

Valószínűségi változók (4. előadás)

események (egy kísérlet eredményéhez **igen vagy nem** tartozik):

- A : holnap lesz csapadék Budapesten
- B : egy véletlenszerűen választott magyar ember Budapesten lakik
- C : egy véletlenszerűen választott magyar ember 500000 forintnál többet keres

valószínűségi változók (egy kísérlet eredményéhez egy **szám** tartozik):

- X : a holnap Budapesten lehulló csapadék mennyisége mm-ben
- Y : egy véletlenszerűen választott magyar ember lakcímének irányítószáma
- Z : egy véletlenszerűen választott magyar ember bruttó havi keresete forintban

Valószínűségi változók: jelölések és definíció

- X : a holnap lehulló csapadék mennyisége mm-ben $\rightarrow \mathbb{P}(X \leq 5)$, azaz mennyi annak valószínűsége, hogy holnap **legfeljebb 5 mm** csapadék esik;
- Y : egy véletlenszerűen választott magyar ember lakcímének irányítószáma $\rightarrow \mathbb{P}(Y = 2092)$, azaz mennyi annak valószínűsége, hogy a kiválasztott ember Budakeszin lakik, **irányítószáma pontosan 2092**
- Z : egy véletlenszerűen választott magyar ember bruttó havi keresete forintban $\rightarrow \mathbb{P}(Z \leq 500000)$, azaz mennyi annak valószínűsége, hogy a kiválasztott ember **legfeljebb bruttó 500000 forintot** keres havonta

Definíció (Valószínűségi változó)

Egy $X : \Omega \rightarrow \mathbb{R}$ függvény **valószínűségi változó**, ha tetszőleges t valós számra teljesül, hogy

$$\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{A},$$

azaz tetszőleges t valós számra a $\mathbb{P}(X \leq t)$ valószínűség értelmes.

Valószínűségi változó: példa

Valakinek három gyereke születik. Legyen X a fiúk száma. Ekkor az összes lehetőség halmaza Ω , és $X : \Omega \rightarrow \mathbb{R}$ függvény az alábbi módon:

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

$$X(LLL) = 0; \quad X(LLF) = X(LFL) = X(FLL) = 1;$$

$$X(FFL) = X(FLF) = X(LFF) = 2; \quad X(FFF) = 3.$$

Az X **valószínűségi változó lehetséges értékei:**

$$\{0, 1, 2, 3\} \rightarrow \text{véges halmaz}$$

A lehetséges értékekhez tartozó valószínűségek, feltéve, hogy a gyerekek egymástól függetlenül $1/2$ valószínűséggel fiúk:

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

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

Diszkrét valószínűségi változó és eloszlása: definíciók

Definíció

Egy $X : \Omega \rightarrow \mathbb{R}$ függvény **valószínűségi változó**, ha tetszőleges t valós számra

$$\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{A},$$

azaz tetszőleges t valós számra a $\mathbb{P}(X \leq t)$ valószínűség értelmes.

Ilyenkor a $\mathbb{P}(X = t)$ valószínűség is tetszőleges t -re értelmes.

Az $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó **diszkrét**, ha **ha lehetséges értékeinek halmaza véges vagy megszámlálhatóan végtelen**.

Diszkrét valószínűségi változó és eloszlása: definíciók

Definíció

Egy $X : \Omega \rightarrow \mathbb{R}$ függvény **valószínűségi változó**, ha tetszőleges t valós számra

$$\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{A},$$

azaz tetszőleges t valós számra a $\mathbb{P}(X \leq t)$ valószínűség értelmes.

Ilyenkor a $\mathbb{P}(X = t)$ valószínűség is tetszőleges t -re értelmes.

Az $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó **diszkrét**, ha **ha lehetséges értékeinek halmaza véges vagy megszámlálhatóan végtelen**.

Legyenek az X **diszkrét valószínűségi változó** lehetséges értékei:

$$\{x_1, x_2, \dots\}, \quad \text{és } p_k = \mathbb{P}(X = x_k) \quad (k = 1, 2, \dots).$$

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

$$p_k \geq 0 \text{ minden } k\text{-ra, és } \sum_{k=1}^{\infty} p_k = 1,$$

azaz a $(p_k)_{k \geq 0}$ sorozat valószínűségeloszlás.

Diszkrét valószínűségi változó eloszlása

Három gyerek. Valakinek három gyereke születik, X a fiúk száma, feltesszük, hogy mind a $2^3 = 8$ lehetőség egyformán valószínű. Ekkor X lehetséges értékei:

$\{0, 1, 2, 3\} \rightarrow$ véges halmaz $\rightarrow X$ **diszkrét**.

