Vizsgadolgozat

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.

Aritmetikai 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. Írjuk fel az alábbi definíciót, illetve állítást:

   (a) Hogyan definiáljuk egy egyszerű valószínűségi változó várható értékét?

   (b) Milyen feltétel esetén, és hogyan fejezhető ki az X és Y valószínűségi változók szorzatának várható értéke $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ segítségével, az előadáson elhangzott állítás szerint?

   (10 pont) Egy egyszerű valószínűségi változó várható értéke:

   $$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{k \in \mathrm{Ran}(X)} k \cdot \mathbb{P}(X = k)$$

   (Ha a szumma indexhalmaza helytelen, de ettől eltekintve a formula helyes, akkor 7 pont.)
   (jegyzet: 3.2.1)
   (4 pont) Ha X és Y független valószínűségi változók, és
   (2 pont) $\mathbb{E}(XY)$, $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik,
   (4 pont) akkor

   $$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

   (jegyzet: 6.1.4) (A feltételek felírásáért akkor jár a pont, ha az állítás is szerepel.)

2. Egy képzeletbeli szervezet három ügynöke épp lehallgatja Xavért és Yvettet. Kikapcsolódásképp mindhárman tippelnek: vajon hány kávét iszik aznap Xavér (jelölés: X), illetve Yvett (jelölés: Y). A tippek a következők: 'A' ügynök szerint $\{X \leq 2\}$, 'B' ügynök szerint $\{Y \leq 2\}$, továbbá 'C' ügynök szerint $\{X \leq 3, Y \leq 3\}$. Tegyük fel, hogy X és Y független, örökifjú eloszlású, nem-konstans valószínűségi változók a pozitív egész számok halmazán, és $\mathbb{E}(X) = \mathbb{E}(Y) = 2$. Mi a valószínűsége, hogy a három tipp közül legalább egy helyes?

(0 pont) Jelölje A, B és C a három szövegben szereplő $\{X \leq 2\}$, $\{Y \leq 2\}$ és $\{X \leq 3, Y \leq 3\}$ eseményeket.
(1 pont) $\mathbb{P}(A \cup B \cup C) = ?$
(2 pont) X és Y geometriai eloszlású
(1 pont) a várható értékekből adódóan $X, Y \sim \text{Geo}\left(\frac{1}{2}\right)$
(2 pont) $\mathbb{P}(A) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$ és ugyanígy $\mathbb{P}(B) = \frac{3}{4}$
(2 pont) Függetlenség miatt: $\mathbb{P}(X \leq 3, Y \leq 3) = \mathbb{P}(X \leq 3)\,\mathbb{P}(Y \leq 3)$
(1 pont) $\mathbb{P}(X \leq 3) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$ és hasonlóan $\mathbb{P}(Y \leq 3) = \frac{7}{8}$
(1 pont) Tehát $\mathbb{P}(C) = \mathbb{P}(X \leq 3, Y \leq 3) = \frac{7}{8} \cdot \frac{7}{8} = \frac{49}{64}$
(1 pont) $\mathbb{P}(A \cap B) = \mathbb{P}(X \leq 2) \cdot \mathbb{P}(Y \leq 2) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$
(1 pont) $\mathbb{P}(A \cap C) = \mathbb{P}(X \leq 2, Y \leq 3) = \mathbb{P}(X \leq 2)\,\mathbb{P}(Y \leq 3) = \frac{3}{4} \cdot \frac{7}{8} = \frac{21}{32}$ és hasonlóan $\mathbb{P}(B \cap C) = \frac{21}{32}$
(2 pont) $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(X \leq 2, X \leq 2) = \mathbb{P}(A \cap B) = \frac{9}{16}$
(Ha a fenti egyenlet a számszerű érték nélkül szerepel, akkor is jár a pont.)
(1 pont) Poincaré-formula alapján:
(3 pont) $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$
(1 pont) Behelyettesítve:

$$\mathbb{P}(A \cup B \cup C) = \frac{3}{4} + \frac{3}{4} + \frac{49}{64} - \frac{9}{16} - \frac{21}{32} - \frac{21}{32} + \frac{9}{16}$$

(1 pont) $= \frac{61}{64} \approx 0{,}9531$
(Ha csak hányados, vagy csak tizedestört alakban van megadva a megoldás, akkor is jár a pont.)

3. Egy nyári táborban szörpivó versenyt rendeznek. A piros csapat összesen 138 korsónyi szörpöt ivott meg. A győzelemhez a kék csapatnak ezt kellene túlteljesítenie. A kék csapatnak 36 tagja van. A csapattagok azonos eloszlású véletlen mennyiségeket isznak, egymástól függetlenül, egyenként átlagosan 4,2 korsónyit, 2 korsónyi szórással.
   (a) Mi a valószínűsége, hogy a kék csapat kikap, azaz összesen kevesebb, mint 138 korsónyit isznak?
   (b) Mekkora kellene legyen 4,2 helyett az átlagos ivókapacitása egy csapattagnak, hogy az a) feladatban kiszámolt valószínűség a felére csökkenjen (azonos szórás mellett)?

