Nevezetes diszkrét eloszlások (5. előadás)

$X : \Omega \to \mathbb{R}$ diszkrét valószínűségi változó eloszlása: a lehetséges értékei és a hozzájuk tartozó valószínűségek.

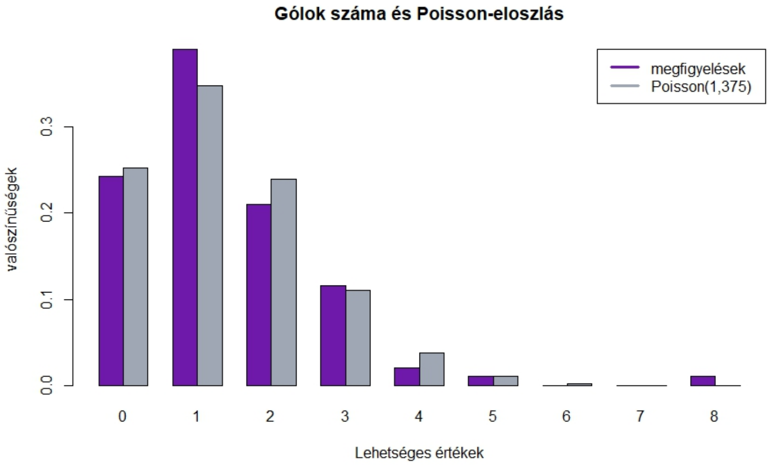
- binomiális eloszlás: n független kísérlet, mindegyik p valószínűséggel sikerül, X: hány sikerült. $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- hipergeometriai eloszlás: visszatevés nélküli mintavételnél a húzott fekete golyók száma; $\mathbb{P}(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$.
- geometriai eloszlás: minden kísérlet p valószínűséggel sikerül; Y: hányadik az első sikeres. $\mathbb{P}(Y = k) = (1-p)^{k-1} p$.

Nevezetes diszkrét eloszlások (5. előadás)

$X : \Omega \to \mathbb{R}$ diszkrét valószínűségi változó eloszlása: a lehetséges értékei és a hozzájuk tartozó valószínűségek.

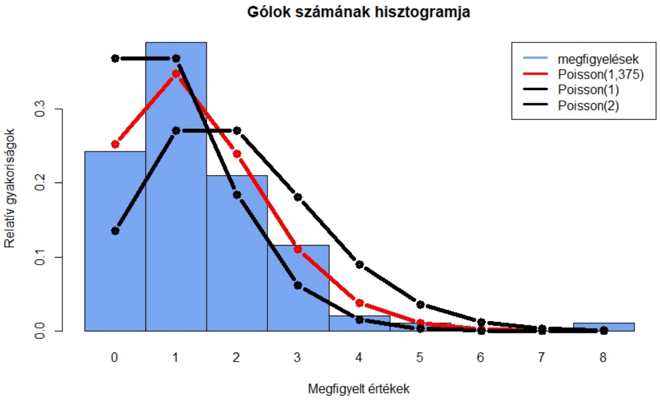
- binomiális eloszlás: n független kísérlet, mindegyik p valószínűséggel sikerül, X: hány sikerült. $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- hipergeometriai eloszlás: visszatevés nélküli mintavételnél a húzott fekete golyók száma; $\mathbb{P}(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$.
- geometriai eloszlás: minden kísérlet p valószínűséggel sikerül; Y: hányadik az első sikeres. $\mathbb{P}(Y = k) = (1-p)^{k-1} p$.
- Poisson-eloszlás: az X valószínűségi változó Poisson-eloszlású λ paraméterrel, ha
$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2 \ldots).$$

Gólok száma



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

Poisson-eloszlás



A gólok számának hisztogramja $n = 95$ mérkőzésen, és különböző paraméterű Poisson-eloszlások

Poisson-eloszlás

Definíció

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

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

Poisson-eloszlás

Definíció

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

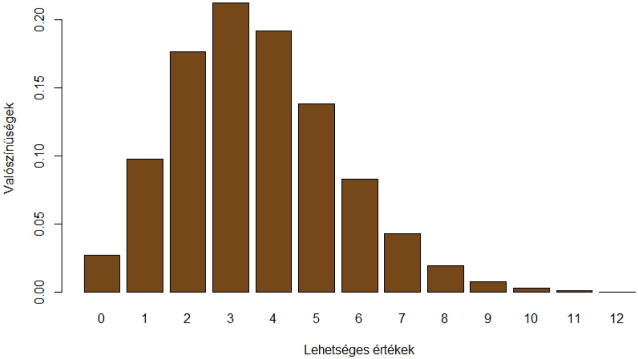
$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, \ldots).$$

Valóban eloszlás:

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) e^{-\lambda} = e^{\lambda} \cdot e^{-\lambda} = 1,$$

az exponenciális eloszlás Taylor-sora szerint.

