A november 30.-i gyakorlat témája

Az előadáson szerepelt annak az állításnak a bizonyítása, hogy a standard normális eloszlás sűrűségfüggvénynek nevezett $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ függvény sűrűségfüggvény, tehát

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = 1.$$

Lássuk be (e tény felhasználásával),

1.) Egy standard normális eloszlású ξ valószínűségi változó, azaz egy $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ sűrűségfüggvénnyel rendelkező valószínűségi változó várható értéke nulla és szórásnégyzete 1.

Megoldás:

$$E\xi = \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = 0,$$

mert az $x\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ integrandus páratlan függvény. Ezért $\operatorname{Var}\xi = E\xi^2$, és parciális integrálással $f(x) = x$ és $g(x) = x\frac{1}{\sqrt{2\pi}}e^{-x^2/2} = \frac{d}{dx}\left(-\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right)$ választással kapjuk, hogy

$$\begin{aligned} E\xi^2 &= \int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = \int_{-\infty}^{\infty} -x\frac{d}{dx}\left(\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right)dx \\ &= \left[-x\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = 1. \end{aligned}$$

2.) Legyen ξ standard normális eloszlású valószínűségi változó. Számoljuk ki a ξ valószínűségi változó negyedik momentumát. Legyen $\bar{\xi}$ egy várható értékű és kettő szórásnégyzetű normális eloszlású valószínűségi változó. Számoljuk ki $\bar{\xi}$ sűrűségfüggvényét és negyedik momentumát.

Megoldás:

$$\begin{aligned} E\xi^4 &= \int_{-\infty}^{\infty} x^4\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = \int_{-\infty}^{\infty} -x^3\frac{d}{dx}\left(\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right)dx \\ &= \left[-x^3\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} 3x^2\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx \\ &= 3\int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = 3. \end{aligned}$$

A $\tilde{\xi} = \frac{\bar{\xi}-1}{\sqrt{2}}$ valószínűségi változó standard normális eloszlású, és $\bar{\xi} = \sqrt{2}(\tilde{\xi}+1)$, ahol $\tilde{\xi}$ standard normális eloszlású, ha $\bar{\xi}$ normális eloszlású 1 várható értékű és 2 szórásnégyzetű valószínűségi változó. Ezért, mint azt az előző órán megtárgyaltuk $\bar{\xi}$ sűrűségfüggvénye $\frac{1}{\sqrt{2}}\varphi\left(\frac{x-1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2\pi}}e^{-(x-1)^2/4}$. Ezért

$$E\bar{\xi}^4 = \int_{-\infty}^{\infty} x^4 \frac{1}{2\sqrt{\pi}} e^{-(x-1)^2/4}\,dx,$$

ahonnan $u = \frac{x-1}{\sqrt{2}}$ helyettesítéssel

$$\begin{aligned}
E\bar{\xi}^4 &= \int_{-\infty}^{\infty} (\sqrt{2}u+1)^4 \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du = 4\int_{-\infty}^{\infty} u^4 \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du \\
&\quad + 8\sqrt{2}\int_{-\infty}^{\infty} u^3 \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du + 12\int_{-\infty}^{\infty} u^2 \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du \\
&\quad + 4\sqrt{2}\int_{-\infty}^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/4}\,du \\
&= 12+0+12+0+1 = 25
\end{aligned}$$

Valójában az $E\bar{\xi}^4$ negyedik momentumot egyszerűbben is kiszámolhattuk volna. $E\bar{\xi}^4 = E(\sqrt{2}\tilde{\xi}+1)^4 = 4E\tilde{\xi}^4 + 8\sqrt{2}E\tilde{\xi}^3 + 12E\tilde{\xi}^2 + 4\sqrt{2}\tilde{\xi} + 1 = 4\cdot 3 + 0 + 12\cdot 1 + 1 = 25.$

3.) Egy szabályos dobókockát és egy szabályos érmét feldobunk 3300 alkalommal egymástól függetlenül. (Az érme és kockadobások eredményei is függetlenek egymástól.) Ha a kockadobás eredménye páros és az érme a fej oldalra esett, akkor annyi forintot nyerünk, amennyi a kockadobás eredménye. Ha az érme az írás oldalra esett vagy a kockadobás eredménye páratlan szám, akkor nem nyerünk, és nem is vesztünk semmit. Mi a valószínűsége annak, hogy az össznyereményünk 3190 és 3520 forint közé esik? Adjunk erre jó közelítő becslést a centrális határeloszlástétel és egy normális eloszlástáblázat segítségével.

