A november 4.-i gyakorlat témája

Tárgyaljuk meg néhány példán keresztül azt, hogy hogyan lehet kiszámítani olyan valószínűségi változók szórásnégyzetét, amelyek előállíthatóak mint egyszerű, de nem feltétlenül független valószínűségi változók összegei. Ebből a célból be kell vezetnünk két valószínűségi változó kovarianciafüggvényének a fogalmát. Ez azt méri, hogy milyen mértékben függ a két valószínűségi változó egymástól.

A kovarianciafüggvény definiciója. *Legyen ξ és η két valószínűségi változó, amelyek mindegyikének létezik szórásnégyzete, azaz $E\xi^2 < \infty$ és $E\eta^2 < \infty$. Ekkor a ξ és η kovariancifüggvényét* $\operatorname{Cov}(\xi,\eta)$*-t a*

$$\operatorname{Cov}(\xi,\eta) = E\left[(\xi - E\xi)(\eta - E\eta)\right]$$

kifejezés definiálja.

Lemma.

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

Valóban,

$$\begin{aligned}\operatorname{Cov}(\xi,\eta) &= E\left[(\xi - E\xi)(\eta - E\eta)\right] = E\left[(\xi\eta - \eta E\xi - \xi E\eta + E\xi E\eta)\right]\\ &= E\xi\eta - E\eta E\xi - E\xi E\eta + E\xi E\eta = E\xi\eta - E\xi E\eta.\end{aligned}$$

1.) ξ és η két független valószínűségi változó, akkor $\operatorname{Cov}(\xi,\eta) = 0$.

Megoldás: Láttuk, hogy $\operatorname{Cov}(\xi,\eta) = E\xi\eta - E\xi E\eta$. Viszont, ha ξ és η független valószínűségi változók, akkor $E\xi\eta = E\xi E\eta$.

Az alábbi eredmény megadja, hogy hogyan tudjuk kifejezni valószínűségi változók összegének a szórásnégyzetét viszonylag egyszerű módon.

Tétel. *Ha $\xi_1,\dots,\xi_n$ valószínűségi változók ugyanazon a valószínűségi mezőn, amelyek mindegyikének létezik szórásnégyzete, akkor*

$$\begin{aligned}\operatorname{Var}\left(\sum_{j=1}^n \xi_j\right) &= \sum_{j=1}^n \operatorname{Var}\xi_j + 2\sum_{1\le j<k\le n}(E\xi_j\xi_k - E\xi_j E\xi_k)\\ &= \sum_{j=1}^n \operatorname{Var}\xi_j + 2\sum_{1\le j<k\le n}(\operatorname{Cov}(\xi_j,\xi_k)\\ &= \sum_{j=1}^n \operatorname{Var}\xi_j + \sum_{1\le j,k\le n,\ j\ne k}(\operatorname{Cov}(\xi_j,\xi_k).\end{aligned}$$

Bizonyítás.

$$\mathrm{Var}\left(\sum_{j=1}^n \xi_j\right) = E\left(\sum_{j=1}^n \xi_j\right)^2 - \left(E\sum_{j=1}^n \xi_j\right)^2 = E\left(\sum_{j=1}^m\sum_{k=1}^n \xi_j\xi_k\right) - \left(\sum_{j=1}^m\sum_{k=1}^n E\xi_j E\xi_k\right)$$
$$= \left(\sum_{j=1}^m\sum_{k=1}^n E\xi_j\xi_k\right) - \left(\sum_{j=1}^m\sum_{k=1}^n E\xi_j E\xi_k\right)$$
$$= \sum_{j=1}^n \left(E\xi_j^2 - (E\xi_j)^2\right) + 2\sum_{1\le j<k\le n} (E\xi_j\xi_k - E\xi_j E\xi_k).$$

és $E\xi_j^2 - (E\xi_j)^2 = \mathrm{Var}\,\xi_j$, $E\xi_j\xi_k - E\xi_j E\xi_k = \mathrm{Cov}\,(\xi_j, \xi_k)$.

