

Valószínűségszámítás előadás, 4. hét, 2020. szeptember 30.
Valószínűségi változók, várható érték, szórás, binomiális eloszlás

1. Valószínűségi változók

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 változókkal kapcsolatban az alábbi típusú kérdések fordulnak elő gyakran.

- 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

Ez vezet el a definícióhoz: úgy definiáljuk a valószínűségi változót, hogy a $\mathbb{P}(X \leq t)$ fogalom értelmes legyen, azaz $\{X \leq t\}$ egy esemény legyen, benne legyen az események \mathcal{A} halmazában.

1.1. Definíció (Valószínűségi változó). Egy $X : \Omega \rightarrow \mathbb{R}$ függvény **valószínűségi változó**, ha tetszőleges t valós számra teljesül, hogy

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

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

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:

$$\begin{aligned} \Omega &= \{FFF, FFL, FLF, FLL, LFF, LFL, LLF, LLL\}; \\ X(LLL) &= 0; \quad X(LLF) = X(LFL) = X(FLL) = 1; \\ X(FFL) &= X(FLF) = X(LFF) = 2; \quad X(FFF) = 3. \end{aligned}$$

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:

$$\begin{aligned} \mathbb{P}(X = 0) &= 1/8, & \mathbb{P}(X = 1) &= 3/8, \\ \mathbb{P}(X = 2) &= 3/8, & \mathbb{P}(X = 3) &= 1/8. \end{aligned}$$

1.1. Diszkrét valószínűségi változó eloszlása

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

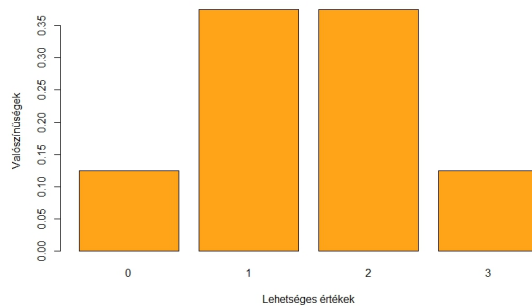
Legyenek az X **diszkrét valószínűségi változó** lehetséges értékei:

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

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.



1. ábra. A fiúk számának eloszlása három gyerek közül

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

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

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:

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

Az 1. ábrán láthatjuk is ezt az eloszlást, a $\mathbb{P}(X = k)$ valószínűséget ábrázolva a k függvényében.

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

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

A fiúk számának eloszlása: a lehetséges értékek:

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

és a hozzájuk tartozó valószínűségek:

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

1.2. Valószínűségi változó eloszlása általánosan

A valószínűségi változók eloszlását általánosabb esetben az alábbi módon írhatjuk le.

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

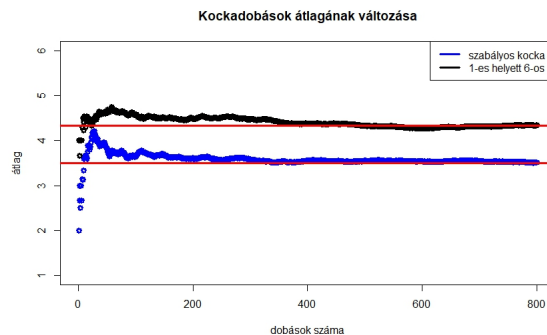
$$Q_X([a, b]) = \mathbb{P}(X \in [a, b]) = \mathbb{P}(a \leq X \leq b).$$

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

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

2. Diszkrét valószínűségi változók várható értéke

A diszkrét valószínűségi változót az eloszlásával írhatjuk le részletesen, de gyakran szükségünk van egyéb jellemzőire is. Az egyik legfontosabb jellemző a várható érték, ami egyetlen szám, és, ahogy később látni fogjuk, teljesíti például azt, hogy a valószínűségi változó (megfelelő feltételek mellett, megfelelő értelemben) nagy valószínűséggel nem esik túlságosan messze tőle. Sőt: a 2. ábrán két dobókocka esetében tekintettünk 800 – 800 dobást, és k függvényében ábráztuk az első k dobás átlagát. Látható, hogy ez átlag egy-egy adott értékhez közeledik, ez szintén a várható érték lesz ezekben az esetekben.



2. ábra. A dobások átlagának változása a dobások számának növelésével

- 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:

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

- Ha sokszor dobnánk, az átlag is egyre közelebb lenne a fent kiszámított értékhez (erről szól majd a nagy számok törvénye), emiatt is fontos ez a fogalom a statisztikában is.

2.1. 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:

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

Ha az X -nek csak véges sok lehetséges értéke van, akkor csak véges sok tagból áll az összeg, ekkor mindig létezik a várható érték.

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

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

2.1. Állítás (A várható érték tulajdonságai). A várható értékre az alábbi tulajdonságok jellemzők.

- (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

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

Gyakran előfordul, hogy egy valószínűségi változó egy függvényének a várható értékére vagyunk kíváncsiak. Ez lesz hasznos például a szórás kiszámításakor, de gondolhatunk például hasznosságfüggvényre is: X nyemremény (vagy bármilyen pénzüsszeg) hasznossága $g(X)$ lehet egy ember számára, ekkor ennek a várható értéke is érdekes.

2.2. 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

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

A teljes valószínűség tételének a várható érték számítására is van következménye: ha egy teljes eseményrendszer eseményeire feltételesen ismerjük egy valószínűségi változó várható értékét, és az egyes események valószínűsége is ismert, akkor összerakható a feltétel nélküli várható érték is,

2.2. 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.

