7 Paskaita. Absoliučiai tolydžiųjų atsitiktinių dydžių apžvalga

Diskrečiųjų atsitiktinių dydžių pasiskirstymo funkcijų grafikai turi trūkių. Jei X yra diskretusis atsitiktinis dydis, o a yra jo reikšmė, $P(X = a) = p$, tai taške $x = a$ pasiskirstymo funkcija $F_X(x)$ turės trūkį, kurio dydis yra p, t. y.

$$F(x+\varepsilon) - F(x-\varepsilon) \to p, \varepsilon \to 0, \varepsilon > 0.$$

Jeigu atsitiktinio dydžio X pasiskirstymo funkcija yra visuose taškuose tolydi, tai tokį atsitiktinį dydį vadinsime tolydžiuoju. Tolydžiųjų atsitiktinių dydžių pavyzdžiai – tolygusis, eksponentinis, normalusis.

Pavyzdys. Tolygusis atsitiktinis dydis

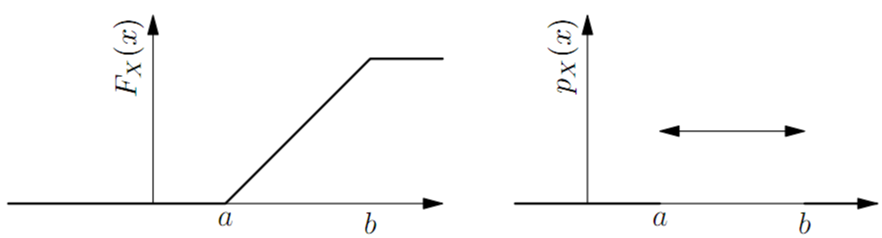
Tarkime, atsitiktinio dydžio ξ reikšmė – atsitiktinai parinktas intervalo $[a, b], a < b$, skaičius. Jeigu visi skaičiai turi vienodas galimybes būti parinkti, tai pasiskirstymo funkciją surasime pasinaudoję geometriniu tikimybių apibrėžimu

$$F_\xi(x) = \begin{cases} 0, & \text{jei } x \le a, \\ \dfrac{x-a}{b-a}, & \text{jei } a < x < b, \\ 1, & \text{jei } x \ge b. \end{cases}$$

Kai $x \ne a, b$ pasiskirstymo funkcija turi išvestinę. Kai $x \notin [a, b]$, tai $F'_\xi(x) = 0$, kai $x \in [a, b]$, tai $F'_\xi(x) = 1/(b-a)$.

Apibrėžimas. *Sakysime, kad atsitiktinis dydis yra tolygiai pasiskirstęs intervale* $[a, b]$*, jei jis turi tankį*

$$p_\xi(x) = \begin{cases} 0, & \textit{jei } \notin [a, b], \\ \dfrac{1}{b-a}, & \textit{jei } \in (a, b), \end{cases}$$



1 pav.: Tolygiojo skirstinio pasiskirstymo ir tankio funkcijos.

Pavyzdys. Eksponentinis atsitiktinis dydis

Radžio atomas skyla, virsdamas radono atomu ir išspinduliuodamas alfa dalelę:

$$Ra \to Rn + \alpha.$$

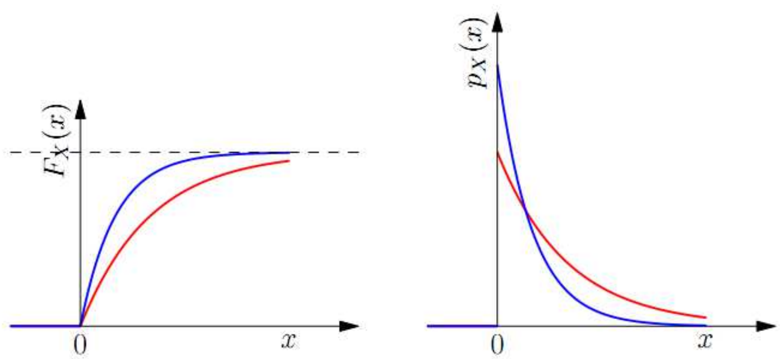
Tarkime ξ yra laikotarpis nuo atomo stebėjimo pradžios iki skilimo momento. Tuomet atsitiktinio dydžio ξ pasiskirstymo funkcijos reikšmė taške x parodo tikimybę, kad atomas skils per laikotarpį x:

