Villamosmérnök A4

9. gyakorlat (2012. 11. 12-13.)

Kétdimenziós normális eloszlás, várható érték és szórás tulajdonságai

Egy kétdimenziós (X, Y) valószínűségi változó függvényének várható értéke, sűrűségfüggvénnyel kifejezve:

$$\mathbb{E}(t(X,Y)) = \iint_{\mathbb{R}^2} t(x,y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

ahol $t : \mathbb{R}^2 \to \mathbb{R}$.

Kovariancia: $\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$; és a **kovariancia mátrix** (kétdimenziós eset):

$$\Sigma = \begin{pmatrix} \mathbb{V}\mathrm{ar}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathbb{V}\mathrm{ar}(Y) \end{pmatrix}$$

Korrelációs együttható: $\mathrm{r}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathbb{SD}(X)\mathbb{SD}(Y)}$.

Kétdimenziós normális eloszlás
Standard kétdimenziós normális eloszlású egy (U, V) pár, ha sűrűségfüggvénye

$$\varphi(u,v) := \frac{1}{2\pi}\exp\left(-\frac{u^2+v^2}{2}\right) \qquad (u, v \in \mathbb{R}).$$

(U, V) zérus várható érték vektorú, egységmátrix kovarianciájú pár. Azt mondjuk, hogy az $(X, Y) = (U, V)\mathbf{A} + \mu$ pár kétdimenziós normális eloszlású $\mu \in \mathbb{R}^2$ várható érték vektorral és $\Sigma \in \mathbb{R}^{2\times 2}$ (invertálható) kovariancia mátrixszal, ha (U, V) standard normális pár, és $\mathbf{A}^\top\mathbf{A} = \Sigma$, ahol $\mathbf{A} \in \mathbb{R}^{2\times 2}$.
Egy kétdimenziós normális (X, Y) pár sűrűségfüggvénye:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2r\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\},$$

ahol $\mathbb{E}X = \mu_X$, $\mathbb{E}Y = \mu_Y$, $\mathbb{SD}(X) = \sigma_X$, $\mathbb{SD}(Y) = \sigma_Y$, és $r = r(X, Y)$ X és Y korrelációs egyyüthatója. Kétdimenziós normális eloszlás esetében a perem- illetve feltételes eloszlások is normálisok.

1. Határozzuk meg az alábbi mennyiségeket:

   (a) $\mathrm{r}(X, Y)$, $\mathbb{SD}(X + Y)$, $\mathbb{SD}(2X - 3Y)$;

   (b) $\mathbb{E}(X - 3Y)^2$, $\mathbb{E}(XY + Y^2)$;

   feltéve, hogy $\mathrm{Cov}(X, Y) = 2$, $\mathbb{SD}(X) = 3$, $\mathbb{SD}(Y) = 4$, $\mathbb{E}X = -1$, $\mathbb{E}Y = 1$.

   Megoldás:

   (a) $\mathrm{r}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathbb{SD}(X)\mathbb{SD}(Y)} = \frac{2}{3\times 4} = \frac{1}{6}$; $\mathbb{SD}(X+Y) = \sqrt{\mathbb{V}\mathrm{ar}(X+Y)} = \sqrt{\mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y) + 2\mathrm{Cov}(X,Y)} = \sqrt{9+16+2\times 2} = \sqrt{29}$; $\mathbb{SD}(2X-3Y) = \sqrt{\mathbb{V}\mathrm{ar}(2X-3Y)} = \sqrt{4\mathbb{V}\mathrm{ar}(X) + 9\mathbb{V}\mathrm{ar}(Y) - 2\mathrm{Cov}(2X,3Y)} = \sqrt{4\times 9 + 9\times 16 - 2\times 2\times 3\times 2} = \sqrt{204}$;

   (b) $\mathbb{E}(X-3Y)^2 = \mathbb{E}(X^2) - 6\mathbb{E}(XY) + 9\mathbb{E}(Y^2) = 10 - 6\times 1 + 9\times 17 = 157$; $\mathbb{E}(XY+Y^2) = \mathbb{E}(XY) + \mathbb{E}(Y^2) = 1 + 17 = 18$; mivel $\mathbb{E}(X^2) = \mathbb{V}\mathrm{ar}(X) + (\mathbb{E}(X))^2 = 9 + 1 = 10$, $\mathbb{E}(Y)^2 = \mathbb{V}\mathrm{ar}(Y) + (\mathbb{E}(Y))^2 = 16 + 1 = 17$, illetve $\mathbb{E}(XY) = \mathrm{Cov}(X,Y) + \mathbb{E}(X)\mathbb{E}(Y) = 2 + (-1)\times 1 = 1$.

2. Két kockával dobva, mi a dobott számok

   (a) összegének és különbségének

   (b) minimumának és maximumának;

   (c) összegének és különbség négyzetének;

   kovarianciája? Számítsuk ki a perem várható értékeket és szórásokat is!

   Megoldás: Jelölje X, Y a két független kockadobás eredményét. Ekkor

   (a) $\mathrm{Cov}(X+Y, X-Y) = \mathbb{E}\left[(X+Y)(X-Y)\right] - \mathbb{E}(X+Y)\mathbb{E}(X-Y) = 0$, mivel $(X+Y)(X-Y) = X^2 - Y^2$ és X-, Y-nak ugyanaz az eloszlása. Így a kovariancia (és így a korreláció is) 0, DE: $X+Y$ nem független $X-Y$-től! $\mathbb{E}(X+Y) = 2\times\frac{7}{2} = 7$ és függetlenség miatt $\mathbb{V}\mathrm{ar}(X\pm Y) = \mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y) = 2\times\frac{35}{12} = \frac{35}{6}$.

   (b) Vegyük a két valószínűségi változó együttes eloszlását (ezt a 8. feladatsorban már meghatároztuk de) ismétlésként ez:

   $$\mathbb{P}(\min(X,Y) = i, \max(X,Y) = j) = \begin{cases} \frac{1}{36}, & \text{ha } i = j; \\ \frac{1}{18}, & \text{ha } i < j. \\ 0, & \text{egyébként;} \end{cases}$$

Ebből

$$\mathbb{E}(\min(X,Y)\max(X,Y)) = \sum_{i=1}^{6}\sum_{j=i}^{6} ij\,\mathbb{P}(\min(X,Y)=i, \max(X,Y)=j) =$$
$$= 1 + \sum_{i=1}^{5}\left(\frac{i^2}{36} + \frac{i}{18}\sum_{j=i+1}^{6} j\right) = 1 + \sum_{i=1}^{5}\frac{i^2}{36} + \frac{i}{18}(21 - i(i+1)/2) =$$
$$= 1 + \frac{1}{36}\sum_{i=1}^{5} 42i - i^3 = 1 + \frac{45}{4} = \frac{7^2}{2^2} = \frac{49}{4} = 12.25.$$

A marginálisok várható értéke: $\mathbb{E}(\min(X,Y)) = 1\frac{1+5\times 2}{36} + 2\frac{1+4\times 2}{36} + 3\frac{1+3\times 2}{36} + 4\frac{1+2\times 2}{36} + 5\frac{1+1\times 2}{36} + 6\frac{1+0\times 1}{36} = \frac{11+18+21+20+15+6}{36} = \frac{91}{36} \approx 2.53$. Valamint $\mathbb{E}(\max(X,Y)) = 1\frac{1}{36} + 2\frac{3}{36} + 3\frac{5}{36} + 4\frac{7}{36} + 5\frac{9}{36} + 6\frac{11}{36} = \frac{161}{36} \approx 4.47$. Azaz $\mathrm{Cov}(\max(X,Y), \min(X,Y)) = \frac{49}{4} - \frac{91}{36}\frac{161}{36} = \frac{1225}{1296} \approx 0.945$.
Szórásokhoz a második momentumokat számoljuk ki: $\mathbb{E}(\min(X,Y)^2) = 1^2\frac{11}{36} + 2^2\frac{9}{36} + 3^2\frac{7}{36} + 4^2\frac{5}{36} + 5^2\frac{3}{36} + 6^2\frac{1}{36} = \frac{301}{36}$, azaz $\mathbb{V}\mathrm{ar}(\min(X,Y)) = \frac{301}{36} - \frac{91^2}{36^2} = \frac{2555}{1296} \approx 1.97$.
$\mathbb{E}(\max(X,Y)^2) = 1^2\frac{1}{36} + 2^2\frac{3}{36} + 3^2\frac{5}{36} + 4^2\frac{7}{36} + 5^2\frac{9}{36} + 6^2\frac{11}{36} = \frac{791}{36}$, azaz $\mathbb{V}\mathrm{ar}(\max(X,Y)) = \frac{791}{36} - \frac{161^2}{36^2} = \frac{2555}{1296} \approx 1.97$.
Másik megoldás: (rövidebb és érdemes legyűrni)
A következő azonosságokat vegyük észre: $\max(X,Y) + \min(X,Y) = X + Y$, illetve $\max(X,Y) - \min(X,Y) = |X - Y|$, így

$$\mathrm{Cov}(\max(X,Y), \min(X,Y)) = \mathrm{Cov}\left(\frac{1}{2}(X+Y+|X-Y|), \frac{1}{2}(X+Y-|X-Y|)\right) =$$
$$= \frac{1}{4}(\mathbb{V}\mathrm{ar}(X+Y) - \mathbb{V}\mathrm{ar}(|X-Y|)) = \frac{1}{4}(\mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y) - \mathbb{E}(X-Y)^2 + (\mathbb{E}|X-Y|)^2) =$$
$$= \frac{1}{4}(\mathbb{E}|X-Y|)^2 = \frac{1}{4}\frac{35^2}{18^2} = \frac{1225}{1296} \approx 0.945,$$

