A március 14.-i gyakorlat feladatai

1. Egy szabályos pénzdarabot feldobunk 100-szor egymás után. Számoljuk ki a fejdobások számának a várható értékét és szórásnégyzetét. Számítsuk ki ezt egyrészt direkt módon felírva és kiszámolva a várható értéket és szórásnégyzetet kifejező összeget másrészt egyszerűbben a független valószínűségi változók szórásnégyzetét kifejező képlet segítségével.

   Megoldás: A fejdobások száma 0 és 100 között van. Annak valószínűsége, hogy k fejdobás következik be, $0 \le k \le 100$, $p_k = \binom{100}{k} 2^{-100}$. Ezért a dobások számának várható értéke $E\xi = \sum\limits_{k=0}^{100} k \binom{100}{k} 2^{-100}$, ahol ξ jelöli a fejdobások számát megadó valószínűségi változót. Továbbá $E\xi^2 = \sum\limits_{k=0}^{100} k^2 \binom{100}{k} 2^{-100}$, a szórásnégyzet pedig $\operatorname{Var}\xi = E\xi^2 - (E\xi)^2$. Ezeket az összegeket közvetlenül kiszámíthatjuk. Valóban, $k\binom{100}{k} = 100\binom{99}{k-1}$, $E\xi = \sum\limits_{k=0}^{100} k\binom{100}{k} 2^{-100} = 50 \sum\limits_{k=0}^{99} k \binom{99}{k} 2^{-99} = 50\left(\frac{1}{2}+\frac{1}{2}\right)^{99} = 50$. Továbbá, $k^2\binom{100}{k} = [k(k-1)+k]\binom{100}{k} = 100 \cdot 99 \cdot \binom{98}{k-2} + 100 \cdot \binom{100}{99}$, $E\xi^2 = 2^{-100} \cdot 100 \cdot 99 \cdot \sum\limits_{k=0}^{98} \binom{98}{k} + 2^{-100} \cdot 100 \cdot \sum\limits_{k=0}^{99} \binom{99}{k} = 2^{-100}\left(100 \cdot 99 \cdot (1+1)^{98} + 100 \cdot (1+1)^{99}\right) = 25{\cdot}99{+}50$, $\operatorname{Var}\xi = E\xi^2 - (E\xi^2) = 25$. Valójában a vizsgált várható értéket és szórásnégyzetet 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 $\xi = \sum\limits_{j=1}^{100} \xi_j$, $E\xi = \sum\limits_{j=1}^{100} E\xi_j$, $\operatorname{Var}\xi = \sum\limits_{j=1}^{100} \operatorname{Var}\xi_j$. Mivel $E\xi_j = \frac{1}{2}$, $E\xi_j^2 = \frac{1}{2}$, $\operatorname{Var}\xi_j = \frac{1}{4}$, $1 \le j \le 100$, ezért $E\xi = \frac{1}{2} \cdot 100 = 50$, $\operatorname{Var}\xi = \frac{1}{4} \cdot 100 = 25$.

2. Egy szabályos dobókockát feldobunk 100-szor egymás után. Számoljuk ki a dobáseredmények összegének várható értékét és szórásnégyzetét.

   Megoldás: Legyenek $\eta_1, \dots, \eta_{100}$ *független* valószínűségi változók, amelyekre $P(\eta_j = k) = \frac{1}{6}$, $1 \le j \le 100$, $1 \le k \le 6$. Ekkor a kiszámítandó várható érték és szórásnégyzet $E\sum\limits_{j=1}^{100} \eta_j = \sum\limits_{j=1}^{100} E\eta_j$, $\operatorname{Var}\sum\limits_{j=1}^{100} \eta_j = \sum\limits_{j=1}^{100} \operatorname{Var}\eta_j$. (A második reláció felhasználja a tekintett valószínűségi változók függetlenségét.) $E\eta_j = \frac{1}{6}(1+2+3+4+5+6) = 3.5$, $E\eta_j^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$, $\operatorname{Var}\eta_j = \frac{35}{12}$, $E\xi = 350$, $\operatorname{Var}\xi = \frac{3500}{12}$.

