Zárthelyi dolgozat Valószínűségszámítás és statisztikából, 2024.10.22

A zh 6. feladata 2026/27-ben is vizsgaanyag, de előreláthatólag zh utáni anyag, vizsgafelkészüléshez ajánljuk.

1. A $[0,1]$ intervallumon egymás után (egymástól függetlenül) találomra választunk két számot. Mennyi a valószínűsége, hogy a két szám közül az elsőként választott szám a nagyobb és ez a szám is kisebb 0,75-nél, ugyanakkor a másodikként választott szám értéke is nagyobb 0,5-nél?

2. Az A, B és C eseményekről tudjuk, hogy páronként függetlenek (azaz bármely kettő független egymástól), $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{3}$, továbbá $\frac{1}{9}$ annak a valószínűsége, hogy a három esemény közül egyik sem következik be. Mennyi a valószínűsége, hogy mindhárom esemény bekövetkezik? Együttesen független-e A, B és C?

3. Egy cégnek két irodája van, az egyik Prágában, a másik pedig Isztambulban, ahol a cég dolgozóinak fele-fele dolgozik. A prágai irodában dolgozók egyharmadának kedvenc itala a kávé, másik egyharmadának az ásványvíz, harmadik egyharmadának pedig a sör. Az isztambuli irodában dolgozók felének kedvenc itala szintén a kávé, másik felének pedig a tea. (Minden dolgozónak pontosan egy kedvenc itala van, a felsoroltak közül.)
   Véletlenszerűen kiválasztjuk a cég egy dolgozóját.

   (a) Mennyi a valószínűsége, hogy a dolgozó kedvenc itala a kávé?

   (b) Feltéve, hogy a dolgozó kedvenc itala a kávé, mi a valószínűsége, hogy Isztambulban dolgozik?

   (c) Feltéve, hogy a dolgozó kedvenc itala nem a kávé, mi a valószínűsége, hogy a sör az?

4. Egy épületben négy lift van. Minden lift mindennap $\frac{1}{2}$ valószínűséggel nem működik, függetlenül attól, hogy a többi lift közül aznap melyikek működnek, és függetlenül attól is, hogy a többi napon melyik liftek működnek (ideértve az adott liftet is).

   (a) Mennyi a valószínűsége, hogy egy adott napon legfeljebb két lift nem működik?

   (b) Egy adott naptól kezdve mindennap megfigyeljük, hogy melyik napon melyik liftek működnek, nevezzük ezt az első napnak. Jelölje Y azon nap sorszámát, amikor először fordul elő, hogy egyik lift sem működik.

      i. Milyen eloszlású Y és milyen paraméterrel/paraméterekkel?
      Az erre a kérdésre adott választ kivételesen nem kell indokolni.

      ii. Határozzuk meg Y szórását és Y^2 várható értékét.

      iii. Határozzuk meg a $\mathbb{P}(Y > 2024 | Y > 2022)$ feltételes valószínűség értékét.

5. Egy három vízszintesen elhelyezett képkockából álló kép kockáit egymástól függetlenül 1/2 valószínűséggel feketére és 1/2 valószínűséggel fehérre színeztük. Jelölje X a fehér képkockák számát, Y pedig azt, hogy hány képkockából áll a legnagyobb összefüggő fekete terület.
   (Például az ábrán látható szituációban X és Y értéke is 1.)

   (a) Határozzuk meg X és Y együttes eloszlását (lehetőleg táblázatos formában, de elég $\mathbb{P}(X = k, Y = \ell)$-t megadni minden olyan k-ra és ℓ-re, ahol ez pozitív).

   (b) Határozzuk meg X és Y peremeloszlásait.

   (c) Határozzuk meg $\mathbb{E}(XY)$ értékét.

(Minden válaszhoz indoklást is várunk, táblázatos megadás esetén is.)

6. * Az X valószínűségi változó $f_X(x)$ sűrűségfüggvényének értéke $\dfrac{1}{x^4}$ a $(-\infty, -c)$ és (c, ∞) intervallumokon, ahol $c > 0$ egy pozitív valós szám. Tudjuk továbbá, hogy f_X a $[-c, c]$ intervallumon konstans, és hogy folytonos a teljes valós számegyenesen. Határozzuk meg c értéket és X várható értékét. (Azt felhasználhatjuk bizonyítás nélkül, hogy $\mathbb{E}(|X|) < \infty$.)

Eloszlás neve	Jelölés	$\mathrm{Ran}(X)$	$\mathbb{P}(X = k)$ vagy $F_X(t)$	$f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
indikátor	$\mathbf{1}(p)$	$\{0, 1\}$	$p,\ 1-p$		p	$p(1-p)$
binomiális	$\mathrm{B}(n; p)$	$\{0, 1, ..., n\}$	$\binom{n}{k} p^k (1-p)^{n-k}$		np	$np(1-p)$
Poisson	$\mathrm{Pois}(\lambda)$	$\{0, 1, ...\}$	$\frac{\lambda^k}{k!} e^{-\lambda}$		λ	λ
geometriai	$\mathrm{Geo}(p)$	$\{1, 2, ...\}$	$(1-p)^{k-1} p$		$\frac{1}{p}$	$\frac{1-p}{p^2}$
egyenletes	$\mathrm{U}(a; b)$	$(a; b)$	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponenciális	$\mathrm{Exp}(\lambda)$	$\mathbb{R}^+$	$1 - e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Tudnivalók: A zárthelyi időtartama 90 perc. Számológépet lehet használni. A számszerű megoldásokat 4 értékes jegyre kerekítsük. A teljes pontszám eléréséhez a megoldás menete is szükséges, beleértve az egyes lépéseknél felhasznált tulajdonságok és tételek jelzését. Hivatkozni csak az előadáson elhangzottakra lehet. A zárthelyi első 30 percében nem lehet a termet elhagyni.

Pótló zárthelyi dolgozat

2024. november 5.

A pótzh 5.+6. feladata 2026/27-ben is vizsgaanyag, de előreláthatólag zh utáni anyag, vizsgafelkészüléshez ajánljuk.

1. A $[0,1]$ intervallumon egymás után (egymástól függetlenül) találomra választunk két számot. Mi a valószínűsége, hogy a két szám közül legalább az egyik nagyobb, mint $\frac{1}{5}$?

2. Egy forró nyári napon egy dél-pesti autóbuszjáraton sok busz közlekedik, amelyek mindegyike (egymástól függetlenül) $\frac{1}{3}$ valószínűséggel Ikarus 412-es, $\frac{1}{3}$ valószínűséggel gázüzemű Van Hool és $\frac{1}{3}$ valószínűséggel Mercedes Conecto típusú. Az Ikarus 412-eseken eleve nincs klímaberendezés. A másik két típusú buszon van klímaberendezés, de csak a buszok egy részén működik: tíz darab gázüzemű Van Hool buszból átlagosan csak nyolcon, tíz darab Mercedes Conectóból pedig átlagosan kilencen működik.
Véletlenszerűen kiválasztunk egy buszt erről az autóbuszjáratról ezen a napon.

   (a) Mi a valószínűsége, hogy a buszon működő klímaberendezés van?

   (b) Feltéve, hogy a buszon működő klímaberendezés van, mi a valószínűsége, hogy a busz Mercedes Conecto típusú?

   (c) Függetlenek-e a következő események: {a busz gázüzemű Van Hool típusú} és {a buszon működő klímaberendezés van}?

3. Egy egyetemi kar tudományos diákköri konferenciájára 150 hallgató jelentkezett a jelentkezési határidőig.

   (a) A tapasztalatok szerint a jelentkezett hallgatók mindegyike 95% valószínűséggel ad be ténylegesen pályaművet a konferenciára, a többi hallgatótól függetlenül (és csak az adhat be pályaművet, aki jelentkezett). Mi a valószínűsége, hogy a jelentkezett 150 hallgató közül legfeljebb kettő lesz, aki nem ad be pályaművet?

   (b) Erről a konferenciáról a legsikeresebb hallgatók továbbjutnak az országos tudományos diákköri konferenciára, amelyen nagyon sok hallgató vesz részt, de csak egy nagyon kis részük nyer első díjat. Jelöljük az itt első díjban részesülők számát Y-nal. Tudjuk, hogy Y várható értéke 6 fő. Határozzuk meg Y szórásnégyzetét és a $\mathbb{P}(Y > 2)$ valószínűséget.

4. Az X és Y valószínűségi változók együttes eloszlását az alábbi táblázat tartalmazza.

Y \ X	0	1	2
0	$\frac{1}{4}$	0	0
1	0	$\frac{1}{8}$	$\frac{3}{8}$
2	0	$\frac{1}{16}$	$\frac{3}{16}$

   (a) Határozzuk meg az X és az Y valószínűségi változó peremeloszlását és várható értékét.

   (b) Határozzuk meg $\mathbb{E}(XY)$-t.

   (c) Igaz-e, hogy
   $$\mathbb{P}(Y = 1|X = 2) = \mathbb{P}(Y = 1|X = 1)?$$
   És az igaz-e, hogy
   $$\mathbb{P}(X = 2|Y = 2) = \mathbb{P}(X = 2|Y = 1)?$$

   (d) Független-e X és Y?

Minden választ indokoljunk! A részfeladatokat bármilyen sorrendben megoldhatjuk.
A feladatsor a túloldalon folytatódik!

5. Egy nógrádi kistelepülésen az utóbbi években néha medvét észlelnek, egyszerre mindig csak egy medve érkezik magányosan. Egy adott időpontban elkezdjük megfigyelni, hogy mikor érkezik a következő medve, nevezzük ezt a $t = 0$ időpontnak. Jelölje X az ettől az időponttól a következő medve észleléséig eltelő időt (évben mérve). A megfigyelések szerint X folytonos, örökifjú eloszlást követ, 5 év várható értékkel.

   (a) Határozzuk meg X szórását.

   (b) Határozzuk meg X sűrűségfüggvényét a $t = 1$ helyen és X eloszlásfüggvényét a $t = -1$ helyen. *(Segít a nevezetes eloszlás táblázat.)*

   (c) Mennyi a valószínűsége, hogy a $t = 0$ időponttól kezdve a következő medve észleléséig több, mint 8 évig kell várni, feltéve, hogy több, mint 6 évig kell rá várni?

6. * Egy X folytonos valószínűségi változó sűrűségfüggvénye

$$f_X(x) = \begin{cases} e^x, & \text{ha } x < a, \\ 0, & \text{egyébként,} \end{cases}$$

   ahol $a \in \mathbb{R}$ egy alkalmasan választott (nem feltétlenül pozitív) valós szám (és $e \approx 2{,}71\ldots$).

   (a) Határozzuk meg a értékét.

   (b) Jelölje F_X az X valószínűségi változó eloszlásfüggvényét. Mely $x \in \mathbb{R}$ pontokra igaz, hogy $f_X(x) = F_X(x)$?

   (c) Határozzuk meg az $Y = e^X$ valószínűségi változó eloszlásfüggvényét és várható értékét.

Eloszlás neve	Jelölés	$\mathrm{Ran}(X)$	$\mathbb{P}(X = k)$ vagy $F_X(t)$	$f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
indikátor	$\mathbf{1}(p)$	$\{0,1\}$	$p,\ 1-p$		p	$p(1-p)$
binomiális	$\mathrm{B}(n;p)$	$\{0,1,...,n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$		np	$np(1-p)$
Poisson	$\mathrm{Pois}(\lambda)$	$\{0,1,...\}$	$\frac{\lambda^k}{k!}e^{-\lambda}$		λ	λ
geometriai	$\mathrm{Geo}(p)$	$\{1,2,...\}$	$(1-p)^{k-1}p$		$\frac{1}{p}$	$\frac{1-p}{p^2}$
egyenletes	$\mathrm{U}(a;b)$	$(a;b)$	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponenciális	$\mathrm{Exp}(\lambda)$	$\mathbb{R}^+$	$1-e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Tudnivalók: A zárthelyi időtartama 90 perc. Számológépet lehet használni. A számszerű megoldásokat 4 értékes jegyre kerekítsük. A teljes pontszám eléréséhez a megoldás menete is szükséges, beleértve az egyes lépéseknél felhasznált tulajdonságok és tételek jelzését. Hivatkozni csak az előadáson elhangzottakra lehet. A zárthelyi első 30 percében nem lehet a termet elhagyni.

Vizsgadolgozat

Valószínűségszámítás és statisztikából (VISZAB04, ÚJ Valszám), 2024.12.19.

2026/27-ben nem vizsgaanyag: 5., 6. feladat.

1. Válaszoljuk meg az alábbi kérdéseket az előadáson elhangzott definíciók és 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?

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?

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 $X_1, \dots, X_{50}$ (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.)

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.

5. $n \in \mathbb{N}$ esetén legyenek $X_1, \dots, X_n$ 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 $x \in \{2, 3, 4, \dots\}$, $0 < \vartheta < 1$ pedig egy ismeretlen paraméter.

   (a) Határozzuk meg az $(X_1, \dots, X_n)$ fae. minta egy adott $(x_1, \dots, x_n)$ realizációja esetén a likelihood-függvény $L_\vartheta(x_1, \dots, x_n)$ 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. A 6. feladatot lásd a túloldalon!

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.

Vizsgadolgozat

Valószínűségszámítás és statisztikából (VISZAB04, ÚJ Valszám), 2025.01.09.

Módosított változat, de 2026/27-ben már ennek a változatnak sem vizsgaanyag a 6c. feladata.

1. Legyen X és Y két valószínűségi változó pozitív és véges szórással.

   (a) Hogyan jelöljük X és Y korrelációját és hogyan definiáljuk ezt a korrelációt X és Y kovarianciája és szórásai segítségével?

   (b) Amennyiben X és Y függetlenek, milyen összefüggés áll fenn ezen két valószínűségi változó szórásnégyzetei és az összegük szórásnégyzete között?

   Megjegyzés: A szórásra, szórásnégyzetre és kovarianciára használhatjuk az előadásról ismert jelöléseket, ezen jelöléseket és a szórás(négyzet) és kovariancia jelentését nem kell megmagyarázni. A (b) részben a lehető legegyszerűbb helyes képletet adjuk meg.

2. Pista a sarki péknél minden reggel egymástól függetlenül 1/3 valószínűséggel sós kiflit vásárol, amely 150 forintba kerül, 1/3 valószínűséggel pogácsát, amely 300 forintba, és 1/3 valószínűséggel kakaós csigát, amely 450 forintba kerül. Minden reggel pontosan egy pékárut vásárol. 40 egymást követő napon át megvizsgáljuk, hogy Pista mennyit költ az egyes reggeleken a péknél, jelölje X_i azt, hogy hány forintot költ Pista az i-edik reggelen, ahol $i = 1, \dots, 40$.

   (a) Határozzuk meg X_1 várható értékét és szórását.

   (b) Jelölje X a Pista által a péknél a 40 nap alatt elköltött pénz összegét. Közelítsük alkalmas approximációval a $\mathbb{P}(X \geq 12\,500)$ valószínűséget. (A valószínűség pontos értékét nem kell meghatározni.)

3. Egy budapesti lakótelepi lakásban időnként előfordulnak áramszünetek. Tudjuk, hogy 3/4 annak a valószínűsége, hogy egy adott évben előfordul legalább egy áramszünet. Tegyük fel, hogy az egyes években egymástól függetlenül fordul elő vagy nem fordul elő áramszünet.

   (a) Jelölje Y tíz egymást követő év közül azon évek számát, amelyekben előfordul legalább egy áramszünet. Mennyi Y várható értéke?

   (b) Mi a valószínűsége, hogy 10 egymást követő évből legalább 8 évben lesz legalább egy áramszünet?

   (c) Egy adott pillanatban elkezdjük mérni, hogy mennyi idő telik el a következő áramszünetig. Jelölje ezt a várakozási időt (évben mérve) X. Feltehetjük, hogy X eloszlása folytonos és örökifjú, valamint a fentiek alapján $\mathbb{P}(X \leq 1) = 3/4$. Határozzuk meg X szórását.

   (d) Feltéve, hogy 4 évnél többet kell várni az első áramszünetre, mi a valószínűsége, hogy 5 évnél is többet kell várni rá?

4. Az (X, Y) folytonos valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f_{X,Y}(x,y) = \begin{cases} c(x+y), & \text{ha } 1 < x < 2 \text{ és } 1 < y < 2, \\ 0, & \text{egyébként,} \end{cases}$$

   ahol $c \in \mathbb{R}$ egy alkalmasan választott valós szám.

   (a) Határozzuk meg c értékét.

   (b) Határozzuk meg X sűrűségfüggvényét.

   (c) Határozzuk meg Y várható értékét.

A feladatsor a túloldalon folytatódik!

5. Egy rövid menetidejű repülőjáraton a menetidő jó közelítéssel normális eloszlású $\sqrt{2}/100$ óra szórással. Hat alkalommal megmértük ezen a járaton a menetidőt és (órában mérve, tizedestört alakban) az alábbi eredményeket kaptuk:

$$x_1 = 0{,}97 \quad x_2 = 0{,}98 \quad x_3 = 0{,}99 \quad x_4 = 1{,}01 \quad x_5 = 1{,}00 \quad x_6 = 0{,}99.$$

Feltehetjük, hogy az egyes alkalmakkor mért menetidők egymástól függetlenek.
(*A szórás az eredeti vizsgafeladatban ismeretlen volt, ezt módosítottam.*)

   (a) Adjunk a minta alapján 0,95 szintű konfidenciaintervallumot a mintaelemek várható értékére. Adjuk meg a minta korrigált empirikus szórását is.

   (b) Döntsük el $\varepsilon = 0{,}02$ terjedelmű próbával, hogy a mintaelemek várható értéke megegyezik-e 1 órával. Adjuk meg H_0-t, H_1-et, a próbastatisztika értékét, azt, hogy mi alapján döntünk, és magát a döntést is.

   (c) Adjuk meg (indoklással) a minta empirikus eloszlásfüggvényének értékét az $x = 0{,}995$ helyen.

   (d) Adjuk meg a minta korrigált empirikus szórását.

6. * *(A többdimenziós normális eloszlás fogalmának hiányában átdolgozott feladat.)*
Legyenek Z és W független, standard normális eloszlású változók és legyen $X := \sqrt{5}Z + 2W - 1$, $Y := \sqrt{5}Z - 2W + 2$.

   (a) Határozzuk meg az $(X, Y)^T$ valószínűségi vektorváltozó várható érték vektorát és kovarianciamátrixát. *Ez az eredeti vizsgafeladatnak nem volt része, de az előadás jelenlegi anyaga alapján is megoldható és nem 6-os nehézségű, valamint kell a (b) és (c) kitűzéséhez.*

   (b) Legyen $U = X + Y$ és $V = X - Y$. Határozzuk meg az $\begin{pmatrix} U \\ V \end{pmatrix}$ valószínűségi vektorváltozó várható érték vektorát és kovarianciamátrixát.

   (c) Ez alapján adjuk meg az $\mathbb{E}(V|U)$ regressziót.

Vizsgadolgozat

Valószínűségszámítás és statisztikából (VISZAB04, ÚJ Valszám), 2025.01.16.

2026/27-ben nem vizsgaanyag az 1c. és a 3b. feladat.

1. Válaszoljuk meg az alábbi kérdéseket az előadáson elhangzott definíciók alapján.

   (a) Legyenek A, B események. Mikor mondjuk azt definíció szerint, hogy A és B függetlenek?

   (b) Mikor nevezünk egy valószínűségi változót egyszerűnek?

   (c) Legyen (X, Y) folytonos valószínűségi vektorváltozó. Hogyan nevezzük és hogyan jelöljük azt a függvényt, amelynek értékét az (x, y) pontban $f_X(x) \neq 0$ esetén az

   $$\frac{f_{X,Y}(x,y)}{f_X(x)}$$

   képlettel definiáljuk, $f_X(x) = 0$ esetén pedig definíció szerint 0 az értéke?

   (d) Legyen $x \in \mathbb{R}$. A (c) részfeladatban szereplő szituációban hogyan számíthatjuk ki $f_X(x)$ értékét, ha $f_{X,Y}(x, y)$ értékét ismerjük minden $y \in \mathbb{R}$ esetén?

2. Legyenek $X_1, \ldots, X_{100}$ független, azonos eloszlású valószínűségi változók, amelyek sűrűségfüggvénye

   $$f_{X_1}(x) = \begin{cases} \frac{3}{7}x^2, & \text{ha } 1 < x < 2, \\ 0, & \text{egyébként.} \end{cases}$$

   (a) Bizonyítsuk be, hogy f_{X_1} valóban sűrűségfüggvény.

   (b) Határozzuk meg X_1 várható értékét és szórását.

   (c) Legyen $X = X_1 + \ldots + X_{100}$. Közelítsük alkalmas approximációval a $\mathbb{P}(X > 160)$ valószínűséget. (A valószínűség pontos értékét nem kell meghatározni.)
   Pótfeladat: Ha a (b)-t nem sikerült megoldani, tegyük fel, hogy $\mathbb{E}(X_1) = 1{,}6100$ *és* $\mathbb{D}(X_1) = 1/4$ *és oldjuk meg a (c)-t úgy.*

3. Legyen X és Y két független és a $(0; 1)$ intervallumon egyenletes eloszlású valószínűségi változó, valamint $U = X + Y$ és $V = XY$.

   (a) Határozzuk meg az U és V valószínűségi változók kovarianciáját és a várható értékeiket, valamint U szórásnégyzetét.
   Segítség: A feladat integrálás nélkül is megoldható (a tanult azonosságok használatával).

   (b) Mi V regressziós egyenese U-ra?

4. Feldobunk egy szabályos kockát, majd egy szabályos érmét annyiszor, amennyit a kocka mutat. Jelölje Y a dobott fejek számát és X a kockadobás értékét.

   (a) Várhatóan hány fejet dobunk, feltéve, hogy a kocka 3-at mutat (és miért)?

   (b) Mennyi $\mathbb{E}(Y|X)$?

   (c) Határozzuk meg Y várható értékét.

   (d) Lássuk be, hogy X és Y nem függetlenek.

A részfeladatokat tetszőleges sorrendben megoldhatjuk. A feladatsor a túloldalon folytatódik!

5. 7 darab, a BME campusán élő idősebb nyárfa magasságát megmértük és (méterben mérve) az alábbi eredményeket kaptuk:

   $x_1 = 23{,}1 \qquad x_2 = 21{,}5 \qquad x_3 = 22{,}0 \qquad x_4 = 22{,}8 \qquad x_5 = 21{,}2 \qquad x_6 = 20{,}9 \qquad x_7 = 22{,}5.$

   Egy kertészeti szakkönyv szerint az idősebb nyárfák magassága jó közelítéssel normális eloszlású, 1 méter szórással. Ezt az állítást elfogadva

   (a) adjunk 0,98 szintű konfidenciaintervallumot a minta alapján a mintaelemek várható értékére,

   (b) döntsük el $\varepsilon = 0{,}04$ terjedelmű próbával, hogy eléri-e a mintaelemek várható értéke a 22,5 métert. Adjuk meg H_0-t, H_1-et, a próbastatisztika értékét, azt, hogy mi alapján döntünk, és magát a döntést is.

   (c) Határozzuk meg a minta empirikus mediánját.

6. * $N \geq 2$ ember felírja a nevét egy-egy cetlire és ezeket egy kalapba rakják, majd mindenki véletlenszerűen húz egy cetlit a kalapból (visszatevés nélkül).

   (a) Az N ember között szerepel Anna és Béla. Mi a valószínűsége, hogy ők ketten egymás nevét húzzák?

   (b) Mi a várható értéke a *transzpozíciók* számának, azaz azon párok számának, akik egymás nevét húzzák?

Vizsgadolgozat

Valószínűségszámítás és statisztikából (VISZAB04, ÚJ Valszám), 2025.01.23.

2026/27-ben nem vizsgaanyag: 5b., 5c., 6. Az 5d. ugyan vizsgaanyag, de ezt sem kérdeznénk, mert a b. és a c. részek nélkül túl időigényes.

1. Válaszoljuk meg az alábbi kérdéseket az előadáson elhangzott definíciók és állítások alapján. Legyen $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező és legyenek $X, Y \colon \Omega \to \mathbb{R}$ valószínűségi változók.

   (a) Mikor mondjuk azt definíció szerint, hogy X és Y függetlenek?

   (b) Ha X és Y független valószínűségi változók, hogyan számíthatjuk ki a két változó szorzatának várható értékét a változók várható értékei segítségével, ha ezek a várható értékek mind léteznek? (Az $\mathbb{E}$ jelölést magyarázat nélkül használhatjuk a szokásos módon.)

   (c) Legyen $g \colon \mathbb{R} \to \mathbb{R}$ függvény és tegyük fel, hogy X folytonos és $g(X)$ várható értéke létezik. Jelölje $x \mapsto f_X(x)$ az X sűrűségfüggvényét. Hogyan számíthatjuk ki ekkor $g(X)$ várható értékét?

2. Egy szabályos dobókocka feldobását az első **hatosnál kisebb** eredményig ismételjük. Jelölje X az ehhez szükséges dobások számát (beleértve az első nem hatos dobást is).

   (a) Mi a valószínűsége, hogy a kockát legalább háromszor fel kell dobnunk?

   (b) Határozzuk meg X szórását és X^2 várható értékét.

   (c) Feltéve, hogy hét dobás után még nem sikerült egyszer sem hatosnál kisebbet dobnunk, mi a valószínűsége, hogy az utána következő két dobásban sem sikerül?

   (d) Definiáljuk ugyanebben a kísérletben az Y valószínűségi változót a következőképpen: $Y = 1$, ha a második dobás eredménye nem hatos és $Y = 0$ egyébként (ha nincs második dobás, Y értéke akkor is 0). Független-e X és Y?

3. Egy izzógyár mérnökei megfigyelték, hogy minél több izzót gyártanak egy napon, annál kisebb valószínűséggel lesz hibás egy-egy izzó. Ha egy napon $n \geq 1$ izzót készítenek, akkor minden izzó egymástól függetlenül $\frac{1}{2n}$ valószínűséggel lesz hibás.

   (a) Ha egy napon $n = 10$ izzót gyártanak, akkor mennyi az aznap gyártott hibás izzók számának várható értéke és szórásnégyzete (és miért)?

   (b) $n = 10$ esetén mi a valószínűsége, hogy 4-nél több, de 8-nál kevesebb izzó lesz hibás a 10 izzó közül? (A valószínűség pontos értékét adjuk meg.)

   (c) Tegyük fel, hogy egy napon n izzót gyártanak, ahol n nagyon nagy. Közelítsük alkalmas approximációval annak a valószínűségét, hogy az adott napon pontosan 3 izzó lesz hibás. (A valószínűség pontos értékét itt nem kell meghatározni.)

4. Legyenek $X \sim \mathrm{N}(1; 9)$ és $Y \sim \mathrm{N}(2; 16)$ független, normális eloszlású valószínűségi változók.

   (a) Mennyi $\mathbb{P}(X > 0{,}5)$?

   (b) Határozzuk meg $X + 2Y$ és $X - 4Y$ kovarianciáját és korrelációját.

   (c) Milyen eloszlású a $2X + Y$ valószínűségi változó és milyen paraméterekkel (és miért)?

5. Az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2}, & \text{ha } 0 < x < 2 \text{ és } 1 + x < y < 1 + 2x, \\ 0, & \text{különben.} \end{cases}$$

   (a) Mi Y értékkészlete? *Ehhez a részfeladathoz kivételesen nem kell indoklás, a többihez igen.*

(b) Határozzuk meg az $\mathbb{R} \to [0, \infty), y \mapsto f_{Y|X}(y|x)$ feltételes sűrűségfüggvényt minden olyan $x \in \mathbb{R}$ esetén, amelyre $f_X(x) > 0$. *A feladat a túloldalon folytatódik!*

(c) Ez alapján határozzuk meg az $\mathbb{E}(Y|X)$ regressziót.