ahol rendre a következőket használtuk: a fenti azonosságot, a kovariancia szimmetria tulajdonságát ($\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$) és linearitását ($\mathrm{Cov}(aX+bY, cZ+dU) = ac\mathrm{Cov}(X,Z) + ad\mathrm{Cov}(X,U) + bc\mathrm{Cov}(Y,Z) + bd\mathrm{Cov}(Y,U)$ minden X,Y,Z,U valószínűségi változóra és $a,b,c,d \in \mathbb{R}$ számra), X,Y függetlenségét, a variancia definícióját és azt, hogy X és Y azonos eloszlású (így pl. $\mathbb{E}X = \mathbb{E}Y = 7/2$); végül $\mathbb{E}|X-Y| = 1\times\frac{10}{36} + 2\times\frac{8}{36} + 3\times\frac{6}{36} + 4\times\frac{4}{36} + 5\times\frac{2}{36} = \frac{35}{18}$.
De ebből az is adódik, hogy $\mathbb{E}(\min(X,Y)) = \frac{1}{2}(\mathbb{E}X + \mathbb{E}Y - \mathbb{E}|X-Y|) = \frac{7}{2} - \frac{35}{36} = \frac{91}{36}$. $\mathbb{E}(\max(X,Y)) = \mathbb{E}(X+Y-\min(X,Y)) = 7 - \frac{91}{36} = \frac{161}{36}$. Szórásokat hasonlóan lehet számolni, fenti azonosságokból az is adódik, hogy $\mathbb{V}\mathrm{ar}(\max(X,Y)) = \mathbb{V}\mathrm{ar}(\min(X,Y))$.

(c) $\mathrm{Cov}(X+Y,(X-Y)^2) = \mathrm{Cov}(X+Y, X^2-2XY+Y^2) = \mathrm{Cov}(X,X^2) + \mathrm{Cov}(Y,Y^2) + \mathrm{Cov}(X,Y^2) + \mathrm{Cov}(Y,X^2) - 2\mathrm{Cov}(X,XY) - 2\mathrm{Cov}(Y,XY)$. Szimmetria és függetlenség miatt: $\mathrm{Cov}(X,X^2) = \mathrm{Cov}(Y,Y^2)$, $\mathrm{Cov}(Y,X^2) = \mathrm{Cov}(X,Y^2) = 0$, és $\mathrm{Cov}(X,XY) = \mathrm{Cov}(Y,XY)$. Emiatt (és megint X,Y függetlenségét használva): $\mathrm{Cov}(X+Y,(X-Y)^2) = 2(\mathbb{E}X^3 - \mathbb{E}X\mathbb{E}X^2) - 4(\mathbb{E}X^2\mathbb{E}Y - (\mathbb{E}X)^2\mathbb{E}Y)$. Tudjuk korábbról, hogy $\mathbb{V}\mathrm{ar}(X) = \frac{35}{12}$, $\mathbb{E}X = \mathbb{E}Y = \frac{7}{2}$ és $\mathbb{E}X^3 = (1^3+2^3+3^3+4^3+5^3+6^3)\frac{1}{6} = \frac{147}{2}$. Összerakva ezeket: $\mathrm{Cov}(X+Y,(X-Y)^2) = 2\left(\frac{147}{2} - \frac{7}{2}\frac{91}{6}\right) - 4\frac{35}{12}\frac{7}{2} = 0$. Ez meglepetés volt, de tartsuk észben továbbra is: $X+Y$ nem független $X-Y$-tól!
$\mathbb{E}(X-Y)^2 = \mathbb{E}(X^2) - 2\mathbb{E}(XY) + \mathbb{E}(Y^2) = 2\mathbb{E}(X^2) - 2(\mathbb{E}X)^2 = 2\mathbb{V}\mathrm{ar}(X) = \mathbb{V}\mathrm{ar}(X+Y) = \frac{35}{6} \approx 5.83$.
$\mathbb{E}(X-Y)^4 = 2\mathbb{E}(X^4) - 8\mathbb{E}(X^3)\mathbb{E}(X) + 6\mathbb{E}(X^2)\mathbb{E}(X^2) = 2\times\frac{2275}{6} - 8\times\frac{147}{2}\frac{7}{2} + 6\times\frac{91^2}{6^2} = \frac{161}{2} = 80.5$ (ellenőrizzük), így $\mathbb{V}\mathrm{ar}((X-Y)^2) = \mathbb{E}(X-Y)^4 - (\mathbb{E}(X-Y)^2)^2 = \frac{1673}{36} \approx 46.47$.

3. Legyenek X_1, X_2, X_3, X_4 páronként korrelálatlan valószínűségi változók, melyek közös szórása 1. Számoljuk ki a következő párok korrelációs együtthatóit:

(a) $X_1 + X_2$ és $X_2 + X_3$;

(b) $X_1 + X_2$ és $X_3 + X_4$

Megoldás: A kovariancia linearitását használjuk mindkét esetben:

(a) $\mathrm{Cov}(X_1+X_2, X_2+X_3) = \mathrm{Cov}(X_1,X_2) + \mathrm{Cov}(X_1,X_3) + \mathrm{Cov}(X_2,X_2) + \mathrm{Cov}(X_2,X_3) = 0 + 0 + \mathbb{V}\mathrm{ar}(X_2) + 0 = 1$, mivel páronként korrelátlanok voltak az X_i'k és 1 volt a szórásuk.

(b) Hasonlóan az előbbiekhez: $\mathrm{Cov}(X_1+X_2, X_3+X_4) = \mathrm{Cov}(X_1,X_2) + \mathrm{Cov}(X_1,X_3) + \mathrm{Cov}(X_2,X_3) + \mathrm{Cov}(X_3,X_4) = 0+0+0+0 = 1$.

4. Szabályos kockával dobunk tízszer, majd felírjuk hányszor dobtunk 6-ot, ezt jelölje X, illetve hányszor dobtunk 1-et, ezt meg Y jelöli. Mi (X,Y) kovarianciája, korrelációs együtthatója?

Megoldás: A feladat megoldásának a kulcsa az, hogy az $X|Y=j$ feltételes eloszlás binomiális $n = 10-j$, illetve $p = \frac{1}{5}$ paraméterrel (ahol $0 \le j \le 10$ egy rögzített természetes szám). Ennek a várható értéke tehát:

$np = \frac{10-j}{5}$. Azaz

$$\begin{aligned}\mathbb{E}(XY) &= \sum_{i,j=0}^{10} ij\mathbb{P}(X=i, Y=j) = \sum_{j=0}^{10} j\left[\sum_{i=0}^{10-j} i\mathbb{P}(X=i \,|\, Y=j)\right]\mathbb{P}(Y=j) = \\ &= \sum_{j=0}^{10} j\frac{10-j}{5}\mathbb{P}(Y=j) = 2\mathbb{E}Y - \frac{1}{5}\mathbb{E}Y^2 = 2\times\frac{10}{6} - \frac{1}{5}\frac{25}{6} = \frac{5}{2},\end{aligned}$$

ahol kihasználtuk az XY valószínűségi változó várható értékére vonatkozó képletet, a feltételes eloszlás definícióját, valamint azt, hogy a marginálisok is binomiális eloszlásúak $n = 10$, illetve $p = \frac{1}{6}$ paraméterrel, azaz $\mathbb{E}X = \mathbb{E}Y = \frac{10}{6}$ és $\mathbb{E}X^2 = \mathbb{E}Y^2 = 10 \times \frac{1}{6}\frac{5}{6} + \frac{10^2}{6^2} = \frac{25}{6}$. Vagyis:

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y = \frac{5}{2} - \frac{10^2}{6^2} = -\frac{5}{18},$$

azaz negatívan korrelált X és Y, ami intuitíve világos is, hiszen minél több a 6-os dobások száma, annál kevesebb 1-es dobást látunk nagy valószínűséggel.

5. Legyenek X és Y független valószínűségi változók közös μ várható értékkel, de különböző σ_X és σ_Y szórásokkal. μ értékét nem tudjuk, de egy mintavétel alapján az X és Y súlyozott átlagával szeretnénk megbecsülni. Azaz: μ értékére a $\lambda X + (1-\lambda)Y$ becslést adjuk, valamilyen $0 < \lambda < 1$ paraméterrel. Hogyan válasszuk λ-t, hogy a becslésünk szórása minimális legyen?

   Megoldás: A feladat kívánsága az, hogy a $\sigma(\lambda) := \mathbb{V}\mathrm{ar}(\lambda X + (1-\lambda)Y)$ mennyiséget, mint λ függvényét minimalizáljuk. Ehhez használjuk a variancia (szórásnégyzet) megfelelő tulajdonságait, illetve X, Y függetlenségét, azaz $\sigma(\lambda) = \lambda^2\mathbb{V}\mathrm{ar}(X) + (1-\lambda)^2\mathbb{V}\mathrm{ar}(Y) = \lambda^2(\sigma_X + \sigma_Y) - 2\lambda\sigma_Y + \sigma_Y$. Ez egy mosolygós parabola λ-ban, minimumhelye a két zérushelyének átlaga, azaz $\lambda_{\min} = \frac{\sigma_Y}{\sigma_X + \sigma_Y}$.

