Vizsgadolgozat, 2020. január 9.
Megoldás

Tanszéki általános alapelvek

A pontozási útmutató célja, hogy a javítók a dolgozatokat egységesen értékeljék. Ezért az útmutató minden feladat (legalább egy lehetséges) megoldásának főbb gondolatait, és az ezekhez rendelt részpontszámokat közli. Az útmutatónak nem célja a feladatok teljes értékű megoldásának részletes leírása; a leírt lépések egy maximális pontszámot érő megoldás vázlatának tekinthetők.

Az útmutatóban feltüntetett részpontszámok csak akkor járnak a megoldónak, ha a kapcsolódó gondolat egy áttekinthető, világosan leírt és megindokolt megoldás egy lépéseként szerepel a dolgozatban. Így például az anyagban szereplő ismeretek, definíciók, tételek puszta leírása azok alkalmazása nélkül nem ér pontot (még akkor sem, ha egyébként valamelyik leírt tény a megoldásban valóban szerephez jut). Annak mérlegelése, hogy az útmutatóban feltüntetett pontszám a fentiek figyelembevételével a megoldónak (részben vagy egészében) jár-e, teljes mértékben a javító hatásköre.

Részpontszám jár minden olyan ötletért, részmegoldásért, amelyből a dolgozatban leírt gondolatmenet alkalmas kiegészítésével a feladat hibátlan megoldása volna kapható. Ha egy megoldó egy feladatra több, egymástól lényegesen különböző megoldást is elkezd, akkor legfeljebb az egyikre adható pontszám. Ha mindegyik leírt megoldás vagy megoldásrészlet helyes vagy helyessé kiegészíthető, akkor a legtöbb részpontot érő megoldáskezdeményt értékeljük. Ha azonban több megoldási kísérlet között van helyes és (lényeges) hibát tartalmazó is, továbbá a dolgozatból nem derül ki, hogy a megoldó melyiket tartotta helyesnek, akkor a kevesebb pontot érő megoldáskezdeményt értékeljük (akkor is, ha ez a pontszám 0). Az útmutatóban szereplő részpontszámok szükség esetén tovább is oszthatók. Az útmutatóban leírttól eltérő jó megoldás természetesen maximális pontot ér.

Artimetikai hiba esetén elszámolásonként 1-1 pont vonandó le a feladatokból. Ez alól kivétel, ha az elszámolás lényegesen egyszerűsíti vagy módosítja a feladat felépítését. Ilyen esetekben azon feladatrészekért, amik az elszámolás okán fel sem merültek, nem jár pont.

1. Bélának két szabálytalan érme van a zsebében: egy olyan, ami $\frac{1}{3}$ eséllyel esik a fej oldalára, illetve egy olyan, aminek mindkét oldala írás. Véletlenszerűen előveszi az egyik érmét, és dob vele négyszer. Feltéve, hogy páros sok írást lát, mi az esélye, hogy az első érmét húzta?

   (2 pont) $A_1 = \{\text{első pénzérmét választjuk}\}$, $A_2 = \{\text{második pénzérmét választjuk}\}$,
   $B = \{\text{páros sok írást dobunk}\}$ (Ha az eseményekre jelölést nem vezet be, de a nekik megfelelő szövegek következetesen vannak használva a továbbiakban, szintén jár a pont.)
   (1 pont) $\mathbb{P}(A_1|B) =?$
   (4 pont) $\mathbb{P}(B|A_1) = \left(\frac{1}{3}\right)^4 + \binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^4 = \frac{41}{81} \approx 0{,}5062$
   (2 pont) $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \frac{1}{2}$ és $\mathbb{P}(B|A_2) = 1$
   (1 pont) Mivel A_1 es A_2 teljes esemenyrendszert alkot, ezert
   (1 pont) a teljes valószínűség tétele alapján (Ha rögtön az összetett Bayes-tétel van hivatkozva, akkor is jár a pont.)
   (4 pont) $\mathbb{P}(B) = \mathbb{P}(B|A_1)\,\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\,\mathbb{P}(A_2) = \frac{61}{81} \approx 0.7531$
   (2 pont) így a Bayes-tétel alapjan
   (2 pont) $\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1)\,\mathbb{P}(A_1)}{\mathbb{P}(B)}$
   (1 pont) $\mathbb{P}(A_1|B) = \frac{41}{122} \approx \underline{\underline{0.3361}}$

2. Válasszunk ki egyenletesen véletlenszerűen egy (U, V) pontot a $[0; 1] \times [0; 1]$ egységnégyzetből. Legyen $X = \sqrt{U}$ és $Y = \sqrt{V}$.
   a) Határozzuk meg X eloszlásfüggvényét.
   b) Határozzuk meg $X + Y$ sűrűségfüggvényét.

