# Valószínűségszámítás és matematikai statisztika (valószínűségszámítás rész)

Barczy Mátyás, Pap Gyula

Debreceni Egyetem, Szegedi Tudományegyetem

2014

# Ajánlott irodalom:


Denkinger Géza
*Valószínűségszámítás*
Nemzeti Tankönyvkiadó, 1997.

Denkinger Géza
*Valószínűségszámítási Gyakorlatok*
Nemzeti Tankönyvkiadó, 1999.

Fazekas István
*Valószínűségszámítás*
Debrecen, Kossuth Egyetemi Kiadó, 2000.

Pap Gyula
*Valószínűségszámítás előadáskövető anyagok*
`http://www.math.u-szeged.hu/~papgy/`

# Ajánlott irodalom:


- SOLT GYÖRGY
  *Valószínűségszámítás*
  Műszaki Könyvkiadó, 1969.

- BOGNÁR JÁNOSNÉ, MOGYORÓDI JÓZSEF, PRÉKOPA ANDRÁS, RÉNYI ALFRÉD, SZÁSZ DOMOKOS
  *Valószínűségszámítás feladatgyűjtemény*
  Nemzeti Tankönyvkiadó, 2001.

- N. SHIRYAEV
  *Probability, 2nd edition*
  Springer-Verlag, 1995.

# 1. Véletlen kísérletek, eseményalgebrák

## Véletlen események

Valószínűsége **véletlen eseményeknek** van,

- melyekről nem tudjuk előre megmondani, hogy bekövetkeznek-e, vagy sem;
- és amelyek
  - vagy **véletlen jelenségek** megfigyelésével kapcsolatosak (amikor a körülményeket nem tudjuk befolyásolni),
  - vagy pedig **véletlen kimenetelű kísérletekkel** kapcsolatosak (amikor befolyásolni tudjuk a körülményeket).

**Példák:**

- Minőségellenőrzés: $n$ termékből kiválasztunk $m$ darabot ($m \leqslant n$), és megszámoljuk, hogy hány selejtes van; **lehetséges kimenetelek:** az $\Omega := \{0, 1, 2, \dots, m\}$ halmaz elemei;
- Hagyományos lottó: megjelölünk 5 számot 90-ből, és megszámoljuk, hogy hány találatunk van; **lehetséges kimenetelek:** az $\Omega := \{0, 1, 2, 3, 4, 5\}$ halmaz elemei;
- Ragályos fertőzés terjedése, csapadékmennyiség alakulása, szeizmográf mozgása, sorhosszúság alakulása pénztáraknál, szerencsejátékok, tőzsdei áringadozások.

# 1. Véletlen kísérletek, eseményalgebrák

- **Elemi események:** a kísérlet/megfigyelés lehetséges kimenetelei.
- **Eseménytér:** az elemi események halmaza; jelölés: $\Omega$.
- **Esemény:** az eseménytér bizonyos $A \subset \Omega$ részhalmaza, amit úgy értünk, hogy ha az $\omega \in \Omega$ elemi esemény következik be, akkor
  - $\omega \in A$ esetén bekövetkezik az $A$ esemény is,
  - $\omega \notin A$ esetén az $A$ esemény nem következik be.
- **Biztos esemény:** amely mindig bekövetkezik; be lehet azonosítani az $\Omega \subset \Omega$ részhalmazzal.
- **Lehetetlen esemény:** amely sohasem következik be; be lehet azonosítani az $\emptyset \subset \Omega$ üres részhalmazzal.

## Logikai műveletek eseményekkel

- Minden $A$ eseménnyel kapcsolatban tekinthetjük az $A$ **ellentett (komplementer) eseményét:** ez pontosan akkor következik be, amikor az $A$ esemény nem következik be; jelölése: $\overline{A}$.
- Az $A$ és $B$ események **összege (uniója)** az az esemény, amely pontosan akkor következik be, amikor az $A$ és $B$ események közül legalább az egyik bekövetkezik; jelölése: $A+B$ vagy $A \cup B$.
- Az $A$ és $B$ események **szorzata (metszete)** az az esemény, amely pontosan akkor következik be, amikor az $A$ és $B$ események mindegyike bekövetkezik; jelölése: $A \cdot B$ vagy $A \cap B$.
- Az $A$ és $B$ események **különbsége** az az esemény, mely pontosan akkor következik be, amikor az $A$ esemény bekövetkezik, a $B$ esemény pedig nem; jelölése: $A - B$ vagy $A \setminus B$.

## A logikai műveletek tulajdonságai

- kommutativitás: $$A + B = B + A,$$ $$A \cdot B = B \cdot A.$$
- asszociativitás: $$A + (B + C) = (A + B) + C,$$ $$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$
- idempotencia: $$A + A = A,$$ $$A \cdot A = A.$$
- disztributivitás: $$A \cdot (B + C) = (A \cdot B) + (A \cdot C),$$ $$A + (B \cdot C) = (A + B) \cdot (A + C).$$
- de Morgan-féle azonosságok: $$\overline{A + B} = \overline{A} \cdot \overline{B},$$ $$\overline{A \cdot B} = \overline{A} + \overline{B}.$$
- különbség: $$A - B = A \cdot \overline{B}.$$

# 1. Véletlen kísérletek, eseményalgebrák

## Diszjunkt események

Azt mondjuk, hogy az $A$ és $B$ események **diszjunktak (kizárják egymást)**, ha egyszerre nem következhetnek be.

Az $A$ és $B$ események akkor és csak akkor diszjunktak, ha $A \cdot B = \emptyset$.

## $\Rightarrow$

Azt mondjuk, hogy az $A$ esemény **maga után vonja** a $B$ eseményt, ha az $A$ esemény bekövetkezése esetén mindig bekövetkezik a $B$ esemény is; jelölése: $A \Rightarrow B$.

A következő állítások ekvivalensek:

- $A \Rightarrow B$;
- $A \subset B$;
- $\overline{B} \subset \overline{A}$;
- $\overline{B} \Rightarrow \overline{A}$.

## Eseményalgebra

Egy $\Omega$ eseménytér bizonyos eseményeiből álló $\mathcal{A}$ rendszert **eseményalgebrának** nevezünk, ha tartalmazza a biztos eseményt, és zárt a komplementerképzésre és a véges unióképzésre.

Például $\Omega$ összes részhalmazainak $\mathcal{A} := 2^{\Omega}$ rendszere.

Ha $\mathcal{A} \subset 2^{\Omega}$ eseményalgebra, akkor $\mathcal{A}$ tartalmazza a lehetetlen eseményt is, és zárt a különbségképzésre és a véges metszetképzésre is.

Természetes az a feltevés, hogy egy kísérlettel kapcsolatos események rendszere eseményalgebrát alkot.

## $\sigma$-algebra

Egy eseményalgebrát $\sigma$**-algebrának** nevezünk, ha zárt a megszámlálható unióképzésre.

**Példák:**

1. Egy pénzdarab feldobása esetén

$$\Omega = \{\text{fej}, \text{írás}\}.$$

De lehet a fejhez a 0, az íráshoz pedig az 1 számot hozzárendelni, és így

$$\Omega = \{0, 1\}.$$

Nyilván

$$\mathcal{A} = 2^{\Omega} = \{\emptyset, \{0\}, \{1\}, \Omega\}.$$

Ekkor az elemi események száma: $|\Omega| = 2$, az összes események száma pedig $|2^{\Omega}| = 4$.

2. *n*-szer **egymás után dobva** egy pénzdarabbal:

$$\Omega = \{\omega = (a_1, a_2, \ldots, a_n) : a_1, a_2, \ldots, a_n \in \{0, 1\}\}.$$

Ekkor $|\Omega| = 2^n$, $|2^\Omega| = 2^{2^n}$.

Ha $n$ darab egyforma pénzdarabot **egyidőben dobunk fel**, akkor is lehet ugyanezt az eseményteret tekinteni, hiszen a kísérlet kimenetelét nem változtatja meg, ha megszámozzuk a pénzdarabokat. De lehet csak a megkülönböztethető kimenetelekre szorítkozni: ezek száma $n+1$. Az első eseménytér általában alkalmasabb, mert például szabályos pénzdarab esetén az elemi események egyforma esélyűek!

3. Egy zsákban $n$ különböző színű golyó van. Kihúzunk ezek közül $k$ darabot; négy lehetőség van aszerint, hogy **visszatevéssel** vagy **visszatevés nélkül** húzunk (az utóbbi esetben $k \leqslant n$ szükséges), és aszerint, hogy a **sorrend számít** vagy a **sorrend nem számít.**

   Ez a kísérlet ekvivalens azzal a kísérlettel, amikor $n$ rekeszbe helyezünk el $k$ tárgyat; az előbbi négy lehetőség annak felel meg, hogy egy rekeszbe több tárgy is kerülhet vagy csak egy, illetve a tárgyak meg vannak különböztetve, vagy nem.

| | sorrend számít (variáció) | sorrend nem számít (kombináció) | |
|---|---|---|---|
| visszatevés nélkül (ismétlés nélkül) | $\frac{n!}{(n-k)!}$ | $\binom{n}{k}$ | egy rekeszbe legfeljebb egy tárgy kerülhet |
| visszatevéssel (ismétléses) | $n^k$ | $\binom{n+k-1}{k}$ | egy rekeszbe több tárgy is kerülhet |
| | a tárgyak különbözőek | a tárgyak nem különböznek | |

- Ha $n$ különböző elem közül húzunk visszatevés nélkül úgy, hogy a sorrend számít, és kihúzzuk az összes $n$ elemet (ami azzal ekvivalens, hogy $n$ elemet sorbaállítunk; ezeket **permutációknak** nevezzük), akkor a lehetőségek száma

$$|\Omega| = n! := 1 \cdot 2 \cdot \dots \cdot n,$$

hiszen az első húzásnál még $n$ lehetőség van, a másodiknál $n-1$, stb., és ezek szorzata adja az eredményt.

- Ha $n$ különböző elem közül $k$ elemet húzunk visszatevés nélkül (ahol $k \leqslant n$) úgy, hogy a sorrend számít (ezeket **ismétlés nélküli variációknak** nevezzük), akkor a lehetőségek száma

$$|\Omega| = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!},$$

amit az előzőhöz hasonló gondolatmenettel bizonyíthatunk.

- Ha $n$ különböző elem közül $k$ elemet húzunk visszatevéssel úgy, hogy a sorrend számít (ezeket **ismétléses variációknak** nevezzük), akkor a lehetőségek száma
$$|\Omega| = n^k,$$
hiszen minden húzásnál $n$ lehetőség van.
- Ha $n$ különböző elem közül $k$ elemet húzunk visszatevés nélkül (ahol $k \leqslant n$) úgy, hogy a sorrend nem számít (ezeket **ismétlés nélküli kombinációknak** nevezzük), akkor a lehetőségek száma
$$|\Omega| = \binom{n}{k} := \frac{n!}{k!\,(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!},$$
hiszen a megfelelő ismétlés nélküli variációkat úgy lehet megkapni, hogy a kihúzott $k$ elemet az összes lehetséges módon sorbarakjuk; ezek száma pedig mindig $k!$.

- Ha $n$ különböző elem közül $k$ elemet húzunk visszatevéssel úgy, hogy a sorrend nem számít (ezeket **ismétléses kombinációknak** nevezzük), akkor a lehetõségek száma
$$|\Omega| = \binom{n+k-1}{k}.$$
Ezt úgy lehet belátni, hogy a kísérlet kimeneteleihez egyértelműen hozzá lehet rendelni egy olyan sorozatot, mely $n-1$ darab egyesből és $k$ darab nullából áll, mégpedig úgy, hogy az első egyes elé írt nullák száma (ami 0 is lehet) jelenti az első fajta elemből húzottak számát, az első és második egyes közé írt nullák száma jelenti a második fajta elemből húzottak számát, stb., az $(n-1)$-edik egyes után írt nullák száma jelenti az $n$-edik fajta elemből húzottak számát; az ilyen nulla–egy sorozatok száma pedig nyilván $\binom{n+k-1}{k}$, hiszen azt kell megmondani, hogy az $n+k-1$ hely közül melyik $k$ helyre kerüljön nulla.

4. Adva van $n$ kártya; ezeket osztjuk szét $k$ játékos között úgy, hogy sorban $n_1, n_2, \ldots, n_k$ kártyát kapjanak, ahol $n_1 + n_2 + \cdots + n_k = n$, és az egy játékoshoz kerülő lapok sorrendje nem számít (ezeket **ismétléses permutációknak** nevezzük). Ekkor az eseménytér elemeinek száma
$$|\Omega| = \frac{n!}{n_1!\, n_2! \cdots n_k!},$$
hiszen a kártyák $n!$ számú permutációit úgy lehet ezekből a leosztásokból megkapni, hogy az egy játékoshoz került $n_1$, $n_2$, $\ldots$, $n_k$ kártyát tetszőleges sorrendbe helyezzük.
5. Addig dobálunk egy érmével, míg az első fejet sikerül elérni. Ekkor
$$\Omega = \{\mathrm{f}, \mathrm{if}, \mathrm{iif}, \mathrm{iiif}, \ldots, \mathrm{i}_\infty\},$$
ahol $\mathrm{i}_\infty$ azt a lehetséges kimenetelt jelöli, amikor csak írást dobunk a végtelenségig.

### Gyakoriság, relatív gyakoriság

Ha egy $A$ eseménnyel kapcsolatban $n$ darab véletlen, független kísérletetet hajtunk végre, akkor $A$ **gyakorisága** az a szám, ahányszor $A$ bekövetkezik; ez egy **véletlen mennyiség**, melynek lehetséges értékei: $0, 1, \ldots, n$; jelölése: $k_n(A)$.

Az $A$ esemény **relatív gyakorisága:** $r_n(A) := \dfrac{k_n(A)}{n}$; ez is **véletlen mennyiség**, melynek lehetséges értékei: $0, \dfrac{1}{n}, \dfrac{2}{n}, \ldots, \dfrac{n}{n} = 1$;

**Tapasztalat:**
ha $n$-et növeljük, azaz egyre több kísérletet hajtunk végre, akkor az $A$ esemény relatív gyakorisága egyre kisebb kilengésekkel ingadozik egy $\mathrm{P}(A)$ szám körül; amit majd $A$ **valószínűségének** hívunk.

## Relatív gyakoriság tulajdonságai

- $0 \leqslant r_n(A) \leqslant 1$ tetszőleges $A$ esemény esetén.
- $r_n(\emptyset) = 0$, $r_n(\Omega) = 1$.
- ha $A$ és $B$ egymást kizáró események, akkor
$$r_n(A \cup B) = r_n(A) + r_n(B).$$
- ha $A_1, A_2, \ldots$ páronként egymást kizáró események, akkor
$$r_n\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} r_n(A_j).$$
- $r_n(\overline{A}) = 1 - r_n(A)$ tetszőleges $A$ esemény esetén.
- ha $A \subset B$ események, akkor $r_n(A) \leqslant r_n(B)$.

## Valószínűségi mező (Kolmogorov)

$(\Omega, \mathcal{A}, \mathrm{P})$ hármas, ahol

- $\Omega$ egy nemüres halmaz (az eseménytér);
- $\mathcal{A} \subset 2^{\Omega}$ az $\Omega$ bizonyos részhalmazaiból álló $\sigma$-algebra (az események rendszere);
- $\mathrm{P} : \mathcal{A} \to \mathbb{R}$ olyan leképezés (halmazfüggvény), melyre
  1. $\mathrm{P}(A) \in [0, 1]$ tetszőleges $A \in \mathcal{A}$ esetén,
  2. $\mathrm{P}(\Omega) = 1$,
  3. ha $A_1, A_2, \ldots \in \mathcal{A}$ páronként diszjunktak, akkor
$$\mathrm{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathrm{P}(A_j).$$
     (Ezt a tulajdonságot $\sigma$**-additivitásnak** nevezzük).

Egy $A$ esemény esetén a $\mathrm{P}(A)$ számot az $A$ **valószínűségének**, a $\mathrm{P} : \mathcal{A} \to \mathbb{R}$ leképezést pedig **valószínűségeloszlásnak** (valószínűségi mértéknek) nevezzük.

## A valószínűségeloszlások tulajdonságai

- $\mathrm{P}(\emptyset) = 0$. (Hiszen ha $\mathrm{P}(\emptyset) > 0$ volna, akkor a $\sigma$-additivitásban $A_1 = A_2 = \ldots = \emptyset$ választással ellentmondásra jutnánk.)
- Ha $A_1, A_2, \ldots, A_n \in \mathcal{A}$ páronként diszjunktak, akkor
$$\mathrm{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mathrm{P}(A_j).$$
(Használjuk a $\sigma$-additivitást $A_{n+1} = A_{n+2} = \ldots = \emptyset$ esetére, és alkalmazzuk azt, hogy $\mathrm{P}(\emptyset) = 0$.)
Ezt a tulajdonságot **véges additivitásnak** nevezzük.
- $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$.
(Hiszen $\Omega = A \cup \overline{A}$ diszjunkt felbontás, így a véges additivitással $1 = \mathrm{P}(\Omega) = \mathrm{P}(A \cup \overline{A}) = \mathrm{P}(A) + \mathrm{P}(\overline{A})$.)

