Bayesovsko zaključivanje za proporciju binomne razdiobe

Perić, Josipa

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Josipa Perić

BAYESOVSKO ZAKLJUČIVANJE ZA PROPORCIJU BINOMNE RAZDIOBE

Diplomski rad

Voditelj rada:
doc. dr. sc. Snježana Lubura Strunjak

Zagreb, rujan 2022.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Mami i tati, za ljubav, pomoć, podršku, za sve

Sadržaj

Uvod

Možete li zamisliti bilo kakvo istraživanje bez primjene statistčkih metoda? Već samo prikupljanje podataka, razna mjerenja i prebrojavanja, prikazi rezultata te ono najbitnije, statistička analiza podataka i donošenje zaključaka, neizbježno zahtijevaju primjenu statistike. Upravo zato je statistika jedna od najrasprostranjenijih grana matematike. Činjenica je da je u modernom svijetu sve više podataka, a u svom tom bogatstvu informacija leže i brojna pitanja koja čekaju na odgovore. Stoga se lako nađemo u nedoumici pri odabiru statističke metode kojom bismo obradili podatke.

Poznajemo dvije vrste pristupa statističkoj analizi: frekvencionistički i Bayesovski. Razlika je u načinu interpretacije vjerojatnosti. Prvi, klasični i većini dobro poznat pristup, vjerojatnost smatra limesom relativnih frekvencija ponavljanog pokusa. Međutim, ako promatramo događaje ili fenomene koji se rijetko pojavlju, ovaj prisup nije moguć. Tada se zbog nedovoljne količine podataka oslanjamo na teorijske rezultate. S druge strane, Bayesovski pristup vjerojatnost smatra subjektivnom. Promatra je kao stupanj uvjerenja koji se mijenja dolaskom novih informacija. Nadalje, jedna od osnovnih razlika Bayesovske statistike u odnosu na frekvencionističku je ta što se u Bayesovskoj statistici parametri u statističkom modelu smatraju slučajnim varijablama, dok se u klasičnom pristupu tretiraju kao fiksne vrijednosti. U Bayesovskoj statistici i prije nego što promotrimo podatke imamo subjektivna vjerovanja o parametru koja su sadržana u apriornoj distribuciji parametra. Na temelju podataka te korištenjem Bayesova teorma, revidiramo uvjerenja o parametru. Time dobivamo aposteriornu distribuciju parametra.

Bayesovska statistika je dobila ime po Thomasu Bayesu (1702.-1761.), engleskom matematičaru i teologu koji je prvi dokazao specijalni slučaj teorema koji nosi njegovo ime: Bayesov teorem. Međutim, Bayes nikad nije objavio ono što je postalo njegovo najpoznatije postignuće.

Njegov prijatelj Richard Price pronašao je Bayesove spise te posthumno, 1763., Royal Society objavljuje njegov rad *Essay Towards Solving a Problem in the Doctrine of Chances*. Zanimljivo, bez znanja o radu Thomasa Bayesa, francuski matematičar i astronom Pierre-Simon Laplace otkriva Bayesov teorem (1774.), ali u općenitijem obliku.

Bayesovska statistika je bila zapostavljena dugi niz godina jer nije bilo moguće izračunati izraze koji se javljaju u Bayesovu teoremu, ali razvojem računala dolazi do naglog porasta popularnosti Bayesovske statistike. Ona se danas primjenjuje u raznim znanstvenim disciplinama.

Cilj ovog rada je dati pregled teorije na kojoj se zasniva Bayesovska statistika, usporediti frekvencionistički i Bayesovski pristup statistici te opisati metode koje vode do Bayesovskih zaključaka o proporciji binomne razdiobe.

Poglavlje 1

Osnovni pojmovi vjerojatnosti i statistike

Za početak ćemo uvesti osnovne pojmove koji će nam biti potrebni u daljnjem radu. Većina teorema i definicija preuzeto je iz [5] i [6] .

1.1 Vjerojatnosni prostor

Definicija 1.1.1. *Neka je Ω neprazan skup. Familija podskupova $\mathcal{F}$ od Ω zove se σ-algebra, ako vrijede sljedeća tri svojstva:*

i) $\Omega \in \mathcal{F}$;

ii) Ako je $A \in \mathcal{F}$, onda je i $A^c \in \mathcal{F}$ (zatvorenost na komplement);

iii) Ako su $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, onda je i $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ (zatvorenost na prebrojive unije) .

Definicija 1.1.2. *Neka je Ω neprazan skup i $\mathcal{F}$ σ-algebra događaja. Vjerojatnost na izmjerivom prostoru $(\Omega, \mathcal{F})$ je funkcija $\mathbb{P}$: $\mathcal{F} \to [0, 1]$ koja zadovoljava sljedeća tri aksioma:*

(A1) Za sve $A \in \mathcal{F}$, $P(A) \geq 0$ (nenegativnost);

(A2) $P(\Omega) = 1$ (normiranost);

(A3) Za svaki niz A_i, $i \in \mathbb{N}$, po parovima disjunktnih događaja $A_i \in \mathcal{F}$ ($A_i \cap A_j = \emptyset$ za $i \neq j$) vrijedi:

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

Uređena trojka $(\Omega, \mathcal{F}, \mathbb{P})$ zove se vjerojatnosni prostor.

Elemente σ-algebre $\mathcal{F}$ zovemo događaji, a broj $\mathbb{P}(A)$, A $\in \mathcal{F}$, zovemo vjerojatnost događaja A.

Propozicija 1.1.3. *(Osnovna svojstva vjerojatnosti)*

a) $\mathbb{P}(\emptyset) = 0$;

b) Za svaki $n \in \mathbb{N}$ i za svaki konačan niz $A_1, \ldots, A_n$ po parovima disjunktnih događaja iz $\mathcal{F}$ vrijedi:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

To se svojstvo naziva konačna aditivnost.

c) $\mathbb{P}(A^c)$=1 $- \mathbb{P}(A)$;

d) $\mathbb{P}(A \setminus B)$=$\mathbb{P}(A) - \mathbb{P}(A \cap B)$;

e) Ako su A,B $\in \mathcal{F}$ i $A \subseteq B$, onda je $\mathbb{P}(A) \leq \mathbb{P}(B)$. Ovo svojstvo zovemo monotonost.

f) Za vjerojatnost unije ne nužno disjunktnih događaja vrijedi:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

g) (Sylvestrova formula) To je generalizacija prethodnog rezultata na više od dva događaja. Neka su $A_1, A_2, \ldots, A_n$ događaji u $\mathcal{F}$. Tada vrijedi:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \mathbb{P}(A_{i1} \cap A_{i2} \cap \ldots \cap A_{ik})$$

h) ($\sigma -$ subaditivnost) Neka je $(A_i)_{i \in \mathbb{N}}$ niz događaja iz $\mathcal{F}$. Tada vrijedi:

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

Red na desnoj strani može konvergirati u $+\infty$.

Definicija 1.1.4. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnostni prostor.*

a) Događaji A,B $\in \mathcal{F}$ su nezavisni ako vrijedi:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

b) Familija događaja $(A_i)_{i\in I}$ (konačna, prebrojiva ili neprebrojiva) je nezavisna ako vrijedi:

$$\mathbb{P}\Bigg(\bigcap_{i\in\mathcal{F}} A_i\Bigg) = \prod_{i\in\mathcal{F}} \mathbb{P}(A_i),$$

za svaki konačan podskup $F \subseteq I$.

c) Familija događaja $(A_i)_{i\in I}$ je po parovima nezavisna ako za sve i,j $\in$ I, i $\neq$j, vrijedi $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$.

Nezavisni događaji najčešće se javljaju u primjenama kao rezultati nezavisnih pokusa.

1.2 Slučajne varijable

Vjerojatnosni prostor predstavlja matematički model za slučajni pokus, a on ima više mogućih ishoda. Ishode slučajnog pokusa zovemo događajima. Intuitivno, slučajni pokus je svaki pokus (proces, opažanje, mjerenje) kojemu se ishod ne može sa sigurnošću predvidjeti. Sada ćemo se upoznati sa slučajnim varijablama, a to su funkcije koje svakom elementarnom događaju pridružuju neki broj.

