1. 0-tól 9-ig felírjuk az egészeket egy-egy cetlire, bedobjuk egy kalapba, majd véletlenszerűen húzunk egyet. Megfigyeljük az alábbi eseményeket A=1,2,3; B=1,4,8; C=3,4,5,6; D=2,6,7,8 (ahol a számsorok azt jelentik, hogy a megfelelő esemény pontosan akkor következik be, ha a számok közül húztuk az egyiket).

   Tekintsük a $C \setminus (B + D)$ és az $A + B$ eseményeket. Erről az eseménypárról döntsük el, hogy kizáró-e illetve független-e!

   Adjunk meg az A, B, C és D események segítségével egy kizáró eseménypárt úgy, hogy csak az összeg és szorzat műveleteket használhatjuk (egyik esemény se legyen lehetetlen esemény, sem biztos esemény).

   (1pont) $C \setminus (B + D) = 3, 5$
   (1pont) $A + B = 1, 2, 3, 4, 8$
   (2pont) Kizáró eseménypár, aminek az együttes bekövetkezése lehetetlen esemény.
   (1pont) Nem kizáróak, mert 3-as.
   (3pont) Független két esemény, ha a szorzatuk valószínűsége a valószínűségek szorzata.
   (2pont) Klasszikus valószínűségi mezőn gondolkozunk, így minden egyes szám valószínűsége $\frac{1}{10}$.
   (2pont) $P(C \setminus (B + D)) \cdot P(A + B) = \frac{1}{10}$
   (2pont) $P((C \setminus (B + D)) \cdot (A + B)) = \frac{1}{10}$
   (1pont) tehát függetlenek.
   (4pont) Egy jó példa kizáróra (pl. A és CD)
   (1pont) Ezek jók, mert kizáróak/ a szorzatuk a lehetetlen.

2. Választunk egy (X, Y) vektort az alábbi sűrűségfüggvény szerint: $f_{X,Y}(x, y) = \frac{1}{4}x \cdot y$ ahol $0 < x, y < 2$. Mi annak a valószínűsége, hogy szerkeszthető olyan háromszög, aminek oldalainak hosszai éppen X,Y és 1?

   (2pont) Akkor szerkeszthető (S), ha teljesülnek a háromszög egyenlőtlenségek: $X + 1 > Y$, $Y + 1 > X$, $X + Y > 1$
   (1pont) Egy esemény valószínűsége megegyezik a sűrűségfüggvény megfelelő tartományon vett integráljával.
   (1pont) Az egyenlőtlenségek Y-ra rendezve: $X + 1 > Y$, $Y > X - 1$, $Y > 1 - X$
   (1pont) Azaz $P(S) = \int_0^2 \int_{max\{1-x,x-1\}}^{min\{2,x+1\}} \frac{1}{4}xydydx =$
   (2pont) $= \int_0^2 \left[\frac{1}{8}xy^2\right]_{max\{1-x,x-1\}}^{min\{2,x+1\}} dx$
   (2pont) ha $x < 1$ akkor $1 - x > 0$, $x - 1 < 0$, azaz $max\{1 - x, x - 1\} = 1 - x$
   (2pont) ha $x > 1$ akkor $1 - x < 0$, $x - 1 > 0$, azaz $max\{1 - x, x - 1\} = x - 1$
   (1pont) ha $x < 1$ akkor $min\{2, x + 1\} = x + 1$
   (1pont) ha $x > 1$ akkor $min\{2, x + 1\} = 2$
   (3pont) Így $P(S) = \int_0^1 \left[\frac{1}{8}xy^2\right]_{1-x}^{x+1} dx + \int_1^2 \left[\frac{1}{8}xy^2\right]_{x-1}^{2} dx =$
   (1pont) $= \int_0^1 \frac{1}{8}x(x^2 + 2x + 1 - 1 + 2x - x^2)dx + \int_1^2 \frac{1}{8}x(4 - x^2 + 2x - 1)dx =$
   (1pont) $= \int_0^1 \frac{1}{2}x^2dx + \int_1^2 \frac{3x-x^3+2x^2}{8}dx =$
   (1pont) $= \frac{1}{6} + \left[\frac{\frac{3}{2}x^2-\frac{1}{4}x^4+\frac{2}{3}x^3}{8}\right]_1^2 =$
   (1pont) $= \frac{1}{6} + \frac{\frac{3}{2}4-\frac{1}{4}16+\frac{2}{3}8}{8} - \frac{\frac{3}{2}-\frac{1}{4}+\frac{2}{3}}{8} = \frac{1}{6} + \frac{3}{4} - \frac{1}{2} + \frac{2}{3} - \frac{3}{16} + \frac{1}{32} - \frac{1}{12} = 0{,}8438$

3. A baseball-ütők X hosszának várható értéke 100, szórása 10 (centiméterben mérve). A baseball-labdák Y tömegének várható értéke 200, szórása 20 (grammban mérve). Egy pontos Durrbele200 ütés esetén a labda röppályájának az A legnagyobb magassága $p + 2q$ méter, a B hossza pedig $2p + q$, ha p volt az ütő hossza centiméterben és q a labda tömege grammban. Mi lesz A és B kovarianciája és korrelációs együtthatója (ha véletlenszerűen választott ütővel ütnek, az ütőtől függetlenül, véletlenszerűen választott labdát)? Független-e egymástól A és B?

