

Zadaci za vježbu: osnove vjerojatnosti i uvjetna vjerojatnost

Vježbe 4

Napomena. Svi broježani podaci i modeli pokusa u ovom listu hipotetski su, osim izričito navedenih ugrađenih skupova podataka. Za svaki slučajni izbor navedite prostor mogućih ishoda ili jasno definirajte događaje. Pri simulacijama postavite `set.seed(104)`, koristite `sample()` i prema potrebi `replicate()`. Relativnu frekvenciju događaja možete izračunati kao sredinu logičkog vektora. Za zadatke na papiru dovoljan je običan kalkulator; potrebne tablične vrijednosti navedene su uz zadatke. U radu u Ru koristite osnovnu instalaciju bez dodatnih paketa. Grafove opremite oznakama osi i jedinicama, a rezultate kratko protumačite.

1 Zadaci za rješavanje olovkom i papirom

Zadatak 1. Preklapanje dvaju obilježja.

Među 40 listova njih 18 ima oštećen rub, 12 ima pjege, a 5 ima oba obilježja. Jedan list biramo tako da svih 40 listova ima jednaku vjerojatnost izbora.

- (a) Dopunite tablicu 2×2 za prisutnost odnosno odsutnost obilježja.
- (b) Odredite vjerojatnost da list ima barem jedno, točno jedno ili nijedno obilježje.
- (c) Jesu li događaji “oštećen rub” i “pjege” nezavisni? Provjerite računom.

Zadatak 2. Raspored mjerenja.

Četiri različita uzorka, A, B, C i D, mjere se u slučajnom redoslijedu; svi redoslijedi jednako su vjerojatni.

- (a) Koliko je mogućih redoslijeda? Kolika je vjerojatnost da A prethodi B?
- (b) Odredite vjerojatnost da su A i B neposredno jedan uz drugoga, u bilo kojem poretku.
- (c) Ako znamo da je A prvi, kolika je vjerojatnost da su A i B susjedni? Objasnite zašto se prostor mogućih redoslijeda smanjio.

Zadatak 3. Dvostruka kontrola.

Dva neovisna senzora provjeravaju isti uzorak. Kada uzorak sadrži tvar, senzor A daje pozitivan signal s vjerojatnošću 0.8, a senzor B s 0.6. U ovom zadatku znamo da uzorak sadrži tvar; nezavisnost se odnosi na rezultate senzora u toj situaciji.

- (a) Navedite četiri moguća para rezultata i njihove vjerojatnosti.
- (b) Odredite vjerojatnost barem jednog i točno jednog pozitivnog signala.
- (c) Ako je barem jedan signal pozitivan, kolika je vjerojatnost da su oba pozitivna?

Zadatak 4. Kontrolni uzorak među mjerenjima.

Na stalku je šest različitih bočica: dvije kontrolne i četiri s nepoznatim uzorcima. Na slučajan način biramo tri bočice bez vraćanja.

- (a) Odredite vjerojatnost da izbor sadrži barem jednu kontrolnu bočicu, brojenjem jednako vjerojatnih tročlanih podskupova.
- (b) Ako izbor sadrži barem jednu kontrolnu bočicu, kolika je vjerojatnost da sadrži obje?
- (c) Jesu li događaji “odabrana je prva kontrolna bočica” i “odabrana je druga kontrolna bočica” nezavisni? Obrazložite računom.

Zadatak 5. Pozitivan signal i rijetka pojava.

U hipotetskoj populaciji uzoraka 10% sadrži promatranu tvar. Test daje pozitivan signal u 90% uzoraka koji sadrže tvar i u 5% uzoraka koji je ne sadrže.

- (a) Nacrtajte stablo vjerojatnosti i izračunajte vjerojatnost pozitivnog signala.
- (b) Odredite vjerojatnost prisutnosti tvari ako je signal pozitivan.
- (c) Ponovite račun ako tvar sadrži samo 1% uzoraka, uz ista svojstva testa. Objasnite zašto pozitivan signal nema jednako značenje u objema populacijama.

2 Zadaci za rad u R-u

Zadatak 1. Svi mogući parovi oznaka.

Dvije bočice neovisno dobivaju slučajnu oznaku iz skupa $\{1, 2, 3, 4, 5, 6\}$, pri čemu su oznake jednako vjerojatne.

- (a) Naredbom `expand.grid(a = 1:6, b = 1:6)` popišite sve uređene parove.
- (b) Odredite vjerojatnosti događaja A : zbroj oznaka je barem 9, B : oznake su jednake, te $A \cap B$ i $A \cup B$.
- (c) Izračunajte $P(A \mid B)$ i provjerite nezavisnost događaja. Prikažite točke (a, b) , a pripadnost događaju A označite bojom.

Zadatak 2. Vraćanje mijenja model.

Osam različitih sjemenki ima oznake 1–8; sjemenke 1, 2 i 3 su obojene.

- (a) Simulirajte 10000 odabira po dvije sjemenke bez vraćanja i odredite relativne frekvencije dviju obojenih te barem jedne obojene sjemenke.
- (b) Ponovite simulaciju s vraćanjem. Izračunajte i teorijske vjerojatnosti za oba načina izbora.
- (c) Prikažite teorijske i simulirane vrijednosti u tablici te objasnite razliku između modela i odstupanja zbog konačnog broja simulacija.

Zadatak 3. Događaj uz poznat uvjet.

U 100 uzoraka vode zabilježeno je:

	Tvar prisutna	Tvar odsutna
Signal pozitivan	18	8
Signal negativan	2	72

Slučajno se bira jedan od tih 100 uzoraka.

- (a) Unesite tablicu u matricu i izračunajte rubne relativne frekvencije.
- (b) Izračunajte $P(\text{pozitivan} \mid \text{prisutan})$ i $P(\text{prisutan} \mid \text{pozitivan})$ te objasnite razliku nazivnika.
- (c) Simulirajte 20000 izbora s vraćanjem iz te konačne populacije (četiri vrste uzoraka birajte s odgovarajućim vjerojatnostima) i provjerite obje uvjetne vjerojatnosti relativnim frekvencijama.

Zadatak 4. Pojava ponovljene oznake.

Četiri uzorka neovisno dobivaju jednako vjerojatnu oznaku iz skupa 1–10.

- (a) Koristeći komplement izračunajte vjerojatnost da barem dva uzorka imaju istu oznaku. Ne zbrajajte vjerojatnosti svih parova bez provjere preklapanja.
- (b) Simulirajte 10000 dodjela oznaka. Za otkrivanje ponavljanja upotrijebite `anyDuplicated()` ili `length(unique(...))`.
- (c) Ponovite za 2, 3, 4, 5 i 6 uzoraka te nacrtajte teorijsku i simuliranu vjerojatnost u ovisnosti o broju uzoraka.

Zadatak 5. Isti test u različitim populacijama.

Test ima $P(+ | T) = 0.9$ i $P(+ | T^c) = 0.05$, gdje T označava prisutnost tvari.

- (a) Za udjele prisutnosti od 0.001 do 0.5 u koracima 0.001 izračunajte $P(T | +)$ i nacrtajte ga u ovisnosti o udjelu $P(T)$.
- (b) Odredite najmanji udio na toj mreži za koji pozitivan signal znači vjerojatnost prisutnosti tvari barem 0.8.
- (c) Provjerite rezultat rješavanjem odgovarajuće nejednakosti na papiru i objasnite moguću razliku zbog mreže koraka 0.001.