Definicija 1.2.1. *Neka je $\mathbb{R}$ skup svih realnih brojeva, te neka je $\mathcal{B}$ σ-algebra generirana familijom svih otvorenih skupova na $\mathbb{R}$.*

i) $\mathcal{B}$ zovemo σ-algebrom Borelovih skupova na $\mathbb{R}$, ili Borelovom σ-algebrom, a njene elemente nazivamo Borelovi skupovi.

ii) Funkcija $g : \mathbb{R} \to \mathbb{R}$ je Borelova funkcija ako je $g^{-1}(B) \in \mathcal{B}$ za svaki B iz $\mathcal{B}$, odnosno $g^{-1}(\mathcal{B}) \subseteq \mathcal{B}$.

Za k-dimenzionalne realne prostore, Borelovu σ-algebru generiranu svim otvorenim pravokutnicima na $\mathbb{R}^k$ označavamo sa $\mathcal{B}(\mathbb{R}^k)$.

Definicija 1.2.2. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor. Izmjerivo preslikavanje X: $(\Omega, \mathcal{F}) \to (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ za koje vrijedi $X^{-1}(B) \in \mathcal{F}$ za svaki $B \in \mathcal{B}(\mathbb{R}^k)$ naziva se slučajna varijabla (za k=1) ili slučajni vektor (za k$\geq$ 2).*

Definicija 1.2.3. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor,* $(\mathbb{R}, \mathcal{B})$ *izmjeriv prostor i X slučajna varijabla na* Ω*. Funkciju* $\mathbb{P}_X : \mathcal{B} \to [0, 1]$ *definiranu s*

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B))$$

nazivamo vjerojatnosnom mjerom induciranom s X, a vjerojatnosni prostor $(\Omega, \mathcal{B}, \mathbb{P}_X)$ *zovemo vjerojatnosnim prostorom induciranim s X. Lako se vidi da je* $\mathbb{P}_X$ *vjerojatnost, odnosno vjerojatnosna mjera na* $\mathcal{B}$*.*

Definicija 1.2.4. *Neka je X slučajna varijabla definirana na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$*. Funkcija distribucije od X je funkcija* $F_X : \mathbb{R} \to [0, 1]$ *definirana s*

$$F_X(x) = \mathbb{P}_X(\langle -\infty, x]) = \mathbb{P}(X \in \langle -\infty, x]) = \mathbb{P}(X \leq x), x \in \mathbb{R}.$$

Teorem 1.2.5. *Neka je X slučajna varijabla definirana na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$ *te neka je F pripadna funkcija distribucije. Tada vrijedi*

i) F je neopadajuća;

ii) F je neprekidna zdesna u svakoj točki $x \in \mathbb{R}$*;*

ii) F ima limes slijeva u svakoj točki $x \in \mathbb{R}$*;*

iv) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ *i* $F(\infty) = \lim_{x \to +\infty} F(x) = 1$.

Definicija 1.2.6. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor. Funkcija* $X : \Omega \to \mathbb{R}$ *zove se diskretna slučajna varijabla ako postoji konačan ili prebrojiv skup* $D = \{x_1, x_2, \ldots\} \subset \mathbb{R}$ *takav da je* $\mathbb{P}_X(D) = \mathbb{P}(X \in D) = 1$*.*

Diskretne slučajne varijable obično zadajemo tako da zadamo skup $D = \{x_1, x_2, \ldots\}$ i brojeve $p_n = \mathbb{P}(X = x_n)$.

Za diskretnu slučajnu varijablu koja može poprimiti vrijednosti $x_1, x_2, \ldots$, funkcija vjerojatnosti (gustoće) je funkcija za koju vrijedi:

$$f(x_i) \geq 0,$$

$$\sum_{i=1}^{n} f(x_i) = 1,$$

$$f(x_i) = \mathbb{P}(X = x_i).$$

Nadalje, kumulativna funkcija distribucije od X je funkcija $F_X : \mathbb{R} \to \mathbb{R}$,
$F_X(x) = \mathbb{P}(X \leq x)$.
Vrijedi:

$$F_X(x) = \sum_{x_i \leq x} f(x_i).$$

Definicija 1.2.7. *Slučajna varijabla $X : \Omega \to \mathbb{R}$ je apsolutno neprekidna ako postoji funkcija $f : \mathbb{R} \to [0, \infty\rangle$ takva da za sve $x \in \mathbb{R}$ vrijedi:*

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

Funkcija f se zove funkcija gustoće od X.

Napomenimo da je integral na desnoj strani u gornjoj definiciji tzv. Lebesgueov integral kojeg nećemo definirati u ovom radu. U slučaju da funkcija f ima najviše prebrojivo mnogo prekida, onda se taj integral podudara s klasičnim Riemannovim integralom. Općenito, ako je f Riemann integrabilna, onda je ona i Lebesgue integrabilna.

Definicija 1.2.8. *Neka su $X_1, \dots, X_n$ slučajne varijable na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Kažemo da su $X_1, \dots, X_n$ nezavisne ako za proizvoljne $B_i \in \mathcal{B}, i = 1, \dots, n$ vrijedi:*

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}\left(\bigcap_{i=1}^{n} (X_i \in B_i)\right) = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i).$$

Definicija 1.2.9. *Neka je X apsolutno neprekidna slučajna varijabla s funkcijom gustoće f. Ako je*

$$\int_{-\infty}^{+\infty} |x| f(x)dx < \infty,$$

onda postoji matematičko očekivanje od X koje definiramo s

$$\mathbb{E}(X) := \int_{-\infty}^{+\infty} x f(x)dx,$$

te ga najčešće označavamo s μ.

Definicija 1.2.10. *Neka je X slučajna varijabla na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$ *i* $r > 0$.

i) $\mathbb{E}(X^r)$ *zovemo r-ti moment od X, a* $\mathbb{E}|X^r|$ *zovemo r-ti apsolutni moment od X. Dogovorno vrijedi* $\mathbb{E}(X^0) = \mathbb{E}|X^0| = 1$.

ii) Neka $\mathbb{E}X$ *postoji, tj. konačno je. Tada* $\mathbb{E}((X - \mathbb{E}X)^r)$ *zovemo r-ti centralni moment od X, a* $\mathbb{E}(|X - \mathbb{E}X|^r)$ *zovemo r-ti apsolutni centralni moment od X.*

iii) Varijanca od X, koju označavamo s Var(X) ili σ^2, *je drugi centralni moment od X, dakle:*
$$Var(X) = \mathbb{E}\left((X - \mathbb{E}X)^2\right).$$
Pozitivan drugi korijen iz varijance zovemo standardna devijacija od X i označavamo sa σ.

Koristeći definiciju matematičkog očekivanja i varijance, dobivamo:

$$Var(X) = \mathbb{E}((X - \mathbb{E}X)^2) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx.$$

Kvadriranjem prethodnog izraza dobivamo:

$$Var(X) = \int_{-\infty}^{+\infty} (x^2 - 2\mu x + \mu^2) f(x)dx.$$

Upotrijebimo li linearnost integrala dobit ćemo:

$$Var(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - 2\mu \int_{-\infty}^{+\infty} x f(x)dx + \mu^2 \int_{-\infty}^{+\infty} f(x)dx,$$

iz čega slijedi često korišten izraz za varijancu:

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2.$$

1.3 Neprekidni slučajni vektori

Sada ćemo teoriju proširiti na slučajne vektore kako bismo mogli definirati pojam uvjetne funkcije gustoće koju koristimo u neprekidnoj verziji Bayesovog teorema.

Definicija 1.3.1. *Neka su X i Y dvije slučajne varijable definirane na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$. *Uređeni par (X,Y) nazivamo (dvodimenzionalni) slučajni vektor.*

Definicija 1.3.2. *Neka je (X,Y) slučajan vektor definiran na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathbb{P})$. *Funkcija distribucije slučajnog vektora (X,Y) je funkcija* $F : \mathbb{R}^2 \to [0,1]$ *definirana formulom:*

$$F(x,y) = \mathbb{P}((X,Y) \leq (x,y)), \qquad (x,y) \in \mathbb{R}^2.$$

Definicija 1.3.3. *Slučajan vektor (X,Y) je apsolutno neprekidan ako postoji* $f : \mathbb{R}^2 \to [0,\infty\rangle$ *takva da za sve (x,y)* $\in \mathbb{R}^2$ *vrijedi*

$$F(x,y) = \mathbb{P}((X,Y) \leq (x,y)) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t,u)dudt.$$

Funkciju f nazivamo funkcija gustoće slučajnog vektora (X,Y).