(d) Független-e X és Y?

6. * Legyen $n \in \mathbb{N}$ és legyenek $X_1, \ldots, X_n$ független, azonos eloszlású, $\mathrm{U}(-\vartheta; \vartheta)$ eloszlású valószínűségi változók, ahol $\vartheta > 0$ egy ismeretlen paraméter.

(a) Az $(X_1, \ldots, X_n)$ minta rögzített $(x_1, \ldots, x_n)$ realizációja esetén határozzuk meg a likelihood-függvény $L_\vartheta(x_1, \ldots, x_n)$ értékét.

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

Zárthelyi dolgozat Valószínűségszámítás és statisztikából, 2024. október 22.

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. A $[0,1]$ intervallumon egymás után (egymástól függetlenül) találomra választunk két számot. Mennyi a valószínűsége, hogy a két szám közül az elsőként választott szám a nagyobb és ez a szám is kisebb 0,75-nél, ugyanakkor a másodikként választott szám értéke is nagyobb 0,5-nél?

   Megoldás.
   (2 pont) A véletlen kísérlet megegyezik egy véletlen pont választásával a $[0,1] \times [0,1]$ négyzetben.
   (1 pont) Ez az Ω eseménytér.
   (3 pont) Geometriai valószínűségi mező esetén a valószínűség $= \frac{T_{\text{kedvező}}}{T_\Omega}$. ($T_\Omega$ helyett $T_{\text{összes}}$ vagy hasonló is jó.)
   (1 pont) $T_\Omega = 1$ (indoklás nélkül is OK).
   (4 pont) (x, y)-nal jelölve a véletlen pontot a négyzetben, a kedvező terület a négyzetnek az a része, ahol $0{,}5 < y < x < 0{,}75$.
   (Ha nincs meg mind a három egyenlőtlenség, de kettő jó, akkor 2 pont, ha csak egy jó, akkor 1 pont. Ne vonjunk le pontot, ha nem szigorú egyenlőtlenségeket ír.)
   (3 pont) jó ábra
   (2 pont) A kedvező terület tehát a négyzetnek az $y = x$, $y = 0{,}5$ és $x = 0{,}75$ egyenletű egyenesek közötti része (elég, ha ez az ábráról derül ki, pl. besatírozza a kedvező területet).
   (1 pont) Ez egy derékszögű egyenlőszárú háromszög 0,25 hosszú befogókkal,
   (2 pont) tehát a területe $\frac{0{,}25 \cdot 0{,}25}{2} = 0{,}03125 = \dfrac{1}{32}$.
   (Lehet másra is hivatkozni, pl. a kedvező terület egy bizonyos négyzet területének fele stb.)
   (1 pont) Tehát a keresett valószínűség is 0,03125 (ha egyértelmű, hogy ezt a számot szánja végeredménynek, akkor ne vonjunk le pontot, ha nem hangsúlyozza, hogy le kell osztani az összes területtel, vagyis 1-gyel).

   Megjegyzés: Mivel a geometriai valószínűségi mezős feladatok megoldásait a gyakorlatokon is ritkán írtuk fel ennyire részletesen, ebben a feladatban azokra a lépésekre is jár a pont, amiket a hallgató nem írt fel expliciten, de a későbbi megoldásból egyértelműen kiderül, hogy jól

használta őket. Pl. ha kihagyja az első lépést, de utána a megfelelő módon rajzol ábrát/számol területeket, akkor ez a 2 pont is jár, vagy ha nem írja fel, hogy mi Ω, de később világos, hogy jól határozza meg a területét (és nem csak véletlenül jön ki neki 1 megoldásként), akkor arra is jár pont stb. stb. Ha nincs ábra, de a szövegből/számításokból kiderül, hogy a megfelelő síkidomok területét számolja, akkor az ábrára is adjunk pontot. Ha azonban a későbbi lépés, ami a hiányzó jó, de expliciten nem szereplő lépést használná, hibás, akkor sajnos a hiányzó lépésre sem lehet pontot adni.
Ezen belül konkrétan: ha nem rajzol ábrát, de nemcsak a $0{,}5 < y < x < 0{,}75$ egyenlőtlenséglánc van meg (amire fent külön járt pont), hanem az is, hogy geometriailag hogy néz ki/hogy számolható ki a kedvező terület, akkor helyes számolási mód (nem puszta helyes végeredmény!) esetén elfogadható az ábra hiányában is annak megnevezése, hogy a kedvező terület egy derékszögű egyenlőszárú háromszög. Hibás számolás esetén azonban az ábrára adandó pontszám is elvész.

2. Az A, B és C eseményekről tudjuk, hogy páronként függetlenek (azaz bármely kettő független egymástól), $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{3}$, továbbá $\frac{1}{9}$ annak a valószínűsége, hogy a három esemény közül egyik sem következik be. Mennyi a valószínűsége, hogy mindhárom esemény bekövetkezik? Együttesen független-e A, B és C?

Megoldás. (13 pont az első, 7 a második kérdésre)
(1 pont) A feladat első részében a keresett valószínűség $\mathbb{P}(A \cap B \cap C)$.
(2 pont) A feladat szerint $\frac{1}{9} = \mathbb{P}(\overline{A \cup B \cup C}) = 1 - \mathbb{P}(A \cup B \cup C)$ (a két lépés összevonható pontlevonás nélkül),
(1 pont) tehát $\mathbb{P}(A \cup B \cup C) = \frac{8}{9}$ (ez az 1 pont akkor is jár, ha csak a későbbiekben derül ki, hogy ezzel számol).
(2 pont) A függetlenségek/páronkénti függetlenség miatt $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$, $\mathbb{P}(A \cap C) = \mathbb{P}(A)\,\mathbb{P}(C)$ és $\mathbb{P}(B \cap C) = \mathbb{P}(B)\,\mathbb{P}(C)$ (ha nem írja ki, hogy a páronkénti függetlenség miatt, de az megvan, hogy a metszet valószínűsége a valószínűségek szorzata, akkor ne vonjunk le pontot. De ha nem szerepel itt mind a három egyenlőtlenség és később sem derül ki, hogy mind igaz, akkor max. 1 pont),
(1 pont) tehát mindhárom kettes metszet valószínűsége $\frac{1}{9}$.
(1 pont) A Poincaré-formula szerint
(0 pont) $\frac{8}{9} =$
(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) $= \frac{1}{3} + \frac{1}{3} + \frac{1}{3} - \frac{1}{9} - \frac{1}{9} - \frac{1}{9} + \mathbb{P}(A \cap B \cap C)$ (ha nem írja ki ennyire részletesen a számokat, de kijön egy helyes numerikus eredmény és korábban már szerepeltek az egyes valószínűségek, az rendben van),
(1 pont) amiből $\mathbb{P}(A \cap B \cap C) = \frac{2}{9}$.

Innen látszik, hogy a feladat sajnos nem jó, mert $A \cap B$ valószínűsége kisebb, mint $A \cap B \cap C$-é, tehát nem léteznek ilyen A, B, C események. Mivel erre a feladat kitűzésekor nem jöttünk rá, elfogadjuk a lenti megoldást, amely úgy tesz, mint ha léteznének. Ha valaki a $\mathbb{P}(A \cap B \cap C)$ kiszámolása alapján rájön, hogy ez a probléma fennáll, akkor az együttes függetlenség kiszámolására járó 7 pontot automatikusan megkapja (vagy ha már azelőtt rájön, hogy $\mathbb{P}(A \cap B \cap C)$-t végigszámolta volna, és kellően megindokolja, hogy mi a probléma, akkor is jár a feladat hátralevő részére a maximális pontszám) .

(3 pont) Az együttes függetlenséghez elég azt ellenőrizni, hogy $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$ teljesül-e (ha ez csak impliciten derül ki, de megemlíti az egyenlet mindkét oldalát, akkor ebből jár 2 pont, és ha világos, hogy azt nézi meg, hogy egyenlőek-e, akkor 3 pont jár),
(1 pont) mivel azt tudjuk, hogy páronként függetlenek (vagy: mivel tudjuk, hogy a kettes metszetekre teljesül a függetlenség definíciója; ezt persze egyesével is ki lehet fejezni).

Ha valaki belátja, hogy $\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$, és ebből helyesen következtet arra, hogy emiatt A, B, C nem lehet együttesen független, akkor nem kell a páronkénti függetlenségre hivatkozni. De például ha valaki azt mondja, hogy $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$ ekvivalens az együttes függetlenséggel, vagy (önmagában, a páronkénti függetlenség nélkül) következik belőle az együttes függetlenség, az nem jó.
(2 pont) $\mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot = \frac{1}{27}$,
(0 pont) és láttuk, hogy $\mathbb{P}(A \cap B \cap C) = \frac{2}{9}$,
(1 pont) tehát nem együttesen függetlenek (az is jó, ha azt írja, hogy „nem függetlenek").
A feladatnak ebben a részében nem adható részpontszám arra, ha úgy tippeli be, hogy nem együttesen függetlenek, hogy nem jelenik meg a megoldásban A, B és C (vagy esetleg a komplementereik) hármas metszete.
Lehetséges további jó megoldás pl.: megnézi, hogy $\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C})$ megegyezik-e $\mathbb{P}(\overline{A})\,\mathbb{P}(\overline{B})\,\mathbb{P}(\overline{C})$-vel (ami ekvivalens a hármas metszetre vonatkozó feltétellel, mert ha A, B, C páronként függetlenek, akkor $\overline{A}, \overline{B}, \overline{C}$ is azok). Ekkor nem kell $\mathbb{P}(A \cap B \cap C)$ értéke, mert $\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) = \frac{1}{9}$ a feladat szerint és $\mathbb{P}(\overline{A}) = \mathbb{P}(\overline{B}) = \mathbb{P}(\overline{C}) = 1 - 1/3 = 2/3$.
Jó úgy is, hogy kiszámolja pl. $\mathbb{P}(C|A \cap B)$-t és kijön neki, hogy ez nem $\mathbb{P}(C)$, ezért C nem lehet független $A \cap B$-től és így az együttes függetlenség sem teljesülhet.
Ilyen megoldásoknál a teljes pontszámhoz kell az itt megadott teljes indoklás (ami bővebb, mint a standard megoldás esetén).

3. Egy cégnek két irodája van, az egyik Prágában, a másik pedig Isztambulban, ahol a cég dolgozóinak fele-fele dolgozik. A prágai irodában dolgozók egyharmadának kedvenc itala a kávé, másik egyharmadának az ásványvíz, harmadik egyharmadának pedig a sör. Az isztambuli irodában dolgozók felének kedvenc itala szintén a kávé, másik felének pedig a tea. (Minden dolgozónak pontosan egy kedvenc itala van, a felsoroltak közül.)
Véletlenszerűen kiválasztjuk a cég egy dolgozóját.

   (a) Mennyi a valószínűsége, hogy a dolgozó kedvenc itala a kávé?

   (b) Feltéve, hogy a dolgozó kedvenc itala a kávé, mi a valószínűsége, hogy Isztambulban dolgozik?

   (c) Feltéve, hogy a dolgozó kedvenc itala nem a kávé, mi a valószínűsége, hogy a sör az?

   Megoldás.
   Megjegyzés: Ez egy viszonylag hosszú feladat 20 pontért (ellentétben pl. az 1. és a 2. feladatban), ezért bizonyos lépésekre, amikért más feladatban külön-külön is járhatna pont, itt csak összevonva jár.
   (0 pont) Jelölje K azt az eseményt, hogy a kiválasztott dolgozó kedvenc itala a kávé, S azt, hogy a sör (a többi italra nem lesz szükségünk külön-külön), P azt, hogy a kiválasztott dolgozó Prágában dolgozik és I azt, hogy Isztambulban.

   (a) (9 pont)
   (1 pont) A keresett valószínűség $\mathbb{P}(K)$ (elég, ha ez impliciten derül ki).
   (1 pont) A P és I események teljes eseményrendszert alkotnak,
   (1 pont) mert egymás komplementerei (ehelyett az is jó, ha ellenőrzi, hogy kizáróak és az uniójuk Ω; ha az egyik eseményt eleve a másik komplementereként jelöli és úgy mondja ki, hogy teljes eseményrendszert alkotnak, akkor további indoklás nélkül is jár a pont).
   (1 pont) A feladat szerint $\mathbb{P}(P) = \mathbb{P}(I) = 1/2$
   (1 pont) és $\mathbb{P}(K|P) = 1/3$, $\mathbb{P}(K|I) = 1/2$),
   (1 pont) így a teljes valószínűség tétele (elég: TVT) szerint
   (2 pont) $\mathbb{P}(K) = \mathbb{P}(P)\,\mathbb{P}(K|P) + \mathbb{P}(I)\,\mathbb{P}(K|I)$.
   (0 pont) $= \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{6} + \frac{1}{4}$
   (1 pont) $= \dfrac{5}{12}$.

(b) (5 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(I|K)$.
(1 pont) Bayes-tétellel (ez a pont a tétel megnevezésére jár)
(1 pont) $\mathbb{P}(I|K) = \frac{\mathbb{P}(K|I)\,\mathbb{P}(I)}{\mathbb{P}(K)}$ (természetesen az is jó, ha itt a nevezőre újra felírja a TVT-s képletet)
(1 pont) $= \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{5}{12}}$ (az (a) feladat alapján, tehát nem kell indokolni)
(1 pont) $= \dfrac{3}{5}$.
Nem tilos a Bayes-tétel használata nélkül kiszámolni $\mathbb{P}(I|K)$-t. Ekkor valahogy ki kell számolni $\mathbb{P}(I \cap K)$-t, aminél lehet a szorzási szabályra hivatkozni vagy a feltételes valószínűség definíciója alapján (amire nem muszáj hivatkozni) újra levezetni a szorzási szabálynak megfelelő számítást. Ilyen esetben a Bayes-tételre való hivatkozás helyett a szorzási szabályra való hivatkozásra, illetve ilyen hivatkozás hiányában a (konkrét példában) a szorzási szabálynak megfelelő számítás levezetésére jár az 1 pont.

(c) (6 pont – Bayes-tétellel is megoldható, amire szintén hivatkozni kell, itt ezt nem részletezzük)
(1 pont) $\mathbb{P}(\overline{K}) = 1 - \mathbb{P}(K) = \frac{7}{12}$ (az (a) feladat alapján, tehát nem kell rá hivatkozni),
(1 pont) A keresett valószínűség $\mathbb{P}(S|\overline{K}) = \frac{\mathbb{P}(S\cap\overline{K})}{\mathbb{P}(\overline{K})}$
(1 pont) $= \frac{\mathbb{P}(S)}{7/12}$. (Elég, ha impliciten derül ki, hogy $S \cap \overline{K}$ ugyanaz, mint S, és az is OK, ha eleve S-sel kezdi $S \cap \overline{K}$ helyett.)
(1 pont) $\mathbb{P}(S) = \mathbb{P}(S|P)\,\mathbb{P}(P)$ (itt most nem jár külön pont a szorzási szabályra való hivatkozásra)
(1 pont) $= \frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6}$,
(1 pont) vagyis $\mathbb{P}(S|\overline{K}) = \frac{1/6}{7/12} = \dfrac{2}{7}$.

4. Egy épületben négy lift van. Minden lift mindennap $\frac{1}{2}$ valószínűséggel nem működik, függetlenül attól, hogy a többi lift közül aznap melyikek működnek, és függetlenül attól is, hogy a többi napon melyik liftek működnek (ideértve az adott liftet is).

   (a) Mennyi a valószínűsége, hogy egy adott napon legfeljebb két lift nem működik?

   (b) Egy adott naptól kezdve mindennap megfigyeljük, hogy melyik napon melyik liftek működnek, nevezzük ezt az első napnak. Jelölje Y azon nap sorszámát, amikor először fordul elő, hogy egyik lift sem működik.

      i. Milyen eloszlású Y és milyen paraméterrel/paraméterekkel?
      Az erre a kérdésre adott választ kivételesen nem kell indokolni.

      ii. Határozzuk meg Y szórását és Y^2 várható értékét.

      iii. Határozzuk meg a $\mathbb{P}(Y > 2024|Y > 2022)$ feltételes valószínűség értékét.

Megoldás.
(a) (7 pont)
(1 pont) A nem működő liftek számát jelöljük X-szel, ekkor X binomiális eloszlású (itt az indoklásra nem jár külön pont)
(1 pont) $n = 4$ és $p = 1/2$ paraméterrel.
(1 pont) A keresett valószínűség $\mathbb{P}(X \leq 2)$
(1 pont) $= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)$
(2 pont) $= \binom{4}{0}\cdot(1/2)^4 + \binom{4}{1}\cdot(1/2)\cdot(1/2)^3 + \binom{4}{2}\cdot(1/2)^2\cdot(1/2)^2$ (a $\binom{4}{0}$ elhagyható és a $\binom{4}{1}$ helyett jó az is, hogy 4, a $\binom{4}{2}$ helyett az, hogy $\frac{4\cdot3}{2}$, és ha azt írja, hogy 6, jelöljük meg, hogy ez így hiányos, de ne vonjunk le pontot, illetve lehet minden tagnál $(1/2)^4$-t írni),
(1 pont) $= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16} = 0{,}6875$ (szabad négy értékes jegyre kerekíteni és az is jó, ha

eleve numerikus értékekkel számol – ha kiderül, hogy hogyan számolt, beleértve azt is, hogy $\binom{4}{2} = 6$. Elég közönséges tört alakban megadni a végeredményt.).
Természetesen lehet azt a Z val. változót is tekinteni, amelyik a működő liftek számát jelöli, ekkor $Z \sim \mathrm{B}(4; 1/2)$ és $\mathbb{P}(Z = k) = \mathbb{P}(X = 4 - k)$, ha $k \in \{0, 1, 2, 3, 4\}$.
Nem kötelező a binomiális eloszlásra hivatkozni, meg lehet magyarázni kombinatorikai alapon is, hogy miért ezek a valószínűségek lesznek. Ekkor azonban kell egy megfelelő eseménytér és arra is hivatkozni kell, hogy a valószínűségi mező klasszikus (vagy: kedvező/összes képlettel számolhatunk) és el kell mondani, hogy hány eleme van; erre jár az 1+1 pont a binomiális eloszlás és paramétere helyett.

(b) (összesen 13 pont)
(i) (2 pont):
(1 pont) Y geometriai eloszlású
(1 pont) $p = (1/2)^4 = 1/16 (= 0{,}0625)$ paraméterrel. (Ha végig kerekített tizedestört alakokkal számol és sok/jelentős a kerekítési hiba, akkor vonjunk le 1 pontot.)
(ii) (6 pont):
(1 pont) $\mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}$
(1 pont) $= \sqrt{\frac{1-p}{p^2}}$
(1 pont) $= \sqrt{\frac{15/16}{1/16^2}} = \sqrt{15 \cdot 16} = \sqrt{240} \approx 15{,}4919$ (elég, ha az utóbbi kettő közül egy szerepel, és 15,5 már jó).
Ne vonjunk le pontot, ha nem írja oda az $\sqrt{\frac{1-p}{p^2}}$-et általános p-re, de kiderül, hogy a konkrét p-re ezt használja. Ha azonban rossz képletbe helyettesít be és az általános képlet hiányzik, akkor az utolsó két pontból egyik sem jár.
(1 pont) $\mathbb{E}(Y^2) = \mathbb{D}^2(Y) + \mathbb{E}(Y)^2$.
(1 pont) $\mathbb{E}(Y)^2 = \frac{1}{p^2} = 16^2 = 256$,
(erre itt csak 1 pont jár, mert különben túl sok lenne, de nem baj, ha nem szerepel általános p-re az eredmény, sőt a 256 a 16^2 és az $1/p^2$ nélkül is elfogadható)
(1 pont) így $\mathbb{E}(Y^2) (= 240 + 256) = 496$.
(iii) (5 pont)
(1 pont) arra az ötletre, hogy az örökifjúságra kell hivatkozni, még ha ezt nem is sikerül megfelelően kivitelezni.
(1 pont) $\mathbb{P}(Y > 2024 | Y > 2022) = \mathbb{P}(Y > 2)$ (a geometriai eloszlás örökifjúsága miatt – ennek indoklására az előző pont jár)
(1 pont) $= 1 - \mathbb{P}(Y \leq 2) = 1 - \mathbb{P}(Y = 1) - \mathbb{P}(Y = 2)$ (a két lépés összevonható)
(1 pont) $= 1 - p - p(1-p) = 1 - \frac{1}{16} - \frac{1}{16} \cdot \frac{15}{16}$ (ehhez az 1 ponthoz elég, ha csak a konkrét p-re szerepel a számolás)
(1 pont) $= \frac{256-16-15}{256} = \dfrac{225}{256} = 0{,}87890625$ (elég négy értékes jegyre kerekítve).

A végeredményeket elfogadjuk közönséges tört és tizedestört alakban is, kivéve az (ii)-nél a szórást, ahol a végeredmény irracionális szám. Normálisan levonnánk 1 pontot azért, ha nincs numerikus érték a szórásnál, de itt túl kicsi az összpontszám, ezért kivételesen nem vonjuk le.

5. Egy három vízszintesen elhelyezett képkockából álló kép kockáit egymástól függetlenül 1/2 valószínűséggel feketére és 1/2 valószínűséggel fehérre színeztük. Jelölje X a fehér képkockák számát, Y pedig azt, hogy hány képkockából áll a legnagyobb összefüggő fekete terület.
(Például az ábrán látható szituációban X és Y értéke is 1.)

(a) Határozzuk meg X és Y együttes eloszlását (lehetőleg táblázatos formában, de elég $\mathbb{P}(X = k, Y = \ell)$-t megadni minden olyan k-ra és ℓ-re, ahol ez pozitív).

(b) Határozzuk meg X és Y peremeloszlásait.

(c) Határozzuk meg $\mathbb{E}(XY)$ értékét.

(Minden válaszhoz indoklást is várunk, táblázatos megadás esetén is.)

Megoldás.

(a) (10 pont)

Y \ X	0	1	2	3	p_Y
0	0	0	0	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{1}{8}$	$\frac{3}{8}$	0	$\frac{1}{2}$
2	0	$\frac{1}{4}$	0	0	$\frac{1}{4}$
3	$\frac{1}{8}$	0	0	0	$\frac{1}{8}$
p_X	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	(1)

(6 pont) Tökéletes együttes eloszlás (az indoklásra és a peremeloszlásokra majd később jár pont; a táblázatban szereplő nullákat csak akkor kell figyelembe venni, ha táblázatban adja meg az együttes eloszlást).
Részpontszám a következőkre adható:

- Az öt nem 0 valószínűség közül ahányat eltalált, annyi pont (akkor is, ha a többi elem nem stimmel, pl. a 0-k helyére nem mind 0-kat írt).
- Ezen felül +1 pont, ha az összes olyan elemre, aminek 0-nak kéne lennie, azt írja, hogy 0 (vagy ha nem táblázattal csinálja, akkor egyikről sem említi, hogy pozitív). Ennél az nem baj, ha máshova is ír 0-t.
- Ha az együttes eloszlás táblázatában szereplő számok összege nem 1, akkor viszont az eddigieken felül összesen -1 pont, ha pedig a táblázatban negatív vagy 1-nél nagyobb is szám előfordul, akkor még összesen további -1 pont (kivéve ha ezektől az (a) teljes pontszáma negatívvá válna; abban az esetben az (a)-ra 0 pont).

Az indoklásra az alábbiak szerint jár pont:
(1 pont) arra, hogy ha $X = 0$, akkor Y mindenképpen 3 kell, hogy legyen, és ha $X = 3$, akkor Y mindenképpen 0 kell, hogy legyen. (Ez még önmagában nem igazán teljes indoklás, de pont már jár érte. A rendes indoklás az, hogy ha $X = 0$, akkor nincs fehér képkocka, tehát mindhárom képkocka fehér, amelyek összefüggő területet alkotnak, így $Y = 3$. Ha pedig $X = 3$, akkor három egybefüggő fekete képkocka van, vagyis nincs fehér képkocka.) Ezt nyilván fordítva is el lehet mondani: ha $Y = 0$, abból $X = 3$, és ha $Y = 3$, abból $X = 0$ következik.
(1 pont) arra, hogy ha $X = 2$, akkor Y mindenképpen 1 kell, hogy legyen. (Ennyire is jár a pont; a rendes indoklás az, hogy ha $X = 2$, akkor egyetlen fekete képkocka van, így szükségképpen $Y = 1$.)
(2 pont) $\mathbb{P}(X = 1, Y = 2) = 2/8 = 1/4$ és $\mathbb{P}(X = 1, Y = 1) = 1/8$ megindoklására. Itt szükséges valami érdemi esetszétválasztás: ha $X = 1$, akkor ez az egy fehér képkocka vagy a kép szélén van (ennek a valószínűsége 2/8) és ekkor a két fekete kocka egymás mellé kerül, azaz $Y = 2$ lesz, vagy a kép közepén van (ennek a valószínűsége 1/8) és ekkor a fehér képkocka két részre vágja a fekete területet, ezért $Y = 1$ lesz.
Ha lerajzolja a lehetséges eseteket és leírja, hogy mikor mennyi X és Y értéke, az bőven elég.

(b) (4 pont)
Helyes peremeloszlások az előbbi táblázatban.

(2 pont) $p_X(k)$-t a k-nak megfelelő oszlop összegeként kapja minden k-ra (részpontszám adható, ha ez csak részben sikerül).
(2 pont) $p_Y(\ell)$-t az k-nek megfelelő sor összegeként kapja minden ℓ-re (részpontszám adható, ha ez csak részben sikerül).
Ezeket a pontokat akkor is megadjuk, ha hibás együttes eloszlás táblázatból számolja a peremeloszlásokat. De ha a $p_X(k)$-k vagy a $p_Y(l)$-ek összege nem 1 és ezt nem veszi észre, akkor legfeljebb 1–1 pont. Ha valamelyik peremeloszlásban negatív vagy 1-nél nagyobb $p_X(k)$ vagy $p_Y(l)$ értékek is megjelennek, akkor pedig arra a peremeloszlásra 0 pont.

(c) (6 pont)
1. megoldás: XY eloszlásával számolva.
(1 pont) $\mathrm{Ran}(XY) = \{0, 1, 2\}$ (ha olyan halmazt ír, ami ezt a három elemet tartalmazza és még esetleg néhány olyan elemet, amik előállnak a $0, 1, 2, 3$ számok közül kettő szorzataként, azt is el lehet fogadni). Elég, ha ez a későbbiekben impliciten derül ki.
(1 pont) $\mathbb{P}(XY = 1) = \mathbb{P}(X = 1, Y = 1) = 1/8$,
(2 pont) $\mathbb{P}(XY = 2) = \mathbb{P}(X = 1, Y = 2) + \mathbb{P}(X = 2, Y = 1) = 3/8 + 1/4 = 5/8$,
(0 pont) $\mathbb{P}(XY = 0) = 1 - 1/8 - 5/8 = 1/4$.
(1 pont) Így $\mathbb{E}(XY) = 1 \cdot 1/8 + 2 \cdot 5/8 (+0 \cdot 1/4)$
(1 pont) $= \dfrac{11}{8}$.
Ha $\mathbb{P}(XY = 0)$ jó, de semmi másra nem járna pont a (c)-ben, akkor erre lehet adni 1 pontot.
2. megoldás: eloszlástranszformáció-szerűen számolva.
(1 pont) Megadja az $\mathbb{E}(XY) = \sum_{k=0}^{3} \sum_{\ell=0}^{3} k\ell\, \mathbb{P}(X = k, Y = \ell)$ képletet vagy valami hasonlót; a lehetetlen (k, ℓ) párokat persze nem kötelező belevenni. Elég, ha impliciten derül ki, hogy ezt a képletet használja (ha a következő konkrét képlet jó, akkor ez a pont is jár).
(3 pont) $\mathbb{E}(XY) = 1 \cdot \mathbb{P}(X = Y = 1) + 2 \cdot \mathbb{P}(X = 2, Y = 1) + 2 \cdot \mathbb{P}(X = 1, Y = 2)$ (a 0-s szorzatokat is oda lehet írni, de nem kötelező)
(1 pont) $= 1/8 + 2 \cdot 3/8 + 2 \cdot 1/4$
(1 pont) $= \dfrac{11}{8}$.
Ha az megvan, hogy XY melyik cellákban 0, de semmi másra nem járna pont a (c)-ben, akkor erre lehet adni 1 pontot.