6. Egy kaszinóban a következő játékot játszák: két játékos (egymástól függetlenül) gondol egy nemnegatív 1-nél kisebb véletlen számra. Ezt követően, Tőlük függetlenül, a Bank is gondol egy ilyen véletlen számra, majd az a játékos nyer, akinek a száma nagyobb volt a Bank számánál. Legyen X és Y az egyes játékosok nyerésének Bernoulli valószínűségi változója. Mi (X, Y) kovarianciája, korrelációs együtthatója? Értelmezze a kapott eredményt!

   Megoldás: Jelölje a három – egymástól függetlenül – sorsolt $(0,1)$-beli valós számot: X_1 (1. játékos), X_2 (2. játékos), illetve X_3 (Bank). A szövegben szereplő X (Y) valószínűségi változó expliciten kiírva:

$$X\ (Y) = \begin{cases} 1, & \text{ha } X_1 > X_3 \ (X_2 > X_3); \\ 0, & \text{egyébként.} \end{cases}$$

   Szimmetria miatt $\mathbb{P}(X_1 > X_3) = \mathbb{P}(X_2 > X_3)$, így egyrészt X és Y szórása megegyezik, azaz: $\mathbb{SD}(Y) = \mathbb{SD}(X) = \sqrt{\mathbb{P}(X_1 > X_3)(1 - \mathbb{P}(X_1 > X_3))}$, másrészt kovarianciájuk:

$$\mathrm{Cov}(X,Y) = \mathbb{P}(X_1 > X_3, X_2 > X_3) - \mathbb{P}(X_1 > X_3)\mathbb{P}(X_2 > X_3) = \mathbb{P}(X_1 > X_3, X_2 > X_3) - \mathbb{P}(X_1 > X_3)^2.$$

   Kiszámoljuk a különbség mindkét tagját; X_1, X_2, X_3 együttes sűrűségfüggvénye az egységkockán konstans 1, egyébként zérus (3 független egyenletes szorzata), ezért

$$\begin{aligned}&\mathbb{P}(X_1 > X_3, X_2 > X_3) = \iiint\limits_{x_1 > x_3, x_2 > x_3} f_{X_1,X_2,X_3}(x_1,x_2,x_3)\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int\limits_{x_3=0}^{1}\int\limits_{x_1=x_3}^{1}\int\limits_{x_2=x_3}^{1} \mathrm{d}x_2\mathrm{d}x_1\mathrm{d}x_3 = \\ &= \int_{x_3=0}^{1}(1-x_3)^2\,\mathrm{d}x_3 = \left.\frac{-(1-x_3)^3}{3}\right]_0^1 = \frac{1}{3}, \text{ és} \\ &\mathbb{P}(X_1 > X_3) = \iint\limits_{x_1 > x_3} f_{X_1,X_3}(x_1,x_3)\,\mathrm{d}x_1\mathrm{d}x_3 = \int\limits_{x_3=0}^{1}\int\limits_{x_1=x_3}^{1}\mathrm{d}x_1\mathrm{d}x_3 = \int\limits_{x_3=0}^{1}(1-x_3)\,\mathrm{d}x_3 = \left.\frac{-(1-x_3)^2}{2}\right]_0^1 = \frac{1}{2}.\end{aligned}$$

   azaz $\mathrm{Cov}(X,Y) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$, így $\mathrm{r}(X,Y) = \frac{1/12}{(1/2)^2} = \frac{1}{3}$. Vagyis pozitívan korrelált a nyerések esélye, ami természetes, hiszen feltéve, hogy az egyik játékos nyert, az csak növeli az esélyét annak, hogy a második játékos is nyerni fog.

7. Egy kisváros durván négyzet alakú, a négyzet oldalai 3 kilométer hosszúak. A város $(0,0)$ középpontjában van a kórház, és a város utcái négyzetháló szerűek. Ezért, ha a város (x, y) pontján történik egy baleset, a mentőnek $|x| + |y|$ távolságot kell megtennie a balesettől a kórházig. Ha egy baleset a városon belül egyenletes eloszlású helyen következik be, számoljuk ki a betegszállítás várható hosszát! Hogyan változik az eredmény, ha a távolság négyzetével arányos hosszú utat kell megtennie a mentőnek?

   Megoldás: A feladat szövege alapján a baleset helye modellezhető egy (X, Y) párral, melyek együttes eloszlása egyenletes a $[-1.5, 1.5] \times [-1.5, 1.5]$ négyzeten, azaz az együttes sűrűségfüggvény itt $\frac{1}{3^2}$, egyébként zérus és az is rögtön adódik (mivel téglalapon vagyunk), hogy X és Y (a marginálisok) függetlenek egymástól, mindkettő egyenletes eloszlású a $[-1.5, 1.5]$ intervallumon. Azaz a betegszállítás várható hossza:

$\mathbb{E}(|X|+|Y|) = \mathbb{E}|X| + \mathbb{E}|Y| = 2\mathbb{E}|X| = 2\int_{-1.5}^{1.5} |x|\frac{1}{3}\,\mathrm{d}x = \frac{2}{3}1.5 \times 1.5 = \frac{3}{2} = 1.5$, utóbbi levezetésben rendre a következőket használtuk: a várható érték lineáris (összeg várható értéke a várható értékek összege), X és Y azonos eloszlású (a függetlenség itt nem kell, az ahhoz kellett, hogy X és Y is egyenletes). Az utolsó kérdésre pedig $\mathbb{E}(X^2+Y^2)$-et kell kiszámolni (ami épp $2\mathbb{V}\mathrm{ar}(X)$), tehát $\mathbb{E}X^2 = \int_{-1.5}^{1.5} x^2\frac{1}{3}\,\mathrm{d}x = \left.\frac{x^3}{9}\right]_{x=-1.5}^{1.5} = \frac{3}{4}$, azaz a keresett érték $2 \times \frac{3}{4} = \frac{3}{2} = 1.5$, ami csodának csodájára épp az előbbi eredménnyel esik egybe.

8. Legyen U egyenletes eloszlású véletlen szám a $(-1,1)$ intervallomból. Mi a várható értéke a $\sqrt{U}$, U^2, illetve U^3 valószínűségi változónak?

Megoldás: Tudjuk, hogy $F_U(u) = \begin{cases} 0, & \text{ha } u \le -1; \\ \frac{u+1}{2}, & \text{ha } -1 < u \le 1; \\ 1, & \text{egyébként.} \end{cases}$.
Mivel $-1 \le U \le 1$, ezért $\sqrt{U}$ nem valós, várható értékét nem tudjuk meghatározni.
A többit viszont kiszámoljuk: $\mathbb{E}(U^2) = \int_{u=-1}^{1} u^2\frac{1}{2}\,\mathrm{d}u = \left.\frac{u^3}{6}\right]_{u=-1}^{1} = \frac{1}{3}$.
$\mathbb{E}(U^3) = \int_{u=-1}^{1} u^3\frac{1}{2}\,\mathrm{d}u = \left.\frac{u^4}{8}\right]_{u=-1}^{1} = 0$. Avagy még egyszerűbben U^3 eloszlása ugyanaz, mint $(-U)^3$-é, ezért 0 a várható értéke.
Másik megoldás: (hosszadalmasabb, de azért tanulságos)
U^2 eloszlásának a tartója $(0,1)$, U^3-é $(-1,1)$, előbbi eloszlásfüggvénye: $F_{U^2}(x) = \mathbb{P}(U^2 < x) = \mathbb{P}(-\sqrt{x} < U < \sqrt{x}) = F_U(\sqrt{x}) - F_U(-\sqrt{x}) = \sqrt{x}$, ha $0 \le x \le 1$, egyébként ha $x < 0$, $F_{U^2}(x) = 0$, és ha $x > 1$, akkor pedig $F_{U^2}(x) = 1$. Azaz U^2 sűrűségfüggvénye $f_{U^2}(x) = \frac{1}{2\sqrt{x}}$, ha $0 < x < 1$, egyébként pedig zérus, így $\mathbb{E}(U^2) = \int_{x=0}^{1} xf_{U^2}(x)\,\mathrm{d}x = \frac{1}{2}\int_{x=0}^{1}\sqrt{x}\,\mathrm{d}x = \left.\frac{x^{3/2}}{3}\right]_{x=0}^{1} = \frac{1}{3}$.
U^3 eloszlásfüggvénye: $F_{U^3}(x) = \mathbb{P}(U^3 < x) = \mathbb{P}(U < x^{1/3}) = F_U\left(x^{1/3}\right) = \frac{x^{1/3}+1}{2}$, ha $-1 \le x \le 1$, egyébként ha $x < -1$, $F_{U^3}(x) = 0$, és ha $x > 1$, akkor pedig $F_{U^3}(x) = 1$. Azaz U^3 sűrűségfüggvénye $f_{U^3}(x) = \frac{1}{6x^{2/3}}$, ha $-1 < x < 1$, egyébként pedig zérus, így $\mathbb{E}(U^3) = \int_{x=-1}^{1} xf_{U^3}(x)\,\mathrm{d}x = \frac{1}{6}\int_{x=-1}^{1} x^{1/3}\,\mathrm{d}x = \left.\frac{x^{4/3}}{8}\right]_{x=-1}^{1} = 0$.

9. Magyarországon egy ember egy véletlen $\sqrt{Z}$ ideig él, ahol Z exponenciális eloszlású valószínűségi változó valamilyen ismeretlen λ paraméterrel. Mi λ értéke, ha átlagosan 67 évig élünk?

