg Student-eloszlású  $\nu$  szabadságfokkal, ahol

$$\nu = \frac{1}{\frac{c^2}{n-2} + \frac{(1-c)^2}{m-2}} \quad \text{és} \quad c = \frac{\frac{s_{y_1.x_1}^2}{Q_{x_1}}}{\frac{s_{y_1.x_1}^2}{Q_{x_1}} + \frac{s_{y_2.x_2}^2}{Q_{x_2}}}, \quad n \leq m. \quad (7.95)$$

**Megjegyzés:** Regressziós egyenesekre vonatkozóan sokszor csupán ezt a párhuzamossági kérdést vizsgálják. Ha elfogadtuk a párhuzamossági hipotézist, akkor a regressziós együttható most már a két minta adataiból a következőképpen becsülhető:

Hátra van még az egyenesek azonosságára vonatkozó döntés kérdése. Ezt további két részre kell bontanunk.

1. Ha  $\bar{x}_1 \neq \bar{x}_2$ , akkor a közös regressziós együttható megadható

$$\frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} \quad (7.96)$$

formában is.

2. Ha  $\bar{x}_1 = \bar{x}_2$ , akkor az  $\bar{y}_1 = \bar{y}_2$  biztosítja a regressziós egyenesek azonosságát, így azt kell megvizsgálnunk. A szokásos kétmintás t-próba alkalmas ennek eldöntésére.

Ha nem szükséges a párhuzamosság, akkor közvetlenül alkalmazható a következő F-próba:

$$\hat{F} = \frac{n + m - 4}{2} \frac{Q_{y..T} - Q_{y_1.x_1} - Q_{y_2.x_2}}{Q_{y_1.x_1} + Q_{y_2.x_2}} > F_{(2, n+m-4; 0.05)}, \quad (7.97)$$

ahol  $Q_{y..T}$  az egyesített mintára vonatkozó érték.

**7.3. Példa:** Számítások:  $n = 270$ ,  $m = 111$ ,  $Q_{y..T} = 12.9031$ ,  $Q_{y_1.x_1} = 7.7952$ ,  $Q_{y_2.x_2} = 2.7720$ ,

$$\hat{F} = 41.6683 > 3.0197 = F_{(2, 377, 0.05)}. \quad (7.98)$$

## 7.7. Robusztus statisztika.

Az eddig kidolgozott módszerek főleg olyanok voltak, amelyek valamilyen értelemben optimálisak, ha a feltételezett paraméteres modell pontosan leírja a megfigyelések eloszlását. Természetes kérdésként merül fel azonban az alkalmazhatóság, hiszen ezek a modellek pontosan szinte sohasem igazak. Ennek okait négy csoportba foglalták össze: (1) a nagy hibák előfordulása, (2) a kerekítés és osztályozás, (3) a modell csak közelítőleg érvényes, (4) eltekintve az eloszlásra vonatkozó feltevésektől

a függetlenségi feltétel (vagy valamilyen más korrelációs struktúra) csak közelítőleg teljesülhet.

A nagy hibák előfordulását pedig gyakorlatilag további két csoportba lehet osztani: (1) ritkán előforduló kiugró értékek (outliers), (2) nagyobb százalékban előforduló szennyeződések (contaminations).

Ezek a hibák általában nem hagyhatók figyelmen kívül a gyakorlatban, mivel még nagyon enyhe eltérések is teljesen elronthatják az "optimális" becslés viselkedését.

Maga a robusztusság (robustness) kifejezést G.E.P. Box (1953) használta először.

**Robusztusság:** *Sok statisztikai módszer, beleértve a valószínűségi szinteket, függ a feltételek pontosságától, pl. a vizsgált változó normális eloszlású-e. Ha a feltételek változására az eredmények csak kissé befolyásolódnak, pl. ha egy próba szignifikancia pontjai csak kissé változnak, ha a populáció lényegesen eltér a normálistól, akkor a próbát robusztusnak nevezzük. Még általánosabb értelemben egy statisztikai eljárás robusztus, ha nem nagyon érzékeny azokra a feltételekre, amelyektől függ.*

**Definíció:** Legyen  $\rho : \mathcal{X} \times \Theta \rightarrow \mathbf{R}$ . M-becslésnek nevezzük azokat a  $T_n$  becsléseket, amelyek minimalizálják a

$$\sum_{i=1}^n \rho(\xi_i, \vartheta) \quad (7.99)$$

összeget  $\vartheta$ -ra nézve adott minta esetén.

**Megjegyzés:** Nagyon sokszor azonosítják az M-becsléseket a  $\rho(x, \vartheta)$  függvény deriváltjai alapján (ha léteznek) felírt egyenletekkel. Tegyük fel, hogy a

$$\psi_m(x, \vartheta) = \frac{\partial \rho(x, \vartheta)}{\partial \vartheta_m} \quad (m = 1, 2, \dots, k) \quad (7.100)$$

parciális deriváltak léteznek, s ekkor az M-becslésekre teljesül, hogy

$$\sum_{i=1}^n \psi_m(\xi_i, \vartheta) = 0 \quad (m = 1, 2, \dots, k). \quad (7.101)$$

A rövidség kedvéért sokszor csak a  $\psi$  függvényeket használjuk az M-becslések definiálására.

**A helyparaméter becslései.**

1. Ha  $\psi(x) = x$ , akkor a megfelelő M-becslés az átlag, amelyik nem robusztus.
2. Ha  $\psi(x) = \text{sgn}(x)$ , akkor az M-becslés a medián, amely robusztus.
3. A Huber-féle becslés:

$$\psi_b(x) = x \cdot \min \left\{ 1, \frac{b}{|x|} \right\}, \quad (7.102)$$

ahol  $0 < b < +\infty$ , amelyhez tartozó becslés robusztus. Az optimális robusztus eset a normális  $F = \Phi$  eloszlásfüggvényre.

4. Sajnos a helyparaméter M-becslései általában nem skálainvariánsak, ezért szükséges a skálaparamétert valamilyen módszerrel megbecsülni. Ha gyorsan akarjuk a skálaparamétert becsülni, akkor az általánosan alkalmazott a medián abszolút eltérés (MAD) konstansszorosa, azaz

$$s = C \cdot \text{med}\{|\xi_i - \text{med}\{\xi_j\}|\} = C \cdot \text{MAD}\{\xi_i\}, \quad (7.103)$$

ahol  $\text{med}\{\xi_i\}$  a minta mediánját jelöli, míg  $C$  azt biztosítja, hogy a becslés konzisztens legyen. Például  $F = \Phi$  esetén

$$C = (\Phi^{-1}(\frac{3}{4}))^{-1}. \quad (7.104)$$

5. Általában a helyparaméter meghatározására, a paramétert meghatározó - egyenlet megoldására a következő rekurzív algoritmusok javasoltak:

(a) *Newton-módszer*:

$$T_n^{(m+1)} = T_n^{(m)} + \frac{s \sum_{i=1}^n \psi \left( \frac{\xi_i - T_n^{(m)}}{s} \right)}{\sum_{i=1}^n \psi' \left( \frac{\xi_i - T_n^{(m)}}{s} \right)}. \quad (7.105)$$

(b) *H-módszer (módosított Newton-módszer):*

$$T_n^{(m+1)} = T_n^{(m)} + \frac{s \sum_{i=1}^n \psi \left( \frac{\xi_i - T_n^{(m)}}{s} \right)}{n}. \quad (7.106)$$

(c) *Súlyozott legkisebb négyzetek módszere:*

$$T_n^{(m+1)} = \frac{\sum_{i=1}^n \xi_i w_i}{\sum_{i=1}^n w_i} = T_n^{(m)} + \frac{s \sum_{i=1}^n \psi \left( \frac{\xi_i - T_n^{(m)}}{s} \right)}{\sum_{i=1}^n w_i}, \quad (7.107)$$

ahol

$$w_i = \frac{\psi \left( \frac{\xi_i - T_n^{(m)}}{s} \right)}{\frac{\xi_i - T_n^{(m)}}{s}}, \quad (i = 1, 2, \dots, n). \quad (7.108)$$

A módszerekhez javasolt kiinduló érték:

$$T_n^{(0)} = \text{med}\{\xi_i\}. \quad (7.109)$$

### A skálaparaméter becslései.

1. A maximum-likelihood becslés esetén

$$\psi(x) = -x \frac{f'(x)}{f(x)} - 1. \quad (7.110)$$

Tudjuk, hogy a regularitási feltételek mellett az ehhez tartozó becslésnek a legkisebb az aszimptotikus szórásnégyzete. De például  $F = \Phi$  esetén  $\psi(x) = x^2 - 1$ , amely nem robusztus.

2. A medián abszolút eltérés viszont robusztus, amelynél

$$\psi(x) = \text{sgn}(|x| - 1). \quad (7.111)$$

3. A Huber-féle becslés a skálaparaméterre a helyparaméterre vonatkozó becslés alapján készült, azaz

$$\psi(x) = \psi_b^2(x) - \int \psi_b^2(y) d\Phi(y), \quad (7.112)$$

ahol  $\psi_b(x)$  a helyparaméter becsléséhez bevezetett  $\psi$  függvény. Ez robusztus.

4. A skálaparaméter becslésére csak egy általánosan jól használható algoritmust javasol a szakirodalom:

$$[s_n^{(m+1)}]^2 = \frac{1}{(n-1)\beta} \sum_{i=1}^n \psi_b^2\left(\frac{\xi_i - \vartheta}{s_n^{(m)}}\right) [s_n^{(m)}]^2, \quad (7.113)$$

ahol  $\beta = \int \psi_b^2(x) d\Phi(x)$  és  $\vartheta$  a helyparaméter becslése. A skálaparaméterre kiinduló értéként javasolt az

$$s_n^{(0)} = 1.483 \cdot MAD. \quad (7.114)$$

## 7.8. Kiugró érték.

### 7.4. Példa:

#### 7.1. táblázat: Enyhe és extrém kiugró érték.

|     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 30  | 305 | 409 | 470 | 522 | 585 | 640 | 766 | 860  |
| 171 | 306 | 411 | 480 | 527 | 592 | 656 | 792 | 869  |
| 184 | 322 | 436 | 482 | 548 | 592 | 668 | 792 | 918  |
| 201 | 322 | 437 | 487 | 550 | 607 | 707 | 794 | 925  |
| 212 | 336 | 439 | 494 | 559 | 616 | 709 | 802 | 953  |
| 250 | 346 | 441 | 495 | 560 | 618 | 719 | 818 | 991  |
| 265 | 351 | 444 | 499 | 570 | 621 | 737 | 830 | 1000 |
| 270 | 370 | 448 | 503 | 572 | 629 | 739 | 832 | 1005 |
| 272 | 390 | 451 | 514 | 574 | 637 | 752 | 843 | 1068 |
| 289 | 404 | 453 | 521 | 578 | 638 | 758 | 858 | 1441 |

median =  $\frac{n+1}{2}$  -dik legnagyobb adat

= a 45-dik és 46-dik átlaga =  $(559+560)/2=559.5$ .

Alsó kvartilis =  $0.25(n+1)$  -dik adat

=22.75-dik adat =  $411 + 0.75(436 - 411) = 429.75$ .

Felső kvartilis =  $0.75(n+1)$  -dik adat

=68.25-dik adat =  $739 + 0.25(752 - 739) = 742.25$ .

Kvartilis terjedelem = Felső kvartilis - Alsó kvartilis = 312.5

Alsó belső határ = Alsó kvartilis -  $1.5 \star$  Kvartilis terjedelem = -39

Felső belső határ = Felső kvartilis +  $1.5 \star$  Kvartilis terjedelem = 1211.

Alsó külső határ = Alsó kvartilis -  $3 \star$  Kvartilis terjedelem = -507.75.

Felső külső határ = Felső kvartilis +  $3 \star$  Kvartilis terjedelem = 1679.75.

Enyhe kiugró érték = 1441.

