Matematika A4 - Valószínűségszámítás 9. gyakorlat
Többváltozós eloszlások

Többváltozós diszkrét eloszlások

- **Együttes súlyfüggvény:** $p(x_i, y_j) = \mathbb{P}(X = x_i, Y = y_j)$
- X és Y független: $\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_j) \cdot \mathbb{P}(Y = y_j)$ minden i, j-re
- Peremsúlyfüggvény: $\mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j)$, $\mathbb{P}(Y = y_j) = \sum_{i=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j)$

1. feladat Először egy (szabályos) kockával dobunk, majd annyi (igazságos) érmével, ahányast a kockával dobtunk. Mi a valószínűsége, hogy a kockával 4-est dobunk és az érmékkel 2 fejet kapunk? Mi a valószínűsége, hogy 5 fejet kapunk? Adjuk meg az együttes-, perem- és feltételes eloszlásokat is!

Megoldásvázlat Az X (perem)eloszlása - a kockadobás eredménye - egyenletes az $\{1, 2, 3, 4, 5, 6\}$ halmazon. A dobott fejek száma (Y) függ a kockadobás eredményétől, így $Y|X = i$ binomiális $n = i$, $p = \frac{1}{2}$ paraméterrel, ahol $1 \leq i \leq 6$ egész szám. Ebből az együttes eloszlás:

$$\mathbb{P}(X = i, Y = j) = \mathbb{P}(Y = j|X = i)\mathbb{P}(X = i) = \binom{i}{j} 2^{-i} \cdot \frac{1}{6},$$

ha $0 \leq j \leq i$ egész.

Azaz

$$\mathbb{P}(Y = j) = \sum_{i=j}^{6} \mathbb{P}(X = i, Y = j) = \frac{1}{6j!} \sum_{i=j}^{6} i(i-1)\cdots(i-j+1) \cdot 2^{-i}.$$

A másik feltételes eloszlás:

$$\mathbb{P}(X = i|Y = j) = \frac{\mathbb{P}(X = i, Y = j)}{\mathbb{P}(Y = j)} = \frac{i(i-1)\cdots(i-j+1) \cdot 2^{-i}}{\sum_{k=j}^{6} k(k-1)\cdots(k-j+1) \cdot 2^{-k}}.$$

Tehát a kérdésekre a válaszok rendre:

- $p(X = 4, Y = 2) = \binom{4}{2} 2^{-4} \frac{1}{6} = \frac{1}{16}$,
- $\mathbb{P}(Y = 2) = \frac{1}{12} \sum_{i=2}^{6} i(i-1)2^{-i} = \frac{33}{128}$.

2. feladat Két kockával dobva, mi a dobott számok

(a) összegének és különbségének

(b) minimumának és maximumának

az együttes eloszlása? Függetlenek-e a marginálisok egymástól? Legyen (X, Y) a részfeladatoknak megfelelő valószínűségi változó pár. Számítsuk ki $\mathrm{COV}(X, Y)$ értékét!

Megoldásvázlat Jelölje U, illetve V a két független kockadobás eredményét.

(a) Ha $2 \leq n \leq 12$, $-5 \leq m \leq 5$ és $n > m$, akkor

$$\mathbb{P}(U+V = n, U-V = m) = \mathbb{P}\left(U = \frac{n+m}{2}, V = \frac{n-m}{2}\right) = \mathbb{P}\left(U = \frac{n+m}{2}\right) \cdot \mathbb{P}\left(V = \frac{n-m}{2}\right),$$

ami $\frac{1}{36}$, ha $n + m$ és $n - m$ is páros, egyébként pedig 0.

Nem független $U + V$ az $U - V$-től. Ezt a tényt például mutatja az $n = 10$, $m = -5$ választás, hiszen ekkor az együttes valószínűség: $\mathbb{P}(U + V = n, U - V = m) = 0$, de $\mathbb{P}(U + V = n) = \frac{1}{12}$ és $\mathbb{P}(U - V = -5) = \frac{1}{36}$.

Végül

$$\mathrm{COV}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}[(U+V)(U-V)] - \mathbb{E}(U+V)\mathbb{E}(U-V) = 0,$$

mert U, V azonos eloszlású és $(U+V)(U-V) = U^2 - V^2$ (tehát korrelálatlanok de nem függetlenek).

(b) A két valószínűségi változó (min. és max.) együttes eloszlása: ha $1 \le n, m \le 6$, akkor

$$\mathbb{P}(\min(U,V) = n, \max(U,V) = m) = \begin{cases} \frac{1}{36}, & \text{ha } n = m, \\ \frac{1}{18}, & \text{ha } n < m, \\ 0, & \text{egyébként.} \end{cases}$$

Definíció alapján kiszámolható, hogy $\mathbb{E}(XY) = 12,25$, $\mathbb{E}(\min(U,V)) = \frac{91}{36}$, $\mathbb{E}(\max(U,V)) = \frac{161}{36}$. Ebből $\mathrm{COV}(X,Y) = \frac{2555}{1296}$, így nem függetlenek.

