**A március 23.-i gyakorlat feladatai**

Tekintünk néhány feladatot, amelyek független eseményekkel foglalkoznak. Előtte azonban felelevenítünk néhány fontos fogalmat és eredményt.

**Két esemény függetlenségének a definiciója.** *Azt mondjuk, hogy egy $A$ és $B$ esemény független, ha $P(A \cap B) = P(A)P(B)$.*

**Több esemény függetlenségének definiciója:** *Az $A_1, \dots, A_n$ események akkor (teljesen) függetlenek, ha az $\{1, \dots, n\}$ indexhalmaz minden $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$ részhalmazára*

$$P(A_{j_1} \cap \cdots \cap A_{j_k}) = P(A_{j_1}) \cdots P(A_{j_k}).$$

*Események végtelen $A_1, A_2, \dots$ sorozata akkor és csak akkor független, ha tetszőleges pozitív $n$ egész számra az $A_1, \dots, A_n$ események függetlenek.*

Speciálisan $n = 3$ esetben ez a definició a következőt jelenti: Az $A$, $B$ és $C$ események akkor függetlenek, ha a következő azonosságok *mindegyike* teljesül:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \\ P(A \cap B \cap C) &= P(A)P(B)P(C). \end{aligned}$$

*Megjegyzés:* Alább bevezetjük események páronkénti függetlenségének az irodalomban szintén használt fogalmát. Fontos, hogy események páronkénti függetlenségének és (teljes) függetlenségének a fogalmát meg tudjuk különböztetni egymástól.

**Események páronkénti függetlenségének a definiciója.** *Legyen $A_1, A_2, \dots,$ események (véges vagy végtelen) sorozata egy valószínűségi mezőn. Azt mondjuk, hogy ezek az események páronként függetlenek, ha tetszőleges különböző számokból álló $(j, k)$, $j \neq k$, indexekre $P(A_j \cap A_k) = P(A_j)P(A_k)$.*

Megmutatandó, hogy az események (teljes) függetlenségének definiciójában szereplő feltételek mindegyike fontos tekintsük a következő feladatot.

1.) Adjunk példát egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőre azon három $A_1$, $A_2$ és $A_3$ eseményre, amelyekre $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$, de az $A_1$, $A_2$ és $A_3$ események nem függetlenek.
*Egy lehetséges konstrukció:* Legyen $\Omega = \{1, 2, 3, 4, 5\}$, $\mathcal{A}$ az $\Omega$ halmaz összes részhalmazaiból álló $\sigma$-algebra, $P(\{1\}) = x$, $P(\{2\}) = P(\{3\}) = P(\{4\}) = y$, $P(\{5\}) = 1 - x - 3y$, alkalmas $x$ és $y$ számokkal, $P(A) = \sum\limits_{u \in A} P(\{u\})$ minden $A \in \mathcal{A}$ halmazra. Definiáljuk az $A_1 = \{1, 2\}$, $A_2 = \{1, 3\}$ és $A_3 = \{1, 4\}$ halmazokat. Ekkor $A_1 \cap A_2 \cap A_3 = \{1\}$, ezért $P(A_1 \cap A_2 \cap A_3) = x$. Másrészt $P(A_1) = P(A_2) = P(A_3) = x + y$. Válasszuk meg az $x$ és $y$ számokat úgy, hogy $x = (x + y)^3$. Egy lehetőség erre, $x + y = \frac{1}{3}$, és ekkor $x = (x + y)^3 = \frac{1}{27}$, $y = \frac{8}{27}$,

továbbá $P(\{5\}) = \frac{2}{27}$. Ebben a példában $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$. Viszont $A_1 \cap A_2 = \{1\} = A_1 \cap A_2 \cap A_3$, így nyílván $P(A_1 \cap A_2) \neq P(A_1)P(A_2)$. Tehát a függetlenség nem teljesül.

