A november 19.-i előadáshoz kapcsolódó gyakorlat feladatai

1. Péter és Pál a következő játékot játszhatja. Egymás után 10 000 alkalommal feldobnak egy szabályos kockát. Ha a kocka páros oldalra esik, akkor Péter kap a banktól annyi forintot, amennyi a dobás értéke, ha a dobás értéke hárommal osztható akkor Pál kap a banktól annyi forintot, amennyi a dobás értéke. (Páratlan dobás esetén Péter nem nem nyer semmit, és hasonlóan hárommal nem osztható dobás esetén Pál nem nyer semmit. Mi annak a valószínűsége, hogy Péter legalább 19 800 és Pál legalábbb 15 200 forintot nyer? Hogyan tudjuk ezt közelítőleg jól kiszámolni a több-dimenziós centrális határeloszlástétel és egy számítógép segítségével?

   Megoldás: Jelölje Y_j Péter és Z_j Pál nyereményét a j-ik lépésben, $1 \le j \le 10\ 000$, és vezessük be az $S_j = \sum_{j=1}^{10\ 000} Y_j$, és $T_j = \sum_{j=1}^{10\ 000} Z_j$, valószínűségi változókat. Akkor minket a $P(S \ge 19\ 800, T \ge 15\ 200)$ valószínűség érdekel. Ezt a több-dimenziós centrális határeloszlástétel segÍtségével akarjuk megtenni. Ahhoz, hogy ezt megtegyük, számoljuk ki a (független és egyforma eloszlású) (Y_j, Z_j) vektorok várható érték vektorát és kovariancia mátrixát. $EY_j = \frac{1}{6}(2+4+6) = 2$, mert Péter Y_j nyereménye $\frac{1}{6}$ valószínűséggel lesz 2, 4 vagy 6 forint, és $\frac{1}{2}$ valószínűséggel 0 forint. Hasonlóan $EZ_j = \frac{1}{6}(3+6) = \frac{3}{2}$, $EY_j^2 = \frac{1}{6}(4+16+36) = \frac{28}{3}$, $EZ_j^2 = \frac{6}{(}9+36) = \frac{15}{2}$, $\operatorname{Var} Y_j = EY_j^2 - (EY^j)^2 = \frac{16}{3}$. $\operatorname{Var} Z_j = EZ_j^2 - (EZ_j)^2 = \frac{21}{4}$. Végül $EY_jZ_j = \frac{1}{6} \cdot 36 = 6$, mert $Y_jZ_j = 36$, ha hatos dobás történik, aminek a valószínűsége $\frac{1}{6}$, és a maradék $\frac{5}{6}$ valószínűségi esemény bekövetkezésekor nulla, ahonnan $\operatorname{Cov}(Y_j, Z_j) = EY_jZ_j - EY_jEZ_j = 6 - 3 = 3$.

   A fentiek alapján

$$P(S \ge 19\ 800, T \ge 15\ 200) = P\left(\frac{S - ES}{\sqrt{10\ 000}} \ge -2, \frac{T - ET}{\sqrt{10\ 000}} \ge 2\right)$$

   és a több-dimenziós centrális határeloszlástétel segítségével a fenti valószínűség annak valószínűségével közelíthető, hogy $P(\eta_1 \ge -2, \eta_2 \ge 2)$ egy olyan normális eloszlású (η_1, η_2) véletlen vektorra, melyre $E\eta_1 = E\eta_2 = 0$, és e vektor kovariancia mátrixa a $D = \begin{pmatrix} \frac{16}{3}, 3 \\ 3, \frac{21}{4} \end{pmatrix}$. Számítsuk ki ennek a normális eloszlású vektornak a sűrűségfüggvényét. Ennek érdekében először kiszámítjuk a D mátrix D^{-1} inverzét. Azt kapjuk, hogy $\det D = 19$, $D^{-1} = \dfrac{1}{19}\begin{pmatrix} \frac{21}{4}, -3 \\ -3, \frac{16}{3} \end{pmatrix}$, ahonnan $-\frac{1}{2}(x,y)D^{-1}(x,y)^* = -\frac{1}{38}\left(\frac{21}{4}x^2 + \frac{16}{3}y^2 - 6xy\right)$. Ezért a keresett sűrűségfüggvény $f(x,y) = \frac{1}{2\pi\sqrt{\det D}} e^{-(x,y)D^{-1}(x,y)^*/2}$ alakját felhasználva kapjuk, hogy