Definicija 1.3.4. *Neka je (X,Y) slučajan vektor s funkcijom gustoće* $f = f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$. *Ako je* $y \in \mathbb{R}$ *takav da je* $f_Y(y) > 0$, *onda definiramo:*

i) Uvjetnu funkciju gustoće

$$f(\cdot|y) = f_{X|Y}(\cdot|y) : \mathbb{R} \to \mathbb{R}$$

slučajne varijable X uz uvjet Y=y formulom

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}. \tag{1.1}$$

ii) Uvjetnu razdiobu $\mathbb{P}_{X|Y} : \mathcal{B} \to [0,1]$ *od X uz dano Y=y kao*

$$\mathbb{P}_{X|Y}(X \in B | Y = y) = \int_B f_{X|Y}(x|y)dx. \tag{1.2}$$

Ako u (1.2) uvrstimo $B = \langle -\infty, x]$, $x \in \mathbb{R}$, dobivamo

$$\mathbb{P}_{X|Y}(X \leq x | Y = y) = \int_{-\infty}^{x} f_{X|Y}(t|y)dt\,,$$

stoga za svaki $x \in \mathbb{R}$ vrijedi

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(t|y)dt, \tag{1.3}$$

čime smo dobili uvjetnu distribuciju od X uz dano Y=y.

Definicija 1.3.5. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor i neka je X slučajna varijabla na Ω s funkcijom distribucije F. Slučajni uzorak duljine n za X je niz od n nezavisnih, jednako distribuiranih slučajnih varijabli $X_1, X_2, \ldots, X_n$ kojima je distribucija jednaka (populacijskoj) razdiobi varijable X. Realizaciju slučajnog uzorka, odnosno opažene vrijednosti x_i od X_i, $i = 1, 2, \ldots, n$ zovemo uzorkom.*

Definicija 1.3.6. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor i neka je $\mathcal{P}$ neka familija vjerojatnosti na $(\Omega, \mathcal{F})$. Uređena trojka $(\Omega, \mathcal{F}, \mathcal{P})$ naziva se statistička struktura. Familija $\mathcal{P}$ često je parametrizirana i zapisuje se u obliku :*

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}.$$

Definicija 1.3.7. *Statistika na statističkoj strukturi $(\Omega, \mathcal{F}, \mathcal{P})$ je svaka slučajna varijabla (ili slučajni vektor za $k \geq 2$) $T : \Omega \to \mathbb{R}^k$ takva da za neki $n \in \mathbb{N}$ postoji n-dimenzionalni slučajni vektor $(X_1, X_2, \ldots, X_n)$ na $(\Omega, \mathcal{F}, \mathcal{P})$ te izmjeriva funkcija $t : \mathbb{R}^n \to \mathbb{R}^k$ takva da je $T = t(X_1, X_2, \ldots, X_n)$.*

Definicija 1.3.8. *Neka je $(x_1, x_2, \ldots, x_n)$ jedna realizacija slučajnog uzorka X=$(X_1, X_2. \ldots, X_n)$ sa zajedničkom funkcijom gustoće $f(x \mid \theta), x \in \mathbb{R}^n$ te neka je $\theta \in \Theta$ nepoznati parametar. Funkcija vjerodostojnosti je funkcija $L: \Theta \to \mathbb{R}$ definirana s*

$$L(\Theta) := \prod_{i=1}^{n} f(x_i|\theta).$$

Slučajni uzorak duljine n odgovara nizu n nezavisnih mjerenja nekog slučajnog svojstva u populaciji koje se opisuje slučajnom varijablom X.

1.4 Bayesov teorem

Sada ćemo uvesti Bayesov teorem, odnosno njegovu diskretnu i neprekidnu verziju te sve potrebne pojmove. Glavna razlika u odnosu na klasične metode statistike koje poznajemo jest u tome da parametre u ovom slučaju promatramo kao slučajne varijable, dok su u klasičnom frekvencionističkom pristupu parametri fiksirani. Bayesovski pristup statistici temelji se na subjektivnoj vjerojatnosti, što ćemo često naglašavati tokom daljnjeg rada. Teorem nam pomaže da revidiramo svoja predviđanja o nekom događaju na temelju novih saznanja o tom događaju.

DISKRETNA VERZIJA

Pretpostavimo da je Ω konačan ili prebrojiv skup.

Definicija 1.4.1. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor te* $B \in \mathcal{F}$ *događaj takav da je* $\mathbb{P}(B) > 0$. *Uvjetna vjerojatnost događaja A uz dano B definira se formulom:*

$$\mathbb{P}(A|B) = \mathbb{P}_B(A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, A \in \mathcal{F}. \tag{1.4}$$

Propozicija 1.4.2. *Neka je* $B \in \mathcal{F}$ *takav da je* $\mathbb{P}(B) > 0$. *Tada je* $\mathbb{P}_B : \mathcal{F} \to [0, 1]$ *vjerojatnost na* $(\Omega, \mathcal{F})$.

Dokaz. Dokaz slijedi iz definicije, provjeravanjem aksioma funkcije vjerojatnosti $\mathbb{P}_B$. □

Definicija 1.4.3. *Događaji A,B* $\in \mathcal{F}$ *su uvjetno nezavisni uz dano* $C \in \mathcal{F}$, $\mathbb{P}(C) > 0$, *ako vrijedi:* $\mathbb{P}(A \cap B\,|C) = \mathbb{P}(A|C)\mathbb{P}(B|C)$.

Događaj B još možemo zapisati kao $B = (A \cap B) \cup (A^c \cap B)$. Budući da su $(A \cap B)$ i $(A^c \cap B)$ disjunktni događaji, imamo :

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B).$$

Uvrstimo u (1.4) i te dobivamo:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)}. \tag{1.5}$$

Analogno definiramo uvjetnu vjerojatnost događaja B uz dano A:

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

Sada izrazimo $\mathbb{P}(A \cap B)$:

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\,\mathbb{P}(A).$$

Slično možemo napisti $\mathbb{P}(A^c \cap B)$ te uvrstimo li to u (1.5) dobijemo:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B|A)\,\mathbb{P}(A) + \mathbb{P}(B|A^c)\,\mathbb{P}(A^c)}. \tag{1.6}$$

Time smo dobili Bayesovu formulu za jedan događaj, a možemo ju poopćiti na proizvoljnu familiju disjunktnih događaja.

Definicija 1.4.4. *Neka je $(H_i)_{i\in I}$ konačna ili prebrojiva familija događaja iz $\mathcal{F}$ takva da je:*

i) $\mathbb{P}(H_i) > 0$ *za sve* $i \in I$;

ii) $H_i \cap H_j = \emptyset$ *za* $i \neq j$;

iii) $\cup_{i\in I}^{+\infty} H_i = \Omega$.

Takvu familiju zovemo potpun sustav događaja, a događaje H_i *zovemo hipotezama.*

Propozicija 1.4.5. *(Formula potpune vjerojatnosti) Neka je $(H_i)_{i\in I}$ potpun sustav događaja. Tada za svaki $A \in \mathcal{F}$ vrijedi*

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(H_i)\mathbb{P}(A|H_i). \tag{1.7}$$

Dokaz. Računamo:

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}\Bigg(A \cap \Bigg(\bigcup_{i\in I} H_i\Bigg)\Bigg)\\ &= \mathbb{P}\Bigg(\bigcup_{i\in I}(A \cap H_i)\Bigg)\\ &= \sum_{i\in I} \mathbb{P}(A \cap H_i)\\ &= \sum_{i\in I} \mathbb{P}(H_i)\mathbb{P}(A|H_i),\end{aligned}$$

gdje redom koristimo činjenice da je $\cup_{i\in I} H_i = \Omega$, σ-aditivnost te identitet $\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B)$.

