Vizsgadolgozat

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 $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező.

   (a) Ha $A, B \in \mathcal{F}$ események, mikor mondjuk azt definíció szerint, hogy A és B függetlenek?

   (b) Ha $X, Y \colon \Omega \to \mathbb{R}$ valószínűségi változók, mikor mondjuk azt definíció szerint, hogy X és Y függetlenek?

   (c) Ha $X, Y \colon \Omega \to \mathbb{R}$ független valószínűségi változók, hogyan számíthatjuk ki a szorzatuk várható értékét a két változó várható értékei segítségével (feltéve, hogy mindhárom említett várható érték létezik)?

Megoldás.
(a) (5 pont) Ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.
(b) (7 pont)
(5 pont) Ha az $\{X < x\}$ és $\{Y < y\}$ események függetlenek
(2 pont) minden $x, y \in \mathbb{R}$ esetén.
Természetesen meg lehet fogalmazni úgy is, hogy $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\,\mathbb{P}(Y < y)$, minden $x, y \in \mathbb{R}$ esetén. Eloszlásfüggvény-jelöléssel is OK.
Ha „$<$” helyett „$\leq$”, „$\geq$” vagy „$>$” szerepel, akkor max. 5 pont a 6-ból, és a kvantorokra jár még max. 2. Ha egyenlőség, akkor max. 2+2 pont. Ha olyan definíciót ír, ami együttesen folytonos esetben ekvivalens a függetlenséggel (pl. $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ minden $x, y \in \mathbb{R}$ esetén), akkor is max. 2+2 pont.
(c) (8 pont)
(4 pont) Ha páronként kizáróak (szintén OK: páronként diszjunktak, szintén OK képlettel, ha odaírja, hogy csak a nem egyenlő indexekre vonatkozik – ha nem írja oda, akkor -1 pont. Ha szöveggel írja és nem írja oda, hogy páronként, akkor max. 2 pont)
(4 pont) és az uniójuk Ω (szintén OK: az uniójuk valószínűsége 1. Ha azt írja, hogy a valószínűségek összege 1, arra max. 3 pont, ha odaírja, hogy páronként kizáróak, és ha nem írja oda, akkor max. 1 pont).
Részpontok csak minimális szintaktikai hiba esetén járnak.

2. Az X és Y valószínűségi változók együttes eloszlását az alábbi táblázat tartalmazza, ahol a c és d pozitív valós számok értékét úgy kell megválasztani, hogy $\mathbb{E}(X) = \mathbb{E}(Y) = 1$ teljesüljön.

$Y \backslash X$	0	2
0	$\frac{1}{8}$	$\frac{1}{8}$
1	c	d
2	$\frac{1}{8}$	$\frac{1}{8}$

(a) Határozzuk meg c és d értékét.

(b) Független-e X és Y?

(c) Határozzuk meg az $X + Y$ valószínűségi változó szórásnégyzetét.

Ha az (a) részt nem sikerül megoldani, a (b) és a (c) részt paraméteresen is megoldhatjuk. Ugyanez vonatkozik a 4-es feladatra is.

Megoldás.
(a) (8 pont)
(4 pont) Meghatározzuk a peremeloszlásokat (azaz X és Y perem-súlyfüggvényeit; nem muszáj táblázatosan). X súlyfüggvényét a megfelelő oszlopösszegek, Y-ét pedig a megfelelő sorösszegek adják.

$Y \backslash X$	0	2	p_Y
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
1	c	d	$c+d$
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
p_X	$\frac{1}{4}+c$	$\frac{1}{4}+d$	$(\frac{1}{2}+c+d=1)$

Indokolni nem kell, de ha nem hibátlan a táblázat, akkor részpontok csak akkor járnak, ha van valami szöveges indoklás vagy a számolásból látható, hogy alapvetően jól csinálja, csak elszámolja. 1 pont arra, hogy X peremeloszlását hogyan kapjuk, 1 pont arra, hogy Y-ét hogyan, és 1–1 pont arra, hogy tényleg jól ki is számolja ezeket.
(1 pont) $1 = \mathbb{E}(X) = 2 \cdot (\frac{1}{4} + d) = \frac{1}{2} + 2d$.
(2 pont) $1 = \mathbb{E}(Y) = 1 \cdot (c + d) + 2 \cdot \frac{1}{4} = c + d + \frac{1}{2}$.
(0 pont) Innen $c + d = \frac{1}{2}$ és $c + d = 2d$.
(1 pont) Tehát $c = d = \frac{1}{4}$.

