

3. Gyakorlat

Valószínűségi változók, diszkrét valószínűségi változók, nevezetes diszkrét eloszlások

Valószínűségi változók

Definíció (Valószínűségi változó) Legyen $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező. Ekkor az $X : \Omega \rightarrow \mathbb{R}$ függvényt valószínűségi változónak nevezzük, amennyiben minden $x \in \mathbb{R}$ esetén $\{X < x\} \in \mathcal{F}$. (Vagyis egy valószínűségi változó minden elemi eseményhez egy valós számot rendel.)

Diszkrét valószínűségi változók

Definíció (Diszkrét valószínűségi változó) Az X valószínűségi változó diszkrét, amennyiben az értékkészlete megszámlálható (véges vagy megszámlálhatóan végtelen).

Definíció (Súlyfüggvény) Az X valószínűségi változó súlyfüggvénye $p : \text{Ran}(X) \rightarrow [0, 1]$

$$p(x) = \mathbb{P}(X = x).$$

Nevezetes diszkrét eloszlások

- **Bernoulli:** $X \sim \text{Ber}(p)$ ($0 < p < 1$)

Egy kísérletet végzünk, ami p valószínűséggel sikeres. Ekkor X az 1-es értéket veszi fel, ha a kísérlet sikeres, egyébként nullát.

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

- **Binomiális:** $X \sim \text{Bin}(n, p)$ ($n \in \mathbb{N}^+, 0 < p < 1$)

Végezzünk n darab kísérletet, amik egymástól függetlenül azonos p valószínűséggel sikeresek. Ekkor X a sikeres kísérletek száma.

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n.$$

- **Geometriai:** $X \sim \text{Geo}(p)$ ($0 < p < 1$)

Egy kísérletet, ami p valószínűséggel sikeres addig ismételünk egymástól függetlenül, amíg sikeres nem lesz. Ekkor X a szükséges próbálkozások száma.

$$p(k) = (1-p)^{k-1} p, \quad 1 \leq k < \infty.$$

- **Poisson:** $X \sim \text{Poi}(\lambda)$ ($\lambda > 0$)

Amennyiben n elég nagy és p elég kicsi, akkor $\text{Bin}(n, p)$ jól közelíthető $\text{Poi}(np)$ eloszlással.

$$p(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad 0 \leq k < \infty.$$

Tétel Legyen $X \sim \text{Bin}(n, p_n)$. Amennyiben $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda > 0$, akkor minden $k \in \mathbb{N}$ esetén

$$\lim_{n \rightarrow \infty} \mathbb{P}(X = k) = \mathbb{P}(Y = k),$$

ahol $Y \sim \text{Poi}(\lambda)$.

Feladatok

1. Dobjunk fel egy szabályos érmét egymástól függetlenül kétszer. Legyen X a fejek száma.
 - (a) Írjuk fel az eseményteret és X -et, mint $X : \Omega \rightarrow \mathbb{R}$ függvényt.
 - (b) Milyen eloszlást követ X ?

Megoldás

- (a) Egy jó felírása az eseménytérnek

$$\Omega = \{FF, IF, FI, II\}.$$

Ekkor az X valószínűségi változó, mint $X : \Omega \rightarrow \mathbb{R}$ függvény.

$$X(FF) = 2, \quad X(IF) = 1, \quad X(FI) = 1, \quad X(II) = 0.$$

- (b) Mivel egy $\frac{1}{2}$ valószínűséggel sikeres (fejet dobunk) kísérletet egymástól függetlenül 2-szer ismétlünk meg, így amennyiben X a sikeres kísérletek száma, akkor