Ahogy láttuk:

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

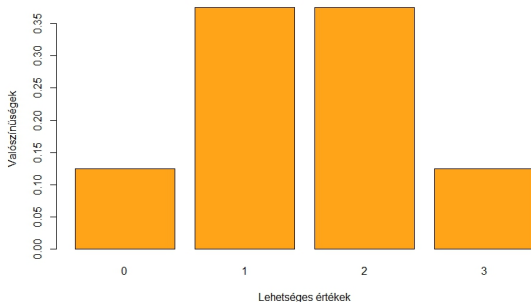
Mindezek alapján X **eloszlása** az alábbi sorozat:

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

Szabályos kockadobás. Egyszer dobunk szabályos dobókockával, jelölje Y a dobott számot. Ekkor Y **diszkrét**, és az **eloszlása**:

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

Példa: a fiúk számának eloszlása



A fiúk számának eloszlása: a lehetséges értékek:

$$\{0, 1, 2, 3\}$$

és a hozzájuk tartozó valószínűségek:

$$1/8, \quad 3/8, \quad 3/8, \quad 1/8.$$

Valószínűségi változó eloszlása

Definíció (Valószínűségi változó)

Egy $X : \Omega \rightarrow \mathbb{R}$ függvény **valószínűségi változó**, ha tetszőleges t valós számra

$$\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{A},$$

vagyis tetszőleges t valós számra a $\mathbb{P}(X \leq t)$ valószínűség értelmes.

Nem feltétlenül diszkrét $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó **eloszlása**: Q_X mérték, melyre

$$Q_X(B) = \mathbb{P}(X \in B),$$

ahol $B \subseteq \mathbb{R}$ megfelelő feltételeket teljesítő halmaz (Borel-halmaz).

Például: $B = [a, b]$ intervallum esetén

$$Q_X([a, b]) = \mathbb{P}(X \in [a, b]) = \mathbb{P}(a \leq X \leq b).$$

Vagy $B = (-\infty, t]$ félegyenes esetén

$$Q_X((-\infty, t]) = \mathbb{P}(X \in (-\infty, t]) = \mathbb{P}(X \leq t).$$

Házi feladat október 2., 10:15-ig

- Hanna minden nap a többiektől függetlenül $1/2$ valószínűséggel 4-es, $1/2$ valószínűséggel hatos villamossal jön egyetemre. Feltéve, hogy egy hónap 20 tanítási napja alatt pontosan nyolcszor jött négyessel, mennyi a valószínűsége, hogy ebben a hónapban az első tanítási napon négyes villamossal érkezett?
- Tegyük fel, hogy a négyes villamos $0,01$ valószínűséggel, a hatos villamos $0,02$ valószínűséggel hibásodik meg a Hanna által megtett úton. Egy adott napon mennyi a valószínűsége, hogy az a villamos, amivel Hanna utazik, meghibásodik?

Házi feladat október 2., 10:15-ig: megoldás

Hanna minden nap a többtől függetlenül $1/2$ valószínűséggel 4-es, $1/2$ valószínűséggel hatos villamossal jön egyetemre. Feltéve, hogy egy hónap 20 tanítási napja alatt pontosan nyolcszor jött négyessel, mennyi a valószínűsége, hogy ebben a hónapban az első tanítási napon négyes villamossal érkezett?

A : Hanna az első tanítási napon négyessel érkezik

B : a 20 nap alatt nyolcszor jön négyessel

Az A -nak B -re vonatkozó feltételes valószínűsége a visszatevéses mintavételre vonatkozó összefüggések alapján:

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{2} \cdot \binom{19}{7} \cdot \left(\frac{1}{2}\right)^{19}}{\binom{20}{8} \left(\frac{1}{2}\right)^{20}} = \\ &= \frac{\binom{19}{7}}{\binom{20}{8}} = \frac{19!}{7! \cdot 12!} \cdot \frac{8! \cdot 12!}{20!} = \frac{8}{20} = 0,4.\end{aligned}$$

A számlálóban: $A \cap B$ esetén az első napon négyessel jön, a többi 19 napon pedig pontosan hétszer jön négyessel, és ezek egymástól függetlenek. A nevezőben: a 20 napból pontosan 8-szor jön négyessel.

Házi feladat október 2., 10:15-ig: megoldás

Tegyük fel, hogy a négyes villamos 0,01 valószínűséggel, a hatos villamos 0,02 valószínűséggel hibásodik meg a Hanna által megtett úton. Egy adott napon mennyi a valószínűsége, hogy az a villamos, amivel Hanna utazik, meghibásodik?

A: meghibásodik a villamos

B_1 : Hanna négyessel jön; B_2 : Hanna hatossal jön.