3. feladat Kockával dobunk addig, amíg először páratlan szám jön! Ehhez szükséges dobások számát jelölje N és legyen X az első páratlan dobás értéke. Mi az eloszlása X-nek, illetve N-nek? Független-e X N-től? Adjuk meg az együttes eloszlásukat is!

Megoldásvázlat Az első páratlan dobás értéke: $\{1,3,5\}$ halmazból kerül ki egyenlő valószínűséggel. N eloszlása pedig geometriai, a siker valószínűsége: $\frac{1}{2}$. X N-től független, hiszen semmilyen információt nem ad az a tény, hogy N-edikre dobtunk páratlant arra vonatkozóan, hogy mi a páratlan dobás értéke, avagy megfordítva: az, hogy az első páratlan dobás mi lett, nem befolyásolja azt, hogy ezt hanyadikra dobtuk ki. Emiatt együttes eloszlásuk szorzat-eloszlás:

$$\mathbb{P}(X = x, N = n) = \mathbb{P}(X = x)\mathbb{P}(N = n) = \frac{1}{3}\left(\frac{1}{2}\right)^n \qquad x = 1,3,5,\ n = 1,2,\dots$$

Többváltozós folytonos eloszlások

Együttes eloszlásfüggvény: $F(x,y) := \mathbb{P}(X < x, Y < y)$. Tulajdonságai:

- $0 \le F(x,y) \le 1$ minden $x, y \in \mathbb{R}$ esetén, minden változójában balról folytonos és monoton növő
- $\lim\limits_{x\to-\infty} F(x,y) = \lim\limits_{y\to-\infty} F(x,y) = 0, \quad \lim\limits_{(x,y)\to(\infty,\infty)} F(x,y) = 1$

Együttes sűrűségfüggvény: $f(x,y) := \frac{\partial^2 F(x,y)}{\partial x \partial y}$. Tulajdonságai:

- $0 \le f(x,y)$ minden $x, y \in \mathbb{R}$ esetén
- $\iint\limits_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$

Ebből tudunk valószínűségeket számolni: $\mathbb{P}((X,Y) \in A) = \iint\limits_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y$.

Peremeloszlás- és sűrűségfüggvény (marginális):

- $F_X(x) = \lim\limits_{y\to\infty} F(x,y)$, $F_Y(y) = \lim\limits_{x\to\infty} F(x,y)$
- $f_X(x) = \frac{\partial F_X(x)}{\partial x} = \int\limits_{\mathbb{R}} f(x,y)\,\mathrm{d}y$, $f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = \int\limits_{\mathbb{R}} f(x,y)\,\mathrm{d}x$
- X és Y független, ha $F(x,y) = F_X(x) \cdot F_Y(y)$, vagyis $f(x,y) = f_X(x) \cdot f_Y(y)$ minden x, y esetén

Feltételes eloszlás- és sűrűségfüggvény:

- $F_{X|Y}(x|y) = \mathbb{P}(X < x \,|\, Y = y) = \frac{1}{f_Y(y)} \cdot \frac{\partial F(x,y)}{\partial y}$, $F_{Y|X}(y|x) = \mathbb{P}(Y < y \,|\, X = x) = \frac{1}{f_X(x)} \cdot \frac{\partial F(x,y)}{\partial x}$
- $f_{X|Y}(x|y) = \frac{\partial F_{X|Y}(x|y)}{\partial x} = \frac{f(x,y)}{f_Y(y)}$, $f_{Y|X}(y|x) = \frac{\partial F_{Y|X}(y|x)}{\partial y} = \frac{f(x,y)}{f_X(x)}$
- Folytonos Bayes-formula: $f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y) \cdot f_Y(y)}{\int\limits_{\mathbb{R}} f_{X|Y}(x|y) \cdot f_Y(y)\,\mathrm{d}y}$

4. feladat Legyen $f(x,y)=\frac{c}{y}$, ha $0<x<1$ és $x<y<\frac{1}{x}$, egyébként pedig $f(x,y)=0$.

(a) Határozzuk meg c értékét úgy, hogy $f(x,y)$ egy sűrűségfüggvény legyen!

(b) Határozzuk meg a peremeloszlás- és sűrűségfüggvényeket!

(c) Mik a feltételes eloszlás- és sűrűségfüggvények?

(d) Mennyi $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X|Y=y)$, $\mathbb{E}(Y|X=x)$, $\mathbb{E}(XY)$, $\mathrm{COV}(X,Y)$?

(e) Független-e X és Y?

