3 Paskaita. Sąlyginės tikimybės ir nepriklausomi įvykiai

3.1 Sąlyginė tikimybė

Pavyzdys. Metame kauliuką. Žinome, kad iškritusių akučių skaičius yra lyginis, tačiau tikslaus iškritusio skaičiaus nematome. Kokia tikimybė, kad iškritusių akučių skaičius yra nedidesnis nei 5?

Nagrinėkime elementariųjų įvykių erdvę $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6\}, P(w_i) = 1/6, 1 \leq i \leq 6$. Įvykis $A = \{w_2, w_4, w_6\}$ – "iškritusių akučių skaičius lyginis", įvykis $B = \{w_1, w_2, w_3, w_4, w_5\}$ – iškritusių akučių skaičius neviršija 5. Turime apskaičiuoti įvykio B tikimybę, jei žinome (papildomai), kad įvyko A. Įvykus įvykiui A, žinome, kad galimos baigtys yra $\{w_2, w_4, w_6\}$. Kadangi daugiau informacijos apie eksperimento rezultatą neturime, visos trys baigtys yra vienodai tikėtinos. Tarp jų mums plankūs (t.y., priklausantys aibei B) yra w_2, w_4. Jie sudaro $2/3$ visų galimų baigčių. Sakome, kad sąlyginė įvykio B tikimybė su sąlyga (kai yra įvykęs įvykis) A yra $2/3$. Žymime $P(B|A) = 2/3$.

Toliau laikysime, kad nagrinėjami įvykiai yra tikimybinės erdvės $(\Omega, \mathcal{A}, P)$ įvykių $\sigma-$algebros $\mathcal{A}$ elementai.

Apibrėžimas. *Tarkime $A, B \in \mathcal{A}$ ir $P(A) > 0$. Įvykio B sąlygine tikimybe su sąlyga A vadinsime santykį*

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Pavyzdyje $P(A \cap B) = \frac{2}{6} = \frac{1}{3}, P(A) = \frac{1}{2}$. Apskaičiavę sąlyginę tikimybę pagal apibrėžimą, gausime atsakymą $2/3$.

Pastabos. Teisingos lygybės

$$P(\Omega|A) = 1, \text{ nes } P(\Omega|A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1,$$

$$P(A|A) = 1, \text{ nes } P(A|A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

Teorema. Tikimybių sandaugos teorema. *Tarkime, kad įvykiai $A_i \in \mathcal{A}, i = 1, 2, \ldots, n$ yra tokie, kad $P(A_1 \cap A_2 \cap \ldots \cap A_n) > 0$. Tuomet teisinga lygybė*

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap \ldots \cap A_{n-1}). \tag{1}$$

Kai $n = 2$ iš sąlyginės tikimybės apibrėžimo turime

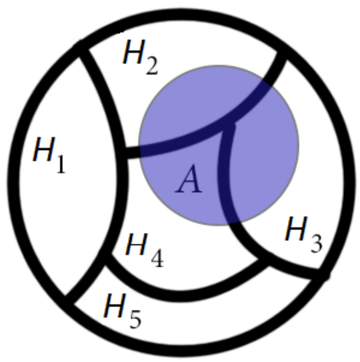
$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1).$$

Teoremos įrodymas remiasi matematinės indukcijos principu.

Pavyzdys. Piniginėje yra 40 monetų. Tarp jų 10 padirbtų. Iš eilės traukiame 4 monetas. Traukdami monetą laikome, kad visos tuo metu piniginėje esančios monetos turi vienodus šansus būti ištrauktos. Kokia tikimybė, kad pirma ištraukta moneta yra tikra, antra netikra, trečia ir ketvirta abi yra tikros?

