Sztochasztika alapjai informatikusoknak feladatgyűjtemény

Kevei Péter

2024. március 21.

Tartalomjegyzék

1. Alapfogalmak **1**

1.1. Események 1
1.2. Szitaformula 5
1.3. Kombinatorikus valószínűség 10
1.4. Geometriai valószínűség 16
1.5. Feltételes valószínűség, Bayes tétele 17
1.6. Függetlenség 30

2. Véletlen változók **33**

2.1. Diszkrét véletlen változók 33
2.2. Nevezetes diszkrét eloszlások 43
2.3. Folytonos véletlen változók 50
2.4. Nevezetes folytonos eloszlások 54
2.5. Vektorváltozók 59
2.6. De Moivre–Laplace tétel 67

3. Statisztika **74**

3.1. Alapstatisztikák, pontbecslések 74
3.2. Konfidenciaintervallumok, próbák 82

Előszó

A feladatgyűjtemény az SZTE TTIK informatikus hallgatóinak tartott *A sztochasztika alapjai* kurzushoz készült. Minden témakörből több feladat részletes megoldása szerepel. A feladatok többsége típusfeladat. Ilyen nehézségű, ehhez hasonló feladatok a zh dolgozatban, vizsgán előkerülhetnek. Az érdeklődő hallgatók számára egy-egy nehezebb feladat is szerepel megoldással vagy anélkül, melyeket $*$-gal (vagy $**$-gal) jelöltem. A nehezebb feladatok kidolgozott megoldásainak megértése mindenképpen hasznos.

A feladatok között van olyan, amit más feladatgyűjteményből, tankönyvből vettem át, ezeket általában jelöltem. Sok feladat korábbi évek dolgozatfeladatai közül való, és vannak saját feladatok is.

1. Alapfogalmak

1.1. Események

1.1.1. Egy szabályos érmét kétszer feldobunk. Adjuk meg a kísérlet egy matematikai modelljét! Adjuk meg azt az eseményt, hogy (i) dobunk fejet; (ii) két fejet dobunk!

Oldjuk meg a feladatot, ha az érme cinkelt, és a fejdobás valószínűsége p!

Megoldás. Kétszer dobjuk fel az érmét, ezért van egy első és egy második dobás. Tehát a lehetséges kimenetelek

$$\Omega = \{(F,F),(F,I),(I,F),(I,I)\}.$$

Figyeljünk a jelölésre: (F,I) azt jelenti, hogy az első dobás fej, a második írás, azaz ez egy vektor, aminek van első és van második komponense. A $\{\}$-zárójel halmazt jelöl, amiben az elemeknek nincsen sorrendje. Ez egy fontos formalizmus.

A lehetséges kimenetelek száma 4. Ekkor az események halmazát választhatjuk a hatványhalmaznak, azaz az Ω összes részhalmaza esemény. Ezek

$$\begin{aligned}\mathcal{A} = 2^{\Omega} = \Big\{&\emptyset, \{(F,F)\}, \{(F,I)\}, \{(I,F)\}, \{(I,I)\},\\ &\{(F,F),(F,I)\}, \{(F,F),(I,F)\}, \{(F,F),(I,I)\},\\ &\{(F,I),(I,F)\}, \{(F,I),(I,I)\}, \{(I,F),(I,I)\},\\ &\{(F,F),(F,I),(I,F)\}, \{(F,F),(F,I),(I,I)\},\\ &\{(F,F),(I,F),(I,I)\}, \{(F,I),(I,F),(I,I)\},\\ &\{(F,F),(F,I),(I,F),(I,I)\}\Big\}\end{aligned}$$

Egy n elemű halmaznak 2^n részhalmaza van, azaz $|2^{\Omega}| = 2^{|\Omega|}$ (milyen praktikus jelölés), azaz esetünkben $|\mathcal{A}| = 2^4 = 16$. Ez még kis Ω esetén is elég nagy, úgyhogy többször nem írom ezeket ki.

Az az esemény, hogy dobunk fejet azt jelenti, hogy vagy az első, vagy (nem kizáró vagy!) a második dobás fej. Azaz

$$A = \{\text{dobunk fejet}\} = \{(F,F),(F,I),(I,F)\}.$$

Azaz a megfelelő halmaznak 3 eleme van, ez egy összetett esemény.

Az az esemény, hogy két fejet dobunk az egyetlen (F,F) kimenetelt tartalmazza, azaz

$$B = \{\text{két fejet dobunk}\} = \{(F,F)\},$$

ez egy elemű, tehát elemi esemény.

Vegyük észre, hogy eddig nem kellett az, hogy az érme cinkelt-e vagy szabályos. Csak a lehetséges kimenetelek kellettek, azaz a fontos, hogy hányszor dobjuk fel az érmét. Szabályos érme esetén a fej és az írás valószínűsége egyaránt 1/2, ahonnan látjuk, hogy mind a négy lehetséges kimenetel egyformán valószínű. Tehát klasszikus valószínűségi mezőnk van, így $\mathbf{P} : \mathcal{A} \to [0,1]$,

$$\mathbf{P}(A) = \frac{|A|}{4}.$$

Cinkelt érme esetén p a fej valószínűsége, $1-p$ az írásé. Ekkor, a (F,F) valószínűsége p^2, hiszen mindkétszer fejet kaptunk, azaz

$$\mathbf{P}(\{(F,F)\}) = p^2.$$

Kicsit körülményes a formalizmus, de ilyen. Valószínűsége eseménynek van. Tehát annak az (F,F) elemi eseménynek a valószínűségét vizsgáljuk. Mivel a valószínűség additív halmazfüggvény, elég az elemi események valószínűségét megadni. Ezek

$$\mathbf{P}(\{(F,I)\}) = p \cdot (1-p),\ \mathbf{P}(\{(I,F)\}) = p \cdot (1-p),\ \mathbf{P}(\{(I,I)\}) = (1-p)^2.$$

□

1.1.2. Adjuk meg a lottóhúzást leíró valószínűségi mezőt! Mennyi annak a valószínűsége, hogy pont 3 találatunk lesz? Mennyi a valószínűsége, hogy lesz találatunk?

Megoldás. A klasszikus ötöslottón az $1, 2, \dots, 90$ számok közül húznak ki véletlenszerűen ötöt. Azaz a kísérlet lehetséges kimenetelei számötösök. Mivel a számok kihúzási sorrendje nem számít, ezért növekvő sorrendbe rendezzük az öt számot. Tehát

$$\Omega = \{(a_1, a_2, \dots, a_5) : 1 \le a_1 < a_2 < \dots < a_5 \le 90\}.$$

Egy 90 elemű halmaz 5 elemű részhalmazainak száma

$$|\Omega| = \binom{90}{5} = \frac{90 \cdot 89 \cdot \ldots \cdot 86}{5!} = 43\,949\,268.$$

Véges alaphalmaz esetén az események halmazát választhatjuk a hatványhalmaznak, azaz

$$\mathcal{A} = 2^{\Omega} = \{\emptyset, \{(1,2,3,4,5)\}, (1,2,3,4,6), \dots, \Omega\}.$$

Mivel bármely számötöst ugyanakkora valószínűséggel húznak ki, ez egy *klasszikus valószínűségi mező*, azaz $\mathbf{P} : \mathcal{A} \to [0,1]$,

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|}.$$

Egy szelvénnyel játszunk, a számaink: $1 \le a_1 < a_2 < a_3 < a_4 < a_5 \le 90$. (Játszhatunk az $1, 2, 3, 4, 5$ számokkal is.) Jelölje A azt az eseményt, hogy pontosan 3 találatunk lesz. Ez úgy lehetséges, hogy az $a_1, \dots, a_5$ számok közül választunk 3-at, amit $\binom{5}{3}$-féleképp tehetünk meg, és a nem megjelölt 85 szám közül ($\{1, \dots, 90\} \setminus \{a_1, \dots, a_5\}$ halmazból) kiválasztunk 2 számot. Ez

$$\binom{5}{3} \cdot \binom{85}{2}$$

lehetőség, azaz ennyi a kedvező esetek száma. Tehát a valószínűség

$$\mathbf{P}(A) = \frac{\binom{5}{3} \cdot \binom{85}{2}}{\binom{90}{5}}.$$

Jelölje B azt az eseményt, hogy lesz találatunk. Ez lehet úgy, hogy pontosan 1, pontosan 2, ..., pontosan 5 találatunk van. Ehelyett azt számoljuk, hogy nem lesz találatunk, mert ez „egyféleképpen lehet" úgy, hogy a 85 nem megjelölt számból húznak ki 5-öt. Ez

$$\binom{85}{5}$$

lehetőség. Tehát a valószínűség

$$\mathbf{P}(B) = 1 - \mathbf{P}(B^c) = 1 - \frac{\binom{85}{5}}{\binom{90}{5}}.$$

$\square$

1.1.3. Fejezzük ki az A, B, C halmazokkal az alábbi eseményeket!

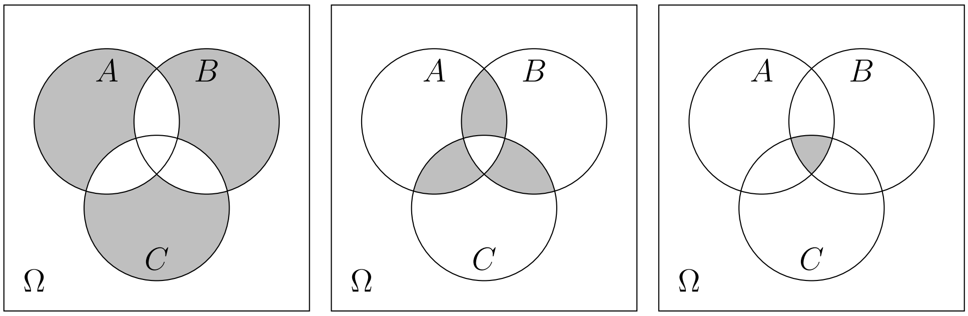
(a) Az A, B, C események közül pontosan $k \in \{1, 2, 3\}$ következik be.

(b) Az A, B, C események közül legalább k következik be.

(c) Az A, B, C események közül legfeljebb k következik be.

Megoldás. Csak az (a) részt csináljuk meg. Ha pontosan egy esemény következik be, akkor valamelyik bekövetkezik *és* a másik kettő nem, azaz ez az esemény

$$\{\text{pontosan 1 következik be}\} = (A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C).$$



1. ábra. $k = 1, 2, 3$

Hasonlóan

$$\{\text{pontosan 2 következik be}\} = (A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C),$$

és ha mindhárom bekövetkezik az a legegyszerűbb, hiszen

$$\{\text{pontosan 3 következik be}\} = A \cap B \cap C.$$

□

1.1.4. Három kockával dobva mennyi a valószínűsége, hogy a dobott számok összege 4? Adjuk meg a kísérletet leíró valószínűségi mezőt!

Megoldás. Három kockával dobunk, ezért a lehetséges kimenetelek halmaza

$$\Omega = \{(i, j, k) : 1 \leq i, j, k \leq 6\}.$$

(Mindig különböztessük meg a kockákat, hiszen azok úgyis különbözők!) Egy $\omega = (i, j, k) \in \Omega$ esetén i jelöli az első kockán dobott számot, j a másodikon dobott számot, k pedig a harmadikon dobott számot. Világos, hogy $|\Omega| = 6^3$. Az is világos, hogy minden kimenetel egyformán valószínű, ezért egy klasszikus valószínűségi mezőnk van.

A dobott számok összege úgy lehet 4, ha két 1-est és egy 2-est dobunk, azaz

$$A = \{\text{az összeg 4}\} = \{(1, 1, 2), (1, 2, 1), (2, 1, 1)\}.$$

Ez egy összetett esemény, aminek a valószínűsége

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{6^3} = \frac{1}{72}.$$

□

1.1.5. Egy szabályos dobókockával egyszer dobunk. Adjuk meg a kísérletet leíró valószínűségi mezőt! Adjuk meg azt az eseményt és annak valószínűségét, hogy (i) páros számot dobunk; (ii) prímet dobunk; (iii) hatost dobunk!

1.1.6. Egy szabályos dobókockával kétszer dobunk. Adjuk meg a kísérletet leíró valószínűségi mezőt! Adjuk meg azt az eseményt és annak valószínűségét, hogy (i) két hatost dobunk; (ii) dobunk hatost; (iii) mindkétszer páratlan számot dobunk!

1.1.7. Egy szabályos érmét tízszer feldobunk. Adjuk meg a kísérlet egy matematikai modelljét! Adjuk meg azt az eseményt és annak valószínűségét, hogy (i) nem dobunk fejet; (ii) az első dobás fej; (iii) pontosan 5 fejet dobunk!

Oldjuk meg a feladatot, ha az érme cinkelt, és a fejdobás valószínűsége p!

1.1.8. Igazoljuk az alábbi formulák helyességét!

(a) $A \circ B = (A \cup B) - A \cap B$;

(b) $A \cap (B - C) = A \cap B - A \cap C$;

(c) $A - (A - (B - C)) = A \cap B \cap C^c$;

(d) $A \cup B = A \circ B \circ (A \cap B)$.

1.2. Szitaformula

1.2.1. Egy pénzügyi befektető cég három cégbe fekteti pénzét, melyek rendre $0,19$, $0,25$, illetve $0,28$ valószínűséggel mennek csődbe az elkövetkező öt évben. Annak a valószínűsége, hogy az első és a második cég is csődbe megy $0,05$, hogy az első és a harmadik is csődbe megy $0,1$, míg hogy a második és a harmadik is becsődöl annak is $0,1$. Annak az esélye, hogy mindhárom cég becsődöl 2%. Mennyi a valószínűsége, hogy

(a) az első vagy a második cég csődbe megy?

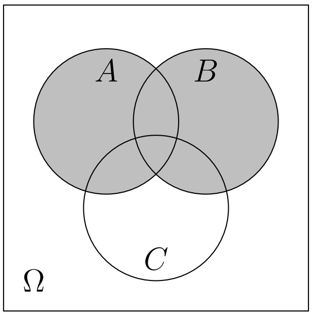
(b) egyik cég sem megy csődbe?

(Szűcs Gábor feladata)

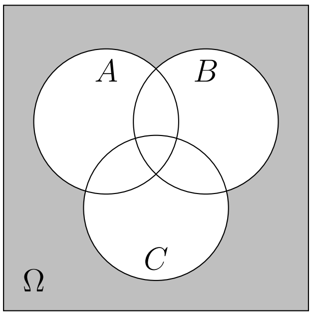
Megoldás. Ez egy egyszerű szitaformulás feladat. Jelölje A, B, illetve C azt az eseményt, hogy az első, második, illetve harmadik cég csődbe megy az első öt évben. A feladat megadja a egyes, kettes és hármas metszetek valószínűségét.

(a) Az, hogy az első vagy a második cég csődbe megy a $A \cup B$ esemény. Ennek a valószínűsége

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) = 0,19 + 0,25 - 0,05 = 0,39.$$



2. ábra. Az első vagy a második csődbe megy.



3. ábra. Egyik sem megy csődbe.

(b) Az, hogy egyik cég sem megy csődbe, az a $(A \cup B \cup C)^c$ esemény. A szitaformula szerint

$$\begin{aligned}\mathbf{P}(A \cup B \cup C) &= \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) \\ &\quad - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) \\ &\quad + \mathbf{P}(A \cap B \cap C) \\ &= 0,19 + 0,25 + 0,28 - 0,05 - 0,1 - 0,1 + 0,02 \\ &= 0,49.\end{aligned}$$

□

1.2.2. A Faluvégi Kurta Kocsma előtt 5 bicikli áll. Záróra előtt egymás után jön ki az 5 tulajdonos, és mindegyikük véletlenszerűen választ egy kerékpárt. Mennyi a valószínűsége, hogy senki sem a saját biciklijén jutott haza?

Megoldás. Számozzuk meg az embereket 1-től 5-ig, és jelölje A_i azt az eseményt, hogy az i-edik a saját biciklijét választja. Ekkor az az A esemény, hogy *valaki* a saját biciklijét választja, az $A_1 \cup A_2 \cup \ldots \cup A_5$, hiszen ez az az esemény, hogy *vagy* az 1-es a saját biciklijét választja, *vagy* a 2-es a saját biciklijét választja, ..., vagy az 5-ös a saját biciklijét választja. Az, hogy *senki* nem a saját biciklijét választja, az éppen az A esemény ellentettje. Tehát

$$\mathbf{P}(\text{senki nem a saját biciklijén jut haza }) = \mathbf{P}(A^c) = 1 - \mathbf{P}(A).$$

Az unió valószínűségét a szitaformulával tudjuk meghatározni. Eszerint

$$\begin{aligned}\mathbf{P}(A_1 \cup \ldots \cup A_5) = \mathbf{P}(A_1) + \ldots + \mathbf{P}(A_5) \\ &- [\mathbf{P}(A_1 \cap A_2) + \ldots + \mathbf{P}(A_4 \cap A_5)] \\ &+ \mathbf{P}(A_1 \cap A_2 \cap A_3) + \ldots + \mathbf{P}(A_3 \cap A_4 \cap A_5) \\ &- [\mathbf{P}(A_1 \cap \ldots \cap A_4) + \ldots + \mathbf{P}(A_2 \cap \ldots \cap A_5)] \\ &+ \mathbf{P}(A_1 \cap \ldots \cap A_5).\end{aligned}$$

A metszetek valószínűsége nyilván csak attól függ, hogy hány esemény metszetét veszem, hiszen annak a valószínűsége, hogy az 1-es és a 2-es a saját biciklijén megy haza, ugyanannyi, mint annak a valószínűsége, hogy a 3-as és az 5-ös a sajátján megy.

Számítsuk ki a metszetek valószínűségét! Az A_1 esemény azt jelenti, hogy az 1-es a saját biciklijét választja, a többiekre nincs megszorításunk. Ennek a valószínűsége $1/5$, hiszen 5 bicikli közül választhat. Az $A_1 \cap A_2$ esemény azt jelenti, hogy az 1-es *és* a 2-es *is* a saját biciklijét választja, ennek az esélye $1/(5 \cdot 4)$, hiszen az első 5 bicikli közül választhat, a második már csak 4 közül, és csak 1 lehetőség kedvező. Hasonlóan,

$$\begin{aligned}\mathbf{P}(A_1 \cap A_2 \cap A_3) &= \frac{1}{5 \cdot 4 \cdot 3}, \\ \mathbf{P}(A_1 \cap A_2 \cap A_3 \cap A_4) &= \frac{1}{5 \cdot 4 \cdot 3 \cdot 2}, \\ \mathbf{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) &= \frac{1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}.\end{aligned}$$

Vegyük észre, hogy ha 4-en a saját biciklijüket választják, akkor szükségképpen az 5-ödik is, tehát

$$A_1 \cap A_2 \cap A_3 \cap A_4 = A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5.$$

Annyit kell még látni, hogy 2-es metszet $\binom{5}{2}$ van, 3-as metszet $\binom{5}{3}$, 4-es $\binom{5}{4}$, és 5-ös $\binom{5}{5}$. Összegezve

$$\begin{aligned}&\mathbf{P}(A_1 \cup \ldots \cup A_5) \\ &= 5 \cdot \frac{1}{5} - \binom{5}{2}\frac{1}{5 \cdot 4} + \binom{5}{3}\frac{1}{5 \cdot 4 \cdot 3} - \binom{5}{4}\frac{1}{5 \cdot 4 \cdot 3 \cdot 2} + \binom{5}{5}\frac{1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}.\end{aligned}$$

□

1.2.3 (∗). Egy urnában k-féle színű golyó van, mindegyik színűből ugyanannyi darab. Egyenként húzunk a golyókból úgy, hogy minden húzás után visszatesszük a kihúzott golyót, és minden húzásnál bármelyik golyó ugyanolyan valószínűséggel kerülhet kihúzásra.

(a) Mennyi annak a q_n valószínűsége, hogy legalább n húzás kellett ahhoz, hogy minden szín előforduljon?

(b) Mennyi annak a p_n valószínűsége, hogy n húzás során minden szín előfordult, és ez az n-edik húzásnál következik be először (vagyis az első $(n-1)$ húzás során csak $(k-1)$ szín fordult elő) ?

([1] 1.2.15)

Megoldás. Ez egy kicsit nehezebb szitaformulás feladat.

(a) Az, hogy legalább n húzás kell, hogy minden szín előforduljon pontosan azt jelenti, hogy $n-1$ húzás után még nem volt minden szín. Jelölje $A_{i,n-1} = A_i k$ azt az eseményt, hogy az első $n-1$ húzás során nem volt i színű golyó. Ekkor az az esemény, hogy $n-1$ húzás után nem volt minden szín, pontosan azt jelenti, hogy az A_1, A_2, ..., A_k események közül legalább egy bekövetkezett, azaz

$$C_n := \{\text{legalább } n \text{ húzás kell}\} = \cup_{i=1}^k A_i.$$

Az A_i események nem kizáróak, ezért az unió valószínűségét szitaformulával határozhatjuk meg. Eszerint

$$\begin{aligned}\mathbf{P}(A_1 \cup \ldots \cup A_k) &= \mathbf{P}(A_1) + \ldots + \mathbf{P}(A_k)\\ &\quad - [\mathbf{P}(A_1 \cap A_2) + \ldots + \mathbf{P}(A_{k-1} \cap A_k)]\\ &\quad \ldots\\ &\quad \pm \mathbf{P}(A_1 \cap \ldots \cap A_k)\\ &= \sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} \mathbf{P}(A_1 \cap \ldots \cap A_j),\end{aligned}$$

ahol az utolsó egyenlőségnél felhasználtuk, hogy a j-es metszetek valószínűségei megegyeznek (hát persze, ugyanakkora valószínűséggel nem volt sem piros sem kék, mint sárga meg zöld. A metszetek valószínűségeit könnyű meghatározni. Valóban

$$\mathbf{P}(A_1) = \frac{(k-1)^{n-1}}{k^{n-1}},$$

hiszen az összes eset k^{n-1}, mert az $n-1$ húzás során mindig k-féle golyót kaphatunk, és a kedvező esetek száma meg $(k-1)^{n-1}$, hiszen 1-es színű golyót nem húztam, bármi más lehetett. Hasonlóan,

$$\mathbf{P}(A_1 \cap A_2) = \frac{(k-2)^{n-1}}{k^{n-1}},$$

hiszen ekkor már sem 1-es sem 2-es színű golyót nem húzhattam. Általánosan

$$\mathbf{P}(A_1 \cap A_2 \cap \ldots \cap A_j) = \frac{(k-j)^{n-1}}{k^{n-1}}.$$

Ezt visszahelyettesítve

$$\begin{aligned}\mathbf{P}(\cup_{j=1}^{k} A_j) &= \sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} \mathbf{P}(A_1 \cap \ldots \cap A_j) \\ &= \sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} \frac{(k-j)^{n-1}}{k^{n-1}}.\end{aligned}$$

(b) Legyen D_n az az esemény, hogy pontosan n húzás kellett. Könnyű látni, hogy

$$D_n = C_n \setminus C_{n+1},$$

hiszen ha pontosan n kellett, akkor legalább n kellett, de nem kellett $n+1$. Nyilván $C_n \supset C_{n+1}$, hiszen ha legalább $n+1$ húzás kellett, akkor legalább n, ezért

$$\mathbf{P}(D_n) = \mathbf{P}(C_n) - \mathbf{P}(C_{n+1}).$$

□

1.2.4. A Jonas Brothers nevű együttes újra összeáll és koncertet adnak. A PepsiCo cég a következő ötlettel áll elő: a kólásüvegeik kupakjában elrejtik a banda egy-egy tagjának a nevét és azok között, akik összegyűjtik mindhárom nevet kisorsolnak egy VIP belépőt. Kevin neve a kupakok felén szerepel, Joe-val a kupakok egyharmadában találkozhatunk és Nick a legritkább, neve átlagosan minden hatodik kupakban szerepel. Mennyi a valószínűsége, hogy 5 kólát vásárolva sikerül kigyűjtenünk a három testvért? (Segítség: a kupakokra gondoljunk úgy mintha egy zsákból húznánk egy nevet, melyben Kevin háromszor, Joe kétszer, Nick pedig egyszer szerepel.)

1.3. Kombinatorikus valószínűség

1.3.1. Hét törpe közül Hófehérke leültet ötöt egy kör alakú asztalhoz. Tegyük fel, hogy az összes lehetséges elrendezés egyformán valószínű. Mennyi a valószínűsége, hogy Morgó és Kuka nem kerül egymás mellé?

Megoldás. Minden lehetséges választás egyformán valószínű, ezért klasszikus a valószínűségi mező. Tehát kedvező (vagy kedvezőtlen) és összes esetet kell számolni. Először mindig az összes esetet számoljuk.

Hófehérke $\binom{7}{5}$-féleképp választhat ki 7 törpe közül 5-öt, akik leülnek. Ők 5!-féleképpen ülhetnek le, hiszen az első székre 5, a másodikra 4, …, az ötödikre 1 törpe ülhet. Figyeljünk, most megkülönböztettük a székeket, nem csak a szomszédság számít! Tehát az összes esetek száma

$$\binom{7}{5} \cdot 5!.$$

Mindig meg kell gondolni, hogy a kedvező, vagy a kedvezőtlen eseteket könnyebb összeszámolni. Most talán a kedvezőtlent. Ekkor mindketten az 5 leültetett törpe között vannak, és a maradék 3 törpét Hófehérke $\binom{5}{3}$-féleképp választhatja ki. Az ültetésnél Morgó 5 helyre ülhet, Kuka pedig 2-re, hiszen Morgó mellé kell ültetni. A többi 3 törpét 3!-féleképp lehet leültetni. Összesen

$$\binom{5}{3} \cdot 5 \cdot 2 \cdot 3!$$

a rossz esetek száma. A keresett valószínűség

$$\begin{aligned}
&\mathbf{P}(\text{Morgó és Kuka nem ül egymás mellé}) \\
&= 1 - \mathbf{P}(\text{Morgó és Kuka egymás mellé ül}) \\
&= 1 - \frac{\binom{5}{3} \cdot 5 \cdot 2 \cdot 3!}{\binom{7}{5} \cdot 5!} \\
&= 1 - \frac{5}{21} = \frac{16}{21}.
\end{aligned}$$

□

1.3.2. Tíz pár cipőből véletlenül kiválasztunk négy darabot. Mekkora a valószínűsége, hogy nem lesz egy pár sem?

Megoldás. Tíz pár cipő az 20 db cipő, ezért az összes esetek száma

$$\binom{20}{4} = \frac{20 \cdot 19 \cdot 18 \cdot 17}{4!} = 4845.$$

Számoljuk a kedvező eseteket. Mivel nincs pár, összesen 4 pár cipőből választok ki egy-egy cipőt, ezt $\binom{10}{4}$-féleképpen tehetem meg, majd mindegyik párból a bal vagy jobb cipőt választhatom 2^4-féleképpen. Összesen

$$\binom{10}{4} \cdot 2^4 = 3360.$$

lehetőség. Tehát a keresett valószínűség

$$\mathbf{P}(\text{nincs pár}) = \frac{\binom{10}{4} \cdot 2^4}{\binom{20}{4}} = \frac{224}{323}.$$

□

1.3.3. Egy vendéglőben az egyik asztalnál 9 vendég ül. Négyen kólát, hárman sört rendeltek, ketten pedig ásványvizet rendeltek. A kissé feledékeny pincér emlékszik, hogy miből mennyit rendeltek, de azt már elfelejtette, hogy ki mit kért. Ezért véletlenszerűen osztja ki az italokat. Mekkora a valószínűsége, hogy mindenki azt kapja, amit rendelt?

Megoldás. A vendégeknek az italokat 9!-féleképpen oszthatja ki a pincér. Ekkor megkülönböztettük a kólákat! A kedvező leosztások száma

$$4! \cdot 3! \cdot 2,$$

hiszen a kólákat 4!-féleképpen oszthatja ki jó, a söröket 3!-féleképpen, a vizeket pedig 2-féleképpen. Tehát a keresett valószínűség

$$\mathbf{P}(\text{ mindenki azt kapja amit rendelt }) = \frac{4! \cdot 3! \cdot 2}{9!}.$$

2. megoldás. Nem kell megkülönböztetnünk az azonos italokat. Ekkor az összes esetek száma

$$\frac{9!}{4! \cdot 3! \cdot 2},$$

hiszen a kólákat 4!-féleképpen, a söröket 3!-féleképp, a vizeket 2-féleképp permutálhatjuk. Ha az összes eseteket így számoljuk, akkor kedvező eset csak 1 van, amikor mindenki azt kapja amit kért. Tehát a keresett valószínűség

$$\mathbf{P}(\text{ mindenki azt kapja amit rendelt }) = \frac{4! \cdot 3! \cdot 2}{9!},$$

ami persze ugyanaz, mint az előbb. Többféleképpen számolhatjuk az eseteket, de arra kell nagyon ügyelni, hogy ugyanúgy számoljuk az összes esetet, mint a kedvezőt. □

1.3.4. A Bajnokok Ligájában 2017-ben három spanyol csapat jutott a 8 közé: az Atlético Madrid, a Barcelona és a Real Madrid. Sorsolással határozták meg a negyeddöntők párosítását (itt már nincs kiemelés, és azonos nemzet csapatai is összekerülhetnek). Mennyi volt a sorsolás előtt a valószínűsége annak, hogy a negyeddöntőben

(a) Barcelona – Real Madrid párharc lesz?

