Vizsgadolgozat, 2020. január 29.
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. Hókuszpók egy törpcsapdával 1 és 2 cm közötti törpöt szeretne fogni, hogy azt fogyaszthassa koktélparadicsom helyett a reggeli mellé. A csapdával egyszerre csak egy törpöt tud elkapni. A kifogott törp magasságának (cm-ben vett) sűrűségfüggvénye: $f(x) = \frac{x}{4}$, ha $1 \leq x \leq 3$, és 0 egyébként. Jelölje A, B és C rendre azon eseményeket, hogy az első, második illetve harmadik alkalommal kifogott törp megfelelő méretű. Tegyük fel, hogy ezek az események együttesen függetlenek. Határozzuk meg $\mathbb{P}(A \cup B \cup C)$-t.

   (0 pont) Jelölje X az egy alkalommal kifogott törp magasságát.
   (2 pont) $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \mathbb{P}(1 < X < 2)$ (a $\mathbb{P}(X < 2)$, illetve $<$ helyett $\leq$ esetén szintén jár a teljes pont)
   (3 pont) $\mathbb{P}(1 < X < 2) = \int_1^2 f_X(x)\mathrm{d}x$ (Elírással nem fogadható el; kivétel, ha határokként a és b szerepel 1 és 2 helyett, avagy ha $\mathbb{P}(X < 2)$ van felírva.)
   (2 pont) $= \int_1^2 \frac{x}{4}\mathrm{d}x = \left[\frac{x^2}{8}\right]_1^2 = \frac{2^2}{8} - \frac{1^2}{8} = \frac{3}{8}$

   I.megoldás:
   (2 pont) $\mathbb{P}(A \cup B \cup C) = 1 - \mathbb{P}(\overline{A \cup B \cup C})$
   (1 pont) de Morgan-azonosság alapján:
   (3 pont) $= 1 - \mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C})$
   (1 pont) A, B, C együttesen független események, így $\overline{A}, \overline{B}, \overline{C}$ is együttesen független, ezért
   (3 pont) $= 1 - \mathbb{P}(\overline{A})\,\mathbb{P}(\overline{B})\,\mathbb{P}(\overline{C})$
   (1 pont) $\mathbb{P}(\overline{A}) = \mathbb{P}(\overline{B}) = \mathbb{P}(\overline{C}) = 1 - \frac{3}{8} = \frac{5}{8}$
   (2 pont) $\mathbb{P}(A \cup B \cup C) = 1 - \left(\frac{5}{8}\right)^3 = 1 - \frac{125}{512} = \frac{387}{512} \approx \underline{\underline{0{,}7559}}$

   II.megoldás:
   (2 pont) Szita-formula alapján:
   (4 pont) $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$

(2 pont) Mivel A, B, C együttesen független, ezért
(3 pont) $= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A)\,\mathbb{P}(B) - \mathbb{P}(A)\,\mathbb{P}(C) - \mathbb{P}(B)\,\mathbb{P}(C) + \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$
(2 pont) $= 3 \cdot \frac{3}{8} - 3 \cdot \left(\frac{3}{8}\right)^2 + \left(\frac{3}{8}\right)^3 = \frac{387}{512} \approx \underline{\underline{0{,}7559}}$

2. Legyen $c > 0$ valós szám, és X olyan folytonos valószínűségi változó, aminek eloszlásfüggvénye

$$F_X : x \mapsto \begin{cases} 0 & \text{ha } x \le c, \\ \ln(x) - \ln(c) & \text{ha } c < x \le ce, \\ 1 & \text{ha } ce < x, \end{cases}$$

ahol e nem ismeretlen, hanem az Euler-féle konstans, $\approx 2{,}718$.
a) Tegyük fel, hogy $\mathbb{E}(X) = e - 1$. Határozzuk meg c-t.
b) Határozzuk meg X szórásnégyzetét.

