A május 13.-i előadáshoz kapcsolódó gyakorlat feladatai

1. Lásssuk be a következő állítást:
Legyenek $Z^{(j)} = \left(Z_1^{(j)}, \dots, Z_k^{(j)}\right)$, $1 \le j \le n$, véletlen k-dimenziós vektorok ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor a $Z^{(1)} + \dots + Z^{(n)}$ összeg várható értéke megegyezik a $Z^{(j)}$ vektorok várható értékeinek az összegével, azaz

$$E\left(Z^{(1)} + \dots + Z^{(n)}\right) = EZ^{(1)} + \dots + EZ^{(n)}.$$

Ha a $Z^{(j)}$, $1 \le j \le n$, véletlen vektorok függetlenek, akkor a kovariancia mátrix is additív, azaz, ha a $Z^{(j)}$ mátrix kovariancia mátrixa a D_j mátrix, $1 \le j \le n$, akkor a $Z_{(1)} + \dots + Z^{(n)}$ véletlen összeg kovariancia mátrixa a $D_1 + \dots + D_n$ mátrix. Ha egy $Z = (Z_1, \dots, Z_k)$, véletlen vektor várható értéke $M = (M_1, \dots, M_k)$, kovariancia mátrixa a D $k \times k$ méretű mátrix, a tetszőleges valós szám, akkor az $aZ = (aZ_1, \dots, aZ_k)$, véletlen vektor várható értéke aM, kovariancia mátrixa pedig az $a^2 D$ kovariancia mátrix. Legyen továbbá $x = (x_1, \dots, x_k$ tetszőleges k-dimenziós vektor. Ekkor $E(Z + x) = EZ + x$, a $Z + x$ vektor kovariancia mátrixa pedig megegyezik a Z vektor kovariancia mátrixával.
Megoldás: A feladatban kimondott állítások következnek az egy-dimenziós esetben bizonyított állításokból. Az egy-dimenziós várható érték additivitából következik, hogy $E\left(Z^{(1)} + \dots + Z^{(n)}\right) = EZ^{(1)} + \dots + EZ^{(n)}$. Ha a $Z^{(j)}$ vektorok függetlenek, akkor tetszőleges j és k indexre,

$$\begin{aligned}\operatorname{Cov}\left(\sum_{p=1}^n Z_j^{(p)}, \sum_{q=1}^n Z_k^{(q)}\right) &= E\left(\sum_{p=1}^n \left(Z_j^{(p)} - EZ_j^{(p)}\right) \sum_{q=1}^n \left(Z_k^{(q)} - EZ_k^{(q)}\right)\right) \\ &= E\left(\sum_{p=1}^n \left(Z_j^{(p)} - EZ_j^{(p)}\right)\left(Z_k^{(p)} - Z_k^{(p)}\right)\right) = \sum_{p=1}^n \operatorname{Cov}\left(Z^{(p)}, Z^{(q)}\right),\end{aligned}$$

mert $E\left(Z_j^{(p)} - EZ_j^{(p)}\right)\left(Z_k^{(q)} - Z_k^{(q)}\right) = E\left(Z_j^{(p)} - EZ_j^{(p)}\right) E\left(Z_k^{(q)} - Z_k^{(q)}\right) = 0$, ha $p \ne q$ a függetlenség miatt. Ez az azonosság pedig azt jelenti, hogy független véletlen vektorok kovariancia mátrixa additív.
Az, hogy $E(Z + x) = EZ + x$, $EaZ = aEZ$, $\operatorname{Cov}(Z_p + x, Z_q + y) = \operatorname{Cov}(Z_p, Z_q)$, $\operatorname{Cov}(aZ_p, aZ_q) = a^2 \operatorname{Cov}(Z_p, Z_q)$, tetszőleges $Z = (Z_1, \dots, Z_k)$ vektorra és $1 \le p, q \le k$ indexekre következik az egydimenziós esetben már bizonyított állításokból. Ezek az azonosságok tartalmazzák a feladat korábban nem bizonyított állításait.
2. Mutassunk példát két korrelálatlan ξ és η valószínűségi változóra, amelyek nem függetlenek.
Megoldás: Sok egyszerű példát adhatunk. Tekintsük például a következő példát: Legyen ξ egyenletes eloszlású valószínűségi változó a $\left[-\frac{1}{2}, \frac{1}{2}\right]$ intervallumban, Ekkor a ξ és $\eta = \xi^2$ valószínűségi változók korrelálatlanok, de nem függetlenek. Valóban, $E\xi = 0$, $E\eta = E\xi^2 = \dfrac{1}{12}$, $E\xi\eta = E\xi^3 = 0$, $\operatorname{Cov}(\xi, \eta) = E\xi\eta - E\xi E\eta = 0$. Másrészt

ξ és η nem függetlenek, sőt az η valószínűségi változó a ξ valószínűségi változó determinisztikus függvénye. a felhasznált azonosságok ellenőrzését egyszerűsítheti az az észrevétel, hogy ξ valószínűségi változó sűrűségfüggvénye páros függvény. Viszont, ha egy valószínűségi változó sűrűségfüggvénye egy páros $f(x)$, függvény, akkor $E\xi = \int xf(x)\,dx = 0$, $E\xi^3 = \int x^3 f(x)\,dx$.