**7.5. Példa:** Kiugró érték (Grubbs).

A rendezett minta:

$$\xi_1^* \leq \xi_2^* \leq \dots \leq \xi_n^*. \quad (7.115)$$

Felső oldali kiugró statisztika:

$$T_n = \frac{\xi_n^* - \bar{\xi}}{s_n^*}. \quad (7.116)$$

Alsó oldali kiugró statisztika:

$$T_1 = \frac{\bar{\xi} - \xi_1^*}{s_n^*}. \quad (7.117)$$

**7.6. Példa:** Adatok: 10.2, 9.5, 10.1, 10.3, 9.8, 9.9, 11.9, 10.0

$$\bar{X} = 10.2, \quad s_n^* = 0.726, \quad T_n = \frac{11.9 - 10.2}{0.726} = 2.34 > 2.22$$

## HIPOTÉZISVIZSGÁLAT

---

Kritikus értékek (0.01 szint):

$$\begin{aligned} n = 7, & \quad 2.10; & n = 8, & \quad 2.22; & n = 9, & \quad 2.32; & n = 10, & \quad 2.41; \\ n = 11, & \quad 2.48; & n = 12, & \quad 2.55; & n = 13, & \quad 2.61; & n = 14, & \quad 2.66. \\ n = 90, & \quad 3.563 \end{aligned}$$

$$\frac{1441 - 576}{237.6} \approx 3.641 > 3.563$$

0.05 szint:  $n = 90$ , 3.171

**7.7. Példa:** Grubbs-féle statisztikák:

$$G_1 = \frac{|\bar{X} - \xi_j^*|}{s_n^*}, \quad G_2 = \frac{\xi_n^* - \xi_1^*}{s_n^*}, \quad G_3 = 1 - \left( \frac{(n-3) \cdot s_{n-2}^{*2}}{(n-1) \cdot s_n^{*2}} \right). \quad (7.118)$$

Adatok:

47.876 47.997 48.065 48.118 48.151 48.211 48.251 48.559

48.634 48.711 49.005 49.166 49.484

$$\bar{x} = 48.479, \quad s_n^* = 0.498,$$

$|x_i - \bar{x}|$  : 0.603, 0.482, 0.414, 0.361, 0.328, 0.268, 0.228, 0.080, 0.155, 0.232, 0.526, 0.687, 1.005.

$s_n^{*2} = 0.123$ , így a statisztikák értékei és a kritikus értékek ( $\alpha = 0.05$ )

$$\begin{aligned} G_1 &= \frac{49.484 - 48.479}{0.498} = 2.02 < 2.331, \\ G_2 &= \frac{49.484 - 47.876}{0.498} = 3.23 < 4.00, \\ G_3 &= 1 - \frac{10 \cdot 0.123}{12 \cdot 0.498^2} = 0.587 < 0.6705. \end{aligned} \quad (7.119)$$

Kritikus értékek a függelék táblázatában (Grubbs).

## 8. VÉLETLEN SZÁMOK, SZIMULÁCIÓ

### 8.1. Monte Carlo módszerek.

**Definíció:** *Monte Carlo módszereknek* nevezzük matematikai feladatok megoldásának véletlen mennyiségek modellezését felhasználó numerikus módszereit (Szobol).

A véletlen számok fontos szerepet játszanak a véletlen helyzetek generálásában (pénzérme, dobókocka, szerencsejátékok), amelyek segítenek a tapasztalatszerzésben a valószínűségről és annak törvényszerűségeiről. Ne felejtjük el, hogy a valószínűség-számítás fogalmai, tételei feltételezik, hogy az elemzés tömegjelenségre vonatkozik.

A véletlen számok legfontosabb alkalmazása a szimuláció, amely lehetővé teszi a tapasztalatszerzést a véletlenről, drága kísérletek modellezését, más módszerrel nehezen kiszámítható értékek meghatározását stb.

Kezdetben ehhez vagy egyszerű kísérleteket (kockadobás) végeztek vagy előre-gyártott táblázatokot alkalmaztak. De az összetettebb jelenségek lefutásának vizsgálatához 200...500 kockadobás is szükséges. További problémákat vet fel az egyenletesség és a különböző típusú "kockák" elkészítése. Ma már legtöbbször számítógépes algoritmusokat használunk.

A szimuláció alapvető problémái: egy determinisztikus számítógépen közelítjük a véletlent. Diszkrétel közelítünk folytonosat vagy fordítva. Végtelen feladat korlátos modell, véges szimuláció. A tapasztalatszerzésben fontos szerepet játszanak a sztochasztikus játékok. Néhány egyszerű játék:

**Ingadozás:**  $4 \times 4$  mezőre 8 fehér és 8 fekete golyót helyezünk. A fej vagy írás dönt arról, hogy fehér vagy fekete golyót helyettesítünk ellenkező színű golyóval.

(Gyorsítás: a körülzárt golyók egyszerre változtatják a színüket.)

**Egyensúly:**  $8 \times 8$  mezőre 32 fehér és 32 fekete golyót helyezünk. Két oktaéder feldobásával meghatározunk egy mezőt, az itt található golyót levesszük és egy ellenkező színűt rakunk a helyére.

**Minden vagy semmi:** a játék az előző elrendezésben indul, a szabály azonban most más: a megcímzett golyó "megduplázódik", vagyis leveszünk egy ellenséges golyót, és a saját színűt tesszük ennek a helyére is.

## 8.2. Pszeudovéletlen számok.

Azokat az  $x_1, x_2, \dots, x_n$  számokat, amelyeket egy adott algoritmus alapján számítottunk ki, és a véletlen számok helyett használhatók, pszeudovéletlen számoknak nevezzük. A generálásuknak és ellenőrzésüknek (egyenletesség, véletlenszerűség) külön elmélete alakult ki. Ezzel itt nem foglalkozunk. A legtöbb magasszintű számítógépi programozási nyelv elég jó generátort tartalmaz beépített eljárásként. Azért ajánlatos az ellenőrzés. Itt most két egyszerű pszeudovéletlenszám generátort adunk meg.

**8.1. Példa:**  $x_1 = 1, \quad x_{n+1} \equiv 125x_n \pmod{8192}.$

**8.2. Példa:**  $x_1 = 1, \quad x_{n+1} \equiv 16807x_n \pmod{2147483647}.$

**Megjegyzés:** Diszkrét egyenletes (klasszikus valószínűségi mező) eloszlást közeleltítenek a megadott rekurzív algoritmusok. Az első még számítógép nélkül is jól használható.

A számítógépi algoritmusok legtöbbször (valójában mindig diszkrétet, hiszen véges a számábrázolásuk) a  $[0, 1]$  intervallumon egyenletes eloszlást próbálják közelíteni, mert ebből különböző módszerek segítségével – a tanult eloszlások tulajdonságainak felhasználásával – más eloszlású véletlen számokat tudunk előállítani.

Pl. az egyenletes eloszlás adja a geometriai valószínűségi mező elméleti alapját. Továbbá, ha  $F$  szigorúan monoton növekvő eloszlásfüggvény és  $\xi \sim F$  eloszlású, akkor

$\eta = F(\xi)$  egyenletes eloszlású a  $[0, 1]$  intervallumon. Fordítva, ha  $\xi \sim U(0, 1)$ , akkor  $\eta = F^{-1}(\xi)$  éppen  $F$  eloszlású.

- 8.3. Példa:**
1. Ha  $\xi \sim U(0, 1)$ , akkor  $\eta = (b - a)\xi + a \sim U(a, b)$ .
  2. Ha  $\xi \sim U(0, 1)$ , akkor  $\eta = -\frac{1}{\lambda} \ln(\xi) \sim \text{Exp}(\lambda)$ .
  3. Ha  $\xi \sim U(0, 1)$ , akkor  $\eta = \text{tg}(\pi(\xi - 0.5))$  standard Cauchy eloszlású.
  4. Ha  $\xi \sim U(0, 1)$ , akkor  $\eta = \Phi^{-1}(\xi)$  standard normális eloszlású.

Milyen eloszlásúak lehetnek az 1. fejezetben megadott adatok?

**8.4. Példa:** Ha  $\xi_i \sim U(0, 1)$  ( $i = 1, \dots, 12$ ), akkor  $\eta = \sum_{i=1}^{12} \xi_i - 6$  közelítőleg standard normális eloszlású.

**8.1. FELADAT:** Lewis Carroll *Pillow Problems* c. híres könyvében az 58. probléma a következő: számoljuk ki, mennyi annak a valószínűsége, hogy a síkon véletlenül kiválasztott három pont tompaszögű háromszöget alkot. Adjunk becslést erre a valószínűségre úgy, hogy szimulációt végzünk a számítógépen!

**Megoldás:** (*A lehetséges egyik változat.*) Legyen a háromszög leghosszabb oldala  $AB$ , amelynek a hosszát jelölje  $c$ . Az  $AB$  oldal azon oldalán (félsíkon), ahol a háromszög fekszik rajzoljuk meg az  $AFB$  félkört (középpontja  $AB$  felezőpontja). Továbbá  $A$  és  $B$  középponttal és  $c$  sugárral rajzoljuk meg azokat a  $BDC$ ,  $AEC$  íveket, amelyek a  $C$  pontban metszik egymást. Nyilvánvaló, hogy a háromszög csak az  $ABDCE$  tartományban fekszik. Ha benne van az  $AFB$  félkörben, akkor tompaszögű, egyébként nem. Tehát

$$\frac{\text{az } AFB \text{ félkör területe}}{\text{az } ABDCE \text{ területe}} = \frac{\pi \frac{c^2}{8}}{2 \frac{c^2 \pi}{6} - \frac{\sqrt{3}}{4} c^2} = \frac{3\pi}{8\pi - 6\sqrt{3}} \approx 0.6394. \quad (8.1)$$

Ha egy számítógépes környezetben  $n = 10^6$  szimulációt végzünk, akkor a tompaszögű háromszögek aránya 0.7249 (a standard hiba 0.00045).  $\square$

### 8.3. A közelítő integrálás hibája.

Az egyszerű Monte-Carlo módszer esetén a hibabecslés jellemzésére általában a szórást használjuk.

Legyen  $h$  egy tetszőleges valós függvény, amely esetén az

$$\int_{-\infty}^{\infty} h^2(x) dF(x) \quad (8.2)$$

létezik. Ez szükséges és elégséges feltétele, hogy az  $\eta = h(\xi)$  valószínűségi változó, ahol  $\xi$   $F$  eloszlásfüggvényű, szórásnégyzete létezzen. Továbbá legyen

$$E(h(\xi)) = \mu, \quad \text{és} \quad D^2(h(\xi)) = \sigma^2, \quad (8.3)$$

akkor a  $\xi_1, \xi_2, \dots, \xi_n$  minta esetén ( $\xi_i$   $F$  eloszlású) a hibabecslés szórásnégyzete

$$D^2(\varepsilon) = \frac{1}{n^2} D^2(h(\xi_1) + h(\xi_2) + \dots + h(\xi_n)) = \frac{\sigma^2}{n}. \quad (8.4)$$

Ebből leolvashatjuk a Monte-Carlo módszer egy igen lényeges tulajdonságát: ha a mintaelemek számát növeljük a hiba illetve a jellemzését adó szórás csak  $\sqrt{n}$  arányában csökken. Látszólag ez azt jelenti, hogy azok a jó becslések, amelyeknek kicsi a szórása. De azzal, hogy a robusztus tulajdonságok nem változnak meg egy konstans tényező hatására az következik, hogy más szempontból kell összehasonlítani az integrálási tulajdonságokat, illetve érzékenységeket. Ezeket a további vizsgálatokat célszerű úgy elvégezni, hogy a szórással legyenek egyenlőek a becsléseknél. Legyen ez a közös érték 1, s az ilyen egyenletet nevezzük *kanonikus egyenletnek*.

8.2. FELADAT: Hány darab véletlen számot kell generálni ahhoz, hogy az

$$I = \int_0^{\frac{\pi}{2}} \sin x dx \quad (8.5)$$

integrált megbecsüljük úgy, hogy a becslés abszolút hibája legfeljebb  $I$  0.1% legyen legalább 0.99 valószínűséggel?

**Megoldás:** Tudjuk, hogy

$$\frac{2}{\pi}I = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \sin x dx = \int_0^{\frac{\pi}{2}} \sin x \frac{2}{\pi} dx = E(\sin \xi), \quad (8.6)$$

ahol  $\xi \sim U\left(0, \frac{\pi}{2}\right)$ . Tehát  $I$  egy közelítő értéke

$$I_n = \frac{\pi}{2n} \sum_{i=1}^n \sin \xi_n, \quad (8.7)$$

ahol  $\xi_n$  pszeudovéletlenszám a  $\left(0, \frac{\pi}{2}\right)$  intervallumból. Felhasználva, hogy

$$\frac{I_n - I}{D(I_n)} \sim N(0, 1), \quad (8.8)$$

ahol

$$d^2(I_n) = \frac{\pi^2}{4n} D^2(\sin \xi) = \frac{\pi^2 - 8}{8n} \quad (8.9)$$

kapjuk, hogy  $n \approx 1550579$ .  $\square$

#### 8.4. Kombinációk, permutációk generálása.

Legyen  $n, k \in \mathbf{N}$ ,  $n \geq k$ . Adjunk olyan módszert, melynek segítségével ismétlés nélkül felsorolhatjuk  $n$  elem  $k$ -adosztályú kombinációit!