   Házi feladat:

   Egy szabályos dobókockát feldobunk 100-szor egymás után. Tekintsük a páros értékű dobások dobáseredményeinek összegét és számítsuk ki annak szórásnégyzetét.

3. Legyen egy urnában 20 piros és 30 fehér golyó. Húzzunk ki 20 golyót visszatevés nélkül. Számoljuk ki a kihúzott piros golyók számának várható értékét és szörásnégyzetét.

   Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 20$, valószínűségi változókat: $\xi_j(\omega) = 1$, ha a j-ik húzás eredménye piros, $\xi_j(\omega) = 0$, ha a j-ik húzás eredménye fehér. Ekkor a $\xi = \sum_{j=1}^{20} \xi_j$ összeg várható értékét és szórásnégyzetét kell kiszámolnunk. Továbbá $E\xi_j = E\xi_1 = \frac{2}{5}$, $\operatorname{Var}\xi_j = \operatorname{Var}\xi_1 = \frac{2}{5} - \frac{4}{25} = \frac{6}{25}$, $E\xi_j\xi_k - E\xi_j E\xi_k = E\xi_1\xi_2 - E\xi_1 E\xi_2 = \frac{2}{5} \cdot \frac{19}{49} - \frac{4}{25} = -\frac{6}{1245}$, ha $j \ne k$. Innen a 20 dobásban kihúzott piros golyók számának várható értéke $20E\xi_1 = 20 \cdot \frac{2}{5} = 8$, és szórásnégyzete $20\operatorname{Var}\xi_1 + 20 \cdot 19 \cdot \operatorname{Cov}(\xi_1, \xi_2) = 20 \cdot \frac{6}{25} - 380 \cdot \frac{6}{1245} = \frac{144}{49}$.

   Házi feladat:

   Legyen egy urnában 10 piros és 20 fehér golyó. Húzzunk ki 20 golyót, és minden húzás után a kihúzott golyót visszadobjuk egy másik azonos színű golyóval együtt. Számoltuk a kihúzott piros golyók számának várható értékét és szórásnégyzetét.

   Megjegyzés: Használjuk az előző gyakorlat eredményét arról, hogy mi a valószínűsége annak, hogy egy adott húzásban piros golyót húzunk vagy két különböző húzás mindegyikében piros golyót húzunk.

4. Egy szabályos dobókockát feldobunk tízszer. Számoljuk ki a dobásösszeg harmadik hatványának a várható értékét.

   Megoldás: Vezessük be a következő valószínűségi változókat: $\xi_j(\omega) = k$, ha a j-ik dobás eredménye k, $1 \le j \le 10$, $1 \le k \le 6$. Ekkor az $E\left(\sum_{j=1}^{10} \xi_j(\omega)\right)^3$ várható értéket kell kiszámítanunk. Ennek érdekében tekintsük a $\left(\sum_{j=1}^{10} \xi_j(\omega)\right)^3$ kifejezést és értsük meg milyen tagokat kapunk, ha elvégezzük a beszorzásokat. Egyrészt megjelenik 10 ξ_j^3 alakú kifejezés, és $E\xi_j^3 = E\xi_1^3$ minden ilyen tagra. Ezenkívül megjelenik $3 \cdot 10 \cdot 9$ $\xi_j^2\xi_k$, $j \ne k$, alakú kifejezés, mert a lehetséges (j,k) párokat $10 \cdot 9$ módon választhatjuk ki, és a k (csak egyszer szereplő tényező) három helyen szerepelhet. Továbbá minden ilyen tagra $E\xi_j^2\xi_k = E\xi_j^2 E\xi_k = E\xi_1^2 E\xi_1$. Továbbá $10 \cdot 9 \cdot 8$ módon jelenhet meg $\xi_j\xi_k\xi_l$ alakú tag, ahol a j, k és l indexek mind különbözőek, és ezekre $E\xi_j\xi_k\xi_l = (E\xi_1)^3$. Másfajta tag nem jelenik meg a szorzatban. Innen $\left(\sum_{j=1}^{10} \xi_j(\omega)\right)^3 = 10E\xi_1^3 + 270E\xi_1(E\xi_1)^2 + 720\,(E\xi_1)^3 = 4410 + 14332.5 + 3087 = 218295.5$, mert $E\xi_1 = 3.5$, $E\xi_1^2 = \frac{91}{6}$ és $E\xi_1^3 = 441$.

