Verjetnost 1

Napisal: Jon Pascal Miklavčič

Predgovor

Zapiski v tej skripti so bili osnovanih na rokopisih predavanj iz študijskih let pred letom 2023 in doknčno dopolnjeni z zapiski iz predavanj iz leta 2024.

Kazalo

1	Neformalni uvod v verjetnost					
2	2 Aksiomatična definicija verjetnosti					
3	B Pogojna verjetnost					
4	Zaporedja neodvisnih ponovitev poskusa	17				
5	Slučajne spremenljivke					
	5.1 Diskretne slučajne spremenljivke	24				
	5.2 Zvezne slučajne spremenljivke	27				

Poglavje 1

Neformalni uvod v verjetnost

Začetki verjetnosti so v 17. stoletju, iz iger na srečo (kartanje, kockanje, ...):

- 17. stol.: Fermant, Pascal, Bernulli;
- 18./19. stol.: Laplace, Poisson, Čebišev, Markov;
- 20. stol.: Kolmogorov.

Izvajamo poskus in opazujemo določen pojav, ki ga imenujemo *dogodek*. Ta se lahko zgodi ali ne.

Zgled. Poskus je met kocke. Da pade šestica, da pade sodo število pik pa sta dogodka.

Poskus ponovimo n-krat. Opazujemo dogodek A. S $k_n(A)$ označimo frekvenco dogodka A, t.j. število tistih ponovitev poskusa, pri katerih se je dogodek A zgodil. Naj bo $f_n(A) = \frac{k_n(A)}{n}$ relativna frekvenca dogodka A. Dokazati je mogoče, da zporedje $\{f_n(A)\}_n$ konvergira k nekemu številu $p \in [0,1]$; $f_n(A) \xrightarrow{n \to \infty} p$. Dobimo:

Statistično definicijo verjetnosti:

$$P(A) := p$$

Pogosto lahko verjetnost določimo vnaprej in sicer s:

Klasično definicijo verjetnosti:

$$P(A) := \frac{\# \text{ ugodnih izidov za dogodek } A}{\# \text{ vseh izidov}}$$

pri pogoju, da imajo vsi izidi enake možnosti.

Zgled. Met poštene kocke:

$$P(\text{sodo število pik}) = \frac{3}{6} = \frac{1}{2}$$

 \Diamond

Zgled. Kolikšna je verjetnost, da pri metu dveh poštenih kock znaša vsota pik 7? Možne vsote so: $2, 3, 4, \ldots, 12$. Opazimo, da je vseh vsot 11 in od tega 1 ugodna. Ali to pomeni, da $P(A) = \frac{1}{11}$. Ne! Izidi niso enkaoverjetni.

Na primer 2 lahko dobimo samo kot 2 = 1 + 1, 5 pa kot 5 = 2 + 3 = 1 + 4 = 4 + 1 = 3 + 2.

Torej vsi možni izidi, bodo urejeni pari (x, y), kjer $x, y \in [1, 6] \subset \mathbb{N}$

Vseh izidov je torej 36 in od tega je 6 ugodnih. Torej $P(A) = \frac{6}{36} = \frac{1}{6}$.

Če je izidov neskončno, si lahko pomagamo s **Geometrijsko definicijo verjetno**sti.

Zgled. Osebi se dogovorita za srečanje med 10. in 11. uro. Čas prihoda je slučajen. Vsak od njiju po prihodu čaka največ 20 minut. Če v tem času drugega ni, odide. Najdlje čaka do 11. ure. Kolišna je vrejetnost srečanja?

Čas začnemo šteti ob 10. uri. Vsi izidi so urejeni pari $(x,y) \in [0,1] \times [0,1]$. Ugodni izidi so $|x-y| \leq \frac{1}{3}$. Torej:

1)
$$x \ge y : x - \frac{1}{3} \le y$$

2) $x \le y : y - x \le \frac{1}{3} \iff y \le x + \frac{1}{3}$

Torej je

$$P(\text{srečanja}) = \frac{1 - (\frac{2}{3})^2}{1} = \frac{5}{9}$$

 \Diamond

Teorija mere se ukvarja z splošnim zapisom geometrijske definicije.

Zgled. Vzamemo $m, n \in \mathbb{N}$, m > n. n krogljic slučajno razporedimo v m posod. Kolikšna je verjetnost dogodka, da so vse krogljice v prvih n posodah, v vsaki ena?

To je pomankljivo zastavljena naloga. Ne vemo namreč, ali med seboj krogljice razlikujemo, ali ne. Za dodatno predpostavko se ponujajo 3 možnosti:

1) krogljice razlikujemo:

Število vseh izidov v tem primeru je ravno število $variacij\ m$ elementov na n mestih $s\ ponavljanjem$. Za vsako od n-tih kroglic imamo m možnosti, torej je vseh možnosti $m \cdot m \cdots m = m^n$.

Število ugodnih izidov pa je ravno število permutacij n krogljic v prvih n posodah. Torej je ugodnih možnosti $n(n-1)\dots 2\cdot 1=n!$.

Torej je

$$P(A) = \frac{n!}{m^n}$$

2) krogljic ne razlikujemo:

V vsaki posodi je lahko več krogljic. Število vseh izidov je ravno število kombinacij s ponavljanjem. Število kombinacij m elementov s ponavljanjem na n mestih je:

$$\binom{n+m-1}{n} = \binom{n+m-1}{m-1}$$

Postavimo n krogljic in med njih razporedimo m-1 črtic, ki predstavljajo stene posod:

$$\left|\underbrace{\circ \left| \circ \left| \circ \circ \right| \circ \circ \cdots \circ \circ \right|}_{n \text{ krogljic, } m-1 \text{ črtic}}\right|$$

Na n+m-1 mestih moramo določiti n krogljic. Ugoden izid je samo eden:

Torej je

$$P(A) = \frac{1}{\binom{n+m-1}{n}}$$

3) krogljic ne razlikujemo, v vsaki posodi je kvečjemu ena krogljica:

Število vseh izidov je ravno število kombinacij brez ponavljanja $\binom{m}{n}$. Ugoden izid je eden.