(b) (5 pont)
(1 pont) Akkor lesznek függetlenek, ha $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y)$ teljesül (súlyfüggvény-jelöléssel is OK)
(1 pont) minden $x \in \mathrm{Ran}(X)$ és $y \in \mathrm{Ran}(Y)$ esetén.
(1 pont) Behelyettesítve az (a)-ban kapott értékeket, az alábbi táblázatot kapjuk:

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

(1 pont) Valami minimális indoklást ír, akár konkrét számokkal, amiből (elég, ha impliciten) kiderül, hogy tudja, hogy $\mathbb{P}(X = x, Y = y)$ a $p_X(x)$-hez tartozó oszlop és a $p_Y(y)$-hoz tartozó sor kereszteződésénél látható valószínűség.
(1 pont) A feltétel minden x és y esetén teljesül (ahol x lehet 0 vagy 2, y pedig $0, 1$ vagy 2 –

ezt nem kell odaírni), tehát függetlenek.
Paraméteres számolás esetén az (a)-beli táblázat alapján az a válasz jön ki, hogy ha $c = d = \frac{1}{4}$, akkor függetlenek, különben meg nem.

(c) (7 pont)
(1 pont) Mivel függetlenek,
(1 pont) $\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$.
(1 pont) Tudja, hogy $\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, és hasonlóan Y-ra.
(0 pont) $\mathbb{E}(X)^2 = \mathbb{E}(Y)^2 = 1$ az (a)-beli feltétel alapján.
(1 pont) $\mathbb{E}(X^2) = (0^2 \cdot \frac{1}{2} +) 2^2 \cdot \frac{1}{2} = 2$.
(0 pont) Tehát $\mathbb{D}^2(X) = 1$.
(2 pont) $\mathbb{E}(Y^2) = (0^2 \cdot \frac{1}{4} +) 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} = \frac{3}{2}$.
(0 pont) Tehát $\mathbb{D}^2(Y) = \frac{1}{2}$.
Ehelyett azt is lehet, hogy 1 pontért észreveszi, hogy Y binomiális eloszlású $n = 2$ és $p = \frac{1}{2}$ paraméterrel (nem kell indokolni), és ez alapján további 1 pontért: a szórásnégyzete $np(1-p) = \frac{1}{2}$.
(1 pont) Tehát $\mathbb{D}^2(X+Y) = \frac{3}{2}$ (ez az 1 pont jár az összes részletszámításra a második momentumok kiszámolásán kívül, beleértve X és Y szórásnégyzetének meghatározását).

Paraméteres számolás esetén $\mathbb{E}(X^2) = 1 + 4d$, $\mathbb{E}(Y^2) = c + d + 1$.
Viszont ha nem jött rá a (b)-ben, hogy függetlenek (akár azért, mert c és d értékét rosszul határozta meg az (a)-ban), akkor a (c)-t elvileg a $\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y) + 2\mathrm{cov}(X,Y)$ képlettel kell megoldani. 1 pont jár erre a képletre, és sajnos csak 1 pont a kovariancia meghatározására, ami csak akkor jár, ha tényleg sikeresen és kellő indoklással meghatározza – kell hozzá, hogy $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$ és $\mathbb{E}(XY) = \frac{1}{4} \cdot 2 \cdot 1 + \frac{1}{8} \cdot 2 \cdot 2 = 1$. Ebben az esetben nehezebbé vált a feladat.
Ha viszont nem ezt a képletet használja, hanem azt akarja használni, hogy $\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$, az nem jó, ezért ekkor a (c)-ből az első 2 pont nem jár, a többit viszont megszerezheti, mivel azok kellenének a helyes megoldáshoz is. Ebből az első 2 pontból 1-et csak akkor lehet szerezni, ha tudja, hogy függetlenek, de nem hivatkozik erre a képlet alkalmazásánál.
Természetesen meg szabad határozni $X+Y$ eloszlását is (konvolúciós képlettel vagy akár kézzel) és az alapján a szórásnégyzetét, erre is meg lehet kapni a 7 pontot.