*Leszámlálási módszer:* Felsoroljuk azokat az  $n$  hosszúságú bináris sorozatokat, melyekben  $k$  darab 1-es van és a többi 0-ás.

Kiindulás:  $(1, \dots, 1, 0, \dots, 0)$  – az első  $k$  db 1-es a többi 0.

Tekintsünk tetszőleges  $\tilde{\alpha} = (\alpha_1, \dots, \alpha_n)$  bináris sorozatot és megmutatjuk, hogy melyik a leszámolásban utána következő  $\tilde{\alpha}' = (\alpha'_1, \dots, \alpha'_n)$  sorozat.

Megkeressük a legkisebb olyan  $j$  számot, melyre  $\alpha_j = 1$  és  $\alpha_{j+1} = 0$ . Ekkor legyen  $\alpha'_j = 0$  és  $\alpha'_{j+1} = 1$  (jobbra mozgatjuk az egyest) és az összes,  $j$ -nél kisebb koordinátán található egyest annyival balra mozgatjuk, amennyire csak lehetséges.

11000  $\rightarrow$  10100  $\rightarrow$  01100  $\rightarrow$  10010  $\rightarrow$  01010  $\rightarrow$  00110  $\rightarrow$  10001  $\rightarrow$  01001  $\rightarrow$  00101  $\rightarrow$  00011

**8.3. FELADAT:** Ha az alaphalmazunkat  $n$  elem  $k$ -adosztályú kombinációi alkotják, akkor hogyan válasszunk ki ezek közül egyet véletlenszerűen?

*Kombinatorikai megfeleltetési módszer:*

Tetszőleges  $m$ -nek ( $0 \leq m < \binom{n}{k}$ ) egyértelműen megfeleltethető egy

$$\tilde{\beta}(m) = (\beta_1, \dots, \beta_k) \quad (8.10)$$

vektor, melyre

$$\beta_1 > \beta_2 > \dots > \beta_k; \quad m = \binom{\beta_1}{k} + \binom{\beta_2}{k-1} + \dots + \binom{\beta_k}{1}. \quad (8.11)$$

$\beta_i$  értékeinek meghatározása:  $\beta_1$  a legnagyobb egész szám, amelyre

$$m \geq \binom{\beta_1}{k}. \quad (8.12)$$

Ha a  $\beta_1, \dots, \beta_i$  számokat már meghatároztuk, akkor  $\beta_{i+1}$  a legnagyobb egész szám, amelyre

$$m - \binom{\beta_1}{k} - \binom{\beta_2}{k-1} - \dots - \binom{\beta_i}{k-i+1} \geq \binom{\beta_{i+1}}{k-i}. \quad (8.13)$$

Az egyértelműség a  $\beta_i$ -k meghatározásából rögtön adódik.

$\tilde{\alpha}(m) = (\alpha_1, \dots, \alpha_n)$  meghatározása: Álljanak az  $\tilde{\alpha}_m$ -ben az egyesek a  $\beta_1 + 1, \beta_2 + 1, \dots, \beta_k + 1$  számú koordinátákon.

**8.5. Példa:**  $n = 6, k = 3, m = 17$ , azaz határozzuk meg a 17. kombinációt!

**Megoldás:**  $\binom{6}{3} = 20, 17 \geq \binom{5}{3} = 10, 10 \geq \binom{4}{2} = 6, 1 = \binom{1}{1} = 1,$

$\tilde{\beta}(17) = (5, 4, 1), \tilde{\alpha}(17) = (010011),$  A kombináció:  $\{2, 5, 6\}.$

**Megjegyzés:** 1. Ezek alapján viszonylag könnyű azonos részhalmazok közül véletlenszerűen kiválasztani egyet, azaz például visszatevés nélküli mintavételt generálni.

2. Lehetőségünk van bizonyos tulajdonságú részhalmazok felsorolására és megszámlálására. Például kis mintás esetben a Wilcoxon-próba kritikus értékeinek meghatározására.

8.4. *FELADAT:* Adjunk olyan módszert, melynek segítségével ismétlés nélkül felsorolhatjuk  $n$  elem permutációit!

**Megoldás:** Legyen  $m$  nemnegatív egész szám és  $n = n(m)$  a legkisebb olyan egész, melyre  $m < n!$ . Először megmutatjuk, hogy pontosan egyféleképpen feleltethető meg  $m$ -nek egy olyan

$$\tilde{\alpha}(m) = (\alpha_1, \alpha_2, \dots, \alpha_{n-1}) \quad (8.14)$$

vektor, amelyre

$$m = \alpha_1 \cdot 1! + \alpha_2 \cdot 2! + \dots + \alpha_{n-1} \cdot (n-1)!, \quad (8.15)$$

ahol  $0 \leq \alpha_i \leq i$  ( $i = 1, 2, \dots, n-1$ ).

Legyen

$$\alpha_{n-1} = \left\lfloor \frac{m}{(n-1)!} \right\rfloor, \quad (8.16)$$

és legyenek az  $\alpha_{n-1}, \alpha_{n-2}, \dots, \alpha_{n-i}$  együtthatók már definiáltak. Ekkor

$$\alpha_{n-(i+1)} = \left\lfloor \frac{m - \alpha_{n-1}(n-1)! - \alpha_{n-2}(n-2)! - \dots - \alpha_{n-i}(n-i)!}{(n-i-1)!} \right\rfloor. \quad (8.17)$$

Az előállítás egyértelműségét indirekt úton bizonyítjuk be. Feleljen meg valamely  $m$ -nek két vektor:  $\tilde{\alpha}(m) = (\alpha_1, \alpha_2, \dots, \alpha_{n-1})$  és  $\tilde{\beta}(m) = (\beta_1, \beta_2, \dots, \beta_{n-1})$ . Legyen  $j$  a legnagyobb olyan koordináta száma, melyben a vektorok különböznek. Az általánosítás megszorítása nélkül tekinthetjük úgy, hogy  $\alpha_j < \beta_j$ . Ekkor

$$0 = \sum_{i=1}^{n-1} \beta_i i! - \sum_{i=1}^{n-1} \alpha_i i! \geq (\beta_j - \alpha_j) j! - \sum_{i=1}^{j-1} \alpha_i i! \geq j! - \sum_{i=1}^{j-1} i \cdot i! > 0. \quad (8.18)$$

Ellentmondásra jutottunk.

Példa  $\tilde{\alpha}(m)$  vektorokra:  $\tilde{\alpha}(15) = (1, 1, 2)$ ,  $\tilde{\alpha}(23) = (1, 2, 3)$ ,  $\tilde{\alpha}(37) = (1, 0, 2, 1)$ .

*Algoritmus:* Minden  $m$  egész számnak  $0 \leq m < n!$  megfeleltetjük az  $1, \dots, n$  egész számok egy permutációját. Az  $\tilde{\alpha}(m) = (\alpha_1, \alpha_2, \dots, \alpha_{n-1})$  vektor segítségével megmutatjuk az  $1, \dots, n$  számok helyét a permutációban. Az 1 számot az  $1 + \alpha_{n-1}$  számú helyre tesszük. Tegyük fel, hogy az  $1, \dots, i$ , ( $i = 1, \dots, n-1$ ), számokat már elhelyeztük.  $n-i$  darab hely üresen maradt. A szabad helyeket megszámozzuk az  $1, 2, \dots, n-i$  számokkal. Az  $i+1$  számot az  $1 + \alpha_{n-i-1}$  számú helyre rakjuk.. Ha az  $1, 2, \dots, n-1$  számokat elhelyeztük, akkor csak egy hely marad a permutációban az  $n$  számára. Példa:  $m = 15$ ,  $n = 4$ ,  $\tilde{\alpha}(15) = (1, 1, 2)$ , a permutáció:  $(4, 2, 1, 3)$ .  $\square$

**Megjegyzés:** Ezek alapján viszonylag könnyű a permutációk közül véletlenszerűen kiválasztani egyet, azaz például elkészíteni egy szimulációs programot a 2. fejezetben vizsgált "Szinbád" típusú problémákra..

## 9. FÜGGELÉK

### 9.1. Jelölések, rövidítések.

$\mathbf{N}$  – a természetes számok halmaza (pozitív egészek).

$\mathbf{Z}$  – az egész számok halmaza.

$\mathbf{Q}$  – a racionális számok halmaza.

$\mathbf{R}$  – a valós számok halmaza.

$\mathbf{R}^2 = \{(x, y) | x, y \in \mathbf{R}\}$ .

$A \subset B$  – az  $A$  részhalmaza a  $B$ -nek.

$A \cap B$  – az  $A$  és  $B$  halmaz közös része.

$A \cup B$  – az  $A$  és  $B$  halmaz összes eleme egy halmazban.

$\overline{A}$  – az alaphalmaz  $A$  halmazon kívüli elemei.

$A \setminus B = A \cap \overline{B}$ .

$k \ll n$  –  $k$  sokkal kisebb, mint  $n$ .

$[a]$  –  $a$  egész része.

$a := b$  –  $a$  legyen egyenlő  $b$ .

$f(\cdot) : D \rightarrow R$  – az  $f$  leképezés,  $D$  az értelmezési tartomány, a "pont" a változót helyettesíti.

$f(D)$  – az  $f$  leképezés értékkészlete.

$F(a+0)$  – a jobboldali határérték, azaz  $\lim_{x \rightarrow a+0} F(x)$ .

$F(a-0)$  – a baloldali határérték, azaz  $\lim_{x \rightarrow a-0} F(x)$ .

$\exp(x) = e^x$ .

Q.E.D. – quod erat demonstrandum – ezt kellett bizonyítani.

$\square$  – egy levezetés vagy magyarázat vége.

$(a) \Rightarrow (b)$  –  $(a)$ -ból következik  $(b)$ .

$\exists$  – létezik olyan.

$\forall$  – bármely (minden egyes).

$a \equiv b \pmod{m}$  –  $a - b$  osztható  $m$ -mel.

## 9.2. Halmazelméleti alapfogalmak.

A halmaznak és a halmaz elemének fogalmát csak axiomatikus módszerrel lehet definiálni, ezért szemléletünk alapjám ismertnek tekintjük a fogalmakat. Az  $a \in A$  azt jelöli, hogy  $a$  eleme az  $A$  halmaznak. A halmazt elemeinek, vagy elemei tulajdonságainak felsorolásával adjuk meg.

**Definíció:** Az  $A$  halmazt a  $B$  halmaz részhalmazának nevezzük, ha bármely  $a \in A$  esetén  $a \in B$  is igaz. Jele:  $A \subset B$ .

**Definíció:** Az  $A$  halmaz egyenlő a  $B$  halmazzal, ha  $A$ -nak és  $B$ -nek ugyanazok az elemei. Jele:  $A = B$ .

**Megjegyzés:**  $A = B$ , akkor és csak akkor, ha  $A \subset B$  és  $B \subset A$ .

**Definíció:** Ha egy halmaznak nincs eleme üres halmaznak nevezzük. Jele:  $\emptyset$ .

**Definíció:** Két halmaz uniójának nevezzük azoknak az elemeknek a halmazát, amelyek a két halmaz közül legalább az egyiknek elemei.

Jele:  $A \cup B$ , azaz  $A \cup B = \{x | x \in A \text{ vagy } x \in B\}$ .

**Definíció:** Két halmaz metszetének nevezzük azoknak az elemeknek a halmazát, amelyek mindkét halmaznak elemei.

Jele:  $A \cap B$ , azaz  $A \cap B = \{x | x \in A \text{ és } x \in B\}$ .

**Definíció:** Két halmaz különbségének nevezzük a kisebbítendő halmaznak azokat az elemeit, amelyek a kivonandónak nem elemei.

Jele:  $A \setminus B$ , azaz  $A \setminus B = \{x | x \in A, \text{ de } x \notin B\}$ .

**Definíció:** Ha  $B \subset A$ , akkor az  $A \setminus B$  halmazt a  $B$  halmaz  $A$ -ra vonatkozó komplementer halmazának (komplementerének) nevezzük. Jele:  $\overline{B}$ .

**Megjegyzés:** Ha egyértelmű, hogy a komplementer melyik halmazra vonatkozik, akkor egyszerűen a  $\overline{B}$  jelölést használjuk.

**Definíció:** Az  $A$  és  $B$  halmazok diszjunktak, ha  $A \cap B = \emptyset$ .

**Definíció:** Az  $\{(a, b) | a \in A, b \in B\}$  halmazt két halmaz Descartes-féle szorzatának nevezzük. Jele:  $A \times B$ .

