Valószínűségszámítás gyakorlat

2020. december 1.

1. Egy szabályos kockával N-szer dobunk, ahol $N \sim \text{Poisson}(\lambda)$. Jelölje X_1 az egyesek, X_2 a kettesek számát. Adjuk meg az együttes eloszlást!

2. Legyen az (X, Y) véletlen vektorváltozó eloszlása egyenletes az egységkörben. Határozzuk meg az Y feltételes sűrűségfüggvényét az $X = x$ feltétel mellett! Számítsuk ki az $\mathbf{E}[Y^2|X = x]$ feltételes várható értéket.

Megoldás. Jelölje $B = \{(x, y) : x^2 + y^2 \leq 1\}$ a zárt egységkört. Mivel (X, Y) egyenletes eloszlású, ezért az együttes h sűrűségfüggvény

$$h(x,y) = \begin{cases} \frac{1}{\pi}, & (x,y) \in B, \\ 0, & \text{különben.} \end{cases}$$

Innen a peremeloszlások sűrűségei

$$f(x) = \int_{-\infty}^{\infty} h(x,y)\mathrm{d}y = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2}, & x \in [-1,1], \\ 0, & \text{különben,} \end{cases}$$

és a szimmetria miatt $f(x) = g(x)$. A feltételes sűrűségfüggvény

$$f_{Y|X}(y|x) = \frac{h(x,y)}{f(x)} = \frac{1}{2\sqrt{1-x^2}}, \quad y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}],$$

azaz feltéve, hogy $X = x$ az Y eloszlása egyenletes a $[-\sqrt{1-x^2}, \sqrt{1-x^2}]$ intervallumon. Nyilván

$$\mathbf{E}[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\mathrm{d}y = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} y \frac{1}{2\sqrt{1-x^2}}\mathrm{d}y = 0,$$

és

$$\mathbf{E}[Y^2|X = x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x)\mathrm{d}y = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} y^2 \frac{1}{2\sqrt{1-x^2}}\mathrm{d}y = \frac{1-x^2}{3}.$$

$\square$

3. Legyenek X, Y, Z független exponenciális eloszlású véletlen változók, λ, μ, ν paraméterekkel. Határozzuk meg a $\mathbf{P}(X > Y)$, $\mathbf{P}(X > Y > Z)$ valószínűségeket.

Megoldás. Jelölje f az Y sűrűségfüggvényét. A teljes valószínűség tétele szerint

$$\begin{aligned}\mathbf{P}(X > Y) &= \int_0^\infty \mathbf{P}(X > y | Y = y) f(y) \mathrm{d}y \\ &= \int_0^\infty e^{-\lambda y} \mu e^{-\mu y} \mathrm{d}y \\ &= \frac{\mu}{\lambda + \mu} \int_0^\infty (\lambda + \mu) e^{-(\lambda+\mu)y} \mathrm{d}y \\ &= \frac{\mu}{\lambda + \mu}.\end{aligned}$$

Ezt már meghatároztuk korábban, kicsit más számolással.

Hasonlóan

$$\begin{aligned}\mathbf{P}(X > Y > Z) &= \int_0^\infty \mathbf{P}(X > y > Z | Y = y) f(y) \mathrm{d}y \\ &= \int_0^\infty e^{-\lambda y} (1 - e^{-\nu y}) \mu e^{-\mu y} \mathrm{d}y \\ &= \frac{\mu}{\lambda + \mu} \int_0^\infty (\lambda + \mu) e^{-(\lambda+\mu)y} \mathrm{d}y \\ &\quad - \frac{\mu}{\lambda + \mu + \nu} \int_0^\infty (\lambda + \mu + \nu) e^{-(\lambda+\mu+\nu)y} \mathrm{d}y \\ &= \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \nu + \mu}.\end{aligned}$$

□

4. Legyenek X, Y független 1 paraméterű exponenciális eloszlású véletlen változók. Legyen $U = X \wedge Y$, $V = X \vee Y$. Határozzuk meg a maximum minimumra vett feltételes sűrűségét, és fordítva, azaz adjuk meg a $g_{U|V}(u|v)$, $g_{V|U}(v|u)$ feltételes sűrűségeket! Ismerjünk rá a kapott eloszlásokra!