(2 pont) $f_X(x) = F'_X(x)$ (szövegesen is jár a teljes pont), tehát
(2 pont) $f_X(x) = \frac{1}{x}$, ha $c < x < ce$,
(1 pont) és 0 egyébként
(2 pont) $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x$ (Ha a határok c és ce, szintén jár a teljes pont.)
(2 pont) $= \int_c^{ce} x\frac{1}{x}\mathrm{d}x = \int_c^{ce} 1\mathrm{d}x = [1]_c^{ce} = c(e-1)$
(1 pont) $\Rightarrow c = \underline{\underline{1}}$
(2 pont) $\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$
(3 pont) $\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\,\mathrm{d}x$
(2 pont) $= \int_1^e x^2\frac{1}{x}\,\mathrm{d}x = \int_1^e x\,\mathrm{d}x = \left[\frac{x^2}{2}\right]_1^e = \frac{e^2-1}{2}$
(2 pont) $\Rightarrow \mathbb{D}^2(X) = \frac{e^2-1}{2} - (e-1)^2$
(1 pont) $= \frac{1}{2}(e-1)(3-e) \approx \underline{\underline{0{,}2420}}$

3. Babett egy cégnél IT hardware support feladatot lát el. Tegyük fel, hogy minden nap 18 problémát old meg. Egy probléma lényeges, ha nem oldható meg az adott készülék újraindításával. Az egyes problémák egymástól függetlenül, $\frac{1}{3}$ valószínűséggel lényegesek.
a) Milyen eloszlású az 1 nap alatt megoldott lényeges problémák száma?
b) Közelítőleg mi a valószínűsége, hogy 100 nap alatt 580-nál több, de 620-nál kevesebb lényeges problémát kell megoldania? (Az egyes napokon fellépő lényeges problémák darabszámai egymástól függetlenek.)

(1 pont) Jelölje X az egy nap alatt fellépő lényeges problémák számát.
(2 pont) X binomiális,
(1 pont) paraméterei: $X \sim B(18, \frac{1}{3})$
(1 pont) Jelölje $X_1, \ldots, X_{100}$ az egyes napokon fellépő lényeges problémák számát.
(1 pont) $\mathbb{E}(X_i) = 18 \cdot \frac{1}{3} = 6$
(1 pont) $\mathbb{D}^2(X_i) = 18 \cdot \frac{1}{3} \cdot \frac{2}{3} = 4$ azaz $\mathbb{D}(X_i) = 2$
(1 pont) $\mathbb{P}(580 < \sum_{i=1}^{100} X_i < 620) = ?$
(2 pont) $= \mathbb{P}(\sum_{i=1}^{100} X_i < 620) - \mathbb{P}(\sum_{i=1}^{100} X_i \le 580)$
(1 pont) $= \mathbb{P}(\sum_{i=1}^{100} X_i - 100 \cdot 6 < 20) - \mathbb{P}(\sum_{i=1}^{100} X_i - 100 \cdot 6 \le -20)$
(1 pont) $= \mathbb{P}\left(\frac{\sum_{i=1}^{100} X_i - 600}{\sqrt{100} \cdot 2} < 1\right) - \mathbb{P}\left(\frac{\sum_{i=1}^{100} X_i - 600}{\sqrt{100} \cdot 2} \le -1\right)$
(2 pont) A centrális határeloszlás-tétel miatt
(2 pont) $\frac{\sum_{i=1}^{100} X_i - 600}{20}$ közelítőleg standard normális eloszlású. (Ha az a) feladat megoldása nélkül van levezetve az előző 5 pontozási sor, azokra akkor is jár a teljes pont.)
(1 pont) Ezért a keresett valószínűség: $\approx \Phi(1) - \Phi(-1)$
(2 pont) $= 2\Phi(1) - 1 \approx 2 \cdot 0{,}8413 - 1$
(1 pont) $= \underline{\underline{0{,}6826}}$

4. Legyen $X \sim \mathrm{Exp}(1)$ valószínűségi változó. Az X kiértékelése után választunk egy Y számot a $[0; e^X]$ intervallumon egyenletesen.
a) Határozzuk meg a $\mathbb{P}(Y < X \mid X = x)$ feltételes valószínűséget.
b) Határozzuk meg a $\mathbb{P}(Y < X)$ valószínűséget.