B_1, B_2 **teljes eseményrendszer**, hiszen pontosan az egyik következik be, és pozitív a valószínűségük. A **teljes valószínűség tétele** alkalmazható:

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) = 0,01 \cdot 0,5 + 0,02 \cdot 0,5 = \mathbf{0,015}.$$

Diszkrét valószínűségi változók várható értéke

- Szerencsejáték igazságos ára: a nyeremény várható értéke, például ha $1/1000$ valószínűséggel 1000000 forintot nyerünk (különben semmit), akkor 1000 .
- Annyit nyerünk, amennyit egy szabályos dobókockával dobunk. A nyeremény várható értéke:

$$\frac{1}{6}(1 + 2 + \dots + 6) = \frac{21}{6} = 3,5.$$

- Ha szabálytalan a kocka, például az 1 helyett is 6 van:

$$\frac{1}{6}(6 + 2 + 3 + 4 + 5 + 6) = \frac{1}{6}(2 + 3 + 4 + 5) + \frac{1}{3} \cdot 6 = \frac{13}{3} = 4,33.$$

Diszkrét valószínűségi változók várható értéke

- Szerencsejáték igazságos ára: a nyeremény várható értéke, például ha $1/1000$ valószínűséggel 1000000 forintot nyerünk (különben semmit), akkor 1000 .
- Annyit nyerünk, amennyit egy szabályos dobókockával dobunk. A nyeremény várható értéke:

$$\frac{1}{6}(1 + 2 + \dots + 6) = \frac{21}{6} = 3,5.$$

- Ha szabálytalan a kocka, például az 1 helyett is 6 van:

$$\frac{1}{6}(6 + 2 + 3 + 4 + 5 + 6) = \frac{1}{6}(2 + 3 + 4 + 5) + \frac{1}{3} \cdot 6 = \frac{13}{3} = 4,33.$$

A lehetséges értékeket megszorozzuk a hozzájuk tartozó valószínűséggel, és ezeket összeadjuk.

Diszkrét valószínűségi változók várható értéke

Definíció (Várható érték, diszkrét eset)

Legyen $X : \Omega \rightarrow \mathbb{R}$ olyan diszkrét valószínűségi változó, melynek eloszlása $(x_1, p_1), (x_2, p_2), \dots$, azaz $\mathbb{P}(X = x_i) = p_i$, ahol $i = 1, 2, \dots$. Ekkor X várható értéke:

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

Diszkrét valószínűségi változók várható értéke

Definíció (Várható érték, diszkrét eset)

Legyen $X : \Omega \rightarrow \mathbb{R}$ olyan diszkrét valószínűségi változó, melynek eloszlása $(x_1, p_1), (x_2, p_2), \dots$, azaz $\mathbb{P}(X = x_i) = p_i$, ahol $i = 1, 2, \dots$. Ekkor X várható értéke:

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

Példa: három gyerek. Legyen X a fiúk száma a három gyerek közül. Ekkor

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

Példa: szabályos kockadobás. Legyen Y egy szabályos dobókockával dobott szám. Ekkor

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

Diszkrét valószínűségi változók várható értéke

- (elfajult eloszlás) Ha $X = c$ fennáll 1 valószínűséggel: $\mathbb{E}(X) = c \cdot \mathbb{P}(X = c) = c$.
- (korlátosság) Ha $a \leq X \leq b$ valamely $a < b$ számokra, akkor $a \leq \mathbb{E}(X) \leq b$.
- (egyenletes eloszlás) Ha az x_1, x_2, \dots, x_n számok mindegyikének $1/n$ a valószínűsége, akkor a várható érték a számok számtani közepe: $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$.
- (indikátor) Legyen \mathbb{I}_A az A esemény indikátora, vagyis 1, ha A bekövetkezik, és 0 különben. Ekkor $\mathbb{E}(\mathbb{I}_A) = 1 \cdot \mathbb{P}(\mathbb{I}_A = 1) = \mathbb{P}(A)$.
- (összeg) Ha X, Y valószínűségi változók és $X + Y$ várható értéke létezik, akkor

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Függvény várható értéke

Állítás

Legyen $X : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi változó, melynek lehetséges értékei x_1, x_2, \dots , továbbá $\mathbb{P}(X = x_k) = p_k$ teljesül $k \geq 1$ esetén. Legyen továbbá $g : \mathbb{R} \rightarrow \mathbb{R}$ függvény. Ekkor

$$\mathbb{E}(g(X)) = g(x_1)p_1 + g(x_2)p_2 + g(x_3)p_3 \dots,$$

ha ez a várható érték létezik.

Például: legyen Y egy szabályos dobókockával dobott szám. Ekkor $g(x) = x^2$ -tel:

$$\mathbb{E}(Y^2) = 1 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6} = 15,17.$$

Függvény várható értéke

Állítás

Legyen $X : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi változó, melynek lehetséges értékei x_1, x_2, \dots , továbbá $\mathbb{P}(X = x_k) = p_k$ teljesül $k \geq 1$ esetén. Legyen továbbá $g : \mathbb{R} \rightarrow \mathbb{R}$ függvény. Ekkor

$$\mathbb{E}(g(X)) = g(x_1)p_1 + g(x_2)p_2 + g(x_3)p_3 \dots,$$

ha ez a várható érték létezik.