$$X \sim \text{Bin}\left(2, \frac{1}{2}\right).$$

2. Dobjunk két 10 oldalú dobókockával, jelölje az eredményeiket X és Y . Mennyi $\mathbb{P}(X \leq Y)$?

Megoldás Vegyük észre, hogy kizáró eseményekre bontással

$$\mathbb{P}(X \leq Y) = \mathbb{P}(X < Y) + \mathbb{P}(X = Y). \quad (1)$$

Továbbá mivel X és Y szerepe teljesen szimmetrikus

$$\mathbb{P}(X < Y) = \mathbb{P}(Y < X). \quad (2)$$

Másfelől az eseményteret diszjunkt eseményekre bontva és kihasználva a (2)-es azonosságot

$$\begin{aligned} 1 &= \mathbb{P}(X < Y) + \mathbb{P}(Y < X) + \mathbb{P}(X = Y) \\ &= 2\mathbb{P}(X < Y) + \mathbb{P}(X = Y). \end{aligned} \quad (3)$$

Továbbá könnyen számolható, hogy

$$\mathbb{P}(X = Y) = \mathbb{P}(\text{mindkét kockán ugyanaz a szám szerepel}) = \frac{10}{10 \cdot 10} = \frac{1}{10},$$

amit visszahelyettesítve a (3)-as egyenletbe

$$\mathbb{P}(X < Y) = \frac{9}{20}.$$

A fenti két eredményt behelyettesítve az (1)-es egyenletbe, azt kapjuk, hogy

$$\mathbb{P}(X \leq Y) = \frac{11}{20}.$$

3. Egy boltban izzókat árulnak. Az izzók 1%-a hibás. Ha veszünk 100 darabot, akkor
 - (a) mekkora eséllyel lesz legfeljebb három hibás?
 - (b) hány lesz közülük rossz a legnagyobb valószínűséggel?

Megoldás

- (a) Vegyük észre, hogy 100 darab kísérletet végzünk (megnézzük, hogy az adott izzó hibás vagy nem) egymástól függetlenül. Egy kísérlet akkor "sikeres", ha az adott izzó hibás, ami 0.01 valószínűséggel következik be. Amennyiben X -el jelöljük a hibás izzók számát, akkor ez alapján

$$X \sim \text{Bin}(100, 0.01).$$

A keresett valószínűség ekkor

$$\begin{aligned} \mathbb{P}(X \leq 3) &= \sum_{k=0}^3 \mathbb{P}(X = k) \\ &= \sum_{k=0}^3 \binom{100}{k} 0.01^k 0.99^{100-k} \\ &\approx 0.9816 \end{aligned}$$

- (b) Azt kell megkeresni, hogy $\mathbb{P}(X = k)$ milyen $0 \leq k \leq 100$ esetén lesz maximális. Keressük meg, hogy milyen k esetén csökken a valószínűség, ha eggyel növeljük k értékét.

$$\begin{aligned} \mathbb{P}(X = k+1) &< \mathbb{P}(X = k) \\ \binom{100}{k+1} 0.01^{k+1} 0.99^{100-(k+1)} &< \binom{100}{k} 0.01^k 0.99^{100-k} \\ \frac{100!}{(k+1)!(100-(k+1))!} 0.01^{k+1} 0.99^{100-(k+1)} &< \frac{100!}{k!(100-k)!} 0.01^k 0.99^{100-k} \\ (100-k)0.01 &< 0.99(k+1) \\ 0 &< k. \end{aligned}$$

Vagyis minden $1 \leq k$ esetén csökkeni fog a valószínűség, ha növeljük k -t. A maximumot tehát $k = 0$ vagy $k = 1$ esetén veszi fel. Továbbá látható, hogy

$$\mathbb{P}(X = 0) = \underbrace{\binom{100}{0} 0.01^0 0.99^{100}}_{0.99^{100}} < \underbrace{\binom{100}{1} 0.01^1 0.99^{99}}_{0.99^{99}} = \mathbb{P}(X = 1),$$

így az izzók között 1 darab lesz rossz a legnagyobb valószínűséggel.

4. Hányszor dobjunk egy kockával, hogyha azt akarjuk, hogy 0.5-nél ne legyen kisebb annak a valószínűsége, hogy a 6-os dobások száma legalább kettő?