Megoldás: Vezessük be a következő ξ_j, és η_j $1 \le j \le 3300$, valószínűségi változókat: $\xi_j = 2$, ha a j-ik kockadobás eredménye 2, $\xi_j = 4$, ha a j-ik kockadobás eredménye 4, $\xi_j = 6$, ha a j-ik kockadobás eredménye 6, $\xi_j = 0$, ha a j-ik kockadobás eredménye 1, 3 vagy 5. Legyen $\eta_j = 1$, ha a j-ik érmedobás eredménye fej, és $\eta_j = 0$, ha a j-ik érmedobás írás. Legyen $\zeta_j = \xi_j\eta_j$. Ekkor a j-ik dobásnál a nyereményünk ζ_j lesz, $1 \le j \le 3300$, és a $P\left(3190 \le \sum\limits_{j=1}^{3300} \zeta_j \le 3520\right)$ valószínűséget kell jól megbecsülnünk. Ennek érdekében számoljuk ki a független ζ_j valószínűségi változók várható értékét és szórásnégyzetét. $E\zeta_j = E\xi_j\eta_j = E\xi_j E\eta_j = \frac{1}{6}(2+4+6)\frac{1}{2} = 1$, $E\zeta_j^2 = E\xi_j^2 E\eta_j^2 = \frac{1}{6}(4+16+36)\cdot\frac{1}{2} = \frac{14}{3}$, $\operatorname{Var}\zeta_j^2 = E\zeta_j^2 - (E\zeta_j)^2 = \frac{11}{3}$. Innen a centrális határeloszlástétel alapján

$$P\left(3190 \le \sum_{j=1}^{3300} \zeta_j \le 3520\right) = P\left(\frac{-110}{\sqrt{3300\frac{11}{3}}} \le \frac{\sum\limits_{j=1}^{3300}\zeta_j - \sum\limits_{j=1}^{3300} E\zeta_j}{\sqrt{\sum\limits_{j=1}^{3300}\operatorname{Var}\xi_j}} \le \frac{220}{\sqrt{3300\frac{11}{3}}}\right)$$

$$\sim \Phi(2) - \Phi(-1) = 0.9772 + 0.8413 - 1 = 0.9285.$$

4.) Egy pénzdarabról ellenőrizni akarjuk, hogy igaz-e az a hipotézis, amely szerint ez az érme legalább $\frac{3}{4}$ valószínűséggel esik a fej és legfeljebb $\frac{1}{4}$ valószínűséggel az írás oldalára. Ennek érdekében feldobjuk a pénzdarabot 30 000 alkalommal, és a következő döntési szabályt hozzuk. Választunk egy k számot, és akkor fogadjuk el a hipotézist helyesnek, ha legalább k fejdobás történt. Legalább mekkorának kell válassztanunk ezt a k számot, ha azt akarjuk, hogy egy a hipotézist teljesítő pénzdarab esetén legalább 0.9 valószínűséggel döntsünk úgy, hogy a hipotézis teljesül?

Megoldás: Vezessük be a következő valószínűségi változókat: $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 30\,000$, $S = S_{30000} = \sum\limits_{j=1}^{30\,000} \xi_j$. Ha a fejdobás eredményének valószínűsége pontosan $\frac{3}{4}$, $E\xi_j = \frac{3}{4}$, $E\xi_j^2 = \frac{3}{4}$, $\operatorname{Var}\xi_j = E\xi_j^2 - (E\xi_j)^2 = \frac{3}{16}$, $ES = 30\,000 E\xi_j = 22\,500$, $\operatorname{Var} S = 30\,000 \operatorname{Var}\xi_j = 5625 = 75^2$. Innen és a centrális határeloszlástételből,

$$P(S > k) = P\left(\frac{S - ES}{\sqrt{\operatorname{Var} S}} > \frac{k - 22\,500}{75}\right) = 1 - P\left(\frac{S - ES}{\sqrt{\operatorname{Var} S}} \le \frac{k - 22\,500}{75}\right)$$
$$\sim 1 - \Phi\left(\frac{k - 22\,500}{75}\right).$$