(0 pont) Jelölje $X_1, \ldots X_{36}$ az egyes csapattagok fogyasztását.
(1 pont) $\mathbb{P}(\sum_{i=1}^{36} X_i < 138) = ?$
(4 pont) Sztenderdizálunk:

$$\mathbb{P}\left(\frac{\sum_{i=1}^{36} X_i - 36 \cdot 4{,}2}{\sqrt{36} \cdot 2} < \frac{138 - 36 \cdot 4{,}2}{\sqrt{36} \cdot 2}\right)$$

(0 pont) Mivel X_i-k egymástól független, azonos eloszlású val. változók, ezért
(2 pont) a centrális határeloszlás-tétel miatt
(3 pont) $\frac{\sum_{i=1}^{36} X_i - 36 \cdot 4{,}2}{\sqrt{36} \cdot 2}$ közelítőleg standard normális eloszlású.
(2 pont) Tehát a keresett mennyiség: $\Phi\left(\frac{138 - 36 \cdot 4{,}2}{\sqrt{36} \cdot 2}\right) = \Phi(-1{,}1)$
(1 pont) $= 1 - \Phi(1{,}1)$
(1 pont) $= 1 - 0{,}8643 = \underline{\underline{0{,}1357}}$
(0 pont) A b) feladat megoldásához legyen a fenti helyett $\mathbb{E}(X_1) = m$ ismeretlen, és
(1 pont) $\mathbb{P}(\sum_{i=1}^{36} X_i < 138) = 0{,}1357 \cdot \frac{1}{2} = 0{,}06785$
(1 pont) A korábbi gondolatmenetet felhasználva, a fenti egyenlet bal oldala a következőképp írható:

$$\mathbb{P}\left(\frac{\sum_{i=1}^{36} X_i - 36 \cdot m}{\sqrt{36} \cdot 2} < \frac{138 - 36 \cdot m}{\sqrt{36} \cdot 2}\right)$$

(1 pont) Tehát $\Phi\left(\frac{138 - 36m}{12}\right) = 0{,}06785$
(1 pont) Átrendezve: $1 - \Phi\left(\frac{138 - 36m}{12}\right) = 1 - 0{,}06785$ ahol a jobb oldal $= 0{,}93215$ míg a bal oldal $\Phi\left((-1) \cdot \frac{138 - 36m}{12}\right)$

(1 pont) Tehát $(-1)\cdot\frac{138-36m}{12} = \Phi^{-1}(0{,}93215) \approx 1,49$.
(1 pont) Átrendezve: $m \approx \underline{\underline{4{,}33}}$.

4. Legyen $\lambda > 0$ valós szám, $X \sim \mathrm{Pois}(\lambda)$ és $Y \sim \mathrm{Exp}\left(\frac{1}{\sqrt{\lambda}}\right)$. Tegyük fel, hogy $\mathrm{corr}(X, 2Y) = \lambda$, továbbá Y-nak az X-re vett lineáris regressziója $0{,}01\cdot X + c$ alakú, valamilyen c valós számra. Határozzuk meg λ és c értékét.