Megoldás Vegyük észre, hogy mivel egy kockával $\frac{1}{6}$ valószínűséggel dobunk 6-ost, így amennyiben ezt n -szer egymástól függetlenül megismételjük, akkor amennyiben X a hatos dobások száma:

$$X \sim \text{Bin}\left(n, \frac{1}{6}\right).$$

A feladat azt kérdezi, hogy mi legyen n értéke, ha azt akarjuk, hogy

$$\frac{1}{2} \leq \mathbb{P}(2 \leq X).$$

Ezt átrendezve a komplementer esemény segítségével, azt kapjuk, hogy

$$\begin{aligned} \frac{1}{2} &\leq 1 - \mathbb{P}(X < 2) \\ &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - \binom{n}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^n - \binom{n}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{n-1} \\ &= 1 - \left(\frac{5}{6}\right)^n - \frac{n}{6} \left(\frac{5}{6}\right)^{n-1} \\ &= 1 - \left(\frac{5}{6}\right)^n \left(1 + \frac{n}{5}\right) \end{aligned}$$

Itt érdemes behelyettesíteni. Így látható, hogy $n = 9$ -re még nem teljesül az egyenlőtlenség, de $n = 10$ -re már igen. Azt kaptuk, hogy 10-szer kell dobunk.

5. Hét közben (hétfőtől péntekig) 0.25 valószínűséggel késik a reggeli vonat.
- Mennyi a valószínűsége, hogy hét közben legfeljebb két napon késik?
 - Legyen jó hét az, amelyen legfeljebb két napon késik. Mennyi a valószínűsége, hogy 4 hétből legalább 3 hét jó hét lesz?

Megoldás

- (a) Mivel egy napon egymástól függetlenül 0.25 valószínűséggel késik a vonat, így amennyiben X jelöli a hétfőtől péntekig késő vonatok számát, akkor

$$X \sim \text{Bin}(5, 0.25).$$

Ekkor a keresett valószínűség

$$\begin{aligned} \mathbb{P}(X \leq 2) &= \sum_{k=0}^2 \mathbb{P}(X = k) \\ &= \sum_{k=0}^2 \binom{5}{k} 0.25^k 0.75^{5-k} \\ &= \frac{459}{512} \approx 0.8965 \end{aligned}$$

- (b) Az előző részben kiszámoltuk, hogy mekkora a valószínűsége, hogy egy hét jó lesz, így amennyiben Y jelöli a jó hetek számát 4 hétből, akkor

$$Y \sim \text{Bin}\left(4, \frac{459}{512}\right).$$

Innen a keresett valószínűség

$$\begin{aligned} \mathbb{P}(3 \geq Y) &= \mathbb{P}(Y = 3) + \mathbb{P}(Y = 4) \\ &= \binom{4}{3} \left(\frac{459}{512}\right)^3 \left(\frac{53}{512}\right)^1 + \binom{4}{4} \left(\frac{459}{512}\right)^4 \left(\frac{53}{512}\right)^0 \\ &\approx 0.9442. \end{aligned}$$

6. Egy társasjátékban két kockával dobok egyszerre. Egy olyan mezőre kerültem, ahonnan csak akkor tudok továbblépni, ha a dobott számok között szerepel a hatos. Mekkora valószínűséggel tudok két próbálkozáson belül továbblépni?

Megoldás Annak a valószínűsége, hogy szerepel hatos

$$\mathbb{P}(\text{dobunk hatost}) = 1 - \mathbb{P}(\text{nem dobunk hatost}) = 1 - \frac{5^2}{6^2} = \frac{11}{36}.$$

Addig próbálkozunk egymástól függetlenül, amíg a dobott számok között nem szerepel a hatos. Amennyiben X jelöli a szükséges próbálkozások számát, akkor

$$X \sim \text{Geo}\left(\frac{11}{36}\right).$$

A keresett valószínűség

$$\begin{aligned} \mathbb{P}(X \leq 2) &= \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \frac{11}{36} + \frac{25}{36} \cdot \frac{11}{36} \\ &= \frac{671}{1296} \approx 0.5177. \end{aligned}$$

7. Találomra választunk egymástól függetlenül két számot a $[0, 1]$ intervallumon. Ezt addig ismételjük egymástól függetlenül, amíg az egyik szám több mint kétszerese nem lesz a másiknak. Mennyi annak a valószínűsége, hogy ez legfeljebb 3 lépésben sikerül.