2.3. 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

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

Ha csak véges sok eseményből áll a teljes eseményrendszer, akkor ez egy véges összeg lesz.

3. Valószínűségi változók 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.

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

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

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

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

Van olyan valószínűségi változó, melynek várható értéke létezik, de a szórása nem. Viszont ha egy valószínűségi változónak létezik a szórása, akkor a várható értéke is.

A szórás gyakran nem a definíció, hanem az alábbi állítás alapján számíthatjuk ki.

3.1. 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

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

3.1. Egész értékű valószínűségi változó szórása

3.2. 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

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

Példa. 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:

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

Példa. 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: $D(X) = \sqrt{\frac{n^2-1}{12}}$.

4. Binomiális eloszlás

A diszkrét eloszlások közül vannak bizonyos típusúak, melyek gyakran előfordulnak az alkalmazásoknál, ezeket érdemes külön megvizsgálni. Ezek közé tartozik a binomiális eloszlás.

Nézzünk először egy példát (következő oldal).

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 kettő hiányoznak a csapatból?

(A) (B) (C) (D) (E) (F)

0,03	0,03	0,97	0,97	0,97	0,97	→	$0,03^2 \cdot 0,97^4$
0,03	0,97	0,03	0,97	0,97	0,97	→	$0,03^2 \cdot 0,97^4$
0,03	0,97	0,97	0,03	0,97	0,97	→	$0,03^2 \cdot 0,97^4$
...							
0,97	0,03	0,03	0,97	0,97	0,97	→	$0,03^2 \cdot 0,97^4$
...							
0,97	0,97	0,97	0,97	0,03	0,03	→	$0,03^2 \cdot 0,97^4$

szorzás

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

Itt tehát $\binom{6}{2}$ -féleképpen választhatjuk ki a két hiányzót, és bárhogy is választottuk őket, annak a valószínűsége, hogy pontosan ők hiányoznak, $0,03^2 \cdot 0,97^4$, felhasználva, hogy a hiányzások függetlenek, így annak a valószínűsége, hogy az egyes emberek hiányoznak (nem hiányoznak), a külön-külön vett valószínűségek szorzata.

A fenti helyzetet így általánosíthatjuk:

- 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 kérdezzük 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.

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 sikerül: $\binom{n}{k}$.
- 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:

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

Így juthatunk el a binomiális eloszlás definíciójához.

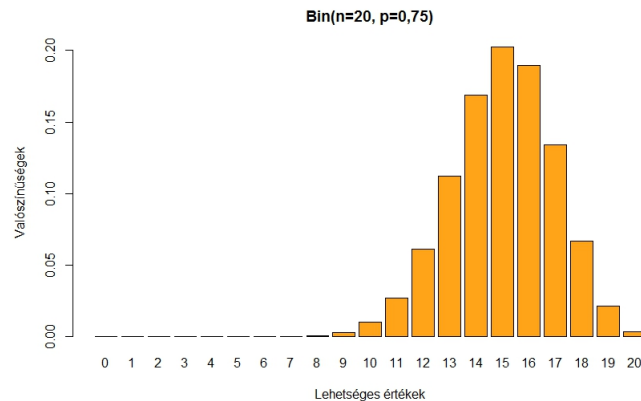
4.1. Definíció (Binomiális eloszlás). Az X valószínűségi változó **binomiális eloszlású** n renddel és p paraméterrel, ha lehetséges értékei:

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

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: $\text{Bin}(n, p)$.



3. ábra. 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. Várható érték: 15, szórás: 1,93.

Példa. 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:

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

4.1. Állítás. 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**:

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

Bizonyítás. Csak a várható értékre vonatkozó részt bizonyítjuk. Definíció szerint

$$\mathbb{E}(X) = \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k},$$

a jobb oldalon látható összeget azonban nem kell kiszámítanunk, ha felhasználjuk a várható érték egyik tulajdonságát, az additivitást.

A binomiális eloszlásra úgy is gondolhattunk, mint a sikeres kísérletek számára, n független, p valószínűséggel bekövetkező eseményből. Tekintsük az alábbi indikátor valószínűségi változókat minden $j = 1, 2, \dots, n$ -re:

$$\mathbb{I}_j = \begin{cases} 1, & \text{ha a } j. \text{ kísérlet sikerül;} \\ 0, & \text{különben.} \end{cases}$$

Ekkor az \mathbb{I}_j indikátorok összege éppen X lesz. Így a 2.1. állítás alapján

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{j=1}^n \mathbb{I}_j\right) = \sum_{j=1}^n \mathbb{E}(\mathbb{I}_j) = \sum_{j=1}^n 1 \cdot \mathbb{P}(\mathbb{I}_j = 1) = np. \quad \square$$

Példa. 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**:

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

Házi feladat október 7., szerda, 8:00-ig Egy liftben 15 ember száll be a földszinten. Mindenki egymástól függetlenül kiválasztja az $1, 2, \dots, 10$ emeletek valamelyikét, mindegyiket azonos valószínűséggel választva. Legyen X az olyan emeletek száma, ahol nem száll ki senki. Számítsuk ki X várható értékét. (segíthet, ha először 10 helyett 1, 2, 3 emelet esetén végezzük el a számítást).

Készítsünk számítógépes szimulációt, ami a fenti helyzetet modellezi, vagyis kisorsolja, hogy hol szállnak ki az emberek, és kiszámítja X értékét. Ismételjük meg ezt a kísérletet 100-szor. Mennyi lett a 100 kísérletből kapott X értékek átlaga? Hasonlítsuk ezt össze a kiszámított várható értékkel.