Matematika A4 - Valószínűségszámítás 10. gyakorlat
Többváltozós eloszlások, feltételes várható érték és variancia, vegyes eloszlás

- Kovariancia: $\mathrm{COV}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$, ahol $\mathbb{E}(XY)=\iint\limits_{\mathbb{R}^2} f(x,y)\cdot xy\,\mathrm{d}x\,\mathrm{d}y$
- Feltételes várható érték: $\mathbb{E}(X|Y=y)=\int\limits_{-\infty}^{\infty} f_{X|Y}(x|y)\cdot x\,\mathrm{d}x$
- Feltételes variancia: $\mathbb{D}^2(X|Y=y)=\mathbb{E}\left(X^2|Y=y\right)-(\mathbb{E}(X|Y=y))^2$

1. 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?

(f) Számoljuk ki a feltételes szórásnégyzeteket!

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

(f)
- Láttuk, hogy $(X|Y = y)$ eloszlása egyenletes, így ebből a szórásnégyzet is könnyen meghatározható:
  $0 < y < 1$ esetén $\mathbb{D}^2(X|Y = y) = \frac{y^2}{12}$, míg $1 < y$ esetén $\mathbb{D}(X^2|Y = y) = \frac{1}{12y^2}$.
- $\mathbb{D}^2(Y|X = x)$ számolásához szükségünk van $\mathbb{E}\left(Y^2|X = x\right)$-ra:

$$\mathbb{E}\left(Y^2|X = x\right) = \int\limits_{x}^{1/x} y^2 \cdot \frac{-1}{2y\ln x}\,\mathrm{d}y = \cdots = \frac{x^2 - \frac{1}{x^2}}{4\ln x}$$

  Innen

$$\mathbb{D}^2(Y|X = x) = \mathbb{E}\left(Y^2|X = x\right) - (\mathbb{E}(Y|X = x))^2 = \frac{x^2 - \frac{1}{x^2}}{4\ln x} - \left(\frac{x - \frac{1}{x}}{\ln x}\right)^2.$$

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

3. 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:

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

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

  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 $v \geq 1$ eseteket külön kell kezelni. Így $v \in (0,1)$ esetén:

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

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

$v \in (1,\infty)$ esetén:

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

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

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 $f_V(v) = 0$.

4. 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 $f_X(x) = \frac{1}{30}$, ha $0 < x < 30$, másrészt $f_{Y|X}(y|x) = \frac{1}{10}$, 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 $f_Y(y) = \int\limits_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$, vagyis

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

Ebből a feltételes sűrűségfüggvény számolható a $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$ 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 $(X|Y=y) \sim \mathrm{UNI}\left(\frac{y-10}{3}, 30\right)$.

5. feladat Legyen (X,Y) eloszlása egyenletes azon a négyszögön, aminek a csúcsai

(a) $(1;1)$, $(3;1)$, $(3;2)$, $(1;2)$,

(b) $(1;0)$, $(0;1)$, $(-1;0)$, $(0;-1)$.

- Adjuk meg (X,Y) együttes sűrűségfüggvényét!
- Mik a marginális sűrűségfüggvények és várható értékek?
- Független-e X és Y?
- Mik a feltételes sűrűségfüggvények, várható értékek, varianciák?
- Mennyi $\mathbb{E}(XY)$ és $\mathrm{COV}(X,Y)$?

Megoldásvázlat

(a)