Megoldás Mivel egy kísérletet (két számot választunk) egymástól függetlenül ismételünk, az első sikerig, így amennyiben X a szükséges választások száma, akkor

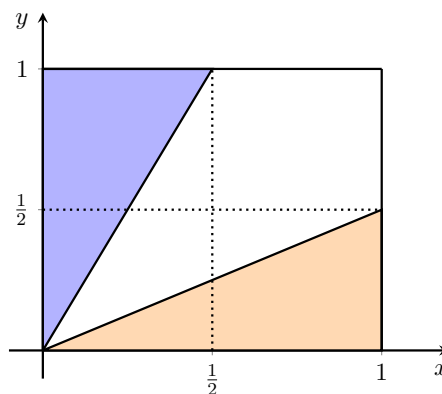
$$X \sim \text{Geo}(p),$$

ahol p annak a valószínűsége, hogy egy lépésben az egyik szám több mint kétszerese lesz a másiknak. Az, hogy két számot választunk egymástól függetlenül egyenletesen a $[0, 1]$ intervallumról modellezhető pontosan úgy, hogy egy pontot választunk a síkon egyenletesen az $\Omega = [0, 1] \times [0, 1]$ négyzetről.

A kísérlet akkor lesz sikeres, ha a pont az alábbi részhalmazába esik Ω -nak

$$\begin{aligned} A &= \{(x, y) \in \Omega : x > 2y \text{ vagy } y > 2x\} \\ &= \{(x, y) \in \Omega : x > 2y\} \cup \{(x, y) \in \Omega : y > 2x\}. \end{aligned}$$

Ábrázolva



Ekkor p nem lesz más, mint A területének és Ω területének a hányadosa, mivel úgynevezett geometria valószínűségi mezővel dolgozunk (ennek nincs köze, ahhoz, hogy X geometriai eloszlást követ). Vagyis

$$p = \frac{\text{terület}(A)}{\text{terület}(\Omega)} = \frac{\frac{1}{2}}{1} = \frac{1}{2}.$$

Azt kaptuk, hogy

$$X \sim \text{Geo}\left(\frac{1}{2}\right).$$

A keresett valószínűség ekkor

$$\begin{aligned} \mathbb{P}(X \leq 3) &= \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) \\ &= \frac{1}{2} + \left(1 - \frac{1}{2}\right) \frac{1}{2} + \left(1 - \frac{1}{2}\right)^2 \frac{1}{2} \\ &= \frac{7}{8}. \end{aligned}$$

8. Egy szabályos pénzérmét addig dobunk fel újra és újra, amíg meg nem kapjuk a második fejet is. Mennyi annak a valószínűsége, hogy az első fej után a második fejig ugyanannyi dobásra van szükség, mint amennyi az elsőig kellett?

Megoldás Legyen X az első fejig és Y az első fej után a második fejig szükséges próbálkozások száma. Ekkor tudjuk, hogy

$$X \sim \text{Geo}\left(\frac{1}{2}\right) \quad \text{és} \quad Y \sim \text{Geo}\left(\frac{1}{2}\right).$$

A keresett valószínűség ekkor

$$\mathbb{P}(X = Y).$$

Bontsuk fel kizáró eseményekre

$$\mathbb{P}(X = Y) = \sum_{k=1}^{\infty} \mathbb{P}(X = Y = k) = \sum_{k=1}^{\infty} \mathbb{P}(\{X = k\} \cap \{Y = k\}).$$

Vegyük észre, hogy az $\{X = k\}$ és $\{Y = k\}$ események függetlenek, hiszen az, hogy k dobás kell a második fejhez az első fej után független attól, hogy az első fej a k -adik dobásra sikerült.

