Valószínűségszámítás előadás, 12. hét, 2020. december 2.

Együttes sűrűségfüggvények

1. Együttes sűrűségfüggvény

Emlékeztetőül: sok esetben nem egyetlen valószínűségi változó viselkedését vizsgáljuk, hanem több valószínűségi változó **együttes viselkedését**. Például:

- egy véletlen folyamat (tőzsdeindex, egy folyó vízállása, egy ország népessége) különböző időpontokban;
- egy ember (vagy ország, cég stb.) több különböző jellemzője (például egy ember életkora, jövedelme és kiadásai);
- egy méréssorozatban a különböző mérések során megfigyelt értékek (például egy mérést tízszer megismételve tíz különböző valószínűségi változót kapunk).

Valószínűségi változók együttesét **valószínűségi vektorváltozónak** nevezzük. Ez állhat **összefüggő** (mint az első két esetben) vagy **független** (mint tipikusan a harmadik esetben) valószínűségi változókból is.

1.1. Definíció. *Az*

$$\underline{X} = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$$

függvény ***valószínűségi vektorváltozó****, ha $X_1, X_2, \dots, X_n$ valószínűségi változók.*

Ha $\underline{X}$ valószínűségi vektorváltozó, akkor az X_i valószínűségi változó eloszlását az $\underline{X}$ i. ***peremeloszlásának*** *nevezzük.*

Az $\underline{X}$ valószínűségi vektorváltozó ***diszkrét****, ha értékkészlete véges vagy megszámlálhatóan végtelen.*

- 1000 embert megkérdezünk a havi jövedelméről. Legyen X_i az i. megkérdezett jövedelme. Ekkor $(X_1, X_2, \dots, X_{1000})$ valószínűségi vektorváltozó.
- $(X_1, X_2, \dots, X_{20})$, ha X_j a gdp értéke $2000+j$ évben (általában a hasonló, időben véletlenül változó folyamatokat *idősoroknak* nevezik).

A valószínűségi változók esetén a sűrűségfüggvény sok hasznos információt hordozott, ki tudtuk számítani események valószínűségét, a valószínűségi változó várható értékét, szórását például (azokban az esetekben, amikor létezik a valószínűségi változó sűrűségfüggvénye). Kérdés, hogy ezt a fogalmat hogyan terjeszthetjük ki a többváltozós esetre, vagyis valószínűségi vektorváltozókra. Ehhez először idézzük fel az együttes eloszlásfüggvény fogalmát.

1.2. Definíció. *Az $\underline{X} = (X_1, \dots, X_n)$ valószínűségi vektorváltozó* ***együttes eloszlásfüggvénye*** *az $F : \mathbb{R}^n \to [0,1]$ függvény, melyre*

$$F(\underline{t}) = F(t_1, \dots, t_n) = \mathbb{P}(X_1 \le t_1, X_2 \le t_2, \dots, X_n \le t_n),$$

ha $(t_1, \dots, t_n) \in \mathbb{R}^n$ valós számok.

Például:

- egy véletlenszerűen választott embert megkérdezünk a havi jövedelméről (X_1), a havi kiadásairól (X_2), és az életkoráról (X_3);
- ekkor (X_1, X_2, X_3) valószínűségi vektorváltozó, és
- ha eloszlásfüggvénye F, akkor például

  $$F(200000, 150000, 40) = \mathbb{P}(X_1 \le 200000, X_2 \le 1500000, X_3 \le 40)$$

  annak valószínűsége, hogy egy véletlenszerűen választott ember havi jövedelme legfeljebb 200000 (forint), havi kiadása legfeljebb 150000 (forint), életkora pedig legfeljebb 40 (év);

1.3. Definíció. *Az $\underline{X} = (X_1, \ldots, X_n)$ valószínűségi vektorváltozó* **abszolút folytonos**, *ha van olyan $f : \mathbb{R}^n \to \mathbb{R}$ függvény, melyre*

$$F(t_1, \ldots, t_n) = \int_{-\infty}^{t_n} \ldots \int_{-\infty}^{t_1} f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

teljesül minden $(t_1, \ldots, t_n) \in \mathbb{R}^n$ esetén. Ilyenkor az f függvényt az $(X_1, X_2, \ldots, X_n)$ **együttes sűrűségfüggvényének** *nevezzük.*