Sprendimas. Pažymime įvykius: A_1 – pirma moneta tikra, A_2 - antra moneta netikra, A_3 – trečia moneta tikra, A_4 – ketvirta moneta tikra. Mus dominantis įvykis yra $B = A_1 \cap A_2 \cap A_3 \cap A_4$. Aišku, kad $P(A_1) = 30/40$. Ištraukus tikrą monetą, lieka 39 monetos, tarp kurių 10 padirbtų. Todėl tikimybė, kad antroji moneta yra netikra, lygi $P(A_2|A_1) = 10/39$. Tikimybė, kad trečioji moneta tikra, žinant, jog pirmoji buvo tikra, o antroji padirbta, yra $P(A_3|A_1 \cap A_2) = 29/38$. Tikimybė, kad ketvirtoji moneta tikra, jei primuoju ir trečiuoju traukimu papuolė tikros monetos, o antruoju netikra, yra $P(A_3|A_1 \cap A_2) = 28/37$. Pritaikę (1) formulę, gauname

$$P(B) = P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{30}{40} \cdot \frac{10}{39} \cdot \frac{29}{38} \cdot \frac{28}{37} \approx 0,111.$$



1 pav.: Pilnosios tikimybės formulės Veno diagrama.

Teorema. (Pilnosios tikimybės formulė). *Tarkime, kad įvykiai $H_i \in \mathcal{A}, i = 1, 2, \ldots, k$ yra tokie, kad $H_i \cap H_j = \emptyset$ visiems $i \neq j$. Jei rinkinys $H_1, H_2, \ldots H_k$ tenkina nelygybes $P(H_i) > 0, i = 1, 2, \ldots, k$ ir $H_1 \cup H_2 \cup \ldots \cup H_k = \Omega$, tai bet kuriam įvykiui $A \in \mathcal{A}$ teisinga lygybė*

$$P(A) = \sum_{i=1}^{k} P(H_i) P(A|H_i). \tag{2}$$

Jei skaitus rinkinys $H_1, H_2, \ldots$ tenkina nelygybes $P(H_i) > 0, i = 1, 2, \ldots$ ir $\cup_{i=1}^{\infty} H_i = \Omega$, tai bet kuriam įvykiui $A \in \mathcal{A}$ teisinga lygybė

$$P(A) = \sum_{i=1}^{\infty} P(H_i) P(A|H_i). \tag{3}$$

Įrodymas. Iš aibių lygybės

$$A = A \cap \Omega = A \cap (\cup_{i \geq 1} H_i) = \cup_{i \geq 1}(A \cap H_i)$$

ir tikimybinio mato adityvumo ($\sigma-$adityvumo) išplaukia

$$P(A) = \sum_{i \geq 1} P(A \cap H_i),$$

nes $(A \cap H_i) \cap (A \cap H_j) = \emptyset$, kai $i \neq j$. Iš sąlyginės tikimybės apibrėžimo $P(A \cap H_i) = P(H_i)P(A|H_i)$. Iš čia išplaukia (2), jei aibių rinkinys baigtinis ir (3), jei aibių rinkinys skaitus.

Pavyzdys. Turime 5 dėžes su vaisiais. Dvi dėžės turi po 2 obuolius ir kriaušę, kitos dvi dėžės turi po 3 obuolius ir kriaušę. Penktoji dėžė turi 10 kriaušių. Sakysime, kad turime po dvi dėžes pirmosios ir antrosios rūšies ir vieną dėžę trečiosios rūšies. Atsitiktinai parinkę dėžę iš jos atsitiktinai išrinkome vaisių. Kokia tikimybė, kad tai obuolys? Laikome, kad visos dėžės turi vienodus šansus būti išrinktomis.