□

Teorem 1.4.6. *(Bayesov teorem) Neka je $(H_i)_{i\in I}$ potpun sustav događaja na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Tada za svaki $A \in \mathcal{F}$ takav da je $\mathbb{P}(A) > 0$ vrijedi:*

$$\mathbb{P}(H_j|A) = \frac{\mathbb{P}(H_j)\mathbb{P}(A|H_j)}{\sum_{i\in I} \mathbb{P}(H_i)\mathbb{P}(A|H_i)}. \tag{1.8}$$

Dokaz. Računamo:

$$\begin{aligned}\mathbb{P}(H_j|A) &= \frac{\mathbb{P}(H_j \cap A)}{\mathbb{P}(A}\\ &= \frac{\mathbb{P}(H_j)\mathbb{P}(A|H_j)}{\mathbb{P}(A)}\end{aligned}$$

$$= \frac{\mathbb{P}(H_j)\mathbb{P}(A|H_j)}{\sum_{i\in I}\mathbb{P}(H_i)\mathbb{P}(A|H_i)}$$

U prvoj jednakosti koristili smo definiciju uvjetne vjerojatnosti, u drugoj jednakosti svojstvo $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A)$, a u posljednjoj jednakosti formulu potpune vjerojatnosti (1.7). □

Vjerojatnosti $\mathbb{P}(H_i)$ zovemo *apriorne vjerojatnosti* događaja H_i, a one predstavljaju informacije koje imamo o događajima H_i prije nego što znamo da se događaj A dogodio. Vjerojatnosti $\mathbb{P}(H_i|A)$ zovemo *aposteriorne vjerojatnosti* događaja H_i uz dano A. One predstavljaju informacije koje imamo o događajima H_i kada znamo da se događaj A dogodio.

NEPREKIDNA VERZIJA

Bayesov teorem često koristimo i u neprekidnom obliku stoga uvodimo i tu verziju teorema.

Teorem 1.4.7. *(Neprekidna formula potpune vjerojatnosti) Neka je B proizvoljan Borelov skup i neka su X i Y neprekdine slučajne varijable. Tada vrijedi*

$$\mathbb{P}(X \in B) = \int_{\mathbb{R}} \mathbb{P}(X \in B|Y = y)\, f_y(y)\, dy. \tag{1.9}$$

Dokaz. Računamo:

$$\mathbb{P}(X \in B) = \int_{\mathbb{R}}\int_B f_{X,Y}(x,y)\, dxdy = \int_{\mathbb{R}}\int_B f_{X|Y}(x|y)\, f_y(y)\, dxdy$$

$$= \int_{\mathbb{R}}\left(\int_B f_{X|Y}(x|y)\, dx\right) f_y(y)\, dy = \int_{\mathbb{R}} \mathbb{P}(X \in B|Y = y)\, f_y(y)\, dy.$$

□

Teorem 1.4.8. *(Neprekidna verzija Bayesovog teorema) Neka su X i Y neprekidne slučajne varijable. Tada za svaki $x \in \mathbb{R}$ i svaki $y \in \mathbb{R}$ takav da je $f_Y(y) > 0$ vrijedi :*

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)\, f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)\, f_X(x)}{\int_{\mathbb{R}} f_{Y|X}(y|x)\, f_X(x)\, dx}. \tag{1.10}$$

Dokaz. Koristeći definiciju uvjetne funkcije gustoće dobivamo

$$f_{X,Y}(x,y) = f_{Y|X} f_X(y|x)\, f_X(x).$$

Slijedi:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)\, f_X(x)}{f_Y(y)},$$

pri čemu zadnja jednakost slijedi korištenjem neprekidne verzije formule potpune vjerojatnosti. □

Uočimo da je analogon apriornih vjerojatnosti, odnosno $\mathbb{P}(H_i)$ upravo funkcija gustoće $f_X(x)$, dok je analogon aposteriornih vjerojatnosti uvjetna funkcija gustoće $f_{X|Y}(x|y)$.

1.5 Primjeri važnih distribucija

Najvažnije svojstvo slučajnih varijabli, kada govorimo o njihovoj primjeni, je njihova distribucija opisana gustoćom ili funkcijom distribucije. Sada ćemo navesti neke slučajne varijable koje smo koristili u daljnjem radu te njihove karakteristike i interpretaciju.

DISKRETNE DISTRIBUCIJE

Diskretne slučajne varijable interpretiramo kao broj nečega, odnosno kao rezultat nekog procesa prebrojavanja (npr. broj uspjeha i slično).

Bernoullijeva distribucija

Bernoullijeva slučajna varijabla X prati je li rezultat slučajnog pokusa bio "uspjeh" ili "neuspjeh". Preciznije, X poprima vrijednost 1 ako se dogodio uspjeh, inače poprima vrijednost 0. Dakle, vrijednosti koje poprima slučajna varijabla X su 0 ili 1. Označimo sa $\theta = \mathbb{P}(X = 1)$ vjerojatnost uspjeha. Primjetimo da nužno vrijedi $\theta \in [0, 1]$. Tada je funkcija vjerojatnosti od X jednaka:

$$f_X(x) = \theta^x (1-\theta)^{1-x},$$

te vrijedi

$$\mathbb{E}[X] = \theta,$$

$$Var[X] = \theta(1-\theta).$$

Primjerice, recimo da se u pokusu bacanja simetrične kocke dogodi uspjeh ako padne šestica. Tada slučajna varijabla ima vrijednost 1 kada se okrene šestica, a inače je jednaka 0, odnosno ima Bernoullijevu razdiobu s vjerojatnošću uspjeha $\frac{1}{6}$.

Binomna distribucija

Zamislimo slučajni pokus koji se sastoji od n nezavisnih Bernoullijevih pokusa. Svaki pokus koji ima samo dva moguća ishoda (”uspjeh” i ”neuspjeh”) je Bernoullijev. U svakom pokusu vjerojatnost uspjeha je θ. Nezavisnost znači da ishod svakog od njih ne ovisi o ishodima ostalih pokusa. Tada slučajna varijabla X, koja je jednaka ukupnom broju uspjeha u tih n pokusa, ima binomnu razdiobu s parametrima n i θ $(0 \leq \theta \leq 1)$. Vrijedi:

$$f_X(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x},$$

$$\mathbb{E}[X] = n\theta,$$

$$Var[X] = n\theta(1-\theta).$$

Iz navedene interpretacije binomne slučajne varijable slijedi da se ona može prikazati kao zbroj n nezavisnih jednako distribuiranih Bernoullijevih slučajnih varijabli.

Poissonova distribucija

Poissonova distribucija modelira broj slučajnih događaja koji se realiziraju tijekom nekog vremenskog intervala. Diskretna slučajna varijabla ima Poissonovu distribuciju s parametrom λ, $\lambda > 0$, i pišemo $X \sim P(\lambda)$, ako je njena funkcija gustoće oblika:

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}.$$

Vrijedi:

$$\mathbb{E}[X] = Var[X] = \lambda.$$

NEPREKIDNE DISTRIBUCIJE

Uniformna distribucija

Kažemo da neprekidna slučajna varijabla X ima uniformnu razdiobu na intervalu $\langle a, b\rangle$ ako joj je funkcija gustoće:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{za } x \in \langle a, b\rangle \\ 0 & \text{inače.} \end{cases}$$

Vrijedi:

$$\mathbb{E}[X] = \frac{a+b}{2},$$

$$Var[X] = \frac{(b-a)^2}{12}.$$

Gama distribucija

Kažemo da slučajna varijabla X ima Gama distribuciju s parametrima $\alpha > 0$ i $\lambda > 0$, i pišemo $X \sim \Gamma(\alpha, \lambda)$, ako je strogo pozitivna i njena funkcija gustoće je:

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{za } x > 0 \\ 0 & \text{inače,} \end{cases}$$

gdje je $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$ gama funkcija. Svojstva gama funkcije su:

i) $\Gamma(1) = 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ za $\alpha > 1$, odakle slijedi da je $\Gamma(n) = (n-1)!$ za prirodan broj n;

ii) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Vrijedi:

$$\mathbb{E}[X] = \frac{\alpha}{\lambda},$$

$$Var[X] = \frac{\alpha}{\lambda^2}.$$

Beta distribucija

Neprekidna slučajna varijabla X ima Beta distribuciju s parametrima $a > 0$ i $b > 0$ te pišemo $X \sim B(a, b)$, ako prima vrijednosti u intervalu $\langle 0, 1\rangle$ i njena funkcija gustoća je:

$$f_X(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} & \text{za } 0 < x < 1 \\ 0 & \text{inače.} \end{cases}$$

Primijetite da vrijedi:

$$\int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Desnu stranu te jednakosti označavamo sa $B(a, b)$. Funkcija koja paru pozitivnih brojeva (a, b) pridružuje tu vrijednost zove se Beta funkcija. Vrijedi:

$$\mathbb{E}[X] = \frac{a}{a+b},$$

$$Var[X] = \frac{ab}{(a+b)^2(a+b+1)}.$$

Uniformna razdioba na $\langle 0, 1\rangle$ je poseban slučaj Beta razdiobe kada su parametri a=b=1.