(2 pont) $\mathrm{corr}(X, 2Y) = \frac{\mathrm{cov}(X,2Y)}{\mathbb{D}(X)\,\mathbb{D}(2Y)}$
(Ha más változókkal, pl X és Y szerepel helyesen a korreláció definíciójak, szintén jár a pont.)
(1 pont) $\mathrm{cov}(X, 2Y) = 2\mathrm{cov}(X, Y)$
(1 pont) $\mathbb{D}(2Y) = 2\,\mathbb{D}(Y)$
(1 pont) $\mathbb{D}(X) = \sqrt{\lambda}$, (mert $X \sim \mathrm{Pois}(\lambda)$),
(A Poisson eloszlásra történő explicit hivatkozás nélkül is jár a pont.)
(1 pont) $\mathbb{D}(Y) = \left(\frac{1}{\sqrt{\lambda}}\right)^{-1} = \sqrt{\lambda}$, (mert $Y \sim \mathrm{Exp}\left(\frac{1}{\sqrt{\lambda}}\right)$)
(Az exponenciális eloszlásra történő explicit hivatkozás nélkül is jár a pont.)
(1 pont) A fentiekből tehát $\lambda = \mathrm{corr}(X, 2Y) = \frac{\mathrm{cov}(X,Y)}{\mathbb{D}(X)\cdot\mathbb{D}(Y)}$
(1 pont) $= \frac{\mathrm{cov}(X,Y)}{\sqrt{\lambda}\cdot\sqrt{\lambda}}$
(1 pont) Átrendezve $\mathrm{cov}(X, Y) = \lambda\cdot\sqrt{\lambda}\cdot\sqrt{\lambda} = \lambda^2$
(1 pont) lineáris regresszió: $\beta\cdot X + \alpha$, ahol
(2 pont) $\beta = \frac{\mathrm{cov}(X,Y)}{\mathbb{D}^2(X)}$
(2 pont) $\alpha = \mathbb{E}(Y) - \beta\cdot\mathbb{E}(X) = \mathbb{E}(Y) - \frac{\mathrm{cov}(X,Y)}{\mathbb{D}^2(X)}\,\mathbb{E}(X)$
(1 pont) $\beta = \frac{\lambda^2}{\lambda} = \lambda$
(1 pont) $\beta = 0{,}01$ (Megj.: a lineáris regresszió ha létezik, akkor egyértelmű)
(1 pont) $\Rightarrow \lambda = \underline{\underline{0{,}01}}$
(1 pont) $\mathbb{E}(X) = \lambda$ és $\mathbb{E}(Y) = \sqrt{\lambda}$
(1 pont) $c = \alpha = \sqrt{\lambda} - \lambda\cdot\lambda = \sqrt{0{,}01} - 0{,}01\cdot 0{,}01$
(1 pont) $= \underline{\underline{0{,}0999}}$

5. Legyen (X, Y) folytonos valószínűségi vektorváltozó, aminek együttes sűrűségfüggvénye:

$$f_{X,Y}(x, y) = \begin{cases} x + y & \text{ha } 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{egyébként.} \end{cases}$$

Határozzuk meg X sűrűségfüggvényét, az $\mathbb{E}(Y \mid X)$ regressziót, illetve az $\mathbb{E}(Y)$ várható értéket.

(2 pont) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y$
(Ha a fenti egyenlet implicit jelenik meg az f_X kiszámolásában, akkor is jár a pont.)
(1 pont) $= \int_0^1 (x + y)\,\mathrm{d}y$
(1 pont) $= [xy + \frac{y^2}{2}]_{y=0}^1 = x + \frac{1}{2}$
(1 pont) ha $0 < x < 1$, különben 0.
(3 pont) $\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y\cdot f_{Y|X}(y \mid x)\,\mathrm{d}y$
(2 pont) $f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$ (ahol $f_X(x) \neq 0$)
(1 pont) $= \frac{x+y}{x+\frac{1}{2}}$
(1 pont) Tehát $\mathbb{E}(Y \mid X = x) = \int_0^1 y\cdot\frac{x+y}{x+\frac{1}{2}}\,\mathrm{d}y$
(2 pont) $= \frac{1}{x+\frac{1}{2}}\int_0^1 (xy + y^2)\,\mathrm{d}y = \frac{1}{x+\frac{1}{2}}\left[x\frac{y^2}{2} + \frac{y^3}{3}\right]_{y=0}^1 = \frac{\frac{1}{2}x+\frac{1}{3}}{x+\frac{1}{2}}$, (ha $0 < x < 1$ és 0 egyébként)
(1 pont) A val. változó X-et visszahelyettesítve: $\mathbb{E}(Y \mid X) = \frac{\frac{1}{2}X+\frac{1}{3}}{X+\frac{1}{2}} = \frac{3X+2}{6X+3}$

1. megoldás $\mathbb{E}(Y)$-ra:
(2 pont) $\mathbb{E}(Y) = \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x)\,\mathrm{d}x$
(Ha a fenti egyenlet implicit jelenik meg az f_X kiszámolásában, akkor is jár a pont.)
(1 pont) $= \int_0^1 \frac{\frac{1}{2}x+\frac{1}{3}}{x+\frac{1}{2}}(x + \frac{1}{2})\,\mathrm{d}x$

(1 pont) $= \int_0^1 \left(\frac{1}{2}x + \frac{1}{3}\right) \mathrm{d}x = \left[\frac{x^2}{4} + \frac{x}{3}\right]_0^1 = \frac{1}{4} + \frac{1}{3}$
(1 pont) $= \frac{7}{12} \approx 0{,}5833$

2. megoldás $\mathbb{E}(Y)$-ra:
(2 pont) $\mathbb{E}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \cdot f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$
(Ha a fenti egyenlet implicit jelenik meg az f_X kiszámolásában, akkor is jár a pont.)
(1 pont) $= \int_0^1\int_0^1 y(x+y)\,\mathrm{d}x\,\mathrm{d}y$
(1 pont) $= \int_0^1 \left[y\frac{x^2}{2} + xy^2\right]_{x=0}^1 \mathrm{d}y = \int_0^1 (\frac{1}{2}y + y^2)\,\mathrm{d}y = \left[\frac{y^2}{4} + \frac{y^3}{3}\right]_0^1 = \frac{1}{4} + \frac{1}{3}$
(1 pont) $= \frac{7}{12} \approx 0{,}5833$
(Ha csak hányados, vagy csak tizedestört alakban van megadva a megoldás, akkor is jár a pont.)

