

8. Gyakorlat

Konvolúció és eloszlástranszformációk

Konvolúció

Állítás Legyen X és Y két független valószínűségi változó. Ekkor

- Amennyiben X és Y is diszkrét (egész értékkészlettel), akkor $X + Y$ súlyfüggvénye

$$\mathbb{P}(X + Y = m) = \sum_{k=-\infty}^{\infty} \mathbb{P}(X = k) \mathbb{P}(Y = m - k).$$

- Amennyiben X és Y is folytonos, akkor $X + Y$ sűrűségfüggvénye

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s) f_Y(t - s) ds.$$

Állítás Igazak az alábbiak

- Amennyiben $X \sim \text{Bin}(n, p)$ és $Y \sim \text{Bin}(m, p)$ függetlenek, akkor $X + Y \sim \text{Bin}(n + m, p)$.
- Amennyiben $X \sim \text{Poi}(\lambda)$ és $Y \sim \text{Poi}(\mu)$ függetlenek, akkor $X + Y \sim \text{Poi}(\lambda + \mu)$.
- Amennyiben $X_1, X_2, \dots, X_n \sim \text{Exp}(\lambda)$ függetlenek, akkor $X_1 + X_2 + \dots + X_n \sim \text{Gamma}(n, \lambda) = \text{Erlang}(n, \lambda)$.
- Amennyiben $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ és $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ függetlenek, akkor $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Eloszlástranszformációk

Állítás Legyen X valószínűségi változó és $g : \mathbb{R} \rightarrow \mathbb{R}$ függvény. Továbbá legyen $Y = g(X)$. Ekkor Y eloszlásfüggvénye és sűrűségfüggvénye (ha g differenciálható és X folytonos), amennyiben

- g szigorúan monoton nő

$$F_Y(y) = F_X(g^{-1}(y)) \quad \text{és} \quad f_Y(y) = \frac{1}{g'(g^{-1}(y))} \cdot f_X(g^{-1}(y)).$$

- g szigorúan monoton csökken

$$F_Y(y) = 1 - F_X(g^{-1}(y)) \quad \text{és} \quad f_Y(y) = -\frac{1}{g'(g^{-1}(y))} \cdot f_X(g^{-1}(y)).$$

Megjegyzés A fenti transzformált várható értéke (ez csak emlékeztető):

- Amennyiben X diszkrét

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \text{Ran}(X)} g(x) \cdot \mathbb{P}(X = x)$$

- Amennyiben X folytonos

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx.$$

Feladatok

1. Legyenek X, Y, Z független $\text{Uni}(0, 1)$ eloszlású valószínűségi változók. Határozzuk meg az alábbi valószínűségi változók sűrűségfüggvényét.

- (a) $X+Y$
(b) $X+Y+Z$

Megoldás

- (a) Mivel $\text{Ran}(X + Y) = [0, 2]$, ezért, amennyiben $t \notin [0, 2]$, akkor

$$f_{X+Y}(t) = 0.$$

Legyen $t \in [0, 2]$. Ekkor a konvolúciós képlet szerint

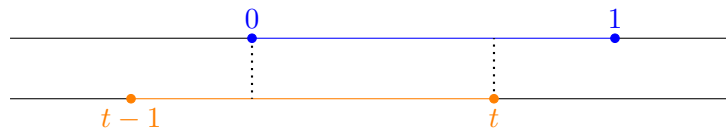
$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s)f_Y(t-s) ds,$$

ahol a sűrűségfüggvények

$$f_X(s) = \begin{cases} 1 & \text{ha } 0 \leq s \leq 1 \\ 0 & \text{egyébként} \end{cases} \quad \text{és} \quad f_Y(t-s) = \begin{cases} 1 & \text{ha } t-1 \leq s \leq t \\ 0 & \text{egyébként} \end{cases}$$

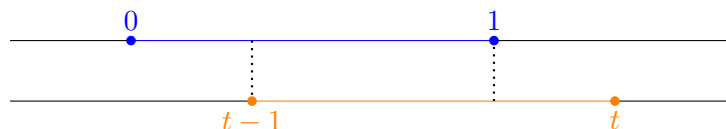
Vizsgáljuk meg, hogy hol nem lesz nulla a szorzat. Két esetet kell megkülönböztetni.

- Amennyiben $0 \leq t < 1$, akkor



$$f_{X+Y}(t) = \int_0^t 1 ds = t.$$

- Amennyiben $1 \leq t \leq 2$, akkor



$$f_{X+Y}(t) = \int_{t-1}^1 1 ds = 2 - t.$$

Azt kaptuk, hogy a sűrűség

$$f_{X+Y}(t) = \begin{cases} t & \text{ha } 0 \leq t < 1 \\ 2 - t & \text{ha } 1 \leq t \leq 2 \\ 0 & \text{egyébként} \end{cases}$$

