Vizsgadolgozat, 2020. január 23.
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. Borbála azt a feladatot kapta a valószínűségszámítás vizsgáján, hogy számolja ki az $\mathbb{E}(XY)$ várható értéket az X és Y diszkrét valószínűségi változók együttes eloszlása alapján. Sajnos a vizsgalapjára rálottyent a kólával kevert energiaitalos kávéja, így az együttes eloszlásból csak a következő látszódik:

$Y \backslash X$	-1	0	1
3	$\frac{1}{9}$	$\frac{2}{9}$	
4	$\frac{1}{18}$	$\frac{1}{9}$	

ahol $\mathrm{Ran}(X) = \{-1, 0, 1\}$ és $\mathrm{Ran}(Y) = \{3, 4\}$. Borbála biztosan tudja (korábban leellenőrizte), hogy X és Y függetlenek. Meg tudjuk-e válaszolni a feladatot pusztán ennyi információ birtokában? Ha igen, válaszoljuk is meg; ha nem, bizonyítsuk be, hogy nem lehetséges.

(0 pont) Jelölés: $\mathbb{P}(X = 1, Y = 3) = a$, $\mathbb{P}(X = 1, Y = 4) = b$.
(2 pont) A táblázatban szereplő valószínűségek összege 1, ezért
(1 pont) $\frac{1}{9} + \frac{2}{9} + \frac{1}{18} + \frac{1}{9} + a + b = 1$
(1 pont) $\Rightarrow a + b = \frac{1}{2} \Rightarrow b = \frac{1}{2} - a$
(4 pont) $\mathbb{P}(X = k, Y = 3) + \mathbb{P}(X = k, Y = 4) = \mathbb{P}(X = k)$ ill. $\sum_{k\in\{-1,0,1\}} \mathbb{P}(X = k, Y = l) = \mathbb{P}(Y = l)$
Ha a táblázatban ki vannak számolva a sor- és oszlopösszegek, de ezek nincsenek összefüggésbe hozva X és Y eloszlásával, legfeljebb 2 pont. Ha a későbbiekből az is kiderül, hogy sor- és oszlopösszegek a $\mathbb{P}(X = k)$ és $\mathbb{P}(Y = l)$ valószínűségek, akkor jár a teljes pont.
(2 pont) X és Y függetlensége miatt $\mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k)\,\mathbb{P}(Y = l)$

Konkrét k és l esetére felírva is jár a pont.
(1 pont) $\Rightarrow \frac{1}{9} = \mathbb{P}(X=-1, Y=3) = \mathbb{P}(X=-1)\,\mathbb{P}(Y=3) = \left(\frac{1}{9}+\frac{1}{18}\right)\cdot\left(\frac{1}{9}+\frac{2}{9}+a\right)$
Más (k,l) párra alkalmazva a fenti megfigyeléseket, ugyanúgy jár a teljes pont.
(2 pont) $\Rightarrow a = \frac{1}{3} \Rightarrow b = \frac{1}{6}$
(2 pont) $\mathbb{E}(XY) = \sum_{k\in\mathrm{Ran}(X)} \sum_{l\in\mathrm{Ran}(Y)} k\cdot l\cdot \mathbb{P}(X=k, Y=l)$
(4 pont) $\mathbb{E}(XY) = (-1)\cdot 3\cdot\frac{1}{9} + 0\cdot 3\cdot\frac{2}{9} + 1\cdot 3\cdot\frac{1}{3} + (-1)\cdot 4\cdot\frac{1}{18} + 0\cdot 4\cdot\frac{1}{9} + 1\cdot 4\cdot\frac{1}{6}$
(1 pont) $= \underline{\underline{\frac{10}{9}}} \approx 1,111$
AVAGY (az utolsó 6 pont helyett)
(3 pont) A függetlenség miatt $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.
(3 pont) A peremeloszlásokból: $\mathbb{E}(X) = (-1)\cdot\frac{1}{6} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{2} = \frac{1}{3}$, $\mathbb{E}(Y) = 3\cdot\frac{2}{3} + 4\cdot\frac{1}{3} = \frac{10}{3}$.
(1 pont) Tehát $\mathbb{E}(XY) = \frac{1}{3}\cdot\frac{43}{12} = \underline{\underline{\frac{10}{9}}} \approx 1,111$

