I Vjerojatnost

0 Uvod – Što je to "vjerojatnost"?

Svi imaju nekakvu intuiciju o tome što je vjerojatnost; npr.

(a) vjerojatnost da padne pismo kod bacanja simetričnog novčića je ... 50 % ("50-50")

(b) vjerojatnost da na kocki dobijemo neki od brojeva 5 ili 6 je ... 33.3̇ %

Kako smo došli do ovih brojeva? → SIMETRIJA

(a) imamo 2 ishoda (P ili G) koji su zbog simetrije "jednako vjerojatni" → vjerojatnost za P je $\frac{1}{2} = 50\%$.

(b) imamo 6 ishoda (1, 2, ..., 6) —— 11 —— → vjerojatnost za 5 ili 6 (2 ishoda) je $\frac{2}{6} = \frac{1}{3} = 33.\dot{3}\%$.

1. ideja za definiciju vjerojatnosti:
Pretp. da procedura ("pokus") sa slučajnim ishodom ima n mogućih ishoda, te da su oni zbog simetrije svi jednako vjerojatni. Ako je A događaj koji sadrži r takvih ishoda (u (a) A = palo je P, $n=2, r=1$, (b) A = palo je 5 ili 6, $n=6, r=2$) definiramo da je

vjerojatnost od A jednaka $\frac{r}{n}$, a pišemo

$$\mathbb{P}(A) = \frac{r}{n}.$$

Uočimo, naime, da je

- uvijek je $0 \le \mathbb{P}(A) \le 1$, (jer $0 \le r \le n$),
- ako A sadrži sve ishode, $\mathbb{P}(A) = \frac{n}{n} = 1$,
- ako A ne sadrži niti jedan ishod, $\mathbb{P}(A) = \frac{0}{n} = 0$,
- ako je A "jako vjerojatan", tj. r jako velik (relativno na n), $\mathbb{P}(A)$ je blizu 1,
- ako je A "jako malo vjerojatan", $\mathbb{P}(A)$ je blizu 0.

Problem s ovakvom definicijom:

1° Mnoge procedure ne sadrže u sebi simetriju.

2° Zaključili smo da je $\mathbb{P}(\text{palo } 5 \text{ ili } 6) = \frac{1}{3}$, bez da smo ijednom bacili kocku?!

Asimptotski pristup vjerojatnosti:

Pretp. da bacamo nesavršenu kocku $\rightarrow \mathbb{P}(\text{palo je } 5 \text{ ili } 6) := ?$

Nema simetrije, ali ako kocku bacamo n puta za jako veliki n, te ako je $r(n)$ broj bacanja u kojima smo dobili 5 ili 6, simetrija među bacanjima sugerira da je

$$\mathbb{P}(5 \text{ ili } 6) \approx \frac{r(n)}{n}.$$

nadalje, kad bi zaista bacali ovu kocku puno puta, uvijek bi se omjer $\frac{r(n)}{n}$ "stabilizirao" oko približno iste vrijednosti kako povećavamo n. Npr., kod savršene kocke bi zaista imali

$$\frac{r(n)}{n} \approx \frac{1}{3} \text{ za velike } n. \quad [\text{7ZVB}]$$

2. definicija vjerojatnosti:
Ako se procedura može ponoviti veliki broj puta n, pod uvijek istim uvjetima, te ako se događaj A dogodio u $r(n)$ od n slučajeva,

$$\mathbb{P}(A) :\approx \frac{r(n)}{n}.$$

Opet imamo $0 \le \mathbb{P}(A) \le 1$, $\mathbb{P}(A) = 0$ ako se A ne može dogoditi (jer $r(n) = 0$), tj. $\mathbb{P}(A) = 1$ ako se A uvijek događa ($r(n) = n$).

Problemi:

(1°) Ovo je samo aproksimacija
[npr. kod bacanja novčića i A = palo je P, možemo dobiti za $n = 10000$, $\frac{r(n)}{n} = 0.5005 \neq 0.5$]

(2°) Neke procedure se ne mogu ponavljati puno puta, ili čak više od jedan put, npr. rezultat utakmice ili izbora.

Pristup preko "pravedne cijene"

Pretp. da igrate igru u kojoj dobivate 1 kn ako se dogodi neki događaj A. Koja je pravedna cijena za ovu igru?

Npr. Ako se baca simetričan novčić te A = palo je P, pravedna cijena je 0.5 kn jer ako puno puta igrate ovu igru vaš dobitak će u prosjeku biti

$$\frac{1 \cdot r(n)}{n} = \frac{r(n)}{n} \approx \frac{1}{2} = \mathbb{P}(A).$$

Ako umjesto 1 kn dobivate d kn ($d \geq 0$, $d \in \mathbb{R}$), pravednu cijenu je ... $d \cdot \frac{1}{2} = d \cdot \mathbb{P}(A)$.

Općenito, ako je $\mathbb{P}(A) = p \in [0,1]$, pravedna cijena igre je $d \cdot p$. → Ovo je povezano s pojmom tzv. matematičkog očekivanja!

Vrijedi i obrat: ako ste se s nekim dogovorili da je cijena gornje igre (uz $d=1$) jednaka $p \in [0,1]$, tada ste implicitno definirali $\mathbb{P}(A) := p$.

→ npr. to je slučaj u klađenju.

↳ za razliku od prva dva pristupa, ovaj je ipak subjektivan!

Slično je i sa sljedećim tvrdnjama:

(a) $\mathbb{P}$(sutra pada kiša) = 60 %, ili

(b) $\mathbb{P}$(postoji život izvan Zemlje) = 20 %.

Naš cilj = nije dati odgovor na pitanje: što je vjerojatnost? (to je više filozofija), već dati (preciznu) matematičku teoriju vjerojatnosti.

Ona će biti u skladu s našom intuicijom o vjerojatnosti u jednostavnim slučajevima (npr. kada postoji očita simetrija), ali će nam dati odgovore i u slučajevima kada nas intuicija ne služi baš najbolje. To mogu biti i jednostavni primjeri:

(a) Rijetka bolest pogađa 1 od 100,000 osoba.

Test koji imate je 99% pouzdan.

[tj. s tom vjerojatnosti pokazuje točan rezultat.]

Ako je vaš test pozitivan, kolika je vjerojatnost da zaista imate bolest? ... manje od 0.1%!

→ pojmovi uvjetne vjerojatnosti i Bayesovog teorema.

(b) U grupi od 23 slučajno odabranih ljudi, kolika je vjerojatnost da su barem dvoje rođeni na isti dan u godini? ... 50.7%

U grupi od 70 ljudi? ... 99.9%!

[$\binom{23}{2}$ i $\binom{70}{2}$ su jako veliki!]

→ "problem rođendana"

(c) tzv. Monty Hall problem: Nasumično odaberete jedna od ukupno troja vrata – iza jednih je auto, iza ostala dva su ovce. Netko tko zna što je iza kojih vrata, otvori jedna od preostala dva vrata, ali nikad ona iza kojeg je auto. Trebate li promijeniti vaš izbor?

.... Da!

[vjerojatnost je $\frac{1}{3}$ za vaša vrata, a $\frac{2}{3}$ za druga!]