A március 4.-i előadáshoz kapcsolódó gyakorlat feladatai.

Feladatok:

Láttuk, hogyan lehet egyszerűen kiszámítani a $A_1 \cup \cdots \cup A_n$ alakú események valószínűségét, ha az A_j, $1 \le j \le n$, események függetlenek. Ezek a számolások természetesen kihasználták a tekintett események függetlenségét. A következőben azt tárgyaljuk meg, hogy amennyiben nincs feltétlenül függetlenség a tekintett események között, de ki tudjuk számolni az $A_{j_1} \cap \cdots \cap A_{j_s}$ alakú események valószínűségét, akkor egy úgynevezett szita formula segítségével ki tudjuk számítani az $A_1 \cup \cdots \cup A_n$ események valószínűségét is. Továbbá megmutatjuk, hogy ez lehetővé teszi érdekes feladatok megoldását. A következő feladatot fogjuk tárgyalni:

1. Egy estélyen megjelenik n házaspár. Egy táncmester, aki nem tudja, hogy kik házastársak és kik nem véletlen módon párba rendezi a férfiakat és nőket a tánc előtt. Mi a valószínűsége annak, hogy egyetlen házaspár sem táncol együtt? Mi ennek a valószínűségnek a határértéke, ha $n \to \infty$?

   Megoldás: Definiáljuk a következő A_j eseményeket:

   $$A_j = \text{a } j\text{-ik házaspár egütt táncol}, \quad 1 \le j \le n.$$

   Ekkor minket a $P(\overline{A_1 \cup \cdots \cup A_n}) = 1 - P(A_1 \cup \cdots \cup A_n)$ valószínűség érdekel. Vegyük észre, hogy a $P(A_{j_1} \cap \cdots \cap A_{j_k}) = \dfrac{(n-k)!}{n!}$ azonosság igaz. Ugyanis az összes lehetséges párbaállítások száma $n!$, míg az olyan párbaállítások száma, amelyben a j_1-ik, j_2-ik, ..., j_k-ik házaspár egy párba kerül $(n-k)!$. Továbbá érvényes a következő az irodalomban szita-formulának nevezett eredmény, amelyet külön fogunk tárgyalni.

Szita formula. *Legyenek $A_1, \dots, A_n$ tetszőleges események egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor*

$$P(A_1 \cup \cdots \cup A_n) = S_1 - S_2 + S_3 - \cdots + (-1)^{n+1} S_n,$$

ahol

$$S_k = \sum_{1 \le j_1 < \cdots < j_k \le n} P(A_{j_1} \cap \cdots \cap A_{j_k}).$$

A szita-formula segítségével kapjuk, hogy n házaspár esetében

$$S_k = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

Innen adódik, hogy a minket érdeklő valószínűség n házaspár esetén

$$\begin{aligned} P(\overline{A_1 \cup \cdots \cup A_n}) &= 1 - P(A_1 \cup \cdots \cup A_n) \\ &= 1 - S_1 + S_2 + \cdots + (-1)^n S_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}, \end{aligned}$$

és ezért

$$P(\overline{A_1 \cup \cdots \cup A_n}) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \to \frac{1}{e} \quad \text{ha } n \to \infty.$$

2. Legyenek $\xi_1, \dots, \xi_k$ független, diszkrét valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Legyenek $A_1 \subset \{x_1, x_2, \dots\}, A_2 \subset \{x_1, x_2, \dots\}, \dots, A_k \subset \{x_1, x_2, \dots\}$ tetszőleges halmazok. Ekkor

$$P(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k) = P(\xi_1 \in A_1) P(\xi_2 \in A_2) \cdots P(\xi_k \in A_k).$$

   Továbbá mutassuk meg, hogy tetszőleges $\{j_1, \dots, j_s\} \subset \{1, \dots, k\}$ indexhalmazra a $\xi_{j_1}, \dots, \xi_{j_s}$ valószínűségi változók függetlenek.