$$P(S \ge 19\ 800, T \ge 15\ 200) \sim \int_{-2}^{\infty} \int_{2}^{\infty} \frac{1}{2\pi\sqrt{19}} \exp\left\{-\frac{1}{38}\left(\frac{21}{4}x^2 + \frac{16}{3}y^2 - 6xy\right)\right\} dx\, dy.$$

2.) Minden $x > 0$-ra

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\varphi(x) < 1 - \Phi(x) < \frac{1}{x}\varphi(x),$$

ahol $\Phi(x)$ és $\varphi(x)$ a standard normális eloszlás és sűrűségfüggvény.

Megoldás: Rarciális integrálással kapjuk, hogy

$$\begin{aligned} 1 - \Phi(x) &= \int_x^\infty e^{-u^2/2}\,du == -\int_x^\infty \frac{1}{u}\frac{de^{-u^2/2}}{du}\,du = \frac{1}{x}e^{-x^2/2} - \int_x^\infty \frac{1}{u^2}e^{-u^2/2}\,du \\ &= \frac{1}{x}e^{-x^2/2} - \frac{1}{x^3}e^{-x^2/2} + \int_x^\infty \frac{3}{u^4}e^{-u^2/2}\,du\,, \end{aligned}$$

és használjuk fel, hogy $\int_x^\infty \frac{1}{u^2}e^{-u^2/2}\,du \ge 0$, és $\int_x^\infty \frac{3}{u^4}e^{-u^2/2}\,du \ge 0$. Innen látható az állítás.

3. Legyenek $\xi_1, \dots, \xi_n$ független exponenciális eloszlású valószínűségi változók $\lambda > 0$ paraméterrel, $S_k = \sum\limits_{j=1}^{k} \xi_j$, $1 \le k \le n$. A $\left(\frac{S_1}{S_n}, \dots, \frac{S_{n-1}}{S_n}\right)$ vektor független az S_n valószínűségi változótól, és e vektor elemei a $[0,1]$ intervallumban független, egyenletes eloszlású valószínűségi változókból álló rendezett mintát alkotnak.

Ezt a feladatot úgy fogjuk megoldani, hogy kiszámoljuk a definiált véletlen vektor sűrűségfüggvényét, és ellenőrizzük, hogy ez olyan tulajdonságú, mint ahogy azt állítottuk. Ahhoz, hogy ezt a programot végrehajtsuk szükségünk van arra, hogy tudjuk bizonyos transzformációk során hogyan változik egy integrál. Felidézem azt az eredményt, amelyikre szükségünk van. Ezt a szükségesnél általánosabb formában fogalmazom meg. A most vizsgált problémában elegendő lenne invertálható leképezéseket vizsgálni, ahol minden pontnak egyetlen őse képe van. A szükséges eredmény megfogalmazása érdekében először felidézem a Jacobian foglmát.

Jacobian definiciója. *Legyen $y_k = T_k(x_1, \dots, x_n)$, $1 \le k \le n$, az n-dimenziós tér egy tartományának síma transzformáltja az n-dimenziós tér egy másik tartományába. E transzformáció $\mathcal{J}(\mathbf{T}(x_1, \dots, x_n))$ Jakobiánja egy $(x_1, \dots, x_n)$ pontban a*

$$\left(\frac{\partial T_k(x_1, \dots, x_k)}{\partial x_n}\right), \quad 1 \le l, k \le n,$$

$n \times n$-es (az $(x_1, \dots, x_n)$ pontban vett derivált) mátrix determinánsának az abszolut értéke.

