Diszkrét valószínűségi változók

Legyen egy X diszkrét valószínűségi változó (v.v.) lehetséges kimeneteli: $x_1, x_2, \ldots$. Ekkor a $p_i = P(X = x_i)$ valószínűséggekkel az X v.v. <u>eloszlását</u> határozzuk meg.

Példák:

1. Kockadobás: X: eredmény:
$$P(X = 1) = \frac{1}{6}, P(X = 2) = \frac{1}{6}, \ldots, P(X = 6) = \frac{1}{6}$$
2. Kétgyerekes családok, X: fiúk száma:
$$P(X = 0) = \frac{1}{4}, \quad P(X = 1) = \frac{2}{4}, \quad P(X = 2) = \frac{1}{4}$$
3. Háromgyerekes családok, X: fiúk száma:
$$P(X = 0) = \frac{1}{8}, \quad P(X = 1) = \frac{3}{8}, \quad P(X = 2) = \frac{3}{8},$$
$$P(X = 3) = \frac{1}{8}$$

Egy pénzérmével addig dobunk, amíg fejet nem dobunk. X: dobások száma:

$$P(X = k) = \frac{1}{2^k}, \quad \text{ha} \quad k = 1, 2, \ldots$$

Az eloszlásfüggvényre teljesül, hogy $\sum_i P(X = x_i) = 1$.

Az X valószínűségi változó esetén az $F(x) = P(X < x)$ függvényt az X <u>eloszlásfüggvényének</u> hívjuk.

Könnyen ellenőrizhető az eloszlásfüggvény alábbi tulajdonságai:

1. $F(x)$ monoton nő
2. $\lim_{x \to -\infty} F(x) = 0$ és $\lim_{x \to \infty} F(x) = 1$

Az X v.v. lehetséges kimenetelei közül a legvalószínűbbet az X <u>módusz</u>ának hívjuk.

Példa: Legyen az X v.v. a kétgyerekes családokban a fiúk száma. Ekkor X módusza a 2.

1.6 Várható érték és szórás

Egy szabályos kockával dobjunk 6000-szer és átlagoljuk ki az eredményeket, ekkor minden bizonnyal a következőt kapjuk:

$$\text{átlag} =$$

$$\frac{1 * (\text{1-esek száma}) + 2 * (\text{2-esek száma}) + \cdots + 6 * (\text{6-osok száma})}{6000}$$

$$\approx \frac{1 * 1000 + 2 * 1000 + \cdots + 6 * 1000}{6000} = 3,5$$

A fenti számításban

$$\frac{\text{1-esek száma}}{6000} \approx P(X = 1),$$

$$\frac{\text{2-esek száma}}{6000} \approx P(X = 2), \ldots$$

tehát

$$\text{átlag} \approx 1P(X = 1) + 2P(X = 2) + \cdots + 6P(X = 6)$$

Legyen X egy diszkrét v.v. $x_1, x_2, \ldots$ kimenetelekkel. Ekkor a $\sum_k x_k P(X = x_k)$ összeget az X v.v. <u>várható értékének</u> hívjuk.
Jelölés: $E(X)$.

A várható érték a következőképpen szemlétethető: tekintsük azt a tömegpontrenszert, ahol az x_i pont súlya $P(X = x_i)$. Ekkor a várható érték az első momentum.
Példa:

1. Kétgyerekes család: X: fiúk száma. Ekkor

$$E(X) = 0P(X = 0) + 1P(X = 1) + 2P(X = 2) = 1 * \frac{2}{4} + 2 * \frac{1}{4} = 1$$

2. Háromgyerekes család: X: fiúk száma. Ekkor

$$E(X) = 0P(X = 0) + 1P(X = 1) + 2P(X = 2) + 3P(X = 3) =$$

$$1 * \frac{3}{8} + 2 * \frac{3}{8} + 3\frac{1}{8} = 1,5$$

Várható érték tulajdonságai:

1. Ha $a, b \in \mathbb{R}$, akkor $E(aX + b) = aE(X) + b$
2. Ha X és Y v.v.-k, akkor $E(X + Y) = E(X) + E(Y)$

Legyen X lehetséges kimenetelei: $x_1, x_2, \ldots$ és Y lehetséges kimenetelei: $y_1, y_2, \ldots$. Az X és Y v.v.-k <u>függetlenek</u>, ha a $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$ minden i,j-re