**Definíció:** Relációnak nevezzük két halmaz Descartes-féle szorzatának egy részhalmazát.

**Definíció:**  $S \subset A \times A$  relációt ekvivalenciának nevezzük, ha

1. minden  $a \in A$  esetén  $(a, a) \in S$ , (reflexív)
2. ha  $(a, b) \in S$ , akkor  $(b, a) \in S$ , (szimmetrikus)
3. ha  $(a, b) \in S$  és  $(b, c) \in S$ , akkor  $(a, c) \in S$ , (transzitiv).

**Definíció:** Osztályozásnak nevezzük egy halmaz olyan nem üres részhalmazainak a halmazát, amelyek diszjunktak, és egyesítésük az eredeti halmaz. A szóbanforgó részhalmazokat osztályoknak nevezzük.

**9.1. TÉTEL:** Egy halmaz osztályozásánál az osztályok önmagukkal való Descartes-féle szorzatainak egyesítése ekvivalencia. Fordítva: egy halmazon adott

ekvivalenciánál az egymással relációban levő elemek olyan részhalmazok, amelyek osztályozást jelentenek.

Néhány egyszerű azonosság (műveleti tulajdonság):

$$A \cup B = B \cup A, \quad A \cap B = B \cap A, \quad \overline{\overline{A}} = A, \quad (9.1)$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}, \quad (\text{De Morgan}) \quad (9.2)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad (9.3)$$

$$B \setminus A = B \cap \overline{A}, \quad B \setminus \bigcap_{i=1}^n A_i = \bigcup_{i=1}^n (B \setminus A_i). \quad (9.4)$$

### 9.3. A függvény fogalma, számosságok.

**Definíció:** Az  $F \subset A \times B$  reláció függvény, ha bármely  $a \in A$  legfeljebb egy  $(a, b) \in F$  rendezett párban szerepel.

Jele: pl.  $F$ ,  $F : A \rightarrow B$ .

**Definíció:** Az  $F$  függvény  $A$  alaphalmazát a függvény értelmezési tartományának nevezzük, ha bármely  $a \in A$  szerepel valamelyik  $(a, b) \in F$  párban.

Jele:  $\text{Dom}(F)$ .

**Definíció:** Az  $F$  függvény  $B$  alaphalmazát a függvény értékkészletének nevezzük, ha bármely  $b \in B$  szerepel valamelyik  $(a, b) \in F$  párban.

Jele:  $\text{Range}(F)$ .

**Definíció:** Az  $F : A \rightarrow B$  függvényt kölcsönösen egyértelműnek vagy invertálhatónak nevezzük, bármely  $b \in B$  legfeljebb egy  $(a, b) \in F$  rendezett párban szerepel.

**Definíció:** Az invertálható  $F : A \rightarrow B$  függvény elemeiben a sorrend felcserélésével előálló  $(b, a)$  alakú párok halmazát az  $F$  függvény inverz függvényének nevezzük.

Jele:  $F^{-1}$ , vagy  $F^{-1} : B \rightarrow A$ .

**Definíció:** Legyen az  $F : A \rightarrow B$  függvény  $B$  értékkészlete egyenlő a  $G : B \rightarrow C$  függvény értelmezési tartományával. Ekkor a

$$G(F(a)) = \{(a, c) | a \in A, (a, b) \in F, (b, c) \in G\} \quad (9.5)$$

függvényt összetett függvénynek nevezzük.

**Megjegyzés:** Az általunk használt számhalmazok között fennáll a következő kapcsolat:

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}. \quad (9.6)$$

**Definíció:** Két halmazt azonos számosságúnak nevezünk, ha létezik  $F : A \rightarrow B$ -re ( $\text{Range}(F) = B$ ) kölcsönösen egyértelmű függvény, azaz létezik  $F^{-1} : B \rightarrow A$ . Az  $A$  halmaz számosságát  $|A|$ -val jelöljük.

9.2. *TÉTEL:* A természetes számok és a racionális számok számossága megegyezik.

9.3. *TÉTEL:* A  $[0, 1] \subset \mathbf{R}$  halmaz és a természetes számok számossága nem egyezik meg.

**Definíció:** Az  $A$  halmaz számossága nagyobb, mint a  $B$  halmaz számossága, ha számosságuk nem egyenlő, de  $\exists C \subset A$  úgy, hogy  $C$  és  $B$  számossága egyenlő.

**Megjegyzés:** Az előzőek alapján  $|\mathbf{N}| < |\mathbf{R}|$ .

**Definíció:** A  $\mathbf{N}$  számosságát megszámlálhatóan végtelennek,  $\mathbf{R}$  számosságát pedig kontinuum számosságnak nevezzük. A véges számosságokat megszámlálhatónak nevezzük.

## 9.4. Kombinatorikai alapok.

A matematika azon területe, amely egy véges halmaz elemeinek csoportosításával, kiválasztásával, sorrendbe rakásával illetve az elemek megszámlálásával foglalkozik.

Legyen

$$A = \{a_1, a_2, \dots, a_n\}, \quad B = \{b_1, b_2, \dots, b_m\}, \quad (9.7)$$

ahol  $|A| = n$  és  $|B| = m$ . Ekkor

$$|A \times B| = n \cdot m. \quad (9.8)$$

Ezt szokás a kombinatorika *szorzási alapelvének* tekinteni. Nyilvánvalóan

$$\max\{n, m\} \leq |A \cup B| \leq n + m, \quad (9.9)$$

amelyben a baloldalon egyenlőség van, ha  $A \subset B$  vagy  $B \subset A$ . Míg a jobboldalon akkor van egyenlőség, ha  $A \cap B = \emptyset$ . Ezt szokás a kombinatorika *összeadási alapelvének* nevezni.

Az  $A$  halmaz  $k$ -szoros ( $k \in \mathbf{N}$ ) Descartes-szorzatát röviden  $A^k$ -nal jelöljük.

**Definíció:** Ha  $x \in A^k$ , akkor az  $x$  elemet az  $A$  halmazhoz tartozó  $n$  elem  $k$ -ad osztályú ismétléses variációjának nevezzük.

**Megjegyzés:** Másképpen az  $A$  halmazból  $k$ -szor választunk visszatevéssel és a kihúzás sorrendje is számít. Az  $n$  elem  $k$ -ad osztályú ismétléses variációinak a száma

$$|A^k| = n^k \quad (9.10)$$

a szorzási alapelv alapján.

Legyen

$$x \in (A^2 \setminus \{(a, a) | a \in A\}), \quad (9.11)$$

akkor  $x$  egy olyan rendezett pár, amelyben az elemek azonos halmazból vannak, de nem ismétlődhetnek. Ekkor

$$|A^2 \setminus \{(a, a) | a \in A\}| = n^2 - n = n(n - 1). \quad (9.12)$$

**Definíció:** Ha  $x \in A^k$  ( $k \leq n$ ) és az  $x$  elemei mind különbözőek, akkor az  $x$  elemet az  $A$  halmazhoz tartozó  $n$  elem  $k$ -ad osztályú variációjának nevezzük.

**Megjegyzés:** Másképpen az  $A$  halmazból  $k$ -szor választunk visszatevés nélkül és a kihúzás sorrendje is számít. Az  $n$  elem  $k$ -ad osztályú variációinak a száma

$$n \cdot (n-1) \cdots (n-k+1) \quad (9.13)$$

a szorzási alapelv és az előző másod osztályú eset alapján.

**Definíció:** Ha  $k = n$ , akkor a variációt *permutációnak* nevezzük.

**Megjegyzés:** A permutáció lényegében az  $A$  halmaz elemeinek a felsorolása, egy sorrendje. Felfogható, mint egy  $g$  függvény, ahol  $g : A \rightarrow A$ -ra. A permutációk száma

$$n \cdot (n-1) \cdots 1 = n!. \quad (9.14)$$

**Definíció:** Egy  $n$  elemű halmaz  $k$  elemű részhalmazát  $n$  elem  $k$ -ad osztályú *kombinációjának* nevezzük.

**Megjegyzés:** Másképpen az  $A$  halmazból  $k$ -szor választunk visszatevés nélkül és a kihúzás sorrendje sem számít. Az  $n$  elem  $k$ -ad osztályú kombinációinak a száma

$$\frac{n(n-1) \cdots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}, \quad (9.15)$$

hiszen  $k!$  darab olyan variáció van, amelyik ugyanazt az egy kombinációt adja.

**9.4. TÉTEL: (binomiális)** Ha  $a, b \in \mathbf{R}$  és  $n \in \mathbf{N}$ , akkor

$$(a+b)^n = \sum_{k=1}^n \binom{n}{k} a^k b^{n-k}. \quad (9.16)$$

**Megjegyzés:**

$$\sum_{k=1}^n \binom{n}{k} = 2^n. \quad (9.17)$$

**Definíció:** Ha egy  $n$  elemű halmazból visszatevéssel kiválasztunk  $k$  elemet úgy, hogy a sorrend nem számít, akkor azt  $n$  elem  $k$ -ad osztályú *ismétléses kombinációjának* nevezzük.

**Megjegyzés:** Tekintsük a következő problémákat:

1. Hány megoldása van az

$$x_1 + x_2 + \dots + x_n = k$$

egyenletnek, ha  $n, k \in \mathbf{N}$ , és  $x_i$  ( $i = 1, 2, \dots, n$ ) nemnegatív egész?

2. Adott  $n$  darab doboz és  $k$  darab egyforma golyó. Hányféleképpen helyezhetjük el a golyókat a dobozokba, ha egy dobozba több golyó is kerülhet, és csak az számít, hogy egy dobozban a végén hány golyó van és az nem, hogy mikor került bele?

3. Van  $n + 1$  darab egyforma pálcikánk és  $k$  darab egyforma karikánk. Helyezzük el a pálcikákat és a karikákat egy egyenesre úgy, hogy az első és az utolsó pálcika. Hányféleképpen lehetséges ez?

**Megoldás:** Kezdjük a 3. problémával. Valójában  $n + k - 1$  darab ( $n - 1$  pálcika,  $k$  karika) elemet kell elhelyeznünk. Ha megadjuk a kraikák helyét, akkor a maradék helyeken lesznek a pálcikák. Az egyformaság pedig azt jelenti, hogy nem kell figyelniük a sorrendre. Tehát az elhelyezések száma  $n + k - 1$  elem  $k$ -ad osztályú kombinációinak a számával egyezik meg, azaz

$$\binom{n + k - 1}{k}. \quad (9.18)$$

Ha a 2. probléma esetén szorosan egymás mellé tesszük a dobozokat és kivesszük a dobozok alját, akkor  $n + 1$  darab dobozzal vagy másképpen pálcika, s a 2. probléma megegyezik a 3. problémával.

Jelölje a 2. problémában az  $i$ -edik dobozba került golyók számát  $x_i$  ( $i = 1, 2, \dots, n$ ), ami azt jelenti, hogy minden dobozokba való elhelyezés megad egy megoldást az egyenlethez.

Tekintsük át még egyszer a 2. problémát. Minden golyó elhelyezésekor az  $n$  doboz egyikét választhatjuk és minden alkalommal ugyanazon az  $n$  doboz közül választunk. Tehát a  $k$  darab golyó elhelyezése  $n$  dobozba pontosan megegyezik  $n$  elem  $k$ -ad osztályú ismétléses kombinációjával.  $\square$

### 9.5. Néhány alapvető matematikai fogalom és állítás.

**Definíció:** Az  $f : [a, b] \rightarrow \mathbf{R}$  ( $a, b \in \mathbf{R}$  függvény abszolút folytonos, ha  $\forall \varepsilon > 0$  esetén  $\exists \delta > 0$ , hogy ha

$$\sum_{i=1}^n |b_i - a_i| < \delta, \quad (9.19)$$

akkor

$$\sum_{i=1}^n |f(b_i) - f(a_i)| < \varepsilon. \quad (9.20)$$

**Definíció:** Egy tulajdonság majdnem mindenütt teljesül, ha annak a mértéke, ahol nem teljesül 0.

**Definíció:** Legyen  $0 < \alpha < \infty$  esetén

$$\Gamma(\alpha + 1) = \int_0^\infty t^\alpha e^{-t} dt. \quad (9.21)$$

Ezt nevezzük *Gamma*-függvénynek.

**9.5. TÉTEL:** A *Gamma*-függvényre teljesül, hogy

- (1)  $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$ ,
- (2)  $\Gamma(n + 1) = n!$ , ha  $n \in \mathbf{N}$ ,
- (3)  $\ln \Gamma$  konvex a  $(0, \infty)$  intervallumon.

**Megjegyzés:** Néhány speciális érték  $\Gamma(1) = 1$ ,  $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ .

**Definíció:** Legyen  $U$  egy intervallum (zárt, nyílt, félig zárt). Az  $f : U \rightarrow \mathbf{R}$  konvex függvény, ha

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b), \quad (9.22)$$

ahol  $a, b \in U$ ,  $\lambda + \mu = 1$ ,  $\lambda \geq 0$  és  $\mu \geq 0$ .