3. Az egyetem nagy előadóiban található táblák mozgatógombjai egy adott pillanatban egymástól függetlenül 1/8 valószínűséggel üzemképtelenek.

   (a) Tegyük fel, hogy a HŐ jelű épületben 10 ilyen mozgatógomb van. Mennyi a valószínűsége, hogy ezen 10 gomb közül legalább három üzemképtelen? (A valószínűség pontos értékét adjuk meg.)

   (b) Tegyük fel, hogy a JÉG jelű épületben 100 ilyen gomb van. Közelítsük alkalmas approximációval annak a valószínűségét, hogy a 100 gomb közül legalább 10, de legfeljebb 15 darab üzemképtelen. (A valószínűség pontos értékét itt nem kell meghatározni.)

Megoldás.
(a) (9 pont)
(1 pont) Az üzemképtelen gombok számát jelöljük X-szel, ekkor X binomiális eloszlású
(1+1 pont) $n = 10$ és $p = 1/8$ paraméterekkel (elég annyi, hogy $\mathrm{Bin}(10; 1/8)$; indoklást itt nem várunk).
(1 pont) A keresett valószínűség $\mathbb{P}(X \geq 3)$
(1 pont) $= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2)$
(3 pont) $= 1 - (1 - 1/8)^{10} - 10 \cdot 1/8 \cdot (1 - 1/8)^9 - \binom{10}{2} \cdot (1/8)^2 \cdot (1 - 1/8)^8$ (a 10-es szorzó helyett persze jó $\binom{10}{1}$ is, és az első tagot is be lehet szorozni $\binom{10}{0}$-val; elég, ha impliciten derül ki, hogy