2.) Definiálunk egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt, azon három $A_1$, $A_2$, $A_3$-mal jelölt eseményt, amelyek páronként függetlenek, azaz $P(A_1 \cap A_2) = P(A_1)P(A_2)$, $P(A_1 \cap A_3) = P(A_1)P(A_3)$ és $P(A_2 \cap A_3) = P(A_2)P(A_3)$, de nem teljesül a $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ azonosság, tehát ezek az események nem függetlenek.
*Megoldás:* Álljon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőben $\Omega$ 4 pontból, a jobb szemlélhetőség kedvéért legyenek ezek az $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$ pontok, álljon $\mathcal{A}$ az $\Omega$ halmaz összes részhalmazából, és legyen $P(\{(1,1)\}) = P(\{(1,2)\}) = P(\{(2,1)\}) = P(\{(2,2)\}) = \frac{1}{4}$. Tekintsük az $A_1 = \{(1,1),(1,2)\}$ $A_2 = \{(1,1),(2,1)\}$ és $A_3 = \{(1,1),(2,2)\}$ halmazokat. Ekkor teljesülnek a $P(A_1 \cap A_2) = P(A_1)P(A_2)$, $P(A_1 \cap A_3) = P(A_1)P(A_3)$ és $P(A_2 \cap A_3) = P(A_2)P(A_3)$ azonosságok, mert $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$, és mivel $A_1 \cap A_2 = A_1 \cap A_2 = A_2 \cap A_3 = \{(1,1)\}$, $P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}$. Másrészt, $P(A_1 \cap A_2 \cap A_3) \neq P(A_1)P(A_2)P(A_3)$, mert $P(A_1 \cap A_2 \cap A_3) = P(\{(1,1)\}) = \frac{1}{4}$, és $P(A_1)P(A_2)P(A_3) = \frac{1}{8}$.

**Lemma.** *Ha $A_1, \dots, A_n$ független események, és bevezetjük az $A_j^1 = A_j$ és $A_j^{-1} = \Omega \setminus A_j$ jelöléseket, akkor tetszőleges $\varepsilon_j = \pm 1$, $1 \le j \le n$, sorozatra az $A_1^{\varepsilon_1}, \dots, A_n^{\varepsilon_n}$ események függetlenek.*

**Lemma.** *Legyenek $A_1, \dots, A_n, B_1, \dots, B_m$ események függetlenek egymástól. Tetszőleges olyan $C$ eseményre, amelyik előállítható az $A_1, \dots, A_k$ halmazokból véges sok metszet, unió és komplementerképzés segítségével igaz, hogy a $B_1, \dots, B_m$ és $C$ halmazok függetlenek.*

3.) Egy 100 tagú társaság minden egyes tagja egymástól függetlenül $\frac{1}{1000}$ valószínűséggel betegszik meg. Mi annak a valószínűsége, hogy a társaságnak lesz beteg tagja? A kapott eredményről mit mondhatunk? Az nagyon nagy, (majdnem 1) vagy nagyon kicsi (majdnem nulla) érték?
*Megoldás:* Jelölje $A_j$ azt az eseményt, hogy a társaság $j$-ik megbetegszik meg. Ekkor a $P(A_j) = \frac{1}{1000}$, és az $A_j$ események függetlenek. Számoljuk ki először a minket érdeklő esemény komplementerének, azaz annak az eseménynek a valószínűségét, hogy mindenki egészséges. Ez az $\bigcap_{j=1}^{1000} (\Omega \setminus A_j)$ esemény. Mivel $P(\Omega \setminus A_j) = 1 - \frac{1}{1000}$, és az $A_j$ események függetlenségéből következik az $\Omega \setminus A_j$ események függetlensége is, ezért $P\left(\bigcap_{j=1}^{1000} (\Omega \setminus A_j)\right) = \left(1 - \frac{1}{1000}\right)^{100}$. Innen a minket érdeklő

esemény valószínűsége $1-\left(1-\frac{1}{1000}\right)^{100}$.

Végül jegyezzük meg, hogy $\left(1-\frac{1}{1000}\right)^{100}=\left(\left(1-\frac{1}{1000}\right)^{1000}\right)^{1/10}\sim e^{-1/10}$. Miért? Ez a szám nagyon közel van az egyhez, ezért a minket érdeklő valószínűség értéke kicsi.