Ha semmi nem jön össze a megoldásból, de az legalább egyértelműen kiderül, hogy egy 8 elemű eseménytérről van szó ÉS klasszikus valószínűséggel kell számolni, akkor lehet adni összesen az egész feladatra max. 2 pontot. De csak akkor, ha semmi másra nem járna pont a feladatban, és ebből a 2 pontból is levonunk 1–1 pontot a nem 0 összegekre és a negatív vagy 1-nél nagyobb valószínűségekre (úgy, hogy az összpontszám legalább 0 legyen).

6. * Az X valószínűségi változó $f_X(x)$ sűrűségfüggvényének értéke $\dfrac{1}{x^4}$ a $(-\infty, -c)$ és (c, ∞) intervallumokon, ahol $c > 0$ egy pozitív valós szám. Tudjuk továbbá, hogy f_X a $[-c, c]$ intervallumon konstans, és hogy folytonos a teljes valós számegyenesen. Határozzuk meg c értékét és X várható értékét. (Azt felhasználhatjuk bizonyítás nélkül, hogy $\mathbb{E}(|X|) < \infty$.)

Megoldás (12 pont c-re, 8 a várható értékre).
(2 pont) A folytonossághoz az kell, hogy $f_X(x) = 1/c^4$ legyen, ha $-c < x < c$. (Nem kell külön indokolni.)
(1 pont) Ahhoz, hogy f_X sűrűségfüggvény legyen, az kell, hogy f_X nemnegatív legyen (erről elég ennyit mondani, nem kell indokolni, hogy miért teljesül)
(1 pont) és teljesüljön, hogy $1 = \int_{-\infty}^{\infty} f_X(x)\mathrm{d}x$
(3 pont) $= \int_{-\infty}^{-c} \frac{1}{x^4}\mathrm{d}x + \int_{-c}^{c} \frac{1}{c^4}\mathrm{d}x + \int_{c}^{\infty} \frac{1}{x^4}\mathrm{d}x$.
(Ha nincs meg neki, hogy $f_X(x) = 1/c^4$, ha $-c < x < c$, de felírja ugyanezt paraméteresen a megfelelő konstanssal, akkor is jár ez a 3 pont. Részpontszám adható.)

(2 pont) $\int_{-\infty}^{-c} \frac{1}{x^4}\mathrm{d}x = \left[-\frac{1}{3}x^{-3}\right]_{-\infty}^{-c} = -\frac{1}{3}(-c)^{-3} = \frac{1}{3c^3}$ (1 a primitív függvényre, 1 az azt követő számolásra),
(1 pont) és hasonlóan $\int_{c}^{\infty} \frac{1}{x^4}\mathrm{d}x = \frac{1}{3c^3}$.
(1 pont) $\int_{-c}^{c} \frac{1}{c^4}\mathrm{d}x = \frac{2c}{c^4} = \frac{2}{c^3}$.
(1 pont) Így tehát $(2 + 2 \cdot \frac{1}{3})\frac{1}{c^3} = 1$, vagyis $c^3 = 8/3$, tehát $c = \frac{2}{\sqrt[3]{3}} \approx 1{,}3867$.
Kb. teljes megoldás a várható értékre:
(1 pont) $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x$,
(1 pont) mivel $\mathbb{E}(|X|) < \infty$ (ennyi elég, de bővebben: mivel $\mathbb{E}(|X|) < \infty$, X várható értékét értelmezzük és a fenti képlettel számolhatjuk ki).
(1 pont) $= \int_{-\infty}^{-\frac{2}{\sqrt[3]{3}}} \frac{x}{x^4}\mathrm{d}x + \int_{-\frac{2}{\sqrt[3]{3}}}^{\frac{2}{\sqrt[3]{3}}} \frac{x}{\frac{16}{3^{4/3}}}\mathrm{d}x + \int_{\frac{2}{\sqrt[3]{3}}}^{\infty} \frac{x}{x^4}\mathrm{d}x$ (paraméteresen/korábban kapott rossz c-értékkel is OK, és az előző ponttal összevonható, tehát elég a konkrét esetre vonatkozó képlet, de ha rossz, akkor az előző pont csak akkor jár, ha az általános képlet minimális szintaktikai hibától eltekintve jó)
(1 pont) $\int_{-\infty}^{-\frac{2}{\sqrt[3]{3}}} \frac{x}{x^4}\mathrm{d}x = \int_{-\infty}^{-\frac{2}{\sqrt[3]{3}}} \frac{1}{x^3}\mathrm{d}x = \left[-\frac{1}{2}\frac{1}{x^2}\right]_{-\infty}^{-\frac{2}{\sqrt[3]{3}}} = -\frac{3^{2/3}}{8} + 0$. (Numerikusan is jó: $-\frac{3^{2/3}}{8} \approx -0{,}26001$.)
(1 pont) Hasonlóan adódik, hogy $\int_{-\frac{2}{\sqrt[3]{3}}}^{\infty} \frac{x}{x^4}\mathrm{d}x$ pont ennek az ellentettje, tehát $\frac{3^{2/3}}{8}$.
(2 pont) Végül $\int_{-\frac{2}{\sqrt[3]{3}}}^{\frac{2}{\sqrt[3]{3}}} \frac{x}{c^4} = \left[\frac{x^2}{2\cdot\frac{16}{3^{4/3}}}\right]_{-\frac{2}{\sqrt[3]{3}}}^{\frac{2}{\sqrt[3]{3}}} = 0$.
(1 pont) Tehát $\mathbb{E}(X) = 0$.

Valószínűsítjük, hogy aki meg tudja oldani ezt a feladatot, az előbb-utóbb a szimmetriára fog hivatkozni és nem fogja végigszámolni az egészet..

Ezen belül 1. lehetőség: a 3 konkrét integrál kiszámítása után jön rá, hogy a szimmetriát kell használni és 0 lesz az eredmény.
(2 pont) Az első 2 pont: mint a *kb. teljes megoldásnál.*
(3 pont) A szimmetria miatt 0 lesz az eredmény (ez a 3 pont további indoklás nélkül is jár, de a további pontok nem. Ha nem hivatkozik a szimmetriára sem, vagy semmi ezzel ekvivalens dologra, akkor számolás nélkül összesen max. 1 pont arra, hogy 0 lesz a várható érték.)
Alternatíva erre a 3 pontra: $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x = \lim_{N\to\infty} \int_{-N}^{N} f_X(x)\mathrm{d}x = \lim_{N\to\infty} 0 = 0$ (az utolsó előtti lépés elhagyható). Ennél a verziónál is meg kell említeni, hogy ez azért biztos, hogy jó, mert $\mathbb{E}(|X|) < \infty$, amire lentebb adunk pontot.
(1 pont) Az első és a harmadik integrál egyenlő egymással, mert a 0-ra szimmetrikus intervallumokon integrálunk és az integrandus páratlan függvény/a primitív függvény páros (ezt valahogy értelmesen megfogalmazza, számolni nem kell. Ez jó úgy is, hogy a $-\infty$-től 0-ig és 0-tól ∞-ig terjedő réeszre bontja szét az integrálást, ekkor a következő 2 pont is jár).
(1 pont) A középső integrál 0, mert szintén egy 0-ra szimmetrikus intervallumon integrálunk és az integrandus páratlan függvény/a primitív függvény páros.
(1 pont) Végül $\mathbb{E}(X)$ értelmes/létezik, mert $\mathbb{E}(|X|) < \infty$.
(0 pont) Így $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x = 0$.
2. lehetőség: már az általános képlet felírása után észreveszi, hogy a szimmetriát lehet használni és 0 lesz az eredmény. Ekkor a szimmetriát valahogy indokolni kell, ha nem is írja fel a 3 szakaszt, pl. a legegyszerűbb itt is az, ha a $-\infty$-től 0-ig és 0-tól ∞-ig terjedő részre bontja szét az integrálást.

Eloszlás neve	Jelölés	$\mathrm{Ran}(X)$	$\mathbb{P}(X=k)$ vagy $F_X(t)$	$f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
indikátor	$\mathbf{1}(p)$	$\{0,1\}$	$p,\ 1-p$		p	$p(1-p)$
binomiális	$\mathrm{B}(n;p)$	$\{0,1,...,n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$		np	$np(1-p)$
Poisson	$\mathrm{Pois}(\lambda)$	$\{0,1,...\}$	$\frac{\lambda^k}{k!}e^{-\lambda}$		λ	λ
geometriai	$\mathrm{Geo}(p)$	$\{1,2,...\}$	$(1-p)^{k-1}p$		$\frac{1}{p}$	$\frac{1-p}{p^2}$
egyenletes	$\mathrm{U}(a;b)$	$(a;b)$	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponenciális	$\mathrm{Exp}(\lambda)$	$\mathbb{R}^+$	$1-e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Pótló zárthelyi dolgozat
2024. november 5.

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. A $[0,1]$ intervallumon egymás után (egymástól függetlenül) találomra választunk két számot. Mi a valószínűsége, hogy a két szám közül legalább az egyik nagyobb, mint $\frac{1}{5}$?

   Megoldás.
   1. megoldás: geometriai valószínűségi mezővel.
   (2 pont) A véletlen kísérlet megegyezik egy véletlen pont választásával a $[0,1]\times[0,1]$ négyzetben.
   (1 pont) Ez az Ω eseménytér.
   (3 pont) Geometriai valószínűségi mező esetén a valószínűség $=\frac{T_{\text{kedvező}}}{T_\Omega}$. ($T_\Omega$ helyett $T_{\text{összes}}$ vagy hasonló is jó.)
   (1 pont) $T_\Omega=1$ (indoklás nélkül is OK).
   (4 pont) (x,y)-nal jelölve a véletlen pontot a négyzetben, a kedvező terület a négyzetnek az a része, ahol $x>\frac{1}{5}$ vagy $y>\frac{1}{5}$ teljesül. (Ha csak az egyik van meg a két egyenlőség közül, vagy a„vagy" kapcsolat nem világos nemcsak a szövegből, de a megoldás további részéből sem, akkor max. 2 pont. Ha csak apróbb pontatlanságok vannak, akkor lehet 3 pont. Ne vonjunk le pontot, ha nem szigorú egyenlőtlenségeket ír).
   (3 pont) jó ábra
   (1 pont) $T_{\text{kedvező}}=T_{\text{összes}}-T_{\text{kedvezőtlen}}$ (az is OK, ha $1-T_{\text{kedvezőtlen}}$t ír)
   (2 pont) A kedvezőtlen terület az, amelyik az $x=\frac{1}{5}$ egyenletű egyenestől balra és az $y=\frac{1}{5}$ egyenletű egyenestől lefelé van (ha az ábra alapján egyértelműen kiderül, hogy ezt tekinti kedvezőnek, akkor ez a 2 pont szöveges leírás nélkül is jár), ami egy négyzet (elég, ha a későbbiekben azonosítható, hogy a téglalap területe alapján és jól számol).
   (2 pont) Ennek a területe $\frac{1}{5}\cdot\frac{1}{5}=\frac{1}{25}$.
   (1 pont) Tehát a keresett valószínűség $\dfrac{24}{25}$ (ha egyértelmű, hogy ezt a számot szánja végeredménynek, akkor ne vonjunk le pontot, ha nem hangsúlyozza, hogy le kell osztani az összes területtel, vagyis 1-gyel).

Szabad egyből a kedvező területet kiszámolni, ebben az esetben a $T_{\text{kedvező}} = T_{\text{összes}} - T_{\text{kedvezőtlen}}$ helyett a kedvező terület kiszámolására jár még 1 pont. A kedvező területet nem kell körülírni, elég, ha az ábrában helyesen szerepel vagy a számolásnál azonosítható, hogy ezt a területet számolja ki.

2. megoldás: Poincaré-formulával és függetlenséggel.
(0 pont) Jelölje X az elsőként és Y a másodikként választott számot.
(2 pont) A keresett valószínűség $\mathbb{P}(X > 1/5 \text{ vagy } Y > 1/5) = \mathbb{P}(\{X > 1/5\} \cup \{Y > 1/5\})$ (az utóbbi önmagában is elég, de ha nem derül ki, hogy unióval kellene számolni, akkor ebből a 2 pontból maximum 1 jár. Az uniónál kisebb szintaktikai hibákért, pl. hiányzó halmazjelölésekért ne vonjunk le pontot.)
(1 pont) A Poincaré-formula szerint
(3 pont) $\mathbb{P}(\{X > 1/5\} \cup \{Y > 1/5\}) = \mathbb{P}(X > 1/5) + \mathbb{P}(Y > 1/5) - \mathbb{P}(\{X > 1/5\} \cap \{Y > 1/5\})$ (utóbbi valószínűség jó $\mathbb{P}(X > 1/5, Y > 1/5)$ alakban megadva is, az összeg helyett pedig jó a $2\,\mathbb{P}(X > 1/5)$, később jár majd 1 pont annak a megindoklására, hogy $\mathbb{P}(Y > 1/5) = \mathbb{P}(X > 1/5)$).
(2+1 pont) $\mathbb{P}(\{X > 1/5\} \cap \{Y > 1/5\}) = \mathbb{P}(X > 1/5)\,\mathbb{P}(Y > 1/5)$, mert függetlenek (2 pont az egyenlőségre, 1 az indoklásra),
(2+1 pont) és $\mathbb{P}(Y > 1/5) = \mathbb{P}(X > 1/5)$, mert X és Y azonos eloszlásúak (2 pont az egyenlőségre, bárhol is derül ki a feladat során, és 1 az indoklásra. Azt, hogy azonos eloszlásúak, bárhogy megindokolhatja, ami erre utal, pl. az is jó, hogy mindkét számot találomra a [0,1]-en/ugyanúgy választjuk. Ehelyett az is jó, ha kiszámolja a két valószínűséget külön-külön).
(3 pont) X és Y is egyenletes eloszlásúak a $(0, 1)$ intervallumon. Vagy: valahogy leírja, hogy X és Y egy-egy intervallumba való esésének valószínűsége az intervallumhosszal arányos (azaz egydimenziós geometriai valószínűségi mezőről van szó – utóbbi önmagában is elég), ahol az összes hosszúság 1.
(3 pont) Ezért $\mathbb{P}(X > 1/5) = 1 - \mathbb{P}(X < 1/5) = \frac{1-0}{5-0} = \frac{4}{5}$. (Nem kell ennyire részletesen kiírni, de derüljön ki, hogy az eloszlásfüggvény táblázatban szereplő képletét használja, vagy integrálhatja a sűrűségfüggvényt.) Vagy: $(1/5, 1)$ egy $4/5$ hosszúságú szakasz, az ebbe esés valószínűsége (az egydimenziós geometriai val. mező „kedvező hosszúság/összes hosszúság" képlete szerint) $(4/5)/1 = 4/5$.
(1 pont) Tehát a keresett valószínűség $\frac{4}{5} + \frac{4}{5} - \frac{16}{25}$
(1 pont) $= \dfrac{24}{25}$.
Szintén jó, ha ezzel a módszerrel és komplementer valószínűséggel számol. Ebben az esetben azt kell használni, hogy egy mínusz a keresett valószínűség $\mathbb{P}(X < 1/5$ és $Y < 1/5)$, amit fel lehet írni metszetként, majd a fenti megoldáshoz hasonlóan kiszámolni. Így tulajdonképpen könnyebb is: nem kell a Poincaré-formula. Az arra adandó 1+3 pont helyett 1+3 pont arra jár, hogy térjünk át a komplementer valószínűségre, majd alkalmazzuk helyesen a vonatkozó De Morgan-azonosságot. A legelső 2 pontból 1 a keresett vagy a komplementer valószínűség bármilyen helyes azonosítására (és/vagy vagy metszet/unió) és 1 az és/vagy és metszet/unió közti áttérésre jár, ahol megint elég a metszettel/unióval kezdeni.

2. Egy forró nyári napon egy dél-pesti autóbuszjáraton sok busz közlekedik, amelyek mindegyike (egymástól függetlenül) $\dfrac{1}{3}$ valószínűséggel Ikarus 412-es, $\dfrac{1}{3}$ valószínűséggel gázüzemű Van Hool és $\dfrac{1}{3}$ valószínűséggel Mercedes Conecto típusú. Az Ikarus 412-eseken eleve nincs klímaberendezés. A másik két típusú buszon van klímaberendezés, de csak a buszok egy részén működik: tíz darab gázüzemű Van Hool buszból átlagosan csak nyolcon, tíz darab Mercedes Conectóból pedig átlagosan kilencen működik.
Véletlenszerűen kiválasztunk egy buszt erről az autóbuszjáratról ezen a napon.

   (a) Mi a valószínűsége, hogy a buszon működő klímaberendezés van?

(b) Feltéve, hogy a buszon működő klímaberendezés van, mi a valószínűsége, hogy a busz Mercedes Conecto típusú?

(c) Függetlenek-e a következő események: {a busz gázüzemű Van Hool típusú} és {a buszon működő klímaberendezés van}?

Megoldás. (a) (9 pont)
(0 pont) Jelölje K azt az eseményt, hogy a kiválasztott buszon működő klímaberendezés van, I azt, hogy a busz Ikarus 412-es, V azt, hogy gázüzemű Van Hool és M azt, hogy Mercedes Conecto típusú.
(1 pont) A keresett valószínűség $\mathbb{P}(K)$.
(1 pont) A feladat szövege szerint $\mathbb{P}(V) = \mathbb{P}(M) = 1/3$ (és $\mathbb{P}(I) = 1/3$, de erre nem igazán lesz szükség),
(1 pont) $\mathbb{P}(K|V) = \frac{8}{10} = \frac{4}{5}$ és $\mathbb{P}(K|M) = \frac{9}{10}$
(1 pont) és $\mathbb{P}(K|I) = 0$. (Ezt nem muszáj ideírni; ha később egyértelmű, hogy ennek megfelelően számít, akkor is jár ez a pont.)
(1 pont) Az I, M, V események teljes eseményrendszert alkotnak,
(1 pont) mert kizáróak és lefedik az eseményteret (vagy ezzel ekvivalens indoklás, ennél jobban nem kell indokolni). Alternatív lehetőség: az eseményteret megszorítjuk a K esemény szempontjából releváns $\Omega' = M \cup V$-re, és ezen az új eseménytérek az M és V események teljes eseményrendszert alkotnak. [Ezen az eseménytéren a $\mathbb{P}(\cdot|M \cup V)$ a releváns valószínűségi mérték, ami alapján a végén vissza kellene térni $\mathbb{P}(\cdot)$-re, de erre nem kell hivatkozni.]
Ha azt állítja, hogy M és V teljes eseményrendszer, akkor az utóbbi 2 pontból csak 1 jár, de különben ne vonjunk le ezért több pontot. Ha bármiből kiderül, hogy I nélkül csak M és V nem teljes eseményrendszer, de $\mathbb{P}(K|I) = 0$ miatt úgy lehet vele számolni, mintha az lenne, akkor ne vonjuk le ezt az 1 pontot.
(1 pont) Így a teljes valószínűség tétele (elég: TVT) szerint
(1 pont) $\mathbb{P}(K) = \mathbb{P}(V)\,\mathbb{P}(K|V) + \mathbb{P}(M)\,\mathbb{P}(K|M)(+\,\mathbb{P}(I)\,\mathbb{P}(K|I))$
(1 pont) $= \frac{1}{3}(\frac{8}{10} + \frac{9}{10}(+0)) = \frac{17}{30}$.

(b) (5 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(M|K)$.
(1 pont) Bayes-tétellel (jó az „egyszerű Bayes-tétel” is; még akkor is fogadjuk el, hogy „egyszerű”, ha esetleg odaírja az I-s tagot)
(2 pont) $\mathbb{P}(M|K) = \frac{\mathbb{P}(M)\,\mathbb{P}(K|M)}{\mathbb{P}(K)}$ (a nevező jó $\mathbb{P}(V)\,\mathbb{P}(K|V) + \mathbb{P}(M)\,\mathbb{P}(K|M)(+\,\mathbb{P}(I)\,\mathbb{P}(K|I))$ alakban is, de már az (a)-ban kiszámoltuk az értékét)
(1 pont) $= \frac{\frac{1}{3}\cdot\frac{9}{10}}{\frac{17}{30}}$
(0 pont) $= \frac{9}{17}$.
Ha már az a-ban szerepelt a $\frac{1}{3}\cdot\frac{9}{10}$ szorzat, akkor itt elég a szorzatukat ($\frac{9}{30}$) közölni. Ha $\mathbb{P}(K\cap M)$-et ír a számlálóba (ami nem a Bayes-tételben szereplő alak), de nem indokolja, hogy ez a szorzási szabály miatt annyi, amennyi, akkor -1 pont.

(c) (6 pont)
Önmagában annak a betippelésére, hogy nem függetlenek, nem jár pont.
(1 pont) Nem függetlenek, mert... (valahogy megindokolja, és ha nem is feltétlenül jól, az indoklás legalább szintaktikailag értelmes, és hihető, hogy komolyan gondolta – kétes esetben a javító dönt)
(2 pont) K és V függetlenséghez az kellene, hogy $\mathbb{P}(K \cap V) = \mathbb{P}(K)\,\mathbb{P}(V)$ teljesüljön.
(0 pont) Korábban láttuk, hogy $\mathbb{P}(V) = 1/3$ és kiszámoltuk, hogy $\mathbb{P}(K) = 17/30$.
(2 pont) $\mathbb{P}(K\cap V) = \mathbb{P}(K|V)\,\mathbb{P}(V) = \frac{8}{30}$ (ha ez korábban már szerepelt, nem kell megismételni)
(0 pont) a szorzási szabály miatt. (Erre itt már külön nem adunk pontot, mert a közzétett régi zh-kban sem adtunk hasonló helyzetben.)
(1 pont) Ez azonban nem egyezik meg $\mathbb{P}(V)\,\mathbb{P}(K) = \frac{17}{90}$-nel (, tehát tényleg nem függetlenek).

Ha valaki arra hivatkozik, hogy 10 busz van és arányokkal számol, akkor az összes TVT-re/Bayes-tételre és feltételeik ellenőrzésére járó pontot elveszti és még -2 pontot kap (feltéve, hogy a pontszáma a végén pozitív marad), de a többi pontot megkaphatja. Aki arra hivatkozik, hogy n busz van és arányokkal számol, ahol n valami egész szám, az -1 pontot kap (feltéve, hogy a pontszáma a végén pozitív marad), és az indoklásokra akkor kaphat pontot, ha megindokolja egyesével, hogy klasszikus a valószínűségi mező, vagy ha leír egy feltételes valószínűséget, akkor a szorzási szabályra hivatkozik ott, ahol a metszet valószínűsége valójában abból származik. Itt szigorúbbak vagyunk mint a zh-ban a kávés feladatnál, mert ott a feladatban tényleg arányok szerepeltek, itt azonban valószínűségek, és hangsúlyozzuk, hogy sok busz van.

3. Egy egyetemi kar tudományos diákköri konferenciájára 150 hallgató jelentkezett a jelentkezési határidőig.

   (a) A tapasztalatok szerint a jelentkezett hallgatók mindegyike 95% valószínűséggel ad be ténylegesen pályaművet a konferenciára, a többi hallgatótól függetlenül (és csak az adhat be pályaművet, aki jelentkezett). Mi a valószínűsége, hogy a jelentkezett 150 hallgató közül legfeljebb kettő lesz, aki nem ad be pályaművet?

   (b) Erről a konferenciáról a legsikeresebb hallgatók továbbjutnak az országos tudományos diákköri konferenciára, amelyen nagyon sok hallgató vesz részt, de csak egy nagyon kis részük nyer első díjat. Jelöljük az itt első díjban részesülők számát Y-nal. Tudjuk, hogy Y várható értéke 6 fő. Határozzuk meg Y szórásnégyzetét és a $\mathbb{P}(Y > 2)$ valószínűséget.

Megoldás.
(a) (7 pont)
(1 pont) A pályaművet be nem adó jelentkezők számát jelöljük X-szel, ez binomiális eloszlású
(1 pont) $n = 150$ és $p = 0{,}05$ valószínűséggel (több indoklás nem kell).
(1 pont) A keresett valószínűség $\mathbb{P}(X \leq 2)$
(1 pont) $= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)$
(2 pont) $= (\binom{150}{0} \cdot 0{,}05^0 \cdot) 0{,}95^{150} + \binom{150}{1} \cdot 0{,}05 \cdot 0{,}95^{149} + \binom{150}{2} \cdot 0{,}05^2 \cdot 0{,}95^{148}$ (a $\binom{150}{1}$ helyett 150 is jó természetesen)
(1 pont) $\approx 0{,}01815$ (ha valaki megadja közönséges tört alakban, az is jó, de mi erre itt nem vállalkozunk).
Megjegyzés: természetesen jó úgy is, hogy Y a beadott pályaművek száma, ami $B(150; 0{,}95)$ eloszlású $\mathbb{P}(Y \geq 148)$-at keressük és $\mathbb{P}(X = k) = \mathbb{P}(Y = 150 - k)$ minden $k \in \{0, 1, \ldots, 150\}$-re.

(b) (13 pont)
(2 pont) Y Poisson-eloszlású (nem kell indokolni, mert a kontextusból kiderül, hogy azzal érdemes modellezni)
(1 pont) és $6 = \mathbb{E}(Y)$
(1 pont) $= \lambda$ az eloszlás paramétere.
(1 pont) Tehát $\mathbb{D}^2(Y) = \lambda$
(1 pont) $= 6$ (ha egyből azt mondja, hogy $\mathbb{D}^2(Y) = 6$, indoklás nélkül, akkor az utóbbi 2 pontból csak 1 jár).
(1 pont) $\mathbb{P}(Y > 2) = 1 - \mathbb{P}(Y \leq 2)$
(1 pont) $= 1 - \mathbb{P}(Y = 0) - \mathbb{P}(Y = 1) - \mathbb{P}(Y = 2)$
(1+1+1 pont) $= 1 - e^{-\lambda} - \lambda e^{-\lambda} - (\lambda^2/2) \cdot e^{-\lambda}$
(2 pont) $(= 1 - (1 + 6 + 18)e^{-6}) \approx 0{,}9380$. (Ha itt egyből csak a numerikus érték szerepel és az kb. jó, akkor jár a 2 pont. Kisebb numerikus hiba esetén 1 pont jár, hibás végeredmény esetén indoklás nélkül 0 pont.)
Ha elfelejti a $\mathbb{P}(Y = 0)$-t, akkor azért -1 pont és a kiszámolásának hiányáért is -1, de a többit el lehet fogadni.

