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. Válaszoljuk meg az alábbi kérdéseket az előadáson elhangzott állítások alapján.

   (a) Legyen X valószínűségi változó és $x \in \mathbb{R}$. Hogyan definiáljuk és hogyan jelöljük X eloszlásfüggvényének x pontbeli értékét? (A definíciót elég megadni, nem kell megmagyarázni a benne szereplő kifejezéseket.)

   (b) Legyen $F\colon \mathbb{R} \to [0,1]$ függvény. Milyen feltételek teljesülése esetén létezik olyan valószínűségi változó, amelynek F az eloszlásfüggvénye?

   Megoldás.
   (a) (6 pont)
   (2 pont) Jelölése: $F_X(x)$.
   (4 pont) Értéke: $\mathbb{P}(X < x)$.
   ($\mathbb{P}(X \leq x)$-re 2 pont. Ha nem x, hanem más betű van, akkor nem igazán a feladatot válaszolja meg, de ha egyébként jó, amit ír, akkor az egész (a)-ban összesen 1 pont levonás.)
   (b) (14 pont)
   (4 pont) Ha F balról folytonos (nem kell képlettel kiírni, de ha kiírja képlettel, az jó, bármilyen értelmes és azonosítható jelölés OK a bal oldali határértékre),
   (4 pont) monoton növő (nem kell odaírni, hogy nem feltétlenül szigorúan. Ha viszont azt írja, hogy szigorúan, akkor max. 1 pont),
   (3 pont) $F(x)$ határértéke $x \to -\infty$ esetén 0
   (3 pont) és $x \to \infty$ esetén 1.
   Ha nem meggyőző, hogy „és” kapcsolatra gondol ezek között, azért le lehet vonni 1 pontot. De ha az „és” kapcsolat megvan, csak nem minden feltétel van meg vagy nem minden feltételt helyesen ír, azért nem vonunk le pontot.

2. Egy lakótelepi fagyizóban pár perccel zárás előtt még négy gombóc fagyi áll rendelkezésre. Két kamasz fiú, Sándor és Tamás egymástól függetlenül 1/2–1/2 valószínűséggel odajön a fagyizóba zárás előtt. Ha csak az egyikük megy oda, akkor 3 gombóc fagyit vesz. Ha mindketten

odamennek, akkor testvériesen osztozva 2–2 gombóc fagyit vesznek. Aki nem megy oda, az értelemszerűen nem vesz fagyit. Jelölje S a Sándor és T a Tamás által megvásárolt fagyigombócok számát.

(a) Határozzuk meg S és T együttes eloszlását (lehetőleg táblázatos formában, de elég $\mathbb{P}(S = k, T = \ell)$-t megadni minden olyan (k, ℓ) párra, amelyre ez pozitív).
Az (a) részhez kivételesen nem kell indoklás. Minden más választ indokoljunk!

(b) Határozzuk meg S és T peremeloszlásait.

(c) Határozzuk meg S és T kovarianciáját és korrelációját.

(d) Független-e S és T?

Megoldás.

$T \backslash S$	0	2	3	p_T
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$	0	$\frac{1}{4}$
3	$\frac{1}{4}$	0	0	$\frac{1}{4}$
p_S	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	(1)

(a) (4 pont)
(4 pont) Tökéletes együttes eloszlás táblázat.
Nem tökéletes megoldás esetén 1–1 pont jár minden olyan pozitív elemre a táblázatban, amely helyesen szerepel. Továbbá ha minden 0 helyén 0 áll, akkor arra jár még 1 pont (akkor is, ha esetleg más helyen is 0-k állnak, ahol nem azoknak kellene állniuk).
Ha a táblázatban szereplő számok összege nem 1, azért -1 pont, ha a táblázatban negatív vagy 1-nél nagyobb számok is szerepelnek, azért darabonként -2 pont (persze az összpontszám nem mehet negatívba). Természetesen nem baj, ha a táblázatban üres sorok és/vagy oszlopok is szerepelnek.
(0 pont) (Indoklás: ha sem Sándor, sem Tamás nem megy el a fagyizóba, akkor $S = T = 0$. Ha csak Sándor megy, akkor $S = 3, T = 0$, ha pedig csak Tamás, akkor $S = 0, T = 3$. Végül ha mindketten mennek, akkor $S = T = 2$. És minden lehetséges kimenetel valószínűsége $(1/2)^2 = 1/4$.)