   (2pont) $A = X + 2Y$ és $B = 2X + Y$
   (2pont) $cov(A, B) = E(AB) - EAEB =$
   (1pont) $= E((X+2Y)(2X+Y)) - E(X+2Y)E(2X+Y) = E(2X^2+2Y^2+5XY) - EXE(2X) - E(2Y)EY - 5EXEY =$
   (2pont) függetlenek szorzata szétbontható: $= 2E(X^2) + 2E(Y^2) + 5EXEY - 2(EX)^2 - 2(EY)^2 - 5EXEY =$
   (2pont) Steiner-formula szerint: $= 2\sigma^2(X) + 2\sigma^2(Y) =$
   (1pont) $= 200 + 800 = 1000$
   (4pont) függetlenség miatt szórásnégyzetek összegződnek: $\sigma A = \sqrt{100 + 1600} \approx 41{,}2311$, $\sigma B = \sqrt{400 + 400} \approx 28{,}2843$
   (3pont) $R(A, B) = \frac{cov(A,B)}{\sigma A \sigma B} = \frac{1000}{1166{,}1904} = 0{,}8574$
   (2pont) Ha a korrelációjuk nem 0, akkor nem lehetnek függetlenek,
   (1pont) ezért nem függetlenek.

   2.megoldás:
   (2pont) $A = X + 2Y$ és $B = 2X + Y$
   (1pont) $cov(A, B) = cov(X + 2Y, 2X + Y) =$
   (2pont) a kovariancia bilineáris:
   (1pont) $cov(X, 2X) + cov(X, Y) + cov(2Y, 2X) + cov(2Y, Y)$.
   (1pont) $cov(X, Y) = cov(2Y, 2X) = 0$, mert függetlenek
   (1pont) $cov(X, 2X) = 2 \cdot cov(X, X) =$
   (1pont) $2\sigma^2 X = 200$ és $cov(2Y, Y) = 2\sigma^2 Y = 800$
   (1pont) Tehát $cov(A, B) = 200 + 800 = 1000$.
   (4pont) függetlenség miatt szórásnégyzetek összegződnek: $\sigma A = \sqrt{100 + 1600} \approx 41{,}2311$, $\sigma B = \sqrt{400 + 400} \approx 28{,}2843$
   (3pont) $R(A, B) = \frac{cov(A,B)}{\sigma A \sigma B} = \frac{1000}{1166{,}1904} = 0{,}8574$
   (2pont) Ha a korrelációjuk nem 0, akkor nem lehetnek függetlenek,
   (1pont) ezért nem függetlenek.

4. Egy pincészetben munkanapokon átlagosan 100 liter bort mérnek ki, 20 szórással (a napok teljesen függetlenek és azonos eloszlásúak). Az évből hátralevő 50 munkanap alatt 4750 liter bort kellene eladniuk ahhoz, hogy felülmúlják a tavalyi teljesítményt. Mi annak a valószínűsége, hogy ez sikerülni fog?

(2pont) Legyen az egy napi fogyasztás eloszlása X. Ekkor ezzel azonos eloszlású $X_1, ..., X_{50}$ változók összegeként mérhető a hátralevő fogyasztás.
(1pont) $EX = 100$ és $\sigma X = 20$
(2pont) $P(\sum_1^{50} X_i > 4750) =?$
(2pont) CHT értelmében az összeg normálissal közelíthető:
(2pont) $P(\sum_1^n X_i < t) \approx \Phi(\frac{t-nEX}{\sqrt{n}\sigma X})$
(4pont) $P(\sum_1^{50} X_i < 4750) \approx \Phi(\frac{4750-5000}{100\sqrt{2}}) =$
(2pont) $= \Phi(-\frac{5}{2\sqrt{2}}) \approx \Phi(-1{,}77) =$
(3pont) $= 1 - \Phi(1{,}77) = 1 - 0{,}9616 = 0{,}0384$
(2pont) Azaz $P(\sum_1^{50} X_i > 4750) = 1 - 0{,}0384 = 0{,}9616$

5. 2019-szer dobunk egy szabályos kockával, és figyeljük az 1-esek X számát, és a 6-osok Y számát. Adjuk meg az X és Y vetületi (perem-) eloszlásait, X feltételes eloszlását az $Y = \ell$ feltétel mellett, valamint az $E(X|Y)$ regressziót!