4.) Tekintsük a következő valószínűségi mezőt. $\Omega=\{1,\dots,n\}$, ahol $n$ valamely pozitív egész szám, $\mathcal{A}$ az $\Omega$ összes részhalmazaiból álló $\sigma$-algebra,

$$P(A)=\frac{\text{az } A \text{ halmaz számossága}}{n}.$$

Legyen $n=p_1^{r_1}\cdots p_k^{r_k}$ az $n$ szám prímtényezős felbontása, és definiáljuk a következő $A_j$ eseményeket: $A_j=\{m\colon m \text{ osztható a } p_j \text{ számmal}\}$, $1\le j\le k$. Legyen $B=\{m\colon m \text{ relatív prim az } n \text{ számhoz képest}\}$. Mutassuk meg, hogy

a. Az $A_1,\dots,A_k$ események függetlenek.

b. $P(B)=\prod\limits_{j=1}^{k}\left(1-\frac{1}{p_j}\right)$, azaz összesen $n\cdot\prod\limits_{j=1}^{k}\left(1-\frac{1}{p_j}\right)$ $n$-nél kisebb és az $n$-hez képest relatív prim van.

*Megoldás:* $A_j=\left\{p_j,2p_j,\dots,\frac{n}{p_j}p_j\right\}$ egy $\frac{n}{p_j}$ számból álló halmaz, ezért $P(A_j)=\frac{1}{p_j}$, $1\le j\le k$. Az $A_{j_1}\cap A_{j_2}\cap\dots\cap A_{j_s}$ halmaz az $n$-nél kisebb $p_{j_1}\cdots p_{j_s}$ számmal osztható számokból áll minden $1\le j_1<j_2<\dots<j_s\le k$ sorozatra, ezért számossága $\frac{n}{p_{j_1}\cdots p_{j_s}}$, és $P(A_{j_1}\cap A_{j_2}\cap\dots\cap A_{j_s})=\frac{1}{p_{j_1}\cdots p_{j_s}}$. Ez azt jelenti, hogy $P(A_{j_1}\cap A_{j_2}\cap\dots\cap A_{j_s})=P(A_{j_1})P(A_{j_2})\cdots P(A_{j_s})$ minden $1\le j_1<j_2<\dots<j_s\le k$ sorozatra, ezért az $A_1$, $A_2$, $\dots$ $A_k$ halmazok függetlenek.

Végül $B=\Omega\setminus(A_1\cup\dots\cup A_k)=\bigcap\limits_{j=1}^{k}(\Omega\setminus A_j)$. Ezért és az $A_j$ események függetlensége miatt $P(B)=\prod\limits_{j=1}^{k}P(\Omega\setminus A_j)=\prod\limits_{j=1}^{k}\left(1-\frac{1}{p_j}\right)$, ahonnan következik a $B$ halmaz számosságára megadott képlet.

5.) Véletlenül meghívunk 30 embert. Tegyük fel, hogy az egyes embereknek egymástól függetlenül van születésnapjuk, és minden ember esetében $\frac{1}{365}$ annak a valószínűsége, hogy az év valamely napján született. Mi annak a valószínűsége, hogy van két ember a társaságban, akiknek ugyanaznap van a születésnapjuk?

Általánosabban, van $n$ urna, amelyekbe bedobunk egymástól függetlenül $k$ golyót úgy, hogy mindegyik golyó egyforma valószínűséggel esik az egyes urnákba. Mi annak a valószínűsége, hogy van olyan urna amelybe legalább két golyó esik? Érdekel minket továbbá ennek a valószínűségnek a viselkedése, ha mind az $n$ mind a $k$ szám nagy, és a $k=k(n)$ számnak megfelelő a nagyságrendje. Lássuk be, hogy a fenti

