Verovatnoca i statistika

U literaturi koja obradjuje verovatnocu susrecemo se sa pojmovima eksperimenta, prostora ishoda, dogadja i funkcije verovatnoce. Eksperimentom mozemo smatrati proceduru koja se moze ponavljati (u istim uslovima) i koja ima dobro definisane ishode. Skup svih ishoda nazivamo **prostorom ishoda**. Podskup prostora ishoda nazivamo **dogadjaj**. Funkciju kojom svakom ishodu pridruzujemo verovatnocu nazivamo **funkcijom verovatnoce**.

Na primer, bacanje fer novcica mozemo smatrati eksperimentom sa mogucim ishodima pismo (P) ili glava (G). Ako novcic bacamo tri puta, prostor ishoda je skup $\Omega = \{PPP, PPG, PGP, PGG, GPP, GPG, GGP, GGG\}$. Svaki ishod je jednako verovatan, sa verovatnocom 1/8. Dogadjaj bi mogao biti skup ishoda $E = \{GGP, GPG, PGG\}$ u kojima se glava pojavljuje tacno dva puta. Verovatnoca ovog dogadjaja je 3/8.

In [ ]:

```

```

U opstem slucaju prostor ishoda moze biti *diskretan* ili *kontinualan*.

Formalno, za diskretan prostor ishoda Ω, funkcija $P : \Omega \to R$ je **funkcija verovatnoće** ukoliko zadovoljava uslove:

- $0 \leq P(w) \leq 1$
- $\sum_{w \in \Omega} P(w) = 1$

Za dogadjaje A, L i R koji pripadaju istom prostoru ishoda Ω i funkciju verovatnoce P definisanu na njemu vaze:

- princip komplementarnosti: $P(A^c) = 1 - P(A)$
- princip ukljucivanja i iskljucivanja: $P(L \cup R) = P(L) + P(R) - P(L \cap R)$

In [ ]:

```

```

Uslovna verovatnoca

Neka su A i B dogadjaji koji pripadaju istom prostoru ishoda Ω. Uslovna verovatnoca dogadjaja A pri uslovu B se definise kao $P(A|B) = \frac{P(A \cap B)}{P(B)}$, uz ogranicenje $P(B) \neq 0$.

In [ ]:

```

```

Posmatrajmo bacanje novcica tri puta. Neka je A dogadjaj koji predstavlja dobitak tri glave $\{GGG\}$, a B dogadjaj koji predstavlja da je pri prvom bacanju dobijena glava $\{GGG, GGP, GPG, GPP\}$. Kako je $P(A) = \frac{1}{8}$, $P(B) = \frac{1}{2}$ i $A \cap B = A$, mozemo izracunati da je $P(A|B) = \frac{1}{4}$.

Bajesova formula

Neka su A i B dogadjaji koji pripadaju istom prostoru ishoda Ω. Tada vazi:

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

In [ ]:

Ova formula nam prakticno pomaze da izracunamo verovatnoce P(B|A) i P(A|B) ukoliko je jedna od njih poznata.

In [ ]:

Nezavnisnost dogadjaja

Neka su A i B dogadjaji koji pripadaju istom prostoru ishoda Ω. Za njih kazemo da su nezavnisni ako vazi $P(A \cap B) = P(A)P(B)$.

In [ ]:

Posmatrajmo bacanje novcica dva puta. Neka je A dogadjaj koji oznacava da je pri prvom bacanju dobijena glava, a B dogadjaj koji oznacava da je pri drugom bacanju dobijena glava. Kako je presek dogadjaja A i B dogadjaj u kojem su glave dobijene oba puta i kako vazi $P(A \cap B) = \frac{1}{4} = P(A)P(B)$, mozemo zakljuciti da su dogadjaji A i B nezavisni.

In [ ]:

Posmatrajmo bacacanje novcica 3 puta. Neka je A dogadjaj u koji oznacava da je prilikom prvog bacanja dobijena glava, a B dogadjaj koji oznacava da su ukupno pale dve glave. Kako dogadjaj A objedinjuje ishode $\{GPP, GPG, GGP, GGG\}$, a dogadjaj B ishode $\{GGP, GPG, PGG\}$, to je $A \cap B$ skup ishoda $\{GPG, GGP\}$ i mozemo zakljuciti da ne vazi $P(A \cap B) = P(A)P(B)$ i da dogadjaji A i B nisu nezavisni.