5. Legyenek X, Y független azonos eloszlású véletlen változók, f sűrűségfüggvénnyel. Legyen $U = X \wedge Y$, $V = X \vee Y$. Határozzuk meg a maximum minimumra vett feltételes sűrűségét, és fordítva, azaz adjuk meg a $g_{U|V}(u|v)$, $g_{V|U}(v|u)$ feltételes sűrűségeket!

Megoldás. Mivel $U \leq V$, így az $f_{U,V}(u, v)$ együttes sűrűség 0, ha $v < u$. Ha $v \geq u$, akkor

$$\begin{aligned}\mathbf{P}(U \leq u, V \leq v) &= 2\mathbf{P}(X \leq u, Y \in (u, v]) + \mathbf{P}(X \leq u, Y \leq u) \\ &= 2F(u)(F(v) - F(u)) + F(u)^2 \\ &= F(u)(2F(v) - F(u)),\end{aligned}$$

ahol F a közös eloszlásfüggvény. Így

$$f_{U,V}(u,v) = \frac{\partial^2}{\partial u \partial v}\mathbf{P}(U \le u, V \le v) = 2f(u)f(v),\ u \le v.$$

A marginális sűrűséget innen integrálással, vagy direkt számolással meghatározhatjuk. Kapjuk, hogy

$$f_U(u) = \int_{-\infty}^{\infty} f(u,v)\mathrm{d}v = \int_u^{\infty} 2f(u)f(v)\mathrm{d}v = 2f(u)[1-F(u)],$$

és

$$f_V(v) = \int_{-\infty}^{\infty} f(u,v)\mathrm{d}u = \int_{-\infty}^{v} 2f(u)f(v)\mathrm{d}v = 2f(v)F(v).$$

A feltételes sűrűségek

$$g_{V|U}(v|u) = \frac{f_{U,V}(u,v)}{f_U(u)} = \frac{f(v)}{1-F(u)},\ v \ge u,$$
$$g_{U|V}(u|v) = \frac{f_{U,V}(u,v)}{f_V(v)} = \frac{f(u)}{F(v)},\ u \le v.$$

Vegyük észre, hogy a $g_{U|V}(\cdot|v)$ sűrűség egy olyan F eloszlású változó sűrűsége, amiről feltesszük, hogy v-nél kisebb, a $g_{V|U}(\cdot|u)$ pedig egy olyan F eloszlású változó sűrűsége, amiről feltesszük, hogy u-nál nagyobb. Hát persze, pontosan ezt kellett kapjuk. $\square$

6. Egyenletes eloszlás szerint választok egy p értéket a $[0,1]$ intervallumon, majd gyártok egy olyan érmét, mely p valószínűséggel ad fejet. Jelölje X annak a dobásnak a sorszámát, mikor először dobok fejet. Adjuk meg X eloszlását és várható értékét. Ugyanez lesz a várható érték, ha egy olyan érmét dobálok, mely $\mathbf{E}(p)$ valószínűséggel ad fejet?

Megoldás. Nyilván X lehetséges értékei $1, 2, \ldots$, így

$$\begin{aligned}\mathbf{P}(X=k) &= \int_0^1 \mathbf{P}(X=k|p=x)f(x)\mathrm{d}x \\ &= \int_0^1 (1-x)^{k-1}x\mathrm{d}x = \int_0^1 x^{k-1}(1-x)\mathrm{d}x \\ &= \frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}.\end{aligned}$$

A várható érték

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} k\mathbf{P}(X=k) = \sum_{k=1}^{\infty} \frac{1}{k+1} = \infty.$$

$\square$

7. Legyen $G = (V, E)$ egy egyszerű gráf, és $v_1, v_2, \ldots, v_n$ a csúcsok egy sorrendje. Minden csúcsra feldobunk egy szabályos érmét, ha az fej akkor a csúcs az A halmazba, különben a B halmazba kerül. Határozzuk meg az A és B halmazok közt futó élek számának várható értékét!

