VJEROJATNOST I STATISTIKA

Dodatni zadaci za 7. i 8. tjedan (14., 15. i 16. predavanje)

Primjeri neprekinutih razdioba

Eksponencijalna razdioba

1. Vrijeme do pojave prvog e-maila na serveru je eksponencijalna slučajna varijabla s očekivanjem 4 sekunde. Ako se nijedan e-mail nije pojavio u prve 4 sekunde, izračunajte vjerojatnost da se neće pojaviti ni u sljedeće 4 sekunde.

Rješenje:

Neka je X slučajna varijabla:

$$X \; = \; \text{vrijeme do pojave e-maila (u sekundama).}$$

Tada znamo da vrijedi $P(X > b \mid X > a) = P(X > b - a)$ za $b > a > 0$. Naime,

$$P(X > b \mid X > a) = \frac{P(X > b, X > a)}{P(X > a)} = \frac{P(X > b)}{P(X > a)} = \frac{e^{-\lambda b}}{e^{-\lambda a}} = P(X > b - a).$$

Iz teksta vidimo da je $E(X) = 4$, a budući da je očekivanje eksponencijalne slučajne varijable s parametrom λ jednako $\frac{1}{\lambda}$, slijedi $\lambda = \frac{1}{4}$, odnosno $X \sim \mathcal{E}(\frac{1}{4})$. Računamo vjerojatnost

$$P(X > 8 \mid X > 4) = P(X > 4) = e^{-4 \cdot \frac{1}{4}} = e^{-1}.$$

2. Vrijeme ispravnog rada računala je slučajna varijabla s eksponencijalnom razdiobom. Vjerojatnost ispravnog rada računala tijekom jedne godine je jednaka 0.9.

(a) Kolika je vjerojatnost da će to računalo raditi ispravno tijekom 2 godine?

(b) Kolika je vjerojatnost da će od 15 takvih računala u računarskom praktikumu njih barem 13 raditi ispravno tijekom 2 godine?

Rješenje: Označimo s X slučajnu varijablu

$$X \; = \; \text{vrijeme ispravnog rada računala (u godinama).}$$

Zadana je vjerojatnost $P(X > 1) = 0.9$ odakle slijedi $e^{-\lambda} = 0.9$.

(a) Vrijedi

$$P(X > 2) = 1 - P(X < 2) = 1 - F(2) = 1 - (1 - e^{-2\lambda}) = e^{-2\lambda} = (e^{-\lambda})^2 = 0.81.$$

(b) Neka je

$$Y \; = \; \text{broj računala koji će raditi ispravno tijekom 2 godine.}$$

Znamo da se Y ravna po binomnoj razdiobi gdje je $n = 15$ i $p = 0.81$. Odnosno vrijedi

$$Y \sim \mathcal{B}(15, 0.81).$$

Stoga je konačno rješenje:

$$P(Y \geq 13) = \sum_{k=13}^{15} \binom{n}{k} \cdot p^k (1-p)^{n-k} = 0.436.$$

3. Slučajna varijabla X ima eksponencijalnu razdiobu s očekivanjem 2. Odredite očekivanje slučajne varijable $Y = |2 - X|$.

Rješenje:

Iz teksta zadatka slijedi: $E(X) = 2$, odnosno $\lambda = \frac{1}{2}$. Radi se o eksponencijalnoj razdiobi $X \sim \mathcal{E}(\frac{1}{2})$. Za eskponencijalnu razdiobu znamo funkciju gustoće:

$$f(x) = \lambda e^{-\lambda x},\ x > 0.$$

Slučajna varijabla $Y = |2 - X|$ poprima vrijednosti u skupu $[0, \infty\rangle$. Označimo

$$Y = \psi(X) = |2 - X|\,.$$

Traži se da izračunamo očekivanje

$$E(Y) = \int_{-\infty}^{\infty} y g(y)\,\mathrm{d}y, \tag{1}$$

a da bismo to napravili prvo moramo odrediti funkciju gustoće $g(y)$ slučajne varijable Y. Pokazat ćemo dva moguća načina dobivanja funkcije g.