(2 pont) Jelölje Z az Y értékét az $\{X = x\}$ feltétel esetén. $Z \sim \mathrm{U}(0; e^x)$
(2 pont) Az $\mathrm{U}(a; b)$ egyenletes eloszlás eloszlásfüggvénye $\mapsto \frac{z}{b-a}$ ha $a < z \le b$, 0, ha $z \le a$ és 1, ha $z > b$, ezért Z eloszlásfüggvénye: $F_Z(z) = \frac{1}{e^x - 0} z$, ha $0 < x \le e^x$, illetve 0, ha $x < 0$. (Ha a 0, 1 esetek teljesen lemaradnak, legfeljebb 1 pont. Akkor is adható teljes pontszám, ha nincs megvizsgálva mikor teljesül az $x \le e^x$ egyenlőtlenség.)
(3 pont) $\mathbb{P}(Y < X \mid X = x) = \mathbb{P}(Z < x) = F_Z(x) = \frac{x}{e^x}$
(2 pont) A teljes valószínűség tétele szerint
(3 pont) $\mathbb{P}(Y < X) = \int_{-\infty}^{\infty} \mathbb{P}(Y < X | X = x) f_X(x)\, \mathrm{d}x$
(2 pont) $= \int_0^\infty \frac{x}{e^x} \cdot e^{-x}\, \mathrm{d}x = \frac{1}{2} \int_0^\infty x 2e^{-2x}\, \mathrm{d}x$
(4 pont) $= \frac{1}{2}\,\mathbb{E}(W)$, ahol $W \sim \mathrm{Exp}(2)$, tehát $\mathbb{E}(W) = \frac{1}{2}$.
Avagy parciálisan integrálva $= \left[\frac{1}{2}x(-e^{-2x})\right]_0^\infty - \int_0^\infty \frac{1}{2}(-e^{-2x})\mathrm{d} = 0 - \left[\frac{1}{4}e^{-2x}\right]_0^\infty = \frac{1}{4}$
(2 pont) Ezért $P(Y < X) = \frac{1}{4} = \underline{\underline{0{,}25}}$

5. Legyen $(X, Y) \sim N\left(\underline{0}, \underline{\underline{\Sigma}}\right)$ kétdimenziós normális eloszlású valószínűségi vektorváltozó, ahol

$$\underline{\underline{\Sigma}} = \begin{bmatrix} 9 & 1 \\ 1 & 4 \end{bmatrix}.$$

a) Milyen $c > 0$ valós számot válasszunk, hogy $\mathrm{corr}(X + cY, X - 2cY) = 0$ teljesüljön?
b) Tegyük fel, hogy c teljesíti az a) feladat feltételét. Független lesz-e ekkor $X + cY$ és $X - 2cY$? Indokoljunk!

(2 pont) $\mathrm{corr}(X + cY, X - 2cY) = 0 \Leftrightarrow \mathrm{cov}(X + cY, X - 2cY) = 0$
(2 pont) $\mathrm{cov}(X + cY, X - 2cY) = \mathrm{cov}(X, X) + \mathrm{cov}(X, -2cY) + \mathrm{cov}(cY, X) + \mathrm{cov}(cY, -2cY)$
(2 pont) $= \mathrm{cov}(X, X) - c \cdot \mathrm{cov}(X, Y) - 2c^2 \mathrm{cov}(Y, Y)$ (Ha $\mathrm{cov}(X, X)$ helyett $\mathbb{D}^2(X)$-el történik a számolás, szintén jár a teljes pont.)
(2 pont) $\mathrm{cov}(X, X) = 9$, $\mathrm{cov}(Y, Y) = 4, \mathrm{cov}(X, Y) = 1$
(1 pont) $\Rightarrow 9 - c - 8c^2 = 0 \Rightarrow c_1 = 1, c_2 = -\frac{9}{8}$,
(1 pont) $c > 0$, ezért $c = \underline{\underline{1}}$
(3 pont) Kétdimenziós normális eloszlású valószínűségi vektorváltozó lineáris transzformáltja szintén kétdimenziós normális eloszlású.
(3 pont) Az $(X + cY, X - 2cY)$ vektor valóban lineáris transzformáltja (X, Y)-nak, hiszen

$$\begin{bmatrix} 1 & c \\ 1 & -2c \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X + cY \\ X - 2cY \end{bmatrix}$$

(3 pont) A kétdimenziós normális eloszlás koordinátái pontosan akkor függetlenek, ha kovarianciájuk 0. (Ha az állítás korrelációval van kimondva, szintén jár a teljes pont.)
(1 pont) Tehát $X + cY$ és $X - 2cY$ függetlenek.

6. Legyenek X_1, X_2 független, $\mathrm{Exp}(1)$ eloszlású valószínűségi változók, és jelölje Y az $X_1 + X_2$ valószínűségi változót. Adjuk meg $\mathbb{E}(X_1 + 2X_2 \mid Y)$-t.