- (b) Vegyük észre, hogy $\text{Ran}(X + Y + Z) = [0, 3]$. Ezért amennyiben $t \notin [0, 3]$, akkor

$$f_{X+Y+Z}(t) = 0.$$

Legyen $t \in [0, 3]$ ekkor a konvolúciós képlet szerint

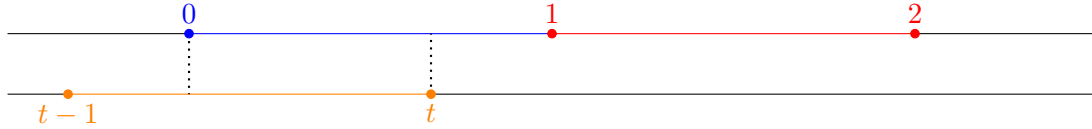
$$f_{X+Y+Z}(t) = \int_{-\infty}^{\infty} f_{X+Y}(s)f_Z(t-s) ds,$$

ahol

$$f_{X+Y}(s) = \begin{cases} s & \text{ha } 0 \leq s < 1 \\ 2-s & \text{ha } 1 \leq s \leq 2 \\ 0 & \text{egyébként} \end{cases} \quad \text{és} \quad f_Z(t-s) = \begin{cases} 1 & \text{ha } t-1 \leq s \leq t \\ 0 & \text{egyébként} \end{cases}$$

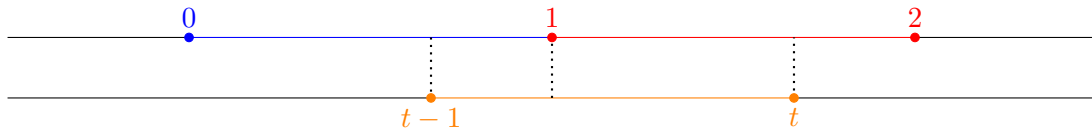
Ekkor az alábbi három esetet kell megkülönböztetni.

- Amennyiben $0 \leq t < 1$, akkor



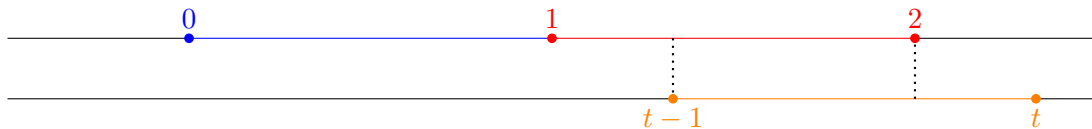
$$f_{X+Y+Z}(t) = \int_0^t s \, ds = \left[\frac{1}{2} s^2 \right]_{s=0}^{s=t} = \frac{1}{2} t^2.$$

- Amennyiben $1 \leq t < 2$, akkor



$$f_{X+Y+Z}(t) = \int_{t-1}^1 s \, ds + \int_1^t (2-s) \, ds = \left[\frac{1}{2} s^2 \right]_{s=t-1}^{s=1} + \left[2s - \frac{1}{2} s^2 \right]_{s=1}^{s=t} = \frac{3}{4} - \left(t - \frac{3}{2} \right)^2.$$

- Amennyiben $2 \leq t \leq 3$, akkor

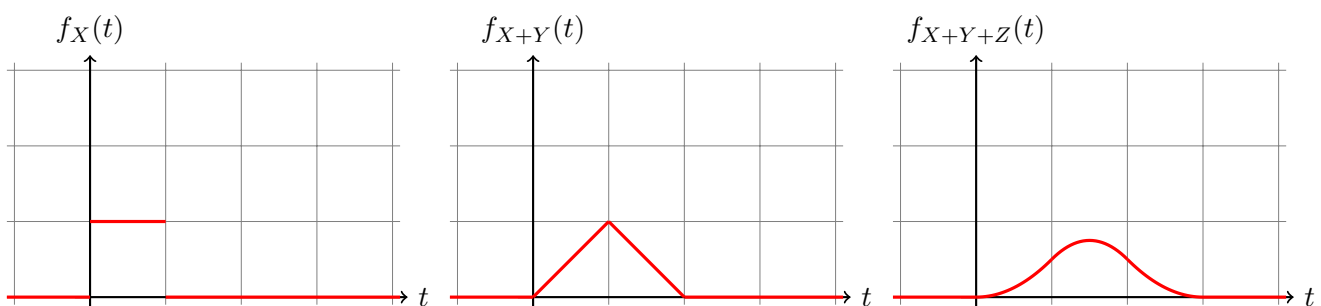


$$f_{X+Y+Z}(t) = \int_{t-1}^2 (2-s) \, ds = \left[2s - \frac{1}{2} s^2 \right]_{s=t-1}^{s=2} = \frac{1}{2} (t-3)^2.$$

Azt kaptuk, hogy

$$f_{X+Y+Z}(t) = \begin{cases} \frac{1}{2} t^2 & \text{ha } 0 \leq t < 1 \\ \frac{3}{4} - \left(t - \frac{3}{2} \right)^2 & \text{ha } 1 \leq t < 2 \\ \frac{1}{2} (t-3)^2 & \text{ha } 2 \leq t \leq 3 \\ 0 & \text{egyébként} \end{cases}$$

Megjegyzés: Ábrázoljuk X , $X+Y$ és $X+Y+Z$ sűrűségfüggvényeit.