Megoldás: Z exponenciális eloszlású, azaz ha $x \ge 0$, $\mathbb{P}(Z > x) = \exp(-\lambda x)$, egyébként 1. Határozzuk meg $\sqrt{Z}$ eloszlását. Ha $x \ge 0$, akkor $\mathbb{P}(\sqrt{Z} > x) = \mathbb{P}(Z > x^2) = \exp(-\lambda x^2)$, egyébként 1, mivel Z (így $\sqrt{Z}$) sem negatív. Azaz $\sqrt{Z}$ sűrűségfüggvénye $f_{\sqrt{Z}}(x) = \lambda 2x\exp(-\lambda x^2)$, ha $x \ge 0$, egyébként zérus. Ebből:

$$\begin{aligned}\mathbb{E}(\sqrt{Z}) &= \int_{x=0}^{+\infty} xf_{\sqrt{Z}}(x)\,\mathrm{d}x = \lambda 2\int_{x=0}^{+\infty} x^2\exp(-\lambda x^2)\,\mathrm{d}x = \lambda\int_{x=-\infty}^{+\infty} x^2\exp(-\lambda x^2)\,\mathrm{d}x = \\ &= \sqrt{\pi\lambda}\int\limits_{x=-\infty}^{+\infty} x^2\frac{1}{\sqrt{2\pi}\frac{1}{\sqrt{2\lambda}}}\exp\left(-\frac{x^2}{2\left(\frac{1}{\sqrt{2\lambda}}\right)^2}\right)\mathrm{d}x = \sqrt{\pi\lambda}\frac{1}{2\lambda} = \frac{\sqrt{\pi}}{2\sqrt{\lambda}}.\end{aligned}$$

A 3. egyenlőség előbb abból következik, hogy az integrandus pozitív integrálható páros függvény. Ellenőrizzük, hogy a 4. egyenlőségnél ekvivalens átalakításokat végeztünk csak. Végül az utolsó előtti egyenlőség abból következik, hogy a képletben utoljára felírt integrál tulajdonképpen egy 0 várható értékű és $\frac{1}{\sqrt{2\lambda}}$ szórású normális eloszlású valószínűségi változónak a második momentuma, ami – a 0 várható értékűség miatt – a szórásnégyzet, azaz $\frac{1}{2\lambda}$.
Végül λ meghatározásához kell $67 = \frac{\sqrt{\pi}}{2\sqrt{\lambda}}$, azaz $\lambda = \frac{\pi}{134^2} \approx 0.000175$.
Másik megoldás: (ha tudunk integrálni, akkor ez a rövidebb)

$$\mathbb{E}(\sqrt{Z}) = \int_{z=0}^{+\infty}\sqrt{z}\lambda\exp(-\lambda z)\,\mathrm{d}z \underset{y:=\sqrt{z}}{=} 2\lambda\int_{y=0}^{+\infty} y^2\exp(-\lambda y^2)\,\mathrm{d}y,$$

ahol az integrálhelyettesítésnél $y^2 = z$, azaz $2y\,\mathrm{d}y = \mathrm{d}z$. Innentől már ismerős a dolog, a megoldás hasonló az előbbiekhez. Az első megoldás tehát egy szép valszámos szemléltetése annak, hogy mi zajlik le egy integrálhelyettesítésnél.

10. Az (X,Y) együttes sűrűségfüggvénye az alábbiak valamelyike:

(a) $f(x,y) = \begin{cases} 60xy^2, & \text{ha } 0 \le x \le 1,\ 0 \le y \le 1-x; \\ 0, & \text{egyébként.} \end{cases}$

(b) $g(x,y) = \begin{cases} \frac{2}{x}\exp(-2x), & \text{ha } 0 < x,\ 0 \le y \le x; \\ 0, & \text{egyébként.} \end{cases}$

(c) $h(x,y) = \begin{cases} \frac{1}{y}\exp(-y-x/y), & \text{ha } 0 < x, y; \\ 0, & \text{egyébként.} \end{cases}$

Határozzuk meg X, Y várható értékét, szórását illetve kovarianciájukat!

Megoldás:

(a) Kovarianciához kellenek a várható értékek, ezekkel kezdjük:

$$\begin{aligned}\mathbb{E}X &= \int_{x=0}^{1}\int_{y=0}^{1-x} x\,60xy^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^2\int_{y=0}^{1-x} y^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^2\frac{(1-x)^3}{3}\,\mathrm{d}x = \\ &= 20\int_{x=0}^{1}(-x^5+3x^4-3x^3+x^2)\,\mathrm{d}x = 20\left(-\frac{1}{6}+3\frac{1}{5}-3\frac{1}{4}+\frac{1}{3}\right) = \frac{20}{60} = \frac{1}{3}.\end{aligned}$$

$$\begin{aligned}\mathbb{E}Y &= \int_{x=0}^{1}\int_{y=0}^{1-x} y\,60xy^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x\int_{y=0}^{1-x} y^3\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x\frac{(1-x)^4}{4}\,\mathrm{d}x = \\ &= 15\int_{x=0}^{1}(x^5-4x^4+6x^3-4x^2+x)\,\mathrm{d}x = 15\left(\frac{1}{6}-4\frac{1}{5}+6\frac{1}{4}-4\frac{1}{3}+\frac{1}{2}\right) = \frac{15}{30} = \frac{1}{2}.\end{aligned}$$

$$\begin{aligned}\mathbb{E}(XY) &= \int_{x=0}^{1}\int_{y=0}^{1-x} xy\,60xy^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^2\int_{y=0}^{1-x} y^3\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^2\frac{(1-x)^4}{4}\,\mathrm{d}x = \\ &= 15\int_{x=0}^{1}(x^6-4x^5+6x^4-4x^3+x^2)\,\mathrm{d}x = 15\left(\frac{1}{7}-4\frac{1}{6}+6\frac{1}{5}-4\frac{1}{4}+\frac{1}{3}\right) = \frac{15}{105} = \frac{1}{7}.\end{aligned}$$

Azaz $\mathrm{Cov}(X,Y) = \frac{1}{7} - \frac{1}{3}\frac{1}{2} = -\frac{1}{42}$. A szórásnégyzetekhez meghatározzuk a második momentumokat:

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{x=0}^{1}\int_{y=0}^{1-x} x^2\,60xy^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^3\int_{y=0}^{1-x} y^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x^3\frac{(1-x)^3}{3}\,\mathrm{d}x = \\ &= 20\int_{x=0}^{1}(-x^6+3x^5-3x^4+x^3)\,\mathrm{d}x = 20\left(-\frac{1}{7}+3\frac{1}{6}-3\frac{1}{5}+\frac{1}{4}\right) = \frac{20}{140} = \frac{1}{7}.\end{aligned}$$

$$\begin{aligned}&\mathbb{E}(Y^2) = \int_{x=0}^{1}\int_{y=0}^{1-x} y^2\,60xy^2\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x\int_{y=0}^{1-x} y^4\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{1} 60x\frac{(1-x)^5}{5}\,\mathrm{d}x = \\ &= 12\int_{x=0}^{1}(-x^6+5x^5-10x^4+10x^3-5x^2+x)\,\mathrm{d}x = 12\left(-\frac{1}{7}+5\frac{1}{6}-10\frac{1}{5}+10\frac{1}{4}-5\frac{1}{3}+\frac{1}{2}\right) = \frac{12}{42} = \frac{2}{7}.\end{aligned}$$

Végül: $\mathbb{SD}(X) = \sqrt{\mathbb{V}\mathrm{ar}(X)} = \sqrt{\frac{1}{7}-\left(\frac{1}{3}\right)^2} = \frac{\sqrt{2}}{\sqrt{63}} = \frac{\sqrt{2}}{3\sqrt{7}}$, és $\mathbb{SD}(Y) = \sqrt{\mathbb{V}\mathrm{ar}(Y)} = \sqrt{\frac{2}{7}-\left(\frac{1}{2}\right)^2} = \frac{1}{\sqrt{28}} = \frac{1}{2\sqrt{7}}$. Korrelációs együttható: $\mathrm{r}(X,Y) = -\frac{1/42}{(\sqrt{2}/(3\sqrt{7}))(1/(2\sqrt{7}))} = -\frac{1}{\sqrt{2}}$.

