Matematika A4 - Valószínűségszámítás 8. 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 $\Leftrightarrow \mathbb{P}(X = x_i,\, Y = y_j) = \mathbb{P}(X = x_i) \cdot \mathbb{P}(Y = y_j)$ minden i, j-re
- Peremsúlyfüggvény (marginális):

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_i,\, Y = y_j), \quad \mathbb{P}(Y = y_j) = \sum_{i=1}^{\infty} \mathbb{P}(X = x_i,\, Y = y_j)$$

- Feltételes súlyfüggvény:

$$\mathbb{P}(X = x_i \,|\, Y = y_j) = \frac{\mathbb{P}(X = x_i,\, Y = y_j)}{\mathbb{P}(Y = y_j)}, \quad \mathbb{P}(Y = y_j \,|\, X = x_i) = \frac{\mathbb{P}(Y = y_j,\, X = x_i)}{\mathbb{P}(X = x_i)}$$

- Kovariancia: $\mathrm{COV}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$
- Ha X és Y független, akkor $\mathrm{COV}(X, Y) = 0$, de ez visszafele nem igaz

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 a kockadobás eredményének és a fejek számának az együttes-, a perem- és a feltételes eloszlásait!

Megoldásvázlat Legyen X a kockadobás eredménye, Y pedig a dobott fejek száma!

- X (perem)eloszlása egyenletes az $\{1, 2, 3, 4, 5, 6\}$ halmazon, azaz X peremsúlyfüggvénye:

$$\mathbb{P}(X = i) = \frac{1}{6} \text{ minden } i \in \{1, 2, 3, 4, 5, 6\} \text{ esetén.}$$

- A dobott érmék száma függ a kockadobás eredményétől, így Y-nak az X-re vett feltételes eloszlását tudjuk könnyen megmondani. Az Y változó feltételes eloszlása az $\{X = i\}$ feltétel mellett (rövidebben $(Y|X = i)$) binomiális $n = i$, $p = \frac{1}{2}$ paraméterekkel, ahol $1 \leq i \leq 6$ egész szám. Vagyis Y feltételes súlyfüggvénye X-re nézve:

$$\mathbb{P}(Y = j \,|\, X = i) = \binom{i}{j} \cdot 2^{-i}, \quad \text{ha } 0 \leq j \leq i \leq 6,\ i \neq 0 \text{ egészek.}$$

- A fentiekből az együttes eloszlás:

$$\mathbb{P}(X = i,\, Y = j) = \mathbb{P}(Y = j \,|\, X = i) \cdot \mathbb{P}(X = i) = \binom{i}{j} \cdot 2^{-i} \cdot \frac{1}{6}, \quad \text{ha } 0 \leq j \leq i \leq 6,\ i \neq 0 \text{ egészek.}$$

- Így Y peremeloszlása:

$$\mathbb{P}(Y = j) = \sum_{i=j}^{6} \mathbb{P}(X = i, Y = j) = \frac{1}{6} \sum_{i=\max(j,1)}^{6} \binom{i}{j} \cdot 2^{-i}, \quad \text{ha } j \in \{0, 1, \ldots, 6\}.$$

- Végül pedig a másik feltételes eloszlás:

$$\mathbb{P}(X = i \,|\, Y = j) = \frac{\mathbb{P}(X = i,\, Y = j)}{\mathbb{P}(Y = j)} = \frac{\binom{i}{j} \cdot 2^{-i}}{\sum\limits_{k=\max(j,1)}^{6} \binom{k}{j} \cdot 2^{-k}}, \quad \text{ha } 0 \leq j \leq i \leq 6,\ i \neq 0 \text{ egészek.}$$

Tehát a feladat kérdéseire a válaszok:

- $\mathbb{P}(X = 4,\, Y = 2) = \binom{4}{2} \cdot 2^{-4} \cdot \frac{1}{6} = \frac{1}{16}$,
- $\mathbb{P}(Y = 5) = \sum\limits_{i=5}^{6} \mathbb{P}(X = i,\, Y = 5) = \frac{1}{6} \cdot \frac{1}{2^5} + \frac{1}{6} \cdot 6 \cdot \frac{1}{2^6} = \frac{1}{48}$.

2. feladat Két kockával dobva, mi a dobott számok

(a) összegének és különbségének (ahol mindig ugyanannak a kockának az eredményéből vonjuk ki a másikat, vagyis a különbség lehet negatív)

(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) Legyen n és m egész, melyekre $2 \le n \le 12$ és $-5 \le m \le 5$. Ekkor

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

ami $1/36$, ha $n+m$ és $n-m$ is páros (azaz n és m is páros vagy mindkettő páratlan) és $2 \le n-m \le 12$, $2 \le n+m \le 12$, 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 $\mathbb{P}(U+V=n)=1/12$ és $\mathbb{P}(U-V=-5)=1/36$.