Használva a függetlenséget és kihasználva, hogy X és Y geometriai eloszlású $\frac{1}{2}$ paraméterrel

$$\begin{aligned} \mathbb{P}(X = Y) &= \sum_{k=1}^{\infty} \mathbb{P}(\{Y = k\} \cap \{X = k\}) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(X = k) \mathbb{P}(Y = k) \\ &= \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{1}{2^k} \\ &= \sum_{k=1}^{\infty} \frac{1}{4^k} \\ &= \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} \\ &= \frac{1}{3}. \end{aligned}$$

9. Lovas gátversenyen a ló az akadályok mindegyikét egymástól függetlenül, azonos valószínűséggel veri le. A pályán számtalan akadály akad. Ha 5% annak a valószínűsége, hogy a lovas hibátlanul teljesít egy kört, mennyi az esélye, hogy legfeljebb három akadályt ver le?

Megoldás Legyen X a levert akadályok száma. Mivel a pályán nagyon sok akadály van, amelyeket egyenként egymástól függetlenül azonos kis valószínűséggel ver le a lovas, így X jól közelíthető Poisson eloszlással

$$X \sim \text{Poi}(\lambda), \quad \text{ahol } \lambda > 0.$$

Mivel a feladat szerint annak a valószínűsége, hogy egy kört hibátlanul teljesít a lovas $5\% = \frac{1}{20}$, illetve ez az esemény pontosan $X = 0$ azt kapjuk, hogy

$$\begin{aligned} \frac{1}{20} &= \mathbb{P}(X = 0) \\ &= e^{-\lambda} \cdot \frac{\lambda^0}{0!} \\ &= e^{-\lambda} \end{aligned}$$

Mindkét oldal logaritmusát véve azt kapjuk, hogy

$$\lambda = \ln(20).$$

Innen a keresett valószínűség

$$\begin{aligned} \mathbb{P}(X \leq 3) &= \sum_{k=0}^3 \mathbb{P}(X = k) \\ &= \sum_{k=0}^3 e^{-\ln(20)} \cdot \frac{\ln(20)^k}{k!} \\ &= \frac{1}{20} \sum_{k=0}^3 \frac{\ln(20)^k}{k!} \\ &\approx 0.6472. \end{aligned}$$

10. Átlagosan hány mazsolának kell egy sütiben lennie, ha azt kívánjuk elérni, hogy egy véletlenszerűen választott sütiben legalább 99% eséllyel legyen (legalább egy szem) mazsola?

Megoldás Legyen X a sütiben található mazsolák száma. Mivel feltehető, hogy sok mazsola található és egy mazsola egymástól függetlenül kis valószínűséggel kerül egy sütibe, így

$$X \sim \text{Poi}(\lambda).$$

Ekkor a feladat X átlagos értékét kérdezi (várható érték - következő gyakorlaton részletesebben), ami egy Poisson eloszlás esetén λ .

A feltétel szerint

$$\begin{aligned} \frac{99}{100} &\leq \mathbb{P}(X \geq 1) \\ &= 1 - \mathbb{P}(X < 1) \\ &= 1 - \mathbb{P}(X = 0) \\ &= 1 - e^{-\lambda}. \end{aligned}$$

Ekkor átrendezve

$$e^{-\lambda} \leq \frac{1}{100}.$$

Ekkor kihasználva, hogy $x \mapsto \ln(x)$ monoton függvény, azt kapjuk, hogy

$$\begin{aligned} -\lambda &\leq -\ln(100) \\ \ln(100) &\leq \lambda \end{aligned}$$

Mivel $\ln(100) \approx 4.6052$, így átlagosan legalább ennyi mazsolát kell tenni egy sütibe.

11. Egy futóversenyen a pályát sajnos kullancsal fertőzött területen át vezették. Kiderült, hogy a versenyzők közül 300-an találtak magukban pontosan egy kullancsot, 75-en pedig kettőt. Ezek alapján becsüljük meg, hogy körülbelül hányan indultak a versenyen.