Például: legyen Y egy szabályos dobókockával dobott szám. Ekkor $g(x) = x^2$ -tel:

$$\mathbb{E}(Y^2) = 1 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6} = 15,17.$$

Nem mindig létezik ez a várható érték. Például (szentpétervári paradoxon): Z geometriai eloszlású $p = 1/2$ paraméterrel (azaz $p_k = \frac{1}{2^k}$, $g(x) = 2^x$. Ekkor

$$\mathbb{E}(g(Z)) = 2^1 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2^2} + 2^3 \cdot \frac{1}{2^3} + \dots = 1 + 1 + 1 + \dots = \infty.$$

Teljes várható érték tétele

Definíció (Feltételes várható érték)

Legyen X diszkrét valószínűségi változó, melynek lehetséges értékei x_1, x_2, \dots , és B pozitív valószínűségű esemény. Ekkor az X -nek a B -re vonatkozó feltételes várható értéke:

$$\mathbb{E}(X|B) = \sum_{k=1}^{\infty} x_k \cdot \mathbb{P}(X = x_k|B) = \sum_{k=1}^{\infty} x_k \cdot \frac{\mathbb{P}(\{X = x_k\} \cap B)}{\mathbb{P}(B)}.$$

Például: két szabályos kockadobásnál az első dobás feltételes várható értéke, feltéve, hogy az összeg 6:

Teljes várható érték tétele

Definíció (Feltételes várható érték)

Legyen X diszkrét valószínűségi változó, melynek lehetséges értékei x_1, x_2, \dots , és B pozitív valószínűségű esemény. Ekkor az X -nek a B -re vonatkozó feltételes várható értéke:

$$\mathbb{E}(X|B) = \sum_{k=1}^{\infty} x_k \cdot \mathbb{P}(X = x_k|B) = \sum_{k=1}^{\infty} x_k \cdot \frac{\mathbb{P}(\{X = x_k\} \cap B)}{\mathbb{P}(B)}.$$

Például: két szabályos kockadobásnál az első dobás feltételes várható értéke, feltéve, hogy az összeg 6: három.

Állítás

Legyen X véges várható értékű valószínűségi változó, B_1, B_2, \dots pedig teljes eseményrendszer (azaz uniójuk Ω , és bármely kettő metszete üres). Ekkor

$$\mathbb{E}(X) = \mathbb{E}(X|B_1)\mathbb{P}(B_1) + \mathbb{E}(X|B_2)\mathbb{P}(B_2) + \dots = \sum_{k=1}^{\infty} \mathbb{E}(X|B_k)\mathbb{P}(B_k).$$

Valószínűségi változó szórása

Lehetséges motiváció: nem mindegy, hogy a buszok ütemesen (szabályosan) tíz percenként érkeznek, vagy a követési idő várható értéke tíz perc, de hol öt, hol tizenöt percenként jönnek; egy mérőeszköztől a mérési hiba, vagyis a mérés bizonytalansága is fontos, például nem mindegy, hogy adott pontosság eléréséhez hány mérést kell átlagolni.

Definíció (Szórásnégyzet (variancia))

Legyen $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor X szórásnégyzete:

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

Valószínűségi változó szórása

Lehetséges motiváció: nem mindegy, hogy a buszok ütemesen (szabályosan) tíz percnként érkeznek, vagy a követési idő várható értéke tíz perc, de hol öt, hol tizenöt percnként jönnek; egy mérőeszköztől a mérési hiba, vagyis a mérés bizonytalansága is fontos, például nem mindegy, hogy adott pontosság eléréséhez hány mérést kell átlagolni.

Definíció (Szórásnégyzet (variancia))

Legyen $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor X szórásnégyzete:

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

Definíció (Szórás (standard deviation))

Legyen $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik. Ekkor X szórásnégyzete:

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

A szórás kiszámítása

Állítás

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

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

A szórás kiszámítása

Állítás

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

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

Bizonyítás.

$$\begin{aligned} D^2(X) &= \mathbb{E}\left((X - \mathbb{E}X)^2\right) = \mathbb{E}\left(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\right) = \\ &= \mathbb{E}(X^2) - \mathbb{E}(2X\mathbb{E}(X)) + \mathbb{E}(X)^2 = \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2, \end{aligned}$$

ahol felhasználtuk, hogy összeg várható értéke a várható értékek összege, illetve hogy a konstans szorzó kiemelhető.

A szórás kiszámítása

Állítás

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

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

Bizonyítás.

$$\begin{aligned} D^2(X) &= \mathbb{E}\left((X - \mathbb{E}X)^2\right) = \mathbb{E}\left(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\right) = \\ &= \mathbb{E}(X^2) - \mathbb{E}(2X\mathbb{E}(X)) + \mathbb{E}(X)^2 = \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2, \end{aligned}$$

ahol felhasználtuk, hogy összeg várható értéke a várható értékek összege, illetve hogy a konstans szorzó kiemelhető.

