6.

Primjeri neprekinutih razdioba

Sadržaj poglavlja

Neke su razdiobe važnije od drugih, jer su važnije situacije u kojima se one pojavljuju. Među svim neprekinutim razdiobama izdvojit ćemo dvije koje se (uz jednoliku radiobu) najčešće pojavljuju.

6.1. Eksponencijalna razdioba

Eksponencijalna razdioba javlja se u problemima vezanim za vrijeme ispravnog rada uređaja čije se karakteristike ne mijenjaju u vremenu: vrijeme ispravnog rada žarulje, vrijeme do prvog poziva (pa i između dva uzastopna poziva) u telefonskoj centrali, vrijeme do ulova prve ribe, do pojave neke nesreće i općenito vrijeme do pojave nekog događaja čija je vjerojatnost pojavljivanja u svakom kratkom intervalu jednake duljine jednaka.

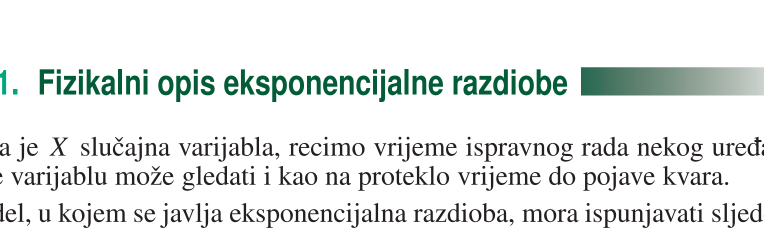
Eksponencijalna razdioba, definicija

Kažemo da slučajna varijabla X ima **eksponencijalnu razdiobu** s parametrom $\lambda > 0$ ako ona poprima pozitivne vrijednosti s gustoćom

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0. \tag{1}$$

Pišemo $X \sim \mathscr{E}(\lambda)$. Njezina je funkcija razdiobe

$$F(x) = 1 - e^{-\lambda x}, \qquad x > 0.$$



Sl. 6.1. Funkcija razdiobe i gustoće eksponencijalne razdiobe

6.1.1. Fizikalni opis eksponencijalne razdiobe

Neka je X slučajna varijabla, recimo vrijeme ispravnog rada nekog uređaja. Na tu se varijablu može gledati i kao na proteklo vrijeme do pojave kvara.

Model, u kojem se javlja eksponencijalna razdioba, mora ispunjavati sljedeće uvjete:

Teorem 6.1. Eksponencijalna razdioba, izvod iz modela pojavljivanja

Pretpostavimo da je uvjetna vjerojatnost pojavljivanja događaja u nekom vrlo kratkom intervalu $(x, x+\Delta x)$, ako je poznato da se događaj nije pojavio do trenutka x, proporcionalna duljini tog podintervala i ne ovisi o vrijednosti x:

$$\boldsymbol{P}(X < x+\Delta x \mid X > x) = \lambda \Delta x + r, \qquad \frac{r}{\Delta x} \to 0 \quad \text{kad} \quad \Delta x \to 0.$$

Tada je X slučajna varijabla distribuirana po eksponencijalnom zakonu $\mathscr{E}(\lambda)$.

Dokaz. Po formuli za uvjetnu vjerojatnost možemo napisati

$$\boldsymbol{P}(X > x+\Delta x) = \boldsymbol{P}(X > x) \cdot \boldsymbol{P}(X > x+\Delta x \mid X > x).$$

Međutim, kako je

$$\boldsymbol{P}(X > x+\Delta x \mid X > x) = 1 - \boldsymbol{P}(X < x+\Delta x \mid X > x) = 1 - \lambda\Delta x - r,$$

dobivamo

$$\boldsymbol{P}(X > x+\Delta x) = \boldsymbol{P}(X > x)(1 - \lambda\Delta x - r).$$

Definirajmo funkciju

$$Q(x) := \boldsymbol{P}(X > x) = 1 - F(x).$$

Ona zadovoljava relaciju

$$Q(x+\Delta x) - Q(x) = Q(x)[-\lambda\Delta x - r]$$

te je

$$\frac{dQ(x)}{dx} = \lim_{\Delta x \to 0} \frac{Q(x+\Delta x) - Q(x)}{\Delta x} = \lim_{\Delta x \to 0} Q(x)\left[-\lambda - \frac{r}{\Delta x}\right] = -\lambda Q(x).$$

Rješenje ove diferencijalne jednadžbe je

$$Q(x) = Ce^{-\lambda x}.$$

Konstantu C određujemo iz početnog uvjeta: $Q(0) = \boldsymbol{P}\{X > 0\} = 1$. Odavde $C = 1$ i zato

$$F(x) = 1 - Q(x) = 1 - e^{-\lambda x},$$

te X zaista ima eksponencijalnu razdiobu $\mathscr{E}(\lambda)$. ◄

6.1.2. Veza s Poissonovom razdiobom

Neka Poissonova slučajna varijabla Z mjeri broj pojavljivanja nekog događaja u nekom jediničnom vremenskom razdoblju. Označimo parametar te razdiobe s λ. (Taj je parametar jednak očekivanom broju realizacija varijable Z unutar tog jediničnog intervala.)

Općenitije, neka slučajna varijabla Z_x mjeri broj realizacija događaja u intervalu $[0, x]$. To je također Poissonova varijabla, s očekivanjem λx.

Definirajmo sada slučajnu varijablu X kao vrijeme do prvog pojavljivanja događaja. Pokažimo da X ima eksponencijalnu razdiobu.

Neka je x fiksno vrijeme nakon početnog trenutka, izraženo u jedinicama vremena. Slučajna varijabla X će poprimiti vrijednost manju od x ukoliko se u intervalu $[0, x]$ realizira barem jedan događaj, tj. ako Z_x poprimi vrijednost veću od nule:

$$\boldsymbol{P}(X < x) = \boldsymbol{P}(Z_x > 0) = 1 - \boldsymbol{P}(Z_x = 0) = 1 - e^{-\lambda x}.$$

Vidimo da X ima eksponencijalnu razdiobu s parametrom λ.

6.1.3. Očekivanje i disperzija eksponencijalne razdiobe

Odredimo Laplaceov transformat, karakterističnu funkciju, očekivanje, disperziju i centralne momente eksponencijalne razdiobe.