1. način. Koristimo formulu

$$g(y) = f(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|,\ y = \psi(x), \tag{2}$$

gdje su f i g funkcije gustoće slučajnih varijabli X i Y te ψ neka injektivna funkcija.

Konkretno u našem slučaju moramo napraviti restrikciju funkcije

$$y = \psi(x) = |2 - x|$$

na dva intervala kako bismo došli do injektivnih funkcija. Jasno je da se radi o intervalima $\langle -\infty, 2\rangle$ i $\langle 2, +\infty\rangle$.

- $x \in \langle -\infty, 2\rangle$ $(y \in \langle 0, +\infty\rangle)$
  Tada vrijedi $y = \psi(x) = |2 - x| = 2 - x \implies x = 2 - y,\quad \frac{\mathrm{d}x}{\mathrm{d}y} = -1$. Iz formule (2) slijedi:

$$g_1(y) = f(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = f(2-y) = \begin{cases} \lambda e^{-\lambda(2-y)}, & 2 - y > 0 \\ 0, & 2 - y < 0 \end{cases} = \begin{cases} \frac{1}{2} e^{\frac{1}{2}y - 1}, & 0 < y < 2 \\ 0, & y > 2 \end{cases}.$$

- $x \in \langle 2, +\infty\rangle$ $(y \in \langle 0, +\infty\rangle)$
  Tada vrijedi $y = \psi(x) = |2 - x| = x - 2 \implies x = 2 + y,\quad \frac{\mathrm{d}x}{\mathrm{d}y} = 1$. Iz formule (2) slijedi:

$$g_2(y) = f(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = f(2+y) = \begin{cases} \lambda e^{-\lambda(2+y)}, & 2 + y > 0 \\ 0, & 2 + y < 0 \end{cases} = \frac{1}{2} e^{-\frac{1}{2}y - 1},\ y > 0.$$

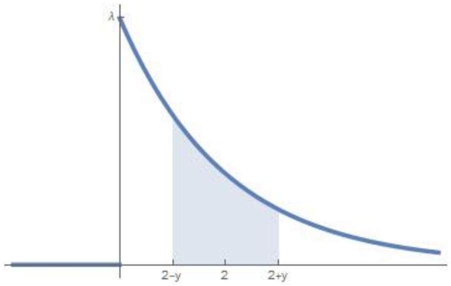
U konačnici dobivamo traženu funkciju gustoće kao zbroj funkcija g_1 i g_2:

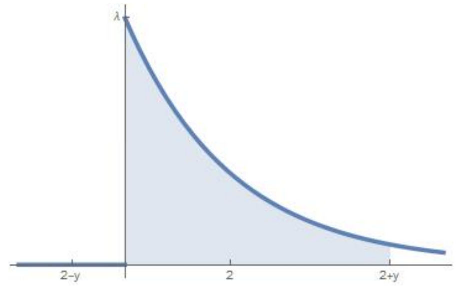
$$g(y) = g_1(y) + g_2(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{2} e^{-1}\left[e^{\frac{y}{2}} + e^{-\frac{y}{2}}\right], & 0 < y < 2, \\ \frac{1}{2} e^{-\frac{1}{2}y - 1}, & y > 2. \end{cases}$$

2. način. Računamo funkciju razdiobe G slučajne varijable Y:

$$G(y) = P(Y < y) = P(|2-X| < y) = P(-y < X-2 < y) = P(2-y < X < 2+y) = \int_{2-y}^{2+y} f(x)\,\mathrm{d}x,$$

gdje je $y \in [0, +\infty\rangle$. Razlikujemo dva slučaja.

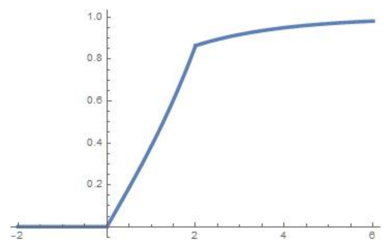




- Za $2 - y \geq 0$ vrijedi $G(y) = \int_{2-y}^{2+y} \lambda e^{-\lambda x}\, \mathrm{d}x = -e^{-\lambda x}\Big|_{x=2-y}^{2+y} = e^{-\lambda(2-y)} - e^{-\lambda(2+y)}$.
- Za $2 - y < 0$ vrijedi $G(y) = \int_{0}^{2+y} \lambda e^{-\lambda x}\, \mathrm{d}x = -e^{-\lambda x}\Big|_{x=0}^{2+y} = 1 - e^{-\lambda(2+y)}$.