Slucajna promenljiva (slucajna velicina)

Neka je Ω prostor ishoda. Diskretna slucajna promenljiva X je funkcija $X : \Omega \rightarrow R$.

In [ ]:

Posmatrajmo bacanje dveju kockica. Neka i oznavaca vrednost dobijenu na prvoj kockici, a j vrednost dobijenu na drugoj kockici. Funkcija

$$X(i,j) = \begin{cases} 500, & \text{ako je } i+j=7 \\ -100, & \text{inace} \end{cases}$$

je primer jedne slucajne promenljive.

In [ ]:

Matematicko ocekivanje

Neka je X diskretna slucajna velicina koja uzima vrednosti x_1, x_2, ..., x_n redom sa verovatnocama p_1, p_2, ..., p_n. Matematicko ocekivanje $E(X)$ slucajne velicine X se definise kao

$$E(X) = \sum_{i=1}^{n} p_i \cdot x_i = p_1 x_1 + p_2 x_2 + \ldots + p_n x_n$$

.

Posmatrajmo slucajnu velicinu X koja predstavlja zbir vrednosti prilikom bacanja dveju kockica. Vrednosti koje ova slucajna velicina uzima i njima pripadajuce verovatnoce mozemo zapisati tablicno:

2	3	4	5	6	7	8	9	10	11	12
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Njeno matematicko ocekivanje je

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12$$

In [ ]:

Neka su X i Y dve slucajne velicine definisane na istom prostoru ishoda i a i b konstante. Tada vaze osobine:

- $E(X+Y) = E(X) + E(Y)$
- $E(aX+b) = aE(x) + b$
- Ako je funkcija h definisanu za vrednosti x_1, x_2, ..., x_n koje uzima slucajna velicina X, tada je i $h(X)$ slucajna velicina sa matematickim ocekivanjem $E(h(X)) = \sum_{i=1}^{n} p_i \cdot h(x_i)$.

Posmatrajmo isti problem bacanja dveju kockica. Neka je X_1 slucajna velicina koja predstavlja vrednost dobijenu na prvoj kockici, a X_2 slucajna velicina koja predstavlja vrednost dobijenu na drugoj kockici. Za slucajnu velicinu X iz prethodnog primera vazi $X = X_1 + X_2$ pa se trazeno matematicko ocekivanje moze dobiti i malo jednostavnije $E(X) = E(X_1) + E(X_2)$. Kako je $E(X_1) = E(X_2) = 3.5$ (proveriti!), dobijamo $E(X) = 7$.

In [ ]:

```
```

Varijansa i standardna devijacija diskretne slučajne promenljive

Varijansa je mera koja ukazuje na to koliko su vrednosti diskretne slučajne promenljive raspršene (engl. spread) oko srednje vrednosti tj. ocekivanja.

Neka je X slučajna veličina sa očekivanjem $E(X) = \mu$, `varijansa slučajne veličine` $Var(X)$ se definiše kao

$$Var(X) = E((X - \mu)^2) = \sum p(x_i)(x_i - \mu)^2$$

.

`Standardna devijacija` σ slučajne veličine X se definiše kao koren varijanse $\sqrt{Var(X)}$. Ona se izračunava u istim merama kao i slučajna veličina X.

Varijansa se često obeležava sa σ^2.

Svojstva varijanse:

- Ako su X i Y nezavisne slučajne veličine, tada je $Var(X + Y) = Var(X) + Var(Y)$.
- Za konstante a i b važi $Var(aX + b) = a^2 Var(X)$.
- $Var(X) = E(X^2) - E(X)^2$

In [ ]:

```
```

Pretpostavimo da treba izračunati srednju vrednost, varijansu i standardnu devijaciju slučajne veličine X koja uzima vrednosti 1, 3 i 5 redom sa verovatnoćama $1/4$, $1/4$ i $1/2$.

Kako je $E(X) = 1 \cdot 0.25 + 3 \cdot 0.25 + 5 \cdot 0.5 = 3.5$, veličina $X - E(X)$ uzima vrednosti -2.5, 0.5, i 1.5, a veličina $(X - E(X))^2$ vrednosti 6.25, 0.25 i 2.25. Stoga je varijansa $0.25 \cdot 6.25 + 0.25 \cdot 0.25 + 0.5 \cdot 2.25 = 2.75$, a standardna devijacija 1.65.

In [ ]:

```
```