Valószínűségszámítás

levelezős minta ZH

VEMKMA1344B – L1
2023.05.09.

NÉV: ______________________________

NEPTUN KÓD: ______________________

---

1. Az ötös lottón 90 számból 5-öt húznak ki. Ha kitöltünk 1 db lottószelvényt, mekkora az esélye, hogy kevesebb, mint két találatunk lett?
2. Választunk két számot a $[-1, 3]$ intervallumról egymástól függetlenül a geometriai valószínűség szerint. Mennyi az esélye, hogy mindkét szám negatív?
3. A személyi jövedelemadó bevallások 40%-át könyvelők, 35%-át munkahelyi munkaügyi alkalmazottak és 25%-át maguk az adóalanyok készítik el. A könyvelők által készített bevallások 0.05, a munkaügyi alkalmazottak által készített bevallások 0.07, az adóalanyok által készített bevallások 0.3 valószínűséggel tartalmaznak hibát.
   (a) Kiválasztva egy bevallást mennyi az esélye, hogy hibát tartalmaz?
   (b) Ha a kiválasztott bevallás tartalmaz hibát, mennyi az esélye, hogy az adóalany készítette?
4. Magyar kártyából visszatevéssel húzunk 3 lapot. Legyen X a húzott piros lapok száma. Adjuk meg X eloszlását! Mennyi X várható értéke és szórása?
5. Addig dobunk egy kockával, amíg először hatost kapunk. Legyen X a szükséges dobások száma. Adjuk meg X eloszlását! Mennyi az esélye, hogy legfeljebb 3 dobás szükséges? Tegyük fel, hogy az első 10 dobás egyike sem volt hatos. Mennyi az esélye, hogy az első 12 dobásból sem lesz hatos?
6. Legyen $X \sim N(0, 1)$ eloszlású val. változó. Adjunk olyan 0-ra szimmetrikus intervallumot, amibe X értéke 0.8 valószínűséggel beleesik!
7. 100-szor elgurítunk egy szabályos kockát. A centrális határeloszlás-tétel segítségével adja meg közelítőleg annak az esélyét, hogy a gurítások összege legalább 330, de kevesebb, mint 365 lesz!

Megoldások

1. Az ötös lottón 90 számból 5-öt húznak ki. Ha kitöltünk 1 db lottószelvényt, mekkora az esélye, hogy kevesebb, mint két találatunk lett?

   Megoldás. Tekintsük az alábbi val. változót: $X =$ a találatok száma. Könnyen végiggondolható (gyakorlaton is kiszámoltuk), hogy ebben az esetben X egy tetszőleges k értéket az alábbi p_k valószínűséggel vesz fel:

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

   A feladat arra kíváncsi, hogy mennyi $P(X < 2)$. Diszkrét eloszlásról lévén szó,

   $$P(X < 2) = P(X = 0) + P(X = 1) = p_0 + p_1 = \frac{\binom{5}{0}\binom{85}{5}}{\binom{90}{5}} + \frac{\binom{5}{1}\binom{85}{4}}{\binom{90}{5}}.$$

   □

2. Választunk két számot a $[-1, 3]$ intervallumról egymástól függetlenül a geometriai valószínűség szerint. Mennyi az esélye, hogy mindkét szám negatív?

   Megjegyzés. VIGYÁZAT! a ”geometriai valószínűség” nem összekeverendő a geometriai eloszlással, ami sok egyforma független kísérletnél az első sikeres kimenetel bekövetkeztének helye, ill. annak valószínűsége. Itt inkább arról van szó, hogy egy valahány dimenziós tér (egyenes, sík, stb.) pontjai az elemi események, maguk az események ponthalmazok, a ponthalmaz megfelelő dimenziós térfogata (hossz, terület, térfogat, stb.) arányos az esemény valószínűségével.

   Megoldás. Legyen a két véletlen szám $X, Y \in [-1, 3]$. Mivel X, Y függetlenek, két merőleges koordinátatengely mentén mérjük fel őket, és a számegyenesről választott két pont helyett sík egyetlen $(X, Y) \in [-1, 3] \times [-1, 3]$ pontjának tekintjük mint rendezett párt.

   A keresett valószínűség a kedvező tulajdonságú (X, Y) pontok halmaza (jelölje ezt ”A”), és az $\Omega = [-1, 3] \times [-1, 3]$ teljes eseménytér területének az aránya. Itt tehát

   $$A = \{(X, Y) \colon X < 0 \text{ és } Y < 0\} = [-1, 0) \times [-1, 0),$$

   $$P(A) = \frac{t(\text{jó})}{t(\text{összes})} = \frac{t(A)}{t(\Omega)} = \frac{1^2}{4^2} = \frac{1}{16} = 0.0625$$

   □

3. A személyi jövedelemadó bevallások 40%-át könyvelők, 35%-át munkahelyi munkaügyi alkalmazottak és 25%-át maguk az adóalanyok készítik el. A könyvelők által készített bevallások 0.05, a munkaügyi alkalmazottak által készített bevallások 0.07, az adóalanyok által készített bevallások 0.3 valószínűséggel tartalmaznak hibát.