Megjegyzés: az x_1, x_2, \dots, x_n számok tapasztalati szórása

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2}.$$

Diszkrét valószínűségi változó szórása

Állítás

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

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

Állítás (A szórás kiszámítása egész értékek esetén)

Legyen X olyan diszkrét valószínűségi változó, melyre $\mathbb{E}(X^2)$ létezik, és melynek lehetséges értékei nemnegatív egészek. Ekkor

$$D^2(X) = \sum_{k=0}^{\infty} k^2 \mathbb{P}(X = k) - \mathbb{E}(X)^2 = \sum_{k=0}^{\infty} k^2 \mathbb{P}(X = k) - \left[\sum_{k=0}^{\infty} k \mathbb{P}(X = k) \right]^2.$$

Diszkrét valószínűségi változó szórása

Legyen továbbra is X a fiúk száma három gyerek közül:

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

Diszkrét esetben így számolhatunk:

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

Ebből és a korábbi számolásból

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

Végül pedig a fiúk számának szórása:

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

A kockadobás szórása

Legyen X egy szabályos dobókockával dobott szám. Ekkor

$$\mathbb{E}(X^2) = \frac{1}{6} \cdot 1^2 + \frac{1}{6} \cdot 2^2 + \frac{1}{6} \cdot 3^2 + \frac{1}{6} \cdot 4^2 + \frac{1}{6} \cdot 5^2 + \frac{1}{6} \cdot 6^2 = \frac{91}{6}.$$

Másrészt

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

Ebből

$$D^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = 2,92.$$

A kockadobás szórása: $D(X) = \sqrt{2,92} = 1,71$.

Általában n oldalú dobókocka esetén: $D(X) = \sqrt{\frac{n^2-1}{12}}$.

Binomiális eloszlás: példa

Egy munkahelyi csapatban **hatan** dolgoznak együtt.

Tegyük fel, hogy egy tetszőleges napon **egymástól függetlenül** mindannyian **$p = 0,03$** valószínűséggel **hiányoznak**.

Mennyi a valószínűsége, hogy egy munkanapon **pontosan ketten hiányoznak a csapatból?**

Binomiális eloszlás: példa

Egy munkahelyi csapatban **hatan** dolgoznak együtt.
Tegyük fel, hogy egy tetszőleges napon **egymástól függetlenül**
mindannyian **$p = 0,03$** valószínűséggel **hiányoznak**.

Mennyi a valószínűsége, hogy egy munkanapon
pontosan ketten hiányoznak a csapatból?



néhány jó lehetőség és a valószínűsége:

0,03	0,03	0,97	0,97	0,97	0,97
0,03	0,97	0,03	0,97	0,97	0,97
0,03	0,97	0,97	0,03	0,97	0,97

...

0,97	0,03	0,03	0,97	0,97	0,97
------	------	------	------	------	------

...

0,97	0,97	0,97	0,97	0,03	0,03
------	------	------	------	------	------

$$\rightarrow 0,03^2 \cdot 0,97^4$$

$$\rightarrow 0,03^2 \cdot 0,97^4$$

$$\rightarrow 0,03^2 \cdot 0,97^4$$

$$\rightarrow 0,03^2 \cdot 0,97^4$$

$$\rightarrow 0,03^2 \cdot 0,97^4$$

szorzás

Binomiális eloszlás: példa

Egy munkahelyi csapatban **hatan** dolgoznak együtt.
Tegyük fel, hogy egy tetszőleges napon **egymástól függetlenül**
mindannyian **$p = 0,03$** valószínűséggel **hiányoznak**.

Mennyi a valószínűsége, hogy egy munkanapon
pontosan ketten hiányoznak a csapatból?

$$\boxed{0,97} \boxed{0,03} \boxed{0,03} \boxed{0,97} \boxed{0,97} \boxed{0,97} \rightarrow 0,03^2 \cdot 0,97^4$$

a jó lehetőségek száma, azaz
hányféleképpen választhatjuk ki a két hiányzót:

egy jó lehetőség valószínűsége:

tehát a valószínűség:

Binomiális eloszlás: példa

Egy munkahelyi csapatban **hatan** dolgoznak együtt.
Tegyük fel, hogy egy tetszőleges napon **egymástól függetlenül**
mindannyian $p = 0,03$ valószínűséggel **hiányoznak**.