Megoldás. Legyenek a gráf élei $e_1, \ldots, e_k$. Ekkor az A és B között futó élek száma

$$X = \sum_{i=1}^{k} I_i,$$

ahol $I_i = 1$, ha e_i él, 0, különben. Nyilván $I_i = 1$ pontosan akkor, ha az e_i él két végpontja különböző halmazokba esik. Tehát

$$\mathbf{E}I_i = \mathbf{P}(I_i = 1) = \frac{1}{2}.$$

Innen

$$\mathbf{E}X = k\mathbf{E}I_1 = \frac{k}{2}.$$

□

8. Mutassuk meg, hogy tetszőleges hurokélmentes gráfból páros gráfot kaphatunk legfeljebb az élek felének elhagyásával!

9. A Texpo áruházakban az i-edik kasszánál egy vásárló percben számolva i paraméterű exponenciális időt tölt el. A kiszolgálási idők az egyes kasszáknál egymástól függetlenek.

- András éppen üresen találja az 1-es kasszát. Mennyi a valószínűsége, hogy 2 percen belül végez?
- Andrással pontosan egyidőben Béla beáll az ugyancsak üres 2-es kasszához. Mennyi a valószínűsége, hogy mindketten 2 percen belül végeznek? Mennyi a valószínűsége, hogy Béla 2 percen belül végez, de András nem?
- Mennyi a valószínűsége, hogy valamelyikük 2 percen belül végez? Határozzuk meg a hamarabb végző kiszolgálási idejének eloszlásást! Tehát András és Béla kiszolgálási idejének a minimumára vagyunk kíváncsiak.
- Mennyi a valószínűsége, hogy András végez hamarabb?

Megoldás. (a) Legyen X András kiszolgálási ideje. Ekkor X exponenciális eloszlású 1 paraméterrel, ezért az a valószínűség, hogy András 2 percen belül végez

$$\mathbf{P}(X \leq 2) = 1 - e^{-1 \cdot 2} = 1 - e^{-2} = 0,86.$$

(b) Legyen Y Béla kiszolgálási ideje. Ez exponenciális eloszlású 2 paraméterrel, és független X-től. Ezért az a valószínűség, hogy mindketten 2 percen belül végeznek

$$\mathbf{P}(X \leq 2, Y \leq 2) = \mathbf{P}(X \leq 2) \cdot \mathbf{P}(Y \leq 2) = (1 - e^{-2}) \cdot (1 - e^{-4}) = 0,84.$$

Ha Béla 2 percen belül végez, András pedig nem, akkor $Y \leq 2$ és $X > 2$, aminek a valószínűsége

$$\mathbf{P}(X > 2, Y \leq 2) = \mathbf{P}(X > 2) \cdot \mathbf{P}(Y \leq 2) = e^{-2} \cdot (1 - e^{-4}) = 0,13.$$

(c) Jelölje Z az X és Y közül a kisebbet, azaz $Z = \min\{X, Y\}$. Ekkor Z nemnegatív értékeket vesz fel. Valamely $z > 0$ esetén

$$\begin{aligned}\mathbf{P}(Z \leq z) &= 1 - \mathbf{P}(Z > z) = 1 - \mathbf{P}(X > z, Y > z)\\ &= 1 - \mathbf{P}(X > z) \cdot \mathbf{P}(Y > z) = 1 - e^{-z} \cdot e^{-2z}\\ &= 1 - e^{-3z}.\end{aligned}$$

Ez pontosan azt jelenti, hogy Z exponenciális eloszlású 3 paraméterrel.

Lényegében beláttuk (1 és 2 helyett λ és μ-t írva), hogy független exponenciálisok minimuma exponenciális, és a paraméter a paraméterek összege. Ez az egyszerű tény fontos lesz Sztochasztikus modellek kurzuson.