4. Az X és Y valószínűségi változók együttes eloszlását az alábbi táblázat tartalmazza.

Y \ X	0	1	2
0	$\frac{1}{4}$	0	0
1	0	$\frac{1}{8}$	$\frac{3}{8}$
2	0	$\frac{1}{16}$	$\frac{3}{16}$

(a) Határozzuk meg az X és az Y valószínűségi változó peremeloszlását és várható értékét.

(b) Határozzuk meg $\mathbb{E}(XY)$-t.

(c) Igaz-e, hogy
$$\mathbb{P}(Y=1|X=2)=\mathbb{P}(Y=1|X=1)?$$
És az igaz-e, hogy
$$\mathbb{P}(X=2|Y=2)=\mathbb{P}(X=2|Y=1)?$$

(d) Független-e X és Y?

Minden választ indokoljunk! A részfeladatokat bármilyen sorrendben megoldhatjuk.

Megoldás.
(a) (7 pont)
(2 pont) X peremeloszlását a megfelelő oszlopösszegek adják: $\mathbb{P}(X=0)=1/4$, $\mathbb{P}(X=1)=\frac{3}{16}$, $\mathbb{P}(X=2)=\frac{9}{16}$, ez OK akár indoklás nélkül is, ha az eredmény jó.
(2 pont) Hasonlóan Y peremeloszlását a megfelelő oszlopösszegek adják, $\mathbb{P}(Y=0)=\mathbb{P}(Y=2)=1/4$, $\mathbb{P}(Y=1)=1/2$.
Ha odaírja, hogy hogy jönnek ki a peremeloszlások, de az eredmények nem jók, akkor max. 1–1 pont.
Ha nem indokol és egy adott peremeloszlásból csak 2 érték jó, akkor a 2-ből 1 pont. Ha max. 1 érték jó, akkor 0 pont.
Ha X és Y peremeloszlását felcseréli, akkor max. 2 pont.
(1 pont) arra, hogy a várható értékről **legalább az egyik val. változó esetében** valahogy leírja, hogy hogyan számolja ki, ott van a képlet általánosan vagy a konkrét számokkal (pl. $\mathbb{E}(X)=(0\cdot\frac{1}{4})+1\cdot\frac{3}{16}+2\cdot\frac{9}{16}$ vagy hasonlóan $\mathbb{E}(Y)=1\cdot\frac{1}{2}+2\cdot\frac{1}{4}=1$).
(1 pont) $\mathbb{E}(X)=\frac{21}{16}$ (itt már csak a végeredményre jár a pont, tizedestört alakban is jó: 1,3125),
(1 pont) $\mathbb{E}(Y)=1$.

(b) (6 pont)
1. megoldás: XY eloszlásával számolva.
(1 pont) $\mathrm{Ran}(XY)=\{0,1,2,4\}$ (ha olyan halmazt ír, ami ezt a négy elemet tartalmazza és még esetleg néhány olyan elemet, amik előállnak a $0,1,2$ számok közül kettő szorzataként, azt is el lehet fogadni). Elég, ha ez a későbbiekben impliciten derül ki.
(1 pont) $\mathbb{P}(XY=1)=\mathbb{P}(X=1,Y=1)=1/8$,
(1 pont) $\mathbb{P}(XY=2)=\mathbb{P}(X=1,Y=2)+\mathbb{P}(X=2,Y=1)=1/16+3/8=7/16$,
(1 pont) $\mathbb{P}(XY=4)=\mathbb{P}(X=2,Y=2)=3/16$ (0 pont) $\mathbb{P}(XY=0)=1/4$.
(1 pont) Így $\mathbb{E}(XY)=1\cdot 1/8+2\cdot 7/16+4\cdot 3/16$
(1 pont) $=\frac{28}{16}=\frac{7}{4}$.
Ha $\mathbb{P}(XY=0)$ jó, de semmi másra nem járna pont a (b)-ben, akkor erre lehet adni 1 pontot.
2. megoldás: eloszlástranszformáció-szerűen számolva.
(1 pont) Megadja az $\mathbb{E}(XY)=\sum_{k=0}^{2}\sum_{\ell=0}^{2}k\ell\,\mathbb{P}(X=k,Y=\ell)$ képletet vagy valami hasonlót; a lehetetlen (k,ℓ) párokat persze nem kötelező belevenni. Elég, ha impliciten derül ki, hogy ezt a képletet használja (ha a következő konkrét képlet jó, akkor ez a pont is jár).
(3 pont) $\mathbb{E}(XY)=1\cdot\mathbb{P}(X=Y=1)+2\cdot\mathbb{P}(X=2,Y=1)+2\cdot\mathbb{P}(X=1,Y=2)+4\,\mathbb{P}(X=$

$2, Y = 2$) (a 0-s szorzatokat is oda lehet írni, de nem kötelező)
(1 pont) $= 1/8 + 2 \cdot 3/8 + 2 \cdot 1/16 + 4 \cdot 3/16$
(1 pont) $= \frac{7}{4}$. (A $\frac{7}{4}$ jó nem egyszerűsített alakban és tizedestört alakban is.)

(c) (5 pont)
Önmagában arra, hogy igazak az állítások, nem jár pont.
(1 pont) ha azt írja, hogy $\mathbb{P}(Y = 1|X = 2) = \mathbb{P}(Y = 1|X = 1)$, és van valami indoklás, ami ha nem is teljes, de helyes (pl. legalább megjelenik a feltételes valószínűség definíciója helyesen).
(1 pont) $\mathbb{P}(Y = 1|X = 2) = \frac{\mathbb{P}(Y=1,X=2)}{\mathbb{P}(X=2)}$ és $\mathbb{P}(Y = 1|X = 1) = \frac{\mathbb{P}(Y=1,X=1)}{\mathbb{P}(X=1)}$ (jó metszettel leírva is),
(1 pont) ahol $\frac{\mathbb{P}(Y=1,X=1)}{\mathbb{P}(X=1)} = \frac{1/8}{3/16} = \frac{2}{3}$ és $\frac{\mathbb{P}(Y=1,X=2)}{\mathbb{P}(X=2)} = \frac{3/8}{9/16} = \frac{2}{3}$, tehát tényleg egyenlőek.
(1 pont) ha $\mathbb{P}(X = 2|Y = 2) = \mathbb{P}(X = 2|Y = 1)$-re is azt mondja, hogy igaz, és van valami (legalább részben) helyes indoklás rá, pl. hogy hasonlóan számol, mint az előző feladatban. Itt a feltételes valószínűség definícióját nem kell még egyszer leírni,
(1 pont) $\mathbb{P}(X = 2|Y = 2) = \frac{\mathbb{P}(X=2,Y=2)}{\mathbb{P}(Y=2)} = \frac{3/16}{1/4} = 3/4$ és $\mathbb{P}(X = 2|Y = 1) = \frac{\mathbb{P}(X=2,Y=1)}{\mathbb{P}(Y=1)} = \frac{3/8}{1/2} = 3/4$, tehát tényleg egyenlőek.
Ha jó eseményekkel számol, de nem jók a részletszámítások, akkor lehet részpontot adni (a két különböző számolásra járó 2 pontból 1-et).
(d) (2 pont)
(2 pont) Nem függetlenek, mert például $\mathbb{P}(X = 1, Y = 0) \neq \mathbb{P}(X = 1)\,\mathbb{P}(Y = 0)$. (Az utóbbi szorzat értékét nem kell kiszámolni.)
Vagy: nem függetlenek, mert a korábbi eredmények alapján $\mathbb{E}(XY) \neq \mathbb{E}(X)\,\mathbb{E}(Y)$.
Indoklás nélkül a (d)-re nem jár pont (és rossz indoklással sem). Részpontot legfeljebb kisebb szintaktikai hibák esetén adunk.

5. Egy nógrádi kistelepülésen az utóbbi években néha medvét észlelnek, egyszerre mindig csak egy medve érkezik magányosan. Egy adott időpontban elkezdjük megfigyelni, hogy mikor érkezik a következő medve, nevezzük ezt a $t = 0$ időpontnak. Jelölje X az ettől az időponttól a következő medve észleléséig eltelő időt (évben mérve). A megfigyelések szerint X folytonos, örökifjú eloszlást követ, 5 év várható értékkel.

   (a) Határozzuk meg X szórását.

   (b) Határozzuk meg X sűrűségfüggvényét a $t = 1$ helyen és X eloszlásfüggvényét a $t = -1$ helyen. *(Segít a nevezetes eloszlás táblázat.)*

   (c) Mennyi a valószínűsége, hogy a $t = 0$ időponttól kezdve a következő medve észleléséig több, mint 8 évig kell várni, feltéve, hogy több, mint 6 évig kell rá várni?

Megoldás. (a) (6 pont)
(1 pont) X exponenciális eloszlású,
(1 pont) mert ha egy eloszlás folytonos és örökifjú, akkor csak exponenciális lehet (valami olyasmi is elég, hogy „folytonos és örökifjú ⇒ exponenciális),
(0 pont) és a szóráshoz az eloszlás λ paraméterét kell meghatározni.
(1 pont) $\mathbb{E}(X) = \frac{1}{\lambda} = 5$,
(1 pont) tehát $\lambda = \frac{1}{5}$.
(1 pont) Így $\mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}$ (helyes szórásnégyzet esetén elég, ha $\sqrt{25}$-öt ír, de hibás szórásnégyzet esetén csak akkor jár részpont, ha odaírja a képletet általánosan is).
(1 pont) $= \frac{1}{\lambda} = 5$.
(b) (4 pont)
(1 pont) $f_X(1) = (1/5) \cdot e^{-1/5 \cdot 1}$ (a nevezetes eloszlás táblázat alapján, ezért nem kell indokolni – de ehhez a ponthoz még elég az is, ha általános λ-ra és t-re írja be a képletet)
(1 pont) $\approx 0{,}1637$.

(1 pont) $F_X(-1) = \mathbb{P}(X < -1)$
(1 pont) $= 0$ (mert $\mathrm{Ran}(X) = \mathbb{R}^+$ benne van a táblázatban, tehát nem kell indokolni).
Az is jó, ha nem írja fel az $F_X(-1)$ jelentését, de odaírja, hogy az X értékkészletén kívül van a -1.
Az is jó, ha azt írja, hogy $F_X(-1) = \int_{-\infty}^{-1} f_X(x)\mathrm{d}x$, és megállapítja, hogy ez az integrál 0, mert a konstans 0 függvényt integrálja. Ha 0-nál kisebb vagy 1-nél nagyobb számot ír be numerikus végeredménynek, akkor az eloszlásfüggvényre semmiképpen nem jár részpont, még akkor sem, ha $\mathbb{P}(X < -1)$ szerepel!
(c) (10 pont)
(2 pont) A keresett valószínűség $\mathbb{P}(X > 8|X > 6)$,
(2+1 pont) ami $\mathbb{P}(X > 2)$-vel egyenlő az örökifjúság miatt (nem kell az exponenciális eloszlásra hivatkozni, mert az örökifjúság a feladatban szerepelt, így ez járhat akkor is, ha nem jön rá, hogy exponenciális az eloszlás).
(1 pont) $\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 2) = 1 - \mathbb{P}(X < 2)$ (nem baj, ha a középső lépés hiányzik, a folytonosságra majd a következő feladatban jár pont)
(1 pont) mert X folytonos [val. változó] (vagy: mert X eloszlásfüggvénye folytonos, vagy: mert X-nek van sűrűségfüggvénye).
(2 pont) $1-\mathbb{P}(X < 2) = 1-F_X(2) = e^{-1/5\cdot 2}$ (nem kell, hogy itt megjelenjen az F_X kifejezés, de indoklásként legalább $e^{-\lambda x}$ szerepeljen általánosan vagy azonosíthatóan konkrét λ-ra és x-re)
(1 pont) $\approx 0{,}6703$.

6. * Egy X folytonos valószínűségi változó sűrűségfüggvénye

$$f_X(x) = \begin{cases} e^x, & \text{ha } x < a, \\ 0, & \text{egyébként,} \end{cases}$$

ahol $a \in \mathbb{R}$ egy alkalmasan választott (nem feltétlenül pozitív) valós szám (és $e \approx 2{,}71\ldots$).

   (a) Határozzuk meg a értékét.

   (b) Jelölje F_X az X valószínűségi változó eloszlásfüggvényét. Mely $x \in \mathbb{R}$ pontokra igaz, hogy $f_X(x) = F_X(x)$?

   (c) Határozzuk meg az $Y = e^X$ valószínűségi változó eloszlásfüggvényét és várható értékét.

Megoldás.
(a) (6 pont)
(1 pont) Ahhoz, hogy f_X sűrűségfüggvény legyen, az kell, hogy nemnegatív legyen (ez itt nem probléma, mert e^x minden $x \in \mathbb{R}$ esetén pozitív, így ezt elég egyszer megemlíteni és nem kell többször ellenőrizni)
(1 pont) és az integrálja $-\infty$-től ∞-ig 1 legyen.
(1 pont) Tehát annak kell teljesülnie, hogy $1 = \int_{-\infty}^{a} e^x \mathrm{d}x$
(1 pont) $[e^x]_{-\infty}^{a}$
(1 pont) $= e^a - 0 = e^a$ (a -0 hiányáért nem kell levonni pontot).
(1 pont) Tehát $a = 0$.

(b) (6 pont)
(1 pont) $F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t$ (mert X folytonos val. változó – ezt nem muszáj odaírni).
(1 pont) Ha $x \leq 0$, akkor $F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t = \int_{-\infty}^{x} e^t \mathrm{d}t$ (ezért az 1 pontért elég a két integrálból az egyik)
(1 pont) $= [e^t]_{-\infty}^{x} = e^x$.
(1 pont) Tehát $f_X(x) = F_X(x)$ teljesül, ha $x < 0$,
(1 pont) és $F_X(x) \neq f_X(x)$, ha $x = 0$ (mert $f_X(0) = 0$ a feladat szerint – ezt nem kell odaírni. Az is jó, ha az $x = 0$-t a következő esetnél kezeli).

(1 pont) Ha $x > 0$, akkor $F_X(x) = 1$ (ez OK további indoklás nélkül is), tehát $F_X(x) \neq f_X(x)$. (Itt részpontot nem lehet szerezni.)
(0 pont) Így $f_X(x) = F_X(x)$ pontosan akkor teljesül, ha $x < 0$.

(c) (8 pont)
(1 pont) $\mathrm{Ran}(Y) = (0,1)$, ezért $F_Y(x) = 0$, ha $x \leq 0$ és $F_Y(x) = 1$, ha $x \geq 1$ (az értékkészletet nem muszáj megemlíteni, de csak akkor jár az 1 pont, ha az $x \leq 0$ és az $x \geq 1$ eset is szerepel).
(1 pont) Ha $0 < x < 1$, akkor $F_Y(x) = \mathbb{P}(Y < x) = \mathbb{P}(e^X < x)$ (a $\mathbb{P}(Y < x)$ elhagyható)
(1 pont) $\mathbb{P}(X < \ln x)$
(1 pont) az exponenciális függvény/logaritmusfüggvény szigorú monotonitása miatt,
(1 pont) ami $F_X(\ln x) = e^{\ln x} = x$ (OK az F_X jelölés nélkül is).
(2 pont) Innen látjuk, hogy Y egyenletes eloszlású a $(0,1)$ intervallumon. (További indoklás nem kell, mert benne van a nevezetes eloszlás táblázatban.)
(1 pont) Ezért $\mathbb{E}(Y) = \dfrac{1}{2}$. (Ehhez sem kell indoklás, még ananyi sem, hogy ez $\frac{a+b}{2}$.)
Ha az $x = 0$-t vagy az $x = 1$-et a középső szakaszhoz sorolja és úgy kezeli, az sem probléma, sőt az sem, ha $x = 0$-ra és $x = 1$-re két (egyenlő) F_Y-értéket ad meg, hiszen F_Y folytonos függvény. De ha valamelyik pontban nem ad meg értéket, akkor (összesen) -1 pt.

Alternatív megoldások a várható érték kiszámítására:

- transzformált várható értékés képlettel: $\mathbb{E}(Y) = \int_{-\infty}^{\infty} e^x f_X(x)\mathrm{d}x = \int_{-\infty}^{0} e^{2x}\mathrm{d}x = [e^{2x}/2]_{-\infty}^{0} = \dfrac{1}{2}$. Ez jó az eloszlásfüggvény megadása nélkül is 3 pontért,
- vagy meghatározza Y sűrűségfüggvényét (ami $\mathrm{U}(0;1)$ sűrfv., a fenti F_Y deriváltja, illetve a 0-ban és 1-ben 0) és az alapján $\mathbb{E}(Y) = \int_{-\infty}^{\infty} x f_Y(x)\mathrm{d}x = \int_0^1 x\mathrm{d}x = \ldots = \dfrac{1}{2}$. Ez tovább tart, mint a többi megoldás, de sajnos szintén csak 2 pontot ér összesen, amiből maximum 1 jár a sűrűségfüggvény meghatározására. Az is jó, ha a sűrűségfüggvény alapján jön rá, hogy $Y \sim \mathrm{U}(0;1)$, ha az eloszlásfüggvény alapján nem jött rá, és ezután olvassa ki a táblázatból a várható értéket/hivatkozik az ismert képletre.

Eloszlás neve	Jelölés	$\mathrm{Ran}(X)$	$\mathbb{P}(X = k)$ vagy $F_X(t)$	$f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
indikátor	$\mathbf{1}(p)$	$\{0,1\}$	$p,\ 1-p$		p	$p(1-p)$
binomiális	$\mathrm{B}(n;p)$	$\{0,1,...,n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$		np	$np(1-p)$
Poisson	$\mathrm{Pois}(\lambda)$	$\{0,1,...\}$	$\frac{\lambda^k}{k!}e^{-\lambda}$		λ	λ
geometriai	$\mathrm{Geo}(p)$	$\{1,2,...\}$	$(1-p)^{k-1}p$		$\frac{1}{p}$	$\frac{1-p}{p^2}$
egyenletes	$\mathrm{U}(a;b)$	$(a;b)$	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponenciális	$\mathrm{Exp}(\lambda)$	$\mathbb{R}^+$	$1-e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Tudnivalók: A zárthelyi időtartama 90 perc. Számológépet lehet használni. A számszerű megoldásokat 4 értékes jegyre kerekítsük. A teljes pontszám eléréséhez a megoldás menete is szükséges, beleértve az egyes lépéseknél felhasznált tulajdonságok és tételek jelzését. Hivatkozni csak az előadáson elhangzottakra lehet. A zárthelyi első 30 percében nem lehet a termet elhagyni.

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 \ 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: $\mathbb{E}(ST) = \sum_{k \in \operatorname{Ran}(S)} \sum_{\ell \in \operatorname{Ran}(T)} k\ell\, \mathbb{P}(S = k, T = \ell) = 2 \cdot 2 \cdot \mathbb{P}(S = T = 2)$.)
(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 $X_1, \dots, X_n$ 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 $x \in \{2, 3, 4, \dots\}$, $0 < \vartheta < 1$ pedig egy ismeretlen paraméter.

   (a) Határozzuk meg az $(X_1, \dots, X_n)$ fae. minta egy adott $(x_1, \dots, x_n)$ realizációja esetén a likelihood-függvény $L_\vartheta(x_1, \dots, x_n)$ 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) $L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n p_\vartheta(x_i)$ (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 $l_\vartheta(x_1, \dots, x_n) = \ln L_\vartheta(x_1, \dots, x_n)$ (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) $\frac{\partial}{\partial\vartheta} l_\vartheta(x_1, \dots, x_n) = \frac{2n}{\vartheta} - \sum_{i=1}^n \frac{x_i-2}{1-\vartheta}$ (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) $\frac{\partial^2}{\partial\vartheta^2} l_\vartheta(x_1, \dots, x_n) = -\frac{2n}{\vartheta^2} - \sum_{i=1}^n \frac{x_i-2}{(1-\vartheta)^2}$
(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 $\vartheta_*(x_1, \dots, x_n) = \frac{2n}{\sum_{i=1}^n x_i} = \frac{2}{\overline{x_n}}$ (megint csak elég az egyik a kettő közül)
(1 pont) és így a maximum likelihood becslés $\vartheta_*(X_1, \dots, X_n) = \frac{2n}{\sum_{i=1}^n X_i} = \frac{2}{\overline{X_n}}$ (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 $\vartheta_*(X_1, \dots, X_n)$ 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

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

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 $\frac{1}{x} \ge 1$, vagyis $x \le 1$, 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 $y\le 1/2$, 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) $f_Y(y)=\int_0^{1/y}\frac{x^2}{2}\mathrm{d}x=\left[x^3/6\right]_0^{1/y}=\frac{1}{6y^3}$, ha $y>1/2$ (Ebből 1 pont jár a számolásra és 1 a végeredményre.)
(1+1 pont) és $f_Y(y)=\int_0^2\frac{x^2}{2}\mathrm{d}x=\left[x^3/6\right]_0^2=\frac{4}{3}$, ha $0<y<1/2$. (Hasonlóan.)
(0 pont) Ha $y\le 0$, akkor $f_Y(y)=0$.
(2 pont) Tehát $\mathbb{P}(Y<1)=\int_0^{1/2}\frac{4}{3}\mathrm{d}y+\int_{1/2}^1\frac{1}{6y^3}\mathrm{d}y=\frac{2}{3}+\left[-\frac{1}{12y^2}\right]_{y=1/2}^1=\frac{2}{3}-\frac{1}{12}+\frac{1}{3}=\frac{11}{12}$.

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).

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. Legyen X és Y két valószínűségi változó pozitív és véges szórással.

   (a) Hogyan jelöljük X és Y korrelációját és hogyan definiáljuk ezt a korrelációt X és Y kovarianciája és szórásai segítségével?

   (b) Amennyiben X és Y függetlenek, milyen összefüggés áll fenn ezen két valószínűségi változó szórásnégyzetei és az összegük szórásnégyzete között?

   Megjegyzés: A szórásra, szórásnégyzetre és kovarianciára használhatjuk az előadásról ismert jelöléseket, ezen jelöléseket és a szórás(négyzet) és kovariancia jelentését nem kell megmagyarázni. A (b) részben a lehető legegyszerűbb helyes képletet adjuk meg.
   Megoldás.
   (a)
   (3 pont) Jelölés: $\operatorname{corr}(X, Y)$
   (7 pont) $= \frac{\operatorname{cov}(X,Y)}{\mathbb{D}(X)\mathbb{D}(Y)}$.
   (b)
   (10 pont) $\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$.

   Pontozási instrukciók:
   (a)
   Ha valahogy szövegesen megfogalmazza a képletet, az is jó, ilyen esetben nem kellenek a cov és $\mathbb{D}$ jelölések. De ha a szórás helyett a szórásnégyzet gyökét írja oda, akkor max. 5 pont, mert a kérdésben szórás volt. Ha nem az előadáson szereplő, de értelmes jelölést használ, például Var vagy σ, akkor max. 4 pont, attól függően, hogy szemantikailag és szintaktikailag mennyire értelmes, amit ír.
   Ha nem jó a képlet, de a $\operatorname{cov}(X, Y)$ jelölése jó, arra jár max. 1 pont, és a $\mathbb{D}(X)$ vagy $\mathbb{D}(Y)$ jelölésre (elég az egyik) is jár max. 1 pont.

(b)
A megoldás szövegesen is elfogadható, pl. X és Y összegének szórásnégyzete X és Y szórásnégyzeteinek összege. Ha valahogy úgy fogalmazza meg, hogy X és Y közül nem szerepel benne mindkettő, akkor max. 9 pont.
Ha odaírja a $+2\mathrm{cov}(X,Y)$ tagot is, akkor max. 8 pont.
Kisebb szintaktikai hibák, nyilvánvalóan csak sajtóhibának számító elírások esetén lehet részpontot adni. Ha kihagyja a bal oldalt, akkor max. 6 pont. Ha kihagyja a jobb oldalt, akkor 0 pont (+ lásd a következő sort).
Ha nem jó a képlet, de a szórásnégyzet jelölése helyes (nem csak annyi, hogy $\mathbb{D}^2$, hanem $\mathbb{D}^2(X), \mathbb{D}^2(Y), \mathbb{D}^2(X+Y)$ közül legalább az egyik ott van), arra jár max. 1 pont.

2. Pista a sarki péknél minden reggel egymástól függetlenül 1/3 valószínűséggel sós kiflit vásárol, amely 150 forintba kerül, 1/3 valószínűséggel pogácsát, amely 300 forintba, és 1/3 valószínűséggel kakaós csigát, amely 450 forintba kerül. Minden reggel pontosan egy pékárut vásárol. 40 egymást követő napon át megvizsgáljuk, hogy Pista mennyit költ az egyes reggeleken a péknél, jelölje X_i azt, hogy hány forintot költ Pista az i-edik reggelen, ahol $i = 1, \dots, 40$.

   (a) Határozzuk meg X_1 várható értékét és szórását.

   (b) Jelölje X a Pista által a péknél a 40 nap alatt elköltött pénz összegét. Közelítsük alkalmas approximációval a $\mathbb{P}(X \geq 12\,500)$ valószínűséget. (A valószínűség pontos értékét nem kell meghatározni.)

Megoldás.
(a) (8 pont)
(1+1 pont) $\mathbb{E}(X_1) = \frac{1}{3}(150+300+450) = 300$ (1 pont a számolásra, ahol elég csak az általános vagy csak a konkrét képlet, de ha hiányzik az általános képlet és rossz a behelyettesítés, akkor ez az 1 pont nem jár; 1 pont a végeredményre).
(1+1 pont) $\mathbb{D}(X_1) = \sqrt{\mathbb{D}^2(X_1)} = \sqrt{\mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2}$ (a két lépés összevonható),
(1+1 pont) ahol $\mathbb{E}(X_1^2) = \frac{1}{3}(150^2 + 300^2 + 450^2) = 105\,000$ (pontozás az első momentumhoz hasonlóan),
(1+1 pont) tehát $\mathbb{D}(X) = \sqrt{105\,000 - 300^2} = \sqrt{15000} \approx 122{,}4745$ (itt már további számolás nem kell, ha az összes részeredmény jó, viszont ha hiányzik a numerikus végeredmény, akkor max. 1 pont. Minimális kerekítési hibáért itt ne vonjunk le pontot).