Laplaceov transformat eksponencijalne razdiobe $\mathscr{E}(\lambda)$ iznosi

$$\begin{aligned} f^*(s) &= \int_0^\infty e^{-sx} f(x)dx = \int_0^\infty e^{-sx}\lambda e^{-\lambda x}dx \\ &= \lambda \int_0^\infty e^{-(s+\lambda)x}dx = -\frac{\lambda}{s+\lambda}e^{-(s+\lambda)x}\Big|_{x=0}^{x=\infty} = \frac{\lambda}{s+\lambda} \end{aligned}$$

(Pri uvrštavanju gornje granice uvažavamo da je f^* definirana za $s > 0$.)

Karakterističnu funkciju eksponencijalne razdiobe možemo napisati pomoću poznate veze: $\vartheta(t) = f^*(-it)$:

$$\vartheta(t) = \frac{\lambda}{\lambda - it}.$$

Očekivanje eksponencijalne razdiobe izračunat ćemo iz veze momenata i Laplaceovog transformata:

$$\boldsymbol{E}(X) = -f^{*\prime}(0) = -\left(\frac{\lambda}{s+\lambda}\right)'\Big|_{s=0} = \frac{1}{\lambda}.$$

Ishodišni moment reda n računamo ovako:

$$\begin{aligned} \boldsymbol{E}(X^n) &= (-1)^n f^{*(n)}(0) = (-1)^n \left(\frac{\lambda}{s+\lambda}\right)^{(n)}\Big|_{s=0} \\ &= \frac{n!\lambda}{(s+\lambda)^{n+1}}\Big|_{s=0} = \frac{n!}{\lambda^n}. \end{aligned}$$

Sad za **disperziju** vrijedi

$$\boldsymbol{D}(X) = \boldsymbol{E}(X^2) - (\boldsymbol{E}X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

— ◆ —

Iz ovog izvoda vidimo da je očekivanje eksponencijalne razdiobe recipročna vrijednost njezinog parametra: što je parametar razdiobe veći, to je manje njezino očekivanje. Interesantno je da *vjerojatnost realizacije* varijable prije očekivanja ne ovisi o parametru razdiobe. Naime, vrijedi:

$$\boldsymbol{P}(X < \boldsymbol{E}(X)) = \boldsymbol{P}(X < \frac{1}{\lambda}) = F(\frac{1}{\lambda}) = 1 - e^{-\lambda\cdot(1/\lambda)} = 1 - e^{-1}.$$

Primjer 6.1.

Neki je uređaj u godini dana bio 10 puta u kvaru. Kolika je vjerojatnost da će prvi mjesec sljedeće godine raditi ispravno?

► Neka je X slučajna varijabla: vrijeme do prvog kvara. X ima eksponencijalnu razdiobu. Na temelju podataka, njezino očekivanje je $\boldsymbol{E}(X) = \frac{12}{10}$, pa je parametar razdiobe $\lambda = \frac{10}{12}$. Sada vrijedi

$$\begin{aligned} \boldsymbol{P}(X > 1) &= 1 - \boldsymbol{P}(X < 1) = 1 - \left(1 - e^{-\frac{10}{12}\cdot 1}\right) \\ &= e^{-5/6} = 0.435. \end{aligned}$$ ◄

Primjer 6.2.

Vrijeme ispravnog rada nekog uređaja je slučajna varijabla distribuirana po eksponencijalnom zakonu s očekivanjem 2 mjeseca. Kolika je vjerojatnost da će se uređaj pokvariti tijekom

A. prvog mjeseca,

B. drugog mjeseca,

C. drugog mjeseca, ako je poznato da se tijekom prvog mjeseca nije nalazio u kvaru.

► Po uvjetu zadatka je $X \sim \mathscr{E}(\frac{1}{2})$.

A. $\boldsymbol{P}(X < 1) = 1 - e^{-\frac{1}{2}} = 0.393$,

B. $\boldsymbol{P}(1 < X < 2) = (1 - e^{-\frac{1}{2}\cdot 2}) - (1 - e^{-\frac{1}{2}\cdot 1}) = e^{-\frac{1}{2}} - e^{-1} = 0.239$,

C. $$\begin{aligned} \boldsymbol{P}(1 < X < 2 \mid X > 1) &= \frac{\boldsymbol{P}(1 < X < 2,\ X > 1)}{\boldsymbol{P}(X > 1)} \\ &= \frac{\boldsymbol{P}(1 < X < 2)}{\boldsymbol{P}(X > 1)} = \frac{e^{-\frac{1}{2}} - e^{-1}}{e^{-\frac{1}{2}}} = 1 - e^{-\frac{1}{2}} = 0.393. \end{aligned}$$

Primijeti da se prva i treća vjerojatnost podudaraju. ◄

6.1.4. Odsustvo pamćenja

Svojstvo iz prethodnog primjera može se poopćiti. Za sve $x, t > 0$ vrijedi

$$\boldsymbol{P}(X < x+t \mid X > t) = \boldsymbol{P}(X < x).$$

što možemo riječima iskazati na način: eksponencijalna razdioba nema pamćenja. Evo primjera: ako je očekivano vrijeme do ulova prve ribe jedan sat, i ako je prvih 50 minuta proteklo bez ikakvog ulova, tada očekivano vrijeme do ulova te ribe ostaje (na žalost) i dalje jedan sat, bez obzira na proteklo vrijeme.

Ovo svojstvo karakterizira eksponencijalnu razdiobu.

Teorem 6.2. Karakterizacija eksponencijalne razdiobe

Neka za slučajnu varijablu X koja uzima samo pozitivne vrijednosti za sve pozitivne x i t vrijedi

$$\boldsymbol{P}(X < x+t \mid X > t) = \boldsymbol{P}(X < x). \tag{2}$$

Tada X ima eksponencijalnu razdiobu.

