

Példa: a dobott számok összege (4. előadás)

Két szabályos dobókockával dobunk. Ekkor Ω :

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

Legyen $Y : \Omega \rightarrow \mathbb{R}$ a dobott számok összege, mint **valószínűségi változó**:

$$Y(11) = 2, Y(12) = Y(21) = 3, \dots,$$

$$Y(61) = Y(52) = \dots = Y(16) = 7, \dots, Y(66) = 12.$$

Példa: a dobott számok összege (4. előadás)

Két szabályos dobókockával dobunk. Ekkor Ω :

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

Legyen $Y : \Omega \rightarrow \mathbb{R}$ a dobott számok összege, mint **valószínűségi változó**:

$$Y(11) = 2, Y(12) = Y(21) = 3, \dots,$$

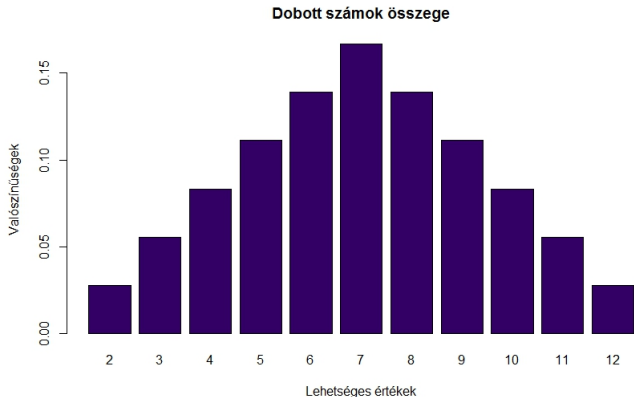
$$Y(61) = Y(52) = \dots = Y(16) = 7, \dots, Y(66) = 12.$$

Ekkor például

$$\mathbb{P}(Y = 2) = 1/36, \mathbb{P}(Y = 3) = 2/36, \mathbb{P}(Y = 7) = 1/6, \mathbb{P}(X = 8) = 5/36.$$

Példa: a dobott számok összege

Két szabályos dobókockával dobunk. A dobott számok összegének eloszlása:



Két kockadobás összegének eloszlása: lehetséges értékek és valószínűségek

Diszkrét valószínűségi változó eloszlása

Definíció (Valószínűségi változó)

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

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

Definíció (Diszkrét valószínűségi változó)

Az $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változó **diszkrét**, ha véges sok vagy megszámlálható sok értéket vesz fel.

Definíció (Diszkrét valószínűségi változó eloszlása)

Legyenek az X valószínűségi változó lehetséges értékei $x_1, x_2, \dots \in \mathbb{R}$, és legyen

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

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

Ilyenkor mindig teljesül, hogy $p_k \geq 0$ minden k -ra, és $\sum_{k=1}^{\infty} p_k = 1$, azaz a $(p_k)_{k \geq 0}$ sorozat **valószínűségeloszlás** (röviden: eloszlás).

Diszkrét valószínűségi változó eloszlása

Példa: három gyerek. Korábbi példa: X a fiúk száma három gyerek közül, mind a $2^3 = 8$ lehetőség egyformán valószínű.

Ekkor X lehetséges értékei: 0, 1, 2, 3.

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

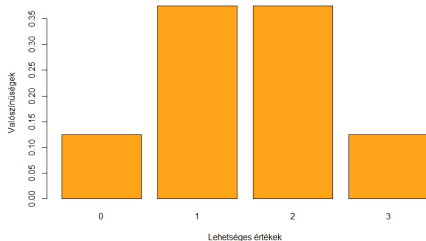
Mindezek alapján X eloszlása az alábbi sorozat:

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

Példa: szabályos kockadobás. Egyszer dobunk szabályos dobókockával, jelölje Y a dobott számot. Ekkor Y 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 gyerekek számának eloszlása.