(b) (12 pont)
(4 pont) Sztenderdizálunk (ezt szöveggel nem kell odaírni):

$$\mathbb{P}(X \geq 12\,500) = \mathbb{P}\left(\frac{X - 40\,\mathbb{E}(X_1)}{\sqrt{40}\,\mathbb{D}(X_1)} \geq \frac{12\,500 - 40\,\mathbb{D}(X_1)}{\sqrt{40}\cdot\mathbb{D}(X_1)}\right) = \mathbb{P}\left(\frac{X - 40\cdot 300}{\sqrt{40}\cdot 122{,}4745)} \geq \frac{12\,500 - 40\cdot 300}{\sqrt{40}\cdot 122{,}4745}\right).$$

Ennek a 4 pontnak a megoszlása: 2 pont, ha tudja, hogy hogyan kell sztenderdizálni a val. változót (általános $\mathbb{E}(X_1)$-re és $\mathbb{D}^2(X_1)$-re), 1 pont, ha jól helyettesíti be a képletbe az adatokat és 1 pont, ha ugyanezeket a transzformációkat a jobb oldalakra is megfelelően végrehajtja.
40 helyett az általános képlet felírásánál természetesen jó az n is, ha később kiderül, hogy $n = 40$.
A két lépés összevonható pontlevonás nélkül, ha azonosítható, hogy az előbbi képletet használja (ide értve azt is, hogy a $\sqrt{40}$ nem csak numerikus értékkel szerepel!)
(1 pont) $= \mathbb{P}(\frac{X - 12\,000}{774{,}5967} \geq 0{,}6455)$ (ez az előző lépéssel összevonható, de csak akkor, ha azt nem vonta össze az azelőttivel)
(1 pont) ami a centrális határeloszlás-tétel (elég: CHT; de Moivre–Laplace viszont nem jó!) miatt
(1 pont) $\mathbb{P}(Y \geq 0{,}6455)$-tel közelíthető, ahol $Y \sim \mathrm{N}(0;1)$.
(1 pont) arra, hogy valahol hivatkozik arra, hogy az X független, azonos eloszlású valószínűségi

változók összege (konkrétan az X_i-ké, de ezt nem muszáj explicite kimondani. Ha a függetlenség megvan, csak az azonos eloszlás hiányzik, és egyébként jól csinálja, akkor ne vonjuk le ezt a pontot). Emiatt helyes a sztenderdizálás (emiatt $40\,\mathbb{E}(X_1)$ az X val. változó várható értéke és $\sqrt{40}\,\mathbb{D}(X_1)$ a szórása) és emiatt alkalmazható a CHT, de elég, ha egyszer hivatkozik rá.
(1 pont) Tehát a keresett valószínűség közelítőleg $\mathbb{P}(Y \geq 0{,}6455) = 1 - \mathbb{P}(Y < 0{,}6455) = 1 - \Phi(0{,}6455)$ (a kettő közül elég az egyik)
(1 pont) $\approx 1 - \Phi(0{,}65)$
(1 pont) $\approx 1 - 0{,}7422$
(1 pont) $= 0{,}2578$.

3. Egy budapesti lakótelepi lakásban időnként előfordulnak áramszünetek. Tudjuk, hogy 3/4 annak a valószínűsége, hogy egy adott évben előfordul legalább egy áramszünet. Tegyük fel, hogy az egyes években egymástól függetlenül fordul elő vagy nem fordul elő áramszünet.

   (a) Jelölje Y tíz egymást követő év közül azon évek számát, amelyekben előfordul legalább egy áramszünet. Mennyi Y várható értéke?

   (b) Mi a valószínűsége, hogy 10 egymást követő évből legalább 8 évben lesz legalább egy áramszünet?

   (c) Egy adott pillanatban elkezdjük mérni, hogy mennyi idő telik el a következő áramszünetig. Jelölje ezt a várakozási időt (évben mérve) X. Feltehetjük, hogy X eloszlása folytonos és örökifjú, valamint a fentiek alapján $\mathbb{P}(X \leq 1) = 3/4$. Határozzuk meg X szórását.

   (d) Feltéve, hogy 4 évnél többet kell várni az első áramszünetre, mi a valószínűsége, hogy 5 évnél is többet kell várni rá?

Megoldás.
(a) (3 pont)
(1 pont) Y binomiális eloszlású
(1 pont) $n = 10$ és $p = 3/4$ paraméterekkel (nem kell indokolni az eddigieket, csak kimondani).
(1 pont) Ezért $\mathbb{E}(Y) = np = 10 \cdot \frac{3}{4} = 7{,}5$. (Itt már elég a végeredmény, ha megvannak a paraméterek; np szerepel a táblázatban.)
(b) (5 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(Y \geq 8)$
(1 pont) $= \mathbb{P}(Y = 8) + \mathbb{P}(Y = 9) + \mathbb{P}(Y = 10)$
(2 pont) $= \binom{10}{8} \cdot (3/4)^8 \cdot (1/4)^2 + \binom{10}{9} \cdot (3/4)^9 \cdot (1/4) + \binom{10}{10}(3/4)^{10}$ ($\binom{10}{k}$ helyett jó a $\binom{10}{n-k}$ is, a $\binom{10}{10}$ elhagyható és a $\binom{10}{9}$ helyett elég az, hogy 10)
(1 pont) $\approx 0{,}5256$. (Nem kell részletezni, de az első tag numerikus értéke 0,2816, a másodiké 0,1877, a harmadiké pedig 0,0563.)
(c) (9 pont)
(1 pont) X exponenciális eloszlású,
(1 pont) mert folytonos és örökifjú eloszlás csak exponenciális lehet (elég annyi, hogy folytonos és örökifjú $\Rightarrow$ exponenciális [eloszlású]).
(0 pont) A kérdések megválaszolásához az eloszlás λ paraméterét kell meghatározni.
(1+1 pont) $\mathbb{P}(X \leq 1) = \mathbb{P}(X < 1) = 1 - e^{-\lambda}$, mert X (vagy: X eloszlásfüggvénye) folytonos.
Megjegyzés javítóknak: Ha csak annyit ír, hogy $\mathbb{P}(X \leq 1) = 1 - e^{-\lambda}$, esetleg ugyanezt eloszlásfüggvénnyel kifejezve, akkor a 2 pontból 1 jár. A második pont akkor jár, ha hivatkozik a folytonosságra, viszont ebben az esetben elég $\mathbb{P}(X \leq 1) = 1 - e^{-\lambda}$ is, $\mathbb{P}(X \leq 1) = \mathbb{P}(X < 1)$-et nem muszáj kiírni.
(1 pont) Tehát $1 - e^{-\lambda} = 3/4$, vagyis $e^{-\lambda} = 1/4$ (elég az utóbbi),
(1+1 pont) vagyis $-\lambda = \ln(1/4)$, azaz $\lambda = \ln 4 \approx 1{,}3863$. (1 pont a végeredményre, 1 a számolásra. Itt még elég az $\ln 4$ vagy $-\ln(1/4)$, a numerikus végeredmény csak a későbbiekben kell, viszont ha van helyes végeredmény, akkor az elég $\ln 4$ nélkül is.)

(1 pont) $\mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}$
(1 pont) $= \sqrt{\frac{1}{\lambda^2}} \approx 0{,}7213$ (az $1/\lambda$ vagy $1/\ln 4$ elég, ha impliciten derül ki, viszont kell a numerikus végeredmény).
(d) (3 pont)
(1 pont) A keresett (feltételes) valószínűség $\mathbb{P}(X > 5|X > 4)$
(1 pont) $= \mathbb{P}(X > 1)$ az örökifjúság miatt (ez a pont csak akkor jár, ha hivatkozik az örökifjúságra, de ha az eredmény jó, akkor a továbbiakra még lehet pontot kapni).
(1 pont) $= 1 - \mathbb{P}(X \leq 1) = \frac{1}{4}$.
(Ugye $\mathbb{P}(X \leq 1)$ értékét tudjuk a (c) rész szövegéből, ezért indoklásként elég a komplementer valószínűségre hivatkozni, a végeredmény és az indoklás együtt 1 pont. Ha már a (c)-re adott megoldásában is megjelent valahol, hogy $\mathbb{P}(X > 1) = 1/4$, akkor OK a végeredmény újabb indoklás nélkül is.
Természetesen azt is lehet mondani, hogy $\mathbb{P}(X \leq 1) = 1 - e^{-\lambda}$. Itt már ne vonjunk le pontot, ha nem hivatkozik a folytonosságra, sőt akkor se, ha csak azt írja, hogy $\mathbb{P}(X > 1) = 1 - \mathbb{P}(X < 1)$ vagy $\mathbb{P}(X > 1) = 1 - F_X(1)$ vagy $\mathbb{P}(X > 1) = e^{-\lambda}$.)

4. Az (X, Y) folytonos valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f_{X,Y}(x,y) = \begin{cases} c(x+y), & \text{ha } 1 < x < 2 \text{ és } 1 < y < 2, \\ 0, & \text{egyébként,} \end{cases}$$

ahol $c \in \mathbb{R}$ egy alkalmasan választott valós szám.

(a) Határozzuk meg c értékét.

(b) Határozzuk meg X sűrűségfüggvényét.

(c) Határozzuk meg Y várható értékét.

Megoldás.
(a) (7 pont)
(1 pont) $f_{X,Y}$ akkor lesz sűrűségfüggvény, ha nemnegatív (az is jó, ha azt írja, hogy c-nek kell nemnegatívnak lennie, és nem baj, ha nem ellenőrzi a végén még egyszer, hogy ez teljesül-e)
(1 pont) és $1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$
(1 pont) $= \int_1^2 \int_1^2 c(x+y)\mathrm{d}x\mathrm{d}y$ (az előző lépéssel összevonható, ha kiderül, hogy 1-nek kell lennie az integrálnak)
(1 pont) $= c\int_1^2 [x^2/2 + xy]_{x=1}^2 \mathrm{d}y = c\int_1^2 \frac{3}{2} + y \ \mathrm{d}y$
(1 pont) $= c[\frac{3}{2}y + \frac{y^2}{2}]_{y=1}^2$
(1 pont) $= c(\frac{3}{2} + \frac{3}{2}) = 3c$,
(1 pont) tehát $c = \frac{1}{3}$.
(b) (6 pont)
(1 pont) $x \in \mathbb{R}$ esetén $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$,
(1 pont) $\int_1^2 \frac{1}{3}(x+y)\mathrm{d}y$
(1 pont) $= \frac{1}{3}[xy + y^2/2]_{y=1}^2$
(1 pont) $= \frac{1}{3}(x + \frac{3}{2}) = \frac{1}{3}x + \frac{1}{2}$,
(1 pont) ha $x \in (1,2)$ (ha kihagyja, hogy a fenti képletek csak $x \in (1,2)$ esetén érvényesek, de egyébként jól számol, akkor ezt az 1 pontot vonjuk le),
(1 pont) és $f_X(x) = 0$ egyébként.
Megjegyzés csak javítóknak a pontozáshoz: Ha a határokat máshogy sorolja be a két szakaszhoz, azért ne vonjunk le pontot, de ha valamelyik határponto(ka)t nem sorolja be sehova, akkor vonjunk le egyet. Bármilyen nemnegatív, minden pontban definiált függvényt elfogadunk, ami a fenti f_X-szel majdnem mindenütt megegyezik.
(c) (7 pont)

1. megoldás.
(1 pont) Y-nak ugyanaz a sűrűségfüggvénye, mint X-nek
(ezért az 1 pontért elég, ha ez impliciten derül ki, viszont ha Y helyett X várható értékét kezdi el kiszámolni és Y-ra nincs kellő utalás, akkor a (c)-re max. 3 pont),
(1 pont) mert $f_{X,Y}$ szimmetrikus (x-ben és y-ban).
Természetesen az is jó, ha kiszámolja x szerinti integrálással és úgy hozza ki ugyanezt az eredményt. Akkor is max. 2 pont, de ha kihagyja, hogy $(1,2)$-n kívül 0 a sűrűségfüggvény, azért ne vonjunk le még egy pontot abban az esetben, ha $(1,2)$-n jó képletet ad meg.
(1+1 pont) $\mathbb{E}(Y) = \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y = \int_1^2 y(\frac{1}{3}y + \frac{1}{2})\mathrm{d}y$ (a két lépés összevonható)
(0 pont) $= \int_1^2 \frac{1}{3}y^2 + \frac{1}{2}y\mathrm{d}y$
(1 pont) $= [y^3/9 + y^2/4]_{y=1}^2$
(1 pont) $= \frac{8}{9} + \frac{4}{4} - \frac{1}{9} - \frac{1}{4}$
(1 pont) $= \frac{55}{36} (\approx 1{,}5278)$.
2. megoldás.
(1 pont) $\mathbb{E}(Y) = \int_{-\infty}^{\infty} y f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$
(1 pont) $= \int_1^2 \int_1^2 \frac{1}{3}xy + \frac{1}{3}y^2\mathrm{d}x\mathrm{d}y$
(1 pont) $= \int_1^2 [\frac{1}{6}x^2y + \frac{1}{3}y^2x]_{x=1}^2\mathrm{d}y$
(1 pont) $= \int_1^2 \frac{4}{6}y + \frac{2}{3}y^2 - \frac{1}{6}y - \frac{1}{3}y^2\mathrm{d}y = \int_1^2 \frac{1}{2}y + \frac{1}{3}y^2\mathrm{d}y$
(3 pont) ugyanúgy, mint az 1. megoldás utolsó 3 pontjánál. (Természetesen az is OK, ha először y és aztán x szerint integrál, pontozás analóg módon.)

5. Egy rövid menetidejű repülőjáraton a menetidő jó közelítéssel normális eloszlású $\sqrt{2}/100$ óra szórással. Hat alkalommal megmértük ezen a járaton a menetidőt és (órában mérve, tizedestört alakban) az alábbi eredményeket kaptuk:

$$x_1 = 0{,}97 \quad x_2 = 0{,}98 \quad x_3 = 0{,}99 \quad x_4 = 1{,}01 \quad x_5 = 1{,}00 \quad x_6 = 0{,}99.$$

Feltehetjük, hogy az egyes alkalmakkor mért menetidők egymástól függetlenek.
(*A szórás az eredeti vizsgafeladatban ismeretlen volt, ezt módosítottam.*)

(a) Adjunk a minta alapján 0,95 szintű konfidenciaintervallumot a mintaelemek várható értékére. Adjuk meg a minta korrigált empirikus szórását is.

(b) Döntsük el $\varepsilon = 0{,}02$ terjedelmű próbával, hogy a mintaelemek várható értéke megegyezik-e 1 órával. Adjuk meg H_0-t, H_1-et, a próbastatisztika értékét, azt, hogy mi alapján döntünk, és magát a döntést is.

(c) Adjuk meg (indoklással) a minta empirikus eloszlásfüggvényének értékét az $x = 0{,}995$ helyen.

(d) Adjuk meg a minta korrigált empirikus szórását.

Megoldás.
(a) (7 pont)
(1 pont) a konfidenciaintervallum (ismert szórás esetére vonatkozó) helyes általános képletére; ha jó konkrét paraméterekkel írja fel, az is jó.
(1 pont) $n = 6$ (ez a mintaelemszám),
(1 pont) $\overline{x_6} = \frac{1}{6}\sum_{i=1}^6 x_i = 0{,}99$ (csak 1 pont jár rá, ezért OK indoklás nélkül is, $\overline{x_n}$ jelöléssel is jó),
(1 pont) a szükséges kvantilis pedig $u_{\varepsilon/2} = u_{0{,}025}$
(1 pont) $= 1{,}96$.
(1+1 pont) A képletbe behelyettesítve a konfidenciaintervallum:

$$\left[\overline{x_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{x_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\right] = \left[0{,}99 - \frac{1{,}96 \cdot 0{,}0141}{\sqrt{6}}, 0{,}99 - \frac{1{,}96 \cdot 0{,}0141}{\sqrt{6}}\right] = [0{,}9786, 1{,}0014].$$

Az általános képletre itt nem jár pont. 1 pont arra, hogy behelyettesíti a képletbe a kapott értékeket, és 1 pont a numerikus végeredményre. Ha már mindent kiszámolt és utána csak a numerikus végeredményt adja meg, akkor ne vonjunk le pontot.
Megjegyzés: ha a $\frac{\sqrt{2}}{100}$ pontos értékkel számolunk, akkor is ez az eredmény jön ki 4 tizedesjegyre kerekítve.

(b) (7 pont)
Először a jó megoldást közlöm, a hibás alternatíva vagy próbatípus esetén járó részpontokat lásd lentebb.
(1 pont) $H_0\colon \mu = 1$ vs. $H_1\colon \mu \neq 1$,
(1 pont) tehát mivel a szórás ismeretlen, a kétoldali, egymintás t-próba képletét kell használni. (Ha ezt nem írja oda, de jól alkalmazza ezt a próbát, akkor ezt az 1 pontot ne vonjuk le.)
(1+1 pont) A próbastatisztika értéke $u(= u(x_1, \ldots, x_6)) = \frac{\overline{x_n}-\mu_0}{\sigma}\sqrt{n} = \frac{0{,}99-1}{0{,}0141}\sqrt{6} \approx -1{,}7372$. (1 pont arra, hogy jól behelyettesít a táblázatban megadott képletbe, ideértve azt is, hogy $\mu_0 = 1$. Természetesen jó n helyett 6-ot írni mindenhova. 1 pont pedig arra, hogy kiszámolja a helyes numerikus eredményt. Ha az utóbbi megvan és a szükséges mennyiségek értékei μ_0-t kivéve már az (a)-ban kiderültek, akkor adjuk meg az első pontot is, akkor is, ha nem írja oda, hogy $\mu_0 = 1$).
(1 pont) H_0-t akkor fogadjuk el/nem utasítjuk el, ha $|u| \leq u_{\varepsilon/2} = u_{0{,}01}$ (bármilyen, ezzel ekvivalens alak jó, az is, ha azt írja, hogy mikor utasítjuk el. Ne vonjunk le pontot, ha rossz helyre teszi az egyenlőséget.)
(1 pont) $= 2{,}33$.
(1 pont) Tehát H_0-t elfogadjuk (azaz azt mondjuk, hogy a mintaelemek várható értéke megegyezik 1 órával, illetve nincs elég bizonyítékunk ahhoz, hogy azt mondjuk hogy ez nem igaz).
Pontozás a (b)-ben hibás próbatípus, illetve alternatíva esetén:
Ha egyoldali u-próbát hajt végre $H_0\colon \mu \geq 1$ vs. $H_1\colon \mu < 1$ alternatívával, akkor max. 6 pont.
Ha egyoldali u-próbát hajt végre $H_0\colon \mu \leq 1$ vs. $H_1\colon \mu > 1$ alternatívával, az még rosszabb, mert az alacsony átlag alapján nem gyanakodhatunk arra, hogy a várható érték 1-nél nagyobb, így ekkor max. 5 pont.

(c) (3 pont)
(2 pont) Megadja az empirikus eloszlásfüggvény adott pontbeli értékének definícióját: $F_n^*(x)$ az x-nél kisebb mintaelemek száma, osztva n-nel. Jó szövegesen is, a $\frac{1}{n}|\{i \in \{1, \ldots, n\}\colon x_i < x\}|$ definícióval és a rendezett mintaelemekkel kifejezett ($F_n^*(x) = \frac{i}{n}$, ha $x_i^* < x \leq x_{i+1}^*$) definícióval is. Utóbbi esetben ne vonjunk le pontot, ha a határpontokban nem jó a definíció (hanem pl. jobbról folytonos) vagy ha hiányzik az $x < x_1^*$ és az $x \geq x_n^*$ eset, mert ezek itt nem relevánsak. Kisebb szintaktikai hibák esetén lehet 1 pontot adni a 2-ből.
Az is jó, ha eleve a konkrét n-re és x-re mondja el ugyanezt, de ha nincs általános képlet és a konkrét esetben rosszul írja fel, akkor 0 pont.
(1 pont) Esetünkben ez $F_6^*(0{,}995) = \frac{1}{6} \cdot 4 = \frac{2}{3}$. (Ezért az egy pontért elég a végeredmény is, ha van előtte valami indoklás vagy legalább minimális számolás. Tizedestört alakban és nem egyszerűsített $\frac{4}{6}$ alakban is jó.)

d) (3 pont)
(0 pont) $s_n^* = \sqrt{(s_n^*)^2}$
(1 pont) $= \sqrt{\frac{1}{5}\sum_{i=1}^{6}(x_i - 0{,}99)^2}$ (ez akkor is OK, ha esetleg csak $(s_n^*)^2$-et számolja ki és s_n^*-ot nem, viszont az általános képlet puszta lemásolására nem jár pont, csak ha $n-1$, n és $\overline{x_n}$ képletét is behelyettesíti)
(2 pont) $= \sqrt{\frac{1}{5}(0{,}02^2 + 0{,}01^2 + 0 + 0{,}02^2 + 0{,}01^2 + 0)} = \sqrt{\frac{1}{5} \cdot \frac{4+1+4+1}{10\ 000}} = \sqrt{1/5000} = \frac{\sqrt{2}}{100} = 0{,}0141$ (ebből 1 pont arra, hogy van valami részletszámolás, ahol a különbségek négyzetei valamilyen értelmes formában megjelennek, nem kell ennyire részletesen és a 0-kat el lehet hagyni,

1 pont pedig a numerikus végeredményre jár),

6. * *(A többdimenziós normális eloszlás fogalmának hiányában átdolgozott feladat.)*
Legyenek Z és W független, standard normális eloszlású változók és legyen $X := \sqrt{5}Z+2W-1$, $Y := \sqrt{5}Z - 2W + 2$.

   (a) Határozzuk meg az $(X, Y)^T$ valószínűségi vektorváltozó várható érték vektorát és kovarianciamátrixát. *Ez az eredeti vizsgafeladatnak nem volt része, de az előadás jelenlegi anyaga alapján is megoldható és nem 6-os nehézségű, valamint kell a (b) és (c) kitűzéséhez.*

   (b) Legyen $U = X + Y$ és $V = X - Y$. Határozzuk meg az $\begin{pmatrix} U \\ V \end{pmatrix}$ valószínűségi vektorváltozó várható érték vektorát és kovarianciamátrixát.

   (c) Ez alapján adjuk meg az $\mathbb{E}(V|U)$ regressziót.

Megoldás.
(a) (pontozás nélkül)
$\mathbb{E}(X) = \mathbb{E}(\sqrt{5}Z + 2W - 1) = \sqrt{5}\,\mathbb{E}(Z) + 2\,\mathbb{E}(W) - 1 = -1$ a várható érték linearitása miatt, hasonlóan $\mathbb{E}(Y) = \mathbb{E}(\sqrt{5}Z - 2W + 2) = 2$, tehát a várható érték vektor $(-1, 2)^T$.
$\mathbb{D}^2(X) = \mathbb{D}^2(\sqrt{5}Z+2W-1) = \mathbb{D}^2(\sqrt{5}Z+2W) = \mathbb{D}^2(\sqrt{5}Z)+\mathbb{D}^2(2W) = 5\,\mathbb{D}^2(Z)+4\,\mathbb{D}^2(W) = 9$, ahol az összegre bontásnál W és Z függetlenségét használtuk.
Hasonló indoklással: $\mathbb{D}^2(Y) = \mathbb{D}^2(\sqrt{5}Z - 2 + 2) = \mathbb{D}^2(\sqrt{5}Z - 2W) = \mathbb{D}^2(\sqrt{5}Z) + \mathbb{D}^2(-2W) = 5\,\mathbb{D}^2(Z) + 4\,\mathbb{D}^2(W) = 9$.
Továbbá $\mathrm{cov}(X, Y) = \mathrm{cov}(\sqrt{5}Z + 2W - 1, \sqrt{5}Z - 2W + 2) = \mathrm{cov}(\sqrt{5}Z + 2W, \sqrt{5}Z - 2W) = \sqrt{5}\cdot\sqrt{5}\mathrm{cov}(Z, Z)+\sqrt{5}\cdot(-2)\cdot\mathrm{cov}(Z, W)+2\cdot\sqrt{5}\cdot\mathrm{cov}(W, Z)+2\cdot(-2)\cdot\mathrm{cov}(W, W) = 5+0+0-4 = 1$, ahol a kovariancia eltolás-invarianciáját („additív konstans nem számít"), bilinearitását és az utolsó lépésben Z és W függetlenségét használtuk.
Így a kovarianciamátrix $\begin{pmatrix} 9 & 1 \\ 1 & 9 \end{pmatrix}$.
(b) (12 pont)
(1 pont) $\mathbb{E}(U) = \mathbb{E}(X)+\mathbb{E}(Y)$ (ha kimondja, hogy a várható érték linearitását használja, akkor egyből mehet a $-1 + 2$ anélkül, hogy odaírná, hogy ez $\mathbb{E}(X) + \mathbb{E}(Y)$-nal egyenlő)
(1 pont) $= -1 + 2 = 1$.
(1 pont) $\mathbb{E}(V) = \mathbb{E}(X) - \mathbb{E}(Y) = -3$ (itt már nem jár még egyszer pont az indoklásra, elég a végeredmény).
(1 pont) $\mathbb{D}^2(U) = \mathbb{D}^2(X) + \mathbb{D}^2(Y) + 2\mathrm{cov}(X, Y)$
(1 pont) $= 9 + 9 + 2 = 20$.
(2 pont) $\mathbb{D}^2(V) = \mathbb{D}^2(X)+\mathbb{D}^2(Y)-2\mathrm{cov}(X, Y)$ (ha ez a képlet jól szerepel, akkor nem kell jobban részletezni. Ha U szórásnégyzetének kiszámolásánál már szerepelt az összeg szórásnégyzetének képlete és most jól csinálja, akkor elég konkrét számokkal, de ha előjelhiba van és hiányzik az általános képlet, akkor ez a 2 pont sem jár. Ha azt írja, hogy $\mathbb{D}^2(V) = \mathbb{D}^2(X) + \mathbb{D}^2(-Y) + 2\mathrm{cov}(X, -Y)$ és később rontja el, akkor itt 1 pont jár, a másik nem, és később lehet részpontokat adni).
(1 pont) $= 9 + 9 - 2 = 16$.
(2 pont) $\mathrm{cov}(U, V) = \mathrm{cov}(X+Y, X-Y) = \mathbb{D}^2(X)+\mathrm{cov}(X, Y)-\mathrm{cov}(X, Y)-\mathbb{D}^2(X)$ (itt 1 pont jár arra, hogy a kovariancia bilinearitását használja – nem kell megnevezni – és 1 összesen arra, hogy valaminek az önmagával vett kovarianciája a szórásnégyzet és a kovariancia szimmetrikus. A két lépés összevonható, itt is összevontuk)
(0 pont) $= 0$.
Megjegyzés: A 8. feladatsoron szerepelt, hogy ha X és Y azonos szórásúak, akkor $X + Y$ és $X - Y$ kovarianciája 0. Ez itt alkalmazható, hiszen $\mathbb{D}(X) = \mathbb{D}(Y) = 3$, de mivel ez az állítás csak gyakorlaton szerepelt (előadáson nem és Mészáros Szabolcs jegyzetében sincs benne), ezért

ha valaki erre hivatkozik, akkor a fenti 2 pontból csak 1 jár.
(1+1 pont) Tehát a várható érték vektor $\begin{pmatrix} 1 \\ -3 \end{pmatrix}$ és a kovarianciamátrix $\begin{pmatrix} 20 & 0 \\ 0 & 16 \end{pmatrix}$.
Itt nem kell az általános képlet, ha jól helyettesít be a kapott eredmények alapján. Ha nem tudja kiszámolni az értékeket, de felírja, hogy milyen mennyiségeknek kellene szerepelnie a várható érték vektor és a kovarianciamátrix adott helyein *a konkrét U és V esetén*, akkor ezt az 1–1 pontot meg lehet adni.

(c) (8 pont)
(2 pont) U és V is független normális eloszlású val. változók lineáris kombinációjaként áll elő (nem kell tovább indokolni; ha azt írja, hogy független normálisok összegeként, azt is elfogadjuk). *Innen 1. megoldás*:
(2 pont) Adott független normális eloszlású val. változók lineáris kombinációi esetén [vagy: valahogy ugyanerre képlettel helyesen hivatkozik] a korrelálatlanságból következik a függetlenség. (Azt nem kell még jobban a javító szájába rágni, hogy U és V korrelálatlanok, ha ez már kijött az (a)-ban. A korrelálatlanság helyett természetesen arra is elég hivatkozni, hogy a kovariancia 0.)
Ha csak annyit ír, hogy a korrelálatlanság miatt függetlenek, de már korábban megjelent valahogy, hogy U és V is független normálisok lineáris kombinációja, azt fogadjuk el. Ha ez nem szerepelt, akkor max. 1 pont a kettőből.
(2 pont) $\mathbb{E}(V|U) = \mathbb{E}(V)$
(1 pont) $= -3$,
(1 pont) mert függetlenek. (Ha már írta, hogy függetlenek, csak most nem indokolja az előző lépést, akkor ezt az 1 pontot ne vonjuk le. De ha rosszul indokolja, akkor vonjuk le).