Vegyük észre, hogy X sűrűségfüggvénye nem folytonos, $X + Y$ sűrűségfüggvénye ugyan folytonos, de nem differenciálható. Azonban $X + Y + Z$ sűrűségfüggvénye már differenciálható. Továbbá mivel a sűrűségfüggvény az eloszlásfüggvény deriválva, ezért X eloszlásfüggvénye 0-szor, $X + Y$ -nek 1-szer, $X + Y + Z$ -nek 2-szer folytonosan differenciálható. Általánosan is igaz, hogy amennyiben egy valószínűségi változóhoz hozzáadunk egy tőle független $\text{Uni}(0, 1)$ eloszlású valószínűségi változót, akkor az legalább eggyel növeli az eloszlás simaságát.

2. Legyen $X \sim \text{Exp}(\lambda)$ és $Y \sim \text{Exp}(\mu)$ függetlenek. Határozzuk meg $X + Y$ sűrűségfüggvényét, ha tudjuk, hogy $\lambda \neq \mu$.

Megoldás Vegyük észre, hogy $\text{Ran}(X + Y) = [0, \infty)$. Ekkor minden $t < 0$ esetén

$$f_{X+Y}(t) = 0.$$

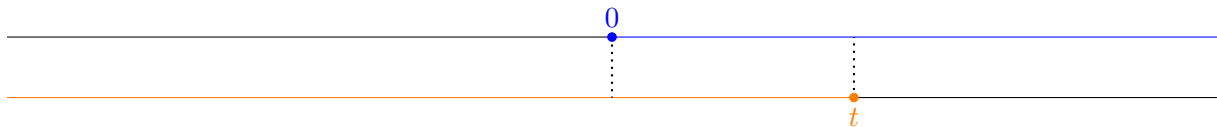
Legyen $0 \leq t$, ekkor

$$f_{X+Y} = \int_0^t \lambda e^{-\lambda s} \cdot \mu e^{-\mu(t-s)} ds = \int_{-\infty}^{\infty} f_X(s) f_Y(t-s) ds,$$

ahol

$$f_X(t) = \begin{cases} \lambda e^{-\lambda s} & \text{ha } 0 \leq s \\ 0 & \text{egyébként} \end{cases} \quad \text{és} \quad f_Y(t) = \begin{cases} \mu e^{-\mu(t-s)} & \text{ha } s \leq t \\ 0 & \text{egyébként} \end{cases}$$

Ekkor megkeresve azon tartományt, ahol mindkét függvény nemnulla, továbbá kihasználva, hogy $\lambda \neq \mu$



$$f_{X+Y} = \int_0^t \lambda e^{-\lambda s} \cdot \mu e^{-\mu(t-s)} ds = \lambda \mu e^{-\mu t} \left[\frac{-e^{(\lambda-\mu)s}}{\lambda - \mu} \right]_{s=0}^{s=t} = \frac{\lambda \mu}{\lambda - \mu} (e^{-\mu t} - e^{-\lambda t}).$$

Azt kaptuk, hogy

$$f_{X+Y}(t) = \begin{cases} \frac{\lambda \mu}{\lambda - \mu} (e^{-\mu t} - e^{-\lambda t}) & \text{ha } 0 \leq t \\ 0 & \text{egyébként} \end{cases}$$

3. Legyen X és Y független $\text{Geom}\left(\frac{1}{4}\right)$ eloszlású valószínűségi változók. Határozzuk meg $X + Y$ súlyfüggvényét. Milyen eloszlást követ $X + Y$?

Megoldás Vegyük észre, hogy $\text{Ran}(X + Y) = \{2, 3, 4, \dots\}$. Legyen $k \in \text{Ran}(X + Y)$, ekkor a konvolúciós formula alapján

$$\mathbb{P}(X + Y = k) = \sum_{m=-\infty}^{\infty} \mathbb{P}(X = m) \mathbb{P}(Y = k - m).$$

Tudjuk, hogy $\text{Ran}(X) = \text{Ran}(Y) = \{1, 2, 3, \dots\}$, ezért amennyiben $m < 1$, akkor

$$\mathbb{P}(X = m) = 0.$$

Továbbá amennyiben $k - m < 1$ (vagy másképpen $k - 1 < m$), akkor

$$\mathbb{P}(Y = k - m) = 0.$$

Ezért a konvolúciós fomula

$$\begin{aligned} \mathbb{P}(X + Y = k) &= \sum_{m=1}^{k-1} \mathbb{P}(X = m) \mathbb{P}(Y = k - m) = \sum_{m=1}^{k-1} \frac{1}{4} \left(\frac{3}{4}\right)^{m-1} \cdot \frac{1}{4} \left(\frac{3}{4}\right)^{k-m-1} \\ &= \frac{1}{4^2} \left(\frac{3}{4}\right)^{k-2} \cdot \sum_{m=1}^{k-1} 1 = \frac{1}{4^2} \left(\frac{3}{4}\right)^{k-2} (k-1) = \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^{k-2} \binom{k-1}{2-1}. \end{aligned}$$

Az utóbbi formából az is látható, hogy

$$X + Y \sim \text{Negbinom}\left(2, \frac{1}{4}\right).$$

Megjegyzés: Általánosan is belátható, hogy amennyiben $X_1, X_2, \dots, X_n$ függetlenek $\text{Geom}(p)$ eloszlással, akkor

$$X_1 + X_2 + \dots + X_n \sim \text{Negbinom}(n, p).$$

Ez igazából a definícióból is látszik, hiszen a negatív binomiális az n -edik siker sorszáma. Ez valóban nem más, mint n darab független geometriai összege.

