

Matematika A4 - Valószínűségszámítás 3. gyakorlat

Valószínűségi változók, súlyfüggvény, nevezetes diszkrét eloszlások

Valószínűségi változók

A **valószínűségi változók** az eseménytéren értelmezett függvények. Egy X valószínűségi változó **diszkrét**, ha csak véges sok vagy megszámlálhatóan végtelen sok értéket vehet fel. Ennek a **súlyfüggvénye** az a p_X függvény, amire $p_X(x_i) = \mathbb{P}(X = x_i)$, ahol $x_1, x_2, \dots$ az X változó lehetséges értékei.

Azt mondjuk, hogy az X és Y valószínűségi változók **függetlenek**, ha $\{X \in A\}$, $\{Y \in B\}$ események függetlenek minden $A, B \subseteq \mathbb{R}$ halmazra.

Nevezetes eloszlások

Bernoulli: X Bernoulli eloszlású p paraméterrel, ha egy p valószínűséggel sikeres kísérlet eredményét jellemzi, azaz p valószínűséggel 1, $1 - p$ valószínűséggel 0 az értéke. Jelölés: $X \sim \text{BER}(p)$.

Binomiális: Ha egy kísérletet n -szer végzünk el egymástól függetlenül, ahol a siker valószínűsége p és X jelöli a sikeres kísérletek számát, akkor X eloszlása binomiális n, p paraméterekkel. Jelölés: $X \sim \text{BIN}(n, p)$ vagy $X \sim \text{B}(n, p)$. A súlyfüggvénye $k = [(n + 1)p]$ -ben maximális.

Ha $X \sim \text{BIN}(n, p)$, akkor n darab, független $\text{BER}(p)$ eloszlású valószínűségi változó összege.

Poisson: Sok, kis valószínűségű, független kísérletből a sikeresek számát Poisson eloszlású valószínűségi változóval jellemezhetjük. Ha a sikeres kísérletek várható száma λ , akkor X valószínűségi változó eloszlása λ -paraméterű Poisson eloszlás. Jelölés: $X \sim \text{POI}(\lambda)$. A súlyfüggvénye $k = [\lambda]$ -ban maximális. Ha n elég nagy és p kicsi, akkor $\text{BIN}(n, p)$ jól közelíthető $\text{POI}(np)$ eloszlással.

Geometriai: Egy kísérletet végezzünk el végtelen sokszor úgy, hogy a kísérletek egymástól függetlenek. Ha X a kísérletek száma az első sikerig (a sikert is beleszámítva), akkor X geometriai eloszlású p paraméterrel, ahol p egy kísérlet sikerességének valószínűsége. Jelölés: $X \sim \text{GEO}(p)$.

Negatív binomiális: Ha X azt adja meg, hogy mikor következik be az r -edik sikeres kísérlet, akkor X negatív binomiális eloszlású p és r paraméterekkel. Jelölés: $X \sim \text{NB}(p, r)$.

Hipergeometriai: Van N termékünk, amiből M hibás. Kihúzzunk n -et, és legyen X a kihúzott hibások száma. Ekkor X hipergeometriai eloszlású N, M, n paraméterekkel. Jelölés: $X \sim \text{HGEO}(N, M, n)$.

X eloszlása	lehetséges értékek	$p_X(k)$ súlyfüggvény
$\text{BER}(p)$	$k = 0, 1$	$p_X(0) = 1 - p, p_X(1) = p$
$\text{BIN}(n, p)$	$k = 0, 1, \dots, n$	$\binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$
$\text{POI}(\lambda)$	$k = 0, 1, 2, \dots$	$\frac{\lambda^k}{k!} \cdot e^{-\lambda}$
$\text{GEO}(p)$	$k = 1, 2, 3, \dots$	$(1 - p)^{k-1} \cdot p$
$\text{NB}(p, r)$	$k = r, r + 1, \dots$	$\binom{k-1}{r-1} p^r (1 - p)^{k-r}$
$\text{HGEO}(N, M, n)$	$\max(0, n - N + M) \leq k \leq \min(n, M)$	$\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$

1. feladat Egy adóhatóság a cégek 5%-át vizsgálja meg évente teljesen véletlenszerűen. Mi a valószínűsége, hogy egy adott céget öt év alatt

(a) legalább egyszer megvizsgálják?