2. megoldás:
(1 pont) Adott független normálisok különböző lineáris kombinációi esetén [vagy: valahogy ugyanerre képlettel helyesen hivatkozik] (ezért az 1 pontért ennyi elég, de természetesen az is jó, ha konkrétan azt írja, hogy $\mathbb{E}(V|U)$ megegyezik V-nek U-ra vett lineáris regressziójával, vagy ha nem szövegesen írja ezt, hanem egyértelműen a lineáris regressziónak megfelelő, helyes képletet akar használni).
Ha csak annyit ír, hogy a regresszió lineáris (vagy ennek megfelelőjét a konkrét esetre nézve), de már korábban megjelent valahogy, hogy U és V is független normálisok lineáris kombinációja, azt fogadjuk el. Ha ez nem szerepelt, akkor ez a pont nem jár.
(1 pont) Tehát $\mathbb{E}(V|U) = \beta U + \alpha$, (ha kiszámolja β-t és α-t helyesen, de nem használja ezt a jelölést, az is OK)
(1+1 pont) ahol $\beta = \frac{\mathrm{cov}(U,V)}{\mathbb{D}^2(U)} = 0$ (ha csak $\beta = 0$ szerepel, akkor max. 1 pont)
(1+1 pont) és $\alpha = \mathbb{E}(V) - \beta\,\mathbb{E}(U) = -3$ (itt az első pont a képletre jár, ami jó konkrét számokkal is, de csak ha azonosítható, hogy ezt a képletet használja, a második pont pedig $\mathbb{E}(V)$ azonosítására és a végeredményre, $\mathbb{E}(U)$ értéke természetesen nem kell).
(0 pont) Tehát $\mathbb{E}(V|U) = -3$. (Ha $\mathbb{E}(V|U) = \beta U + \alpha$ már szerepelt és csak a végeredmény hiányzik, akkor ne vonjunk le pontot.)
Ha regressziós egyenest ad meg, nem pedig magát a lineáris regressziót, akkor 1 pont levonás. De ha megadja a regressziót is, nem baj, ha odaírja az egyenest is.

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 definíciók alapján.

   (a) Legyenek A, B események. Mikor mondjuk azt definíció szerint, hogy A és B függetlenek?

   (b) Mikor nevezünk egy valószínűségi változót egyszerűnek?

   (c) Legyen (X, Y) folytonos valószínűségi vektorváltozó. Hogyan nevezzük és hogyan jelöljük azt a függvényt, amelynek értékét az (x, y) pontban $f_X(x) \neq 0$ esetén az

   $$\frac{f_{X,Y}(x,y)}{f_X(x)}$$

   képlettel definiáljuk, $f_X(x) = 0$ esetén pedig definíció szerint 0 az értéke?

   (d) Legyen $x \in \mathbb{R}$. A (c) részfeladatban szereplő szituációban hogyan számíthatjuk ki $f_X(x)$ értékét, ha $f_{X,Y}(x, y)$ értékét ismerjük minden $y \in \mathbb{R}$ esetén?

Megoldás.
(a)
(4 pont) Ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.
(b)
(3 pont) Ha értékkészlete véges/ha csak véges sok értéket vesz fel.
(c)
(3 pont) $f_{Y|X}(y|x)$-szel jelöljük (nem kell odaírni, hogy $(x, y) \mapsto f_{Y|X}(y|x)$)
(5 pont) és az Y-nak X-re vett (vagy X-re vonatkozó) feltételes sűrűségfüggvényének nevezzük.
(d)
(5 pont) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\mathrm{d}y$.

Megjegyzések javítóknak, részben hibás megoldások értékelése.
(a)
Szövegesen is elfogadható, ha egyértelmű, hogy erre gondol.
Ha az $A \cap B$ helyett más jelölést használ, akkor abban az esetben lehet részpontot adni, ha egyértelmű, hogy a metszetet akarja körülírni, max. 3 pont. Ha uniót vagy arra utaló kifejezést ír, akkor 0 pont!
Ha feltételes valószínűségekkel akarja megadni a definíciót, akkor max. 2 pont, és abból is 1 pont arra jár, hogy odaírja, hogy a feltétel valószínűségének pozitívnak kell lennie.
(b)
A $\mathrm{Ran}(X)$ jelölés itt nem kötelező az értékkészletre, és ha az egyetlen hiba az, hogy valahogy rosszul jelöli, pl. Ran_X, de egyértelműen erre utal, azért ne vonjunk le pontot.
Ha azt mondja, hogy maga a val. változó véges, akkor max. 1 pont.
(c)
Ha csak annyit ír, hogy Y feltételes sűrűségfüggvénye, akkor az 5-ből max. 4 pont; ha csak annyit ír, hogy feltételes sűrűségfüggvény, akkor max. 3 pont.
A következő megjegyzés kizárólag a javítóknak szól, elolvasását még nem vizsgázott hallgatóknak kifejezetten nem ajánlom: Ha más forrásokban szereplő, de az előadásról/jegyzetből nem ismert jelölést használ, – ilyenek bőven vannak, pl. $f_{Y|X}(x,y)$, $f_{Y|X=x}(y|x)$, $f_{Y|X=x}(x,y)$ – akkor a 3-ból max. 2 pont, és 2 csak akkor, ha úgy tűnik, hogy ez a jelölés valamilyen, egyetemi oktató által írott forrásban tényleg szerepel és nincs benne egyéb szintaktikai hiba.
Ha összekeveri X és Y szerepét, akkor a 3-ból max. 1 pont, ami csak a tökéletes $f_{X|Y}(x|y)$ jelölés esetén jár.
(d)
Nem baj, ha nem írja oda, hogy „$f_X(x) =$”.
Kisebb szintaktikai hibák esetén adható részpont.
Ha az integrálási határok nem jók vagy hiányoznak, akkor max. 3 pont.
Ha x szerint integrál vagy nem világos, hogy mi szerint integrál, akkor egyből 0 pont!

2. Legyenek $X_1, \ldots, X_{100}$ független, azonos eloszlású valószínűségi változók, amelyek sűrűségfüggvénye

$$f_{X_1}(x) = \begin{cases} \frac{3}{7}x^2, & \text{ha } 1 < x < 2, \\ 0, & \text{egyébként.} \end{cases}$$

   (a) Bizonyítsuk be, hogy f_{X_1} valóban sűrűségfüggvény.

   (b) Határozzuk meg X_1 várható értékét és szórását.

   (c) Legyen $X = X_1 + \ldots + X_{100}$. Közelítsük alkalmas approximációval a $\mathbb{P}(X > 160)$ valószínűséget. (A valószínűség pontos értékét nem kell meghatározni.)
   Pótfeladat: Ha a (b)-t nem sikerült megoldani, tegyük fel, hogy $\mathbb{E}(X_1) = 1{,}6100$ *és* $\mathbb{D}(X_1) = 1/4$ *és oldjuk meg a (c)-t úgy.*

Megoldás.
(a) (5 pont)
(1 pont) Ahhoz, hogy f_{X_1} sűrűségfüggvény legyen, az kell, hogy nemnegatív legyen (ez itt elég világos, nem kell részletezni, csak megemlíteni)
(1 pont) és teljesüljön, hogy $1 = \int_{-\infty}^{\infty} f_{X_1}(x)\mathrm{d}x$
(1 pont) $= \int_1^2 \frac{3}{7}x^2\mathrm{d}x$ (az előzővel összevonható, ha ezt jól írja)
(1 pont) $= [\frac{1}{7}x^3]_1^2$
(1 pont) $= \frac{8}{7} - \frac{1}{7} = 1$, tehát tényleg teljesül. (Ha jól megadja a primitív függvényt, akkor itt már elég a végeredmény. Ha kihozza, hogy az eredmény 1, nem baj, ha nem kommentálja.)
(b) (8 pont)
(1 pont) $\mathbb{E}(X_1) = \int_{-\infty}^{\infty} x f_{X_1}(x)\mathrm{d}x = \int_1^2 \frac{3x^3}{7}\mathrm{d}x$ (itt csak 1 pont jár a két alakra együtt, ezért ha

a kettőből legalább az egyik megvan helyesen, az elég)
(1 pont) $= [\frac{3}{28}x^4]_1^2$
(1 pont) $= \frac{48}{28} - \frac{3}{28} = \frac{45}{28} (\approx 1{,}6071)$ (pontozás az előző integrálhoz hasonlóan, elég csak közönséges vagy csak tizedes tört alakban is).
(1 pont) $\mathbb{D}(X_1) = \sqrt{\mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2}$ (itt arra, hogy a szórás a szórásnégyzet gyöke, nem jár külön pont, mert elég hosszú a feladat és sok mindenre kell pontot adni; erre és a $\mathbb{D}^2(X_1)$ képletére csak együtt jár pont).
(1 pont) $\mathbb{E}(X_1^2) = \int_{-\infty}^{\infty} x^2 f_{X_1}(x)\mathrm{d}x = \int_1^2 \frac{3x^4}{7}\mathrm{d}x$ (pontozás $\mathbb{E}(X_1)$-éhez hasonlóan)
(1 pont) $= [\frac{3}{35}x^5]_1^2$
(1 pont) $= \frac{96}{35} - \frac{3}{35} = \frac{93}{35} (\approx 2{,}6571)$ (pontozás az $\mathbb{E}(X_1)$ kiszámításának utolsó lépéséhez hasonlóan).
(1 pont) Tehát $\mathbb{D}(X_1) = \sqrt{\frac{93}{35} - (\frac{45}{28})^2} (\approx \sqrt{0{,}0742}) \approx 0{,}2725$.
(Ha már korábban megadta $\mathbb{E}(X_1)$ és $\mathbb{E}(X_1^2)$ értékét és $\mathbb{D}(X)$ általános képletét, akkor itt elég a végeredmény, de kell a numerikus értéke is. Az is OK, ha az általános képletet csak $\mathbb{E}(X_1)^2$ és $\mathbb{E}(X_1^2)$ konkrét értékeivel használja, feltéve, hogy azonosítható, hogy ezt a képletet alkalmazza.)

(c) (7 pont)
Megoldás a (b)-ben kapott, négy értékes jegyre kerekített numerikus értékekkel dolgozva:
(1+1 pont) Sztenderdizálunk (ezt nem kell szövegesen odaírni):

$$\mathbb{P}(X > 160) = \mathbb{P}\Big(\frac{X - 100\,\mathbb{E}(X_1)}{\sqrt{100}\,\mathbb{D}(X_1)} > \frac{160 - 100\,\mathbb{E}(X_1)}{\sqrt{100}\,\mathbb{D}(X_1)} \approx \mathbb{P}\Big(\frac{X - 160{,}71}{\sqrt{100}\cdot 0{,}2725} > \underbrace{\frac{160 - 160{,}71}{\sqrt{100}\cdot 0{,}2725}}_{\approx -0{,}2606}\Big)$$

(ha a második képlet helyes és világos módon az elsőből származik, akkor a két lépés összevonható pontlevonás nélkül),
(1 pont) ami a centrális határeloszlás-tétel (elég: CHT) miatt (De Moivre–Laplace itt nem jó!)
(1 pont) közelítőleg $\mathbb{P}(Y > -0{,}2606)$, ahol $Y \sim \mathrm{N}(0;1)$.
(1 pont) $\mathbb{P}(Y > -0{,}2606) = \mathbb{P}(Y < 0{,}2606)$ (ugyanez kifejezhető Φ-értékekkel is, de nem kötelező)
(1 pont) $\approx \mathbb{P}(Y < 0{,}26) = \Phi(0{,}26)$ (itt sem kötelező a Φ jelölés)
(1 pont) $\approx 0{,}6026$.
A (b)-ben kiszámolt momentumok pontos értékeivel számolva kapott megoldásról:
$\frac{160 - \frac{4500}{28}}{\sqrt{100}\sqrt{\frac{93}{35} - (\frac{45}{28})^2}} \approx -0{,}2621$, tehát két tizedesjegyre való kerekítés után ugyanazt az eredményt kell kiolvasni a normális eloszlás táblázatból, mint a kerekített értékekkel számolva.
A pótmegoldásról:
Ebben az esetben a standardizálás után a jobb oldalon $\frac{160-161}{10\cdot 0{,}25} = -0{,}4$ áll, így a végeredmény $\mathbb{P}(Y < 0{,}4) \approx 0{,}6554$. Pontozás ezeket az adatokat figyelembe véve ugyanúgy, mint az első megoldás esetén.
Ha (c)-ben a pótmegoldással számol, akkor a (b)-re és (c)-re együtt max. 13 pont, akkor is, ha a (b)-ben esetleg jól kiszámolta a várható értéket és szórást, csak nem mert azokkal számolni. De ha a rendes megoldást és a pótmegoldást is jól megadja, akkor nem kell pontot levonni.

3. Legyen X és Y két független és a $(0;1)$ intervallumon egyenletes eloszlású valószínűségi változó, valamint $U = X + Y$ és $V = XY$.

   (a) Határozzuk meg az U és V valószínűségi változók kovarianciáját és a várható értékeiket, valamint U szórásnégyzetét.
   Segítség: A feladat integrálás nélkül is megoldható (a tanult azonosságok használatával).

   (b) Mi V regressziós egyenese U-ra?

Megoldás.
(a) (15 pont)
Az azonosságok megnevezésére itt nem jár külön pont. Elég, ha részletesen megmutatja, hogy hogyan használja őket.
(0 pont) $\operatorname{cov}(U,V) = \operatorname{cov}(X+Y, XY)$
(1 pont) $\operatorname{cov}(X,XY) + \operatorname{cov}(Y,XY)$.
(1 pont) $\operatorname{cov}(X,XY) = \mathbb{E}(X^2Y) - \mathbb{E}(X)\,\mathbb{E}(XY)$. (Ez a pont csak akkor jár, ha a kovariancia általános képletét egy konkrét releváns helyzetben alkalmazza.)
(1 pont) $\mathbb{E}(X^2Y) = \mathbb{E}(X^2)\,\mathbb{E}(Y)$,
(1 pont) mert függetlenek.
(Nem kell részletezni, hogy ha X és Y függetlenek, akkor X^2 és Y is azok, és elég összesen egyszer hivatkozni arra, hogy függetlenség esetén a szorzat várható értéke a várható értékek szorzata).
(1 pont) $\mathbb{E}(Y) = \frac{1+0}{2} = \frac{1}{2}$ (elég a végeredmény)
(1 pont) és $\mathbb{E}(X^2) = \mathbb{D}^2(X) + \mathbb{E}(X)^2$
(1+1 pont) $= \frac{1}{12} + \frac{1}{4} = \frac{1}{3}$ (1 pont arra, ha valahol – a megoldás során bárhol – megjelenik, hogy $\mathbb{D}^2(X) = \mathbb{D}^2(Y) = \frac{1}{12} (= \frac{(1-0)^2}{12})$, indoklás nélkül is; sőt elég, ha az egyik jelenik meg és a másikat csak magyarázat nélkül írja oda; és 1 pont a végeredményre, a számolást nem kell nagyon részletezni).
(1 pont) $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y) = \frac{1}{4}$ (ha megvan a szorzatra bontás, akkor már jár az 1 pont, és ha jól csinálja, akkor nem kell indokolni, hogy $\mathbb{E}(Y) = \mathbb{E}(X)$ [mert X és Y azonos eloszlásúak]).
(1 pont) Tehát $\operatorname{cov}(X,XY) = \frac{1}{3}\cdot\frac{1}{2} - \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{6} - \frac{1}{8} = \frac{4-3}{24} = \frac{1}{24}$. (Ha minden korábbi részeredmény jó és a $\operatorname{cov}(X,XY) = \mathbb{E}(X^2Y) - \mathbb{E}(X)\,\mathbb{E}(XY)$ képlet is megvan, akkor elég a végeredmény is.)
(1 pont) A szimmeria miatt $\operatorname{cov}(Y,XY) = \operatorname{cov}(X,XY) = \frac{1}{24}$. (Indoklás nélkül is OK.)
(0 pont) Tehát $\operatorname{cov}(U,V) = \frac{1}{12}$.
(0 pont) $\mathbb{D}^2(U) = \mathbb{D}^2(X+Y)$
(1 pont) $= \mathbb{D}^2(X) + \mathbb{D}^2(Y)$
(0 pont) $= \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$ (itt már nem jár pont a szórásnégyzetek értékére, mert korábban már adtunk rá),
(1 pont) mert függetlenek.
(1 pont) $\mathbb{E}(U) = \mathbb{E}(X) + \mathbb{E}(Y)$
(1 pont) $= \frac{1}{2} + \frac{1}{2} = 1$ (itt is elég a végeredmény).
(0 pont) $\mathbb{E}(V) = \mathbb{E}(X)\,\mathbb{E}(Y) = \frac{1}{4}$, ahogy már kiszámoltuk. (Erre a feladat során egyszer jár 1 pont és a várható érték szorzatra bomlására is összesen egyszer.)

Összefoglalva a végeredmények: $\operatorname{cov}(U,V) = \frac{1}{12}$, $\mathbb{D}^2(U) = \frac{1}{6}$, $\mathbb{E}(U) = 1$, $\mathbb{E}(V) = \frac{1}{4}$.

Jó úgy is, ha azt írja, hogy $\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y) + 2\operatorname{cov}(X,Y)$, és utána veszi észre, hogy a függetlenség miatt a kovariancia 0. Ekkor a függetlenségre járó pont az utóbbi észrevételre jár.
Természetesen nem tilos a feladatot integrálással megoldani, ekkor a kovariancia kiszámításához kell X és Y együttes sűrűségfüggvénye, és ennek szorzat alakját is indokolni kell. Lehet egyes részfeladatokban is integrálni, ha nem megy másként, például ha $\mathbb{E}(X^2)$-et nem tudja a szórásnégyzet segítségével meghatározni, de integrálással igen.

(b) (5 pont)
(1 pont) A lineáris regresszió képletében szereplő mennyiségek: $\beta = \frac{\operatorname{cov}(U,V)}{\mathbb{D}^2(U)}$
(1 pont) $= \frac{\frac{1}{12}}{\frac{1}{6}} = \frac{1}{2}$ (itt már elég a végeredmény, ha korábban megvolt minden szükséges részletszámítás)
(1 pont) és $\alpha = \mathbb{E}(V) - \beta\,\mathbb{E}(U)$
(1 pont) $= \frac{1}{4} - \frac{1}{2}\cdot 1 = -\frac{1}{4}$ (az előzőhöz hasonlóan).
(1 pont) Tehát a regressziós egyenes $\{(x,y) \in \mathbb{R}^2 \colon y = \frac{1}{2}x - \frac{1}{4}\}$.

Ha nem az egyenest adja meg, hanem csak a $\beta X + \alpha$ lineáris regressziót, akkor az utolsó pont nem jár.
Ha az egyenes helyes általános képletét ($\beta x + \alpha$-val) adja meg, akkor jár, még ha a számszerű értékeket nem is sikerült meghatározni.
A β és α jelölések nem kötelezőek, lehet helyettük más betűket is használni, és semmilyen betűt nem kell használni, ha anélkül is azonosítható, hogy melyik mennyiség honnan származik.

4. Feldobunk egy szabályos kockát, majd egy szabályos érmét annyiszor, amennyit a kocka mutat. Jelölje Y a dobott fejek számát és X a kockadobás értékét.

   (a) Várhatóan hány fejet dobunk, feltéve, hogy a kocka 3-at mutat (és miért)?

   (b) Mennyi $\mathbb{E}(Y|X)$?

   (c) Határozzuk meg Y várható értékét.

   (d) Lássuk be, hogy X és Y nem függetlenek.

   A részfeladatokat tetszőleges sorrendben megoldhatjuk.
   Megoldás.
   (a) (4 pont)
   (1 pont) Várhatóan $3/2 = 1{,}5$ fejet dobunk, ha a kocka 3-at mutat (ez az 1 pont indoklás nélkül is jár).
   (3 pont) Megindokolja rendesen, hogy miért igaz, hogy $\mathbb{E}(Y|X = 3) = \frac{3}{2}$ (ehhez nem kötelező a feltételes várható értékes jelölést használni), ilyen esetben jár az előző 1 pont is.
   Például: ha tudjuk, hogy a kocka 3-at mutat, akkor a fejek száma binomiális eloszlású $n = 3$ és $p = 1/2$ paraméterrel, és nekünk éppen ennek a várható értéke kell, ami $\frac{3}{2}$.
   Ugyanez indikátorok összegével kifejezve: ha tudjuk, hogy a kocka 3-at mutat, minden dobás $1/2$ valószínűséggel lesz fej (tehát egy dobásból várható értékben $1/2$ fej lesz), így a várható érték linearitása miatt a fejek várható száma $3/2$. Ha itt a linearitásra nem hivatkozik és nem ír le olyan képletet sem, ami erre utalna, akkor a 3 pontból max. 2 pont.
   Ha csak annyi indoklást ír, hogy átlagosan/várható értékben a dobások fele lesz fej, akkor tekintsük úgy, hogy az utóbbi megoldást próbálja leírni és adjunk rá 2 pontot a 3-ból. Ha csak annyit ír, hogy minden dobás $1/2$ valószínűséggel fej és ebből következik az eredmény, akkor 1 pontot adjunk a 3-ból.
   (b) (4 pont)
   (1 pont) Ha a kocka k-t mutat (ahol $k \in \{1, \dots, 6\}$), akkor várhatóan $k/2$ fejet dobunk. (Nem kell indokolni, ugyanúgy következik, mint az (a)-ban.)
   (1 pont) Ez azt jelenti, hogy $\mathbb{E}(Y|X = k) = \frac{k}{2}$ ilyen k esetén. (Ha ez megvan, akkor jár az előző 1 pont is. Ha megnevezi, hogy ez regressziós függvény, elég $g(k)$-val jelölni.)
   (1 pont) $\mathbb{E}(Y|X)$-et (definíció szerint) úgy kapjuk, hogy a $k \mapsto g(k) = \mathbb{E}(Y|X = k)$ regressziós függvénybe X-et írunk.
   (1 pont) Tehát $\mathbb{E}(Y|X) = \frac{X}{2}$. Ez az előző lépéssel összevonható abban az esetben, ha a (b) első 2 pontjából legalább egyet megszerzett, különben a 3. pont nem jár.
   (c) (7 pont)
   (2 pont) A teljes várható érték tétele szerint (szintén OK: TVÉT, toronyszabály. Ha TVT-t ír, akkor 0 pont, kivéve, ha utána mégis megjelenik a jó képlet, mert akkor 1 pont)
   (2 pont) $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(g(X))$ (elég a három közül az egyik jelölés, és ha már korábban **expliciten** szerepelt, hogy $g(k) = k/2$, akkor $\mathbb{E}(X/2)$ vagy $\mathbb{E}(X)/2$ is elég; egyébként $\mathbb{E}(X/2)$ vagy $\mathbb{E}(X)/2$ esetén csak 1 pont jár)
   1. megoldás:
   (1 pont) $= \sum_{k=1}^{6} \frac{k}{2} \mathbb{P}(X = k)$ (ez a pont akkor is jár, ha $\frac{k}{2}$ helyett csak $g(k)$-val szerepel a képlet)
   (1 pont) $= \frac{1}{6} \sum_{k=1}^{6} \frac{k}{2}$ (úgy is jó, hogy $\frac{1}{6} \cdot \frac{1+2+\dots+6}{2}$)

(1 pont) $= \frac{7}{4} = 1{,}75$ (elég az egyik alak és részletszámítások sem kellenek).
2. megoldás:
(1 pont) $= \mathbb{E}(X)/2$ (ha ez az alak megvan, akkor jár az előző 2 pont is).
(2 pont) Kockadobás várható értéke $\mathbb{E}(X) = 3{,}5 = \frac{7}{2}$ (ezt elég közölni, nem kell újra kiszámolni), tehát $\mathbb{E}(Y) = \frac{7}{4} = 1{,}75$ (megint elég az egyik alak).
(d) (5 pont)
1. megoldás: a függetlenség definíciója alapján.
Megad egy olyan (k, ℓ) párt, amelyre $\mathbb{P}(X = k, Y = \ell) \neq \mathbb{P}(X = k)\,\mathbb{P}(Y = \ell)$.
Itt 1 pont jár arra, ha egyértelműen ezt akarja csinálni, 1 pont arra, ha jó párt (k, ℓ) választ, és 1–1 pont a három valószínűség értékére, ezekhez nem kell indoklás. A legegyszerűbb olyan példát választani, ahol a jobb oldalon az egyik tényező 0, ekkor a másik tényező értékét nem is kell meghatározni és arra is jár az 1 pont. Például ha $\ell > k$ (több fejet dobunk, mint ahányat a kockával dobtunk, azaz ahány érmével dobunk), akkor mindig ez a helyzet.
Ugyanez a fajta megoldás feltételes valószínűségekkel és együttes/marginális eloszlásfüggvényekkel kifejezve is OK.
2. megoldás: a regresszió tulajdonságai alapján.
(2 pont) Ha X és Y függetlenek lennének, akkor $\mathbb{E}(Y|X) = \mathbb{E}(Y)$ teljesülne.
(3 pont) Ez viszont itt nem teljesül, mert $\mathbb{E}(Y)$ egy konstans, míg $\mathbb{E}(Y|X) = X/2$ [elég az egyik alak] nem (1 valószínűséggel) konstans. Vagy: mert $\mathbb{E}(Y) = 1{,}75$ a (c) alapján, és $\mathbb{E}(Y|X) = X/2$ nem (1 valószínűséggel) ezzel az értékkel egyenlő.

5. 7 darab, a BME campusán élő idősebb nyárfa magasságát megmértük és (méterben mérve) az alábbi eredményeket kaptuk:

$$x_1 = 23{,}1 \qquad x_2 = 21{,}5 \qquad x_3 = 22{,}0 \qquad x_4 = 22{,}8 \qquad x_5 = 21{,}2 \qquad x_6 = 20{,}9 \qquad x_7 = 22{,}5.$$

Egy kertészeti szakkönyv szerint az idősebb nyárfák magassága jó közelítéssel normális eloszlású, 1 méter szórással. Ezt az állítást elfogadva

(a) adjunk 0,98 szintű konfidenciaintervallumot a minta alapján a mintaelemek várható értékére,

(b) döntsük el $\varepsilon = 0{,}04$ terjedelmű próbával, hogy eléri-e a mintaelemek várható értéke a 22,5 métert. Adjuk meg H_0-t, H_1-et, a próbastatisztika értékét, azt, hogy mi alapján döntünk, és magát a döntést is.

(c) Határozzuk meg a minta empirikus mediánját.

