Atsitiktinių dydžių vidurkiai ir kitos skaitinės charakteristikos

Prieš pradedant spręsti 10 užduoties uždavinius, reikėtų susipažinti su paskaitomis apie atsitiktinių dydžių skaitines charakteristikas ir jų savybes.
Trumpai priminsime pagrindines sąvokas. Tegu X yra atsitiktinis dydis, $f(x)$ - kokia nors realioji funkcija. Rasime atsitiktinio dydžio $Z := f(X)$ *vidurkį*.

- Jeigu atsitiktinis dydis X yra *diskretusis*, įgyjantis reikšmes x_k su tikimybėmis $p_k = P(X = x_k)$, tai

$$EZ = \sum_k f(x_k)\, p_k\,. \tag{1}$$

- Jeigu atsitiktinis dydis X yra *tolydusis*, turintis tankio funkciją $p(x)$, tai

$$EZ = \int\limits_{-\infty}^{\infty} f(x)\, p(x)\, dx\,. \tag{2}$$

Atskirais atvejais, kai $f(x) = x$ ir $f(x) = (x - EX)^2$, EZ yra lygus atsitiktinio dydžio X *vidurkiui* EX ir *dispersijai*

$$DX = E(X - EX)^2 = EX^2 - (EX)^2\,. \tag{3}$$

Tegu X ir Y yra atsitiktiniai dydžiai, turintys vidurkius. Tada

$$E(aX + bY) = aEX + bEY\,, \quad a, b \in \mathbb{R}.$$

Jeigu X ir Y yra nepriklausomi ir turi dispersijas, tai $EXY = EX \cdot EY$ ir

$$D(aX + bY) = a^2 DX + b^2 DY\,.$$

Kai X ir Y gali būti ir priklausomi, tada

$$D(X + Y) = DX + DY + 2\, cov(X, Y)\,,$$

čia

$$cov(X, Y) := E[(X - EX)(Y - EY)] = E[XY] - EX \cdot EY \tag{4}$$

yra atsitiktinių dydžių X ir Y *kovariacija*.
Akivaizdu, kad nepriklausomų atsitiktinių dydžių kovariacija lygi 0. Taigi kovariacijos reikšmė atspindi atsitiktinių dydžių ryšio pobūdį. Tačiau X ir Y priklausomybei nusakyti patogesnė yra "normuota" kovariacija, kuri nepriklauso nuo matavimo vienetų ir vadinama *koreliacijos koeficientu*

$$\rho(X, Y) := \frac{cov(X, Y)}{\sqrt{DX \cdot DY}} \in [-1,\ 1]\,. \tag{5}$$

Nepriklausomų atsitiktinių dydžių koreliacijos koeficientas lygus 0. Jeigu $\rho(X, Y) = \pm 1$, tai atsitiktinius dydžius X ir Y sieja tiesinė funkcinė priklausomybė.

1 pavyzdys. Standartinis lošimo kauliukas mėtomas iki k-ojo šešeto pasirodymo, $k = 1,\ 2$. Raskite vidutinį metimų skaičių abiem atvejais.
Pažymėkime X_k metimų skaičių iki k - jo šešeto. Pravers dvi formulės

$$\sum_{m=1}^{\infty} m\, x^{m-1} = \frac{1}{(1-x)^2}, \qquad \sum_{m=1}^{\infty} m(m-1)\, x^{m-2} = \frac{2}{(1-x)^3}, \quad |x| < 1. \tag{6}$$

Nesunku įsitikinti, kad

$$P(X_1 = m) = \left(\frac{5}{6}\right)^{m-1} \frac{1}{6}, \quad m = 1, 2, ...$$

Todėl, pasinaudoję (1) ir (6), gausime

$$EX_1 = \sum_{m=1}^{\infty} m\,P(X_1 = m) = \frac{1}{6}\sum_{m=1}^{\infty} m\left(\frac{5}{6}\right)^{m-1} = 6\,.$$

Analogiškai

$$P(X_2 = m) = C_{m-1}^{1}\frac{1}{6}\left(\frac{5}{6}\right)^{m-2}\frac{1}{6} = (m-1)\left(\frac{5}{6}\right)^{m-2}\frac{1}{36},\quad m = 2, 3, \ldots$$

ir

$$EX_2 = \sum_{m=1}^{\infty} m\,P(X_2 = m) = \frac{1}{36}\sum_{m=1}^{\infty} m(m-1)\left(\frac{5}{6}\right)^{m-2} = 12\,.$$

Šį uždavinį gana paprastai galima išspręsti su bet kokiu $k \geq 1$. Tegul $Y_1 = X_1$. Pažymėkime Y_n metimų skaičių po $n-1$ -jo šešeto iki n-jo, $n \geq 2$. Pastebėsime, kad $Y_1, Y_2, Y_3, \ldots$ yra vienodai pasiskirstę geometriniai atsitiktiniai dydžiai, t.y. $Y_k \sim \mathcal{G}(1/6)$. Todėl $EY_k = 6$. Akivaizdu, kad

$$X_k = Y_1 + Y_2 + \ldots + Y_k\,.$$

Taigi, remiantis vidurkio adityvumo savybe, gauname

$$EX_k = EY_1 + EY_2 + \ldots + EY_k = 6k\,.$$

2 pavyzdys. Atsitiktinio dydžio X tankio funkcija yra

$$p(x) = \begin{cases} 0, & \text{jei } x \notin [0,\ 2], \\ Cx^2, & \text{jei } x \in [0,\ 2]. \end{cases}$$

Raskite DX ir EX^3.