Megmutatható, hogy ha X és Y függetlenek, akkor $E(XY) = E(X)E(Y)$

A következő érdekes kérdés az, hogy az átlagértéket hogyan kapjuk: lehet, hogy mindig az átlagértékhez közeli lesz a v.v., de az is lehet, hogy közben jelentősen eltér. Ezt ellenőrzi az ún. szórás: az X v.v. szórásnégyzete (vagy varianciája):

$$D^2(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2,$$

ahol

$$E(X^2) = \sum_i x_i^2 P(X = x_i)$$

Ekkor a szórás:

$$D(X) = \sqrt{D^2(X)}$$

A szórásnégyzet az előző tömegpontrendszerre nézve az első momentumra vonatkozó második momentumot jelenti.

Példák:

X: kockadobás eredménye. Ekkor tudjuk, hogy $E(X) = 3,5$, továbbá

$$E(X^2) = 1^2 P(X = 1) + 2^2 P(X = 2) + 3^2 P(X = 3) +$$

$$4^2 P(X = 4) + 5^2 P(X = 5) + 6^2 P(X = 6) =$$

$$1 * \frac{1}{6} + 4 * \frac{1}{6} + 9 * \frac{1}{6} + 16 * \frac{1}{6} + 25 * \frac{1}{6} + 36 * \frac{1}{6} =$$

$$\frac{91}{6}$$

Innen

$$D^2(X) = E(X^2) - (E(X))^2 = \frac{91}{6} - 3,5^2 = \frac{35}{12} \Rightarrow$$

$$D(X) = \sqrt{\frac{35}{12}} = 1,708$$

X: kétgyerekes családban a fiúk száma: Tudjuk, hogy $E(X) = 1$ és

$$E(X^2) = 0^2 P(X = 0) + 1^2 P(X = 1) + 2^2 P(X = 2) =$$

$$0 * \frac{1}{4} + 1 * \frac{2}{4} + 4 * \frac{1}{4} = 1,5$$

Innen

$$D^2(X) = E(X^2) - (E(X))^2 = 1,5 - 1^2 = 0,5 \Rightarrow$$

$$D(X) = \sqrt{0,5} = 0,707$$

A szórás tulajdonságai:

1. Ha X valószínűségi változó, akkor tetszőleges $a, b \in \mathbb{R}$ esetén $aX + b$ is v.v. és

$$D(aX + b) = |a|D(X)$$

2. Ha X és Y független v.v-k, akkor

$$D^2(X + Y) = D^2(X) + D^2(Y)$$

Példa: Dobjunk kétszer egy egymás után egy kockával. Jelölje Z a dobott számok összegét! Határozza meg Z várható értékét és szórását!
Megoldás: Jelölje X az első és Y a második dobás eredményét! Ekkor $Z = X + Y$, ezért

$$E(Z) = E(X + Y) = E(X) + E(Y) = 3,5 + 3,5 = 7$$

$$D^2(Z) = D^2(X + Y) = D^2(X) + D^2(Y) = 2 * \frac{35}{12} = 2,41$$

2. Nevezetes diszkrét eloszlások

Egyenletes eloszlás

Az X valószínűségi változó egyenletes az $x_1, x_2, \ldots, x_n$ számokon, ha minden $1 \leq i \leq n$ esetén

$$P(X = x_i) = \frac{1}{n}$$

Ekkor

$$E(X) = \frac{\sum_{i=1}^{n} x_i}{n}$$

Példa: X: kockadobás eredménye: $P(X = i) = \frac{1}{6}$, $i = 1, 2, 3, 4, 5, 6$.