Megoldás. (a) (8 pont)
(1 pont) arra, hogy az ismert szórás esetére vonatkozó (avagy u-kvantilises, u-próbánál szereplő stb.) konfidenciaintervallum képlete kell. Ha jó képletet használ, akkor nem baj, ha ezt nem írja oda.
(1 pont) $n = 7$ (ez a mintaelemszám),
(1 pont) $\overline{x_7} = \frac{1}{7}\sum_{i=1}^{7} x_i = 22$ (csak 1 pont jár rá, ezért OK indoklás nélkül is, $\overline{x_n}$ jelöléssel is jó),
(1 pont) $\sigma = 1$ (ez a mintaelemek szórásának felel meg, de ezt nem kell odaírni. Ha viszont azt állítja, hogy σ a szórásnégyzet, akkor nem jár a pont)
(1 pont) a szükséges u-kvantilis pedig $u_{\varepsilon/2} = u_{0{,}01} = \Phi^{-1}(0{,}99)$ (az ε jelölés nem kell, és elég ha az utóbbi két alakból az egyik szerepel, de derüljön ki, hogy ε-nak itt 0,02 felel meg)
(1 pont) $= 2{,}33$.
(1+1 pont) A képletbe behelyettesítve a konfidenciaintervallum:

$$\left[\overline{x_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{x_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\right] = \left[22 - \frac{1 \cdot 2{,}33}{\sqrt{7}}, 22 + \frac{1 \cdot 2{,}33}{\sqrt{7}}\right] = [21{,}1193, 22{,}8807].$$

Az általános képletre itt nem jár pont. 1 pont arra, hogy behelyettesíti a képletbe a kapott értékeket, és 1 pont a numerikus végeredményre. Ha már mindent kiszámolt és utána csak a numerikus végeredményt adja meg, akkor ne vonjunk le pontot.
Megjegyzés: ha a $\frac{\sqrt{2}}{100}$ pontos értékkel számolunk, akkor is ez az eredmény jön ki 4 tizedesjegyre kerekítve.

Javítóknak: Ha a t-próbás konfidenciaintervallum képletét akarja használni, akkor max. 7 pont. Ebből a 7 pontból 3 jár a korrigált empirikus szórásnégyzet kiszámolására (az előző vizsga 5-ös feladatához hasonlóan), és a kvantilisre járó pont csak akkor jár, ha a szabadsági fok ($n-1=6$) is helyes. Így a mintamegoldásban a σ értékére járó ponton kívül megszerezhető 7 pontból a korrigált empirikus szórás nélkül maximum 4 pont jár, további hibák esetén ebből vonunk le, és ha 0 pontra csökken az összpontszám, akkor az (a)-ra 0 pont és nem kell az (a) javítását folytatni.
Ha a jó képletet akarja használni, de azt hiszi, hogy σ a korrigált empirikus szórást jelöli, akkor s_7^* kiszámolására nem jár pont, a többi 7 pont megszerezhető. Ha az ismeretlen szórásos képletet akarja használni és s_n^* helyére beírja σ értékét, akkor a fentiek szerint max. 4 pont.

(b) (9 pont)
(1 pont) A feladat szövege alapján (nem kell indokolni) az alternatíva $H_0\colon \mu \geq 22{,}5$ vs. $H_1\colon \mu < 22{,}5$ (ahol μ jelöli a mintaelemek várható értékét – ezt sem kell odaírni).
(1 pont) Tehát ez egy egyoldali, egymintás próba (ha ezt nem írja oda, de ennek megfelelő képlettel számol, akkor ezt az 1 pontot meg lehet adni),
(1 pont) és mivel σ ismert, ezért u-próba kell. (Ha ezt nem írja oda, a pontot abban az esetben lehet megadni, ha a későbbiekben egyértelmű, hogy u-próbával számol és világos, hogy tudja, hogy mi σ.)
(1+1 pont) A próbastatisztika értéke $u(= u(x_1,\dots,x_7)) = \frac{\overline{x_n}-\mu_0}{\sigma}\sqrt{n} = \frac{22-22{,}5}{1}\sqrt{7} = \frac{-\sqrt{7}}{2} \approx -1{,}3229$.
Itt 1 pont jár arra, hogy a helyes általános képletet értelmesen használja a már kiszámolt mennyiségekkel, ideértve, hogy tudja, hogy μ_0 a H_0 szerinti várható értéket jelenti. Elég, ha konkrét számokkal szerepel a képlet, ha azonosítható, hogy jó képletet használ. 1 pont pedig a helyes numerikus végeredményre jár, részletszámítások nem kellenek.
(1 pont) H_0-t akkor fogadjuk (nem utasítjuk) el, ha $u \geq -u_\varepsilon = -u_{0{,}04}$
(1 pont) $= -\Phi^{-1}(1-0{,}04) = -\Phi^{-1}(0{,}96)$ (ez az előzővel összevonható, ha helyesen szerepel)
(1 pont) $\approx -1{,}75$. (Ha jó eredményt olvas ki a táblázatból, akkor elég az $u_{\varepsilon/2}$ vagy $u_{0{,}02}$ jelölés, nem kell Φ^{-1}-zel kifejezni.)
(1 pont) Tehát H_0-t elfogadjuk/nem utasítjuk el (vagyis azt mondjuk, hogy a nyárfák magasságának várható értéke eléri a 22,5 métert – de itt nem kell szöveges magyarázat).

Pontozás hibás alternatívák/próbatípusok esetén: Ha kétoldali u-próbát hajt végre $H_0\colon \mu = 22{,}5$ vs.$H_1\colon \mu \neq 22{,}5$ alternatívával, az nem felel meg a feladatnak, hibátlan végrehajtás esetén max. 7 pont. Ha egyoldali u-próbát hajt végre, de $H_0\colon \mu \leq 22{,}5$ vs. $H_1\colon \mu > 22{,}5$ alternatívával, az még kevésbé logikus, max. 8 pont.
Ha az alternatíva jó, de erre (egyoldali, egymintás) t-próbát alkalmaz, akkor max. 7 pont. Az ismert vs. ismeretlen szórás miatt ugyan csak egy pontot kellene levonni, de $t_{f,0{,}04}$ nincs a táblázatban, és ennek fel kellene tűnnie. Hasonlóan, ha kétoldali t-próbát hajt végre, akkor nemcsak a szórással és az alternatívával van probléma, hanem azzal is, hogy $t_{f,0{,}02}$ nincs a táblázatban, ekkor max. 5 pont. Végül ha egyoldali t-próbát hajt végre a $H_0\colon \mu \leq 22{,}5$ vs. $H_1\colon \mu > 22{,}5$ alternatívára, akkor is fennállnak ugyanezek a problémák és az alternatíva még kevésbé logikus, max. 4 pont.
Általában induljunk ki a hallgató által megadott alternatívából és a megoldás további részeinek helyességét az alapján ítéljük meg, hogy az alternatívának megfelel-e. Ha nem egyértelmű az sem, hogy mi az alternatíva, akkor legfeljebb arra járnak részletpontok, hogy a táblázat-

ban/képletekben szereplő egyes kifejezéseket az adott szituációban értelmezi és behelyettesíti/kiolvassa.

(c) (3 pont)
(1 pont) Páratlan mintaelemszám esetén
(1 pont) az empirikus medián a mintának a nagyság szerinti (növekvő/csökkenő) sorrendben a középső eleme,
(1 pont) ami itt 22,0. (Ha odaírja, hogy ez x_3, az nem baj, mert nincsenek egyenlő mintaelemek, de kell a számszerű érték is).
Ha csak a helyes eredményt közli, 1 pont. Természetesen jó rendezett mintás jelöléssel is, de akkor is hivatkozni kell arra, hogy a mintaelemszám páratlan. Ha a második pont nem jár, akkor az első sem.

6. * $N \geq 2$ ember felírja a nevét egy-egy cetlire és ezeket egy kalapba rakják, majd mindenki véletlenszerűen húz egy cetlit a kalapból (visszatevés nélkül).

   (a) Az N ember között szerepel Anna és Béla. Mi a valószínűsége, hogy ők ketten egymás nevét húzzák?

   (b) Mi a várható értéke a *transzpozíciók* számának, azaz azon párok számának, akik egymás nevét húzzák?

Megoldás.
(a) (6 pont)
(1 pont) Az eseménytér modellezhető az N név/elem (vagy az $\{1, \ldots, N\}$ halmaz) permutációinak/sorbarendezéseinek halmazával,
(1 pont) a valószínűségi mező pedig klasszikus (vagy: kedvező esetek száma/összes eset száma képlettel lehet számolni).
Az első pontért nem muszáj a permutációkra hivatkozni, ha később kiderül a megoldásból, hogy az összes eset száma $N!$. A második pontért viszont kell, hogy legyen legalább annyi a megoldásban, hogy „kedvező/összes" vagy „klasszikus valószínűség(i mező)", még akkor is, ha láthatóan így számol.
(1 pont) A kedvező esetek azok, ahol Anna és Béla egymás nevét húzzák, ekkor a maradék $N-2$ ember közül mindenki e közül az $N-2$ ember közül húzza ki valakinek a nevét.
(1 pont) Így tehát $(N-2)!$ a kedvező esetek száma (mivel ha tudjuk, hogy Anna és Béla egymás nevét húzták, akkor a többi ember húzásai egyértelműen meghatározzák a kimenetelt – ezt nem kell leírni).
Ha ez szerepel helyesen és legalább valami minimális indoklást próbál rá adni, – ha nem is olyan részletesen, mint ami itt a mintamegoldásban az előző pontnál szerepelt – akkor jár az előző pont is.
(1 pont) Az összes lehetséges kimenetel száma $N!$,
(1 pont) így a keresett valószínűség $\frac{(N-2)!}{N!} = \frac{(N-2)!}{(N-2)!(N-1)N} = \frac{1}{N(N-1)}$.
(Ha egyértelmű, hogy kedvező/összes képlettel számol, akkor elég a végeredmény. A végeredményt előbb-utóbb egyszerűsíteni kell, de elég a (b)-ben megtenni ezt.)

(b) (14 pont)
(0 pont) A transzpozíciók számát jelölje X, ekkor $\mathbb{E}(X)$-et kell meghatározni.
(3 pont) Egy tetszőleges $\{i, j\}$ párra, ahol $1 \leq i \leq N$ és $1 \leq j \leq N$ és $i \neq j$ (elég, ha az szerepel, hogy $i \neq j$), vezessük be az $X_{i,j} = \mathbb{1}_{\{i \text{ és } j \text{ transzpozíciót alkot}\}}$ valószínűségi változót/indikátorváltozót.
Más szavakkal megfogalmazva (elég az egyik megfogalmazás): legyen $X_{i,j} = 1$, ha i és j transzpozíciót alkot és $X_{i,j} = 0$ egyébként.
VAGY: az $X_{i,j}$-ket tekintsük csak $1 \leq i < j \leq N$ esetén.

VAGY: $X_{i,j}$-k helyett $X_{\{i,j\}}$-ket vezessünk be, ahol $1 \le i \le N$ és $1 \le j \le N$ és $i \ne j$ (és nyilván $\{i,j\} = \{j,i\}$).
(2 pont) Ekkor $X = \sum_{\{i,j\}\colon 1 \le i \le N, 1 \le j \le N} X_{i,j}$, VAGY: $X = \sum_{1 \le i < j \le N} X_{i,j}$.
(VAGY: $X = \sum_{\{i,j\}\colon 1 \le i \le N, 1 \le j \le N} X_{\{i,j\}}$; ezt a fajta megoldást a továbbiakban nem részletezem, mert majdnem ugyanaz, mint az első.)
(1 pont) A várható érték linearitása miatt (képlettel is kifejezhető, elég az egyszerű val. változókra vonatkozó verziójára hivatkozni, de nem elég a konkrét esetre felírni)
(2 pont) $\mathbb{E}(X)$ az összes $\mathbb{E}(X_{i,j})$ összegével egyezik meg (képlettel is jó persze).
(1 pont) $\binom{N}{2} = \frac{N(N-1)}{2}$, a feltételeknek megfelelő $\{i,j\}$ pár van.
(Ezt itt nem kell jobban indokolni. Ha az $1 \le i < j \le N$ alakú (i,j) rendezett párokat nézzük, akkor ezeknek a száma pontosan az N elem ismétlés nélküli kombinációinak száma, ami $\binom{N}{2}$. Ha pedig az $\{i,j\}$ alakú párokat számláljuk le, ahol i és j is 1 és N között lehet és különbözőek, akkor az N elem másodosztályú, ismétlés nélküli variációinak számát kell 2-vel osztani, mert $\{i,j\}$ és $\{j,i\}$ mindig megegyezik egymással, ez tehát $\frac{N(N-1)}{2}$ pár.)
(2 pont) Az (a) feladat alapján bármely két elem $\frac{1}{N(N-1)}$ valószínűséggel alkot transzpozíciót. (Ha csak annyit ír, hogy ez az Annából és Bélából álló párra igaz, arra 1 pont jár a 2-ből. Nem kell hivatkozni a szimmetriára/azonos eloszlásra.)
(2 pont) Ez alapján $\mathbb{E}(X_{i,j}) = \frac{1}{N(N-1)}$ minden vizsgált i,j esetén, mert indikátorváltozó várható értéke az indikátornak megfelelő esemény valószínűsége. (1 pont az indoklásra, 1 pont az eredményre. Természetesen jó úgy is, hogy a várható érték definíciója szerint kiszámolja $X_{i,j}$ várható értékét, ezt nem kell nagyon részletezni, de 1 pontért legyen ott az általános képlet vagy annak a konkrét esetre vonatkozó alakja, a 0-t el lehet hagyni, a másik pont az eredményre jár.)
(1 pont) Tehát $\mathbb{E}(X) = \frac{N(N-1)}{2} \frac{1}{N(N-1)} = \frac{1}{2}$ a keresett várható érték. (Ha minden részeredmény megvolt már, akkor elég a végeredmény, de az legyen kiegyszerűsítve: 1/2 vagy 0,5.)

Ha következetesen rendezett párokkal számol, akkor a (b)-re max. 12 pont. (A rendezett párok száma éppen az ismétlés nélküli variációk száma, azaz $N(N-1)$.)

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 definíciók és állítások alapján. Legyen $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező és legyenek $X, Y \colon \Omega \to \mathbb{R}$ valószínűségi változók.

   (a) Mikor mondjuk azt definíció szerint, hogy X és Y függetlenek?

   (b) Ha X és Y független valószínűségi változók, hogyan számíthatjuk ki a két változó szorzatának várható értékét a változók várható értékei segítségével, ha ezek a várható értékek mind léteznek? (Az $\mathbb{E}$ jelölést magyarázat nélkül használhatjuk a szokásos módon.)

   (c) Legyen $g \colon \mathbb{R} \to \mathbb{R}$ függvény és tegyük fel, hogy X folytonos és $g(X)$ várható értéke létezik. Jelölje $x \mapsto f_X(x)$ az X sűrűségfüggvényét. Hogyan számíthatjuk ki ekkor $g(X)$ várható értékét?

   Megoldás.
   (a)
   (5 pont) Ha az $\{X < x\}$ és az $\{Y < y\}$ események minden $x, y \in \mathbb{R}$ esetén függetlenek.
   (A nívóhalmazok jelentését nem kell megmagyarázni, természetesen jó úgy is, hogy $\{\omega \in \Omega \colon X(\omega) < x\}$. A függetlenség kimondható valószínűségekkel is: $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\,\mathbb{P}(Y < y)$.)
   (b)
   (5 pont) $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.
   (c)
   (10 pont) $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x$.

   Javítási instrukciók:
   (a)
   Ha diszkrétben vagy folytonosban fogalmazza meg az abban az esetben ekvivalens definíciót, akkor max. 2 pont.
   Ha feltételes valószínűségekkel fogalmazza meg, de a jó eseményekkel dolgozik, akkor max. 3

pont. Ha nem a jó eseményekkel dolgozik, de pl. diszkrét esetben pozitív valószínűségű eseményeknél jó, amit ír, akkor max. 1 pont.
Ha nem szerepel a „$\forall x, y \in \mathbb{R}$” kitétel, akkor max. 3 pont, egyéb okból is hibás/hiányos megoldás esetén max. 2 pont, a korábbiak szerint eleve csak 1 pontos megoldás esetén 0 pont.
(b)
$\mathbb{E}(\cdot)$ helyett $\mathbb{E}[\cdot]$ és $\mathbb{E}\cdot$ is jó, de utóbbi csak akkor, ha konzisztensnek és hihetőnek tűnő módon használja. Kétes esetben a javító dönt. $\mathbb{E}\{\cdot\}$-ért legalább egy pont levonás.
(c)
Kisebb szintaktikai hibákért adható részpont. Például ha f_X helyett f van, akkor max. 8 pont, ha f_X-nek nem jó az argumentuma, akkor pedig max. 6 pont. Ha nincsenek integrálási határok, akkor max. 4 pont, ha pedig rosszak, akkor max. 1 pont.

2. Egy szabályos dobókocka feldobását az első **hatosnál kisebb** eredményig ismételjük. Jelölje X az ehhez szükséges dobások számát (beleértve az első nem hatos dobást is).

   (a) Mi a valószínűsége, hogy a kockát legalább háromszor fel kell dobnunk?

   (b) Határozzuk meg X szórását és X^2 várható értékét.

   (c) Feltéve, hogy hét dobás után még nem sikerült egyszer sem hatosnál kisebbet dobnunk, mi a valószínűsége, hogy az utána következő két dobásban sem sikerül?

   (d) Definiáljuk ugyanebben a kísérletben az Y valószínűségi változót a következőképpen: $Y = 1$, ha a második dobás eredménye nem hatos és $Y = 0$ egyébként (ha nincs második dobás, Y értéke akkor is 0). Független-e X és Y?

Megoldás.
(a) (6 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(X \geq 3)$. (Ha ezt nem írja oda expliciten, de később ennek megfelelő képlettel számol, akkor jár ez a pont is.)
(1 pont) X geometriai eloszlású
(1 pont) $p = 5/6$ paraméterrel (nem kell a p jelölés és indoklás sem kell).
(1 pont) Ezért $\mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X = 1) - \mathbb{P}(X = 2)$ (ez csak akkor jár, ha megvan a komplementer valószínűség és az esetszétbontás is)
(1+1 pont) $= 1 - p - p(1-p) = 1 - \frac{5}{6} - \frac{5}{6} \cdot \frac{1}{6} = \frac{1}{36}$.
(A p jelölés itt sem kötelező, elég, ha konkrét számokkal megvan, illetve ha p-vel megvan, azért jár 1 pont a 2-ből, és utána a végeredményt további részletszámolás nélkül elhisszük.)
Az is elfogadható, ha azt írja, hogy $\mathbb{P}(X \geq 3) = \mathbb{P}(X > 2) = (1-p)^2$, ahol az utóbbi két lépés összevonható és p helyett megint OK a konkrét 5/6 paraméterrel is (ha azonosítható, hogy ez a p). Úgy tekintem, hogy a $\mathbb{P}(X > k) = (1-p)^k$ képlet szerepelt előadáson (ha nem is nagyon hangsúlyoztam).
Nem kell komplementer valószínűséggel számolni, ha $\mathbb{P}(X \geq 3)$-at a végtelen mértani sor összegképletével számolja ki. Ekkor a $\mathbb{P}(X \geq 3) = \sum_{k=3}^{\infty} \mathbb{P}(X = k) = p \sum_{k=3}^{\infty} (1-p)^{k-1}$ vagy hasonló képletre még csak 1 pont jár, a további 2 pont csak akkor jár, ha jól alkalmazza a végtelen mértani sor összegképletét és jó végeredményt hoz ki (kizárólag számolási, de nem elvi hiba esetén lehet a 2-ből 1 pontot adni).

(b) (6 pont)
(1 pont) $\mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}$
(1 pont) $= \sqrt{\frac{1-p}{p^2}}$ (az előző lépéssel összevonható)
(1 pont) $= \sqrt{\frac{1/6}{25/36}} = \sqrt{\frac{6}{25}} \approx 0{,}4899$.
(A numerikus végeredmény kell, viszont ha az előző lépés és p jó, akkor részletszámolások nem

kellenek. Ha az előző lépés hiányzik, de vannak jó részletszámolások a konkrét p-re, amik egyértelműen ugyanabból a számolásból származnak, akkor ez a lépés az előzővel összevonható.)
(1 pont) $\mathbb{E}(X^2) = \mathbb{D}^2(X) + \mathbb{E}(X)^2$
(1 pont) $= \frac{6}{25} + \frac{1}{p^2} = \frac{6}{25} + \frac{36}{25}$ (itt az 1 pont az $1/p^2$-re és az abba való behelyettesítésre jár. Jó konkrét p-vel is, ha azonosítható, hogy így csinálja)
(1 pont) $= \frac{42}{25} (= 1{,}68)$. (Elfogadjuk közönséges tört, végtelen szakaszos tizedestört és négy tizedesjegyre kerekített tizedestört alakban is.)

(c) (4 pont)
(1 pont) A keresett (feltételes) valószínűség $\mathbb{P}(X > 9 | X > 7)$ (avagy $\mathbb{P}(X > 7 + 2 | X > 7)$, $\geq$-kkel helyesen megfogalmazva is OK)
(1 pont) A geometriai eloszlás örökifjúsága (elég: az örökifjúság) miatt
(1 pont) ez $\mathbb{P}(X > 2)$-vel egyenlő (az is jó, ha nem írja fel a feltételes valószínűséget, hanem azt mondja, hogy az örökifjúság miatt $\mathbb{P}(X > 2)$-t keressük, de ha az örökifjúság hiányzik, akkor csak ez a pont jár rá és az előző kettő nem),
(1 pont) ami pont ugyanaz, mint az (a)-ban már kiszámított $\mathbb{P}(X \geq 3)$ valószínűség, vagyis $\frac{1}{36}$. (Ha kiszámolja, arra is 1 pont jár.)

(d) (4 pont)
Önmagában annak kimondására, hogy nem függetlenek, nem jár pont.
(1 pont) arra, hogy kimondja, hogy nem függetlenek, és ad hozzá egy indoklást, ami legalább szintaktikailag nagyjából rendben van és hihető, hogy tényleg komolyan gondolja.
(1 pont) arra, hogy azzal akarja belátni, hogy nem függetlenek, hogy van olyan x és y, amelyre $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x)\,\mathbb{P}(Y = y)$.
(2 pont) arra, hogy talál is egy ilyen helyes (x, y) párt és megmondja, hogy mennyi $\mathbb{P}(X = x, Y = y), \mathbb{P}(X = x)$ és $\mathbb{P}(Y = y)$ (indoklás nem kell). Ilyen esetben jár az előző 1 pont is. Ha a bal oldal 0 és ezt odaírja, akkor elég odaírni, hogy a jobb oldalon szereplő egyik szorzótényező sem 0, nem kell kiszámolni ezen tényezők értékeit. Ha a bal oldal nem 0, akkor kellenek a jobb oldali tényezők értékei (indoklás ekkor sem kell).
Pl.: $\mathbb{P}(X = 1, Y = 1) = 0$ (mert ha $Y = 1$, akkor van 2. dobás, tehát nem lehet X értéke 1), $0 \neq \mathbb{P}(X = 1) (= \frac{5}{6})$, $0 \neq \mathbb{P}(Y = 1) (= \mathbb{P}(X = 2) = \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{36})$.
Lehet a függetlenség definíciója szerint is érvelni vagy $\operatorname{cov}(X, Y)$ nem nulla voltával, de ezek csak lényegesen hosszasabbak. Ha nincsenek befejezett és legfeljebb számolási hibákat tartalmazó számítások, akkor ilyen esetekben max. 1 pont.

3. Egy izzógyár mérnökei megfigyelték, hogy minél több izzót gyártanak egy napon, annál kisebb valószínűséggel lesz hibás egy-egy izzó. Ha egy napon $n \geq 1$ izzót készítenek, akkor minden izzó egymástól függetlenül $\frac{1}{2n}$ valószínűséggel lesz hibás.

   (a) Ha egy napon $n = 10$ izzót gyártanak, akkor mennyi az aznap gyártott hibás izzók számának várható értéke és szórásnégyzete (és miért)?

   (b) $n = 10$ esetén mi a valószínűsége, hogy 4-nél több, de 8-nál kevesebb izzó lesz hibás a 10 izzó közül? (A valószínűség pontos értékét adjuk meg.)

   (c) Tegyük fel, hogy egy napon n izzót gyártanak, ahol n nagyon nagy. Közelítsük alkalmas approximációval annak a valószínűségét, hogy az adott napon pontosan 3 izzó lesz hibás. (A valószínűség pontos értékét itt nem kell meghatározni.)

Megoldás.
(a) (6 pont)
(0 pont) Jelöljük X-szel az adott napon gyártott hibás izzók számát.
(1 pont) Ekkor X binomiális eloszlású

(1 pont) $n = 10$ (arra önmagában nem jár pont, hogy $n = 10$, csak arra, ha kiderül, hogy a binomiális eloszlásban a kísérletek számának megfelelő paraméter értéke 10)
(1 pont) és $p = \frac{1}{2\times 10} = \frac{1}{20}$ paraméterekkel (az n és p jelölések nem kötelezőek.),
(1 pont) ezért $\mathbb{E}(X) = np = \frac{1}{2}$ (Ha már n és p értéke is szerepelt, akkor itt elég a végeredmény)
(1 pont) és $\mathbb{D}^2(X) = np(1-p)$
(1 pont) $= 10 \cdot \frac{1}{20} \cdot \frac{19}{20} = \frac{19}{40} = 0{,}475$ (elég csak közönséges tört vagy csak tizedestört alakban is. Ha a konkrét n-nel és p-vel van valami részletszámolás, akkor az előző pont is rendben van).
(b) (6 pont)
(1 pont) A keresett valószínűség $\mathbb{P}(4 < X < 8)$
(1 pont) $= \mathbb{P}(X = 5) + \mathbb{P}(X = 6) + \mathbb{P}(X = 7)$ (ha ez megvan, akkor jár az előző pont is)
(3 pont) $= \binom{10}{5}(\frac{1}{20})^5(\frac{19}{20})^5 + \binom{10}{6}(\frac{1}{20})^6(\frac{19}{20})^4 + \binom{10}{7}(\frac{1}{20})^7(\frac{19}{20})^3$ (az is elég, ha ez n-nel és p-vel kifejezve van meg és már elmondta, mi n és p értéke; az egyes helyes tagokért jár részpont)
(1 pont) $\approx 6{,}369 \cdot 10^{-5} = 0{,}00006369$. (Elég a végeredmény, mivel a számológép tud binomiális együtthatókkal számolni, de egyébként $\mathbb{P}(X = 5) = 6{,}0935 \cdot 10^{-5}$, $\mathbb{P}(X = 6) = 2{,}6726 \cdot 10^{-6}$, $\mathbb{P}(X = 7) = 8{,}0379 \cdot 10^{-8}$, utóbbi a négy értékes jegyre való kerekítésnél el is hagyható. Ha ezek a numerikus értékek megvannak, de az előző részben nincsenek meg a súlyfüggvény értékei, akkor abból a 3 pontból meg lehet adni utólag max. 2-t.)
(c) (8 pont)
(1 pont) arra, hogy Poisson-approximációt kell használni (még ha esetleg nem is sikerül, pl. nem tudja, hogy melyik változót közelítse Poissonnal vagy milyen paraméterű Poissonnal közelítse). Nem baj, ha ezt nem írja oda, feltéve, hogy jól alkalmazza a Poisson-approximációt (ide értve azt is, hogy λ-t elvileg helyesen határozza meg), de ha nem jól csinálja, akkor ez az 1 pont sem jár.
(1 pont) arra, hogy miért azt kell használni: elég valami olyasmi, hogy n nagyon nagy $p(= p_n)$ pedig nagyon kicsi (és a kettő szorzata egy pozitív számhoz tart és egy ilyen paraméterekkel rendelkező binomiális eloszlású val. változót közelítünk, de ezeket nem kötelező odaírni).
(1 pont) A hibás izzók száma egy Y Poisson-eloszlású valószínűségi változóval közelíthető,
(2 pont) amelynek paramétere $\lambda = \lim_{n\to\infty} np_n = \frac{1}{2}$. (A p_n jelölés nem kötelező, sőt ha azt írja, hogy np mindig 1/2, akkor a határértéket sem kötelező odaírni.)
(1 pont) Így tehát a keresett (közelítő) valószínűség $\mathbb{P}(Y = 3)$
(1 pont) $= \frac{\lambda^3}{3!}e^{-\lambda} = \frac{(1/2)^3}{3!}e^{-1/2}$ (elég a két alak közül az egyik. Az $(1/2)^3$ helyett szabad 1/8-ot, a 3! helyett 6-ot írni, de nem elég a puszta végeredmény, ha csak az van, akkor -1 pont)
(1 pont) $\approx 0{,}0126$.