Torej je

$$P(A) = \frac{1}{\binom{m}{n}}$$

POGLAVJE 1. NEFORMALNI UVOD V VERJETNOST

Opomba. V fiziki so krogljice delici (atomi, molekule, ...), posode pa fazna stanja, v katerih so lahko delci. Glede na zgornje primere ločimo:

- 1. Maxwell-Boltzmannovo statistika, ki velja za molekule plina.
- 2. Bose-Einsteinovo statistika, ki velja za delce imenovane bozoni.
- 3. Fermi-Diracovo statistika, ki velja za fermione.

Diracovo izključitveno načelo.

Poglavje 2

Aksiomatična definicija verjetnosti

Imamo prostor vseh izidov oz. $vzorčni \ prostor \ \Omega$ (možna oznaka je tudi \mathcal{G}). Dogodki so nekatere (ne nujno vse) podmnozice Ω .

Zgled. Met kocke. Vzorčni prostor je $\Omega = \{1, 2, 3, 4, 5, 6\}$, dogodki pa so poljubne podmnožice Ω , to je $\mathcal{P}(\Omega) = 2^{\Omega}$. Na primer $A = \{2, 4, 6\}$ je dogodek, da pade sodo število pik.

Računanje z dogodki:

1. Vsota dogodkov oz. unija dogodkov (zgodi se vsaj enden od dogodkov):

$$A + B = A \cup B$$

2. Produkt dogodkov oz. presek dogodkov (zgodita se oba dogodka hkrati):

$$A \cdot B = A \cap B$$

3. Nasprotni dogodek oz. komplement dogodka (dogodek se ne zgodi):

$$\bar{A} = A^c$$

Pravila za računanje z dogodki:

1. idempotentnost: $A \cup A = A = A \cap A$

2. **komutativnost**: $A \cup B = B \cup A$, $A \cap B = B \cap A$

3. asociativnost:

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

4. distributivnost:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

5. de Morganova zakona:

$$(A \cap B)^c = A^c \cup B^c$$
$$(A \cup B)^c = A^c \cap B^c$$

Še več:

$$\left(\bigcap_{i} A_{i}\right)^{c} = \bigcup_{i} A_{i}^{c}, \quad \left(\bigcup_{i} A_{i}\right)^{c} = \bigcap_{i} A_{i}^{c}$$

V splošnem ni vsaka podnožica množice Ω dogodek. Neprazna družina podmnožic (dogodkov) \mathcal{F} v Ω je σ -algerba, če zanjo velja:

- 1. $\Omega \in \mathcal{F}$
- 2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

3.
$$A_1, A_2, \ldots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$$

Elementi v \mathcal{F} so dogodki. Če v točki 3. zahtevamo manj:

$$3.* A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$$

potem pravimo, da je \mathcal{F} algebra.

V algebri imamo potem tudi zaprtost za končne unije: $A_1, A_2, \ldots, A_n \in \mathcal{F} \implies A_1 \cup A_2 \cup \cdots \cup A_n \in \mathcal{F}$. Ker po de-Morganu velja $\bigcap_i A_i = (\bigcup_i A_i^c)^c$, je algebra zaprta za končne preseke, σ -algebra pa celo za števne preseke. Ker velja $A \setminus B = A \cap B^c$, je algebra zaprta za razlike.

Vsaka algebra vsebuje $\{\emptyset, \Omega\}$. Ker je \mathcal{F} neprazna, obstaja $A \in \mathcal{F}$ in zato tudi $\Omega = A \cup A^c \in \mathcal{F}$ in $\emptyset = \Omega^c \in \mathcal{F}$. Tako dobimo, da je $\{\emptyset, \Omega\}$ najmanšja možna $(\sigma$ -)algebra, $\mathcal{P}(\Omega)$ pa največja možna $(\sigma$ -)algebra.

Zgled. Za $A \neq \emptyset \neq \Omega$ je najmanjša $(\sigma$ -)algebra, ki vsebuje A enaka $\{\emptyset, A, A^c, \Omega\}$. Za $\Omega = \{1, 2, 3\}$ in $A = \{1, 2\}$, je potem taka σ -algebra $\{\emptyset, \{3\}, \{1, 2\}, \{1, 2, 3\}\}$. \Diamond

Dogodka A in B sta disjunkta oz. nezdružljiva če je $A \cap B = \emptyset$.

Zaporedje $\{A_i\}_i$ (končno ali števno mnogo) je popoln sistem dogodkov, če velja:

$$\bigcup_{i} A_i = \Omega \quad \text{in} \quad A_i \cap A_j = \emptyset \quad \text{za} \quad i \neq j.$$

Naj bo \mathcal{F} σ -algebra na Ω . Verjetnost na (Ω, \mathcal{F}) je preslikava $P: \mathcal{F} \to \mathbb{R}$ z lastnostmi:

- 1. Za vsak $A \in \mathcal{F}$: P(A) > 0
- 2. $P(\Omega) = 1$
- 3. Za poljubne paroma nezdružljive dogodke $\{A_i\}_{i=1}^{\infty}$ velja števna aditivnost:

$$P\left(\bigcup_{j=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$$

Lastnosti verjetnosti P:

(a) $P(\emptyset) = 0$.