$\binom{10}{2} = 45$, és szabad eleve úgy írni, hogy $\frac{10\cdot 9}{2}$)
(0 pont) $= 1 - 0{,}2631 - 0{,}3758 - 0{,}2416$
(1 pont) $= 0{,}1195$ (elég ennyi a numerikus számolásból, ha megvannak az egyes tagok nem numerikusan).
(b) (11 pont)
Megoldás De Moivre–Laplace-tétellel (az általános CHT-s verzió pontozására lentebb térek ki):
(1 pont) Jelölje Y az üzemképtelen gombok számát, ekkor az (a)-hoz hasonlóan $Y \sim \mathrm{Bin}(100; 1/8)$.
(Ha már szerepelt az (a)-ban a binomiális eloszlás, és itt csak impliciten derül ki, hogy ennek az n-nek és p-nek megfelelően használja a De Moivre–Laplace-tételt, az nem baj).
(1 pont) A keresett valószínűség $\mathbb{P}(10 \le Y \le 15)$,
(1 pont) amit a De Moivre–Laplace-tétellel közelítünk (az nem baj, ha a „De” hiányzik, vagy a két szerző nevét fordított sorrendben írja),
(0 pont) mert Y binomiális eloszlású nagy n-nel és „nem túl kicsi” (n-től nem függő) p-vel.
(0 pont) Sztenderdizálunk (ezt nem kell így odaírni): $\mathbb{P}(10 \le Y \le 15)$
(1 pont) $= \mathbb{P}\left(\frac{10-np}{\sqrt{np(1-p)}} \le \frac{Y-np}{\sqrt{np(1-p)}} \le \frac{15-np}{\sqrt{np(1-p)}}\right)$
(1 pont) $= \mathbb{P}\left(\frac{10-100/8}{\sqrt{700/64}} \le \frac{Y-100/8}{\sqrt{700/64}} \le \frac{15-100/8}{\sqrt{700/64}}\right)$ (az előzővel összevonható, ha látszik, hogy honnan jönnek a számok – hibás számolás esetén indoklás nélkül részpont nem jár)
(numerikusan is OK: $100/8 = 12{,}5$, $700/64 = 175/16 = 10{,}9375$, ennek a négyzetgyöke 3,3072)
(1 pont) $\approx \mathbb{P}(-0{,}7559 \le Y \le 0{,}7559)$
(1 pont) $\approx \mathbb{P}(-0{,}7559 \le Z \le 0{,}7559)$, ahol $Z \sim \mathrm{N}(0; 1)$ [a De Moivre–Laplace-tétel miatt, amire már korábban járt pont]
(1 pont) $\overset{*}{=} \mathbb{P}(Z < 0{,}7559) - \mathbb{P}(Z < -0{,}7559) = \Phi(0{,}7559) - \Phi(-0{,}7559)$ (nem kötelező a Φ-s jelölés, viszont elég, ha azt adja meg),
(1 pont) $\approx 2\Phi(0{,}76) - 1$ (ez a pont együttesen jár arra, hogy a 0,7559-et közelíti a táblázatbeli legközelebbi hellyel és tudja, hogy $\Phi(-x) = 1 - \Phi(x)$)
(1 pont) ami a táblázatból kiolvasva közelítőleg $2 \cdot 0{,}7764 - 1$
(0 pont) $= 0{,}5528$.
(1 pont) A $*$-gal jelölt egyenlőség azért igaz, mert Z (vagy: Z eloszlásfüggvénye) folytonos. (Hivatalosan itt $\mathbb{P}(Z \le 0{,}7559) - \mathbb{P}(Z < -0{,}7559)$-et kellene írni, de az egyenlőség a folytonosság miatt elhagyható.) Ha csak annyit ír indoklásnak, hogy „folytonos”, az akkor fogadható el, ha jól kezeli a határokat, különben részpont sem jár.

Általános CHT-s megoldás esetén annyi a változás, hogy a CHT-re kell hivatkozni a De Moivre–Laplace-tétel helyett, és a binomiális eloszlás várható értéke/szórása helyett azt kell írni, hogy Y 100 darab független, azonos eloszlású indikátor összege, amelyek várható értéke 1/8 és szórásnégyzete 7/64.
Itt 1 pont a binomiális eloszlásra járó pont helyett arra, hogy Y független, azonos eloszlású indikátorok összege. 1 pont a De Moivre–Laplace helyett a CHT-re való hivatkozásra, és 1 pont a sztenderdizálás első lépése helyett valami olyasmire, hogy

$$\mathbb{P}(10 \le Y \le 15) = \mathbb{P}\left(\frac{10 - 10\,\mathbb{E}(X_1)}{\sqrt{10}\cdot\mathbb{D}(X_1)} \le \frac{Y - 10\,\mathbb{E}(X_1)}{\sqrt{10}\cdot\mathbb{D}(X_1)} \le \frac{15 - 10\,\mathbb{E}(X_1)}{\sqrt{10}\cdot\mathbb{D}(X_1)}\right.,$$

ahol $X_1, \ldots, X_{100}$ az indikátorok, amelyek összege Y. Írhat $10\,\mathbb{E}(X_1)$ helyett $\mathbb{E}(Y)$-t és $\sqrt{10}\,\mathbb{D}(X_1)$ helyett $\mathbb{D}(Y)$-t is. A következő 1 pont a várható érték és szórás helyes behelyettesítésére jár. Az ezutáni részek pontozása, illetve a keresett valószínűségre járó 1 pont ugyanúgy megy, mint az előző megoldásnál.
Ha az egyenlőséget már a normális eloszlásra való áttérés előtt elhagyja, akkor a keresett valószínűségre járó 1 pontot le kell vonni, illetve a folytonosságra való hivatkozásra sem jár a pont.

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

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

ahol $c > 0$ egy alkalmasan választott konstans.

(a) Határozzuk meg c értékét.