(b) legalább kétszer megvizsgálják?

Megoldásvázlat Legyen X a céget öt év alatt érintő vizsgálatok száma. Ekkor ennek eloszlása $\text{BIN}(5, 0.05)$, azaz $X \sim \text{BIN}(5, 0.05)$. Így komplementerekkel számolva:

$$(a) \mathbb{P}(X > 0) = 1 - 0.95^5 \approx 0.22$$

$$(b) \mathbb{P}(X > 1) = 1 - 0.95^5 - 5 \cdot 0.95^4 \cdot 0.05 \approx 0.023$$

2. feladat Egy 400 oldalas könyvben 200 db sajtóhiba található. Mennyi a valószínűsége, hogy a 13. oldalon egynél több hiba lesz? Indokoljuk meg a használt modellt!

Megoldásvázlat Minden hiba $\frac{1}{400}$ valószínűséggel esik a 13. oldalra, így a 13. oldalon a hibák száma $\text{BIN}(200, \frac{1}{400})$. Legyen X a 13. oldalon lévő hibák száma! Annak a valószínűsége, hogy több, mint 1 hiba van a 13. oldalon:

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

Másképp: Sok hiba van, amik egymástól függetlenül helyezkednek el az oldalakon és egy adott hiba kis valószínűséggel van épp a 13. oldalon. Ezért az egy oldalon lévő hibák száma jól közelíthető Poisson eloszlással. Mivel egy oldalon átlagosan $\frac{1}{2}$ hiba van, így egy adott oldalon a hibák számának eloszlása közelítőleg $\text{POI}(\frac{1}{2})$. Ekkor

$$\mathbb{P}(X > 1) = 1 - \sum_{k=0}^1 p_X(k) = 1 - \sum_{k=0}^1 e^{-1/2} \cdot \frac{(\frac{1}{2})^k}{k!} = 1 - \frac{3}{2} \cdot e^{-1/2} \approx 0.0902.$$

Megjegyzés: A második megoldásban igazából azt a tételt használtuk, hogy ha n nagy és p kicsi, akkor $\text{BIN}(n, p)$ jól közelíthető $\text{POI}(n \cdot p)$ eloszlással. Most $n = 200$, $p = \frac{1}{400}$, így $\text{BIN}(200, \frac{1}{400})$ eloszlást $\text{POI}(\frac{1}{2})$ eloszlással tudtuk közelíteni.

3. feladat Tegyük fel, hogy egy egyetem 2400 hallgatója éjjel-nappal kétóránként ellenőrzi az email-jét és egy ellenőrzés 1 percet vesz igénybe (tehát $1/120$ annak a valószínűsége, hogy egy adott hallgató éppen az emailjével van elfoglalva). Mi a közelítő valószínűsége, hogy éppen 19, 20 vagy 21 hallgató ellenőrzi az emailjét?

Megoldásvázlat Azon hallgatók száma, akik épp az emailjüket ellenőrzik $\text{BIN}(2400, \frac{1}{120})$ eloszlású, ami jól közelíthető $\text{POI}(20)$ eloszlással. Így a kért közelítő valószínűség:

$$e^{-20} \cdot \left(\frac{20^{19}}{19!} + \frac{20^{20}}{20!} + \frac{20^{21}}{21!} \right) \approx 0.26.$$

4. feladat Egy sportlövő 0.9 valószínűséggel lő 10-est. Egy lövéssorozat 10 lövésből áll és akkor elfogadható számára, ha legalább 9 tizes találat van benne. Mi a valószínűsége, hogy három lövéssorozat mindegyike elfogadható lesz?