Dokaz. V lastnosti 3. vzamamo $A_i = \emptyset$ za vsak i:

$$P\left(\bigcup_{i}\emptyset\right) = P(\emptyset) + P(\emptyset) + P(\emptyset) + \dots = 0 + 0 + 0 + \dots = 0$$

(b) P je $končno\ aditivna$, t.j. za končno mnogo paroma nezdružljivih dogodkov $\{A_i\}_{i=1}^n$ velja:

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

Dokaz. V lastnosti 3. vzamemo $A_{n+1}=A_{n+2}=\cdots=\emptyset$ in upoštevamo lastnost (a).

(c) P je monotona, t.j. velja:

$$A \subseteq B \implies P(A) \le P(B)$$

Še več: iz $A \subseteq B$ sledi $P(B \setminus A) = P(B) - P(A)$.

Dokaz. Ker je $B = A \cup (B \setminus A)$ in $A \cap (B \setminus A) = \emptyset$ je $P(B) = P(A) + P(B \setminus A)$ zaradi lastnosti (b).

(d) $P(A^c) = 1 - P(A)$

Dokaz. V (c) vzamemo $B = \Omega$.

(e) P je zvezna, t.j.:

(i)
$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \implies P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \to \infty} P\left(A_n\right)$$

(ii)
$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots \implies P\left(\bigcap_{i=1}^{\infty} B_i\right) = \lim_{n \to \infty} P\left(B_n\right)$$

Dokaz. (i) Definiramo $C_1 = A$ in $C_i = A_i \setminus A_{i-1}$ za $i = 2, 3, \ldots$ Potem je $A_n = C_1 \cup \ldots \cup C_n$, kjer velja $C_i \cap C_j = \emptyset$ za $i \neq j$ in $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} C_i$. Torej je:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} C_i\right)$$

$$= \sum_{i=1}^{\infty} P(C_i)$$

$$= \lim_{n \to \infty} \sum_{i=1}^{n} P(C_i)$$

$$= \lim_{n \to \infty} P\left(\bigcup_{i=1}^{n} C_i\right)$$

$$= \lim_{n \to \infty} P(A_n)$$

(ii) Ker $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$, je potem $B_1^c \subseteq B_2^c \subseteq B_3^c \subseteq \cdots$. Po (i) potem velja

$$P\left(\bigcup_{i=1}^{c} B_i^c\right) = \lim_{i \to \infty} P\left(B_i^c\right)$$

Toda

$$\bigcup_{i=1}^{\infty} B_i^c = \left(\bigcap_{i=1}^{\infty} B_i\right)^c \implies 1 - P\left(\bigcap_{i=1}^{\infty} B_i\right) = \lim_{i \to \infty} \left(1 - P(B_i)\right)$$

Od koder sledi željena enakost.

Verjetnostni prostor je trojica (Ω, \mathcal{F}, P)

Zgled (Končni ali števni verjetnostni prostor). $\Omega = \{\omega_2, \omega_2, \omega_3, \ldots\}$ končno ali števno mnogo izidov. $\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \ldots$ je popoln sistem dogodkov, neka podmnožica v Ω je končna ali števna unija teh dogodkov. Torej $\mathcal{F} = \mathcal{P}(\Omega)$. Vzamemo:

$$A = \bigcup_{i:\omega_i \in A} \{\omega_0\}$$

Če označimo $P(\{\omega_i\}) = p_i \ge 0$ je $\sum_i p_i = 1$ in $P(A) = \sum_{i:\omega_i \in A} p_i$, $A \subseteq \Omega$. Če ima Ω n elementov in $p_i = \frac{1}{n}$ za $i = 1, 2, \ldots, n$. Potem je $P(A) = \frac{|A|}{n} = \frac{\text{moč}(A)}{n}$. To je klasična definicija verjetnosti.

Zgled (Neskončni neštevni verjetnostni prostor). Primer srečanja dveh oseb, kjer $\Omega = [0,1] \times [0,1]$. Za σ -algebro \mathcal{F} ne moremo vzeti vseh podmnožic, radi pa bi jih vzeli čim več.

 \mathcal{F} naj bo najmanjša σ -algebra, ki vsebje vse odprte pravokotnike $(a,b)\times(c,d)$ (izkaže se, da je isto, če vzamemo zaprte pravokotnike). \mathcal{F} imenujemo Borelova σ -algebra.

Verjetnost definiramo na pravokotnikih kot:

$$P((a,b) \times (c,d)) = (b-a)(d-c)$$

Ni lahko videti, da lahko P razširimo do verjetnosti na \mathcal{F} . P pa ne moremo razširiti na $\mathcal{P}(\Omega)$. Problem je števna aditivnost.

To je geometrijska definicija verjetnosti.

 \Diamond

Poglavje 3

Pogojna verjetnost

Naj bo (Ω, \mathcal{F}, P) verjetnostni prostor. Fiksirajmo dogodek B s P(B) > 0. Pogojna verjetnost dogodka A glede na dogodek B je:

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

Zgled. V posodi sta dve beli in ena črna krogljica. Dvakrat zaporedoma izvlečemo krogljico. Kolikšna je verjetnost, da smo drugič izbrali belo krogljico, če smo prvič izbrali belo?

(a) Krogljice vrečamo.