(b) (2 pont)
(1 pont) S peremeloszlását az együttes táblázat megfelelő oszlopösszegei adják: $\mathbb{P}(S = 0) = \frac{1}{2}$, $\mathbb{P}(S = 2) = \mathbb{P}(S = 3) = \frac{1}{4}$. Ez az egy pont csak akkor jár, ha (az adott táblázathoz képest) a számszerű értékek is jók.
(1 pont) T peremeloszlása ugyanaz, mint S-é.
Külön indoklás itt nem kell.
Ha felcseréli S és T szerepét (p_S-t írja a sorokhoz és p_T-t az oszlopokhoz, pedig S és T úgy szerepel, ahogy a mintamegoldásban van), akkor max. 1 pont (akkor sem 2 pont, ha a táblázat történetesen helyes, mert nem szimmetrikus táblázatnál ez nem lenne jó megoldás).

(c) (13 pont)
(1 pont) $\operatorname{cov}(S, T) = \mathbb{E}(ST) - \mathbb{E}(S)\,\mathbb{E}(T)$ (meg lehet határozni definíció szerint is, de ezt itt nem részletezzük).
(2 pont) Valahogyan meghatározza, hogy $\mathbb{E}(ST) = \frac{1}{4} \cdot 2 \cdot 2 = 1$, ebből 1 pont az indoklás. (Vagy meghatározza ST eloszlását, illetve az egyetlen nem 0 érték valószínűségét: $\mathbb{P}(ST = 4) = \mathbb{P}(S = T = 2) = \frac{1}{4}$, vagy az eloszlástranszformáció-szerű képlettel számol: $\mathbb{E}(ST) = \sum_{k \in \mathrm{Ran}(S)} \sum_{\ell \in \mathrm{Ran}(T)} k\ell\, \mathbb{P}(S = k, T = \ell) = 2 \cdot 2 \cdot \mathbb{P}(S = T = 2)$.)
(1 pont) $\mathbb{E}(S) = 2 \cdot \mathbb{P}(S = 2) + 3 \cdot \mathbb{P}(S = 3)(+0 \cdot \mathbb{P}(S = 0))$ (ez az 1 pont akkor is jár, ha a várható érték általános képletét írja fel helyesen).
(1 pont) $= 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = \frac{5}{4}$.

(1 pont) $\mathbb{E}(T) = \frac{5}{4}$ (ezt vagy ugyanígy lehet kiszámolni, mint az előzőt, vagy arra lehet hivatkozni, hogy S és T azonos eloszlásúak, illetve a szimmetriára). Itt kivételesen nem jár pont az indoklásra, elég a helyes végeredmény.
(0 pont) Tehát $\operatorname{cov}(S,T) = 1 - \frac{5}{4} \cdot \frac{5}{4} = -\frac{9}{16}$.
(1 pont) $\operatorname{corr}(S,T) = \frac{\operatorname{cov}(S,T)}{\mathbb{D}(S)\,\mathbb{D}(T)}$.
(1 pont) $\mathbb{D}(S) = \sqrt{\mathbb{D}^2(S)} = \sqrt{\mathbb{E}(S^2) - \mathbb{E}(S)^2}$. (Ez a pont csak akkor jár, ha mindkét azonosság megvan, viszont összevonva is OK.)
(1+1 pont) $\mathbb{E}(S^2) = 2^2\,\mathbb{P}(S=2) + 3^2\,\mathbb{P}(S=3) = \frac{13}{4}$.
(1 pont) Tehát $\mathbb{D}(S) = \sqrt{\frac{13}{4} - \frac{25}{16}} = \sqrt{\frac{27}{16}}$.
(1 pont) $\mathbb{D}(T) = \sqrt{\frac{27}{16}}$ a szimmetria miatt (vagy ugyanígy kiszámolja).
(1 pont) Így $\operatorname{corr}(S,T) = \frac{-\frac{9}{16}}{\frac{27}{16}} = -\frac{1}{3}$.