Megjegyzés: A fenti tételben két különböző, de ekvivalens módon fejeztük ki valószínűségi változók összegének a kovarianciafüggvényét. Az egyik módon az összegzés a második tagban az $1 \le j < k \le n$ alakú tagokra történt, és az összeget beszoroztuk kettővel. A másik kifejezésben az összegezés az $1 \le j,k \le n$, $j \neq k$, alakú számpárokra történt, és a 2 szorzó nem szerepelt. Ez analóg a következő egyszerű algebrai azonosságokkal: $\left(\sum\limits_{j=1}^n x_j\right)^2 = \sum\limits_{j=1}^n x_j^2 + 2\sum\limits_{1\le j<k\le n} x_jx_k$, és $\left(\sum\limits_{j=1}^n x_j\right)^2 = \sum\limits_{j=1}^n x_j^2 + \sum\limits_{1\le j,k\le n, j\neq k} x_jx_k$.

A fenti eredmény segítségével viszonylag egyszerűen meg tudjuk oldani a következő két feladatot. Ezek közül az első a dolgozatban szerepelt 5. feladat kiegészítése.

2.) Egy szabályos dobókockát feldobunk 101 alkalommal. Arra vagyunk kiváncsiak, hogy hány olyan dobásidőpont volt, (a 2., 3., és így tovább a 101. dobásidőpontok között), amikor a dobás értéke (szigorúan) nagyobb volt, mint a megelőző dobás értéke. Számítsuk ki az ilyen dobásidőpontok számának a várható értékét és szórásnégyzetét.

Megoldás: Vezessük be a következő ξ_k valószínűségi változókat, $1 \le k \le 100$: $\xi_k = 1$, ha a $k+1$-ik kockadobás értéke nagyobb, mint a k-ik kockadobás értéke, $\xi_k = 0$, ha a $k+1$-ik kockadobás értéke kisebb vagy egyenlő, mint a k-ik kockadobás értéke. Ekkor minket a $\xi = \sum\limits_{k=1}^{100} \xi_k$ valószínűségi változó várható értéke és szórásnégyzete érdekel. Viszont $E\xi = \sum\limits_{k=1}^{100} E\xi_k$, $E\xi_k = P(\xi_k = 1$, és $P(\xi_k = \frac{1}{36}(5+4+3+2+1) = \frac{5}{12}$, (Ha a k-ik dobás értéke 1, akkor a $k+1$-ik dobás 5 értéket vehet fel, ha 2 akkor 4-et, ha 3 akkor 3-at sit.) Ezért a keresett várható érték $\frac{500}{12} = \frac{125}{3}$.

A ξ valószínűségi változó szórásnégyzetét kiszámolhatjuk a fenti tétel segítségével. E tétel alkalmazása során ki kell számolnunk az $\mathrm{Var}\,\xi_k$ szórásnégyzeteket, továbbá a $\mathrm{Cov}\,(\xi_j, \xi_k)$, $j \neq k$, kovarianciákat. $\mathrm{Var}\,\xi_k = E\xi_k^2 - (E\xi_k)^2 = P(\xi_k = 1) - P(\xi_k = 1)^2 = \frac{5}{12} - \frac{25}{144} = \frac{35}{144}$. Vegyük észre, hogy amennyiben $|k - j| \ge 2$, akkor ξ_j és ξ_k független valószínűségi változók, ezért $\mathrm{Cov}\,(\xi_j, \xi_k) = 0$ ebben az esetben.

A $\mathrm{Cov}\,(\xi_k, \xi_{k+1}) = E\xi_k\xi_{k+1} - E\xi_k E\xi_{k+1} = P(\xi_k = \xi_{k+1} = 1) - \frac{25}{144}$ azonosság teljesül. A $P(\xi_k = \xi_{k+1} = 1)$ valószínűség kiszámolásához azt kell összeszámolni, hogy hány három hosszúságú szigorúan monoton növekvő sorozat van, amelynek tagjai értéküket 1 és 6 között veszik fel. Ezt kézzel is össze lehetne számolni, de az első két gyakorlaton megbeszéltük, hogy $\binom{6}{3} = 20$ ez a szám. (A hat számból kiválasztunk három különbözőt, és azokat nagyság szerint sorba rakjuk.) Ezért $E\xi_k\xi_{k+1} = \frac{20}{216} = \frac{5}{54}$, $\mathrm{Cov}\,(\xi_k, \xi_{k+1}) = \frac{20}{216} - \frac{25}{144} = \frac{10}{432}$. Innen $\mathrm{Var}\,\xi = \sum\limits_{k=1}^{100} \mathrm{Var}\,\xi_k + 2\sum\limits_{k=1}^{99} \mathrm{Cov}\,(\xi_k, \xi_{k+1}) = \frac{350}{144} + \frac{1980}{432} = \frac{2330}{432}$.