Vseh izidov je 9:

$$B_1B_1$$
 B_1B_2 B_1 Č
 B_2B_1 B_2B_2 B_2 Č
Č B_1 Č B_2 ČČ

$$\begin{split} &P(\text{prvič belo}) = \frac{6}{9} = \frac{2}{3} \\ &P(\text{prvič in drugič belo}) = \frac{4}{9} \\ &P(\text{drugič belo} \mid \text{prvič belo}) = \frac{4/9}{6/9} = \frac{4}{6} = \frac{2}{3} \\ &P(\text{drugič belo}) = \frac{6}{5} = \frac{2}{3} \end{split}$$

(b) Krogljic ne vračamo.

Vseh izidov je 6:

$$B_1B_2$$
 B_1 Č B_2B_1 B_2 Č Č B_1 Č B_2

$$P(\text{prvič belo}) = \frac{4}{6} = \frac{2}{3}$$

$$P(\text{prvič in drugič belo}) = \frac{2}{6} = \frac{1}{3}$$

$$P(\text{drugič belo} \mid \text{prvič belo}) = \frac{2/6}{4/6} = \frac{1}{2} = \frac{2}{4}$$

$$P(\text{drugič belo}) = \frac{4}{6} = \frac{2}{3}$$



Iz definicije pogojne verjetnosti dobimo:

$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

Za 3 dogodke A, B, C, kjer velja $P(B \cap C) > 0$ dobimo:

$$P(A \cap B \cap C) = P(A \mid B \cap C) \cdot P(B \cap C) = P(A \mid B \cap C) \cdot P(B \mid C) \cdot P(C)$$

Če to posplošimo na n dogodkov dobimo:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2 \mid A_1) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

Če desno stran razpišemo, res dobimo:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 \mid A_1) \cdot \dots \cdot P(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

$$= P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \dots \cdot \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})}$$

$$= P(A_1 \cap A_2 \cap \dots \cap A_n)$$

Zgled. V posodi imamo 6 modrih, 5 rdečih in 4 zelene krogljice. Brez vračanja izberemo 3 krogljice. Kolikšna je verjetnost, da so vse rdeče?

Označimo s A_k dogodek, da je k-ta krogljica rdeče za k=1,2,3. Nalogo laho rešimo na dva načina:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$
$$= \frac{5}{15} \cdot \frac{4}{14} \cdot \frac{3}{13} = \frac{2}{7 \cdot 13} = \frac{2}{91}$$

$$P(\text{vse rdeče}) = \frac{\binom{5}{3}}{\binom{15}{3}} = \frac{\frac{5\cdot 4\cdot 3}{3!}}{\frac{15\cdot 14\cdot 13}{3!}} = \frac{2}{91}$$

 \Diamond

Imejmo poskus, ki ga opravimo v 2 korakih (fazah).

- 1. V prvem koraku se zgodi natanko enden izmed paroma nezdružljivih dogodkov H_1, H_2, H_3, \ldots (končno ali števno mnogo).
- 2. V drugem koraku pa nas zanima dogodek A. Izrazimo P(A) z verjetnostmi:

$$P(H_i)$$
 in $P(A | H_i)$ za $i = 1, 2, 3, ...$

Ker je $\{H_i\}_i$ popoln sistem dogodkov, je:

$$A = A \cap \Omega = A \cap \left(\bigcup_{i} H_{i}\right) = \bigcup_{i} A \cap H_{i}$$

in zato

$$P(A) = \sum_{i} P(A \cap H_i) = \sum_{i} P(H_i) \cdot P(A \mid H_i)$$

To je formula za popolno verjetnost.

Zgled. Pri srečolovu je n srečk, od tega je m dobitnih (m < n). Ali imamo večje možnosti za dobitek, če izbiramo prvi ali drugi?

Če izbiramo prvi je:

$$P(\text{dobitka}) = \frac{m}{n}$$

Če izbiramo drugi je:

$$\begin{split} P(\text{dobitka}) &= P(\text{prvi dobi}) \cdot P(\text{dobitka} \mid \text{prvi dobi}) \\ &\quad + P(\text{prvi ne dobi}) \cdot P(\text{dobitka} \mid \text{prvi ne dobi}) \\ &= \frac{m}{n} \cdot \frac{m-1}{n-1} + \frac{n-m}{n} \cdot \frac{m}{n-1} = \frac{m}{n} \end{split}$$



Pogosto nas v dvofaznem poskusu zanima:

$$P(H_k \mid A) = \frac{P(H_k \cap A)}{P(A)} = \frac{P(H_k) \cdot P(A \mid H_k)}{\sum_i P(H_i) \cdot P(A \mid H_i)}$$

To je Bayesova formula.

 \Diamond

Zgled. Test s poligrafom (detektorjem laži). Resnicoljub opravi test s poligrafom z verjetnostjo 0.95. Z enako verjetnostjo poligraf prepozna lažnivca. Izmed 1000 oseb, med kateremi je natanko en lažnivec, slučajno izberemo eno osebo za katero poligraf pravi, da je lažnivec. Kolikšna je pogojna verjetnost, da je oseba zares lažnivec?

Označimo sL dogodek, da je izbrana oseba lažnivec, L_p pa dogodek, da poligraf za osebo pravi, da je lažnivec. Potem:

$$P(L_p \mid L) = 0.95, \quad P(L_p \mid L^c) = 0.05 \text{ in } P(L) = 0.001$$

Zanima nas $P(L \mid L_p)$. Po Bayesovi formuli je:

$$P(L \mid L_p) = \frac{P(L) \cdot P(L_p \mid L)}{P(L) \cdot P(L_p \mid L) + P(L^c) \cdot P(L_p \mid L^c)}$$

$$= \frac{0.55 \cdot 0.001}{0.55 \cdot 0.001 + 0.05 \cdot 0.999}$$

$$= \frac{95}{5090} = \frac{1}{50} = 0.02$$

Dogodka A in B sta neodvisna, če velja:

$$P(A \cap B) = P(A) \cdot P(B)$$

Če je P(B) > 0, to enakost lahko zapišemo kot:

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

Če imamo več dogodkov, so dogodki $\{A_i\}_i$ so neodvisni, če za poljuben končen nabor različnih dogodkov $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ velja:

$$P\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right) = P\left(A_{i_1}\right) \cdot P\left(A_{i_2}\right) \cdots P\left(A_{i_k}\right)$$

Če zahtevamo to le za k=2, torej A_i in A_j sta neodvisna za vsak $i \neq j$, potem rečemo, da so dogodki paroma neodvisni. To je šibkejši pogoj kot neodvisnost.