Poglavlje 2

Bayesovsko zaključivanje za diskretne slučajne varijable

U ovom poglavlju uvodimo Bayesov teorem za diskretne slučajne varijable. Konkretno, promotrit ćemo binomnu i Poissonovu slučajnu varijablu. Vidjet ćemo kako možemo korigirati naša vjerovanja o parametru uzimajući u obzir podatke koji su nam dani. Upravo to predstavlja način na koji dolazimo do statističkih zaključaka koristeći se Bayesovim pristupom statistici.

Dakle, parametar ćemo smatrati slučajnom varijablom X, koja ima moguće vrijednosti $x_1, \ldots, x_I$. Slučajna varijabla Y, koja ovisi o parametru, ima moguće vrijednosti $y_1, \ldots, y_J$. Želimo donijeti zaključke o parametru kojeg predstavlja slučajna varijabla X, a to činimo koristeći opažene vrijednosti $Y = y_j$ i Bayesov teorem. Bayesovski pristup obuhvaća sve moguće uređene parove (x_i, y_j) za $i = 1, \ldots, I$ i $j = 1, \ldots, J$. Treba imati na umu da slučajnu varijablu X i slučajnu varijablu Y ne promatramo na isti način. Događaji $(X = x_1), \ldots, (X = x_I)$ utječu na podjelu među svim mogućim ishodima, ali nikada ne promatramo koji x_i se točno dogodio. Dakle, ono što do sada znamo je da Bayesovski pristup vodi do podataka koje lako možemo svrstati u tablicu kao na slici ispod.

(x_1, y_1)	(x_1, y_2)	$\ldots$	(x_1, y_j)	$\ldots$	(x_1, y_J)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
(x_i, y_1)	(x_i, y_2)		(x_i, y_j)	$\ldots$	(x_i, y_J)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
(x_I, y_1)	(x_I, y_2)	$\ldots$	(x_I, y_j)		(x_I, y_J)

Tablica 2.1: Podaci svrstani u tablicu

Vjerojatnost je definirana u svim točkama koje su predočene u tablici. Sada ćemo djelomično promijeniti našu notaciju. Naime, koristit ćemo $f()$ kako bismo označili vjerojatnosnu distribuciju koja se odnosi na varijablu Y, a $g()$ kako bismo označili vjerojatnosnu distribuciju koja se odnosi na varijablu X, odnosno na parametar. To će nam pomoći da lakše razlikujemo Y i X. Svaku od zajedničkih vjerojatnosti nalazimo koristeći pravilo umnoška:

$$f(x_i\,,\,y_j) = g(x_i) \times f(y_j\,|\,x_i). \tag{2.1}$$

Marginalnu distribuciju od Y nađemo tako da sumiramo odgovarajuće stupce. To vidimo u narednoj tablici. Na lijevoj strani nalazi se apriorna distribucija.

	apriorna distribucija	y_1	$\cdots$	y_j	$\cdots$	y_J
x_1	$g(x_1)$	$f(x_1, y_1)$	$\cdots$	$f\left(x_1, y_j\right)$	$\cdots$	$f(x_1, y_J)$
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
x_i	$g(x_i)$	$f(x_i, y_1)$	$\cdots$	$f\left(x_i, y_j\right)$	$\cdots$	$f(x_i, y_J)$
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
x_I	$g(x_I)$	$f(x_I, y_1)$	$\cdots$	$f\left(x_I, y_j\right)$	$\cdots$	$f(x_I, y_J)$
		$f(y_1)$	. . .	$f\left(y_j\right)$	. . .	$f(y_J)$

Tablica 2.2: Zajedničke i marginalne distribucije od X i Y

Kada promatramo opaženu vrijednost $Y = y_j$, dolazimo do reducirane tablice koja sadrži sve uređene parove j-tog stupca. Vidjet ćemo kako su nam informacije iz tog stupca dovoljne za računanje traženih vjerojatnosti.

(x_1, y_j)
.
.
.
(x_i, y_j)
.
.
.
(x_I, y_j)

Tablica 2.3 : Reducirana tablica uz dano $Y = y_j$

Cilj nam je pronaći aposteriornu vjerojatnosnu funkciju gustoće od X uz dani $Y = y_j$:

$$g(x_i \,|\, y_j) = \frac{g(x_i) \times f(y_j \,|\, x_i)}{\sum_{i=1}^{n_i} g(x_i) \times f(y_j \,|\, x_i)}. \tag{2.2}$$

Objasnimo sada svaki dio formule zasebno:

- Apriorna distribucija diskretne slučajne varijable X dana je apriornom vjerojatnosnom funkcijom $g(x_i)$, za $i = 1, \ldots, n$. Ona predstavlja naše subjektivno vjerovanje o vjerojatnosti svakog x_i prije nego pogledamo dane podatke. Bitno je da to dolazi iz našeg prethodnog iskustva, a ne iz trenutačnih podataka. Jedan od najvećih problema Bayesovske statistike je adekvatan odabir apriorne distribucije; više o tome će biti riječ u narednim poglavljima.

- Nakon što smo opazili $Y = y_j$, računamo funkciju vjerodostojnosti diskretne slučajne varijable X. Funkcija vjerodostojnosti dana je u oznaci $f(y_j \,|\, x_i)$, za $i = 1, \ldots, n$. To je uvjetna vjerojatnosna funkcija od Y uz dano X=x_i izračunata za y_j. Za to vrijeme X varira duž cijelog dopustivog skupa vrijednosti $x_1, \ldots, x_n$.

- Aposteriorna vjerojatnosna distribucija diskretne slučajne varijable je dana aposteriornom vjerojatnosnom funkcijom $g(x_i|y_j)$ evaluiranom u x_i za $i = 1, \ldots, n$, uz dano $Y = y_j$.

Formula (2.2) daje metodu za revidiranje naših vjerovanja o vjerojatnosti mogućih vrijednosti slučajne varijable X.

Sljedeći primjer, kao i svi ostali, preuzeti su iz [1].

Primjer 1. *Neka je dana posuda u kojoj se nalazi 5 kuglica. Jedan dio je crvene boje, a drugi dio zelene boje, pri tome ne znamo koliko kojih ima. Neka slučajna varijabla X bude broj crvenih kuglica u posudi. Moguće vrijednost za X su* $x_i = i$, $i = 0, \ldots, 5$. *Budući da ne znamo broj crvenih kuglica u posudi, pretpostavit ćemo da su sve moguće vrijednosti jednako vjerojatne. Stoga je naša apriorna distribucija od X dana na sljedeći način: g(0)=g(1)=g(2)=g(3)=g(4)=g(5)=*$\frac{1}{6}$. *Na slučajan načn izvlačimo kuglicu iz posude. Neka je slučajna varijabla Y=1 ako smo izvukli crvenu kuglicu, u suprotnom (izvukli smo zelenu) neka je Y=0. Uvjetna opažena distribucija od* $Y \,|\, X$ *je* $\mathbb{P}(Y = 1 \,|\, X = x_i) = \frac{i}{5}$ *i* $\mathbb{P}(Y = 0 \,|\, X = x_i) = \frac{5-i}{5}$.