Megoldás Legyen N a versenyzők száma. Továbbá legyen X az egy versenyzőben található kullancsok száma. Mivel feltehető, hogy a pálya mentén számtalan kullancs található és ezek egymástól függetlenül azonos kis valószínűséggel csípnék bele a versenyzőbe, ezért X jól közelíthető Poisson eloszlással;

$$X \sim \text{Poi}(\lambda), \quad \text{ahol } \lambda > 0.$$

Tudjuk továbbá tudjuk azon versenyzők arányát akikbe pontosan egy, illetve kettő kullancs csípett bele. Ezeknek az aránya az összes futóhoz viszonyítva jó empirikus becslése annak a valószínűségnek, hogy egy futóba pontosan egy, illetve két kullancs található a verseny végén. Így az alábbi egyenleteket kapjuk

$$\begin{aligned} \mathbb{P}(X = 1) &= \frac{300}{N} \\ \mathbb{P}(X = 2) &= \frac{75}{N} \end{aligned}$$

Kihasználva, hogy X Poisson eloszlású

$$\begin{aligned} e^{-\lambda} \cdot \lambda &= \frac{300}{N} \\ e^{-\lambda} \cdot \frac{\lambda^2}{2} &= \frac{75}{N} \end{aligned}$$

A második egyenletet leosztva az elsővel, azt kapjuk, hogy

$$\frac{\lambda}{2} = \frac{75}{300},$$

amiből

$$\lambda = \frac{1}{2}.$$

Végül az első egyenletből kifejezve N -et

$$N = 2 \cdot \sqrt{e} \cdot 300 \approx 998.2.$$

Azt kaptuk, hogy körülbelül 998 futó indult a versenyen.

12. Egy 400 oldalas könyvben 200 db sajtóhiba található. Mennyi a valószínűsége, hogy a 13. oldalon egyenél több hiba lesz?

Megoldás Legyen X a 13. oldalon található hibák száma. Ekkor mivel minden összesen 200 hiba található a könyvben és feltehető, hogy mindegyik hiba egymástól függetlenül azonos valószínűséggel ($\frac{1}{400}$) esik egy oldalra, így

$$X \sim \text{Bin}\left(200, \frac{1}{400}\right).$$

Ekkor a keresett valószínűség

$$\begin{aligned} \mathbb{P}(X > 1) &= 1 - \mathbb{P}(X \leq 1) \\ &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - \left(\frac{399}{400}\right)^{200} - 200 \left(\frac{399}{400}\right)^{199} \frac{1}{400} \\ &\approx 0.0900. \end{aligned}$$

Egy másik jó megoldás, ha X -et Poisson eloszlással közelítjük, hiszen sok hiba van és azok egymástól függetlenül kis valószínűséggel kerülnek egy oldalra. Ekkor

$$X \sim \text{Poi}(\lambda),$$

ahol

$$\lambda = 200 \cdot \frac{1}{400} = \frac{1}{2}.$$

Mivel ebben az esetben egy $\text{Bin}(n, p)$ jól közelíthető $\text{Poi}(n \cdot p)$ eloszlással. Tudjuk azt is, hogy egy Poisson eloszlású valószínűségi változó várható értéke a paramétere, és egy oldalra várhatóan $\frac{200}{400}$ hiba esik. Ekkor a valószínűség, amit keresünk

$$\begin{aligned} \mathbb{P}(X > 1) &= 1 - \mathbb{P}(X \leq 1) \\ &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - e^{-\frac{1}{2}} - \frac{1}{2}e^{-\frac{1}{2}} \\ &= 1 - \frac{1}{\sqrt{e}} - \frac{1}{2\sqrt{e}} \\ &\approx 0.0902. \end{aligned}$$

Látható, hogy a két modell hasonló eredményre vezetett (három tizedesjegyre megegyeznek a számok).