   (a) Kiválasztva egy bevallást mennyi az esélye, hogy hibát tartalmaz?

   (b) Ha a kiválasztott bevallás tartalmaz hibát, mennyi az esélye, hogy az adóalany készítette?

Megoldás. Legyenek:

- A: az adóbevallás tartalmaz hibát, $P(A) =?$
- B_1: az adóbevallást könyvelő készítette, $P(B_1) = 0,4$
- B_2: az adóbevallást munkaügyi alkalmazott készítette, $P(B_2) = 0,35$
- B_3: az adóbevallást az adóalany készítette, $P(B_3) = 0,25$

A jelölést az indokolja, hogy B_1, B_2, B_3 teljes eseményrendszert alkot (páronként kizáró események, amik a teljes eseménytert lefedik), ezeket felhasználva akarjuk majd felírni a *teljes valószínűség tételét*.

VIGYÁZAT! Értelmezzük helyesen a szöveget! Mit jelent pl., hogy " A könyvelők altal készített bevallások 0.05 valószínűséggel tartalmaznak hibát". Mi a 0.05? NEM $P(A \cap B_1)$, hanem az, hogy HA egy bevallást könyvelő készített, AKKOR az 0.05 valséggel hibás, vagyis $P(A|B_1) = 0.05$.

Tehát $P(A|B_1) = 0.05$, $P(A|B_2) = 0.07$, $P(A|B_3) = 0.3$.

(a) Alkalmazzuk a teljes valószínűség tételét A, B_1, B_2, B_3-ra:

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = \\ &= 0,05 \cdot 0,4 + 0,07 \cdot 0,45 + 0,3 \cdot 0,25 \\ &= 0,02 + 0,245 + 0,075 = 0,34. \end{aligned}$$

(b) Bayes tételét alkalmazva az A, B_3 eseményekre azt kapjuk, hogy:

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{0,3 \cdot 0,25}{0,34} = \frac{0,075}{0,34} \approx 0,22.$$

□

4. Magyar kártyából visszatevéssel húzunk 3 lapot. Legyen X a húzott piros lapok száma. Adjuk meg X eloszlását! Mennyi X várható értéke és szórása?

MEGOLDÁS. Vegyük észre, hogy azonos kísérletet ismétlünk $n = 3$-szor egymástól függetlenül, az egyszeri siker valószínűsége $p = 8/32 = 1/4$, ami binomiális eloszláshoz vezet, így tehát $X \sim \text{Binom}(3; 1/4)$.

$$X \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ p_0 & p_1 & p_2 & p_3 \end{pmatrix}, \quad \text{ahol} \quad p_k = \binom{n}{k} p^k (1-p)^{n-k}, (k = 0, 1, 2, 3).$$

Várható értéke

$$E(X) = np = 3 \cdot 0,25 = 0,75,$$

Szórása

$$D(X) = \sqrt{np(1-p)} = \sqrt{3 \cdot 0,25 \cdot 0,75} = 0,75.$$

□

5. Addig dobunk egy kockával, amíg először hatost kapunk. Legyen X a szükséges dobások száma. Adjuk meg X eloszlását! Mennyi az esélye, hogy legfeljebb 3 dobás szükséges? Tegyük fel, hogy az első 10 dobás egyike sem volt hatos. Mennyi az esélye, hogy az első 12 dobásból sem lesz hatos?

MEGOLDÁS. Azonos kísérletet ismétlünk egymástól függetlenül az első siker bekövetkezétig, melynek valószínűsége kísérletenként $p = 1/6$, ez tehát geometriai eloszláshoz vezet, $X \sim \text{Geom}(1/6)$.

$$X \sim \begin{pmatrix} 1 & 2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}, \quad \text{ahol} \quad p_k = (1-p)^{k-1} p, (k = 1, 2, 3, \dots).$$

$$P(X \leq 3) = p_1 + p_2 + p_3 = \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} + \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6} = \frac{36 + 30 + 25}{216} \approx 0,42.$$

A másik kérdéshez felhasználjuk, hogy a geometriai eloszlás *örökifjú*, vagyis

$$P(X > 12 | X > 10) = P(X > 2),$$

Utóbbit pedig a komplementer valószínűségből könnyen ki tudjuk számolni:

$$P(X > 2) = 1 - P(X \leq 2) = 1 - p_1 - p_2 = 1 - \frac{1}{6} - \frac{5}{6} \cdot \frac{1}{6} = \frac{36 - 6 - 5}{36} = \frac{25}{36} \approx 0,69.$$

□

6. Legyen $X \sim N(0,1)$ eloszlású val. változó. Adjunk olyan 0-ra szimmetrikus intervallumot, amibe X értéke 0.8 valószínűséggel beleesik!

Megoldás. Olyan r számot keresünk, amire

$$P(X \in (-r, r)) = 0,8$$

A feltételt kicsit átrendezve

$$P(X \in (-r, r)) = P(|X| < r) = \Phi(r) - \Phi(-r) = 2\Phi(r) - 1,$$

Felhasználva, hogy $\Phi(-r) = 1 - \Phi(r)$.