4. Egy hallgató egy tárgyat próbál teljesíteni az egyetemen. Minden félévben egymástól függetlenül $\frac{1}{3}$ valószínűséggel szerez aláírást. Amennyiben megszerezte az aláírást megpróbálkozik a vizsgával, amit $\frac{1}{4}$ valószínűséggel teljesít sikeresen minden félévben, egymástól függetlenül (ha már megszerezte az aláírást, azt nem kell újra megszereznie a következő félévben). Adjuk meg minden k -ra, hogy mekkora valószínűséggel van k félévre szüksége a hallgatónak a tárgy elvégzéséhez.

Megoldás Legyen X a szükséges szemeszterek száma az aláírás megszerzéséig. Ekkor minden $k \in \text{Ran}(X) = \{1, 2, 3, \dots\}$ esetén a feladat szerint

$$\mathbb{P}(X = k) = \frac{1}{3} \left(\frac{2}{3}\right)^{k-1}.$$

Továbbá legyen Y az aláírás megszerzése után szükséges szemeszterek száma az első sikeres vizsgáig. Ekkor minden $k \in \text{Ran}(Y) = \{0, 1, 2, \dots\}$ (itt nullától kezdődik, mert levizsgálhat sikeresen ugyanabban a szemeszterben, amiben megszerezte az aláírást) esetén a feladat szerint

$$\mathbb{P}(Y = k) = \frac{1}{4} \left(\frac{3}{4}\right)^k.$$

Ekkor vegyük észre, hogy X és Y független, továbbá a szükséges szemeszterek száma a tárgy elvégzéséhez pontosan $X + Y$. Másfelől ekkor nyilván $\text{Ran}(X + Y) = \{0, 1, 2, \dots\}$. Illetve a konvolúciós formula miatt minden $k \in \text{Ran}(X + Y)$ esetén

$$\mathbb{P}(X + Y = k) = \sum_{m=-\infty}^{\infty} \mathbb{P}(X = m) \mathbb{P}(Y = k - m).$$

Továbbá amennyiben $m < 1$, akkor

$$\mathbb{P}(X = m) = 0,$$

illetve amennyiben $k - m < 0$ (vagy másképpen $k < m$), akkor

$$\mathbb{P}(Y = k - m) = 0.$$

Ekkor a formula szerint minden $k \in \text{Ran}(X + Y)$ esetén

$$\begin{aligned} \mathbb{P}(X + Y = k) &= \sum_{m=1}^k \mathbb{P}(X = m) \mathbb{P}(Y = k - m) = \sum_{m=1}^k \frac{1}{3} \left(\frac{2}{3}\right)^{m-1} \cdot \frac{1}{4} \left(\frac{3}{4}\right)^{k-m} \\ &= \frac{1}{12} \left(\frac{3}{4}\right)^k \frac{3}{2} \sum_{m=1}^k \left(\frac{2}{3}\right)^m \left(\frac{3}{4}\right)^{-m} = \frac{1}{8} \left(\frac{3}{4}\right)^k \sum_{m=1}^k \left(\frac{8}{9}\right)^m \\ &= \frac{1}{8} \left(\frac{3}{4}\right)^k \cdot \frac{8}{9} \cdot \frac{1 - \left(\frac{8}{9}\right)^{k-1}}{1 - \frac{8}{9}} = \left(\frac{3}{4}\right)^k \left(1 - \left(\frac{8}{9}\right)^{k-1}\right) \end{aligned}$$

5. Kacsákat pakolunk talicskába. A tojók tömege egyenletes eloszlású 2 és 3 kg között, a gácséroké pedig 2 és 4 kg között. Mi lesz egy tojó és egy gácsér együttes tömegének sűrűségfüggvénye? Hány tojó fér bele a talicskába 90%-os valószínűséggel, ha a talicska teherbírása 1 mázsa?

Megoldás Legyen X a tojó tömege és Y a gácsér tömege (kg-ban). Ekkor a feladat szerint X és Y független és

$$X \sim \text{Uni}(2, 3) \quad \text{és} \quad Y \sim \text{Uni}(2, 4).$$

Ekkor a sűrűségfüggvényeik

$$f_X(x) = \begin{cases} 1 & \text{ha } 2 \leq x \leq 3 \\ 0 & \text{egyébként} \end{cases} \quad \text{és} \quad f_Y(y) = \begin{cases} \frac{1}{2} & \text{ha } 2 \leq y \leq 4 \\ 0 & \text{egyébként} \end{cases}.$$