Zgled. Met tetraedra. $\Omega = \{1, 2, 3, 4\}$ in $P(\{i\}) = \frac{1}{4}$ za i = 1, 2, 3, 4.

Če označimo $A=\{1,2\}, B=\{1,3\}, C=\{1,4\},$ vidimo $A\cap B=A\cap C=B\cap C=\{1\}$ kar implicira $P(A\cap B)=P(A\cap C)=P(B\cap C)=\frac{1}{4}=\frac{1}{2}\cdot\frac{1}{2}.$ To implicira, da so dogodki $\{A,B,C\}$ paroma neodvisni. Toda $P(A\cap B\cap C)=P(\{1\})=\frac{1}{4}\neq\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}$, zato dogodki niso neodvisni.

Trditev 1. Če sta dogodka A in B neodvisna, potem sta neodvisna tudi dogodka A in B^c , dogodka A^c in B, ter dogodka A^c in B^c .

Dokaz. (i) Ker je $A \cap B^c = A \setminus A \cap B$ velja:

$$P(A \cap B^c) = P(A) - P(A \cap B)$$

$$= P(A) - P(A)2P(B)$$

$$= P(A)(1 - P(B))$$

$$= P(A)P(B^c)$$

(ii) Za A^c in B dokažemo podobno.

$$P(A^{c} \cap B^{c}) = 1 - P(A \cup B)$$

$$= 1 - P(A) - P(B) + P(A \cap B)$$

$$= P(A^{c}) - P(B) + P(A) \cdot P(B)$$

$$= P(A^{c}) - P(B)(1 - P(A))$$

$$= P(A^{c}) - P(B)P(A^{c})$$

$$= P(A^{c}) P(B^{c})$$

Poglavje 4

Zaporedja neodvisnih ponovitev poskusa

Imejmo zaporedje n neodvisnih ponovitev poskusa določenega z verjetnostnim prostorom (Ω, \mathcal{F}, P) v katerem je možen dogodek A s P(A) = p. Označimo še $q := P(A^c) = 1 - p$.

Z $A_n(k)$ označimo dogodek, da se v n ponovitvah poskusa, dogodek A zgodi natanko k-krat, za $k=0,1,2,\ldots,n$. Izračunajemo verjetnost:

$$P_n(k) := P\left(A_n(k)\right)$$

 $A_n(k)$ je dijunktna unija $\binom{n}{k}$ dogodkov, da se A zgodi na predpisanih k mestih, na ostalih pa A^c . Verjetnost teh dogodkov je $p^k \cdot q^{n-k}$. Zato je:

$$P_n(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$
 za $k = 0, 1, 2, \dots, n$

To je Bernoullijeva formula.

Zgled. Kaljivost semen je 0.95. Kolikšna je verjetnost, da izmed 1000 semen v zavojčku vzkali točno 950 semen?

Aje dogodek, da seme vzkali. Potem je p=P(A)=0,05, q=0.95, k=50. Zato:

$$P_{1000}(50) = {1000 \choose 50} \cdot 0.05^{50} \cdot 0.95^{950} = 0.05779$$

To je težko izračunati, tudi če bi uporabljali Stirlingovo formulo: $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$. Kjer:

$$a_n \sim b_n \iff \lim_{n \to \infty} \frac{a_n}{b_n} = 1$$

Torej:

$$\lim_{n \to \infty} \frac{\sqrt{2\pi n}}{n!} \left(\frac{n}{e}\right)^n = 1$$

Aproksimativni formuli za $P_n(k)$:

(a) Poissonova formula:

Če je p blizu 0 in n velik, potem je:

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}$$
 kjer je $\lambda = np$

Dokaz.

$$P_n(k) = \frac{n(n-1)\cdots n(n-k+1)}{k!} p^k (1-p)^{n-k}$$

$$= \frac{n(n-1)\cdots (n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots (n-k+1)}{\underbrace{n\cdot n\cdot n\cdot n}} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\approx \frac{\lambda^k}{k!} e^{-\lambda}$$

Zadnja aproksimacija velja v limiti.

Zgled (Kaljivost semen).

$$P_{1000}(50) = {1000 \choose 50} \cdot 0.05^{50} \cdot 0.95^{950}$$

$$\approx \frac{50^{50}}{50!} \cdot e^{-50}$$

$$= \frac{1}{50!} \left(\frac{50}{e}\right)^{50}$$

$$\approx \frac{1}{\sqrt{2\pi \cdot 50}}$$

$$\approx 0.05642$$

(b) Laplaceova lokalna formula:

Za velike n velja:

$$P_n(k) = \frac{1}{\sqrt{2\pi npq}} \cdot e^{-\frac{(k-np)^2}{2npq}}$$

Kasneje boste dokazali še splošnejši izrek (centralni limitni izrek).