**9.6. TÉTEL:** 1. Ha  $f$  és  $g$  konvex függvény és  $\alpha \geq 0$ ,  $\beta \geq 0$ , akkor  $\alpha f + \beta g$  szintén konvex.

2. Véges sok konvex függvény összege is konvex.

3. Konvex függvények egy konvergens sorozatának a (pontonkénti) határa is konvex.

4. Ha  $f : U \rightarrow \mathbf{R}$  konvex függvény és  $a < x < b$ , akkor

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}. \quad (9.23)$$

Ha  $f$  szigorúan konvex, akkor az egyenlőtlenségek is azok.

5. Ha  $f : U \rightarrow \mathbf{R}$  konvex függvény és  $a < c < b$ , akkor létezik a bal- és jobboldali derivált minden  $c$  esetén. Továbbá,  $f'_-$  és  $f'_+$  monoton nemcsökkenő és

$$f'_-(c) \leq f'_+(c). \quad (9.24)$$

Ezenkívül minden  $x \in U$  esetén

$$f(x) \geq f(c) + f'_-(c)(x - c), \quad f(x) \geq f(c) + f'_+(c)(x - c), \quad (9.25)$$

azaz a konvex függvény minden pontjához létezik egyenes ( amely az adott ponton keresztül megy), amely a görbe alatt marad vagy legfeljebb érinti azt.

6. Az  $f : U \rightarrow \mathbf{R}$  konvex függvény folytonos az intervallum minden belső pontjában.

7. Legyen  $U$  nyílt és  $f$  kétszer differenciálható, akkor  $f$  konvex akkor és csak akkor, ha  $f'' > 0$  minden  $x \in U$ .

## 9.6. Táblázatok.

A standard normális eloszlásfüggvénye

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du, \quad x \in \mathbf{R} \quad \Phi(-x) = 1 - \Phi(x). \quad (9.26)$$

Ha  $\xi$  standard normális eloszlású, akkor az  $\eta = \sigma\xi + m$  ( $m \in \mathbf{R}$ ,  $\sigma > 0$ ) valószínűségi változó  $F$  eloszlásfüggvényére jellemző, hogy

$$F(x) = \Phi\left(\frac{x - m}{\sigma}\right). \quad (9.27)$$

### 9.1. táblázat: Néhány standard normális eloszlás érték.

| $x_p$ | $\Phi(x_p) = p$ | $P(m - x_p\sigma < \eta < m + x_p\sigma)$ |
|-------|-----------------|-------------------------------------------|
| 1     | 0.8413447460    | 0.682689492                               |
| 1.96  | 0.9750021049    | 0.950004210                               |
| 2     | 0.9772498680    | 0.954499736                               |
| 3     | 0.9986501020    | 0.997300204                               |
| 4     | 0.9999683288    | 0.999936658                               |
| 6     | 0.9999999990    | 0.999999998                               |

A normális eloszláshoz kapcsolódik a hibafüggvény

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du$$

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x),$$

azaz

$$\Phi(x) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) \right].$$

$$\begin{aligned}
 \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)n!} \\
 &= \frac{2}{\sqrt{\pi}} \left( x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - \dots \right). \\
 \operatorname{erfc}(x) &= \frac{e^{-x^2}}{x\sqrt{\pi}} \left[ 1 + \sum_{n=1}^{\infty} (-1)^n \frac{2n!}{n!(2x)^{2n}} \right] \\
 &= \frac{e^{-x^2}}{x\sqrt{\pi}} \left( 1 - \frac{1}{2x^2} + \frac{3}{4x^4} - \frac{15}{8x^6} + \frac{105}{16x^8} - \dots \right).
 \end{aligned}$$

Az eloszlásfüggvény közelítésére egy  $10^{-9}$  pontosságú polinomiális közelítést alkalmazhatunk. A közelítő polinom:

$$p(x) = c_0 + c_1x + c_2x^2 + \dots + c_8x^8. \quad (9.28)$$

### 9.2. táblázat: A közelítő polinom együtthatói.

| intervallum | $[0, 1.5]$       | $(1.5, 3]$       | $(3, 6]$         |
|-------------|------------------|------------------|------------------|
| $c_0$       | 0.4999999853197  | 0.5300774546729  | -0.1621966195471 |
| $c_1$       | 0.3989437251038  | 0.2799241265723  | 1.8137844596010  |
| $c_2$       | -0.0000232473822 | 0.2005701987176  | -1.2430841874817 |
| $c_3$       | -0.0663495262607 | -0.2504062323459 | 0.4883401215203  |
| $c_4$       | -0.0004071645564 | 0.0949343858651  | -0.1201986229749 |
| $c_5$       | 0.0105643510048  | -0.0131657278224 | 0.0189705569006  |
| $c_6$       | -0.0003504976933 | -0.0009270280158 | -0.0018738388405 |
| $c_7$       | -0.0012947802876 | 0.0004671302299  | 0.0001058586660  |
| $c_8$       | 0.0002619054865  | -0.0000383458376 | -0.0000026175074 |

Az eloszlásfüggvény inverzének a közelítésére egy  $10^{-15}$  pontosságú racionális törtfüggvény közelítést alkalmazhatunk.

### 9.3. táblázat: Standard normális eloszlás inverze (Pascal részlet).

## 9. FEJEZET

---

```
function Invphi(var x:extended):extended; var szi,ni,ui:extended;
begin ui:=x; if (ui<0) or (ui>1) then Halt; if ui>=0.5 then
ui:=1-ui; if ui<(2*1e-15) then ui:=2*1e-15;
ui:=sqrt(-2*ln(ui))-sqrt(ln(4)); if 0.01<ui then begin
szi:=0.2385099062881218351e-3;
szi:=szi*ui+0.7748014532123519149e-2;
szi:=szi*ui+0.9433047102597236601e-1;
szi:=szi*ui+0.5906175347671242813e0;
szi:=szi*ui+0.2052429201482605360e1;
szi:=szi*ui+0.3926527220876257871e1;
szi:=szi*ui+0.3827787912267809326e1;
szi:=szi*ui+0.1475664626635605795e1;szi:=szi*ui+0.0;
ni:=0.2384667219100680462e-3;
ni:=ni*ui+0.7472090148248391360e-2;
ni:=ni*ui+0.8675117268832776800e-1;
ni:=ni*ui+0.5155497927835685221e0;
ni:=ni*ui+0.1685386828574926342e1 ;
ni:=ni*ui+0.3019304632408607270e1 ;
ni:=ni*ui+0.2757985744718406918e1 ;
ni:=ni*ui+1;end
else begin
szi:=0.1389671822546715525e-4;
szi:=szi*ui+0.9933095513250211212e-3;
szi:=szi*ui+0.2132223881469308687e-1;
szi:=szi*ui+0.1971184884114817024e0;
szi:=szi*ui+0.9208235553699620741e0;
szi:=szi*ui+0.2302486886454418763e1;
szi:=szi*ui+0.2934913383940946604e1;
szi:=szi*ui+0.1475663066897793476e1;
szi:=szi*ui+0.2236640681757362082e-6;
ni:=0.1389640654034188922e-4;
ni:=ni*ui+0.9770522217813339426e-3;
ni:=ni*ui+0.2025571989491669521e-1;
ni:=ni*ui+0.1775558927085441912e0;
ni:=ni*ui+0.7785719242838022205e0;
ni:=ni*ui+0.1819506588454068626e1;
ni:=ni*ui+0.2152916059924272000e1;
ni:=ni*ui+1.0;end;
if x<0.5 then szi:=-szi; Invphi:=szi/ni; end;
```

## FÜGGELEK

---

**9.4. táblázat: Standard normális eloszlás értékek.**

| ' x | .00   | .01   | .02   | .03   | .04   | .05   | .06   | .07   | .08   | .09   |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.0 | .5000 | .5040 | .5080 | .5120 | .5160 | .5199 | .5239 | .5279 | .5319 | .5359 |
| 0.1 | .5398 | .5438 | .5478 | .5517 | .5557 | .5596 | .5636 | .5675 | .5714 | .5753 |
| 0.2 | .5793 | .5832 | .5871 | .5910 | .5948 | .5987 | .6026 | .6064 | .6103 | .6141 |
| 0.3 | .6179 | .6217 | .6255 | .6293 | .6331 | .6368 | .6406 | .6443 | .6480 | .6517 |
| 0.4 | .6554 | .6591 | .6628 | .6664 | .6700 | .6736 | .6772 | .6808 | .6844 | .6879 |
| 0.5 | .6915 | .6950 | .6985 | .7019 | .7054 | .7088 | .7123 | .7157 | .7190 | .7224 |
| 0.6 | .7257 | .7291 | .7324 | .7357 | .7389 | .7422 | .7454 | .7486 | .7517 | .7549 |
| 0.7 | .7580 | .7611 | .7642 | .7673 | .7704 | .7734 | .7764 | .7794 | .7823 | .7852 |
| 0.8 | .7881 | .7910 | .7939 | .7967 | .7995 | .8023 | .8051 | .8078 | .8106 | .8133 |
| 0.9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 1.0 | .8413 | .8438 | .8461 | .8485 | .8508 | .8531 | .8554 | .8577 | .8599 | .8621 |
| 1.1 | .8643 | .8665 | .8686 | .8708 | .8729 | .8749 | .8770 | .8790 | .8810 | .8830 |
| 1.2 | .8849 | .8869 | .8888 | .8907 | .8925 | .8944 | .8962 | .8980 | .8997 | .9015 |
| 1.3 | .9032 | .9049 | .9066 | .9082 | .9099 | .9115 | .9131 | .9147 | .9162 | .9177 |
| 1.4 | .9192 | .9207 | .9222 | .9236 | .9251 | .9265 | .9279 | .9292 | .9306 | .9319 |
| 1.5 | .9332 | .9345 | .9357 | .9370 | .9382 | .9394 | .9406 | .9418 | .9429 | .9441 |
| 1.6 | .9452 | .9463 | .9474 | .9484 | .9495 | .9505 | .9515 | .9525 | .9535 | .9545 |
| 1.7 | .9554 | .9564 | .9573 | .9582 | .9591 | .9599 | .9608 | .9616 | .9625 | .9633 |
| 1.8 | .9641 | .9649 | .9656 | .9664 | .9671 | .9678 | .9686 | .9693 | .9699 | .9706 |
| 1.9 | .9713 | .9719 | .9726 | .9732 | .9738 | .9744 | .9750 | .9756 | .9761 | .9767 |
| 2.0 | .9772 | .9778 | .9783 | .9788 | .9793 | .9798 | .9803 | .9808 | .9812 | .9817 |
| 2.1 | .9821 | .9826 | .9830 | .9834 | .9838 | .9842 | .9846 | .9850 | .9854 | .9857 |
| 2.2 | .9861 | .9864 | .9868 | .9871 | .9875 | .9878 | .9881 | .9884 | .9887 | .9890 |
| 2.3 | .9893 | .9896 | .9898 | .9901 | .9904 | .9906 | .9909 | .9911 | .9913 | .9916 |
| 2.4 | .9918 | .9920 | .9922 | .9924 | .9927 | .9929 | .9931 | .9932 | .9934 | .9936 |
| 2.5 | .9938 | .9940 | .9941 | .9943 | .9945 | .9946 | .9948 | .9949 | .9951 | .9952 |
| 2.6 | .9953 | .9955 | .9956 | .9957 | .9959 | .9960 | .9961 | .9962 | .9963 | .9964 |
| 2.7 | .9965 | .9966 | .9967 | .9968 | .9969 | .9970 | .9971 | .9972 | .9973 | .9974 |
| 2.8 | .9974 | .9975 | .9976 | .9977 | .9977 | .9978 | .9979 | .9979 | .9980 | .9981 |
| 2.9 | .9981 | .9982 | .9982 | .9983 | .9984 | .9984 | .9985 | .9985 | .9986 | .9986 |
| 3.0 | .9986 | .9987 | .9987 | .9988 | .9988 | .9989 | .9989 | .9989 | .9990 | .9990 |
| 3.1 | .9990 | .9991 | .9991 | .9991 | .9992 | .9992 | .9992 | .9992 | .9993 | .9993 |
| 3.2 | .9993 | .9993 | .9994 | .9994 | .9994 | .9994 | .9994 | .9995 | .9995 | .9995 |
| 3.4 | .9995 | .9995 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 |
| 3.5 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9998 | .9998 |