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

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

Definíció alapján kiszámolható, hogy

$$\mathbb{E}(XY)=12.25, \quad \mathbb{E}(\min(U,V))=\frac{91}{36}, \quad \mathbb{E}(\max(U,V))=\frac{161}{36}.$$

Ebből $\mathrm{COV}(X,Y)=1225/1296$, 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 az X változó az N változótól? 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: $\frac{1}{2}$. Az X változó független N-től, hiszen semmilyen információt nem ad az a tény, hogy N-edikre dobtunk páratlant arra vonatkozóan, hogy mi a páratlan dobás értéke. Vagy megfordítva: az, hogy az első páratlan dobás mi lett, nem befolyásolja azt, hogy ezt hanyadikra dobtuk ki. Emiatt az együttes súlyfüggvényük a peremsúlyfüggvények szorzataként számolható:

$$\mathbb{P}(X=x,\, N=n) = \mathbb{P}(X=x)\cdot\mathbb{P}(N=n) = \frac{1}{3}\cdot\left(\frac{1}{2}\right)^n, \quad \text{ha } x=1,3,5, \quad n=1,2,\ldots$$

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
- $\lim\limits_{x\to-\infty} F(x,y) = \lim\limits_{y\to-\infty} F(x,y) = 0$, $\lim\limits_{(x,y)\to(\infty,\infty)} F(x,y) = 1$

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 X, Y együttes sűrűségfüggvénye $f(x,y) = c \cdot xy$ a $0 < x < 1$, $x^2 < y < \sqrt{x}$ tartományon (és nulla egyébként)!

(a) Számold ki c-t!

(b) Számold ki a peremsűrűségeket!

(c) Számold ki a feltételes sűrűségeket!

(d) Számold ki $\mathbb{E}(X)$-et és $\mathbb{E}(Y)$-t!

(e) Független-e X és Y? (Indokolj!)

Megoldásvázlat

(a) Olyan c-t keresünk, amire:

$$\int\limits_0^1 \int\limits_{x^2}^{\sqrt{x}} cxy\,\mathrm{d}y\,\mathrm{d}x = 1$$

Ebből $c = 12$ adódik. (Ekkor $f(x,y) \ge 0$ is teljesül.)

(b) Fenti képletek alapján együttes sűrűségfüggvényből számolhatók.

- X peremsűrűségfüggvénye:

$$f_X(x) = \int\limits_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int\limits_{x^2}^{\sqrt{x}} 12xy\,\mathrm{d}y = 6(x^2 - x^5), \text{ ha } x \in (0,1),$$

míg $x \notin (0,1)$ esetén $f_X(x) = 0$.

- Y peremsűrűségfüggvénye:
  Az $f_Y(y)$ számolásához a $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$ képletet tudjuk használni. Ehhez érdemes először felrajzolni a $0 < x < 1$, $x^2 < y < \sqrt{x}$ tartományt, mert most x szerint kell majd integrálni. Első lépésben az ábráról le kell olvasni, hogy y milyen konstansok között mozog (azaz milyen y-nál van a tartomány legalsó és legfelső pontja), ez most 0 és 1. Utána pedig azt kell megvizsgálni, hogy ha fixálok egy y-t, akkor a tartományban ahhoz az y-hoz milyen x értékek jönnek szóba. (Ez persze függhet y-tól.) Az ábráról leolvasható, hogy fix y esetén az x-ek az $y = \sqrt{x}$, azaz $x = y^2$ görbétől az $y = x^2$, azaz $x = \sqrt{y}$ görbéig mennek. Tehát a $0 < x < 1$, $x^2 < y < \sqrt{x}$ tartományt úgy is fel lehet írni, hogy: $0 < y < 1$, $y^2 < x < \sqrt{y}$. Most már tudjuk számolni a peremsűrűségfüggvényt:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x = \int_{y^2}^{\sqrt{y}} 12xy\,\mathrm{d}x = 6(y^2 - y^5), \text{ ha } y \in (0,1),$$

  míg $y \notin (0,1)$ esetén $f_Y(y) = 0$.

  Megjegyzés: Ha észrevesszük, hogy a sűrűségfüggvény szimmetrikus a változókban, akkor a második számolásra nem lett volna szükség, hiszen szimmetriából következik, hogy a peremsűrűségek megegyeznek.

(c) Együttes- és peremsűrűségekből egy osztással számolhatók.

- X feltételes sűrűségfüggvénye Y-ra nézve:
  - Legyen $y \in (0,1)$. Ekkor

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{2x}{y - y^4}, \text{ ha } x \in (y^2, \sqrt{y}),$$

    míg $x \notin (y^2, \sqrt{y})$ esetén $f_{X|Y}(x|y) = 0$.
  - Viszont, ha $y \notin (0,1)$, akkor a feltételes sűrűségfüggvény nem definiált ebben a példában.
- Y feltételes sűrűségfüggvénye X-ra nézve:
  - Legyen $x \in (0,1)$. Ekkor

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{2y}{x - x^4}, \text{ ha } y \in (x^2, \sqrt{x}),$$

    míg $y \notin (x^2, \sqrt{x})$ esetén $f_{Y|X}(y|x) = 0$.
  - Viszont, ha $x \notin (0,1)$, akkor a feltételes sűrűségfüggvény nem definiált ebben a példában.