Narišimo zaporedje $\{P_n(k)\}_{k=0}^n$, kjer je n fiksen. Dobimo:

$$P_n(0) = \binom{n}{0} p^0 q^n = q^n$$

$$P_n(1) = \binom{n}{1} p^1 q^{n-1}$$
:

Kdaj velja $P_n(k) \leq P_n(k+1)$?

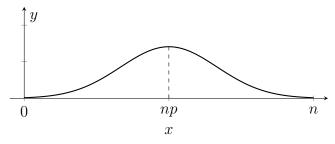
$$\frac{P_n(k+1)}{P_n(k)} \ge 1 \qquad \iff$$

$$\frac{n-k}{k+1} \cdot \frac{p}{q} \ge 1 \qquad \iff$$

$$np - kp \ge (k+1)q = kq + q \qquad \iff$$

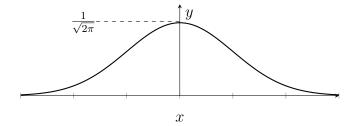
$$np \ge k + q$$

To je premaknjena in raztegnjena "normalna porazdelitev".



Gausova ali normalna porazdelitev:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$



Zgled (Kaljivost semen). p = 0.05, np = 50:

$$P_{1000}(50) \approx \frac{1}{\sqrt{2\pi \cdot 50 \cdot 0.95}} = \frac{1}{\sqrt{95 \cdot \pi}} = 0.05788$$



Poglavje 5

Slučajne spremenljivke

Danemu poskusu priredimo številsko količino, katere vrednost je odvisna od slučaja. Imenujemo jo slučajna spremenljivka.

Zgled.

- 1. Met kocke; število pik je slučajna spremenljivka.
- 2. Streljanje v tarčo; razdalja zadetka od središča tarče je slučajna spremenljivka. Na primer, verjetnost, da je razdalja $\leq x$ je sorazmerna z x^2 .

 \Diamond

Definicija 1. Realna slučajna spremenljivka na verjetnostnem prostoru (Ω, \mathcal{F}, P) je funkcija $X : \Omega \to \mathbb{R}$, z lastnostjo, da je za vsak $x \in \mathbb{R}$ množica:

$$\{\omega \in \Omega : X(\omega) \le x\}$$

v \mathcal{F} , se pravi dogodek.

Oznaka:

$$\{\omega \in \Omega : X(\omega) \le x\} \equiv X^{-1} \left((-\infty, x] \right) \equiv (X \le x)$$

Definicija 2. Naj bo X slučajna spremenljivka. Funkcija $F_X : \mathbb{R} \to \mathbb{R}$, definirana s predpisom:

$$F_X(x) := P(X \le x)$$

se imenuje porazdelitvena funkcija slučajne spremenljivke X.

Lastnosti porazdelitvene funkcije $F = F_X$:

- 1. $0 \le F(x) \le 1$ za vsak $x \in \mathbb{R}$.
- 2. F je naraščujoča funkcija:

$$x_1 < x_2 \implies F(x_1) \le F(x_2)$$

Dokaz. Sledi iz $(X \leq x_1) \subseteq (X \leq x_2)$ in monotonosti preslikave P.

3.

$$\lim_{x \to \infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0$$

Dokaz. Naj bo $\{x_n\}_n\subseteq\mathbb{R}$ ne
omejeno strogo naraščujoče zaporedje. Potem je:

$$\lim_{x \to \infty} F(x) = \lim_{n \to \infty} F(x_n)$$

$$= \lim_{n \to \infty} P(X \le x_n)$$

$$= P\left(\bigcup_{n=1}^{\infty} (X \le x_n)\right) \tag{**}$$

$$= P(\Omega) = 1 \tag{**}$$

kjer smo v (\star) vrstici uporabili zveznost P, $(\star\star)$ vrstici pa $\forall \omega \in \Omega : \exists n : X(\omega) \leq x_n$, torej $\omega \in (X \leq x_n)$.

Podobno se dokaže tudi druga limita.

4. F je z desne zvezna, t.j.:

$$F(x+) = F(x)$$
, kjer je $F(x+) = \lim_{\substack{h>0 \ h \to 0}} F(x+h)$

Dokaz. Naj zaporedje $\{x_n\}\subseteq\mathbb{R}$ strogo pada protix. Potem je:

$$F(x+) = \lim_{n \to \infty} F(x_n)$$

$$= \lim_{n \to \infty} P(X \le x_n)$$

$$= P\left(\bigcap_{n=1}^{\infty} (X \le x_n)\right) \tag{**}$$

$$= P(X \le x)$$

$$= F(x)$$

 (\star) velja, zaradi zveznosti P, $(\star\star)$ pa velja saj je:

$$\bigcap_{n=1}^{\infty} (X \le x_n) = (X \le x)$$

Tukaj je:

⊇: očitna

 \subseteq :

$$\exists \omega \in \bigcap_{n=1}^{\infty} (X \le x_n) : X(\omega) \le x_n \quad \forall n$$
$$\implies X(\omega) \le x, \text{ t.j. } \omega \in (X \le x)$$

5. • $P(x_1 < X \le x_2) = P(X \in (x_1, x_2])$ $= P((X \le x_2) \setminus (X \le x_1))$ $= P(X \le x_2) - P(X \le x_1)$ $= F(x_2) - F(x_1)$

•
$$P(x_1 < X < x_2) = P(X < x_2) - P(X \le x_1)$$

= $F(x_2 -) - F(x_1)$

Saj je:

$$P(X < x) = P\left(\bigcup_{n=1}^{\infty} \left(X \le x - \frac{1}{n}\right)\right) = \lim_{n \to \infty} P\left(X \le x - \frac{1}{n}\right) = F(x-1)$$

•
$$P(x_1 \le X \le x_2) = P(X \le x_2) - P(X < x_1)$$

= $F(x_2) - F(x_1 - x_1)$

•
$$P(x_1 \le X < x_2) = P(X < x_2) - P(X < x_1)$$

= $F(x_2-) - F(x_1-)$

Oglejmo si dva najpomembnejša razreda slučajnih spremenljivk.