(b) Várható értékek:

$$\begin{aligned}\mathbb{E}X &= \int_{x=0}^{+\infty}\int_{y=0}^{x} x\,\frac{2}{x}\exp(-2x)\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{+\infty} 2x\exp(-2x)\,\mathrm{d}x = \\ &\underset{\text{parc. int.}}{=} [-x\exp(-2x)]_{x=0}^{+\infty} + \int_{x=0}^{+\infty}\exp(-2x)\,\mathrm{d}x = 0 + \frac{-1}{2}\exp(-2x)\Big]_{x=0}^{+\infty} = \frac{1}{2}.\end{aligned}$$

$$\mathbb{E}Y = \int_{x=0}^{+\infty}\int_{y=0}^{x} y\,\frac{2}{x}\exp(-2x)\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{+\infty} x\exp(-2x)\,\mathrm{d}x = \ldots = \frac{1}{4}.$$

$$\begin{aligned}\mathbb{E}(XY) &= \int_{x=0}^{+\infty}\int_{y=0}^{x} xy\,\frac{2}{x}\exp(-2x)\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{+\infty} x^2\exp(-2x)\,\mathrm{d}x = \\ &\overset{\text{parc. int.}}{=} \left[\frac{-x^2}{2}\exp(-2x)\right]_{x=0}^{+\infty} + \int_{x=0}^{+\infty} x\exp(-2x)\,\mathrm{d}x = 0 + \frac{1}{4}.\end{aligned}$$

Ebből a kovariancia $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y = \frac{1}{4} - \frac{1}{2}\frac{1}{4} = \frac{1}{8}$. A szórások:

$$\mathbb{E}(X^2) = \int_{x=0}^{+\infty}\int_{y=0}^{x} x^2\,\frac{2}{x}\exp(-2x)\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{+\infty} 2x^2\exp(-2x)\,\mathrm{d}x = \ldots = \frac{1}{2}.$$

$$\mathbb{E}(Y^2) = \int_{x=0}^{+\infty}\int_{y=0}^{x} y^2\,\frac{2}{x}\exp(-2x)\,\mathrm{d}y\mathrm{d}x = \int_{x=0}^{+\infty} \frac{2}{3}x^2\exp(-2x)\,\mathrm{d}x = \ldots = \frac{1}{6}.$$

Végül: $\mathbb{SD}(X) = \sqrt{\mathbb{V}\mathrm{ar}(X)} = \sqrt{\frac{1}{2}-\left(\frac{1}{2}\right)^2} = \frac{1}{2}$, és $\mathbb{SD}(Y) = \sqrt{\mathbb{V}\mathrm{ar}(Y)} = \sqrt{\frac{1}{6}-\left(\frac{1}{4}\right)^2} = \frac{\sqrt{5}}{4\sqrt{3}}$. Korrelációs együttható: $\mathrm{r}(X,Y) = \frac{1/4}{(1/2)(\sqrt{5}/(4\sqrt{3}))} = \frac{2\sqrt{3}}{\sqrt{5}}$.

(c) Várható értékek:

$$\mathbb{E}X = \int_{y=0}^{+\infty}\int_{x=0}^{+\infty} x\,\frac{1}{y}\exp(-y-x/y)\,\mathrm{d}x\mathrm{d}y = \int_{y=0}^{+\infty}\exp(-y)y\,\mathrm{d}y = \ldots = 1.$$

$$\mathbb{E}Y = \int_{y=0}^{+\infty}\int_{x=0}^{+\infty} y\,\frac{1}{y}\exp(-y-x/y)\,\mathrm{d}x\mathrm{d}y = \int_{y=0}^{+\infty}\exp(-y)y\,\mathrm{d}y = \ldots = 1.$$

$$\mathbb{E}(XY) = \int_{y=0}^{+\infty}\int_{x=0}^{+\infty} xy\,\frac{1}{y}\exp(-y-x/y)\,\mathrm{d}x\mathrm{d}y = \int_{y=0}^{+\infty}\exp(-y)y^2\,\mathrm{d}y = \ldots = 2.$$

Így a kovariancia: $\mathrm{Cov}(X,Y) = 2 - 1\times 1 = 1$. Második momentumok:

$$\mathbb{E}(X^2) = \int_{y=0}^{+\infty}\int_{x=0}^{+\infty} x^2\,\frac{1}{y}\exp(-y-x/y)\,\mathrm{d}x\mathrm{d}y = \ldots = \int_{y=0}^{+\infty}\exp(-y)2y^2\,\mathrm{d}y = 4.$$

$$\mathbb{E}(Y^2) = \int_{y=0}^{+\infty}\int_{x=0}^{+\infty} y^2\,\frac{1}{y}\exp(-y-x/y)\,\mathrm{d}x\mathrm{d}y = \ldots = \int_{y=0}^{+\infty}\exp(-y)y^2\,\mathrm{d}y = 2.$$

Ebből a szórások: $\mathbb{SD}(X) = \sqrt{4-1^2} = \sqrt{3}$, $\mathbb{SD}(Y) = \sqrt{2-1^2} = 1$ és végül a korrelációs együttható: $\mathrm{r}(X,Y) = \frac{1}{\sqrt{3}}$.

Tehát egyik esetben sem függetlenek a marginálisok; utóbbi két esetben pozitív korrelációt kaptunk, míg az első esetben negatívat.

11. Egy gyárban két gép egymás mellett működik, de egyik üzemelése vagy épp nem üzemelése sem befolyásolja a másik működését vagy épp nem működését. Mindkettő exponenciális ideig él, és az esetek 90%-ában több, mint 1000 órát üzemelnek.

    (a) Várhatóan meddig lesz üzemképes mindkét gép?

    (b) Várhatóan meddig fog üzemelni legalább az egyik gép?

    (c) A két gép együttesen átlagosan hány órát üzemel?

    Megoldás: Legyen X, Y független exponenciális eloszlású valószínűségi változók, X jelenti az egyik, Y a másik gép működési idejét. Paraméter meghatározása: $\mathbb{P}(X > 1000) = 0.9$, azaz $\exp(-\lambda 1000) = 0.9$, vagyis $\lambda = \frac{\log(10)-\log(9)}{1000} \approx 105\times 10^{-6}$. A továbbiakban a λ jelölést visszük tovább, csak a végén írjuk be a – közelítő – értékét.

    (a) Mindkét pontosan $\min(X,Y)$ ideig üzemel. E valószínűségi változó eloszlása: $\mathbb{P}(\min(X,Y)\geq z) = \mathbb{P}(X\geq z, Y\geq z) = \mathbb{P}(X\geq z)\mathbb{P}(Y\geq z) = \exp(-2\lambda)$, ha $z>0$, egyébként 1, ahol kihasználtuk: X, Y független, nemnegatív, azonos eloszlású változók. Azaz $\min(X,Y)$ exponenciális eloszlású 2λ paraméterrel, vagyis $\mathbb{E}(\min(X,Y)) = \frac{1}{2\lambda}\approx 4762$.

    (b) Tudjuk, hogy $\min(X,Y)+\max(X,Y) = X+Y$. Ezt, és az (a) részt használva: $\mathbb{E}(\max(X,Y)) = \mathbb{E}(X+Y-\min(X,Y)) = \mathbb{E}X + \mathbb{E}Y - \mathbb{E}(\min(X,Y)) = \frac{2}{\lambda} - \frac{1}{2\lambda}\approx 14286$, mivel exponenciális várható értéke, a paraméterének reciproka.
    <u>Másik megoldás</u>: (kicsit hosszabb de megéri)
    Kiszámoljuk $\max(X,Y)$ sűrűségfüggvényét. $\mathbb{P}(\max(X,Y)\leq z) = \mathbb{P}(X\leq z, Y\leq z) = \mathbb{P}(X\leq z)\mathbb{P}(Y\leq z) = (1-\exp(-\lambda z))^2$, ha $0\leq z$; egyébként zérus, azaz $\max(X,Y)$ sűrűségfüggvénye: $f_{\max(X,Y)}(z) = 2(1-\exp(-\lambda z))\lambda\exp(-\lambda z)$, ha $0\leq z$, egyébként zérus; így

    $$\begin{aligned}\mathbb{E}\left(\max(X,Y)\right) &= \int_{z=0}^{+\infty} z\,2(1-\exp(-\lambda z))\lambda\exp(-\lambda z)\,\mathrm{d}z = \\ &= 2\int_{z=0}^{+\infty} z\lambda\exp(-\lambda z)\,\mathrm{d}z - \int_{z=0}^{+\infty} z\,2\lambda\exp(-2\lambda z))\,\mathrm{d}z = \frac{2}{\lambda} - \frac{1}{2\lambda}\approx 14286\end{aligned}$$

    utóbbi egyenlőség parciális integrálással adódik, vagy abból az észrevételből, hogy tulajdonképpen két exponenciális eloszlású változó várható értékét számoljuk ki.

    (c) Az előbbiekből már nyilvánvaló, hogy a válasz: $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y = \frac{2}{\lambda}\approx 19048$.

12. Határozzuk meg 6 egyenletesen véletlenül választott $(0,1)$ szám közül a harmadik legkisebb várható értékét és szórását! Mi a második és negyedik legnagyobb szám kovarianciája, korrelációja?