1.1. Állítás. *Tegyük fel, hogy az $(X_1, X_2, \ldots, X_n)$ valószínűségi vektorváltozó együttes sűrűségfüggvénye f. Ekkor minden (megfelelő) $A \subseteq \mathbb{R}^n$ halmazra*

$$\mathbb{P}((X_1, X_2, \ldots, X_n) \in A) = \int_A f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n.$$

Például, ha $n = 2$, és $A \subseteq \mathbb{R}^2$ egy megfelelő halmaz a síkon, és az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye f, akkor

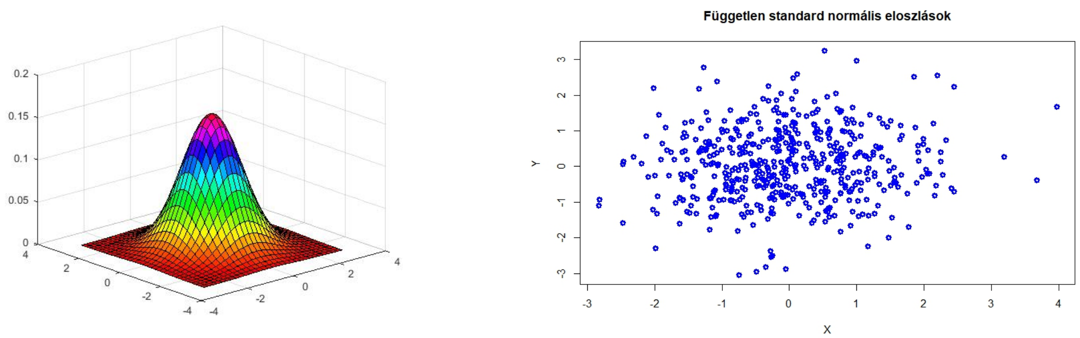
$$\mathbb{P}((X, Y) \in A) = \int_A f(x, y)\, dx\, dy.$$

Ennek következménye az alábbi állítás is, ami annak analógiája, hogy az egyváltozós esetben is a sűrűségfüggvény integrálja 1:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n = 1.$$

Megfordítva, minden olyan n-változós függvény lehet sűrűségfüggvény, ami nemnegatív, és ezt a tulajdonságot (az integrálja 1) teljesíti.

Az 1. ábrán azt figyelhetjük meg, hogy továbbra is igaz marad, hogy a sűrűségfüggvény nagy értékeinek a környékére nagyobb valószínűséggel esnek a megfigyelések, ha az adott sűrűségfüggvényű eloszlásból mintát veszünk. Itt X és Y függetlenek, normális eloszlásúak, így mindkét koordináta nagy valószínűséggel a 0-hoz közel esik, egymástól pedig függetlenek, így adódik a forgásszimmetrikus sűrűségfüggvény, és az együttes eloszlásból is hasonló viselkedés figyelhető meg.