DOKAZ. Označimo $Q(x) := \boldsymbol{P}(X > x) = 1 - F(x)$, gdje je F tražena funkcija razdiobe. Po pretpostavci je $Q(0) = \boldsymbol{P}(X > 0) = 1$. Također, vrijedi $Q'(x) = -f(x) < 0$ za $x > 0$. Označimo $Q'(0) =: -\lambda$. Osnovnu relaciju (2) možemo napisati na način

$$1 - \boldsymbol{P}(X > x+t \mid X > t) = 1 - \boldsymbol{P}(X > x)$$

te je

$$\frac{\boldsymbol{P}(X > x+t, X > t)}{\boldsymbol{P}(X > t)} = \boldsymbol{P}(X > x)$$

ili

$$\boldsymbol{P}(X > x+t) = \boldsymbol{P}(X > t)\boldsymbol{P}(X > x).$$

Tako dobivamo funkcionalnu jednadžbu

$$Q(x+t) = Q(t)Q(x), \qquad \forall t, x > 0 \tag{3}$$

Slučajnu varijablu tražimo u klasi neprekinutih varijabli. Zato možemo pretpostaviti da je Q diferencijabilna funkcija. Deriviranjem relacije (3) po varijabli t dobivamo

$$Q'(x+t) = Q'(t)Q(x).$$

Uvrštavanjem $t = 0$ slijedi

$$Q'(x) = Q'(0)Q(x) = -\lambda Q(x)$$

i odavde dobivamo $Q(x) = Ce^{-\lambda x}$. Kako je $Q(0) = 1$, dobivamo $C = 1$. Zato je

$$F(x) = 1 - Q(x) = 1 - e^{-\lambda x}.$$ ◄

Primjer 6.3.

Vrijeme ispravnog rada elektroničkog elementa opisano je eksponencijalnom razdiobom s parametrom λ. U trenutku $x = T$, došlo je do strujnog udara. Vjerojatnost da element "ne preživi" taj udar (ako je do tog trenutka još uvijek bio ispravan) jednaka je p. Nakon toga, vrijeme daljnjeg ispravnog rada bit će opet eksponencijalna razdioba, s možda promijenjenim parametrom μ

A. Odredi i nacrtaj funkciju razdiobe ove slučajne varijable.

B. Izračunaj njezino očekivanje.

► Na intervalu $0 \leqslant x < T$ vrijedi

$$F(x) = 1 - e^{-\lambda x}.$$

Vrijednost ove funkcije u trenutku T iznosi $1 - e^{-\lambda T}$. Suprotna vjerojatnost, $e^{-\lambda T}$ je vjerojatnost da je element još uvijek ispravan, neposredno prije strujnog udara. Vjerojatnost da će on biti ispravan i nakon tog udara jednaka je umnošku vjerojatnosti

$$(1-p)\cdot e^{-\lambda T},$$

jer je preživljavanje udara nezavisan događaj od preživljavanja u normalnim uvjetima. Dakle, u točki T vrijedi

$$F(T) = 1 - (1-p)e^{-\lambda T}.$$

Zato je

$$\boldsymbol{P}(X > T) = 1 - F(T) = (1-p)e^{-\lambda T}.$$

Označimo ovu vjerojatnost s r. Za $x > T$ sad možemo napisati

$$1 - F(x) = \boldsymbol{P}(X > x) = \boldsymbol{P}(X > x \mid X > T)\cdot \boldsymbol{P}(X > T).$$

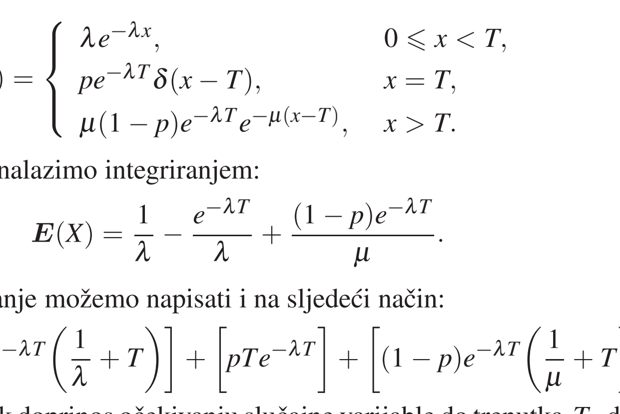
Zbog odsustva pamćenja (nakon trenutka T) uvjetnu vjerojatnost dobivamo ovako:

$$\boldsymbol{P}(X > x \mid X > T) = \boldsymbol{P}(X > x - T) = e^{-\mu(x-T)}.$$

Konačno, imamo za $x > T$

$$F(x) = 1 - e^{-\mu(x-T)}\cdot(1-p)e^{-\lambda T} = 1 - (1-p)e^{-\lambda T - \mu(x-T)}.$$

Graf ove funkcije nacrtan je na slici:



Sl. 6.2.

Iznos skoka u točki T iznosi $pe^{-\lambda T}$.

Slučajna varijabla je mješovitog tipa. Zato će njezina gustoća imati δ-udar u točki T:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & 0 \leqslant x < T, \\ pe^{-\lambda T}\delta(x-T), & x = T, \\ \mu(1-p)e^{-\lambda T}e^{-\mu(x-T)}, & x > T. \end{cases}$$

Očekivanje nalazimo integriranjem:

$$\boldsymbol{E}(X) = \frac{1}{\lambda} - \frac{e^{-\lambda T}}{\lambda} + \frac{(1-p)e^{-\lambda T}}{\mu}.$$

Ovo očekivanje možemo napisati i na sljedeći način:

$$\boldsymbol{E}(X) = \left[\frac{1}{\lambda} - e^{-\lambda T}\left(\frac{1}{\lambda} + T\right)\right] + \left[pTe^{-\lambda T}\right] + \left[(1-p)e^{-\lambda T}\left(\frac{1}{\mu} + T\right)\right].$$

Prvi je pribrojnik doprinos očekivanju slučajne varijable do trenutka T, drugi opisuje doprinos udara, a treći ponašanje slučajne varijable nakon trenutka T. Objasni, ne računanjući očekivanja, kako se dobiva svaki od tih pribrojnika. ◄

6.2. Normalna razdioba

6.2.1. Opis razdiobe

Normalna (govorimo još i Gaussova) razdioba najvažnija je neprekinuta razdioba, ima li se u vidu učestalost i važnost modela u kojima se ona pojavljuje. Razlog tome je što se ta razdioba javlja kao *granična* u svim situacijama kad je slučajna varijabla dobivena kao zbroj velikog broja međusobno nezavisnih pribrojnika.