    Megoldás: Jelölje U_3^* a harmadik legkisebb véletlen számot a hat közül. Az első kérdésre a válasz: béta-eloszlás, amelynek sűrűségfüggvénye:

    $$f_{U_3^*}(x)\,\mathrm{d}x = \mathbb{P}(U_3^*\approx x) = \binom{6}{2}x^2\binom{4}{1}\mathrm{d}x\binom{3}{3}(1-x-\mathrm{d}x)^3,$$

    mindez szóban: ahhoz hogy 6 független egyenletes közül a harmadik durván x legyen (azaz az $(x, x+\mathrm{d}x)$ kicsi intervallumba essen) annak az esélye: sorsoljunk 2 számot amelyek kisebbek x-nél, ennek a valószínűsége: $\binom{6}{2}x^2$, egy szám essen $(x, x+\mathrm{d}x)$-be ennek a valószínűsége: $\binom{4}{2}\mathrm{d}x$, végül a maradék 3 szám az $(x+\mathrm{d}x, 1)$ intervallumba kell, hogy essék, ennek az esélye $(1-x-\mathrm{d}x)^3$. A fenti azonosságból úgy varázsoljuk elő a sűrűségfüggvényt, hogy $\mathrm{d}x$-szel leosztunk, majd 0-ba tartunk vele, így kapjuk, hogy $f_{U_3^*}(x) = \frac{6!}{2!1!3!}x^2(1-x)^3$, ha $0\leq x\leq 1$; egyébként 0. Másik út: az eloszlásfüggvényt közvetlenül kiszámoljuk majd deriváljuk:

    $$\begin{aligned}1 - F_{U_3^*}(x) &= \mathbb{P}(U_3^* > x) = \mathbb{P}(\text{legalább 4 szám nagyobb mint } x) \\ &= \binom{6}{4}x^2(1-x)^4 + \binom{6}{5}x(1-x)^5 + \binom{6}{6}(1-x)^6 \quad (0\leq x\leq 1).\end{aligned}$$

Várható érték és szórás:

$$\begin{aligned}
\mathbb{E}(U_3^*) &= \int_{x=0}^1 x\, f_{U_3^*}(x)\,\mathrm{d}x = \frac{6!}{2!1!3!}\int_{x=0}^1 x^3(1-x)^3\,\mathrm{d}x = 60\int_{x=0}^1 (x^3-3x^4+3x^5-x^6)\,\mathrm{d}x = \ldots = \frac{60}{140} = \frac{3}{7},\\
\mathbb{E}(U_3^*)^2 &= \int_{x=0}^1 x^2\, f_{U_3^*}(x)\,\mathrm{d}x = \frac{6!}{2!1!3!}\int_{x=0}^1 x^4(1-x)^3\,\mathrm{d}x = 60\int_{x=0}^1 (x^4-3x^5+3x^6-x^7)\,\mathrm{d}x = \ldots = \frac{60}{280} = \frac{3}{14}.
\end{aligned}$$

Tehát $\mathbb{SD}(U_3^*) = \sqrt{\frac{3}{14}-\frac{3^2}{7^2}} = \frac{\sqrt{3}}{7\sqrt{2}} \approx 0.175$.
A második és negyedik legnagyobb számot jelölje a (U_2^*, U_4^*) pár, együttes sűrűségfüggvényük:

$$f_{(U_2^*,U_4^*)}(x,y)\,\mathrm{d}x\mathrm{d}y = \mathbb{P}(U_2^*\approx x, U_4^*\approx y) = \binom{6}{1}x\binom{5}{1}\mathrm{d}x\binom{4}{1}(y-x-\mathrm{d}x)\binom{3}{1}\mathrm{d}y\binom{2}{2}(1-y-\mathrm{d}y)^2$$

mindez szóban: egy szám $(0,x)$-be kell, hogy essen ennek esélye $\binom{6}{1}x$, utána $(x, x+\mathrm{d}x)$-be kell egy számnak esnie, ennek $\binom{5}{1}\mathrm{d}x$ a valószínűsége, és így tovább, végül az $(y+\mathrm{d}y, 1)$ intervallumba a maradék két számnak kell esnie: ennek a valószínűsége pedig $(1-y-\mathrm{d}y)^2$. A fenti azonosságból úgy kapjuk meg az együttes sűrűségfüggvényt, hogy $\mathrm{d}x\,\mathrm{d}y$-nal leosztunk, majd 0-ba tartunk velük, így kapjuk, hogy $f_{(U_2^*,U_4^*)}(x,y) = \frac{6!}{2!}x(y-x)(1-y)^2$, ha $0\le x\le y\le 1$; egyébként 0. Ellenőrzésképpen ez utóbbi marginális sűrűségfüggvényei:

$$\begin{aligned}
&f_{U_2^*}(x) = \int_{y=0}^1 f_{(U_2^*,U_4^*)}(x,y)\,\mathrm{d}y = \frac{6!}{2!}x\int_{y=x}^1 (y-x)(1-y)^2\,\mathrm{d}y =\\
&\frac{6!}{2!}x\int_{y=x}^1(-xy^2+2xy-x+y^3-2y^2+y)\,\mathrm{d}y = \frac{6!}{2!}x\left[-x\frac{y^3}{3}+2x\frac{y^2}{2}-xy+\frac{y^4}{4}-2\frac{y^3}{3}+\frac{y^2}{2}\right]_{y=x}^1 =\\
&= \frac{6!}{2!}x\frac{(1-x)^4}{12} = \frac{6!}{1!1!4!}x(1-x)^4 \quad (0\le x\le 1)
\end{aligned}$$

ez 6 egyenletes független szám közül a 2. legnagyobbnak a sűrűségfüggvénye, illetve

$$f_{U_4^*}(y) = \int_{x=0}^1 f_{(U_2^*,U_4^*)}(x,y)\,\mathrm{d}x = \frac{6!}{2!}(1-y)^2\int_{x=0}^y x(y-x)\,\mathrm{d}x = \frac{6!}{2!}(1-y)^2\left[y\frac{x^2}{2}-\frac{x^3}{3}\right]_{x=0}^y = \frac{6!}{3!1!2!}y^3(1-y)^2$$

ahol $0\le y\le 1$; ez pedig éppen a 4. legnagyobb sűrűségfüggvénye.
(U_2^*, U_4^*) kovarianciája, korrelációja:

$$\begin{aligned}
\mathbb{E}(U_2^*) &= \int_{x=0}^1 x\, f_{U_2^*}(x)\,\mathrm{d}x = \frac{6!}{1!1!4!}\int_{x=0}^1 x^2(1-x)^4\,\mathrm{d}x = 30\int_{x=0}^1(x^6-4x^5+6x^4-4x^3+x^2)\,\mathrm{d}x = \ldots = \frac{30}{105} = \frac{2}{7}\\
\mathbb{E}(U_4^*) &= \int_{y=0}^1 y\, f_{U_2^*}(y)\,\mathrm{d}y = \frac{6!}{3!1!2!}\int_{y=0}^1 y^4(1-y)^2\,\mathrm{d}y = 60\int_{y=0}^1(y^6-2y^5+y^4)\,\mathrm{d}y = \ldots = \frac{60}{105} = \frac{4}{7},
\end{aligned}$$

mindezt egyszerűbben is megoldhattuk volna, hiszen $\mathbb{E}U_2^* = \frac{2}{7}\int_{x=0}^1 f_{V_3^*}(x)\,\mathrm{d}x = \frac{2}{7}$, ahol V_3^* 7 független, egyenletes szám közül a 3. legkisebb eloszlása. Ugyanezzel $\mathbb{E}U_4^* = \frac{4}{7}\int_{y=0}^1 f_{V_6^*}(y)\,\mathrm{d}y = \frac{4}{7}$, ahol V_6^* 7 független, egyenletes szám közül a 6. legkisebb eloszlása.

$$\begin{aligned}
&\mathbb{E}(U_2^*U_4^*) = \iint_{[0,1]\times[0,1]} xy\, f_{(U_2^*,U_4^*)}(x,y)\,\mathrm{d}x\mathrm{d}y = \frac{6!}{2!}\int_{y=0}^1\int_{x=0}^y x^2y(y-x)(1-y)^2\,\mathrm{d}x\,\mathrm{d}y = \ldots =\\
&= \frac{6!}{2!}\int_{y=0}^1 \frac{1}{12}y^5(1-y)^2\,\mathrm{d}y = \ldots = \frac{6!}{2!}\frac{1}{12}\frac{1}{2016} = \frac{5}{336}.
\end{aligned}$$

Tehát $\mathrm{Cov}(U_2^*, U_4^*) = \frac{5}{336} - \frac{2}{7}\frac{4}{7} = -\frac{349}{2352} \approx -0.15 < 0$.
Második momentumok, ezeket csak okosan számoljuk ki: $\mathbb{E}(U_2^*)^2 = \frac{3}{28}\int_{x=0}^1 f_{W_4^*}(x)\,\mathrm{d}x = 3/28$, ahol W_4^* a 4. legnagyobb eloszlása 8 egyenletes közül. $\mathbb{E}(U_4^*)^2 = \frac{5}{14}\int_{x=0}^1 f_{W_6^*}(x)\,\mathrm{d}x = 5/14$, ahol W_6^* pedig a 6. legnagyobb eloszlása 8 egyenletes közül.
Azaz U_2^* szórása: $\mathbb{SD}(U_2^*) = \sqrt{3/28-4/49} = \sqrt{5}/14 \approx 0.16$, illetve U_4^* szórása: $\mathbb{SD}(U_4^*) = \sqrt{5/14-16/49} = \sqrt{3}/(7\sqrt{2}) \approx 0.175$.

13. Tegyük fel, hogy egy jólmenő étterem heti összbevétele normális eloszlást követ 1 millió forint várható haszonnal, és 700000 forint szórással. Mi annak a valószínűsége, hogy kevesebb, mint 1.5 millió forint a bevétele két, egymást követő héten? Itt tegyünk még fel függetlenséget! Majd nézzük meg, hogyan változik annak a valószínűsége, hogy a második héten több mint 2 millió forint a bevétel feltéve, hogy az első héten 1.5 millió forint volt a bevétel és a korreláció -0.5! Mi a második hét várható bevétele ugyanezen feltétel mellett? Mennyi a két hét várható összbevétele? Mi a szórás?