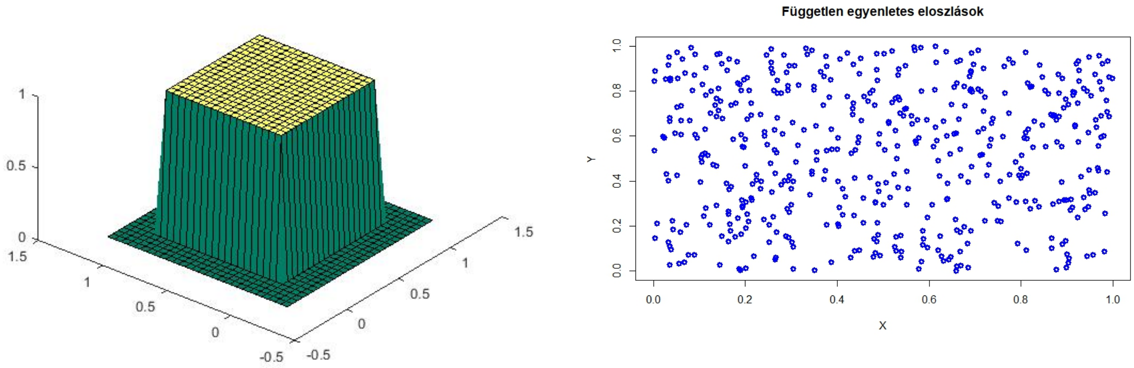


1. ábra. Az (X, Y) együttes sűrűségfüggvénye, illetve ehhez tartozó 500 elemű minta, ahol X és Y függetlenek, $N(0, 1)$ standard normális eloszlásúak

Nézzünk egy másik példát, legyenek X és Y függetlenek, a $[0, 1]$ intervallumon egyenletes eloszlású valószínűségi változók. A kapott (X, Y) párra úgy is gondolhatunk, mintha a $[0, 1] \times [0, 1]$ négyzetből egyenletes eloszlás szerint választanánk egy pontot, úgy, hogy egy $A \subseteq [0, 1] \times [0, 1]$ halmazba annyi valószínűséggel esik a pont, amennyi A területe (általában a négyzet területével osztanánk, de az most 1). Ez látható a 2. ábrán.

2. Peremeloszlások sűrűségfüggvénye

Tegyük fel, hogy az $(X_1, X_2, \ldots, X_n)$ valószínűségi vektorváltozó **együttes sűrűségfüggvénye** f. Hogyan kapható meg például az **első peremeloszlás, azaz X_1 sűrűségfüggvénye**?



2. ábra. (X, Y) együttes sűrűségfüggvénye, ahol X és Y függetlenek és a $[0,1]$ intervallumon egyenletes eloszlásúak, és 500 elemű minta ebből az eloszlásból

2.1. Állítás. *Tegyük fel, hogy az $\underline{X} = (X_1, \ldots, X_n)$ valószínűségi változók együttes sűrűségfüggvénye f. Ekkor az X_j valószínűségi változó sűrűségfüggvénye (melyet f_j-vel jelölünk), azaz a j. peremsűrűség-függvény így kapható meg f-ből:*

$$f_j(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_n) dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_n.$$

Speciálisan $n = 2$-re:

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y) dy; \qquad f_2(y) = \int_{-\infty}^{\infty} f(x,y) dx.$$

Ugyanis az $n = 2$ kétdimenziós esetben, ha az (X, Y) valószínűségi vektorváltozóról van szó, akkor tetszőleges t valós számra

$$\int_{-\infty}^{t} f_1(x)\, dx = \int_{-\infty}^{t} \int_{-\infty}^{\infty} f(x,y) dy dx = \mathbb{P}(X \leq t, Y \in \mathbb{R}) = \mathbb{P}(X \leq t),$$

felhasználva az 1.1. állítást (pontosabban annak $n = 2$-re vonatkozó speciális esetét). Itt az $A = (-\infty, t] \times \mathbb{R})$ halmazra alkalmaztuk az állítást, és észrevettük, hogy az $Y \in \mathbb{R}$ feltétel semmitmondó, mindenképpen teljesül. Ezzel pedig összességében azt ellenőriztük, hogy valóban f_1 lesz az X sűrűségfüggvénye.

Példa. Tegyük fel, hogy az (X, Y) valószínűségi vektorváltozó sűrűségfüggvénye

$$f(x,y) = \begin{cases} c(x+y), & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Határozzuk meg c értékét, az X sűrűségfüggvényét és várható értékét.

Tudjuk, hogy a sűrűségfüggvény integrálja 1:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dx\, dy = c \int_0^1 \int_0^1 x + y\, dx\, dy = c \int_0^1 \left[\frac{x^2}{2} + xy\right]_{x=0}^{1} dy = c \int_0^1 \frac{1}{2} + y\, dy = c\Big(\frac{1}{2} + \frac{1}{2}\Big).$$

Ez $c = 1$ esetén lesz 1, ez lesz a jó választás.

