Vizsgadolgozat, 2020. január 16.
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. Bonifác fonetikus fogyókúrába kezd, azaz minden héten véletlenszerűen kiválasztja a hét egy hétköznapját, és aznap annyi kilométert fut, ahány mássalhangzó van az adott hétköznap nevében. A véletlenszerű választás nem egyenletes: az egyes hétköznapokat olyan arányban választja, ahány magánhangzó van a nap nevében. (Bonifác más napokon nem fut. Hétvégén pláne nem.) Mennyi az egy héten futott kilométerek számának várható értéke és szórásnégyzete?

   A darabszámokat és összegeiket a következő táblázat foglalja össze (pl. "dd" illetve "sz" egy hangnak számít): (2 pont) X: az egy héten futott kilométerek száma

Nap	Hétfő	Kedd	Szerda	Csütörtök	Péntek	$\sum$
Mássalhangzók száma	3	2	3	5	4	17
Magánhangzók száma	2	1	2	3	2	10

(2 pont) Az egyes napokat $\frac{2}{10}$, $\frac{1}{10}$, $\frac{2}{10}$, $\frac{3}{10}$, $\frac{2}{10}$ valószínűségekkel választja.
(1 pont) $\mathbb{P}(X=2)=\frac{1}{10}$, $\mathbb{P}(X=4)=\frac{2}{10}$, $\mathbb{P}(X=5)=\frac{3}{10}$
(1 pont) $\mathbb{P}(X=3)=\frac{2}{10}+\frac{2}{10}=\frac{4}{10}$
(1 pont) $\mathrm{Ran}(X)=\{2,3,4,5\}$ (Ha szummában implicit fel van használva, akkor is jár a pont.)
(2 pont) $\mathbb{E}(X)=\sum_{k\in\mathrm{Ran}(X)} k\cdot\mathbb{P}(X=k)$ (Ha a szumma futó indexe 2-től 5-ig fut, az is 2 pont.)
(3 pont) $=2\cdot\frac{1}{10}+3\cdot\frac{4}{10}+4\cdot\frac{2}{10}+5\cdot\frac{3}{10}$
(1 pont) $=\underline{\underline{3{,}7}}$
(2 pont) $\mathbb{D}^2(X)=\mathbb{E}(X^2)-\mathbb{E}(X)^2$,
(2 pont) $\mathbb{E}(X^2)=\sum_{k\in\mathrm{Ran}(X)} k^2\cdot\mathbb{P}(X=k)$
(2 pont) $=2^2\cdot\frac{1}{10}+3^2\cdot\frac{4}{10}+4^2\cdot\frac{2}{10}+5^2\cdot\frac{3}{10}=14{,}7$
(1 pont) $\mathbb{D}^2(X)=14{,}7-3{,}7^2=\underline{\underline{1{,}01}}$

2. Legyen (X,Y) olyan folytonos valószínűségi vektorváltozó, aminek sűrűségfüggvénye valamilyen $\alpha \in \mathbb{R}$ esetén
$$f_{X,Y}: (x,y) \mapsto \begin{cases} \frac{\alpha}{2x} & \text{ha } 1 < x < 3 \text{ és } 0 < y < x, \\ 0 & \text{egyébként.} \end{cases}$$
Határozzuk meg $\mathrm{cov}(X,Y)$-t.