Megoldásvázlat

(a) $\int\limits_0^1\int\limits_x^{1/x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x=1$ feltételből adódik, hogy $c=1/2$.

(b)
- $f_X(x)=\int\limits_x^{1/x}\frac{1}{2y}\,\mathrm{d}y=-\ln x$, ha $x\in(0,1)$, egyébként $f_X(x)=0$.
- Ebből integrálással $F_X(x)=x(1-\ln x)$, ha $x\in(0,1)$.
- $f_Y(y)=\int\limits_0^{\min(y,1/y)}\frac{1}{2y}\,\mathrm{d}x$, vagyis $f_Y(y)=1/2$, ha $0<y<1$, $f_Y(y)=\frac{1}{2y^2}$, ha $y>1$.
- Ebből integrálással $F_Y(y)=\frac{y}{2}$, ha $0<y<1$ és $F_Y(y)=1-\frac{1}{2y}$, ha $y>1$.

(c)
- $f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$ alapján:
  ha $y\in(0,1)$, akkor $f_{X|Y}(x|y)=\frac{1}{y}$, ha $x\in(0,y)$, egyéb x esetén pedig 0, vagyis
  $$X|Y=y\sim \mathrm{UNI}(0,y),$$
  ha $y>1$, akkor $f_{X|Y}(x|y)=y$, ha $x\in\left(0,\frac{1}{y}\right)$, egyéb x esetén pedig 0, vagyis
  $$X|Y=y\sim \mathrm{UNI}\left(0,\frac{1}{y}\right).$$
- $F_{X|Y}(x|y)=\int\limits_{-\infty}^{x} f_{X|Y}(t|y)\,\mathrm{d}t$, amiből
  ha $y\in(0,1)$, akkor $F_{X|Y}(x|y)=\frac{x}{y}$, ha $0<x<y$
  ha $y>1$, akkor $F_{X|Y}(x|y)=xy$, ha $0<x<\frac{1}{y}$
  (itt az eloszlásfüggvényeket az alapján is fel tudtuk volna írni, hogy tudjuk, hogy a feltételes eloszlás egyenletes).
- $f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{-1}{2y\ln x}$, ha $0<x<1$ és $x<y<\frac{1}{x}$, egyéb y esetén pedig 0
- $F_{Y|X}(y|x)=\int\limits_{-\infty}^{y} f_{Y|X}(t|x)\,\mathrm{d}t=\frac{1}{2}-\frac{\ln y}{2\ln x}$, ha $0<x<1$ és $x<y<\frac{1}{x}$.

(d)
- $\mathbb{E}(X)=\int\limits_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x=\frac{1}{4}$
- $\mathbb{E}(Y)=\int\limits_{-\infty}^{\infty} y f_Y(y)\,\mathrm{d}y=\infty$
- Mivel $X|Y$ feltételes eloszlás egyenletes (l. (c) feladat), így a feltételes várható érték $\frac{y}{2}$, ha $0<y<1$ és $\frac{1}{2y}$, ha $y>1$.
- $\mathbb{E}(Y|X=x)=\int\limits_{-\infty}^{\infty} y f_{Y|X}(y|x)\,\mathrm{d}y=\int\limits_x^{1/x}\frac{-1}{\ln x}\,\mathrm{d}y=\frac{x-\frac{1}{x}}{\ln x}$
- $\mathbb{E}(XY)=\iint\limits_{\mathbb{R}^2} xyf(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int\limits_0^1\int\limits_x^{1/x}\frac{x}{2}\,\mathrm{d}y\,\mathrm{d}x=\cdots=\frac{1}{3}$
- $\mathrm{COV}(XY)$ nem létezik, hiszen $\mathbb{E}(Y)$ sem létezik

(e) Nem, hiszen $f(x,y)\neq f_X(x)f_Y(y)$.

5. feladat A $(0;0)$, $(1;0)$, $(1;1)$ csúcsú háromszögön vett egyenletes eloszlás esetén határozzuk meg $Z = Y/X$ eloszlását!

Megoldásvázlat Z értéke definíciója alapján 0 és 1 között van. Így $z \in (0,1)$ esetén:

$$F_Z(z) = \mathbb{P}(Z < z) = \mathbb{P}(Y < z \cdot X) \overset{(*)}{=} \frac{z/2}{1/2} = z,$$

ahol$(*)$-nál azt használtuk, hogy a háromszög azon részének a területe, ahol $y < z \cdot x$, az $z/2$, míg az egész háromszög területe $1/2$ (egyenletes eloszlás miatt a valószínűségek arányosak a területekkel).

A kapott eloszlásfüggvényből látjuk, hogy $Z \sim \mathrm{UNI}(0,1)$.