Mennyi a valószínűsége, hogy egy munkanapon
pontosan ketten hiányoznak a csapatból?

$$\boxed{0,97} \boxed{0,03} \boxed{0,03} \boxed{0,97} \boxed{0,97} \boxed{0,97} \rightarrow 0,03^2 \cdot 0,97^4$$

a jó lehetőségek száma, azaz
hányféleképpen választhatjuk ki a két hiányzót: $\binom{6}{2}$

egy jó lehetőség valószínűsége: $0,03^2 \cdot 0,97^4$

tehát a valószínűség:

$$\mathbb{P}(\text{pontosan két hiányzó}) = \binom{6}{2} \cdot 0,03^2 \cdot 0,97^4 = 1,2\%.$$

Binomiális eloszlás

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

Például:

- Visszatevéses mintavétel, n húzás, p a fekete golyók aránya.
- Egy felmérésben $n = 1500$ embert kérdezőnk meg, egy adott kérdésre mindenki egymástól függetlenül $p = 0,8$ valószínűséggel válaszol. A válaszok száma binomiális eloszlású.
- Egy biztosító $n = 60000$ ügyfelének mindegyike egymástól függetlenül $p = 0,0001$ valószínűséggel okoz balesetet egy adott évben. A balesetet okozó ügyfelek száma binomiális eloszlású.
- Tegyük fel, hogy a nyár $n = 92$ napjának mindegyikén egymástól függetlenül $p = 0,02$ valószínűséggel lesz jégeső egy adott helyen. A nyári jégesős napok száma binomiális eloszlású.

Binomiális eloszlás

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

Mennyi a valószínűsége, hogy **pontosan k kísérlet sikerül**, azaz $X = k$? Ahogyan a korábbi példában láttuk:

- A jó lehetőségek száma, azaz hányféleképpen választhatjuk ki, hogy melyik k kísérlet sikeres: $\binom{n}{k}$.
- Egy jó lehetőség valószínűsége: $p^k(1 - p)^{n-k}$, hiszen a kísérletek függetlenek, ezért az együttes bekövetkezés (metszet) valószínűsége a valószínűségek szorzata, és k kísérlet sikerül, a többi $n - k$ nem.
- Mivel minden jó lehetőség ugyanolyan valószínű, az $X = k$ valószínűsége a lehetőségek számának és egy lehetőség valószínűségének szorzata:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (k = 0, 1, \dots, n).$$

Binomiális eloszlás: definíció

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

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

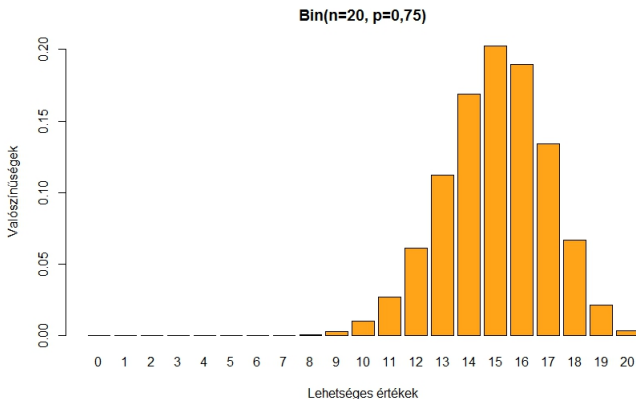
$$0, 1, 2, \dots, n,$$

és minden $0 \leq k \leq n$ egészre

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

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

Példa: binomiális eloszlás



Binomiális eloszlás, $n = 20$, $p = 0,75$. Vízszintes tengely: lehetséges értékek, azaz $k = 0, 1, \dots, 20$, oszlopok magassága: a $\mathbb{P}(X = k)$ valószínűségek.

Példa: binomiális eloszlás

Egy felmérésben $n = 1500$ embert kérdezzük meg. Egy adott kérdésre minden résztvevő **egymástól függetlenül** $p = 0,8$ valószínűséggel válaszol. Jelölje X , hogy hányan válaszoltak erre a kérdésre. Ekkor

- X **binomiális eloszlású** $n = 1500$ renddel és $p = 0,8$ paraméterrel.
- Tetszőleges $0 \leq k \leq 1500$ esetén

$$\begin{aligned}\mathbb{P}(k \text{ válasz}) &= \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \\ &= \binom{1500}{k} 0,8^k \cdot 0,2^{1500-k}.\end{aligned}$$

- Például annak valószínűsége, hogy pontosan $k = 1200$ -an válaszolnak a kérdésre:

$$\mathbb{P}(1200 \text{ válasz}) = \mathbb{P}(X = 1200) = \binom{1500}{1200} 0,8^{1200} \cdot 0,2^{300} = 2,57\%.$$

A binomiális eloszlás várható értéke és szórása

Ha az X valószínűségi változó **binomiális eloszlású** n renddel és p paraméterrel, azaz