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 szórásnégyzetét.

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 $\xi = \sum\limits_{j=1}^{20} \xi_j$ összeg várható értékét és szórásnégyzetét kell kiszámolnunk. Továbbá $E\xi_j = E\xi_1 = \dfrac{2}{5}$, $\mathrm{Var}\,\xi_j = \mathrm{Var}\,\xi_1 = \dfrac{2}{5} - \dfrac{4}{25} = \dfrac{6}{25}$, $E\xi_j\xi_k - E\xi_j E\xi_k = E\xi_1\xi_2 - E\xi_1 E\xi_2 = \dfrac{2}{5}\cdot\dfrac{19}{49} - \dfrac{4}{25} = -\dfrac{6}{1245}$, ha $j \ne k$. Innen a 20 dobásban kihúzott piros golyók számának várható értéke $20 \cdot \dfrac{2}{5} = 8$ és szórásnégyzete $20 \cdot \dfrac{6}{25} - 380 \cdot \dfrac{6}{1245} = \dfrac{144}{49}$.

Házi feladat:

Egy szabályos pénzdarabot feldobunk egymás után 101-szer. Számítsuk ki az egymást követő fej-fej dobások számának a várható értékét és szórásnégyzetét.

Végül tekintsük azt, hogy hogyan lehet kiszámolni független, diszkrét eloszlású valószínűségi változók összegének az eloszlását. Megmutatjuk, hogy két független Poisson eloszlású valószínűségi változónak az összege is Poisson eloszlású, és az összeg paramétere megegyezik az összeadandók paraméterének az összegével. Először definiáljuk magát a Poisson eloszlást, és lássuk be, hogy az valóban eloszlás.

Poisson eloszlás definiciója. *Egy ξ valószínűségi változó λ paraméterű Poisson eloszlású, ha ξ nem negatív egész értékeket vesz fel, és*

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

4.) A fent definiált Poisson eloszlás, valóban eloszás, azaz $P(\xi = k) \ge 0$ minden egész számra, és $\sum\limits_{k=1} P(\xi = k) = 1$.

Megoldás: A $P(\xi = k) \geq 0$ reláció nyilvánvaló. Másrészt,

$$\sum_{k=0}^{\infty} P(\xi = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

5.) Mutassuk meg, hogy ha ξ és η két független, λ illetve μ paraméterű Poisson eloszlású valószínűségi változó, akkor $\xi + \eta$ $\lambda + \mu$ paraméterű Poisson eloszlású valószínűségi változó.

Megoldás:

$$\begin{aligned} P(\xi + \eta = k) &= \sum_{j=0}^{k} P(\xi = j, \eta = k - j) = \sum_{j=0}^{k} P(\xi = j) P(\eta = k - j) \\ &= \sum_{j=0}^{k} \frac{\lambda^j}{j!} e^{-\lambda} \frac{\mu^{k-j}}{(k-j)!} e^{-\mu} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} \lambda^j \mu^{(k-j)} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k. \end{aligned}$$

6.) Legyen ξ és η két független binomiális elsoszlású valószínűségi változó (n, p) illetve (m, p) paraméterekkel, azaz legyen $P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $0 \leq k \leq n$, és $P(\xi = k) = 0$ egyébként, $P(\eta = k) = \binom{m}{k} p^k (1-p)^{m-k}$, $0 \leq k \leq m$, és $P(\xi = k) = 0$ egyébként. Lássuk be, hogy $\xi + \eta$ binomiális eloszlású $(n+m, p)$ paraméterrel, azaz legyen $P(\xi + \eta = k) = \binom{n+m}{k} p^k (1-p)^{n+m-k}$, $0 \leq k \leq n + m$, és $P(\xi = k) = 0$ egyébként.