6. feladat Vegyünk egy két dimenziós (X, Y) eloszlást, amelynek sűrűségfüggvénye $f(x, y) = 4xy$ ha $0 < x < 1$ és $0 < y < 1$, egyébként 0. Számoljuk ki külön-külön $U = XY$ és $V = Y/X$ eloszlásfüggvényét és sűrűségfüggvényét!

Megoldásvázlat

- U értéke definíciója alapján 0 és 1 között van. Így $u \in (0,1)$ esetén:

  $$F_U(u) = \mathbb{P}(U < u) = \mathbb{P}\left(Y < u \cdot \frac{1}{X}\right) \overset{(*)}{=} 1 - \int_u^1 \int_{u/x}^1 f(x,y)\,\mathrm{d}y\,\mathrm{d}x = ... = u^2 - 2u^2 \ln u,$$

  ahol $(*)$-nál azt használtuk, hogy azok a pontok az egységnégyzeten, amelyekre $y < u \cdot \frac{1}{x}$ NEM teljesül az épp $\{(x,y)\,:\, u \le x \le 1, u/x \le y \le 1\}$.

  Ebből deriválással a sűrűségfüggvény $f_U(u) = 4u \ln u$, ha $u \in (0,1)$, egyébként $f_U(u) = 0$.

- V értéke definíciója alapján 0 és ∞ között van. Viszont rajz alapján látjuk, hogy $v < 1$ és $v \ge 1$ eseteket külön kell kezelni. Így $v \in (0,1)$ esetén:

  $$F_V(v) = \mathbb{P}(V < v) = \mathbb{P}\left(Y < vX\right) \overset{(*)}{=} \int_0^1 \int_0^{vx} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = ... = \frac{v^2}{2},$$

  ahol $(*)$-nál azt használtuk, hogy azok a pontok az egységnégyzeten, amelyekre $y < vx$ teljesül az épp $\{(x,y)\,:\, 0 \le x \le 1, 0 \le y \le vx\}$.

  $v \in (1,\infty)$ esetén:

  $$F_V(v) = \mathbb{P}(V < v) = \mathbb{P}\left(Y < uX\right) \overset{(*)}{=} 1 - \int_0^1 \int_0^{y/v} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = ... = 1 - \frac{1}{2v^2},$$

  ahol $(*)$-nál azt használtuk, hogy azok a pontok az egységnégyzeten, amelyekre $y < vx$ NEM teljesül az épp $\{(x,y)\,:\, 0 \le y \le 1, 0 \le x \le y/v\}$.

  Ebből deriválással a sűrűségfüggvény $f_V(v) = v$, ha $v \in (0,1)$, míg $f_V(v) = v^{-3}$, ha $v \in (1,\infty)$ egyébként $f_V(v) = 0$.

7. feladat Egy (egységsugarú) körbe szabályos hatszöget rajzolunk. Mi a valószínűsége, hogy a körben egyenletesen választott pont a hatszögben is benne van?

Megoldásvázlat Mivel az egységsugarú körben egyenletes eloszlás szerint választunk pontot, ezért a szóban forgó valószínűség $\frac{\text{kedvező terület}}{\text{összterület}}$, ahol a kedvező rész a hatszög területe, ami 6 egyenlő (egységnyi) oldalú háromszög területe azaz $3\sqrt{3}/2$ az összterület pedig π, azaz

$$\mathbb{P}(\text{a véletlen pont beleesik a hatszögbe}) = \frac{3\sqrt{3}}{2\pi} \approx 0.827.$$

8. feladat X_1, X_2, X_3 három véletlen pontot jelöl a $[0,3]$ intervallumból. Mi annak a valószínűsége, hogy az X_2 pont X_1 és X_3 között helyezkedik el?

Megoldásvázlat A három véletlen pont legyen független egymástól. Az együttes sűrűségfüggvényük ezért a $[0,3]^3$ kockán $\frac{1}{27}$ egyébként 0. Ekkor

$$\mathbb{P}(X_1 < X_2 < X_3) = \iiint\limits_{0 \le x_1 < x_2 < x_3 \le 3} f(x_1, x_2, x_3)\, \mathrm{d}x_3\, \mathrm{d}x_2\, \mathrm{d}x_1 = \frac{1}{27} \int\limits_0^3 \int\limits_{x_1}^3 \int\limits_{x_2}^3 \mathrm{d}x_3\, \mathrm{d}x_2\, \mathrm{d}x_1 = \cdots = \frac{1}{6}$$

Szimmetria miatt $\mathbb{P}(X_3 < X_2 < X_1) = \frac{1}{6}$, így a kérdezett valószínűség $\frac{1}{3}$.

Megjegyzés: szimmetria miatt azt is tudjuk, hogy X_1, X_2, X_3 minden sorrendje egyforma valószínű, így számolás nélkül is tudhatjuk, hogy $\frac{1}{3}$ valószínűséggel lesz X_2 a középső a három változó közül.