(d) (1 pont)
(1 pont) Nem függetlenek, mert a kovarianciájuk (vagy a korrelációjuk) nem 0.
Vagy: nem függetlenek, mert pl. $0 \neq \mathbb{P}(X=3)\,\mathbb{P}(Y=3) \neq \mathbb{P}(X=Y=3) = 0$.
Helyes indoklás nélkül nem jár a pont!

3. Egy digitális oktatási eszközökkel jól felszerelt iskolának 50 tanterme van, 1-től 50-ig számozva. Egy adott pillanatban megvizsgáljuk, hogy melyik teremben hány HDMI-kábel van. Jelölje X_i az i-edik teremben található HDMI-kábelek számát. Tegyük fel, hogy $X_1, \ldots, X_{50}$ (együttesen) független, Poisson-eloszlású valószínűségi változók $\lambda = 2$ paraméterrel.

   (a) Mi a valószínűsége, hogy az 1-es teremben pontosan két HDMI-kábel van az adott pillanatban?

   (b) Legyen X az iskola azon termeinek száma, amelyekben az adott pillanatban pontosan 2 darab HDMI-kábel van. Határozzuk meg (indoklással!) X várható értékét és szórását.

   (c) Közelítsük alkalmas approximációval a $\mathbb{P}(X \geq 13)$ valószínűséget. (A valószínűség pontos értékét nem szükséges meghatározni, használjuk a normális eloszlás táblázatot.)

Megoldás.
(a) (3 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(X_1 = 2)$
(1 pont) $= (\lambda^2/2)e^{-\lambda} = 2e^{-2}$ (ezért az 1 pontért elég a kettőből az egyik alak)
(1 pont) $\approx 0{,}2707$.
(b) (7 pont)
(1 pont) X binomiális eloszlású,
(1+1+1 pont) amelynek paraméterei $n = 50$ (mert ennyi kísérletet végzünk/ennyi terem van) és $p = 2e^{-2}$ (mert ilyen valószínűséggel sikeres egymástól függetlenül minden kísérlet/ilyen valószínűséggel van egy teremben pontosan 2 HDMI-kábel, a többi teremtől függetlenül). A kétfajta megfogalmazásból az egyik is elég az indoklásra járó 1 ponthoz, a másik 2 pont indoklás nélkül is jár. Ha csak a függetlenség hiányzik az indoklásból, azért ne vonjuk le ezt az 1 pontot.
(1 pont) Tehát $\mathbb{E}(X) = np = 50 \cdot 0{,}2707 = 13{,}535$ (13,5335 jön ki, ha nem az előzőleg kerekített értékkel dolgozunk. Ha jó a végeredmény, akkor OK indoklás nélkül is, np benne van a táblázatban.)
(2 pont) és $\mathbb{D}(X) = \sqrt{n \cdot p \cdot (1-p)} \approx 3{,}1418$ (a 3,1417 a pontos érték, ha nem a korábban kerekített értékkel dolgozunk). Itt egy pont jár a szórásnégyzet képletére plusz a gyökvonásra, egy pedig a számolásra.
Alternatív megoldás: X 50 darab független, azonos eloszlású indikátorok összege, amelyek paramétere a fenti p; indikátorváltozó minden momentuma p, ebből jön ki a várható érték linearitása