(b) lesz spanyol párharc?

Megoldás. (a) Sokféleképpen számolhatjuk az eseteket. Arra kell mindig figyelni, hogy a kedvezőeket ugyanúgy számoljuk össze, mint az összeset. A feladat szempontjából lényegtelen, hogy ki kezd otthon, és az is, hogy hanyadiknak húzták ki az adott párt. Tehát az összes eset

$$\frac{\binom{8}{2}\cdot\binom{6}{2}\cdot\binom{4}{2}\cdot\binom{2}{2}}{4!},$$

hiszen először kiveszem az első párt a 8 csapatból, aztán a maradék 6 csapatból a második párt, majd a 4-ből a harmadik párt, végül a negyedik párt. De azt mondtuk, hogy mindegy a párok sorrendje, ezért leosztom az egészet az összes lehetséges sorrenddel, ami 4!.

Ekkor a kedvező esetek összeszámolásánál rögzítem a Barcelona – Real Madrid párt. A maradék 6 csapatból választom a második párt, majd a maradék 4-ből a harmadikat, végül a negyediket. Most csak 3 párt választottam, ezért 3! a sorrendek száma. Összegezve a kedvező esetek száma:

$$\frac{\binom{6}{2}\cdot\binom{4}{2}\cdot\binom{2}{2}}{3!}.$$

A kérdéses valószínűség

$$\mathbf{P}(\text{Barcelona} - \text{RM}) = \frac{\text{kedvező}}{\text{összes}} = \frac{4}{\binom{8}{2}} = \frac{1}{7}.$$

2. megoldás. Miután kijött az 1/7, gyanús, hogy egyszerűbben is lehetett volna:

– Andrés, most akkor a Madriddal játszunk?

– Nem, Leo. Hét csapat közül sorsolnak egyet nekünk. Bármelyiket egyforma eséllyel kapjuk, tehát 1/7 a valószínűsége, hogy a Madriddal játszunk.

(b) A változatosság kedvéért, most máshogy számoljuk az összes esetet. Mindent figyelünk, hogy kit hanyadiknak húztak, ki kezd otthon. Ekkor az összes esetek száma 8!, hiszen az elsőre kihúzott csapat 8-féle lehet, ... Ekkor a kedvező eseteket is eszerint kell számolni. Mivel 3 spanyol csapat van, ezért ha van spanyol párharc, az csak úgy lehet, hogy két spanyol csapat egymás ellen játszik, a harmadik meg mással. Ezt 3-féleképpen tehetjük meg, a két csapat 8-féleképpen játszhat egymás ellen (4 pár, ki kezd otthon). A maradék 6 csapat pedig 6!-féleképpen rendezhetem el. Összesen

$$3 \cdot 8 \cdot 6!.$$

Tehát a keresett valószínűség

$$\mathbf{P}(\text{spanyol párharc}) = \frac{3 \cdot 8 \cdot 6!}{8!} = \frac{3}{7}.$$

□

1.3.5 (∗)**.** Egy halastóban M aranyhal és K ezüsthal van. Egy horgász addig fogja ki egyesével a halakat, amíg már csak egyszínű hal marad a tóban (tehát vagy csupa aranyhal, vagy csupa ezüsthal). Mennyi a valószínűsége, hogy a Gyuri nevű ezüsthal megússza a horgászkalandot?

Megoldás. Ez egy kicsit nehezebb feladat. Könnyen belefuthatunk egy olyan összeszámlálási feladatba, amit nehéz megoldani.

Képzeljük el, hogy a halakat véletlenszerűen megszámozzuk 1-től $(M + K)$-ig, és a sorszám azt jelöli, hogy a horgász hanyadiknak horgászná ki az adott halat. Nem húzza ki mindet, az utolsó néhány sorszámú hal megússza, de az a véletlentől függ, hogy hányan. Azt kell észrevenni, hogy Gyuri pontosan akkor ússza meg, ha az ő sorszáma nagyobb, mint az utolsó aranyhal sorszáma. Tehát csak arra kell figyelni, hogy az M aranyhalhoz képest Gyuri sorszáma hol van. Tehát Gyuri pontosan akkor ússza meg, ha $M + 1$ hal közül ő az utolsó, aminek a valószínűsége

$$\frac{1}{M+1}.$$

□

1.3.6 (∗)**.** Mekkora a valószínűsége, hogy az ötöslottón kihúzott számok között nem lesznek egymást követők?

Megoldás. Ez egy kicsit nehezebb. Mivel 90 szám közül választunk ki ötöt, így az összes esetek száma

$$\binom{90}{5}.$$

A kedvező eseteket trükkös összeszámolni. A keresett esemény

$$A = \{(i_1, i_2, \dots, i_5) : 1 \le i_1 < i_2 - 1 < i_3 - 2 < i_4 - 3 < i_5 - 4 \le 86\},$$

hiszen az, hogy nincsenek szomszédosak, pontosan azt jelenti, hogy $i_2 > i_1+1$, $i_3 > i_2 + 1$, …, $i_5 > i_4 + 1$, ami átrendezve ugyanaz ami fönt van. Ekkor majdnem kész is vagyunk, hiszen az $(i_1, i_2-1, \dots, i_5-4)$ az $1, \dots, 86$ számok közül kiválasztott számötös, azaz az ilyenek száma

$$\binom{86}{5}.$$

Valójában megadtunk egy bijekciót (egy-egy értelmű leképezést) a nemszomszédos számötösök halmaza és az $1, \dots, 86$ halmaz számötösei között. Tehát a keresett valószínűség

$$\mathbf{P}(\text{nincs szomszédos}) = \frac{\binom{86}{5}}{\binom{90}{5}}.$$

□

1.3.7. Egy sakktáblán találomra elhelyezünk 8 bástyát. Mi a valószínűsége, hogy egyik sem üti a másikat?

1.3.8. Egy hallgató 40 tétel közül 20-at úgy megtanul, hogy abból jelesre tud vizsgázni, a másik 20-ból csak jóra. A vizsgatétel kiválasztásakor 40 tétel közül húz 2 tételt, majd ebből választ egyet és ebből felel. Mennyi a valószínűsége, hogy jelesre vizsgázik? ([1])

1.3.9. Egy urnában csak piros, zöld és kék golyók vannak. A piros golyók száma 18. Egy golyó kihúzása esetén annak a valószínűsége, hogy nem piros golyót húzunk 1/15-del kisebb, mint azé, hogy zöld vagy piros golyót húzunk. Annak a valószínűsége viszont, hogy kék vagy piros golyót húzunk 11/10-szer nagyobb, mint annak a valószínűsége, hogy zöld vagy piros golyót húzunk. Hány zöld és hány kék golyó van az urnában? (2008-as érettségi feladat)

1.3.10. 2008-ban a Bajnokok Ligájában 4 angol (MU, Arsenal, Chelsea, Liverpool), egy olasz (Roma), egy spanyol (Barcelona), egy német (Schalke) és egy török (Fenerbahce) csapat jutott a 8 közé. Sorsolással határozzák meg a negyeddöntők párosítását. Határozzuk meg annak a valószínűségét, hogy a negyeddöntőben

(a) a Roma elkerüli a MU-t!

(b) az angol csapatok elkerülik egymást!

1.3.11. A Derelye pékség polcán 30 db mákos kifli van, melyek között vannak hibás termékek is. Ha két kiflit kiveszünk, akkor annak a valószínűsége, hogy mindkettő jó, $\frac{38}{9}$-szerese annak a valószínűségnek, hogy mindkettő hibás. Hány hibás mákos kifli van? (KöMaL 2023/1)

1.3.12. Az autók rendszámai 2022-ig 3 betűből (26 lehetséges betű közül) és 3 számból álltak. Mennyi a valószínűsége, hogy egy véletlenül kiválasztott rendszám 3 különböző betűből és 3 különböző számból áll? (KöMaL 2023/1)

1.3.13. Egy unatkozó gyakorlatvezető dolgozatiratás során arra lett figyelmes, hogy a csoportjában az összes lány egy sorban ül. A csoportban 10 hallgató van, közülük 3 lány. A teremben 4 sor van és minden sorban 4 hely, és feltesszük, hogy mindenki véletlenszerűen választ helyet, azaz minden leülési konfiguráció egyforma valószínűségű. Mennyi a kérdéses esemény valószínűsége?

1.3.14. Tegyük fel, hogy a bergengóc lakosság 10%-a beteg. Mennyi annak a valószínűsége, hogy 100 bergengócból véletlenszerűen 3-at kiválasztva és őket letesztelve azt találjuk, hogy egyikük sem fertőzött?

Melyik az a legkisebb x, melyre igaz, hogy ha a lakosság x%-a beteg, akkor 1/2-nél kisebb lesz az előbbi valószínűség?

(Segítség: Az, hogy a lakosság x%-a fertőzött azt jelenti, hogy pontosan x fertőzött van a 100-ból.)

1.3.15. Az A, B, C, D, E, F kereskedőcégek mindegyike az öt másik céggel kötött egy-egy üzletet az előző hónapban (bármelyik két cég között pontosan egy üzletkötés jött létre). Az ellenőrző hatóság véletlenszerűen kiválaszt a hat cég előző havi (egymás közötti) üzletkötései közül négyet, és azokat ellenőrzi. Mekkora a valószínűsége, hogy az A vagy a B cég üzletkötései közül is ellenőriznek legalább egyet? (Érettségi 2018)

1.3.16. Egy szabályos kockával 11-szer dobunk. Mennyi a valószínűsége, hogy az egymást követő 1, 2, 3, 4, 5, 6 eredménysorozat nem fordul elő?

1.3.17. Száz alma közül tíz férges. Véletlenül kiválasztva ötöt, mi a valószínűsége, hogy lesz közte férges?

1.3.18. Egy embernek n egyforma kinézetű kulcsa van, melyek közül pontosan egy nyitja az ajtót. Emberünk véletlenül választva sorra próbálja a kulcsokat addig, amíg a jó kulcs elő nem kerül. Valamely $k \in \{1, 2, \ldots, n\}$ esetén mennyi a valószínűsége, hogy a k-adik próbálkozása sikeres, ha

(a) a kipróbált rossz kulcsokat mindig félreteszi?

(b) a kipróbált rossz kulcsokat sose teszi félre?

1.3.19. Mind a 4 pár különböző cipőmet az előszobában levő beépített szekrényben tartom. Az előszobában kiégett a villany, így sötétben keresgélve véletlenszerűen kiveszek 4 cipőt. Mennyi a valószínűsége, hogy pontosan egy összeillő párt veszek ki? Mekkora ugyanez a valószínűség, ha 4 pár egyforma cipőm van?

1.3.20. Az iskolai karácsonyi vásárra készülődve Blanka, Csenge és Dóri feladata az volt, hogy különböző figurákat hajtogassanak színes papírból. Összesen 70 figurát hajtogattak. A figurák kétheted részét Dóri készítette, a maradékot pedig fele-fele arányban Blanka és Csenge. Számítsa ki annak a valószínűségét, hogy a 70 figura közül véletlenszerűen kiválasztott két figurát ugyanaz a lány készített! (Érettségi 2018)

1.3.21. Egy kertész három juhar-, négy tölgy- és öt nyírfát ültet egy sorba véletlen sorrendben, mindegyik fát egyenlő valószínűséggel választva. Mennyi annak a valószínűsége, hogy nem kerül egymás mellé két nyír?

1.3.22. Egyes vidékeken elterjedt a következő babona: egy lány 6 fűszálat fog a markába úgy, hogy azok a kezéből mindkét irányban kiállnak. Egy másik lány mindkét oldalon páronként összecsomózza a fűszálakat. Ha így egy zárt lánc keletkezik, akkor arra következtetnek, hogy a lány a következő évben férjhez megy. Ha a csomózás teljesen véletlenszerűen történik, mennyi a valószínűsége, hogy zárt láncot kapunk? Mi a helyzet $2n$ fűszál esetén? ([1, 1.3.29])

1.4. Geometriai valószínűség

1.4.1. Egy labdát találomra nekirúgunk egy háznak, amely 10 m hosszú és 5 m magas. A házon két 2 m × 1,5 m-es ablak van. Mennyi a valószínűsége, hogy ablakot talál a labda? ([1, 1.4.2])

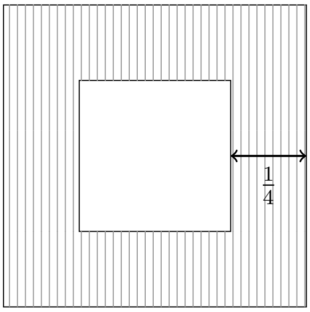
Megoldás. Világos, hogy egy geometriai valószínűségi mezőnk van. A teljes terület $5 \times 10 = 50$ egység (m^2). A két ablak együttes területe $2 \times 2 \times 1{,}5 = 6$ egység. Tehát, annak a valószínűsége, hogy ablakot találunk

$$\mathbf{P}(\text{ ablakot talál }) = \frac{\text{kedvező terület}}{\text{összes terület}} = \frac{6}{50} = \frac{3}{25}.$$

□

1.4.2. Egy négyzet belsejében egyenletes eloszlás szerint választunk egy pontot. Mennyi a valószínűsége, hogy a választott pont közelebb van valamelyik oldalhoz, mint 1/4?

Megoldás. A kísérlet egy geometriai valószínűségi mezőn írható le, ahol az eseménytér az egységnégyzet, az események az egységnégyzet Borel-halmazai (szép halmazai), és valószínűség pedig a terület, azaz $\Omega = [0,1]^2$, $\mathcal{A} = \mathcal{B}([0,1]^2)$, $\mathbf{P}(A) = \mathrm{ter}(A)$. A kedvező síkrész:



A kedvező terület, ami éppen a keresett valószínűség

$$\mathbf{P}(\text{közelebb van mint } 1/4) = 4 \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{4}.$$

$\square$

1.4.3. Borókát az egyéves kislányt, ha betesszük egy 2 méter oldalhosszúságú négyzet alapjú járókába, akkor elindul a legközelebbi oldal felé. Tegyük fel, hogy különböző alkalmak során egymástól függetlenül, egyenletes eloszlás szerinti pontban tesszük be Borókát.

(a) Mekkora a valószínűsége, hogy a járóka déli oldalához megy?

(b) Hány alkalommal kell betenni ahhoz, hogy legalább 90% eséllyel legalább egyszer a déli oldalt válassza?

1.5. Feltételes valószínűség, Bayes tétele

1.5.1. Aladár a pénzét három egyforma borítékban tartja. Az elsőben két ezerforintos, a másodikban egy ezer- és egy kétezerforintos, a harmadikban egy ezer és három kétezerforintos van. Aladár találomra kivesz egy borítékot, és onnan egy bankjegyet. Mennyi a valószínűsége, hogy ezerforintost húzott?

Megoldás. Jelölje A, B, C azt, hogy az első, második, vagy harmadik borítékot választotta Aladár, E pedig az, hogy ezerforintost húzott. Ekkor A, B, C teljes eseményrendszer, hiszen diszjunktak (pontosan egy borítékot választ), és lefedik az eseményteret (választ borítékot). A feladat szerint minden borítékot egyforma valószínűséggel választ, azaz

$$\mathbf{P}(A) = \mathbf{P}(B) = \mathbf{P}(C) = \frac{1}{3},$$

továbbá az egyes borítékokból húzva az ezres valószínűsége

$$\mathbf{P}(E|A) = 1, \quad \mathbf{P}(E|B) = \frac{1}{2}, \quad \mathbf{P}(E|C) = \frac{1}{4}.$$

Tehát (formálisan a teljes valószínűség tételét használjuk)

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(A) \cdot \mathbf{P}(E|A) + \mathbf{P}(B) \cdot \mathbf{P}(E|B) + \mathbf{P}(C) \cdot \mathbf{P}(E|C) \\ &= \frac{1}{3}\left(1 + \frac{1}{2} + \frac{1}{4}\right) = \frac{7}{12}.\end{aligned}$$

□

1.5.2. Az 52 lapos francia kártyából kiosztanak 13 lapot. Legyen A az az esemény, hogy pontosan 2 ászt kaptunk. Határozzuk meg a $\mathbf{P}(A|B_i)$ feltételes valószínűségeket, ha

(a) B_1 azt jelenti, hogy van legalább egy ászunk;

(b) B_2 azt jelenti, hogy a kőr ász nálunk van;

(c) B_3 azt jelenti, hogy a kiosztott lapok közül az első ász;

(d) B_4 azt jelenti, hogy a kiosztott lapok közül az első a kőr ász.

Megoldás. Definíció szerint

$$\mathbf{P}(A|B_i) = \frac{\mathbf{P}(A \cap B_i)}{\mathbf{P}(B_i)},$$

tehát ezeket a valószínűségeket kell meghatározni.

(a) Azt vizsgáljuk, hogy milyen lapokat kaptunk. Mivel 52 lap közül kapunk 13-at, ezért az összes esetek száma

$$\binom{52}{13}.$$

Ha van legalább egy ász, akkor vagy pontosan egy, vagy pontosan 2, vagy pontosan 3, vagy pontosan 4 ászunk van. (A 'legalább'-ot mindig szét kell szedni 'pontosan'-ok uniójára.) Ennél egyszerűbb, ha áttérünk a komplementer számolására, hiszen az azt jelenti hogy nincs ászunk. Tehát a kedvezőtlen/rossz esetek száma

$$\binom{48}{13},$$

hiszen ekkor a 48 nemász lap közül kaphatunk. Tehát

$$\mathbf{P}(B_1) = 1 - \frac{\binom{48}{13}}{\binom{52}{13}}.$$

Az $A \cap B_1$ esemény azt jelenti, hogy A és B_1 is bekövetkezik. Ez $A \cap B_1 = A$, azaz pontosan 2 ászunk van. Ennek a valószínűsége

$$\mathbf{P}(A \cap B_1) = \mathbf{P}(A) = \frac{\binom{4}{2}\binom{48}{11}}{\binom{52}{13}},$$

hiszen az összes eset maradt, a kedvezőnél pedig a 4 ász közül választunk 2-t, a maradék 11 lapot pedig a 48 nemász közül kapjuk. Összegezve

$$\mathbf{P}(A|B_1) = \frac{\mathbf{P}(A \cap B_1)}{\mathbf{P}(B_1)} = \frac{\frac{\binom{4}{2}\binom{48}{11}}{\binom{52}{13}}}{1 - \frac{\binom{48}{13}}{\binom{52}{13}}} = \frac{\binom{4}{2}\binom{48}{11}}{\binom{52}{13} - \binom{48}{13}}.$$

(b) Most is csak az érdekel, hogy milyen lapok vannak nálunk. Tehát az összes eset

$$\binom{52}{13}.$$

A B_2 eseménynél a kedvező esetek azok, amikor a kőr ászt kivettuk magunknak, és a maradék 51 lapból választunk 12-t, azaz

$$\binom{51}{12}.$$

Vagyis a keresett valószínűség

$$\mathbf{P}(B_2) = \frac{\binom{51}{12}}{\binom{52}{13}}.$$

Az $A \cap B_2$ esemény azt jelenti, hogy nálunk van a kőr ász, és még pontosan egy ász van nálunk. A kedvező esetek száma

$$\binom{3}{1}\binom{48}{11},$$

hiszen a 3 ászból választunk egyet, a nemászok közül pedig 11-et. Tehát

$$\mathbf{P}(A \cap B_2) = \frac{\binom{3}{1}\binom{48}{11}}{\binom{52}{13}},$$

ahonnan a keresett feltételes valószínűség

$$\mathbf{P}(A|B_2) = \frac{\mathbf{P}(A \cap B_2)}{\mathbf{P}(B_2)} = \frac{\binom{3}{1}\binom{48}{11}}{\binom{51}{12}}.$$

(c) Most olyan eseményre kérdez rá a feladat, amiben van első lap. Amikor az összes eset $\binom{52}{13}$ akkor van a kezünkben 13 lap, de nincs sorrendje, tehát nincs első lap. Tehát azon a valószínűségi mezőn, amin az (a–b) részben dolgoztunk, a B_3 esemény nem vizsgálható. Olyan mező kell, amiben van első lap, meg 12 másik. Ekkor az összes esetek száma

$$\binom{52}{13} \cdot 13,$$

hiszen kivettem a 13 lapot, és abból kijelöltem, hogy melyik volt az első. (Ez persze ugyanaz, mint $52 \cdot \binom{51}{12}$.) A kedvező esetek száma a B_3 eseménynél

$$4 \cdot \binom{51}{12},$$

hiszen az első lap a 4 ász közül valamelyik, a többi meg akármi lehet. Tehát

$$\mathbf{P}(B_3) = \frac{4 \cdot \binom{51}{12}}{\binom{52}{13} \cdot 13} = \frac{1}{13}.$$

Persze, hiszen annak a valószínűsége, hogy egyetlen lap éppen ász az $\frac{4}{52} = \frac{1}{13}$. Az $A \cap B_3$ esemény azt jelenti, hogy az első lap ász, és még pontosan egy ászt kapunk. A kedvező esetek száma

$$4 \cdot \binom{3}{1} \cdot \binom{48}{11},$$

hiszen az első ász 4-féle lehet, után a 3 megmaradt ászból kell egy, és a 48 nemászból pedig 11. Tehát

$$\mathbf{P}(A \cap B_3) = \frac{4 \cdot \binom{3}{1} \cdot \binom{48}{11}}{\binom{52}{13} \cdot 13},$$

így a keresett feltételes valószínűség

$$\mathbf{P}(A|B_3) = \frac{\mathbf{P}(A \cap B_3)}{\mathbf{P}(B_3)} = \frac{\frac{4 \cdot \binom{3}{1} \cdot \binom{48}{11}}{\binom{52}{13} \cdot 13}}{\frac{1}{13}} = \frac{12 \cdot \binom{48}{11}}{\binom{52}{13}}.$$

(d) Megint van első lap, ezért a (c)-részbeli mezőn vagyunk. Annak a valószínűsége, hogy az első lap a kőr ász 1/52, hiszen 52 lapból választunk egyet, tehát

$$\mathbf{P}(B_4) = \frac{1}{52}.$$

Az $A \cap B_4$ azt jelenti, hogy az első lap a kőr ász, és utána még pontosan egy ászt kapunk. Az összes esetek száma

$$\binom{52}{13} \cdot 13,$$

a kedvező esetek száma pedig

$$1 \cdot \binom{3}{1} \cdot \binom{48}{11},$$

hiszen az első lap csak a kőr ász lehet, után pedig 1 ászt húzunk 3 közül, és a 48 nemász közül 11-et. Így

$$\mathbf{P}(A \cap B_4) = \frac{3 \cdot \binom{48}{11}}{\binom{52}{13} \cdot 13}.$$

A keresett feltételes valószínűség

$$\mathbf{P}(A|B_4) = \frac{\mathbf{P}(A \cap B_4)}{\mathbf{P}(B_4)} = \frac{\frac{3 \cdot \binom{48}{11}}{\binom{52}{13} \cdot 13}}{\frac{1}{52}} = \frac{12 \cdot \binom{48}{11}}{\binom{52}{13}}.$$

□

1.5.3. Egy cukrászdában 3 cukrász A, B és C süt süteményt, és a sütemények 2, 3 illetve 5%-át rontják el. A sütemények 50%-át A, 30 %-át B, 20%-át pedig C készíti. Mennyi a valószínűsége, hogy A sütötte a süteményt, feltéve, hogy az rossz?

Megoldás. Ez tipikus Bayes-tételes feladat. Jelölje A, B, C azt, hogy ki sütötte a süteményt, R pedig azt, hogy rossz a sütemény. Ekkor

$$\mathbf{P}(A) = 0,5, \quad \mathbf{P}(B) = 0,3, \quad \mathbf{P}(C) = 0,2,$$

valamint

$$\mathbf{P}(R|A) = 0,02, \quad \mathbf{P}(R|B) = 0,03, \quad \mathbf{P}(R|C) = 0,05,$$

Arra vagyunk kíváncsiak, hogy egy rossz süteményt milyen valószínűséggel készített A szakács, azaz $\mathbf{P}(A|R)$. Definíció szerint

$$\mathbf{P}(A|R) = \frac{\mathbf{P}(A \cap R)}{\mathbf{P}(R)}.$$

Annak a valószínűsége, hogy egy sütemény rossz

$$\begin{aligned}\mathbf{P}(R) &= \mathbf{P}(A)\cdot\mathbf{P}(R|A)+\mathbf{P}(B)\cdot\mathbf{P}(R|B)+\mathbf{P}(C)\cdot\mathbf{P}(R|C)\\ &= 0,5\cdot 0,02+0,3\cdot 0,03+0,2\cdot 0,05=0,029,\end{aligned}$$

és

$$\mathbf{P}(R\cap A)=\mathbf{P}(A)\cdot\mathbf{P}(R|A)=0,01.$$

Tehát

$$\mathbf{P}(A|R)=\frac{0,01}{0,029}=\frac{10}{29}.$$

□

1.5.4. A koronavírus teszt hatékonysága függ attól, hogy a fertőzött hány napja kapta el a betegséget. Tegyük fel, hogy a fertőzés minden esetben 12 napig tart (ez kb. igaz, ezért van 14 nap karantén). Ha a fertőzött $1-3$ napja kapta el a betegséget, akkor a teszt ezt nem tudja kimutatni, ha $4-5$ napja, akkor $0,5$ valószínűséggel mutatja ki, ha $6-12$ napja, akkor $0,75$ valószínűséggel (ez nagyjából stimmel). Feltehetjük, hogy egy fertőzött egyén az elmúlt 12 nap bármelyikén egyforma valószínűséggel fertőződött meg.

(a) Mekkora a valószínűsége, hogy egy fertőzött egyén tesztje negatív?

(b) Mivel ez így nem túl hatásos, a következő a szokásos eljárás. A vizsgálandó egyént letesztelik, majd karanténba küldik (ahol elvileg nem fertőződhet meg), és két nap múlva újra letesztelik. Csak akkor engedik ki a karanténból, ha mindkét tesztje negatív. Mekkora a valószínűsége, hogy egy fertőzött mindkét tesztje negatív, és még mindig fertőző? (Vegyük észre, hogy a betegünk a második teszt után meggyógyulhatott, ha elég régen fertőződött meg.)

Megoldás. (a) Aszerint kell felbontani az eseményteret, hogy a beteg hány napja fertőződött meg. Jelölje A, B, C azt az eseményt, hogy a beteg $1-3$, $4-5$, $6-12$ napja beteg, T pedig azt, hogy a teszt pozitív. Ekkor

$$\mathbf{P}(A)=\frac{3}{12},\quad \mathbf{P}(B)=\frac{2}{12},\quad \mathbf{P}(C)=\frac{7}{12}.$$

A teszt hatékonysága

$$\mathbf{P}(T|A)=0,\quad \mathbf{P}(T|B)=0,5,\quad \mathbf{P}(T|C)=0,75.$$

Annak a valószínűsége, hogy pozitív a teszt

$$\begin{aligned}\mathbf{P}(T) &= \mathbf{P}(A)\cdot\mathbf{P}(T|A)+\mathbf{P}(B)\cdot\mathbf{P}(T|B)+\mathbf{P}(C)\cdot\mathbf{P}(T|C)\\ &= \frac{1}{4}\cdot 0+\frac{1}{6}\cdot 0,5+\frac{7}{12}\cdot 0,75=0,52,\end{aligned}$$

azaz annak a valószínűsége, hogy egy fertőzött tesztje negatív

$$\mathbf{P}(\text{negatív a teszt}) = 1 - \mathbf{P}(T) = 0,48,$$

ami tényleg nem túl hatékony.