Bernoulli-féle eloszlás

Legyen $0 \leq p \leq 1$. Az X p-paraméterű Bernoulli-féle v.v., ha X két értéket vehet fel: a 0-t és az 1-et. Ezek valószínűségei:

$$P(X = 0) = 1 - p \quad \text{és} \quad P(X = 1) = p$$

Példa: Kockadobás: X a 6-osok száma. Ekkor

$$P(X = 0) = \frac{5}{6}, \quad P(X = 1) = \frac{1}{6}$$

Ha X egy p-paraméterű v.v., akkor

$$E(X) = 0P(X = 0) + 1P(X = 1) = p$$

$$E(X^2) = 0^2P(X = 0) + 1^2P(X = 1) = p,$$

ezért

$$D^2(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p) \quad \Rightarrow$$

$$D(X) = \sqrt{p(1 - p)}$$

Binomiális eloszlás vagy visszatevéses mintavétel

Az X v.v. n és p paraméterű binomiális eloszlású, ha n egymástól független kisérletet végzünk és közben megszámoljuk, hogy egy p valószínűségű esemény hányszor következik be.
Ekkor X v.v. $0 \leq k \leq n$ értékeket vehet fel. Bebizonyítható, hogy

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

Példák: 10-szer dobunk egy pénzérmével. Jelölje X a fejek számát. Ekkor X egy $n = 10$ és $p = 0,5$ binomiális eloszlású v.v, ezért

$$P(6fej) = P(X = 6) = \binom{10}{6} \left(\frac{1}{2}\right)^6 \left(1 - \frac{1}{2}\right)^{10-6}$$

20-szor dobunk egy kockával. Jelölje X a 6-osok számát. Ekkor X egy $n = 20$ és $p = \frac{1}{6}$ binomiális eloszlású v.v, ezért

$$P(3 \text{ db 6-os}) = P(X = 3) = \binom{20}{3} \left(\frac{1}{6}\right)^3 \left(1 - \frac{1}{6}\right)^{20-3}$$

Visszatevéses mintavétel: Adott N tárgy és ezek közül M valamilyen szempontból kitüntetett. Ha ezek közül n-szer húzunk véletlenszerűen, visszatevéssel, akkor a köztük lévő kitüntettek száma n és $p = \frac{M}{N}$ paraméterű binomiális eloszlású v.v.

Példa: Magyar kártyából visszatevéssel húzunk 6 lapot. Jelölje X a köztük lévő tökök számát. Ekkor X v.v. $n = 6$ és $p = \frac{9}{36}$ paraméterű binomiális eloszlású v.v.

Egy nap egy gépsor 10.000 terméket állít elő, amik közül 1.000 hibás. Visszatevéssel kiválasztunk 500 terméket. Mi a valószínűsége, hogy legfeljebb 45 hibásat választottunk ki?

Megoldás: Legyen X az 500 kiválasztott termék között lévő hibásak száma. Ekkor X egy $n = 500$ és $p = \frac{1.000}{10.000} = 0,1$ paraméterű binomiális eloszlású v.v. Ekkor

$$P(\text{legfeljebb 45 hibás}) = P(X \leq 45) =$$

$$\sum_{k=0}^{45} \binom{500}{k} 0,1^k (1 - 0,1)^{500-k} \approx ?$$

Ha X egy n és p paraméterű binomiális eloszlású v.v., akkor ezt fel lehet írni Bernoulli v.v. összegeként a következőképpen: jelölje X_i azt, hogy az i-edik kisérletnél hányszor következett be az esemény. Ekkor

$$P(X_i = 0) = 1-p \text{ és } P(X_i = 1) = p, \quad E(X_i) = p, D^2(X_i) = p(1-p)$$

és

$$X = X_1 + X_2 + \cdots + X_n,$$

ezért

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$$

$$D^2(X) = D^2(X_1+X_2+\cdots+X_n) = D^2(X_1)+D^2(X_2)+\cdots+D^2(X_n) = np(1-p)$$

Tehát $D(X) = \sqrt{np(1-p)}$

Megmutatható, hogy X módusza $\lceil np \rceil$

Hipergeometriai eloszlás (visszatevés nélküli mintavétel)

Legyen N tárgy és ebből M kitüntetett. Visszatevés nélkül húzunk n darabot. Az X v.v. azt mutatja, hogy hány kitüntetettet húztunk. Ekkor megmutatható, hogy $0 \leq k \leq n$ esetén

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

bizonyítható, hogy $E(X) = \frac{M}{N}n$ és X módusza $\lceil \frac{M}{N}n \rceil$