- A téglalap területe 2, így az együttes sűrűségfüggvény $f(x,y) = \frac{1}{2}$, ha $1 < x < 3$ és $1 < y < 2$, míg a téglalapon kívül $f(x,y) = 0$.
- Ha $1 < x < 3$, akkor
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int_1^2 \frac{1}{2}\,\mathrm{d}y = \frac{1}{2}.$$
Vagyis $X \sim \mathrm{UNI}(1,3)$, így $\mathbb{E}(X) = 2$.
Hasonlóan $1 < y < 2$ esetén
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x = \int_1^3 \frac{1}{2}\,\mathrm{d}x = 1.$$
Vagyis $Y \sim \mathrm{UNI}(1,2)$, így $\mathbb{E}(Y) = 1,5$.
- Jól látszódik, hogy $f(x,y) = f_X(x) \cdot f_Y(y)$ minden x, y esetén (a tartományra is figyelve). Tehát X és Y független.
- Függetlenség miatt a feltételes eloszlások megegyeznek a peremeloszlásokkal, így
$$f_{X|Y}(x|y) = f_X(x) = \frac{1}{2}, \qquad \text{ha } 1 < x < 3,\ 1 < y < 2,$$
$$\mathbb{E}(X|Y=y) = \mathbb{E}(X) = 2,\ \mathbb{D}^2(X|Y=y) = \mathbb{D}^2(X) = \frac{1}{3},$$
$$f_{Y|X}(y|x) = f_Y(y) = 1, \qquad \text{ha } 1 < x < 3,\ 1 < y < 2,$$
$$\mathbb{E}(Y|X=x) = \mathbb{E}(Y) = 1,5,\ \mathbb{D}^2(Y|X=x) = \mathbb{D}^2(Y) = \frac{1}{12}.$$
- A szorzat várható értéke és a kovariancia
$$\mathbb{E}(XY) = \iint_{\mathbb{R}^2} xyf(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_1^3 \int_1^2 \frac{xy}{2}\,\mathrm{d}y\,\mathrm{d}x = \cdots = 3$$
$$\mathrm{COV}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$$
(a függetlenség miatt is tudjuk, hogy a kovariancia 0).

(b)

- A négyzet területe 2, így az együttes sűrűségfüggvény $f(x,y) = \frac{1}{2}$ a négyzeten (azaz ha $-1 < x < 0$ és $-1 - x < y < x + 1$ vagy $0 < x < 1$ és $x - 1 < y < 1 - x$), míg a négyzeten kívül $f(x,y) = 0$.
- Az együttes sűrűségfüggvény integrálásából kapjuk, hogy $f_X(x) = x + 1$, ha $-1 < x < 0$ és $f_X(x) = 1 - x$ ha $0 < x < 1$. Egyéb x esetén pedig $f_X(x) = 0$.
Hasonlóan (szimmetria miatt) $f_Y(y) = y+1$, ha $-1 < y < 0$ és $f_Y(y) = 1-y$, ha $0 < y < 1$. Egyéb y esetén pedig $f_Y(y) = 0$.
A várható értékek definícióból számolhatók: $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ (ez abból is következik, hogy a sűrűségfüggvények páros függvények).
- Jól látszódik, hogy $f(x,y) = f_X(x) \cdot f_Y(y)$ nem teljesül minden x, y esetén, így X és Y nem független.
- $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$, így $-1 < y < 0$ esetén $f_{X|Y}(x|y) = \frac{1}{2(y+1)}$, egyébként pedig 0. Vagyis ekkor $(X|Y=y)$ feltételes eloszlása $\mathrm{UNI}(-1-y, y+1)$. Ebből a feltételes várható érték és variancia is könnyen megállapítható:
$$\mathbb{E}(X|Y=y) = 0, \quad \mathbb{D}^2(X|Y=y) = \frac{(y+1)^2}{3}.$$

$0 < y < 1$ esetén hasonlóan kapjuk, hogy $f_{X|Y}(x|y) = \frac{1}{2(1-y)}$, azaz most $(X|Y = y)$ feltételes eloszlása UNI$(y-1, 1-y)$. Ebből a feltételes várható érték és variancia:

$$\mathbb{E}(X|Y = y) = 0, \quad \mathbb{D}^2(X|Y = y) = \frac{(1-y)^2}{3}.$$

Mivel az eloszlás szimmetrikus a változókra nézve, így $f_{Y|X}(y|x)$-t számolva ugyanezeket a feltételes sűrűségfüggvényeket, várható értékeket és szórásnégyzeteket kapjuk, csak X és Y (illetve x és y) szerepét mindenhol fel kell cserélni.

- $\mathbb{E}(XY) = \iint\limits_{\mathbb{R}^2} xyf(x,y)\,\mathrm{d}x\,\mathrm{d}y = \cdots = 0$ és így a kovariancia is 0 (de a változók nem függetlenek).