(b) Most tovább kell osztani az eseményteret, hiszen ha a beteg az első teszt előtt $1-3$ nappal fertőződött meg, akkor a második teszt napján már $3-5$ napja beteg. Az viszont nem mindegy, hogy 3 vagy 5, hiszen más a teszt hatékonysága. Jelölje A_i azt azt eseményt, hogy a beteg az első teszt előtt i nappal betegedett meg, ahol $i = 1, 2, \ldots, 12$. Ekkor

$$\mathbf{P}(A_i) = \frac{1}{12}, \quad i = 1, \ldots, 12.$$

Jelölje N azt az eseményt, hogy mindkét teszt negatív lett és még beteg a második teszt után is. Mivel pontosan 12 napig tart a betegség, ezért ha valaki az első teszt előtt 11 nappal fertőződött meg, akkor a második teszt után már egészséges. Azaz, ekkor nem kell számolgatni,

$$\mathbf{P}(N|A_{11}) = \mathbf{P}(N|A_{12}) = 0.$$

Ha A_3 követekezett be, azaz az első teszt előtt 3 napja fertőződött meg a beteg, akkor az első teszt biztos negatív. A második tesztkor már 5 napja beteg, így az $0,5$ valószínűséggel negatív. Azaz

$$\mathbf{P}(N|A_3) = 1 \cdot 0,5 = 0,5.$$

A többi eset hasonlóan számolható:

$$\begin{aligned}
&\mathbf{P}(N|A_1) = 1,\\
&\mathbf{P}(N|A_2) = \mathbf{P}(N|A_3) = 1 \cdot 0,5 = 0,5,\\
&\mathbf{P}(N|A_4) = \mathbf{P}(N|A_5) = 0,5 \cdot 0,25 = 0,125,\\
&\mathbf{P}(N|A_i) = 0,25 \cdot 0,25 = 0,0625, \quad i = 6, 7, \ldots, 10.
\end{aligned}$$

Összegezve

$$\begin{aligned}
\mathbf{P}(N) &= \sum_{i=1}^{12} \mathbf{P}(A_i)\mathbf{P}(N|A_i)\\
&= \frac{1}{12}\left(1 + 2 \cdot 0,5 + 2 \cdot 0,125 + 5 \cdot 0,0625 + 2 \cdot 0\right) = 0,21.
\end{aligned}$$

Ez már lényegesen jobb, de 100 fertőzöttből 20 még mindig átcsúszik. Azt is vegyük észre, hogy 16 már felgyógyul ($1/6 = \mathbf{P}(A_{11} \cup A_{12})$ a valószínűsége, hogy a második teszt idejére már letelt a 12 nap), tehát a 2 negatív teszt feltétel az valójában 84 betegből szűr ki 64-et. $\square$

1.5.5 (∗). A sztochasztika tanszék egyik oktatója p valószínűséggel szokott bejönni a tanszékre. Ha ismerőseinek azt mondta, hogy aznap bejön, akkor annak a valószínűsége, hogy pontosan k-an keresik telefonon $e^{-\mu}\mu^k/k!$, ha pedig azt mondta, hogy nem, akkor $e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \ldots$, $0 < \lambda < \mu$. Feltéve, hogy k hívás érkezett, mennyi a valószínűsége, hogy aznap bent volt az oktató? Vizsgáljuk a $k \to \infty$ esetet. ([1])

Megoldás. Jelölje A_k azt az eseményt, hogy pontosan k-an keresik telefonon, B pedig azt, hogy bejön. Ekkor

$$\mathbf{P}(B) = p, \quad \mathbf{P}(A_k|B) = \frac{\mu^k}{k!}e^{-\mu}, \quad \mathbf{P}(A_k|B^c) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, 2, \ldots.$$

A kérdés

$$\mathbf{P}(B|A_k) = \frac{\mathbf{P}(B \cap A_k)}{\mathbf{P}(A_k)},$$

ahol

$$\begin{aligned}\mathbf{P}(A_k) &= \mathbf{P}(A_k \cap B) + \mathbf{P}(A_k \cap B^c)\\ &= \mathbf{P}(B)\mathbf{P}(A_k|B) + \mathbf{P}(B^c)\mathbf{P}(A_k|B^c)\\ &= p\frac{\mu^k}{k!}e^{-\mu} + (1-p)\frac{\lambda^k}{k!}e^{-\lambda}.\end{aligned}$$

Tehát

$$\mathbf{P}(B|A_k) = \frac{p\frac{\mu^k}{k!}e^{-\mu}}{p\frac{\mu^k}{k!}e^{-\mu} + (1-p)\frac{\lambda^k}{k!}e^{-\lambda}} = \frac{pe^{-\mu}}{pe^{-\mu} + (1-p)(\lambda/\mu)^k e^{-\lambda}}.$$

Mivel $\lambda < \mu$

$$\lim_{k\to\infty} \mathbf{P}(B|A_k) = 1.$$

Ez valahogy logikus, mert többen keresik akkor, ha azt mondja, hogy bejön. Ha nagyon sokan keresik, akkor egyre valószínűbb, hogy bent van. □

1.5.6. Magyar kártyából kapunk 12 lapot, a másik két játékos 10–10 lapot kap. Mennyi a valószínűsége, hogy pontosan 5 pirosat kapunk, közte a piros hetest? Feltéve, hogy pontosan 5 pirosat kaptunk közte a piros hetest, mennyi a valószínűsége, hogy a másik 3 piros egy kézben van?

1.5.7. *Doppingteszt.* Kifejlesztenek egy új doppingtesztet, mely a doppingolók 99%-ánál pozitív eredményt ad, azonban a nem doppingoló sportolók 1%-nál is tévesen pozitív eredményt ad. Tegyük föl, hogy a sportolók 1%-a doppingol. Mennyi annak a valószínűsége, hogy egy véletlenül kiválasztott sportoló

(a) doppingtesztje pozitív?

(b) doppingolt, ha tudjuk, hogy a doppingtesztje pozitív?

1.5.8. Két pénzérme közül az egyik szabályos, a másik cinkelt, 1/4 valószínűséggel ad fejet. Véletlenszerűen kiválasztjuk az egyiket, majd ezzel kétszer dobunk. Mennyi a valószínűsége, hogy két fejet kapunk? Ha két fejet kapunk, mennyi a valószínűsége, hogy a szabályos érmét választottuk?

1.5.9. Egy hallgató p valószínűséggel tudja a választ egy kérdésre. Ha nem tudja, akkor az n lehetséges válasz közül véletlenül választ egyet. Mennyi legyen a lehetséges válaszok n száma, hogy az oktató legalább $0,9$ valószínűséggel következtethessen arra a hallgató jó válaszából, hogy a hallgató tudta a választ?

1.5.10. Az iskolai karácsonyi vásárra készülődve Blanka, Csenge és Dóri feladata az volt, hogy különböző figurákat hajtogassanak színes papírból. Összesen 70 figurát hajtogattak. A figurák kétheted részét Dóri készítette, a maradékot pedig fele-fele arányban Blanka és Csenge. A Blanka által készített figurák 40%-a volt karácsonyfa, a Csenge által készített figuráknak 60%-a, a Dóri által készített figuráknak pedig 30%-a. Az első vásárló a vásáron Blanka édesanyja volt; ő megvett egy véletlenszerűen kiválasztott karácsonyfa-figurát. Határozza meg annak a valószínűségét, hogy a figurát éppen Blanka készítette! (Érettségi 2018)

1.5.11. Aladár hétfő reggelenként 7:15-kor indul el otthonról, hogy 8:00-ra beérjen az egyetemre. Gyalog kimegy a buszmegállóba, 20 percet buszozik, aztán villamosra száll át, amelyen 15 percet utazik, végül ismét gyalogol az egyetemig. A buszon 0,5, a villamoson pedig 0,2 valószínűséggel hallja meg, ha csörög a mobiltelefonja, míg gyaloglás közben biztosan észreveszi, ha hívják. Ha hétfő reggel 7:15 és 8:00 között egy véletlen időpontban felhívjuk, és felveszi a telefonját, akkor mennyi a valószínűsége, hogy épp villamoson van?

1.5.12. Vándorlásai közben Odüsszeusz egyszer egy hármas útelágazáshoz ért. Tudta, hogy az egyik út Athénba, a másik Mükénébe, a harmadik pedig Spártába vezet. Azt is tudta, hogy az athéniak átlagosan minden harmadik alkalommal mondanak igazat, a mükénéiek minden második alkalommal, a spártaiak pedig becsületesek, sosem hazudnak. Kockadobással döntötte el, melyik utat válassza, egyforma esélyt adva mindegyiknek. Ezután ment, mendegélt, míg egy városba nem ért. Itt az első szembejövőtől megkérdezte, hogy mennyi kettő meg kettő, és azt a választ kapta, hogy négy. Mennyi a valószínűsége, hogy Athénba érkezett?

1.5.13. Bergengóciában tombol a Covid-$\pi^2/6$ járvány. Egy bergengóc élelmiszerboltban 0,5 valószínűséggel sokan, 0,3 valószínűséggel kevesen vannak, 0,2 valószínűséggel pedig üres a bolt. A megfertőződés valószínűsége rendre 0,6, 0,3 és 0,1. Mennyi a valószínűsége annak, hogy egy egészséges bergengóc megfertőződik a bevásárlás során?

1.5.14. A szegedi Gabonakutatóban két új kukorica vetőmagot fejlesztettek ki. Az A típusú strapabíró vetőmag csapadékos, átlagos, és száraz időjárás esetén egyaránt 0,2 valószínűséggel hoz prémium termést. A B típusú vetőmag csapadékos időjárás estén 0,5, átlagos időjárás esetén 0,1 valószínűséggel hoz prémium termést, száraz időben viszont biztosan nem hoz jó termést. A meteorológusok 0,5 valószínűséggel átlagos, 0,2 valószínűséggel sok, 0,3 valószínűséggel pedig kevés csapadékot jósolnak. Melyik vetőmagot válasszuk, ha minél nagyobb valószínűséggel akarunk prémium termést?

1.5.15. Máté az iskolából hazafelé menet betér apukájával a Csiga pékségbe. Máté kedvencei a meggyes és a mákos-meggyes rúd. A meggyes rúdban 0,4 valószínűséggel, a mákos-meggyes rúdban 0,2 valószínűséggel van meggymag. Máté pénzfeldobással dönt, hogy meggyes vagy mákos-meggyes rudat kér. Mennyi a valószínűsége, hogy lesz meggymag a péksüteményben? Máté apukája mákos-meggyes rudat választ. Mennyi a valószínűsége, hogy mindkettejük süteményében lesz meggymag?

1.5.16. Máté nagymamája meggylevest készít a vasárnapi ebédhez. Összesen 5 szem szegfűszeget tesz a levesbe. A levest 4 egyforma adagra osztják. Máté már az első kanál levesében talál egy szegfűszeget. Mennyi a valószínűsége, hogy legalább 3 szegfűszeg lesz a levesében, feltéve, hogy van a levesében szegfűszeg?

1.5.17. Egy könyvelő irodában három alkalmazott van, akik egyenlő arányban osztoznak a munkán. Felvesznek egy negyedik embert is az irodára, hogy csökkentsék a terhelést, így már négyfelé oszlik a munka egyenlő arányban. A tapasztalt könyvelők ritkán, az eseteknek mindössze 3%-ában hibáznak, míg az újonc, aki csak most tanult bele a könyvelésbe 10%-os hibarátával bír. A könyvelők főnöke kézbevesz egy adóbevallást és azt látja, hogy az hibásan van kitöltve, mekkora a valószínűsége, hogy az új alkalmazottat terheli a felelősség?

1.5.18. Egy matematikuscsaládban a fiú megfigyelte, hogy mikor hazaér, az esetek 20%-ában senki nincs otthon, 50%-ában csak az édesanyja, a maradék esetekben pedig mindkét szülője. Tudja, hogy édesanyja az esetek 80%-ában magára zárja az ajtót; ha mindketten otthon vannak, akkor már csak 40%-ban zárják be, de ha nincs otthon senki, szórakozottságból akkor is az esetek

5%-ában nyitva marad az ajtó. A fiú egy délután hazaér, és zárva találja az ajtót. Mekkora a valószínűsége, hogy van otthon valaki?

1.5.19 $(*)$**.** Egy fiú és egy lány megbeszéli, hogy két utca kereszteződésénél találkoznak egy meghatározott időpontban. Elfelejtik megbeszélni, hogy a négy sarok közül melyiknél várnak egymásra. Az útkereszteződés nagyon forgalmas, nem lehet átlátni a többi sarokra. Mindketten pontosan érkeznek, és ha a másik nincs ott, akkor 2,5 perc után átmennek a szomszédos sarkok valamelyikére, 1/2–1/2 valószínűséggel. Ez fél percet vesz igénybe, majd ha megint nem találkoztak, akkor 2,5 perc után megint sarkot váltanak. Először mindketten $1/4$ valószínűséggel választanak sarkot. Természetesen az is találkozásnak számít, ha egymással szembe jönnek az úttesten.

(a) Mennyi a valószínűsége, hogy az első percen belül találkoznak?

(b) Mennyi annak a p_n valószínűsége, hogy az első $3n$ percen belül találkoznak?

(c) Mennyi annak az r_n valószínűsége, hogy pontosan a $3n$-edik percben találkoznak?

(d) Igazoljuk, hogy egy valószínűséggel véges időn belül találkoznak.

([1])

1.5.20 $(**)$**.** A parti tüzérség 1 km távolságban felfedez egy ellenséges cirkálót, és elkezd rá tüzelni, percenként egy lövést adva le. A cirkáló az első lövés leadásakor menekülni kezd 60 km/h sebességgel. A találat valószínűsége x km távolság esetén $0,75\,x^{-2}$. Ha egy lövés talált, akkor még mindig $1/4$ valószínűséggel a cirkáló nem süllyed el, és tovább menekül. Mekkora valószínűséggel menekül el a cirkáló? (Schweitzer, 1950)

Megoldás. Ez egy nehéz feladat.

Jelölje A_n az azt eseményt, hogy a cirkáló túléli az n-edik lövést. Nyilván $A_1 \supset A_2 \supset \ldots$ monoton csökkenő halmazsorozat, és ha A az az esemény, hogy elmenekül, akkor

$$\cap_{n=1}^{\infty} A_n = A.$$

A valószínűség tulajdonságai szerint

$$\mathbf{P}(A) = \mathbf{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbf{P}(A_n).$$

Nyilván

$$\mathbf{P}(A_1) = \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} \quad \text{(nem talál, vagy talál, de nem süllyed).}$$

Hasonlóan,

$$\mathbf{P}(A_n) = \mathbf{P}(A_{n-1})\left(1 - \frac{3}{4n^2} + \frac{3}{4n^2}\frac{1}{4}\right) = \mathbf{P}(A_{n-1})\left(1 - \frac{9}{16n^2}\right).$$

Innen indukcióval

$$\mathbf{P}(A_n) = \prod_{k=1}^{n}\left(1 - \frac{9}{16k^2}\right),$$

és így a limesz esemény valószínűsége

$$\mathbf{P}(A) = \prod_{k=1}^{\infty}\left(1 - \frac{9}{16k^2}\right).$$

A valószínűségszámítás része kész a feladatnak. De ez mennyi?

A Stirling-formula szerint (ezt egy kicsit később belátjuk, de biztos előkerült valamiből korábban)

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \quad \text{amint } n \to \infty.$$

Most egy egyszerű, de elég macerás számolás jön. Vezessük be szemifaktoriális jelölést:

$$(2n+1)!! = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1), (2n)!! = 2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n) = 2^n\, n!.$$

Mivel

$$1 - \frac{9}{16k^2} = \frac{(4k-3)\cdot(4k+3)}{16k^2},$$

így, a Stirling-formulát beírva

$$\begin{aligned}
\prod_{k=1}^{n}((4k-3)\cdot(4k+3)) = \frac{(4n+3)!!}{3(4n+1)} &= \frac{(4n+3)!}{3(4n+1)2^{2n+1}(2n+1)!} \\
&\sim \frac{1}{3(4n+1)2^{2n+1}} \frac{(\frac{4n+3}{e})^{4n+3}\sqrt{2\pi(4n+3)}}{(\frac{2n+1}{e})^{2n+1}\sqrt{2\pi(2n+1)}} \\
&\sim \frac{\sqrt{2}}{3(4n+1)2^{2n+1}e^{2n+2}} \frac{(4n+3)^{4n+3}}{(2n+1)^{2n+1}}.
\end{aligned}$$

Ezt beírva

$$\begin{aligned}
\prod_{k=1}^{n}\left(1 - \frac{9}{16k^2}\right) &\sim \frac{\sqrt{2}}{3(4n+1)2^{2n+1}e^{2n+2}} \frac{(4n+3)^{4n+3}}{(2n+1)^{2n+1}} \cdot \frac{1}{16^n(\frac{n}{e})^{2n}2\pi n} \\
&\sim \frac{\sqrt{2}}{3(4n+1)2^{2n+1}e^{2}} \frac{(4n+3)^{4n+3}}{(2n+1)^{2n+1}n^{2n}} \cdot \frac{1}{16^n 2\pi n}.
\end{aligned}$$

Még gyúrni kell, de minden egyszerű analízis. Egyrészt

$$\lim_{n\to\infty} \frac{(4n+3)^3}{(4n+1)(2n+1)n} = 8.$$

Másrészt

$$\frac{(4n+3)^{4n}}{(2n+1)^{2n}n^{2n}} = 2^{6n}\left[\frac{(n+3/4)^2}{(n+1/2)n}\right]^{2n}.$$

A jobb oldalon szereplő második tényező 1^∞ típusú határérték. Az alapot átírva

$$\frac{(n+3/4)^2}{(n+1/2)n} = \frac{n^2+\frac{3}{2}n+\frac{9}{16}}{n^2+\frac{1}{2}n} = 1 + \frac{n+\frac{9}{16}}{n^2+\frac{1}{2}n} = 1 + \frac{1}{n} + O(n^{-2}).$$

Tehát

$$\left(\frac{(n+3/4)^2}{(n+1/2)n}\right)^n = \left(1+\frac{1}{n}+O(n^{-2})\right)^n \to e.$$

Összegezve,

$$\prod_{k=1}^{n}\left(1-\frac{9}{16k^2}\right) \sim \frac{\sqrt{2}\cdot 8\cdot 2^{6n}}{3\cdot 2^{2n+1}e^2 16^n\, 2\pi}e^2 = \frac{\sqrt{2}\cdot 2}{3\pi} \approx 0,3$$

Tehát

$$\mathbf{P}(\text{megmenekül a cirkáló}) = \frac{2\sqrt{2}}{3\pi}.$$

Ezt máshogy is lehet. Ha p egy valós együtthatós n-ed fokú polinom, és gyökei $x_1, \dots, x_n$, akkor

$$p(x) = p(0)\prod_{i=1}^{n}\left(1-\frac{x}{x_i}\right).$$

A $\sin x$ függvény zérushelyei 0, $k\pi$, $k\in\mathbb{Z}$, ezért a fenti előállítás a $\sin x/x$ függvényre a következőt adja:

$$\begin{aligned}\sin x &= x\prod_{k\in\mathbb{Z}\setminus\{0\}}\left(1-\frac{x}{k\pi}\right)\\ &= x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2\pi^2}\right).\end{aligned}$$

Itt persze csaltunk, mert a $\sin x$ az nem polinom. A végtelen szorzat alak a Weierstrass-féle faktorizációs tétel, ami lehet, hogy lesz majd komplex függvénytanból. Innen az $x = 3\pi/4$ helyettesítéssel

$$\prod_{k=1}^{\infty}\left(1-\frac{9}{16k^2}\right) = \frac{\sin\frac{3}{4}\pi}{\frac{3}{4}\pi} = \frac{\sqrt{2}\cdot 2}{3\pi}.$$

$\square$

1.6. Függetlenség

1.6.1. Legyen A önmagától független esemény. Mutassuk meg, hogy $\mathbf{P}(A) = 0$ vagy 1!

Megoldás. Ehhez csak a definíciót kell alkalmazni, nem túl izgalmas. Az, hogy A független önmagától az azt jelenti, hogy

$$\mathbf{P}(A \cap A) = \mathbf{P}(A) \cdot \mathbf{P}(A).$$

Mivel $A \cap A = A$, ez azt jelenti, hogy $\mathbf{P}(A) = \mathbf{P}(A)^2$. Ebből következik, hogy $\mathbf{P}(A) = 0$ vagy $\mathbf{P}(A) = 1$. $\square$

1.6.2. Francia kártyapakliból véletlenszerűen húzunk egy lapot. Jelölje D azt az eseményt, hogy dámát húzunk, K pedig azt, hogy kőrt. Függetlenek-e D és K? (A francia kártyánál 52 lap van, 4 szín (kőr, káró, pikk, treff), 13 figura (Á, 2 - 10, J, D, K).)

Megoldás. Ekkor $D \cap K$ az az esemény, hogy a kőr dámát húztuk ki, így $\mathbf{P}(D \cap K) = 1/52$. Ugyanakkor $\mathbf{P}(D) = 4/52 = 1/13$ és $\mathbf{P}(K) = 13/52 = 1/4$, azaz a két esemény független. $\square$

1.6.3. Egy dobókockával tízszer dobunk. Jelölje A azt az eseményt, hogy az első 5 dobás során nincs hatos, B pedig azt, hogy tíz dobás közt nincs egyes. Mekkora az A és a B események valószínűsége? Függetlenek-e A és B?

Megoldás. Az A esemény azt jelenti, hogy az első 5 dobás során nincs hatos, így mindig 5 értéket dobhatunk, ezért

$$\mathbf{P}(A) = \left(\frac{5}{6}\right)^5 = \frac{5^5}{6^5}.$$

Hasonlóan, B eseménynél 10 dobás során nincs egyes, így

$$\mathbf{P}(B) = \frac{5^{10}}{6^{10}}.$$

Az $A \cap B$ esemény esetén az első 5 dobásra 4 lehetőségem van, a második ötre pedig 5, így

$$\mathbf{P}(A \cap B) = \frac{4^5 \cdot 5^5}{6^{10}}.$$

Látjuk, hogy $\mathbf{P}(A \cap B) \neq \mathbf{P}(A) \cdot \mathbf{P}(B)$, azaz A és B nem függetlenek. $\square$

1.6.4. (i) Legalább hány lottószelvényt kell kitölteni ahhoz, hogy egy sorsolásnál a telitalálat valószínűsége legalább 1/2 legyen? (ii) Legalább hány hétig kell játszani egyetlen szelvénnyel, hogy annak a valószínűsége, hogy legalább egyszer volt telitalálatunk legalább 1/2 legyen?

Megoldás. (i) Ha k szelvényt veszünk (és mindegyiken más számötöst írunk be), akkor a nyerési esélyünk $k/\binom{90}{5}$. Ahhoz, hogy ez a valószínűség 1/2-nél nagyobb legyen, nyilván az kell, hogy

$$k \geq \frac{1}{2}\binom{90}{5} = 21974634.$$

(ii) Egy héten, annak a valószínűsége, hogy *nem* nyerünk,

$$\mathbf{P}(\text{nincs telitalálatunk 1 héten 1 szelvénnyel}) = 1 - \frac{1}{\binom{90}{5}}.$$

A sorsolások eredményei függetlenek, így annak a valószínűsége, hogy k héten nem nyerünk egyszer sem

$$\mathbf{P}(k \text{ héten 1 szelvénnyel egyszer sincs telitalálatunk}) = \left(1 - \frac{1}{\binom{90}{5}}\right)^k.$$

Az, hogy legalább egyszer telitalálatunk van k hét során, az annak az eseménynek az ellentett eseménye, hogy *egyszer sincs telitalálatunk k hét során*, így

$$\mathbf{P}(\ k \text{ hét során legalább egy telitalálatunk van}) = 1 - \left(1 - \frac{1}{\binom{90}{5}}\right)^k.$$

Amint $k \to \infty$ az utóbbi valószínűség 1-hez tart (hiszen 1-nél kisebb számot hatványozva 0-hoz tart a sorozat). Tehát olyan k-t keresünk, melyre a fenti valószínűség 1/2-nél nagyobb. Azaz

$$\left(1 - \frac{1}{\binom{90}{5}}\right)^k \leq \frac{1}{2}.$$

Logaritmust véve és rendezve (1-nél kisebb szám logaritmusa negatív, és negatív számmal szorozva az egyenlőtlenséget a relációjel fordul), kapjuk, hogy

$$k \geq \frac{\ln 2}{-\ln\left(1 - \frac{1}{\binom{90}{5}}\right)} = 30463311.$$

Azaz legalább ennyi hétig kell játszani. $\square$

1.6.5. András egy év mind az 52 hetében egy-egy ugyanúgy kitöltött szelvénnyel játszik az ötöslottón. Bea ellenben az év utolsó sorsolása előtt vesz 52 szelvényt, és azokat (páronként különbözőféleképpen kitöltve) egyszerre játssza meg. Andrásnak vagy Beának van nagyobb esélye arra, hogy legyen telitalálatos szelvényük? (KöMaL B5256)

1.6.6. A vezeték nélküli hálózati kapcsolatot létrehozó egységek (wifi routerek) 3%-a 2 éven belül meghibásodik (ezt úgy tekinthetjük, hogy 0,03 annak a valószínűsége, hogy egy készülék meghibásodik 2 év alatt). A meghibásodott eszközt garanciálisan kicserélik. Az iskola 20 ilyen eszközt vásárolt. Mennyi a valószínűsége annak, hogy 2 év alatt legfeljebb egy hibásodik meg a vásárolt eszközök közül? (Érettségi 2018)

1.6.7. Zajos csatornán akarunk 0–1 jelsorozatot küldeni. A csatorna minden bitet egymástól függetlenül $p > 0$ valószínűséggel elront. Mennyi a valószínűsége, hogy egy 3 bites üzenet hiba nélkül átmegy?

1.6.8. Háromszor feldobunk egy szabályos érmét. Jelentse A azt az eseményt, hogy a dobások között fej és írás is előfordul, B pedig azt, hogy legfeljebb egy írás fordul elő. Függetlenek-e A és B?

1.6.9. A sztochasztika tanszék egyik oktatója az óra során elhasznált kréták kis darabjait 1 méter távolságból a szemetesbe dobja. A dobások egymástól függetlenül 95% eséllyel találnak. Ha egy gyakorlat alatt pontosan 6-szor próbálkozik, mennyi a valószínűsége, hogy mindig talál? Hányadik gyakorlaton lesz először 0.5-nél nagyobb annak a valószínűsége, hogy (az összes eddigi dobása közül) legalább egyszer nem talál?

1.6.10. Bence egy labdát rugdos egy háznak, amely 10 m hosszú és 5 m magas. A házon két 2 m $\times$ 1,5 m-es ablak van. Mennyi a valószínűsége, hogy 5 rúgásból egyszer sem talált ablakot?

1.6.11. Egy dobókockával n-szer dobunk. Jelölje A azt az eseményt, hogy az első m dobás során nincs hatos, B pedig azt, hogy az n dobás közt nincs egyes, $m < n$. Mekkora az A és a B események valószínűsége? Igazoljuk, hogy $\mathbf{P}(A \cap B) < \mathbf{P}(A)\mathbf{P}(B)$.

1.6.12 (∗). Válasszunk találomra az $1, 2, \ldots, n$ számok közül úgy, hogy mindegyiket $1/n$ valószínűséggel választjuk. Jelölje A_p azt az eseményt, hogy a választott szám p-vel osztható.

(a) Igazoljuk, hogy ha p_1 és p_2 relatív prím és $p_1 p_2 | n$, akkor A_{p_1} és A_{p_2} függetlenek.

(b) Igazoljuk, hogy

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

ahol $\varphi(n)$ az Euler–féle függvény, azaz $\varphi(n)$ az n-nél kisebb n-hez relatív prím pozitív egészek száma.

1.6.13 $(*)$**.** A szultán birodalmának mind az 1024 matematikusát börtönbe záratta. Mindegyikük csak a saját réztalléros érméjét tarthatta meg. A matematikusok tudják, hányan vannak, de semmiféle módon nem képesek kommunikálni egymással. A szultán a születésnapján nagy kegyesen a következő játékot ajánlotta a matematikusoknak: az udvaron egyenként vagy 0-t, vagy 1-et mondanak. Ha a mondott számok összege 1, akkor szabadon bocsátja őket. (A matematikusok nem adhatnak jelet egymásnak, nem tudják, hogy őket hányadiknak vitték ki, vagy hogy az előttük az udvaron lévők mit csináltak.) Mekkora eséllyel szabadulhatnak ki a matematikusok? (KöMaL B.5018)

2. Véletlen változók

2.1. Diszkrét véletlen változók

2.1.1. Egy fiókban három pár kesztyű van összekeveredve: az egyik pár fekete, a másik szürke, a harmadik piros. (A három pár kesztyű csak a színében különböző.) A fiókból egyesével elkezdjük kihúzni a kesztyűket úgy, hogy húzás előtt nem nézzük meg a kesztyű színét, és a kihúzott kesztyűket nem tesszük vissza a fiókba. Addig folytatjuk a húzást, amíg lesz két azonos színű kesztyűnk. Határozza meg annak a hat eseménynek a valószínűségét, hogy ehhez 1, 2, 3, 4, 5, illetve 6 kesztyű kihúzására lesz szükség, majd számítsa ki a húzások számának várható értékét és szórását!