(b) Határozzuk meg X és Y kovarianciáját (X és Y várható értékeinek meghatározására járnak részpontok).

(c) Egyenletes eloszlású-e Y a $(0; 1)$ intervallumon?

Megoldás.
(a) (5 pont)
(0 pont) $f_{X,Y}$ akkor lesz együttes eloszlásfüggvény, ha nemnegatív (kikötöttük, hogy $c > 0$, ezért erre nem jár pont)
(1 pont) és $1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x$ (tetszőleges sorrendben lehet integrálni)
(1 pont) $= \int_0^c\int_0^1 x\mathrm{d}y\mathrm{d}x$ (az előzővel összevonható, de ha nem jó, akkor az előző hiányában részpont nem jár)
(1 pont) $= \int_0^c x\mathrm{d}x$ (nem kell ennél jobban részletezni a konstans 1 függvény integrálását dy)
(1 pont) $= [x^2/2]_{x=0}^c$
(0 pont) $= c^2/2$.
(1 pont) Innen $c^2 = 2$, tehát $c = \sqrt{2}$ (nem kell indoklás; c-nek nemnegatívnak kell lennie, ezért $c = -\sqrt{2}$ nem jó, de ezt nem kell leírni; ha valaki azt írja, hogy az is jó, akkor ez a pont nem jár).
(b) (11 pont)
Standard megoldás:
(1 pont) $\mathrm{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$.
(1 pont) $\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x$
(1 pont) $= \int_0^{\sqrt{2}}\int_0^1 x^2y\mathrm{d}y\mathrm{d}x$ ((az előzővel összevonható, de ha nem jó, akkor az előző hiányában részpont nem jár)
(1 pont) $= \int_0^{\sqrt{2}} x^2[y^2/2]_{y=0}^1\mathrm{d}x$
(0 pont) $= \int_0^{\sqrt{2}} x^2/2\mathrm{d}x$
(1 pont) $= [x^3/6]_{x=0}^{\sqrt{2}}$
(1 pont) $= (\sqrt{2})^3/6 = \dfrac{2\sqrt{2}}{6}$. (Mindkét alakban jó. Numerikus érték itt nem kell, mert szimbolikusan is könnyen ki fog jönni, hogy a kovariancia 0, de egyébként 0,4714 az értéke.)
(1 pont) Hasonlóan $\mathbb{E}(X) = \int_0^{\sqrt{2}}\int_0^1 x^2\mathrm{d}y\mathrm{d}x$ (itt már nem jár pont az általános alakra, de elég az is, ha utána jól folytatódik, pl. jó primitív függvénnyel jó határokon)
(1 pont) $= \int_0^{\sqrt{2}} x^2\mathrm{d}x = [x^3/3]_{x=0}^{\sqrt{2}} = (\sqrt{2})^3/3 = \dfrac{2\sqrt{2}}{3}$ (ez nagyon hasonló számolás az $\mathbb{E}(XY)$-hoz, ezért csak 1 pont jár rá, és elég, ha minimális indoklás van, pl. az itt szereplő lépésekből legalább 2, amiből a primitív függvényt vissza lehet követni és a végeredmény megvan).
(1 pont) És szintén hasonlóan $\mathbb{E}(Y) = \int_0^{\sqrt{2}}\int_0^1 xy\mathrm{d}y\mathrm{d}x$
(1 pont) $= \int_0^{\sqrt{2}} x[y^2/2]_{y=0}^1\mathrm{d}x = \int_0^{\sqrt{2}} x/2\mathrm{d}x$ (összevonható, itt már nem jár külön pont az y szerinti integrálásra, amire $\mathbb{E}(XY)$-nál járt)
(1 pont) $= [x^2/4]_{x=0}^{\sqrt{2}} = \dfrac{1}{2}$.
(0 pont) Tehát $\mathrm{cov}(X, Y) = 0$.