   Megoldás: A függetlenség miatt

$$\begin{aligned} &P(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k) \\ &\quad = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \cdots \sum_{x_k \in A_k} P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_k = x_k) \\ &\quad = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \cdots \sum_{x_k \in A_k} P(\xi_1 = x_1) P(\xi_2 = x_2) \cdots P(\xi_k = x_k). \end{aligned}$$

   Másrészt

$$\begin{aligned} &P(\xi_1 \in A_1) P(\xi_2 \in A_2) \cdots P(\xi_k \in A_k) \\ &\quad = \sum_{x_1 \in A_1} P(\xi_1 = x_1) \sum_{x_2 \in A_2} P(\xi_2 = x_2) \cdots \sum_{x_k \in A_k} P(\xi_k = x_k). \end{aligned}$$

   Elvégezve a beszorzásokat kapjuk a feladat első állítását.

   A második állítást megkapjuk az első speciális eseteként $A_{j_u} = \{x_{j_u}\}$, $1 \le u \le s$ és $A_j = \{x_1, x_2, \dots\}$, $u \in \{1, \dots, k\} \setminus \{j_1, \dots, j_s\}$ választással.

A következő feladat megfogalmazása előtt felidézzük az alábbi definiciót.

Halmaz indikátorfüggvényének a definiciója. *Legyen adva egy $A \in \mathcal{A}$ esemény egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Az A halmaz indikátorfüggvényén azt a $\chi_A(\omega)$ valószínűségi változót értjük, amelyre $\chi_A(\omega) = 1$, ha $\omega \in A$, és $\chi_A(\omega) = 0$, ha $\omega \notin A$.*

3. Legyenek $A_1, \dots, A_k$ események egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Az $A_1, \dots, A_k$ események akkor és csak akkor függetlenek, ha azok $\chi_{A_1}(\omega), \dots, \chi_{A_k}(\omega)$ indikátorfüggvényei függetlenek.

   Megoldás: Ha a $\chi_{A_1}(\omega), \dots, \chi_{A_k}(\omega)$ indikátorfüggvények függetlenek, akkor ezek tetszőleges részhalmaza is független az előző feladat eredménye szerint. Ezért minden $\{j_1, \dots, j_s\} \subset \{1, \dots, k\}$ indexhalmazra

$$\begin{aligned} P(A_{j_1} \cap \cdots \cap A_{j_s}) &= P\left(\chi_{A_{j_1}} = 1, \dots, \chi_{A_{j_s}} = 1\right) \\ &= P\left(\chi_{A_{j_1}} = 1\right) \cdots P\left(\chi_{A_{j_s}} = 1\right) = P(A_{j_1}) \cdots P(A_{j_s}). \end{aligned}$$

Ha az $A_1, \dots, A_k$ események függetlenek, akkor egyes A_j eseményeket azok komplementerével helyettesítve ismét független eseményeket kapunk. Felírva az összes ilyen relációt, megkapjuk azokat az azonosságokat, amelyek a $\chi_{A_1}(\omega), \dots, \chi_{A_k}(\omega)$ indikátorfüggvények függetlenségét jelentik.

4. Véletlenül meghívunk 30 embert. Tegyük fel, hogy az egyes embereknek egymástól függetlenül van születésnapjuk, és minden ember esetében $\dfrac{1}{365}$ 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 ξ_j azt a valószínűségi változót, hogy a j-ik embernek az év hanyadik napján van a születésnapja. Ekkor a ξ_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 $\sum\limits_{j=1}^{k-1} \log\left(1-\frac{j}{n}\right) \sim -\sum\limits_{j=1}^{k-1} \frac{j}{n} = -\frac{(k-1)k}{2n}$, ahonnan $\log \prod\limits_{j=1}^{k(n)-1} \left(1-\frac{j}{n}\right) \to -\frac{\alpha^2}{2}$, ha $n\to\infty$, és $\frac{k(n)}{\sqrt{n}} \to \alpha$. Ez a számolás precizzé tehető, ha felhasználjuk például azt az egyenlőtlenséget, amely szerint

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

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

5. Egy urnában F fehér és P piros golyó van. Kihúzunk véletlenül egy golyót, majd visszadobunk R, $R \ge 0$ ugyanolyan színű golyót az urnába. Minden húzás során a korábbiaktól függetlenül az urnában levő egyes golyókat egyforma valószínűséggel húzzuk ki. Lássuk be, hogy annak a valószínűsége, hogy a j-ik húzásban piros golyót húzunk megegyezik annak valószínűségével, hogy az első húzásban húzunk piros golyót. Annak valószínűsége, hogy a j-ik és k-ik húzásban húzunk piros golyót, $j \ne k$, megegyezik annak valószínűségével, hogy az első és második húzásban húzok piros golyót.