$$F_\xi(x) = \begin{cases} 0, & \text{jei } x \le 0, \\ 1 - e^{-\lambda x}, & \text{jei } x > 0, \end{cases}$$

čia $\lambda > 0$ – parametras, priklausantis nuo fizinių radžio savybių. Tikimybė, kad atomas nesuskils per laikotarpį x yra $G(x) = e^{-\lambda x}$. Eksponentinio atsitiktinio dydžio tankis:

$$p_\xi(x) = \begin{cases} 0, & \text{jei } x \le 0, \\ \lambda e^{-\lambda x}, & \text{jei } x > 0, \end{cases}$$

Eksponentinį atsitiktinį dydį žymėsime $X \sim \mathcal{E}(\lambda)$. Tegu atsitiktinis dydis X reiškia laiko intervalų tarp tam tikrų įvykių ilgį, $X \sim \mathcal{E}(\lambda)$. Tuomet atsitiktinis dydis Y, kuris reiškia įvykių skaičių per laikotarpį (per laiko vienetą), turi Puasono skirstinį su tuo pačiu parametru λ: $Y \sim \mathcal{P}(\lambda)$.



2 pav.: Eksponentinio skirstinio pasiskirstymo ir tankio funkcijos.

Pavyzdys. Atsitiktinis dydis X yra pasiskirstęs pagal eksponentinį dėsnį $X \sim \mathcal{E}(2)$. Raskite atsitiktinio dydžio $Y = X^2$ pasiskirstymo funkciją ir tankį.

$F_Y(x) = P(X^2 < x) = P(X < \sqrt{x}) = 1 - e^{-2\sqrt{x}}$, kai $x > 0$.

$p_Y(x) = F'_Y(x) = \dfrac{e^{-2\sqrt{x}}}{\sqrt{x}}$.

Pavyzdys. Pareto atsitiktinis dydis

Maždaug prieš šimtą metų ekonomistas Vilfredas Pareto nustatė, kad parinkus tam tikrus skaičius $C > 0, \alpha > 0$ gyventojų, kurių metinės pajamos ne mažesnės kaip x, dalį galima gana tiksliai vertinti dydžiu C/x^α. Tikimybių teorijos kalba tai reikštų štai ką: jeigu X yra atsitiktinai parinkto gyventojo pajamas reiškiantis dydis, tai $P(X \ge x) \approx \dfrac{C}{x^\alpha}$, kai $x \ge x_0$, čia x_0 – tam tikras fiksuotas skaičius. Kad būtų paprasčiau, tarkime, kad $C = 1, x_0 = 1$ ir rašykime tikslias lygybes. Taigi

$$P(X \ge x) = \frac{1}{x^\alpha}, F_X(x) = 1 - P(X \ge x) = 1 - \frac{1}{x^\alpha}.$$

Akivaizdu, kad dydis turi tankį, kai $x > 1$: $p_X(x) = \alpha/x^{\alpha+1}$.

Apibrėžimas. *Jei atsitiktinis dydis X turi tankį*

$$p_X(x) = \begin{cases} 0, & \text{jei } x < 1, \\ \dfrac{\alpha}{x^{\alpha+1}}, & \text{jei } x \geq 1, \end{cases}$$

čia $\alpha > 0$, tai X vadinsime Pareto atsitiktiniu dydžiu.