Pretpostavimo da je prvo izvučena crvena kuglica, stoga će jedini stupac koji je potreban u izračunu aposteriorne funkcije distribucije biti upravo stupac gdje je $y_j = 1$.

U sljedećim tablicama vidimo postupak koji vodi do traženog rezultata korištenjem formule (2.2):

$$f(x_i \,|\, y_j) = \mathbb{P}(X = x_i \,|\, Y = y_j) = \frac{g(x_i) \times f(y_j \,|\, x_i)}{\sum_{i=1}^{n_i} g(x_i) \times f(y_j \,|\, x_i)}$$

x_i	apriorne vjerojatnosti	$y_j = 0$	$y_j = 1$	aposteriorne vjerojatnosti
0	1/6	$\frac{1}{6} \times \frac{5}{5} = \frac{5}{30}$	$\frac{1}{6} \times \frac{0}{5} = 0$	0
1	1/6	$\frac{1}{6} \times \frac{4}{5} = \frac{4}{30}$	$\frac{1}{6} \times \frac{1}{5} = \frac{1}{30}$	$\frac{1}{30}/\frac{1}{2} = \frac{1}{15}$
2	1/6	$\frac{1}{6} \times \frac{3}{5} = \frac{3}{30}$	$\frac{1}{6} \times \frac{2}{5} = \frac{2}{30}$	$\frac{2}{30}/\frac{1}{2} = \frac{2}{15}$
3	1/6	$\frac{1}{6} \times \frac{2}{5} = \frac{2}{30}$	$\frac{1}{6} \times \frac{3}{5} = \frac{3}{30}$	$\frac{3}{30}/\frac{1}{2} = \frac{3}{15}$
4	1/6	$\frac{1}{6} \times \frac{1}{5} = \frac{1}{30}$	$\frac{1}{6} \times \frac{4}{5} = \frac{4}{30}$	$\frac{4}{30}/\frac{1}{2} = \frac{4}{15}$
5	1/6	$\frac{1}{6} \times \frac{0}{5} = \frac{0}{30}$	$\frac{1}{6} \times \frac{5}{5} = \frac{5}{30}$	$\frac{5}{30}/\frac{1}{2} = \frac{5}{15}$
$f(y_j)$		$\frac{15}{30}$	$\frac{15}{30} = \frac{1}{2}$	

Tablica 2.4: Aposteriorne vjerojatnosti za $X \,|\, Y = 1$

x_i	apriorne vjerojatnosti	funkcija vjerodostojnosti	umnožak 2. i 3. stupca	aposteriorne vjerojatnosti
0	1/6	$\frac{0}{5}$	$\frac{1}{6} \times \frac{0}{5} = 0$	0
1	1/6	$\frac{1}{5}$	$\frac{1}{6} \times \frac{1}{5} = \frac{1}{30}$	$\frac{1}{30} \times \frac{1}{2} = \frac{1}{15}$
2	1/6	$\frac{2}{5}$	$\frac{1}{6} \times \frac{2}{5} = \frac{2}{30}$	$\frac{2}{30} \times \frac{1}{2} = \frac{2}{15}$
3	1/6	$\frac{3}{5}$	$\frac{1}{6} \times \frac{3}{5} = \frac{3}{30}$	$\frac{3}{30} \times \frac{1}{2} = \frac{3}{15}$
4	1/6	$\frac{4}{5}$	$\frac{1}{6} \times \frac{4}{5} = \frac{4}{30}$	$\frac{4}{30} \times \frac{1}{2} = \frac{4}{15}$
5	1/6	$\frac{5}{5}$	$\frac{1}{6} \times \frac{5}{5} = \frac{5}{30}$	$\frac{5}{30} \times \frac{1}{2} = \frac{5}{15}$
$f(y_j)$			$\frac{15}{30} = \frac{1}{2}$	

Tablica 2.5: Kraći način za pronalazak aposteriornih vjerojatnosti za $X \,|\, Y = 1$

Prema tome, jednostavniji način za pronalazak aposteriorne funkcije distribucije je korištenje samo tog stupca koji odgovara našoj opservaciji.

2.1 Dva ekvivalentna načina korištenja Bayesovog teorema

Možemo imati više od jednog skupa podataka pomoću kojeg želimo donijeti zaključke o parametru. Moguće je da nam čak nisu dostupni istovremeno. Trebamo li čekati novi skup podataka te ga kombinirati s prethodnim i onda koristiti Bayesov teorem na novodobivenom skupu podataka? To bi značilo da svaki put iznova trebamo zagrebati u podatke što na kraju rezultira s jako puno posla. Drugi manje zahtjevan način bio bi da koristimo aposteriorne vjerojatnosti koje nam je dao prvi skup podataka kao apriorne vjerojatnosti za analiziranje drugog skupa podataka. Promatrajući ta dva pristupa, dolazimo do zaključka da oni vode istim konačnim rezultatima. To je značajna prednost u odnosu na frekvencionističku statistiku u kojoj moramo reanalizirati novodobiveni skup svaki put kad se pojave novi podaci.

Analiza opažanja sekvencijalno

Pretpostavimo da smo na slučajan način izvukli drugu kuglicu iz posude bez da smo prethodno vratili prvu. Pretpostavimo da je druga izvučena kuglica zelene boje, stoga je $Y = 0$. Želimo naći aposteriornu vjerojatnost od X uz dana dva opažanja od Y. Naime, izvukli smo prvo crvenu kuglicu, a zatim zelenu. Koristit ćemo analizu opažanja pomoću Bayesova teorema, primjenivši ga na podatke koje imamo nakon prvog izvlačenja kuglice, a zatim na podatke nakon drugog izvlačenja. Aposteriorne vjerojatnosti nakon prvog izvlačenja koristimo kao apriorne u drugom izvlačenju. Rezultate možemo vidjeti u narednoj tablici:

x_i	apriorne vjerojatnosti	funkcija vjerodostojnosti	umnožak 2. i 3. stupca	aposteriorne vjerojatnosti
0	0	??	0	$0/\frac{1}{3} = 0$
1	$1/15$	$\frac{4}{4}$	$\frac{1}{15}$	$\frac{1}{15} / \frac{1}{3} = \frac{1}{5}$
2	$2/15$	$\frac{3}{4}$	$\frac{1}{10}$	$\frac{1}{10} / \frac{1}{3} = \frac{6}{20}$
3	$3/15$	$\frac{2}{4}$	$\frac{1}{10}$	$\frac{1}{10} / \frac{1}{3} = \frac{6}{20}$
4	$4/15$	$\frac{1}{4}$	$\frac{1}{15}$	$\frac{1}{15} / \frac{1}{3} = \frac{1}{5}$
5	$5/15$	$\frac{0}{4}$	0	$0 / \frac{1}{3} = 0$
			$\frac{1}{3}$	1.00

Tablica 2.6: Aposteriorne vjerojatnosti nakon druge opservacije

Analiza opažanja istovremeno u jednom koraku

Alternativno, mogli bismo promatrati izvlačenja iz posude istovremeno te zatim revidirati početnu apriornu distribuciju samo jednom. Na početku smo u istoj situaciji kao u prethodnom pristupu. Naime imamo istu apriornu distribuciju, a ona daje jednake vjerojatnosti svim mogućim ishodima slučajne varijable X. Stoga je apriorna funkcija distribucije dana sa g(x) = $\frac{1}{6}$ za x = 0, . . . , 5. Neka su Y_1 i Y_2 ishodi prvog i drugog izvlačenja, respektivno. Vjerojatnost drugog izvlačenja ovisi o kuglicama koje su ostale u posudi nakon prvog izvlačenja. Koristeći pravilo umnoška (2.1), opažajuća uvjetna vjerojatnost je:

$$f(y_1, y_2 \,|\, x) = f(y_1 \,|\, x) \times f(y_2 \,|\, y_1, x). \tag{2.3}$$

Prva kuglica je bila crvena, druga zelena, stoga nam je dovoljan stupac u kojem je

$$y_{j1}, y_{j2} = 1, 0.$$

Vidimo da smo na kraju dobili istu aposteriornu distribuciju kao i u slučaju kada smo aposterironu distribuciju prvog izvlačenja koristili za apriornu distribuciju drugog izvlačenja. Dakle, nema razlike između spomenuta dva pristupa. Sljedeća tablica prikazuje postupak koji smo proveli.