Dobili smo

$$G(y) = \begin{cases} 0, & y \in \langle -\infty, 0], \\ e^{\frac{1}{2}y-1} - e^{-\frac{1}{2}y-1}, & y \in \langle 0, 2], \\ 1 - e^{-\frac{1}{2}y-1}, & y \in \langle 2, +\infty\rangle. \end{cases}$$

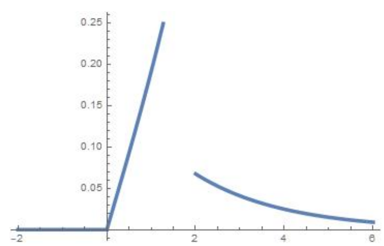


Za funkciju gustoće g slučajne varijable Y u točkama neprekinutosti vrijedi jednakost

$$g(y) = \frac{dG(y)}{\mathrm{d}y}.$$

Zbog toga je

$$g(y) = \begin{cases} 0, & y \in \langle -\infty, 0\rangle, \\ \frac{1}{2}e^{-1}\left[e^{\frac{y}{2}} - e^{-\frac{y}{2}}\right], & y \in \langle 0, 2\rangle, \\ \frac{1}{2}e^{-\frac{1}{2}y-1}, & y \in \langle 2, +\infty\rangle. \end{cases}$$



Naposljetku računamo očekivanje po formuli (1):

$$E(Y) = \int_0^2 yg(y)\, \mathrm{d}y + \int_2^\infty yg(y)\, \mathrm{d}y = \int_0^2 \frac{1}{2}ye^{-1}\left[e^{\frac{y}{2}} + e^{-\frac{y}{2}}\right] \mathrm{d}y + \int_2^\infty \frac{1}{2}ye^{-\frac{1}{2}y+1}\, \mathrm{d}y = \frac{4}{e}.$$

Normalna razdioba

4. Vrijeme koje student provede na putu od kuće do fakulteta je slučajna varijabla s normalnom razdiobom $\mathcal{N}(a, \sigma^2)$ s očekivanjem $a = 50$ minuta. Student ima predavanje u 8:15 i kreće iz kuće u 7:20. Ako je vjerojatnost da će stići na fakultet u vremenskom intervalu od 8:05 do 8:15 jednaka 0.383, kolika je vjerojatnost da će kasniti na predavanje više od 5 min?

Rješenje: Neka je

$$X \;=\; \text{vrijeme koje student provede na putu.}$$

Tada X ima normalnu razdiobu s očekivanjem $a = 50$, tj. $X \sim \mathcal{N}(50, \sigma^2)$. Ako će student stići u intervalu od 8:05 do 8:15 znači da će na putu provesti od 45 do 55 min. Prvi korak je centriranje i normiranje slučajne varijable X kako bismo dobili slučajnu varijablu s jediničnom normalnom razdiobom. Vrijedi:

$$X^* = \frac{X-50}{\sigma} \sim \mathcal{N}(0,1).$$

Slijedi:

$$P(45 < X < 55) = P\left(\frac{45-50}{\sigma} < \frac{X-50}{\sigma} < \frac{55-50}{\sigma}\right) = P\left(-\frac{5}{\sigma} < \frac{X-50}{\sigma} < \frac{5}{\sigma}\right) = 0.383.$$

Za funkciju razdiobe Φ jedinične normalne slučajne varijable vrijedi

$$\Phi(u) = \frac{1}{2}\left[1 + \Phi^*(u)\right],$$

gdje je

$$\Phi^*(u) = \frac{1}{\sqrt{2\pi}}\int_{-u}^{u} e^{-\frac{1}{2}t^2}dt = \int_{-u}^{u}\phi(t)dt = P(-u < X^* < u).$$

Čitamo odgovarajuću vrijednost iz tablice:

$$\Phi^*\left(\frac{5}{\sigma}\right) = 0.383 \implies \left(\frac{5}{\sigma}\right) \approx 0.50 \implies \sigma \approx 10.$$