Egy lehetséges formális indoklása annak, hogy ξ és η nem független a következő: Legyen $0 < a < 1$ tetszőleges szám. Ekkor $\{\omega\colon \eta < a^2\} = \{\omega\colon |\xi| < a\}$. Ezért $P(\xi < a, \eta < a^2) = P(\xi < a)$, azaz $P(\xi < a, \eta < a^2) \neq P(\xi < a)P(\eta < a^2)$.

3. Legyen $\eta = (\eta_1, \dots, \eta_k)$ k-dimenziós, valószínűségi változó $m = (m_1, \dots, m_k)$ várható értékkel és $D = (d_{j,l})$, $1 \le j, l \le k$, kovariancia mátrix-szal. Legyen $B = (b_{j,l})$, $1 \le j, l \le k$, valamely $k \times k$ méretű mátrix. Mutassuk meg, hogy az $\eta B = (\eta_1, \dots, \eta_k)B$ k-dimenziós véletlen vektor várható érték vektora mB és kovariancia mátrixa B^*DB.

   Megoldás: Az $E\eta B$ vektor r-ik eleme

$$E\left(\sum_{j=1}^k \eta_j b_{j,r}\right) = \left(\sum_{j=1}^k b_{j,r} E\eta_j\right) = \left(\sum_{j=1}^k m_j b_{j,r}\right),$$

   és ez az mB vektor r-ik eleme. Az ηB vektor kovariancia mátrixának r-ik sorában és s-ik oszlopában álló elem

$$\begin{aligned} f_{r,s} = \mathrm{Cov}\,\left(\sum_{j=1}^k \eta_j b_{j,r}, \sum_{l=1}^k \eta_l b_{l,s}\right) &= E\left(\sum_{j=1}^k (\eta_j - E\eta_j) b_{j,r} \sum_{l=1}^k (\eta_l - E\eta_l) b_{l,s}\right) \\ = \sum_{j=1}^k \sum_{l=1}^k b_{j,r} b_{l,s} E(\eta_j - E\eta_j)(\eta_l - E\eta_l)) &= \sum_{j=1}^k \sum_{l=1}^k b_{j,r} b_{l,s} \mathrm{Cov}\,(\eta_j, \eta_l) \\ = \sum_{j=1}^k \sum_{l=1}^k E b_{j,r} b_{l,s} d_{j,l} &= \sum_{j=1}^k \sum_{l=1}^k E b^*_{r,j} d_{j,l} b_{l,s}, \end{aligned}$$

   ahol $b^*(r,j) = b_{j,r}$. Ez a kifejezés viszont megegyezik a B^*DB mátrix r-ik sorában és s-ik oszlopában álló elemmel.

4. Definiáljuk a következő $(\Omega, \mathcal{B}, \mathbf{P})$ valószínűségi mezőt: $\Omega = [0,1]$, $\mathcal{B}$ a Borel σ-algebra $[0,1]$-en, és $\mathbf{P}$ a Lebesgue mérték. Definiáljuk a következő ξ és η valószínűségi változókat ezen a mezőn: $\xi(x) = \Phi^{-1}(x)$,

$$\eta(x) = \begin{cases} \xi(1-x) & \text{ha } 0 \le x < \frac{1}{2} \\ \xi\left(x - \frac{1}{2}\right) & \text{ha } \frac{1}{2} \le x \le 1 \end{cases}.$$

   Lássuk be, hogy ξ és η normális eloszlású valószínűségi változók, de a (ξ, η) vektor nem normális eloszlású valószínűségi vektor.

   Megoldás: A ξ és η valószínűségi változók azonos eloszlásúak. Továbbá,

$$P(\xi > x) = \lambda(\Phi(x), 1]) = 1 - \Phi(x)\,.$$

Az, hogy (ξ,η) vektor nem normális eloszlású következik például a $P(\xi+\eta=0)=\frac{1}{2}$ azonosságból. Miért?