Megoldás. Ez egyszerű, pár éve érettségi feladat volt. Jelölje ξ azt, hogy hányszor kellett húzni, hogy legyen egy pár. Világos, hogy ξ lehetséges értékei 2,3,4. Tehát
$$\mathbf{P}(\xi = 1) = \mathbf{P}(\xi = 5) = \mathbf{P}(\xi = 6) = 0.$$
Elsőre húzunk valamit. Ha $\xi = 2$, akkor a második kihúzott kesztyű éppen az első párja. Mivel 5 kesztyű maradt, ennek a valószínűsége
$$\mathbf{P}(\xi = 2) = \frac{1}{5}.$$
Ha $\xi = 3$, akkor másodikra nem húztuk ki az első párját, ennek a valószínűsége 4/5, harmadikra viszont párt húztunk. A harmadik húzás előtt 4 kesztyű van, ebből 2 jó, tehát a pár valószínűsége 1/2. Összegezve
$$\mathbf{P}(\xi = 3) = \frac{4}{5} \cdot \frac{2}{4} = \frac{2}{5}.$$

Itt valójában egy feltételes valószínűséges gondolatmenetet követtünk. Úgy is gondolkodhatunk, hogy az első húzás után van 5 kesztyű, és kétszer húzunk. Mivel fontos a sorrend az összes esetek száma $5 \cdot 4 = 20$. A kedvező eseteknél a második húzás nem pár, erre van 4 lehetőségünk, míg a harmadik húzásra párt húzunk, de már két kesztyűt húztunk ki, úgyhogy a jó esetek száma 2. Összegezve a kedvező esetek száma $4 \cdot 2 = 8$. Persze azt kaptuk mint korábban, mert $\frac{8}{20} = \frac{2}{5}$.

Végül

$$\mathbf{P}(\xi = 4) = 1 - \mathbf{P}(\xi = 2) - \mathbf{P}(\xi = 3) = \frac{2}{5}.$$

Persze ezt is kiszámolhatjuk a korábbiak szerint.

Tehát a valószínűségeloszlás:

$$\mathbf{P}(\xi = 2) = \frac{1}{5}, \quad \mathbf{P}(\xi = 3) = \frac{2}{5}, \quad \mathbf{P}(\xi = 4) = \frac{2}{5}.$$

A várható érték

$$\mathbf{E}(\xi) = 2 \cdot \frac{1}{5} + 3 \cdot \frac{2}{5} + 4 \cdot \frac{2}{5} = \frac{16}{5}.$$

A második momentum

$$\mathbf{E}(\xi^2) = 2^2 \cdot \frac{1}{5} + 3^2 \cdot \frac{2}{5} + 4^2 \cdot \frac{2}{5} = \frac{54}{5},$$

ahonnan a szórásnégyzet

$$\mathbf{D}^2(\xi) = \mathbf{E}(\xi^2) - (\mathbf{E}(\xi))^2 = \frac{54}{5} - \frac{256}{25} = \frac{14}{25},$$

és a szórás $\mathbf{D}(\xi) = \sqrt{14}/5$. □

2.1.2. Egy szabályos érmét feldobunk egymás után háromszor. Jelölje ξ a dobott fejek és a dobott írások számának a különbségét. Határozzuk meg ξ eloszlását és eloszlásfüggvényét (készítsünk ábrát is)! Mi a valószínűsége, hogy ξ a $[-\sqrt{2}, \sqrt{2}]$ intervallumba esik? Adjuk meg ξ várható értékét és szórását!

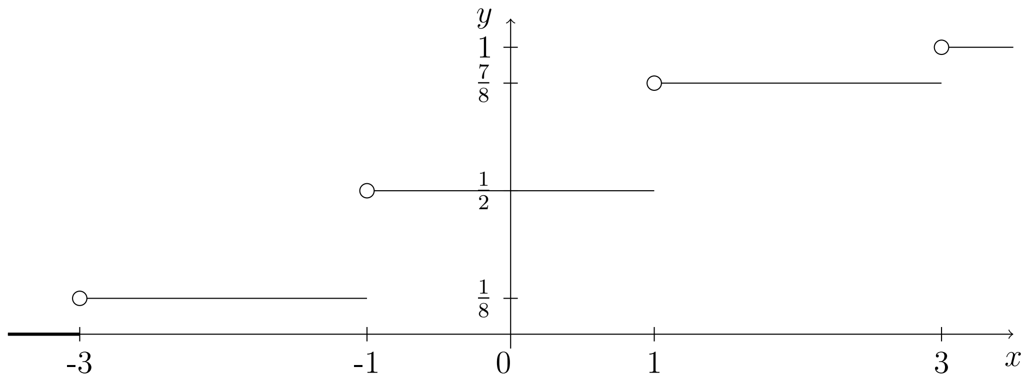
Megoldás. Világos, hogy ξ lehetséges értékei 3 (ha 3 fejet kapunk), 1 (2 fej, 1 írás), -1 (1 fej, 2 írás), és -3 (3 írás). A megfelelő valószínűségek

$$\mathbf{P}(\xi = 3) = \frac{1}{8} = \mathbf{P}(\xi = -3)$$
$$\mathbf{P}(\xi = 1) = \frac{3}{8} = \mathbf{P}(\xi = -1).$$

Ez ξ eloszlása. Az eloszlásfüggvény formulával

$$F(x) = \mathbf{P}(\xi < x) = \begin{cases} 0, & x \le -3, \\ \frac{1}{8}, & -3 < x \le -1, \\ \frac{4}{8}, & -1 < x \le 1, \\ \frac{7}{8}, & 1 < x \le 3, \\ 1, & 3 < x, \end{cases}$$

melynek grafikonja



Ha $\xi \in [-\sqrt{2}, \sqrt{2}]$, akkor $\xi = \pm 1$, tehát

$$\mathbf{P}(\xi \in [-\sqrt{2}, \sqrt{2}]) = \frac{6}{8} = \frac{3}{4}.$$

Végül, a várható érték a definíció alapján

$$\mathbf{E}(\xi) = \sum_{i=1}^{n} x_i \mathbf{P}(\xi = x_i) = (-3) \cdot \frac{1}{8} + (-1) \cdot \frac{3}{8} + 1 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = 0.$$

Azaz, $\mathbf{E}(\xi) = 0$, hát persze, hiszen szimmetrikus az eloszlás. A szórásnégyzet

$$\mathbf{D}^2(\xi) = \mathbf{E}(\xi^2) = \sum_{i=1}^{n} x_i^2 \mathbf{P}(\xi = x_i) = (-3)^2 \cdot \frac{1}{8} + (-1)^2 \cdot \frac{3}{8} + 1^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = 3,$$

tehát a szórás $\mathbf{D}(\xi) = \sqrt{3}$. $\square$

2.1.3. Ötöslottón egy szelvénnyel játszva határozzuk meg a találataink számának eloszlását!

Megoldás. Ezt már láttuk, csak akkor azt kérdeztük, hogy mennyi a valószínűsége, hogy telitalálatunk lesz, vagy pontosan három találatunk lesz.

Jelölje ξ a találataink számát. Világos, hogy ennek lehetséges értékei $0, 1, \ldots, 5$. Ha pontosan k találatunk van, akkor azt az 5 kihúzott szám közül $\binom{5}{k}$-féleképp választhatom ki, a maradék $5-k$ számot a nemkihúzottak közül $\binom{85}{5-k}$-féleképp választhatom ki. Az összes eset $\binom{90}{5}$. A valószínűségi mező klasszikus, tehát

$$\mathbf{P}(\xi = k) = \frac{\binom{5}{k}\binom{85}{5-k}}{\binom{90}{5}}, \quad k = 0, 1, \ldots, 5.$$

Ez a találatok számának az eloszlása. □

2.1.4. Egy urnában 101 golyó van, közülük pontosan három piros. A golyókat visszatevés nélkül egyesével kihúzzuk. Jelölje ξ a második piros sorszámát. Adjuk meg ξ eloszlását!

Megoldás. A második piros sorszáma lehet $2, 3, \ldots, 100$, hiszen csak az első és az utolsó helyen nem lehet. Ha a második piros a k-adik, akkor az elsőre van $k-1$ lehetőségünk, a harmadikra pedig $101-k$ lehetőségünk, hiszen csak annyi a megkötés, hogy az első a második előtt van, a harmadik meg utána. Tehát a kedvező esetek száma $(k-1)\cdot(101-k)$, míg az összes esetek száma $\binom{101}{3}$, hiszen ennyiféleképpen tudjuk a 3 pirosat elhelyezni. Tehát

$$\mathbf{P}(\xi = k) = \frac{(k-1)\cdot(101-k)}{\binom{101}{3}}, \quad k = 2, 3, \ldots, 100.$$

□

2.1.5. Bence addig rugdos egy labdát egy 10 m hosszú és 5 m magas háznak, amíg el nem talál egy ablakot. A házon két 2 m × 1,5 m-es ablak van. Jelölje ξ az első ablakot találó rúgás sorszámát. Adjuk meg ξ eloszlását!

2.1.6. Egy csomag francia kártyát megkevertünk, majd egyesével kihúzzuk a lapokat. Adjuk meg a második ász helyének eloszlását!

2.1.7. Egy hallgató 26 tétel közül 16-ot jelesre tud, 8-at jóra, 2-t viszont csak közepesre. A vizsgatételek kiválasztásakor 26 tétel közül húz 1-et. Adjuk meg a kapott jegy értékének eloszlását, várható értékét és szórását!

2.1.8. Egy embernek n egyforma kinézetű kulcsa van, melyek közül pontosan egy nyitja az ajtót. Emberünk véletlenül választva sorra próbálja a kulcsokat addig, amíg a jó kulcs elő nem kerül. Valamely $k \in \{1, 2, \ldots, n\}$ esetén mennyi a valószínűsége, hogy a k-adik próbálkozása sikeres, ha

(a) a kipróbált rossz kulcsokat mindig félreteszi?

(b) a kipróbált rossz kulcsokat sose teszi félre?

Várhatóan hányadik próbálkozása sikeres?

2.1.9. A Chelsea–Manchester City Bajnokok Ligája döntő játékvezetője Mateu Lahoz lesz. Lahoz a fontos mérkőzéseken 0,5 valószínűséggel nem oszt ki piroslapot, pontosan 1 és pontosan 2 piroslapot egyaránt 0,2 valószínűséggel ad, míg 0,1 valószínűséggel 3 pirosat oszt ki. 3-nál több pirosat nem ad. Adjuk meg és ábrázoljuk a szombaton kiosztandó piroslapok számának eloszlásfüggvényét! Számoljuk ki a várható értékét és szórását!

2.1.10. Máté nagymamája meggylevest készít a vasárnapi ebédhez. Összesen 5 szem szegfűszeget tesz a levesbe. A levest 4 egyforma adagra osztják. Adjuk meg Máté levesében található szegfűszegek számának eloszlását, várható értékét, és szórását!

2.1.11. Mind a 4 pár különböző cipőmet az előszobában levő beépített szekrényben tartom. Az előszobában kiégett a villany, így sötétben keresgélve véletlenszerűen kiveszek 4 cipőt. Jelölje ξ a kivett összeillő párok számát! Adjuk meg ξ eloszlását, várható értékét és szórását!

2.1.12. Száz alma közül tíz férges. Véletlenül kiválasztunk ötöt! Adjuk meg a férges almák számának eloszlását!

2.1.13. Két játékos 3 győzelemig tartó kő-papír-olló párbajt játszik. Tegyük fel, hogy mindketten minden menetben véletlenszerűen (egymástól és a korábbi mutatásoktól függetlenül), $1/3 - 1/3 - 1/3$ eséllyel választják ki, hogy mit mutatnak. Adjuk meg a menetek számának várható értékét! (KöMaL B.5164.)

2.1.14. Máté esténként sokszor belefeledkezik a játékba, emiatt többszöri felszólítás után sem jön vacsorázni. Édesapja egyre nyomatékosabban szól, így annak a valószínűsége, hogy a k-adik felszólításra Máté abbahagyja a játékot, az $1 - 1/(k+1)$, $k = 1, 2, \ldots$. Adjuk meg a szükséges felszólítások számának eloszlását és várható értékét!

2.1.15 ($*$). Határozzuk meg az ötöslottón kihúzott legnagyobb szám eloszlását, várható értékét és szórását!

2.1.16 ($*$). Egy urnában van 20 piros és 30 fehér golyó. Húzzunk ki 20 golyót visszatevés nélkül. Számoljuk ki a kihúzott piros golyók számának várható értékét és szórásnégyzetét!

Megoldás. Jelölje ξ a kihúzott piros golyók számát. Világos, hogy ξ diszkrét véletlen változó, melynek lehetséges értékei $0, 1, 2, \ldots, 20$. Számoljuk ki a $\mathbf{P}(\xi = k)$ valószínűségeket. Klasszikus valószínűségi mezőn vagyunk, az összes eset száma annyi, ahányféleképpen ki tudunk választani 50 golyóból 20-at. Azaz $\binom{50}{20}$. Az $\{\xi = k\}$ esemény azt jelenti, hogy pontosan k db piros golyót húztunk, azaz a 20 pirosból k-t, a 30 fehérből $20-k$-t. Ezt $\binom{20}{k} \cdot \binom{30}{20-k}$ féleképpen tehetjük meg. Tehát

$$\mathbf{P}(\xi = k) = \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}}, \quad k = 0, 1, \ldots, 20.$$

Ez az ξ véletlen változó eloszlása. Ő a negatív binomiális eloszlás.

Innen a várható értéket meghatározhatjuk a definíció alapján. Eszerint

$$\mathbf{E}(\xi) = \sum_{k=0}^{20} k\mathbf{P}(\xi = k) = \sum_{k=1}^{20} k \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}}.$$

Ezt kellene zárt alakra hozni. Vegyük észre, hogy

$$k\binom{20}{k} = 20\binom{19}{k-1}. \tag{$*$}$$

Valóban, egy 20 fős társaságból kell kiválasztani egy k fős bizottságot, és annak a vezetőjét. Ezt számoljuk meg a bal és jobb oldalon is. A bal oldalon először kiválasztjuk a k fős csapatot, és utána annak a vezetőjét, míg a jobb oldalon először a vezetőt választjuk ki, és utána a maradék 19 főből választunk még $k-1$-et. Ezt a formulát felhasználva

$$\sum_{k=1}^{20} k\binom{20}{k} \cdot \binom{30}{20-k} = \sum_{k=1}^{20} 20\binom{19}{k-1} \cdot \binom{30}{20-k}.$$

Az összegben megjelenő $\binom{19}{k-1} \cdot \binom{30}{20-k}$ kifejezés pontosan olyan, mint amit korábban kaptunk, csak most összesen 49 golyó közül, melyből 19 piros és 30 fehér, választunk ki 19-et. Tehát a fenti összeg az összes kiválasztást adja meg, azaz

$$\sum_{k=1}^{20} \binom{19}{k-1} \cdot \binom{30}{20-k} = \sum_{\ell=0}^{19} \binom{19}{\ell} \cdot \binom{30}{19-\ell} = \binom{49}{19}.$$

Ezt visszahelyettesítve a várható értékre kapott formulába

$$\mathbf{E}(\xi) = 20\frac{\binom{49}{19}}{\binom{50}{20}} = 20\frac{2}{5} = 8.$$

Itt felhasználtuk azt is, hogy $50\binom{49}{19} = 20\binom{50}{20}$, ami éppen $(*)$ csak más számokkal.

A szórás meghatározásához először a második momentum kell. Ez

$$\mathbf{E}(\xi^2) = \sum_{k=0}^{20} k^2 \mathbf{P}(\xi = k) = \sum_{k=0}^{20} k^2 \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}}.$$

Ez még az előbbinél is bonyolultabb formula. Használjuk, hogy $k^2 = k(k-1) + k$, így a fenti összegben megjelenik a már meghatározott várható érték. A $(*)$ formulához hasonlóan

$$k(k-1)\binom{20}{k} = 20 \cdot 19 \cdot \binom{18}{k-2}.$$

A bal- és jobboldalon is azt számoljuk össze, hogy hányféleképp lehet egy 20 fős társaságból kiválasztani egy k fős bizottságot, annak a vezetőjét, és egy helyettest. Ezt a formulát felhasználva

$$\begin{aligned}\sum_{k=2}^{20} k(k-1)\binom{20}{k}\binom{30}{20-k} &= \sum_{k=2}^{20} 20 \cdot 19 \cdot \binom{18}{k-2}\binom{30}{20-k} \\ = 20 \cdot 19 \sum_{\ell=0}^{18} \binom{18}{\ell}\binom{30}{18-\ell} &= 20 \cdot 19 \binom{48}{18}.\end{aligned}$$

Itt használtuk, hogy $k(k-1) = 0$ ha $k = 0$ vagy $k = 1$, ezért elegendő 2-től összegezni, valamint a várható érték kiszámításánál is használt formulát. Most 18 piros és 30 fehér golyó közül veszünk visszatevés nélkül 18-at. Ezt visszahelyettesítve a második momentum formulájába

$$\begin{aligned}\mathbf{E}(\xi^2) &= \sum_{k=0}^{20} k^2 \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}} \\ &= \sum_{k=0}^{20} k(k-1) \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}} + \sum_{k=0}^{20} k \frac{\binom{20}{k} \cdot \binom{30}{20-k}}{\binom{50}{20}} \\ &= 20 \cdot 19 \frac{\binom{48}{18}}{\binom{50}{20}} + \mathbf{E}(\xi) = 20 \cdot 19 \frac{20 \cdot 19}{50 \cdot 49} + 8.\end{aligned}$$

Az utolsó egyenlőségnél felhasználtuk, hogy $20 \cdot 19\binom{50}{20} = 50 \cdot 49\binom{48}{18}$, amit már láttunk. Tehát a szórásnégyzet

$$\mathbf{D}^2(\xi) = \mathbf{E}(\xi^2) - (\mathbf{E}(\xi))^2 = \frac{(20 \cdot 19)^2}{50 \cdot 49} + 8 - 64 = \frac{144}{49},$$

és így $\mathbf{D}(\xi) = \frac{12}{7}$. $\square$

2. Megoldás. Megadunk egy másik megoldást, aminek az ötletét később is használni fogjuk. Jelölje megint ξ a kihúzott piros golyók számát. Ekkor persze ξ diszkrét véletlen változó, de most nem határozzuk meg az eloszlását. Ehelyett ξ-et felírjuk egyszerűbb változók összegeként. A módszer: *írjuk fel indikátorok összegeként!*

Legyen $i = 1, 2, \ldots, 20$ esetén

$$I_i = \begin{cases} 1, & \text{ha az } i\text{-ediknek húzott golyó piros,} \\ 0, & \text{különben.} \end{cases}$$

Ekkor $\xi = \sum_{i=1}^{20} I_i$. Egy indikátorváltozó mindig két értéket vehet fel, 0-t és 1-et. Ezért várható értéke a definíció szerint

$$\mathbf{E}(I_i) = 0 \cdot \mathbf{P}(I_i = 0) + 1 \cdot \mathbf{P}(I_i = 1) = \mathbf{P}(I_i = 1).$$

Tehát azt kell meghatároznunk, hogy mennyi a valószínűsége, hogy az i-edik golyó piros. Ez egyszerű, hiszen az i-edik helyre 50 golyó közül választhatunk, abból 20 piros, tehát tetszőleges i esetén

$$\mathbf{P}(I_i = 1) = \mathbf{P}(\, i\text{-edik golyó piros}\,) = \frac{20}{50} = \frac{2}{5}.$$

Tehát

$$\mathbf{E}(\xi) = \mathbf{E}\left(\sum_{i=1}^{20} I_i\right) = \sum_{i=1}^{20} \mathbf{E}(I_i) = 20 \cdot \frac{2}{5} = 8.$$

Így talán egyszerűbb volt a számolás, mint az előbb. Annyit használtunk, hogy *összeg várható értéke az a várható értékek összege*. Ez mindig teljesül, tetszőleges véletlen változók esetén.

A szórásnégyzet meghatározása is hasonlóan történik, de itt már többet kell számolni. A kovariancia tulajdonságai szerint

$$\begin{aligned} \mathbf{D}^2(\xi) = \mathbf{Cov}(\xi, \xi) = \mathbf{Cov}\left(\sum_{i=1}^{20} I_i, \sum_{i=1}^{20} I_i\right) \\ = \sum_{i=1}^{20} \sum_{j=1}^{20} \mathbf{Cov}(I_i, I_j) = \sum_{i=1}^{20} \mathbf{D}^2(I_i) + 2 \sum_{1 \le i < j \le 20} \mathbf{Cov}(I_i, I_j). \end{aligned}$$

Vegyük észre, hogy I_i eloszlása nem függ i-től, azaz annak a valószínűsége, hogy elsőre pirosat húzunk, pontosan ugyanaz, mint annak a valószínűsége, hogy 7.-re. Ezt talán úgy látjuk legkönnyebben, ha a hetedik húzásra nem úgy gondolunk, ami előtt már volt hat másik, hanem mint a húsz húzás közül az egyikre. Azaz ekkor semmi más nem érdekel minket, csak a hetedik húzás. Ugyanígy látjuk, hogy (I_1, I_2) együttes eloszlása ugyanaz, mint (I_i, I_j) együttes eloszlása tetszőleges $i \neq j$ esetén. Azaz, annak a valószínűsége, hogy az első két húzás piros, pontosan ugyanaz, mint annak a valószínűsége, hogy a hetedik és tizenharmadik piros. Ezek szerint $\mathbf{D}^2(I_i) = \mathbf{D}^2(I_1)$ és $\mathbf{Cov}(I_i, I_j) = \mathbf{Cov}(I_1, I_2)$. Tehát, folytatva a szórásnégyzet kiszámolását

$$\mathbf{D}^2(\xi) = 20\mathbf{D}^2(I_1) + 20 \cdot 19\mathbf{Cov}(I_1, I_2).$$

A definíció szerint

$$\mathbf{D}^2(I_1) = \mathbf{E}(I_1^2) - (\mathbf{E}(I_1))^2 = \frac{2}{5} - \left(\frac{2}{5}\right)^2 = \frac{2}{5} \cdot \frac{3}{5},$$

és

$$\begin{aligned}\mathbf{Cov}(I_1, I_2) &= \mathbf{E}(I_1 I_2) - \mathbf{E}(I_1) \cdot \mathbf{E}(I_2) \\ &= \mathbf{P}(I_1 = 1, I_2 = 1) - \left(\frac{2}{5}\right)^2 = \frac{20 \cdot 19}{50 \cdot 49} - \left(\frac{2}{5}\right)^2,\end{aligned}$$

hiszen annak a valószínűsége, hogy az első és második golyó is piros

$$\frac{20 \cdot 19}{50 \cdot 49}.$$

Összegezve,

$$\mathbf{D}^2(\xi) = 20\frac{2}{5} \cdot \frac{3}{5} + 20 \cdot 19 \cdot \left(\frac{20 \cdot 19}{50 \cdot 49} - \left(\frac{2}{5}\right)^2\right) = \frac{144}{49},$$

és így $\mathbf{D}(\xi) = \frac{12}{7}$.

2.1.17. Egy urnában van 20 piros és 30 fehér golyó. Húzzunk ki 20 golyót visszatevéssel. Számoljuk ki a kihúzott piros golyók számának várható értékét és szórásnégyzetét!

Megoldás. Ezt a feladatot is megoldhatjuk mindkét fenti módszerrel. Mivel visszatevéssel húzunk, ezért minden húzás előtt pontosan ugyanaz van az urnában, azaz igazából egy kísérletet ismétlek 20-szor, ahol a siker valószínűsége 2/5, hiszen a piros golyókat figyelem, annak a kihúzása lesz most a siker. Tehát, ha η jelöli a kihúzott pirosak számát, akkor η lehetséges értékei $0, 1, \ldots, 20$, és

$$\mathbf{P}(\eta = k) = \binom{20}{k} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{20-k}.$$

Vagyis η binomiális eloszlású 20 és $2/5$ paraméterekkel. Binomiális eloszlás várható értékét, szórását már számoltuk,

$$\mathbf{E}(\eta) = 20 \cdot \frac{2}{5} = 8, \quad \mathbf{D}^2(\eta) = 20 \cdot \frac{2}{5} \cdot \frac{3}{5} = \frac{24}{5}.$$

Felbonthatom η-t is indikátorok összegére. Ha $J_i = 1$, ha az i-edik golyó piros, 0 különben, akkor

$$\eta = J_1 + \ldots + J_{20}.$$

Eszerint a várható érték és a szórásnégyzet is könnyen kiszámolható. Most a változók függetlenek, ezért az összeg szórásnégyzete megegyezik a szórásnégyzetek összegével.

Az előző feladattal összevetve nagyon fontos különbség, hogy itt $J_1, \ldots, J_{20}$ *független* véletlen változók, hiszen az első 5 húzás semmilyen módon nem befolyásolja a hatodikat. Ezzel ellentétben, ha visszatevés nélkül húzunk, akkor $I_1, \ldots, I_{20}$ nem függetlenek (persze a kovarianciájuk sem 0), mert ha az első 5 húzás piros, akkor a hatodik húzás előtt az urnában már csak 15 piros van és 30 fehér.

Vegyük észre, hogy a várható érték mindkét modell esetén ugyanaz, míg $\mathbf{D}^2(\xi) < \mathbf{D}^2(\eta)$, azaz visszatevés nélkül a szórásnégyzet jóval kisebb. Gondoljuk végig, mért természetes ez! □

2.1.18. Egy urnában egy piros és egy fehér golyó van. Visszatevéssel húzunk az urnából, minden húzás után még egy piros golyót teszünk az urnába. Jelölje ξ annak a kísérletnek a sorszámát, amikor először húztunk fehéret. Adjuk meg ξ eloszlását, várható értékét!

2.1.19. Jelölje S_n n elem véletlen permutációja során a fixpontok számát. Határozzuk meg S_n várható értékét és szórását!

2.1.20. Egy halastóban N hal van. Kihalászunk M halat, megjelöljük őket, és visszaeresztjük a tóba. Bizonyos idő elteltével, miután jól elkeveredtek, kihalászunk n-et. Ezek között legyen a megjelöltek száma ξ. A teljes halállomány N meghatározására az $Mn/(\xi + 1)$ becslést használjuk. Számítsuk ki ennek a várható értékét és szórását! Miért nem a logikusabb Mn/ξ becslést használjuk?

2.1.21 (∗). Földobunk n-szer egy szabályos pénzérmét. Határozzuk meg az F-I, I-F váltások számának eloszlását!

2.1.22 (∗). Egy városban 200 taxi közlekedik. Telefonon taxit rendelünk, és ha van szabad taxi, akkor a központ a legközelebbit hozzánk küldi. Feltesszük, hogy a taxik egymástól függetlenül, egyenletes eloszlás szerint helyezkednek el a városban, és mindegyik egymástól függetlenül $2/3$ valószínűséggel foglalt. Továbbá egy taxi helyzete a városon belül független attól, hogy foglalt-e vagy sem. Mennyi annak a valószínűsége, hogy a legközelebbi szabad taxi 1 km-es körzetünkben legyen (mely nem nyúlik ki a városból), feltéve, hogy van szabad taxi? A város területe $28,26\,\mathrm{km}^2$. ([1])

2.2. Nevezetes diszkrét eloszlások

2.2.1. A sztochasztika alapjai kurzus 400 hallgatójának mindegyike bemegy az első előadásra. Ezt követően minden hallgató minden további előadás előtt feldob egy szabályos pénzérmét. Ha fejet kap, akkor bemegy a következő előadásra, ha írást akkor nem, és utána már egyetlen előadásra sem megy be. Véletlen Vince a 400 hallgató egyike. Mennyi a valószínűsége, hogy Vince az összes előadásra bemegy? Várhatóan hány előadáson vesz részt? Mennyi a valószínűsége, hogy a 12. (utolsó) előadáson lesz hallgató? Várhatóan hány hallgató lesz a 2., 3., utolsó előadáson? Mennyi a valószínűsége, hogy az utolsó előadáson pontosan 2 hallgató van?

Megoldás. Vince egy szabályos érmét dobál legfeljebb 12-szer. Jelölje ξ annak a dobásnak a sorszámát, amikor először írást dob, ha csupa fejet, akkor legyen $\xi = 12$. Ekkor Vince pontosan ξ db előadáson vesz részt. (Ha pl. $\xi = 1$, akkor elsőre fejet dob, így a második előadásra már nem megy be.) Ha $k \leq 11$, akkor a $\xi = k$ esemény pontosan azt jelenti, hogy az első $k-1$ dobás fej, a k-adik írás, így

$$\mathbf{P}(\xi = k) = \frac{1}{2^k}, \quad k = 1, 2, \ldots, 11.$$

A $\xi = 12$ akkor lehet az, hogy mind a 12 dobás fej vagy az is lehet, hogy az első 11 fej, és a 12. írás. Ez azt jelenti, hogy az első 11 fej, így

$$\mathbf{P}(\xi = 12) = \frac{1}{2^{12}} + \frac{1}{2^{12}} = \frac{1}{2^{11}}.$$

Az, hogy Vince az összes előadásra bemegy pontosan azt jelenti, hogy $\xi = 12$. Ez éppen azt jelenti, hog az első 11 dobás mindegyike fej, ennek a valószínűsége $(1/2)^{11}$, tehát

$$\mathbf{P}(\text{Vince az összes előadásra bemegy}) = \mathbf{P}(\xi = 12) = \frac{1}{2^{11}}.$$

A látogatott előadások számának várható értéke $\mathbf{E}(\xi)$, ami definíció szerint

$$\mathbf{E}(\xi) = \sum_{k=1}^{11} \frac{k}{2^k} + \frac{12}{2^{11}} = 2 - \frac{1}{2^{11}} \approx 2.$$

Mind a 400 hallgató úgy jár el mint Vince. Legyenek $\xi_1, \xi_2, \dots, \xi_{400}$ az egyes hallgatókhoz tartózó előadásszámok, azaz független, azonos eloszlású véletlen változók, melyek közös eloszlása az, mint ξ eloszlása. Pontosan akkor van hallgató az utolsó előadáson, ha van olyan $1 \le k \le 400$, melyre $\xi_k = 12$, vagyis, ha $\max_{1\le k\le 400} \xi_k = 12$. Ennek valószínűsége, a függetlenséget és a korábbiakat felhasználva

$$\begin{aligned}\mathbf{P}(\max_{1\le k\le 400} \xi_k \ge 12) &= 1 - \mathbf{P}(\max_{1\le k\le 400} \xi_k < 12)\\ &= 1 - \mathbf{P}(\xi_1 < 12, \xi_2 < 12, \dots, \xi_{400} < 12)\\ &= 1 - (\mathbf{P}(\xi_1 < 12))^{400}\\ &= 1 - \left(1 - \frac{1}{2^{11}}\right)^{400} = 0{,}18.\end{aligned}$$

Tehát 0,18 annak a valószínűsége, hogy az utolsó előadáson lesz hallgató.