A fiúk számának eloszlása: a lehetséges értékek és a hozzájuk tartozó valószínűségek

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 $X \subseteq \mathbb{R}$ 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

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

Binomiális eloszlás

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

Binomiális eloszlás

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

Például:

- Visszatevéses mintavétel
- Egy felmérésben $n = 1500$ embert kérdezőnk meg, egy adott kérdésre mindenki a többiekétől függetlenül $p = 0,8$ valószínűséggel válaszol. A válaszok száma binomiális eloszlású.
- Egy biztosító $n = 60000$ ügyfele közül mindenki a többiekétől 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 a többiekétől függetlenül $p = 0,02$ valószínűséggel lesz jégeső egy adott helyen. A nyári jégesők száma binomiális eloszlású.

Binomiális eloszlás

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

Binomiális eloszlás

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

Definíció

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

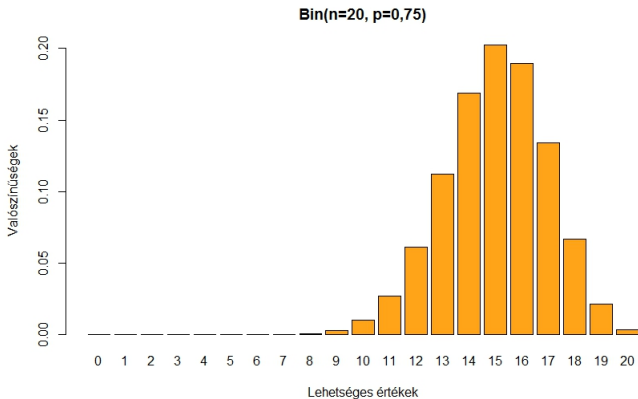
$$0, 1, 2, \dots, n,$$

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, $n = 20$, $p = 0,75$.

Példa: binomiális eloszlás

- Egy felmérésben $n = 1500$ embert kérdezőnk meg, egy adott kérdésre mindenki a többiektől függetlenül $p = 0,8$ valószínűséggel válaszol. Ha X jelöli a válaszadók számát, akkor tetszőleges $0 \leq k \leq 1500$ esetén

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

Például

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

- Visszatevéses mintavételnél a húzott fekete golyók száma (N golyó, ebből M fekete, n -szer húzunk visszatevéssel) binomiális eloszlású n renddel és $p = M/N$ paraméterrel:

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

Hipergeometriai eloszlás

Definíció

Legyenek N, M, n pozitív egészek úgy, hogy $1 \leq n \leq M \leq N$. Az X valószínűségi változó **hipergeometriai eloszlású**, ha

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

Hipergeometriai eloszlás

Definíció

Legyenek N, M, n pozitív egészek úgy, hogy $1 \leq n \leq M \leq N$. Az X valószínűségi változó **hipergeometriai eloszlású**, ha

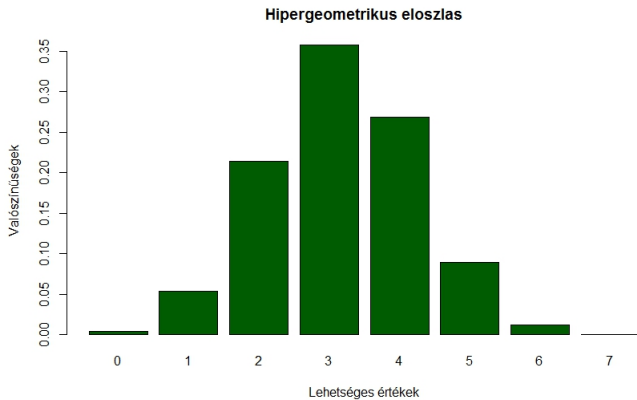
$$\mathbb{P}(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad (k = 0, 1, \dots, n).$$

Példa: visszatevés nélküli mintavételnél a húzott fekete golyók száma (N golyó, ebből M fekete, n -szer húzunk visszatevés nélkül).