valószínűségnek van határértéke, ha $n \to \infty$, $\dfrac{k}{\sqrt{n}} \to \alpha$ valamilyen $0 \le \alpha < \infty$ számmal, és határozzuk meg ezt a határértéket.
*Megoldás:* Jelölje $\xi_j$ azt a valószínűségi változót, hogy a $j$-ik embernek az év hanyadik napján van a születésnapja. Ekkor a $\xi_j$, $1 \le j \le 30$, valószínűségi változók függetlenek, $P(\xi_j = l) = \dfrac{1}{365}$, $1 \le j \le 30$, $1 \le l \le 365$, és $P(\xi_j \neq \xi'_j$ ha $j \neq j')$ annak a valószínűsége, hogy mindenkinek különböző nap van a születésnapja. Ez a valószínűség viszont

$$\frac{365 \cdot 364 \cdots (365-30+1)}{365^k} = \prod_{j=1}^{29} \left(1 - \frac{j}{365}\right),$$

mert annak valószínűlége, hogy az első ember születésnapja az $l_1$-ik, a másodiké az $l_2$-ik és így tovább a $k$-ik ember születésnapja az $l_k$-ik napon van $\dfrac{1}{365^k}$, tetszőleges $1 \le l_j \le 365$, $1 \le j \le 30$ számok esetén, és ezeket a számokat $\prod\limits_{j=0}^{k-1}(365-j)$ módon választhatjuk úgy, hogy mindegyik $l_j$ szám különböző legyen. Így annak a valószínűsége, hogy van két ember akinek ugyanazon a napon van a születésnapja $1 - \prod\limits_{j=1}^{29}\left(1 - \dfrac{j}{365}\right)$.
Hasonlóan, annak valószínűsége, hogy ha $k$ golyót dobunk $n$ urnába az adott módon, akkor van olyan urna, amelyikbe legalább két golyó esik $1 - \prod\limits_{j=1}^{k-1}\left(1 - \dfrac{j}{n}\right)$.
Adjunk jó közelítést a $\log \prod\limits_{j=1}^{k-1}\left(1 - \dfrac{j}{n}\right) = \sum\limits_{j=1}^{k-1} \log\left(1 - \dfrac{j}{n}\right)$ kifejezésre, ha $n \to \infty$, $\dfrac{k(n)}{\sqrt{n}} \to \alpha$. Heurisztikus érvelés szerint mivel $\log\left(1 - \dfrac{j}{n}\right) \sim -\dfrac{j}{n}$ a $\log(1+x)$ függvény Taylor sorfejtése szerint, ezért $\sum\limits_{j=1}^{k-1} \log\left(1 - \dfrac{j}{n}\right) \sim -\sum\limits_{j=1}^{k-1} \dfrac{j}{n} = -\dfrac{(k-1)k}{2n}$, ahonnan $\log \prod\limits_{j=1}^{k(n)-1}\left(1 - \dfrac{j}{n}\right) \to -\dfrac{\alpha^2}{2}$, ha $n \to \infty$, és $\dfrac{k(n)}{\sqrt{n}} \to \alpha$. Ez a számolás precizzé tehető, ha felhasználjuk például azt az egyenlőtlenséget, amely szerint

$$\left|\log\left(1 - \frac{j}{n}\right) + \frac{j}{n}\right| \le \frac{2j^2}{n^2} \le \frac{\text{const.}}{n}, \quad \text{ha } n \text{ elég nagy és } \frac{j}{\sqrt{n}} \le \alpha + 1,$$

ami szintén következik a $\log(1+x)$ Taylor sorfejtéséből. Miért? Innen kapjuk, hogy $1 - \prod\limits_{j=1}^{k(n)-1}\left(1 - \dfrac{j}{n}\right) \to 1 - e^{-\alpha^2/2}$, ha $n \to \infty$, és $\dfrac{k(n)}{\sqrt{n}} \to \alpha$.

*Házi feladat:*

Egy szabályos pénzdarabot feldobunk százszor. Számoljuk ki a három egymást követő fej-dobásból álló részsorozatok várható értékét és szórásnégyzetét.

*Házi feladat:*

Adott két urna, mind a kettőben 10 piros és 20 fehér golyó. Kihúzunk 10 alkalommal mind a két urnából egy-egy golyót, az elsőből visszatevéssel, a másodikból visszatevés nélkül. Számítsuk ki azon húzáspárok számának a várható értékét és szórásnégyzetét, amelyekben azonos színű golyót húztunk.