5. Legyenek $\xi_1,\dots,\xi_n$ független a $\left[-\frac{1}{2},\frac{1}{2}\right]$ intervallumban egyenletes eloszlású valószínűségi változók. Mutassuk meg, hogy a $\sum\limits_{j=1}^{n}\xi_j$ és $\sum\limits_{j=1}^{n}\xi_j^2$ összegek normalizáltjainak azaz az $\sqrt{\frac{12}{n}}\sum\limits_{j=1}^{n}\xi_j$ és $\sqrt{\frac{180}{n}}\sum\limits_{j=1}^{n}\left(\xi_j^2-\frac{1}{12}\right)$ valószínűségi változóknak az együttes eloszlása a két-dimenziós standard normális eloszláshoz konvergál, ha $n\to\infty$.
Megoldás: $E\xi=0$, $E\xi^2=\frac{1}{12}$, $\operatorname{Var}\xi=\frac{1}{12}$, $\operatorname{Var}\xi^2=E\xi^4-(E\xi^2)^2=\frac{1}{80}-\frac{1}{144}=\frac{1}{180}$. Továbbá $\operatorname{Cov}(\xi,\xi^2)=E\xi^3-E\xi E\xi^2=0$. Ezért a $\left(\sqrt{12}\xi_j,\sqrt{180}\left(\xi_j^2-\frac{1}{12}\right)\right)$, $j=1,2,\dots$, véletlen vektorok függetlenek, nulla várható értékkel és az identitás kovariancia mátrix-szal. Innen, és a több-dimenziós centrális határeloszlástételből következik a feladat állítása.

6. Legyen (ξ,η) normális eloszlású vektor $m=(m_1,m_2)=(E\xi,E\eta)$ várható értékkel és
$$\begin{pmatrix}\sigma_{1,1}, & \sigma_{1,2}\\ \sigma_{2,1}, & \sigma_{1,1}\end{pmatrix}=\begin{pmatrix}E\xi^2-(E\xi)^2, & E\xi\eta-E\xi E\eta\\ E\xi\eta-E\xi E\eta, & E\eta^2-(E\eta)^2\end{pmatrix}$$
kovarianciamátrix-szal. Ekkor létezik a ξ valószínűségi változónak $\xi=a\eta+\zeta$ alakú előállítása alkalmas a konstanssal, és az η valószínűségi változótól független ζ normális eloszlású valószínűségi változóval. Ez azt jelenti, hogy ha (ξ,η) két-dimenziós normális eloszlású valószínűségi változó, akkor az első koordináta kifejezhető mint a második kooordináta konstansszorosának és egy a második koordinátától független normális eloszlású valószínűségi változó összege. A kívánt a konstanst explicit módon megadhatjuk az $a=\frac{\sigma_{1,2}}{\sigma_{2,2}}$ képlet segítségével.
Hogy általánosítható a fenti állítás abban az esetben, ha ξ és η vektorváltozók is lehetnek?
Megoldás: A $\zeta=\xi-\frac{\sigma_{1,2}}{\sigma_{1,1}}\eta$ valószínűségi változó független az η valószínűségi változótól. Ehhez a több-dimenziós normális eloszlás tulajdonságai alapján elég ellenőrizni, hogy $\operatorname{Cov}(\zeta,\eta)=0$. Innen következik a feladat állítása.
Az az eset, amikor $\xi=(\xi_1,\dots,\xi_s)$, $\eta=(\eta_1,\dots,\eta_p)$, és $(\xi_1,\dots,\xi_s,\eta_1,\dots,\eta_p)$ egy $s+p$ dimenziós normális eloszlású valószínűségi változó hasonlóan tárgyalható. Lássuk be, hogy létezik olyan $\mathbf{A}$ mátrix, amelyre η és $\xi-\eta\mathbf{A}$ függetlenek. Ennek érdekében lássuk be először, hogy létezik olyan $\mathbf{U}$ unitér mátrix amelyre $\eta\mathbf{U}=(\bar\eta_1,\dots,\bar\eta_p)=\bar\eta$ vektor koordinátái függetlenek. Ugyanis, ha az η véletlen vektor D kovarianciamátrixát $D=\mathbf{U}^*\Lambda\mathbf{U}$ alakban írjuk, ahol $\mathbf{U}$ unitér Λ pedig diagonális mátrix, akkor az $\bar\eta=\eta\mathbf{U}$ véletlen normális eloszlású vektor kovarianciamátrixa Λ, ahonnan következik, hogy az $\bar\eta$ mátrix koordinátái függetlenek. Legyen $\bar\xi_r=\xi_r-\sum\limits_{k=1}^{p}\frac{E\xi_r\bar\eta_k}{E\bar\eta_k^2}\bar\eta_k$, $r=1,\dots,s$. Ezt mátrixjelöléssel írjuk $\bar\xi=\xi-\bar\eta\mathbf{B}$ formában.