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

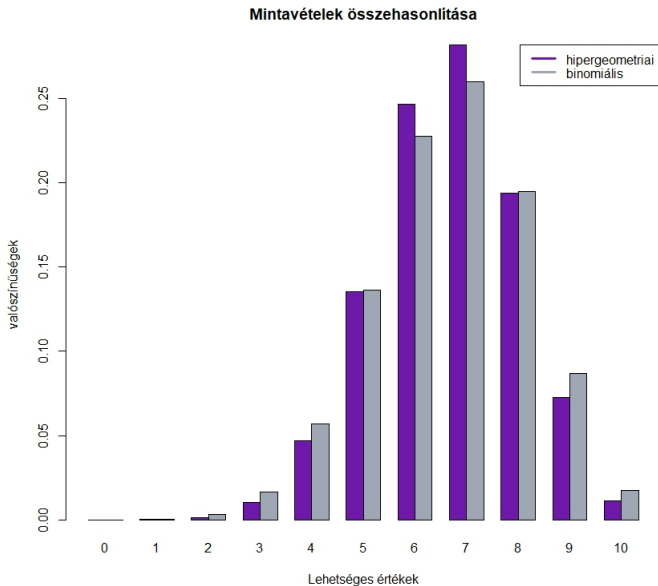
$$\mathbb{P}(X = k) = \mathbb{P}(k \text{ találat}) = \frac{\binom{5}{k} \binom{85}{5-k}}{\binom{90}{5}} \quad k = 0, 1, 2, 3, 4, 5.$$

Példa: a hipergeometriai eloszlás

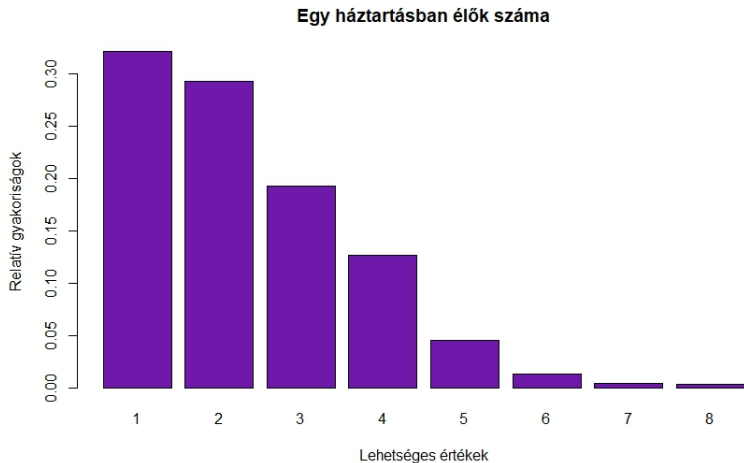


ábra: Hipergeometriai eloszlás, $N = 20$, $M = 9$, $n = 7$.

A hipergeometriai és binomiális eloszlás összehasonlítása



Megfigyelések



Egy háztartásban élők számának hisztogramja (forrás: KSH, 2011)

Geometriai eloszlás

Egy közvéleménykutatásban mindenki a többiektől függetlenül $0,2$ valószínűséggel válaszol egy adott kérdésre. Jelölje Y , hogy hány embert kell megkérdezni, míg találunk egy válaszadót.

$$\mathbb{P}(Y = 1) = \mathbb{P}(\text{az első ember válaszol}) = 0,2;$$

$$\mathbb{P}(Y = 2) = \mathbb{P}(\text{az első nem válaszol, a második igen}) =$$

Geometriai eloszlás

Egy közvéleménykutatásban mindenki a többiektől függetlenül $0,2$ valószínűséggel válaszol egy adott kérdésre. Jelölje Y , hogy hány embert kell megkérdezni, míg találunk egy válaszadót.

$$\mathbb{P}(Y = 1) = \mathbb{P}(\text{az első ember válaszol}) = 0,2;$$

$$\mathbb{P}(Y = 2) = \mathbb{P}(\text{az első nem válaszol, a második igen}) = 0,8 \cdot 0,2;$$

$$\mathbb{P}(Y = 3) = \mathbb{P}(\text{az első kettő nem válaszol, a harmadik igen}) =$$

