Valószínűségszámítás

Földtudomány BSc szak, 2016/2017. őszi félév

Backhausz Ágnes

agnes@cs.elte.hu

Tartalomjegyzék

1. Bevezetés

Célok:

- véletlen folyamatok modellezése;
- kísérletekből, felmérésekből származó adatok elemzése;
- ismeretlen mennyiségek becslése a mérések alapján;
- hipotézisek ellenőrzése vagy cáfolata;
- múltbeli adatok alapján a jövőbeli folyamatok előrejelzése.

Alkalmazási területek:

- élő és élettelen természettudományok, társadalomtudományok: kísérleti eredmények értelmezése; véletlen folyamatok (pl. időjárás) előrejelzése;
- gazdasági folyamatok elemzése, biztosítás– és pénzügyi matematika.

Példák

- Két szabályos dobókockával dobunk. Mennyi a valószínűsége, hogy a dobott számok összege 7?
- Mennyi a valószínűsége, hogy decemberben pontosan 6 olyan nap lesz, amikor a hőmérséklet fagypont alá süllyed?
- Csomagot várunk, a futár 8 és 12 óra között érkezik. Mennyi a valószínűsége, hogy még 11 óra előtt megérkezik?
- Mennyi a valószínűsége, hogy a holnapi csapadékmennyiség nem haladja meg a 10 mm-t?
- Mennyi a valószínűsége, hogy egy véletlenszerűen kiválasztott ember balkezes? Annak, hogy pontosan $0, 1, 2, \ldots$ testvére van? Annak, hogy a testmagassága 190 cm és 200 cm közé esik?

2. A Kolmogorov-féle valószínűségi mező

1. Definíció. *Az $(\Omega, \mathcal{A}, \mathbb{P})$ hármas Kolmogorov-féle valószínűségi mező, ha*

- Ω *nem üres halmaz;*
- $\mathcal{A} \subseteq \mathcal{P}(\Omega)$*, azaz minden* $A \in \mathcal{A}$*-ra* $A \subseteq \Omega$ *úgy, hogy*
  - *(i)* $\Omega \in \mathcal{A}$*;*
  - *(ii) ha* $A_1, A_2, \ldots \in \mathcal{A}$*, akkor* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ *(azaz megszámlálható sok* $\mathcal{A}$*-beli elem uniója is* $\mathcal{A}$*-beli);*
  - *(iii) ha* $A \in \mathcal{A}$*, akkor* $\Omega \setminus A \in \mathcal{A}$ *(azaz* $\mathcal{A}$*-beli halmazok komplementere is* $\mathcal{A}$*-beli.*
- $\mathbb{P} : \mathcal{A} \to [0, 1]$ *függvény, melyre*
  - *(i)* $\mathbb{P}(\Omega) = 1$*;*
  - *(ii) ha* $A_1, A_2, \ldots \in \mathcal{A}$ *és minden* $1 \leq i < j$*-re* $A_i \cap A_j = \emptyset$*, akkor*

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Elnevezések:

- Ω: eseménytér vagy elemi események halmaza.
- Ω elemei ($\omega \in \Omega$): elemi események.
- $\mathcal{A}$: események halmaza (vagy események σ-algebrája).
- $\mathcal{A}$ elemei ($A \in \mathcal{A}$): események.
- $\mathbb{P}$: valószínűség (probability).
- Ω esemény neve: biztos esemény.
- $\emptyset$ (üres halmaz) esemény neve: lehetetlen esemény.
- $A \in \mathcal{A}$ és $B \in \mathcal{A}$ kizáró események, ha $A \cap B = \emptyset$, azaz nincs olyan $\omega \in A$, melyre $\omega \in B$ is teljesül.

2. Definíció (Diszkrét valószínűségi mező). *Az $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mező* **diszkrét***, ha Ω véges vagy megszámlálhatóan végtelen.*

Példa: két szabályos kockadobás Két szabályos dobókockával dobunk, egy pirossal és egy kékkel.

- $\Omega = \{11, 12, 13, \ldots, 65, 66\}$, ahol például 13 azt jelenti, hogy a piros dobás 1, a kék 3 (ez különbözik 31-től). $|\Omega| = 6 \cdot 6 = 36$, vagyis 36 darab elemi esemény (dobássorozat) van.
- $\mathcal{A}$: az Ω összes részhalmaza. $|\mathcal{A}| = 2^{36}$.
- $\mathbb{P}(A) = \frac{|A|}{36}$ minden $A \in \mathcal{A}$-ra. Például:

$$A = \{16, 25, 34, 43, 52, 61\} = \{\text{az összeg } 7\}.$$

11	12	13	14	15	**16**
21	22	23	24	**25**	26
31	32	33	**34**	35	36
41	42	**43**	44	45	46
51	**52**	53	54	55	56
61	62	63	64	65	66

Tehát $\mathbb{P}(A) = \mathbb{P}(\text{az összeg } 7) = 6/36 = 1/6$.

Példa: gyerekek nemének eloszlása

Valakinek három gyermeke születik.

- A gyerekek neme a következőképpen alakulhat (születési sorrendben):

$$\Omega = \{LLL, LLF, LFL, LFF, FLL, FLF, FFL, FFF\}.$$

  Ennek 8 eleme van, feltételezzük, hogy mind a nyolc lehetőség egyformán valószínű.

- $\mathcal{A}$: az Ω összes részhalmaza. $|\mathcal{A}| = 2^8$.
- Például legyen $A \subset \Omega$ az az esemény, hogy a gyerekek között pontosan egy lány van: $A = \{LFF, FLF, FFL.\}$ Ennek valószínűsége: $\mathbb{P}(A) = 3/8$.
- Legyen B az az esemény, hogy a középső gyerek fiú:

$$B = \{LFL, LFF, FFL, FFF\}.$$

  Ennek valószínűsége: $\mathbb{P}(B) = 1/2$.

2.1. Klasszikus valószínűségi mező

3. Definíció. *Tegyük fel, hogy az* $(\Omega, \mathcal{A}, \mathbb{P})$ *Kolmogorov-féle valószínűségi mezőre az alábbiak teljesülnek:*

- Ω *véges sok elemből áll:* $\Omega = \{\omega_1, \ldots, \omega_n\}$.
- $\mathcal{A}$ *az* Ω *összes részhalmazából áll.*
- $\mathbb{P}(\omega) = \frac{1}{n}$ *minden* $\omega \in \Omega$*-ra, azaz az elemi események valószínűsége egyenlő.*

Ilyenkor $(\Omega, \mathcal{A}, \mathbb{P})$ **klasszikus valószínűségi mező***.*

1. Állítás. *Klasszikus valószínűségi mezőben tetszőleges* $A \in \mathcal{A}$ *eseményre*

$$\mathbb{P}(A) = \frac{A \textit{ elemeinek száma}}{\Omega \textit{ elemeinek száma}} = \frac{|A|}{|\Omega|}.$$

Példa: két kockadobás, gyerekek neme.

Nem klasszikus valószínűségi mező: a decemberi fagyos napok száma. Az összes lehetőség $0, 1, \ldots, 31$, ez véges sok. Viszont nem jó feltételezés, hogy minden lehetőség (pl. 5 és 30) egyformán valószínű.

2.2. A valószínűség elemi tulajdonságai

Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ Kolmogorov-féle valószínűségi mező.

4. Definíció. *Legyenek* $A, B \in \mathcal{A}$ *események.*

- *Unió:* $A \cup B = \{\omega \in \Omega : \omega \in A \textit{ vagy } \omega \in B\}$.
- *Metszet:* $A \cap B = \{\omega \in \Omega : \omega \in A \textit{ és } \omega \in B\}$.
- *Komplementer/ellentett esemény:* $\overline{A} = \{\omega \in \Omega : \omega \notin A\}$.
- A *maga után vonja* B*-t, ha minden* $\omega \in A$*-ra* $\omega \in B$ *is teljesül. Jelölés:* $A \subseteq B$.

2. Állítás (Komplementer valószínűsége). *Ha* $A \in \mathcal{A}$ *esemény, akkor* $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.

3. Állítás (Tartalmazás). *Ha az* $A, B \in \mathcal{A}$ *eseményekre* $A \subseteq B$*, akkor* $\mathbb{P}(A) \leq \mathbb{P}(B)$.

3. Függetlenség

5. Definíció. *Az* $A, B \in \mathcal{A}$ *események* **függetlenek**, *ha*

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

Példa. Két szabályos dobókockával dobunk, egy pirossal és egy kékkel.

A: a piros kockával hármat dobunk.

B: a kék kockával ötöt dobunk.

C: a dobott számok összege 6.

A és B függetlenek: $\mathbb{P}(A \cap B) = 1/36 = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

A és C nem függetlenek, illetve B és C sem függetlenek.

$\mathbb{P}(A \cap C) = 1/36 \neq 1/6 \cdot 5/36 = \mathbb{P}(A) \cdot \mathbb{P}(C)$.

Példa. Tegyük fel, hogy annak valószínűsége, hogy holnap Budapesten esik az eső, $0,2$. Annak valószínűsége, hogy holnap Torontóban esik az eső, $0,1$. Annak valószínűsége, hogy holnap Budapesten és Torontóban is esik az eső, $0,02$. Ekkor az a két esemény, hogy Budapesten (B), illetve Torontóban (T) esik az eső, *független*, hiszen

$$\mathbb{P}(B \cap T) = \mathbb{P}(\text{mindkét helyen esik az eső}) = 0,02 = 0,2 \cdot 0,1 = \mathbb{P}(B)\mathbb{P}(T).$$

Példa. Annak valószínűsége, hogy holnap Budapesten esik az eső, $0,2$. Annak valószínűsége, hogy holnap Budaörsön esik az eső, $0,15$. Annak valószínűsége, hogy holnap Budapesten és Budaörsön is esik az eső, $0,12$. Ekkor az a két esemény, hogy Budapesten (B), illetve Budaörsön (C) esik az eső, *nem független*, hiszen

$$\mathbb{P}(B \cap C) = 0,12 \neq 0,03 = 0,2 \cdot 0,15 = \mathbb{P}(B)\mathbb{P}(C).$$

6. Definíció (Függetlenség). *Az* $A_1, A_2, \ldots, A_n \in \mathcal{A}$ *véges sok esemény (teljesen) függetlenek, ha minden* $1 \leq k \leq n$*-re és* $1 \leq i_1 < i_2 < \ldots < i_k \leq n$*-re*

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \ldots \mathbb{P}(A_{i_k}).$$

Az $A_1, A_2, \ldots \in \mathcal{A}$ *megszámlálhatóan végtelen sok esemény (teljesen) független, ha minden* $k \in \mathbb{N}$*-re és* $1 \leq i_1 < i_2 < \ldots < i_k$*-ra*

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \ldots \mathbb{P}(A_{i_k}).$$

7. Definíció (Páronkénti függetlenség). *Az* $A_1, A_2, \ldots \in \mathcal{A}$ *események páronként függetlenek, ha minden* $1 \leq i < j$ *esetén* A_i *és* A_j *függetlenek, azaz*

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

Példák függetlenségre:

- Egy helyen öt különböző műszerrel mérjük meg a légszennyezettséget. Legyen C_i az az esemény, hogy az i. műszer mérési hibája a valós érték 10%-ánál több. Feltételezhetjük, hogy a $C_1, \ldots, C_5$ események függetlenek.
- Egy szabályos kockával n-szer dobunk. Legyen A_k az az esemény, hogy a k. dobás kettes ($k = 1, 2, \ldots, n$). Ekkor $A_1, \ldots, A_n$ függetlenek.
- Valakinek n gyereke születik. Legyen B_k az az esemény, hogy a k. gyerek fiú ($k = 1, 2, \ldots, n$). Ekkor feltételezhetjük, hogy $B_1, \ldots, B_n$ függetlenek.

4. Állítás. *Ha az* $A_1, A_2, \ldots, A_n$ *események függetlenek, akkor páronként függetlenek. A páronkénti függetlenségből nem következik a függetlenség.*

Példa. Szabályos dobókockával kétszer dobunk. Legyen

A: az első dobás páros. B: a második dobás páros. C: a két dobás összege páros.

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2.$$
$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = 1/4.$$

Ugyanakkor, ha az első és a második dobás páros, akkor a két dobás összege is biztosan páros, így

$$\mathbb{P}(A \cap B \cap C) = 1/4 \neq 1/2 \cdot 1/2 \cdot 1/2 = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C).$$

Tehát: A, B, C páronként függetlenek, de nem függetlenek.

5. Állítás (Komplementer függetlensége). *Ha az* $A_1, A_2, \ldots, A_n \in \mathcal{A}$ *események függetlenek, akkor az* $\overline{A_1}, A_2, \ldots, A_n \in \mathcal{A}$ *események is függetlenek.*

6. Állítás. *Ha az $A, B \in \mathcal{A}$ események függetlenek, akkor*

- *A és $\overline{B}$ függetlenek;*
- *$\overline{A}$ és B függetlenek;*
- *$\overline{A}$ és $\overline{B}$ függetlenek.*

7. Állítás. *Legyenek $A, B \in \mathcal{A}$ események úgy, hogy $\mathbb{P}(A) = 0$, vagy $\mathbb{P}(A) = 1$. Ekkor A és B függetlenek.*

4. Valószínűségi változók és eloszlásuk

7.1. Definíció (Valószínűségi változó). *Egy $X : \Omega \to \mathbb{R}$ függvény valószínűségi változó, ha tetszőleges $a < b$ valós számokra*

$$\{\omega \in \Omega : a < X(\omega) \leq b\} \in \mathcal{A}.$$

Példa. Valakinek három gyereke születik. Legyen X a fiúk száma. Ekkor

$$\Omega = \{FFF, FFL, FLF, FLL, LFF, LFL, LLF, LLL\};$$

$$X(LLL) = 0; \qquad X(LLF) = X(LFL) = X(FLL) = 1;$$
$$X(FFL) = X(FLF) = X(LFF) = 2; \qquad X(FFF) = 3.$$

4.1. Diszkrét valószínűségi változók eloszlása

7.2. Definíció (Diszkrét valószínűségi változó). *Az $X : \Omega \to \mathbb{R}$ valószínűségi változó diszkrét, ha véges sok vagy megszámlálható sok értéket vesz fel.*

7.3. Definíció (Eloszlás). *Legyenek az X valószínűségi változó lehetséges értékei $x_1, x_2, \ldots \in \mathbb{R}$, és legyen*

$$p_i = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x_i\}) = \mathbb{P}(X = x_i) \qquad (i = 1, 2, \ldots).$$

Ekkor az $(x_1, p_1), (x_2, p_2), \ldots$ sorozat az X valószínűségi változó eloszlása.