## A valószínűségeloszlások tulajdonságai

- Ha $A \Rightarrow B$, azaz $A \subset B$, akkor

$$\mathrm{P}(A) \leqslant \mathrm{P}(B), \qquad \mathrm{P}(B \setminus A) = \mathrm{P}(B) - \mathrm{P}(A).$$

  (Hiszen $A \subset B$ esetén $B = A \cup (B \setminus A)$ diszjunkt felbontás, ezért $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(B \setminus A) \geqslant \mathrm{P}(A).$)
  Az első tulajdonságot **monotonitásnak** nevezzük.

- Tetszőleges $A, B \in \mathcal{A}$ esetén

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B),$$

  hiszen

$$A \cup B = \big[A \setminus (A \cap B)\big] \cup \big[B \setminus (A \cap B)\big] \cup (A \cap B)$$

  diszjunkt felbontás, ezért $A \cap B \subset A$ és $A \cap B \subset B$ miatt

$$\mathrm{P}(A \cup B) = \big[\mathrm{P}(A) - \mathrm{P}(A \cap B)\big] + \big[\mathrm{P}(B) - \mathrm{P}(A \cap B)\big] + \mathrm{P}(A \cap B).$$

## A valószínűségeloszlások tulajdonságai

- Tetszőleges $A, B, C \in \mathcal{A}$ esetén

$$\begin{aligned} \mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) \\ - \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) \\ + \mathrm{P}(A \cap B \cap C), \end{aligned}$$

hiszen

$$\mathrm{P}\left((A \cup B) \cup C\right) = \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}\left((A \cup B) \cap C\right)$$

és

$$\begin{aligned} \mathrm{P}\left((A \cup B) \cap C\right) &= \mathrm{P}\left((A \cap C) \cup (B \cap C)\right) \\ &= \mathrm{P}(A \cap C) + \mathrm{P}(B \cap C) - \mathrm{P}\left((A \cap C) \cap (B \cap C)\right), \end{aligned}$$

ahol $(A \cap C) \cap (B \cap C) = A \cap B \cap C$.
Ez a tulajdonság a **szita-formula** speciális esete 3 eseményre.

## A valószínűségeloszlások tulajdonságai

- **Szita-formula (Poincaré-formula):** Tetszőleges $A_1, \ldots, A_n \in \mathcal{A}$ esetén

$$\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k-1} S_k^{(n)},$$

ahol

$$S_k^{(n)} = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} \mathrm{P}(A_{i_1} \cap \cdots \cap A_{i_k}), \qquad n \in \mathbb{N}.$$

Azaz, események $n$ tagú uniójának a valószínűsége az események legfeljebb $n$ tagú metszeteinek a valószínűségeivel kifejezhető.

## A valószínűségeloszlások tulajdonságai

Továbbá,

$$\sum_{k=1}^{2\ell}(-1)^{k-1}S_k^{(n)} \leqslant \mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{2\ell-1}(-1)^{k-1}S_k^{(n)}, \qquad \ell \in \mathbb{N},$$

ahol $S_k^{(n)} := 0$, ha $k > n$. Ez az egyenlőtlenség azt jelenti, hogy a $\mathrm{P}(A_1 \cup \cdots \cup A_n)$ valószínűséget a szita-formulában kifejező

$$S_1^{(n)} - S_2^{(n)} + S_3^{(n)} - S_4^{(n)} + \cdots$$

előjeles összeg páratlan számú tagot tartalmazó részletösszegei felülről, míg a páros számú tagot tartalmazó részletösszegei alulról közelítik.

# 2. Valószínűség

Legyen $\Omega$ nemüres halmaz. Legyen minden $n \in \mathbb{N}$ esetén $A_n \subset \Omega$.

Ha $A_1 \subset A_2 \subset \dots$ és $A := \bigcup_{n=1}^{\infty} A_n$, akkor azt írjuk, hogy $A_n \uparrow A$.

Ha $A_1 \supset A_2 \supset \dots$ és $A := \bigcap_{n=1}^{\infty} A_n$, akkor azt írjuk, hogy $A_n \downarrow A$.

## A valószínűségeloszlások tulajdonságai

- Tetszőleges $A_1, A_2, \dots \in \mathcal{A}$, $A_n \uparrow A$ esetén
$$\lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}(A).$$
Ezt a tulajdonságot **alulról folytonosságnak** nevezzük.
- Tetszőleges $A_1, A_2, \dots \in \mathcal{A}$, $A_n \downarrow A$ esetén
$$\lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}(A).$$
Ezt a tulajdonságot **felülről folytonosságnak** nevezzük.

# 2. Valószínűség

## Diszkrét valószínűségi mező

$\Omega$ véges vagy megszámlálhatóan végtelen, azaz

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$$

vagy

$$\Omega = \{\omega_1, \omega_2, \ldots\}$$

alakú, és $\mathcal{A} = 2^{\Omega}$.

## Valószínűségek kiszámolása diszkrét valószínűségi mezőben

Tetszőleges $A \in \mathcal{A}$ esemény előáll az

$$A = \bigcup_{i\,:\,\omega_i \in A} \{\omega_i\}$$

diszjunkt felbontás alakjában, így

$$\mathrm{P}(A) = \sum_{i\,:\,\omega_i \in A} \mathrm{P}(\{\omega_i\}).$$

Ezért diszkrét valószínűségi mezőben elég megadni az elemi események valószínűségeit, a

$$p_i := \mathrm{P}(\{\omega_i\}), \qquad i = 1, 2, \ldots$$

számokat ahhoz, hogy tetszőleges esemény valószínűségét ki tudjuk számolni.

Nyilván szükséges az, hogy ezek a $\{p_1, p_2, \ldots\}$ számok nemnegatívak legyenek és összegük 1 legyen, hiszen

$$\sum_i p_i = \sum_i \mathrm{P}(\{\omega_i\}) = \mathrm{P}\left(\bigcup_i \{\omega_i\}\right) = \mathrm{P}(\Omega) = 1.$$

Az is igaz, hogy ha $\{p_1, p_2, \ldots\}$ nemnegatív, 1 összegű számok, akkor a fentieknek megfelelően bevezetett $\mathrm{P} : \mathcal{A} \to \mathbb{R}$ függvény valószínűség.
Ha a $\{p_1, p_2, \ldots\}$ számok nemnegatívak és 1 összegűek, akkor azt mondjuk, hogy **eloszlást** alkotnak.

# 2. Valószínűség

## Egyenletes eloszlás véges halmazon

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\},$$

és az elemi események egyenlő esélyűek, azaz

$$\mathrm{P}(\{\omega_1\}) = \mathrm{P}(\{\omega_2\}) = \ldots = \mathrm{P}(\{\omega_N\}) = \frac{1}{N}.$$

## Valószínűségek véges halmazon egyenletes eloszlás esetén

$$\mathrm{P}(A) = \sum_{i\,:\,\omega_i \in A} \mathrm{P}(\{\omega_i\}) = \frac{1}{N} \sum_{i\,:\,\omega_i \in A} 1 = \frac{|A|}{N},$$

vagyis

$$\mathrm{P}(A) = \frac{\text{kedvező kimenetelek száma}}{\text{összes kimenetelek száma}}.$$

Ez a **valószínűség kiszámításának klasszikus képlete.**

### Klasszikus valószínűségi mező

Olyan $(\Omega, \mathcal{A}, \mathrm{P})$ valószínűségi mező, ahol

- $\Omega$ véges, azaz $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ alakú, ahol $N \in \mathbb{N}$,
- $\mathcal{A} = 2^\Omega$,
- $$\mathrm{P}(\{\omega_1\}) = \mathrm{P}(\{\omega_2\}) = \ldots = \mathrm{P}(\{\omega_N\}) = \frac{1}{N}.$$

Azaz, véges sok elemi esemény van, az elemi események tetszőleges halmaza esemény, és az elemi események egyenlő valószínűségűek.

**Példák:**

1. *Két érmét feldobva mennyi annak a valószínűsége, hogy egy fej és egy írás legyen az eredmény?*

   Ekkor a két érmét megkülönböztetve az

   $$\Omega = \{\mathrm{ff}, \mathrm{fi}, \mathrm{if}, \mathrm{ii}\}$$

   eseményteret kapjuk, amelyben a kimenetelek egyforma valószínűségűek, így az

   $$A = \{\mathrm{fi}, \mathrm{if}\}$$

   esemény valószínűsége

   $$\mathrm{P}(A) = \frac{2}{4} = \frac{1}{2}.$$

**Példák:**

2. *Mennyi a valószínűsége, hogy egy $n$ tagú társaságban van legalább két olyan személy, akiknek ugyanakkor van a születésnapja? (Feltesszük, hogy a szökőnap nem lehet.)*

   Nyilván $n > 365$ esetén (a „skatulya-elv” miatt) ez a biztos esemény, így ekkor a valószínűség 1.

   Ha pedig $n \leqslant 365$, akkor az ellentett eseménnyel számolva

$$\mathrm{P}(A) = 1 - \frac{365 \cdot 364 \cdots (365 - n + 1)}{365^n}$$

$$= 1 - \frac{365!}{(365-n)! \cdot 365^n} \approx \begin{cases} 0.284 & \text{ha } n = 16, \\ 0.476 & \text{ha } n = 22, \\ 0.507 & \text{ha } n = 23, \\ 0.891 & \text{ha } n = 40, \\ 0.970 & \text{ha } n = 50, \\ 0.990 & \text{ha } n = 57. \end{cases}$$

## Egyenletes eloszlás $\mathbb{R}^k$ véges mértékű részhalmazain

$\Omega$: $\mathbb{R}^k$ egy véges (Lebesgue-)mértékű részhalmaza,
$\mathcal{A}$: $\Omega$ Borel-halmazaiból álló $\sigma$-algebra,
P: ,,minden pont egyenlő esélyű", pontosabban egy $A \in \mathcal{A}$ részhalmaz valószínűsége $A$ mértékével arányos, azaz

$$\mathrm{P}(A) = \frac{\mu(A)}{\mu(\Omega)},$$

ahol $\mu$ az illető halmaz $k$-dimenziós (Lebesgue-)mértékét jelöli:

- $k = 1$ esetén hossz,
- $k = 2$ esetén terület,
- $k = 3$ esetén térfogat.

Ez a **valószínűség geometriai kiszámítási módja**.

**Példa:** *Egy egységnyi hosszúságú szakaszt két, találomra kiválasztott ponttal három szakaszra bontunk fel. Mennyi annak a valószínűsége, hogy a három szakaszból háromszöget lehet szerkeszteni?*

Az eredmény a $[0,1]\times[0,1]$ négyzet azon részhalmazának területe, melynek pontjaira fennállnak a következő egyenlőtlenségek:

$$\begin{cases} 0<x<y<1, \\ 1-y<x+(y-x)=y, \\ x<(y-x)+(1-y)=1-x, \\ y-x<x+(1-y), \end{cases} \quad \text{vagy} \quad \begin{cases} 0<y<x<1, \\ 1-x<y+(x-y)=x, \\ y<(x-y)+(1-x)=1-y, \\ x-y<y+(1-x), \end{cases}$$

azaz

$$\begin{cases} 0<x<\frac{1}{2}<y<1, \\ y-x<\frac{1}{2}, \end{cases} \quad \text{vagy} \quad \begin{cases} 0<y<\frac{1}{2}<x<1, \\ x-y<\frac{1}{2}. \end{cases}$$

Ezért a keresett valószínűség $1/4$.

## Feltételes relatív gyakoriság

Ha $n$ független kísérletet végzünk, akkor az $A$ esemény **feltételes relatív gyakorisága azon feltétel mellett, hogy a $B$ esemény bekövetkezett**

$$r_n(A \mid B) := \frac{k_n(A \cap B)}{k_n(B)} = \frac{r_n(A \cap B)}{r_n(B)}.$$

## Feltételes valószínűség

Az $A$ esemény **feltételes valószínűsége** a $B$ feltétel mellett (azaz ha tudjuk, hogy a $B$ esemény bekövetkezett)

$$\mathrm{P}(A \mid B) := \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)},$$

feltéve, hogy $\mathrm{P}(B) > 0$.

**Példák:**

1. *Mennyi a valószínűsége, hogy egy kétgyermekes családban mindkét gyerek fiú, ha feltételezzük, hogy egy gyerek egyenlő valószínűséggel lehet fiú vagy lány, és tudjuk, hogy*
   - *az idősebb gyerek fiú;*
   - *legalább az egyik gyerek fiú ?*

   Ekkor az eseménytér
   $$\Omega = \{\mathrm{FF}, \mathrm{FL}, \mathrm{LF}, \mathrm{LL}\},$$
   melynek elemei egyformán $1/4$ valószínűségűek, ahol pl. FL annak felel meg, hogy az 1. (idősebb) gyerek fiú és a 2. (fiatalabb) gyerek lány. Legyen
   $$\begin{aligned} A &:= \{\text{mindkét gyerek fiú}\} = \{\mathrm{FF}\}, \\ B_1 &:= \{\text{az idősebb gyerek fiú}\} = \{\mathrm{FF}, \mathrm{FL}\}, \\ B_2 &:= \{\text{legalább az egyik gyerek fiú}\} = \{\mathrm{FF}, \mathrm{FL}, \mathrm{LF}\}. \end{aligned}$$
   Nyilván $A \cap B_1 = A \cap B_2 = \{\mathrm{FF}\}$, így
   $$\mathrm{P}(A \mid B_1) = 1/2, \qquad \mathrm{P}(A \mid B_2) = 1/3.$$

2. *Az 52 lapos francia kártyát kiosztjuk 4 embernek, mindenki 13 lapot kap. Tudjuk, hogy az egyik ember 2 ászt kapott. Mennyi a valószínűsége, hogy a másik 2 ász a partnerénél van (csak egy partnere van)?*

   Az összes leosztások száma $\dfrac{52!}{(13!)^4}$, ezek egyforma valószínűségűek. Ezekből $\dbinom{4}{2} \cdot \dbinom{48}{11} \cdot \dfrac{39!}{(13!)^3}$ olyan leosztás van, melynél az első játékos 2 ászt kap, és ezek között pedig

$$\binom{4}{2} \cdot \binom{48}{11} \cdot \binom{37}{11} \cdot \frac{26!}{(13!)^2}$$

   olyan leosztás van, melynél a másik 2 ász a partnerénél van. Tehát a keresett feltételes valószínűség

$$\frac{\binom{4}{2} \cdot \binom{48}{11} \cdot \binom{37}{11} \cdot \frac{26!}{(13!)^2}}{\binom{4}{2} \cdot \binom{48}{11} \cdot \frac{39!}{(13!)^3}} = \frac{2}{19}.$$

**Láncszabály**

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n)$$
$$= \mathrm{P}(A_1)\,\mathrm{P}(A_2 \mid A_1)\,\mathrm{P}(A_3 \mid A_1 \cap A_2) \cdots \mathrm{P}(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}),$$

feltéve, hogy $\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$.

A jobb oldal

$$\mathrm{P}(A_1)\,\frac{\mathrm{P}(A_1 \cap A_2)}{\mathrm{P}(A_1)}\,\frac{\mathrm{P}(A_1 \cap A_2 \cap A_3)}{\mathrm{P}(A_1 \cap A_2)} \cdots \frac{\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n)}{\mathrm{P}(A_1 \cap A_2 \cdots \cap A_{n-1})}$$

**Példa:** *Húzzunk ki a 32 lapos magyar kártyából hármat visszatevés nélkül. Mennyi a valószínűsége, hogy az első és a harmadik kihúzott lap piros, a második pedig nem az?*

Jelölje $i = 1, 2, 3$ esetén $A_i$ azt az eseményt, hogy az $i$-edik húzás eredménye piros. Ekkor

$$\mathrm{P}(A_1) = \frac{8}{32} = \frac{1}{4}, \qquad \mathrm{P}(\overline{A}_2 \mid A_1) = \frac{24}{31}, \qquad \mathrm{P}(A_3 \mid A_1 \cap \overline{A}_2) = \frac{7}{30},$$

így

$$\mathrm{P}(A_1 \cap \overline{A}_2 \cap A_3) = \frac{1}{4} \cdot \frac{24}{31} \cdot \frac{7}{30} = \frac{7}{155}.$$

Persze lehetne használni azt az eseményteret is, amely az első három kihúzott lapból áll a sorrendet is figyelembe véve; ekkor $|\Omega| = 32{\cdot}31{\cdot}30$, és a kimenetelek egyenlő valószínűségűek. Mivel a kedvező esetek száma $8 \cdot 24 \cdot 7$, így a keresett valószínűség $\frac{8\cdot24\cdot7}{32\cdot31\cdot30} = \frac{7}{155}$.