Konačno imamo

$$P(X > 60) = P\left(\frac{X-a}{\sigma} > \frac{60-a}{\sigma}\right) = P\left(\frac{X-50}{10} > \frac{10}{10}\right) = P\left(\frac{X-50}{10} > 1\right) = \frac{1}{2}\left[1 - \Phi^*(1)\right]$$
$$= 0.159.$$

5. Visina čovjeka je slučajna varijabla s normalnom razdiobom $\mathcal{N}(a, \sigma^2)$ gdje je očekivanje $a =$ 174cm. Ako 68.269% ljudi ima visinu između 165cm i 183cm, izračunajte vjerojatnost da je čovjek viši od 170cm.

Rješenje:
Označimo:

$$X \;=\; \text{visina čovjeka},\; X \sim \mathcal{N}(174, \sigma^2).$$

Centriramo i normiramo slučajnu varijablu X kako bismo dobili slučajnu varijablu s jediničnom normalnom razdiobom. Vrijedi:

$$\frac{X-174}{\sigma} \sim \mathcal{N}(0,1).$$

Na skoro identičan način kao u prethodnom zadatku računamo:

$$P(165 < X < 183) = 0.6829 \implies \left(\frac{165-174}{\sigma} < \frac{X-174}{\sigma} < \frac{183-174}{\sigma}\right) = 0.6829.$$

Slijedi da je:

$$\Phi^*\left(\frac{9}{\sigma}\right) = 0.6829 \implies \sigma \approx 9.$$

Konačno:

$$P(X > 170) = P\left(\frac{X-a}{9} > \frac{170-a}{9}\right) = P\left(\frac{X-174}{9} > \frac{170-174}{9}\right) = P\left(\frac{X-174}{9} > -\frac{4}{9}\right)$$
$$= \frac{1}{2}\left[1 - \Phi^*\left(-\frac{4}{9}\right)\right] = 0.67148.$$

6. Masa domaćih jabuka podvrgava se normalnoj razdiobi s očekivanjem 150g i standardnom devijacijom 20g, dok se masa industrijski proizvedenih jabuka podvrgava normalnoj razdiobi s očekivanjem 220g i standardnom devijacijom 5g. Jabučar Jan prodaje jabuke na Dolcu isključivo u paketima od po 4 jabuke i to dvije domaće i dvije industrijski proizvedene jabuke. Odredite vjerojatnost da je masa Janovog paketa između 820g i 1000g.

Rješenje:

Označimo:

$$\begin{aligned} X &= \text{ masa domaće jabuke, } X \sim \mathcal{N}\left(180, 20^2\right), \\ Y &= \text{ masa industrijski proizvedene jabuke, } Y \sim \mathcal{N}\left(220, 5^2\right). \end{aligned}$$

Što se od nas točno traži? Vjerojatnost da je masa Janovog paketa od 2 domaće i 2 industrijski proizvedene jabuke u intervalu $[820, 1000]$.

Prisjetimo se svojstva stabilnosti normalne razdiobe. Za nezavisne slučajne varijable $X_1 \sim \mathcal{N}(a_1, \sigma_1^2)$ i $X_2 \sim \mathcal{N}(a_2, \sigma_2^2)$ te bilo koje realne brojeve s_1 i s_2 vrijedi

$$s_1X_1 + s_2X_2 \sim \mathcal{N}(s_1a_1 + s_2a_, s_1^2\sigma_1^2 + s_2^2\sigma_2^2). \tag{3}$$

Prvo što bi nam moglo pasti na pamet je da masu Janovog paketa opišemo slučajnom varijablom $2X + 2Y$ koja prema svojstvu (3) ima normalnu razdiobu s očekivanjem $2 \cdot 180 + 2 \cdot 220 = 800$ i disperzijom $2^2 \cdot 400 + 2^2 \cdot 25 = 1700$. Ovo je mjesto gdje trebamo posumnjati. Nigdje nije rečeno da su dvije jabuke svake vrste koje jabučar Jan stavlja u paket potpuno identične. Zato masu svake jabuke prestavljamo novom slučajnom varijablom koja ima istu razdiobu kao X odnosno Y. Ako sa X' i X'' označimo nezavisne kopije slučajne varijable X i sa Y' i Y'' nezavisne kopije slučajne varijable Y, onda je masa Janovog paketa dana slučajnom varijablom