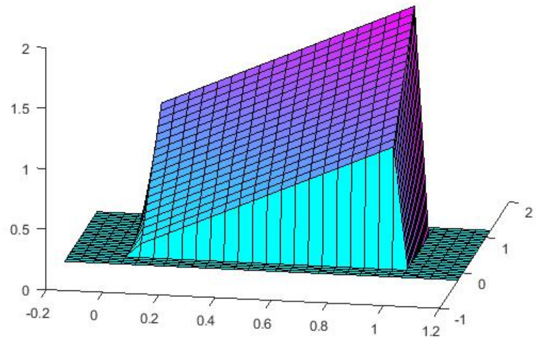
Az X sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y) dy = \int_0^1 x + y\, dy = x + \left[\frac{y^2}{2}\right]_{y=0}^{1} = x + \frac{1}{2},$$

ha $0 \leq x \leq 1$. Különben azonosan nullát integrálunk, a sűrűségfüggvény is 0. Ez másképpen azt jelenti, hogy az (X, Y) valószínűségi vektorváltozó a $[0,1] \times [0,1]$ négyzetből veszi fel az értékeit.

Ebből az X várható értéke a szokásos módon számolható:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^1 x^2 + \frac{x}{2}\, dx = \left[\frac{x^3}{3} + \frac{x^2}{4}\right]_{x=0}^{1} = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$



3. ábra. A $[0,1] \times [0,1]$ négyzeten $x+y$ alakú együttes sűrűségfüggvény

3. Függetlenség

Valószínűségi változók függetlenségét az alábbi módon definiáltuk. A kérdés most az lesz, hogy az együttes sűrűségfüggvényből hogyan olvasható le, ha a két valószínűségi változó független egymástól.

3.1. Definíció (Véges eset). *Azt mondjuk, hogy az $X_1, \dots, X_n : \Omega \to \mathbb{R}$ valószínűségi változók* **függetlenek***, ha*

$$\mathbb{P}(X_1 \le t_1, X_2 \le t_2, \dots, X_n \le t_n) = \mathbb{P}(X_1 \le t_1) \cdot \mathbb{P}(X_2 \le t_2) \dots \mathbb{P}(X_n \le t_n)$$

teljesül tetszőleges $t_1, t_2, \dots, t_n$ valós számokra, azaz

$$F(t_1, t_2, \dots, t_n) = F_1(t_1) \cdot F_2(t_2) \cdot \dots F_n(t_n),$$

ahol F_j az X_j valószínűségi változó eloszlásfüggvénye.

3.2. Definíció (Végtelen eset). *Az $X_1, X_2, X_3 \dots$ valószínűségi változók függetlenek, ha közülük bármely véges sokat kiválasztva független valószínűségi változókat kapunk.*

3.1. Állítás (Függetlenség és sűrűségfüggvény). *Legyen az $\underline{X} = (X_1, \dots, X_n)$ valószínűségi vektorváltozó együttes sűrűségfüggvénye f, továbbá az X_j valószínűségi változó sűrűségfüggvénye f_j minden $j = 1, 2, \dots, n$ esetén. Ezekkel a jelölésekkel: $X_1, X_2, \dots, X_n$ pontosan akkor függetlenek, ha*

$$f(t_1, \dots, t_n) = f_1(t_1) \cdot f_2(t_2) \dots f_n(t_n)$$

teljesül bármely $t_1, t_2, \dots, t_n$ valós számokra.

Például: ha X, Y független standard normális eloszlású valószínűségi változók, akkor együttes sűrűségfüggvényük

$$f(x,y) = f_1(x) f_2(y) = \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} \right) \right) \cdot \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2} \right) \right) = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2} \right).$$

Ez azt jelenti, hogy a sűrűségfüggvény csak az $x^2 + y^2$ mennyiségtől, azaz a pontnak az origótól vett távolságától függ, és minél nagyobb ez a távolság, annál kisebb a sűrűségfüggvény értéke. Ez megfelel az 1. ábrán látott forgásszimmetrikus, csökkenő függvénynek.