# 3. Feltételes valószínűség, független események

## Teljes eseményrendszer

Az eseménytér megszámlálható páronként diszjunkt felbontása eseményekre, azaz események $A_1, A_2, \ldots$ véges vagy megszámlálhatóan végtelen sorozata, melyek egymást páronként kizárják, és uniójuk az egész eseménytér, vagyis

$$A_i \cap A_j = \emptyset \quad \text{ha} \quad i \neq j, \qquad \text{és} \qquad \bigcup_i A_i = \Omega.$$

Egy teljes eseményrendszer eseményei közül mindig pontosan egy következik be, és

$$\sum_i \mathrm{P}(A_i) = 1.$$

## Teljes valószínűség tétele

Ha az $A_1, A_2, \ldots$ pozitív valószínűségű események teljes eseményrendszert alkotnak, akkor tetszőleges $B$ eseményre

$$\mathrm{P}(B) = \sum_i \mathrm{P}(B \mid A_i) \cdot \mathrm{P}(A_i).$$

**Bizonyítás.** Nyilván $B = \bigcup_i (B \cap A_i)$ diszjunkt felbontás, hiszen

$$B = B \cap \Omega = B \cap (\bigcup_i A_i) = \bigcup_i (B \cap A_i),$$

és $i \neq j$ esetén

$$(B \cap A_i) \cap (B \cap A_j) = B \cap A_i \cap A_j = \emptyset,$$

ugyanis $A_i \cap A_j = \emptyset$. Ezért

$$\mathrm{P}(B) = \sum_i \mathrm{P}(B \cap A_i) = \sum_i \mathrm{P}(B \mid A_i) \cdot \mathrm{P}(A_i).$$

**Példa:** *Három gép csavarokat gyárt. A selejt aránya az első gépnél* 1 %*, a másodiknál* 2 %*, a harmadiknál* 3 %*. Az össztermék* 50 %*-át az első gép,* 30 %*-át a második,* 20 %*-át pedig a harmadik állítja elő. Mi a valószínűsége annak, hogy az össztermékből véletlenszerűen választott csavar selejtes?*

Jelölje $B$ azt az eseményt, hogy selejtet húzunk, $i = 1, 2, 3$ esetén pedig $A_i$ azt, hogy a kihúzott csavar az $i$-edik gépen készült. Ekkor

$$\mathrm{P}(B|A_1) = 0.01, \qquad \mathrm{P}(B|A_2) = 0.02, \qquad \mathrm{P}(B|A_3) = 0.03,$$
$$\mathrm{P}(A_1) = 0.5, \qquad \mathrm{P}(A_2) = 0.3, \qquad \mathrm{P}(A_3) = 0.2,$$

így

$$\mathrm{P}(B) = 0.01 \cdot 0.5 + 0.02 \cdot 0.3 + 0.03 \cdot 0.2 = 0.017.$$

# 3. Feltételes valószínűség, független események

## Bayes-formula

Ha $A$ és $B$ pozitív valószínűségű események, akkor

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A) \cdot \mathrm{P}(B \mid A)}{\mathrm{P}(B)}.$$

**Bizonyítás:** $\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ és $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A)$.

## Bayes-tétel

Ha az $A_1, A_2, \ldots$ pozitív valószínűségű események teljes eseményrendszert alkotnak és $\mathrm{P}(B) > 0$, akkor

$$\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(A_i) \cdot \mathrm{P}(B \mid A_i)}{\sum\limits_j \mathrm{P}(B \mid A_j) \cdot \mathrm{P}(A_j)}.$$

**Bizonyítás:** A Bayes-formulával $\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(A_i) \cdot \mathrm{P}(B \mid A_i)}{\mathrm{P}(B)}$. A teljes valószínűség tételével $\mathrm{P}(B) = \sum\limits_j \mathrm{P}(B \mid A_j) \cdot \mathrm{P}(A_j)$.

**Példa:** *Mennyi a feltételes valószínűsége az előző példában annak, hogy az első, második, illetve harmadik gépen gyártották a kiválasztott csavart azon feltétel mellett, hogy az selejtesnek bizonyult?*

$$\mathrm{P}(A_1 \mid B) = \frac{0.5 \cdot 0.01}{0.017} = \frac{5}{17}, \qquad \mathrm{P}(A_2 \mid B) = \frac{6}{17}, \qquad \mathrm{P}(A_3 \mid B) = \frac{6}{17}.$$

## Független események

Azt mondjuk, hogy az $A$ és $B$ események **függetlenek**, ha

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B).$$

Ha $A$, $B$ és $\overline{B}$ **pozitív** valószínűségű események, akkor a következő állítások ekvivalensek:

- $A$ és $B$ függetlenek;
- $\mathrm{P}(A \,|\, B) = \mathrm{P}(A)$;
- $\mathrm{P}(B \,|\, A) = \mathrm{P}(B)$;
- $A$ és $\overline{B}$ függetlenek;
- $\mathrm{P}(A \,|\, B) = \mathrm{P}(A \,|\, \overline{B})$.

Ha $\mathrm{P}(A) = 0$ vagy $\mathrm{P}(A) = 1$, akkor $A$ tetszőleges $B$ eseménytől független.

## Páronként független események

Azt mondjuk, hogy az $A_1, A_2, \ldots$ események **páronként függetlenek**, ha közülük bármely két esemény független.

## (Teljesen) független események

Azt mondjuk, hogy az $A_1, A_2, \ldots$ események **(teljesen) függetlenek**, ha tetszőleges $i_1, i_2, \ldots, i_k$ **páronként különböző** indexekre

$$\mathrm{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathrm{P}(A_{i_1})\,\mathrm{P}(A_{i_2}) \cdots \mathrm{P}(A_{i_k}).$$

Lehetséges, hogy például három esemény páronként független, de nem (teljesen) független.

## Páronként függetlenség $\not\Rightarrow$ függetlenség

Dobjunk fel egy szabályos pénzdarabot kétszer egymás után. Legyen

$$A := \{\text{az első dobás fej}\}, \quad B := \{\text{a második dobás fej}\},$$
$$C := \{\text{a két dobás eredménye különböző}\}.$$

Mivel

$$\mathrm{P}(A) = \frac{1}{2}, \quad \mathrm{P}(B) = \frac{1}{2}, \quad \mathrm{P}(C) = \frac{2}{4} = \frac{1}{2}, \quad \mathrm{P}(A \cap B) = \frac{1}{4},$$
$$\mathrm{P}(A \cap C) = \mathrm{P}(\text{az első dobás fej, a második dobás írás}) = \frac{1}{4},$$
$$\mathrm{P}(B \cap C) = \mathrm{P}(\text{az első dobás írás, a második dobás fej}) = \frac{1}{4},$$

kapjuk, hogy $A$, $B$ és $C$ páronként függetlenek. Azonban,

$$\mathrm{P}(A \cap B \cap C) = \mathrm{P}(\emptyset) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C),$$

így $A$, $B$ és $C$ nem függetlenek.

## Valószínűségi változó / véletlen mennyiség, eloszlásfüggvény

Ha $(\Omega, \mathcal{A}, \mathrm{P})$ valószínűségi mező, akkor a $\xi : \Omega \to \mathbb{R}$ leképezés **valószínűségi változó/ véletlen változó/ véletlen mennyiség**, ha tetszőleges $x \in \mathbb{R}$ esetén $\{\omega \in \Omega : \xi(\omega) < x\} \in \mathcal{A}$. Ekkor az $F_\xi : \mathbb{R} \to [0, 1]$,

$$F_\xi(x) := \mathrm{P}\{\xi < x\}, \quad x \in \mathbb{R},$$

függvényt $\xi$ **(kumulatív) eloszlásfüggvényének** nevezzük.

## Eloszlásfüggvény jellemzése

Egy $F : \mathbb{R} \to [0, 1]$ függvény akkor és csak akkor lehet eloszlásfüggvénye valamely $\xi : \Omega \to \mathbb{R}$ véletlen változónak, ha

1. $F$ monoton növekvő,
2. $F$ balról folytonos,
3. $\lim_{x\to-\infty} F(x) = 0, \qquad \lim_{x\to+\infty} F(x) = 1.$

## Eloszlásfüggvény

Legyen $\xi$ valószínűségi változó $F_\xi$ eloszlásfüggvénnyel. Tetszőleges $a, b \in \mathbb{R}$, $a < b$ esetén

$$\mathrm{P}\{a \leqslant \xi < b\} = F_\xi(b) - F_\xi(a).$$

Tetszőleges $c \in \mathbb{R}$ esetén

$$\mathrm{P}\{\xi < c\} = F_\xi(c) = \lim_{x \uparrow c} F_\xi(x), \qquad \mathrm{P}\{\xi \leqslant c\} = \lim_{x \downarrow c} F_\xi(x) = F_\xi(c+0),$$

így $\mathrm{P}\{\xi = c\}$ az $F_\xi$ ugrása a $c$ pontban, azaz

$$\mathrm{P}\{\xi = c\} = \lim_{x \downarrow c} F_\xi(x) - \lim_{x \uparrow c} F_\xi(x) = F_\xi(c+0) - F_\xi(c).$$

## Diszkrét véletlen változó

A $\xi : \Omega \to \mathbb{R}$ véletlen változó **diszkrét**, ha lehetséges értékeinek halmaza, a $\{\xi(\omega) : \omega \in \Omega\}$ értékkészlet megszámlálható.

A $\xi$ diszkrét véletlen változó **eloszlása** az a $\mathrm{P}_\xi$ mérték a $\xi$ lehetséges értékeinek $X := \{x_1, x_2, \dots\}$ halmazán, melyre $\mathrm{P}_\xi(\{x_i\}) = \mathrm{P}(\{\omega \in \Omega : \xi(\omega) = x_i\})$, $x_i \in X$.

## Diszkrét véletlen változó eloszlásfüggvénye

Egy $\xi$ diszkrét véletlen változó eloszlásfüggvénye olyan lépcsős függvény, mely a lehetséges értékeknél ugrik, és az ugrás nagysága az illető érték valószínűsége. Ha a $\xi$ lehetséges értékeinek halmaza $X := \{x_1, x_2, \dots\}$, akkor

$$F_\xi(x) = \sum_{\{i \,:\, x_i < x\}} \mathrm{P}_\xi(\{x_i\}), \qquad x \in \mathbb{R}.$$

**Példák:**

1. *Két kockát dobva a dobott számok összegét jelölje $\xi$. Határozzuk meg $\xi$ eloszlását!*

   Ekkor $\xi$ diszkrét véletlen változó; lehetséges értékeinek halmaza:

$$X = \{2, 3, \dots, 12\},$$

   eloszlása:

$$\mathrm{P}\{\xi = k\} = \begin{cases} \dfrac{k-1}{36} & \text{ha } 2 \leqslant k \leqslant 7, \\[2ex] \dfrac{13-k}{36} & \text{ha } 7 \leqslant k \leqslant 12. \end{cases}$$

2. **Binomiális eloszlás.**

   $n$ független kísérlet, $n \in \mathbb{N}$
   $A$ esemény, $p := \mathsf{P}(A) \in [0,1]$
   $A$ gyakorisága: $\xi := k_n(A)$ diszkrét véletlen változó;
   $\xi$ lehetséges értékeinek halmaza:

   $$X = \{0, 1, 2, \ldots, n\},$$

   eloszlása

   $$\mathsf{P}\{\xi = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in X,$$

   melyet $(n, p)$ **paraméterű binomiális eloszlásnak** nevezünk.

3. **Elsőrendű negatív binomiális eloszlás.**

kísérlet, $A$ esemény, $p := \mathrm{P}(A) \in (0, 1]$
Addig ismételjük egymás után függetlenül a kísérletet, míg $A$ először bekövetkezik.
$\xi :=$ az ehhez szükséges ismétlések száma;
lehetséges értékeinek halmaza:

$$X = \{1, 2, \dots, \infty\},$$

eloszlása: $k = 1, 2, \dots$ esetén

$$\mathrm{P}\{\xi = k\} = p \cdot (1 - p)^{k-1},$$

így

$$\mathrm{P}\{\xi = \infty\} = 1 - \mathrm{P}\{\xi < \infty\} = 1 - \sum_{k=1}^{\infty} \mathrm{P}\{\xi = k\} = 1 - p \sum_{k=1}^{\infty} (1 - p)^{k-1} = 0.$$

Ekkor $\xi$ eloszlását **elsőrendű** $p$ **paraméterű negatív binomiális eloszlásnak** (vagy geometriai eloszlásnak) nevezzük.

4. **Hipergeometrikus eloszlás.**

   Egy dobozban $M$ piros és $N-M$ fekete golyó van $(M < N)$.
   Visszatevés nélkül húzunk ki $n$ golyót $(n \leqslant N)$.
   $\xi :=$ a kihúzott piros golyók száma;
   lehetséges értékeinek halmaza: olyan $k$ értékek, melyekre teljesül $0 \leqslant k \leqslant n$, $k \leqslant M$, és $n-k \leqslant N-M$,
   eloszlása:

$$\mathsf{P}\{\xi = k\} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

   Ekkor $\xi$ eloszlását $(n, M, N-M)$ **paraméterű hipergeometrikus eloszlásnak** nevezzük.

5. **Poisson eloszlás.**

Mazsolás kalácsot sütünk; 1000 gramm tésztába $n = 50$ darab mazsolát teszünk. Egy szelet súlya 25 gramm, tehát $N = 40$ szelet készül. Minden mazsola egyforma valószínűséggel kerülhet bele bármely szeletbe, és a mazsolák egymástól függetlenül ,,mozognak". Jelölje $\xi$ egy kiválasztott szeletbe kerülő mazsolák számát. Lehetséges értékeinek halmaza $X = \{0, 1, \dots, 50\}$, eloszlása

$$\mathsf{P}\{\xi = k\} = \binom{50}{k} \left(\frac{1}{40}\right)^k \left(1 - \frac{1}{40}\right)^{50-k}, \qquad k \in X,$$

ugyanis annak a valószínűsége, hogy egy konkrét mazsola a kiválasztott szeletbe kerül $1/40$. Tehát $\xi$ eloszlása $n$-edrendű $1/N$ paraméterű binomiális eloszlás.

Mi történik, ha növeljük a tészta mennyiségét, ill. ezzel arányosan a mazsolák számát?

Ha $n$ mazsolát használunk fel $20 \cdot n$ gramm tésztához, akkor $N = 20 \cdot n/25$ szelet készül, így a szóbanforgó binomiális eloszlás $n$-edrendű és $p_n := 1/N = \lambda/n$ paraméterű, ahol $\lambda := 5/4 (= \frac{n}{N})$ az egy szeletre átlagosan jutó mazsolák száma. Ekkor

$$
\begin{aligned}
&\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} \\
&= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!n^k} \lambda^k \left(1-\frac{\lambda}{n}\right)^{n} \left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \frac{1}{k!} \cdot \lambda^k \cdot \mathrm{e}^{-\lambda} \cdot 1 = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}, \qquad k = 0, 1, 2, \ldots
\end{aligned}
$$

Ha egy $\eta$ véletlen változó lehetséges értékei a nemnegatív egész számok és $k = 0, 1, \ldots$ esetén

$$\mathsf{P}(\eta = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},$$

ahol $\lambda > 0$, akkor azt mondjuk, hogy $\eta$ eloszlása $\lambda$ **paraméterű Poisson-eloszlás.**

Valóban diszkrét valószínűségeloszlást adtunk meg, mert a megadott számok pozitívak (így nemnegatívak) és

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \mathrm{e}^{\lambda}\mathrm{e}^{-\lambda} = \mathrm{e}^{0} = 1.$$

## Véletlen változó sűrűségfüggvénye

Ha $(\Omega, \mathcal{A}, \mathrm{P})$ valószínűségi mező, $\xi : \Omega \to \mathbb{R}$ véletlen változó és létezik olyan $f_\xi : \mathbb{R} \to [0, \infty)$ (Borel-mérhető) függvény, melyre

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)\,\mathrm{d}t, \qquad x \in \mathbb{R},$$

akkor az $f_\xi$ függvényt a $\xi$ **sűrűségfüggvényének** nevezzük, és azt mondjuk, hogy a $\xi$ véletlen változó, illetve $\xi$ eloszlása **abszolút folytonos.**
(Abszolút folytonos véletlen változó esetén annak sűrűségfüggvénye nem egyértelműen meghatározott!)