Alternatív megoldás, aminek köszönhetően a számolások nagy részére nincs szükség:
(3 pont) X és Y függetlenek (erre indoklás nélkül nem jár pont, indoklással is csak akkor, ha

az indoklásra legalább 1 pont még jár!),
(1 pont) mert (X, Y) értékkészlete egy téglalap,
(1 pont) aminek az oldalai párhuzamosak a koordinátatengelyekkel
(3 pont) és az együttes sűrűségfüggvény ezen a téglalapon egy csak x-től függő kifejezés (konkrétan: x) és egy csak y-tól függő kifejezés (sőt attól sem függő kifejezés, konkrétan: y) szorzatára bomlik (és ezért megegyezik $f_X(x)$ és $f_Y(y)$ szorzatával – ezt nem kell odaírni),
(1 pont) a téglalapon kívül pedig 0.
(1 pont) A függetlenség miatt
(1 pont) $\mathrm{cov}(X, Y) = 0$.
A függetlenség bármilyen más helyes indoklására is jár 6 pont. Például: kiszámolja a peremsűrűségfüggvényeket (Y-ét amúgy is ki kell a (c) miatt) és megállapítja, hogy az együttes sűrűségfüggvény ezek szorzatára bomlik. Ebből 1 pont jár arra az esetre, amikor x **vagy** y kívül esik X, illetve Y értékkészletén; ekkor az együttes sűrűségfüggvény értéke és a marginális sűrűségfüggvény-értékek szorzata is 0.

(c) (4 pont)
(1 pont) Akkor lesz egyenletes eloszlású Y a $(0; 1)$-en, ha a sűrűségfüggvénye itt konstans ($= 1$ – ezt nem baj, ha nem írja oda), mindenhol máshol pedig 0. Az utóbbi nyilvánvalóan teljesül. (Ezt nem kell indokolni, de ha nem is említi, hogy a $(0; 1)$-en kívül mi van, akkor ez az 1 pont csak abban az esetben jár, ha az is ott van, hogy a konstansnak 1-nek kell lennie, mert különben más olyan sűrűségfüggvény is létezik, ami teljesíti a feltételeket).
(1 pont) Ha $0 < y < 1$, akkor $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\mathrm{d}x = \int_0^{\sqrt{2}} x\mathrm{d}x$ (a két alakból itt már elég az egyik)
(1 pont) $= [x^2/2]_{x=0}^{\sqrt{2}} = 1$.
(1 pont) Tehát Y sűrűségfüggvénye tényleg konstans a $(0; 1)$-en, vagyis Y egyenletes eloszlású ezen az intervallumon.
A sűrűségfüggvény helyett az eloszlásfüggvényt is lehet ellenőrizni, de az bonyolultabb.

5. Szilveszterre veszünk 8 doboz kocsonyát. A kocsonyák dekagrammban mért tömegeiből álló vektort az $(X_1, \ldots, X_8)$ fae. minta írja le, ahol a független mintaelemek egységesen ismeretlen μ dekagramm várható értékkel és 2 dekagramm szórással normális eloszlásúak. A vásárolt kocsonyákat megmérve az alábbi $(x_1, \ldots, x_8)$ realizációt kaptuk:

$$x_1 = 49 \qquad x_2 = 52 \qquad x_3 = 48 \qquad x_4 = 46 \qquad x_5 = 50 \qquad x_6 = 49 \qquad x_7 = 51 \qquad x_8 = 47.$$

(a) Adjunk a minta alapján 96%(=0,96) szintű konfidenciaintervallumot a μ paraméterre.

(b) Döntsük el alkalmasan választott $\varepsilon = 0{,}02$ terjedelmű próbával, hogy a kapott realizáció alapján megegyezik-e 50 dekagrammal egy kocsonya tömegének várható értéke. Adjuk meg H_0-t, H_1-et, a próbastatisztika értékét, azt, hogy mi alapján döntünk, és magát a döntést is.

(c) Jelölje F_8^* a minta empirikus eloszlásfüggvényét az adott realizáció esetén. Mennyi $F_8^*(50{,}5)$ és $F_8^*(52)$?