Példa: Poisson-eloszlás



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

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 sajtóhibák száma egy könyvben;
- egy biztosító 15000 ügyfele által összesen okozott balesetek száma;
- nagyobb tűzesetek száma egy adott időszakban.

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 sajtóhibák száma egy könyvben;
- egy biztosító 15000 ügyfele által összesen okozott balesetek száma;
- nagyobb tűzesetek száma egy adott időszakban.

Legyen $\lambda > 0$ pozitív szám, és $p_n = \lambda/n$ minden $n = 1, 2, \ldots$ egészre. Legyen X Poisson-eloszlású valószínűségi változó λ 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 tetszőleges $k \geq 0$ egészre

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \lim_{n\to\infty} \binom{n}{k} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^{n} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

A Poisson-eloszlás és a binomiális eloszlás kapcsolata

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

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

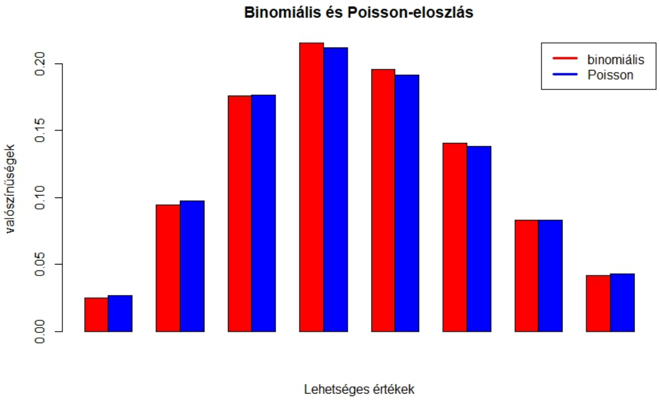
azaz

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

- Y binomiális eloszlás $n = 92$ renddel és $p = 0,0392$ paraméterrel;
- X Poisson-eloszlású $\lambda = 3,61 = 92 \cdot 0,0392 = np$ paraméterrel;

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

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

Alkalmazások

- első példa: lórúgás áldozatainak száma porosz hadseregben
- Poisson-folyamat: időben lejátszódó folyamat, melynél az $[a, b)$ intervallumba eső események száma $\lambda(b-a)$ paraméterű Poisson-eloszlású, és $a < b < c$ esetén az $[a, b)$ és $[b, c)$ intervallumba eső események száma független
- például Poisson-folyamattal lehet modellezni (közelíteni) azt, hogy egy üzletbe mikor érkeznek vásárlók (minden beérkezés egy esemény)

Megjegyzés: ha X, Y függetelnek, Poisson-eloszlásúak λ_1 és λ_2 paraméterrel, akkor $X + Y$ is Poisson-eloszlású, $\lambda_1 + \lambda_2$ paraméterrel.

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:
$$\frac{1}{6}(1+2+\ldots+6)=\frac{21}{6}=3,5.$$
- Ha szabálytalan a kocka, például az 1 helyett is 6 van:
$$\frac{1}{6}(6+2+3+4+5+6)=\frac{1}{6}(2+3+4+5)+\frac{1}{3}\cdot 6=\frac{13}{6}.$$

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:
$$\frac{1}{6}(1+2+\ldots+6) = \frac{21}{6} = 3,5.$$
- Ha szabálytalan a kocka, például az 1 helyett is 6 van:
$$\frac{1}{6}(6+2+3+4+5+6) = \frac{1}{6}(2+3+4+5) + \frac{1}{3}\cdot 6 = \frac{13}{6}.$$

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

Diszkrét valószínűségi változók várható értéke

Definíció (Várható érték, diszkrét eset)

Legyen $X : \Omega \to \mathbb{R}$ *olyan diszkrét valószínűségi változó, melynek eloszlása* $(x_1, p_1), (x_2, p_2), \ldots$*, azaz* $\mathbb{P}(X = x_i) = p_i$*. Ekkor* X *várható értéke:*

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

Diszkrét valószínűségi változók várható értéke

Definíció (Várható érték, diszkrét eset)

Legyen $X : \Omega \to \mathbb{R}$ *olyan diszkrét valószínűségi változó, melynek eloszlása* $(x_1, p_1), (x_2, p_2), \ldots$, *azaz* $\mathbb{P}(X = x_i) = p_i$. *Ekkor* X *várható értéke:*