**9.5. táblázat:  $\chi^2$ -eloszlás értékek.**

| $n$ | 0.005  | 0.010  | 0.025  | 0.050  | 0.950  | 0.975  | 0.990  | 0.995  |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   | .39e-4 | .16e-3 | .98e-3 | .39e-2 | 3.841  | 5.024  | 6.635  | 7.879  |
| 2   | .010   | .020   | .051   | .103   | 5.991  | 7.377  | 9.210  | 10.597 |
| 3   | .0717  | .1148  | .2158  | .3518  | 7.814  | 9.348  | 11.345 | 12.838 |
| 4   | .2069  | .2971  | .4844  | .7107  | 9.487  | 11.143 | 13.277 | 14.860 |
| 5   | .4117  | .5543  | .8312  | 1.1455 | 11.070 | 12.833 | 15.086 | 16.750 |
| 6   | .67573 | .87209 | 1.2373 | 1.6354 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7   | .98926 | 1.2390 | 1.6899 | 2.1673 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8   | 1.3444 | 1.6465 | 2.1797 | 2.7326 | 15.507 | 17.535 | 20.090 | 21.955 |
| 9   | 1.7349 | 2.0879 | 2.7004 | 3.3251 | 16.919 | 19.023 | 21.666 | 23.589 |
| 10  | 2.1559 | 2.5582 | 3.2470 | 3.9403 | 18.307 | 20.483 | 23.209 | 25.188 |
| 11  | 2.6032 | 3.0535 | 3.8157 | 4.5748 | 19.675 | 21.920 | 24.725 | 26.757 |
| 12  | 3.0738 | 3.5706 | 4.4038 | 5.2260 | 21.026 | 23.337 | 26.217 | 28.300 |
| 13  | 3.5650 | 4.1069 | 5.0088 | 5.8919 | 22.362 | 24.736 | 27.688 | 29.819 |
| 14  | 4.0747 | 4.6604 | 5.6287 | 6.5706 | 23.685 | 26.119 | 29.141 | 31.319 |
| 15  | 4.6009 | 5.2293 | 6.2621 | 7.2609 | 24.996 | 27.488 | 30.578 | 32.801 |
| 16  | 5.1422 | 5.8122 | 6.9077 | 7.9616 | 26.296 | 28.845 | 32.000 | 34.267 |
| 17  | 5.6972 | 6.4078 | 7.5642 | 8.6718 | 27.587 | 30.191 | 33.409 | 35.718 |
| 18  | 6.2648 | 7.0149 | 8.2307 | 9.3905 | 28.869 | 31.526 | 34.805 | 37.156 |
| 19  | 6.8440 | 7.6327 | 8.9065 | 10.117 | 30.144 | 32.852 | 36.191 | 38.582 |
| 20  | 7.4338 | 8.2604 | 9.5908 | 10.851 | 31.410 | 34.170 | 37.566 | 39.997 |
| 21  | 8.0337 | 8.8972 | 10.283 | 11.591 | 32.671 | 35.479 | 38.932 | 41.401 |
| 22  | 8.6427 | 9.5425 | 10.982 | 12.338 | 33.924 | 36.781 | 40.289 | 42.796 |
| 23  | 9.2604 | 10.196 | 11.689 | 13.091 | 35.172 | 38.076 | 41.638 | 44.181 |
| 24  | 9.8862 | 10.856 | 12.401 | 13.848 | 36.415 | 39.364 | 42.980 | 45.559 |

# FÜGGELEK

---

| $n$ | 0.005  | 0.010  | 0.025  | 0.050  | 0.950  | 0.975  | 0.990  | 0.995  |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|
| 25  | 10.520 | 11.524 | 13.120 | 14.611 | 37.652 | 40.646 | 44.314 | 46.928 |
| 26  | 11.160 | 12.198 | 13.844 | 15.379 | 38.885 | 41.923 | 45.642 | 48.290 |
| 27  | 11.808 | 12.879 | 14.573 | 16.151 | 40.113 | 43.195 | 46.963 | 49.645 |
| 28  | 12.461 | 13.565 | 15.308 | 16.928 | 41.337 | 44.461 | 48.278 | 50.993 |
| 29  | 13.121 | 14.256 | 16.047 | 17.708 | 42.557 | 45.722 | 49.588 | 52.336 |
| 30  | 13.787 | 14.953 | 16.791 | 18.493 | 43.773 | 46.979 | 50.892 | 53.672 |
| 40  | 20.707 | 22.164 | 24.433 | 26.509 | 55.758 | 59.342 | 63.691 | 66.766 |
| 50  | 27.991 | 29.707 | 32.357 | 34.764 | 67.505 | 71.420 | 76.154 | 79.490 |
| 60  | 35.534 | 37.485 | 40.482 | 43.188 | 79.082 | 83.298 | 88.379 | 91.952 |
| 70  | 43.275 | 45.442 | 48.758 | 51.739 | 90.531 | 95.023 | 100.43 | 104.21 |
| 80  | 51.172 | 53.540 | 57.153 | 60.391 | 101.88 | 106.63 | 112.33 | 116.32 |
| 90  | 59.196 | 61.754 | 65.647 | 69.126 | 113.15 | 118.14 | 124.12 | 128.30 |
| 100 | 67.328 | 70.065 | 74.222 | 77.929 | 124.34 | 129.56 | 135.81 | 140.17 |

**9.6. táblázat:  $t$ -eloszlás értékek.**

| $n$ | 0.9    | 0.95   | 0.975  | 0.99   | 0.995  |
|-----|--------|--------|--------|--------|--------|
| 1   | 3.0777 | 6.3138 | 12.706 | 31.821 | 63.657 |
| 2   | 1.8856 | 2.9200 | 4.3027 | 6.9646 | 9.9248 |
| 3   | 1.6377 | 2.3534 | 3.1824 | 4.5407 | 5.8409 |
| 4   | 1.5332 | 2.1318 | 2.7764 | 3.7469 | 4.6041 |
| 5   | 1.4759 | 2.0150 | 2.5706 | 3.3649 | 4.0321 |
| 6   | 1.4398 | 1.9432 | 2.4469 | 3.1427 | 3.7074 |
| 7   | 1.4149 | 1.8946 | 2.3646 | 2.9980 | 3.4995 |
| 8   | 1.3968 | 1.8595 | 2.3060 | 2.8965 | 3.3554 |
| 9   | 1.3830 | 1.8331 | 2.2622 | 2.8214 | 3.2498 |
| 10  | 1.3722 | 1.8125 | 2.2281 | 2.7638 | 3.1693 |
| 11  | 1.3634 | 1.7959 | 2.2010 | 2.7181 | 3.1058 |
| 12  | 1.3562 | 1.7823 | 2.1788 | 2.6810 | 3.0545 |
| 13  | 1.3502 | 1.7709 | 2.1604 | 2.6503 | 3.0123 |
| 14  | 1.3450 | 1.7613 | 2.1448 | 2.6245 | 2.9768 |
| 15  | 1.3406 | 1.7531 | 2.1314 | 2.6025 | 2.9467 |
| 16  | 1.3368 | 1.7459 | 2.1199 | 2.5835 | 2.9208 |
| 17  | 1.3334 | 1.7396 | 2.1098 | 2.5669 | 2.8982 |
| 18  | 1.3304 | 1.7341 | 2.1009 | 2.5524 | 2.8784 |
| 19  | 1.3277 | 1.7291 | 2.0930 | 2.5395 | 2.8609 |
| 20  | 1.3253 | 1.7247 | 2.0860 | 2.5280 | 2.8453 |
| 21  | 1.3232 | 1.7207 | 2.0796 | 2.5176 | 2.8314 |
| 22  | 1.3212 | 1.7171 | 2.0739 | 2.5083 | 2.8188 |
| 23  | 1.3195 | 1.7139 | 2.0687 | 2.4999 | 2.8073 |
| 24  | 1.3178 | 1.7109 | 2.0639 | 2.4922 | 2.7969 |
| 25  | 1.3163 | 1.7081 | 2.0595 | 2.4851 | 2.7874 |
| 26  | 1.3150 | 1.7056 | 2.0555 | 2.4786 | 2.7787 |
| 27  | 1.3137 | 1.7033 | 2.0518 | 2.4727 | 2.7707 |

# FÜGGELEK

---

| $n$      | 0.9    | 0.95   | 0.975  | 0.99   | 0.995  |
|----------|--------|--------|--------|--------|--------|
| 28       | 1.3125 | 1.7011 | 2.0484 | 2.4671 | 2.7633 |
| 29       | 1.3114 | 1.6991 | 2.0452 | 2.4620 | 2.7564 |
| 30       | 1.3104 | 1.6973 | 2.0423 | 2.4573 | 2.7500 |
| 31       | 1.3095 | 1.6955 | 2.0395 | 2.4528 | 2.7440 |
| 32       | 1.3086 | 1.6939 | 2.0369 | 2.4487 | 2.7385 |
| 33       | 1.3077 | 1.6924 | 2.0345 | 2.4448 | 2.7333 |
| 34       | 1.3070 | 1.6909 | 2.0322 | 2.4411 | 2.7284 |
| 35       | 1.3062 | 1.6896 | 2.0301 | 2.4377 | 2.7238 |
| 36       | 1.3055 | 1.6883 | 2.0281 | 2.4345 | 2.7195 |
| 37       | 1.3049 | 1.6871 | 2.0262 | 2.4314 | 2.7154 |
| 38       | 1.3042 | 1.6860 | 2.0244 | 2.4286 | 2.7116 |
| 39       | 1.3036 | 1.6849 | 2.0227 | 2.4258 | 2.7079 |
| 41       | 1.3025 | 1.6829 | 2.0195 | 2.4208 | 2.7012 |
| 51       | 1.2984 | 1.6753 | 2.0076 | 2.4017 | 2.6757 |
| 61       | 1.2956 | 1.6702 | 1.9996 | 2.3890 | 2.6589 |
| 71       | 1.2936 | 1.6666 | 1.9939 | 2.3800 | 2.6469 |
| 81       | 1.2921 | 1.6639 | 1.9897 | 2.3733 | 2.6379 |
| 91       | 1.2909 | 1.6618 | 1.9864 | 2.3680 | 2.6309 |
| 101      | 1.2900 | 1.6601 | 1.9837 | 2.3638 | 2.6254 |
| 201      | 1.2858 | 1.6525 | 1.9718 | 2.3450 | 2.6005 |
| 301      | 1.2844 | 1.6499 | 1.9679 | 2.3388 | 2.5923 |
| 401      | 1.2837 | 1.6487 | 1.9659 | 2.3357 | 2.5881 |
| 501      | 1.2832 | 1.6479 | 1.9647 | 2.3338 | 2.5857 |
| 601      | 1.2830 | 1.6474 | 1.9639 | 2.3326 | 2.5840 |
| 701      | 1.2828 | 1.6470 | 1.9634 | 2.3317 | 2.5829 |
| 801      | 1.2826 | 1.6468 | 1.9629 | 2.3310 | 2.5820 |
| 901      | 1.2825 | 1.6465 | 1.9626 | 2.3305 | 2.5813 |
| 1001     | 1.2824 | 1.6464 | 1.9623 | 2.3301 | 2.5807 |
| $\infty$ | 1.2816 | 1.6449 | 1.9600 | 2.3263 | 2.5758 |