Válasszuk a k számot úgy, hogy a fenti valószínűség körülbelül 0.9 legyen. Ekkor a $\Phi\left(\frac{k-22\,500}{75}\right) = 0.1$ vagy ami ezzel ekvivalens, a $\Phi\left(\frac{22\,500-k}{75}\right) = 0.9$ egyenletet kell kielégítenünk. A normális eloszlás-táblázat alapján $\frac{22\,500-k}{75} \sim 1.28$, ami azt jelenti, hogy $k = 22\,500 - 75 \times 1.28$ és $p = \frac{3}{4}$ esetén annak valószínűsége, hogy a fejdobások száma nagyobb mint $k = 22\,500 - 75 \times 1.28 = 22\,212$ és $p = \frac{3}{4}$ esetében annak valószínűsége, hogy legalább ennyi fejdobás történik körülbelül 0.9. Ha $p \ge \frac{3}{4}$, akkor ez a valószínűség nagyobb. Ezért a $k = 22\,212$ helyes választás.

5.) Ledobunk egymástól függetlenül 24 000 pontot a $[0, 2]$ intervallumra egyenletes eloszlással, (azaz annak a valószínűsége, hogy egy ledobott pont értéke x-nél kisebb $\frac{x}{2}$-vel egyenlő, ha $0 \le x \le 2$, eggyel egyenlő, ha $x \ge 2$, és nulla, ha $x \le 0$.) Őrizzük meg azokat a ledobott pontokat, melyek értéke 1-nél kisebb, és hagyjuk el azokat, melyek értéke, nagyobb mint egy. Mi annak a valószínűsége, hogy a megőrzött pontok értékeinek az összege 5900 és 6075 közé esik? Adjunk erre a valószínűségre jó közelítő becslést a mellékelt normális eloszlástáblázat segítségével.

Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 24\ 000$, valószínűségi változókat: $\xi_j = x$, ha a j-ik ledobott pont értéke x, és $0 \le x \le 1$, és $\xi_j = 0$, ha a j-ik ledobott pont értéke az $(1, 2]$ intervallumba esik. Ekkor a megőrzött pontok összege $S = \sum\limits_{j=1}^{24\ 000} \xi_j$, továbbá a ξ_j valószínűségi változók függetlenek és egyforma eloszlásúak. Ezért a centrális határeloszástétel segítségével jó becslést tudunk adni a minket érdeklő $P(5900 < S < 6075)$ valószínűségre. Ennek érdekében ki kell számolnunk a ξ_1 valószínűségi változó várható értékét és szórásnégyzetét.

A ξ_1 valószínűségi változó várható értékének és szórásnégyzetének a kiszámolása érdekében vezessük be az η_1 valószínűségi változót, amelyik megegyezik az első ledobott pont értékével, és a következő $h(x)$ függvényt a $[0,2]$ intervallumon: Legyen $h(x) = x$, ha $0 \le x \le 1$, és $h(x) = 0$, ha $1 \le x \le 2$. Ekkor $\xi_1 = h(\eta_1)$, és η_1 sűrűségfüggvénye $f(x) = \frac{1}{2}$, ha $0 \le x \le 2$, $f(x) = 0$, ha $x < 0$, és $x > 2$. Innen $E\xi_1 = Eh(\xi_1) = \int h(x)\,dx = \int_0^1 x\frac{1}{2}\,dx = \frac{1}{4}$, $E\xi_1^2 = Eh(\eta_1)^2 = \int_0^1 x^2\frac{1}{2}\,dx = \frac{1}{6}$, és $\operatorname{Var}\xi_1 = E\xi_1^2 - (E\xi_1)^2 = \frac{1}{6} - \frac{1}{16} = \frac{5}{48}$. Ezért $ES = 6000$, $\operatorname{Var} S = 2500$. Innen

$$
\begin{aligned}
P(5900 < S < 6075) &= P\left(-2 < \frac{S - ES}{\sqrt{\operatorname{Var} S}} < 1.5\right) \\
&\sim \Phi(1.5) - \Phi(-2) = \Phi(1.5) + \Phi(2) - 1.
\end{aligned}
$$