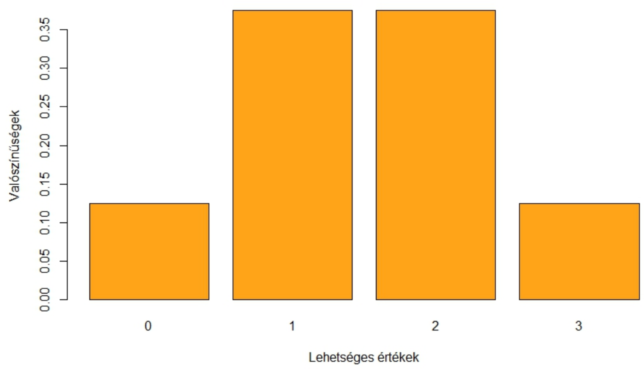
Ezentúl a definícióban szereplő rövidebb, $\mathbb{P}(X = x_i)$ jelölést fogjuk használni.

Vegyük észre, hogy $0 \leq p_i \leq 1$ teljesül minden $i = 1, 2, \dots$ esetén, és $\sum_{i=1}^{\infty} p_i = 1$. Ugyanis $\{X = x_i\}$ teljes eseményrendszer (pontosan egy következik be közülük), tehát az uniójuk Ω, így pedig az 1. definíció $\mathbb{P}$-re vonatkozó részéből következik az állítás.

7.4. Definíció (Valószínűségeloszlás). *Azt mondjuk, hogy a $p_1, p_2, \dots$ sorozat valószínűségeloszlás, ha*

- $p_i \geq 0$ *minden* $i = 1, 2, \dots$ *esetén;*
- $\sum_{i=1}^{\infty} p_i = 1$.

Tehát azt láttuk, hogy ha (x_i, p_i) egy X valószínűségi változó eloszlása, akkor p_i valószínűségeloszlás.



1. ábra.
A fiúk számának lehetséges értékei a három gyerek közül és a hozzájuk tartozó valószínűségek

Példa: három gyerek. Az előző példában: háromszor gyerek születik, X a fiúk száma. Ekkor X lehetséges értékei: $0, 1, 2, 3$, vagyis $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$. Tegyük fel, hogy minden gyerek a többiektől függetlenül $1/2$ valószínűséggel fiú. Eszerint mind a nyolc lehetőség valószínűsége $1/8$, hiszen például $\mathbb{P}(FLF) = 1/2 \cdot 1/2 \cdot 1/2 = 1/8$. Így, a fenti csoportosításból adódóan:

$$\mathbb{P}(X = 0) = 1/8; \quad \mathbb{P}(X = 1) = 3/8; \quad \mathbb{P}(X = 2) = 3/8; \quad \mathbb{P}(X = 3) = 1/8.$$

Mindezek alapján X eloszlása az alábbi sorozat:

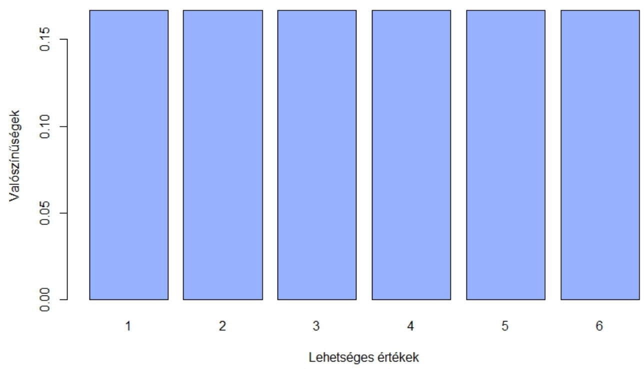
$$(0, 1/8), \qquad (1, 3/8), \qquad (2, 3/8), \qquad (3, 1/8).$$

Ez az eloszlás látható az 1. ábrán.

Példa: szabályos kockadobás. Egyszer dobunk szabályos dobókockával, jelölje Y a dobott számot. Ekkor Y eloszlása (2. ábra):

$$(1, 1/6), \qquad (2, 1/6), \qquad (3, 1/6), \qquad (4, 1/6), \qquad (5, 1/6), \qquad (6, 1/6).$$

A 3. ábrán 1000 (számítógéppel szimulált) szabályos kockadobás után mind a hat lehetséges értékről feltüntettük a relatív gyakoriságát: az előfordulások számának és az összes dobásnak a hányadosát.



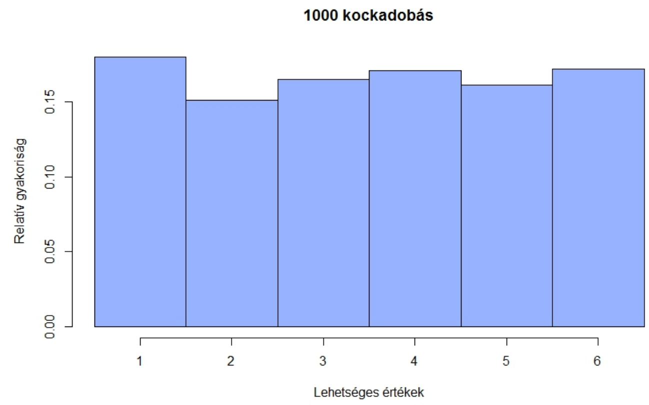
2. ábra.
A szabályos kockadobás lehetséges értékei és a hozzájuk tartozó valószínűségek

Példa: jégeső. Tegyük fel, hogy júliusban $0,7$ valószínűséggel nincs jégeső, $0,2$ valószínűséggel egyszer fordul elő, $0,1$ valószínűséggel pedig kétszer. Legyen Z a júliusi jégesők száma. Ekkor Z eloszlása:

$$(0, 7/10), \qquad (1, 2/10), \qquad (2, 1/10).$$

4.2. Diszkrét valószínűségi változó várható értéke

7.5. Definíció (Várható érték, diszkrét eset). *Legyen $X : \Omega \to \mathbb{R}$ olyan diszkrét valószínűségi változó, melynek eloszlása $(x_1, p_1), (x_2, p_2), \ldots$. Ekkor*



3. ábra. 1000 szabályos kockadobásból készített hisztogram

X várható értéke:

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i p_i, \qquad \textit{ha} \ \ \mathbb{E}(X) = \sum_{i=1}^{\infty} |x_i| p_i < \infty.$$

A definícióból is látszik, hogy a várható érték csak a valószínűségi változó eloszlásától függ.

Példák várható értékre

Fiúk száma. Továbbra is három gyerek születik, egymástól függetlenül $1/2 - 1/2$ valószínűséggel fiúk, illetve lányok. X a fiúk száma. A fenti számolás alapján

$$\mathbb{E}(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2} = 1,5.$$

Kockadobás. Y a dobott számot jelöli. Ekkor Y várható értéke:

$$\mathbb{E}(Y) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2} = 3,5.$$

Jégeső. A fenti példában a júliusi jégeső számának várható értéke:

$$\mathbb{E}(Z) = 0 \cdot 0,7 + 1 \cdot 0,2 + 2 \cdot 0,1 = 0,4.$$

4.3. Diszkrét valószínűségi változó szórása

7.6. Definíció (Szórásnégyzet). *Legyen $X : \Omega \to \mathbb{R}$ diszkrét valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor X szórásnégyzete:*

$$D^2(X) = \mathbb{E}\Big((X - \mathbb{E}X)^2\Big).$$

7.7. Definíció (Szórás). *Legyen $X : \Omega \to \mathbb{R}$ diszkrét valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor X szórásnégyzete:*

$$D(X) = \sqrt{\mathbb{E}\Big((X - \mathbb{E}X)^2\Big)}.$$

8. Állítás. *Legyen X olyan diszkrét valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor*

$$D^2(X) = \mathbb{E}(X^2) - \big[\mathbb{E}(X)\big]^2.$$

Példák szórásnégyzetre

Fiúk száma. X a fiúk száma a három gyerek között. Már láttuk, hogy X lehetséges értékei: $0, 1, 2, 3$, és $\mathbb{E}(X) = 3/2$. A definíció alapján tehát ezt kell kiszámítanunk:

$$D^2(X) = \mathbb{E}\big((X - 3/2)^2\big).$$

Az $X - 3/2$ valószínűségi változó eloszlása (minden lehetséges érték mellett feltüntetjük a valószínűségét):

$$(-3/2; 1/8); \qquad (-1/2; 3/8); \qquad (1/2, 3/8); \qquad (3/2, 1/8).$$

Ebből látszik, hogy az $(X - 3/2)^2$ valószínűségi változó lehetséges értékei csak $9/4$, illetve $1/4$. Az első esetben: $(X - 3/2)^2 = 9/4$ pontosan akkor teljesül, ha $X = 0$ vagy $X = 3$. Ennek valószínűsége $1/8 + 1/8 = 1/4$. Tehát $(X - 3/2)^2$ eloszlása:

$$(1/4, 3/4); \qquad (9/4, 1/4).$$

A 7.5. definíció alapján:

$$D^2(X) = \mathbb{E}\big((X - 3/2)^2\big) = \frac{1}{4} \cdot \frac{3}{4} + \frac{9}{4} \cdot \frac{1}{4} = \frac{12}{16} = \frac{3}{4}.$$

Másik megoldás a 8. állítás alapján. Az X valószínűségi változó lehetséges értékei $0, 1, 2, 3$ voltak. Így az X^2 valószínűségi változó lehetséges értékei $0, 1, 4, 9$. Eloszlása pedig:

$$(0, 1/8); \qquad (1, 3/8); \qquad (4, 3/8); \qquad (9, 1/8).$$

Így a 7.5. definíció szerint

$$\mathbb{E}(X^2) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 4 \cdot \frac{3}{8} + 9 \cdot \frac{1}{8} = \frac{24}{8} = 3.$$

Ebből és a korábbi számolásból

$$D^2(X) = \mathbb{E}(X^2) - \big[\mathbb{E}(X)\big]^2 = 3 - 1,5^2 = 3 - 2,25 = 0,75 = \frac{3}{4}.$$

Végül pedig a fiúk számának szórása:

$$D(X) = \sqrt{\frac{3}{4}} = 0,866.$$

Kockadobás. Most már csak a 8. állítás alapján számolunk. Az Y^2 valószínűségi változó lehetséges értékei: $1, 4, 9, 16, 25, 36$, mindegyiknek $1/6$ a valószínűsége. Tehát

$$\mathbb{E}(Y^2) = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} = \frac{91}{6}.$$

Tehát, mivel már láttuk, hogy $\mathbb{E}(Y) = 3,5 = \frac{7}{2}$:

$$D^2(Y) = \mathbb{E}(Y^2) - \big[\mathbb{E}(Y)\big]^2 = \frac{91}{6} - \frac{49}{4} = \frac{182 - 147}{12} = \frac{35}{12} = 2,9167,$$

amiből a kockadobás értékének szórása ennek négyzetgyöke:

$$D(Y) = 1,7078.$$

Jégeső. A fenti példában a júliusi jégesők száma (amit Z-vel jelöltünk) 0 volt $0,7$ valószínűséggel, 1 volt $0,2$ valószínűséggel, és 2 volt $0,1$ valószínűséggel. Azt is láttuk, hogy $\mathbb{E}(Z) = 0,4$. Tehát most

$$\mathbb{E}(Z^2) = 0^2 \cdot 0,7 + 1^2 \cdot 0,2 + 2^2 \cdot 0,1 = 0,6.$$

Tehát a jégesők számának szórása:

$$D(Z) = \sqrt{\mathbb{E}(Z^2) - \big[\mathbb{E}(Z)\big]^2} = \sqrt{0,6 - 0,4^2} = \sqrt{0,44} = 0,6633.$$

5. Nevezetes diszkrét eloszlások

5.1. Binomiális eloszlás

Az X valószínűségi változó binomiális eloszlású n renddel és p paraméterrel, ha

- n független kísérletet végzünk;
- mindegyik p valószínűséggel sikerül;
- X a sikeres kísérletek száma.

Ezt az eloszláson keresztül a következőképpen tudjuk definiálni.

8. Definíció. *Az X valószínűségi változó binomiális eloszlású n renddel és p paraméterrel, ha lehetséges értékei:*

$$0, 1, 2, \ldots, n,$$

és minden $1 \leq k \leq n$ egészre

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

($n \geq 1$ egész, $0 < p < 1$.) Jelölés: Bin(n, p).