## Sűrűségfüggvény jellemzése

Egy $f : \mathbb{R} \to [0, \infty)$ (Borel-mérhető) függvény akkor és csak akkor lehet sűrűségfüggvénye valamely $\xi : \Omega \to \mathbb{R}$ véletlen változónak, ha

$$\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1.$$

## Abszolút folytonos véletlen változó

Legyen $\xi$ abszolút folytonos véletlen változó $f_\xi$ sűrűségfüggvénnyel. Ekkor $F_\xi$ folytonos, és tetszőleges $a, b \in \mathbb{R}$, $a < b$ esetén

$$\mathsf{P}\{a \leqslant \xi < b\} = F_\xi(b) - F_\xi(a) = \int_a^b f_\xi(t)\,\mathrm{d}t.$$

Általánosabban: tetszőleges $B \subset \mathbb{R}$ (Borel-halmaz) esetén

$$\mathsf{P}\{\xi \in B\} = \int_B f_\xi(t)\,\mathrm{d}t.$$

Tetszőleges $c \in \mathbb{R}$ esetén

$$\mathsf{P}\{\xi = c\} = 0.$$

Ha az $f_\xi$ sűrűségfüggvény folytonos az $x \in \mathbb{R}$ pontban, akkor

$$F_\xi'(x) = f_\xi(x).$$

## Egyenletes eloszlás az $(a, b)$ intervallumon

Ha az $(a, b)$ intervallumon választunk véletlenszerűen egy $\xi$ pontot úgy, hogy egy $A \subset (a, b)$ részhalmazba esés valószínűsége az illető részhalmaz mértékével arányos, akkor $\xi$ eloszlásfüggvénye nyilván

$$F_\xi(x) = \frac{x-a}{b-a}\mathbb{1}_{[a,b)}(x) + \mathbb{1}_{[b,\infty)}(x) = \begin{cases} 0 & \text{ha } x < a, \\ \dfrac{x-a}{b-a} & \text{ha } a \leqslant x < b, \\ 1 & \text{ha } x \geqslant b. \end{cases}$$

Ekkor a $\xi$ véletlen változót **egyenletes eloszlásúnak** nevezzük a $(a, b)$ intervallumon. Továbbá az

$$f_\xi(x) = \frac{1}{b-a}\mathbb{1}_{(a,b)}(x) = \begin{cases} \dfrac{1}{b-a} & \text{ha } a < x < b, \\ 0 & \text{egyébként} \end{cases}$$

függvény sűrűségfüggvénye $\xi$-nek. Jelölés: $\xi \sim U(a, b)$.

## Normális eloszlás

Ha a $\xi$ véletlen változó sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}$$

alakú, ahol $m \in \mathbb{R}$, $\sigma > 0$, akkor azt mondjuk, hogy $\xi$ **normális eloszlású** $(m, \sigma^2)$ **paraméterekkel**. Jelölés: $\xi \sim \mathcal{N}(m, \sigma^2)$.
**Standard normális:** $m = 0$ és $\sigma = 1$, jelölés: $\xi \sim \mathcal{N}(0, 1)$.

Az, hogy $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$, abból következik, hogy

$$\left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_0^{2\pi}\left(\int_0^{\infty} r\mathrm{e}^{-r^2/2}\,\mathrm{d}r\right)\mathrm{d}\varphi = 2\pi\left[-\mathrm{e}^{-r^2/2}\right]_{r=0}^{r=\infty} = 2\pi,$$

## Normális eloszlás

ahol a második lépésben egy integráltranszformációt hajtottunk végre:

$$x = r\cos\varphi, \qquad y = r\sin\varphi, \qquad r \in [0,\infty),\ \varphi \in [0,2\pi).$$

Ezen transzformáció Jacobi mátrixa:

$$\begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix},$$

melynek determinánsa:

$$r\cos^2\varphi + r\sin^2\varphi = r.$$

## Exponenciális eloszlás

Jelölje a $\xi$ véletlen változó egy radioaktív atom élettartamát. Ez rendelkezik az úgynevezett **örökifjú tulajdonsággal:** ha $t, h > 0$, akkor

$$\mathsf{P}\{\xi > t + h \mid \xi > t\} = \mathsf{P}\{\xi > h\},$$

vagyis annak ellenére, hogy tudjuk, hogy az atom már megélt $t$ időt, a még hátralevő élettartam eloszlása éppen olyan, mint a teljes élettartam eredeti eloszlása. Mivel

$$\mathsf{P}\{\xi > t + h \mid \xi > t\} = \frac{\mathsf{P}(\{\xi > t + h\} \cap \{\xi > t\})}{\mathsf{P}\{\xi > t\}},$$

és $\mathsf{P}(\{\xi > t + h\} \cap \{\xi > t\}) = \mathsf{P}\{\xi > t + h\}$, ezért a $G(t) := \mathsf{P}\{\xi > t\}$ **túlélési függvényre** teljesül

$$\frac{G(t+h)}{G(t)} = G(h), \qquad \text{azaz} \qquad G(t+h) = G(t)G(h), \quad t > 0, h > 0.$$

## Exponenciális eloszlás

Be lehet látni, hogy ha $G$ *folytonos*, akkor létezik olyan $\lambda > 0$, hogy

$$G(t) = \mathrm{e}^{-\lambda t}, \qquad \text{ha } t > 0.$$

Mivel $P\{\xi > t\} = 1$, ha $t \leqslant 0$, kapjuk, hogy $\xi$ eloszlásfüggvénye

$$F_\xi(x) = P\{\xi < x\} = 1 - P\{\xi \geqslant x\} = 1 - \lim_{y \uparrow x} G(y)$$

$$= 1 - G(x) = \begin{cases} 0 & \text{ha } x < 0, \\ 1 - \mathrm{e}^{-\lambda x} & \text{ha } x \geqslant 0 \end{cases}$$

alakú, ahol $\lambda > 0$. Ezt az eloszlást $\lambda$ **paraméterű exponenciális eloszlásnak** nevezzük. Van sűrűségfüggvénye:

$$f_\xi(x) = \begin{cases} 0 & \text{ha } x < 0, \\ \lambda e^{-\lambda x} & \text{ha } x \geqslant 0. \end{cases}$$

Jelölés: $\xi \sim Exp(\lambda)$.

## Exponenciális eloszlás

**Bomlási állandó:**

$$
\begin{aligned}
\lim_{h\downarrow 0}\frac{1}{h}\mathrm{P}\{t\leqslant\xi<t+h\,|\,\xi>t\} &= \lim_{h\downarrow 0}\frac{1}{h}\frac{\mathrm{P}(\{t\leqslant\xi<t+h\}\cap\{\xi>t\})}{\mathrm{P}\{\xi>t\}}\\
&= \lim_{h\downarrow 0}\frac{1}{h}\frac{\mathrm{P}\{t\leqslant\xi<t+h\}}{\mathrm{P}\{\xi>t\}}\\
&= \lim_{h\downarrow 0}\frac{1}{h}\frac{(1-e^{-\lambda(t+h)})-(1-e^{-\lambda t})}{e^{-\lambda t}}\\
&= \lim_{h\downarrow 0}\frac{1}{h}(1-\mathrm{e}^{-\lambda h})=\lambda.
\end{aligned}
$$

## Véletlen vektor

$$\xi : \Omega \to \mathbb{R}^k, \quad \text{azaz} \quad \xi = (\xi_1, \ldots, \xi_k),$$

ahol $\xi_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, k$ véletlen változók;
**eloszlásfüggvénye:** $F_\xi : \mathbb{R}^k \to \mathbb{R}$,

$$F_\xi(x_1, \ldots, x_k) := \mathsf{P}\{\xi_1 < x_1, \ldots, \xi_k < x_k\}, \quad (x_1, \ldots, x_k) \in \mathbb{R}^k.$$

## Peremeloszlás függvények

Ha $\xi$ és $\eta$ együttes eloszlásfüggvénye $F_{\xi,\eta}$, akkor
$\xi$ eloszlásfüggvénye $F_\xi(x) = \lim_{y\to\infty} F_{\xi,\eta}(x, y)$, $x \in \mathbb{R}$,
és $\eta$ eloszlásfüggvénye $F_\eta(y) = \lim_{x\to\infty} F_{\xi,\eta}(x, y)$, $y \in \mathbb{R}$.

## Diszkrét véletlen vektor

A $\xi : \Omega \to \mathbb{R}^k$ véletlen vektor **diszkrét**, ha lehetséges értékeinek halmaza, a $\{\xi(\omega) : \omega \in \Omega\}$ értékkészlet megszámlálható.

Ha $(\xi, \eta) : \Omega \to \mathbb{R}^2$ diszkrét, akkor $\xi$ és $\eta$ is diszkrét. Ha $\xi$ és $\eta$ lehetséges értékei $x_1, x_2, \ldots$, illetve $y_1, y_2, \ldots$, akkor $(\xi, \eta)$ lehetséges értékeinek halmaza $\{(x_i, y_j) : i, j = 1, 2, \ldots\}$.
Ha ismerjük $(\xi, \eta)$ eloszlását, azaz a

$$\mathsf{P}\{\xi = x_i,\, \eta = y_j\} \qquad i, j = 1, 2, \ldots$$

valószínűségeket, akkor ki tudjuk számolni $\xi$ és $\eta$ eloszlását is:

$$\mathsf{P}\{\xi = x_i\} = \sum_j \mathsf{P}\{\xi = x_i,\, \eta = y_j\}, \quad \mathsf{P}\{\eta = y_j\} = \sum_i \mathsf{P}\{\xi = x_i,\, \eta = y_j\}.$$

Ezek $(\xi, \eta)$ **peremeloszlásai / marginális eloszlásai.**

A $\xi$ és $\eta$ együttes eloszlását, azaz $(\xi, \eta)$ eloszlását, szokás kontingencia táblázattal is megadni.

## Polinomiális eloszlás

$n$ független kísérletet hajtunk végre egy $A_1, A_2, \ldots, A_r$ teljes eseményrendszerre, $p_i := \mathrm{P}(A_i)$ (ekkor $p_1 + p_2 + \cdots + p_r = 1$). Ekkor $A_i$ gyakorisága: $\xi_i := k_n(A_i)$, és $\xi = (\xi_1, \xi_2, \ldots, \xi_r)$ diszkrét véletlen vektor; lehetséges értékeinek halmaza:

$$\left\{(k_1, k_2, \ldots, k_r) \in \mathbb{Z}^r : k_i \geqslant 0,\ k_1 + k_2 + \cdots + k_r = n\right\},$$

eloszlása

$$\mathrm{P}\{\xi_1 = k_1,\, \xi_2 = k_2, \ldots, \xi_r = k_r\} = \frac{n!}{k_1!k_2!\ldots k_r!}\, p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r},$$

melyet $(n, p_1, p_2, \ldots, p_r)$ **paraméterű polinomiális eloszlásnak** nevezünk. Peremeloszlásai binomiális eloszlások:

$$\mathrm{P}\{\xi_i = k_i\} = \frac{n!}{k_i!(n-k_i)!}\, p_i^{k_i}(1-p_i)^{n-k_i}, \quad k_i = 0, \ldots, n.$$

## Polihipergeometrikus eloszlás

Egy dobozban $N$ darab színes golyó van; a színek száma $r$, az $i$-edik színből $N_i$ golyó van $(N = N_1 + \cdots + N_r)$. Ha visszatevés nélkül húzunk ki $n$ golyót ($n \leqslant N$), és $\xi_i$ jelöli az $i$-edik színből húzott golyók számát, akkor $\xi = (\xi_1, \xi_2, \ldots, \xi_r)$ olyan $(k_1, k_2, \ldots, k_r)$ értékeket vehet fel, melyekre minden $i = 1, 2, \ldots, r$ esetén teljesül $0 \leqslant k_i \leqslant N_i$ és $k_1 + k_2 + \cdots + k_r = n$, továbbá

$$\mathsf{P}\{\xi_1 = k_1,\, \xi_2 = k_2, \ldots, \xi_r = k_r\} = \frac{\binom{N_1}{k_1}\binom{N_2}{k_2}\cdots\binom{N_r}{k_r}}{\binom{N}{n}}.$$

Ekkor $\xi$ eloszlását $(n, N_1, \ldots, N_r)$ **paraméterű polihipergeometrikus eloszlásnak** nevezzük.
Peremeloszlásai hipergeometrikus eloszlások:

$$\mathsf{P}\{\xi_i = k_i\} = \frac{\binom{N_i}{k_i}\binom{N-N_i}{n-k_i}}{\binom{N}{n}}.$$

## Véletlen vektor sűrűségfüggvénye

Ha létezik olyan $f_\xi : \mathbb{R}^k \to [0, \infty)$ (Borel-mérhető) függvény, melyre

$$F_\xi(x_1, \ldots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f_\xi(t_1, \ldots, t_k) \, \mathrm{d}t_1 \ldots \mathrm{d}t_k$$

teljesül minden $(x_1, \ldots, x_k) \in \mathbb{R}^k$ pontban, akkor az $f_\xi$ függvényt $\xi$ **sűrűségfüggvényének** nevezzük, és azt mondjuk, hogy a $\xi$ véletlen vektor, illetve $\xi$ eloszlása **abszolút folytonos.**

Ha $\xi$ abszolút folytonos eloszlású véletlen vektor, akkor eloszlásfüggvénye folytonos.

## Véletlen vektor sűrűségfüggvénye

Legyen $\xi$ abszolút folytonos véletlen vektor $f_\xi$ sűrűségfüggvénnyel. Tetszőleges $a_i, b_i \in \mathbb{R}$, $a_i < b_i$, $i = 1, \ldots, k$, esetén

$$\mathrm{P}\{a_i \leqslant \xi_i < b_i,\, i = 1, \ldots, k\} = \int_{a_1}^{b_1} \ldots \int_{a_k}^{b_k} f_\xi(t_1, \ldots, t_k)\, \mathrm{d}t_1 \ldots \mathrm{d}t_k.$$

Tetszőleges $B \subset \mathbb{R}^k$ (Borel-halmaz) esetén

$$\mathrm{P}\{(\xi_1, \ldots, \xi_k) \in B\} = \underset{B}{\int \cdots \int} f_\xi(t_1, \ldots, t_k)\, \mathrm{d}t_1 \ldots \mathrm{d}t_k.$$

Ha $F_\xi$ $k$-szor folytonosan differenciálható, akkor

$$f_\xi(x_1, \ldots, x_k) = \frac{\partial^k F_\xi(x_1, \ldots, x_k)}{\partial x_1 \ldots \partial x_k}.$$

Ha a $(\xi, \eta)$ véletlen vektornak létezik $f_{\xi,\eta}$ sűrűségfüggvénye, akkor a $\xi$, illetve $\eta$ véletlen változónak is létezik sűrűségfüggvénye, mégpedig

$$f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)\,\mathrm{d}y, \quad \text{m.m. } x \in \mathbb{R},$$

$$f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)\,\mathrm{d}x, \quad \text{m.m. } y \in \mathbb{R}.$$

Ezek $(\xi, \eta)$ **peremeloszlásai / marginális eloszlásai.**

## Kétdimenziós egyenletes eloszlás

A $(\xi, \eta)$ véletlen vektort egy $T \subseteq \mathbb{R}^2$ (Borel-mérhető) halmazon egyenletes eloszlásúnak nevezünk, ha a sűrűségfüggvénye:

$$f_{(\xi,\eta)}(x, y) := \begin{cases} \frac{1}{|T|} & \text{ha } (x, y) \in T, \\ 0 & \text{egyébként,} \end{cases}$$

ahol $|T|$ a $T$ területét jelöli. Ekkor tetszőleges $B \subseteq \mathbb{R}^2$ (Borel-mérhető) részhalmaz esetén

$$P((\xi, \eta) \in B) = \int_B f_{(\xi,\eta)}(x, y) \, \mathrm{d}x\mathrm{d}y = \frac{|B \cap T|}{|T|}.$$

### Független véletlen változók

A $\xi$ és $\eta$ véletlen változókat akkor nevezzük **függetleneknek**, ha tetszőleges $x, y \in \mathbb{R}$ esetén

$$\mathrm{P}\{\xi < x,\, \eta < y\} = \mathrm{P}\{\xi < x\}\,\mathrm{P}\{\eta < y\},$$

azaz $F_{\xi,\eta}(x, y) = F_\xi(x)F_\eta(y)$.