(d) Mivel X és Y függetlenek, ezért együttes sűrűségfüggvényük az egyes sűrűségfüggvények szorzata, azaz

$$h(u, v) = f_X(u) f_Y(v) = \begin{cases} e^{-u} 2e^{-2v}, & \text{ha } u, v > 0,\\ 0, & \text{különben.}\end{cases}$$

Az, hogy András végez hamarabb pontosan azt jelenti, hogy $X < Y$, azaz $(X, Y) \in \{(x, y) : x < y\} = A$. Ennek a valószínűsége

$$\begin{aligned}\mathbf{P}(X < Y) &= \mathbf{P}((X, Y) \in A) = \iint_A h(u, v)\mathrm{d}u\mathrm{d}v\\ &= \int_0^\infty \left(\int_u^\infty 2e^{-u} e^{-2v} \mathrm{d}v \right) \mathrm{d}u\\ &= \int_0^\infty e^{-u} \left[-e^{-2v} \right]_{v=u}^{v=\infty} \mathrm{d}u\\ &= \int_0^\infty e^{-u} e^{-2u} \mathrm{d}u = \frac{1}{3}.\end{aligned}$$

□

10. Válasszunk két számot egymástól függetlenül az egyenletességi hipotézis szerint a $(0,1)$ intervallumból! Adjuk meg a maximum és a minimum együttes eloszlását, és számoljuk ki a kovarianciájukat!

Megoldás. Legyen U_1 és U_2 a két, függetlenül egyenletes eloszlás szerint választott pont, és jelölje X a maximumot, Y pedig a minimumot. Először meghatározzuk X és Y eloszlását külön-külön.

Világos, hogy X lehetséges értékei a $[0,1]$ intervallum elemei, és $x \in [0,1]$ esetén $\{X \leq x\}$ pontosan akkor teljesül, ha mindkét pont a $[0,x]$ intervallumba esik, aminek valószínűsége x^2. Tehát

$$F(x) = \mathbf{P}(X \leq x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ x^2, & \text{ha } x \in [0,1], \\ 1, & \text{ha } x \geq 1. \end{cases}$$

Innen deriválással kapjuk a sűrűségfüggvényt, azaz

$$f(x) = \begin{cases} 2x, & \text{ha } x \in (0,1), \\ 0, & \text{különben.} \end{cases}$$

A várható érték

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = \int_0^1 x \cdot 2x \mathrm{d}x = \frac{2}{3}.$$

Hasonlóan, az Y lehetséges értékei a $[0,1]$ intervallum elemei, és $\{Y \leq y\}$ pontosan akkor teljesül, ha legalább 1 pont a $[0,y]$ intervallumba esik, azaz vagy pontosan 1 esik oda, aminek a valószínűsége

$$2 \cdot y \cdot (1-y),$$

melyik pont esik oda, az a $[0,y]$-ba esik, a másik pedig nem; vagy pontosan 2 pont esik oda, aminek a valószínűsége

$$y^2,$$

hiszen ekkor mindkettő a $[0,y]$-ba esik. Tehát az eloszlásfüggvény

$$G(y) = \mathbf{P}(Y \leq y) = \begin{cases} 0, & \text{ha } y \leq 0, \\ 2y(1-y) + y^2 = 2y - y^2, & \text{ha } y \in [0,1], \\ 1, & \text{ha } y \geq 1. \end{cases}$$

Innen a sűrűségfüggvény

$$g(y) = \begin{cases} 2-2y, & \text{ha } y \in (0,1), \\ 0, & \text{különben.} \end{cases}$$

A várható érték

$$\mathbf{E}(Y) = \int_0^1 y(2-2y)\mathrm{d}y = \frac{1}{3}.$$

Ez heurisztikusan világos, hiszen 2 véletlen pont nagyjából 3 egyforma hosszú intervallumra osztja az egységnyi intervallumot. Tehát a kisebbik pont várható értéke 1/3, a nagyobbiké pedig 2/3, ahogy kiszámoltuk.