és a szórásnégyzet azonosságai (ezeket legalább egy kicsit részletezni kell) alapján az eredmény. Ha nem írja, hogy binomiális, akkor a (c)-ben rendes CHT kell, különben minimum 1 pont levonás!
(c) (10 pont)
(3 pont) Sztenderdizálunk: $\mathbb{P}(X \geq 13)$
$= \mathbb{P}(\frac{X-\mathbb{E}(X)}{\mathbb{D}(X)} \geq \frac{13-\mathbb{E}(X)}{\mathbb{D}(X)})$
(1 pont) $= \mathbb{P}(\frac{X-13{,}5335}{3{,}1417} \geq \frac{13-13{,}5335}{3{,}1417})$ (az előző lépéssel összevonható),
(1 pont) $\approx \mathbb{P}(\frac{X-13{,}5335}{3{,}1417} \geq -0{,}1698)$ (ez is összevonható az előző lépéssel. Itt az előző részfeladatban kapott, négy tizedesjegyre kerekített értékekkel számoltam és nem határoztam meg a kerekítések nélkül kapott eredményt, de elfogadunk minden olyan számot, ami alapján a 0,17 lesz a releváns szám a normális eloszlás táblázatában), (1 pont) ami a de Moivre–Laplace-tétel miatt (vagy: a centrális határeloszlás-tétel/CHT miatt)
(1 pont) közelítőleg $\mathbb{P}(Y \geq -0{,}1698)$, ahol Y standard normális eloszlású,
(1 pont) ami közelítőleg $\approx \mathbb{P}(Y \geq -0{,}17)$ (az előző lépéssel összevonható)
(1 pont) ami pedig $\mathbb{P}(Y \leq 0{,}17) = \mathbb{P}(Y < 0{,}17) = \Phi(0{,}17)$ (elég az egyik jelölés; nem jár külön pont a $\leq$-ról $<$-re való áttérésért, így a középső lépés el is hagyható, és ha az első megvan, akkor az utolsós is. De ha túl hamar tér át, – a normális eloszlás bevezetése előtt – és egyenlőséggel tér át, azért -1 pont. Később OK áttérni, tulajdonképpen a Φ-táblázatból való kiolvasásig bármikor lehet),
(1 pont) ami közelítőleg 0,5675.

4. Legyenek $X \sim \mathrm{Geo}(1/3)$ és $Y \sim \mathrm{Exp}(4)$ független valószínűségi változók. Adjuk meg az $\mathbb{E}(\cos(X)Y + XY^2 - 6X^2 + 5|X)$ regressziót.
Megoldás.
(2 pont) $\mathbb{E}(\cos(X)Y+XY^2-6X^2+5|X) = \mathbb{E}(\cos(X)Y|X)+\mathbb{E}(XY^2|X)+\mathbb{E}(-6X^2|X)+\mathbb{E}(5|X)$
(2 pont) a regresszió (vagy: a **feltételes** várható érték) linearitása miatt.
Ha csak annyit ír, hogy a linearitás miatt, akkor a 2 pont akkor jár, ha kiderül, hogy jól alkalmazza a regresszió linearitását, különben 0 pont. Ha azt írja, hogy a várható érték linearitása miatt, akkor max. 1 pont (akkor is, ha a feladat jól folytatódik). Ha általános képlettel fejezi ki a megfelelő linearitást, az OK.
(1+1 pont) $= \mathbb{E}(\cos(X)Y|X)+\mathbb{E}(XY^2|X)-6\,\mathbb{E}(X^2|X)+5$ (nem kell további indoklás és összevonható az előző lépéssel is)
(3 pont) $= \cos(X)\,\mathbb{E}(Y|X) + X\,\mathbb{E}(Y^2|X) - 6X^2 + 5$, (a $6X^2$-et nem muszáj $6X^2\,\mathbb{E}(1|X)$-ként felírni, szintén összevonható az előző lépéssel, de külön indoklás kell)
(2 pont) mert... (hivatkozik a regresszió megfelelő tulajdonságára, pl. ismert információ kiemelése, taking out what is known, homogenitás, X függvényeit kiemelhetjük az X-re vonatkozó feltételes várható értékből, X függvényei X ismeretében konstansnak számítanak stb. Ha látszik, hogy jól csinálja ÉS próbál indokolni, csak nem nagyon sikerül megneveznie, akkor lehet rá adni max. 1 pontot, ha nincs indoklás vagy sem a számolás, sem az indoklás nem jó, akkor 0 pont.)
(2 pont) $= \cos(X)\,\mathbb{E}(Y) + X\,\mathbb{E}(Y^2) - 6X^2 + 5$,
(2 pont) mert függetlenek (nem kell részletezni, hogy ha Y független X-től, akkor Y^2 is az).
(1 pont) $\mathbb{E}(Y)(= \frac{1}{\lambda}) = \frac{1}{4}$ (nem kell indokolni)
(1 pont) és $\mathbb{E}(Y^2) = \mathbb{D}^2(Y) + \mathbb{E}(Y)^2$
(1 pont) $= \frac{1}{4^2} + (\frac{1}{4})^2$ (itt a pont a helyes szórásnégyzetre jár, $\frac{1}{\lambda^2}$ formában is OK megadni)
(1 pont) $= \frac{2}{16} = \frac{1}{8}$.
(1 pont) Tehát a keresett regresszió $\frac{1}{4}\cos(X) + \frac{1}{8}X - 6X^2 + 5$ (ezzel ekvivalens alak is jó).
Ha egy adott fajta indoklás legalább egyszer jól szerepel, de nem írja oda, hogy máskor is ezt kell használni, viszont máskor is mindig jól csinálja, akkor ne vonjunk le pontot. Ha nem mindig jól csinálja és az indoklás is hiányzik, akkor az indoklásból is -1 pont.
Ha a feltételt elhagyja (olyan esetben, ahol nem lehet elhagyni a függetlenség miatt), azért -2