O dabar susipažinsime su kone pačia svarbiausia tolydžiųjų atsitiktinių dydžių šeima. Nagrinėkime Bernulio schemą, tegu p – sėkmės tikimybė viename bandyme, o n – bandymų skaičius. Tegu bandymų skaičius n vis didėja. Teisinga tokia teorema.

Teorema. Muavro-Laplaso integralinė teorema. *Tegu sėkmės tikimybė Bernulio schemoje lygi p, o ξ_n – sėkmių skaičius po n bandymų. Tada bet kokiam x*

$$P\left(w : \frac{\xi_n(w) - np}{\sqrt{np(1-p)}} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, n \to \infty.$$

Šioje teoremoje pasirodo labai svarbus tikimybių teorijoje skirstinys.

Apibrėžimas. *Jei atsitiktinis dydis X turi tankį*

$$p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, -\infty < x < \infty,$$

tai sakysime, kad jis pasiskirstęs pagal standartinį normalųjį dėsnį. Rašysime $X \sim \mathcal{N}(0,1)$.

Šio teiginio esmė tokia: kai n didelis, tai dydis $\dfrac{\xi_n(w) - np}{\sqrt{np(1-p)}}$ supanašėja su standartiniu normaliuoju atsitiktiniu dydžiu, kurio pasiskirstymo funkcija yra

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

Tegu X yra standartinis normalusis dydis, o $\sigma > 0$ ir μ – realieji skaičiai. Sudarykime naują atsitiktinį dydį $Y = \sigma X + \mu$. Rasime jo pasiskirstymo funkciją.

$$F_Y(u) = P(\sigma X + \mu < u) = P\left(X < \frac{u - \mu}{\sigma}\right) = F_X(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-x^2/2} dx,$$

čia $v = \dfrac{u-\mu}{\sigma}$.

Paskutinį integralą kintamojo keitimu galime pertvarkyti taip:

$$F_X(v) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{u} e^{-(x-\mu)^2/2\sigma^2} dx.$$

Iš gautosios lygybės išplaukia, kad atsitiktinio dydžio Y tankis yra

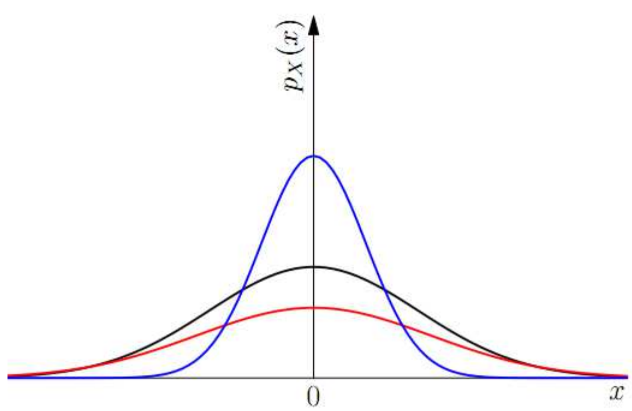
$$p_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}.$$

Apibrėžimas. *Atsitiktinį dydį X, kurio tankis yra*

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$

vadinsime normaliuoju dydžiu, žymėsime $X \sim \mathcal{N}(\mu, \sigma)$.

Taigi normalieji atsitiktiniai dydžiai sudaro didelę šeimą. Standartinis normalusis dydis – tik vienas iš šios šeimos atstovų. Nustatėme, kad tiesiškai transformuojant dydį $X \sim \mathcal{N}(0,1)$ gautasis dydis $Y = \sigma X + \mu$ yra normalusis. Jei $Y \sim \mathcal{N}(\mu, \sigma)$, tai panašiai galime įsitikinti, kad dydis $X = \dfrac{Y-\mu}{\sigma}$ yra standartinis normalusis. Standartinio normaliojo tankio grafikas yra simetriškas tiesės $x = 0$ atžvilgiu. Jeigu μ fiksuosime, o σ mažinsime, tankio kreivė darysis statesnė, jeigu σ didinsime – kreivė lėkštės (žr. pav. 3).