Ekkor $X + Y$ egy tojó és egy gácsér együttes tömege. Vegyük észre, hogy mivel $\text{Ran}(X) = [2, 3]$ és $\text{Ran}(Y) = [2, 4]$, ezért $\text{Ran}(X + Y) = [4, 7]$. Innen amennyiben $t \notin [4, 7]$, akkor

$$f_{X+Y}(t) = 0.$$

Legyen $t \in [4, 7]$. Ekkor a konvolúciós formula szerint

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s)f_Y(t-s) ds,$$

ahol a korábbiak miatt

$$f_X(s) = \begin{cases} 1 & \text{ha } 2 \leq s \leq 3 \\ 0 & \text{egyébként} \end{cases} \quad \text{és} \quad f_Y(t-s) = \begin{cases} \frac{1}{2} & \text{ha } t-4 \leq s \leq t-2 \\ 0 & \text{egyébként} \end{cases}.$$

Az alábbi eseteket kell megkülönböztetni.

- Amennyiben $4 \leq t < 5$, akkor

$$f_{X+Y}(t) = \int_2^{t-2} 1 \cdot \frac{1}{2} ds = \frac{t-2-2}{2} = \frac{1}{2}t - 2.$$

- Amennyiben $5 \leq t < 6$, akkor

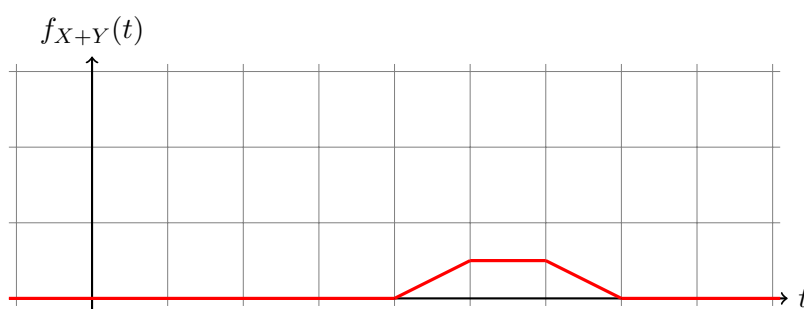
$$f_{X+Y}(t) = \int_2^3 1 \cdot \frac{1}{2} ds = \frac{3-2}{2} = \frac{1}{2}.$$

- Amennyiben $6 \leq t \leq 7$, akkor

$$f_{X+Y}(t) = \int_{t-4}^3 1 \cdot \frac{1}{2} ds = \frac{3-(t-4)}{2} = \frac{7}{2} - \frac{1}{2}t.$$

Azt kaptuk, hogy a $X + Y$ sűrűségfüggvénye

$$f_{X+Y}(t) = \begin{cases} \frac{1}{2}t - 2 & \text{ha } 4 \leq t < 5 \\ \frac{1}{2} & \text{ha } 5 \leq t < 6 \\ \frac{7}{2} - \frac{1}{2}t & \text{ha } 6 \leq t \leq 7 \\ 0 & \text{egyébként} \end{cases}$$



Legyen $X_1, X_2, \dots, X_n$ n darab tojó tömege. Ekkor a feladat szerint függetlenek és $\text{Uni}(2, 3)$ eloszlásúak. Ekkor

$$\mu = \mathbb{E}(X_j) = \frac{2+3}{2} = \frac{5}{2} \quad \text{és} \quad \sigma = \mathbb{D}(X_j) = \frac{3-2}{\sqrt{12}} = \frac{1}{\sqrt{12}}.$$

Ekkor az n darab tojó együttes tömege

$$S_n = X_1 + X_2 + \dots + X_n.$$

Ekkor keressük azt az n -t, amire

$$0.9 \leq \mathbb{P}(S_n \leq 100).$$

Ekkor standardizálás után CHT közelítést alkalmazva

$$0.9 \leq \mathbb{P}\left(\frac{S_n - n \cdot \frac{5}{2}}{\sqrt{n} \cdot \frac{1}{\sqrt{12}}} \leq \frac{100 - n \cdot \frac{5}{2}}{\sqrt{n} \cdot \frac{1}{\sqrt{12}}}\right) \approx \Phi\left(\frac{100 - n \cdot \frac{5}{2}}{\sqrt{n} \cdot \frac{1}{\sqrt{12}}}\right)$$

Mivel $\Phi^{-1}(0.9) \approx 1.285$, ezért az alábbi egyenlőtlenségnek kell teljesülnie

$$1.285 \leq \frac{100 - n \cdot \frac{5}{2}}{\sqrt{n} \cdot \frac{1}{\sqrt{12}}},$$

amit átrendezve

$$n \leq 39.0725.$$

Vagyis nagyjából legfeljebb 39 tojót tehetünk egy talicskába, hogy 90%-os valószínűséggel elbírja őket.

6. Egy augusztusi napon a csillaghullások száma Poisson eloszlású. Naponta átlagosan 10 csillagot látunk. Mennyi a valószínűsége, hogy egymást követő 2 napon együttesen több, mint 25 csillaghullás lesz? Ha két napon együttesen volt 15 csillaghullás, akkor mennyi a valószínűsége, hogy az első napon 3 db volt?