(d) Peremsűrűségből a várható érték definíciója alapján számolható:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x = \int_0^1 6(x^3 - x^6)\,\mathrm{d}x = \frac{9}{14}.$$

  Hasonlóan $\mathbb{E}(Y) = \frac{9}{14}$.

(e) Nem, hiszen az együttes sűrűségfüggvény nem egy téglalap alakú tartományon pozitív. (Így $f(x,y) = f_X(x) \cdot f_Y(y)$ nem teljesülhet minden pontban.)

5. feladat Egy villamosmérnök A4 zh eredménye elsősorban a ~~gyakorlatvezető hangulatától~~ készülésre szánt időtől függ, aminek az eloszlása egyenletes 0 és 30 óra között. Ha a hallgató x órát készült, akkor a zh százalékos eredményének eloszlása szintén egyenletes, de $3x$ és $3x+10$ százalék között. Ha tudjuk egy hallgató zh eredményét, akkor mit mondhatunk a készülésre szánt idejéről? (Vagyis adjuk meg a feltételes sűrűségfüggvényt!)

Megoldásvázlat Legyen X a készülésre szánt idő és Y a zh eredménye százalékban. Tudjuk, hogy $X \sim \mathrm{UNI}(0,30)$ és $(Y|X=x) \sim \mathrm{UNI}(3x, 3x+10)$. Vagyis $f_X(x) = 1/30$, ha $0 < x < 30$, másrészt $f_{Y|X}(y|x) = 1/10$, ha $3x < y < 3x+10$. Ebből

$$f(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = \frac{1}{300}, \quad 0 < x < 30,\ 3x < y < 3x+10.$$

Innen $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$, vagyis a konstans $\frac{1}{300}$ függvényt kell integrálni, de kérdés, hogy milyen intervallumon. Ha lerajzoljuk a $0 < x < 30$, $3x < y < 3x+10$ paralelogrammát, akkor azt kell nézni, hogy egy fix y esetén hol vannak a hozzá tartozó x-ek. Legkönnyebben az látszódik, hogy csak $0 < y < 100$ érdekes, hiszen a paralelogramma legalsó pontja $y = 0$-nál van, legfelső pedig $y = 100$-nál. (Egyébként Y definíciójából is nyilvánvaló, hogy 0 és 100 közötti értékeket vesz fel.)

A kérdés tehát, hogy egy rögzített y esetén hol van a tartomány bal széle (ez lesz x-re az alsó határ, így az integrálási intervallum eleje) és jobb széle (ez meg a felső határ). Az ábráról könnyen látszódik, hogy például $0 < y < 10$ esetén a bal szél az $x = 0$ függőleges egyenes, míg a jobb szél az $y = 3x$, azaz $x = y/3$ egyenes. Így ekkor 0-tól $\frac{y}{3}$-ig kell integrálni a konstans 1/300-at. A többi eset hasonló: $10 < y < 90$ esetén $(y-10)/3$-tól $y/3$-ig kell integrálni, $90 < y < 100$ esetén pedig $(y-10)/3$-tól 30-ig. Ezekből kapjuk az alábbi peremsűrűségfüggvényt:

$$f_Y(y) = \begin{cases} y/900, & \text{ha } y \in (0,10), \\ 1/90, & \text{ha } y \in (10,90), \\ (100-y)/900, & \text{ha } y \in (90,100), \\ 0 & \text{egyébként.} \end{cases}$$

Ebből a feltételes sűrűségfüggvény számolható az $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$ képlettel.

- $0 < y < 10$ esetén $f_{X|Y}(x|y) = \frac{3}{y}$, ha $0 < x < \frac{y}{3}$, azaz $(X|Y=y) \sim \mathrm{UNI}\left(0, \frac{y}{3}\right)$.
- $10 < y < 90$ esetén $f_{X|Y}(x|y) = \frac{3}{10}$, ha $\frac{y-10}{3} < x < \frac{y}{3}$, azaz $(X|Y=y) \sim \mathrm{UNI}\left(\frac{y-10}{3}, \frac{y}{3}\right)$.
- $90 < y < 100$ esetén $f_{X|Y}(x|y) = \frac{3}{100-y}$, ha $\frac{y-10}{3} < x < 30$, azaz $(X|Y=y) \sim \mathrm{UNI}\left(\frac{y-10}{3}, 30\right)$.

Geometriai valószínűségek

6. 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 $Z \sim \mathrm{UNI}(0,1)$.

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

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