Megoldásvázlat Egy lövéssorozatban a 10-es lövések száma $\text{BIN}(10, 0.9)$ eloszlású. Így annak a valószínűsége, hogy egy lövéssorozat sikeres $0.9^{10} + 10 \cdot 0.9^9 \cdot 0.1 \approx 0.736$. Annak a valószínűsége, hogy mindhárom lövéssorozat sikeres $0.736^3 \approx 0.399$.

5. feladat A roulette keréken 38 szám van: 1-től 36-ig, 0 és dupla 0. Ha Kovács úr mindig arra fogad, hogy az eredmény 1-től 12-ig terjed, mi a valószínűsége, hogy

(a) Kovács úr elveszti az első öt fogadását,

(b) először a negyedik fogadáson nyer?

Megoldásvázlat

$$(a) \left(\frac{26}{38}\right)^5$$

$$(b) \left(\frac{26}{38}\right)^3 \cdot \frac{12}{38}$$

6. feladat 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 próbálkozásán belül továbblépni?

Megoldásvázlat Minden próbálkozásnál $\frac{11}{36}$ valószínűséggel járok sikerrel, így a próbálkozások számának az eloszlása $\text{GEO}\left(\frac{11}{36}\right)$. Így a kért valószínűség:

$$\frac{11}{36} + \frac{25}{36} \cdot \frac{11}{36}.$$

7. feladat Az AB-negatív vércsoport gyakorisága 1%. Ha 100 embert véletlenszerűen kiválasztunk, akkor

- (a) mi a közelítő valószínűsége (Poisson eloszlást használva), hogy közülük legalább egynek a vércsoportja AB-negatív?
- (b) mi a pontos valószínűsége, hogy közülük legalább egynek a vércsoportja AB-negatív?

Megoldásvázlat Az AB-vércsoportú emberek száma $\text{BIN}(100, 0.01)$, azaz $\text{POI}(1)$ eloszlással közelíthető. Így

- (a) $1 - e^{-1} \approx 0.63$
- (b) $1 - 0.99^{100} \approx 0.63$

8. feladat Hétközben (hétfőtől péntekig) 0.25 valószínűséggel késik a reggeli vonat.

- (a) Mi a valószínűsége, hogy hétközben legfeljebb két napon késik?
- (b) Legyen jó hét az, amelyen legfeljebb két napon késik. Mi a valószínűsége, hogy 4 hétből legalább 3 hét jó hét lesz?

Megoldásvázlat

- (a) Héten a késések számának eloszlása $\text{BIN}(7, 0.25)$. Annak a valószínűsége, hogy legfeljebb 2-szer késik: $p := 0.75^7 + 7 \cdot 0.75^6 \cdot 0.25 + \binom{7}{2} \cdot 0.75^5 \cdot 0.25^2 \approx 0.9$.
- (b) 4 hét alatt a jó hetek számának eloszlása $\text{BIN}(4, p)$. Így annak a valószínűsége, hogy legalább 3 hét jó: $p^4 + 4 \cdot p^3 \cdot (1 - p) \approx 0.95$.

9. feladat Az A érme feldobásakor 0.4 valószínűséggel kapunk fejet, a B érme feldobásakor ugyanez a valószínűség 0.7. A két érme közül egyet véletlenszerűen kiválasztunk és 10-szer feldobjuk.

- (a) Függetlenek-e a dobások kimenetelei egymástól?
- (b) Mi a valószínűsége, hogy pontosan 7 dobás lesz fej?
- (c) Feltéve, hogy az első dobás fej, mi a valószínűsége, hogy az A érmevel dobunk?
- (d) Feltéve, hogy az első dobás fej, mi a valószínűsége, hogy összesen 7 dobás lesz fej?