5. Egy urnában 10 piros és 20 fehér golyó van. Kihúzunk visszatevés nélkül 10 golyót. A páros sorszámú húzások esetén fehér golyó húzás esetén nyerünk 3 forintot, piros golyó húzás esetén pedig nem nyerünk és nem veszítünk semmit. A páratlan sorszámú húzások esetén piros húzás esetén 2 forintot nyerünk, fehér golyó húzás esetén nem nyerünk és nem veszítünk semmit. Számoljuk ki a nyereményünk várható értékét és szórásnégyzetét.

   Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 10$, valószínűségi változókat: Páros j számokra $\xi_j = 3$, ha a j-ik húzás eredménye piros, $\xi_j = 0$, ha a j-ik húzás eredménye fehér golyó. Páratlan j számokra $\xi_j = 2$, ha a j-ik húzás eredménye piros, $\xi_j = 0$, ha a j-ik húzás eredménye fehér golyó. Ekkor az $S = \sum\limits_{j=1}^{10} \xi_j$ összeg várható értékét és szórásnégyzetét kell kiszámolnunk. Ennek érdekében számoljuk ki az $E\xi_j$ várható értékeket, $\operatorname{Var}\xi_j$ szórásnégyzeteket és $\operatorname{Cov}(\xi_j, \xi_k)$ kovarianciákat. Annak valószínűsége, hogy a j-ik húzás eredménye piros $\frac{1}{3}$, annak valószínűsége, hogy j-ik húzás eredménye fehér $\frac{2}{3}$. Ezért $E\xi_j = 2$, ha j páros, $E\xi_j = \frac{2}{3}$, ha j páratlan, és $ES = 5\left(2 + \frac{2}{3}\right) = 13\frac{1}{3}$.

   A szórásnégyzet kiszámítása érdekében vegyük észre, hogy $E\xi_j^2 = 6$, $\operatorname{Var}\xi_j = 2$, ha j páros, és $E\xi_j^2 = \frac{4}{3}$, $\operatorname{Var}\xi_j = \frac{8}{9}$. Továbbá annak valószínűsége, hogy két j és k indexre $j \ne k$, a j-ik és k-ik húzás mindegyike fehér $\frac{2\cdot 19}{3\cdot 29} = \frac{38}{87}$, annak, hogy mindkét húzás piros $\frac{9}{3\cdot 29} = \frac{3}{29}$, annak, hogy az egyik húzás fehér a másik piros $\frac{2\cdot 10}{3\cdot 29} = \frac{20}{57}$. Ezért $E\xi_j\xi_k = 9 \cdot \frac{2\cdot 19}{3\cdot 29} = \frac{114}{29}$, $\operatorname{Cov}(\xi_j, \xi_k) = \frac{114}{29} - 4 = -\frac{2}{29}$, ha j és k páros, $E\xi_j\xi_k = 4 \cdot \frac{3}{29} = \frac{12}{29}$, $\operatorname{Cov}(\xi_j, \xi_k) = \frac{12}{29} - \frac{4}{9} = -\frac{8}{263}$, ha j és k páratlan, $E\xi_j\xi_k = 6 \cdot \frac{20}{57} = \frac{40}{29}$, $\operatorname{Cov}(\xi_j, \xi_k) = \frac{40}{29} - \frac{4}{3} = \frac{4}{87}$, ha j és k közül az egyik páros, a másik páratlan. Olyan (j,k) pár, $1 \le j,k \le 10$, $j \ne k$, melyre j és k mindegyike páros vagy mindegyike páratlan, összesen 20 van, és olyan (j,k) pár, melyekre az egyik páros, a másik páratlan $2 \cdot 5 \cdot 5 = 50$ van, ezért

$$\sum_{1\le j,k\le 10,\, j\ne k} \operatorname{Cov}(\xi_j, \xi_k) = 20\left(-\frac{2}{29} - \frac{8}{263}\right) + 50 \cdot \frac{4}{87} = \frac{80}{263}.$$