Megoldás: Rögzítsük a jelöléseket: $\mu := 1$, $\sigma := 0.7$ millió forintban számolva. (X,Y) pár pedig az első, illetve második hét (véletlen) összbevétele (millióban), mindkettő tehát normális (μ,σ) paraméterrel. Függetlenséget feltéve (azaz $r = 0$):

$$\begin{aligned}
\mathbb{P}(X<1.5, Y<1.5) &= \mathbb{P}(X<1.5)\mathbb{P}(Y<1.5) = \mathbb{P}\left(\frac{X-1}{0.7}<\frac{5}{7}\right)\mathbb{P}\left(\frac{Y-1}{0.7}<\frac{5}{7}\right) =\\
&= \Phi(5/6)\Phi(5/7) \approx 0.762^2 = 0.581
\end{aligned}$$

Avagy, ha a két hét (teljes) összbevételét tekintjük, akkor: $X-1+Y-1$ eloszlása normális $(0, 0.7\sqrt{2})$ paraméterrel, mivel $z \in \mathbb{R}$-re:

$$\begin{aligned}
f_{X+Y-2}(z)\,\mathrm{d}z &= \mathbb{P}(X-1+Y-1 \approx z) = \int_{t=-\infty}^{+\infty} f_{X-1,Y-1}(t, z-t)\,\mathrm{d}z\,\mathrm{d}t = \\
&= \int_{t=-\infty}^{+\infty} \frac{1}{2\pi\, 0.7^2} \exp\left(-\frac{t^2+(z-t)^2}{2\times 0.7^2}\right)\mathrm{d}z\,\mathrm{d}t = \frac{1}{2\pi\, 0.7^2}\int_{t=-\infty}^{+\infty} \exp\left(-\frac{2t^2-2zt+z^2}{2\times 0.7^2}\right)\mathrm{d}z\,\mathrm{d}t = \\
&= \frac{1}{\sqrt{2\pi}\,0.7\sqrt{2}} \frac{1}{\sqrt{2\pi}\,(0.7/\sqrt{2})} \int_{t=-\infty}^{+\infty} \exp\left(-\frac{2(t-z/2)^2+z^2/2}{2\times 0.7^2}\right)\mathrm{d}z\,\mathrm{d}t = \\
&= \frac{1}{\sqrt{2\pi}\,0.7\sqrt{2}} \exp\left(-\frac{z^2}{2\times(0.7\sqrt{2})^2}\right)\mathrm{d}z.
\end{aligned}$$

Így $X+Y$ $(2, 0.7\sqrt{2})$ paraméterű normális (vesd össze a 16. feladattal), tehát $\mathbb{P}(X+Y<1.5) = \Phi\left(-\frac{0.5}{0.7\sqrt{2}}\right) \approx 0.307$.

Namost tegyük fel, hogy $r=-0.5$. Az (X,Y) pár kétdimenziós (nem-standard) normális, peremeloszlásai is normálisak: $Y|X=x$ $(x=1.5)$ várható értéke: $\mu + r(x-\mu) = 1-0.5(1.5-1) = 3/4 = 0.75$ (ez a feltételes várható érték), szórása $\sigma\sqrt{1-r^2} = 0.7\sqrt{3}/2 \approx 0.606$, azaz behelyettesítve és standardizálva kapjuk, hogy

$$\mathbb{P}(Y>2\,|\,X=1.5) = \mathbb{P}\left(\frac{Y-0.75}{0.606} > 2.06\,|\,X=1.5\right) = 1-\Phi(2.06) \approx 0.02.$$

$\mathbb{E}(X+Y)=2$ az összbevétel, $\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X)+\mathbb{V}\mathrm{ar}(Y)+2\mathrm{Cov}(X,Y) = 2\times 0.7^2 + 2\times(-0.5)0.7^2 = 0.49$ pedig a szórásnégyzet.

14. Budapesten májusban az átlagos hőmérséklet 25 °C, 7 °C szórással, valamint az átlagnyomás 10^5 Pa, 2×10^4 Pa szórással. A hőmérséklet/nyomás változása szoros összhangban van, köztük lévő korreláció 0.7. Írjuk fel a kovariancia mátrixot majd határozzuk meg a következőket:

   (a) Mi a valószínűsége annak, hogy egy nap melegebb lesz, mint 40 °C? És, hogy alacsonyabb a nyomás 6×10^4 Pa-nál?

   (b) Egy nap 20 °C-ot mértünk. Mi annak a valószínűsége, hogy a légnyomás 1.2×10^5 Pa fölött járt? Átlagosan mekkora volt a légnyomás? Mekkora a szórás?

   (c) Feltéve, hogy egy nap 10^5 Pa volt a légnyomás, mi annak a valószínűsége, hogy melegebb volt, mint 35 °C? Átlagosan hány fok volt aznap? Mekkora a szórás?

   Megoldás: Rögzítsük az adatokat: $\mu_X := 25$, $\sigma_X := 7$, $\mu_Y := 10^5$, $\sigma_Y := 2\times 10^4$ és $r := 0.7$, X jelenti az átlag hőmérsékletet, Y pedig a hozzátartozó nyomást. (X,Y) kétdimenziós normális, a kovariancia mátrix:

$$\Sigma = \begin{pmatrix} 49 & 9.28\times 10^4 \\ 9.28\times 10^4 & 4\times 10^8 \end{pmatrix}.$$

   (a) $\mathbb{P}(X>40) = \mathbb{P}\left(\frac{X-25}{7} > \frac{15}{7}\right) = 1-\Phi(15/7) \approx 0.016$; $\mathbb{P}(Y<6\times 10^4) = \mathbb{P}\left(\frac{X-10^5}{2\times 10^4} < -2\right) = 1-\Phi(2) \approx 0.023$;

   (b) $x=20$, $Y\,|\,X=x$ normális eloszlású $\left(\mu_Y + \frac{\sigma_Y}{\sigma_X} r(x-\mu_X), \sigma_Y\sqrt{1-r^2}\right)$ paraméterekkel. Azaz ennek a várható értéke: $10^5 + 2\times 10^4/7(20-25)\times 0.7 = 90000$, és szórása: $2\times 10^4\sqrt{1-0.7^2} \approx 14283$. A keresett valószínűség pedig:

$$\mathbb{P}(Y>1.2\times 10^5\,|\,X=20) = \mathbb{P}\left(\frac{Y-90000}{14283} > 2.1\,|\,X=20\right) = 1-\Phi(2.1) \approx 0.018.$$

   (c) $y=10^5$, $X\,|\,Y=y$ normális eloszlású $\left(\mu_X + \frac{\sigma_X}{\sigma_Y} r(y-\mu_Y), \sigma_X\sqrt{1-r^2}\right)$ paraméterekkel. Azaz ennek a várható értéke: $25 + 7/(2\times 10^4)(10^5-10^5)\times 0.7 = 25$, és szórása: $7\sqrt{1-0.7^2} \approx 5$. A keresett valószínűség pedig:

$$\mathbb{P}(X>35\,|\,Y=10^5) = \mathbb{P}\left(\frac{X-25}{5} > 2\,|\,Y=10^5\right) = 1-\Phi(2) \approx 0.023.$$

15. Magyarországon a felnőtt férfiak testmagassága átlagosan 178 cm, 9 cm szórással, míg testsúlyuk 85 kg, 10 kg szórással. A korrelációs együttható 0.7, azaz minél magasabb valaki, annál súlyosabb is. Írjuk fel a kovariancia mátrixot!

   (a) Mi a valószínűsége annak, hogy egy férfi magasabb 2 méternél? És, hogy nehezebb 100 kg-nál?

   (b) Feltéve, hogy egy férfi 190 cm, mi annak a valószínűsége, hogy könnyebb, mint 60 kg? Átlagosan mekkora súlyú egy ilyen férfi? Mekkora a szórás?

   (c) Feltéve, hogy egy férfi 80 kg, mi annak a valószínűsége, hogy magasabb, mint 180 cm? Várhatóan hány cm magas egy ilyen férfi? Mekkora a szórás?

Megoldás: Rögzítsük az adatokat: $\mu_X := 178$, $\sigma_X := 9$, $\mu_Y := 85$, $\sigma_Y := 10$ és $r := 0.7$, X jelenti a magasságot, Y pedig hozzátartozó testsúlyt. (X, Y) kétdimenziós normális, a kovariancia mátrix:

$$\Sigma = \begin{pmatrix} 81 & 63 \\ 63 & 100 \end{pmatrix}.$$

(a) $\mathbb{P}(X > 200) = \mathbb{P}\left(\frac{X-178}{9} > \frac{22}{9}\right) = 1 - \Phi(22/9) \approx 0.0073$; $\mathbb{P}(Y > 100) = \mathbb{P}\left(\frac{X-85}{10} > 1.5\right) = 1 - \Phi(1.5) \approx 0.067$;

(b) $x = 190$, $Y \mid X = x$ normális eloszlású $\left(\mu_Y + \frac{\sigma_Y}{\sigma_X} r(x - \mu_X), \sigma_Y\sqrt{1-r^2}\right)$ paraméterekkel. Azaz ennek a várható értéke: $85 + 10/9(190 - 178) \times 0.7 \approx 102.11$, és szórása: $10\sqrt{1 - 0.7^2} \approx 7.14$. A keresett valószínűség pedig:

$$\mathbb{P}(Y < 60 \mid X = 190) = \mathbb{P}\left(\frac{Y - 102.11}{7.14} < -5.9 \mid X = 20\right) = 1 - \Phi(5.9) \approx 1.82 \times 10^{-9}.$$

(c) $y = 80$, $X \mid Y = y$ normális eloszlású $\left(\mu_X + \frac{\sigma_X}{\sigma_Y} r(y - \mu_Y), \sigma_X\sqrt{1-r^2}\right)$ paraméterekkel. Azaz ennek a várható értéke: $178 + 9/10(80 - 85) \times 0.7 = 174.85$, és szórása: $9\sqrt{1 - 0.7^2} \approx 6.43$. A keresett valószínűség pedig:

$$\mathbb{P}(X > 180 \mid Y = 80) = \mathbb{P}\left(\frac{X - 174.85}{6.43} > 0.8 \mid Y = 10^5\right) = 1 - \Phi(0.8) \approx 0.212.$$

Érdekességképpen: feltétel nélkül $\mathbb{P}(X > 180) = 1 - \Phi(2/9) \approx 0.412$.