x_i	apriorne vjerojatnost	y_{j1}, y_{j2} 0,0	y_{j1}, y_{j2} 0,1	y_{j1}, y_{j2} 1,0	y_{j1}, y_{j2} 1,1	aposteriorne vjerojatnosti
0	1/6	$\frac{1}{6}\times\frac{5}{5}\times\frac{4}{4}=\frac{20}{120}$	$\frac{1}{6}\times\frac{5}{5}\times\frac{4}{4}=0$	$\frac{1}{6}\times\frac{0}{5}\times\frac{4}{4}=0$	$\frac{1}{6}\times\frac{0}{5}\times\frac{4}{4}=0$	0
1	1/6	$\frac{1}{6}\times\frac{4}{5}\times\frac{3}{4}=\frac{12}{120}$	$\frac{1}{6}\times\frac{4}{5}\times\frac{1}{4}=\frac{4}{120}$	$\frac{1}{6}\times\frac{1}{5}\times\frac{4}{4}=\frac{4}{120}$	$\frac{1}{6}\times\frac{1}{5}\times\frac{0}{4}=0$	$\frac{4}{120}/\frac{20}{120}=\frac{1}{5}$
2	1/6	$\frac{1}{6}\times\frac{3}{5}\times\frac{2}{4}=\frac{6}{120}$	$\frac{1}{6}\times\frac{3}{5}\times\frac{2}{4}=\frac{6}{120}$	$\frac{1}{6}\times\frac{2}{5}\times\frac{3}{4}=\frac{6}{120}$	$\frac{1}{6}\times\frac{2}{5}\times\frac{1}{4}=\frac{2}{120}$	$\frac{6}{120}/\frac{20}{120}=\frac{3}{10}$
3	1/6	$\frac{1}{6}\times\frac{2}{5}\times\frac{1}{4}=\frac{2}{120}$	$\frac{1}{6}\times\frac{2}{5}\times\frac{3}{4}=\frac{6}{120}$	$\frac{1}{6}\times\frac{3}{5}\times\frac{2}{4}=\frac{6}{120}$	$\frac{1}{6}\times\frac{3}{5}\times\frac{2}{4}=\frac{6}{120}$	$\frac{6}{120}/\frac{20}{120}=\frac{3}{10}$
4	1/6	$\frac{1}{6}\times\frac{1}{5}\times\frac{0}{4}=0$	$\frac{1}{6}\times\frac{1}{5}\times\frac{4}{4}=\frac{4}{120}$	$\frac{1}{6}\times\frac{4}{5}\times\frac{1}{4}=\frac{4}{120}$	$\frac{1}{6}\times\frac{4}{5}\times\frac{3}{4}=\frac{12}{120}$	$\frac{4}{120}/\frac{20}{120}=\frac{1}{5}$
5	1/6	$\frac{1}{6}\times\frac{0}{5}\times\frac{0}{4}=0$	$\frac{1}{6}\times\frac{0}{5}\times\frac{4}{4}=0$	$\frac{1}{6}\times\frac{5}{5}\times\frac{0}{4}=0$	$\frac{1}{6}\times\frac{5}{5}\times\frac{4}{4}=\frac{20}{120}$	0
	$f(y_1, y_2)$			$\frac{20}{120}$		1.00

Tablica 2.7: Aposteriorne vjerojatnosti za dane $Y_1 = 1$ i $Y_2 = 0$

2.2 Bayesov teorem za binomnu slučajnu varijablu

Promotrit ćemo primjenu Bayesova teorema kada opažanja dolaze iz binomne razdiobe koju smo opisali u prethodnom poglavlju. Dakle, imamo n nezavnisih izvođenja istog pokusa od kojih svaki može rezultirati uspjehom s vjerojatnošću π te neuspjehom s vjerojatnošću $1 - \pi$. Slučajna varijabla Y broji koliko imamo uspjeha u n nezavisnih pokusa. Sada ćemo postaviti tablicu kako bismo primijenili Bayesov teorem. Sami odlučujemo koje apriorne vjerojatnosti dodjeljujemo parametru π, a ako nemamo nikakvu posebnu ideju kako bismo to mogli učiniti, odlučujemo da je svaka od mogućih vrijednosti jednako vjerojatna. Zajedničku vjerojatnosnu distribucija parametra π i opažene vrijednosti Y tražimo tako da pomnožimo $f(y\,|\,\pi)$ s apriornom vjerojatnosti $g(\pi)$.

Primjer 2. *Neka je $X \sim B(4, \pi)$. Pretpostavimo da su samo tri moguće vrijednosti za parametar, $\pi = 0.4$, 0.5 i 0.6. Pretpostavljamo da su sve jednako vjerojatne. Sve vjerojatnosti koje su u vezi s binomnom distribucijom možemo pronaći u Dodatku B u [1]. Prepostavimo da je Y=3, stoga je za traženje aposteriorne distribucije dovoljan samo taj stupac. U narednoj tablici vidimo dobivene rezultate.*

x_i	apriorne vjerojatnosti	0	1	2	3	4	aposteriorne vjerojatnosti
.4	$\frac{1}{3}$	.0432	.1152	.1152	.0512	.0085	$\frac{.0512}{.2497} = .205$
.5	$\frac{1}{3}$	.0208	.0833	.1250	.0833	.0208	$\frac{.0833}{.2497} = .334$
.6	$\frac{1}{3}$	.0085	.0512	.1152	.1152	.0432	$\frac{.1152}{.2497} = .461$
		.0725	.2497	.3554	.2497	.0725	1.000

Tablica 2.8: Rezultati za distribuciju parametra π uz dano $\mathrm{Y} = 3$

2.3 Karakteristike u računu Bayesova teorema

Množenje svih apriornih vjerojatnosti konstantom ne mijenja rezultat Bayesova teorema

Svaki od izraza *apriorna vjerojatnost* $\times$ *vjerojatnostna funkcija distribucije* su pomnoženi nekom konstantom. Marginalna gustoća je zbroj članova duž stupca koji su svi pomnoženi jednakom konstantom. Stoga aposteriorne vjerojatnosti ostaju iste, bez obzira množili sa konstantom ili ne, jer se u svakom slučaju konstanta pokrati. Bitne su relativne težine koje dodjeljujemo vrijednostima, a ne nužno egzaktne vrijednosti. Ako postoji formula za apriornu vjerojatnost, svaki dio koji ne sadrži parametar može se apsorbirati u konstantu. Na taj način račun postaje dosta jednostavniji što vidimo iz Tablice 2.9. Naime, nastavno na prethodni primjer, svaku aprironu vjerojatnost pomnožimo s 3 što onda svakoj vrijednosti parametra π daje težinu 1.

Množenje funkcije vjerodostojnosti konstantom ne mijenja rezultat Bayesova teorema

Isto kao u prethodnom paragrafu, funkcija vjerodostojnosti može biti pojednostavljena na način da svaki dio, koji ne sadrži parametar, možemo apsorbirati u konstantu, koja će se u konačnom računu pokratiti, te na taj način činimo račun jednostavnijim. Formula za funkciju vjerodostojnosti binomne razdiobe u slučaju Primjera 2. je :

$$f(y\,|\,\pi) = \binom{4}{3}\pi^3(1-\pi)^1. \tag{2.4}$$

Binomni koeficijent $\binom{4}{3}$ ne sadrži parametar, stoga ćemo taj dio zanemariti jer je to konstanta koja će se svakako pokratiti u daljnjem računu. U sljedećoj tablici vidimo rezultate koji objedinjuju dva spomenuta svojstva te vidimo da se aposteriorne vjerojatnosti podudaraju s onima iz prethodne tablice.