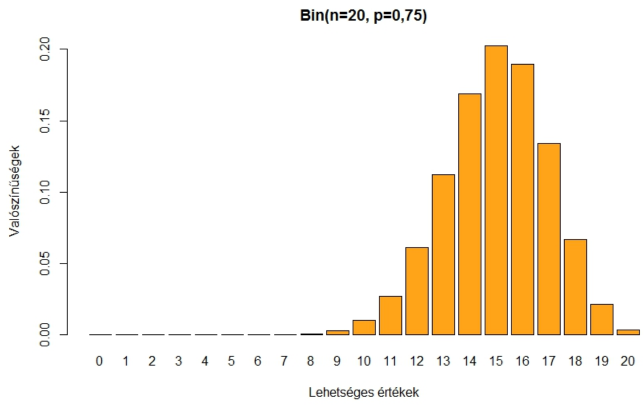
Példa: binomiális eloszlás

- A fiúk száma: 3 gyerek esetén a születendő fiúk száma (X) binomiális eloszlású $n = 3$ renddel és $p = 1/2$ paraméterrel (függetlenséget feltételezve).

  $$\mathbb{P}(\text{pontosan } k \text{ fiú születik}) = \mathbb{P}(X = k) = \binom{3}{k} 0,5^k 0,5^{3-k}.$$

  Például

  $$\mathbb{P}(\text{pontosan 2 fiú születik}) = \mathbb{P}(X = 2) = \binom{3}{2} 0,5^2 0,5^1 = 3 \cdot \left(\frac{1}{2}\right)^3 = 3/8.$$



4. ábra. Binomiális eloszlás, $n = 20$, $p = 0,75$.

- Tegyük fel, hogy júliusban minden nap a többitől függetlenül $0,04$ valószínűséggel jön jégeső. Ekkor a jégesők száma (Y) binomiális eloszlású $n = 31$ renddel és $p = 0,04$ paraméterrel:

$$\mathbb{P}(\text{pontosan } k \text{ napon van jégeső}) = \mathbb{P}(Y = k) = \binom{31}{k} 0,04^k 0,96^{31-k}.$$

9. Állítás (A binomiális eloszlás tulajdonságai). *Legyen X binomiális eloszlású valószínűségi változó n renddel és p paraméterrel.*

(a) *Az $X \sim Bin(n,p)$ binomiális eloszlású valószínűségi változó várható értéke:*
$$\mathbb{E}(X) = np.$$

(b) *Az $X \sim Bin(n,p)$ binomiális eloszlású valószínűségi változó szórása:*
$$D(X) = \sqrt{np(1-p)}.$$

(c) *Az a k érték, melyre $p_k = \mathbb{P}(X = k)$ maximális (vagyis X módusza): $[(n+1)p]$, ahol $[\cdot]$ az egész részt jelöli. Ha $(n+1)p$ egész, akkor az eggyel kisebb k is maximális értéket ad.*

Példa: $X \sim \text{Bin}(3, 1/2)$ a fiúk száma. Ekkor

$$\mathbb{E}(X) = 3 \cdot 1/2 = 3/2; \qquad D(X) = \sqrt{3 \cdot 1/2 \cdot 1/2} = \sqrt{\frac{3}{4}} \approx 0,866.$$

Tegyük fel, hogy egy osztályban mind a 32 gyerek a többiektől függetlenül $1/4$ valószínűséggel balkezes. Jelölje Y a balkezes gyerekek számát. Ekkor Y binomiális eloszlású $n = 32$ renddel és $p = 1/4$ paraméterrel. Ebből az állítás alapján következik, hogy $\mathbb{E}(Y) = np = 8$ és $D(Y) = \sqrt{np(1-p)} = \sqrt{32 \cdot 3/16} = \sqrt{6}$.

5.2. Poisson-eloszlás

9. Definíció. *Legyen $s > 0$. Azt mondjuk, hogy az X valószínűségi változó s paraméterű Poisson-eloszlású, ha lehetséges értékei $k = 0, 1, 2, \ldots$, a hozzájuk tartozó valószínűségek pedig:*

$$\mathbb{P}(X = k) = \frac{s^k}{k!}e^{-s} \qquad (k = 0, 1, \ldots).$$

10. Állítás (A Poisson-eloszlás tulajdonságai). *Legyen X Poisson-eloszlású valószínűségi változó s paraméterrel.*

(a) Ekkor X várható értéke, szórása és szórásnégyzete:

$$\mathbb{E}(X) = s; \qquad D(X) = \sqrt{s}; \qquad D^2(X) = s.$$

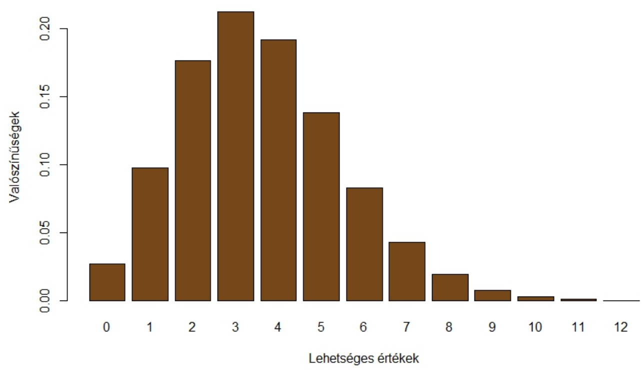
(b) A $\mathbb{P}(X = k)$ valószínűség $k = [s]$ esetén maximális. Ha s egész, az eggyel kisebb k is a legnagyobb értéket adja.

Példa. Tegyük fel, hogy a nyári jégesők száma, Z, Poisson-eloszlású $s = 3,61$ paraméterrel. Ekkor $\mathbb{E}(Z) = s = 3,61$, illetve $D(Z) = \sqrt{s} = \sqrt{3,61}$.

A Poisson-eloszlás és a binomiális eloszlás kapcsolata

A Poisson-eloszlást általában akkor használják, ha sok független, kis valószínűséggel bekövetkező eseménynél a bekövetkező események számát kell tekinteni. Például:

- a földrengések száma egy év alatt;
- a sajtóhibák száma egy könyvben;
- egy biztosító 15000 ügyfele által összesen okozott balesetek száma;
- nagyobb kárt okozó viharok száma egy adott időszakban.



5. ábra.
Poisson-eloszlás, $s = 3,61$, $k = 12$-ig, a nyári jégesők számának közelítésére.

Legyen $s > 0$ pozitív szám, és $p_n = s/n$ minden $n = 1, 2, \ldots$ egészre. Legyen X Poisson-eloszlású valószínűségi változó s paraméterrel, Y_n eloszlása pedig $\mathrm{Bin}(n, p_n)$. Ekkor tetszőleges $k = 0, 1, 2, \ldots$ esetén

$$\lim_{n\to\infty} \mathbb{P}(Y_n = k) = \mathbb{P}(X = k),$$

azaz

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{s^k}{k!} e^{-s} \qquad (k = 0, 1, 2, \ldots).$$

Legyen $s > 0$ pozitív szám, és $p_n = s/n$ minden $n = 1, 2, \ldots$ egészre. Legyen X Poisson-eloszlású valószínűségi változó s paraméterrel, Y_n eloszlása pedig $\mathrm{Bin}(n, p_n)$. Ekkor tetszőleges $k = 0, 1, 2, \ldots$ esetén

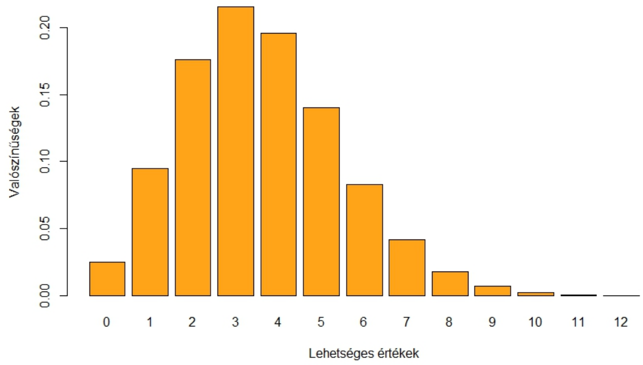
$$\lim_{n\to\infty} \mathbb{P}(Y_n = k) = \mathbb{P}(X = k),$$

azaz

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{s^k}{k!} e^{-s} \qquad (k = 0, 1, 2, \ldots).$$

- X Poisson-eloszlású $s = 3,61$ paraméterrel;
- Y binomiális eloszlás $n = 92$ renddel és $p = 0,0392$ paraméterrel;
- vegyük észre, hogy $\mathbb{E}(X) = s = 3,61 = n \cdot p = \mathbb{E}(Y)$.

k	0	1	2	3	4	5	6	7
$\mathbb{P}(X = k)$	0,027	0,098	0,176	0,212	0,191	0,138	0,083	0,042
$\mathbb{P}(Y = k)$	0,025	0,094	0,176	0,215	0,195	0,14	0,083	0,043



6. ábra.
A nyári jégesők számának közelítése: binomiális eloszlás, $n = 92$, $p = 0,0392$, $k = 12$-ig.

5.3. Geometriai eloszlás

Az Y valószínűségi változó geometriai eloszlású, ha

- független kísérleteket végzünk;
- mindegyik p valószínűséggel sikerül;
- Y: hányadik kísérlet az első sikeres.

Az eloszláson keresztül így tudjuk ezt definiálni.

10. Definíció. *Az Y valószínűségi változó geometriai eloszlású p paraméterrel, ha lehetséges értékei:*

$$1, 2, 3 \dots$$

és minden $1 \leq k$ egészre

$$\mathbb{P}(Y = k) = (1 - p)^{k-1} p.$$

($0 < p < 1$.) Jelölés: Geo(p). Másik elnevezés: Pascal-eloszlás.

Példa: egy szabályos dobókockával dobunk sokszor egymás után. Jelölje Y, hogy hányadik dobásnál kapjuk az első hatost. Ekkor Y geometriai eloszlású $p = 1/6$ paraméterrel. Y lehetséges értékei $k = 1, 2, \dots$, és

$$\mathbb{P}(Y = k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}.$$

11. Állítás (A geometriai eloszlás tulajdonságai). *Legyen Y geometriai eloszlású valószínűségi változó p paraméterrel.*

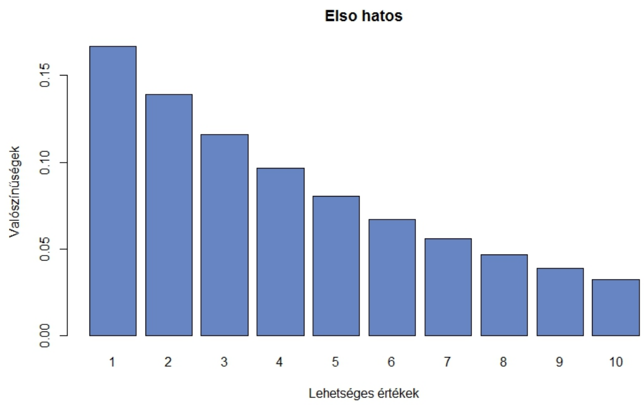
(a)

$$\mathbb{E}(Y) = \frac{1}{p}; \qquad D(Y) = \sqrt{\frac{1-p}{p^2}}.$$

(b) A $\mathbb{P}(Y = k)$ valószínűség $k = 1$-re maximális.

A példában szereplő Y várható értéke és szórása:

$$\mathbb{E}(Y) = 6; \qquad D(Y) = \sqrt{\frac{5/6}{1/36}} = \sqrt{30} = 5,477.$$



7. ábra. Az első hatos eloszlása: geometriai eloszlás, $p = 1/6$, $k = 10$-ig.

5.4. Negatív binomiális eloszlás

- független kísérleteket végzünk;
- mindegyik p valószínűséggel sikerül;
- Z: hányadik kísérlet az r. sikeres.

11. Definíció. *A Z valószínűségi változó negatív binomiális eloszlású p paraméterrel, ha lehetséges értékei:*

$$r, r+1, r+2, \ldots$$

és minden $k \geq r, r+1, r+2\dots$ egészre

$$\mathbb{P}(Z = k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r.$$

($r \geq 1$ egész, $0 < p < 1$.)

$r = 1$-re a negatív binomiális eloszlás megegyezik a geometriai eloszlással.

12. Definíció. *A Z valószínűségi változó negatív binomiális eloszlású p paraméterrel, ha lehetséges értékei:*

$$r, r+1, r+2, \dots$$

és minden $k \geq r, r+1, r+2\dots$ egészre

$$\mathbb{P}(Z = k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r.$$

($r \geq 1$ egész, $0 < p < 1$.)

12. Állítás (A negatív binomiális eloszlás tulajdonságai). *Legyen Z negatív binomiális eloszlású valószínűségi változó p paraméterrel. Ekkor*

$$\mathbb{E}(Z) = \frac{r}{p}; \qquad D(Z) = \sqrt{r\frac{1-p}{p^2}}.$$

5.5. Hipergeometrikus eloszlás

Példa. Egy osztályban 8 balkezes és 25 jobbkezes gyerek van. Tornaórán véletlenszerűen kiválasztanak egy hatfős csapatot (minden hatfős csoportot azonos valószínűséggel választanak). Mennyi a valószínűsége, hogy a kiválasztott csapatban pontosan két balkezes gyerek van?

$$\mathbb{P}(\text{pontosan két balkezes}) = \frac{\binom{8}{2} \cdot \binom{25}{4}}{\binom{33}{6}}.$$

Egy dobozban N golyó van, közülük M fekete, a többi fehér. Visszatevés nélkül kihúznak n darabot (minden húzásnál minden, még a dobozban lévő golyót azonos valószínűséggel választva). Tegyük fel, hogy $n \leq M$ és $n \leq$

$N-M$. Annak valószínűsége, hogy pontosan k darab fekete golyót húznak ki:

$$\mathbb{P}(\text{pontosan } k \text{ fekete}) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}.$$

13. Definíció. *Legyenek N, M, n pozitív egészek úgy, hogy $1 \leq n \leq M \leq N$. Az X valószínűségi változó hipergeometrikus eloszlású N, M és n paraméterekkel ha lehetséges értékei $k = 0, 1, \ldots, n$, és $k = 0, 1, \ldots, n$ esetén*

$$\mathbb{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

Példa: visszatevés nélküli mintavételnél a húzott fekete golyók száma. A fenti példában a kiválasztott balkezes gyerekek száma hipergeometrikus eloszlású: $N = 32, M = 8, n = 6$.