Geometriai eloszlás

Egy közvéleménykutatásban mindenki a többiektől függetlenül $0,2$ valószínűséggel válaszol egy adott kérdésre. Jelölje Y , hogy hány embert kell megkérdezni, míg találunk egy válaszadót.

$$\mathbb{P}(Y = 1) = \mathbb{P}(\text{az első ember válaszol}) = 0,2;$$

$$\mathbb{P}(Y = 2) = \mathbb{P}(\text{az első nem válaszol, a második igen}) = 0,8 \cdot 0,2;$$

$$\mathbb{P}(Y = 3) = \mathbb{P}(\text{az első kettő nem válaszol, a harmadik igen}) = 0,8^2 \cdot 0,2;$$

$$\mathbb{P}(Y = k) = \mathbb{P}(\text{az első } k - 1 \text{ nem válaszol, a } k. \text{ igen}) =$$

Geometriai eloszlás

Egy közvéleménykutatásban mindenki a többiektől függetlenül $0,2$ valószínűséggel válaszol egy adott kérdésre. Jelölje Y , hogy hány embert kell megkérdezni, míg találunk egy válaszadót.

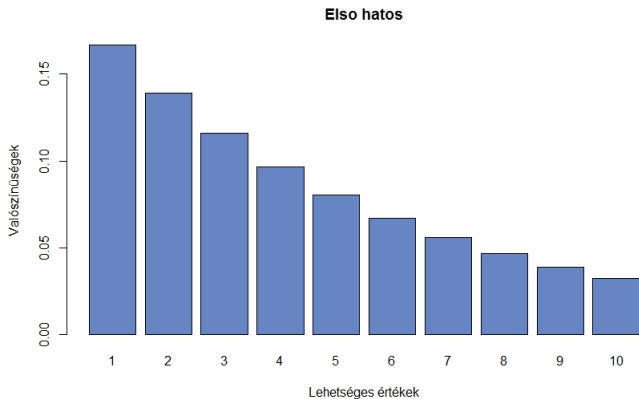
$$\mathbb{P}(Y = 1) = \mathbb{P}(\text{az első ember válaszol}) = 0,2;$$

$$\mathbb{P}(Y = 2) = \mathbb{P}(\text{az első nem válaszol, a második igen}) = 0,8 \cdot 0,2;$$

$$\mathbb{P}(Y = 3) = \mathbb{P}(\text{az első kettő nem válaszol, a harmadik igen}) = 0,8^2 \cdot 0,2;$$

$$\mathbb{P}(Y = k) = \mathbb{P}(\text{az első } k - 1 \text{ nem válaszol, a } k. \text{ igen}) = 0,8^{k-1} \cdot 0,2.$$

Példa: geometriai eloszlás



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

Geometriai eloszlás

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

Geometriai eloszlás

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

Definíció

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

$$1, 2, 3 \dots$$

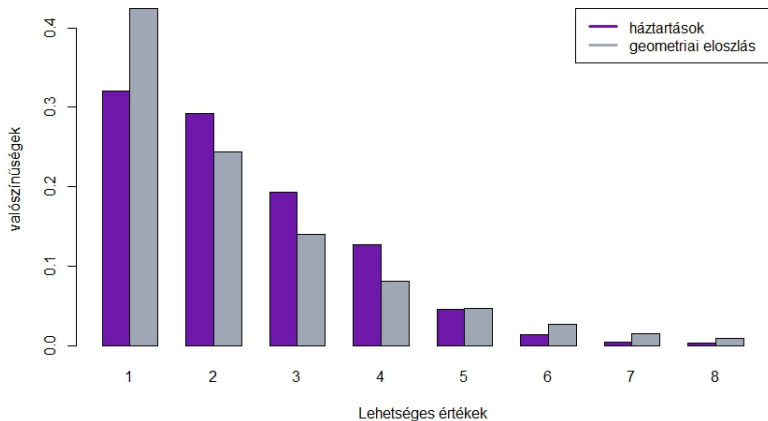
és minden $1 \leq k$ egészre

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

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

Megfigyelések

Háztartások és geometriai eloszlás



Egy háztartásban élők száma és geometriai eloszlás $p = \frac{1}{X} = 0,423$ -vel