Megoldás.
(a) (8 pont)
(0 pont) (A képletgyűjtemény alapján) a konfidenciaintervallum $\left[\overline{x_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{x_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\right]$.
(1 pont) Itt $n = 8$ (ez a mintaelemszám),
(1 pont) σ a szórás (nem kell szövegesen kiírni), vagyis $\sigma = 2$,
(1 pont) a mintaátlag $\overline{x_n} = \frac{1}{8}\sum_{i=1}^{8} x_i$
(1 pont) $= 49$.
(1 pont) A konfidenciaintervallum szintje $0{,}96 = 1 - \varepsilon$, tehát $\varepsilon = 0{,}04$, vagyis $1 - \varepsilon/2 = 0{,}98$ (elég, ha ez később, $u_{\varepsilon/2}$-nél derül ki),
(0 pont) így $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$
(0 pont) $= \Phi^{-1}(0{,}98)$
(1 pont) $\approx 2{,}05$.
(1 pont) Behelyettesítve a konfidenciaintervallum $[49 - \frac{2\cdot 2{,}05}{\sqrt{8}}, 49 + \frac{2\cdot 2{,}05}{\sqrt{8}}]$
(1 pont) $\approx [49 - 1{,}4496, 49 + 1{,}4496] = [47{,}5504, 50{,}4496]$ (elég, ha a végeredmény van meg numerikusan, sőt az előző lépéssel is összevonható az utolsó lépés, ha világos, hogy jól használja az általános képletet).
Elég a képletbe jól láthatóan beírni az adatokat, nem kell mindent külön-külön megnevezni.
(b) (8 pont)
(1 pont) Kétoldali,
(0 pont) egymintás u-próbát használunk.
(Ha jó az alternatíva és az eljárás, akkor nem baj, ha nincs odaírva, hogy kétoldali.)
(1 pont) Az alternatíva $H_0\colon \mu = 50$ vs. $H_1\colon \mu \neq 50$.
(1 pont) H_0-t akkor fogadjuk (nem utasítjuk) el, ha $|u| < u_{\varepsilon/2}$
(1 pont) $= u_{0{,}01} = \Phi^{-1}(0{,}99)$ (elég, ha az utóbbi szerepel)
(1 pont) $= 2{,}33$.
(0 pont) $u(= u(x_1, \ldots, x_8)) = \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n}$
(2 pont) $= \frac{49-50}{2}\sqrt{8} = -\sqrt{2} \approx -1{,}4142$ (itt már elég a végeredmény, ha már kiderült, hogy jó u-t használt; elég, ha impliciten derül ki, hogy $\mu_0 = 50$),
(1 pont) tehát H_0-t elfogadjuk.
Ha egyoldali próbát csinál $H_0\colon \mu \geq 50$ vs. $H_1\colon \mu < 50$ alternatívával, akkor max. 7 pont, ha $H_0\colon \mu \leq 50$ vs. $H_1\colon \mu > 50$ alternatívával, akkor max. 6 pont. Ha t-próbát próbál csinálni (pl. vizsgakurzusos hallgató), ahhoz sok minden hiányzik a táblázatból és nem fog rendesen menni, így kétoldali esetben max. 4 pont, egyoldali esetben az előzőhöz hasonlóan max. 3, illetve 2 pont.
(c) (4 pont)
(2 pont) $F_8^*(x) = \frac{1}{n}\{i \in \{1, \ldots, 8\} |\ x_i < x\}$ (ugyanez jó szövegesen is).

(1 pont) Így $F_8^*(50{,}5) = \frac{6}{8}(= \frac{3}{4} = 0{,}75)$
(1 pont) és $F_8^*(52) = \frac{7}{8}(= 0{,}875)$.
Ha „$<$” helyett „$\leq$”-t ír, akkor az első 2 pontból max. 1 jár és $F_8^*(52)$-re sem jár az 1 pont. Ha általános indoklást ad (és az helyes), akkor utána elég a konkrét esetekben a végeredményt közölni.