Lottósorsolásnál a találatok száma (X) hipergeometrikus eloszlású $N = 90$, $M = 5$, $n = 5$ paraméterekkel:

$$\mathbb{P}(X = k) = \frac{\binom{5}{k}\binom{85}{5-k}}{\binom{90}{5}} \qquad k = 0, 1, 2, 3, 4, 5.$$

13. Állítás (A hipergeometrikus eloszlás tulajdonságai). *Legyen az X valószínűségi változó hipergeometrikus eloszlású N, M és n paraméterekkel.*

(a) Az X valószínűségi változó várható értéke:

$$\mathbb{E}(X) = n \cdot \frac{M}{N}.$$

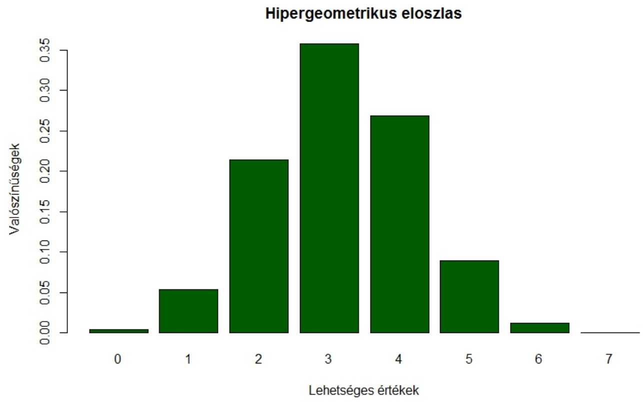
(b) Az X valószínűségi változó szórása:

$$D(X) = \sqrt{n \cdot \frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}}.$$

Példa. Az ötöslottón a találatok számának várható értéke és szórása:

$$\mathbb{E}(X) = \frac{25}{90} = 0,2778; \qquad D(X) = 0,5006.$$

A fenti példában a kiválasztott balkezes gyerekek számának várható értéke: $8 \cdot 6/32 = 1,5$.



8. ábra. Hipergeometrikus eloszlás, $N = 20$, $M = 9, n = 7$.

6. Eloszlásfüggvény és sűrűségfüggvény

13.1. Definíció (Eloszlásfüggvény). *Legyen* $X : \Omega \to \mathbb{R}$ *valószínűségi változó. Ekkor* X *eloszlásfüggvénye az alábbi* $F : \mathbb{R} \to [0,1]$ *függvény:*

$$F(t) = \mathbb{P}(X \le t) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \le t\}) \qquad \textit{minden } t \in \mathbb{R} \textit{ valós számra.}$$

14. Állítás. *Legyenek* $a, b \in \mathbb{R}$ *tetszőleges valós számok. Ekkor*

$$\mathbb{P}(a < X \le b) = F(b) - F(a).$$

Bizonyítás. Ez azonnal adódik az $\mathbb{P}(a < X \le b) = \mathbb{P}(X \le b) - \mathbb{P}(X \le a)$ egyenlőségből és az eloszlásfüggvény definíciójából. □

Példa. Szabályos kockával dobunk. A dobott számot jelölje X. Legyen F az X eloszlásfüggvénye. Ekkor

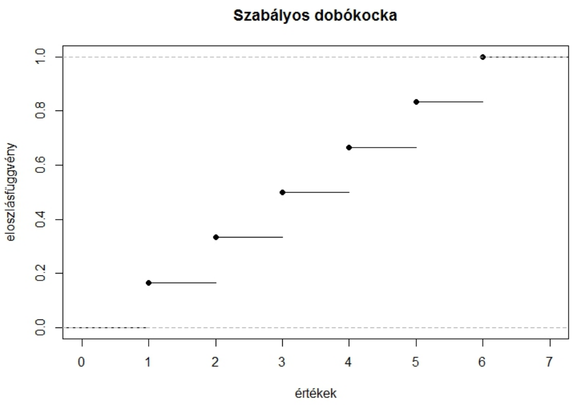
$$F(0) = \mathbb{P}(X \le 0) = 0; \qquad F(1) = \mathbb{P}(X \le 1) = 1/6;$$

$$F(\pi) = \mathbb{P}(X \le \pi) = \mathbb{P}(X \le 3) = 1/2; \qquad F(6) = \mathbb{P}(X \le 6) = 1.$$

1. Tétel (Az eloszlásfüggvény tulajdonságai). *Legyen* X *valószínűségi változó,* F *pedig az eloszlásfüggvénye. Ekkor*

(i) F *monoton növő:* $a < b$ *esetén* $F(a) \le F(b)$.

(ii) $\lim_{t \to -\infty} F(t) = 0; \qquad \lim_{t \to \infty} F(t) = 1.$



9. ábra. Szabályos dobókockával dobott szám eloszlásfüggvénye.

(iii) F jobbról folytonos, azaz minden $t \in \mathbb{R}$ valós számra $\lim_{s \to t-} F(s) = F(t)$.

1. Megjegyzés. *Ha a $G : \mathbb{R} \to [0,1]$ függvény rendelkezik a tételben szereplő $(i)-(iii)$ tulajdonságokkal, akkor van olyan valószínűségi változó, melynek G az eloszlásfüggvénye.*

13.2. Definíció. *Azt mondjuk, hogy az X valószínűségi változó folytonos, ha eloszlásfüggvénye folytonos.*

Egy valószínűségi változó pontosan akkor folytonos, ha $\mathbb{P}(X = t) = 0$ teljesül minden t számra. Így például a kockadobás nem folytonos: $\mathbb{P}(X = 1) = 1/6 \neq 0$. Általánosabban, nincs olyan valószínűségi változó, mely egyszerre diszkrét és folytonos is lenne. Olyan viszont van, ami sem nem diszkrét, sem nem folytonos (például a napi csapadékmennyiség, ami pozitív valószínűséggel nulla, viszont nem megszámlálhatóan végtelen az értékkészlete).

13.3. Definíció (Abszolút folytonosság és sűrűségfüggvény). *Az X valószínűségi változó abszolút folytonos, ha van olyan $f : \mathbb{R} \to \mathbb{R}$ függvény, melyre*

$$\mathbb{P}(X \leq t) = \int_{-\infty}^{t} f(s)ds$$

teljesül minden $t \in \mathbb{R}$ számra. Ilyenkor az f függvényt az X valószínűségi változó sűrűségfüggvényének nevezzük.

15. Állítás. *Legyen az X abszolút folytonos valószínűségi változó, melynek sűrűségfüggvénye f. Ekkor tetszőleges $a < b$ számokra teljesül, hogy*

$$\mathbb{P}(a \le X \le b) = \int_a^b f(s)ds.$$

Példa: lépcsős sűrűségfüggvény. Tegyük fel, hogy a holnapi csapadékmennyiség sűrűségfüggvénye az alábbi:

$$f(s) = \begin{cases} 0,2, & \text{ha } s \in [0,1]; \\ 0,4, & \text{ha } s \in (1,3]; \\ 0 & \text{ha } s < 0 \text{ vagy } s > 2. \end{cases}$$

Jelölje a holnapi csapadékmennyiséget X. Ekkor annak valószínűsége, hogy holnap legfeljebb $0,5$ mm csapadék lesz:

$$\mathbb{P}(0 \le X \le 0,5) = \int_0^{0,5} f(s)ds = \int_0^{0,5} 0,2ds = 0,5 \cdot 0,2 = 0,1.$$

Annak valószínűsége, hogy a holnapi csapadékmennyiség 0,5 mm és 2 mm között lesz:

$$\mathbb{P}(0,5 \le X \le 2) = \int_{0,5}^{2} f(s)ds = \int_{0,5}^{1} 0,2ds + \int_1^2 0,4ds = 0,5 \cdot 0,2 + 1 \cdot 0,4 = 0,5.$$

16. Állítás (Az eloszlásfüggvény és sűrűségfüggvény kapcsolata). *Legyen X abszolút folytonos valószínűségi változó, melynek F az eloszlásfüggvénye.*

(a) Ha f az X sűrűségfüggvénye, akkor minden $t \in \mathbb{R}$ számra

$$F(t) = \int_{-\infty}^{t} f(s)ds.$$

(b) Az $f(t) = F'(t)$ függvény (azokra a t-kre, ahol F differenciálható) az X sűrűségfüggvénye.

17. Állítás (A sűrűségfüggvény tulajdonságai). *Legyen $f : \mathbb{R} \to \mathbb{R}$ az X valószínűségi változó sűrűségfüggvénye. Ekkor*

(i) $f(s) \ge 0$ teljesül "majdnem minden" $s \in \mathbb{R}$-re (például véges vagy megszámlálható sok kivétel lehetséges).

(ii) $\int_{-\infty}^{\infty} f(s)ds = 1$.

2. Megjegyzés. *Ha a $g : \mathbb{R} \to \mathbb{R}$ függvényre teljesül a fenti (i) és (ii) tulajdonság, akkor van olyan X valószínűségi változó, melynek g a sűrűségfüggvénye.*

6.1. Abszolút folytonos valószínűségi változók várható értéke, szórása és momentumai

13.4. Definíció (Várható érték, abszolút folytonos eset). *Legyen X abszolút folytonos valószínűségi változó, melynek sűrűségfüggvénye f. Ekkor X várható értéke:*

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} s \cdot f(s) ds,$$

ha ez az integrál létezik és véges.

13.5. Definíció (Szórásnégyzet és szórás). *Tegyük fel, hogy az X valószínűségi változó abszolút folytonos, és sűrűségfüggvénye f. Ekkor X szórásnégyzete:*

$$D^2(X) = \mathbb{E}\big[(X - \mathbb{E}(X))^2\big],$$

szórása pedig

$$D(X) = \sqrt{\mathbb{E}\big[(X - \mathbb{E}(X))^2\big]},$$

ha ezek a várható értékek léteznek.

18. Állítás (A szórásnégyzet kiszámítása). *A szórásnégyzetet a következőképpen számíthatjuk ki abszolút folytonos X valószínűségi változó esetén:*

$$D^2(X) = \mathbb{E}(X^2) - \big[\mathbb{E}(X)\big]^2 = \int_{-\infty}^{\infty} s^2 f(s) ds - \left[\int_{-\infty}^{\infty} s \cdot f(s) ds\right]^2,$$

ahol f az X sűrűségfüggvénye.

7. Nevezetes abszolút folytonos eloszlások

7.1. Egyenletes eloszlás

13.6. Definíció (Egyenletes eloszlás). *Legyenek $a < b$ valós számok. Azt mondjuk, hogy az X valószínűségi változó egyenletes eloszlású az $[a, b]$ intervallumon, ha sűrűségfüggvénye*

$$f(s) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq s \leq b; \\ 0, & \text{különben.} \end{cases}$$

Vegyük észre, hogy $f(s) \geq 0$ minden s-re, és $\int_{-\infty}^{\infty} f(s)ds = \int_a^b \frac{1}{b-a}ds = 1$, így f valóban lehet sűrűségfüggvény, a 17. állítás és az azt követő megjegyzés alapján.

19. Állítás (Az egyenletes eloszlás tulajdonságai). *Legyen az X valószínűségi változó egyenletes eloszlású az $[a, b]$ intervallumon. Ekkor a következők teljesülnek.*

(i) X eloszlásfüggvénye:

$$F(t) = \mathbb{P}(X \leq t) = \begin{cases} 0, & \text{ha } t \leq a; \\ \frac{t-a}{b-a}, & \text{ha } a < t < b; \\ 1, & \text{ha } t \geq b. \end{cases}$$

(ii) Ha $a \leq c \leq d \leq b$, akkor

$$\mathbb{P}(c \leq X \leq d) = \int_c^d f(s)ds = \int_c^d \frac{1}{b-a}ds = \frac{d-c}{b-a}.$$

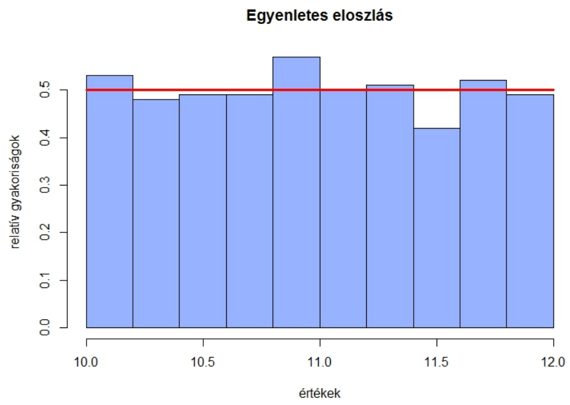
(iii) Az X valószínűségi változó várható értéke és szórása:

$$\mathbb{E}(X) = \frac{a+b}{2}; \qquad D(X) = \frac{b-a}{\sqrt{12}}.$$

Vegyük észre, hogy a várható érték a megadott intervallum közepe, a szórás pedig az intervallum hosszával arányos – hosszabb intervallum esetén nagyobb a szórás.

Példa. Csomagot várunk, a futár 10 és 12 óra között érkezik. Feltesszük, hogy érkezésének időpontja egyenletes eloszlású a $[10, 12]$ intervallumon. Ekkor az előző állítás alapján az alábbiak igazak.

- Annak valószínűsége, hogy 10 és 11 óra között érkezik: $(11-10)/(12-10) = 1/2$.
- Annak valószínűsége, hogy 10:15 és 10:30 között érkezik, $1/8 = 0,125$.
- Érkezési időpontjának várható értéke: $(10+12)/2 = 11$ óra.
- Érkezési időpontjának szórása: $(12-10)/\sqrt{12} = 1/\sqrt{3} = 0,5774$.



10. ábra. $U(10,12)$ sűrűségfüggvénye és 500 elemű minta hisztogramja.