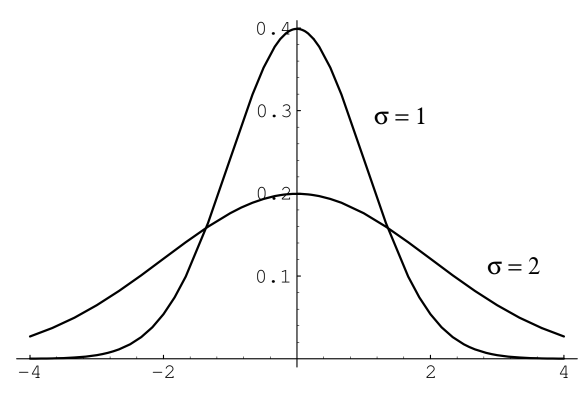
Definicija 6.1. Normalna razdioba

Slučajna varijabla X ima **normalnu razdiobu** s parametrima $a \in \mathbf{R}$ i $\sigma^2 > 0$ ako je X neprekinuta slučajna varijabla s gustoćom

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(x-a)^2}{2\sigma^2}\Big). \tag{4}$$

Pišemo $X \sim \mathcal{N}(a, \sigma^2)$.

Graf ove funkcije zvonolika je krivulja koju nazivamo **Gaussova krivulja**.



Sl. 6.3. Graf gustoće normalne razdiobe, za vrijednost parametra $a = 0$.

6.2.2. Jedinična normalna razdioba

Izaberemo li parametre $a = 0$ i $\sigma = 1$, dobivamo slučajnu varijablu $\mathcal{N}(0,1)$ koja se zove **jedinična normalna razdioba**. Njezinu gustoću označavamo sa

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}. \tag{5}$$

Za pripadnu funkciju razdiobe Φ vrijedi

$$\Phi(u) = \int_{-\infty}^{u} \phi(t)\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}t^2}\,\mathrm{d}t. \tag{6}$$

Ovaj integral nije elementaran, ne može se eksplicitno izraziti pomoću elementarnih funkcija.

Funkcija ϕ je parna, pa ima sljedeća svojstva:

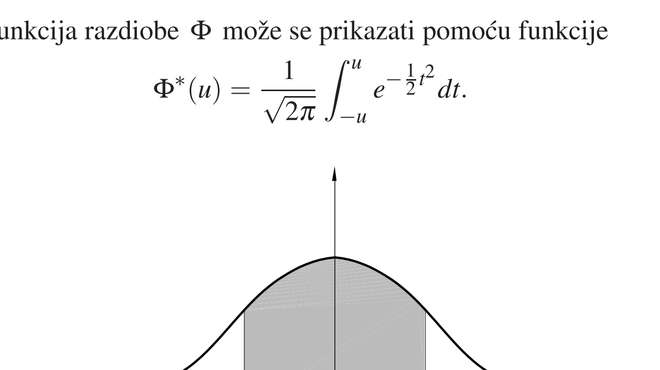
$$\int_{-\infty}^{0} \phi(t)\,\mathrm{d}t = \frac{1}{2}\int_{-\infty}^{\infty} \phi(t)\,\mathrm{d}t = \frac{1}{2},$$
$$\int_{0}^{u} \phi(t)\,\mathrm{d}t = \frac{1}{2}\int_{-u}^{u} \phi(t)\,\mathrm{d}t.$$

Sad možemo napisati

$$\begin{aligned}\Phi(u) &= \int_{-\infty}^{u} \phi(t)\,\mathrm{d}t = \int_{-\infty}^{0} \phi(t)\,\mathrm{d}t + \int_{0}^{u} \phi(t)\,\mathrm{d}t \\ &= \frac{1}{2} + \frac{1}{2}\int_{-u}^{u} \phi(t)dt = \frac{1}{2}\Big[1 + \Phi^*(u)\Big].\end{aligned}$$

Dakle, funkcija razdiobe Φ može se prikazati pomoću funkcije

$$\Phi^*(u) = \frac{1}{\sqrt{2\pi}} \int_{-u}^{u} e^{-\frac{1}{2}t^2} dt. \tag{7}$$



Sl. 6.4. Vrijednost funkcije Φ^* jednaka je površini ispod grafa funkcije, a nad ovim simetričnim intervalom.

U ovoj ćemo knjizi koristiti tabelirane vrijednosti funkcije Φ^*. Ta je funkcija neparna, pa je dovoljno znati njezine vrijednosti za pozitivne vrijednosti od u.

6.2.3. Veza između jedinične i općenite normalne razdiobe

Neka je X jedinična normalna radioba, a bilo koji realni i σ bilo koji pozitivni broj. Definirajmo slučajnu varijablu

$$Y = a + \sigma X.$$

Odredimo njezinu razdiobu! Za pripadnu gustoću g vrijedi

$$y = a + \sigma x, \qquad x = \frac{y-a}{\sigma},$$
$$g(y) = \phi(x) \cdot \left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = \frac{1}{\sigma}\phi\Big(\frac{y-a}{\sigma}\Big).$$

Dobili smo:

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-a)^2}{2\sigma^2}}.$$

Prema tome, varijabla Y ima normalnu razdiobu s parametrima a i σ^2.

Na potpuno isti način dobili bismo i obrnutu vezu. Ako krenemo od normalne varijable $X \sim \mathcal{N}(a, \sigma^2)$, tada slučajna varijabla

$$Y = \frac{X-a}{\sigma}$$

ima jediničnu normalnu razdiobu.

Veza jedinične i opće normalne razdiobe

Jedinična i opća normalna razdioba mogu se dobiti jedna iz druge *linearnom transformacijom*:

$$X \sim \mathcal{N}(0,1) \implies a + \sigma X \sim \mathcal{N}(a, \sigma^2),$$
$$X \sim \mathcal{N}(a, \sigma^2) \implies \frac{X-a}{\sigma} \sim \mathcal{N}(0,1).$$

6.2.4. Karakteristična funkcija

Odredimo karakterističnu funkciju normalne razdiobe $\mathcal{N}(a, \sigma^2)$.