pont. Ha a val. változók helyére behelyettesíti a várható értékeket és számot kap végeredményül, azért is -2 pont, még akkor is, ha közbülső lépésként megjelenik a helyes végeredmény.

5. $n \in \mathbb{N}$ esetén legyenek $X_1, \ldots, X_n$ független, azonos eloszlású valószínűségi változók, amelyek súlyfüggvénye $p_\vartheta(x) = \mathbb{P}_\vartheta(X_1 = x) = (x-1)\vartheta^2(1-\vartheta)^{x-2}$, ahol $x \in \{2, 3, 4, \ldots\}$, $0 < \vartheta < 1$ pedig egy ismeretlen paraméter.

   (a) Határozzuk meg az $(X_1, \ldots, X_n)$ fae. minta egy adott $(x_1, \ldots, x_n)$ realizációja esetén a likelihood-függvény $L_\vartheta(x_1, \ldots, x_n)$ értékét.

   (b) Adjunk a minta alapján maximum likelihood-becslést ϑ-ra.

   A teljes pontszám eléréséhez – az idő rövidsége miatt – nem kell megvizsgálni, hogy a $(0, 1)$ *paramétertartomány határán lehet-e maximumhely.*

   Megoldás.

   (a) (4 pont)

   (1 pont) (Definíció szerint) $L_\vartheta(x_1, \ldots, x_n) = \prod_{i=1}^n p_\vartheta(x_i)$ (az is jó, ha kiírja, hogy mi a szorzat, ...-tal a közepén)

   (1 pont) $= \prod_{i=1}^n (x_i - 1)\vartheta^2(1-\vartheta)^{x_i-2}$

   (1 pont) $= \vartheta^{2n} \prod_{i=1}^n ((x_i - 1)(1-\vartheta)^{x_i-2})$

   (1 pont) $= \vartheta^{2n} \prod_{i=1}^n (x_i - 1) \prod_{i=1}^n (1-\vartheta)^{x_i-2}$.

   Az utolsó lépés az előtte lévővel összevonható. Az is rendben van, ha az utolsó lépést csak a (b)-ben végzi el, akár logaritmizálás után.

   (b) (16 pont)

   (1 pont) arra az ötletre, hogy meg kell határozni a log-likelihood-függvényt (akkor is, ha az nem is igazán sikerül).