Jelölje η_2 a 2. előadáson levő hallgatók számát, η_3 a 3. előadáson levő hallgatók számát, ... A 2. előadáson azok vesznek részt, akik elsőre fejet dobtak, tehát η_2 binomiális eloszlású $n = 400$ és $p_2 = 1/2$ paraméterekkel. Így

$$\mathbf{E}(\eta_2) = 400 \cdot \frac{1}{2} = 200.$$

Hasonlóan, a 3. előadáson résztvevők száma binomiális $n = 400$ és $p_3 = 1/4$ paraméterekkel, hiszen itt azok vannak, akik kétszer fejet dobtak. Így

$$\mathbf{E}(\eta_3) = 400 \cdot \frac{1}{4} = 100.$$

Míg az utolsó előadáson résztvevők száma binomiális $n = 400$, $p_{12} = 1/2^{11}$ paraméterrel, így

$$\mathbf{E}(\eta_{12}) = 400 \cdot \frac{1}{2^{11}} = 0{,}2.$$

Tehát az utolsó előadáson várhatóan 0,2 hallgató lesz. Innen persze másképpen is meghatározható az a valószínűség, hogy az utolsó előadáson lesz hallgató. Ez éppen azt jelenti, hogy $\eta_{12} > 0$, azaz

$$\mathbf{P}(\eta_{12} > 0) = 1 - \mathbf{P}(\eta_{12} = 0) = 1 - \left(1 - \frac{1}{2^{11}}\right)^{400} = 0{,}18,$$

ami ugyanaz, mint amit korábban kaptunk.

Az, hogy az utolsó előadáson pontosan két hallgató van, azt jelenti, hogy $\eta_{12} = 2$. Mivel $n = 400$ nagy és $p_{12} = 2^{-11}$ kicsi, ezért a binomiálist közelíthetjük Poisson-eloszlással, aminek a paraméter $\lambda = 400 \cdot 2^{-11} \approx 0{,}195$. A pontos valószínűség

$$\mathbf{P}(\eta_{12} = 2) = \binom{400}{2} p_{12}^2 (1 - p_{12})^{398} = 0{,}01566,$$

míg a közelítő valószínűség Poisson-közelítéssel

$$\frac{\lambda^2}{2} e^{-\lambda} = 0{,}01569.$$

□

2.2.2. Húsvétra dobta piacra a Kinder Meglepetés új, matematikusfigurákat tartalmazó Kinder tojásait. Átlagosan minden 4-edik tojás rejt matematikusfigurát. Aladár 10 Kinder tojást kapott. Adjuk meg annak a valószínűségét, hogy Aladár matematikusfigurának örülhet! Adjuk meg Aladár matematikusfigurái számának eloszlását, várható értékét!

2.2.3. Három, külsőre egyforma érmével a fejdobás valószínűsége $1/4, 1/2$, és $3/4$. Véletlenszerűen választunk egy érmét, és azzal az első fejig dobunk. Legyen ξ a dobások száma. Adjuk meg ξ eloszlását és várható értékét!

2.2.4. Egy könyvben az egyes oldalakon a sajtóhubák száma egymástól független, Poisson(2) eloszlást követ. Határozzuk meg annak a valószínűségét, hogy a 30. és 31. oldalon sincs hiba! Adjuk meg az ezeken az oldalakon található sajtóhubák várható értékét és szórását!

Megoldás. Először is felidézzük a Poisson-eloszlás definícióját (rosszabb esetben megnézzük a képletgyűjteményben). Eszerint ξ Poisson-eloszlású $\lambda = 2$ paraméterrel, ha

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

Jelölje ξ a 30. oldalon, η a 31. oldalon levő hibák számát. A feladat szerint ezek független, Poisson-eloszlásúak 2 paraméterrel. Az, hogy egyik oldalon sincsen hiba, azt jelenti, hogy $\xi = 0$ és $\eta = 0$. Ennek a valószínűsége, a függetlenség miatt

$$\mathbf{P}(\xi = 0, \eta = 0) = \mathbf{P}(\xi = 0) \cdot \mathbf{P}(\eta = 0) = e^{-2} \cdot e^{-2} = e^{-4}.$$

Tudjuk, hogy λ-paraméterű Poisson várható értéke és szórásnégyzete is λ, ezért

$$\mathbf{E}(\xi + \eta) = \mathbf{E}(\xi) + \mathbf{E}(\eta) = 4,$$

és mivel független változók összegének szórásnégyzete egyenlő a szórásnégyzetek összegével, ezért

$$\mathbf{D}^2(\xi+\eta) = \mathbf{D}^2(\xi) + \mathbf{D}^2(\eta) = 4,$$

azaz $\mathbf{D}(\xi+\eta) = 2$. □

2.2.5. Egy augusztusi éjszakán megfigyelhető csillaghullások száma Poisson–eloszlást követ. Annak a valószínűsége, hogy egy éjszaka egyetlen hullócsillagot sem látunk 0,1. Várhatóan hány hullócsillag figyelhető meg egy éjszaka?

Megoldás. Itt is Poisson-eloszlás van. Most ki kell találnunk az adatokból a paramétert, mert nincs expliciten megadva, mint az előbb. Legyen ξ a megfigyelt hullócsillagok száma egy este. Ekkor $\xi \sim$Poisson(λ). Az, hogy nem látunk hullócsillagot azt jelenti, hogy $\xi = 0$. Azaz, azt tudjuk, hogy

$$\mathbf{P}(\xi = 0) = 0,1.$$

Na de a formula szerint ez éppen $\lambda^0/0! \cdot e^{-\lambda} = e^{-\lambda}$. Tehát $e^{-\lambda} = 0,1$ ahonnan $\lambda = -\log 0,1 = \log 10$ (itt a $\log = \ln$ a természetes alapú logaritmus). A Poisson paramétere tetszőleges pozitív szám lehet, nem kell, hogy egész legyen! Poisson várható értéke a paramétere, azaz

$$\mathbf{E}(\xi) = \lambda = \log 10.$$

□

2.2.6. Egy biztosítótársaság felmérte, hogy egy év során egy családi ház 0,0002 valószínűséggel gyullad ki. Mennyi a valószínűsége, hogy 2008-ban egy faluban, ahol 15000 ház van, négynél kevesebb tűz üt ki? (Közelítsünk Poisson–eloszlással!)

Megoldás. Itt is Poisson-eloszlással kell számolni. A pontos eloszlás binomiális $n = 15000$ és $p = 0{,}0002$ paraméterekkel, de a binomiális jól közelíthető Poisson-eloszlással ha p kicsi és n nagy.

Először meg kell határoznunk a λ paramétert. Mivel a tűz valószínűsége egy háznál 0,0002, és a faluban 15000 ház van, ezért az összes tűz várható értéke

$$15000 \cdot 0{,}0002 = 3.$$

Tehát egy olyan Poisson-eloszlás kell nekünk, aminek a várható értéke 3. A várható érték éppen a paraméter, tehát $\lambda = 3$. Ezek szerint annak a valószínűsége, hogy négynél kevesebb tűz üt ki az

$$\begin{aligned}\mathbf{P}(\xi < 4) &= \mathbf{P}(\xi = 0) + \mathbf{P}(\xi = 1) + \mathbf{P}(\xi = 2) + \mathbf{P}(\xi = 3)\\ &= \sum_{k=0}^{3} \frac{3^k}{k!} e^{-3} = e^{-3}\left(1 + 3 + \frac{9}{2} + \frac{27}{6}\right)\\ &= e^{-3} 13.\end{aligned}$$

□

2.2.7. Egy szövet 100 méterében átlagosan 5 hiba van. Három méteres darabokra vágnak 300 m hosszú szövetet. Várhatóan hány hibátlan darab lesz?

2.2.8. Mikulás minden csomagba egymástól függetlenül véletlen számú szaloncukrot tesz azonos $\lambda > 0$ paraméterű Poisson-eloszlás szerint. Átlagosan 10 csomag között egy van, amelyikben nincsen szaloncukor. Várhatóan hány szaloncukor lesz Máté csomagjában? Máté unokatestvéreivel, Bencével és Lucával együtt kapja meg a csomagját a Mikulástól. Mennyi a valószínűsége, hogy mindhármuk csomagjában lesz szaloncukor?

2.2.9. A Bajnokok Ligája döntőiben az összes gólok száma Poisson-eloszlást követ. A döntők 20%-án nem esik gól. Várhatóan hány gólt láthatunk egy döntőn?

2.2.10. Minden este megiszom egy limonádét, a poharat pedig mosatlanul kinn hagyom az asztalon. A pohár alján mindig marad egy kis cukor, ezért reggelente néhány hangyát találok benne. A pohárban levő hangyák számai az egyes napokon egymástól függetlenek, 3 paraméterű Poisson eloszlást követnek. Határozzuk meg annak a valószínűségét, hogy 5 egymást követő napon mindig találok hangyát a poharamban! Mekkora a valószínűsége, hogy az 5 nap alatt valamikor találok hangyát? Mennyi az 5 nap alatt talált hangyák számának várható értéke és szórása?

2.2.11 (∗). Valamely növényfajta magjaiból álló mintában a hibás magok száma λ paraméterű Poisson–eloszlású véletlen változó. Minden mintát 3 technikus vizsgál meg egymás után, hogy eltávolítsák a hibás magokat. Az i-edik technikus $p_i < 1$ valószínűséggel veszi észre a hibás magokat; döntései az egyes magokra nézve függetlenek, és az egyes technikusok is egymástól függetlenül döntenek. Határozzuk meg az el nem távolított hibás magok eloszlását! ([1])

Megoldás. Azt nézzük, hogy egy technikus után mi a helyzet. Legyen $p = p_1$. Jelölje ξ az eredeti mintában a hibás magok számát, és η a technikus ellenőrzése után megmaradt hibás magok számát. Tudjuk, hogy ξ Poisson-eloszlású λ paraméterrel, azaz lehetséges értékei $0, 1, \ldots$. Világos, hogy η-nak is ezek a lehetséges értékei.

Így nehéz a feladat. Az eredeti hibás magok száma is véletlen. Mi lenne, ha tudnánk a hibás magok számát. Tegyük fel, hogy az eredeti hibás magok száma k, azaz $\xi = k$. Jön az első technikus, minden egyes hibás magot p valószínűséggel vesz észre, azaz $1 - p$ valószínűséggel nem veszi észre. Ez minden hibás magra egymástól független, ezért a bennemaradt hibás magok száma binomiális eloszlású k és $1-p$ paraméterekkel. Tehát, az $\xi = k$ feltétel mellett (ez egy kutya közönséges feltételes valószínűség)

$$\mathbf{P}(\eta = \ell | \xi = k) = \binom{k}{\ell}(1-p)^\ell p^{k-\ell}, \quad \ell = 0, 1, \ldots, k.$$

Akkor alkalmazzuk a teljes valószínűség tételét az $\{\xi = 0\}$, $\{\xi = 1\}$, $\ldots$, teljes eseményrendszerrel. Eszerint

$$\mathbf{P}(\eta = \ell) = \sum_{k=0}^{\infty} \mathbf{P}(\eta = \ell | \xi = k)\mathbf{P}(\xi = k).$$

Írjuk be a Poisson-eloszlás formuláját, és a fent kapott formulát. Kicsit számolunk (persze ha $\eta = \ell$ akkor ξ is legalább ℓ, azaz $\mathbf{P}(\eta = \ell | \xi = k) = 0$, ha $k < \ell$)

$$\begin{aligned}
\mathbf{P}(\eta = \ell) &= \sum_{k=0}^{\infty} \mathbf{P}(\eta = \ell | \xi = k)\mathbf{P}(\xi = k) \\
&= \sum_{k=\ell}^{\infty} \binom{k}{\ell}(1-p)^\ell p^{k-\ell} \frac{\lambda^k}{k!} e^{-\lambda} \\
&= (1-p)^\ell \lambda^\ell e^{-\lambda} \frac{1}{\ell!} \sum_{k=\ell}^{\infty} \frac{1}{(k-\ell)!} (\lambda p)^{k-\ell} \\
&= [\lambda(1-p)]^\ell e^{-\lambda} \frac{1}{\ell!} e^{\lambda p} \\
&= \frac{[\lambda(1-p)]^\ell}{\ell!} e^{-(1-p)\lambda}.
\end{aligned}$$

Az utolsó előtti egyenlőségnél használtuk, hogy az exponenciális függvény Taylor-sora jelent meg ($e^x = \sum_{k=0}^{\infty} x^k/k!$). Látjuk, hogy ez éppen egy $(1-p)\lambda$ paraméterű Poisson-eloszlás.

Tehát az első technikus után maradt hibás magok száma is Poisson. Azaz a második technikus is Poisson darab hibás magot kap, csak a paraméter $(1-p_1)\lambda$. A fentiek szerint a második technikus után maradt hibás magok száma is Poisson, a paraméter pedig $(1-p_1)(1-p_2)\lambda$, végül a harmadik után maradt hibás magok száma is Poisson-eloszlású, $(1-p_1)(1-p_2)(1-p_3)\lambda$ paraméterrel. $\square$

2.2.12 (∗). Mutassuk meg, hogy két független Poisson–eloszlású véletlen változó összege is Poisson–eloszlású!

Megoldás. Itt először is a Poisson-eloszlás definícióját kell tudni. Az ξ Poisson-eloszlású $\lambda > 0$ paraméterrel, ha nemnegatív egész értékű, és

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

Legyenek tehát ξ, η független Poisson-eloszlású véletlen változók $\lambda > 0$ és $\mu > 0$ paraméterekkel. Ekkor a $Z = \xi + \eta$ változó lehetséges értékei is $0, 1, \ldots$, és ha $Z = n$ akkor $\xi = k$ és $\eta = n - k$ valamilyen k-ra, ahol persze $k = 0, 1, \ldots, n$. Vagyis

$$\mathbf{P}(Z = n) = \mathbf{P}(\xi + \eta = n) = \sum_{k=0}^{n} \mathbf{P}(\xi = k, \eta = n - k).$$

Eddig nem használtunk igazából semmit. Mivel ξ, η függetlenek, ezért az utóbbi összeg tagjait szorzattá alakíthatjuk, azaz

$$= \sum_{k=0}^{n} \mathbf{P}(\xi = k) \cdot \mathbf{P}(\eta = n - k).$$

Most használjuk, hogy ξ és η is Poisson, és egy kicsit számolunk. Ezért

$$\begin{aligned}
&= \sum_{k=0}^{n} \frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{\mu^{n-k}}{(n-k)!}e^{-\mu} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \mu^{n-k} \\
&= \frac{(\lambda+\mu)^n}{n!} e^{-(\lambda+\mu)}.
\end{aligned}$$

Itt a binomiális tételt használtuk, és a binomiális együttható alakját, semmi extra. Tehát azt kaptuk, hogy tetszőleges $n = 0, 1, \ldots$ esetén

$$\mathbf{P}(Z = n) = \frac{(\lambda+\mu)^n}{n!} e^{-(\lambda+\mu)},$$

na de ez éppen a $(\lambda+\mu)$ paraméterű Poisson-eloszlás definíciója. Kész vagyunk. $\square$

2.3. Folytonos véletlen változók

2.3.1. Egy permetező szakaszoló szelep napokban mért élettartamának sűrűségfüggvénye

$$f(x)=\begin{cases}\frac{3000}{x^4}, & \text{ha } x\geq 10,\\ 0, & \text{különben.}\end{cases}$$

Mennyi a valószínűsége, hogy a permetező szakaszoló szelep 20 napot túlél? Határozzuk meg a szelep élettartamának eloszlásfüggvényét, várható értékét és szórását!

Megoldás. Jelölje ξ egy permetező szakaszoló szelep napokban mért élettartamát. Világos, hogy ez egy véletlen mennyiség, véletlen változó. A feladat szerint ez egy folytonos véletlen változó f sűrűségfüggvénnyel, ami definíció szerint azt jelenti, hogy

$$\mathbf{P}(\xi<x)=\int_{-\infty}^{x} f(y)\mathrm{d}y.$$

Az, hogy a permetező szakaszoló szelep 20 napot túlél pontosan azt jelenti, hogy $\xi\geq 20$. Ennek a valószínűsége

$$\begin{aligned}\mathbf{P}(\xi\geq 20)&=1-\mathbf{P}(\xi<20)=1-\int_{-\infty}^{20} f(y)\mathrm{d}y\\ &=1-\int_{10}^{20} 3000\,x^{-4}\mathrm{d}x=1-3000\,\left[-3^{-1}x^{-3}\right]_{x=10}^{x=20}\\ &=1+1000\left(\frac{1}{8000}-\frac{1}{1000}\right)=\frac{1}{8}.\end{aligned}$$

Itt annyit használtunk, hogy a sűrűségfüggvény értéke 0, ha $x<10$, és ezért elég a $[10,20]$ intervallumon integrálni, valamint hatványfüggvény integrálját számoltuk ki. Hát persze, ha $\alpha\neq -1$ akkor

$$\int x^{\alpha}\mathrm{d}x=\frac{x^{\alpha+1}}{\alpha+1}.$$

Tehát $\mathbf{P}(\xi\geq 20)=\frac{1}{8}$.

Az eloszlásfüggvény meghatározásához a definíciót használjuk: $F(x) = \mathbf{P}(\xi < x) = \int_{-\infty}^{x} f(y)\mathrm{d}y$. Ha $x < 10$, akkor a konstans 0 függvényt integráljuk, ha $x > 10$ akkor pedig a fenti számoláshoz hasonlóan

$$F(x) = \int_{-\infty}^{x} f(y)\mathrm{d}y = \int_{10}^{x} 3000\, y^{-4}\mathrm{d}y = 3000 \left[-\frac{y^{-3}}{3}\right]_{y=10}^{y=x} = 1 - \frac{1000}{x^3}.$$

Tehát

$$F(x) = \int_{-\infty}^{x} f(y)\mathrm{d}y = \begin{cases} 0, & \text{ha } x \leq 10, \\ 1 - \frac{1000}{x^3}, & \text{ha } x \geq 10. \end{cases}$$

Mivel folytonos az eloszlás, az eloszlásfüggvény is folytonos, ezért mindegy, hogy $x = 10$-re melyik ágon definiáljuk az értéket. Azt is látjuk, hogy az eloszlásfüggvény 10-ig konstans 0, utána szigorúan monoton nő, és végtelenben 1-be tart, mint minden eloszlásfüggvény.

A várható értéket is definíció alapján számoljuk. Eszerint

$$\begin{aligned}\mathbf{E}(\xi) &= \int_{-\infty}^{\infty} x f(x)\mathrm{d}x = \int_{10}^{\infty} x 3000\, x^{-4}\mathrm{d}x \\ &= 3000 \int_{10}^{\infty} x^{-3}\mathrm{d}x = 3000 \left[-\frac{x^{-2}}{2}\right]_{x=10}^{\infty} \\ &= 3000 \left[\frac{1}{200} - \lim_{x\to\infty} \frac{x^{-2}}{2}\right] = 15.\end{aligned}$$

Itt egy végtelenben improprius integrált kellett kiszámolni, ezért ha nagyon precízen akarjuk írni a dolgot akkor a végtelen felső határ helyett $\lim_{N\to\infty} \int_{10}^{N}$ kellene, azaz

$$\int_{10}^{\infty} h(x)\mathrm{d}x = \lim_{N\to\infty} \int_{10}^{N} h(x)\mathrm{d}x.$$

Ezt később is elhagyjuk, és nem akarunk ebben az értelemben nagyon precízek lenni.

A szórásnégyzet kiszámításához először a második momentumot kell meghatározni. Ez

$$\mathbf{E}(\xi^2) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = 3000 \int_{10}^{\infty} x^{-2}\mathrm{d}x = 3000 \left[-x^{-1}\right]_{x=10}^{x=\infty} = 300.$$

És így a szórásnégyzet

$$\mathbf{D}^2(\xi) = \mathbf{E}(\xi^2) - (\mathbf{E}(\xi))^2 = 300 - (15)^2 = 75,$$

azaz a szórás $\mathbf{D}(\xi) = \sqrt{75}$. $\square$

2.3.2. Legyen a ξ véletlen változó sűrűségfüggvénye $f(x) = c/x^2$, ha $x > 1$.

(a) Határozzuk meg c értékét!

(b) Adjuk meg ξ várható értékét (ha létezik)!

(c) Mennyi $\mathbf{P}(\xi > 4)$?

(d) Legyen $Y = 1/\xi$. Adjuk meg Y eloszlás–, és sűrűségfüggvényét!

Megoldás. (a) Ahhoz, hogy egy f függvény sűrűségfüggvény legyen az kell, hogy ő nemnegatív, és integrálja az egyenesen 1. Mivel a sűrűség értéke cx^{-2} ha $x > 1$, különben pedig 0, ez azt jelenti, hogy $c \geq 0$ és

$$1 = \int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{1}^{\infty} cx^{-2}\mathrm{d}x = c\left[-x^{-1}\right]_{x=1}^{\infty} = c.$$

Tehát $c = 1$.

(b) A várható érték definíció szerint

$$\mathbf{E}(\xi) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_{1}^{\infty} x\frac{1}{x^2}\mathrm{d}x = \int_{1}^{\infty} \frac{1}{x}\mathrm{d}x = \infty,$$

azaz a várható érték nem létezik. A várható érték definíciójában feltettük, hogy az $\int_{-\infty}^{\infty} |x|f(x)\mathrm{d}x$ improprius integrál létezik és véges. Most ez nem teljesül. Látjuk, hogy vannak olyan véletlen változók, melyeknek nem véges a várható értéke.

(c) A sűrűségfüggvény tulajdonságai szerint $\mathbf{P}(\xi \in A) = \int_A f(x)\mathrm{d}x$, azaz egy A halmazba esés valószínűsége egyenlő a sűrűségfüggvény A halmazon vett integráljával. Eszerint

$$\mathbf{P}(\xi > 4) = \int_{4}^{\infty} f(x)\mathrm{d}x = \int_{4}^{\infty} x^{-2}\mathrm{d}x = [-x^{-1}]_{x=4}^{\infty} = \frac{1}{4}.$$

Az eloszlásfüggvényt is meghatározzuk. A definíció alapján

$$F(x) = \mathbf{P}(\xi < x) = \int_{-\infty}^{x} f(y)\mathrm{d}y = \begin{cases} 0, & \text{ha } x \leq 1, \\ \int_1^x y^{-1}\mathrm{d}y = 1 - \frac{1}{x}, & \text{ha } x \geq 1. \end{cases}$$

Ez az 1 paraméterű Pareto-eloszlás.

(d) Legyen $\eta = 1/\xi$. Ezzel a definícióval nincs baj, hiszen $\xi \geq 1$. Először meghatározzuk η eloszlásfüggvényét. Mivel $\xi \geq 1$ ezért $\eta = 1/\xi \in [0, 1]$. Ezek szerint

$$G(y) = \mathbf{P}(\eta < y) = \begin{cases} 0, & \text{ha } y \leq 0, \\ 1, & \text{ha } y \geq 1. \end{cases}$$

Az érdekes eset amikor $y \in [0,1]$. Ekkkor

$$\mathbf{P}(\eta < y) = \mathbf{P}(\xi^{-1} < y) = \mathbf{P}(\xi < y^{-1}) = 1 - \mathbf{P}(\xi < y^{-1}) = 1 - F(y^{-1}) = y.$$

Itt felhasználtuk, hogy ξ folytonos véletlen változó, ezért eloszlásfüggvénye is folytonos, valamint beírtuk a (c) részben meghatározott eloszlásfüggvény pontos alakját. Ezek szerint η eloszlásfüggvénye

$$G(y) = \mathbf{P}(\eta < y) = \begin{cases} 0, & \text{ha } y \le 0, \\ y, & \text{ha } y \in [0,1], \\ 1, & \text{ha } y \ge 1. \end{cases}$$

A sűrűségfüggvény az eloszlásfüggvény deriváltja (majdnem mindenütt). Az eloszlásfüggvény $G(y)$ szép differenciálható függvény, kivéve a 0 és az 1 pontokban. Tehát

$$g(y) = G'(y) = \begin{cases} 1, & \text{ha } y \in (0,1), \\ 0, & \text{különben}. \end{cases}$$

Ez az egyenletes eloszlás a $(0,1)$ intervallumon. □

2.3.3. A $(0,1)$ intervallumon találomra kijelölünk három pontot. Határozzuk meg a középső nullától vett távolságának eloszlás- és sűrűségfüggvényét, várható értékét és szórását!

2.3.4. A $[0,1]$ intervallumon kijelölünk három pontot egymástól függetlenül, egyenletes eloszlás szerint. Határozzuk meg a legkisebb eloszlásfüggvényét és várható értékét!

2.3.5. Válasszunk az egységnégyzetben egy pontot véletlenszerűen. Legyen ξ a pontnak a négyzet határától vett távolsága. Adjuk meg ξ eloszlásfüggvényét, várható értékét, szórását!

2.3.6. Előrejelzések alapján a forint/euró árfolyam 2021. 07. 31.-i eloszlásának sűrűségfüggvénye

$$f(x) = \frac{(360-x)^2}{9000}, \quad x \in (330, 360).$$

Adjuk meg annak a valószínűségét, hogy egy euró 350 Ft-nál többe kerül. Adjuk meg a várható értéket!

2.3.7. A labdarúgó világbajnokságon az első gólig eltelt percben mért játékidő sűrűségfüggvénye $f(x) = 4050 \cdot (45+x)^{-3}$, $x > 0$. Mennyi a valószínűsége, hogy az első félidőben gól esik? Mennyi a valószínűsége, hogy az első mérkőzésen nem esik gól? Várhatóan hány percet kell várni az első gólig? (Egy mérkőzés 90 percig tart, 45 perc egy félidő.)

2.3.8. A Bergengóc Élettani Kutatási Alap kifejlesztett egy egydózisú vakcinát a Covid-$\pi^2/6$ vírus ellen. A BÉKA vakcina a beadást követően azonnali védelmet biztosít egy véletlen ξ ideig, ahol ξ hónapokban mért sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{x-5}{8}, & x \in [5, 9], \\ 0, & \text{különben.} \end{cases}$$

Mennyi a valószínűsége, hogy a BÉKA vakcina fél év után még hatásos? Adjuk meg ξ védettségi idő várható értékét és szórását!

2.3.9. Válasszunk két számot egymástól függetlenül az egyenletességi hipotézis szerint a $(-1, 1)$ intervallumból! Adjuk meg a két szám maximumának eloszlásfüggvényét! Számoljuk ki a várható értéket és a szórást is!

2.4. Nevezetes folytonos eloszlások

2.4.1. Egy villanykörte élettartama exponenciális eloszlású, átlagosan 2 évig működik. Mennyi a valószínűsége, hogy legalább egy évig fog működni egy új villanykörte? Mennyi a valószínűsége, hogy legalább még egy évig fog működni egy már fél éve működő? Mennyi időt él meg a villanykörték 90%-a?

Megoldás. Exponenciális eloszlásunk van, először meg kell határozni a paramétert.

Legyen egy villanykörte évben mért élettartama ξ. Ekkor a feladat szerint ξ exponenciális eloszlású és várható értéke 2. Tudjuk (vagy megnézzük a képletgyűjteményben), hogy az exponenciális várható értéke a paraméter reciproka. Azaz, ha $\xi \sim\text{Exp}(\lambda)$, akkor $\mathbf{E}(\xi) = 1/\lambda$. Az, hogy egy villanykörte átlagosan 2 évig működik, éppen azt jelenti, hogy az élettartam várható értéke 2. Azaz $2 = 1/\lambda$, ahonnan kapjuk, hogy $\lambda = 1/2$.