Megoldás: Legyen $0 \leq k \leq n + m$. Ekkor

$$\begin{aligned} P(\xi + \eta = k) &= \sum_{j=0}^{n} P(\xi = j, \eta = k - j) = \sum_{j=0}^{n} P(\xi = j) P(\eta = k - j) \\ &= \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{k-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-(k-j)} \\ &= p^k (1-p)^{n+m-k} \sum_{j=0}^{n} \binom{n}{j} \binom{m}{k-j} = p^k (1-p)^{n+m-k} \binom{n+m}{k}, \end{aligned}$$

mert, mint azt az első gyakorlaton megbeszéltük érvényes a $\sum\limits_{j=0}^{n} \binom{n}{j} \binom{m}{k-j} \binom{n+m}{k}$. Az, hogy $P(\xi + \eta = k) = 0$ reláció teljesül minden egyéb k számra.

Házi feladat (nem kötelező)

Bizonyítsuk be a fenti azonosságot a binomiális eloszlás szemléletes tartalmát felhasználva, azaz konstruáljunk két független (n, p) és (m, p) paraméterű binomiális eloszlású valószínűségi változót, amelyek összegéről látszik, hogy binomiális eloszlású $(n + m, p)$ paraméterrel.

DOLGOZAT FELADATOK

(A programozó matematikus gyakorlaton)

1.) Minden héten 1 000 000 ember lottózik, akik a lottószelvényeiket egymástól függetlenül töltik ki. (90 számból kell eltalálni az öt kihúzott azámot.) Mi annak a valószínűsége, hogy a 20. héten lesz először 5 találat?

2.) Három város valamelyikébe érkeztünk, mindegyikbe egyforma, azaz 1/3 valószínűséggel. Az egyik az igazmondók városa, ahol minden kérdésre 0.9 valószínűséggel válaszolnak igazat, a második a hazugok városa, ahol minden kérdésre 0.8 valószínűséggel hamis választ adnak, a harmadik a figyelmetlenek városa, ahol minden kérdésre a kérdéstől függetlenül 1/2 valószíműséggel igent és 1/2 valószínűséggel nemet válaszolnak. Megkérdezünk egy embert, akivel találkoztunk, hogy az igazmondók városába érkeztünk-e. Azt a választ kapjuk, hogy nem. Mi annak a valószínűsége, hogy az igazmondók városába érkeztünk?

3.) Két szabályos dobókockát feldobunk. Mi annak a valószínűsége, hogy mind a két dobás páros feltéve, hogy legalább az egyik dobás hatos?

4.) Egy szabályos dobókockát feldobunk százszor. Számoljuk a páros értékű dobások öszegének a várható értékét és szórásnégyzetét.

5.) Egy szabályos dobókockát feldobunk 101 alkalommal. Arra vagyunk kiváncsiak, hogy hány olyan dobásidőpont volt, (a 2., 3., és így tovább a 101. dobásidőpontok között), amikor a dobás értéke (szigorúan) nagyobb volt, mint a megelőző dobás értéke. Számítsuk ki az ilyen dobásidőpontok számának a várható értékét.

6.) Mikor mondjuk, hogy három A, B és C esemény független egymástól?

MEGOLDÁSOK

1.) Annak a valószínűsége, hogy valaki egy héten ötös találatot ér el, $\frac{1}{\binom{90}{5}}$, annak a valószínűsége, hogy nem ér el ötös találatot $1-\frac{1}{\binom{90}{5}}$, és mivel a lottózok egymástól függetlenüljátszanak, annak valószínűsége, hogy az 1 000 000 lottozó egyike sem ér el ötös találatot egy adott héten $P=\left(1-\frac{1}{\binom{90}{5}}\right)^{1\,000\,000}$. Továbbá annak valószínűsége, hogy 19 héten keresztül nincs ötös találat, a 20. héten viszont van $P^{19}(1-P)$. Ezért a keresett valószínűség

$$\left(1-\frac{1}{\binom{90}{5}}\right)^{19\,000\,000}\left(1-\left(1-\frac{1}{\binom{90}{5}}\right)^{1\,000\,000}\right).$$

2.) Jelölje A_1 azt az eseményt, hogy az igazmondók városába érkeztünk, A_2 hogy a hazugok, A_3, pedig, hogy a figyelmetlenek városába. Jelölje B azt az eseményt, hogy kérdésünkre nemleges választ kaptunk. Ekkor feltevéseink szerint $P(A_1)=$