Ekkor $(\xi - \bar{\eta}B, \bar{\eta})$ olyan $p+s$ dimenziós normális eloszlású valószínűségi változó, amelynek első s és utolsó p koordinátája korrelálatlan, ezért független. Mivel a $\xi - \bar{\eta}\mathbf{B}$ és $\bar{\eta}$ vektorok függetlenek, ezért $\zeta = \xi - \bar{\eta}\mathbf{B} = \xi - \eta\mathbf{UB}$ független az $\eta = \bar{\eta}\mathbf{U}^*$ vektortól.

7. Legyen ξ normális eloszlású valószínűségi változó 1 szórásnégyzettel és 2 várható értékkel, azaz legyen ξ sűrűségfüggvénye $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-(x-2)^2/2}$. Számítsuk ki az $Ee^{t\xi}$ várható értéket tetszőleges t valós számra.
Első megoldás: Vezessük be az $\eta = \xi - 2$ valószínűségi változót. Ekkor η standard normális eloszlású valószínűségi változó, és korábban láttuk (április 22.-i gyakorlat, 6. feladat), hogy egy ilyen valószínűségi változóra $Ee^{t\eta} = e^{t^2/2}$. Innen következik, hogy $\xi = \eta + 2$, $Ee^{t\xi} = Ee^{t(\eta+2)} = e^{2t}Ee^{t\eta} = e^{2t}e^{t^2/2} = e^{t^2/2+2t}$.
Második megoldás: Számolhatunk hasonlóan, mint tettük azt a standard normális eloszlás esetében.

$$Ee^{t\xi} = \int_{-\infty}^{\infty} e^{tu} \frac{1}{\sqrt{2\pi}} e^{-(u-2)^2/2}\, du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{2t+t(u-2)-(u-2)^2/2-t^2/2+t^2/2}\, du$$
$$= e^{2t+t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(t-(u-2))^2/2}\, du = e^{2t+t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = e^{2t+t^2/2}.$$

8. Legyenek $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók, amelyekre $E\xi_1 = 0$ és $E\xi_1^2 < \infty$. Legyen $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots$. Ekkor pozitív valós számok valamely a_n sorozatára a $\dfrac{S_n}{a_n}$ valószínűségi változók akkor és csak akkor tartanak sztochasztikusan nullához $n \to \infty$ esetén, ha $\lim\limits_{n\to\infty} \dfrac{a_n}{\sqrt{n}} = \infty$.
Megoldás: Ha $\lim\limits_{n\to\infty} \dfrac{a_n}{\sqrt{n}} = \infty$, akkor

$$P\left(\frac{|S_n|}{a_n} > \varepsilon\right) = P\left(\frac{|S_n|}{\sqrt{n \operatorname{Var} \xi_1}} > \frac{\varepsilon a_n}{\sqrt{n \operatorname{Var} \xi_1}}\right) \le P\left(\frac{|S_n|}{\sqrt{n \operatorname{Var} \xi_1}} > K\right)$$

tetszőleges $\varepsilon > 0$ és $K > 0$ számra, ha $n \ge n_0(\varepsilon, K)$. Ezért a centrális határeloszlástétel alapján $P\left(\dfrac{|S_n|}{a_n} > \varepsilon\right) \le \delta$, ha $n \ge n_0 = n_0(\varepsilon, \delta)$, azaz az $\dfrac{S_n}{a_n}$ sorozat sztochasztikusan tart nullához.
Megfordíta, ha az $\dfrac{a_n}{\sqrt{n}}$ sorozat nem tart végtelenhez, akkor létezik a természetes számoknak olyan $n_k \to \infty$ részsorozata, amelyre $\limsup\limits_{k\to\infty} \dfrac{a_{n_k}}{\sqrt{n_k}} \le L$, ezért

$$\liminf_{k\to\infty} P\left(\frac{|S_{n_k}|}{a_{n_k}} > 1\right) = \liminf_{k\to\infty} P\left(\frac{|S_{n_k}|}{\sqrt{n_k \operatorname{Var} \xi_1}} > \frac{a_{n_k}}{\sqrt{n_k \operatorname{Var} \xi_1}}\right)$$
$$\ge \liminf_{k\to\infty} P\left(\frac{|S_{n_k}|}{\sqrt{n \operatorname{Var} \xi_1}} > \frac{L}{\sqrt{\operatorname{Var} \xi_1}}\right) > 0$$

a centrális határeloszlástétel alapján. Ezért ebben az esetben az $\dfrac{S_n}{a_n}$ sorozat nem tar sztochasztikusan nullához.