(A Jacobian szemléletes tartalma: Ez adja meg, hogy az $(x_1, \dots, x_n)$ pont kis környezetének a térfogatát a $\mathbf{T} = (T_1, \dots, T_n)$ transzformáció hányszorosára nagyítja ki.)

Integráltranszformációról szóló képlet. *Legyen adva az n-dimenziós tér egy A tartományának egy síma $y_k = T_k(x_1, \dots, x_n)$, $1 \le k \le n$, transzformáltja az n-dimenziós tér egy másik B tartományába. Legyen továbbá adva a B tartományon egy $f(y_1, \dots, y_n)$*

függvény. Ezen $f(y_1,\dots,y_n)$ függvénynek a $\mathbf{T}=(T_1,\dots,T_n)$ transzformáció által meghatározott ősképén azt az A tartományon értelmezett $g(x_1,\dots,x_n)=\mathbf{T}^{-1}f(x_1,\dots,x_n)$ függvényt értjük az A tartományon, melyre

$$g(x_1,\dots,x_n)=f(T_1(x_1,\dots,x_n),\dots,T_n(x_1,\dots,x_n))$$

minden $(x_1,\dots,x_n)\in A$ pontban. Ekkor

$$\begin{aligned}&\int_B f(y_1,\dots,y_n)\,dy_1\dots dy_n\\ &\qquad=\int_A \frac{\mathbf{T}^{-1}f(x_1,\dots,x_n)}{\sum\limits_{\substack{\text{olyan }(z_1,\dots,z_n)\in A\text{ pontok}\\ \text{melyekre }T_k(z_1,\dots,z_n)=T_k(x_1,\dots,x_n)\\ k=1,\dots,n}}\dfrac{1}{\mathcal{J}(\mathbf{T}(z_1,\dots,z_k))}}\,dx_1\dots dx_n.\end{aligned}$$

A 3. feladat megoldása: Számoljuk ki az $(S_1,\dots,S_n)$ vektor sűrűségfüggvényét.

$$P(S_1<u_1,\dots,S_n<u_n)=P((\xi_1,\dots,\xi_n)\in B)=\int_B \lambda^n e^{-\lambda(y_1+\dots+y_n)}\,dy_1\,\dots\,dy_n,$$

ahol $B=B(u_1,\dots,u_n)=\{(y_1,\dots,y_n)\colon y_j\ge 0, y_1+y_2+\dots+y_j<u_j,\ j=1,\dots,n\}$.
Írjuk át a fenti integrált az $x_j=y_1+\dots+y_j$, $j=1,2,\dots,n$ transzformációval. E transzformáció invertálható. Jacobiánja 1, a B halmazt az $\{0\le x_j<u_j,\ j=1,\dots,n\}$ halmazba képezi, ezért

$$P(S_1<u_1,\dots,S_n<u_n)=\int_{\{0\le x_1\le x_2<\dots\le x_n,\ x_j<u_j,\ j=1,\dots,n\}}\lambda^n e^{-\lambda x_n}\,dx_1\,\dots\,dx_n,$$

ahonnan az $(S_1,\dots,S_n)$ vektor sűrűségfüggvénye $f(x_1,\dots,x_n)=\lambda^n e^{-\lambda x_n}$, ha $0\le x_1\le x_2\le\dots\le x_n$, és $f(x_1,\dots,x_n)=0$ egyébként. Innen az S_n sűrűségfüggvénye

$$f_n(x_n)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\dots,x_n)\,dx_1\dots dx_{n-1}=\frac{\lambda^n}{(n-1)!}x^{n-1}e^{-\lambda x_n},$$

ha $x_n\ge 0$.
Számítsuk ki az $\left(\dfrac{S_1}{S_n},\dots,\dfrac{S_{n-1}}{S_n},S_n\right)$ vektor sűrűségfüggvényét az $(S_1,\dots,S_n)$ véletlen vektor $f(x_1,\dots,x_n)$ sűrűségfüggvényének ismeretében.