Megjegyzés: Egy más módon is kiszámíthatjuk a ξ_1 valószínűségi változó várható értékét. Ennek érdekében vegyük észre, hogy ugyan a ξ_1 valószínűségi változónak nincs sűrűségfüggvénye, és nem diszkrét eloszlású, viszont annak F eloszlásfüggvénye felírható $F(x) = F_1(x) + F_2(x)$ alakban, ahol $F_1(x)$-nek van sűrűségfüggvénye, és ez az $f(x) = \frac{1}{2}$ függvény a $[0,1]$ intervallumon, és $f(x) = 0$, ha $x < 0$ vagy $x > 1$, és $F_2(x)$ olyan mértéket határoz meg, amely a nullába van koncentrálva, és a nulla mértéke $\frac{1}{2}$. Pontosabban, tetszőleges A halmaz valószínűsége $P(A) = \int_A F(\,dx) = \int_A F_1(\,dx) + \int_A F_2(\,dx) = \int_A \frac{1}{2}\,dx + \int_A F_2(\,dx)$, ahol $\int_A F_2(\,dx) = \frac{1}{2}$, ha $0 \in A$, és $\int_A F_2(\,dx) = 0$, ha $0 \notin A$. Ekkor tetszőleges $h(x)$ függvényre $Eh(\xi) = \int h(x)F(\,dx) = \int h(x)F_1(\,dx) + \int h(x)F_2(\,dx) = \int_0^1 \frac{1}{2}h(x)\,dx + \frac{1}{2}h(0)$. Ezért

$$
E\xi_1 = \frac{1}{2}\int_0^1 x\,dx = \frac{1}{4}, \qquad E\xi_1^2 = \frac{1}{2}\int_0^1 x^2\,dx = \frac{1}{6},
$$

$\operatorname{Var}\xi_1 = \frac{1}{6} - \frac{1}{16} = \frac{5}{48}$.

6.) Legyen ξ egyenletes eloszlású valószínűségi változó a $[-1,1]$ intervallumon, azaz legyen sűrűségfüggvénye $f(x) = \frac{1}{2}$, ha $-1 \le x \le 1$, és $f(x) = 0$, ha $x > 1$ vagy $x < -1$. Számoljuk ki a ξ^4 valószínűségi változó várható értékét és szórásnégyzetét.

Megoldás: $\operatorname{Var}\xi^4 = E(\xi^4)^2 - (E\xi^4)^2 = E\xi^8 - (E\xi^4)^2 = \int_{-1}^1 \frac{1}{2}x^8\,dx - \left(\int \frac{1}{2}x^4\,dx\right)^2 = \frac{1}{9} - \left(\frac{1}{5}\right)^2 = \frac{16}{225}$.

Második megoldás. Számoljuk ki ξ^4 $F(x)$ eloszlás és $f(x)$ sűrűségfüggvényét.

$$
F(x) = P(\xi^4 < x) = P(-x^{1/4} < \xi < x^{1/4}) = x^{1/4}, \qquad \text{ha } 0 \le x \le 1
$$

$F(x) = 0$, ha $x \le 0$ és $F(x) = 1$, ha $x \ge 1$. Innen $f(x) = \frac{1}{4}x^{-3/4}$, ha $0 \le x \le 1$, $f(x) = 0$ egyébként. Ezért $E\xi = \int xf(x)\,dx = \int_0^1 \frac{1}{4}x^{1/4}\,dx = \frac{1}{4}\cdot\frac{4}{5} = \frac{1}{5}$, $E\xi^2 = \int x^2 f(x)\,dx = \int_0^1 \frac{1}{4}x^{5/4}\,dx = \frac{1}{4}\cdot\frac{4}{9} = \frac{1}{9}$, $\operatorname{Var}\xi = E\xi^2 - (E\xi)^2 = \frac{1}{9} - \left(\frac{1}{5}\right)^2 = \frac{16}{225}$.