3 pav.: Normaliojo skirstinio tankio funkcijų priklausomybė nuo parametro σ.

Normalųjį skirstinį apibrėšime ir daugiamačiu atveju.

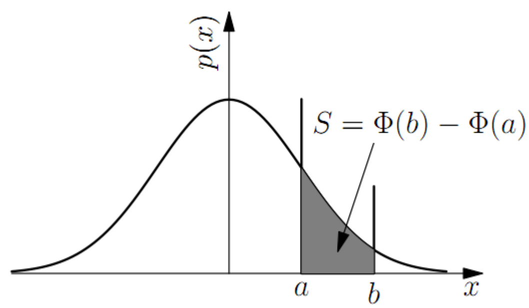
Apibrėžimas. *Sakysime, kad atsitiktinio vektoriaus $\xi : \Omega \to \mathcal{R}^n$ skirstinys yra standartinis normalusis, jei ξ yra absoliučiai tolydusis atsitiktinis vektorius, turintis tankį*

$$p_\xi(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\sum_{k=1}^{n} x_k^2\}.$$

Skirtumas $\Phi(b) - \Phi(a)$ turi paprastą geometrinę prasmę – jis lygus plotui po tankio funkcijos $p(x)$ grafiku, kurį riboja tiesės $y = 0, x = a, x = b$, žr. pav. 4. Savo ruožtu $\Phi(v)$ reikšmė lygi plotui po šiuo grafiku į kairę nuo tiesės $x = v$. Visas plotas po grafiku virš tiesės $y = 0$ lygus 1. Kadangi funkcija $p(x)$ yra lyginė, tai jos grafikas yra simetriškas tiesės $x = 0$ atžvilgiu, todėl $\Phi(0) = 1/2$. Iš šios simetrijos išplaukia, kad plotas po grafiku į kairę nuo tiesės $x = -v (v > 0)$ yra lygus plotui po grafiku į dešinę nuo tiesės $x = v$. Taigi

$$\Phi(-v) = 1 - \Phi(v). \tag{1}$$

Iš (1) matyti, kad pakanka mokėti rasti reikšmes, kai v teigiamas. Funkcija $\Phi(v)$



4 pav.: Standartinio normaliojo skirstinio pasiskirstymo ir tankio funkcijos.

yra labai svarbi tikimybių teorijoje, todėl gerai ištyrinėta. Tikimybių teorijos knygose galite rasti jos reikšmių lenteles, jos reikšmių skaičiavimo komandas rasite bet kurioje kompiuterinėje skaičiuoklėje, turinčioje statistinių funkcijų rinkinį (pav. Excelyje ar OpenOffce skaičiuoklėje).

Muavro-Laplaso teorema apytiksliams skaičiavimams taikoma taip. Jei bandymų skaičius n yra didelis, galime naudotis lygybėmis

$$P\left(\frac{S_n - np}{\sqrt{npq}} < b\right) \approx \Phi(b),$$
$$P\left(\frac{S_n - np}{\sqrt{npq}} \geq a\right) \approx 1 - \Phi(a),$$
$$P\left(a \leq \frac{S_n - np}{\sqrt{npq}} < b\right) \approx \Phi(b) - \Phi(a).$$

7.1 Ribinės teoremos Bernulio schemoje II

Pavyzdys. Dideli lašai, smulkus tinklas

Įsivaizduokime kad lašeliai yra rutuliuko su spinduliu $r = 0,5$ cm formos, lašelių yra $n = 10000$, o tinklą sudaro $10cm \times 10cm$ dydžio kvadratėliai. Dabar subyrės daug lašelių. Kokia tikimybė, kad subyrėjusių lašelių skaičius bus tarp 1900 ir 2000? Vėl pažymėję raidėmis p, q vieno lašelio subyrėjimo ir nesubyrėjimo tikimybes gausime