   (1 pont) A log-likelihood-függvény $l_\vartheta(x_1, \ldots, x_n) = \ln L_\vartheta(x_1, \ldots, x_n)$ (ez a pont akkor is jár, ha ez csak később, impliciten derül ki, de csak akkor, ha a logaritmizálásnak legalább az első lépése helyes)

   (1+1+1 pont) $= 2n \ln \vartheta + \sum_{i=1}^n \ln(x_i - 1) + \sum_{i=1}^n (x_i - 2)\ln(1-\vartheta)$.

   (Ha ln helyett log-ot ír, azért ne vonjunk le pontot. Kevésbé egyszerűsített alak is jó abban az esetben, ha később helyesen tudja deriválni.)

   (1 pont) A log-likelihood függvényt ϑ szerint deriváljuk és megnézzük, hogy hol 0 a derivált. (Ez a pont akkor is jár, ha a deriválás esetleg nem igazán szerepel, feltéve, hogy az is megvan, hogy a deriváltat egyenlővé tesszük 0-val.)

   (1+1 pont) $\frac{\partial}{\partial\vartheta} l_\vartheta(x_1, \ldots, x_n) = \frac{2n}{\vartheta} - \sum_{i=1}^n \frac{x_i-2}{1-\vartheta}$ (más, értelmes jelölés is jó a deriváltra, később a másodikra is)

   (1+2 pont) pontosan akkor 0, ha $\vartheta = \frac{2n}{\sum_{i=1}^n x_i} = \frac{2}{\overline{x}_n}$ (nem kell az átlagjelölést használni, illetve az is elég, ha csak azt használja. Ebből két pont a részletszámításokra jár és egy a helyes végeredményre, utóbbi akár indoklás nélkül is, ha jó volt a deriválás.

   Részletszámítások pl.: a $\frac{2n}{\vartheta} - \sum_{i=1}^n \frac{x_i-2}{1-\vartheta} = 0$ egyenlet jobb oldalára átviszi a szummát és beszorozza mindkét oldalt $\vartheta(1-\vartheta)$-val (1 pont), majd felbontja a zárójeleket és ϑ-t kifejezi (1 pont).).

   (1+1 pont) $\frac{\partial^2}{\partial\vartheta^2} l_\vartheta(x_1, \ldots, x_n) = -\frac{2n}{\vartheta^2} - \sum_{i=1}^n \frac{x_i-2}{(1-\vartheta)^2}$

   (1 pont) ami minden pontban (vagy: a $\frac{2}{\overline{x}_n}$ pontban) kisebb, mint 0, így $\vartheta = \frac{2}{\overline{x}_n}$-ban valóban szigorú lokális maximum van (ezért az 1 pontért elég, ha azt írja, hogy lokális maximum, sőt az is, ha azt írja, hogy maximum).

   (1 pont) Így tehát az adott realizációra vonatkozó maximum likelihood becslés $\vartheta_*(x_1, \ldots, x_n) = \frac{2n}{\sum_{i=1}^n x_i} = \frac{2}{\overline{x}_n}$ (megint csak elég az egyik a kettő közül)

   (1 pont) és így a maximum likelihood becslés $\vartheta_*(X_1, \ldots, X_n) = \frac{2n}{\sum_{i=1}^n X_i} = \frac{2}{\overline{X}_n}$ (itt is elég az egyik, és ha megvan az egyik helyesen, nagy X_i-kkel és kiderül, hogy ezt tekinti ML-becslésnek, akkor nem baj, ha hiányzik a $\vartheta_*(X_1, \ldots, X_n)$ jelölés vagy más szerepel helyette).

Megjegyzés a javításhoz: nem baj, ha nem nézi meg, hogy x_i-k milyen értékeket vehetnek fel. Mivel realizációról van szó, ezeknek mind 1-nél nagyobb egész számoknak kell lenniük. A határokat a feladat szerint nem kellett ellenőrizni.