6. feladat Sűrűségfüggvények-e az alábbiak? Ha igen, adjuk meg a perem- és feltételes sűrűségfüggvényeket, várható értékeket, illetve a kovarianciát!

(a) $f(x,y) = \frac{4}{5}(x + xy + y)$, ha $0 < x < 1, 0 < y < 1$, egyébként $f(x,y) = 0$

(b) $f(x,y) = \lambda^2 e^{-\lambda(x+y)}$, ha $x, y > 0$, egyébként $f(x,y) = 0$

Megoldásvázlat

(a) Ellenőrizhető, hogy $\iint\limits_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$, így ez egy sűrűségfüggvény. Korábbi feladatokban látott képleteket használva számolhatjuk a kérdezett mennyiségeket. (A számolások felét megspórolhatjuk, ha észrevesszük, hogy az eloszlás szimmetrikus a változókban.)

- $f_X(x) = \frac{6x+2}{5}$, ha $0 < x < 1$, egyéb x esetén pedig $f_X(x) = 0$.
  Hasonlóan (szimmetria miatt) $f_Y(y) = \frac{6y+2}{5}$, ha $0 < y < 1$, egyéb y esetén pedig $f_Y(y) = 0$.
  $\mathbb{E}(X) = \mathbb{E}(Y) = \frac{3}{5}$
- Jól látszódik, hogy $f(x,y) = f_X(x) \cdot f_Y(y)$ nem teljesül minden x, y esetén, így X és Y nem független.
- $f_{X|Y}(x|y) = \frac{2(x+xy+y)}{3y+1}$, míg $f_{Y|X}(y|x) = \frac{2(x+xy+y)}{3x+1}$ (mindkettő $0 < x < 1$, $0 < y < 1$ esetén).
  $\mathbb{E}(X|Y = y) = \frac{5y+2}{9y+3}$, $\mathbb{E}(Y|X = x) = \frac{5x+2}{9x+3}$
  $\mathbb{E}\left(X^2|Y = y\right) = \frac{7y+3}{18y+6}$, $\mathbb{E}\left(Y^2|X = x\right) = \frac{7x+3}{18x+6}$
  $\mathbb{D}^2(X|Y = y) = \frac{13y^2+8y+1}{18(3y+1)^2}$, $\mathbb{D}^2(Y|X = x) = \frac{13x^2+8x+1}{18(3x+1)^2}$
- $\mathbb{E}(XY) = \frac{16}{45}$, $\mathrm{COV}(X,Y) = \frac{-1}{225}$.

(b) Ellenőrizhető, hogy $\iint\limits_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$, így ez egy sűrűségfüggvény. Korábbi feladatokban látott képleteket használva számolhatjuk a kérdezett mennyiségeket. (A számolások felét megspórolhatjuk, ha észrevesszük, hogy az eloszlás szimmetrikus a változókban.)

- $f_X(x) = \lambda e^{-\lambda x}$, ha $0 < x$ (egyéb x esetén pedig $f_X(x) = 0$), vagyis $X \sim \mathrm{EXP}(\lambda)$. Hasonlóan $Y \sim \mathrm{EXP}(\lambda)$. Emiatt $\mathbb{E}(X) = \mathbb{E}(Y) = \frac{1}{\lambda}$.
- Jól látszódik, hogy $f(x,y) = f_X(x) \cdot f_Y(y)$ teljesül minden x, y esetén, így X és Y független.
- Függetlenség miatt a feltételes sűrűségfüggvények, várható értékek és szórásnégyzetek megegyeznek a peremeloszlás (EXP(λ)) megfelelő mennyiségeivel:

$$f_{X|Y}(x|y) = \lambda e^{-\lambda x}, \quad f_{Y|X}(y|x) = \lambda e^{-\lambda y} \qquad 0 < x, 0 < y,$$

$$\mathbb{E}(X|Y = y) = \mathbb{E}(Y|X = x) = \frac{1}{\lambda},$$

$$\mathbb{D}^2(X|Y = y) = \mathbb{D}^2(Y|X = x) = \frac{1}{\lambda^2}.$$

- $\mathbb{E}(XY) \overset{\text{függetlenség}}{=} \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{\lambda^2}$, $\mathrm{COV}(X,Y) = 0$.