Nézzük most az együttes eloszlást. Legyen

$$\mathbf{P}(X \le x, Y \le y) = H(x,y)$$

az együttes eloszlásfüggvény. Mivel a lehetséges értékek a halmaza $[0,1] \times [0,1]$, ezért

$$H(x,y) = \begin{cases} 1, & \text{ha } x \ge 1, y \ge 1, \\ \mathbf{P}(X \le x), & \text{ha } y \ge 1, \\ \mathbf{P}(Y \le y), & \text{ha } x \ge 1, \\ 0, & \text{ha } x \le 0 \text{ vagy } y \le 0. \end{cases}$$

Így az érdekes eset amikor $x, y \in [0,1]$. Sőt, mivel $X \ge Y$, így ha $x \le y$ akkor

$$\mathbf{P}(X \le x, Y \le y) = \mathbf{P}(X \le x).$$

Hát persze, ha a maximum kisebb, mint x, akkor a minimum is kisebb, így az a feltétel elhagyható. Legyen tehát $1 \ge x > y \ge 0$. Ekkor az $\{X \le x, Y \le y\}$ esemény akkor következik be, ha mindkét pont $\le x$, és legalább egy pont $\le y$. Ez úgy lehet, ha vagy mindkét pont a $[0,y]$-ba esik, aminek a valószínűsége y^2, vagy az egyik a $[0,y]$-ba, a másik $(y,x]$-be, aminek a valószínűsége $2 \cdot y \cdot (x-y)$ (melyik pont, hova, hova). Tehát

$$\mathbf{P}(X \le x, Y \le y) = y^2 + 2y(x-y) = 2xy - y^2.$$

Összegezve,

$$H(x,y) = \mathbf{P}(X \le x, Y \le y) = \begin{cases} 1, & \text{ha } x \ge 1, y \ge 1, \\ \mathbf{P}(X \le x), & \text{ha } y \ge 1, \\ \mathbf{P}(Y \le y), & \text{ha } x \ge 1, \\ \mathbf{P}(X \le x), & \text{ha } x \le y, \\ 2xy - y^2, & \text{ha } 0 \le y \le x \le 1, \\ 0, & \text{ha } x \le 0 \text{ vagy } y \le 0. \end{cases}$$

Látjuk, hogy kétváltozós eloszlásfüggvény meghatározása nehezebb. A sűrűségfüggvény többváltozós esetben is az eloszlásfüggvény deriváltja, csak most minden változó szerint egyszer kell deriválni. Először x szerint deriválunk, az y-t konstansnak tekintve, majd az eredmény deriváljuk y szerint. Az eloszlásfüggvény alakjából látjuk, hogy csak ott nem 0 a sűrűség, ahol x-től és y-tól is függ az eredmény, azaz

$$h(x,y) = \frac{\partial^2}{\partial x \partial y} = \begin{cases} 2, & \text{ha } 0 \leq y \leq x \leq 1, \\ 0, & \text{különben.} \end{cases}$$

Persze a sűrűség éppen azon a tartomány nem tűnik el, ami az (X,Y) vektor lehetséges értékeit adják, ami most $\{(x,y) : 0 \leq y \leq x \leq 1\}$. Ezek után kiszámolhatjuk a kovarianciát. A formula szerint

$$\mathbf{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y),$$

és a várható értékeket már ismerjük. A szorzat várható értéke, a várható érték tulajdonságainál megismert formula szerint (szorzat, mint kétváltozós függvény)

$$\mathbf{E}(XY) = \int\int h(x,y)xy\mathrm{d}x\mathrm{d}y = \int_0^1 \left(\int_0^x 2xy\mathrm{d}y \right) \mathrm{d}x = \int_0^1 x \cdot x^2 \mathrm{d}x = \frac{1}{4}.$$

Az előbbi integrálásnál figyeljünk a határokra: rögzített $x \in [0,1]$ esetén $y \in [0,x]$-re lesz $h = 2$! Tehát

$$\mathbf{Cov}(X,Y) = \frac{1}{4} - \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{36}.$$

$\square$