### Független véletlen változók

Ha $\xi$ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel és $\eta$ diszkrét véletlen változó $y_1, y_2, \ldots$ lehetséges értékekkel, akkor $\xi$ és $\eta$ függetlensége azzal ekvivalens, hogy

$$\forall i, j \qquad \mathrm{P}\{\xi = x_i,\, \eta = y_j\} = \mathrm{P}\{\xi = x_i\}\,\mathrm{P}\{\eta = y_j\}.$$

Ha létezik $(\xi, \eta)$-nak $f_{\xi,\eta}$ sűrűségfüggvénye, akkor $\xi$ és $\eta$ függetlensége azzal ekvivalens, hogy

$$\forall x, y \in \mathbb{R} \qquad f_{\xi,\eta}(x, y) = f_\xi(x)f_\eta(y).$$

## Valószínűségeloszlások konvolúciója

Ha a $\xi$ és $\eta$ véletlen változók függetlenek, akkor azt mondjuk, hogy $\xi+\eta$ eloszlása a $\xi$ és $\eta$ **eloszlásának konvolúciója.**

## Valószínűségeloszlások konvolúciója

Ha $\xi$ és $\eta$ független diszkrét véletlen változók, és a lehetséges értékeik nemnegatív egész számok, akkor a

$$\{\xi+\eta=k\}=\bigcup_{j=0}^{k}(\{\xi=j\}\cap\{\eta=k-j\})$$

diszjunkt felbontás alapján

$$\mathsf{P}\{\xi+\eta=k\}=\sum_{j=0}^{k}\mathsf{P}\{\xi=j\}\,\mathsf{P}\{\eta=k-j\},\qquad k=0,1,\ldots$$

Ha a $\xi$ és $\eta$ független véletlen változóknak léteznek az $f_\xi$ és $f_\eta$ sűrűségfüggvényei, akkor $\xi+\eta$-nak is létezik sűrűségfüggvénye:

$$f_{\xi+\eta}(x)=\int_{-\infty}^{\infty}f_\xi(u)f_\eta(x-u)\,\mathrm{d}u=\int_{-\infty}^{\infty}f_\xi(x-u)f_\eta(u)\,\mathrm{d}u,\qquad x\in\mathbb{R}.$$

**Példák:**

1. Ha $\xi$ és $\eta$ független binomiális eloszlásúak $(n_1, p)$ illetve $(n_2, p)$ paraméterekkel, ahol $n_1, n_2 \in \mathbb{N}$, $p \in (0,1)$, akkor ezek konvolúciója ismét binomiális eloszlás, mégpedig $(n_1 + n_2, p)$ paraméterekkel:
$$\sum_{\substack{i,j\,:\,i+j=k\\ 0\leqslant i\leqslant n_1,\,0\leqslant j\leqslant n_2}} \binom{n_1}{i} p^i (1-p)^{n_1-i} \binom{n_2}{j} p^j (1-p)^{n_2-j} = \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k}.$$

2. Ha $\xi$ és $\eta$ független normális eloszlásúak $(m_1, \sigma_1^2)$ illetve $(m_2, \sigma_2^2)$ paraméterekkel, akkor ezek konvolúciója ismét normális eloszlás, mégpedig $(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ paraméterekkel:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\frac{(u-m_1)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} \mathrm{e}^{-\frac{(x-u-m_2)^2}{2\sigma_2^2}} \,\mathrm{d}u = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} \mathrm{e}^{-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2+\sigma_2^2)}}.$$

**Intuíció:**

Tekintsünk egy $\xi : \Omega \to \mathbb{R}$ diszkrét véletlen változót $x_1, \ldots, x_N$ lehetséges értékekkel.

$n$ független kísérletet hajtunk végre.

$A_k := \{\xi = x_k\}$ relatív gyakorisága

$$r_n(A_k) \approx \mathrm{P}(A_k) = \mathrm{P}\{\xi = x_k\},$$

ezért az $x_k$ értéket körülbelül $n \cdot \mathrm{P}\{\xi = x_k\}$ esetben kapjuk, így **a megfigyelt értékek átlaga körülbelül**

$$\frac{1}{n}\sum_{k=1}^{N} x_k \cdot n \cdot \mathrm{P}\{\xi = x_k\} = \sum_{k=1}^{N} x_k \cdot \mathrm{P}\{\xi = x_k\}.$$

# 5. Várható érték

## Véletlen változó várható értéke

Ha $\xi : \Omega \to \mathbb{R}$ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel, akkor az

$$\mathsf{E}(\xi) := \sum_{k=1}^{\infty} x_k \cdot \mathsf{P}\{\xi = x_k\}$$

mennyiséget a $\xi$ **várható értékének** nevezzük, amennyiben ez a sor abszolút konvergens, azaz $\sum_k |x_k| \cdot \mathsf{P}\{\xi = x_k\} < \infty$.

Ha $\xi : \Omega \to \mathbb{R}$ egy abszolút folytonos véletlen változó, melynek sűrűségfüggvénye $f_\xi : \mathbb{R} \to [0, \infty)$, akkor az

$$\mathsf{E}(\xi) := \int_{-\infty}^{\infty} x\, f_\xi(x)\, \mathrm{d}x$$

mennyiséget a $\xi$ **várható értékének** nevezzük, amennyiben ez az improprius integrál abszolút konvergens, azaz $\int_{-\infty}^{\infty} |x| f_\xi(x)\, \mathrm{d}x < \infty$.

**Példák:**

1. Ha $\xi$ $n$-edrendű és $p$-paraméterű binomiális eloszlású, akkor
$$\mathsf{E}(\xi) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = np.$$
2. Ha egy egységnyi oldalú négyzetben választunk egyenletes eloszlás szerint egy pontot, és $\xi$ jelöli a pontnak a legközelebbi oldaltól való távolságát, akkor $\mathsf{E}(\xi) = \frac{1}{6}$, hiszen

$$F_\xi(x) = \begin{cases} 0 & \text{ha } x \leqslant 0, \\ 1-(1-2x)^2 & \text{ha } x \in [0, \frac{1}{2}], \\ 1 & \text{ha } x > \frac{1}{2}, \end{cases} \qquad f_\xi(x) = \begin{cases} 4-8x & \text{ha } x \in [0, \frac{1}{2}], \\ 0 & \text{egyébként}, \end{cases}$$

$$\mathsf{E}(\xi) = \int_0^{1/2} x(4-8x)\,\mathrm{d}x = \left[2x^2 - \frac{8}{3}x^3\right]_{x=0}^{x=1/2} = \frac{1}{6}.$$

3. *Az $A$ és $B$ játékosok a következő játékot játsszák. Felváltva dobnak egy szabályos érmét; $A$ kezd, és az nyer, akinek először sikerül fejet dobnia. Az első dobásnál 2–2 forintot tesznek be, és minden dobás előtt duplázzák a tétet, azaz ha az $n$-edik dobásra sikerül fejet dobni és $n$ páratlan, akkor $A$ nyer $2^n$ forintot $B$-től, ha pedig $n$ páros, akkor $B$ nyer $2^n$ forintot $A$-tól. Mennyi az $A$ illetve $B$ játékos várható nyereménye?*

Jelölje $\xi$ az $A$ játékos nyereményét (mely pozitív, ha $A$ nyer, és negatív, ha $A$ veszít). Ekkor $\xi$ lehetséges értékei $2, -4, 8, -16,$ $\ldots$ és $\mathsf{P}\{\xi = 2\} = \frac{1}{2}$, $\mathsf{P}\{\xi = -4\} = \frac{1}{4}$, $\ldots$ Mivel

$$
\begin{aligned}
&|2| \cdot \frac{1}{2} + |-4| \cdot \frac{1}{4} + |8| \cdot \frac{1}{8} + |-16| \cdot \frac{1}{16} + \cdots \\
&= 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + 16 \cdot \frac{1}{16} + \cdots = 1 + 1 + 1 + \cdots = \infty,
\end{aligned}
$$

így $\xi$ várható értéke nem létezik!

4. Legyen $\xi$ olyan véletlen változó, melynek sűrűségfüggvénye $f_\xi(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$ (Cauchy-eloszlás). Ekkor nem létezik $\mathsf{E}(\xi)$, mert $\int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} \, \mathrm{d}x = \infty$.

   Valóban,

$$\int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} \, \mathrm{d}x = 2 \int_{0}^{\infty} \frac{x}{\pi(1+x^2)} \, \mathrm{d}x = \frac{1}{\pi} [\ln(1+x^2)]_0^\infty = \infty.$$

## A várható érték tulajdonságai

Ha $\xi$ és $\eta$ (diszkrét vagy abszolút folytonos) véletlen változók (hogy várható értékük létezik és véges) és $a, b \in \mathbb{R}$, akkor

- $\mathsf{E}(a\xi) = a\,\mathsf{E}(\xi)$ (homogenitás)
- $\mathsf{E}(\xi + \eta) = \mathsf{E}(\xi) + \mathsf{E}(\eta)$ (additivitás)
- $\mathsf{E}(a\xi + b\eta) = a\,\mathsf{E}\,\xi + b\,\mathsf{E}\,\eta$ (linearitás)
- ha $\xi$ és $\eta$ függetlenek, akkor $\mathsf{E}(\xi\eta) = \mathsf{E}(\xi) \cdot \mathsf{E}(\eta)$
- ha $\xi \leqslant \eta$, akkor $\mathsf{E}(\xi) \leqslant \mathsf{E}(\eta)$ (monotonitás)
- ha $\xi \geqslant 0$, akkor $\mathsf{E}(\xi) \geqslant 0$ (nemnegativitás)
- ha $\xi \geqslant 0$ és $\mathsf{E}(\xi) = 0$, akkor $\mathsf{P}(\xi = 0) = 1$
- $|\,\mathsf{E}(\xi)| \leqslant \mathsf{E}(|\xi|)$
- $\mathsf{E}(|\xi\eta|) \leqslant \sqrt{\mathsf{E}(\xi^2)\,\mathsf{E}(\eta^2)}$ (Cauchy-Schwarz egyenlőtlenség)

## Véletlen változó függvényének várható értéke

Legyen $g : \mathbb{R} \to \mathbb{R}$ (Borel-mérhető) függvény.

Ha $\xi$ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel, akkor

$$\mathsf{E}(g(\xi)) = \sum_{k=1}^{\infty} g(x_k) \cdot \mathsf{P}\{\xi = x_k\},$$

amennyiben ez a sor abszolút konvergens.

Ha $\xi$ abszolút folytonos véletlen változó $f_\xi$ sűrűségfüggvénnyel, akkor

$$\mathsf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) \cdot f_\xi(x)\,\mathrm{d}x,$$

amennyiben ez az improprius integrál abszolút konvergens.

## Véletlen vektor függvényének várható értéke

Legyen $g : \mathbb{R}^2 \to \mathbb{R}$ (Borel-mérhető) függvény.

Ha $\xi$ és $\eta$ diszkrét véletlen változók $x_1, x_2, \ldots,$ illetve $y_1, y_2, \ldots$ lehetséges értékekkel, akkor

$$\mathsf{E}(g(\xi,\eta)) = \sum_{k,\ell=1}^{\infty} g(x_k, y_\ell) \cdot \mathsf{P}\{\xi = x_k, \eta = y_\ell\},$$

amennyiben ez a sor abszolút konvergens.

Ha $(\xi,\eta)$ abszolút folytonos véletlen vektor $f_{\xi,\eta}$ sűrűségfüggvénnyel, akkor

$$\mathsf{E}(g(\xi,\eta)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) \cdot f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

amennyiben ez az improprius integrál abszolút konvergens.

# 5. Várható érték

## Véletlen változó varianciája / szórásnégyzete

Legyen $\xi$ egy véletlen változó, hogy $\mathsf{E}(\xi)$ létezik és véges. Ekkor

$$\mathrm{var}(\xi) := \mathsf{D}^2(\xi) := \mathsf{E}\left[(\xi - \mathsf{E}(\xi))^2\right].$$

## Variancia / szórásnégyzet kiszámolása

$$\begin{aligned}\mathrm{var}(\xi) &= \mathsf{E}\left[\xi^2 - 2\xi\,\mathsf{E}(\xi) + (\mathsf{E}(\xi))^2\right] = \mathsf{E}(\xi^2) - 2\cdot\mathsf{E}(\xi)\cdot\mathsf{E}(\xi) + [\mathsf{E}(\xi)]^2 \\ &= \mathsf{E}(\xi^2) - [\mathsf{E}(\xi)]^2,\end{aligned}$$

így ha $\xi$ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel, akkor

$$\mathrm{var}(\xi) = \sum_{k=1}^{\infty} x_k^2 \cdot \mathsf{P}\{\xi = k\} - \left(\sum_{k=1}^{\infty} x_k \cdot \mathsf{P}\{\xi = k\}\right)^2,$$

ha pedig $\xi$ abszolút folytonos véletlen változó $f_\xi$ sűrűségfüggvény-nyel, akkor

$$\mathrm{var}(\xi) = \int_{-\infty}^{\infty} x^2 f_\xi(x)\,\mathrm{d}x - \left(\int_{-\infty}^{\infty} x f_\xi(x)\,\mathrm{d}x\right)^2.$$

# 5. Várható érték

## Variancia / szórásnégyzet tulajdonságai

Ha $\xi$ és $\eta$ véletlen változók és $a, c \in \mathbb{R}$, akkor

- $\mathrm{var}(\xi) \geqslant 0$
- ha $\mathrm{var}(\xi) = 0$, akkor $\mathrm{P}(\xi = \mathrm{E}\,\xi) = 1$
- $\mathrm{var}(\xi + a) = \mathrm{var}(\xi)$ (eltolásinvariancia)
- $\mathrm{var}(c \cdot \xi) = c^2 \cdot \mathrm{var}(\xi)$
- ha $\xi$ és $\eta$ függetlenek, akkor $\mathrm{var}(\xi + \eta) = \mathrm{var}(\xi) + \mathrm{var}(\eta)$ (additivitás)

## Véletlen változó szórása

Legyen $\xi$ egy véletlen változó, hogy $\mathrm{E}(\xi)$ létezik és véges. Ekkor

$$\mathrm{D}(\xi) := \sqrt{\mathrm{D}^2(\xi)} = \sqrt{\mathrm{E}\left[(\xi - \mathrm{E}(\xi))^2\right]} = \sqrt{\mathrm{E}(\xi^2) - (\mathrm{E}(\xi))^2}.$$

## $p$ paraméterű Bernoulli-eloszlás

Tekintsünk egy kísérletet, ebben egy $A$ eseményt, melyre $p := \mathsf{P}(A)$. Ekkor

$$\xi := k_1(A) = \begin{cases} 1 & \text{ha } A \text{ bekövetkezik,} \\ 0 & \text{ha } A \text{ nem következik be} \end{cases}$$

diszkrét véletlen változó; lehetséges értékei: 0 és 1, eloszlása

$$\mathsf{P}\{\xi = 1\} = p, \qquad \mathsf{P}\{\xi = 0\} = 1 - p.$$

Ezt az eloszlást $p$ **paraméterű Bernoulli-eloszlásnak** nevezzük.

Nyilván

$$\begin{aligned} \mathsf{E}(\xi) &= 1 \cdot p + 0 \cdot (1 - p) = p, \\ \mathsf{E}(\xi^2) &= 1^2 \cdot p + 0^2 \cdot (1 - p) = p, \\ \operatorname{var} \xi &= \mathsf{E}(\xi^2) - [\mathsf{E}(\xi)]^2 = p - p^2 = p(1 - p). \end{aligned}$$

## $n$-edrendű, $p$-paraméterű binomiális eloszlás

$A$ esemény, $p := \mathsf{P}(A)$; $n$ független kísérlet; $A$ gyakorisága: $\xi := k_n(A)$ diszkrét véletlen változó; lehetséges értékeinek halmaza: $\{0, 1, 2, \ldots, n\}$, eloszlása

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

Ha

$$\xi_i := \begin{cases} 1 & \text{ha } A \text{ beköv. az } i\text{-edik alkalommal,} \\ 0 & \text{egyébként,} \end{cases} \qquad i = 1, \ldots, n,$$

akkor $\xi = \xi_1 + \cdots + \xi_n$, és $\xi_1, \ldots, \xi_n$ független, $p$ paraméterű Bernoulli-eloszlásúak. Ezért

$$\mathsf{E}(\xi) = \mathsf{E}(\xi_1) + \cdots + \mathsf{E}(\xi_n) = np,$$
$$\operatorname{var}(\xi) = \operatorname{var}(\xi_1) + \cdots + \operatorname{var}(\xi_n) = np(1-p).$$

Legyen $\eta$ **standard normális eloszlású**, azaz normális eloszlású $(0, 1)$ paraméterekkel. Ekkor

$$\mathsf{E}(\eta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot \mathrm{e}^{-x^2/2} \, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \lim_{\substack{K \to -\infty \\ L \to +\infty}} \left[ -\mathrm{e}^{-x^2/2} \right]_{x=K}^{x=L} = 0,$$

$$\begin{aligned} \mathsf{E}(\eta^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \mathrm{e}^{-x^2/2} \, \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \left[ -x\mathrm{e}^{-x^2/2} \right]_{x=-\infty}^{x=\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \, \mathrm{d}x = 1, \end{aligned}$$

$$\operatorname{var}(\eta) = \mathsf{E}(\eta^2) - \big[ \mathsf{E}(\eta) \big]^2 = 1.$$

Ha pedig $\xi$ normális eloszlású $(m, \sigma^2)$ paraméterekkel, akkor

$$\xi = \sigma \cdot \eta + m,$$

ahol

$$\eta = \frac{\xi - m}{\sigma} \qquad \text{(ez } \xi \text{ standardizáltja)}$$

standard normális eloszlású, hiszen $\eta$ eloszlásfüggvénye

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \mathsf{P}\left\{\frac{\xi - m}{\sigma} < x\right\} = \mathsf{P}\{\xi < \sigma \cdot x + m\} = F_\xi(\sigma \cdot x + m),$$

ahol $x \in \mathbb{R}$, így $\eta$ sűrűségfüggvénye

$$f_\eta(x) = F_\eta'(x) = \sigma \cdot F_\xi'(\sigma x + m) = \sigma \cdot f_\xi(\sigma x + m) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2},$$

ahol $x \in \mathbb{R}$, ezért

$$\mathsf{E}(\xi) = \sigma \cdot \mathsf{E}(\eta) + m = m, \qquad \operatorname{var}(\xi) = \sigma^2 \cdot \operatorname{var}(\eta) = \sigma^2.$$