Najprije ćemo odrediti karakterističnu funkciju jedinične normalne razdiobe $\mathcal{N}(0,1)$. Vrijedi

$$\vartheta(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} e^{itx} dx.$$

Deriviramo ovu funkciju i primijenimo parcijalnu integraciju

$$\begin{aligned}\vartheta'(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ixe^{-\frac{1}{2}x^2} e^{itx} dx \\ &= \frac{i}{\sqrt{2\pi}} \Big(-e^{itx}e^{-\frac{1}{2}x^2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} ite^{itx}e^{-\frac{1}{2}x^2} dx\Big) \\ &= -t\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx}e^{-\frac{1}{2}x^2} dx = -t\vartheta(t).\end{aligned}$$

Prema tome, ϑ zadovoljava diferencijalnu jednadžbu

$$\frac{d\vartheta(t)}{dt} = -t\vartheta(t)$$

uz početni uvjet $\vartheta(0) = 1$, koji vrijedi za svaku karakterističnu funkciju.

$$\frac{d\vartheta(t)}{\vartheta(t)} = -t\,dt \implies \vartheta(t) = \vartheta(0)e^{-\frac{1}{2}t^2} = e^{-\frac{1}{2}t^2}.$$

Ako je $X \sim \mathcal{N}(0,1)$, tada $a + \sigma X \sim \mathcal{N}(a, \sigma^2)$ i dobivamo karakterističnu funkciju opće normalne razdiobe $\mathcal{N}(a, \sigma^2)$:

$$\vartheta_{a+\sigma X}(t) = e^{ita}\vartheta_X(\sigma t) = e^{ita}e^{-\frac{1}{2}\sigma^2t^2} = e^{ita-\frac{1}{2}\sigma^2t^2}.$$

6.2.5. Očekivanje i disperzija: parametri normalne razdiobe

Izračunat ćemo očekivanje i disperziju, koristeći karakterističnu funkciju. Neka je $X \sim \mathcal{N}(a, \sigma^2)$. Vrijedi

$$\begin{aligned}&\vartheta(t) = e^{ita-\frac{1}{2}\sigma^2t^2} \\ &\vartheta'(t) = (ia - \sigma^2 t)e^{ita-\frac{1}{2}\sigma^2t^2}, && \vartheta'(0) = ia \\ &\vartheta''(t) = [-\sigma^2 + (ia - \sigma^2 t)^2]e^{ita-\frac{1}{2}\sigma^2t^2}, && \vartheta''(0) = -\sigma^2 - a^2.\end{aligned}$$

Zato je

$$\boldsymbol{E}(X) = -i\vartheta'(0) = a,$$
$$\boldsymbol{D}(X) = -\vartheta''(0) + \vartheta'(0)^2 = \sigma^2.$$

Time smo odredili značenje parametara normalne razdiobe: parametar a jednak je **očekivanju**, a parametar σ^2 **disperziji** normalne razdiobe. Korijen iz disperzije je σ, koji se naziva **standardno odstupanje** ili **standardna devijacija** normalne razdiobe

6.2.6. Računanje vjerojatnosti za jediničnu normalnu razdiobu

Neka je X jedinična normalna razdioba. Pokažimo kako se računaju temeljne vjerojatnosti, da slučajna varijabla poprimi vrijednost unutar nekog intervala.

Računanje vjerojatnosti za jediničnu normalnu razdiobu

Za jediničnu normalnu razdiobu X vrijedi:

$$\boldsymbol{P}(u_1 < X < u_2) = \Phi(u_2) - \Phi(u_1) = \frac{1}{2}\Big[\Phi^*(u_2) - \Phi^*(u_1)\Big]. \tag{8}$$

Posebice, u slučaju simetričnog intervala

$$\boldsymbol{P}(|X| < u) = \frac{1}{2}\Big[\Phi^*(u) - \Phi^*(-u)\Big] = \Phi^*(u). \tag{9}$$

Primjer 6.4.

Neka je X jedinična normalna varijabla. Odredi vjerojatnost događaja

A. $0 < X < 1$; **B.** $-2 < X < 0$; **C.** $-1 < X < 2$;
D. $1 < X < 2$; **E.** $X < 1$; **F.** $X > 1$.

▶

$$\boldsymbol{P}(0 < X < 1) = \tfrac{1}{2}[\Phi^*(1) - \Phi^*(0)] = \tfrac{1}{2}\Phi^*(1) = 0.341,$$
$$\boldsymbol{P}(-2 < X < 0) = \tfrac{1}{2}[\Phi^*(0) - \Phi^*(-2)] = \tfrac{1}{2}\Phi^*(2) = 0.477,$$
$$\boldsymbol{P}(-1 < X < 2) = \tfrac{1}{2}[\Phi^*(2) - \Phi^*(-1)] = \tfrac{1}{2}[\Phi^*(2) + \Phi^*(1)] = 0.819,$$
$$\boldsymbol{P}(1 < X < 2) = \tfrac{1}{2}[\Phi^*(2) - \Phi^*(1)] = 0.136,$$
$$\boldsymbol{P}(X < 1) = \tfrac{1}{2} + \tfrac{1}{2}\Phi^*(1) = 0.841,$$
$$\boldsymbol{P}(X > 1) = \tfrac{1}{2} - \tfrac{1}{2}\Phi^*(1) = 0.159.$$ ◀

6.2.7. Računanje vjerojatnosti općenite normalne razdiobe

Neka je $X \sim \mathcal{N}(a, \sigma^2)$. Tada je $\dfrac{X-a}{\sigma} \sim \mathcal{N}(0,1)$ i funkciju razdiobe F varijable X možemo izraziti uz pomoć funkcije Φ:

$$F(x) = \Phi(u) = \tfrac{1}{2}[1 + \Phi^*(u)], \qquad u = \frac{x-a}{\sigma}.$$

Tako računamo

$$\begin{aligned}\boldsymbol{P}(x_1 < X < x_2) &= \boldsymbol{P}\Big(\frac{x_1-a}{\sigma} < \frac{X-a}{\sigma} < \frac{x_2-a}{\sigma}\Big) \\ &= \Phi\Big(\frac{x_2-a}{\sigma}\Big) - \Phi\Big(\frac{x_1-a}{\sigma}\Big) = \frac{1}{2}\Big[\Phi^*\Big(\frac{x_2-a}{\sigma}\Big) - \Phi^*\Big(\frac{x_1-a}{\sigma}\Big)\Big].\end{aligned}$$

Primjer 6.5.