   (1 pont) $\mathrm{Ran}(X) = \mathrm{Ran}(Y) = [0; 1]$
   (3 pont) Ha $0 < x < 1$, akkor $F_X(x) = \mathbb{P}(X < x) = \mathbb{P}(\sqrt{U} < x) = \mathbb{P}(U < x^2) = x^2$,
   (0 pont) valamint $F_X(x) = 0$, ha $x \le 0$ és $F_X(x) = 1$ ha $x \ge 1$. (Ha ez szerepel, de $\mathrm{Ran}(X) = [0; 1]$ nem, akkor is jár az első 1 pont.)
   (3 pont) Ezért $f_X(x) = 2x$, ha $0 < x < 1$, és 0 egyébként. Ugyanígy Y-ra. (Ha csak X-re van felírva, de Y-ra is van használva, 2 pont.)

(1 pont) $\mathrm{Ran}(X+Y) = [0;2]$ (Ezt helyettesíti, ha szerepel, hogy $f_{X+Y}(z) = 0$, ha $z \le 0$ vagy $z > 2$.)
(3 pont) $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\mathrm{d}x$ (Elírással nem fogadható el.)
(2 pont) $= \int_{\max(0,z-1)}^{\min(1,z)} 2x \cdot 2(z-x)\mathrm{d}x =$ (A két eset, mikor $z < 1$ illetve $z > 1$, külön integrálban is szerepelhet, akkor is jár a pont.)
(1 pont) $\left[2zx^2 - \frac{4}{3}x^3\right]_{\max(0,z-1)}^{\min(1,z)}$
(2 pont) Ha $0 < z < 1$, akkor $\left[2zx^2 - \frac{4}{3}x^3\right]_0^z = \frac{2}{3}z^3$
(2 pont) Ha $1 < z < 2$, akkor $\left[2zx^2 - \frac{4}{3}x^3\right]_{z-1}^1 = -\frac{2}{3}(z^3 - 6z + 4)$
(2 pont) Tehát

$$f_{X+Y}(z) = \begin{cases} \frac{2}{3}z^3 & \text{ha } 0 < z < 1, \\ -\frac{2}{3}(z^3 - 6z + 4) & \text{ha } 1 < z < 2, \\ 0 & \text{egyébként.} \end{cases}$$

3. Bence rendelt n zsák I.Z.-t. Az egyes zsákokban lévő I.Z. mennyisége véletlenszerű: várhatóan 10 kg-ot tartalmaz, de 2 kg szórással. (A különböző zsákokban lévő mennyiségek függetlenek.) Ha a várható $10n$ kg-nál legalább 20 kg-mal több I.Z. érkezik, akkor a 20 kg extrán felüli részt kidobja. Tegyük fel, hogy Bence 0,0336 valószínűséggel dob ki valamennyi I.Z.-t. Hány zsákot rendelhetett?

(2 pont) $X_1, \ldots, X_n$: az egyes zsákok tömege
(1 pont) $\mathbb{E}(X_i) = 10$, $\mathbb{D}(X_i) = 2$
(2 pont) $\mathbb{P}(\sum_{i=1}^n X_i > 10n + 20) = 0{,}0336$
(3 pont) $\mathbb{P}(\sum_{i=1}^n X_i > 10n + 20) = \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - 10n}{2\sqrt{n}} > \frac{20}{2\sqrt{n}}\right)$
(2 pont) A centrális határeloszlás-tétel miatt
(3 pont) $\frac{\sum_{i=1}^n X_i - 10n}{2\sqrt{n}}$ közelítőleg standard normális eloszlású.
(3 pont) Így a fenti valószínűségre azt kapjuk, hogy $0{,}0336 \approx 1 - \Phi\left(\frac{20}{2\sqrt{n}}\right)$
(1 pont) $(\Phi^{-1}(0{,}9664)) \approx 1{,}83$ (kicsit pontosabb értéke 1,8303),
(1 pont) így $\sqrt{n} \approx \frac{10}{1{,}83} \approx 5{,}46$
(2 pont) $n \approx \underline{\underline{30}}$. (Egészre nem kerekítve 1 pont.)