$$Z = X' + X'' + Y' + Y''.$$

Svojstvo (3) se jednostavno proširi i na linearnu kombinaciju više nezavisnih slučajnih varijabli te je

$$Z \sim \mathcal{N}\left(1 \cdot 180 + 1 \cdot 180 + 1 \cdot 220 + 1 \cdot 220, 1^2 \cdot 400 + 1^2 \cdot 400 + 1^2 \cdot 25 + 1^2 \cdot 25\right) = \mathcal{N}(800, 850).$$

Tražena vjerojatnost je

$$\begin{aligned} P(820 < Z < 1000) &= \left(\frac{820 - 800}{\sqrt{850}} < \frac{Z - 800}{\sqrt{850}} < \frac{1000 - 800}{\sqrt{850}}\right) \\ &= \frac{1}{2}\left[\Phi^*(6.8599) - \Phi^*(0.686)\right] = \frac{1}{2}\left[1 - 0.50729\right] = 0.246. \end{aligned}$$

7. Vrijeme izrade T nekog proizvoda je normalna slučajna varijabla s očekivanjem 36 sati. U 20% slučajeva ta je duljina veća od 42 sata.

(a) Odredite parametre razdiobe varijable T.

(b) Odredite simetričan interval oko srednje vrijednosti unutar kojeg se s vjerojatnošću 85% realizira varijabla T.

(c) Što pravilo 3σ govori u ovom slučaju?

(d) Kad se završi izrada jednog proizvoda, započinje se s novim istih karakteristika. Kolika je vjerojatnost da će se dva proizvoda završiti u vremenu manjem od 70 sati?

Rješenje:

(a) Slučajna varijabla T ima normalnu razdiobu $\mathcal{N}(a, \sigma^2)$ gdje je $a = 36$ i σ za sada nepoznat. Zadano je

$$P(T > 42) = 0.2.$$

S druge strane

$$P(T > 42) = P\left(\frac{T-36}{\sigma} > \frac{42-36}{\sigma}\right) = \frac{1}{2}[1 - \Phi^*(\frac{6}{\sigma})].$$

Dobivamo

$$\frac{6}{\sigma} = 0.842 \implies \sigma = 7.125891.$$

(b) Simetričan interval oko srednje vrijednosti je oblika $\langle a - z, a + z\rangle$, za neki $z > 0$. Tražimo z takav da vrijedi

$$P(a - z < T < a + z) = 0.85.$$

Vjerojatnost s lijeve strane jednakosti je jednaka

$$P(a - z < T < a + z) = P\left(-\frac{z}{\sigma} < \frac{T-a}{\sigma} < \frac{z}{\sigma}\right) = \Phi^*\left(\frac{z}{\sigma}\right),$$

što znači da je

$$\Phi^*\left(\frac{z}{\sigma}\right) = 0.85,$$

odnosno

$$\frac{z}{\sigma} = 1.44 \implies z = 10.261.$$

Traženi interval je

$$\langle 25.7387, 46.26128\rangle.$$

(c) Pravilo 3σ kaže da je

$$P(a - 3\sigma < T < a + 3\sigma) = P(14.6223 < T < 57.3777) = 0.9973.$$

(d) Neka je

$$\begin{aligned} T_1 = &\ \text{vrijeme izrade 1. proizvoda}, \quad T_1 \sim \mathcal{N}(a, \sigma^2), \\ T_2 = &\ \text{vrijeme izrade 2. proizvoda}, \quad T_2 \sim \mathcal{N}(a, \sigma^2). \end{aligned}$$

Tada je

$$T_1 + T_2 \sim \mathcal{N}(a + a, \sigma^2 + \sigma^2)$$

i vrijedi

$$P(T_1 + T_2 < 70) = \frac{1}{2}\left[1 - \Phi^*\left(\frac{70-2a}{\sqrt{2}\sigma}\right)\right] = \frac{1}{2}\left[1 - \Phi^*(0.1985)\right] = 0.4215.$$