2. Legyenek X és Y független, Exp(2) eloszlású folytonos valószínűségi változók.
a) Határozzuk meg $-Y$ eloszlásfüggvényét és sűrűségfüggvényét.
b) Határozzuk meg $X-Y$ sűrűségfüggvényét.

(1 pont) $\mathrm{Ran}(-Y) = \mathbb{R}^-$
(1 pont) $F_{-Y}(t) = \mathbb{P}(-Y < t) =$
(1 pont) $= \mathbb{P}(Y > -t)$
(2 pont) Ha $t \le 0$, akkor $= 1 - F_Y(-t) = e^{2t}$,
(1 pont) ha pedig $t > 0$, akkor $F_{-Y}(t) = \mathbb{P}(-Y < t) = 1$. (Ha ez szerepel, de $\mathrm{Ran}(-Y) = \mathbb{R}^-$ nem, akkor is jár az első 1 pont.)
(2 pont) Ezért $f_{-Y}(t) = 2e^{2t}$, ha $t \le 0$,
(1 pont) és 0 egyébként.

I. megoldás:
(1 pont) $\mathrm{Ran}(X) = \mathbb{R}^+$, $\mathrm{Ran}(-Y) = \mathbb{R}^-$, $\mathrm{Ran}(X-Y) = \mathbb{R}$
(3 pont) Mivel X és Y függetlenek, ezért $f_{X+(-Y)}(t) = \int_{-\infty}^{\infty} f_X(s) f_{-Y}(t-s)\mathrm{d}s$ (A függetlenségre való hivatkozás nélkül 2 pont.)
(3 pont) $= \int_{\max(0,t)}^{\infty} 2e^{-2s}\cdot 2e^{2t-2s}\mathrm{d}s =$ (A helyes határok megállapítása 2 pont, a helyes behelyettesítés 1 pont.)
(1 pont) $4e^{2t}\left[\frac{e^{-4s}}{-4}\right]_{\max(0,t)}^{+\infty}$
(1 pont) Ha $t \le 0$, akkor $4e^{2t}\left[\frac{e^{-4s}}{-4}\right]_0^{\infty} = e^{2t}$,
(1 pont) ha $t > 0$, akkor $4e^{2t}\left[\frac{e^{-4s}}{-4}\right]_t^{\infty} = e^{-2t}$.
(1 pont) Tehát

$$f_{X-Y}(t) = \begin{cases} e^{2t} & \text{ha } t \le 0, \\ e^{-2t} & \text{ha } t > 0. \end{cases}$$