Példa: Egy hallgató a 100 tétel közül 20-at nem tud. 10 tételt kap a vizsgán. Akkor megy át, ha legfeljebb 1 tételt nem tud. Mi a a valószínűsége, hogy átmegy a vizsgán a hallgató?
$N = 100$, $M = 20$, $n = 20$
Megoldás: Jelölje X, hogy a hallgató a 10 tétel közül hányat nem tud.

$$P(\text{átmegy}) = P(X = 0 \text{ vagy } 1) = P(X = 0) + P(X = 1) =$$

$$\frac{\binom{80}{10}\binom{20}{0}}{\binom{100}{20}} + \frac{\binom{80}{9}\binom{20}{1}}{\binom{100}{20}}$$

A 90-es lottón kiválasztunk 5 számot. Jelölje X a találataink számát. Ekkor $N = 90$, $M = 5$, $n = 5$

$$p_k = P(X = k) = \frac{\binom{5}{k}\binom{85}{5-k}}{\binom{90}{5}}$$

Kisszámolható:

$$p_0 = 0,746, \quad p_1 = 0,23, \quad p_2 = 0,022$$

$$p_3 = 0,0008, \quad p_4 = 0,00001, \quad p_5 = 2,3 * 10^{-8}$$

Geometriai eloszlás

Feladat: Egy kockával addig dobunk, ameddig hatost nem dobunk. Jelölje X a dobások számát. Ekkor

$$P(X = 3) = \left(\frac{5}{6}\right)^2 \frac{1}{6}$$

Általában legyen adva egy p valószínűségű esemény. Ekkor kisérleteket végzünk egészen addig, ameddig a p valószínűségű esemény bekövetkezik. Jelölje X a kisérletek számát. Ekkor

$$P(X = k) = (1 - p)^{k-1} p.$$

Megmutatható, hogy $E(X) = \frac{1}{p}$

Poisson eloszlás

Tekintsük az n és p paraméterű binomiális eloszlású v.v.-t. Ha n nagy és $p = \frac{\lambda}{n}$, akkor, ha k egy nem nagy szám

$$P(X = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \approx e^{-\lambda} \frac{\lambda^k}{k!}$$

Azt mondjuk, hogy az X v.v. λ paraméterű <u>Poisson-eloszlást</u> követ, ha

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Ekkor $E(X) = \lambda$

Poisson eloszlást akkor alkalmazunk, amikor a véletlen jelenség sok kicsi valószínűségű eseményből tevődik össze.

Példák:

1. Hány sajtóhiba van egy újságban
2. Egy nagyvárosban azon emberek száma, akik megélik a 100 évet
3. Hány piros Suzuki megy át a Petőfi hídon öt perc alatt
4. Kapott spam-ek száma egy nap alatt
5. Földrengések száma egy adott idő alatt
6. Magyarországra becsapódó meteorok száma egy év alatt
7. Súlyos balesetek egy nap alatt Budapesten

Példa: Egy évben átlag kétszer siklik ki a metró. Határozza meg a következő valószínűségeket:

1. Egy évben 3-szor siklik ki a metró
2. Egy évben nem siklik ki a metró
3. Egy évben lesz kisiklás
4. Az első félévben 2 kisiklás lesz

Megoldás: Jelölje X, azt, hogy egy évben hányszor siklik ki a metró. Ekkor X egy $\lambda = 2$ paraméterű Poisson-eloszlású v.v., ezért

$$P(\text{Egy évben háromszor siklik ki a metró}) = P(X = 3) =$$

$$e^{-2} * \frac{2^3}{3!}$$

$$P(\text{Egy évben nem siklik ki a metró}) = P(X = 0) =$$

$$e^{-2} * \frac{2^0}{0!} = e^{-2}$$

$$P(\text{Egy évben lesz kisiklás}) = P(X > 0) = 1 - P(X = 0) =$$

$$1 - e^{-2}$$

Jelölje Y, hogy az első félévben hány kisiklás lesz. Ekkor az átlagos szám 1 lesz, ezért ez egy $\lambda = 1$ paraméterű Poisson-eloszlású v.v. lesz. Így

$$P(\text{Az első félévben 2 kisiklás lesz}) = P(Y = 2) = e^{-1} * \frac{1^2}{2!}$$