I.megoldás:
(2 pont) $\mathbb{E}(X_1 + 2X_2 \mid Y) = \mathbb{E}(X_1 \mid Y) + \mathbb{E}(2X_2 \mid Y)$
(1 pont) $= \mathbb{E}(X_1 \mid Y) + 2\,\mathbb{E}(X_2 \mid Y)$
(6 pont) $\mathbb{E}(X_1 \mid Y) = \mathbb{E}(X_2 \mid Y)$ az X_1 és X_2 valószínűségi változók szimmetrikus szerepe miatt
(2 pont) Tehát $\mathbb{E}(X_1 + 2X_2 \mid Y) = 3\,\mathbb{E}(X_1 \mid Y)$
(4 pont) $Y = \mathbb{E}(Y \mid Y)$ hiszen Y függvénye az Y-nak

(2 pont) $\mathbb{E}(Y \mid Y) = \mathbb{E}(X_1 \mid Y) + \mathbb{E}(X_2 \mid Y)$
(1 pont) $= 2\,\mathbb{E}(X_1 \mid Y)$
(1 pont) $\Rightarrow \mathbb{E}(X_1 \mid Y) = 0{,}5Y$
(1 pont) $\Rightarrow \mathbb{E}(X_1 + 2X_2 \mid Y) = 3 \cdot 0{,}5Y = \underline{\underline{1{,}5Y}}$

II.megoldás:
(2 pont) $\mathbb{E}(X_1 + 2X_2 \mid Y) = \mathbb{E}(X_1 + X_2 \mid Y) + \mathbb{E}(X_2 \mid Y)$
(4 pont) $\mathbb{E}(X_1 + X_2 \mid Y) = \mathbb{E}(Y \mid Y) = Y$
(1 pont) (X_1, X_2) együttes sűrűségfüggvénye az X_1 és X_2 függetlensége miatt:

$$f_{X_1,X_2}(u,v) = f_{X_1}(u) f_{X_2}(v) = e^{-u}e^{-v} = e^{-u-v},$$

ha $u, v > 0$, és 0 egyébként.
(3 pont) $(X_2, X_1 + X_2)$ együttes eloszlásfüggvénye:

$$\mathbb{P}(X_2 < x, X_1 + X_2 < y) = \int\int_{\{(u,v) \mid 0<u<x, 0<u+v<y\}} e^{-u-v} \mathrm{d}u\mathrm{d}v = \int_0^x \int_0^{\max(0,y-u)} e^{-u-v} \mathrm{d}v\mathrm{d}u$$

$$= \int_0^x \left[-e^{-u-v}\right]_{v=0}^{\max(0,y-u)} \mathrm{d}u = \int_0^x \left(-e^{-u-(\max(0,y-u))} + e^{-u-0}\right) \mathrm{d}u$$

$$= \int_0^{\min(x,y)} \left(-e^{-u-(y-u)} + e^{-u}\right) \mathrm{d}u + \int_{\min(x,y)}^x \left(-e^{-u-0} + e^{-u}\right) \mathrm{d}u$$

$$= \left[-ue^{-y} - e^{-u}\right]_{u=0}^{\min(x,y)} + 0 = -\min(x,y)e^{-y} - e^{-\min(x,y)} - (-0-1)$$

$$\text{ha } 0 \le x < y: \qquad = 1 - xe^{-y} - e^{-x}$$

$$\text{ha } x \ge y \ge 0: \qquad = 1 - ye^{-y} - e^{-y}$$

(2 pont) Az együttes sűrűségfüggvényhez a fentit kell mindkét változó szerint deriválnunk. Ez csak akkor lehet nem-nulla, ha $x < y$. Ezen a tartományon pedig:

$$f_{X_2,Y}(x,y) = e^{-y} \qquad \text{ha } 0 < x < y$$

(2 pont) $f_{X_2|Y}(x|y) = \frac{f_{X_2,Y}(x,y)}{f_Y(y)}$
(1 pont) $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x$ (Ha ez a két formula csak együtt szerepel, akkor is jár mindkettő pontja. Ha a határok 0 és y, akkor is jár a pont.)
(1 pont) $f_Y(y) = \int_0^y e^{-y}\mathrm{d}x = [xe^{-y}]_0^y = ye^{-y}$, ha $y > 0$ és 0 egyébként
(1 pont) $f_{X_2|Y}(x|y) = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}$, ha $0 < x < y$, és 0 egyébként.
(1 pont) $\mathbb{E}(X_2 \mid Y = y) = \int_0^y x\frac{1}{y}\mathrm{d}x = \left[\frac{x^2}{2y}\right]_0^y = \frac{1}{2}y$
(1 pont) $\mathbb{E}(X_2 \mid Y) = \frac{1}{2}Y$
(1 pont) Tehát $\mathbb{E}(X_1 + 2X_2 \mid Y) = Y + \frac{1}{2}Y = \underline{\underline{1{,}5Y}}$