Az, hogy egy új villanykörte legalább egy évig jó, azt jelenti, hogy az élettartam nagyobb, mint 1. Tehát a kérdés, hogy mennyi $\mathbf{P}(\xi \geq 1)$ (mivel folytonos eloszlásunk van, mindegy, hogy ≥ 1 vagy > 1). Az exponenciális eloszlásfüggvénye $\mathbf{P}(\xi < x) = 1 - e^{-\lambda x}$, ha $x \geq 0$, és 0 különben, ezért

$$\mathbf{P}(\xi \geq 1) = 1 - \mathbf{P}(\xi < 1) = e^{-\frac{1}{2}} \approx 0{,}6.$$

A második kérdésnél tudjuk, hogy már fél éve működik az izzónk. Ez egy feltételes valószínűség, hiszen van egy részinformációm, nevezetesen $\{\xi > 1/2\}$. A kérdés, hogy mennyi a valószínűsége, hogy még legalább 1 évig él, azaz $\{\xi > 3/2\}$. A feltételes valószínűség definíciója szerint

$$\mathbf{P}(\xi > 3/2|\xi > 1/2) = \frac{\mathbf{P}(\xi > 3/2, \xi > 1/2)}{\mathbf{P}(\xi > 1/2)}$$
$$= \frac{\mathbf{P}(\xi > 3/2)}{\mathbf{P}(\xi > 1/2)} = \frac{e^{-\frac{1}{2}\cdot\frac{3}{2}}}{e^{-\frac{1}{2}\cdot\frac{1}{2}}} = e^{-\frac{1}{2}}.$$

Most lényegében beláttuk, hogy az exponenciális eloszlás örökifjú, ami azt jelenti, hogy tetszőleges $s, t > 0$ esetén $\mathbf{P}(\xi > s+t|\xi > s) = \mathbf{P}(\xi > t)$, azaz ha a változóm már s időt megélt (feltétel!), akkor annak a valószínűsége, hogy még t-t megél az ugyanannyi, mint annak a valószínűsége, hogy t-t megél (nincs feltétel!).

A harmadik kérdésen kicsit el kell mélázni. Az, hogy a időt megél az izzók 90%-a azt jelenti, hogy ha egyszerre becsavarok 100 izzót, akkor a idő után még 90 db világít. Másképpen, 10 égett ki a-ig, azaz annak a valószínűsége, hogy egy izzó élettartam a-nál kisebb, az 0,1. Tehát a kérdés az az a szám, melyre

$$\mathbf{P}(\xi < a) = 0{,}1.$$

(Mivel folytonos eloszlásunk van, ezért mindegy, hogy hol van szigorú egyenlőtlenség, és hol nincs.) Beírva az eloszlásfüggvény képletét ($\lambda = 1/2$, ezt már kiszámoltuk!)

$$0{,}1 = 1 - e^{-\frac{a}{2}},$$

vagyis $a = 2\log\frac{10}{9} \approx 0{,}21$. Azaz 0,21 évet él meg a villanykörték 90%-a. □

2.4.2. A skót bakák mellkasának körmérete $N(88, 10)$ eloszlást követ. Mekkora hányaduk fér bele 84-es zubbonyba?

Megoldás. Végre egy normális eloszlás. Most mindkét paraméter meg van adva. A várható érték $\mu = 88$, a szórásnégyzet $\sigma^2 = 10$. Arra figyeljünk, hogy a szórásnégyzet a 10, nem a szórás! Azt fogjuk használni, hogy ha ξ normális eloszlású μ és σ^2 paraméterekkel, akkor $Z = (\xi - \mu)/\sigma$ standard normális eloszlású, ami meg táblázatba van szedve. Figyeljünk oda, hogy mi σ és mi σ^2!

Legyen ξ egy skót baka mellkasának körmérete. Erre úgy gondolunk, hogy a skót katonák összességéből véletlenszerűen kiválasztunk egyet. Ezzel a jelöléssel a feladat éppen a $\{\xi \leq 84\}$ esemény valószínűségét kérdezi. Ez pedig

$$\mathbf{P}(\xi < 84) = \mathbf{P}\left(\frac{\xi - 88}{\sqrt{10}} < \frac{84 - 88}{\sqrt{10}}\right) = \mathbf{P}(Z < -1{,}29) = \Phi(-1{,}26).$$

Az standard normális eloszlásfüggvény táblázatában negatív x-ek nincsenek. De nem baj, tudjuk, hogy a sűrűség páros függvény, amiből következik, hogy

$$\Phi(-x) = 1 - \Phi(x), \quad x > 0.$$

Tehát

$$\Phi(-1{,}26) = 1 - \Phi(1{,}26) = 1 - 0{,}8962 = 0{,}1038.$$

Azaz a skót bakák kb. 10%-a fér bele egy 84-es zubbonyba. □

2.4.3. Egy munkadarabokat készítő gép 40 cm-re van beállítva. A hiba normális eloszlást követ 0 várható értékkel. Annak a valószínűsége, hogy egy munkadarab nagyobb, mint 40,5 cm, 0,05. Mennyi a szórás?

Megoldás. Jelölje ξ egy munkadarab hibáját (előjellel, tehát lehet negatív is), azaz a munkadarab hossza $40+\xi$ cm. Tudjuk, hogy ξ normális eloszlású. Most nincs megadva mindkét paramétere a normális eloszlásnak. Azt tudjuk, hogy 0 a várható érték, azaz $\mu = 0$. Legyen σ^2 a szórásnégyzet. Azt tudjuk még, hogy

$$\mathbf{P}(\xi > 0{,}5) = 0{,}05,$$

hiszen a munkadarab pontosan akkor nagyobb, mint 40,5, ha a hiba nagyobb, mint 0,5. Megint a normális eloszlás skálázási tulajdonságát használjuk. Eszerint $\xi/\sigma = Z$ standard normális. Tehát

$$\mathbf{P}(\xi > 0{,}5) = \mathbf{P}(\xi/\sigma > 0{,}5/\sigma) = \mathbf{P}(Z > 0{,}5/\sigma) = 1 - \Phi(0{,}5/\sigma) = 0{,}05.$$

Innen adódik, hogy

$$\Phi(0{,}5/\sigma) = 0{,}95$$

Most azt kell megkeresni a normális eloszlás táblázatában, hogy hol veszi fel a 0,95 értéket. Ez valahol 1,64 és 1,65 között történik, legyen 1,65. Tehát azt kaptuk, hogy

$$\frac{0{,}5}{\sigma} = 1{,}65,$$

ahonnan

$$\sigma = 0{,}3.$$

Azaz a szórás 0,3. □

2.4.4. A Texpo áruházakban az i-edik kasszánál egy vásárló percben számolva i paraméterű exponenciális időt tölt el. A kiszolgálási idők az egyes kasszáknál egymástól függetlenek.

(a) András éppen üresen találja az 1-es kasszát. Mennyi a valószínűsége, hogy 2 percen belül végez?

(b) Andrással pontosan egyidőben Béla beáll az ugyancsak üres 2-es kasszához. Mennyi a valószínűsége, hogy mindketten 2 percen belül végeznek? Mennyi a valószínűsége, hogy Béla 2 percen belül végez, de András nem?

(c) Mennyi a valószínűsége, hogy valamelyikük 2 percen belül végez? Határozzuk meg a hamarabb végző kiszolgálási idejének eloszlásást! Tehát András és Béla kiszolgálási idejének a minimumára vagyunk kíváncsiak.

(d) Mennyi a valószínűsége, hogy András végez hamarabb?

Megoldás. (a) Legyen ξ András kiszolgálási ideje. Ekkor ξ exponenciális eloszlású 1 paraméterrel, ezért az a valószínűség, hogy András 2 percen belül végez

$$\mathbf{P}(\xi < 2) = 1 - e^{-1\cdot 2} = 1 - e^{-2} = 0,86.$$

(b) Legyen η Béla kiszolgálási ideje. Ez exponenciális eloszlású 2 paraméterrel, és független ξ-től. Ezért az a valószínűség, hogy mindketten 2 percen belül végeznek

$$\mathbf{P}(\xi < 2, \eta < 2) = \mathbf{P}(\xi < 2) \cdot \mathbf{P}(\eta < 2) = (1 - e^{-2}) \cdot (1 - e^{-4}) = 0,84.$$

Ha Béla 2 percen belül végez, András pedig nem, akkor $\eta < 2$ és $\xi > 2$, aminek a valószínűsége

$$\mathbf{P}(\xi > 2, \eta < 2) = \mathbf{P}(\xi > 2) \cdot \mathbf{P}(\eta < 2) = e^{-2} \cdot (1 - e^{-4}) = 0,13.$$

(c) Jelölje Z az ξ és η közül a kisebbet, azaz $Z = \min\{\xi, \eta\}$. Ekkor Z nemnegatív értékeket vesz fel. Valamely $z > 0$ esetén

$$\begin{aligned}\mathbf{P}(Z < z) &= 1 - \mathbf{P}(Z > z) = 1 - \mathbf{P}(\xi > z, \eta > z)\\ &= 1 - \mathbf{P}(\xi > z) \cdot \mathbf{P}(\eta > z) = 1 - e^{-z} \cdot e^{-2z}\\ &= 1 - e^{-3z}.\end{aligned}$$

Ez pontosan azt jelenti, hogy Z exponenciális eloszlású 3 paraméterrel.

Lényegében beláttuk (1 és 2 helyett λ és μ-t írva), hogy független exponenciálisok minimuma exponenciális, és a paraméter a paraméterek összege. Ez az egyszerű tény fontos lesz Sztochasztikus modellek kurzuson.

(d) Mivel ξ és η függetlenek, ezért együttes sűrűségfüggvényük az egyes sűrűségfüggvények szorzata, azaz

$$h(u, v) = f_\xi(u) f_\eta(v) = \begin{cases} e^{-u} 2e^{-2v}, & \text{ha } u, v > 0,\\ 0, & \text{különben.}\end{cases}$$

Az, hogy András végez hamarabb pontosan azt jelenti, hogy $\xi < \eta$, azaz $(\xi, \eta) \in \{(x, y) : x < y\} = A$. Ennek a valószínűsége

$$\begin{aligned}
\mathbf{P}(\xi < \eta) = \mathbf{P}((\xi, \eta) \in A) &= \iint_A h(u, v)\mathrm{d}u\mathrm{d}v \\
&= \int_0^\infty \left(\int_u^\infty 2e^{-u}e^{-2v}\mathrm{d}v \right) \mathrm{d}u \\
&= \int_0^\infty e^{-u} \left[-e^{-2v} \right]_{v=u}^{v=\infty} \mathrm{d}u \\
&= \int_0^\infty e^{-u}e^{-2u}\mathrm{d}u = \frac{1}{3}.
\end{aligned}$$

□

2.4.5. Frankenstein professzor vámpír denevéreket tenyészt a laboratóriumában. A denevérek tépőfogainak a hossza normális eloszlást követ $\mu = 28$ mm átlaggal és $\sigma = 4$ mm szórással. Frankenstein tudja, hogy azoknak az állatoknak a harapása halálos, akiknek a tépőfogmérete a populáció felső 5%-ába esik. Számítsuk ki, hogy ez hány mm-es fogméretet jelent!

2.4.6. A házimacskák testsúlya jó közelítéssel normális eloszlást követ. A macskák 10%-a könnyebb, mint 1,5 kg, és 20%-a nehezebb, mint 7 kg. Mekkora a 6 kg-nál nehezebb macskák aránya?

2.4.7. Tegyük fel, hogy Ausztriában a munkavállalók keresete normális eloszlást követ. Tudjuk, hogy a munkavállalók fele keres havi 3000 eurót vagy kevesebbet, míg 5%-uk keres 8000 eurónál többet. Egy törvénytervezet szerint változna az adókulcs az 5000 euRónál többet keresők számára. A munkavállalók mekkora hányadát érinti ez a változtatás?

2.4.8. A WHO adatai szerint a kétéves lányok testmagassága normális eloszlást követ 86 cm várható értékkel és 3,2 cm szórással. A kétéves kislányok hány százaléka magasabb, mint 92 cm? A Dobó utcai bölcsőde Maci csoportjában 5 két éves lány van. Mennyi az esélye, hogy van köztük 92 cm-nél magasabb?

2.4.9. Egy telefonfülke előtt állunk, és várjuk, hogy az előttünk beszélő befejezze a beszélgetést. Az illető véletlentől függő ideig beszél, az időtartam sűrűségfüggvénye (percben mérve) $e^{-(x/3)}/3$, $x > 0$.

(a) Mennyi a valószínűsége, hogy a beszélgetés 3 percnél tovább tart?

(b) Mennyi a valószínűsége, hogy a beszélgetés $t + 3$ percnél tovább tart, feltéve, hogy t percnél tovább tart?

2.4.10. Anna 30-ik születésnapjára azt a 6 darabos pohárkészletet kapja nagymamájától, mely már 100 éve a család tulajdona. A poharak élettartamai egymástól függetlenek, exponenciális eloszlást követnek 50 év várható értékkel. Adjuk meg annak a valószínűségét, hogy 50 év múlva Anna sértetlenül adhatja tovább unokájának a családi ereklyét (azaz mind a hat poharat)!

2.4.11. Az EXPO cég wifi routereinek élettartama exponenciális eloszlást követ 3 év várható értékkel. A gyártó 2 év garanciaidőt vállal a termékre. A routerek hány százaléka hibásodik meg garanciaidőn belül? Ha két routert vásárolunk, akkor mekkora annak a valószínűsége, hogy egyik sem romlik el egy éven belül?

2.4.12. Máté a Mikulástól kapott 100 grammos csokoládémikulást egy ültő helyében megeszi. Édesapja megfigyelte, hogy egy 100 grammos csokoládét exponenciális eloszlású idő alatt fogyaszt el, és átlagosan 20 perc alatt végez. Mennyi a valószínűsége, hogy a csokimikulásból még 30 perc után is marad?

Máté unokatestvérével Bencével együtt kapja meg a csomagját. Bence is exponenciális eloszlású idő alatt tünteti el a csokimikulást, ő átlagosan 15 perc alatt. Mennyi a valószínűsége, hogy 10 perc alatt mindketten végeznek a csokimikulással, ha a két csokievés időtartama egymástól független?

2.4.13 (∗). Legyen ξ pozitív egész értékű véletlen változó, melyre teljesül a diszkrét örökifjú tulajdonság, azaz

$$\mathbf{P}(\xi > k + l | \xi > l) = \mathbf{P}(\xi > k).$$

Mutassuk meg, hogy $\xi \sim \mathrm{Geom}(p)$!

2.4.14 (∗). Diszkrét örökifjúból folytonosat. Legyen $\xi_n \sim \mathrm{Geom}(\lambda/n)$. Határozzuk meg ξ_n/n határeloszlását, azaz adjuk meg a

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\xi_n}{n} \le x\right)$$

határértéket minden $x \in \mathbb{R}$ esetén!

2.4.15 (∗). Folytonos örökifjúból diszkrétet. Legyen $\xi \sim \mathrm{Exp}(\lambda)$ eloszlású véletlen változó. Határozzuk meg $\lfloor \xi \rfloor$ eloszlását! (A geometriai eloszlás a diszkrét örökifjú.)

2.5. Vektorváltozók

2.5.1. A megtakarított pénzünket értékpapírba fektetjük, 20 darabot vásárolunk az A vállalat és 10 darabot a B vállalat részvényeiből. Egy év múlva a két vállalat részvényei várható értékben 700 illetve 1500 dollárt érnek, az árfolyamok szórása pedig 20 illetve 80 dollár.

(a) Tegyük fel, hogy a részvények árfolyama független egymástól. Várhatóan mennyit ér portfóliónk egy év múlva? Mennyi a portfólió értékének a szórása?

(b) Tegyük fel, hogy a részvények árfolyama nem független egymástól. Az árfolyamok közötti korrelációs együttható függvényében írjuk fel formulával és ábrázoljuk a portfólió értékének várható értékét és szórásnégyzetét!

(c) Milyen kapcsolat van a korrelációs együttható és a befektetés kockázata között? Ha én egy kockázatkerülő befektető vagyok, akkor pozitív vagy negatív korrelációjú értékpapírokból állítsak össze portfóliót? (Szűcs Gábor feladata)

Megoldás. Jelölje ξ az A vállalat, η pedig a B vállalat részvényének értékét egy év múlva. A feladat szerint $\mathbf{E}(\xi) = 700$, $\mathbf{D}(\xi) = 20$, $\mathbf{E}(\eta) = 1500$, és $\mathbf{D}(\eta) = 80$. Egy év múlva a portfóliónk értéke $20\xi + 10\eta$.

(a) Az összeg várható értéke *mindig* a várható értékek összege, az összeg szórásnégyzete pedig akkor egyenlő a szórásnégyzetek összegével, ha a változók függetlenek (na jó, korrelálatlanság is elég). Tehát, *mindig*

$$\mathbf{E}(20\xi + 10\eta) = 20\mathbf{E}(\xi) + 10\mathbf{E}(\eta) = 29000,$$

és ha ξ és η függetlenek, akkor

$$\begin{aligned}\mathbf{D}^2(20\xi + 10\eta) &= \mathbf{D}^2(20\xi) + \mathbf{D}^2(10\eta) = (20)^2\mathbf{D}^2(\xi) + (10)^2\mathbf{D}^2(\eta)\\ &= 160000 + 640000 = 800000,\end{aligned}$$

azaz $\mathbf{D}(20\xi + 10\eta) = 400\sqrt{5}$.

(b) Láttuk,

$$\mathbf{E}(20\xi + 10\eta) = 20\mathbf{E}(\xi) + 10\mathbf{E}(\eta) = 29000,$$

azaz a várható érték nem függ a korrelációtól.

A szórásnégyzet (variancia) az általános esetben, használva a korreláció $\rho(\xi, \eta) = \mathbf{Cov}(\xi, \eta)/(\mathbf{D}(\xi)\mathbf{D}(\eta))$ definícióját, és hogy a kovariancia olyan mint a szorzás, bilineáris

$$\begin{aligned}\mathbf{D}^2(20\xi + 10\eta) &= \mathbf{Cov}(20\xi + 10\eta, 20\xi + 10\eta)\\ &= \mathbf{Cov}(20\xi, 20\xi) + 2\mathbf{Cov}(20\xi, 10\eta) + \mathbf{Cov}(10\eta, 10\eta)\\ &= \mathbf{D}^2(20\xi) + 2\mathbf{D}(\xi)\mathbf{D}(\eta)\rho + \mathbf{D}^2(10\eta)\\ &= 400^2 + 2 \cdot 400 \cdot 800 \cdot \rho + 800^2\\ &= 400^2(5 + 4\rho).\end{aligned}$$

Ez egyszerű lineáris függvény. Figyeljünk, hogy $\rho \in [-1, 1]$.

(c) Nyilván a kockázatos befektetés azért kockázatos, mert nagyon ingadozik a várható értéke körül, azaz a szórása nagyobb. Tehát, ha én kockázatkerülő vagyok, akkor minél kisebb szórásnégyzetű portfóliót szeretnék összerakni. Ezt úgy érhetem el, ha ρ a korreláció negatív. □

2.5.2. Egy vállalat egy hónapra eső profitja a havi teljes bevétel és a havi teljes kiadás különbségeként áll elő, ahol a bevétel és a kiadás is véletlen változó. A bevétel várható értéke 120 millió forint 30 millió forint szórással, míg a kiadás várható értéke 80 millió forint 20 millió forint szórással. Határozzuk meg az egy hónapra jutó profit várható értékét és szórását akkor, ha a bevétel és a kiadás független, illetve akkor, ha a közöttük lévő korrelációs együttható 0,8. A korreláció függvényében írjuk fel formulával és ábrázoljuk grafikonon a profit várható értékét és szórásnégyzetét! (Szűcs Gábor feladata)

2.5.3. Bence és Luca testvérek. Ebéd után mindketten véletlentől függő ideig alszanak. Bence esetében ez átlagosan 2 óra, a szórás 30 perc, míg Luca esetében 1,5 óra, 20 perc szórással. Határozzuk meg Bence és Luca együttes (Bence + Luca) alvásának várható értékét és szórását, ha az alvásmennyiségek egymástól függetlenek, illetve akkor, ha a közöttük lévő korrelációs együttható 0,5.

2.5.4. A Real Madrid 2018/2019-es idényben az egy mérkőzésen lőtt góljainak száma Poisson-eloszlást követ $\lambda = 3$ paraméterrel, míg a kapott gólok száma Poisson-eloszlást követ $\mu =$0,7 paraméterrel. Adjuk meg a Real Madrid egy mérkőzésén esett összes gól számának várható értékét és szórásnégyzetét abban az esetben ha

(a) a lőtt és kapott gólok száma függetlenek;

(b) a lőtt és kapott gólok számának korrelációs együtthatója 0,4.

2.5.5. A 2022-es labdarúgó világbajnokság nyitómérkőzésén a rendező Katar játszik Ecuador ellen. Tegyük fel, hogy Katar egy mérkőzésen lőtt góljai számánák várható értéke és szórása 0,5, míg a kapott gólok számának várható értéke 2,7, szórása 2. Adjuk meg a nyitómérkőzésen eső összes gólok számának várható értékét és szórását, ha

(a) a lőtt és kapott gólok száma függetlenek;

(b) a lőtt és kapott gólok számának korrelációs együtthatója 0,4.

2.5.6. Három, külsőre egyforma érmével a fejdobás valószínűsége $1/4, 1/2$, és $3/4$. Véletlenszerűen választunk egy érmét, és azzal kétszer dobunk. Legyen η a fej valószínűsége a választott érmén, ξ a dobott fejek száma. Adjuk meg az együttes eloszlást!

Megoldás. Világos, hogy ξ lehetséges értékei 0,1,2, míg η lehetséges értékei $1/4, 1/2, 3/4$. A szorzási szabállyal kapjuk

$$\begin{aligned}\mathbf{P}\left(\eta = \frac{1}{4}, \xi = 0\right) &= \mathbf{P}\left(\eta = \frac{1}{4}\right) \cdot \mathbf{P}\left(\xi = 0 | \eta = \frac{1}{4}\right) \\ &= \frac{1}{3} \cdot \left(\frac{3}{4}\right)^2 = \frac{9}{48}.\end{aligned}$$

A többi eset hasonlóan számolható, pl.

$$\mathbf{P}\left(\eta = \frac{1}{2}, \xi = 1\right) = \frac{1}{3} \cdot 2\left(\frac{1}{2}\right)^2 = \frac{1}{6}.$$

Minden valószínűséget kiszámolva az alábbi táblázatot kapjuk, ami megadja a (ξ, η) együttes eloszlását:

$\eta \backslash \xi$	0	1	2	$\sum$
$\frac{1}{4}$	$\frac{9}{48}$	$\frac{6}{48}$	$\frac{1}{48}$	$\frac{1}{3}$
$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{3}$
$\frac{3}{4}$	$\frac{1}{48}$	$\frac{6}{48}$	$\frac{9}{48}$	$\frac{1}{3}$
$\sum$	$\frac{14}{48}$	$\frac{20}{48}$	$\frac{14}{48}$	1

A táblázat utolsó oszlopában megjelenik η eloszlása, míg az utolsó sorban ξ eloszlása, ezek a peremeloszlások. □

2.5.7. Egy szabályos dobókockával 10-szer dobunk. Jelülje ξ az egyesek, η a kettesek számát! Adjuk meg (ξ, η) véletlen vektorváltozó eloszlását és várható értékét!

Megoldás. Már láttuk, hogy külön-külön ξ és η binomiális eloszlást követnek 10 és 1/6 paraméterekkel. Így

$$\mathbf{E}(\xi) = \mathbf{E}(\eta) = \frac{10}{6}.$$

A (ξ, η) vektor lehetséges értékei olyan (k, ℓ) számpárok, melyekre $k, \ell \in \{0, 1, \ldots, 10\}$ és $k + \ell \leq 10$. Ez világos, hiszen az egyesek és kettesek száma összesen legfeljebb 10. Az eloszlás

$$\begin{aligned}\mathbf{P}((\xi, \eta) = (k, \ell)) &= \mathbf{P}(\xi = k, \eta = \ell) \\ &= \binom{10}{k}\binom{10-k}{\ell}\left(\frac{1}{6}\right)^k \left(\frac{1}{6}\right)^\ell \left(\frac{4}{6}\right)^{10-k-\ell},\end{aligned}$$

hiszen először kiválasztom a 10 dobás közül a k db egyest, majd a $10-k$ közül az ℓ db kettest. Ennél a $k+\ell$ dobásnál nincs választásom, míg a maradék $10-k-\ell$ dobás nem lehet sem egyes, sem kettes, így 4 lehetőség maradt, innen jön a $(4/6)^{10-k-\ell}$. $\square$

2.5.8. Egy urnában van 1 piros, 2 kék, és 3 zöld golyó. Visszatevés nélkül húzunk 3 golyót. Jelölje ξ a pirosak, η a kékek számát! Adjuk meg a (ξ, η) véletlen vektor eloszlását és várható értékét!

2.5.9. Két szabályos kockával játszunk. Jelölje ξ az első kockával dobott számot és η a dobott számok nagyobbikát. Adjuk meg az együttes eloszlást és az összeg várható értékét, szórását!

Megoldás. Mind az ξ, mind az η véletlen változó lehetséges értékei $1, 2, \ldots,$ 6. Tehát ξ, η diszkrét véletlen változók. Az (ξ, η) vektorváltozó lehetséges értékei számpárok. Mivel $\eta \geq \xi$, hiszen η a két szám közül a nagyobb, ezért a lehetséges értékek halmaza

$$\{(i, j) \,:\, 1 \leq i \leq j \leq 6\}.$$

Először meghatározzuk a $\mathbf{P}((\xi, \eta) = (1, 1))$ valószínűséget. Mivel ξ az első kockával dobott szám, η pedig a két szám nagyobbika, ezért csak úgy lehet mindekettő 1, ha mindkét kockával 1-et dobtunk. Mivel az összes eset száma $6 \cdot 6 = 36$, így

$$\mathbf{P}((\xi, \eta) = (1, 1)) = \mathbf{P}(\xi = 1, \eta = 1) = \frac{1}{36}.$$

Most vizsgáljuk a $\mathbf{P}(\xi = 1, \eta = 2)$ valószínűséget. Az első kockával megint 1-et dobtunk, és a dobott számok maximuma 2. Ez csak úgy lehet, ha az első kockával 1-est, a másodikkal 2-est dobtunk. Megint 1 kedvező eset, ezért

$$\mathbf{P}((\xi, \eta) = (1, 2)) = \mathbf{P}(\xi = 1, \eta = 2) = \frac{1}{36}.$$

Látjuk, hogy tetszőleges $j = 1, 2, \ldots, 6$ esetén

$$\mathbf{P}((\xi, \eta) = (1, j)) = \mathbf{P}(\xi = 1, \eta = j) = \frac{1}{36}.$$

Nézzük a $\mathbf{P}(\xi = 2, \eta = 2)$ valószínűséget. (Persze, ha $\xi = 2$ akkor $\eta \geq 2$.) Ez azt jelenti, hogy az első kockával 2-est dobtunk, és a dobott számok nagyobbika 2. Ez úgy lehet, ha a második kockával 1-est vagy 2-est dobtunk. Azaz, most két kedvező eset van,

$$\mathbf{P}((\xi, \eta) = (2, 2)) = \mathbf{P}(\xi = 2, \eta = 2) = \frac{2}{36}.$$

Ha $\xi = 2$ és $\eta = 3$, akkor az első kockával 2-est dobtunk, a nagyobbik szám pedig 3, tehát a második kockával 3-ast dobtunk. Ez megint egy eset, tehát

$$\mathbf{P}(\xi = 2, \eta = 3) = \frac{1}{36}.$$

Ugyanígy tetszőleges $j \geq 3$ esetén

$$\mathbf{P}(\xi = 2, \eta = j) = \frac{1}{36}.$$

Most már látjuk az általános képet. Ha $\xi = i$, $\eta = i$, akkor az azt jelenti, hogy az első kockával i-t dobtunk, és a nagyobbik szám i, tehát a második kockával az $1, 2, \ldots, i$ számok valamelyikét dobtuk. Ez i lehetőség, azaz

$$\mathbf{P}(\xi = i, \eta = i) = \frac{i}{36}.$$

Ha pedig $\xi = i$ és $\eta = j > i$, akkor az első szám i, a nagyobbik j, és mivel ez nem lehet az első, ezért a második j. Tehát

$$\mathbf{P}(\xi = i, \eta = j) = \frac{1}{36}.$$

Röviden

$$\mathbf{P}(\xi = i, \eta = j) = \begin{cases} \frac{i}{36}, & \text{ha } i = j, \\ \frac{1}{36}, & \text{ha } i < j. \end{cases}$$

Ez az együttes eloszlás.