II. megoldás:
(1 pont) $f_{X-Y}(t) = F'_{X-Y}(t)$
(1 pont) $\mathrm{Ran}(X) = \mathrm{Ran}(Y) = \mathbb{R}^+$, $\mathrm{Ran}(X-Y) = \mathbb{R}$ (Ha kimondva nem szerepel, de később a határoknál felhasználásra kerül, akkor is jár az 1 pont.)
(1 pont) $F_{X-Y}(t) = \mathbb{P}(X-Y<t) = \mathbb{P}(X-t<Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{y+t} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$
(1 pont) Mivel X és Y függetlenek, ezért $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$, $\forall x,y\in\mathbb{R}$, így
(1 pont) $= \int_{\max(0,-t)}^{\infty}\int_0^{y+t} 2e^{-2x}2e^{-2y}\mathrm{d}x\mathrm{d}y$
(1 pont) $= \int_{\max(0,-t)}^{\infty} 2e^{-2y}\int_0^{y+t} 2e^{-2x}\mathrm{d}x\mathrm{d}y = \int_{\max(0,-t)}^{\infty} 2e^{-2y}\left[-e^{-2x}\right]_0^{y+t}\mathrm{d}y$
(1 pont) $= \int_{\max(0,-t)}^{\infty} -2e^{-2t}e^{-4y} + 2e^{-2y}\mathrm{d}y = \left[\frac{1}{2}e^{-2t}e^{-4y} - e^{-2y}\right]_{\max(0,-t)}^{\infty}$
(1 pont) Ha $t \le 0$, akkor $\left[\frac{1}{2}e^{-2t}e^{-4y} - e^{-2y}\right]_{-t}^{\infty} = \frac{1}{2}e^{2t}$,
(1 pont) ha $t < 0$, akkor $\left[\frac{1}{2}e^{-2t}e^{-4y} - e^{-2y}\right]_0^{\infty} = 1 - \frac{1}{2}e^{-2t}$

(1 pont) Tehát

$$F_{X-Y}(t) = \begin{cases} \frac{1}{2}e^{2t} & \text{ha } t \leq 0, \\ 1 - \frac{1}{2}e^{-2t} & \text{ha } t > 0. \end{cases}$$

(1 pont) Amiből

$$f_{X-Y}(t) = \begin{cases} e^{2t} & \text{ha } t \leq 0, \\ e^{-2t} & \text{ha } t > 0. \end{cases}$$

3. A „Jar-jar” gyár befőttesüvegeket gyárt. Egy üveg gyártási idejének várható értéke 1 időegység, szórása 0,2. A cég 1000 db üveg legyártásához 1010 egység időt kér. (Az egyes üvegek gyártásához szükséges idők függetlenek.)
a) Közelítőleg mekkora a valószínűsége, hogy a cég le tudja gyártani az 1000 db üveget az ígért 1010 egység idő alatt?
b) A projektmenedzser javaslatára a cég olcsóbb alapanyagra vált, aminek következtében egy üveg gyártási idejének szórása 0,48-ra növekszik. Közelítőleg mennyi időt kérjen a cég, ha az 1000 db üveget ugyanakkora valószínűséggel szeretné időben legyártani, mint korábban?

(1 pont) Az egyes üvegek legyártásához szükséges idők: $X_1, \ldots, X_{1000}$, $\mathbb{E}(X_i) = 1$, $\mathbb{D}(X_i) = 0{,}2$.
(1 pont) $\mathbb{P}\left(\sum_{i=1}^n X_i < 1010\right) = ?$
(1 pont) $\mathbb{P}\left(\sum_{i=1}^n X_i < 1010\right) = \mathbb{P}(\sum_{i=1}^n X_i - 1000 < 10) =$
(1 pont) $= \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - 1000}{\sqrt{1000} \cdot 0{,}2} < \frac{10}{\sqrt{1000} \cdot 0{,}2}\right)$
(2 pont) A centrális határeloszlás-tétel miatt
(2 pont) $\frac{\sum_{i=1}^n X_i - 1000}{\sqrt{1000} \cdot 0{,}2}$ közelítőleg standard normális eloszlású.
(2 pont) Ezért a keresett valószínűség: $\approx \Phi\left(\frac{10}{\sqrt{1000} \cdot 0{,}2}\right) = \Phi(1.5811) \approx \underline{\underline{0{,}9429}}$.
(3 pont) Ha az előző számolást és érvelést megismételjük $\mathbb{D}(X_i) = 0{,}2$ helyett $\mathbb{D}(Y_i) = 0{,}48$ értékű szórásra és 1010 helyett x-re, akkor $\frac{\sum_{i=1}^n Y_i - 1000}{\sqrt{1000} \cdot 0{,}48}$ közelítőleg standard normális eloszlású. (Ha a standardizálás oka csak az a) részben szerepel, és itt nincs rá utalás, hogy miért ezeket a lépéseket végezzük, akkor ebből a 3 pontból legfeljebb 2 adható.)
(2 pont) $\mathbb{P}\left(\frac{\sum_{i=1}^n Y_i - 1000}{\sqrt{1000} \cdot 0{,}48} < \frac{x - 1000}{\sqrt{1000} \cdot 0{,}48}\right) = \Phi\left(\frac{x-1000}{\sqrt{1000} \cdot 0{,}48}\right)$
(2 pont) A kérdés, hogy milyen x-re lesz $\Phi\left(\frac{x-1000}{\sqrt{1000} \cdot 0{,}48}\right) = \Phi\left(\frac{10}{\sqrt{1000} \cdot 0{,}2}\right)$
(1 pont) azaz $\frac{x-1000}{\sqrt{1000} \cdot 0{,}48} = \frac{10}{\sqrt{1000} \cdot 0{,}2}$
(2 pont) átrendezve: $x = \frac{10 \cdot 0{,}48}{0{,}2} + 1000 = \underline{\underline{1024}}$