Megoldás Legyen X_1 és X_2 az első és a második nap észlelt csillaghullások száma. Ekkor feltehető, hogy X_1 és X_2 független, illetve

$$X_1, X_2 \sim \text{Poi}(10).$$

Ekkor a két nap alatt látott csillagok száma

$$X_1 + X_2 \sim \text{Poi}(10 + 10) = \text{Poi}(20).$$

Ekkor annak a valószínűsége, hogy két egymás követő nap, több mint 25 csillaghullás lesz

$$\mathbb{P}(X_1 + X_2 > 25) = 1 - \mathbb{P}(X_1 + X_2 \leq 24) = 1 - \sum_{k=0}^{24} e^{-20} \frac{20^k}{k!} \approx 0.1568.$$

Továbbá annak a valószínűsége, hogy az első napon 3 darab csillaghullás volt, feltéve, hogy az első két napon 15 csillaghullás volt

$$\begin{aligned} \mathbb{P}(X_1 = 3 \mid X_1 + X_2 = 15) &= \frac{\mathbb{P}(X_1 = 3, X_1 + X_2 = 15)}{\mathbb{P}(X_1 + X_2 = 15)} = \frac{\mathbb{P}(X_1 = 3, X_2 = 12)}{\mathbb{P}(X_1 + X_2 = 15)} \\ &= \frac{\mathbb{P}(X_1 = 3) \cdot \mathbb{P}(X_2 = 12)}{\mathbb{P}(X_1 + X_2 = 15)} = \frac{e^{-10} \frac{10^3}{3!} \cdot e^{-10} \frac{10^{12}}{12!}}{e^{-20} \frac{20^{15}}{15!}} = \frac{455}{32768} \approx 0.0139, \end{aligned}$$

ahol kihasználtuk, hogy X_1 és X_2 független, így $\mathbb{P}(X_1 = 3, X_2 = 12) = \mathbb{P}(X_1 = 3) \cdot \mathbb{P}(X_2 = 12)$.

7. Los Angelesben az éves eső normális eloszlású, várható értéke 30 cm, szórása 8 cm. Mi a valószínűsége, hogy 2 egymást követő év együttes esőzése meghaladja a 63 cm-t? Mennyi a valószínűsége, hogy a következő év esőzése több, mint 8 cm-rel meghaladja az ideit? (különböző évek esőzései függetlenek egymástól.)

Megoldás Legyen X_1 és X_2 az első, illetve a második évben lehullott csapadék mennyisége (centiméterben). Ekkor a feladat szerint X_1 és X_2 független és

$$X_1, X_2 \sim \mathcal{N}(30, 64).$$

Ekkor az első két évben leesett csapadék mennyisége

$$X_1 + X_2 \sim \mathcal{N}(30 + 30, 64 + 64) = \mathcal{N}(60, 128).$$

Ekkor annak a valószínűsége, hogy az első két évben lehullott csapadék meghaladja a 63 centimétert

$$\mathbb{P}(X_1 + X_2 > 63) = \mathbb{P}\left(\frac{X_1 + X_2 - 60}{\sqrt{128}} > \underbrace{\frac{63 - 60}{\sqrt{128}}}_{\approx 0.27}\right) \approx 1 - \Phi(0.27) \approx 1 - 0.6064 = 0.3936.$$

Másfelől a következő évben leesett eső többete az ideihez képest (kihasználva, hogy $-X_1 \sim \mathcal{N}(-30, 64)$)

$$X_2 - X_1 \sim \mathcal{N}(30 - 30, 64 + 64) = \mathcal{N}(0, 128).$$

Ekkor annak a valószínűsége, hogy jövőre több, mint 8 centiméterrel meghaladja a eső az ideit

$$\mathbb{P}(X_2 - X_1 > 8) = \mathbb{P}\left(\frac{X_2 - X_1}{\sqrt{128}} > \underbrace{\frac{8}{\sqrt{128}}}_{\approx 0.71}\right) \approx 1 - \Phi(0.71) \approx 1 - 0.7611 = 0.2389.$$

8. Ha $X \sim \text{Geom}(p)$, akkor mi az alábbi változók súlyfüggvénye, várható értéke, szórása?

(a) $Y = 3X - 2$

(b) $Z = e^X$

Megoldás

(a) Vegyük észre, hogy $\text{Ran}(Y) = \{3k - 2 : k \in \mathbb{N}^+\}$. Ekkor amennyiben $k \in \mathbb{N}^+$

$$\mathbb{P}(Y = 3k - 2) = \mathbb{P}(3X - 2 = 3k - 2) = \mathbb{P}(X = k) = (1 - p)^{k-1} \cdot p,$$

vagyis amennyiben $m = 3k - 2$, akkor $k = \frac{m+2}{3}$, ezért Y súlyfüggvénye $m \in \text{Ran}(Y)$ esetén

$$\mathbb{P}(Y = m) = (1 - p)^{\frac{m+2}{3}-1} \cdot p.$$

Továbbá Y várható értéke

$$\mathbb{E}(Y) = \mathbb{E}(3X - 2) = 3 \cdot \mathbb{E}(X) - 2 = \frac{3}{p} - 2,$$

illetve szórása

$$\mathbb{D}(Y) = \mathbb{D}(3X - 2) = |3| \cdot \mathbb{D}(X) = 3 \cdot \frac{\sqrt{1-p}}{p}.$$