6. * Az X folytonos valószínűségi változó eloszlásfüggvénye

$$F_X(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \frac{x^2}{4}, & \text{ha } 0 < x < 2, \\ 1, & \text{ha } x \geq 2, \end{cases}$$

valamint X ismeretében Y egyenletes eloszlású a $(0; 1/X)$ intervallumon.
Adjuk meg a $\mathbb{P}(Y < 1)$ valószínűséget.
Megoldás.
(1 pont) Meghatározzuk X sűrűségfüggvényét, amit úgy kapunk meg, hogy F_X-et deriváljuk, ahol pedig F_X nem differenciálható, ott legyen 0 (vagy bármilyen pozitív szám). Elég, ha ez impliciten derül ki, és akkor is jár ez a pont, ha csak azzal a résszel foglalkozik, ahol nem 0 a sűrűségfüggvény, vagy ha nem veszi észre, hogy a szakaszhatárokon nem differenciálható. De ha nem jó a derivált, akkor ez a pont sem jár.
(1+1 pont) Vagyis

$$f_X(x) = \begin{cases} \frac{x}{2}, & \text{ha } 0 < x < 2, \\ 0, & \text{egyébként.} \end{cases}$$

(Javítóknak: minden olyan függvényre, ami nemnegatív és ezzel a képlettel majdnem mindenütt megegyezik, jár a 2 pont. A 0 részre csak akkor jár pont, ha egyértelműen kiderül, hogy a 0-ig és a 2-től is 0 a sűrűségfüggvény értéke.)

Innen 1. megoldás, folytonos TVT-vel.
(2 pont) A folytonos TVT (teljes valószínűség tétele folytonos esetben) szerint [ha eszébe jut, hogy ezt kellene használni, de nem tudja felírni használható alakban, akkor max. 1 pont]
(1 pont) $\mathbb{P}(Y < 1) = \int_{-\infty}^{\infty} \mathbb{P}(Y < 1|X = x) f_X(x) \mathrm{d}x$
(2 pont) $= \int_0^2 \frac{x}{2} \mathbb{P}(Y < 1|X = x) \mathrm{d}x$ (az előző lépéssel összevonható, ha kiderül, hogy a folytonos TVT-nek megfelelő képletet használ).
(2 pont) Ha $\frac{1}{x} < 1$, vagyis $x > 1$ (itt elég a két megfogalmazás közül az egyik), akkor $\mathbb{P}(Y < 1|X = x) = 1$ (mivel Y mindig 0 és $1/X$ közé esik, de bővebb indoklás nem kell).
(2 pont) Ha $\frac{1}{x} \geq 1$, vagyis $x \leq 1$, akkor

$$\mathbb{P}(Y < 1|X = x) = \frac{1}{1/x} = x$$

(ha jó a végeredmény és az, hogy mikor van ez az eset, akkor ez a 2 pont jár indoklás nélkül is és az indoklásra majd a következő 2 pont jár. A határokat a két esethez tetszőlegesen be lehet sorolni, akár mindkettőhöz is.),
(2 pont) mert $X = x$ ismeretében Y egyenletes eloszlású $(0; 1/x)$-en (tehát ez alapján a feltételes eloszlásfüggvény $F_{Y|X}(1|x) = \frac{1-0}{1/x-0}$, de ezt nem muszáj így felírni); vagy ugyanez feltételes sűrűségfüggvénnyel megfogalmazva: $f_{Y|X}(y|x) = \frac{1}{1/x} = x$ ilyen x-re, ha $0 < y < 1/x$ és 0 egyébként (ha ezt felírja, akkor nem muszáj az egyenletességre hivatkozni, benne van a feladatban); vagy egydimenziós geometriai valószínűségi mezővel (kedvező terület: 1, összes: $1/x$) érvelve.
(2 pont) Tehát $\mathbb{P}(Y < 1) = \int_0^1 \frac{x^2}{2} \mathrm{d}x + \int_1^2 \frac{x}{2} \mathrm{d}x$
(2 pont) $= \left[\frac{x^3}{6}\right]_0^1 + \left[\frac{x^2}{4}\right]_1^2$
(2 pont) $= \frac{1}{6} + \frac{4}{4} - \frac{1}{4} = \frac{11}{12}$.