(2 pont) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 1$
(2 pont) $= \int_1^3 \int_0^x \frac{\alpha}{2x}\mathrm{d}y\mathrm{d}x$
(2 pont) $= \frac{\alpha}{2}\int_1^3\int_0^x x^{-1}\mathrm{d}y\mathrm{d}x = \frac{\alpha}{2}\int_1^3 \left[yx^{-1}\right]_0^x \mathrm{d}x = \frac{\alpha}{2}\int_1^3 1\mathrm{d}x = \alpha$
(1 pont) ezért $\alpha = 1$
(3 pont) $\mathbb{E}(XY) = \int_1^3\int_0^x \frac{1}{2}xyx^{-1}\mathrm{d}y\mathrm{d}x = \int_1^3 \left[\frac{y^2}{4}\right]_0^x \mathrm{d}x = \int_1^3 \frac{x^2}{4}\mathrm{d}x = \left[\frac{x^3}{12}\right]_1^3 = \frac{27}{12} - \frac{1}{12} = \frac{13}{6}$
(3 pont) $\mathbb{E}(X) = \int_1^3\int_0^x \frac{1}{2}xx^{-1}\mathrm{d}y\mathrm{d}x = \int_1^3 \left[\frac{1}{2}\right]_0^x \mathrm{d}x = \int_1^3 \frac{x}{2}\mathrm{d}x = \left[\frac{x^2}{4}\right]_1^3 = \frac{9}{4} - \frac{1}{4} = 2$
(A képlet és a peremsűrűségfüggvény a helyes várható érték nélkül legfeljebb 2 pontot ér.)
(3 pont) $\mathbb{E}(Y) = \int_1^3\int_0^x \frac{1}{2}yx^{-1}\mathrm{d}y\mathrm{d}x = \int_1^3 \left[\frac{y^2}{4x}\right]_0^x \mathrm{d}x = \int_1^3 \frac{x}{4}\mathrm{d}x = \left[\frac{x^2}{8}\right]_1^3 = \frac{9}{8} - \frac{1}{8} = 1$
(A képlet és a peremsűrűségfüggvény a helyes várható érték nélkül legfeljebb 2 pontot ér.)
(3 pont) $\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$
(1 pont) $= \frac{13}{6} - 2\cdot 1 = \frac{1}{6} \approx \underline{\underline{0{,}1667}}$

3. Egy gépnek egymás után 50 utasítást kell végrehajtania, amelyek véletlenszerű ideig tartanak. Az egyes utasítások végrehajtásához szükséges idők egymástól függetlenek és azonos eloszlásúak. Az egy utasításhoz szükséges idő várható értéke éppen az egy utasításhoz szükséges idő szórásának a duplája. Mennyi a valószínűsége, hogy az összes utasítás végrehajtásához szükséges idő több, mint 10%-kal haladja meg az összes utasításhoz szükséges idő várható értékét?

(1 pont) Az egyes utasításokhoz szükséges idők: $X_1, X_2, \ldots, X_{50}$.
(1 pont) Ha $m = \mathbb{E}(X_1)$ és $s = \mathbb{D}(X_1)$ akkor $m = 2\cdot s$
(2 pont) $\mathbb{E}(\sum_{i=1}^{50} X_i) = \sum_{i=1}^{50}\mathbb{E}(X_i) = 50m$
(2 pont) $\mathbb{P}(\sum_{i=1}^{50} X_i > 1{,}1\cdot 50m) = ?$ (Ha a kérdésben még $\mathbb{E}(\sum_{i=1}^{50} X_i)$ szerepel, akkor is jár a pont.)
(1+1 pont) $= \mathbb{P}(\sum_{i=1}^{50} X_i - 50m > 5m) = \mathbb{P}\left(\frac{\sum_{i=1}^{50} X_i - 50m}{s\sqrt{50}} > \frac{5m}{s\sqrt{50}}\right)$
(1+1 pont) $= \mathbb{P}\left(\frac{\sum_{i=1}^{50} X_i - 50m}{s\sqrt{50}} > \frac{10s}{s\sqrt{50}}\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{50} X_i - 50m}{s\sqrt{50}} > \sqrt{2}\right)$
(1 pont) $= 1 - \mathbb{P}\left(\frac{\sum_{i=1}^{50} X_i - 50m}{s\sqrt{50}} \le \sqrt{2}\right)$
(2 pont) A centrális határeloszlás-tétel miatt
(3 pont) $\frac{\sum_{i=1}^{50} X_i - 50m}{s\sqrt{50}}$ közelítőleg standard normális eloszlású.
(2 pont) Ezért a keresett valószínűség $\approx 1 - \Phi(\sqrt{2}) = 1 - \Phi(1{,}4142)$
(2 pont) $\approx 1 - 0{,}9207 = \underline{\underline{0{,}0793}}$