7.2. Normális eloszlás

13.7. Definíció (Normális eloszlás). *Legyen m valós, σ pedig pozitív szám. Azt mondjuk, hogy az Y valószínűségi változó normális eloszlású m várható értékkel és σ^2 szórásnégyzettel, ha sűrűségfüggvénye*

$$f(s) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(s-m)^2}{2\sigma^2} \right) \qquad (s \in \mathbb{R}).$$

Jelölése: $Y \sim N(m, \sigma^2)$.

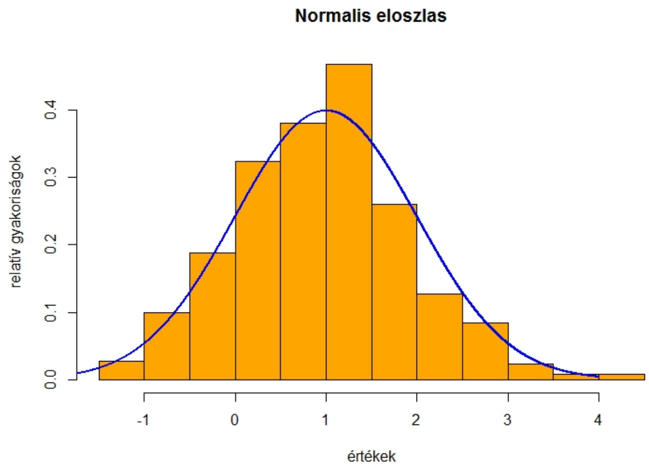
Ez az f valóban sűrűségfüggvény, és hogy az így megadott Y-ra $\mathbb{E}(Y) = m$ és $D^2(Y) = \sigma^2$ (ezeket nem bizonyítjuk).

13.8. Definíció (Standard normális eloszlás). *Azt mondjuk, hogy a Z valószínűségi változó standard normális eloszlású, ha normális eloszlású $m = 0$ várható értékkel és $\sigma^2 = 1$ szórásnégyzettel, azaz $Z \sim N(0,1)$, és sűrűségfüggvénye:*

$$f(s) = \frac{1}{\sqrt{2\pi}} e^{-s^2/2} \qquad (s \in \mathbb{R}).$$

A standard normális eloszlás eloszlásfüggvényét Φ-vel jelöljük, vagyis

$$\Phi(t) = \mathbb{P}(Y \le t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-s^2/2} ds.$$



11. ábra.
A standard normális eloszlás sűrűségfüggvénye és 500 elemű minta hisztogramja

20. Állítás (A normális eloszlás tulajdonságai). *Ha az Y valószínűségi változó normális eloszlású m várható értékkel és σ^2 szórásnégyzettel, valamint $a < b$ valós számok, akkor*

(i)

$$\mathbb{P}(a < Y < b) = \mathbb{P}(a \leq Y \leq b) = \\ = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b \exp\left(- \frac{(s-m)^2}{2\sigma^2} \right) ds = \\ = \Phi\left(\frac{b-m}{\sigma} \right) - \Phi\left(\frac{a-m}{\sigma} \right).$$

(ii) $\mathbb{P}(Y < b) = \mathbb{P}(Y \leq b) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^b \exp\left(- \frac{(s-m)^2}{2\sigma^2} \right) ds = \Phi\left(\frac{b-m}{\sigma} \right).$

(iii) $\mathbb{P}(a < Y) = \mathbb{P}(a \leq Y) = \frac{1}{\sqrt{2\pi}\sigma} \int_a^\infty \exp\left(- \frac{(s-m)^2}{2\sigma^2} \right) ds = 1 - \Phi\left(\frac{a-m}{\sigma} \right).$

(iv) *Y várható értéke m, szórása σ.*

(v) *Legyenek u, v valós számok. Ekkor $uY+v$ is normális eloszlású valószínűségi változó, $um+v$ várható értékkel és $u^2\sigma^2$ szórásnégyzettel.*

Példa. Tegyük fel, hogy a holnapi csúcshőmérséklet (Celsius-fokban mérve) normális eloszlású $m = 2$ várható értékkel és $\sigma^2 = 9$ szórásnégyzettel. Jelöljük ezt a valószínűségi változót Y-nal. Ekkor annak valószínűsége, hogy a holnapi csúcshőmérséklet 0 és 4 °C között lesz:

$$\begin{aligned}
\mathbb{P}(0 < Y < 4) = \mathbb{P}(Y < 4) - \mathbb{P}(Y < 0) &= \Phi\left(\frac{4-2}{3}\right) - \Phi\left(-\frac{2}{3}\right) = \\
&= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{2}{3}\right) = \Phi\left(\frac{2}{3}\right) - \left(1 - \Phi\left(\frac{2}{3}\right)\right) \\
&= 2\Phi\left(\frac{2}{3}\right) - 1 = 2 \cdot 0,7486 - 1 = 0,4972,
\end{aligned}$$

felhasználva a normális eloszlás táblázatát (vagy bármilyen statisztikai szoftvert) és a

$$\Phi(-x) = 1 - \Phi(x)$$

tulajdonságot (mely a standard normális eloszlás sűrűségfüggvényének 0-ra vonatkozó szimmetriájából adódik).

7.3. Exponenciális eloszlás

13.9. Definíció (Exponenciális eloszlás). *Legyen $b > 0$ pozitív szám. Azt mondjuk, hogy az X valószínűségi változó exponenciális eloszlású b paraméterrel, ha sűrűségfüggvénye:*

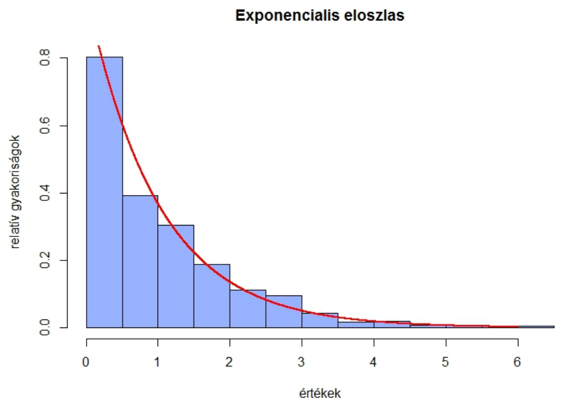
$$f(s) = \begin{cases} be^{-bs}, & \text{ha } s > 0; \\ 0 & \text{különben.} \end{cases}$$

21. Állítás (Az exponenciális eloszlás tulajdonságai). *Legyen X exponenciális eloszlású $b > 0$ paraméterrel. Ekkor a következők teljesülnek.*

(i) X eloszlásfüggvénye:

$$F(t) = \mathbb{P}(X \leq t) = \mathbb{P}(X < t) = \int_{-\infty}^{t} f(s)ds = \begin{cases} 1 - e^{-bs}, & \text{ha } s > 0; \\ 0 & \text{különben.} \end{cases}$$

(ii) X várható értéke: $\mathbb{E}(X) = 1/b$, szórása: $D(X) = 1/b$.



12. ábra. Exp(1) sűrűségfüggvénye és 500 elemű minta hisztogramja.

Példa. Tegyük fel, hogy egy radioaktív részecske élettartamának elolszása (másodpercben mérve) 10 paraméterű exponenciális eloszlású valószínűségi változó. Jelöljük ezt X-szel. Ekkor annak valószínűsége, hogy a részecske több mint 1 másodpercig életben marad:

$$\mathbb{P}(X > 1) = 1 - F(1) = 1 - (1 - e^{-10\cdot 1}) = e^{-10} = 0,0000454.$$

Annak valószínűsége, hogy a részecske élettartama $0,1$ és $0,2$ másodperc közé esik:

$$\begin{aligned}\mathbb{P}(0,1 \leq X \leq 0,2) = F(0,2) - F(0,1) = 1 - e^{-10\cdot 0,2} - (1 - e^{-10\cdot 0,1}) = \\ = 1/e - 1/e^2 = 0,2325.\end{aligned}$$

A részecske élettartamának várható értéke: $\mathbb{E}(X) = 1/10 = 0,1$.

8. Feltételes valószínűség

Kérdés. Valakiről annyit tudunk, hogy három gyermeke van. Mennyi a valószínűsége, hogy a középső gyermeke fiú? És ha elárulja, hogy pontosan egy fia van, mennyi annak valószínűsége, hogy a középső gyermeke fiú?

Kérdés. Mennyi annak valószínűsége, hogy egy véletlenszerűen kiválasztott embernek magas a vérnyomása? Mennyi annak valószínűsége, hogy magas a vérnyomása, ha megmérjük a testtömegét, és megtudjuk róla, hogy túlsúlyos (de nem elhízott)?

13.10. Definíció (Feltételes valószínűség). *Legyenek $A, B \in \mathcal{A}$ események, és tegyük fel, hogy $\mathbb{P}(B) > 0$. Az A esemény B-re vonatkozó feltételes valószínűsége:*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Az első esetben: A: a középső gyerek fiú; B: pontosan egy fiú van a három gyerek között. Ekkor $A \cap B$ azt jelenti, hogy a középső gyerek fiú, és pontosan egy fiú van, vagyis a másik kettő lány. Mindebből, az egyszerűség kedvéért feltételezve, hogy a fiúk és lányok aránya népességben megegyezik, és a testvérek neme között nincs összefüggés (vagyis mind a nyolc lehetőség: FFF, FFL, FLF, ..., LLL egyformán valószínű):

$$\mathbb{P}(A) = \mathbb{P}(\text{a középső fiú}) = \frac{1}{2}.$$

Ugyanakkor a feltételes valószínűséget így számolhatjuk ki:

$$\mathbb{P}(\text{a középső fiú}|\text{egy fiú van}) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(LFL)}{\mathbb{P}(FLL, LFL, LLF)} = \frac{\frac{1}{8}}{\frac{3}{8}} = 1/3.$$

A második esetben (nagyjából valós adatokkal): a magyar lakosság 25 %-ának van magas vérnyomása (A esemény). A magyar lakosság 30%-a túlsúlyos (de nem elhízott), ez a B esemény. Azok aránya, akik túlsúlyosak (de nem elhízottak) és magasvérnyomás-betegségben is szenvednek, 15 %. Tehát:

$$\mathbb{P}(\text{magas vérnyomás}) = 0,25.$$

$$\mathbb{P}(\text{magas vérnyomás}|\text{túlsúly}) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0,15}{0,3} = 0,5.$$

Példa feltételes valószínűségre

Holnap $0,4$ valószínűséggel nem lesz csapadék, $0,15$ valószínűséggel lesz eső és hó is, $0,25$ valószínűséggel csak eső, $0,2$ valószínűséggel pedig csak hó. Feltéve, hogy lesz eső, mennyi a valószínűsége, hogy havazni is fog?

A esemény: havazni fog. B esemény: lesz holnap eső. Ekkor $\mathbb{P}(B) = 0,15 + 0,25 = 0,4$. Másrészt $A \cap B$ azt jelenti, hogy eső és hó is lesz. Így:

$$\mathbb{P}(\text{havazni fog}|\text{lesz eső}) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0,15}{0,4} = 0,375.$$

3. Megjegyzés. *$0 \le \mathbb{P}(A|B) \le 1$ teljesül minden $A, B \in \mathcal{A}$ eseményre, ha $\mathbb{P}(B) > 0$. Szintén ha $\mathbb{P}(B) > 0$: az, hogy A és B függetlenek, ugyanakkor teljesül, mint hogy $\mathbb{P}(A|B) = \mathbb{P}(A)$.*

8.1. Teljes valószínűség tétele

13.11. Definíció (Teljes eseményrendszer). *A $B_1, B_2, \ldots \in \mathcal{A}$ (véges vagy megszámlálható sok) esemény együttesét teljes eseményrendszernek nevezzük, ha*

(i) $\bigcup_{i=1}^{\infty} B_i = \Omega$;

(ii) $B_i \cap B_j = \emptyset$ *teljesül minden* $1 \le i < j$*-re;*

(iii) $\mathbb{P}(B_i) > 0$ *minden* $i = 1, 2, \ldots$*-re.*

Példa.

B_1: holnap nem lesz csapadék;

B_2: holnap lesz csapadék, de nem több 5 mm-nél;

B_3: a holnapi csapadékmennyiség több mint 5 mm, de kevesebb mint 10 mm;

B_4: a holnapi csapadékmennyiség meghaladja a 10 mm-t.

Ekkor B_1, B_2, B_3, B_4 teljes eseményrendszer (közülük pontosan az egyik következik be).

2. Tétel (Teljes valószínűség tétele). *Legyen $A \in \mathcal{A}$ tetszőleges esemény, $B_1, B_2, \ldots$ pedig teljes eseményrendszer. Ekkor*

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Bizonyítás.

1. lépés: Felhasználva, hogy a 13.11. definíció (i) része szerint $\bigcup_{i=1}^{\infty} B_i = \Omega$:

$$\mathbb{P}(A) = \mathbb{P}\Bigg(A \cap \Big(\bigcup_{i=1}^{\infty} B_i\Big)\Bigg) = \mathbb{P}\Bigg(\bigcup_{i=1}^{\infty}(A \cap B_i\Bigg).$$

2. lépés: Az $A \cap B_i$, $i = 1, 2, \ldots$ események páronként kizáróak (vagyis semelyik kettő nem következhet be egyszerre). Ugyanis

$$(A \cap B_i) \cap (A \cap B_j) \subseteq B_i \cap B_j = \emptyset \qquad (1 \leq i < j),$$

hiszen a 13.11. definíció (ii) része szerint a B_i események is páronként kizáróak.