4. A képzeletbeli Szürrealéziában a vasúttársaság örökifjú stratégiát követ. Ez sajnos nem a vonatok belső állapotára utal, hanem azt jelenti, hogy egy vonat késésének ideje örökifjú eloszlású, folytonos, pozitív értékű valószínűségi változó. Tegyük fel, hogy egy késés átlagos ideje 3 fertályóra. A vasúttársaság a késés mértékének függvényében panaszleveleket kap. Rögzített x fertályóra késés esetén a panaszlevelek egymástól független, azonos, de egyenként kis valószínűségű események. Ha egy vonat x fertályórát késik, akkor annak a valószínűsége, hogy egy panaszlevelet sem kapnak e^{-100x}.
a) Mi a panaszlevelek számának (feltételes) várható értéke, ha rögzített x idejű késéssel számolunk.
b) Várhatóan hány levelet kap a vasúttársaság egy adott vonat késése miatt?

a) (1 pont) X: adott vonat késése, Y: panaszlevelek száma
(Opcionálisan definiálható: Z a rögzített $X = x$ esetén a panaszlevelek száma)
(1 pont) $\mathbb{E}(Y \mid X = x) = ?$ (vagy $\mathbb{E}(Z) = ?$)
(2 pont) Rögzített x késés esetén Y Poisson eloszlású (avagy $Z \sim \mathrm{Pois}(\mu)$)
ebből max. 1 pont, ha nem szerepel az x mint feltétel (vagy a Z korábban definiálva)
(3 pont) $\mathbb{P}(Y = 0 \mid X = x) = e^{-100x}$ (avagy $\mathbb{P}(Z = 0) = e^{-100x}$)
ebből max. 2 pont, ha nem szerepel az x mint feltétel (vagy a Z korábban definiálva)
(2 pont) $\mathbb{E}(Y \mid X = x) = 100x$, mert rögzített x késés esetén Y eloszlása $\mathrm{Pois}(100x)$
(avagy $\mathbb{E}(Z) = 100x$, mert $Z \sim \mathrm{Pois}(100x)$)
b) (1 pont) $\mathbb{E}(Y) = ?$
(1 pont) $X \sim \mathrm{Exp}(\lambda)$
(1 pont) $\mathbb{E}(X) = 3 \Rightarrow \lambda = \frac{1}{3}$
(1 pont) A teljes várható érték tétele szerint
(3 pont) $\mathbb{E}(Y) = \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x)\mathrm{d}x$ (avagy $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X))$)
(2 pont) $= \int_0^\infty 100x \cdot \lambda e^{-\lambda x}\mathrm{d}x$ (avagy $\mathbb{E}(Y \mid X) = 100X$ ezért $= \mathbb{E}(100X) = 100\,\mathbb{E}(X)$)
(2 pont) $= 100\lambda\Big(\left[x\frac{e^{-\lambda x}}{-\lambda}\right]_0^\infty - \int_0^\infty \frac{e^{-\lambda x}}{-\lambda}\mathrm{d}x\Big) = 100\lambda\left[\frac{e^{-\lambda x}}{-\lambda^2}\right]_0^\infty = \frac{100}{\lambda} = \underline{\underline{300}}$