Neka je $X \sim \mathcal{N}(2,4)$. Odredi vjerojatnost događaja

A. $0 < X < 4$; **B.** $2 < X < 6$; **C.** $X > 4$.

▶ Ako X ima opću normalnu razdiobu, tada ćemo sa $\widetilde{X}$ označavati pripadnu jediničnu normalnu slučajnu varijablu, $\widetilde{X} = \dfrac{X-a}{\sigma}$.

A.

$$\begin{aligned}\boldsymbol{P}(0 < X < 4) &= \boldsymbol{P}\Big(\frac{0-2}{2} < \frac{X-2}{2} < \frac{4-2}{2}\Big) \\ &= \boldsymbol{P}(-1 < \widetilde{X} < 1) = \Phi^*(1) = 0.683.\end{aligned}$$

B.

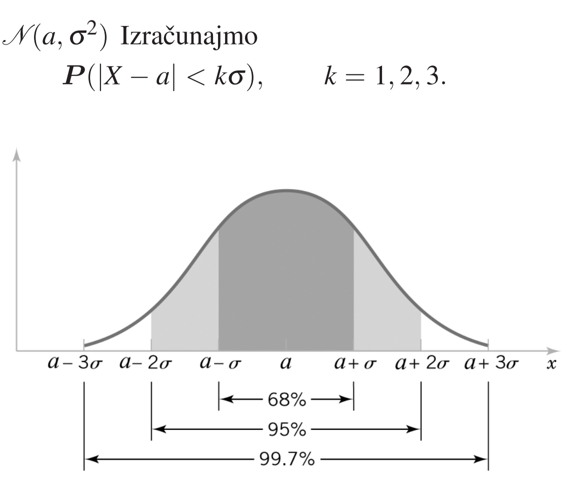
$$\begin{aligned}\boldsymbol{P}(2 < X < 6) &= \boldsymbol{P}\Big(\frac{2-2}{2} < \frac{X-2}{2} < \frac{6-2}{2}\Big) \\ &= \boldsymbol{P}(0 < \widetilde{X} < 2) = \tfrac{1}{2}\Phi^*(2) = 0.477.\end{aligned}$$

C. Na isti način $\boldsymbol{P}(X > 4) = \frac{1}{2}(1 - \Phi^*(1)) = 0.158$. ◀

6.2.8. Pravilo 3σ

Neka je $X \sim \mathcal{N}(a, \sigma^2)$ Izračunajmo

$$\boldsymbol{P}(|X-a| < k\sigma), \qquad k = 1,2,3.$$



Sl. 6.5.

Imamo

$$\boldsymbol{P}(|X-a| < k\sigma) = \boldsymbol{P}(-k\sigma < X - a < k\sigma) = \boldsymbol{P}(-k < \widetilde{X} < k) = \Phi^*(k).$$

Vrijedi $\Phi^*(1) = 0.6827$, $\Phi^*(2) = 0.9545$, $\Phi^*(3) = 0.9973$. Dakle, s vjerojatnošću 99.73% (odnosno, praktički sigurno) normalna varijabla uzima vrijednosti unutar intervala $(a - 3\sigma, a + 3\sigma)$. Ta se činjenica naziva *pravilo tri sigma*.

Primjer 6.6.

Sistematska greška održavanja zrakoplova na danoj visini je 100 m, a slučajna greška je normalna varijabla s odstupanjem 200 m. Odredi vjerojatnost da

A. zrakoplov leti kroz koridor širine 500 m,
B. zrakoplov leti iznad tog koridora,
ako je zrakoplov usmjeren da leti sredinom koridora.

▶ Označimo sa X slučajnu varijablu: odstupanje zrakoplova od sredine koridora. Ta je varijabla zbroj sistematske i slučajne pogreške. Stoga je $X \sim \mathcal{N}(100, 200^2)$.

$$\begin{aligned}\boldsymbol{P}(-250 < X < 250) &= \boldsymbol{P}\Big(\frac{-250-100}{200} < \frac{X-100}{200} < \frac{250-100}{200}\Big) \\ &= \tfrac{1}{2}[\Phi^*(\tfrac{3}{4}) - \Phi^*(-\tfrac{7}{4})] = \tfrac{1}{2}[\Phi^*(0.75) + \Phi^*(1.75)] = 0.733 \\ \boldsymbol{P}(250 < X) &= \boldsymbol{P}\Big(\frac{250-100}{200} < \frac{X-100}{200}\Big) \\ &= \tfrac{1}{2}[1 - \Phi^*(0.75)] = 0.5 - 0.273 = 0.227\end{aligned}$$ ◀

-

6.2.9. Stabilnost normalne razdiobe

Već smo upoznali svojstvo stabilnosti pojedinih razdioba (recimo, kod binomne i Poissonove razdiobe), kad zbroj nezavisnih slučajnih varijabli ima razdiobu istog tipa.

Normalna razdioba je jedina koja ima pojačano svojstvo stabilnosti.

Teorem 6.3. ■ Stabilnost normalne razdiobe

Neka su X_1 i X_2 nezavisne slučajne varijable s normalnim razdiobama
$$X_1 \sim \mathcal{N}(a_1, \sigma_1^2), \quad X_2 \sim \mathcal{N}(a_2, \sigma_2^2)$$
i s_1, s_2 bilo koji realni brojevi. Tada vrijedi
$$s_1X_1 + s_2X_2 \sim \mathcal{N}(s_1a_1 + s_2a_2, s_1^2\sigma_1^2 + s_2^2\sigma_2^2).$$

Dokaz. Karakteristične funkcije slučajnih varijabli X_1 i X_2 su
$$\vartheta_{X_k}(t) = \exp(a_kti - \tfrac{1}{2}\sigma_k^2t^2), \qquad k = 1, 2.$$
Zato je
$$\vartheta_{s_kX_k}(t) = \exp(s_ka_kti - \tfrac{1}{2}s_k^2\sigma_k^2t^2), \qquad k = 1, 2.$$
pa dobivamo
$$\vartheta_{s_1X_1+s_2X_2}(t) = \exp((s_1a_1 + s_2a_2)ti - \tfrac{1}{2}(s_1^2\sigma_1^2 + s_2^2\sigma_2^2)t^2),$$
što je karakteristična funkcija normalne razdiobe $\mathcal{N}(s_1a_1+s_2a_2, s_1^2\sigma_1^2+s_2^2\sigma_2^2)$. ◀

— ◆ —

Može se dokazati, u što se ovdje ne možemo upuštati, da ovo svojstvo *karakterizira normalnu razdiobu.* To znači da je normalna razdioba jedina stabilna na uzimanje linearnih kombinacija nezavisnih pribrojnika.