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (k = 0, 1, \dots, n),$$

akkor X **várható értéke**, illetve **szórása**:

$$\mathbb{E}(X) = np; \quad D(X) = \sqrt{np(1 - p)}.$$

Példa: binomiális eloszlás

Egy felmérésben $n = 1500$ embert kérdezzük meg. Egy adott kérdésre minden résztvevő **egymástól függetlenül** $p = 0,8$ valószínűséggel válaszol. Jelölje X , hogy hányan válaszoltak erre a kérdésre. Ekkor

- A válaszadók számának **várható értéke**:

$$\mathbb{E}(X) = np = 1500 \cdot 0,8 = \mathbf{1200}.$$

- A válaszadók számának **szórása**:

$$D(X) = \sqrt{np(1-p)} = \sqrt{1500 \cdot 0,8 \cdot 0,2} = \mathbf{15,5}.$$

Poisson-eloszlás: definíció

- ha n független kísérletből mindegyik p valószínűséggel sikeres, ahol n „nagy” és p „kicsi”: $\lambda = np$ a sikeres kísérletek számának várható értéke;
- annak valószínűsége, hogy pontosan k kísérlet sikeres, $\frac{\lambda^k}{k!}e^{-\lambda}$ -val közelíthető (a számolás és az ábra alapján);
- ez a kifejezés könnyebben kezelhető, például mert csak egy paraméterrel rendelkezik;
- ez a gyakran használt Poisson-eloszlás **ritkán bekövetkező események számának** modellezésére.

Definíció

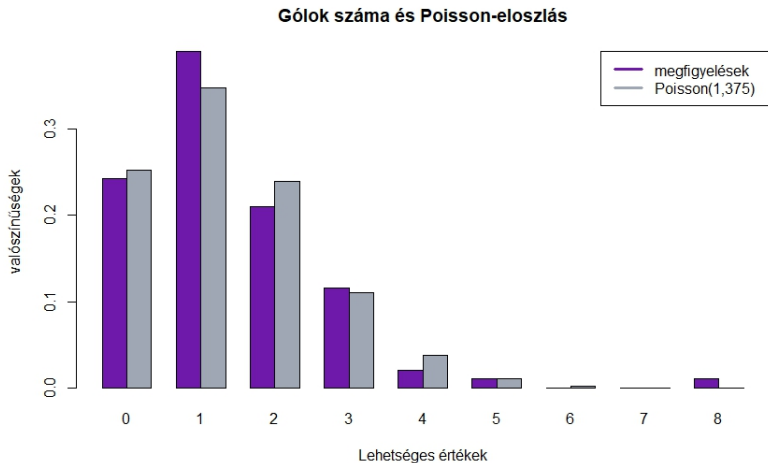
Legyen $\lambda > 0$. Az X valószínűségi változó λ paraméterű Poisson-eloszlású, ha lehetséges értékei:

$$k = 0, 1, 2, \dots, \text{ és ekkor } \mathbb{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

A λ paraméterű Poisson-eloszlás **várható értéke** és **szórása**:

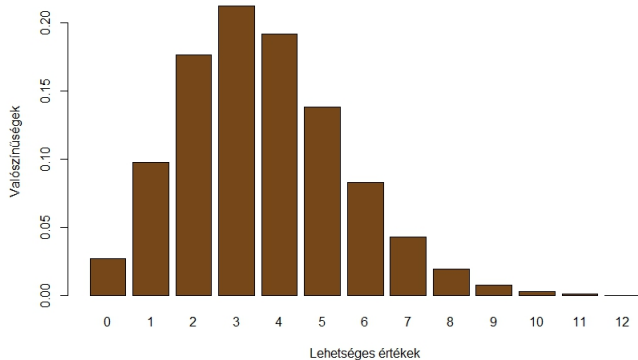
$$\mathbb{E}(X) = \lambda; \quad \mathbb{D}(X) = \sqrt{\lambda}.$$

Gólok száma



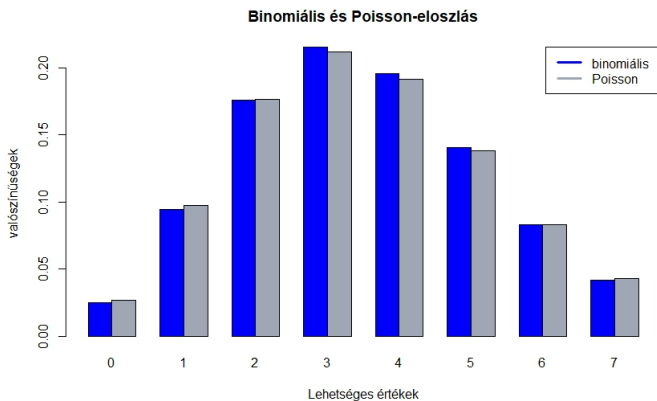
Gólok száma ($n = 95$ mérkőzésből) és Poisson-eloszlás $\lambda = \bar{X} = 1,375$ -tel