## Véletlen változók kovarianciája

Ha $\xi$ és $\eta$ véletlen változók, hogy $\mathsf{E}(\xi^2) < \infty$ és $\mathsf{E}(\eta^2) < \infty$, akkor

$$\operatorname{cov}(\xi, \eta) := \mathsf{E}\big[(\xi - \mathsf{E}\,\xi)(\eta - \mathsf{E}\,\eta)\big].$$

## Kovariancia kiszámítása és addíciós képlet

Ha $\xi$ és $\eta$ véletlen változók, hogy $\mathsf{E}(\xi^2) < \infty$ és $\mathsf{E}(\eta^2) < \infty$, akkor

$$\begin{aligned}\operatorname{cov}(\xi, \eta) &= \mathsf{E}(\xi\eta) - \mathsf{E}(\xi)\,\mathsf{E}(\eta),\\ \operatorname{var}(\xi \pm \eta) &= \operatorname{var}(\xi) \pm 2\operatorname{cov}(\xi, \eta) + \operatorname{var}(\eta).\end{aligned}$$

Valóban,

$$\begin{aligned}\operatorname{cov}(\xi, \eta) &= \mathsf{E}(\xi\eta - \xi\,\mathsf{E}\,\eta - \eta\,\mathsf{E}\,\xi + \mathsf{E}\,\xi\,\mathsf{E}\,\eta)\\ &= \mathsf{E}(\xi\eta) - \mathsf{E}(\xi)\,\mathsf{E}(\eta) - \mathsf{E}(\eta)\,\mathsf{E}(\xi) + \mathsf{E}(\xi)\,\mathsf{E}(\eta)\\ &= \mathsf{E}(\xi\eta) - \mathsf{E}(\xi)\,\mathsf{E}(\eta).\end{aligned}$$

## Véletlen változók korrelációs együtthatója

Feltéve, hogy $0 < \mathrm{var}(\xi) < \infty$ és $0 < \mathrm{var}(\eta) < \infty$,

$$\mathrm{corr}(\xi, \eta) := \frac{\mathrm{cov}(\xi, \eta)}{\sqrt{\mathrm{var}(\xi) \cdot \mathrm{var}(\eta)}}.$$

Ha $\mathrm{corr}(\xi, \eta) = 0$, azaz $\mathrm{cov}(\xi, \eta) = 0$, akkor azt mondjuk, hogy $\xi$ **és** $\eta$ **korrelálatlanok.**

Ha

$$\mathrm{corr}(\xi, \eta) > 0, \quad \text{azaz} \quad \mathrm{cov}(\xi, \eta) > 0,$$

akkor $\xi$ **és** $\eta$ **pozitívan korreláltak**, ha pedig

$$\mathrm{corr}(\xi, \eta) < 0, \quad \text{azaz} \quad \mathrm{cov}(\xi, \eta) < 0,$$

akkor $\xi$ **és** $\eta$ **negatívan korreláltak.**

Ha $\xi$ és $\eta$ függetlenek, akkor $\xi$ és $\eta$ korrelálatlanok, de ez fordítva általában nem igaz (lásd a következő példát).

**Példa:** Legyen a $(\xi, \eta)$ véletlen vektor egyenletes eloszlású a $(-1, 0)$, $(0, -1)$, $(0, 1)$ és $(1, 0)$ pontokon, azaz

$$\mathsf{P}\{\xi=-1, \eta=0\}=\mathsf{P}\{\xi=0, \eta=-1\}=\mathsf{P}\{\xi=0, \eta=1\}=\mathsf{P}\{\xi=1, \eta=0\}=\frac{1}{4}.$$

Ekkor $\mathsf{E}(\xi) = \mathsf{E}(\eta) = 0$ és $\mathsf{E}(\xi\eta) = 0$ miatt

$$\operatorname{cov}(\xi, \eta) = \mathsf{E}(\xi\eta) - \mathsf{E}(\xi) \cdot \mathsf{E}(\eta) = 0,$$

azaz $\xi$ és $\eta$ korrelálatlanok, viszont a peremeloszlások

$$\mathsf{P}\{\xi = -1\} = \mathsf{P}\{\xi = 1\} = \frac{1}{4}, \quad \mathsf{P}\{\xi = 0\} = \frac{1}{2},$$
$$\mathsf{P}\{\eta = -1\} = \mathsf{P}\{\eta = 1\} = \frac{1}{4}, \quad \mathsf{P}\{\eta = 0\} = \frac{1}{2},$$

ezért $\xi$ és $\eta$ nem függetlenek, hiszen például

$$\mathsf{P}\{\xi = 1, \, \eta = 0\} = \frac{1}{4}, \qquad \mathsf{P}\{\xi = 1\} \cdot \mathsf{P}\{\eta = 0\} = \frac{1}{8}.$$

## Kovariancia és korrelációs együttható tulajdonságai

- $\operatorname{var}(\xi) = \operatorname{cov}(\xi, \xi)$
- Ha $\xi$ és $\eta$ véletlen változók, akkor

$$\operatorname{cov}(\xi, \eta) = \operatorname{cov}(\eta, \xi), \qquad \operatorname{corr}(\xi, \eta) = \operatorname{corr}(\eta, \xi) \qquad \text{(szimmetria)}$$

- Ha $\xi_1, \dots, \xi_n$ és $\eta_1, \dots, \eta_m$ véletlen változók és $a_1, \dots, a_n, b_1, \dots, b_m \in \mathbb{R}$, akkor

$$\operatorname{cov}\left(\sum_{i=1}^{n} a_i \xi_i, \sum_{j=1}^{m} b_j \eta_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \operatorname{cov}(\xi_i, \eta_j) \qquad \text{(bilinearitás)}$$

- $|\operatorname{corr}(\xi, \eta)| \leqslant 1$, és $|\operatorname{corr}(\xi, \eta)| = 1$ akkor és csak akkor, ha valamely $a \neq 0$ és $b$ valós számokkal $\mathsf{P}\{\eta = a \cdot \xi + b\} = 1$ teljesül; itt $a > 0$ illetve $a < 0$ aszerint, hogy $\operatorname{corr}(\xi, \eta) = 1$ illetve $\operatorname{corr}(\xi, \eta) = -1$.

## Momentumok, ferdeség, csúcsosság / lapultság

Legyen $\xi$ véletlen változó, $k$ pozitív egész. Ekkor

- ***k*-adik momentum:** $\mathrm{E}(\xi^k)$
- ***k*-adik centrális momentum:** $\mathrm{E}\left[(\xi - \mathrm{E}\,\xi)^k\right]$
- ***k*-adik abszolút momentum:** $\mathrm{E}\left(|\xi|^k\right)$
- ***k*-adik abszolút centrális momentum:** $\mathrm{E}\left[|\xi - \mathrm{E}\,\xi|^k\right]$
- **ferdeség:** $\dfrac{\mathrm{E}\left[(\xi - \mathrm{E}\,\xi)^3\right]}{(\mathrm{E}\,[(\xi - \mathrm{E}\,\xi)^2])^{3/2}} = \dfrac{\mathrm{E}\left[(\xi - \mathrm{E}\,\xi)^3\right]}{(\mathrm{var}(\xi))^{3/2}}$
- **csúcsosság:** $\dfrac{\mathrm{E}\left[(\xi - \mathrm{E}\,\xi)^4\right]}{(\mathrm{E}\,[(\xi - \mathrm{E}\,\xi)^2])^2} - 3 = \dfrac{\mathrm{E}\left[(\xi - \mathrm{E}\,\xi)^4\right]}{(\mathrm{var}(\xi))^2} - 3$

Tehát $\mathrm{E}(\xi)$ az első momentum, $\mathrm{var}(\xi)$ pedig a második (abszolút) centrális momentum.

Ha $a, b \in \mathbb{R}$ és $a \neq 0$, akkor $\xi$ és $a\xi + b$ ferdesége, illetve $\xi$ és $a\xi + b$ csúcsossága megegyezik.

**Példa:** Legyen $\eta$ standard normális eloszlású. Ekkor $\mathsf{E}(\eta) = 0$, $\operatorname{var}(\eta) = 1$,

$$\mathsf{E}(\eta^3) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 \mathrm{e}^{-x^2/2} \, \mathrm{d}x = 0,$$

$$\begin{aligned}
\mathsf{E}(\eta^4) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 \mathrm{e}^{-x^2/2} \, \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \Big[ -x^3 \mathrm{e}^{-x^2/2} \Big]_{x=-\infty}^{x=\infty} + \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \mathrm{e}^{-x^2/2} \, \mathrm{d}x = 3\,\mathsf{E}(\eta^2) = 3,
\end{aligned}$$

ezért $\eta$ ferdesége és csúcsossága is 0.

Ha $\xi$ normális eloszlású $(m, \sigma^2)$ paraméterekkel, akkor $\xi = \sigma\eta + m$, ahol $\eta$ standard normális eloszlású, ezért $\xi$ ferdesége és csúcsossága is 0.

# Várható érték

## Véletlen mátrix

$$\xi = (\xi_{i,j})_{1\leqslant i\leqslant k,\,1\leqslant j\leqslant \ell} : \Omega \to \mathbb{R}^{k\times\ell},$$

ahol $\xi_{i,j} : \Omega \to \mathbb{R}$, $i \in \{1,\ldots,k\}$, $j \in \{1,\ldots,\ell\}$ véletlen változók.

## Véletlen mátrix várható érték mátrixa

Legyen $\xi = (\xi_{i,j})_{1\leqslant i\leqslant k,\,1\leqslant j\leqslant \ell} : \Omega \to \mathbb{R}^{k\times\ell}$ véletlen mátrix.
Ha tetszőleges $i \in \{1,\ldots,k\}$ és $j \in \{1,\ldots,\ell\}$ esetén $\mathsf{E}(|\xi_{i,j}|) < \infty$, akkor $\xi$ **várható érték mátrixa**

$$\mathsf{E}(\xi) := \big(\mathsf{E}(\xi_{i,j})\big)_{1\leqslant i\leqslant k,\,1\leqslant j\leqslant \ell} \in \mathbb{R}^{k\times\ell}.$$

## Várható érték mátrix tulajdonságai

Legyen $\xi = (\xi_{i,j})_{1\leqslant i\leqslant k,\,1\leqslant j\leqslant \ell} : \Omega \to \mathbb{R}^{k\times\ell}$ véletlen mátrix, melyre tetszőleges $i \in \{1,\ldots,k\}$ és $j \in \{1,\ldots,\ell\}$ esetén $\mathsf{E}(|\xi_{i,j}|) < \infty$.
Ha $A \in \mathbb{R}^{r\times k}$ és $B \in \mathbb{R}^{\ell\times s}$, akkor $\mathsf{E}(A\xi B) = A\,\mathsf{E}(\xi)B$.

# Várható érték

## Véletlen vektor kovarianciamátrixa (szórásmátrixa)

Legyen $\xi = (\xi_1, \dots, \xi_d) : \Omega \to \mathbb{R}^d$ véletlen vektor. Ha $\mathsf{E}(\|\xi\|^2) < \infty$, azaz $\mathsf{E}(\xi_1^2) < \infty, \dots, \mathsf{E}(\xi_d^2) < \infty$, akkor $\xi$ **kovarianciamátrixa**

$$\operatorname{cov}(\xi, \xi) := \mathsf{E}\left[(\xi - \mathsf{E}(\xi))(\xi - \mathsf{E}(\xi))^\top\right] \in \mathbb{R}^{d \times d},$$

melynek elemei $\operatorname{cov}(\xi_i, \xi_j) := \mathsf{E}\left[(\xi_i - \mathsf{E}(\xi_i))(\xi_j - \mathsf{E}(\xi_j))\right]$, $1 \leqslant i, j \leqslant d$.

## Kovarianciamátrix tulajdonságai

Legyen $\xi = (\xi_1, \dots, \xi_d) : \Omega \to \mathbb{R}^d$ véletlen vektor, $\mathsf{E}(\|\xi\|^2) < \infty$.

- $\operatorname{cov}(\xi, \xi)$ szimmetrikus: $\operatorname{cov}(\xi, \xi)^\top = \operatorname{cov}(\xi, \xi)$.
- $\operatorname{cov}(\xi, \xi)$ pozitív szemidefinit, azaz $\forall\, x \in \mathbb{R}^d$ esetén
$$x^\top \operatorname{cov}(\xi, \xi) x = \langle \operatorname{cov}(\xi, \xi) x, x \rangle = \sum_{i=1}^d \sum_{j=1}^d \operatorname{cov}(\xi_i, \xi_j) x_i x_j \geqslant 0.$$
- Ha $A \in \mathbb{R}^{r \times d}$ és $b \in \mathbb{R}^r$, akkor
$\operatorname{cov}(A\xi + b, A\xi + b) = A \operatorname{cov}(\xi, \xi) A^\top$.

## Diszkrét véletlen változó feltételes eloszlása, feltételes várható értéke, feltételes varianciája

Legyen $A$ egy pozitív valószínűségű esemény. Ha $\xi$ diszkrét véletlen változó $\mathsf{P}\{\xi = x_k\}$ $(k = 1, 2, \ldots)$ eloszlással, akkor $\xi$-nek az $A$-ra vonatkozó **feltételes eloszlása**

$$\mathsf{P}\{\xi = x_k \,|\, A\}, \qquad k = 1, 2, \ldots$$

**feltételes várható értéke**

$$\mathsf{E}(\xi \mid A) := \sum_k x_k \cdot \mathsf{P}\{\xi = x_k \,|\, A\}$$

amennyiben ez a sor abszolút konvergens, azaz $\sum_k |x_k| \cdot \mathsf{P}\{\xi = x_k \,|\, A\} < \infty$, **feltételes varianciája** pedig

$$\mathrm{var}(\xi \mid A) := \mathsf{E}\left[(\xi - \mathsf{E}(\xi \mid A))^2 \mid A\right] = \mathsf{E}(\xi^2 \mid A) - \left[\mathsf{E}(\xi \mid A)\right]^2$$

$$= \sum_k x_k^2 \cdot \mathsf{P}\{\xi = x_k \,|\, A\} - \left(\sum_k x_k \cdot \mathsf{P}\{\xi = x_k \,|\, A\}\right)^2$$

amennyiben a $\sum_k x_k^2 \, \mathsf{P}(\xi = x_k \,|\, A)$ sor konvergens.

# 6. Feltételes várható érték

- Ha $\xi$ diszkrét véletlen változó, akkor a $\mathsf{P}\{\xi = x_k \mid A\}$, $k = 1, 2, \ldots$ számok eloszlást alkotnak, hiszen nemnegatívak, és összegük 1:

$$\sum_k \mathsf{P}\{\xi = x_k \mid A\} = \frac{1}{\mathsf{P}(A)} \sum_k \mathsf{P}(\{\xi = x_k\} \cap A)$$

$$= \frac{1}{\mathsf{P}(A)} \mathsf{P}\left(\bigcup_k (\{\xi = x_k\} \cap A)\right) = \frac{1}{\mathsf{P}(A)} \mathsf{P}\left(\left(\bigcup_k \{\xi = x_k\}\right) \cap A\right)$$

$$= \frac{1}{\mathsf{P}(A)} \mathsf{P}(\Omega \cap A) = 1.$$

- Ha $\mathsf{E}(|\xi|) < \infty$, akkor tetszőleges pozitív valószínűségű $A$ esemény esetén $\mathsf{E}(|\xi| \mid A) < \infty$.
- Speciálisan, ha $A$ egy olyan esemény, melyre $\mathsf{P}(A) = 1$ (pl.: $A = \Omega$ esetén), akkor $\xi$-nek az $A$-ra vonatkozó feltételes eloszlása, várható értéke, ill. varianciája megegyezik $\xi$ eloszlásával, várható értékével, ill. varianciájával.