(3pont) X és Y eloszlása: $Bin(2019, \frac{1}{6})$
(1pont) $(X|Y = \ell)$ eloszlása: $Bin(2019 - \ell, \frac{1}{5})$, mert
(8pont) indoklás. Pl.:
mivel $R_{(X|Y=\ell)} = \{0, 1, \ldots, 2019 - \ell\}$, (1pont)
és $P(X = k|Y = \ell) = \frac{P(X=k,Y=\ell)}{P(Y=\ell)} =$ (2pont)
$= \frac{\binom{2019}{k}\left(\frac{1}{6}\right)^k\binom{2019-k}{\ell}\left(\frac{1}{6}\right)^\ell\left(\frac{4}{6}\right)^{2019-\ell-k}}{\binom{2019}{\ell}\left(\frac{1}{6}\right)^\ell\left(\frac{5}{6}\right)^{2019-\ell}} =$ (3pont)
$= \binom{2019-\ell}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{2019-\ell-k}, \ k = 0, 1, \ldots, 2019 - \ell$ (2pont)
(2pont) A regressziós függvény $E(X|Y = \ell) = \sum_{x \in R_{(X|Y=\ell)}} xP(X = x|Y = \ell) =$
(2pont) $= \sum_{k=0}^{2019-\ell} k\binom{2019-\ell}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{2019-k-\ell}$,
(3pont) ami a $Bin(2019 - \ell, \frac{1}{5})$ eloszlású $(X|Y = \ell)$ valószínűségi változó várható értéke, azaz $\frac{2019-\ell}{5}$
(1pont) Így a regresszió $E(X|Y) = \frac{2019-Y}{5}$

6. Ha András eltör egy botot, akkor a töréspont a bot középső harmadán egyenletesen helyezkedik el. András most a kezébe vesz egy 15 centi hosszú botot, és azt eltöri, majd a bal kezében levő darabot ismét eltöri. A második törés után a bal kezében maradó darab hossza 5 centi hosszú. Adjuk meg ennek ismeretében az első törés utáni baloldali darab hosszának feltételes sűrűségfüggvényét!

(1pont) Az első töréspont bal oldalról $X \in U(5, 10)$
(2pont) A második töréspont $Y \in U(\frac{X}{3}, \frac{2X}{3})$
(2pont) $f_{X|Y}(x|5) =?$
(3pont) A folytonos Bayes tétel segíthet: $f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|t)f_X(t)dt}$
(1+1pont) $f_X(t) = \frac{1}{5}$, ahol $5 < t < 10$
(1+1pont) $f_{Y|X}(y|t) = \frac{3}{t}$, ahol $5 < t < 10$ és $\frac{t}{3} < y < \frac{2t}{3}$
(1+2pont) $f_{Y|X}(5|t) = \frac{3}{t}$, ahol $5 < t < 10$, $t < 15$, $\frac{15}{2} < t$, azaz $\frac{15}{2} < t < 10$
(2pont) $f_{X|Y}(x|5) = \frac{\frac{3}{x}\frac{1}{5}}{\int_{\frac{15}{2}}^{10} \frac{3}{t}\frac{1}{5}dt}$, ahol $\frac{15}{2} < x < 10$
(2+1pont) $f_{X|Y}(x|5) = \frac{\frac{3}{5x}}{\frac{3}{5}(ln(10)-ln(\frac{15}{2}))} = \frac{1}{xln(\frac{4}{3})} \left(\approx 3{,}4761\frac{1}{x}\right)$, ahol $\frac{15}{2} < x < 10$

2. megoldás:
(1pont) Az első töréspont bal oldalról $X \in U(5, 10)$
(2pont) A második töréspont $Y \in U(\frac{X}{3}, \frac{2X}{3})$
(2pont) $f_{X|Y}(x|5) =?$
(1+1pont) $f_X(x) = \frac{1}{5}$, ahol $5 < x < 10$,
(1+1pont) $f_{Y|X}(y|x) = \frac{3}{x}$, ahol $5 < x < 10$ és $\frac{x}{3} < y < \frac{2x}{3}$,
(2pont) így az együttes sűrűségfüggvény: $f_{X,Y}(x, y) = f_X(x) \cdot f_{Y|X}(y|x) =$
(1+1pont) $= \frac{3}{5x}$, ahol $5 < x < 10$ és $\frac{x}{3} < y < \frac{2x}{3}$.
(2pont) Ebből $f_Y(5) = \int_{-\infty}^{\infty} f_{X,Y}(x, 5)dx =$
(2pont) $= \int_{\frac{15}{2}}^{10} \frac{3}{5x}dx = \frac{3}{5}(ln(10) - ln(\frac{15}{2})) = \frac{3}{5}ln(\frac{4}{3})$,
(2+1pont) és $f_{X|Y}(x|5) = \frac{f_{X,Y}(x,5)}{f_Y(5)} = \frac{\frac{3}{5x}}{\frac{3}{5}ln(\frac{4}{3})} = \frac{1}{xln(\frac{4}{3})}$, ahol $\frac{15}{2} < x < 10$.