Sprendimas. Pažymėkime įvykius $H_i, i = 1, 2, 3$ – "parinkta i-os rūšies dėžė", A – "ištrauktas obuolys". Įvykių H_i tikimybės $P(H_1) = P(H_2) = 2/5, P(H_3) = 1/5$. Tikimybė ištraukti obuolį iš pirmos rūšies dėžės yra $P(A|H_1) = 2/3$, iš antros rūšies dėžės – $P(A|H_2) = 3/4$, iš trečiosios rūšies dėžės – $P(A|H_3) = 0$. Pritaikę formulę (2), gauname

$$P(A) = P(A|H_1)P(H_1)+P(A|H_2)P(H_2)+P(A|H_3)P(H_3) = \frac{2}{3}\cdot\frac{2}{5}+\frac{3}{4}\cdot\frac{2}{5}+0 = \frac{17}{30}.$$

Teiginys. (Bajeso formulė).*Tarkime, kad įvykiai $A, B \in \mathcal{A}$ ir $P(A) > 0, P(B) > 0$. Tuomet teisinga lygybė*

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}. \tag{4}$$

Įrodymas. Įrodymas išplaukia iš sąlyginės tikimybės apibrėžimo

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A).$$

Teorema. (Bajeso teorema).*Tarkime, kad įvykiai $H_i \in \mathcal{A}, i = 1, 2, \dots$ yra tokie, kad $H_i \cap H_j = \emptyset$ visiems $i \neq j$ ir $P(H_i) > 0$ kiekvienam i. Be to, $\cup_{i \geq 1} H_i = \Omega$. Tuomet bet kuriam įvykiui $A \in \mathcal{A}, P(A) > 0$ ir bet kuriam k teisinga lygybė*

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{\sum_{i \geq 1} P(H_i)P(A|H_i)}. \tag{5}$$

Įrodymas. Įrodymas išplaukia iš pilnosios tikimybės formulės (3) ir Bajeso formulės (4).

Teoremoje esančios formulės (5) vadinamos Bajeso arba hipotezių tikrinimo formulėmis. Įvykiai H_i vadinami hipotezėmis, tikimybės $P(H_i)$ – apriorinėmis hipotezių tikimybėmis, $P(H_i|A)$ – aposteriorinėmis (t.y. po bandymo) hipotezių tikimybėmis. Formulėmis (5) tikslinga naudotis, kai reikia pasirinkti vieną iš kelių alternatyvų arba hipotezių. Tarkime, kad žinome, jog įvyko vienas įvykis iš nesutaikomų įvykių šeimos $H_i, i \in I$, t.y. teisinga viena iš kelių hipotezių. Kuris iš įvykių įvyko mes nežinome, tačiau turime papildomos informacijos, kad įvyko įvykis A. Tarkime, reikia nuspręsti, kuria hipoteze H_i vadovautis priimant sprendimą apie tolesnius veiksmus. Mažiausia tikimybė suklysti bus tada, jei savo sprendimą grįsime ta hipoteze, kurios aposteriorinė tikimybė $P(H_i|A)$ didžiausia.

Pavyzdys. Turime 5 dėžes su rutuliais. Dvi dėžės turi po 2 baltus ir 3 juodus rutulius, kitos dvi dėžės turi po 1 baltą ir 4 juodus rutulius. Penktoji dėžė turi 4 baltus ir 1 juodą rutulį. Atsitiktinai parinkę dėžę (visos penkios dėžės vienodai tikėtinos) iš jos atsitiktinai traukiamas rutulys (visi dėžės rutuliai turi vienodus šansus būti ištrauktais). Kokia tikimybė, kad pasirinkta dėžė buvo trečios rūšies, jei žinome, kad ištrauktas rutulys buvo baltas?

Sprendimas. Pažymėkime įvykius $H_i, i = 1, 2, 3$ – "pasirinkta i-os rūšies dėžė", A – "ištrauktas baltas rutulys". Reikia apskaičiuoti tikimybę $P(H_3|A)$. Pasinaudoję Bajeso teorema, gauname

$$\begin{aligned} P(H_3|A) &= \frac{P(H_3)P(A|H_3)}{P(A|H_1)P(H_1) + P(A|H_2)P(H_2) + P(A|H_3)P(H_3)} \\ &= \frac{\frac{1}{5} \cdot \frac{4}{5}}{\frac{2}{5} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{1}{5} + \frac{1}{5} \cdot \frac{4}{5}} = \frac{2}{5}. \end{aligned}$$

Pavyzdys. Šeimoje yra 2 vaikiai. Kokia tikimybė, kad antrasis vaikas yra berniukas, jei žinome, kad a) pirmasis yra berniukas, b) mažiausiai vienas yra berniukas? Laikysime, kad berniuko ir mergaitės gimimo tikimybės vienodos.