7.) Legyen a ξ valószínűségi változó sűrűségfüggvénye $f(x) = e^{-|x|}$, ha $|x| \ge 1$, és $f(x) = \frac{1}{2} - \frac{1}{e}$, ha $|x| \le 1$. Mutassuk meg, hogy $f(x)$ valóban sűrűségfüggvény, és számoljuk ki a ξ valószínűségi változó várható értékét és szórásnégyzetét.

Megoldás: Annak belátásához, hogy $f(x)$ sűrűségfüggvény azt kell ellenőrizni, hogy egyrészt $f(x) \geq 0$ minden $-\infty < x < \infty$ számra, másrészt $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Az első állítás nyílvánvaló. A második állítás is igaz, mivel $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{-1} e^x\,dx + \int_{-1}^{1} \left(\frac{1}{2} - \frac{1}{e}\right)\,dx + \int_{1}^{\infty} e^{-x}\,dx = \frac{1}{e} + 2\left(\frac{1}{2} - \frac{1}{e}\right) + \frac{1}{e} = 1.$
Mivel $f(x)$ páros függvény, ezért $E\xi = \int_{-\infty}^{\infty} xf(x)\,dx = 0$, és $\mathrm{Var}\,\xi = E\xi^2 - (E\xi)^2 = E\xi^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx$. Ezért

$$\begin{aligned} Var\xi &= \int_{-\infty}^{-1} x^2 e^x\,dx + \int_{-1}^{1} \left(\frac{1}{2} - \frac{1}{e}\right) x^2\,dx + \int_{1}^{\infty} x^2 e^{-x}\,dx \\ &= 2\int_{1}^{\infty} x^2 e^{-x}\,dx + 2\int_{0}^{1} \left(\frac{1}{2} - \frac{1}{e}\right) x^2\,dx. \end{aligned}$$

Mivel parciális integrálással kapjuk, hogy

$$\begin{aligned} 2\int_{1}^{\infty} x^2 e^{-x}\,dx &= \left[-2x^2 e^{-x}\right]_1^{\infty} + \int_{1}^{\infty} 4x e^{-x}\,dx \\ &= \frac{2}{e} - \left[4xe^{-x}\right]_1^{\infty} + 4\int_{1}^{\infty} e^{-x}\,dx = \frac{2}{e} + \frac{4}{e} + \frac{1}{e} = \frac{7}{e}, \end{aligned}$$

és

$$2\int_{0}^{1} \left(\frac{1}{2} - \frac{1}{e}\right) x^2\,dx = \frac{1}{3}\left(1 - \frac{2}{e}\right)\left[x^3\right]_0^1 = \left(\frac{1}{3} - \frac{2}{3e}\right),$$

ezért $\mathrm{Var}\,\xi = \frac{1}{3} + \frac{19}{3e}$.

Házi feladat

Ledobunk a $[0,3]$ intervallumra 24 000 pontot egyenletes eloszlással, azaz egy ledobott pont egy valószínűséggel a $[0,3]$ intervalumba esik, és annak valószínűsége, hogy a ledobott pont egy $[a,b] \subset [0,3]$ intervallumba esik $\frac{b-a}{3}$. Ha egy pont a $[0,1]$ intervallumba esik, akkor felírjuk a ledobott pont helyének pontos értékét egy jegyzőkönyvbe, ha a pont az $(1,2]$ intervalumba esik, akkor az 1 számot, ha pedig a $(2,3]$ intervallumba esik, akkor a 2 számot írjuk a jegyzőkönyvbe. Adjunk becslést a mellékelt normális eloszlástáblázat segítségével annak valószínűségére, hogy a jegyzőkönyvbe írt 24 000 szám összege 27 900 és 28 150 közé esik.