3. lépés: A valószínűség (ii) tulajdonsága (az 1. definíció) megszámlálható sok páronként kizáró esemény uniójának valószínűsége a valószínűségek összege, így az 1. lépést folytatva, majd felhasználva a feltételes valószínűség definícióját (a 13.10. definíció):

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^{\infty}(A \cap B_i\right) = \sum_{i=1}^{\infty}\mathbb{P}(A \cap B_i) = \sum_{i=1}^{\infty}\mathbb{P}(A|B_i)\mathbb{P}(B_i). \qquad \square$$

Példa. Hanna sátorozni megy Badacsonyba. Ha van csapadék, de nem több 5 mm-nél, akkor 15% valószínűséggel ázik be a sátra. Ha több mint 5 mm, de kevesebb mint 10 mm csapadék lesz, akkor 35% valószínűséggel. Végül, ha több mint 10 mm csapadék lesz, akkor 60% valószínűséggel ázik be a sátor. Az előrejelzés szerint 10% valószínűséggel lesz csapadék, de nem több 5 mm-nél, 30% valószínűséggel a holnapi csapadékmennyiség több mint 5 mm, de kevesebb mint 10 mm, és 20% annak valószínűsége, hogy a csapadékmennyiség holnap meghaladja a 10 mm-t. Mennyi annak valószínűsége, hogy holnap beázik Hanna sátra?

Az előző példában már láttuk, hogy az ott felsorolt B_1, B_2, B_3, B_4 események teljes eseményrendszert alkotnak. Legyen az A esemény az, hogy Hannának beázik a sátra. Ezekre alkalmazzuk a teljes valószínűség tételét:

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \mathbb{P}(A|B_3)\mathbb{P}(B_3) + \mathbb{P}(A|B_4)\mathbb{P}(B_4) \\ &= 0 \cdot 0,3 + 0,15 \cdot 0,1 + 0,35 \cdot 0,3 + 0,6 \cdot 0,2 = 0,24 = 24\%.\end{aligned}$$

8.2. Bayes-tétel

3. Tétel (Bayes-tétel). *Legyen $A \in \mathcal{A}$ olyan esemény, melyre $\mathbb{P}(A) > 0$, $B_1, B_2, \ldots$ pedig teljes eseményrendszer. Ekkor minden $k = 1, 2, \ldots$-ra teljesül, hogy*

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{\infty}\mathbb{P}(A|B_i)\mathbb{P}(B_i)}.$$

Bizonyítás. Használjuk a feltételes valószínűség definícióját (13.10. definíció), mindkétszer:

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)}.$$

Mivel A esemény, $B_1, B_2, \ldots$ pedig teljes eseményrendszer, ezután használhatjuk a teljes valószínűség tételét (2. tétel) a $\mathbb{P}(A)$ valószínűség kifejezésére, így megkapjuk az állítást. $\square$

Példa. Másnap Hanna a beázott sátorról küld képeket. Mennyi annak valószínűsége, hogy Badacsonyban több mint 10 mm eső esett ezen a napon?

A kérdés a B_4 esemény (több mint 10 mm eső) valószínűsége, feltéve, hogy az A esemény (a sátor beázik) bekövetkezett, azaz $\mathbb{P}(B_4|A)$. Már láttuk, hogy $\mathbb{P}(A) > 0$, B_1, B_2, B_3, B_4 pedig továbbra is teljes eseményrendszer, így alkalmazhatjuk Bayes tételét:

$$\begin{aligned}\mathbb{P}(B_4|A) &= \frac{\mathbb{P}(A|B_4)\mathbb{P}(B_4)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \ldots + \mathbb{P}(A|B_4)\mathbb{P}(B_4)} = \\ &= \frac{0,6 \cdot 0,2}{0 \cdot 0,3 + 0,15 \cdot 0,1 + 0,35 \cdot 0,3 + 0,6 \cdot 0,2} = \frac{0,12}{0,24} = 0,5.\end{aligned}$$

4. Megjegyzés. *Szorzási szabály: ha $A_1, A_2, \ldots, A_n$ események, és teljesül, hogy $\mathbb{P}(A_1 \cap \ldots \cap A_n) > 0$, akkor*

$$\mathbb{P}(A_1 \cap \ldots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)\ldots\mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1}).$$

22. Állítás. Az exponenciális eloszlás örökifjú tulajdonsága. *Legyen X exponenciális eloszlású valószínűségi változó (bármilyen paraméterrel), és legyenek $s, t > 0$ pozitív számok. Ekkor*

$$\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t).$$

Bizonyítás. Legyen X paramétere $b > 0$. Az exponenciális eloszlás eloszlásfüggvénye, azaz a 21. állítás szerint

$$\mathbb{P}(X > u) = 1 - \mathbb{P}(X \leq u) = 1 - F(u) = 1 - (1 - e^{-bu}) = e^{-bu}.$$

Ezt és a feltételes valószínűség definícióját felhasználva:

$$\mathbb{P}(X > s + t | X > s) = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{e^{-b(s+t)}}{e^{-bs}} = e^{-bt} = \mathbb{P}(X > t). \qquad \square$$

9. Valószínűségi változók együttes eloszlása

13.12. Definíció (Valószínűségi vektorváltozó). *Az $\underline{X} = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ függvény valószínűségi vektorváltozó, ha tetszőleges $a_i < b_i$ $(i = 1, 2, \ldots, n)$ valós számokra teljesül, hogy*

$$\{\omega \in \Omega : a_1 < X_1(\omega) \leq b_1, a_2 < X_2(\omega) \leq b_2, \ldots, a_n < X_n(\omega) \leq b_n\} \in \mathcal{A}.$$

Ha $\underline{X}$ valószínűségi vektorváltozó, akkor az X_i valószínűségi változó eloszlását az $\underline{X}$ i. peremeloszlásának nevezzük.

Az $\underline{X}$ valószínűségi vektorváltozó diszkrét, ha értékkészlete véges vagy megszámlálhatóan végtelen.

9.1. Együttes eloszlás– és sűrűségfüggvény

13.13. Definíció. *Az $\underline{X} = (X_1, \ldots, X_n)$ valószínűségi vektorváltozó együttes eloszlásfüggvénye az $F : \mathbb{R}^n \to [0, 1]$ függvény, melyre*

$$F(\underline{t}) = F(t_1, \ldots, t_n) = \mathbb{P}(X_1 \leq t_1, X_2 \leq t_2, \ldots, X_n \leq t_n), \text{ ha } (t_1, \ldots, t_n) \in \mathbb{R}^n.$$

13.14. Definíció. *Az $\underline{X} = (X_1, \ldots, X_n)$ valószínűségi vektorváltozó abszolút folytonos, ha van olyan $f : \mathbb{R}^n \to \mathbb{R}$ függvény, melyre*

$$F(t_1, \ldots, t_n) = \int_{-\infty}^{t_1} \ldots \int_{-\infty}^{t_n} f(s_1, \ldots, s_n) ds_1 \ldots ds_n.$$

teljesül minden $(t_1, \ldots, t_n) \in \mathbb{R}^n$ esetén. Ilyenkor az f függvényt az $\underline{X}$ együttes sűrűségfüggvényének nevezzük.

9.2. Valószínűségi változók függetlensége

13.15. Definíció. *Azt mondjuk, hogy az $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ valószínűségi változók függetlenek, ha*

$$\mathbb{P}(X_1 \leq t_1, X_2 \leq t_2, \ldots, X_n \leq t_n) = \mathbb{P}(X_1 \leq t_1) \cdot \mathbb{P}(X_2 \leq t_2) \ldots \mathbb{P}(X_n \leq t_n)$$

teljesül tetszőleges $t_1, t_2, \ldots, t_n$ valós számokra.

Végtelen hosszú sorozatra a definíció így módosul. Az $X_1, X_2, X_3 \ldots$ valószínűségi változók függetlenek, ha közülük bármely véges sokat kiválasztva független valószínűségi változókat kapunk.

Független valószínűségi változókra példa:

- Két kockadobásnál az elsőként (X_1) és másodikként dobott szám (X_2).
- A holnapi csapadékmennyiség Budapesten és Torontóban.
- Két találomra választott ember testmagassága.

Nem független valószínűségi változókra példa:

- Két kockadobásnál az első szám és a két dobott szám összege.
- A holnapi csapadékmennyiség Budapesten és Budaörsön.
- Két testvér testmagassága.

23. Állítás. *Tegyük fel, hogy az $X_1, \dots, X_n$ valószínűségi változók függetlenek, és $g_1, \dots, g_n : \mathbb{R} \to \mathbb{R}$ függvények. Ekkor a $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ valószínűségi változók is függetlenek. Továbbá, ha $h : \mathbb{R}^k \to \mathbb{R}$, akkor a $h(X_1, \dots, X_k), X_{k+1}, \dots, X_n$ valószínűségi változók is függetlenek.*

24. Állítás (Függetlenség és sűrűségfüggvény). *Tegyük fel, hogy az $\underline{X} = (X_1, \dots, X_n)$ valószínűségi vektorváltozó abszolút folytonos, együttes sűrűségfüggvénye f, továbbá az X_i valószínűségi változó sűrűségfüggvénye f_i minden $i = 1, 2, \dots, n$ esetén. Ezekkel a jelölésekkel: $X_1, X_2, \dots, X_n$ pontosan akkor függetlenek, ha*

$$f(t_1, \dots, t_n) = f_1(t_1) \cdot f_2(t_2) \dots f_n(t_n)$$

teljesül bármely $t_1, t_2, \dots, t_n$ valós számokra.

25. Állítás (Függetlenség egész értékű esetben). *Tegyük fel, hogy az $\underline{X} = (X_1, \dots, X_n)$ valószínűségi vektorváltozó diszkrét, sőt minden peremeloszlása egész értékű. Az $X_1, X_2, \dots, X_n$ valószínűségi változók pontosan akkor függetlenek, ha*

$$\mathbb{P}(X_1 = k_1, \dots, X_n = k_n) = \mathbb{P}(X_1 = k_1) \cdot \mathbb{P}(X_2 = k_2) \dots \mathbb{P}(X_n = k_n)$$

teljesül bármely $k_1, k_2, \dots, k_n$ egész számokra.

Példa. Egy szabályos dobókockát feldobunk n-szer, jelölje X_i az i. dobás értékét ($i = 1, \ldots, n$). Ekkor $(X_1, X_2, \ldots, X_n)$ függetlenek. Például

$$\mathbb{P}(X_1 = 3, X_2 = 4) = \mathbb{P}(X_1 = 3) \cdot \mathbb{P}(X_2 = 4) = \frac{1}{36};$$
$$\mathbb{P}(X_1 = 3, X_2 = 4, X_3 = 2) = \mathbb{P}(X_1 = 3) \cdot \mathbb{P}(X_2 = 4) \cdot \mathbb{P}(X_3 = 2) = \frac{1}{6^3}.$$

Vegyük észre, hogy ez összhangban van azzal, hogy az első két dobás eredménye 36, a az első három dobás eredménye $6^3 = 216$-féle dobássorozat lehet.

Példa. Tegyük fel, hogy az X és Y valószínűségi változók függetlenek, és a következőket tudjuk:

$$\mathbb{P}(X = 0) = 1/2; \quad \mathbb{P}(X = 2) = 1/2; \quad \mathbb{P}(Y = 1) = 1/3; \quad \mathbb{P}(Y = 4) = 2/3.$$

Ekkor a függetlenségből következik, hogy

$$\mathbb{P}(X = 0, Y = 4) = \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 4) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

10. Várható érték, szórás, kovariancia, korreláció

10.1. A várható érték tulajdonságai

26. Állítás. *Legyenek $X, Y, X_1, X_2, \ldots, X_n$ olyan valószínűségi változók, melyeknek várható értéke létezik vagy a 7.5. definíció, vagy a 13.4. definíció értelmében. Ekkor a következők teljesülnek.*

(a) *Ha $a \le X \le b$ teljesül 1 valószínűséggel valamely a, b valós számokra, akkor $a \le \mathbb{E}(X) \le b$.*

(b) **Konstanssal szorzás.** *Ha $c \in \mathbb{R}$, akkor*

$$\mathbb{E}(c \cdot X) = c \cdot \mathbb{E}(X).$$

(c) **Összeg várható értéke.** $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

(d) **Szorzat várható értéke független esetben.** *Ha az X és Y valószínűségi változók függetlenek, akkor $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$.*

(e) *Legyen $g : \mathbb{R}^n \to \mathbb{R}$ függvény. Ekkor, ha $X_1, \ldots, X_n$ diszkrét valószínűségi változók, akkor*

$$\mathbb{E}(g(X_1, \ldots, X_n)) = \sum_{(t_1,\ldots,t_n)} g(t_1, \ldots, t_n)\mathbb{P}(X_1 = t_1, \ldots, X_n = t_n),$$

ha a jobb oldal abszolút konvergens, és az összegzés az $X_1, X_2, \ldots, X_n$ lehetséges értékeire történik. Ha $X_1, \ldots, X_n$ abszolút folytonos valószínűségi változók f együttes sűrűségfüggvénnyel, akkor

$$\mathbb{E}(g(X_1, \ldots, X_n)) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(t_1, \ldots, t_n) f(t_1, \ldots, t_n) dt_1 \ldots dt_n,$$

ha a jobb oldal abszolút konvergens.

Gyakran előfordul, hogy függvényként a k. hatványra emelést használjuk.

13.16. Definíció (Momentumok). *Legyen X olyan diszkrét vagy abszolút folytonos valószínűségi változó, melyre X^k várható értéke létezik. Ekkor az X valószínűségi változó k. momentuma:*

$$\mathbb{E}(X^k) \qquad (k = 1, 2, \ldots).$$

Az állítás (e) része alapján, ha X diszkrét valószínűségi változó, akkor

$$\mathbb{E}(X^k) = \sum_{i=1}^{\infty} u_i^k \mathbb{P}(X = u_i),$$

ahol $u_1, u_2, \ldots$ az X lehetséges értékei. Ha X abszolút folytonos, és sűrűségfüggvénye f, akkor viszont így számolhatjuk a k. momentumot:

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} t^k f(t) dt.$$

10.2. A szórásnégyzet tulajdonságai

27. Állítás. *Legyenek $X, Y, X_1, X_2, \ldots, X_n$ olyan valószínűségi változók, melyeknek szórása létezik. Ekkor a következők teljesülnek.*

(a) **Nemnegativitás.** $D(X) \geq 0$.