Példa: Poisson-eloszlás



Poisson-eloszlás, $\lambda = 3,61$, $k = 12$ -ig

A Poisson-eloszlás és binomiális eloszlás kapcsolata



A binomiális eloszlás $n = 92$ renddel és $p = 0,0392$ paraméterrel, illetve a Poisson-eloszlás $\lambda = np = 3,61$ paraméterrel

Poisson-eloszlás: példa

Az X valószínűségi változó λ **paraméterű Poisson-eloszlású**, ha lehetséges értékei:

$$k = 0, 1, 2, \dots, \text{ és ekkor } \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

A λ paraméterű Poisson-eloszlás **várható értéke** és **szórása**:

$$\mathbb{E}(X) = \lambda; \quad \mathbb{D}(X) = \sqrt{\lambda}.$$

Példa. Tegyük fel, hogy egy városban az egy nap alatt bekövetkező autóbalesetek száma Poisson-eloszlású, és **várható értéke 3,61**. Ekkor az egy nap alatt bekövetkező autóbalesetek számának **szórása**:

$$D(X) = \sqrt{3,61} = \mathbf{1,9}.$$

Annak valószínűsége, hogy pontosan 5 baleset lesz:

$$\mathbb{P}(X = 5) = \frac{3,61^5}{5!} e^{-5} = \mathbf{14\%}.$$

Binomális és Poisson-eloszlás az R-ben

- ha X binomiális eloszlású n renddel és p paraméterrel, akkor

$$\mathbb{P}(X = k) : \quad \text{dbinom}(k, \text{size} = n, \text{prob} = p)$$

és

$$\mathbb{P}(X \leq k) : \quad \text{pbinom}(k, \text{size} = n, \text{prob} = p).$$

- Továbbá az

```
minta<-rbinom(r, size=n, prob=p)
```

eredményeképpen a „minta” vektorba r darab *független* adott binomiális eloszlású valószínűségi változó kerül.

- ha X Poisson-eloszlású λ paraméterrel, akkor

$$\mathbb{P}(X = k) : \quad \text{dpois}(k, \text{lambdas} = \lambda)$$

és

$$\mathbb{P}(X \leq k) : \quad \text{ppois}(k, \text{lambdas} = \lambda).$$

- Továbbá az

```
minta<-rpois(r, lambdas=\lambda)
```

eredményeképpen a „minta” vektorba r darab *független* λ paraméterű eloszlású valószínűségi változó kerül.

Programozós házi feladat október 9., 10:15-ig (3 pont)

- Tegyük fel, hogy egy város n lakosa közül mindenki egymástól függetlenül p valószínűséggel támogat egy adott pártot. Válasszunk tetszőleges $n \geq 1000$ és $0,1 < p < 0,9$ számokat. Milyen eloszlású a pártot támogatók száma? Mennyi a várható értéke, mennyi a szórása?
- Sorsoljunk 500 elemű mintát a támogatók számának eloszlásából. Készítsünk a kapott adatsorból hisztogramot.
- Határozzuk meg a kapott adatsor átlagát (`mean(...)`) és (korrigált tapasztalati) szórását (`sd(...)`). Hasonlítsuk össze a kapott értékeket az első kérdésben kiszámolt várható értékkel és szórással.

Programozós házi feladat október 9., 10:15-ig (3 pont)

- Tegyük fel, hogy egy város n lakosa közül mindenki egymástól függetlenül p valószínűséggel támogat egy adott pártot. Válasszunk tetszőleges $n \geq 1000$ és $0,1 < p < 0,9$ számokat. Milyen eloszlású a pártot támogatók száma? Mennyi a várható értéke, mennyi a szórása?
- Sorsoljunk 500 elemű mintát a támogatók számának eloszlásából. Készítsünk a kapott adatsorból hisztogramot.
- Határozzuk meg a kapott adatsor átlagát (`mean(...)`) és (korrigált tapasztalati) szórását (`sd(...)`). Hasonlítsuk össze a kapott értékeket az első kérdésben kiszámolt várható értékkel és szórással.

Számolás házi feladat október 9., 10:15-ig (3 pont)

Tegyük fel, hogy ha az A és B focicsapat egymás ellen játszik, akkor az A csapat által rúgott gólok száma Poisson-eloszlású $1,2$ várható értékkel, a B csapat által rúgott gólok száma Poisson-eloszlású $1,8$ várható értékkel, továbbá ez a két valószínűségi változó egymástól független.

- Az $A - B$ labdarúgómérkőzésen mennyi az összes gól számának várható értéke?
- Feltéve, hogy egy mérkőzésen összesen 5 gól születik, mennyi a valószínűsége, hogy ebből k darabot szerzett az A csapat? Itt $k = 0, 1, 2, 3, 4, 5$ lehet.
- Ábrázoljuk a kapott feltételes valószínűségeket k függvényében (kézzel készített ábra is megfelelő).