Alternatív megoldás az utóbbi 17 pontért, folytonos TVT nélkül. Az integrálási határok tekintetében jóval bonyolultabb és kellemetlenebb!

5 pont kapható további számolások nélkül a megoldás menetének ötletére az alábbiak szerint – természetesen az is OK, ha a másik irányból kezdi a megoldást és a végén jut el a célhoz, ekkor is jár ebből az 5 pontból, amit menet közben megcsinál. (Ha nem világos, hogy hogyan akar eljutni a végeredményhez, akkor pontozás a fenti tanszéki alapelvek szerint.)
(1 pont) $\mathbb{P}(Y < 1) = F_Y(1)$
(1 pont) $= \int_{-\infty}^{1} f_Y(y)\mathrm{d}y$ (ha ez megvan, akkor az előző pont is jár az eloszlásfüggvényre való hivatkozás nélkül is)
(1 pont) $= \int_{0}^{1} f_Y(y)\mathrm{d}y$ (még ez is összevonható az előzővel, az pedig az azelőttivel).
(1 pont) $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x$,
(1 pont) ahol $f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x)$.
További pontokért a konkrét számolások:
(2 pont) $f_{Y|X}(y|x) = \frac{1}{1/x} = x$, ha $0 < x < 2$ és $0 < y < 1/x$ (ezt továbbra sem kell nagyon indokolni; a feladat szövegéből és a nevezetes eloszlás táblázatból következik, hogy az $\mathrm{U}(0; 1/x)$ eloszlásnak ez a sűrűségfüggvénye) és $f_{Y|X}(y|x) = 0$ egyébként. (Ha nem írja oda, hogy hol 0 a feltételes sűrűségfüggvény, de a későbbiekben úgy számol, hogy abból kiderül, akkor ne vonjunk le ebből a 2 pontból.)
(1 pont) tehát $f_{X,Y}(x,y) = \frac{x^2}{2}$, ha $0 < x < 2$ és $0 < y < 1/x$,
(1 pont) és $f_{X,Y}(x,y) = 0$ egyébként.
(1+1 pont) A $0 < x < 2$ és $0 < y < 1/x$ feltétel alapján ha y adott, akkor x lehetséges értékei 0 és $1/y$ között lehetnek. Ha $y \leq 1/2$, akkor ez azt jelenti, hogy x minden lehetséges értéke (0-tól 2-ig) előfordulhat. (Elég, ha ezek az integrálási határokból, impliciten derülnek ki.)
(1+1 pont) $f_Y(y) = \int_0^{1/y} \frac{x^2}{2}\mathrm{d}x = \Big[x^3/6\Big]_0^{1/y} = \frac{1}{6y^3}$, ha $y > 1/2$ (Ebből 1 pont jár a számolásra és 1 a végeredményre.)
(1+1 pont) és $f_Y(y) = \int_0^2 \frac{x^2}{2}\mathrm{d}x = \Big[x^3/6\Big]_0^2 = \frac{4}{3}$, ha $0 < y < 1/2$. (Hasonlóan.)
(0 pont) Ha $y \leq 0$, akkor $f_Y(y) = 0$.
(2 pont) Tehát $\mathbb{P}(Y < 1) = \int_0^{1/2} \frac{4}{3}\mathrm{d}y + \int_{1/2}^{1} \frac{1}{6y^3}\mathrm{d}y = \frac{2}{3} + \Big[-\frac{1}{12y^2}\Big]_{y=1/2}^{1} = \frac{2}{3} - \frac{1}{12} + \frac{1}{3} = \frac{11}{12}$.

Természetesen az is jó, ha nem f_Y-t határozza meg, hanem úgy számol, hogy

$$\mathbb{P}(Y < 1) = F_Y(1) = \int_{-\infty}^{1}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

Ennek felírására jár az alternatív megoldás első két pontja helyett 2 pont. Az utolsó 8 pont pedig a megfelelő integrálási határok meghatározására és az integrálások elvégzésére (az alternatív megoldás utolsó 8 pontjához hasonlóan).