Primjer 6.7.

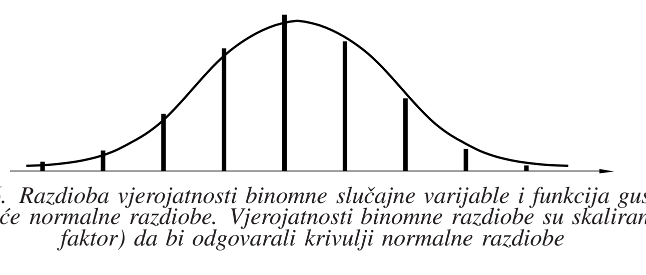
Težina proizvoda A podvrgava se normalnom zakonu s parametrima $a_1 = 40$ p, $\sigma_1 = 5$ p, a težina proizvoda B normalnom zakonu s parametrima $a_2 = 100$ p, $\sigma_2 = 10$ p. U jedan omot pakiraju se po dva proizvoda A i dva proizvoda B. Odredi vjerojatnost da se težina tako načinjenog paketa kreće između 270 p i 300 p. (Težinu omota zanemari.)

▶ Neka su X_A i X_B slučajne varijable koje opisuju težinu proizvoda A, odnosno B. Prema uvjetima je
$$X_A \sim \mathcal{N}(40, 5^2), \qquad X_B \sim \mathcal{N}(100, 10^2).$$
Označimo s Y težinu paketa. On sadrži dva proizvoda A i dva proizvoda B, pa bismo mogli napisati $Y = 2X_A + 2X_B$. No, taj zapis nije ispravan. Naime, mi ne računamo *dvostruku* težinu proizvoda A, već zbroj dviju težina koje su nezavisne jedna od druge, a imaju istu razdobu. Zato je ispravno napisati
$$Y = X_A' + X_A'' + X_B' + X_B''.$$
ovdje su X_A' i X_A'' *nezavisne kopije* slučajne varijable X_A, dakle, nezavisne slučajne varijable koje imaju istu razdiobu kao i X_A. Slično vrijedi za X_B' i X_B''.

Prema teoremu 6.3, vrijedit će
$$Y \sim \mathcal{N}(40 + 40 + 100 + 100, 25 + 25 + 100 + 100) = \mathcal{N}(280, 250).$$
Trebamo izračunati vjerojatnost $\boldsymbol{P}(270 < Y < 300)$:
$$\boldsymbol{P}\left(\frac{270 - 280}{\sqrt{250}} < Y^* < \frac{300 - 280}{\sqrt{250}}\right) = \boldsymbol{P}(-0.632 < Y^* < 1.265)$$
$$= \tfrac{1}{2}(\Phi^*(1.265) + \Phi^*(0.632)) = \tfrac{1}{2}(0.794 + 0.473) = 0.634\,. \quad ◀$$

6.2.10. Aproksimacija binomne razdiobe normalnom

Neka je $X \sim \mathscr{B}(n, p)$ binomna slučajna varijabla. Razdioba ove varijable za velike brojeve n nalikuje funkciji gustoće normalne varijable $Y \sim \mathcal{N}(np, npq)$. Za fiksni n, kvaliteta aproksimacije je bolja što je p bliži $\frac{1}{2}$.



Sl. 6.6. Razdioba vjerojatnosti binomne slučajne varijable i funkcija gustoće odgovarajuće normalne razdiobe. Vjerojatnosti binomne razdiobe su skalirane (za isti faktor) da bi odgovarali krivulji normalne razdiobe

Označimo s f gustoću varijable Y. Tada vrijedi

Teorem 6.4. ■ Lokalni teorem Moivre-Laplacea

Vjerojatnosti realizacija binomne slučajne varijable $\mathscr{B}(n, p)$ mogu se aproksimirati pomoću funkcije gustoće normalne varijable $\mathcal{N}(np, npq)$:
$$\boldsymbol{P}(X = m) \approx \boldsymbol{P}(m - \tfrac{1}{2} < Y < m + \tfrac{1}{2}) \approx f(m) = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(m-np)^2}{2npq}}.$$

Primjer 6.8.

Novčić je bačen 40 puta. Koliki je najvjerojatniji broj pojavljivanja pisma? S kojom vjerojatnošću?

▶ Broj pojavljenih pisma je slučajna varijabla s razdiobom $X \sim \mathscr{B}(40, \frac{1}{2})$. Najvjerojatniji broj realizacija ove varijable je 20. Po lokalnom teoremu, vrijednost $\boldsymbol{P}(X = 20)$ možemo aproksimirati formulom
$$\boldsymbol{P}(X = 20) \approx \frac{1}{\sqrt{2\pi \cdot 40 \cdot \frac{1}{2} \cdot \frac{1}{2}}} e^{-\frac{(20-40\cdot\frac{1}{2})^2}{2\cdot 40\cdot\frac{1}{2}\cdot\frac{1}{2}}} = \frac{1}{\sqrt{20\pi}} = 0.12616\,.$$
Točna vrijednost iznosi $\binom{40}{20} \cdot \frac{1}{2^{40}} = 0.12537$. Uočite kvalitetu ove aproksimacije. ◀

6.2.11. Centralni granični teorem – Teorem Moivre-Laplacea

Jedan od najvažnijih zakona prirode, koji uvodi red u kaotičnu slučajnost, jest *centralni granični teorem.* Iskaz teorema i dokaz odložit ćemo do 9. poglavlja. Ovdje ćemo se zadovoljiti s iskazom najvažnije posebnog slučaja tog važnog teorema, koji se odnosi na slučajne varijable koje imaju binomnu razdiobu.