Pirmiausiai rasime konstantą C. Pasinaudoję tankio funkcijos savybe, gausime

$$\int_{-\infty}^{\infty} p(x)\,dx = C\int_0^2 x^2\,dx = C\,\frac{x^3}{3}\bigg|_0^2 = \frac{8C}{3} = 1.$$

Taigi $C = 3/8$ ir

$$EX = \int_{-\infty}^{\infty} x\,p(x)\,dx = \frac{3}{8}\int_0^2 x^3\,dx = \frac{3}{8}\,\frac{x^4}{4}\bigg|_0^2 = \frac{3}{2}\,.$$

Analogiškai

$$EX^2 = \int_{-\infty}^{\infty} x^2\,p(x)\,dx = \frac{3}{8}\int_0^2 x^4\,dx = \frac{12}{5}\,,\quad EX^3 = \int_{-\infty}^{\infty} x^3\,p(x)\,dx = \frac{3}{8}\int_0^2 x^5\,dx = 4\,.$$

Naudodamiesi (3), randame dispersiją

$$DX = EX^2 - (EX)^2 = \frac{12}{5} - \left(\frac{3}{2}\right)^2 = \frac{3}{20}\,.$$

3 pavyzdys. Spindulio R skritulyje atsitiktinai pasirenkamas taškas. Koks jo vidutinis atstumas iki centro?

Pažymėkime X taško atstumą iki skritulio centro. Rasime atsitiktinio dydžio X pasiskirstymo funkciją $F(x)$. Tegul $0 \leq x \leq R$. Prisiminę geometrinę tikimybę, gausime

$$F(x) = P(X < x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

Taigi

$$F(x) = \begin{cases} 0, & \text{jei } x < 0, \\ \frac{x^2}{R^2}, & \text{jei } x \in [0,\ R], \\ 1, & \text{jei } x > R. \end{cases}$$

Dabar jau nesunkiai randame atsitiktinio dydžio X tankio funkciją $p(x)$ ir vidurkį EX

$$p(x) = F'(x) = \begin{cases} \frac{2x}{R^2}, & \text{jei } x \in [0,\ R], \\ 0, & \text{jei } x \notin [0,\ R]. \end{cases}$$

$$EX = \int_{-\infty}^{\infty} x\,p(x)\,dx = \frac{2}{R^2}\int_0^R x^2\,dx = \frac{2}{R^2}\,\frac{x^3}{3}\bigg|_0^R = \frac{2}{3}R.$$

4 pavyzdys. Atsitiktinis dydis X pasiskirstęs pagal normalųjį dėsnį $X \sim \mathcal{N}(-3;4)$. Raskite atsitiktinio dydžio $Y = 5X^2 - X + 7$ vidurkį.

Pagal vidurkio savybes

$$EY = 5\,EX^2 - EX + 7.$$

Yra žinoma, kad $EX = -3$ ir $DX = 4$. Pasinaudoję (3) lygybe, gausime

$$EX^2 = DX + (EX)^2 = 4 + (-3)^2 = 13.$$

Taigi

$$EY = 5 \cdot 13 + 3 + 7 = 75.$$

5 pavyzdys. Tegul X ir Y yra atsitiktiniai dydžiai, o jų bendrasis dvimatis skirstinys nusakytas lentele

$Y \backslash X$	-1	2	$P(Y=j)$
1	$0,3$	$0,1$	$0,4$
3	0	$0,6$	$0,6$
$P(X=i)$	$0,3$	$0,7$	1

Raskite X ir Y koreliacijos koeficientą $\rho(X,Y)$.

Kovariacija ir koreliacijos koeficientas skaičiuojami naudojant (4) ir (5) formules. Duotasis dvimatis skirstinys nusako ir vienmačius X ir Y skirstinius. Taigi

$$\begin{aligned} &EX = -1 \cdot 0,3 + 2 \cdot 0,7 = 1,1; \quad EX^2 = (-1)^2 \cdot 0,3 + 2^2 \cdot 0,7 = 3,1; \\ &EY = 1 \cdot 0,4 + 3 \cdot 0,6 = 2,2; \quad EY^2 = 1^2 \cdot 0,4 + 3^2 \cdot 0,6 = 5,8; \\ &DX = 3,1 - 1,1^2 = 1,89; \quad DY = 5,8 - 2,2^2 = 0,96. \end{aligned}$$

Vidurkis $E[XY]$ skaičiuojamas pagal turimą dvimatį skirstinį:

$$E[XY] = (-1) \cdot 1 \cdot 0,3 + 2 \cdot 1 \cdot 0,1 + (-1) \cdot 3 \cdot 0 + 2 \cdot 3 \cdot 0,6 = 3,5.$$

Dabar jau randame kovariaciją ir koreliacijos koeficientą

$$\begin{aligned} cov(X,Y) &= E[XY] - EX \cdot EY = 3,5 - 1,1 \cdot 2,2 = 1,08; \\ \rho(X,Y) &= \frac{cov(X,Y)}{\sqrt{DX \cdot DY}} = \frac{1,08}{\sqrt{1,89 \cdot 0,96}} \approx 0.80178. \end{aligned}$$