Sprendimas. Elementariųjų įvykių erdvė $\Omega = \{B_1B_2, B_1M_2, M_1B_2, M_1M_2\}$.

1. $P(B_2|B_1) = \dfrac{P(B_2 \cap B_1)}{P(B_1)} = \frac{1/4}{1/2} = \dfrac{1}{2}.$

2. $P(B_2|B_1 \cup B_2) = \dfrac{P(B_2 \cap (B_1 \cup B_2))}{P(B_1 \cup B_2)} = \dfrac{P(B_2)}{P(B_1 \cup B_2)} = \frac{1/4}{3/4} = \dfrac{1}{3}.$

3.2 Nepriklausomi įvykiai. Bernulio schema

Tarkime, $(\Omega, \mathcal{A}, P)$ yra tikimybinė erdvė, $A, B \in \mathcal{A}$. Jei tikimybės $P(A)$ ir $P(A|B)$ skiriasi, natūralu manyti, jog įvykiai A ir B priklauso vienas nuo kito, o jei šios tikimybės sutampa – kad jie yra nepriklausomi. Iš lygybės $P(A) = P(A|B)$ ir sąlyginės tikimybės apibrėžimo išplaukia $P(A \cap B) = P(A)P(B)$.

Apibrėžimas. *Atsitiktinius įvykius A ir B vadinsime nepriklausomais, jei*

$$P(A \cap B) = P(A)P(B). \tag{6}$$

Pavyzdys. Mėtome lošimo kauliuką. Pažymėkime įvykius A – "iškritusių akučių skaičius dalijasi iš 2", A – "iškritusių akučių skaičius dalijasi iš 3", C –

"iškritusių akučių skaičius didesnis už 2". Nesunku patikrinti, kad įvykiai A, B ir A, C yra nepriklausomi, o įvykiai B, C – priklausomi, nes

$$A = \{w_2, w_4, w_6\}, B = \{w_3, w_6\}, C = \{w_3, w_4, w_5, w_6\}.$$

$$P(A) = 1/2, P(B) = 1/3, P(C) = 2/3.$$

$$P(A \cap B) = 1/6, P(A \cap C) = 1/3, P(B \cap C) = 1/3.$$

Kaip suformuluoti kelių įvykių nepriklausomumą? Tarkime, kad turime tris įvykius. Galbūt pakanka pareikalauti, kad bet kurie du įvykiai iš trijų būtų nepriklausomi? Žemiau esantis pavyzdys rodo, kad to nepakanka.

Pavyzdys. Tarkime, kad ruletės skritulys padalytas į keturis ketvirčius, kurie sužymėti 0, 1, 2, 3, lošėjų yra trys. Jei ruletės strėlė sustoja ties ketvirčiu 0, laimi visi, jei 1 – tik pirmasis ir t.t. Nagrinėkime įvykius A_1, A_2, A_3, kurie atitinkamai reiškia, kad laimėjo pirmasis, antrasis ir trečiasis. $P(A_1) = P(A_2) = P(A_3) = 1/2$. Nesunku įsitikinti, kad kiekviena pora iš šio įvykių trejeto yra nepriklausoma. Tačiau, jeigu žinome, kad įvyko įvykiai A_1 ir A_2, iš to iš karto galima padaryti išvadą, kad įvyko ir A_3. Taigi sakyti, kad visi trys įvykiai tarpusavyje nepriklausomi, negalime.

Nagrinėkime įvykių rinkinį $\mathcal{U} \subset \mathcal{A}$.