6.* Legyen $X \sim N(0,1)$ és legyen a tőle független U valószínűségi változó értéke $\frac{1}{2}$ valószínűséggel $+1$ és $\frac{1}{2}$ valószínűséggel -1. Definiáljuk az $Y = X \cdot U$ valószínűségi változót.

(a) Igaz-e, hogy $Y \sim N(0,1)$? (Tipp: határozzuk meg a $\mathbb{P}(Y < y)$ valószínűségeket minden $y \in \mathbb{R}$-re, az $\{U = 1\}$ és $\{U = -1\}$ teljes eseményrendszer felhasználásával.)
(b) Igaz-e, hogy (X, Y) kétdimenziós normális eloszlású?

(0 pont) $\{U = 1\}$ és $\{U = -1\}$ teljes eseményrendszer
(1 pont) így a teljes valószínűség tétele szerint:
(2 pont) $\mathbb{P}(Y < y) = \mathbb{P}(Y < y \mid U = 1) \cdot \mathbb{P}(U = 1) + \mathbb{P}(Y < y \mid U = -1) \cdot \mathbb{P}(U = -1)$
(2 pont) $\mathbb{P}(Y < y \mid U = 1) = \mathbb{P}(X \cdot 1 < y \mid U = 1)$
(2 pont) $= \mathbb{P}(X < y)$ (hiszen X és U függetlenek)
(1 pont) Hasonlóan, $\mathbb{P}(Y < y \mid U = -1) = \mathbb{P}(X \cdot (-1) < y \mid U = -1)$
(1 pont) $= \mathbb{P}(X > -y \mid U = -1) = 1 - \mathbb{P}(X < -y \mid U = -1) = 1 - \mathbb{P}(X < -y)$
(Ha $\mathbb{P}(Y < z \mid U = -1)$ van meghatározva, de $\mathbb{P}(Y < z \mid U = 1)$ nincs, akkor az $U = 1$ esethez tartozó magasabb pontok adandók a lépésekre.)
(1 pont) Behelyettesítve: $\mathbb{P}(Y < y) = \mathbb{P}(X < y) \cdot \mathbb{P}(U = 1) + \big(1 - \mathbb{P}(X < -y)\big) \cdot \mathbb{P}(U = -1)$
(1 pont) ahol $\mathbb{P}(X < y) = \Phi(y)$ és $\mathbb{P}(X < -y) = \Phi(-y)$ hiszen $X \sim N(0,1)$
(1 pont) $\Phi(-y) = 1 - \Phi(y)$ (hiszen a standard normális eloszlás szimmetrikus)
(1 pont) Tehát $\mathbb{P}(Y < y) = \Phi(y)\frac{1}{2} + \big(1 - (1 - \Phi(y))\big)\frac{1}{2}$
(1 pont) $= \Phi(y)$, vagyis $Y \sim N(0,1)$ igaz
(2 pont) Kiszámolható, hogy $\mathrm{cov}(X, Y) = 0$, hiszen

$$\mathbb{E}(X^2U) = \mathbb{E}(X^2 \cdot (+1) \mid U = 1)\frac{1}{2} + \mathbb{E}(X^2 \cdot (-1) \mid U = -1)\frac{1}{2} = \frac{1}{2}\,\mathbb{E}(X^2) - \frac{1}{2}\,\mathbb{E}(X^2) = 0$$

az előző gondolatmenet szerint, továbbá $\mathbb{E}(X) = \mathbb{E}(Y) = 0$.
(1 pont) Tehát ha (X, Y) kétdimenziós normális eloszlású, akkor kétdimenziós standard normális kell legyen.
(2 pont) Ugyanakkor $\mathbb{P}(|X| = |Y|) = 1$, mivel $|Y| = |XU| = |X| \cdot |U| = |X| \cdot 1$
(1 pont) Ami ellentmond annak, hogy kétdimenziós standard normális eloszlású lenne (X, Y), hiszen az $\{(x,y) \in \mathbb{R}^2 \mid |x| = |y|\}$ halmaz nulla területű a síkon, így 0 valószínűséggel esik ebbe a halmazba az (X, Y) értéke.