4. Legyen (X,Y) olyan valószínűségi vektorváltozó, aminek sűrűségfüggvénye $f_{X,Y}(x,y) = e^{-y}$ ha $y \ge x \ge 0$, és 0 egyébként. Határozzuk meg az $\mathbb{E}(Y|X)$ regressziót.

(2 pont) $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$
(2 pont) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$ (Ha ez a két formula csak együtt szerepel, akkor is jár mindkettő pontja. Ha a határok x és ∞, akkor is jár a pont.)
(2 pont) $f_X(x) = \int_x^{\infty} e^{-y}\mathrm{d}y = \left[-e^{-y}\right]_x^{\infty} = 0 - (-e^{-x}) = e^{-x}$
(1 pont) ha $x > 0$ és 0 egyébként
(2 pont) $f_{Y|X}(y|x) = \frac{e^{-y}}{e^{-x}} = e^{x-y}$
(1 pont) ha $y \ge x > 0$ és 0 egyébként
(2 pont) $\mathbb{E}(Y \mid X)$ meghatározásához az $\mathbb{E}(Y \mid X = x)$ függvényre van szükségünk. (Ha ezt a következtetés csak implicit van használva, akkor is jár a pont.)

(2 pont) $\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|x)\mathrm{d}y$
(2 pont) $= \int_{x}^{\infty} y \cdot e^{x-y}\mathrm{d}y$
(2 pont)$= e^x \Big([y \cdot (-e^{-y})]_x^{\infty} - \int_x^{\infty}(-e^{-y})\mathrm{d}y\Big) = e^x\left(0 - (-xe^{-x})\right) - e^x\left[e^{-y}\right]_x^{\infty} = x + 1$
(2 pont) ezért $\mathbb{E}(Y \mid X) = \underline{\underline{X + 1}}$

5. Válasszunk egyenletesen véletlenszerűen egy X számot a $\left[0; \frac{1}{3}\right]$ intervallumból. Ezután készítünk egy dobókockát, ami éppen úgy szabálytalan, hogy mindegyik páros számot X és mindegyik páratlan számot $\frac{1}{3} - X$ eséllyel adja eredményül. Ezzel a kockával dobunk kétszer.
a) Az $\{X = x\}$ feltétel esetén mekkora a valószínűsége, hogy két egyformát dobunk ($0 < x < \frac{1}{3}$)?
b) Nem tudván X értékét, mi a valószínűsége, hogy két egyformát dobunk?