(b) Vegyük észre, hogy $\text{Ran}(Z) = \{e^k : k \in \mathbb{N}^+\}$. Ekkor amennyiben $k \in \mathbb{N}^+$

$$\mathbb{P}(Z = e^k) = \mathbb{P}(e^X = e^k) = \mathbb{P}(X = k) = (1 - p)^{k-1} \cdot p,$$

vagyis amennyiben $m = e^k$, akkor $k = \ln(m)$, ezért Z súlyfüggvénye $m \in \text{Ran}(Z)$ esetén

$$\mathbb{P}(Z = m) = (1 - p)^{\ln(m)-1} \cdot p.$$

Továbbá Z várható értéke

$$\mathbb{E}(Z) = \mathbb{E}(e^X) = \sum_{k=1}^{\infty} e^k \cdot (1-p)^{k-1} p = e \cdot p \sum_{k=0}^{\infty} (e(1-p))^k,$$

ezért

$$\mathbb{E}(Z) = \begin{cases} \frac{e \cdot p}{(1-e(1-p))} & \text{ha } p > 1 - e^{-1} \\ \infty & \text{egyébként} \end{cases}.$$

Továbbá Z második momentuma

$$\mathbb{E}(Z^2) = \mathbb{E}(e^{2X}) = \sum_{k=1}^{\infty} e^{2k} (1-p)^{k-1} p = \frac{p}{(1-p)} \sum_{k=1}^{\infty} (e^2(1-p))^k$$

amiből

$$\mathbb{E}(Z^2) = \begin{cases} \frac{e^2 \cdot p}{(1-e^2(1-p))} & \text{ha } p > 1 - e^{-2} \\ \infty & \text{egyébként} \end{cases}.$$

Amennyiben $p > 1 - e^{-2}$, akkor a variancia

$$\mathbb{D}^2(Z) = \frac{e^2 \cdot p}{(1-e^2(1-p))} - \left(\frac{e \cdot p}{(1-e(1-p))} \right)^2 = \frac{(e-1)^2 e^2 p(1-p)}{(1-e(1-p))^2 (1-e^2(1-p))}.$$

Ebben az esetben a szórás

$$\mathbb{D}(Z) = \frac{e(1-e)}{1-e(1-p)} \cdot \sqrt{\frac{p(1-p)}{(1-e^2(1-p))}}.$$

9. Legyen $X \sim \text{Exp}(\lambda)$ és $Y = X^2$ Adjuk meg Y sűrűségfüggvényét és várható értékét.

Megoldás Az Y eloszlásfüggvénye

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y).$$

Ekkor

(i) Ha $y < 0$, akkor

$$F_Y(y) = 0.$$

Ekkor a sűrűség

$$f_Y(y) = \frac{d}{dy} F_Y(y) = 0.$$

(ii) Ha $0 \leq y$, akkor mivel $\text{Ran}(X) = [0, \infty)$

$$F_Y(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(X \leq \sqrt{y}) = F_X(\sqrt{y}).$$

Ekkor a sűrűség

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(\sqrt{y}) = F'_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \lambda e^{-\lambda\sqrt{y}} \cdot \frac{1}{2\sqrt{y}}.$$

Vagyis Y sűrűségfüggvénye

$$f_Y(y) = \begin{cases} \frac{\lambda}{2\sqrt{y}} e^{-\lambda\sqrt{y}} & \text{ha } 0 < y \\ 0 & \text{egyébként} \end{cases}$$

Az Y várható értéke

$$\mathbb{E}(Y) = \mathbb{E}(X^2) = \mathbb{D}^2(X) + \mathbb{E}(X)^2 = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2}.$$

10. Legyen $X \sim \text{Uni}(0, 1)$. Adjuk meg az alábbi valószínűségi változók sűrűségfüggvényeit.

$$(a) Y = \sqrt{2X} \quad (b) V = \ln\left(\frac{1}{X}\right) \quad (c) Z = \arctan(X)$$

Megoldás Használjuk minden esetben, hogy $\text{Ran}(X) = [0, 1]$. Illetve, hogy

$$f_X(x) = \begin{cases} 1 & \text{ha } 0 \leq x \leq 1 \\ 0 & \text{egyébként} \end{cases}$$

(a) Mivel $Y = \sqrt{2X}$, ezért $\text{Ran}(Y) = [0, \sqrt{2}]$. Ekkor $0 \leq y \leq \sqrt{2}$ esetén

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\sqrt{2X} \leq y) = \mathbb{P}\left(X \leq \frac{y^2}{2}\right) = F_X\left(\frac{y^2}{2}\right)$$

Amiből a sűrűség

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y^2}{2}\right) = F'_X\left(\frac{y^2}{2}\right) \cdot y = f_X\left(\frac{y^2}{2}\right) \cdot y = y.$$