A 2. ábrához is független valószínűségi változók tartoztak. A 3. ábrán a két koordináta nem független, a sűrűségfüggvény nem szorzat alakú.

4. Kovariancia és korreláció kiszámítása

Két valószínűségi változó összefüggőségének mértékét a kovariancia és korrelációs együttható segítségével tudtuk mérni, pontosabban ez az egyik mérőszám, amit erre használnak. Kérdés, hogy az együttes sűrűségfüggvény (ha nem szorzat alakú), hogyan használható a kovariancia és korrelációs együttható kiszámítására. Ehhez az alábbi általános állítás lesz segítségünkre, melynek egyváltozós formájával már találkoztunk.

4.1. Állítás. *Legyen $(X_1, \dots, X_n)$ abszolút folytonos valószínűségi vektorváltozó, melynek sűrűségfüggvénye f. Legyen továbbá $g : \mathbb{R}^n \to R$ megfelelő tulajdonságokkal rendelkező függvény. Ekkor*

$$\mathbb{E}(g(X_1, \dots, X_n)) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, x_2, \dots, x_n) f(x_1, x_2, \dots, x_n)\, dx_1\, dx_2 \dots dx_n.$$

Speciálisan, ha $n = 2$ és (X, Y) valószínűségi vektorváltozó és f a sűrűségfüggvénye, akkor

$$\mathbb{E}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy.$$

Példa. Tegyük fel, hogy az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \le x \le 1 \text{ és } 0 \le y \le 1; \\ 0 & \text{különben.} \end{cases}$$

Számítsuk ki X és Y korrelációs együtthatóját:

$$R(X, Y) = \frac{\operatorname{cov}(X, Y)}{D(X)D(Y)} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{D(X)D(Y)}.$$

Először a szorzat várható értékét határozzuk meg. A fenti állítást $n = 2$-re és a $g(x, y) = xy$ függvényre alkalmazva:

$$\begin{aligned} \mathbb{E}(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f(x, y)\, dy\, dx = \int_0^1 \int_0^1 xy(x + y) dy dx = \\ &= \int_0^1 \int_0^1 x^2 y dy dx + \int_0^1 \int_0^1 xy^2 dy dx = \int_0^1 \frac{x^2}{2} dx + \int_0^1 \frac{x}{3} dx = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}, \end{aligned}$$

felhasználva, hogy $\int_0^1 x^k dx = [x^{k+1}/(k+1)]_{x=0}^1 = 1/(k+1)$.

Az X valószínűségi változó sűrűségfüggvényét és várható értékét már korábban meghatároztuk (a korábbi példában ugyanez az együttes sűrűségfüggvény szerepelt). A sűrűségfüggvényből:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_1(x) dx = \int_0^1 x^2 \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^3 + \frac{x^2}{2} dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$$

Minthogy a szimmetria miatt Y-nal hasonlóképpen számolhatunk, összességében azt kapjuk, hogy

$$\mathbb{E}(XY) = \frac{1}{3}; \qquad \mathbb{E}(X) = \mathbb{E}(Y) = \frac{7}{12}; \qquad \mathbb{E}(X^2) = \mathbb{E}(Y^2) = \frac{5}{12}.$$

Az X és Y korrelációs együtthatója:

$$\begin{aligned} R(X, Y) &= \frac{\operatorname{cov}(X, Y)}{D(X)D(Y)} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{D(X)D(Y)} = \\ &= \frac{1/3 - (7/12)^2}{\sqrt{5/12 - (7/12)^2} \cdot \sqrt{5/12 - (7/12)^2}} = \frac{1/3 - (7/12)^2}{5/12 - (7/12)^2} = -0,091. \end{aligned}$$

Ez nagyon gyenge negatív korrelációt jelent.

Házi feladat december 9., szerda, 8:00-ig Legyen az (X, Y) valószínűségi változó együttes sűrűségfüggvénye $c(e^{-x} + y^2)$, ha $x > 0$ és $0 \le y \le 2$, és 0 különben. Határozzuk meg c értékét és az $\mathbb{P}(X \le 3, Y \le 1)$ esemény valószínűségét.