5. Legyen $(X, Y) \sim N(\underline{0}, \underline{\underline{\Sigma}})$ kétdimenziós normális eloszlású valószínűségi vektorváltozó. Tegyük fel, hogy $\mathbb{D}^2(X) = \mathbb{D}^2(Y)$, X és Y korrelációja 0,5, és $\det(\underline{\underline{\Sigma}}) = 12$.
a) Határozzuk meg $\underline{\underline{\Sigma}}$-t.
b) Hány százalék az esélye, hogy X nagyobb, mint 3,6?
(2 pont) $\underline{\underline{\Sigma}} = \begin{bmatrix} \mathbb{D}^2(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(Y,X) & \mathbb{D}^2(Y) \end{bmatrix}$
(1 pont) $\mathrm{cov}(Y, X) = \mathrm{cov}(X, Y)$ (Ha csak implicit módon, de használva van, akkor is 1 pont.)
(2 pont) $\det(\underline{\underline{\Sigma}}) = \mathbb{D}^2(X)\,\mathbb{D}^2(Y) - \mathrm{cov}(X,Y)^2$
(2 pont) $\mathrm{corr}(X, Y) = \frac{\mathrm{cov}(X,Y)}{\mathbb{D}(X)\,\mathbb{D}(Y)}$, ezért
(2 pont) $\det(\underline{\underline{\Sigma}}) = \mathbb{D}^2(X)\,\mathbb{D}^2(Y) \cdot (1 - \mathrm{corr}(X,Y)^2) = \mathbb{D}^4(X)(1 - 0{,}5^2)$
(vagy bármilyen ezzel analóg számolás, ami a fentiekből levezet $\det(\underline{\underline{\Sigma}})$ és $\mathbb{D}(X)$ közti összefüggést)
(1 pont) $\Rightarrow 12 = \det(\underline{\underline{\Sigma}}) = \frac{3}{4}\mathbb{D}^4(X)$
(1 pont) $\Rightarrow \mathbb{D}(X) = 2$
(2 pont) $\mathrm{cov}(X,Y) = \mathrm{corr}(X,Y) \cdot \mathbb{D}(X)\,\mathbb{D}(Y) = 0{,}5 \cdot 2 \cdot 2 = 2$
(1 pont) Tehát $\underline{\underline{\Sigma}} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$
(1 pont) $\mathbb{P}(X > 3{,}6) = ?$
(1 pont) $= 1 - \mathbb{P}(X < 3{,}6)$
(1 pont) Normális eloszlás komponensei is normálisak, ezért
(2 pont) $= 1 - \Phi\left(\frac{3{,}6 - \mathbb{E}(X)}{\mathbb{D}(X)}\right)$
(1 pont) $\approx \underline{\underline{0{,}0359}} = 3{,}59\%$

6.* Legyen (U, V) valószínűségi vektorváltozó. Tegyük fel, hogy $\mathbb{D}^2(U)$, $\mathbb{D}^2(V)$ és $\mathrm{cov}(U, V)$ pozitív valós számok. Legyen V lineáris regressziója U-ra $\beta_1 U + \alpha_1$, és U lineáris regressziója $-V$-re $\beta_2(-V) + \alpha_2$. Tegyük fel, hogy az $\{(x, y) \in \mathbb{R}^2 \mid y = \beta_1 x + \alpha_1\}$ egyenes merőleges az $\{(x, y) \in \mathbb{R}^2 \mid y = \beta_2 x + \alpha_2\}$ egyenesre. Mit mondhatunk V eloszlásáról, ha tudjuk, hogy U exponenciális eloszlású? És ha feltesszük, hogy $\alpha_1 = \alpha_2$?
(3 pont) $\beta_1 = \frac{\mathrm{cov}(U,V)}{\mathbb{D}^2(U)}$, és $\beta_2 = \frac{\mathrm{cov}(U,-V)}{\mathbb{D}^2(-V)}$
(2 pont) Az egyenesek pontosan akkor merőlegesek, ha $\beta_1 \cdot \beta_2 = -1$
(2 pont) $\beta_1 \cdot \beta_2 = -1$ pontosan akkor teljesül, ha $-\frac{\mathrm{cov}(U,V)^2}{\mathbb{D}^2(U)\,\mathbb{D}^2(V)} = -1$.
(2 pont) $\Rightarrow \mathrm{corr}(U, V) = \pm 1$
(1 pont) $\mathrm{cov}(U, V) > 0$, ezért $\mathrm{corr}(U, V) > 0$, tehát $+1$.
(3 pont) $\mathrm{corr}(U, V) = 1 \Rightarrow V = bU + a$ valamilyen $b, a \in \mathbb{R}$ esetén, ahol $b > 0$.
(2 pont) Ha U exponenciális, akkor V egy exponenciális eloszlású val. változó lin. transzformáltja.
(2 pont) Előadás alapján exp. eloszlású valószínűségi változó pozitív konstansszorosa is exp. eloszlású.
(1 pont) $\alpha_1 = \mathbb{E}(V) - \frac{\mathrm{cov(U,V)}}{\mathbb{D}^2(U)}\,\mathbb{E}(U) = b\,\mathbb{E}(U) + a - b\,\mathbb{E}(U) = a$
(1 pont) $\alpha_2 = \mathbb{E}(U) - \frac{\mathrm{cov(U,-V)}}{\mathbb{D}^2(-V)}\,\mathbb{E}(-V) = \mathbb{E}(U) + \frac{1}{b}(-b\,\mathbb{E}(U) - a) = -\frac{a}{b}$
(1 pont) Ha $\alpha_1 = \alpha_2$, akkor $b \neq -1$ miatt $a = 0$, tehát $V \sim \mathrm{Exp}\left(\frac{\lambda}{b}\right)$, ha $U \sim \mathrm{Exp}(\lambda)$.