Azt kaptuk, hogy Y sűrűségfüggvénye

$$f_Y(y) = \begin{cases} y & \text{ha } 0 \leq y \leq \sqrt{2} \\ 0 & \text{egyébként} \end{cases}$$

(b) Mivel $V = \ln\left(\frac{1}{X}\right)$, ezért $\text{Ran}(V) = [0, \infty)$. Ekkor $v \geq 0$ esetén

$$\begin{aligned} F_V(v) &= \mathbb{P}(V \leq v) = \mathbb{P}\left(\ln\left(\frac{1}{X}\right) \leq v\right) = \mathbb{P}\left(\frac{1}{X} \leq e^v\right) \\ &= \mathbb{P}(X \geq e^{-v}) = 1 - \mathbb{P}(X \leq e^{-v}) = 1 - F_X(e^{-v}). \end{aligned}$$

Amiből a sűrűség

$$f_V(v) = \frac{d}{dv} F_V(v) = \frac{d}{dv} (1 - F_X(e^{-v})) = -F'_X(e^{-v}) \cdot (-e^{-v}) = f_X(e^{-v}) \cdot e^{-v} = e^{-v}.$$

Azt kaptuk, hogy V sűrűségfüggvénye

$$f_V(v) = \begin{cases} e^{-v} & \text{ha } 0 \leq v \\ 0 & \text{egyébként} \end{cases}$$

Megjegyezhetjük, hogy ekkor $V \sim \text{Exp}(1)$.

(c) Mivel $Z = \arctan(X)$, ezért $\text{Ran}(Z) = [0, \frac{\pi}{4}]$. Ekkor $0 \leq z \leq \frac{\pi}{4}$ esetén

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(\arctan(X) \leq z) = \mathbb{P}(X \leq \tan(z)) = F_X(\tan(z)).$$

Ekkor a sűrűség

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} F_X(\tan(z)) = F'_X(\tan(z)) \cdot \frac{1}{\cos^2(z)} = f_X(\tan(z)) \cdot \frac{1}{\cos^2(z)} = \frac{1}{\cos^2(z)}.$$

Azt kaptuk, hogy Z sűrűségfüggvénye

$$f_Z(z) = \begin{cases} \frac{1}{\cos^2(z)} & \text{ha } 0 \leq z \leq \frac{\pi}{4} \\ 0 & \text{egyébként} \end{cases}$$

11. Legyen X és Y független standard normális eloszlású valószínűségi változók. Mi lesz $X^2 + Y^2$ eloszlása?

Megoldás Legyen $X \sim \mathcal{N}(0, 1)$. Ekkor keressük meg X^2 sűrűségfüggvényét. Az eloszlásfüggvény

$$F_{X^2}(t) = \mathbb{P}(X^2 \leq t) = \begin{cases} \Phi(\sqrt{t}) - \Phi(-\sqrt{t}) & \text{ha } t > 0 \\ 0 & \text{egyébként} \end{cases},$$

amiből a sűrűségfüggvény

$$f_{X^2}(t) = \frac{d}{dt} F_{X^2}(t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp(-\frac{t}{2}) \frac{1}{\sqrt{t}} & \text{ha } t > 0 \\ 0 & \text{egyébként.} \end{cases}$$

Ekkor amennyiben $Y \sim \mathcal{N}(0, 1)$ független X -től, akkor $\text{Ran}(X^2 + Y^2) = (0, \infty)$, ezért amennyiben $t \leq 0$, akkor

$$f_{X^2+Y^2}(t) = 0.$$

Továbbá amennyiben $t > 0$, akkor $X^2 + Y^2$ sűrűségfüggvénye a t helyen, mivel Y^2 sűrűségfüggvénye megegyezik X^2 sűrűségfüggvényével, így a konvolúciós formula alapján

$$\begin{aligned} f_{X^2+Y^2}(t) &= \int_{-\infty}^{\infty} f_{X^2}(s) f_{Y^2}(t-s) ds = \int_0^t \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s} \frac{1}{\sqrt{s}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-s)} \frac{1}{\sqrt{t-s}} \\ &\quad \frac{1}{2\pi} e^{-\frac{1}{2}t} \int_0^t \frac{1}{\sqrt{s(t-s)}} ds = \frac{1}{2\pi} e^{-\frac{1}{2}t} \int_0^1 \frac{1}{\sqrt{u(1-u)}} du = \frac{1}{2\pi} e^{-\frac{1}{2}t} \cdot \pi = \frac{1}{2} e^{-\frac{1}{2}t}. \end{aligned}$$

Azt kaptuk, hogy

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

Ami pontosan azt jelenti, hogy

$$X^2 + Y^2 \sim \text{Exp}\left(\frac{1}{2}\right).$$

Megjegyzés Amennyiben $X_1, X_2, \dots, X_n$ független $\mathcal{N}(0, 1)$ eloszlású valószínűségi változók, akkor $X_1^2 + X_2^2 + \dots + X_n^2 \sim \chi^2(n)$ khi-négyszet eloszlást követ n szabadsági fokkal (ez a statisztikában egy fontos eloszlás). Az $n = 2$ egy speciális eset, ekkor pont exponenciális lesz.