(b) *$D^2(X) = 0$ akkor és csak akkor, ha $\mathbb{P}(X = c) = 1$ valamilyen $c \in \mathbb{R}$ számra.*

(c) **Konstans hozzáadása** *$D^2(X + b) = D^2(X)$ tetszőleges $b \in \mathbb{R}$ számra.*

(d) **Konstanssal való szorzás.** *$D^2(a \cdot X) = a^2 D^2(X)$, és $D(a \cdot X) = |a|D(X)$ tetszőleges $a \in \mathbb{R}$ számra.*

(e) **Összeg szórásnégyzete független esetben.** *Ha X és Y függetlenek, akkor $D^2(X+Y) = D^2(X)+D^2(Y)$. Általánosabban, ha $X_1, X_2, \ldots, X_n$ függetlenek, akkor $D^2(X_1 + \ldots + X_n) = D^2(X_1) + \ldots + D^2(X_n)$.*

Példa. Tegyük fel, hogy X és Y független normális eloszlású valószínűségi változók, és X várható értéke 3, szórása 2, Y várható értéke 10, szórása 5. Ekkor az összegük várható értéke és szórásnégyzete:

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 13; \qquad D^2(X + Y) = D^2(X) + D^2(Y) = 29.$$

A különbségük várható értéke és szórása:

$$\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y) = -7; \qquad D^2(X - Y) = D^2(X) + D^2(Y) = 29.$$

A $2X + 3Y$ valószínűségi változó várható értéke és szórásnégyzete:

$$\mathbb{E}(2X+3Y) = 2\mathbb{E}(X)+3\mathbb{E}(Y) = 36; \qquad D^2(2X+3Y) = 4D^2(X)+9D^2(Y) = 241.$$

Példa: Az átlag várható értéke és szórása

28. Állítás. *Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független, azonos eloszlású valószínűségi változók (vagyis eloszlásfüggvényük megegyezik). Ekkor*

$$\mathbb{E}(X_1 + \ldots + X_n) = n\mathbb{E}(X_1); \qquad D(X_1 + \ldots + X_n) = \sqrt{n}D(X_1).$$

29. Állítás. *Legyenek $X_1, X_2, \ldots, X_n$ független, azonos eloszlású valószínűségi változók. Ekkor*

$$\mathbb{E}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \mathbb{E}(X_1); \qquad D\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{D(X_1)}{\sqrt{n}}.$$

10.3. A kovariancia és tulajdonságai

13.17. Definíció (kovariancia). *Legyenek X és Y olyan valószínűségi változók, melyeknek szórása létezik. Ekkor az X és Y kovarianciája:*

$$\operatorname{cov}(X,Y) = \mathbb{E}\big[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))\big].$$

30. Állítás. *Legyenek $X, Y, Z, X_1, \ldots, X_n$ olyan valószínűségi változók, melyek szórása létezik. Ekkor a következők teljesülnek.*

(a) **A kovariancia kiszámítása.** $\operatorname{cov}(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y)$.

(b) **Szimmetria.** $\operatorname{cov}(X,Y) = cov(Y,X)$.

(c) **Konstanssal való kovariancia.** $\operatorname{cov}(X,c) = 0$, *ha* $c \in \mathbb{R}$.

(d) **Kapcsolat a szórásnégyzettel.** $\operatorname{cov}(X,X) = D^2(X)$.

(e) **Linearitás.** *Egyrészt* $\operatorname{cov}(X+Y,Z) = \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z)$, *másrészt tetszőleges* $c \in \mathbb{R}$ *számra* $\operatorname{cov}(cX,Y) = c \cdot \operatorname{cov}(X,Y)$.

(f) **Függetlenséggel való kapcsolat.** *Ha az X és Y valószínűségi változók függetlenek, akkor* $\operatorname{cov}(X,Y) = 0$.

(g) **Összeg szórásnégyzete.** $D^2(X+Y) = D^2(X) + D^2(Y) + 2\operatorname{cov}(X,Y)$. *Továbbá*

$$D^2\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} D^2(X_i) + 2\sum_{i<j} \operatorname{cov}(X,Y).$$

(h) **Különbség szórásnégyzete** $D^2(X-Y) = D^2(X) + D^2(Y)$.

Példa. Legyen X Poisson-eloszlású valószínűségi változó 2 paraméterrel. Ekkor

$$\operatorname{cov}(X+3, 2\cdot X) \overset{(e)}{=} 2\operatorname{cov}(X+3,X) \overset{(e)}{=} 2\operatorname{cov}(X,X) + 2\operatorname{cov}(3,X) =$$
$$\overset{(c,d)}{=} 2D^2(X) = 2 \cdot 2 = 4,$$

ahol felhasználtuk a Poisson-eloszlás szórásnégyzetére vonatkozó összefüggést (10. állítás (a) rész).

Példa. Legyenek X és Y független standard normális eloszlású valószínűségi változók. Ekkor

$$\operatorname{cov}(2X-5Y, X+Y) \overset{(e)}{=} 2\operatorname{cov}(X,X) + 2\operatorname{cov}(X,Y) - 5\operatorname{cov}(X,Y) - 5\operatorname{cov}(Y,Y)$$
$$\overset{(c,d)}{=} 2D^2(X) - 5D^2(Y) = -3.$$

13.18. Definíció (Korrelálatlanság). *Ha az X, Y valószínűségi változók kovarianciája 0, akkor azt mondjuk, hogy X és Y korrelálatlanok.*

31. Állítás (Függetlenség és korrelálatlanság). *(i) Ha az X és Y valószínűségi változók függetlenek és szórásuk létezik, akkor korrelálatlanok, azaz* $\mathrm{cov}(X, Y) = 0$.

(ii) Vannak olyan U, V valószínűségi változók, melyek nem függetlenek, de korrelálatlanok, azaz $\mathrm{cov}(U, V) = 0$.

Az állítás (ii) részére példa a következő. Legyen X és Y két szabályos kockadobás, ezek függetlenek. Legyen továbbá $U = X + Y, V = X - Y$. Ekkor

$$\begin{aligned}\mathrm{cov}(U, V) &= \mathrm{cov}(X + Y, X - Y) = \\ &\overset{(e,d)}{=} D^2(X) - \mathrm{cov}(X, Y) + \mathrm{cov}(X, Y) - D^2(X) = \\ &\overset{(f)}{=} D^2(X) - D^2(Y) = 0.\end{aligned}$$

Ugyanakkor U és V nem függetlenek, például

$$0 = \mathbb{P}(U = 11, V = 0) \neq \mathbb{P}(U = 11) \cdot \mathbb{P}(V = 0) = \frac{3}{36} \cdot \frac{1}{6}.$$

10.4. A korrelációs együttható

13.19. Definíció (Korrelációs együttható). *Legyenek X és Y olyan valószínűségi változók, melyek szórásnégyzete létezik. Ekkor X és Y korrelációs együtthatója:*

$$R(X, Y) = \begin{cases} \frac{\mathrm{cov}(X,Y)}{D(X)D(Y)}, & \text{ha } D(X) > 0, D(Y) > 0; \\ 0, & \text{ha } D(X) = 0 \text{ vagy } D(Y) = 0. \end{cases}$$

32. Állítás (A korrelációs együttható tulajdonságai). *Legyenek X és Y olyan valószínűségi változók, melyek szórása létezik.*

(i) Ekkor teljesül, hogy
$$|R(X, Y)| \leq 1.$$

(ii) Legyen $a > 0$ valós szám, b tetszőleges valós szám. Ekkor

$$R(X, aX + b) = 1 \text{ és } R(X, -aX + b) = -1.$$

(iii) Tegyük fel, hogy $|R(X,Y)| = 1$. Ekkor léteznek olyan a és b valós számok, hogy az $Y = aX + b$ egyenlet 1 valószínűséggel teljesül.

Példa. Kétszer dobunk szabályos dobókockával. Kiszámítjuk az először dobott számnak és a dobott számok összegének korrelációs együtthatóját.

Ehhez legyen X az először dobott szám, Y a másodszor dobott szám. A kérdés X és $X+Y$ korrelációs együtthatója. Felhasználva, hogy X és Y függetlenek, és így egyrészt a kovarianciájuk nulla, másrészt az összegük szórásnégyzete a szórásnégyzeteik összege (27. állítás):

$$\begin{aligned} R(X, X+Y) &= \frac{\operatorname{cov}(X, X+Y)}{D(X)D(X+Y)} = \frac{\operatorname{cov}(X,X) + \operatorname{cov}(X,Y)}{D(X)\sqrt{D^2(X) + D^2(Y)}} \\ &= \frac{D^2(X)}{D(X)\cdot\sqrt{2}\cdot D(X)} = \frac{1}{\sqrt{2}} > 0. \end{aligned}$$

14. Definíció (Standardizálás.). *Legyen X valószínűségi változó, melynek szórása létezik. Ekkor X standardizáltjának az*

$$X^* = \frac{X - \mathbb{E}(X)}{D(X)}$$

valószínűségi változót nevezzük.

5. Megjegyzés. *A standardizált várható értéke 0, a szórása 1, és az is teljesül, hogy $R(X,Y) = R(X^*, Y^*)$.*

11. A nagy számok törvényei

14.1. Definíció (Sztochasztikus konvergencia). *Legyen $X_1, X_2, \ldots$ valószínűségi változók sorozata. Azt mondjuk, hogy ez a sorozat sztochasztikusan konvergál az Y valószínűségi változóhoz, ha minden $\varepsilon > 0$-ra*

$$\mathbb{P}(|X_n - Y| > \varepsilon) \to 0$$

teljesül $n \to \infty$ esetén.

15. Definíció. *Elnevezés: azt mondjuk, hogy az X és Y valószínűségi változók azonos eloszlásúak, ha eloszlásfüggvényük megegyezik, azaz minden $t \in \mathbb{R}$-re $\mathbb{P}(X \le t) = \mathbb{P}(Y \le t)$.*

4. Tétel (A nagy számok gyenge törvénye). *Legyenek $X_1, X_2, \ldots$ olyan valószínűségi változók, melyek függetlenek, és azonos eloszlásúak (vagyis eloszlásfüggvényük megegyezik). Tegyük fel, hogy $D(X_1)$ létezik. Ekkor*

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to \mathbb{E}(X_1)$$

sztochasztikusan $n \to \infty$ esetén.

Ez tehát azt jelenti, hogy ha $\overline{X}_n = (X_1 + X_2 + \ldots + X_n)/n$ jelöli az első n valószínűségi változó átlagát, akkor minden $\varepsilon > 0$ esetén

$$\mathbb{P}(|\overline{X}_n - \mathbb{E}(X_1)| > \varepsilon) \to 0 \qquad (n \to \infty).$$

15.1. Definíció (1 valószínűségű konvergencia). *Legyen $X_1, X_2, \ldots$ valószínűségi változók sorozata. Azt mondjuk, hogy ez a sorozat 1 valószínűséggel konvergál a Z valószínűségi változóhoz, ha*

$$\mathbb{P}(\{\omega \in \Omega : X_n(\omega) \to Z(\omega)\}) = 1.$$

6. Megjegyzés. *Ha $X_n \to Z$ teljesül 1 valószínűséggel, akkor $X_n \to Z$ sztochasztikusan is. Ennek a megfordítása viszont nem igaz (van olyan sztochasztikusan konvergens sorozat, mely nem 1 valószínűséggel konvergens).*

5. Tétel (A nagy számok erős törvénye). *Legyenek $X_1, X_2, \ldots$ valószínűségi változók, melyek függetlenek és azonos eloszlásúak. Tegyük fel még, hogy $\mathbb{E}(X_1)$ létezik. Ekkor*

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to \mathbb{E}(X_1)$$

teljesül 1 valószínűséggel $n \to \infty$ esetén.

11.1. Egyenlőtlenségek

33. Állítás (Markov-egyenlőtlenség). *Legyen $t > 0$ tetszőleges pozitív szám, X pedig olyan véges várható értékű valószínűségi változó, mely csak nemnegatív értékeket vesz fel, vagyis melyre $X \geq 0$ teljesül. Ekkor*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

34. Állítás (Csebisev-egyenlőtlenség). *Legyen X véges szórású valószínűségi változó, $s > 0$ pozitív szám. Ekkor*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq s) \leq \frac{D^2(X)}{s^2}.$$

35. Állítás. *Legyen X véges szórású valószínűségi változó, $s > 0$ pozitív szám. Ekkor*

$$\mathbb{P}(|X - \mathbb{E}(X)| < s) \geq 1 - \frac{D^2(X)}{s^2}.$$

12. Centrális határeloszlástétel

15.2. Definíció (Eloszlásbeli konvergencia). *Legyen $X_1, X_2, \ldots$ valószínűségi változók sorozata, X_i eloszlásfüggvénye F_i ($i = 1, 2, \ldots$ esetén). Legyen továbbá Y valószínűségi változó, melynek eloszlásfüggvénye F. Azt mondjuk, hogy az $(X_n)_{n \in \mathbb{N}}$ sorozat tart Y-hoz eloszlásban, ha*

$$F_n(t) \to F(t) \qquad (n \to \infty)$$

teljesül minden olyan $t \in \mathbb{R}$-re, melyre F folytonos t-ben.