Azt az egyenletet kell tehát megoldani, hogy

$$\begin{aligned} 2\Phi(r) - 1 &= 0,8 \\ 2\Phi(r) &= 1,8 \\ \Phi(r) &= 0,9 \\ r &= 1,29, \end{aligned}$$

az utolsó sorban r visszafejtésénél igénybe véve a normális eloszlás táblázatát. □

Megjegyzés. Mondhattuk volna rögtön az elején azt is, hogy a *k-szor szigma szabályt* fogjuk alkalmazni a sztenderd normális eloszlásra, ami szerint $P(|X| < r) = 2\Phi(r) - 1$.

7. 100-szor elgurítunk egy szabályos kockát. A centrális határeloszlás-tétel segítségével adja meg közelítőleg annak az esélyét, hogy a gurítások összege legalább 330, de kevesebb, mint 365 lesz!

Megjegyzés. A CHT rávilágít, miért a (sztenderd) normális eloszlás a lefgontosabb eloszlás, amiről tanultunk. Segítségével TETSZŐLEGES független, azonos eloszlású val. változók összege, átlaga közelíthető egy normális eloszlással, melynek átlaga/szórása megegyezik a soktagú összegével. Mennyi az a sok? Tetszőleges eloszlás esetén $n \geq 100$ már elég.

Tegyük fel, hogy X_i-k független, azonos eloszlásúval. változók, várható értékük $E(X_i) = m$, szórásuk $D(X_i) = s$, $(i = 1, 2, \dots)$

$$\begin{aligned}
\text{ÖSSZEG:} \qquad Y &= X_1 + X_2 + \dots + X_n \\
E(Y) &= E(X_1) + E(X_2) + \dots + E(X_n) = nm, \\
D^2(Y) &= D^2(X_1) + D^2(X_2) + \dots + D^2(X_n) = ns^2, \\
D(Y) &= \sqrt{n} \cdot s
\end{aligned}$$

$$\begin{aligned}
\text{ÁTLAG:} \qquad Y &= \frac{X_1 + X_2 + \dots + X_n}{n} \\
E(Y) &= \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} = m, \\
D^2(Y) &= \frac{1}{n^2}\left(D^2(X_1) + D^2(X_2) + \dots + D^2(X_n)\right) = \frac{s^2}{n}, \\
D(Y) &= \frac{s}{\sqrt{n}}.
\end{aligned}$$

Ha az X_i-k független, normális eloszlásúak voltak, akkor az összeg és átlag is az, az imént kiszámolt paraméterekkel.

Ha X_i-k TETSZŐLEGES, de azon belül azonos eloszlásúak m várható értékkel és s szórással, akkor nagy n-re a CHT szerint az összeg ill. az átlag jól közelíthető a megfelelő NORMÁLIS ELOSZLÁSSAL. Speciel az eloszlásfüggvényeikre fennáll

$$F_{összeg}(y) = \Phi\left(\frac{y - nm}{\sqrt{n} \cdot s}\right),$$

$$F_{átlag}(y) = \Phi\left(\frac{y - m}{\frac{s}{\sqrt{n}}}\right).$$

Megoldás. 100-szor gurítunk a kockával, tehát $n = 100$. Világos, hogy függetlenek a kockadobások, egyetlen kockadobás várható értéke

$$m = E(X_i) = \frac{7}{2} = 3,5,$$

szórása s, amit azonban a szórásnégyzetből tudunk kifejezni:

$$s^2 = E(X_i^2) - (E(X_i))^2 = \frac{1^2 + 2^2 + \cdots 6^2}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{364 - 294}{24} = \frac{70}{24} \approx 2,92$$

$$s = \sqrt{\frac{70}{24}} \approx 1,708$$

A CHT értelmében $Y = X_1 + X_2 + \cdots + X_{100} \sim N(nm, \sqrt{n} \cdot s) = N(350\ ;\ 17,1)$ jó közelítés. A kérdés, hogy mennyi $P(330 \leq Y < 365)$?

Az $N(350\ ;\ 17,1)$ paraméteerű normális eloszlású v.v. eloszlásfüggvényét F-fel jelölve fennáll, hogy

$$F(y) = \Phi\left(\frac{y - nm}{\sqrt{n} \cdot s}\right)$$

Így tehát

$$\begin{aligned}
P(330 \leq Y < 365) &= F(365) - F(330) \\
&= \Phi\left(\frac{365 - 350}{\sqrt{n} \cdot s}\right) - \Phi\left(\frac{330 - 350}{\sqrt{n} \cdot s}\right) \\
&= \Phi\left(\frac{365 - 350}{17,1}\right) - \Phi\left(\frac{330 - 350}{17,1}\right) \\
&= \Phi\left(\frac{15}{17,1}\right) - \Phi\left(\frac{-20}{17,1}\right) \\
&= \Phi(0,877) + \Phi(1,169) - 1 \\
&= 0,8078 + 0,8790 - 1 \\
&= 0,6868,
\end{aligned}$$

ahol az utolsó előtti sorban Φ értékeit a normális eloszlás táblázatából határoztuk meg. $\square$