A peremeloszlásokat úgy kapjuk meg, hogy rögzítjük az egyik változó értékét, és kiösszegzünk a másik összes lehetséges értékére (folytonos esetben ugyanez, csak összegzés helyett, a sűrűséget kell integrálni). Azaz

$$\mathbf{P}(\xi = i) = \sum_{j=i}^{6} \mathbf{P}(\xi = i, \eta = j) = \frac{i}{36} + \sum_{j=i+1}^{6} \frac{1}{36} = \frac{6}{36} = \frac{1}{6}.$$

Ez persze nem nagy meglepetés, ehhez nem kellett volna ennyit számolni. A szabályos kockával $1/6$ a valószínűsége minden lehetséges értéknek. Az η eloszlása pedig

$$\mathbf{P}(\eta = j) = \sum_{i=1}^{j} \mathbf{P}(\xi = i, \eta = j) = \sum_{i=1}^{j-1} \frac{1}{36} + \frac{j}{36} = \frac{2j-1}{36}.$$

Innen a várható értékek definíció szerint számolhatók. Valóban

$$\mathbf{E}(\xi) = \sum_{i=1}^{6} \frac{1}{6} \cdot i = \frac{1}{6} \frac{6 \cdot 7}{2} = \frac{7}{2},$$

és

$$\mathbf{E}(\eta) = \sum_{j=1}^{6} \frac{2j-1}{36} \cdot j = \frac{1}{36}\left(2\frac{6\cdot 7\cdot 13}{6} - \frac{6\cdot 7}{2}\right) = \frac{182-21}{36} = \frac{161}{36}.$$

Az összeg várható értéke az a várható értékek összege, azaz

$$\mathbf{E}(\xi+\eta) = \frac{7}{2} + \frac{161}{36} = \frac{126+161}{36} = \frac{287}{36}.$$

Az összeg szórásnégyzetének kiszámítása macerásabb, hiszen ξ és η nyilván nem függetlenek (hát persze, a lehetséges értékek $1, 2, \ldots, 6$, ugyanakkor $\xi \leq \eta$). Az összeg szórásnégyzete akkor egyenlő a szórásnégyzetek összegével, ha *a változók függetlenek* (vagy korrelálatlanok). Általában, a kovariancia tulajdonságai szerint (vegyük észre, hogy a kovariancia úgy viselkedik, mint egy szorzás; pontosabban ő egy bilineáris funkcionál)

$$\begin{aligned}\mathbf{D}^2(\xi+\eta) &= \mathbf{Cov}(\xi+\eta, \xi+\eta)\\ &= \mathbf{Cov}(\xi,\xi) + 2\mathbf{Cov}(\xi,\eta) + \mathbf{Cov}(\eta,\eta)\\ &= \mathbf{D}^2(\xi) + 2\mathbf{Cov}(\xi,\eta) + \mathbf{D}^2(\eta).\end{aligned}$$

A szórásnégyzetek definíció alapján számolhatók

$$\begin{aligned}\mathbf{D}^2(\xi) = \mathbf{E}(\xi^2) - (\mathbf{E}(\xi))^2 &= \sum_{i=1}^{6} \frac{1}{6}\cdot i^2 - \left(\frac{7}{2}\right)^2\\ &= \frac{6\cdot 7\cdot 13}{6\cdot 6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}.\end{aligned}$$

Ugyanígy

$$\mathbf{D}^2(\eta) = \mathbf{E}(\eta^2) - (\mathbf{E}(\eta))^2,$$

és

$$\mathbf{E}(\eta^2) = \sum_{j=1}^{6} \frac{2j-1}{36}\cdot j^2 = \frac{1}{36}\left(2\left(\frac{6\cdot 7}{2}\right)^2 - \frac{6\cdot 7\cdot 13}{6}\right) = \frac{791}{36},$$

tehát

$$\mathbf{D}^2(\eta) = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296}.$$

Maradt a kovariancia. A kovariancia tulajdonságai szerint

$$\mathbf{Cov}(\xi,\eta) = \mathbf{E}(\xi\eta) - (\mathbf{E}(\xi)\cdot\mathbf{E}(\eta)).$$

A várható értékeket már kiszámoltuk, tehát már csak a szorzat várható értéke kell. A szorzat egy kétváltozós függvény, tehát a várható érték tulajdonságai szerint

$$\mathbf{E}(\xi\eta) = \sum_{i,j} \mathbf{P}(\xi = i, \eta = j) \cdot i \cdot j.$$

Ez egy 21 tagú összeg, ki kell számolni:

$$\begin{aligned}
\mathbf{E}(\xi\eta) &= \sum_{i=1}^{6} \sum_{j=i}^{6} \mathbf{P}(\xi = i, \eta = j) \cdot i \cdot j \\
&= \sum_{i=1}^{6} i \left(i \cdot \frac{i}{36} + \frac{1}{36} \frac{6 \cdot 7 - i(i+1)}{2} \right) \\
&= \frac{1}{36} \sum_{i=1}^{6} i \cdot \left(i^2 + 21 - \frac{i(i+1)}{2} \right) \\
&= \frac{1}{36} \sum_{i=1}^{6} \left[\frac{i^3}{2} - \frac{i^2}{2} + 21 \cdot i \right] \\
&= \frac{1}{36} \left[\frac{1}{2} \left(\frac{6 \cdot 7}{2} \right)^2 - \frac{1}{2} \frac{6 \cdot 7 \cdot 13}{6} + 21 \cdot \frac{6 \cdot 7}{2} \right] = \frac{154}{9}
\end{aligned}$$

Tehát

$$\mathbf{Cov}(\xi, \eta) = \frac{154}{9} - \frac{7}{2} \cdot \frac{161}{36} = \frac{35}{24}.$$

A kovariancia pozitivitása azt jelenti, hogy ha az egyik változó nagy, akkor a másik is hajlamos nagyobb lenni. Ez intuitíven világos ebben a példában.

$\square$

2.5.10. A Jonas Brothers nevű együttes újra összeáll és koncertet adnak. A PepsiCo cég a következő ötlettel áll elő: a kólásüvegeik kupakjában elrejtik a banda egy-egy tagjának a nevét és azok között, akik összegyűjtik mindhárom nevet kisorsolnak egy VIP belépőt. Kevin neve a kupakok felén szerepel, Joe-val a kupakok egyharmadában találkozhatunk és Nick a legritkább, neve átlagosan minden hatodik kupakban szerepel. Veszünk 10 kólát. Adjuk meg a Joe feliratú kupakok számának várható értékét és szórását! Adjuk meg a Joe feliratú és a Kevin feliratú kupakok számának kovarianciáját és korrelációs együtthatóját!

2.6. De Moivre–Laplace tétel

2.6.1. Az FC Barcelona passzolási hatékonysága 2019. áprilisában $p = 0{,}85$ (azaz egy passz ekkora valószínűséggel sikeres). Adjuk meg annak a valószínűségét, hogy a Liverpool elleni 525 passzból legalább 460 sikeres.

Megoldás. Ez a legklasszikusabb példa a de Moivre–Laplace-tétel alkalmazására. Jelölje S a sikeres passzok számát. Mivel összesen 525 passz van, és mindegyik passz egymástól függetlenül $0,85$ valószínűséggel sikeres, ezért S eloszlása binomiális $n = 525$ és $p = 0{,}85$ paraméterekkel. A de Moivre–Laplace-tétel szerint

$$\frac{S - np}{\sqrt{np(1-p)}}$$

közelítőleg standard normális eloszlású, azaz

$$\mathbf{P}\left(a \leq \frac{S - np}{\sqrt{np(1-p)}} \leq b\right) \approx \Phi(b) - \Phi(a).$$

A pontos állítás az, hogy ha $n \to \infty$, akkor az $\approx$ helyett $=$ van. Ha csak felső korlát van, akkor $a = -\infty$ és $\Phi(-\infty) = 0$, ha csak alsó, akkor $b = \infty$ és $\Phi(\infty) = 1$. A feladat kérdése a $\mathbf{P}(S \geq 460)$ valószínűség. Egyszerű átalakítással elérjük, hogy $(S - np)/\sqrt{np(1-p)}$ jelenjen meg. Az $np = 525 \cdot 0{,}85 \approx 446$ és $\sqrt{np(1-p)} \approx 8{,}2$ értékeket beírva

$$\begin{aligned}\mathbf{P}(S \geq 460) &= \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \geq \frac{460 - np}{\sqrt{np(1-p)}}\right) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \geq 1{,}7\right) \\ &\approx 1 - \Phi(1{,}7) = 1 - 0{,}955 = 0{,}045.\end{aligned}$$

Tehát a keresett valószínűsége 0,045. □

2.6.2. Magyarországon, és mindenütt a világon, több fiúgyermek születik, mint lány. Az újszülöttek 52%-a fiú, 48%-a lány. Nevezzük *lányos napoknak/heteknek* azokat a napokat/heteket, amikor több lány születik, mint fiú.

(a) Szegeden naponta 9 gyermek születik. Mennyi a pontos valószínűsége, hogy Szegeden egy adott nap lányos nap? Milyen eloszlású az egy héten bekövetkezett lányos napok száma? Várhatóan hány lányos nap van egy héten?

(b) Budapesten naponta 100 gyermek születik. Mennyi a közelítő (normális közelítés, de Moivre–Laplace-tétel) valószínűsége, hogy Budapesten egy adott nap lányos nap?

(c) Egész Magyarországon egy héten 2500 gyermek születik. Mennyi a lányos hét bekövetkezésének közelítő valószínűsége? Várhatóan hány lányos hét lesz 2020-ban? Adjuk meg annak a közelítő valószínűségét (Poisson-közelítés), hogy legalább 3 lányos hét lesz 2020-ban.

Megoldás. (a) Egy nap 9 gyermek születik. Annak a valószínűsége, hogy egy gyermek lány 0,48. Ezek az események egymástól függetlenek. Ezért, ha ξ jelöli a lányok számát, akkor ξ binomiális eloszlású $n = 9$ és $p = 0,48$ paraméterekkel. Ha ξ a lányok száma, akkor $9 - \xi$ a fiúké, és pontosan akkor születik több lány, ha $\xi \geq 5$. Ennek a valószínűsége

$$\mathbf{P}(\xi \geq 5) = \sum_{k=5}^{9} \mathbf{P}(\xi = k) = \sum_{k=5}^{9} \binom{9}{k} (0,48)^k \, (0,52)^{9-k} = 0,45.$$

Tehát annak a valószínűsége, hogy egy nap lányos nap, 0,45. A egyes napokon történt születések egymástól függetlenek, ezért az egy héten bekövetkezett lányos napok száma binomiális eloszlást követ, $n_1 = 7$ és $p_1 = 0,45$ paraméterekkel. A binomiális várható értéke $n_1 \cdot p_1 = 3,15$. Várhatóan 3,15 lányos nap van egy héten.

(b) Jelölje most ξ_B a Budapesten egy napon született lányok számát. Ekkor ξ_B binomiális $n = 100$ és $p = 0,48$ paraméterekkel. Akkor lesz több lány, ha $\xi_B \geq 51$. A de Moivre–Laplace-tétel szerint (most $np = 48$ és $\sqrt{np(1-p)} = 5$)

$$\mathbf{P}\left(\xi_B \geq 51\right) = \mathbf{P}\left(\frac{\xi_B - np}{\sqrt{np(1-p)}} \geq \frac{51 - np}{\sqrt{np(1-p)}}\right)$$
$$\approx \mathbf{P}(Z > 0,6) = 1 - \Phi(0,6) = 0,27.$$

Itt Z standard normális véletlen változó. Tehát közelítőleg 27% a lányos nap valószínűsége. A pontos valószínűség

$$\sum_{k=51}^{100} \binom{100}{k} (0,48)^k (0,52)^{100-k} = 0,30.$$

Egy héten várhatóan $0,27 \cdot 7 = 1,89$ lányos nap van.

(c) Ez ugyanaz, mint az előbb csak $n = 2500$ és $p = 0,48$ paraméterekkel. Legyen ξ_M a Magyarországon egy héten született lányok száma. Akkor van lányos hét, ha $\xi_M \geq 1251$. Most $np = 1200$, $\sqrt{np(1-p)} = 25$, így

$$\mathbf{P}\left(\xi_M \geq 1251\right) = \mathbf{P}\left(\frac{\xi_M - np}{\sqrt{np(1-p)}} \geq \frac{1251 - np}{\sqrt{np(1-p)}}\right)$$
$$\approx \mathbf{P}(Z > 2,04) = 1 - \Phi(2,04) = 0,02.$$

Tehát a magyarországi lányos hét valószínűsége 0,02. Egy évben 52 hét van, így az egy évben levő lányos hetek száma binomiális eloszlású $n = 52$ és $p = 0,02$ paraméterekkel. Mivel p kicsi n pedig nagy, ezt közelíthetjük Poisson-eloszlással. Mivel a várható érték megegyezik a paraméterrel, ezért $\lambda = 52 \cdot 0,02 = 1,04$. Legyen tehát η Poisson-eloszlású $\lambda = 1,04$ paraméterrel. Annak a valószínűsége, hogy legalább 3 lányos hét lesz egy évben

$$\mathbf{P}(\eta \geq 3) = 1 - \mathbf{P}(\eta \leq 2) = 1 - \left(e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2}{2} e^{-\lambda}\right) = 0,09.$$

□

2.6.3. Chicago és Los Angeles között két vasútvonal van, melyek mindegyikén egy-egy vonat közlekedik. Mindkét vonat egyidőben indul, lényegében egyformán kényelmes és k személyes. Tegyük fel, hogy 1000 utas egymástól függetlenül $1/2-1/2$ valószínűséggel választ vonatot. Legalább mekkora legyen az ülőhelyek k száma, hogy 0,01-nél kisebb legyen annak a valószínűsége, hogy lesz olyan utas, akinek nem jut ülőhely? ([2, 186.o])

Megoldás. Ez is de Moivre–Laplace. Jelölje S az A társasággal utazók számát. Ekkor a B-vel utazók száma $1000 - S$. Mivel $n = 1000$ ember egymástól függetlenül $1/2 - 1/2$ valószínűséggel dönt A ill. B mellett, ezért S binomiális eloszlású $n = 1000$ és $p = 1/2$ paraméterekkel. Legyen k az ülőhelyek száma (mindkét vonaton). Az, hogy lesz olyan, akinek nem jut hely, azt jelenti, hogy vagy az A társaságnál túl sokan vannak, vagy a B-nél. Azaz, vagy $S > k$ (A vonat betelik) vagy $1000 - S > k$ (B vonat betelik). A kérdés a legkisebb olyan k érték, melyre

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) \leq 0{,}01.$$

Vegyük észre, hogy $k \geq 500$ kell legyen, és ekkor csak az egyik vonat telhet be, azaz a fenti valószínűségben szereplő két esemény egymást kizáró, ahonnan

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) = \mathbf{P}(S > k) + \mathbf{P}(S < 1000 - k).$$

Használjuk a de Moivre–Laplace-tételt. Beírjuk az $np = 500$, $\sqrt{np(1-p)} = 5\sqrt{10} \approx 15{,}8$ értékeket. Ekkor

$$\mathbf{P}(S > k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{k - np}{\sqrt{np(1-p}}\right) \approx 1 - \Phi((k - 500)/15{,}8),$$

és ugyanígy (csak felhasználjuk, hogy $\Phi(-x) = 1 - \Phi(x)$ ha $x > 0$, és hogy $500 - k \leq 0$)

$$\mathbf{P}(S < 1000 - k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{1000 - k - np}{\sqrt{np(1-p}}\right)$$
$$\approx 1 - \Phi((k - 500)/15{,}8).$$

Azaz, keressük azt a legkisebb k értéket melyre

$$2\left[1 - \Phi((k - 500)/15{,}8)\right] \leq 0{,}01.$$

Átrendezve,

$$\Phi((k - 500)/15{,}8) \geq 0{,}995.$$

Kikeressük a normális eloszlás táblázatából, hogy hol veszi fel Φ a 0,995 értéket. Ez 2,57, azaz

$$\frac{k - 500}{15{,}8} \geq 2{,}57,$$

amit átrendezve

$$k \geq 540{,}6.$$

Mivel k egész, ezért 541 a legkisebb olyan k, amire a feltétel teljesül. □

2.6.4 (∗)**.** Budapesten meg akarják állapítani a dohányosok p arányát. Ehhez kiválasztanak n egyént úgy, hogy minden választásnál mindenki ugyanakkora valószínűséggel kerül kiválasztásra, és csak ezek közt nézik meg a dohányosok k számát. Legalább mekkora legyen az n, hogy a kapott $p' = k/n$ arány legalább 0,95 valószínűséggel legfeljebb 0,005 hibával közelítse a valódi p arányt, akármi is $p \in (0, 1)$? ([2, 187.o])

Megoldás. Ez már érdekesebb feladat, ugyanis semmi nincs megadva. Világos, hogy a feladat nagyon fontos, ugyanis a közvéleménykutatásokhoz pontosan ilyen típusú kérdést kell feltenni. Van valami ismeretlen valószínűség p, ami azt mutatja meg, hogy az emberek ilyen aránya szavazna az A párt jelöltjére. Nem tudjuk mi a p, de erről szeretnénk valamit mondani. Hány embert kell megkérdezni, hogy valami okosat mondhassunk?

Jelölje S a dohányzók számát a megkérdezettek között. Ekkor S binomiális eloszlású véletlen változó n (meghatározandó, de ismert) és p (ismeretlen) paraméterekkel. Világos, hogy az ismeretlen p értékre az S/n becslést adjuk. Elég összetett a kérdés, kicsit el kell rajta gondolkodni. A becslés hibája $|S/n - p|$. Azt akarjuk, hogy ez nagy valószínűséggel (0,95) kicsi legyen (0,005-nél kisebb), azaz olyan n értéket keresünk, amire

$$\mathbf{P}\left(\left|\frac{S}{n} - p\right| < 0{,}005\right) \geq 0{,}95.$$

(Annak a valószínűsége, hogy a hiba 0,005-nél kisebb, legalább 0,95.) A neheze megvan. Használjuk a de Moivre–Laplace-tételt. Eszerint

$$\mathbf{P}\left(a \le \frac{S-np}{\sqrt{np(1-p)}} \le b\right) \approx \Phi(b) - \Phi(a).$$

Tehát be kell erőltetni az $(S-np)/\sqrt{np(1-p)}$ kifejezést. Tegyük meg:

$$\begin{aligned}
\mathbf{P}\left(\left|\frac{S}{n} - p\right| < 0{,}005\right) &= \mathbf{P}\left(\left|\frac{S-np}{n}\right| < 0{,}005\right) \\
&= \mathbf{P}\left(\left|\frac{S-np}{\sqrt{np(1-p)}}\right| < 0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) \\
&\approx \Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - \Phi\left(-0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) \\
&= 2\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1.
\end{aligned}$$

Annyit használtunk, hogy $|x| \le a$ pontosan akkor, ha $-a < x < a$, és hogy $\Phi(-x) = 1 - \Phi(x)$. Ezek szerint az kell, hogy

$$2\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1 \ge 0{,}95,$$

azaz

$$\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) \ge 0{,}975.$$

A táblázatból kikeresve azt kapjuk, hogy

$$0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}} \ge 1{,}96.$$

Átrendezve

$$n \ge 392^2 \cdot p(1-p). \tag{$*$}$$

Na de nem ismerjük p értékét. Úgy kell n-et választani, hogy a fenti egyenlőtlenség minden p-re igaz legyen. Tehát válasszuk p-t úgy, hogy a jobb oldal maximális legyen. Ez $p = 1/2$-nél van, értéke $1/4$. Tehát, ha

$$n \ge 396^2\frac{1}{4} = 38416,$$

akkor $(*)$ teljesül minden $p \in [0,1]$ esetén. □

2.6.5. Marc Elsberg Sosem elég című regényében Fitzroy Peel a következő játékot ajánlja. 100 ponttal indulunk. Egy szabályos érmével dobunk százszor. Ha az eredmény fej, a pontszámunk 50%-kal nő, ha írás 40%-kal csökken. Akkor nyerünk száz dobás után több pontunk van, mint 100. Mennyi a nyerés esélye?

Megoldás. Ha k-szor dobunk fejet, akkor a pontszámunk $1.5^k(0.6)^{100-k}$-szorosára változott. Ez pontosan akkor nagyobb 1-nél, ha

$$k \geq 100 \cdot \frac{\log \frac{5}{3}}{\log \frac{5}{2}} \approx 55.7$$

Jelölje S_n a fejek számát. Ekkor $S_n \sim \text{Binomiális}(100, 0.5)$. Tehát a nyerés valószínűsége

$$\mathbf{P}(S_n \geq 56) = \mathbf{P}\left(\frac{S_n - np}{\sqrt{np(1-p)}} \geq 1.2\right) \approx 1 - \Phi(1.2) = 0.115.$$

A feladat érdekessége, hogy a pontszám várható értéke minden lépésben 1,05-szörösére nő, így 100 dobás után a várható pontszám 131-szeresére nő. $\square$

2.6.6. A gyerekek 30%-a vásott, ők virgácsot is kapnak. Mikulásnál már csak 2230 csomag van, ezek közül 680-ban van virgács. Éppen egy olyan falu fölött repül, ahol 2100 gyerek van. Mennyi a valószínűsége, hogy minden gyereknek jut megfelelő mikuláscsomag?

2.6.7. Bergengóciában engedélyezték a gyorsan terjedő Covarπ vírus elleni új oltóanyagot. A lakosság egy része bizalmatlan az oltással szemben, és az előzetes felmérések szerint a bergengócok 30%-a nem oltatja be magát (azaz ekkora a valószínűsége, hogy egy bergengóc nem oltatja be magát). Legalább mennyi oltást rendeljen egy 2100 lakosú falu háziorvosa, hogy 99%-os valószínűséggel mindenkinek jusson oltóanyag?

2.6.8. Egy szabályos dobókockát feldobunk 200-szor. Jelölje S_n a dobott hatosok számát. Adjuk meg pontosan, majd a de Moivre– Laplace tétellel közelítve a $\mathbf{P}(30 < S_n \leq 40)$ valószínűséget!

2.6.9. Egy étteremben kétféle menü közül lehet választani. A vendégek 5/6 valószínűséggel A menüt, 1/6 valószínűséggel B menüt választanak. Egy adott napon 500 vendég érkezik. A vendéglős 420 A és 100 B menüt készített elő. Feltételezve, hogy a vendégek egymástól függetlenül választanak, mi a valószínűsége, hogy mindenkinek jut olyan menü, amilyet kér?

2.6.10. Egy szerencsejátékon a nyerési esélyed $1/11$. Ha nyersz, visszakapod a feltett tétet és még nyereményként annak kilencszeresét. Elegendő sok kezdőtőkével indulva ezer alkalommal felteszel 1–1 petákot. Mi a valószínűsége, hogy ezer játszma után még legalább annyi pénzed van, mint kezdetben volt?

2.6.11. Az utóbbi években felmerült a mozilátogatókban az igény arra, hogy eredeti hanggal feliratosan is megnézhessék a filmeket. Egy friss felmérés szerint az emberek $2/7$-e választaná a feliratos filmet a szinkronizált változattal szemben. A szegedi Cinema City vezetősége minket kért fel arra, hogy segítsünk dönteni Christopher Nolan The Dark Knight Rises című filmjének premierje kapcsán. A premierre 1000 látogatót várnak, és a filmet 9 egyenként 130 fős teremben vetítik. Hány teremben kell feliratosan vetíteni a filmet, hogy a lehető legnagyobb valószínűséggel tudjon minden látogató arra a változatra beülni, amire szeretne? Mekkora ez a valószínűség?

2.6.12. Egy általános iskolában egy és két forintosok gyűjtését hirdetik meg a pénzérmék bevonása előtti fél évben. Megkérik az oda járó diákokat és szüleiket, hogy az otthoni felesleges apórópénzüket az iskolának adják, hogy az így befolyt összegből játszóteret építhessenek az iskolaudvaron. A játszótér megépítéséhez 1,5 millió Ft-ra van szükségük. A gyűjtés során egymillió darab pénzérmét adományoztak az iskolának. Ha ezen pénzérmék mindegyike a többitől függetlenül $1/2 - 1/2$ valószínűséggel egy illetve két forintos, akkor mennyi a közelítő valószínűsége, hogy az igazgatónak legfeljebb 1000 Ft-tal kell hozzájárulnia a játszótér megépüléséhez?

2.6.13. A héten jelenik meg George R. R. Martin új könyve Winds of Winter címmel. A könyvesboltunkba rendeltünk a keményborítós kiadásból 180-et a puhakötésűből pedig 240-et. Tapasztalataink alapján az emberek 40%-a választja a tartósabb, de valamivel drágább keményborítású kiadást. Ha 400 vevőre számítunk, akik egymástól függetlenül döntenek, akkor milyen valószínűséggel tudunk mindenkit kiszolgálni?

2.6.14. Az OMSZ heves vihar miatt másodfokú riasztást adott ki Tolna megyére. Faddon 2100 családnak van a FIZET biztosítónál házbiztosítása. A korábbi évek tapasztalata alapján a FIZET biztosító vezetősége tudja, hogy heves vihar esetén az egyes házakban, egymástól függetlenül 0,3 valószínűséggel keletkezik kár. Becsüljük meg annak a valószínűségét, hogy a biztosítóhoz legalább 651 kártérítési igény érkezik!

2.6.15. Kávézót szeretnénk nyitni egy város forgalmas utcáján. A közelben van egy kávézó, ahol naponta 300 ember megfordul. Mi ezen emberek egy részét szeretnénk elcsábítani, hogy hozzánk térjenek be. Az eltérő hangulat, árak és üzletstratégia alapján úgy gondoljuk, hogy az emberek 27%-át sikerül erre rávennünk. Hány férőhelyesre tervezzük a kávézónkat, ha azt szeretnénk, hogy 90%-os biztonsággal minden betérőnek jusson szabad hely?

3. Statisztika

3.1. Alapstatisztikák, pontbecslések

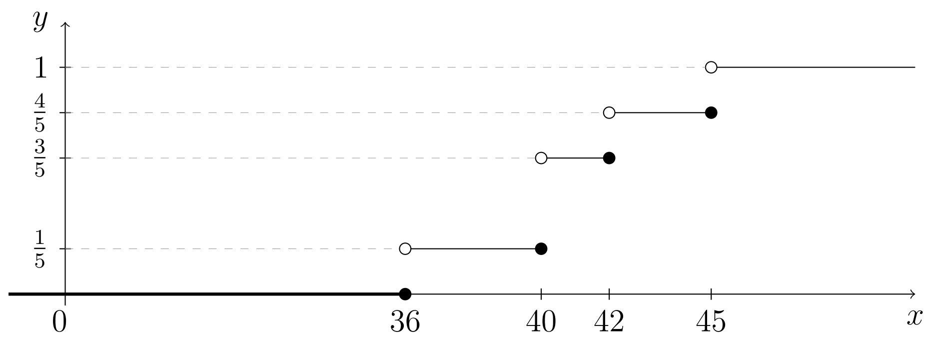
3.1.1. Egy játék fizikai terhelhetőségére elvégzett tesztek kg-ban a következő eredményeket adták: 40, 45, 40, 42, 36. A minta alapján adjuk meg a terhelhetőség empirikus eloszlásfüggvényét, mintaátlagát, korrigált/korrigálatlan empirikus szórásnégyzetét és a mediánt!

Megoldás. Ez egy $n = 5$ elemű minta, ahol a mintaelemek $x_1 = 40$, $x_2 = 45$, $x_3 = 40$, $x_4 = 42$, $x_5 = 36$. A jelölés kicsit szokatlan. A véletlen mintát mindig ξ-kel (vagy nagy X-ekkel) jelöljük, ugyanakkor egy konkrét mintát ami a kezünkben van, a véletlen minta egy *realizációját* kis x-ekkel jelöljük.

A definíciókat kell tudni. Az empirikus eloszlásfüggvény

$$F_n(x) = \frac{1}{n} \left| \{ i : \xi_i < x \} \right|,$$

azaz rögzített x-re megszámoljuk, hogy hány x-nél kisebb mintaelemünk van, és ezt elosztjuk a minta elemszámával. Vagyis az empirikus eloszlásfüggvény a mintaelemekben ugrik $1/n$-et (ha több mintaelem egyenlő, akkor annyiszor $1/n$-et, ahányszor az adott értéket felveszi). Most 5 elemű a minta, azaz $n = 5$. A legkisebb mintaelem 36, addig függvény 0, 36-ban pedig 1/5-öt ugrik. A 40-et kétszer veszi fel, így ott 2/5-öt ugrik, vagyis 3/5-re ugrik föl, stb. Tehát így néz ki az empirikus eloszlásfüggvény:



Formulával,

$$F_5(x) = \begin{cases} 0, & \text{ha } x \leq 36, \\ \frac{1}{5}, & \text{ha } x \in (36, 40], \\ \frac{3}{5}, & \text{ha } x \in (40, 42], \\ \frac{4}{5}, & \text{ha } x \in (42, 45], \\ 1, & \text{ha } x > 45. \end{cases}$$

A mintaátlag az

$$\overline{x}_5 = \frac{1}{5}\sum_{i=1}^{5} x_i = \frac{1}{5}(40 + 45 + 40 + 42 + 36) = 40,6$$

az empirikus szórásnégyzet

$$v_5 = \frac{1}{5}\sum_{i=1}^{5}(x_i - \overline{x}_5)^2 = 8,64,$$

míg a korrigált empirikus szórásnégyzet

$$v_5^* = \frac{1}{4}\sum_{i=1}^{5}(x_i - \overline{x}_5)^2 = 10,8.$$

A medián a középső mintaelem. Most 5 mintaelem van, ennek tényleg van közepe, így a medián értéke 40. Ha páros a mintaelemszám, akkor a középső két elemnek kell a számtani közepét venni. □

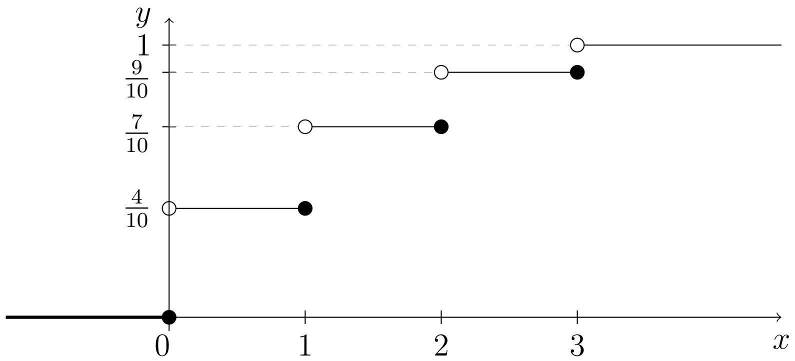
3.1.2. Egy almáskertben véletlenszerűen, egymástól függetlenül találhatók fertőzött fák. Tíz egyforma nagy, egyenként három sorból álló ültetvényben rendre 0, 3, 0, 1, 0, 0, 2, 1, 1, 2 beteg fát találtak.