# 6. Feltételes várható érték

*Mi a két szabályos dobókockával dobott számok eltérésének feltételes eloszlása azon feltétel mellett, hogy a dobott számok összege $\ell$ ?*

Jelölje a dobott számokat $\xi$ és $\eta$. Nyilván $\ell \in \{2, 3, \ldots, 12\}$ és

$$\mathsf{P}(\xi + \eta = \ell) = \begin{cases} \frac{\ell-1}{36} & \text{ha } 2 \leqslant \ell \leqslant 7, \\ \frac{13-\ell}{36} & \text{ha } 7 \leqslant \ell \leqslant 12. \end{cases}$$

Továbbá $|\xi - \eta|$ lehetséges értékei 0,1,2,3,4,5, és $\ell \in \{2, \ldots, 6\}$ esetén a szóbanforgó feltételes eloszlás valószínűségei:

$$\mathsf{P}(|\xi-\eta| = 0 \mid \xi+\eta = 2) = 1, \qquad \mathsf{P}(|\xi-\eta| = 1 \mid \xi+\eta = 3) = 1,$$

$$\mathsf{P}(|\xi-\eta| = 0 \mid \xi+\eta = 4) = \frac{1}{3}, \qquad \mathsf{P}(|\xi-\eta| = 2 \mid \xi+\eta = 4) = \frac{2}{3},$$

$$\mathsf{P}(|\xi-\eta| = 1 \mid \xi+\eta = 5) = \frac{1}{2}, \qquad \mathsf{P}(|\xi-\eta| = 3 \mid \xi+\eta = 5) = \frac{1}{2},$$

$$\mathsf{P}(|\xi-\eta| = 0 \mid \xi+\eta = 6) = \frac{1}{5}, \qquad \mathsf{P}(|\xi-\eta| = 2 \mid \xi+\eta = 6) = \frac{2}{5},$$

$$\mathsf{P}(|\xi-\eta| = 4 \mid \xi+\eta = 6) = \frac{2}{5}.$$

## Abszolút folytonos véletlen változó feltételes eloszlása, feltételes várható értéke

Egy tetszőleges $\xi$ véletlen változónak az $A$ pozitív valószínűségű eseményre vonatkozó **feltételes eloszlásfüggvénye**

$$F_{\xi|A}(x) := \mathsf{P}\{\xi < x \mid A\}, \qquad x \in \mathbb{R}.$$

Ha létezik olyan nemnegatív (Borel-mérhető) $f_{\xi|A}$ függvény, melyre

$$F_{\xi|A}(x) = \int_{-\infty}^{x} f_{\xi|A}(u)\,\mathrm{d}u$$

teljesül tetszőleges $x \in \mathbb{R}$ esetén, akkor az $f_{\xi|A}$ függvényt $\xi$-nek az $A$-ra vonatkozó **feltételes sűrűségfüggvényének** nevezzük. Ekkor $\xi$-nek az $A$-ra vonatkozó **feltételes várható értéke**

$$\mathsf{E}(\xi \mid A) := \int_{-\infty}^{\infty} x \cdot f_{\xi|A}(x)\,\mathrm{d}x$$

amennyiben ez az improprius integrál abszolút konvergens, azaz $\int_{-\infty}^{\infty} |x| \cdot f_{\xi|A}(x)\,\mathrm{d}x < \infty$.

Ha létezik az $f_{\xi|A}(\cdot)$ feltételes sűrűségfüggvény, akkor $f_{\xi|A}(\cdot)$ sűrűségfüggvény, hiszen nemnegatív, Borel-mérhető és $\int_{-\infty}^{\infty} f_{\xi|A}(u)\,\mathrm{d}u = 1$.

## Abszolút folytonos véletlen változó feltételes varianciája

$\xi$-nek az $A$-ra vonatkozó **feltételes varianciája:**

$$\mathrm{var}(\xi \mid A) := \mathrm{E}\left[(\xi - \mathrm{E}(\xi \mid A))^2 \mid A\right] = \mathrm{E}(\xi^2 \mid A) - \left[\mathrm{E}(\xi \mid A)\right]^2$$
$$= \int_{-\infty}^{\infty} x^2 \cdot f_{\xi|A}(x)\,\mathrm{d}x - \left(\int_{-\infty}^{\infty} x \cdot f_{\xi|A}(x)\,\mathrm{d}x\right)^2$$

amennyiben $\int_{-\infty}^{\infty} x^2 \cdot f_{\xi|A}(x)\,\mathrm{d}x < \infty$.

# 6. Feltételes várható érték

**Példa:** Legyen $\xi$ standard normális eloszlású, és $A := \{\xi \geqslant 0\}$. Ekkor $\mathrm{P}(A) = 1/2$, és

$$F_{\xi|A}(x) = \frac{\mathrm{P}(0 \leqslant \xi < x)}{\mathrm{P}(\xi \geqslant 0)} = \begin{cases} 0 & \text{ha } x \leqslant 0, \\ 2\,\mathrm{P}(0 \leqslant \xi < x) & \text{ha } x > 0. \end{cases}$$

Ha $x > 0$, akkor tehát

$$F_{\xi|A}(x) = 2(\Phi(x) - \Phi(0)) = 2\Phi(x) - 1 = \sqrt{\frac{2}{\pi}} \int_0^x \mathrm{e}^{-u^2/2}\,\mathrm{d}u.$$

Így $\xi$-nek az $A$-ra vonatkozó feltételes sűrűségfüggvénye:

$$f_{\xi|A}(x) = \begin{cases} \sqrt{\frac{2}{\pi}}\mathrm{e}^{-x^2/2} & \text{ha } x > 0, \\ 0 & \text{ha } x \leqslant 0. \end{cases}$$

Továbbá, ha $\eta := |\xi|$, akkor

$$F_\eta(x) = \mathrm{P}(|\xi| < x) = \begin{cases} 0 & \text{ha } x \leqslant 0, \\ \mathrm{P}(-x < \xi < x) & \text{ha } x > 0, \end{cases}$$

$$= \begin{cases} 0 & \text{ha } x \leqslant 0, \\ 2\,\mathrm{P}(0 \leqslant \xi < x) & \text{ha } x > 0, \end{cases} = F_{\xi|A}(x), \qquad x \in \mathbb{R},$$

azaz $\xi$-nek az $A$-ra vett feltételes eloszlása megegyezik $|\xi|$ eloszlásával.

# 6. Feltételes várható érték

## Abszolút folytonos véletlen változóra vonatkozó feltételes sűrűségfüggvény

Legyen $(\xi, \eta)$ abszolút folytonos véletlen változó $f_{\xi,\eta}$ sűrűségfüggvénnyel. Ekkor $\xi$ **feltételes sűrűségfüggvénye az $\eta = y$ feltételre nézve:**

$$f_{\xi|\eta}(x|y) := \begin{cases} \dfrac{f_{\xi,\eta}(x,y)}{f_\eta(y)} & \text{ha } f_\eta(y) \neq 0, \\ 0 & \text{ha } f_\eta(y) = 0, \end{cases} \qquad x \in \mathbb{R},$$

ahol $f_\eta$ az $\eta$ sűrűségfüggvénye.

$\xi$ **feltételes eloszlásfüggvénye az $\eta = y$ feltételre nézve:**

$$F_{\xi|\eta}(x|y) := \int_{-\infty}^{x} f_{\xi|\eta}(u|y)\,\mathrm{d}u, \qquad x \in \mathbb{R},$$

## Abszolút folytonos véletlen változóra vonatkozó feltételes várható érték, feltételes variancia

$\xi$ **feltételes várható értéke az** $\eta = y$ **feltételre nézve:**

$$\mathsf{E}(\xi \mid \eta = y) := \int_{-\infty}^{\infty} x \cdot f_{\xi|\eta}(x|y)\,\mathrm{d}x,$$

feltéve, hogy a jobb oldalon álló integrál abszolút konvergens,

$\xi$ **feltételes varianciája az** $\eta = y$ **feltételre nézve:**

$$\begin{aligned}\mathrm{var}(\xi \mid \eta = y) &:= \mathsf{E}\left[(\xi - \mathsf{E}(\xi \mid \eta = y))^2 \mid \eta = y\right] \\ &= \mathsf{E}(\xi^2 \mid \eta = y) - \left[\mathsf{E}(\xi \mid \eta = y)\right]^2 \\ &= \int_{-\infty}^{\infty} x^2 \cdot f_{\xi|\eta}(x|y)\,\mathrm{d}x - \left(\int_{-\infty}^{\infty} x \cdot f_{\xi|\eta}(x|y)\,\mathrm{d}x\right)^2\end{aligned}$$

feltéve, hogy $\int_{-\infty}^{\infty} x^2 \cdot f_{\xi|\eta}(x|y)\,\mathrm{d}x < \infty$.

## Teljes eseményrendszerre vonatkozó teljes várható érték tétel

Ha az $A_1, A_2, \ldots$ pozitív valószínűségű események teljes eseményrendszert alkotnak, $\xi$ véletlen változó és $\mathrm{E}(|\xi|) < \infty$, akkor

$$\mathrm{E}(\xi) = \sum_k \mathrm{E}(\xi \mid A_k) \cdot \mathrm{P}(A_k).$$

**Bizonyítás.** Legyen $\xi$ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel. Ekkor

$$\sum_k \mathrm{E}(\xi \mid A_k) \cdot \mathrm{P}(A_k) = \sum_k \sum_j x_j \, \mathrm{P}\{\xi = x_j \mid A_k\} \cdot \mathrm{P}(A_k)$$
$$= \sum_k \sum_j x_j \, \mathrm{P}(\{\xi = x_j\} \cap A_k) = \sum_j x_j \sum_k \mathrm{P}(\{\xi = x_j\} \cap A_k)$$
$$= \sum_j x_j \, \mathrm{P}\{\xi = x_j\} = \mathrm{E}(\xi).$$

Abszolút folytonos $\xi$ véletlen változó esetén hasonlóan.

**Példa:**

Szabályos dobókockával addig dobunk, míg az első 6-os megjelenik.
Legyen $\xi :=$ az ehhez szükséges dobások száma.
Ekkor $\xi$ geometriai eloszlású $\frac{1}{6}$-paraméterrel, így $\mathsf{E}(\xi) = 6$.

Az alábbiakban a teljes várható érték tétele alapján is meghatározzuk $\mathsf{E}(\xi)$-t.
Legyen $A_k := \{\text{az első dobás } k\}$, $k = 1, \ldots, 6$, pozitív (1/6) valószínűségű eseményekből álló teljes eseményrendszer.
A teljes várható érték tétele alapján:

$$\mathsf{E}(\xi) = \sum_{k=1}^{6} \mathsf{E}(\xi \mid A_k) \cdot \mathsf{P}(A_k)$$

Megmutatjuk, hogy

$$\mathsf{E}(\xi \mid A_k) = \begin{cases} 1 + \mathsf{E}(\xi) & \text{ha } 1 \leqslant k \leqslant 5, \\ 1 & \text{ha } k = 6. \end{cases}$$

Nyilván, $\mathrm{P}(\xi = 1 \mid A_6) = 1$ és így $\mathrm{E}(\xi \mid A_6) = 1 \cdot 1 = 1$.
Ha $1 \leqslant k \leqslant 5$, akkor

$$\mathrm{E}(\xi \mid A_k) = \mathrm{E}(1 + \xi - 1 \mid A_k) = 1 + \mathrm{E}(\xi - 1 \mid A_k) = 1 + \mathrm{E}(\xi),$$

ahol az utolsó lépés abból következik, hogy $\xi - 1$-nek az $A_k$-ra ($k = 1, 2, 3, 4, 5$) vonatkozó feltételes eloszlása megegyezik $\xi$ eloszlásával, ugyanis

$$\mathrm{P}(\xi - 1 = n \mid A_k) = \frac{\mathrm{P}(\xi - 1 = n, A_k)}{\mathrm{P}(A_k)} = \frac{\frac{1}{6}\,\mathrm{P}(\xi = n)}{\frac{1}{6}} = \mathrm{P}(\xi = n), \quad n \in \mathbb{N}.$$

Ezért

$$\mathrm{E}(\xi) = \frac{1}{6}\big(1 + 5(1 + \mathrm{E}(\xi))\big),$$

amiből

$$\mathrm{E}(\xi) = 6.$$

Az $\mathrm{E}(\xi \mid A_k)$, $1 \leqslant k \leqslant 5$, feltételes várható értékeket (is) számolhatjuk direktben (definíció alapján).

Ha $1 \leqslant k \leqslant 5$, akkor $\mathrm{P}(\xi = 1 \mid A_k) = 0$ és

$$\mathrm{P}(\xi = n \mid A_k) = \frac{5^{n-2}}{6^{n-1}}, \quad n = 2, 3, \ldots .$$

Így $k = 1, \ldots, 5$ esetén

$$\begin{aligned}
\mathrm{E}(\xi \mid A_k) &= \sum_{n=2}^{\infty} n \frac{5^{n-2}}{6^{n-1}} = \frac{1}{5} \sum_{n=2}^{\infty} n \left(\frac{5}{6}\right)^{n-1} = \frac{1}{5} \sum_{n=2}^{\infty} (x^n)' \Big|_{x=5/6} \\
&= \left(\sum_{n=2}^{\infty} x^n\right)' \Big|_{x=5/6} = \frac{1}{5} \left(\frac{x^2}{1-x}\right)' \Big|_{x=5/6} \\
&= \frac{1}{5} \frac{2x - x^2}{(1-x)^2} \Big|_{x=5/6} = 7,
\end{aligned}$$

ahogy vártuk.

## Markov-egyenlőtlenség

Ha a $\xi$ véletlen változó nemnegatív, akkor tetszőleges $\varepsilon > 0$ esetén

$$\mathsf{P}\{\xi \geqslant \varepsilon\} \leqslant \frac{\mathsf{E}(\xi)}{\varepsilon}.$$

**Bizonyítás:** Ha $\xi$ diszkrét véletlen változó nemnegatív $x_1, x_2, \dots$ lehetséges értékekkel, akkor

$$\mathsf{E}(\xi) = \sum_{k=1}^{\infty} x_k \cdot \mathsf{P}\{\xi = x_k\} \geqslant \sum_{k\,:\,x_k \geqslant \varepsilon} x_k \cdot \mathsf{P}\{\xi = x_k\}$$

$$\geqslant \varepsilon \sum_{k\,:\,x_k \geqslant \varepsilon} \mathsf{P}\{\xi = x_k\} = \varepsilon \cdot \mathsf{P}\{\xi \geqslant \varepsilon\}.$$

Ha $\xi$ abszolút folytonos nemnegatív valószínűségi változó $f_\xi$ sűrűségfüggvénnyel, akkor $f_\xi(x) = 0$ ha $x \leqslant 0$, ezért

$$\mathsf{E}(\xi) = \int_0^{\infty} x f_\xi(x)\,\mathrm{d}x \geqslant \int_\varepsilon^{\infty} x f_\xi(x)\,\mathrm{d}x \geqslant \varepsilon \int_\varepsilon^{\infty} f_\xi(x)\,\mathrm{d}x = \varepsilon \cdot \mathsf{P}\{\xi \geqslant \varepsilon\}.$$

## Csebisev-egyenlőtlenség

Tetszőleges $\xi$ véletlen változó, melyre $\mathsf{E}(\xi)$ létezik és véges, és $\varepsilon > 0$ esetén

$$\mathsf{P}\{|\xi - \mathsf{E}(\xi)| \geqslant \varepsilon\} \leqslant \frac{\operatorname{var}(\xi)}{\varepsilon^2}.$$

**Bizonyítás:** A Markov-egyenlőtlenséget alkalmazva

$$\mathsf{P}\{|\xi - \mathsf{E}(\xi)| \geqslant \varepsilon\} = \mathsf{P}\{(\xi - \mathsf{E}(\xi))^2 \geqslant \varepsilon^2\} \leqslant \frac{\mathsf{E}\left[(\xi - \mathsf{E}(\xi))^2\right]}{\varepsilon^2} = \frac{\operatorname{var}(\xi)}{\varepsilon^2}.$$

## Bernoulli-féle nagy számok törvénye

$n$ független kísérlet

$A$ esemény, $p := \mathrm{P}(A)$

$A$ gyakorisága: $k_n(A)$

Ekkor tetszőleges $\varepsilon > 0$ esetén

$$\lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{k_n(A)}{n} - p\right| \geqslant \varepsilon\right\} = 0.$$

Szimbolikusan:

$$\frac{k_n(A)}{n} \xrightarrow{\mathrm{P}} p \qquad \text{ha} \;\; n \to \infty,$$

és azt mondjuk, hogy $\frac{k_n(A)}{n}$ **sztochasztikusan konvergál** $p$-hez, amint $n \to \infty$.

## Nagy számok gyenge törvénye

Legyenek $\xi_1, \xi_2, \ldots$ páronként korrelálatlan, azonos eloszlású véletlen változók, melyeknek véges a második momentumuk. Ekkor tetszőleges $\varepsilon > 0$ esetén

$$\lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{\xi_1 + \cdots + \xi_n}{n} - \mathrm{E}(\xi_1)\right| \geqslant \varepsilon\right\} = 0.$$

Szimbolikusan:

$$\frac{\xi_1 + \cdots + \xi_n}{n} \xrightarrow{\mathrm{P}} \mathrm{E}(\xi_1) \qquad \text{ha } n \to \infty.$$

**Bizonyítás:** Legyen $S_n := \xi_1 + \cdots + \xi_n$, $n \in \mathbb{N}$.
Ekkor $\mathrm{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{k=1}^n \mathrm{E}(\xi_k) = \mathrm{E}(\xi_1)$
és $\mathrm{var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathrm{var}(S_n) = \frac{1}{n^2}\sum_{k=1}^n \mathrm{var}(\xi_k) = \frac{\mathrm{var}(\xi_1)}{n}$,
ezért a Csebisev-egyenlőtlenséggel
$\mathrm{P}\left\{\left|\frac{S_n}{n} - \mathrm{E}(\xi_1)\right| \geqslant \varepsilon\right\} \leqslant \frac{\mathrm{var}(\xi_1)}{n\varepsilon^2} \to 0$ ha $n \to \infty$. ■

Mivel $k_n(A) = \xi_1 + \cdots + \xi_n$, ahol $\xi_1, \ldots, \xi_n$ független $p$ paraméterű Bernoulli-eloszlásúak, ezért ebből következik a Bernoulli-féle nagy számok törvénye.