Neka X ima binomnu razdiobu, $X \sim \mathscr{B}(n, p)$. Očekivanje ove varijable je np, a disperzija npq. Zato varijabla $\dfrac{X - np}{\sqrt{npq}}$ ima očekivanje 0 i disperziju 1.

Teorem 6.5. ■ Teorem Moivre-Laplace

Neka je $X \sim \mathscr{B}(n, p)$. Za veliki n razdioba slučajne varijable $\dfrac{X - np}{\sqrt{npq}}$ može se aproksimirati jediničnom normalnom razdiobom:
$$\boldsymbol{P}\left(x_1 < \frac{\mathscr{B}(n, p) - np}{\sqrt{npq}} < x_2\right) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{1}{2}u^2} du. \tag{10}$$
Pišemo $\mathscr{B}(n, p) \approx \mathcal{N}(np, npq)$, za dovoljno veliki n.

Aproksimacija je dobra već za relativno male vrijednosti broja n, npr. za $n \geqslant 10$. Nadalje, aproksimacija je to bolja (i točnija za male vrijednosti od n) što je vrijednost parametra p bliža $\frac{1}{2}$.

Primjer 6.9.

Izračunaj vjerojatnost da se u 10 000 bacanja ispravnog novčića broj pisama nalazi između 4 950 i 5 100.

▶ Označimo sa X broj pojavljenih pisama. Tada je $X \sim \mathscr{B}(10000, \frac{1}{2})$. Međutim, vjerojatnost događaja $(4950 < X < 5100)$ ne može se jednostavno izračunati, zbog mukotrpnog računanja binomnih koeficijenata i velikog broja pribrojnika. Stoga X aproksimiramo normalnom razdiobom $\mathcal{N}(5000, 2500)$. Zbog velikog broja $n = 10000$ aproksimacija će biti iznimno dobra.
$$\boldsymbol{P}(4950 < X < 5100) = \boldsymbol{P}\left(\frac{4950 - 5000}{\sqrt{2500}} < \frac{X - 5000}{\sqrt{2500}} < \frac{5100 - 5000}{\sqrt{2500}}\right)$$
$$= \tfrac{1}{2}[\Phi^*(2) + \Phi^*(1)] = 0.477 + 0.341 = 0.818. \quad ◀$$

Primjer 6.10.

Vjerojatnost rođenja dječaka približno je jednaka 0.515. Kolika je vjerojatnost da među 10 000 novorođene djece bude više djevojčica?

▶ Broj novorođene muške djece je slučajna varijabla $X \sim \mathscr{B}(10000, 0.515)$. Ponovo ćemo aproksimirati X normalnom razdiobom s parametrima $a = 10000\cdot 0.515$, $\sigma^2 = 10000\cdot 0.515\cdot 0.485$, $X \sim \mathcal{N}(5150, 2497.75)$. Tražimo vjerojatnost događaja $\{X < 5000\}$:
$$\boldsymbol{P}(X < 5000) = \frac{1}{2}\left[1 + \Phi^*\left(\frac{5000 - 5150}{\sqrt{2497.75}}\right)\right] = \tfrac{1}{2}[1 - \Phi^*(3.01)] = 0.0013. \quad ◀$$

Primjer 6.11.

Kako se određuje vjerojatnost rođenja dječaka? Ako je vjerovati podacima, od 1871 do 1900 u Švicarskoj se rodio 1 359 671 dječak i 1 285 086 djevojčica. Odredimo interval unutar kojeg se s vjerojatnošću 0.997 nalazi vjerojatnost p rođenja muškog djeteta.

▶ Kao vrijednost vjerojatnosti p uzimamo relativnu frekvenciju rođenja muškog djeteta
$$p_0 = \frac{m}{n} = \frac{1\,359\,671}{2\,644\,757} = 0.5141.$$
Broj m novorođene muške djece ravna se po binomnoj razdiobi s parametrima p_0 i $n = 2\,644\,717$. Možemo ga aproksimirati sa zakonom $\mathcal{N}(np_0, np_0q_0)$. Stoga je razdioba varijable p
$$p = \frac{m}{n} \sim \mathcal{N}\left(p_0, \sqrt{\frac{p_0q_0}{n}}\right).$$
Označimo $\sigma = \sqrt{\dfrac{p_0q_0}{n}}$. Po pravilu 3σ vrijedi $P(|p - p_0| < 3\sigma) = 0.997$. Primijetimo još da je uvijek $p_0q_0 < \frac{1}{4}$. Stoga
$$3\sigma = 3\sqrt{\frac{p_0q_0}{n}} \leqslant \frac{3}{\sqrt{4n}} = 0.00092.$$
Dobili smo
$$0.5132 < p < 0.5150$$
s vjerojatnošću od barem 0.997. ◀

Primjer 6.12.

Vjerojatnost pojavljivanja događaja pri jednom pokusu je 0.3. S kojom vjerojatnošću možemo tvrditi da učestalost tog događaja u 100 pokusa leži u granicama 0.2 – 0.4?

▶ Broj pojavljivanja događaja u 100 pokusa je slučajna varijabla $X \sim \mathscr{B}(100, 0.3)$ koju možemo aproksimirati normalnom $\mathcal{N}(30, 21)$.
$$\boldsymbol{P}(20 \leqslant X \leqslant 40) = \boldsymbol{P}\left(\frac{19.5 - 30}{\sqrt{21}} < \frac{X - 30}{\sqrt{21}} < \frac{40.5 - 30}{\sqrt{21}}\right)$$
$$= \Phi^*\left(\frac{10.5}{\sqrt{21}}\right) = \Phi^*(2.29) = 0.978.$$
Da bismo povećali točnost rezultata kod aproksimacije diskretne razdiobe neprekidnom, donju smo granicu s 20 smanjili na 19.5, a gornju granicu s 40 povećali na 40.5. Ta se korekcija primjenjuje za relativno male vrijednosti od n. Naime, događaju $(X = k)$ za diskretnu varijablu X odgovara događaj $(k - \frac{1}{2} < X < k + \frac{1}{2})$ kod aproksimacije neprekinutom varijablom.

Pretpostavimo da gornji pokus ponavljamo 1000 puta. Izračunaj vjerojatnost da učestalost događaja bude u istim granicama i usporedi rezultate. ◀