

1. del: Osnove verjetnosti

1) Definicija verjetnosti.

Zanima nas verjetnost, da pri enem metu kocke pade šestica.

- Kako to lahko ugotovimo (empirično)?

- Kako to lahko izračunamo (teoretično)?

Za simulacijo empirične ocene verjetnosti, *da na kocki pade šestica* bomo uporabljali aktivne spletne strani "Virtual Laboratories in Probability and Statistics" (<http://www.math.uah.edu/stat/bernoulli/index.xhtml>) in odprite poskus (**Applet**) Basic Coin Experiment.

- nastavite vrednost $p=0.17$ (pribl. $1/6$)
- enkrat vrzite kocko (nastavite $n=1$, kliknite na ▶), rdeča barva označuje, da je pri metu padla šestica (zelena vse ostale)

empirična ocena verjetnosti za šestico: _____

- Simulirajte met 10 kock (nastavite $n=10$, kliknite na ▶).

empirična ocena verjetnosti za šestico: _____

- Simulirajte met 100 kock (nastavite $n=100$, kliknite na ▶).

empirična ocena verjetnosti za šestico: _____

- Kakšna bi bila odstopanja med empiričnimi in teoretičnimi vrednostmi, če bi poizkus ponovili še večkrat (npr. 10.000 krat)?

2) Pogojna verjetnost.

S spletno anketo med študenti smo dobili naslednje podatke o igranju videoigric:

Število

		Igrice		Vsota
		da	ne	
Spol	moski	21	22	43
	zenski	12	85	97
Vsota		33	107	140

Kolikšna je verjetnost, da...	Formula	Izračun
je študent moškega spola?		
študent igra videoigrice (ne glede na spol)?		
videoigrice igra študent, ki je moškega spola?		
je študent moškega spola in igra videoigrice?		
je študent moškega spola ali igra videoigrice?		
je študent, ki igra videoigrice, moškega spola?		

- Poznamo samo verjetnosti, da je študent moškega spola $P(M)$, in, da nekdo igra videoigrice $P(I)$ (za oba spola skupaj). Ali bi lahko izračunali verjetnost, da
 - je študent moškega spola in igra videoigrice?
 - je študent moškega spola ali igra videoigrice?

Ali sta spol in igranje videoigric **neodvisni spremenljivki**?

3) Pogojna verjetnost.

Pri predmetu Statistika v medicini imajo študenti možnost opravljati izpit na rednem in izrednem roku. Redni rok je takoj po končanih predavanjih in vajah, izredni pa čez eno leto. V preteklosti je bil delež študentov, ki so opravili izpit v prvem poizkusu 48%, na redni rok pa je prišlo 40% študentov. Odgovorite na spodnja vprašanja.

- Ob predpostavki, da sta zgornja dogodka neodvisna, izračunajte pričakovani delež študentov, ki na rednem roku opravijo izpit.
- Ali se vam zdi predpostavka o neodvisnosti dogodkov smiselna? Komentirajte.
- Kako bi odgovorili na prvo vprašanje, če veste, da izmed študentov, ki pridejo na redni rok, 75% študentov opravi izpit v prvem poizkusu?
- Kakšen delež vseh študentov opravi izpit v prvem poskusu na izrednem roku?
- Kakšen delež študentov, ki opravlja izpit v izrednem roku, opravi izpit v prvem poskusu?

4) Formula o popolni verjetnosti.

Z obsežnimi raziskavami so ugotovili, da neko zdravilo A v tednu dni ozdravi 0,7 bolnikov, zdravilo B 0,6 bolnikov, brez zdravljenja (s placebo) je delež ozdravljenih v enakem času le 0,2. Načrtujemo raziskavo, v kateri bomo opazovali neželene učinke opisanih zdravil. V ta namen bomo bolnike naključno razdelili v tri enako velike skupine (zdravilo A, zdravilo B, placebo). Kakšna je verjetnosti ozdravitve posameznega bolnika (kakšen delež ozdravljenih bolnikov pričakujemo)?

Rešitev:

$$P(Z / A) = 0,7$$

$$P(Z / B) = 0,7$$

$$P(Z / placebo) = 0,2$$

$$P(A) = 1/3$$

$$P(B) = 1/3$$

$$P(placebo) = 1/3$$

$$\begin{aligned} P(Z) &= P(Z / A) * P(A) + P(Z / B) * P(B) + P(Z / placebo) * P(placebo) = \\ &= 0,7 * (1/3) + 0,6 * (1/3) + 0,2 * (1/3) = 0,5 \end{aligned}$$

Verjetnost, da je ozdravel slučajno izbrani bolnik, za katerega ne vemo, ali je dobil zdravilo A, zdravilo B ali pa placebo, je 0,5.

Pri izračunu smo upoštevali povezavo skupne in pogojne verjetnosti ter načelo izračuna skupne verjetnosti izključujočih se dogodkov.

2. del: Diagnostični testi

5) Diagnostični test.

Odprite datoteko **DiagTest.sav**, kjer so podani rezultati diagnostičnega testa. Podatke predstavite s pomočjo **kontingenčne tabele** (postopek **Analyze** → **Descriptive Statistics** → **Crosstab...**)

Count

	Zajetje		Total
	cisto	fekalije	
Test -			
+			
Total			

Kolikšna je verjetnost, da...	Formula	Izračun
imajo zajetja pozitiven test?		
imajo zajetja negativen test?		
imajo zajetja s pozitivnim testom fekalije?		
so zajetja z negativnim testom čista?		

- Koliko zajetij so določili napačno?
- Koliko lažno pozitivnih in koliko lažno negativnih zajetij je bilo v raziskavi?

- **Občutljivost** testa (sensitivity)

- **Specifičnost** testa (specificity)

- **Pozitivna napovedna vrednost** testa (positive predictive value)

- **Negativna napovedna vrednost** testa (negative predictive value)

- Kolikšna je pogostost prisotnosti fekalij na našem vzorcu?

- Ali mislite, da je vzorec, na podlagi katerega so ocenili diagnostični test, **slučajni vzorec iz populacije zajetij**? Če ni, zakaj so se raziskovalci odločili, da bodo izbrali vzorec, ki ni slučajni?

- Katere vrednosti bi se spremenile in kako, če bi bila pogostost prisotnosti fekalij v populaciji zajetij 0,01?

- Ali bi uporabili ta test rutinsko (testiramo vsako zajetje)?

- Ali bi uporabili ta test rutinsko za zajetja, v okolici katerih ljudje zbolevalo (za ta zajetja ocenite, da je verjetnost prisotnosti fekalij 0,75)?

Povzetek

- 6) Recimo, da je verjetnost nekega dogodka 0,9. Kaj se nam bo zgodilo pri enem poskusu? Kaj pa če poizkusimo velikokrat?

- 7) Kako na vzorcu ocenimo verjetnost?

- 8) Razložite neodvisnost dogodkov. Kaj pomeni, da sta dve spremenljivki neodvisni? Razložite na konkretnem primeru 2 spremenljivk.

- 9) Kaj pomeni slučajni (reprezentativni vzorec) iz populacije? Navedite primer vzorca, ki to je, in primer, ki to ni.

- 10) Kaj pove občutljivost testa in kaj pozitivna napovedna vrednost testa? V čem je razlika? Katera izmed vrednosti je za zajetje bolj informativna?