$$q = \frac{(10 - 2r)^2}{10^2} = 0,9^2 = 0,81, \quad p = 1 - q = 0,19.$$

Naudotis Bernulio formule šiuo atveju dar nepatogiau:

$$P(1900 \leq S_n < 2000) = \sum_{n=1900}^{1999} C_n^m p^m q^{n-m}$$

dar nepatogiau. Taigi turėtume apskaičiuoti net 200 dėmenų! Pasinaudosime Muavro - Laplaso teorema. Kadangi $np = 1900, \quad \sqrt{npq} \approx 39,23$,

$$P(1900 \leq S_n < 2000) \approx P\left(\frac{1900-1900}{39,23} \leq \frac{S_n - np}{\sqrt{npq}} < \frac{2000-1900}{39,23}\right)$$

$$= P\left(0 \leq \frac{S_n - np}{\sqrt{npq}} < 2,549\right) \approx 0,9936 - 0,5 = 0,4936.$$

Muavro - Laplaso teorema yra atskiras centrinės ribinės teoremos atvejis. Ši teorema – tikra tikimybių teorijos viršukalnė.

Kuo panašūs ir kuo skiriasi abu pavyzdžiai? Panašūs visų pirma tuo, kad bandymų skaičius abiem atvejais didelis. O skiriasi sėkmės tikimybės reikšmėmis. Pirmajame pavyzdyje bandymų daug, o sėkmės tikimybė maža. Antrajame – bandymų daug, bet sėkmės tikimybė nėra maža. Kai bandymų daug, skaičiuojant Bernulio schemos tikimybes, galima pasinaudoti apytikslėmis formulėmis, kurias pateikia ribinės teoremos – teiginiai apie tikimybių elgesį, kai bandymų skaičius didėja.

Pavyzdys. Keleivis žygiuoja keliu taip: prieš žengdamas meta monetą, jeigu atvirsta herbas – žengia žingsnį į dešinę, jeigu skaičius – į kairę. Moneta atvirsta herbu su tikimybe p= 0,55. Kokia tikimybė, kad žengęs $n = 100$ žingsnių jis atsidurs toliau nei už penkių žingsnių į dešinę nuo kelionės pradžios taško? Kokia tikimybė, kad žengęs $n = 100$ žingsnių jis atsidurs toliau nei už penkių žingsnių į kairę nuo kelionės pradžios taško? Kokia tikimybė, kad jis nuo pradžios taško bus nutolęs ne daugiau kaip per tris žingsnius?

1. Herbas turi atvirsti dažniau nei skaičius bent 5 metimais, t.y. herbo atvirtimų skaičius X turi būti didesnis nei 53.

$$P(X \geq 53) = 1 - \Phi\left(\frac{53 - 100 \cdot 0,55}{\sqrt{100 \cdot 0,55 \cdot 0.45}}\right) = 1 - \Phi(-0,4) = 1 - (1 - \Phi(0,4)) \approx 0,655.$$

2. Herbas turi atvirsti mažiau nei skaičius bent 5 metimais, t.y. herbo atvirtimų skaičius X turi būti mažesnis nei 47.

$$P(X < 47) = \Phi\left(\frac{47 - 100 \cdot 0,55}{\sqrt{100 \cdot 0,55 \cdot 0,45}}\right) = \Phi(-1,61) =\approx 0,054.$$

3. Herbo ir skaičiaus atvirtimų skaičiaus skirtumas turi būti ne didesnis už 3, t.y. herbas atvirs nuo 49 iki 51 karto

$$P(49 \leq X < 52) = \Phi\left(\frac{52-55}{4,98}\right) - \Phi\left(\frac{49-55}{4,98}\right) = \Phi(-0,6) - \Phi(-1,2) = \Phi(1,2) - \Phi(0,6)$$
$$\approx 0,885 - 0,276 = 0,159.$$