   Megoldás: Vegyük észre, hogy egy olyan húzássorozatnak a valószínűsége, amely N piros és M fehér golyót tartalmaz csak az N és M számtól függ, nem függ a különböző színű golyók húzásának a sorrendjétől. Valóban, egy ilyen húzássorozat valószínűségét fel tudjuk írni. Ez

$$P(N,M) = \frac{P(P+R-1)\cdots(P+(N-1)(R-1))}{\prod\limits_{k=0}^{N+M-1}(P+P-k(R-1))} \cdot F(F+R-1)\cdots(F+(M-1)(R-1))$$

(Miért?), és innen következik az állítás. Innen látható, hogy annak valószínűsége, hogy pontosan N piros és M fehér golyót húzunk és az első húzás eredménye piros megegyezik annak valószínűségével, hogy pontosan N piros és M fehér golyót húzunk és a j-ik húzás eredménye piros. Ugyanis mind a két valószínűség

$$\binom{N+M-1}{M} P(N,M).$$

Innen következik, hogy annak valószínűsége, hogy az első húzás piros megegyezik annak valószínűségével, hogy a j-ik húzás piros. Hasonlóan látható, hogy annak a valószínűsége, hogy a j-ik és k-ik húzás piros megegyezik annak valószínűségével, hogy az első és a második húzás eredménye piros. Ezután a keresett valószínűségeket könnyen kiszámíthatjuk. Annak valószínűsége, hogy a j-ik húzás piros $\frac{P}{P+F}$, és $\dfrac{P(P+R-1)}{(F+P)(F+P+R-1)}$ annak a valószínűsége, hogy a j-ik és k-ik húzás piros, $j \neq k$,

Megjegyzés: A most vizsált modell $R=0$ esetén a visszatevés nélküli, $R=1$ esetén a visszatevéses urnamodellt adja mint speciális esetet.

6. Egy urnában z zöld és s sárga golyó van. Egymás után kihúzunk négy golyót úgy, hogy minden húzás után a golyót visszadobjuk az urnába, és vele együtt az urnába dobunk 2 ellenkező színű golyót. Mi a valószínűsége egy zöld, zöld, zöld, sárga húzássorozatnak?

   Megoldás: Számoljuk ki annak valószínűséségét, hogy az első húzás eredménye Z=(zöld), annak feltételes valószínűségét, hogy a második húzás Z, feltéve, hogy az első húzás Z, annak a feltételes valószínűségét, hogy a harmadik húzás eredménye Z feltéve, hogy előtte Z, Z és annak feltételes valószínűségét, hogy a negyedik húzás eredménye S feltéve, hogy előtte Z, Z, Z húzás volt. Ez a valószínűség, illetve feltételes valószínűségek $\dfrac{z}{z+s}$, $\dfrac{z}{z+s+2}$, $\dfrac{z}{z+s+4}$, $\dfrac{s+6}{z+s+6}$. A keresett valószínűség $\dfrac{z}{z+s} \cdot \dfrac{z}{z+s+2} \cdot \dfrac{z}{z+s+4} \cdot \dfrac{s+6}{z+s+6}$.

7. Tekintsük az előző feladatban bevezetett urnamodellt. Számítsuk ki annak a valószínűségét, hogy az első húzás eredménye zöld, és annak, hogy a második húzás eredménye zöld.

   Megoldás: Annak valószínűsége, hogy az első húzás eredménye zöld $\dfrac{z}{z+s}$. Az az esemény, hogy a második húzás zöld úgy fordulhat elő hogy vagy egy zöld, zöld vagy egy sárga, zöld húzássorozat jelenik meg. A második keresett valószínűség ezen két húzássorozat valószínűségének az összege, tehát $\dfrac{z}{z+s} \cdot \dfrac{z}{z+s+2} + \dfrac{s}{z+s} \cdot \dfrac{z+2}{z+s+2} = \dfrac{z^2+s(z+2)}{(z+s)(z+s+2)}$.