(1 pont) Y_1: első dobás, Y_2: második dobás, a $\{X = x\}$ feltétel esetén. (Ha nincs jelölés bevezetve a dobásokra, de a használt események értelmesek, szintén jár a pont.)
(2 pont) $\mathbb{P}(Y_1 = Y_2) = \sum_{k=1}^{6} \mathbb{P}(Y_1 = k, Y_2 = k)$
(2 pont) $= \sum_{k=1}^{6} \mathbb{P}(Y_1 = k)\,\mathbb{P}(Y_2 = k)$, mert függetlenek.
(2 pont) $= \left(\frac{1}{3} - x\right)^2 + x^2 + \left(\frac{1}{3} - x\right)^2 + x^2 + \left(\frac{1}{3} - x\right)^2 + x^2$, hiszen $\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_2 = 1) = 1 - x$, és hasonlóan $k = 2, 3, 4, 5, 6$ esetén.
(1 pont) Tehát $\mathbb{P}(Y_1 = Y_2) = \underline{\underline{3x^2 + 3\left(\frac{1}{3} - x\right)^2}} = 6x^2 - 2x + \frac{1}{3}$
Ha valószínűségek helyett feltételes valószínűségekkel szerepel a számolás (a feltétel $\{X = x\}$), és konzekvensen van használva a feltétel, szintén jár a teljes pont. Nem konzekvens használat esetén az Y_1, Y_2 definíciójára nem adható pont.
(1 pont) Jelölje A azt az eseményt, hogy a két dobás megegyezik. (Ha erre nincs jelölés bevezetve, de a használt események értelmesek, szintén jár a pont.) A kérdés: $\mathbb{P}(A) =?$
(1 pont) $\mathbb{P}(A|X = x) = 3x^2 + 3\left(\frac{1}{3} - x\right)^2$
(1 pont) A teljes valószínűség tétele szerint
(3 pont) $\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|X = x) f_X(x)\mathrm{d}x$
(2 pont) $f_X(x) = 3$ ha $0 < x < \frac{1}{3}$ és 0 egyébként
(2 pont) $= \int_0^{\frac{1}{3}} \left(3x^2 + 3\left(\frac{1}{3} - x\right)^2\right) \cdot 3\,\mathrm{d}x$
(1 pont) $= \left[x^3 - \left(\frac{1}{3} - x\right)^3\right]_0^{\frac{1}{3}} \cdot 3 = \left(\left(\frac{1}{27} - 0\right) - \left(0 - \frac{1}{27}\right)\right) \cdot 3$
(1 pont) $= \frac{2}{9} \approx \underline{\underline{0{,}2222}}$

6.* Legyen (X, Y) folytonos valószínűségi vektorváltozó, aminek sűrűségfüggvénye

$$f_{X,Y}(x, y) = \frac{1}{8\pi} e^{-\frac{1}{8}\left((x-3)^2 + (y-3)^2\right)} \qquad (x, y \in \mathbb{R}).$$

Határozzuk meg $\mathbb{D}^2(XY)$-t.
(2 pont) (X, Y) kétdimenziós normális, jelölés: $(X, Y) \sim N(\underline{\mu}, \underline{\underline{\Sigma}})$
(1 pont) ahol $\underline{\mu} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ és
(2 pont) $\underline{\underline{\Sigma}} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$
(2 pont) hiszen $\det(\underline{\underline{\Sigma}}) = 4 \cdot 4$ és $\underline{\underline{\Sigma}}^{-1} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$
(3 pont) X és Y függetlenek, mivel $\mathrm{cov}(X, Y) = 0$
AVAGY (a fenti pontok helyett):
(4 pont) $f_{X,Y}(x, y) = \frac{1}{8\pi} e^{-\frac{1}{8}\left((x-3)^2+(y-3)^2\right)} = \frac{1}{\sqrt{2\pi}\sqrt{4}} e^{-\frac{1}{8}(x-3)^2} \cdot \frac{1}{\sqrt{2\pi}\sqrt{4}} e^{-\frac{1}{8}(y-3)^2}, \forall x, y \in \mathbb{R}$,
(3 pont) ezért $X, Y \sim N(3; 4)$,
(3 pont) továbbá X és Y függetlenek, mert $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y), \forall x, y \in \mathbb{R}$

(2 pont) $\mathbb{D}^2(XY) = \mathbb{E}(X^2Y^2) - E(XY)^2$
(1 pont) Függetlenség miatt:
(2 pont) $= \mathbb{E}(X^2)\,\mathbb{E}(Y^2) - \mathbb{E}(X)^2\,\mathbb{E}(Y)^2$
(2 pont) $\mathbb{E}(X^2) = \mathbb{D}^2(X) + \mathbb{E}(X)^2$
(1 pont) $= 4 + 3^2 = 13$,
(1 pont) hasonlóan $\mathbb{E}(Y^2) = 13$
(1 pont) Tehát $\mathbb{D}^2(XY) = 13 \cdot 13 - 3^2 \cdot 3^2 = \underline{\underline{88}}$