$P(A_2) = P(A_3) = \frac{1}{3}$, $P(B|A_1) = 0.1$ (az igazmondók városában hazudtak), $P(B|A_2) = 0.2$ (a hazugok városában igazat mondtak), $P(A_3|B) = 0.5$. Minket a $P(A_1|B)$ feltételes valószínűség érdekel. Ez

$$P(B|A_1) = \frac{P(A_1 \cap B)}{P(A_1)} = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$
$$= \frac{\frac{1}{3} \cdot 0.1}{\frac{1}{3} \cdot 0.1 + \frac{1}{3} \cdot 0.2 + \frac{1}{3} \cdot 0.5} = \frac{1}{8}.$$

3.) A keresett feltételes valószínűség megegyezik annak a valószínűségével, hogy az egyik dobás hatos a másik dobás pedig páros osztva annak a valószínűségével, hogy volt egy hatos dobás. Az, hogy az egyik dobás hatos a másik pedig páros azt jelenti, hogy a (2,6), (4,6), (6,6), (6,2), (6,4) dobások valamelyike következik be, ennek valószínűsége pedig $\frac{5}{36}$. Annak valószínűsége, hogy lesz hatos dobás a valószínűsége pedig $\frac{11}{36}$. Ez például úgy látható, hogy annak valószínűsége, hogy nem lesz hatos dobás $\frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36}$, a komplementer esemény valószínűsége pedig $\frac{11}{36}$. Így a keresett feltételes valószínűség értéke $\frac{\frac{5}{36}}{\frac{11}{36}} = \frac{5}{11}$.

4.) Vezessük be a következő ξ_k valószínűségi változókat, $1 \le k \le 100$: $\xi_k = 2$, ha a k-ik kockadobás értéke 2, $\xi_k = 4$, ha a k-ik kockadobás értéke 4, $\xi_k = 6$, ha a k-ik kockadobás értéke 6, $\xi_k = 0$, ha a k-ik kockadobás értéke 1, 3 vagy 5. Ekkor minket a $\xi = \sum\limits_{k=1}^{100} \xi_k$ valószínűségi változó várható értéke és szórásnégyzete érdekel. Viszont $E\xi = \sum\limits_{k=1}^{100} E\xi_k$ $\operatorname{Var}\xi = \sum\limits_{k=1}^{100} \operatorname{Var}\xi_k$, $E\xi_k = \frac{1}{6}(2+4+6) = 2$, $\operatorname{Var}\xi_k = E\xi_k^2 - (E\xi_k)^2$, $E\xi_k^2 = \frac{1}{6}(4+16+36) = \frac{28}{3}$, $\operatorname{Var}\xi_k = 163$ minden $1 \le k \le 100$ számra. Innen következik, hogy $E\xi = 100 \cdot 2 = 200$, $\operatorname{Var}\xi = 100 \cdot \frac{16}{3} = \frac{1600}{3}$.

5.) Vezessük be a következő ξ_k valószínűségi változókat, $1 \le k \le 100$: $\xi_k = 1$, ha a $k+1$-ik kockadobás értéke nagyobb, mint a k-ik kockadobás értéke, $\xi_k = 0$, ha a $k+1$-ik kockadobás értéke kisebb vagy egyenlő, mint a k-ik kockadobás értéke. Ekkor minket a $\xi = \sum\limits_{k=1}^{100} \xi_k$ valószínűségi változó várható értéke érdekel. Viszont $E\xi = \sum\limits_{k=1}^{100} E\xi_k$, $E\xi_k = P(\xi_k = 1$, és $P(\xi_k = \frac{1}{36}(5+4+3+2+1) = \frac{5}{12}$, (Ha a k-ik dobás értéke 1, akkor a $k+1$-ik dobás 5 értéket vehet fel, ha 2 akkor 4-et, ha 3 akkor 3-at sit.) Ezért a keresett várható érték $\frac{500}{12} = \frac{125}{3}$.

6.) Az A, B és C események akkor függetlenek, ha a következő azonosságok mindegyike teljesül:

$$P(A \cap B) = P(A)P(B)$$
$$P(A \cap C) = P(A)P(C)$$
$$P(B \cap C) = P(B)P(C)$$
$$P(A \cap B \cap C) = P(A)P(B)P(C).$$