6. * Legyenek U és V független, standard normális eloszlású valószínűségi változók. Az X és Y valószínűségi változókról tudjuk, hogy $X = aU+bV+c$ és $Y = dU+eV+f$ alkalmasan választott a, b, c, d, e, f valós számokra, valamint Y regressziós egyenese X-re $\{(x, y) \in \mathbb{R}^2 \colon y = \frac{1}{2}x + 3\}$ és $\mathbb{E}(X|Y) = \frac{3}{2}Y - 1$. Tudjuk azt is, hogy X és Y szórása is pozitív.

   (a) Mi X regressziós egyenese Y-ra (és miért)?

   (b) Határozzuk meg X és Y várható értékét.

   (c) Határozzuk meg c és f értékét.

   (d) Adjuk meg az $\mathbb{E}(\sin(X)Y|X)$ regressziót.

   (*Segítség*: Az (a), (b) és (d) rész megoldásához a, b, c, d, e, f értékét nem kell ismerni.)

   Megoldás.
   (a) (6 pont)
   (1 pont) Mivel X és Y is ugyanazon független normálisok (konkrétan U és V) lineáris/affin kombinációi (meg lehet fogalmazni konkrétabban is),
   (2 pont) a regresszió megegyezik a lineáris regresszióval.
   (1 pont) Ez azt jelenti, hogy $\mathbb{E}(X|Y) = \beta X + \alpha$, ahol β és α az X-nek Y-ra vett lineáris regressziójának megfelelő kifejezések (nem kell kiírni őket). (Az előzővel összevonható.)
   (1 pont) Tehát $\beta = \frac{3}{2}$ és $\alpha = -1$.
   (1 pont) A regressziós egyenes így $\{(y, x) \in \mathbb{R}^2 \colon x = \frac{3}{2}y - 1\}$. (Ha X és Y szerepét felcseréli, akkor ez az egy pont nem jár.) (Ez a lépés az előzővel összevonható, és ebben az esetben ha X és Y szerepét felcseréli, akkor a 2 pontból max. 1 jár.)
   (b) (5 pont)
   (1 pont) A lineáris regresszió képlete alapján $(\alpha =)3 = \mathbb{E}(Y) - \frac{1}{2}\,\mathbb{E}(X)(= \mathbb{E}(Y) - \beta\,\mathbb{E}(X))$. (Az általános képleteket nem muszáj odaírni.)
   (1 pont) Hasonlóan (az (a) alapján) $-1 = \mathbb{E}(X) - \frac{3}{2}\,\mathbb{E}(Y)$.
   (3 pont) A kapott lineáris egyenletrendszert megoldva (sok részletszámítás nem kell, csak látszódjon, hogy tényleg ő oldotta meg) $\mathbb{E}(X) = 14$ és $\mathbb{E}(Y) = 10$.
   (c) (5 pont)
   (1 pont) $\mathbb{E}(X) = a\,\mathbb{E}(U) + b\,\mathbb{E}(V) + c$ (a várható érték linearitása miatt – nem kell odaírni)
   (1 pont) $= c$,
   (1 pont) mivel U és V standard normális eloszlásúak, tehát a várható értékük 0.
   (1 pont) Hasonlóan $\mathbb{E}(Y) = f$.
   (1 pont) Tehát $c = 14$ és $f = 10$.
   (d) (4 pont)
   (1 pont) $\mathbb{E}(\sin(X)Y|X) = \sin(X)\,\mathbb{E}(Y|X)$,
   (1 pont) mert... (valahogyan hivatkozik a regresszió megfelelő tulajdonságára, pl. ismert információ kiemelése, taking out what is known stb.)
   (1 pont) és $\mathbb{E}(Y|X)$ (megint csak a regresszió linearitása miatt – nem kell újra odaírni) $\beta X + \alpha$-val egyezik meg, ahol $\beta = \frac{1}{2}$ és $\alpha = 3$ az Y-nak X-re vett lineáris regressziójának megfelelő kifejezések.
   (1 pont) Így $\mathbb{E}(\sin(X)Y|X) = \frac{1}{2}X\sin(X) + 3\sin(X)$ (nem baj, ha nem bontja fel a zárójelet).