   Innen

$$\begin{aligned}\operatorname{Var} S &= \sum_{j=1}^{10} \operatorname{Var}\xi_j + \sum_{1\le j,k\le 10,\, j\ne k} \operatorname{Cov}(\xi_j, \xi_k)\\ &= 5\left(2 + \frac{8}{9}\right) + \frac{80}{263} = \frac{130}{9} + \frac{80}{263} = \frac{3850}{263}.\end{aligned}$$

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

   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 $ES = E\left(\sum\limits_{j=1}^{99} E\xi_j\right) = \dfrac{99}{4}$, 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.)

A szórásnégyzet kiszámításában viszont figyelembe kell vennünk azt, hogy nem csupa független valószínűségi változó összegét vizsgáljuk. Használjuk a szórásnégyzet kiszámolásánál a következő formulát.

$$\operatorname{Var} S = \operatorname{Var}\left(\sum_{j=1}^{99} \xi_j\right) = \sum_{j=1}^{99} \operatorname{Var} \xi_j + 2 \sum_{1\le j<k\le 99} \operatorname{Cov}(\xi_j, \xi_k).$$

Továbbá $\operatorname{Cov}(\xi_j, \xi_k) = 0$, ha $k \ge j+2$, mert ebben az esetben ξ_j és ξ_k függetlenek, és $E\xi_j\xi_{j+1} = \frac{1}{8} - \frac{1}{16} = \frac{1}{16}$ minden $1 \le j \le 98$ számra. Ugyanis $E\xi_j\xi_{j+1} = \frac{1}{8}$, mivel $\xi_j\xi_{j+1} = 1$, ha a j-ik, $j+1$-ik és $j+2$-ik dobások mindegyike fej, aminek valószínűsége $\frac{1}{8}$, és $\xi_j\xi_{j+1} = 0$ egyébként. Továbbá $E\xi_j E\xi_{j+1} = \frac{1}{16}$. Ezenkívül $\operatorname{Var} \xi_j = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$. Innen $\operatorname{Var} S = 99 \cdot \frac{3}{16} + 2 \cdot \frac{98}{16} = \frac{493}{16}$.

7. 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 $P(\xi_j \ne \xi'_j$ ha $j \ne j')$ 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 $1-\prod_{j=1}^{29}\left(1-\frac{j}{365}\right)$.

Hasonlóan, annak valószínűsége, hogy ha k golyót dobunk n urnába az adott módon, akkor van olyan urna, amelyikbe legalább két golyó esik $1-\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)$.

Adjunk jó közelítést a $\log\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)=\sum_{j=1}^{k-1}\log\left(1-\frac{j}{n}\right)$ kifejezésre, ha $n\to\infty$, $\frac{k(n)}{\sqrt{n}}\to\alpha$. Heurisztikus érvelés szerint mivel $\log\left(1-\frac{j}{n}\right)\sim-\frac{j}{n}$ a $\log(1+x)$ függvény Taylor sorfejtése szerint, ezért $\sum_{j=1}^{k-1}\log\left(1-\frac{j}{n}\right)\sim-\sum_{j=1}^{k-1}\frac{j}{n}=-\frac{(k-1)k}{2n}$, ahonnan $\log\prod_{j=1}^{k(n)-1}\left(1-\frac{j}{n}\right)\to-\frac{\alpha^2}{2}$, 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 $1-\prod_{j=1}^{k(n)-1}\left(1-\frac{j}{n}\right)\to1-e^{-\alpha^2/2}$, ha $n\to\infty$, és $\frac{k(n)}{\sqrt{n}}\to\alpha$.