4. Legyen $Z \sim U(0; 2)$ folytonos valószínűségi változó.
a) Határozzuk meg a $\operatorname{cov}(Z^2, 3Z - 2)$ mennyiséget.
b) Határozzuk meg Z^2 lineáris regresszióját $3Z - 2$-re.

(1 pont) $f_Z(z) = \frac{1}{2}$ ha $0 < z < 2$ és 0 egyébként.
(2 pont) $\operatorname{cov}(3Z - 2, Z^2) = 3\operatorname{cov}(Z, Z^2)$
(2 pont) $\operatorname{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$ avagy $\operatorname{cov}(Z, Z^2) = \mathbb{E}(Z^3) - \mathbb{E}(Z)\,\mathbb{E}(Z^2)$ (avagy ugyanez Z helyett $3Z - 2$-vel)
(1 pont) Egy $g(X)$ transzformált folytonos valószínűségi változó várható értéke: $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x$. (Ha ez bárhol szerepel, vagy impliciten következetesen fel van használva, akkor is jár a teljes pont.)
(2 pont) $\mathbb{E}\left(Z^3\right) = \int_0^2 z^3 \frac{1}{2}\mathrm{d}z = \left[\frac{1}{8}z^4\right]_0^2 = 2$ (avagy $\mathbb{E}\left((3Z - 2)Z^2\right) = \frac{10}{3}$)
(1 pont) $\mathbb{E}(Z^2) = \int_0^2 z^2 \frac{1}{2}\mathrm{d}z = \left[\frac{1}{6}z^3\right]_0^2 = \frac{4}{3}$ (avagy $\mathbb{E}(Z^2) = \mathbb{D}^2(Z) + \mathbb{E}^2(Z) = \frac{1}{3} + 1 = \frac{4}{3}$)
(1 pont) $\mathbb{E}(Z) = \frac{0+2}{2} = 1$ (vagy $\mathbb{E}(3Z-2) = \int_0^2 (3z-2)\frac{1}{2}\mathrm{d}z = 1$, vagy $\mathbb{E}(3Z-2) = 3\,\mathbb{E}(Z) - 2 = 1$)
(1 pont) Tehát $\operatorname{cov}(Z, Z^2) = 2 - \frac{4}{3} \cdot 1 = \frac{2}{3}$ és így $\operatorname{cov}(3Z - 2, Z^2) = 3 \cdot \frac{2}{3} = \underline{\underline{2}}$
(1 pont) Jelölje Z^2 lineáris regresszióját $3Z - 2$-re $\beta \cdot (3Z - 2) + \alpha$