Függetlenség

Definíció

Az X és Y diszkrét valószínűségi változók függetlenek, ha az X minden lehetséges x_k értékére és az Y minden lehetséges y_l értékére

$$\mathbb{P}(X = x_k, Y = y_l) = \mathbb{P}(X = x_k) \cdot \mathbb{P}(Y = y_l) \quad (k, l = 1, 2, \dots)$$

teljesül.

Például: X az első dobás, Y a második dobás egy kockával.

$$\mathbb{P}(X = 3, Y = 5) = \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 5) = \frac{1}{36}.$$

Függetlenség

Definíció

Az X és Y diszkrét valószínűségi változók függetlenek, ha az X minden lehetséges x_k értékére és az Y minden lehetséges y_l értékére

$$\mathbb{P}(X = x_k, Y = y_l) = \mathbb{P}(X = x_k) \cdot \mathbb{P}(Y = y_l) \quad (k, l = 1, 2, \dots)$$

teljesül.

Például: X az első dobás, Y a második dobás egy kockával.

$$\mathbb{P}(X = 3, Y = 5) = \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 5) = \frac{1}{36}.$$

Definíció

Az X és Y valószínűségi változók függetlenek, ha tetszőleges t_1, t_2 számokra

$$\mathbb{P}(X \leq t_1, Y \leq t_2) = \mathbb{P}(X \leq t_1) \cdot \mathbb{P}(Y \leq t_2)$$

teljesül.

Függetlenség

Definíció (Véges eset)

Azt mondjuk, hogy az $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ valószínűségi változók **függetlenek**, ha

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

teljesül tetszőleges t_1, t_2, \dots, t_n valós számokra.

Definíció (Végtelen eset)

Az $X_1, X_2, X_3 \dots$ valószínűségi változók függetlenek, ha közülük bármely véges sokat kiválasztva független valószínűségi változókat kapunk.

Állítás (Egész értékű eset)

Az $X_1, \dots, X_n : \Omega \rightarrow \mathbb{Z}$ valószínűségi változók pontosan akkor függetlenek, ha

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

teljesül tetszőleges k_1, k_2, \dots, k_n egész számokra.

Függetlenség

Független valószínűségi változókra példa:

- Két kockadobásnál az elsőként és másodikként dobott szám.
- A holtapi csapadékmennyiség Budapesten és New Yorkban.
- Két találomra választott ember jövedelme.

Nem független valószínűségi változókra példa:

- Két kockadobásnál az első szám és a két dobott szám összege.
- A holtapi csapadékmennyiség Budapesten és Budaörsön.
- Két találomra választott ember jövedelme öt év múlva.

Feltételes eloszlás

Legyen X diszkrét valószínűségi változó, lehetséges értékei x_1, x_2, \dots , a hozzájuk tartozó valószínűségek: $\mathbb{P}(X = x_k) = p_k$. Legyen A pozitív valószínűségű esemény. Ekkor az X -nek az A eseményre vonatkozó feltételes eloszlása:

$$q_k = \mathbb{P}(X = x_k | A) = \frac{\mathbb{P}(\{X_k = x_k\} \cap A)}{\mathbb{P}(A)}.$$

Ekkor

$$\sum_{k=1}^{\infty} q_k = \sum_{k=1}^{\infty} \frac{\mathbb{P}(\{X_k = x_k\} \cap A)}{\mathbb{P}(A)} = 1,$$

azaz a (q_k) sorozat is valószínűségeloszlás.

Másrészt az $\{X = x_k\}$, $k = 1, 2, \dots$ események teljes eseményrendszert alkotnak.