(a) Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot és az empirikus szórásnégyzetet!

(b) Tegyük fel, hogy a beteg fák száma Poisson-eloszlást követ. Adjunk maximum likelihood becslést és momentumbecslést az egy sorban található fák számának várható értékére!

Megoldás. Az (a) részhez a definíciókat kell tudni. Az empirikus eloszlásfüggvény

$$F_{10}(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \frac{4}{10}, & \text{ha } x \in (0, 1], \\ \frac{7}{10}, & \text{ha } x \in (1, 2], \\ \frac{9}{10}, & \text{ha } x \in (2, 3], \\ 1, & \text{ha } x > 3. \end{cases}$$



A mintaátlag

$$\overline{x}_{10} = \frac{1}{10}\sum_{i=1}^{10} x_i = 1,$$

és a korrigálatlan és korrigált empirikus szórásnégyzet

$$v_{10} = \frac{1}{10}\sum_{i=1}^{10}(x_i - \overline{x}_{10})^2 = 1$$

$$v_{10}^* = \frac{1}{9}\sum_{i=1}^{10}(x_i - \overline{x}_{10})^2 = \frac{10}{9}.$$

(b) Tegyük fel, hogy az adatok Poisson-eloszlásból jönnek, ahol a paraméter $\lambda > 0$ ismeretlen. A megadott realizáció valószínűsége, a függetlenség és a Poisson-eloszlás definícióját felhasználva

$$\begin{aligned}
&\mathbf{P}(\xi_1 = 0, \xi_2 = 3, \xi_3 = 0, \xi_4 = 1, \xi_5 = 0, \xi_6 = 0,\\
&\quad \xi_7 = 2, \xi_8 = 1, \xi_9 = 1, \xi_{10} = 2)\\
&= \prod_{i=1}^{10} \mathbf{P}(\xi_i = x_i) = \prod_{i=1}^{10} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\\
&= \left(\prod_{i=1}^{10} \frac{1}{x_i!}\right) \lambda^{\sum_i x_i} e^{-10\,\lambda}\\
&= \left(\prod_{i=1}^{10} \frac{1}{x_i!}\right) \lambda^{10} e^{-10\,\lambda}.
\end{aligned}$$

Ez annak a valószínűsége, hogy ha adott λ-paraméterű Poisson-eloszlásból jönnek az adataink, akkor a konkrét realizáció bekövetkezik. Fordítsuk meg! Tudjuk, hogy a konkrét realizáció bekövetkezett (persze, megszámoltuk a fákat). Melyik λ paraméter a legvalószínűbb? Az, amelyre a fenti valószínűség a legnagyobb. Ez a *maximum likelihood* módszer. Tehát a

$$\lambda^{10}e^{-10\lambda}$$

kifejezés maximumát kell megtalálnunk, ahol $\lambda > 0$. Ez pont ott maximális, ahol a logaritmusa maximális, azaz a

$$\ell(\lambda) = 10\log\lambda - 10\lambda$$

függvény. Ezt lederiválva

$$\ell'(\lambda) = \frac{10}{\lambda} - 10.$$

Látjuk, hogy a függvény monoton nő $\lambda = 1$-ig, ott maximuma van, utána csökken. Tehát a paraméter ML becslése

$$\widehat{\lambda}_{ML} = 1.$$

Ez persze éppen a $\overline{x}_{10}$ mintaátlag. A levezetés ugyanaz, mint az általános esetben az előadásjegyzetben.

A momentumbecslésnél azt a paraméter választjuk, amelyhez tartozó elméleti várható érték megegyezik a mintaátlaggal. Na de a Poisson várható értéke éppen λ, tehát a momentumbecslés éppen a mintaátlag lesz,

$$\widehat{\lambda}_m = 1.$$

A momentumbecslés és a ML becslés sok ismert eloszlás esetében megegyezik. Ilyen a Poisson, exponenciális, normális. Az egyenletes eloszlásnál a kettő nem ugyanaz. $\square$

3.1.3. Legyenek $\xi_1, \ldots, \xi_n$ független azonos eloszlású véletlen változók véges μ várható értékkel és $\sigma > 0$ szórással. Igazoljuk, hogy $\hat{\mu}_1 = \xi_1$ torzítatlan becslése a várható értéknek! Igazoljuk, hogy $\hat{\mu}_2 = (\xi_1 + \xi_2)/2$ is torzítatlan becslése a várható értéknek!

Megoldás. Az, hogy a becslés torzítatlan azt jelenti, hogy a várható értéke megegyezik a becsült paraméterrel. Ez teljesül, hiszen

$$\mathbf{E}(\hat{\mu}_1) = \mathbf{E}(\xi_1) = \mu,$$

és

$$\mathbf{E}(\hat{\mu}_2) = \mathbf{E}\left(\frac{\xi_1 + \xi_2}{2}\right) = \mu.$$

Azaz mindkét becslés torzítatlan. $\square$

3.1.4. Egy alkatrészekből álló sokaság 6 mintapéldányának következő volt a hónapokban mért teljes élettartama: 39, 45, 67, 50, 50, 60. Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot, és a korrigált/korrigálatlan empirikus szórásnégyzetet!

3.1.5. Egy adott típusú izzó élettartamára öt mérés alapján a következő adataink vannak: 2,3, 4, 1,7, 3,2, 2,8.

(a) Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot és az empirikus szórásnégyzetet!

(b) Tegyük fel, hogy a háttéreloszlás exponenciális ismeretlen paraméterrel. Adjunk ML és momentumbecslést a paraméterre!

3.1.6. Egy adatszerverre a lekérdezések exponenciális időközönként érkeznek, ahol ismeretlen paraméterrel. Az időközökre percben mérve a következő adatokat kaptuk: 1,94, 0,33, 2,51, 5,27, 1,73, és 0,61. Adjunk becslést a paraméterre a maximum likelihood és a momentumbecslés alkalmazásával!

Megoldás. Folytonos eloszlás esetén a maximum likelihood becslésnél a sűrűségfüggvények szorzatát kell venni, és azt maximalizálni az adott realizációra. A λ-paraméterű exponenciális sűrűségfüggvénye $\lambda e^{-\lambda x}$. A likelihood függvény ezek szorzata, mint λ függvénye a realizációt beírva az x-ek helyére, azaz ha $x_1, \ldots, x_n$ az adott realizáció, akkor

$$L_{x_1,\ldots,x_n}(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}.$$

Ezt kell maximalizálni λ-ban. Véve a logaritmust (ez azért jó, mert a logaritmus szigorúan monoton függvény, ezért a maximumhely nem változik, viszont a nagy szorzatból nagy összeget csinál, amit könnyebb deriválni),

$$\ell(\lambda) = \log L_{x_1,\ldots,x_n}(\lambda) = n \log \lambda - \lambda \sum_{i=1}^{n} x_i.$$

Deriváljuk

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i.$$

Ez pozitív, ha $\lambda < 1/\overline{x}_n$, és negatív ha $\lambda > 1/\overline{x}_n$, tehát

$$\hat{\lambda}_{ML} = \frac{1}{\overline{x}_n} = \frac{n}{\sum_{i=1}^{n} x_i}$$

maximumhely. Ez a paraméter maximum likelihood becslése.

A momentumbecslés egyszerűbb. Egy λ-paraméterű exponenciális eloszlás várható értéke $1/\lambda$. A momentumbecslésnél azt a paramétert választjuk, amihez tartozó elméleti várható érték megegyezik a mintaátlaggal. Azaz keressük az a λ paramétert, melyhez tartozó várható érték, ami $1/\lambda$ éppen az $\overline{x}_n$ mintaátlag. Vagyis

$$\hat{\lambda}_m = \frac{n}{\sum_{i=1}^{n} x_i}.$$

Ez ugyanaz, mint a ML becslés. Most ez

$$\hat{\lambda}_{ML} = \hat{\lambda}_m = 2{,}065.$$

□

3.1.7. Legyenek $\xi_1, \ldots, \xi_n$ független véletlen változók, $f(x) = \frac{2x}{3\theta^2}$, $\theta \leq x \leq 2\theta$, sűrűségfüggvénnyel. Adjunk becslést θ-ra momentum módszerrel és ML módszerrel is!

Megoldás. Először megadjuk a paraméter momentum becslését. Ehhez a várható értéket kell meghatározni. Definíció szerint

$$\begin{aligned}\mathbf{E}_\theta(\xi) &= \int_{-\infty}^{\infty} x f_\theta(x) \mathrm{d}x = \int_{\theta}^{2\theta} x \frac{2x}{3\theta^2} \mathrm{d}x \\ &= \frac{2}{3\theta^2} \left[\frac{x^3}{3} \right]_{x=\theta}^{x=2\theta} = \frac{14\theta}{9}.\end{aligned}$$

Tehát úgy kell választani a $\hat{\theta}$ becslést, hogy a hozzá tartozó elméleti várható érték éppen a mintaátlag legyen, azaz

$$\frac{14\hat{\theta}}{9} = \overline{\xi}_n,$$

azaz a paraméter momentum becslése

$$\hat{\theta} = \frac{9}{14} \overline{\xi}_n.$$

Térjünk rá a maximum likelihood becslésre. A minta együttes sűrűségfüggvénye (ne feledkezzünk el az intervallumról, hiszen pont az függ a paramétertől)

$$f_\theta(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_\theta(x_i) = \frac{2^n}{3^n \theta^{2n}} \prod_{i=1}^{n} x_i I(\theta \leq x_i \leq 2\theta).$$

Rögzített $x_1, \ldots, x_n$ realizáció esetén ennek kell meghatározni a maximumhelyét, mint θ függvénye, azaz

$$L(\theta) = \frac{2^n}{3^n \theta^{2n}} \prod_{i=1}^{n} x_i I(\theta \le x_i \le 2\theta).$$

Ezt egy kicsit nézegessük, ne deriváljunk ész nélkül. Látjuk, hogy minél kisebb θ annál nagyobb θ^{-2n}, a többi meg konstans. Na de figyelni kell, hogy θ olyan, hogy $\theta \le \min x_i$, persze, hiszen minden mintaelem legalább θ, és $2\theta \ge \max x_i$, hiszen minden mintaelem legfeljebb 2θ. Tehát a legkisebb olyan θ paraméter kell, ami (1) $\theta \le \min x_i$ és (2) $\theta \ge \max x_i/2$. Összegezve a maximum likelihood becslés

$$\widehat{\theta}_{ML} = \frac{1}{2} \max_{1 \le i \le n} x_i.$$

□

3.1.8. Családok jövedelmét egy olyan skálán mérjük, ahol $\xi = 1$ a létminimumnak felel meg. Feltételezzük, hogy a jövedelem eloszlása $f(x) = \frac{\theta}{x^{\theta+1}}$, $x \ge 1$, sűrűségfüggvénnyel adható meg. Adjunk maximum likelihood becslést θ-ra, ha 10 véletlenszerűen választott család jövedelme: 1,53, 2,76, 19,65, 4,16, 7,31, 1,21, 254,2, 5,43, 1,12, 1,63.

3.1.9. Augusztusban 5 éjszakán át figyeltük meg a hullócsillagok számát. A következő mintát kaptuk: 4, 3, 7, 2, 4. A hullócsillagok száma egy este Poisson-eloszlású. Adjunk ML becslést az eloszlás paraméterére!

3.1.10. Egy céllövő ismeretlen p valószínűséggel talál el egy célpontot. Adjunk ML-becslést p-re, ha az első sikeres lövés k-adikra következett be. Torzítatlan-e a kapott becslés? A második sikeres lövésre további ℓ lövésig kellett várni. Ezt figyelembe véve adjunk ML-becslést p-re!

3.1.11. A müncheni Oktoberfesten minden évben az első hordó sört München főpolgármestere nyitja ki. Ez úgy történik, hogy a hordóba egy sörcsapolót üt bele az Oberbürgermeister. Tegyük fel, hogy minden ütés után p valószínűséggel indul meg a sör. Dieter Reiter főpolgármesternek 2018. szeptember 22-én a 3. ütésre sikerült kinyitni a hordót. Ezek alapján adjunk maximum likelihood becslést p-re!

3.1.12. Egy gyárban a termékek minőségét úgy ellenőrzik, hogy minden nap n terméket vizsgálnak meg. Az adott napi összes gyártmányt akkor fogadják el, ha minden megvizsgált gyártmány jó. Azt tapasztalták, hogy m nap alatt összesen x-szer fogadták el a napi gyártmányokat. Adjunk maximum likelihood becslést annak a valószínűségére, hogy egy termék selejtes!

3.1.13 (∗). Mendel törvényei szerint egy növény AA, AB, BB genotípusa rendre θ^2, $2\theta(1-\theta)$, $(1-\theta)^2$ arányban fordul elő. Egy területen a három gyakoriságra x_{AA}, x_{AB}, és x_{BB} adódott. Adjunk maximum likelihood becslést θ-ra! Mutassuk meg, hogy a kapott becslés torzítatlan!

Megoldás. Legyen n a megvizsgált növények száma, és jelölje ξ az AA genotípusúak, η az AB genotípusúak számát. Ekkor persze a BB genotípusúak száma $n-\xi-\eta$. Ha θ a valódi paraméter, akkor

$$\mathbf{P}_\theta(\xi = k, \eta = \ell) = \binom{n}{k}\binom{n-k}{\ell}(\theta^2)^k(2\theta(1-\theta))^\ell((1-\theta)^2)^{n-k-\ell},$$

$k+\ell \le n$, azaz (ξ,η) trinomiális eloszlású. Valóban, az n növény közül először kiválasztom a k db AA-t, aztán a maradék $n-k$ közül az ℓ db AB-t, a többi pedig BB. Az AA valószínűsége θ^2, az AB-é $2\theta(1-\theta)$, míg a BB-é $(1-\theta)^2$. Nyilván, ha arányokat számolunk, akkor k/n, ℓ/n a lehetséges értékek. Azaz ha ξ_r és η_r jelöli a megfelelő arányokat, akkor

$$\mathbf{P}_\theta(\xi_r = k/n, \eta_r = \ell/n) = \binom{n}{k}\binom{n-k}{\ell}(\theta^2)^k(2\theta(1-\theta))^\ell((1-\theta)^2)^{n-k-\ell},$$

$k+\ell \le n$. Adott x_{AA} és x_{AB} realizáció esetén $k = nx_{AA}$ és $\ell = nx_{AB}$ értékeket írva

$$\begin{aligned}
&\mathbf{P}_\theta(\xi_r = x_{AA}, \eta_r = x_{AB})\\
&= \binom{n}{nx_{AA}}\binom{n-nx_{AA}}{nx_{AB}}(\theta^2)^{nx_{AA}}(2\theta(1-\theta))^{nx_{AB}}((1-\theta)^2)^{n-nx_{AA}-nx_{AB}}\\
&= \binom{n}{nx_{AA}}\binom{n-nx_{AA}}{nx_{AB}}2^{nx_{AB}}\left[\theta^{x_{AA}+x_{AB}/2}(1-\theta))^{x_{AB}/2+x_{BB}}\right]^{2n}.
\end{aligned}$$

Ezt kell maximalizálni θ-ban. Látjuk, hogy elég a szögletes zárójelen belüli értéket maximalizálni θ-ban, hiszen a szorzó nem függ θ-tól és az n-edik hatvány szigorúan monoton. Felhasználva, hogy $x_{AA}+x_{AB}+x_{BB} = 1$, a maximalizálandó függvény

$$h(\theta) = \theta^u(1-\theta)^{1-u},$$

ahol $u = x_{AA} + x_{AB}/2$. Deriváljuk h-t

$$h'(\theta) = \left(\frac{u}{\theta} - \frac{1-u}{1-\theta}\right)\theta^u(1-\theta)^{1-u}.$$

Látjuk, hogy $h'(\theta) = 0$ akkor és csak akkor ha $\theta = u$, és ez valóban maximum hely. Azaz a θ ML becslése

$$\widehat{\theta}_{ML} = x_{AA} + \frac{x_{AB}}{2}.$$

Az, hogy ez a becslés torzítatlan azt jelenti, hogy

$$\mathbf{E}(\xi + \frac{1}{2}\eta) = n\theta.$$

Az együttes eloszlás nem is kell ehhez. Nyilván ξ és η is binomiális eloszlásúak, n és θ^2, illetve n és $2\theta(1-\theta)$ paraméterekkel. Tehát a várható érték

$$\mathbf{E}(\xi + \frac{1}{2}\eta) = \mathbf{E}(\xi) + \frac{1}{2}\mathbf{E}(\eta) = n\theta^2 + \frac{1}{2}n2\theta(1-\theta) = n\theta,$$

amint állítottuk. □

3.1.14. A színtévesztés leggyakoribb fajtája az X kromoszómához kapcsoltan, nemhez kötötten öröklődik. Emiatt a férfiaknál a gyakoriság p, míg a nőknél csak p^2. Adjunk maximum likelihood becslést p-re, az alapján, hogy M férfiból m, N nőből pedig n volt színtévesztő!

3.2. Konfidenciaintervallumok, próbák

3.2.1. Egy játék fizikai terhelhetőségére elvégzett tesztek kg-ban a következő eredményeket adták: 40, 45, 40, 42, 36. Tegyük fel, hogy a háttérváltozó normális eloszlást követ 2 szórással. Adjunk 95%-os konfidenciaintervallumot a várható értékre!

Megoldás. Tegyük fel, hogy a háttéreloszlás normális, ismeretlen μ várható értékkel, és ismert $\sigma = 2$ szórással. (Figyeljünk oda, hogy ez a szórás, és nem a szórásnégyzet.) Az, hogy az adatok normális eloszlásúak egy teljesen természetes feltevés, ami általában teljesül. A centrális határeloszlás-tétel miatt van ez. A statisztikai probléma az az, hogy határozzuk meg a várható értéket. A szórás ismerete bizonyos esetekben indokolható: méréseknél ismerjük a mérőeszköz tulajdonságait. Nekünk az ismert szórás azért kényelmes, mert ez a legegyszerűbb eset.

Tegyük fel, hogy $\xi_1, \xi_2, \ldots, \xi_n$ független normális véletlen változók ismeretlen μ várható értékkel és $\sigma = 2$ szórással. Persze az ismeretlen μ várható értékre a becslésünk a mintaátlag, azaz

$$\widehat{\mu} = \overline{\xi}_n.$$

Ehhez nem kellett volna ennyi mindent tanulni. A kérdés az, hogy ez mennyire pontos. Adjunk meg egy olyan intervallumot, amire teljesül az, hogy az *igazi paraméter* 0,95 valószínűséggel beleesik ebbe az intervallumba. Mivel a normális eloszlás szimmetrikus a várható értékére, ezért természetes az intervallumot $[\widehat{\mu}-\delta, \widehat{\mu}+\delta]$ alakban keresni, ahol $\delta > 0$. Közben tartsuk észben, hogy ez az intervallum a véletlentől függ, hiszen $\widehat{\mu}$ egy véletlen mennyiség. Tehát úgy szeretnénk megválasztani δ értékét, hogy a $\mu \in [\widehat{\mu}-\delta, \widehat{\mu}+\delta]$ esemény valószínűsége 0,95 legyen. Átírva

$$\widehat{\mu}-\mu = \frac{1}{n}\sum_{i=1}^{n}\xi_i - \mu = \frac{1}{n}\sum_{i=1}^{n}(\xi_i-\mu) = \frac{\sigma}{n}\sum_{i=1}^{n}\frac{\xi_i-\mu}{\sigma}.$$

Az $(\xi_i-\mu)/\sigma$ változó normális eloszlású, a várható értéke $\mu-\mu=0$, szórása $\sigma/\sigma=1$. Azaz ő standard normális. Ezt láttuk a normális eloszlás tulajdonságainál. Azt kell felhasználnunk, hogy *független normális eloszlású véletlen változók összege normális eloszlású.* Ezt nem bizonyítjuk sem itt, sem az előadáson. Mivel függetlenek, várható értékük és szórásnégyzetük is összeadódik. Ezt az állítást bizonyítottuk tetszőleges változókra. Összegezve

$$\widehat{\mu}-\mu$$

véletlen változó normális eloszlású, várható értéke 0, szórásnégyzete pedig σ^2/n. Ezek szerint

$$\frac{\sqrt{n}}{\sigma}(\widehat{\mu}-\mu)$$

standard normális véletlen változó. Tehát

$$\begin{aligned}\mathbf{P}(|\widehat{\mu}-\mu| \le \delta) &= \mathbf{P}\left(\frac{\sqrt{n}}{\sigma}(\widehat{\mu}-\mu) \in [-\delta\sqrt{n}/\sigma, \delta\sqrt{n}/\sigma]\right)\\ &= \Phi(\delta\sqrt{n}/\sigma) - \Phi(-\delta\sqrt{n}/\sigma)\\ &= 2\Phi(\delta\sqrt{n}/\sigma) - 1.\end{aligned}$$

Ez utóbbi kell 0,95 legyen, azaz

$$\Phi(\delta\sqrt{n}/\sigma) = 0,975.$$

A normális eloszlás táblázatából kapjuk, hogy

$$\frac{\delta\sqrt{n}}{\sigma} = \Phi^{-1}(0,975) = 1,96$$

vagyis

$$\delta = \frac{\sigma}{\sqrt{n}}1,96 = 1,75.$$

Most használjuk csak a konkrét adatainkat. Eszerint $n=5$ és $\overline{x}_n = 40,6$, ezért a 95%-os konfidenciaintervallum [40,6−1,75, 40,6+1,75] = [38,85, 42,35].

Tehát azt kaptuk, hogy a valódi várható érték, ami determinisztikus de ismeretlen, 95% valószínűséggel beleesik a [38,85, 42,35] (véletlen!) intervallumba.

Ez ugyanaz a levezetés, mint az előadásjegyzetben a konfidenciaintervallum résznél, csak konkrét számokkal. □

3.2.2. Egy véletlen változó értékeit megfigyelve a következő statisztikai mintát kapjuk: 6,5, 7,3, 5,4, 6,5, 2,1. Tegyük fel, hogy a háttéreloszlás normális, ismert $\sigma = 2$ szórással és ismeretlen μ várható értékkel. Teszteljük 5%-os szignifikanciaszinten azt a nullhipotézist, hogy az elméleti várható érték 8.

Megoldás. Egy u-próbát kell végrehajtanunk, hiszen normális eloszlás várható értékét teszteljük ismert szórás mellett. A fenti minta alapján becsült várható érték, a mintaátlag $\overline{x}_5 = 5,56$. A kérdés, hogy elhisszük-e azt a feltevést, hogy a valódi várható érték 8? Persze gyanús a dolog, de ha a valódi várható érték tényleg 8 lenne, akkor sem várnánk, hogy a mintaátlag egy 5 elemű minta esetén éppen 8.

A nullhipotézis az amit felteszünk. Jelen esetben az, hogy a valódi várható érték $\mu_0 = 8$. Az alternatív-, vagy ellenhipotézis, hogy $\mu_0 \neq 8$. A próbastatisztikát már láttuk a konfidenciaintervallum konstruálásánál is,

$$u = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma} = -2,72.$$

Ha a nullhipotézis igaz, akkor ez az u statisztika standard normális eloszlású. Ezt is láttuk a konfidenciaintervallumnál. Tehát azt kell eldönteni, hogy elhisszük-e egy standard normális eloszlású véletlen változóról, hogy értéke $-2,72$. Ehhez választunk egy α szignifikanciaszintet, ez most 0,05. Keressük azt az u_α küszöbértéket, amelyre teljesül, hogy $\mathbf{P}(|Z| > u_\alpha) = \alpha$, ahol Z standard normális. A normális eloszlás szimmetriája miatt

$$\mathbf{P}(|Z| > u_\alpha) = \mathbf{P}(Z > u_\alpha) + \mathbf{P}(Z < -u_\alpha) = 2(1 - \Phi(u_\alpha)),$$

tehát $u_\alpha = \Phi^{-1}(1 - \alpha/2)$. Esetünkben $u_{0,05} = \Phi^{-1}(0,975) = 1,96$. Mivel $|u| = 2,72 > 1,96$, ezért ha igaz lenne a nullhipotézis, akkor a bekövetkezett esemény valószínűsége 0,05-nél kisebb, vagyis egy valószínűtlen esemény következett be. Ekkor 5%-os szignifikanciaszinten *elvetjük* a nullhipotézist. Azt mondjuk, hogy az elfogadási tartomány a $[-1,96,\ 1,,96]$ intervallum. Ha a próbastatisztika értéke ide esik, elfogadjuk a nullhipotézist, különben elvetjük. □

3.2.3. A Liverpool 2020. február 18. - március 11. között lejátszott 6 mérkőzésén a Liverpool labdabirtoklása (%-ban kifejezve) 71, 68, 68, 58, 70,

72. Tegyük fel, hogy a labdabirtoklás %-os értéke normális eloszlást követ ismeretlen μ várható értékkel és ismert $\sigma = 5$ szórással. Teszteljük 5%-os szignifikanciaszinten azt a nullhipotézist, hogy a $\mu = 65$.

3.2.4. Egy játék fizikai terhelhetőségére elvégzett tesztek kg-ban a következő eredményeket adták: 40, 45, 40, 42, 36, 41, 43. A 2019. január 1. óta hatályos EU-s ISO-31415 szabvány szerint a gyermekjátékoknak terhelhetősége legalább 40 kg kell legyen. Tegyük fel, hogy a terhelhetőség normális eloszlást követ ismeretlen μ várható értékkel és ismert $\sigma = 3.1$ szórással. Teszteljük 1%-os szignifikanciaszinten azt a nullhipotézist, hogy a $\mu = 40$.

3.2.5. Az FC Barcelona 2019. április 16. - május 1. között lejátszott 5 mérkőzésén a Barcelona labdabirtoklása (%-ban kifejezve) 47, 62, 77, 50, 66.

(a) Adjuk meg a mintához tartozó mintaátlagot, a tapasztalati szórásnégyzetet és az empirikus eloszlásfüggvényt!

(b) Tegyük fel, hogy a labdabirtoklás %-os értéke normális eloszlást követ ismeretlen μ várható értékkel és ismert $\sigma = 10$ szórással. Teszteljük 5%-os szignifikanciaszinten azt a nullhipotézist, hogy a $\mu = 60$.

3.2.6. Az FC Barcelona 2019. április 16. - május 1. között lejátszott 5 mérkőzésén a Barcelona rúgott góljainak száma 3, 1, 2, 2, 3.

(a) Adjuk meg a mintához tartozó mintaátlagot, a tapasztalati szórásnégyzetet és az empirikus eloszlásfüggvényt!

(b) Tegyük fel, hogy a lőtt gólok száma Poisson-eloszlást követ, ismeretlen λ paraméterrel. Adjunk ML és momentumbecslést λ értékére!

3.2.7. A hároméves Máté reggelente véletlen időpontban ébred. Az elmúlt 5 napban az ébredésekre 6:20, 6:55, 6:30, 7:15, és 6:40 adódott.

(a) Adjuk meg a mintához tartozó mintaátlagot, a tapasztalati szórásnégyzetet és az empirikus eloszlásfüggvényt!

(b) Tegyük fel, hogy Máté ébredésének időpontja normális eloszlást követ ismeretlen μ várható értékkel és $\sigma = 20$ szórással. Adjunk 95%-os konfidenciaintervallumot az ébredés várható értékére!

Hivatkozások

[1] Bognár Jánosné, Mogyoródi József, Prékopa András, Rényi Alfréd, Szász Domokos: *Valószínűségszámítási feladatgyűjtemény.* Negyedik kiadás, Typotex 2001.

[2] William Feller: Bevezetés a valószínűségszámításba és alkalmazásaiba I.