Csebisev-egyenlőtlenséggel $\mathrm{P}\left\{\left|\frac{k_n(A)}{n} - p\right| \geqslant \varepsilon\right\} \leqslant \frac{p(1-p)}{n\cdot\varepsilon^2}$.

**Példa.** Hányszor kell egy szabályos érmét feldobni, hogy a fejek számának relatív gyakorisága legalább 0.95 valószínűséggel 0.1-nél kevesebbel térjen el a valószínűségtől ($1/2$-től)?

Jelölje a szükséges dobások számát $n$. Olyan $n$-et keresünk, melyre:

$$\mathrm{P}\left\{\left|\frac{k_n(A)}{n} - \frac{1}{2}\right| < 0,1\right\} \geqslant 0.95 \quad \Leftrightarrow \quad \mathrm{P}\left\{\left|\frac{k_n(A)}{n} - \frac{1}{2}\right| \geqslant 0,1\right\} \leqslant 0.05.$$

A fentiek alapján

$$\mathrm{P}\left\{\left|\frac{k_n(A)}{n} - \frac{1}{2}\right| \geqslant 0,1\right\} \leqslant \frac{1}{4\cdot(0.1)^2\cdot n}.$$

Így a követelmény biztosan teljesül, ha

$$\frac{1}{4\cdot(0.1)^2\cdot n} \leqslant 0.05, \qquad \text{azaz } n \geqslant 500,$$

vagyis az érmével legalább 500-szor kell dobni.

## Nagy számok erős törvénye

Legyenek $\xi_1, \xi_2, \ldots$ (teljesen) független, azonos eloszlású véletlen változók, hogy $\mathrm{E}(|\xi_1|) < \infty$. Ekkor

$$\mathrm{P}\left\{\lim_{n\to\infty} \frac{\xi_1 + \cdots + \xi_n}{n} = \mathrm{E}(\xi_1)\right\} = 1.$$

# 8. Centrális határeloszlás-tételek

Jelölje $\varphi_{m,\sigma^2}$ az $(m, \sigma^2)$ paraméterű ($m \in \mathbb{R}$, $\sigma > 0$) normális eloszlás sűrűségfüggvényét, azaz $\varphi_{m,\sigma^2}(x) := \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}$, $x \in \mathbb{R}$.

## Lokális határeloszlás-tétel

$n$ független kísérlet, $A$ esemény, $p := \mathrm{P}(A) \in (0, 1)$,

$A$ gyakorisága: $k_n(A)$

Ekkor

$$\mathrm{P}\{k_n(A) = k\} = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\approx \varphi_{np,np(1-p)}(k) = \frac{1}{\sqrt{2\pi np(1-p)}} \mathrm{e}^{-\frac{(k-np)^2}{2np(1-p)}}.$$

Pontosabban, ha $(\psi_n)_{n=1}^{\infty}$ egy olyan valós számsorozat, melyre $\psi_n = o(n^{2/3})$ ha $n \to \infty$, azaz $\lim\limits_{n\to\infty} n^{-2/3}\psi_n = 0$, akkor

$$\sup_{k\,:\,|k-np|\leqslant\psi_n} \left| \frac{\mathrm{P}\{k_n(A) = k\}}{\varphi_{np,np(1-p)}(k)} - 1 \right| \to 0 \qquad \text{ha } n \to \infty.$$

# 8. Centrális határeloszlás-tételek

Jelölje $\Phi$ a standard normális eloszlás eloszlásfüggvényét, azaz

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-u^2/2} \, \mathrm{d}u, \qquad x \in \mathbb{R}.$$

## Moivre-Laplace-tétel (”globális” alak)

$n$ független kísérlet

$A$ esemény, $p := \mathrm{P}(A) \in (0, 1)$

$A$ gyakorisága: $k_n(A)$

Ekkor $\lim\limits_{n\to\infty} \mathrm{P}\left\{ \frac{k_n(A) - np}{\sqrt{np(1-p)}} < x \right\} = \Phi(x)$, $x \in \mathbb{R}$, sőt

$$\lim_{n\to\infty} \sup_{-\infty \leqslant a < b \leqslant +\infty} \left| \mathrm{P}\left\{ \frac{k_n(A) - np}{\sqrt{np(1-p)}} \in [a, b) \right\} - (\Phi(b) - \Phi(a)) \right| = 0.$$

**Példa.** Hányszor kell egy szabályos érmét feldobni, hogy a fejek számának relatív gyakorisága legalább 0.95 valószínűséggel 0.1-nél kevesebbel térjen el a valószínűségtől (1/2-től)?

A fejek számának $k_n(A)$ gyakoriságára tetszőleges $a > 0$ esetén fennáll, hogy

$$\lim_{n\to\infty} \mathrm{P}\left\{ \frac{k_n(A) - n\frac{1}{2}}{\sqrt{n\frac{1}{2}\left(1-\frac{1}{2}\right)}} \in [-a, a) \right\} = \Phi(a) - \Phi(-a),$$

így

$$\lim_{n\to\infty} \mathrm{P}\left\{ \frac{k_n(A)}{n} - \frac{1}{2} \in \left[ -\frac{a}{2\sqrt{n}}, \frac{a}{2\sqrt{n}} \right) \right\} = 2\Phi(a) - 1.$$

Ezért elég nagy $n$-re:

$$\mathrm{P}\left\{ \left| \frac{k_n(A)}{n} - \frac{1}{2} \right| < \frac{a}{2\sqrt{n}} \right\} \approx 2\Phi(a) - 1.$$

Mivel a

$$2\Phi(a) - 1 = 0.95, \qquad \frac{a}{2\sqrt{n}} = 0.1$$

összefüggésekből $a \approx 1.96$ (táblázat alapján) és $n \approx 96$ adódik, így

$$\mathrm{P}\left\{\left|\frac{k_n(A)}{n} - \frac{1}{2}\right| \leqslant 0,1\right\} \approx 0.95 \qquad \text{ha } n \approx 96.$$

Tehát elég *körülbelül* 96-szor dobni egy szabályos érmével ahhoz, hogy a fejek számának relatív gyakorisága *körülbelül* 0.95 valószínűséggel 0.1-nél kevesebbel térjen el a valószínűségtől ($1/2$-től).

## Centrális határeloszlás-tétel ("globális" alak)

Legyenek $\xi_1, \xi_2, \ldots$ teljesen független, azonos eloszlású véletlen változók, melyeknek véges a második momentumuk. Legyen

$$S_n := \xi_1 + \cdots + \xi_n, \qquad n \in \mathbb{N}.$$

Ekkor

$$\lim_{n\to\infty} \mathsf{P}\left\{ \frac{S_n - \mathsf{E}(S_n)}{\sqrt{\mathrm{var}(S_n)}} < x \right\} = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-u^2/2}\, \mathrm{d}u, \qquad x \in \mathbb{R},$$

azaz $\frac{S_n - \mathsf{E}(S_n)}{\sqrt{\mathrm{var}(S_n)}}$ eloszlásban konvergál a standard normális eloszláshoz, ha $n \to \infty$. Szimbolikusan: $\frac{S_n - \mathsf{E}(S_n)}{\sqrt{\mathrm{var}(S_n)}} \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}(0,1)$, ha $n \to \infty$.

Mivel $k_n(A) = \xi_1 + \cdots + \xi_n$, ahol $\xi_1, \ldots, \xi_n$ független $p$ paraméterű Bernoulli-eloszlásúak, ezért ebből következik a Moivre-Laplace-tétel (első része).

## $(n, M, N - M)$ paraméterű hipergeometrikus eloszlás

Egy dobozban $M$ piros és $N - M$ fekete golyó van ($M < N$). Visszatevés nélkül húzunk ki $n$ golyót ($n \leqslant N$), és $\xi$ jelöli a kihúzott piros golyók számát. Ekkor $\xi$ olyan $k$ értékeket vehet fel, melyre teljesül $0 \leqslant k \leqslant n$, $k \leqslant M$, és $n - k \leqslant N - M$, eloszlása

$$\mathsf{P}\{\xi = k\} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

Ha $\xi_i := \begin{cases} 1 & \text{ha az } i\text{-edik golyó piros,} \\ 0 & \text{egyébként,} \end{cases}$ akkor $\xi = \xi_1 + \cdots + \xi_n$,

és a $\xi_1, \ldots, \xi_n$ véletlen változók $\frac{M}{N}$-paraméterű Bernoulli-eloszlásúak, DE NEM FÜGGETLENEK! Például $i \neq j$ esetén

$$\mathsf{P}\{\xi_i = 1, \, \xi_j = 1\} = \frac{M(M-1)}{N(N-1)}, \qquad \mathsf{P}\{\xi_i = 1\} = \mathsf{P}\{\xi_j = 1\} = \frac{M}{N}.$$

$$\mathsf{E}(\xi) = \mathsf{E}(\xi_1) + \cdots + \mathsf{E}(\xi_n) = n \cdot \frac{M}{N},$$

$$\begin{aligned}\operatorname{var}(\xi) &= \operatorname{cov}(\xi, \xi) = \operatorname{cov}\left(\sum_{i=1}^n \xi_i, \sum_{j=1}^n \xi_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \operatorname{cov}(\xi_i, \xi_j) = \sum_{i=1}^n \operatorname{var}(\xi_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \operatorname{cov}(\xi_i, \xi_j),\end{aligned}$$

ahol

$$\operatorname{var}(\xi_i) = \frac{M}{N}\left(1 - \frac{M}{N}\right), \qquad \operatorname{cov}(\xi_i, \xi_j) = \frac{M(M-N)}{N^2(N-1)},$$

így

$$\operatorname{var}(\xi) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}$$

## $p$ paraméterű elsőrendű negatív binomiális eloszlás

$A$ esemény, $p := \mathsf{P}(A)$, melyre $0 < p < 1$. Jelölje $\xi :=$ az $A$ első bekövetkezéséhez szükséges független kísérletek számát; lehetséges értékei: $1, 2, \ldots, \infty$, eloszlása: $\mathsf{P}\{\xi = \infty\} = 0$, és

$$\mathsf{P}\{\xi = k\} = p \cdot (1-p)^{k-1}, \qquad k = 1, 2, \ldots.$$

$$\mathsf{E}(\xi) = \sum_{k=1}^{\infty} k \cdot p(1-p)^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1},$$

ahol $q := 1 - p$, és

$$\sum_{k=1}^{\infty} kq^{k-1} = \sum_{k=1}^{\infty} (q^k)' = \left(\sum_{k=0}^{\infty} q^k\right)' = \left(\frac{1}{1-q}\right)' = \frac{1}{(1-q)^2},$$

így

$$\mathsf{E}(\xi) = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}.$$

$$\mathsf{E}(\xi^2) = \sum_{k=1}^{\infty} k^2 \cdot p(1-p)^{k-1} = \sum_{k=1}^{\infty} \big[k + k(k-1)\big] \cdot p(1-p)^{k-1}$$
$$= p \sum_{k=1}^{\infty} k q^{k-1} + pq \sum_{k=1}^{\infty} k(k-1) q^{k-2},$$

ahol

$$\sum_{k=1}^{\infty} k(k-1) q^{k-2} = \sum_{k=1}^{\infty} (q^k)'' = \left(\sum_{k=0}^{\infty} q^k\right)'' = \left(\frac{1}{1-q}\right)'' = \frac{2}{(1-q)^3},$$

így

$$\mathsf{E}(\xi^2) = \frac{1}{p} + pq \cdot \frac{2}{(1-q)^3} = \frac{2-p}{p^2}.$$

Végül

$$\operatorname{var}(\xi) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

## $\lambda$ paraméterű Poisson-eloszlás

$$\mathsf{P}\{\xi = k\} = \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}, \qquad k = 0, 1, \ldots, \qquad \text{ahol } \lambda > 0.$$

$$\mathsf{E}(\xi) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \mathrm{e}^{-\lambda} \sum_{\ell=0}^{\infty} \frac{\lambda^{\ell+1}}{\ell!} = \mathrm{e}^{-\lambda}\lambda\mathrm{e}^{\lambda} = \lambda,$$

$$\mathsf{E}(\xi^2) = \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} = \sum_{k=0}^{\infty} \big[k + k(k-1)\big] \cdot \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}$$

$$= \lambda + \mathrm{e}^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} = \lambda + \mathrm{e}^{-\lambda} \sum_{\ell=0}^{\infty} \frac{\lambda^{\ell+2}}{\ell!} = \lambda + \lambda^2,$$

$$\operatorname{var}(\xi) = \lambda + \lambda^2 - \lambda^2 = \lambda.$$

## Egyenletes eloszlás az $\{1, 2, \ldots, N\}$ halmazon

$$\mathsf{P}\{\xi = k\} = \frac{1}{N}, \qquad k = 1, 2, \ldots, N.$$

$$\mathsf{E}(\xi) = \sum_{k=1}^{N} k \cdot \frac{1}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2},$$

$$\mathsf{E}(\xi^2) = \sum_{k=1}^{N} k^2 \cdot \frac{1}{N} = \frac{N(N+1)(2N+1)}{6N} = \frac{(N+1)(2N+1)}{6},$$

$$\mathrm{var}(\xi) = \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} = \frac{N^2-1}{12}.$$

## Egyenletes eloszlás az $(a, b)$ intervallumon

$$F_\xi(x) = \begin{cases} 0 & \text{ha } x \leqslant a, \\ \dfrac{x-a}{b-a} & \text{ha } a < x \leqslant b, \\ 1 & \text{ha } x > b. \end{cases} \qquad f_\xi(x) = \begin{cases} \dfrac{1}{b-a} & \text{ha } a \leqslant x \leqslant b \\ 0 & \text{egyébként} \end{cases}$$

$\xi = a + (b-a)\cdot\eta$, ahol $\eta$ egyenletes eloszlású a $(0, 1)$ intervallumon, hiszen

$$\mathrm{P}\{\eta < y\} = \mathrm{P}\left\{\frac{\xi - a}{b-a} < y\right\} = \mathrm{P}\{\xi < a+(b-a)y\} = \begin{cases} 0 & \text{ha } y \leqslant 0, \\ y & \text{ha } 0 \leqslant y \leqslant 1, \\ 1 & \text{ha } y \geqslant 1. \end{cases}$$

$$\mathsf{E}(\eta) = \int_0^1 y \, \mathrm{d}y = \left[\frac{y^2}{2}\right]_{y=0}^{y=1} = \frac{1}{2},$$

$$\mathsf{E}(\eta^2) = \int_0^1 y^2 \, \mathrm{d}y = \left[\frac{y^3}{3}\right]_{y=0}^{y=1} = \frac{1}{3},$$

$$\operatorname{var}(\eta) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12},$$

ezért

$$\mathsf{E}(\xi) = a + (b-a)\,\mathsf{E}(\eta) = a + \frac{b-a}{2} = \frac{a+b}{2},$$

$$\operatorname{var}(\xi) = (b-a)^2 \operatorname{var}(\eta) = \frac{(b-a)^2}{12}.$$

## $\lambda$ paraméterű exponenciális eloszlás

$$f_\xi(x) = \begin{cases} 0 & \text{ha } x < 0, \\ \lambda e^{-\lambda x} & \text{ha } x \geqslant 0, \end{cases} \qquad \text{ahol } \lambda > 0.$$

$$\mathsf{E}(\xi) = \int_0^\infty x\lambda e^{-\lambda x}\,dx = \left[-xe^{-\lambda x}\right]_{x=0}^{x=\infty} + \int_0^\infty e^{-\lambda x}\,dx = \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_{x=0}^{x=\infty} = \frac{1}{\lambda}$$

$$\begin{aligned} \mathsf{E}(\xi^2) &= \int_0^\infty x^2\lambda e^{-\lambda x}\,dx = \left[-x^2e^{-\lambda x}\right]_{x=0}^{x=\infty} + 2\int_0^\infty xe^{-\lambda x}\,dx \\ &= \frac{2}{\lambda}\int_0^\infty x\lambda e^{-\lambda x}\,dx = \frac{2}{\lambda^2}, \end{aligned}$$

$$\operatorname{var}(\xi) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

## $(m, \sigma^2)$ paraméterű normális eloszlás

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}, \qquad \text{ahol } m \in \mathbb{R}, \ \sigma > 0.$$

$\mathsf{E}(\xi) = m$, $\operatorname{var}(\xi) = \sigma^2$, ferdesége 0, csúcsossága 0.

Eloszlásfüggvénye:

$$F_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(u-m)^2}{2\sigma^2}} \,\mathrm{d}u, \qquad x \in \mathbb{R}.$$

Továbbá $\xi = \sigma \cdot \eta + m$, ahol $\eta$ standard normális eloszlású, azaz paraméterei $m = 0$, $\sigma = 1$, így eloszlásfüggvénye

$$\Phi(x) := F_\eta(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-u^2/2} \,\mathrm{d}u, \qquad x \in \mathbb{R},$$

melynek értékei táblázatokban megtalálhatók.