Megoldásvázlat Legyen A és B az az esemény, hogy A vagy B érmevel dobunk. E , F legyen az az esemény, hogy az első, illetve a második dobás fej. C legyen az, hogy pontosan 7 fejet dobtunk.

- (a) $\mathbb{P}(E) = \mathbb{P}(F) = 0.55$ (teljes valószínűség tételével A , B teljes eseményrendszerre nézve). Hasonlóan $\mathbb{P}(E \cap F) = 0.325 \neq \mathbb{P}(E) \cdot \mathbb{P}(F)$, így nem függetlenek.
- (b) Teljes valószínűség tételével (A , B teljes eseményrendszerre nézve) és használva, hogy a fejek száma binomiális eloszlású $n = 10$ és az érmétől függően $p = 0.4$ vagy $p = 0.7$ paraméterekkel: $\mathbb{P}(C) \approx 0.15$.
- (c) Bayes-tétellel: $\mathbb{P}(A | E) = \frac{4}{11}$.

- (d) $\mathbb{P}(C|E) = \frac{\mathbb{P}(C \cap E)}{\mathbb{P}(E)}$, ahol $\mathbb{P}(C \cap E)$ teljes valószínűség tétellel számolható (azt használva, hogy $C \cap E$ az az esemény, hogy az első dobás fej, majd a maradék 9-ből pontosan 6):

$$\begin{aligned}\mathbb{P}(C \cap E) &= \mathbb{P}(C \cap E | A) \cdot \mathbb{P}(A) + \mathbb{P}(C \cap E | B) \cdot \mathbb{P}(B) \\ &= 0.4 \cdot \binom{9}{6} \cdot 0.4^6 \cdot 0.6^3 \cdot 0.5 + 0.7 \cdot \binom{9}{6} \cdot 0.7^6 \cdot 0.3^3 \cdot 0.5 \approx 0.108.\end{aligned}$$

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

Megoldásvázlat Legyen X egy sütiben a mazsolák száma! Ekkor X eloszlása jól közelíthető Poisson eloszlással valamilyen λ paraméterrel. Ekkor $\mathbb{P}(X \geq 1) = 1 - e^{-\lambda}$. Nekünk olyan λ kell, amire ez a valószínűség legalább 0.99, azaz

$$1 - e^{-\lambda} \geq 0.99.$$

Ezt átrendezve $\lambda \geq -\ln 0.01 \approx 4.61$. Ha a mazsolák számának eloszlása $\text{POI}(\lambda)$, akkor várhatóan λ mazsola kerül egy sütibe, így átlagosan legalább 4.61 mazsolát kell beletenni.

11. feladat Egy futóversenyt sajnos kullancsok fertőzött területen át vezettek. Kiderült, hogy a versenyzők közül 300-an találtak magukban egy, 75-en pedig két kullancsot. Ennek alapján becsüljük meg, hogy körülbelül hányan indultak a versenyen.

Megoldásvázlat Egy versenyzőben a kullancsok száma $\text{POI}(\lambda)$ eloszlású ismeretlen λ -ra. Ekkor, ha N az indulók száma, akkor a feltételekből $N \cdot \lambda \cdot e^{-\lambda} = 300$ és $N \cdot \frac{\lambda^2}{2} \cdot e^{-\lambda} = 75$. Ebből $\lambda = 0.5$ és $N \approx 989$.

12. feladat Egy osztályban 22 tanuló van. Egy órára 8-an nem készültek, és 7-en felelnek. Adjuk meg a készületlen felelők számának eloszlását! Mennyi a valószínűsége, hogy pontosan 2 készületlen felelő lesz?

Megoldásvázlat A készületlen felelők számának eloszlása $\text{HGEO}(22, 8, 7)$. Így annak a valószínűsége, hogy pontosan 2 készületlen felelő lesz:

$$\frac{\binom{8}{2} \cdot \binom{14}{5}}{\binom{22}{7}}$$