$$\begin{aligned}&P\left(\frac{S_1}{S_n}<u_1,\dots,\frac{S_{n-1}}{S_n}<u_{n-1},S_n<x\right)\\ &\quad=P(S_1<S_nu_1,\dots,S_{n-1}<S_nu_{n-1},S_n<x)\\ &\quad=\int_{\{0\le y_j<u_jy_n,\,j=1,\dots,n-1,\,0\le y_n<x\}}f(y_1,\dots,y_n)\,dy_1\dots dy_n.\end{aligned}$$

Végezzük el az $x_j = y_j y_n$, $j = 1, 2, \dots, n-1$ és $x_n = y_n$ helyettesítésket. E transzformáció Jacobiánja x_n^{n-1}, és

$$P\left(\frac{S_1}{S_n} < u_1, \dots, \frac{S_{n-1}}{S_n} < u_{n-1}, S_n < x\right)$$
$$= \int_{\{0 \le x_j < u_j,\, j=1,\dots,n-1,\, 0 \le x_n < x\}} x_n^{n-1} f\left(\frac{x_1}{x_n}, \dots, \frac{x_{n_1}}{x_n}, x_n\right) dx_1 \dots dx_n.$$

Innen a keresett sűrűségfüggvény $f(x_1, \dots, x_{n-1}, x_n) = \lambda^n e^{-\lambda x_n} x_n^{n-1}$, ha $0 \le x_1 < x_1 < \dots < x_{n-1} \le 1$, és $x_n \ge 0$. Ez azt jelenti, hogy $f(x_1, \dots, x_{n-1}, x_n) = g(x_1, \dots, x_{n-1}) h(x_n)$, ahol $g(x_1, \dots, x_{n-1}) = (n-1)!$, ha $0 \le x_1 \le \dots \le x_{n-1} \le 1$, $g(x_1, \dots, x_{n-1}) = 0$, egyébként, $h(x_n) = \dfrac{x_n^{n-1}}{(n-1)!} e^{-\lambda x_n}$, ha $x_n \ge 0$, $h(x_n) = 0$, ha $x_n < 0$. Ez azt jelenti, hogy az $\left(\dfrac{S_1}{S_n}, \dots, \dfrac{S_{n-1}}{S_n}\right)$ vektor és S_n valószínűségi változók függetlenek, a tekintett vektor eloszlása megegyezik $n-1$ független, a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változóból készített rendezett minta eloszlásával, az S_n vektor eloszlása pedig n független λ paraméterű exponenciális eloszlású valószínűségi változó összegének az eloszlása. (Azt, hogy $h(x_1, \dots, x_{n-1})$ egy $n-1$ elemű egyenletes eloszlású minta sűrűségfüggvénye megbeszéltük egy korábbi gyakorlaton.)

Házi feladat:

Legyenek $\xi_1, \dots, \xi_n$ független, a $[0,1]$ intervallumon egyenletes eloszlású valószínűségi változók, és jelölje $\xi_1^* = \max\limits_{1 \le j \le n} \xi_j$ a maximumukat. Határozzuk meg a (ξ_1, ξ_1^*) vektor eloszlását. Lássuk be, hogy ezt természetes két részre osztani annak megfelelően, hogy $\xi_1 = \xi_1^*$ vagy $\xi_1 \ne \xi_1^*$. Az egyik mérték az átlóra van koncentrálva és ott sűrűségfüggvénye x^{n-1}, a másik résznek az egész négyzeten van sűrűségfüggvénye, és az $(n-1)y^{n-2}$, ha $0 \le x \le y \le 1$ és nulla különben.