16. Legyen a (X, Y) pár kétdimenziós normális eloszlású, r korrelációval. Mi az eloszlása $U = X + Y$-nak és $V = X - Y$-nak? Független-e U a V-től? Számoljuk ki a várható értékeket és szórásokat is!

Megoldás: Legyen $\mathbf{A} \in \mathbb{R}^{2\times 2}$ és $\mu \in \mathbb{R}^2$ olyan, hogy $(X, Y) = (P, Q)\mathbf{A} + \mu$, ahol $\mathbf{A}^\top\mathbf{A} = \Sigma$ (Σ-ban a paramétereket: r, σ_X, σ_Y jelöli) és P, Q standard normális pár (egymástól függetlenek 0 várható értékkel és 1 szórással). Ekkor $U = (a_{11}+a_{12})P+(a_{21}+a_{22})Q+\mu_X+\mu_Y$, illetve $V = (a_{11}-a_{12})P+(a_{21}-a_{22})Q+\mu_X-\mu_Y$. Tehát (U, V) is kétdimenziós normális eloszlású, így a marginálisai és feltételes eloszlásai is normálisak. Mivel U és V is normális, pontosan akkor függetlenek, ha korrelálatlanok: $\mathrm{Cov}(U, V) = \mathrm{Cov}(X + Y, X - Y) = \mathbb{V}\mathrm{ar}(X) - \mathbb{V}\mathrm{ar}(Y)$, azaz, ha $\sigma_X = \sigma_Y$, akkor függetlenek, egyébként függőek. Valamint $\mathbb{E}(X \pm Y) = \mu_X \pm \mu_Y$; $\mathbb{V}\mathrm{ar}(X \pm Y) = \sigma_X + \sigma_Y \pm 2r\sigma_X\sigma_Y$.

17. Legyen U és Z független valószínűségi változók, előbbi egyenletes a $(0, 2\pi)$ intervallumban, utóbbi exponenciális, 1 várható értékkel. Mi az eloszlása az $X = \sqrt{2Z}\cos(U)$, illetve $Y = \sqrt{2Z}\sin(U)$ valószínűségi változónak? Független-e X Y-tól?

Megoldás: A koordináták közötti kapcsolat: $\sqrt{X^2 + Y^2} = \sqrt{2Z}$, illetve $\arctan(Y/X) = U$. [mivel U egyenletes $(\cos(U), \sin(U))$ egy egyenletes pont az origó körüli egységkörvonalon; $\sqrt{2Z}$ pedig egy véletlen sugár ezért valami polárkoordinátás helyettesítésre gondolunk] Mivel Z független U-tól, együttes sűrűségfüggvényük: $f_{Z,U}(z, u) = \exp(-z)\frac{1}{2\pi}$, ha $z \geq 0$ és $0 \leq u \leq 2\pi$, egyébként 0.

$$f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y = \mathbb{P}(\sqrt{2Z}\cos(U) \approx x,\ \sqrt{2Z}\sin(U) \approx y) = \mathbb{P}(Z \approx (x^2 + y^2)/2,\ U \approx \arctan(y/x)) =$$
$$\underset{z:=(x^2+y^2)/2;\ u:=\arctan(y/x)}{=} f_{Z,U}(z, u)\ \mathrm{d}z\,\mathrm{d}u = \frac{1}{2\pi}\exp(-(x^2 + y^2)/2)\,\mathrm{d}x\,\mathrm{d}y$$

az utolsó egyenlőségben használtuk a $\mathrm{d}z\mathrm{d}u = |\det(J(x, y))|\mathrm{d}x\mathrm{d}y$ többváltozós differenciálhelyettesítési szabályt, ahol $J(x, y)$ a $z = (x^2 + y^2)/2$, $u = \arctan(y/x)$ helyettesítéshez tartozó Jacobi-mátrix, azaz

$$J(x, y) = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 2y \\ \frac{1}{1+(x/y)^2}\frac{1}{y} & \frac{1}{1+(x/y)^2}\frac{-x}{y^2} \end{bmatrix},$$

aminek a determinánsa: $\frac{-2x^2-2y^2}{x^2+y^2} = -1$. Tehát a kettővel előbbi képletben, ha $\mathrm{d}x\mathrm{d}y$-nal átosztunk, kapjuk: $f_{X,Y}(x, y) = \exp(-(x^2 + y^2)/2)/(2\pi)$, azaz (X, Y) független standard normális pár.
Megjegyzés: Ha már tudjuk, hogy X és Y normális, akkor függetlenségük garantált, ha kovarianciájuk 0, de $\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}Z\mathbb{E}(\sin(2U)) - (\mathbb{E}\sqrt{2Z})^2\mathbb{E}(\sin(U))\mathbb{E}(\cos(U))$, ahol használtuk U, Z függetlenségét, illetve a $2\sin(x)\cos(x) = \sin(2x)$ azonosságot. Az előbbi mennyiség valóban zérus, mivel:

$$\mathbb{E}(\sin(2U)) = \int_0^{2\pi} \sin(2x)\frac{1}{2\pi}\,\mathrm{d}x = \frac{-\cos(2x)}{4\pi}\Bigg]_{x=0}^{2\pi} = \frac{-\cos(4\pi) + \cos(0)}{4\pi} = \frac{-1+1}{4\pi} = 0$$
$$\mathbb{E}(\sin(U)) = \int_0^{2\pi} \sin(x)\frac{1}{2\pi}\,\mathrm{d}x = \frac{-\cos(x)}{4\pi}\Bigg]_{x=0}^{2\pi} = \frac{-\cos(2\pi) + \cos(0)}{4\pi} = \frac{-1+1}{4\pi} = 0.$$

18. Hogyan generálna le kétdimenziós normális eloszlású véletlen pontokat a síkon, melyek várható értéke (μ_1, μ_2), szórása (σ_1, σ_2), korrelációs együtthatója pedig r. Függetlenek-e a koordináták, ha $r = 0$?

Megoldás: Vegyünk egy (U, V) standard normális véletlen párt (azaz U, V standard normális véletlen szám, egymástól függetlenül). Adottak a $\mu_1, \mu_2 \in \mathbb{R}$; $\sigma_1, \sigma_2 \in \mathbb{R}^+$ és $-1 < r < 1$ számok. Kell: egy

(X,Y) pár, hogy $\mathbb{E}X = \mu_1$, $\mathbb{E}Y = \mu_2$, $\mathbb{SD}(X) = \sigma_1$, $\mathbb{SD}(Y) = \sigma_2$, illetve $\mathrm{Cov}(X,Y) = r\sigma_1\sigma_2$, és persze (X,Y) kétdimenziós normális. Ezt lineáris transzformációval érjük el a standard (U,V) párból. Legyen $X = \sigma_1 U + \mu_1$, ezzel az egyik marginális már megvan, a másikat $Y := aU + bY + \mu_2$ alakban keressük. Y várható értéke be van lőve, már csak $a, b \in \mathbb{R}$ – egyelőre ismeretlen – paramétereket kell úgy megválasztani, hogy Y szórása, illetve X, Y kovarianciája megfelelő legyen. Erre pedig a következő adódik: a függetlenség miatt $\mathbb{V}\mathrm{ar}(Y) = \mathbb{V}\mathrm{ar}(aU + bV) = a^2 + b^2 = \sigma_2^2$, illetve $\mathrm{Cov}(X,Y) = \mathrm{Cov}(\sigma_1 U + \mu_1, aU + bV + \mu_2) = \sigma_1 a = r\sigma_1\sigma_2$, azaz $a = r\sigma_2$, illetve $b = \sigma_2\sqrt{1-r^2}$. Ezzel $(X = \sigma_1 U + \mu_1;\ Y = r\sigma_2 U + \sigma_2\sqrt{1-r^2}V + \mu_2)$ kétdimenziós normális a megfelelő paraméterekkel. U, V-t úgy generáljuk, hogy veszünk két független véletlen számot $(0,1)$-ből, ezeket jelölje Z_1, Z_2, így $(\Phi^{-1}(Z_1), \Phi^{-1}(Z_2))$ standard normális pár, ebből pedig az előbbi transzformációval kapunk kétdimenziós normálisat. Ha $r = 0$, a koordináták függetlenek, mert normális esetben korrelálatlanságból következik a függetlenség, és ez konkrétan meg is jelenik, ugyanis a fenti transzformációban: $r = 0$ esetén Y csak V-től függ, U-tól nem.

Megjegyzés: anélkül, hogy használnánk a Φ inverzét, standard normális párt az előző, 17. feladat alapján is generálhatunk, legyen Z_1, Z_2 két véletlen szám $(0,1)$-ből. Ekkor $-\log(Z_1)$ exponenciális eloszlású 1 várható értékkel ezt kell szorozni 2-vel majd gyököt vonni, végül a $(\cos(2\pi\, Z_2), \sin(2\pi\, Z_2))$ véletlen vektorral szorozni, az eredmény standard normális pár.