   Megjegyzés: Ebben a feladatban olyan urnamodellt tekintettünk, amelyben eltér annak a valószínűsége, hogy az első húzásban illetve annak a valószínűsége, hogy a második húzásban húzunk zöld golyót.

8. Feldobunk egy szabályos pénzdarabot 100-szor egymás után. Számítsuk ki a fejdobások számának várható értékét.

   Megoldás: A fejdobások száma 0 és 100 között van. Annak valószínűsége, hogy k fejdobás következik be, $0 \le k \le 100$, $p_k = \dbinom{100}{k} 2^{-100}$. Ezért a dobások számának várható értéke $E\xi = \sum\limits_{k=0}^{100} k \dbinom{100}{k} 2^{-100}$, ahol ξ jelöli a fejdobások számát megadó valószínűségi változót. Ezt az összeget közvetlenül kiszámíthatjuk. Valóban, mivel

$k\binom{100}{k} = 100\binom{99}{k-1}$, ezért

$$E\xi = \sum_{k=0}^{100} k\binom{100}{k}2^{-100} = 50\sum_{k=0}^{99}\binom{99}{k}2^{-99} = 50\left(\frac{1}{2}+\frac{1}{2}\right)^{99} = 50.$$

Valójában a vizsgált várható értéket egyszerűbben is kiszámíthatjuk. Vezessük be a ξ_j valószínűségi változókat, $\xi_j = 1$, ha a j-ik dobás eredménye fej $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 100$. Ekkor a fejdobások száma $\xi = \sum\limits_{j=1}^{100} \xi_j$, $E\xi = \sum\limits_{j=1}^{100} E\xi_j$. Mivel $E\xi_j = \frac{1}{2}$, $1 \le j \le 100$, ezért $E\xi = \frac{1}{2} \cdot 100 = 50$,

9.) Feldobunk egy szabályos dobókockát 100-szor egymás után. Tekintsük a dobások eredményeinek összegét. Számítsuk ki ennek az összegnek a várható értékét.

Megoldás: Definiáljuk az $\eta_1, \dots, \eta_{100}$ *független* valószínűségi változókat úgy, hogy η_j és a a j-ik dobás értéke megegyezik. Ekkor $P(\eta_j = k) = \frac{1}{6}$, $1 \le j \le 100$, $1 \le k \le 6$, és minket a $\xi = \sum\limits_{j=1}^{100} \eta_j$ valószínűségi változó várható értéke érdekel. Ez könnyen kiszámolható, mert $E\xi = E\sum\limits_{j=1}^{100} \eta_j = \sum\limits_{j=1}^{100} E\eta_j$, $E\eta_j = \frac{1}{6}(1+2+3+4+5+6) = 3.5$. Innen $E\xi = 350$.

Házi feladat:

Feldobunk egy szabályos dobókockát 100-szor egymás után. Tekintsük a páros értékű dobások eredményeinek összegét. Számítsuk ki ennek az összegnek a várható értékét.

10. Feldobunk egy szabályos pénzdarabat 100-szor egymás után. Tekintsük az egymást követő fej-fej dobássorozatok számát, és számítsuk ki annak várható értékét.

Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 99$, valószínűségi változókat: $\xi_j = 1$, ha a j-ik és $j+1$-ik dobások mindegyikének eredménye fej, $\xi_j = 0$ egyébként. Vegyük észre, hogy minket az $S = \sum\limits_{j=1}^{99} \xi_j$ valószínűségi változó várható értéke érdekel. Ezért $ES = E\left(\sum\limits_{j=1}^{99} \xi_j\right) = \sum\limits_{j=1}^{99} E\xi_j = \dfrac{99}{4}$, mivel $E\xi_j = \dfrac{1}{4}$. (Érdemes megjegyezni, hogy az ebben a feladatban tekintett ξ_j valószínűségi változók nem függetlenek, de a függetlenségre nincs szükség a várható érték additivításához.)