## 9. FEJEZET

### 9.7. táblázat: $F$ -eloszlás értékek.

$n$  – a számáló és  $m$  – a nevező szabadsági foka

| $m \backslash n$ | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1                | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 236.8 | 238.9 | 240.5 | 241.9 |
| 2                | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 |
| 3                | 10.13 | 9.552 | 9.277 | 9.117 | 9.013 | 8.941 | 8.887 | 8.845 | 8.812 | 8.786 |
| 4                | 7.709 | 6.944 | 6.591 | 6.388 | 6.256 | 6.163 | 6.094 | 6.041 | 5.999 | 5.964 |
| 5                | 6.608 | 5.786 | 5.409 | 5.192 | 5.050 | 4.950 | 4.876 | 4.818 | 4.772 | 4.735 |
| 6                | 5.987 | 5.143 | 4.757 | 4.534 | 4.387 | 4.284 | 4.207 | 4.147 | 4.099 | 4.060 |
| 7                | 5.591 | 4.737 | 4.347 | 4.120 | 3.972 | 3.866 | 3.787 | 3.726 | 3.677 | 3.637 |
| 8                | 5.318 | 4.459 | 4.066 | 3.838 | 3.687 | 3.581 | 3.500 | 3.438 | 3.388 | 3.347 |
| 9                | 5.117 | 4.256 | 3.863 | 3.633 | 3.482 | 3.374 | 3.293 | 3.230 | 3.179 | 3.137 |
| 10               | 4.965 | 4.103 | 3.708 | 3.478 | 3.326 | 3.217 | 3.135 | 3.072 | 3.020 | 2.978 |
| 11               | 4.844 | 3.982 | 3.587 | 3.357 | 3.204 | 3.095 | 3.012 | 2.948 | 2.896 | 2.854 |
| 12               | 4.747 | 3.885 | 3.490 | 3.259 | 3.106 | 2.996 | 2.913 | 2.849 | 2.796 | 2.753 |
| 13               | 4.667 | 3.806 | 3.411 | 3.179 | 3.025 | 2.915 | 2.832 | 2.767 | 2.714 | 2.671 |
| 14               | 4.600 | 3.739 | 3.344 | 3.112 | 2.958 | 2.848 | 2.764 | 2.699 | 2.646 | 2.602 |
| 15               | 4.543 | 3.682 | 3.287 | 3.056 | 2.901 | 2.790 | 2.707 | 2.641 | 2.588 | 2.544 |
| 16               | 4.494 | 3.634 | 3.239 | 3.007 | 2.852 | 2.741 | 2.657 | 2.591 | 2.538 | 2.494 |
| 17               | 4.451 | 3.592 | 3.197 | 2.965 | 2.810 | 2.699 | 2.614 | 2.548 | 2.494 | 2.450 |
| 18               | 4.414 | 3.555 | 3.160 | 2.928 | 2.773 | 2.661 | 2.577 | 2.510 | 2.456 | 2.412 |
| 19               | 4.381 | 3.522 | 3.127 | 2.895 | 2.740 | 2.628 | 2.544 | 2.477 | 2.423 | 2.378 |
| 20               | 4.351 | 3.493 | 3.098 | 2.866 | 2.711 | 2.599 | 2.514 | 2.447 | 2.393 | 2.348 |
| 25               | 4.242 | 3.385 | 2.991 | 2.759 | 2.603 | 2.490 | 2.405 | 2.337 | 2.282 | 2.236 |
| 30               | 4.171 | 3.316 | 2.922 | 2.690 | 2.534 | 2.421 | 2.334 | 2.266 | 2.211 | 2.165 |
| 35               | 4.121 | 3.267 | 2.874 | 2.641 | 2.485 | 2.372 | 2.285 | 2.217 | 2.161 | 2.114 |
| 40               | 4.085 | 3.232 | 2.839 | 2.606 | 2.449 | 2.336 | 2.249 | 2.180 | 2.124 | 2.077 |

# FÜGGELEK

| $m \backslash n$ | 11    | 12    | 13    | 14    | 15    | 16    | 17    | 18    | 19    | 20    |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1                | 243.0 | 243.9 | 244.7 | 245.4 | 245.9 | 246.5 | 246.9 | 247.3 | 247.7 | 248.0 |
| 2                | 19.40 | 19.41 | 19.42 | 19.42 | 19.43 | 19.43 | 19.44 | 19.44 | 19.44 | 19.45 |
| 3                | 8.763 | 8.745 | 8.729 | 8.715 | 8.703 | 8.692 | 8.683 | 8.675 | 8.667 | 8.660 |
| 4                | 5.936 | 5.912 | 5.891 | 5.873 | 5.858 | 5.844 | 5.832 | 5.821 | 5.811 | 5.803 |
| 5                | 4.704 | 4.678 | 4.655 | 4.636 | 4.619 | 4.604 | 4.590 | 4.579 | 4.568 | 4.558 |
| 6                | 4.027 | 4.000 | 3.976 | 3.956 | 3.938 | 3.922 | 3.908 | 3.896 | 3.884 | 3.874 |
| 7                | 3.603 | 3.575 | 3.550 | 3.529 | 3.511 | 3.494 | 3.480 | 3.467 | 3.455 | 3.445 |
| 8                | 3.313 | 3.284 | 3.259 | 3.237 | 3.218 | 3.202 | 3.187 | 3.173 | 3.161 | 3.150 |
| 9                | 3.102 | 3.073 | 3.048 | 3.025 | 3.006 | 2.989 | 2.974 | 2.960 | 2.948 | 2.936 |
| 10               | 2.943 | 2.913 | 2.887 | 2.865 | 2.845 | 2.828 | 2.812 | 2.798 | 2.785 | 2.774 |
| 11               | 2.818 | 2.788 | 2.761 | 2.739 | 2.719 | 2.701 | 2.685 | 2.671 | 2.658 | 2.646 |
| 12               | 2.717 | 2.687 | 2.660 | 2.637 | 2.617 | 2.599 | 2.583 | 2.568 | 2.555 | 2.544 |
| 13               | 2.635 | 2.604 | 2.577 | 2.554 | 2.533 | 2.515 | 2.499 | 2.484 | 2.471 | 2.459 |
| 14               | 2.565 | 2.534 | 2.507 | 2.484 | 2.463 | 2.445 | 2.428 | 2.413 | 2.400 | 2.388 |
| 15               | 2.507 | 2.475 | 2.448 | 2.424 | 2.403 | 2.385 | 2.368 | 2.353 | 2.340 | 2.328 |
| 16               | 2.456 | 2.425 | 2.397 | 2.373 | 2.352 | 2.333 | 2.317 | 2.302 | 2.288 | 2.276 |
| 17               | 2.413 | 2.381 | 2.353 | 2.329 | 2.308 | 2.289 | 2.272 | 2.257 | 2.243 | 2.230 |
| 18               | 2.374 | 2.342 | 2.314 | 2.290 | 2.269 | 2.250 | 2.233 | 2.217 | 2.203 | 2.191 |
| 19               | 2.340 | 2.308 | 2.280 | 2.256 | 2.234 | 2.215 | 2.198 | 2.182 | 2.168 | 2.155 |
| 20               | 2.310 | 2.278 | 2.250 | 2.225 | 2.203 | 2.184 | 2.167 | 2.151 | 2.137 | 2.124 |
| 25               | 2.198 | 2.165 | 2.136 | 2.111 | 2.089 | 2.069 | 2.051 | 2.035 | 2.021 | 2.007 |
| 30               | 2.126 | 2.092 | 2.063 | 2.037 | 2.015 | 1.995 | 1.976 | 1.960 | 1.945 | 1.932 |
| 35               | 2.075 | 2.041 | 2.012 | 1.986 | 1.963 | 1.942 | 1.924 | 1.907 | 1.892 | 1.878 |
| 40               | 2.038 | 2.003 | 1.974 | 1.948 | 1.924 | 1.904 | 1.885 | 1.868 | 1.853 | 1.839 |

## 9. FEJEZET

---

| $m \backslash n$ | 30    | 40    | 50    | 60    | 70    | 80    | 90    | 100   | 110   | 120   |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1                | 250.1 | 251.1 | 251.8 | 252.2 | 252.5 | 252.7 | 252.9 | 253.0 | 253.2 | 253.3 |
| 2                | 19.46 | 19.47 | 19.48 | 19.48 | 19.48 | 19.48 | 19.48 | 19.49 | 19.49 | 19.49 |
| 3                | 8.617 | 8.594 | 8.581 | 8.572 | 8.566 | 8.561 | 8.557 | 8.554 | 8.551 | 8.549 |
| 4                | 5.746 | 5.717 | 5.699 | 5.688 | 5.679 | 5.673 | 5.668 | 5.664 | 5.661 | 5.658 |
| 5                | 4.496 | 4.464 | 4.444 | 4.431 | 4.422 | 4.415 | 4.409 | 4.405 | 4.401 | 4.398 |
| 6                | 3.808 | 3.774 | 3.754 | 3.740 | 3.730 | 3.722 | 3.716 | 3.712 | 3.708 | 3.705 |
| 7                | 3.376 | 3.340 | 3.319 | 3.304 | 3.294 | 3.286 | 3.280 | 3.275 | 3.271 | 3.267 |
| 8                | 3.079 | 3.043 | 3.020 | 3.005 | 2.994 | 2.986 | 2.980 | 2.975 | 2.970 | 2.967 |
| 9                | 2.864 | 2.826 | 2.803 | 2.787 | 2.776 | 2.768 | 2.761 | 2.756 | 2.751 | 2.748 |
| 10               | 2.700 | 2.661 | 2.637 | 2.621 | 2.610 | 2.601 | 2.594 | 2.588 | 2.584 | 2.580 |
| 11               | 2.570 | 2.531 | 2.507 | 2.490 | 2.478 | 2.469 | 2.462 | 2.457 | 2.452 | 2.448 |
| 12               | 2.466 | 2.426 | 2.401 | 2.384 | 2.372 | 2.363 | 2.356 | 2.350 | 2.345 | 2.341 |
| 13               | 2.380 | 2.339 | 2.314 | 2.297 | 2.284 | 2.275 | 2.267 | 2.261 | 2.257 | 2.252 |
| 14               | 2.308 | 2.266 | 2.241 | 2.223 | 2.210 | 2.201 | 2.193 | 2.187 | 2.182 | 2.178 |
| 15               | 2.247 | 2.204 | 2.178 | 2.160 | 2.147 | 2.137 | 2.130 | 2.123 | 2.118 | 2.114 |
| 16               | 2.194 | 2.151 | 2.124 | 2.106 | 2.093 | 2.083 | 2.075 | 2.068 | 2.063 | 2.059 |
| 17               | 2.148 | 2.104 | 2.077 | 2.058 | 2.045 | 2.035 | 2.027 | 2.020 | 2.015 | 2.011 |
| 18               | 2.107 | 2.063 | 2.035 | 2.017 | 2.003 | 1.993 | 1.985 | 1.978 | 1.973 | 1.968 |
| 19               | 2.071 | 2.026 | 1.999 | 1.980 | 1.966 | 1.955 | 1.947 | 1.940 | 1.935 | 1.930 |
| 20               | 2.039 | 1.994 | 1.966 | 1.946 | 1.932 | 1.922 | 1.913 | 1.907 | 1.901 | 1.896 |
| 25               | 1.919 | 1.872 | 1.842 | 1.822 | 1.807 | 1.796 | 1.787 | 1.779 | 1.773 | 1.768 |
| 30               | 1.841 | 1.792 | 1.761 | 1.740 | 1.724 | 1.712 | 1.703 | 1.695 | 1.689 | 1.683 |
| 35               | 1.786 | 1.735 | 1.703 | 1.681 | 1.665 | 1.652 | 1.643 | 1.635 | 1.628 | 1.623 |
| 40               | 1.744 | 1.693 | 1.660 | 1.637 | 1.621 | 1.608 | 1.597 | 1.589 | 1.582 | 1.577 |

**9.8. táblázat: Grubbs-féle kritikus értékek.**

|     | 95%   |       |        | 99%   |       |        |
|-----|-------|-------|--------|-------|-------|--------|
| $n$ | G(1)  | G(2)  | G(3)   | G(1)  | G(2)  | G(3)   |
| 3   | 1.153 | 2.00  | -      | 1.155 | 2.00  | -      |
| 4   | 1.463 | 2.43  | 0.9992 | 1.492 | 2.44  | 1.0000 |
| 5   | 1.672 | 2.75  | 0.9817 | 1.749 | 2.80  | 0.9965 |
| 6   | 1.822 | 3.01  | 0.9436 | 1.944 | 3.10  | 0.9814 |
| 7   | 1.938 | 3.22  | 0.8980 | 2.097 | 3.34  | 0.9560 |
| 8   | 2.032 | 3.40  | 0.8522 | 2.221 | 3.54  | 0.9250 |
| 9   | 2.110 | 3.55  | 0.8091 | 2.323 | 3.72  | 0.8918 |
| 10  | 2.176 | 3.68  | 0.7695 | 2.410 | 3.88  | 0.8586 |
| 12  | 2.285 | 3.91  | 0.7004 | 2.550 | 4.13  | 0.7957 |
| 13  | 2.331 | 4.00  | 0.6705 | 2.607 | 4.24  | 0.7667 |
| 15  | 2.409 | 4.17  | 0.6182 | 2.705 | 4.43  | 0.7141 |
| 20  | 2.557 | 4.49  | 0.5196 | 2.884 | 4.79  | 0.6091 |
| 25  | 2.663 | 4.73  | 0.4505 | 3.009 | 5.03  | 0.5320 |
| 30  | 2.745 | 4.89  | 0.3992 | 3.103 | 5.19  | 0.4732 |
| 35  | 2.811 | 5.026 | 0.3595 | 3.178 | 5.326 | 0.4270 |
| 40  | 2.866 | 5.150 | 0.3276 | 3.240 | 5.450 | 0.3896 |
| 50  | 2.956 | 5.350 | 0.2797 | 3.336 | 5.650 | 0.3328 |
| 60  | 3.025 | 5.500 | 0.2450 | 3.411 | 5.800 | 0.2914 |
| 70  | 3.082 | 5.638 | 0.2187 | 3.471 | 5.938 | 0.2599 |
| 80  | 3.130 | 5.730 | 0.1979 | 3.521 | 6.030 | 0.2350 |
| 90  | 3.171 | 5.820 | 0.1810 | 3.563 | 6.120 | 0.2147 |
| 100 | 3.207 | 5.900 | 0.1671 | 3.600 | 6.200 | 0.1980 |
| 110 | 3.239 | 5.968 | 0.1553 | 3.632 | 6.268 | 0.1838 |
| 120 | 3.267 | 6.030 | 0.1452 | 3.662 | 6.330 | 0.1716 |
| 140 | 3.318 | 6.137 | 0.1288 | 3.712 | 6.437 | 0.1519 |