5.1 Diskretne slučajne spremenljivke

Slučajna spremenljivka X je diskretno porazdeljena, če je njena zaloga vrednosti končna ali števna množica števil $\{x_1, x_2, x_3, \ldots\}$.

Tedaj vpeljemo verjetnostno funkcijo:

$$p_k = P(X = x_k)$$
 za $k = 1, 2, 3, ...$

in shemo:

$$X: \left(\begin{array}{ccc} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{array}\right), \quad \sum_k p_k = 1$$

Tukaj je $\{(X = x_k)\}_k$ popoln sistem dogodkov.

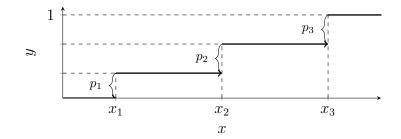
Porazdelitvena funkcija je:

$$F(x) = P(X \le x) = P\left(\bigcup_{k: x_k \le x} (X = x_k)\right) = \sum_{k: x_k \le x} p_k$$

F je odsekoma konstantna.

Na primer za $x_1 < x_2 < x_3$ in $p_1 + p_2 + p_3 = 1$:

$$X: \left(\begin{array}{ccc} x_1 & x_2 & x_3 \\ p_1 & p_2 & p_3 \end{array}\right)$$



Pomembnejše diskretne porazdeitve:

1. Enakomerna diskretna porazdelitev na n točkah x_1, x_2, \dots, x_n :

$$X: \left(\begin{array}{cccc} x_1 & x_2 & \dots & x_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{array}\right)$$

Zgled. Met poteven kocke, n = 6:

$$X: \left(\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array}\right)$$

 \Diamond

2. Bernoullijeva porazdelitev $Ber(p), p \in (0,1), q = 1 - p$:

$$x: \left(\begin{array}{cc} 0 & 1 \\ q & p \end{array}\right)$$

(X = 1) je dogodek, da se dogodek A zgodi.

(X = 0) je dogodek, da se dogodek A ne zgodi.

Indikatorska funkcija:

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & w \in A \\ 0, & w \notin A \end{cases}$$

3. Binomska porazdelitev $Bin(n, p), n \in \mathbb{N}, p \in (0, 1)$:

$$X: \left(\begin{array}{ccccc} 0 & 1 & 2 & 3 & \cdots & n \\ p_0 & p_1 & p_2 & p_3 & \cdots & p_n \end{array}\right)$$

$$p_k = P(X = k) = \binom{n}{k} p^k q^{n-k}$$
, kjer je $q = 1 - p$

(X=k)je dogodek, da se dogodek Azgodi natanko $k\text{-}\mathrm{krat}$ vnponovitvah poskusa.

Zgled. Mečemo kocko. X je število šestic v n metih. $X \sim \text{Bin}(n, \frac{1}{6})$.

4. Poissonova porazdelitev $Poi(\lambda), \lambda > 0$:

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ za } k = 0, 1, 2, \dots$$

Res velja:

$$\sum_{k=0}^{\infty} p_k = e^{-\lambda} \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) = e^{-\lambda} e^{\lambda} = 1$$

Zgled. Število klicev v telefonskem omrežju v minuti. n je število naročnikov. Recimo, da je $n=10^6$ in vsak se z verjetnostjo p odloči, da bo klical. $X \sim \text{Bin}(n,p) \approx \text{Poi}(np)$. Primer je tudi število napačnih črk v časopisu.

5. Geometrijska porazdelitev $geo(p), p \in (0,1)$:

Ponavljanjo poskus in opazujmo dogodek A s P(a) = p. X je število potrebnih ponovitev, da se zgodi dogedek A prvič.

(X = k) je dogodek, da se A zgodi prvič k k-ti ponovitvi poskusa:

$$\underbrace{\bar{A}\bar{A}\bar{A}\dots\bar{A}}_{k-1}A$$

$$p_k = P(X = k) = pq^{k-1}$$
, za $k = 1, 2, 3, ...$

Res velja:

$$\sum_{k=1}^{\infty} p_k = p \sum_{k=1}^{\infty} q^{k-1} = p \frac{1}{1-q} = 1$$

Zgled. Mečemo kocko. X je število potrebnih metov, da pade šestica prvič. $X \sim \text{geo}(\frac{1}{6})$.

6. Pascalova oz. nenegativna binomska porazdelitev $\operatorname{Pas}(m,p), m \in \mathbb{N}, p \in (0,1)$: X je število potrebnih ponovitev poskusa, da se dogodek A zgodi m-krat. (X=k) je dogodek, da se A zgodi m-tič v k-ti ponovitvi.

$$\underbrace{AA\bar{A}A\bar{A}\dots\bar{A}A}_{m-1}\underbrace{A}_{k-m}\underbrace{A}$$

$$p_k = P(x = k) = \binom{k-1}{m-1} p^m q^{k-m}, \text{ za } k = m, m+1, m+2, \dots$$

 $\sum_{k=m}^{\infty} p_k = 1$ se dokaže računsko z (m-1)-kratnim odvajanjem vrste $1+q+q^2+\cdots=\frac{1}{1-q}$ ali pa direktno z uporabo vrste $(1-q)^{-m}=\sum_{j=0}^{\infty} {-m \choose j} (-q)^j$.