Ha CHT-t akar használni (de nem nagyon akar összejönni), akkor a (c)-ben maximum 3 pont. Önmagában arra az ötletre, hogy használjuk a CHT-t, nem jár pont, és arra sem, hogy miért használhatjuk és hogy hivatkozunk rá (hiszen rossz). Viszont jár 1 pont arra, ha jól standardizál (vagyis a várható értéket vonja le, aztán pedig a szórással oszt le, ide értve a jobb oldal átalakítását). 1 pont arra, ha n és p behelyettesítésével – akár konkrét n-re és p-re, akár határértékben helyes határértékkel – kap egy numerikus eredményt a standardizált jobb oldalra. Még 1 pont, ha normális eloszlással közelíti a standardizált val. változót ÉS nem a $\mathbb{P}(X = 3)$-at akarja a normális eloszlásból kiolvasni, hanem valamilyen, folytonos val. változó esetén is viszonylag értelmes valószínűséget, például $\mathbb{P}(2{,}5 < X < 3{,}5)$-et vagy $\mathbb{P}(2 < X < 4)$-et (standardizálás után).

4. Legyenek $X \sim \mathrm{N}(1; 9)$ és $Y \sim \mathrm{N}(2; 16)$ független, normális eloszlású valószínűségi változók.

   (a) Mennyi $\mathbb{P}(X > 0{,}5)$?

   (b) Határozzuk meg $X + 2Y$ és $X - 4Y$ kovarianciáját és korrelációját.

   (c) Milyen eloszlású a $2X + Y$ valószínűségi változó és milyen paraméterekkel (és miért)?

Megoldás.

(a) (6 pont)
(1+1 pont) Sztenderdizálunk (ezt nem muszáj szövegesen odaírni):

$$\mathbb{P}(X > 0{,}5) = \mathbb{P}(\frac{X - \mathbb{E}(X)}{\mathbb{D}(X)} > \frac{0{,}5 - \mathbb{E}(X)}{\mathbb{D}(X)} = \mathbb{P}(\frac{X-1}{3} > \frac{0{,}5-1}{3})$$

(a két lépés összevonható, ha a második jó)
(1 pont) $= \mathbb{P}(Z > -1/6)$, ahol $Z \sim \mathrm{N}(0;1)$ (az 1/6 helyett 4 értékes jegyre kerekített numerikus eredmény is jó)
(1 pont) $= \mathbb{P}(Z < 1/6) = \Phi(1/6)$ (elég a kettő közül az egyik, és nem kell a folytonosságra hivatkozni, mert a sűrűségfüggvény szimmetriája miatt a két integrál ugyanaz)
VAGY: $= 1 - \mathbb{P}(Z \leq -1/6) = 1 - \Phi(-1/6) = 1 - (1 - \Phi(1/6)) = \Phi(1/6)$ (a Φ jelölés nem kötelező, és a folytonosságra itt sem jár pont, mert az egészre összesen 1 pont jár, emiatt bármelyik lépésben szabad váltani a $\leq$ és a $<$ között)
(1 pont) $\approx \Phi(0{,}17)$
(1 pont) $\approx 0{,}5675$.
(Ne vonjunk le pontot, ha csak az a baj, hogy valamelyik $\approx$ helyett =-t ír. Az utolsó két lépés összevonható, ha előtte az 1/6 már megjelent ÉS jó a kiolvasott végeredmény.)
(b) (11 pont)
(2 pont) $\mathrm{cov}(X+2Y, X-4Y) = \underbrace{\mathrm{cov}(X,X)}_{=\mathbb{D}^2(X)} + 2\mathrm{cov}(X,Y) - 4\mathrm{cov}(X,Y) - 8\underbrace{\mathrm{cov}(Y,Y)}_{=\mathbb{D}^2(Y)}$.
(Ebből 1 pont, ha a bilinearitást jól használja, ehhez elég 4 helyes tagra bontani az összeget további egyszerűsítések nélkül. Még 1 pont, ha a konstans szorzókat is jól emeli ki, a szimmetriát jól használja és tudja, hogy egy val. változó önmagával vett kovarianciája a szórásnégyzete.
(1 pont) $\mathrm{cov}(X,Y) = 0$,
(1 pont) mert függetlenek.
(1 pont) Ezért (mivel a szórásnégyzet a normális eloszlás második paramétere – erre nem jár pont, mert benne van a nevezetes eloszlás táblázatban) $\mathrm{cov}(X+2Y, X-4Y) = 9 - 8 \cdot 16 = -119$.
(Ha minden szükséges részeredmény szerepelt már, akkor itt elég a végeredmény.)
(0 pont) A korreláció meghatározásához most kiszámítjuk $X+2Y$ és $X-4Y$ szórásait is.
(1 pont) Mivel X és Y függetlenek,
(1 pont) $\mathbb{D}^2(X+2Y) = \mathbb{D}^2(X) + \mathbb{D}^2(2Y)$
(1 pont) $= \mathbb{D}^2(X) + 4\,\mathbb{D}^2(Y) = 9 + 64 = 73$ (Ha a sor eleji lépés megvan, akkor ez a pont az előzővel összevonható, de ha a végeredmény nem jó, akkor csak az előző pont jár.
Ha pedig az előző lépés végéig megvannak a részeredmények és a végeredmény jó, akkor higgyük el, hogy így számol, de rossz végeredmény esetén ez az egy pont nem jár).
(0 pont) Hasonlóan (itt már nem muszáj a függetlenségre hivatkozni)
(1 pont) $\mathbb{D}^2(X-4Y) = \mathbb{D}^2(X) + \mathbb{D}^2(-4Y)$
(1 pont) $= \mathbb{D}^2(X) + 16\,\mathbb{D}^2(Y) = 9 + 16 \cdot 16 = 265$ (ha minden korábbi részletszámolás jó, akkor itt már elég a végeredmény).
(0 pont) arra, hogy a szórás a szórásnégyzet négyzetgyöke. Erre már a 2. (b)-ben adtunk pontot. Aki ott nem kapott rá pontot és itt leírta ezt, utólag kérhet a 2. feladatra 1 pontot.
(1 pont) Tehát $\mathrm{corr}(X+2Y, X-4Y) = \frac{\mathrm{cov}(X+2Y, X-4Y)}{\mathbb{D}(X+2Y)\,\mathbb{D}(X-4Y)} = \frac{-119}{\sqrt{73} \cdot \sqrt{265}} \approx -0{,}8556$ (ha már minden részeredmény megvan, akkor a korreláció általános képlete után elég a numerikus végeredmény, akkor is OK, ha csak a szórásnégyzetek vannak meg és nem a szórások. Az általános képlet helyett pedig elég a konkrét esetbeli képlet, ha azonosítható a benne szereplő számok szerepe).
(c) (3 pont)
(1 pont) Független normálisok összege normális és a paraméterek összeadódnak. Normális lineáris transzformáltja normális (vagy konstans – ezt nem kell odaírni).
Ezekre összesen 1 pont jár. Ezek általános paraméterekkel is kifejezhetőek, de ez a pont csak akkor jár, ha hivatkozik a függetlenségre is. Ha nem írja le, hogy a paraméterek összeadódnak, de később összeadja őket, akkor ez a pont is jár. Ha nem hivatkozik a lineáris transzformációra,

de kimondja, hogy $2X$ normális eloszlású és minden más rendben van, akkor is adjuk meg ezt az 1 pontot.
(1 pont) Így $2X \sim \mathrm{N}(2;36)$, mivel az eloszlás első paramétere $\mathbb{E}(2X) = 2\,\mathbb{E}(X) = 2$, a második $\mathbb{D}^2(2X) = 4\,\mathbb{D}^2(X) = 36$ és
(indoklás nélkül ez az 1 pont nem jár, de ha annyit odaír, hogy az első paraméter a várható érték és a második a szórás, és jó numerikus eredményt ad, akkor fogadjuk el ezért az 1 pontért a megoldást)
(1 pont) $2X + Y \sim \mathrm{N}(4;52)$.
Az, hogy $2X$ milyen eloszlású, nem szükséges, amennyiben szerepel, hogy normális lineáris transzformáltja normális; ebben az esetben elég a $2X$ várható értékére és szórásnégyzetére vonatkozó indoklás (ami összevonható ugyanezzel a $2X + Y$-ra vonatkozóan, ha részletesen indokol).

5. Az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2}, & \text{ha } 0 < x < 2 \text{ és } 1 + x < y < 1 + 2x, \\ 0, & \text{különben.} \end{cases}$$

   (a) Mi Y értékkészlete? *Ehhez a részfeladathoz kivételesen nem kell indoklás, a többihez igen.*
   (b) Határozzuk meg az $\mathbb{R} \to [0,\infty), y \mapsto f_{Y|X}(y|x)$ feltételes sűrűségfüggvényt minden olyan $x \in \mathbb{R}$ esetén, amelyre $f_X(x) > 0$.
   (c) Ez alapján határozzuk meg az $\mathbb{E}(Y|X)$ regressziót.
   (d) Független-e X és Y?

Megoldás. (a)
(2 pont) $\mathrm{Ran}(Y) = (1,5)$.
(Kisebb szintaktikai hibákért ne vonjunk le pontot, azért se, ha nyílt vagy félig nyílt intervallumot ír.)
(Indoklás nem kell, de a teljesség kedvéért ideírom: az együttes sűrűségfüggvény képletéből látható (X,Y) értékkészlete az $y = 1 + x$, $y = 1 + 2x$ és $x = 2$ egyenletű egyenesek közé eső [tompaszögű] háromszög [és mivel itt konstans a sűrűségfüggvény és mindenhol máshol 0, ezért (X,Y) egyenletes eloszlású ezen a háromszögön]. Ebben a háromszögben a pontok Y koordinátája lehetséges értékeinek halmaza $(1,5)$.)

(b) (7 pont)
(1 pont) Ha $f_X(x) > 0$, akkor $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$ (ez OK az $f_X(x) > 0$ kitétel nélkül is).
(1 pont) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$,
(1 pont) ami $x \in (0,2)$ esetén a következővel egyenlő: $= \int_{1+x}^{2+x} \frac{1}{2}\mathrm{d}y$
(1 pont) $= \left[\frac{y}{2}\right]_{y=1+x}^{1+2x}$
(1 pont) $= \frac{1+2x-(1+x)}{2} = \frac{x}{2}$ (elég a végeredmény, ha a jó primitív függény megvolt),
(1 pont) $x \notin (0,2)$ esetén pedig 0-val egyenlő (mert a konstans 0 függvényt integráljuk y szerint, de ezt nem kell odaírni).
Úgy is jó, ha $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$-t csak az $f_X(x) > 0$ (avagy $x \in (0,2)$) esetben írja oda és csak közli, hogy minden más esetben 0 a sűrűségfüggvény értéke, de az utóbbi 1 pontért annak ki kell derülnie a megoldásból (legalább impliciten), hogy pontosan a $(0,2)$ az a halmaz, ahol a feltételes sűrűségfüggvényt a feladat szerint meg kell határozni. Indoklás nem kell arra ebben az esetben sem, hogy miért 0 az $f_X(x)$ értéke, ha $x \notin (0;2)$.
(1 pont) Tehát $f_X(x) > 0$ esetén $f_{Y|X}(y|x) = \frac{1/2}{x/2} = \frac{1}{x}$ (ha minden részeredményt leírt már, akkor itt elég a végeredmény).
Ha észreveszi, hogy (X,Y) egyenletes eloszlású a háromszögön, akkor szabad a feltételes sűrűségfüggvény szemléletes jelentésével – és geometriai valószínűségi mezős számolással – érvelni. Ezt itt nem részletezzük.

(c) (7 pont)
1. (sztenderd) megoldás.
(1+1 pont) $\mathbb{E}(Y|X) = g(X)$, ahol $g(x) = \mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\mathrm{d}y$ (1 pont a regressziós függvény x-beli értékének képletére, 1 pont pedig arra, hogy a regressziós függvény képletébe írva kapjuk a regressziót. Az $\mathbb{E}(Y|X)$ jelölésnek valahogy meg kell jelennie a megoldásban, mert ezt a regressziót kérdezzük a feladatban, de a g vagy az $\mathbb{E}(Y|X = x)$ jelölések lehet helyett lehet bármi más is, ami nem helytelen)
(1 pont) $= \int_{1+x}^{1+2x} \frac{1}{x} y\mathrm{d}y$ (az előző lépéssel összevonható)
(0 pont) $= \frac{1}{x} \int_{1+x}^{1+2x} y\mathrm{d}y$
(1 pont) $= \frac{1}{x}[y^2/2]_{y=1+x}^{y=1+2x}$
(1 pont) $= \frac{(1+2x)^2-(1+x)^2}{2x}$
(1 pont) $= \frac{(2+3x)x}{2x} = 1 + \frac{3}{2}x$ [két négyzet különbsége —- nem kell indoklás]
(avagy: az előző pontnál lévő kifejezés $= \frac{1+4x+4x^2-1-2x-x^2}{2x} = 3x^2 + 2x2x = 1 + \frac{3}{2}x$).
Ha kijön bármilyen helyes képlet a primitív függvény alapján, de nincs egyszerűsítve, akkor az utolsó pont nem jár, de addig OK.
Megjegyzés: Ez a képlet csak akkor érvényes, ha $f_X(x) > 0$, de ezt nem kell odaírni, mert a regresszió szempontjából az $f_X(x) = 0$ eset érdektelen (mert $\mathbb{E}(Y|X)$ csak egy valószínűséggel egyértelműen van definiálva). Ne vonjunk le még egy pontot, ha a (b)-ben már levontunk azért, mert nem mondja meg, hogy $f_X(x)$ mikor pozitív.
(1 pont) Tehát $\mathbb{E}(Y|X) = 1 + \frac{3}{2}X$ (ha bonyolultabb formában megadott regressziós függvénybe helyettesíti X-et, az is OK).

2. (intuitívabb, kevesebb algebrai átalakítást, de cserébe némi ötletet és indoklást is igénylő) megoldás.
(3 pont) A feltételes sűrűségfüggvény fenti képletéből látható, hogy $x \in (0, 2)$ (avagy $f_X(x) > 0$) esetén $X = x$-et feltéve Y egyenletes eloszlású az $(1 + x, 1 + 2x)$ intervallumon. (A feltételes sűrűségfüggvény konstans ott, ahol nem 0 – de ezt nem kell jobban indokolni, mert az egyenletes eloszlás benne van a táblázatban. De a feltételes sűrűség- vagy eloszlásfüggvényre hivatkozni kell. Pontatlan megfogalmazás esetén lehet részpontot adni.)
(1 pont) Egyenletes eloszlás várható értéke a vonatkozó intervallum középpontja (képlettel is jó: $(a + b)/2$ az $\mathrm{U}(a; b)$ esetén),
(1 pont) ezért $\mathbb{E}(Y|X = x) = \frac{1+x+1+2x}{2} = 1 + \frac{3}{2}x := g(x)$.
(1 pont) arra, hogy $\mathbb{E}(Y|X) = g(X)$ (elég az egyik fajta jelölés, a g nem kötelező, és ha jó az $\mathbb{E}(Y|X = x)$, akkor utána elég $\mathbb{E}(Y|X)$-re a végeredményt megadni és akkor is jár ez a pont)
(1 pont) $= 1 + \frac{3}{2}X$.
2b. megoldás: mint a 2., de $X = x$ elhagyásával, közvetlenül X-re feltételezve.
Itt azt kell mondani, hogy ha X értékét ismerjük (és valahogy kiderül, akár a (b)-ből, hogy csak az az eset releváns, amikor a $(0, 2)$ intervallumba estik), akkor $Y \sim \mathrm{U}(1 + X; 1 + 2X)$. Erre már jár 4 pont, és utána 3 pont arra, hogy emiatt $\mathbb{E}(Y|X) = 1 + \frac{3}{2}X$, amiből 1 pont az indoklás (intervallum közepe/általános képlet stb.).

(d) (4 pont)
Annak a puszta közlésére, hogy nem függetlenek, nem jár pont. Teljesen rossz indoklással sem jár pont rá.
1. [ajánlott] megoldás.
(2 pont) Ha X és Y függetlenek, akkor $\mathbb{E}(Y|X) = \mathbb{E}(Y)$.
(1 pont) Tehát nem lehetnek függetlenek, mert ez itt nem teljesül,
(1 pont) mert... (valahogy megindokolja, pl.: $\mathbb{E}(Y)$ egy konstans, – az értékét nem kell kiszámolni – míg a kapott $\mathbb{E}(Y|X) = 1 + \frac{3}{2}X$ nem [egy valószínűséggel] konstans).
2. megoldás.
(2 pont) Ha X és Y függetlenek, akkor $f_{Y|X}(y|x) = f_Y(y)$ minden olyan esetben, amikor $f_X(x) > 0$ (ha csak az utóbbi kitétel hiányzik, azért ne vonjunk le pontot, viszont arra nem le-

het hivatkozni, hogy azért nem függetlenek, mert $f_X(x) = 0$ esetén $f_{Y|X}(y|x)$ is nulla és emiatt nem egyezik meg $f_Y(y)$-nal).
(1 pont) Ez itt viszont nem teljesülhet,
(1 pont) mert... (valahogy megindokolja, pl.: $f_{Y|X}(y|x) = 1/x$ jól láthatóan függ x-től, nem kell túl részletes indoklás).
3. megoldás.
(4 pont) Kiszámolja $f_Y(y)$-t (x szerint integrálja az együttes sűrűségfüggvényt) és mutat egy olyan (x, y) párt, amelyre $f_{X,Y}(x, y) \neq f_X(x)f_Y(y)$. Ebben az esetben kell $f_Y(y)$ képlete, vagy legalább annyi számítás, ami megmutatja, hogy az egyenlőség nem teljesülhet.
4. megoldás.
(4 pont) A 3.-hoz hasonló, de sűrűségfüggvények helyett eloszlásfüggvényekkel: valamilyen (x, y) pontot mutat, amire $F_{X,Y}(x, y) \neq F_X(x)F_Y(y)$ (vagy a 2. megoldáshoz hasonlót csinál feltételes eloszlásfüggvénnyel). Ez lényegesen több számolást igényel, viszont nem feltétlenül a 3. megoldás szerint kapott sűrűségfüggvényeket kell integrálni, pl. $F_Y(y) = \mathbb{P}(Y < y)$ a folytonos TVT segítségével is megkapható, X-re való feltételezéssel, a már kiszámolt feltételes sűrűségfüggvénnyel. Továbbá geometriai valószínűségi mezős érveléssel egyszerűbben is megkapható (rajzoljuk fel az $\{Y < y\}$ eseményt).
5. megoldás.
(4 pont) Kiszámítja X és Y kovarianciáját és megállapítja, hogy az nem 0, ezért X és Y nem lehet független. Itt az ötletre max. 1 pont jár és max. még 1 pont a kovariancia képletére és $\mathbb{E}(XY)$ képletére folytonos esetben együtt, a többi ponthoz kell a számolás.

6. * Legyen $n \in \mathbb{N}$ és legyenek $X_1, \dots, X_n$ független, azonos eloszlású, $\mathrm{U}(-\vartheta; \vartheta)$ eloszlású valószínűségi változók, ahol $\vartheta > 0$ egy ismeretlen paraméter.

   (a) Az $(X_1, \dots, X_n)$ minta rögzített $(x_1, \dots, x_n)$ realizációja esetén határozzuk meg a likelihood-függvény $L_\vartheta(x_1, \dots, x_n)$ értékét.

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

   Megoldás. Ez a példa nagyon hasonló a jegyzetkiegészítésben szereplő 14.2.10. Példához és a honlapon közzétett 2024.01.25-i vizsga 6-os feladatához.
   (a) (7 pont)
   (3+1 pont) (Az egyenletes eloszlás tulajdonságai miatt) az f_ϑ sűrűségfüggvény

$$f_\vartheta(x) = \begin{cases} \frac{1}{2\vartheta}, & \text{ha } x \in (-\vartheta, \vartheta), \\ 0, & \text{különben.} \end{cases}.$$

   (Az 1. sorban 1 pont a képletre, 1–1 a határokra, 1 pont a 2. sorra. Ha a $-\vartheta$ vagy a ϑ pontot a másik szakaszhoz sorolja, azért ne vonjunk le pontot, de ha nem sorolja be őket, akkor -1 pont Ha a ϑ konstans szorzója (2) rossz, de minden más jó, akkor csak 1 pontot vonunk le.)
   (3 pont) Így a likelihood-függvény $(X_1, \dots, X_n)$ bármely $(x_1, \dots, x_n)$ realizációja esetén

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^{n} f_\vartheta(x_i) = \prod_{i=1}^{n} \frac{1}{2\vartheta} = \frac{1}{(2\vartheta)^n}.$$

   Ebből 1 pont az első lépésre, 1 a másodikra és 1 a harmadikra, viszont jó végeredmény esetén az utóbbi kettő összevonható. Az is OK, ha valahol a (b)-ben írja át a produktumot hatvánnyá.
   Megjegyzés: Nem muszáj odaírni, hogy ez a képlet akkor érvényes, ha minden x_i a $(-\vartheta, \vartheta)$ intervallumban van, mivel ez a realizáció definíciójából következik. (Fentebb az f_ϑ sűrűségfüggvénynél viszont fontos, hogy az $\frac{1}{2\vartheta}$ képlet csak a $(-\vartheta, \vartheta)$ intervallumon érvényes.)
   (b) (13 pont)
   (3 pont) Az intervallum hosszát mindaddig csökkenthetjük (és ezzel a likelihood-függvényt növelhetjük), amíg az $(x_1, \dots, x_n)$ az $(X_1, \dots, X_n)$ egy realizációja marad

(arra az ötletre, hogy az intervallum hossza legyen minimális, még ha nem is világos, hogy milyen értelemben és hogyan kell ezt elérni, már lehet adni 1 pontot. A realizáció tulajdonság említése helyett jó az is, hogy „amíg az összes x_i a $(-\vartheta;\vartheta)$ intervallumban marad”),
(2 pont) azaz amíg

$$x_n^* = \max\{x_1,\dots,x_n\} < \vartheta$$

(Elég a maximumos és x_n^* jelölések közül az egyik. Itt az egyszerűség kedvéért a legvégéig a rendezett mintás jelölést használjuk, de végig jó a maximumos értelemszerűen. Ha $\leq$-vel írja, az nem baj, hasonlóan a későbbiekben is.)
(2 pont) ÉS

$$x_1^* = \min\{x_1,\dots,x_n\} > -\vartheta.$$

Ha nem világos a megoldásból az „ÉS” kapcsolat, akkor az utóbbi 4 pontból max. 2 jár, kivéve ha a megoldás későbbi része tisztázni tudja, hogy a hallgató „ÉS”-re gondolt.
(2 pont) Így a likelihood-függvény éppen a határesetben lesz maximális, amikor vagy x_n^*, vagy $-x_1^*$ egyezik meg ϑ-val,
(1 pont) úgy, hogy a kettő közül a másik a $(-\vartheta,\vartheta)$ intervallum belsejébe essen (vagy szintén pont a határra, ami 0 valószínűséggel fordul elő, ezért nem kell megemlíteni),
(1 pont) vagyis ha $\vartheta = \max\{-x_1^*, x_n^*\}$. (Szövegesen is jó.)
Ha megvan a jó végeredmény, akkor az utóbbi 4 pontból jár minimum 3 pont akár indoklás nélkül is, és csak azt kell indokolni, hogy a likelihood-függvény a vizsgált határesetben lesz maximális és ezért ezt az esetet választjuk.
(1 pont) Ezért az $(x_1,\dots,x_n)$ realizációhoz tartozó maximum likelihood becslésünk

$$\vartheta_*(x_1,\dots,x_n) = \max\{x_n^*, -x_1^*\} = \max\{\max\{x_1,\dots,x_n\}, -\min\{x_1,\dots,x_n\}\}$$

(elég az egyik fajta jelölés; ártatlanabb szintaktikai hibáért ne vonjunk le pontot, csak súlyosabbért),
(1 pont) a maximum likelihood becslésünk pedig

$$\vartheta_*(X_1,\dots,X_n) = \max\{X_n^*, -X_1^*\} = \max\{\max\{X_1,\dots,X_n\}, -\min\{X_1,\dots,X_n\}\}.$$

Megjegyzés: $-x_1^*$ helyett végig lehet $|x_1^*|$-ot írni. Ugyanis bár előfordulhat, hogy x_1^* pozitív, de ilyen esetben az összes x_i pozitív és $x_n^* > x_1^* = |x_1^*| > -x_1^*$, így $\max\{|x_1^*|, x_n^*\} = \max\{-x_1^*, x_n^*\} = x_n^*$, különben pedig $-x_1^* = |x_1^*|$, vagyis bár $-x_1^*$ nem ugyanaz a statisztika, mint $|x_1^*|$, de a kapott ML-becslés minden realizáció esetén ugyanaz a két esetben. Ugyanígy végiggondolható, hogy ha x_n^* helyett $|x_n^*|$-ot írunk, az nem változtat az ML-becslésen, mert ha $x_n^* < 0$, akkor az összes x_i negatív és $-x_1^* = |x_1^*| > |x_n^*| > x_n^*$, ezért $\max\{-x_1^*, x_n^*\}$ és a $\max\{-x_1^*, |x_n^*|\}$ képlet is $-x_1^*$-ot adja. Ezeket a maximális pontszámért nem kell végiggondolni. Elfajuló eset, amikor minden mintaelem 0, de ez 0 valószínűséggel fordul elő, akármi is $\vartheta > 0$, ezért ezzel nem kell foglalkozni (és a kapott végeredmény ilyenkor is jó, csak a $(-\vartheta;\vartheta)$ intervallum hosszát 0-val becsüljük).

További megjegyzések a javításhoz:

- Ha a maximumot összekeveri a minimummal vagy az előjeleket rontja el, de konzisztensen csinálja, akkor összesen 3–3 pont levonás.
- Ha arra hivatkozik, hogy ez ugyanaz, mint a jegyzetkiegészítésbeli példa, és ebből kihozza $-x_1^*$-ot végeredménynek, akkor max. 4 pont. Ha arra, hogy ez ugyanaz, mint a tavalyi vizsgapélda és ebből kihozza x_n^*-ot végeredménynek, akkor max. 3 pont (vizsgafeladatra nem szabad hivatkozni).
- Ha a b)-ben log-likelihood-függvényes módszerrel próbálkozik, a log-likelihood-függvény helyes levezetéséért 2 pont jár, ϑ szerinti parciális deriválására 2 pont. Annak megállapítására, hogy a deriváltnak sehol nincs zérushelye ott, ahol $(x_1,\dots,x_n)$ realizáció, jár 2 pont. Ez a 6 pont akkor is jár, ha nem vesz logaritmust, hanem az eredeti likelihood-függvényt deriválja (4 pont) és megállapítja, hogy nincs zérushely (2 pont). Annak megállapítására, hogy emiatt meg kell nézni a határokat, jár 1 pont. Ezután attól a ponttól kezdve, ahol rá kell jönni, hogy vagy x_n^*-nak, vagy $-x_1^*$-nak kell megegyeznie ϑ-val, pontozás a fenti mintamegoldáshoz hasonlóan.