(2 pont) $\beta = \frac{\operatorname{cov}(Z^2, 3Z-2)}{\mathbb{D}^2(3Z-2)}$
(1 pont) $\mathbb{D}^2(3Z-2) = \mathbb{D}^2(3Z) = 9\,\mathbb{D}^2(Z)$
(1 pont) $= 9\left(\mathbb{E}(Z^2) - \mathbb{E}(Z)^2\right)$
(1 pont) $= 9\left(\frac{4}{3} - 1^2\right) = 3 \Rightarrow \beta = \underline{\underline{\frac{2}{3}}} \approx 0,6667$
(2 pont) $\alpha = \mathbb{E}(Z^2) - \beta \cdot \mathbb{E}(3Z-2)$
(1 pont) $= \frac{4}{3} - \frac{2}{3} \cdot 1 = \underline{\underline{\frac{2}{3}}} \approx 0,6667$

5. Bertold részt vesz egy nyereményjátékban. A nyeremény nagyságának meghatározásához egy szabálytalan érmével dobhat addig, amíg fejet nem kap. Jelölje X azt, hogy hanyadik próbálkozásra kapott először fejet. A nyereményének az $\{X = k+1\}$ eseményre vett feltételes várható értéke $\left(\frac{3}{2}\right)^k$. Tegyük fel, hogy X várható értéke $\frac{3}{2}$. Határozzuk meg a nyeremény várható értékét.

(1 pont) Y: nyeremény értéke, $\mathbb{E}(Y) = ?$
(2 pont) $\mathbb{E}(Y \mid X = k+1) = \left(\frac{3}{2}\right)^k$
(3 pont) $X \sim \mathrm{Geo}(p)$, ahol p a fejdobás valószínűsége
(2 pont) $\mathbb{E}(X) = \frac{1}{p} = \frac{3}{2} \Rightarrow p = \frac{2}{3}$.
(3 pont) A teljes várható érték tétele szerint
(4 pont) $\mathbb{E}(Y) = \sum_{i=1}^{\infty} \mathbb{E}(Y \mid X = x_i)\,\mathbb{P}(X = x_i)$ (avagy $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X))$) (Az utóbbi képletre a regresszió megadásával együtt jár a 4 pont.)
(2 pont) $= \sum_{k=0}^{\infty} \left(\frac{3}{2}\right)^k \left(\frac{1}{3}\right)^k \frac{2}{3}$
(2 pont) $= \frac{2}{3} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = \frac{2}{3} \cdot 2$
(1 pont) $= \underline{\underline{\frac{4}{3}}} \approx 1,333$
Ha a $k = 0$ eset nincs kezelve, azért 2 pont levonás jár.

6.* Legyenek X és Y független, folytonos valószínűségi változók. Igaz-e, hogy

$$\mathbb{E}(F_X(Y)) = \mathbb{P}(X < Y)\,?$$

Ha igen, bizonyítsuk be; ha nem, adjunk ellenpéldát.

(2 pont) $\mathbb{P}(X < Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$
(3 pont) Mivel X és Y függetlenek, ezért $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$, $\forall x, y \in \mathbb{R}$
(2 pont) Így $\mathbb{P}(X < Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_X(x) f_Y(y)\mathrm{d}x\mathrm{d}y$
(1 pont) A transzformált valószínűségi változó várható értékére vonatkozó tétel miatt
(3 pont) $\mathbb{E}(F_X(Y)) = \int_{-\infty}^{+\infty} F_X(y) f_Y(y)\mathrm{d}y$
(4 pont) $= \int_{-\infty}^{+\infty} \mathbb{P}(X < y) f_Y(y)\mathrm{d}y = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{y} f_X(x)\mathrm{d}x\right) f_Y(y)\mathrm{d}y$ (Ha az első formula nem szerepel, akkor is 4 pont. Ha csak az első formula szerepel, akkor 1 pont.)
(3 pont) $= \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_X(x) f_Y(y)\mathrm{d}x\mathrm{d}y$
(2 pont) A két oldal tehát ugyanazt adja, így az állítás igaz.