Geometrijska porazdelitev je poseben primer Pascalove: Pas(1, p) = geo(p)

Zgled. Mečemo kocko. X je število potrebnih meto, da šestica pade m krat. $X \sim \operatorname{Pas}(m, \frac{1}{6})$.

7. Hipergeometrijska porazdelitev $\operatorname{Hip}(n; M, N), n \leq \min\{M, N - M\}$:

V posodi imamo M belih in (N-M) črnih krogljic. Slučajno izvlečemo n krogljic. X naj bo število belih krogljic med izvlečenimi:

$$p_k = P(x = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

Ker je $\{(X=k)_{k=0}^n$ popol
n sistem dogodkov, je $\sum_{k=0}^n p_k = 1$. Torej velja:

$$\sum_{k=0}^{n} \binom{M}{k} \binom{N-M}{n-k} = \binom{N}{n}$$

Zgled. V jezeru je N rib, od tega je M krapov. Ulovimo n rib. Naj bo X število ujetih krapov. $X \sim \text{Hip}(n; M, N)$.

Vzemimo, da je $n \ll \min\{M, N-M\}$. Tedaj ne naredimo velike napake, če krogljice vrečamo. Tedaj je $X \sim \text{Bin}(n, \frac{M}{N})$ ozirioma $p_k = \binom{n}{k} \left(\frac{M}{N}\right)^k \left(\frac{N-M}{N}\right)^{n-k}$ za $k = 0, 1, \ldots, n$.

5.2 Zvezne slučajne spremenljivke

Slučajna spremenljivka X je zvezno porazdeljena, če obstaja nenegativna integrabilna funkcija p_X , imenovana gostota verjetnosti, da za vsak $x \in \mathbb{R}$ velja:

$$F_X(x) = \int_{-\infty}^x p_X(t) dt$$

Pogosto se gostota označuje s f_X .

Tedaj je $F_X = F$ zvezna funkcija. Toda obstajajo tudi zvezne porazdelitvene funkcije, ki nimajo gostote.

Ker je $\lim_{x\to\infty} F(x) = 1$, je $\int_{-\infty}^{\infty} p(t) dt = 1$.

Če je p zvezna v toči x, potem je F odvedljiva v x in velja F'(x) = p(x).

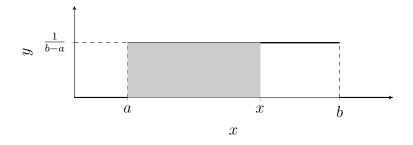
Za vsak $x \in \mathbb{R}$ velja P(X = x) = F(x) - F(x-) = 0. Če je $x_1 \le x_2$, je potem:

$$P(x_1 \le X_1 \le x_2) = F(x_2) - F(x_1 - 1) = \int_{x_1}^{x_2} p(t) dt$$

Nekatere pomembnejše zvezne porazdelitve:

1. Enakomerna zvezna porazdelitev na [a, b], a < b:

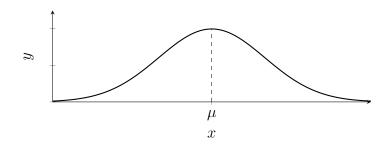
$$p(x) = \begin{cases} \frac{1}{b-a}, & \text{\'e } a \le x \le b \\ 0, & \text{sicer} \end{cases}$$



$$F(x) = \int_{-\infty}^{x} p(t) dt = \begin{cases} 0, & \text{\'e } x \le a \\ \frac{x-a}{b-a}, & \text{\'e } a \le x \le b \\ 1, & \text{\'e } x \ge b \end{cases}$$

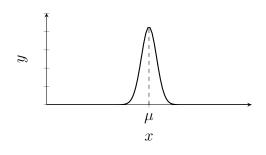
2. Normalna ali Gaussova porazdelitev $N(\mu, \sigma), (\mu \in \mathbb{R}, \sigma > 0)$:

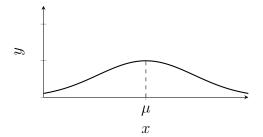
$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Če je σ majhen:

Če je σ velik:





$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2}} dt \stackrel{\star}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}s^{2}} ds = \frac{1}{2} + \phi\left(\frac{x-\mu}{\sigma}\right)$$

Kjer smo v enakosti \star uvedli novo spremenljivko $s=\frac{t-\mu}{\sigma}.$

Tukaj je:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{2\pi t^2} dt$$

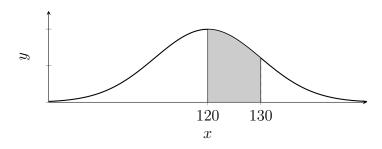
N(0,1): standardna normalna porazdelitev:

$$p(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Laplaceova formula: za velike n je:

$$Bin(n, p) \approx N(np, \sqrt{npq})$$

Zgled. Sistolični krvni tlak je an populciji približno normalno porazdeljen:

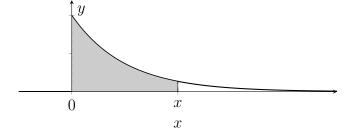


delež ljudi, ki imajo tlak med 120 in 130 mm Hg.

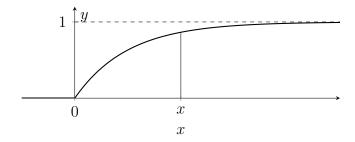
3. Eksponenta porazdelitev $\text{Exp}(\lambda), \lambda > 0$:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{ \'e } x \ge 0\\ 0, & \text{ sicer} \end{cases}$$

 \Diamond



$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{ \'e } x \ge 0\\ 0, & \text{ sicer} \end{cases}$$



Zgled. Radioaktivni razpad - potreben čas, da se nekaj zgodi.