Apibrėžimas. *Sakome, kad rinkinio $\mathcal{U}$ įvykiai yra nepriklausomi, jei bet kuriems rinkinio $\mathcal{U}$ elementams $A_1, A_2, \ldots, A_k \in \mathcal{U}$ yra teisinga lygybė*

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_1)P(A_2)\ldots P(A_k). \tag{7}$$

Teiginys. *Bet kuriam atsitiktiniam įvykiui $A \in \mathcal{A}$ teisingi šie teiginiai*

1. *Įvykiai A ir $\emptyset$ yra nepriklausomi;*
2. *Įvykiai A ir Ω yra nepriklausomi;*
3. *Jei įvykiai A ir B yra nepriklausomi, tai įvykiai A ir $\overline{B}$ yra nepriklausomi.*

Pažymėkime $A^1 = A, A^0 = \overline{A}$. Toliau įrodysime, kad įvykių poros nepriklausomumui nedaro įtakos, jei vieną ar kitą įvykį arba juos abu keičiame priešingais.

Teorema. *Jei egzistuoja indeksai i_0, j_0, kad įvykiai A^{i_0}, B^{j_0} yra nepriklausomi, tai A^i, B^j yra nepriklausomi su bet kokia indeksų pora $i, j \in \{0, 1\}$.*

Įrodymas. Pakaks parodyti, kad iš įvykių A ir B yra nepriklausomumo išplaukia, kad įvykiai A ir $\overline{B}$ yra nepriklausomi.

$$P(\overline{A} \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = P(B)(1 - P(A)) = P(\overline{A}).$$

Teorema. Borelio-Cantelli lema. *Jei sistemą $\mathcal{S} = \{A_\lambda \in \mathcal{A}, \lambda \in \Lambda\}$, sudaro tarpusavyje nepriklausomi įvykiai, tai su bet kokiais $i_\lambda \in \{0, 1\}$ sistemą $\mathcal{S}* = \{A_\lambda^{i_\lambda}, \lambda \in \Lambda\}$ taip pat sudaro tarpusavyje nepriklausomi įvykiai.*

Pavyzdys. Tikimybė, kad šaulys pataikys yra 1/3. Kokia tikimybė, kad iššovus 10 kartų visi šūviai bus sėkmingi? Kokia tikimybė, kad bent vienas šūvis bus sėkmingas? Laikome, kad skirtingų šūvių rezultatai nepriklauso vienas nuo kito.

Sprendimas. Įvykį A – "visi 10 šūvių sėkmingi" atitinka įvykių sankirta $A = \cap_{i=1}^{10} A_i$, kur A_i žymi įvykį, kad i-asis šūvis sėkmingas. Kadangi šūvių rezultatai nepriklausomi, tai įvykiai $A_1, \ldots, A_{10}$ sudaro nepriklausomų įvykių rinkinį. Todėl

$$P(A) = P(\cap_{i=1}^{10} A_i) = P(A_1)P(A_2)\ldots P(A_{10}) = (1/3)^{10}.$$

Įvykio B – "bent vienas šūvis sėkmingas" tikimybę patogu skaičiuoti tokiu būdu. Pradžioje randame jam priešingo įvykio B – "visi 10 šūvių nesėkmingi" tikimybę. Tuomet mus dominanti tikimybė $P(B) = 1 - P(\overline{B})$. Įvykį $\overline{B}$ atitinka įvykių sankirta $\overline{B} = \cap_{i=1}^{10} B_i$, kur B_i žymi įvykį, kad i-asis šūvis nesėkmingas. Kadangi šūvių rezultatai nepriklausomi, tai įvykiai $B1, \ldots, B_{10}$ sudaro nepriklausomų įvykių rinkinį. Todėl

$$P(\overline{B}) = P(\cap_{i=1}^{10} B_i) = P(B_1)P(B_2)\ldots P(B_{10}) = (2/3)^{10}, P(B) = 1 - (2/3)^{10}.$$