6. Tétel (Centrális határeloszlástétel). *Legyenek $X_1, X_2, \ldots$ független azonos eloszlású valószínűségi változók, melyeknek szórása létezik. Használjuk a következő jelöléseket: $\mathbb{E}(X_1) = m$ és $D(X_1) = s$. Legyen Y standard normális valószínűségi változó: $Y \sim N(0, 1)$. Ekkor*

$$\frac{X_1 + X_2 + \ldots + X_n - n \cdot m}{s\sqrt{n}} \to Y$$

eloszlásban $n \to \infty$ esetén. Vagyis tetszőleges $a < b$ valós számokra teljesül, hogy

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq \frac{X_1 + X_2 + \ldots + X_n - n \cdot m}{s\sqrt{n}} < b\right) = \mathbb{P}(a \leq Y < b),$$

azaz (mivel Y standard normális eloszlású):

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq \frac{X_1 + X_2 + \ldots + X_n - n \cdot m}{s\sqrt{n}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Ez utóbbi összefüggést így is átfogalmazhatjuk:

$$\mathbb{P}(nm + as\sqrt{n} \leq X_1 + X_2 + \ldots + X_n < nm + bs\sqrt{n}) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$$

teljesül $n \to \infty$ esetén.

Példa. Legyenek $X_1, X_2, \ldots$ független $s = 0,5$ paraméterű exponenciális eloszlású valószínűségi változók. Ezek azonos eloszlásúak, véges szórásúak, és

$$\mathbb{E}(X_1) = \frac{1}{s} = 2; \qquad D(X_1) = \frac{1}{s} = 2.$$

A nagy számok erős törvénye szerint

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = 2$$

1 valószínűséggel.

A centrális határeloszlástétel szerint tetszőleges $a < b$ számokra

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq \frac{\sum_{j=1}^{n} X_j - 2n}{2\sqrt{n}} < b \right) = \int_a^b e^{-x^2/2} dx = \Phi(b) - \Phi(a).$$

Átrendezve:

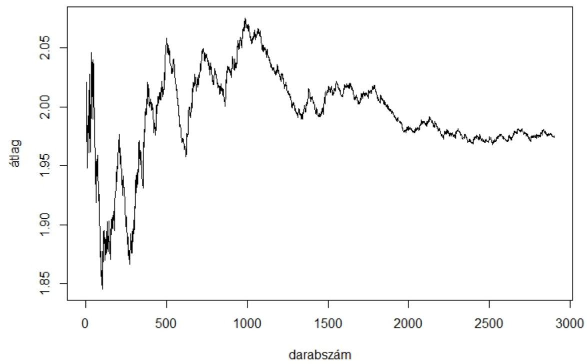
$$\lim_{n\to\infty} \mathbb{P}\left(2a\sqrt{n} \leq \sum_{j=1}^{n} X_j - 2n < 2b\sqrt{n} \right) = \int_a^b e^{-x^2/2} dx = \Phi(b) - \Phi(a).$$

$$\lim_{n\to\infty} \mathbb{P}\left(2n + 2a\sqrt{n} \leq \sum_{j=1}^{n} X_j < 2n + 2b\sqrt{n} \right) = \int_a^b e^{-x^2/2} dx = \Phi(b) - \Phi(a).$$

$$\lim_{n\to\infty} \mathbb{P}\left(2 + \frac{2a}{\sqrt{n}} \leq \frac{\sum_{j=1}^{n} X_j}{n} < 2 + \frac{2b}{\sqrt{n}} \right) = \int_a^b e^{-x^2/2} dx = \Phi(b) - \Phi(a).$$

Tehát például $a = -1$ és $b = 1$ választással

$$\lim_{n\to\infty} \mathbb{P}\left(2 - \frac{2}{\sqrt{n}} \leq \frac{\sum_{j=1}^{n} X_j}{n} < 2 + \frac{2}{\sqrt{n}} \right) = \int_{-1}^{1} e^{-x^2/2} dx = \Phi(1) - \Phi(-1)$$
$$= \Phi(1) - [1 - \Phi(1)] = 2\Phi(1) - 1 = 2 \cdot 0,8413 - 1 = 0,6826.$$



13. ábra.
Az átlag változása a darabszám függvényében független $\exp(0,5)$ eloszlású mintánál

Ugyanakkor például

$$\lim_{n\to\infty} \mathbb{P}\left(1,99993 \le \frac{\sum_{j=1}^{n} X_j}{n} < 2,00008\right) = 1.$$

A 13. ábrán a következő látható. $X_1, X_2, \ldots, X_{3000}$ független exponenciális eloszlású valószínűségi változók $s = 0,5$ paraméterrel, mint előbb. Minden $n = 100, 101, \ldots, 3000$-re kiszámítjuk az $\frac{1}{n}(\sum_{j=1}^{n} X_j)$ átlagot, és ezt ábrázoljuk n függvényében. Az átlag $\mathbb{E}(X_1) = 1/s = 2$-höz konvergál a nagy számok erős törvénye szerint 1 valószínűséggel. Az ábrán az átlag nem megy nagyon közel a kettőhöz, de elég nagy n-re $1,95$ és $2,05$ közé esik. Az átlag szórása $n = 3000$-re $2/\sqrt{3000} = 0,0365$ a 29. állítás szerint.

13. Valószínűségi változók összege

13.1. Konvolúció

36. Állítás. *Legyenek X és Y olyan független valószínűségi változók, melyek lehetséges értékei egész számok. Ekkor*

$$\mathbb{P}(X + Y = k) = \sum_{i=-\infty}^{\infty} \mathbb{P}(X = i)\mathbb{P}(Y = k - i).$$

37. Állítás. *Legyenek X és Y egymástól független, abszolút folytonos valószínűségi változók. Legyen az X sűrűségfüggvénye f, az Y-é pedig g. Ekkor az $X+Y$ valószínűségi változó is abszolút folytonos, és sűrűségfüggvénye:*

$$h_{X+Y}(t) = \int_{-\infty}^{\infty} f(s)g(t-s)ds.$$

13.2. Nevezetes eloszlások összege

38. Állítás (Binomiális eset). *Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független valószínűségi változók, úgy, hogy X_i eloszlása binomiális m_i renddel és p paraméterrel ($i = 1, \ldots, n$). Ekkor $X_1 + \ldots + X_n$ eloszlása binomiális eloszlás $m_1 + m_2 + \ldots + m_n$ renddel és p paraméterrel.*

39. Állítás (Geometriai eset). *Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független valószínűségi változók, úgy, hogy X_i geometriai eloszlású p paraméterrel ($i = 1, \ldots, n$). Ekkor $X_1 + \ldots + X_n$ eloszlása negatív binomiális eloszlás n renddel és p paraméterrel.*

40. Állítás (Negatív binomiális eloszlás). *Legyenek $X_1, \ldots, X_n$ független valószínűségi változók, úgy, hogy X_i negatív binomiális eloszlású m_i renddel és p paraméterrel ($i = 1, \ldots, n$). Ekkor $X_1 + \ldots + X_n$ eloszlása negatív binomiális eloszlás $m_1 + m_2 + \ldots + m_n$ renddel és p paraméterrel.*

41. Állítás (Poisson-eloszlás). *Legyenek $X_1, \ldots, X_n$ független valószínűségi változók, úgy, hogy X_i Poisson-eloszlású s_i paraméterrel ($i = 1, \ldots, n$). Ekkor*

(a) $X_1 + X_2$ Poisson-eloszlású $s_1 + s_2$ paraméterrel.

(b) $X_1 + X_2 + \ldots + X_n$ Poisson-eloszlású $s_1 + \ldots + s_n$ paraméterrel.

42. Állítás (Normális eloszlás). *Legyenek $X_1, \ldots, X_n$ független valószínűségi változók, úgy, hogy X_i normális eloszlású m_i várható értékkel és σ_i^2 szórásnégyzettel ($i = 1, \ldots, n$). Ekkor*

(a) $X_1 + X_2$ normális eloszlású $m_1 + m_2$ várható értékkel és $\sigma_1^2 + \sigma_2^2$ szórásnégyzettel.

(b) $X_1 + X_2 + \ldots + X_n$ normális eloszlású $m_1 + \ldots + m_n$ várható értékkel és $\sigma_1^2 + \ldots + \sigma_n^2$ szórásnégyzettel.

14. További nevezetes abszolút folytonos eloszlások

15.3. Definíció (gamma-függvény). *Ha $a > 0$ pozitív szám, legyen*

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt.$$

Parciális integrálással kiszámítható, hogy $\Gamma(a) = (a-1)\Gamma(a-1)$ minden $a > 1$-re, és így $\Gamma(n) = (n-1)!$, ha n pozitív egész.

15.4. Definíció (gamma-eloszlás). *Legyenek a és λ pozitív számok. Azt mondjuk, hogy az X valószínűségi változó gamma-eloszlású a renddel és λ paraméterrel, ha sűrűségfüggvénye*

$$f(t) = \begin{cases} \frac{\lambda^a t^{a-1}}{\Gamma(a)} e^{-\lambda t}, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

15.5. Definíció (χ^2-eloszlás). *Legyenek $X_1, X_2, \ldots, X_q$ független standard normális eloszlású valószínűségi változók. Az*

$$Y = X_1^2 + X_2^2 + \ldots + X_n^2$$

valószínűségi változó eloszlását q szabadsági fokú χ^2-eloszlásnak nevezzük. Ennek sűrűségfüggvénye:

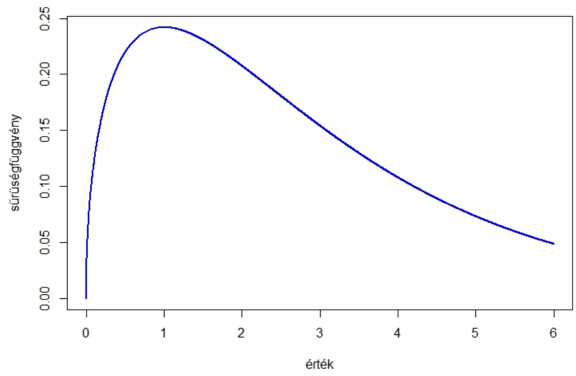
$$f_1(t) = \begin{cases} \frac{t^{q/2-1}}{2^{q/2}\Gamma(q/2)} e^{-t/2}, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

A q szabadsági fokú χ-négyzet eloszlás megegyezik az $a = q/2$ rendű és $\lambda = 1/2$ paraméterű Γ-eloszlással.

15.6. Definíció (F-eloszlás). *Legyenek m, n pozitív egészek, $X_1, \ldots, X_m$, $Y_1, Y_2, \ldots, Y_n$ pedig független standard normális eloszlású valószínűségi változók. Ekkor az*

$$F = \frac{n(X_1^2 + X_2^2 + \ldots + X_m^2)}{m(Y_1^2 + Y_2^2 + \ldots + Y_n^2)}$$

valószínűségi változó eloszlását m, n paraméterű F-eloszlásnak nevezzük.



14. ábra. Az $n = 3$ szabadsági fokú χ^2-eloszlás sűrűségfüggvénye

15.7. Definíció (t-eloszlás). *Legyenek $X_1, X_2, \ldots, X_n$ és Y független standard normális eloszlású valószínűségi változók. Ekkor a*

$$Z = \frac{Y}{\sqrt{(X_1^2 + X_2^2 + \ldots + X_n^2)/n}}$$

valószínűségi változó eloszlását n szabadsági fokú t-eloszlásnak (vagy Student-eloszlásnak) nevezzük.

15.8. Definíció (Cauchy-eloszlás). *Az $n = 1$ szabadsági fokú t-eloszlást Cauchy-eloszlásnak nevezzük. Vagyis, ha X és Y független standard normális eloszlású valószínűségi változók, akkor X/Y Cauchy-eloszlású. Ennek sűrűségfüggvénye:*

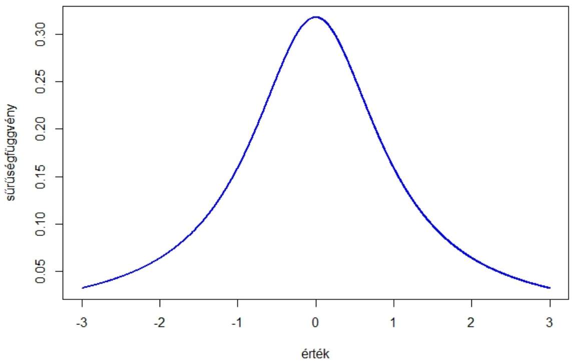
$$f_2(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2}.$$

A Cauchy-eloszlásnak sem várható értéke, sem szórása nem létezik.

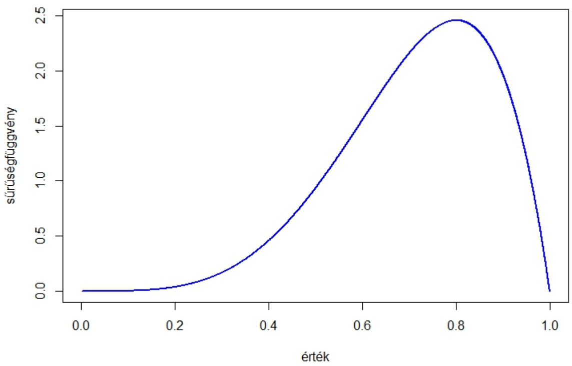
15.9. Definíció (beta-eloszlás). *Legyenek $a, b > 1$ számok. Azt mondjuk, hogy az U valószínűségi változó beta-eloszlású a és b paraméterekkel, ha sűrűségfüggvénye*

$$f_3(t) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^{a-1}(1-t)^{b-1}, & t \in [0, 1]; \\ 0, & t < 0 \text{ vagy } t > 1. \end{cases}$$

Vegyük észre, hogy a sűrűségfüggvény csak a $[0, 1]$ intervallumon vesz fel pozitív értékeket, vagyis a beta-eloszlásból sorsolt értékek mindig 0 és 1 közé esnek.



15. ábra. A Cauchy-eloszlás sűrűségfüggvénye



16. ábra. A beta-eloszlás sűrűségfüggvénye $a = 5$ és $b = 2$ paraméterekkel