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

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 $x_1, x_2, \ldots, x_n$ számok mindegyikének $1/n$ a valószínűsége, akkor a várható érték a számok számtani közepe: $\overline{x} = \frac{1}{n}(x_1 + \ldots + x_n)$.
- (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
$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Nevezetes diszkrét eloszlások várható értéke

- binomiális eloszlás n renddel és p paraméterrel: np
  például: 100 gyerek közül mindenki a többiektől függetlenül $0,2$ valószínűséggel balkezes. A balkezesek várható száma: $100 \cdot 0,2 = 20$
- hipergeometrikus eloszlás M, N, n paraméterekkel: $\frac{M}{N} \cdot n$
- geometriai eloszlás p paraméterrel: $1/p$
  például: szabályos dobókockával az első hatos dobáshoz szükséges dobások számának várható értéke 6
- Poisson-eloszlás λ paraméterrel: λ

Függvény várható értéke

Állítás

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

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

ha ez a várható érték létezik.

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

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

Függvény várható értéke

Állítás

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

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

ha ez a várható érték létezik.

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

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

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

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

Teljes várható érték tétele

Definíció (Feltételes várható érték)

Legyen X diszkrét valószínűségi változó, melynek lehetséges értékei $x_1, x_2, \ldots$, 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 $x_1, x_2, \ldots$, 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, $B_1, B_2, \ldots$ pedig teljes eseményrendszer (azaz uniójuk Ω, és bármely kettő metszete üres). Ekkor

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

Valószínűségi vektorváltozó

Gyakran több mennyiséget egyszerre vizsgálunk, vagy ugyanazt többször megmérjük. Például: véletlenszerűen választott ember havi jövedelme, kiadása rezsire, illetve élelmiszerre; vagy egy helyen öt különböző műszerrel megmérjük a hőmérsékletet.

Definíció (Valószínűségi vektorváltozó)

Egy $\underline{X} : \Omega \to \mathbb{R}^n$ *függvény valószínűségi vektorváltozó, ha tetszőleges* $t_1, t_2, \ldots, t_n$ *számokra*

$$\{\omega \in \Omega : X_1(\omega) \leq t_1, X_2(\omega) \leq t_2, \ldots, X_n(\omega) \leq t_n\} \in \mathcal{A},$$

azaz a $\mathbb{P}(X_1 \leq t_1, X_2 \leq t_2, \ldots, X_n \leq t_n)$ *valószínűség értelmes.*

Itt X értékei n dimenziós vektorok, vagyis n hosszú számsorozatok, és X_k ennek a k. koordinátáját (elemét) jelöli.

Peremeloszlás

Valószínűségi vektorváltozó együttes eloszlása: $\underline{x}_k$ lehetséges értékek és $p_k = \mathbb{P}(X = x_k)$ valószínűségek $(k = 1, 2, \ldots)$.

Az $\underline{X} = (X_1, X_2, \ldots, X_n)$ függvény pontosan akkor valószínűségi vektorváltozó, ha minden koordinátája valószínűségi változó. Ez alapján:

Definíció

Legyen $\underline{X} = (X_1, X_2, \ldots, X_n)$ valószínűségi vektorváltozó. Ekkor az X_k valószínűségi változó eloszlását az $\underline{X}$ valószínűségi vektorváltozó k. peremeloszlásának nevezzük.

Az együttes eloszlás meghatározza a peremeloszlásokat. Fordítva nem, kivéve, ha a koordináták függetlenek.

Együttes eloszlás: példa

Kétszer dobunk szabályos kockával. Legyen X az első dobás, Y pedig a dobott számok közül a nagyobb. Ekkor az (X, Y) valószínűségi vektorváltozó együttes eloszlása:

X/Y	1	2	3	4	5	6	összesen
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
2	0	$1/18$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
3	0	0	$1/12$	$1/36$	$1/36$	$1/36$	$1/6$
4	0	0	0	$1/9$	$1/36$	$1/36$	$1/6$
5	0	0	0	0	$5/36$	$1/36$	$1/6$
6	0	0	0	0	0	$1/6$	$1/6$
összesen	$1/36$	$1/18$	$1/12$	$1/9$	$5/36$	$1/6$	1

A peremeloszlások:

X: $(1, 1/6), (2, 1/6), (3, 1/6), \ldots, (6, 1/6)$

Y: $(1, 1/36), (2, 1/18), \ldots, (6, 1/6)$

Házi feladat október 16-ig

Legyenek X és Y szabályos kockával dobott számok. Határozzuk meg az alábbi mennyiségeket:

1. $\mathbb{E}(X+Y)$
2. $\mathbb{E}(X^2)$
3. $\mathbb{E}(X^2) - \mathbb{E}(X)^2$
4. $\mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2$