---

## IRODALOMJEGYZÉK

- [1] Bognár J-né, Mogyoródi J., Prékopa A., Rényi A., Szász D. (1971): *Valószínűségszámítás feladatgyűjtemény*, Tankönyvkiadó, Budapest.
- [2] Csákány A., Vajda F. (1985): *Játékok számítógéppel*, Műszaki Könyvkiadó, Budapest.
- [3] Deák I. (1986): *Véletlenszámgenerátorok és alkalmazásaik*, Akadémiai Kiadó, Budapest.
- [4] W. Feller (1978): *Bevezetés a valószínűségszámításba és alkalmazásaiba*, Műszaki Könyvkiadó, Budapest.
- [5] Fritz J. (1971): *Bevezetés az információelméletbe*, Tankönyvkiadó, Budapest.
- [6] G.P. Gavrillov, A.A. Szapozsenko (1981): *Diszkrét matematikai feladatgyűjtemény*, Műszaki Könyvkiadó, Budapest.
- [7] B.V. Gnedenko (1975): *The Theory of Probability*, Mir Publishers, Moscow.
- [8] R.V. Hogg, A.T. Craig (1969): *Introduction to Mathematical Statistics*, The MacMillan Company, New York.
- [9] Jordán K. (1956): *Fejezetek a klasszikus valószínűségszámításból*, Akadémiai Kiadó, Budapest
- [10] A.N. Kolmogorov (1982): *A valószínűségszámítás alapfogalmai*, Gondolat, Budapest.
- [11] Lukács O. (1987): *Matematikai statisztika*, Műszaki Könyvkiadó, Budapest.
- [12] D.E. Mansfield, M. Bruckheimer (1973): *Matematika új felfogásban III.*, Gondolat, Budapest.

- [13] J. Nievergelt, J.C. Farrar, E.M. Reingold (1977): *Matematikai problémák megoldásának számítógépes módszerei*, Műszaki Könyvkiadó, Budapest.
- [14] M.M. Rao (1984): *Probability Theory with Applications*, Academic Press, Orlando, Florida.
- [15] Rényi A. (1954): *Valószínűségszámítás*, Tankönyvkiadó, Budapest.
- [16] Solt Gy. (1973): *Valószínűségszámítás*, Műszaki Könyvkiadó, Budapest.
- [17] I.M. Szobol (1981): *A Monte-Carlo módszerek alapjai*, Műszaki Könyvkiadó, Budapest
- [18] Vincze I. (1980): *Matematikai statisztika*, Tankönyvkiadó, Budapest.
- [19] Ja.B. Zeldovics, A.D. Miskisz (1978): *Az alkalmazott matematika elemei*, Gondolat, Budapest.

## VÉLETLEN ÉRTÉKEK I-III. BEMUTATÁSA

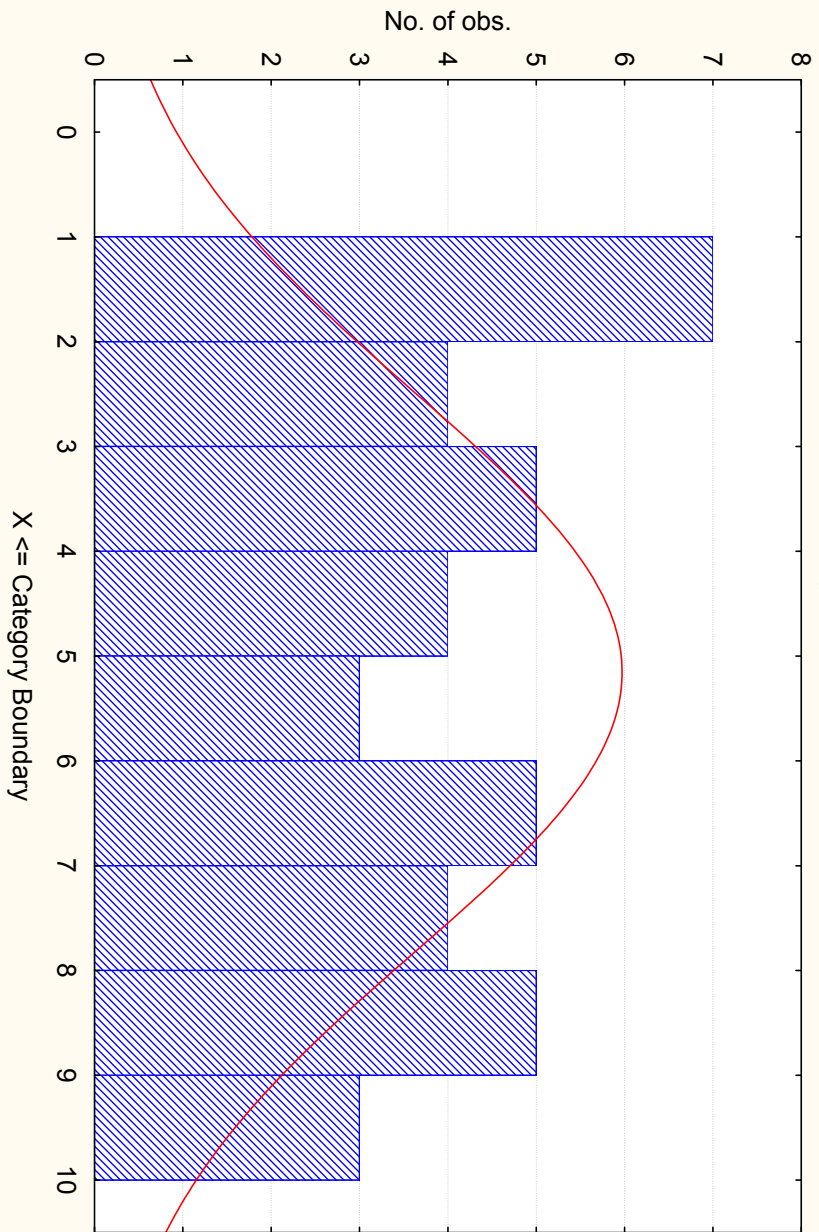
3.88, 7.70, 1.65, 6.42, 4.29, 3.01, 1.41, 6.54, 3.58, 9.05, 1.12, 1.74, 6.17, 3.38, 2.50, 4.01, 2.96, 8.48, 3.76, 5.75, 6.45, 7.89, 1.77, 2.32, 8.83, 4.10, 9.18, 4.99, 1.30, 7.17, 2.00, 9.24, 2.25, 7.80, 8.60, 5.84, 5.05, 8.56, 8.49, 6.95.

-0.340, 28.53, -12.38, -24.89, -6.636, -19.16, -4.626, 1.275, 837.7, -11.83, 21.33, -54.07, 33.24, 1.238, -123.4, 13.61, -65.22, 16.04, 129.9, 54.14, -16.92, -3.643, 88.06, -5.665, -89.20, 8.75, -5.429, -6.107, 411.4, 8.04, -5.791, 18.23, 1.357, -21.18, 51.81, 105.8, -14.51, 10.78, 7.38, -21.65, -12.96, -3.379, 17.48, 4.310, 15.23, -7.661, 1.853, -16.84, -19.26, -16.47.

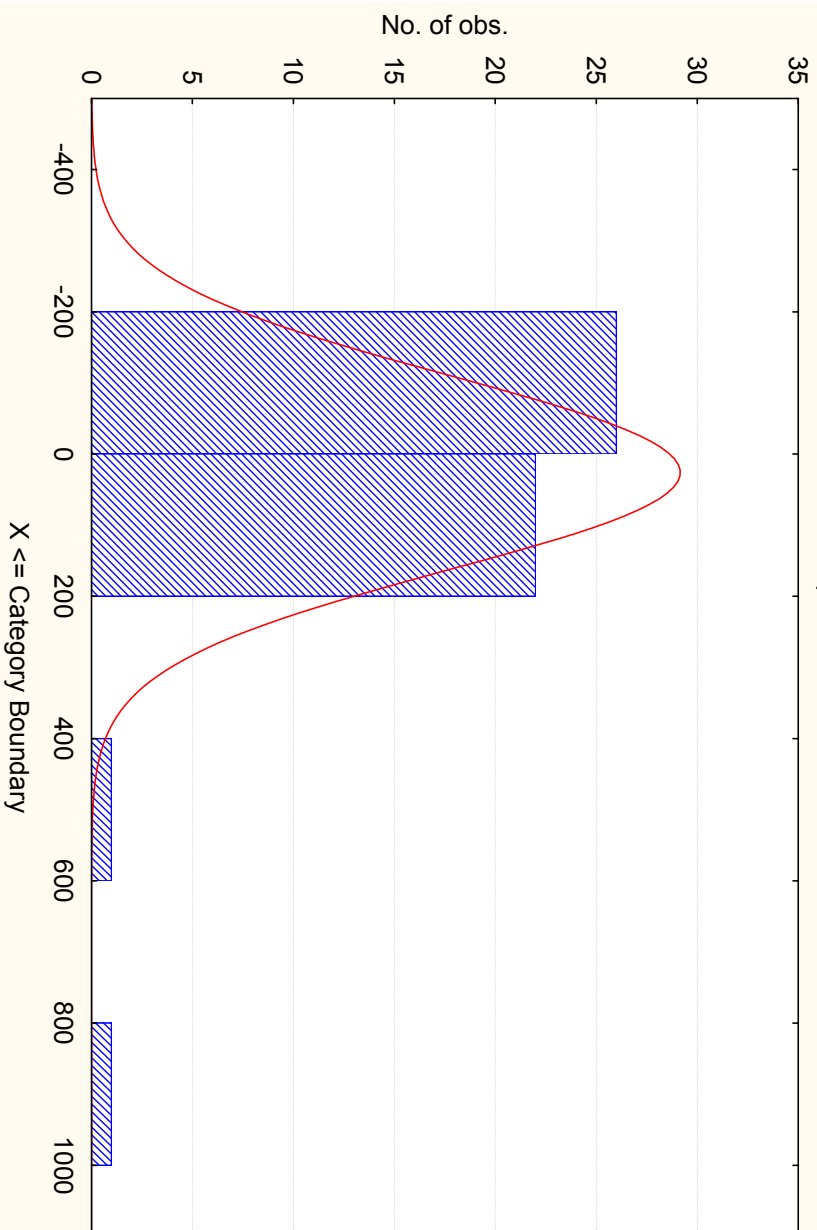
0.082, 0.668, 1.872, -1.494, 1.336, -2.547, -0.575, -3.587, -2.121, -0.030, -1.927, 2.916, 4.542, -2.921, 2.640, -1.718, 2.772, 1.398, -2.021, 2.310, 0.373, 0.153, 0.118, -2.551, 1.691, 2.121, 0.499, 0.711, -2.982, 2.799, -0.569, -3.025, 1.649, -4.036, 0.314, 3.044, 3.482, 0.732, -2.225, -0.469, -1.882, -0.481, -2.581, 0.747, -2.886, -3.271, 1.317, 0.780, 1.803, -2.012, -4.149, -3.657, -0.214, -2.372, 0.997, -1.300, -2.129, -2.567, -0.223, 1.452.

| $n$ | $\bar{\xi}$ | $med\{\xi_i\}$ | min    | max   | $s_n^*$ | $x_{0.25}$ | $x_{0.75}$ |
|-----|-------------|----------------|--------|-------|---------|------------|------------|
| 40  | 5.1545      | 5.02           | 1.120  | 9.24  | 2.6717  | 2.73       | 7.75       |
| 50  | 25.965      | -1.8595        | -123.4 | 837.7 | 136.81  | -14.51     | 16.04      |
| 60  | -0.320      | -0.122         | -4.419 | 4.542 | 2.161   | -2.1771    | 1.367      |

Histogram: Var1  
K-S d=,10465, p> .20; Lilliefors p> .20  
— Expected Normal



Histogram: Var2  
K-S d=,33880, p<,01 ; Lilliefors p<,01  
— Expected Normal



Histogram: Var3  
K-S d=,11512, p> .20; Lilliefors p<,05  
— Expected Normal