π	apriorne vjerojatnosti (prop.)	funkcija vjerodostojnosti (prop.)	$2.stupac \times 3.stupac$	aposteriorne vjerojatnosti
.4	1	$.4^3 \times .6^1 = .0384$	.0384	$\frac{.0384}{.1873} = .205$
.5	1	$.5^3 \times .5^1 = .0625$	.0625	$\frac{.0625}{.1873} = .334$
.6	1	$.6^3 \times .4^1 = .0864$	.0864	$\frac{.0864}{.1873} = .461$
			.1873	1.000

Tablica 2.9: Pojednostavljena tablica za računanje distribucije parametra π uz dano Y=3

2.4 Bayesovsko zaključivanje za Poissonovu slučajnu varijablu

Sada ćemo proučiti primjenu Bayesova teorema na podatke koji dolaze iz Poissonove distribucije s parametrom μ. Radi se o diskretnom parametru koji ima moguće vrijednosti $\mu_1, \ldots, \mu_I$. Do sada smo naučili da ne moramo raditi tablicu sa svim mogućim opservacijama, nego da je dovoljno promatrati stupac u kojem se nalazi točno opažena vrijednost. Dakle, općenito, koraci postavljanja tablice za primjenu Bayesova teorema su:

- Napravimo tablicu sa stupcima za: vrijednosti parametra, apriorne vjerojatnosti, funkciju vjerodostojnosti, umnožak prethodna dva stupca i na kraju imamo stupac aposterioronih vjerojatnosti.
- U prva tri stupca unesemo odgovarajuće vrijednosti. Funkciju vjerodostojnosti očitamo iz tablice koju možemo naći u Dodatku B u [1] za točno opaženu vrijednost y. Također, možemo korisiti i odgovarajuću formulu.
- Pomnožimo svaki element iz stupca apriornih vjerojatnosti s odgovarajućim elementom u idućem stupcu i to unesemo u 4. stupac.
- Svaki član 4. stupca podijelimo sa sumom tog cijelog stupca i dobivene vrijednosti unesemo u stupac aposteriornih vjerojatnosti.

Primjer 3. *Neka je $Y \sim P(\mu)$. Pretpostavimo da imamo samo 4 moguće vrijednosti za μ, 1, 1.5, 2 i 2.5 te da vjerujemo da su 1.5 i 2 duplo vjerojatnije od prve i zadnje. Neka je opservacija y=2. Uvrstimo li y=2 u formulu :*

$$f(y\,|\,\mu) = \frac{\mu^y\, e^{-\mu}}{y\,!},$$

dobit ćemo funkciju vjerodostojnosti. Alternativno, mogli smo pogledati u Dodatak B u [1]. U tablici vidimo dobivene rezultate:

μ	apriorne vjerojatnosti (proporcionalne)	funkcija vjerodostojnosti (proporcionalna)	$2.stupac \times 3.stupac$	aposteriorne vjerojatnosti
1.0	$\frac{1}{6}$	$\frac{1.0^2 e^{-1.0}}{2!} = .1839$	.0307	$\frac{.0307}{.2473} = .124$
1.5	$\frac{1}{3}$	$\frac{1.5^2 e^{-1.5}}{2!} = .2510$	.0837	$\frac{.0837}{.2473} = .338$
2.0	$\frac{1}{3}$	$\frac{2.0^2 e^{-2.0}}{2!} = .2707$	.0902	$\frac{.0902}{.2473} = .365$
2.5	$\frac{1}{6}$	$\frac{2.5^2 e^{-2.5}}{2!} = .2565$	.0428	$\frac{.0428}{.2473} = .173$
			.2473	1.000

Tablica 2.10 : Pojednostavljena tablica za računanje distribucije parametra π uz dano Y=2

Poglavlje 3

Bayesovsko zaključivanje za proporciju binomne razdiobe

Vrlo često imamo jako veliku populaciju gdje proporcija populacije, π, ima nekakav atribut. Na primjer, populaciju mogu činiti registrirani birači koji žive u nekom gradu, a atribut koji ih karakterizira je "planira glasati za kandidata A za gradonačelnika". Uzmimo slučajan uzorak iz populacije i neka je Y broj onih birača koji imaju traženi atribut, odnosno planiraju glasati za kandidata A. Dakle, vidimo da brojimo uspjehe u n nezavisnih pokusa gdje svaki pokus ima dva moguća ishoda, uspjeh ili neuspjeh. Uspjeh, u ovom slučaju, znači da glasač planira glasati za kandidata A, a vjerojatnost tog događaja je π, upravo proporcija populacije koja ima taj atribut. Uvjetna distribucija opservacije Y, odnosno ukupnog broja uspjeha u n pokusa, uz dani parametar π, ima binomnu razdiobu s parametrima n i π. Funkcija vjerojatnosti za y uz dani π je :

$$f(y\,|\,\pi) = \binom{n}{y} \pi^y \,(1-\pi)^{n-y} \quad za \;\; y = 1, \ldots, n.$$

Ovdje držimo π fiksnim i gledamo vjerojatnosnu distribuciju od y duž svih mogućih vrijednosti. U slučaju da pogledamo istu formulu, ali držimo y fiksnim, a pustimo da π varira duž svih mogućih vrijednosti, dobivamo funkciju vjerodostojnosti, danu identičnom formulom, ali parametar koji varira se promijenio :

$$f(y\,|\,\pi) = \binom{n}{y} \pi^y \,(1-\pi)^{n-y} \quad za \;\; 0 \leq \pi \leq 1. \tag{3.1}$$

Bayesov teorem lako možemo izreći na sljedeći način : *aposteriorna distribucija je proporcionalna apriornoj distribuciji pomnoženoj s funkcijom vjerodostojnosti* , odnosno:

$$g(\pi|y) \propto g(\pi) \times f(y|\pi).$$

Na taj način smo dobili uvid u aposteriorne distribucije, odnosno gustoće, ali ne i egzaktnu gustoću jer nismo osigurali da površina ispod grafa te funkcije bude 1. Potrebno je skalirati prethodni rezultat određenom konstantom koju u ovom slučaju dobijemo integriranjem $g(\pi) \times f(y \mid \pi)$. Dakle, računamo :

$$g(\pi|y) = \frac{g(\pi) \times f(y \mid \pi)}{\int_0^1 g(\pi) \times f(y \mid \pi)\, d\pi}$$

Ovisno o tome koju apriornu distribuciju smo odabrali, integral koji je potrebno izračunati neće nužno biti u zatvorenoj formi ili se neće dati izračunat egzaktno. U tim slučajevima upotrebljavamo razne numeričke metode integriranja. Radi toga ćemo se upoznati s nekim apriornim distribucijama koje znatno olakšavaju račun ili čak omogućavaju da izbjegnemo integriranje.

3.1 Korištenje uniformne apriorne distribucije

Ukoliko nemamo nikakvu ideju o tome kakav bi parametar π mogao biti, možemo odabrati apriornu distribuciju koja ne favorizira niti jednu vrijednost u odnosu na drugu. U tim slučajevima biramo uniformnu apriornu distribuciju:

$$g(\pi) = 1 \qquad za \;\; 0 \leq \pi \leq 1.$$

Raspis aposteriorne distribucije u slučaju uniformne apriorne distribucije:

$$g(\pi|y) = \frac{g(\pi) f(y|\pi)}{\int_0^1 g(\pi) \times f(y|\pi)\, d\pi}$$

$$= \frac{\binom{n}{y}\pi^y (1-\pi)^{n-y}}{\int_0^1 \binom{n}{y}\pi^y (1-\pi)^{n-y} d\pi}$$

$$= \frac{\binom{n}{y}\pi^y (1-\pi)^{n-y}}{\binom{n}{y}\frac{\Gamma(y+1)\Gamma(n-y+1)}{\Gamma(n+2)} \int_0^1 \frac{\Gamma(n+2)}{\Gamma(y+1)\Gamma(n-y+1)}\pi^y (1-\pi)^{n-y}\, d\pi} \tag{3.2}$$

$$= \frac{\Gamma(n+2)}{\Gamma(y+1)\,\Gamma(n-y+1)} \pi^y (1-\pi)^{n-y},$$

odnosno:

$$g(\pi \,|\, y) \propto \pi^y\ (1-\pi)^{n-y} \qquad za\ \ 0 \le \pi \le 1.$$

Prethodni izraz predstavlja Beta distribuciju s parametrima a i b (do na konstantu normalizacije koja osigurava da ukupna vjerojatnost iznosi 1), gdje je $a = y+1$ i $b = n-y+1$.

Također, uočimo kako smo u (3.2) podintegralni izraz namjestili da dobijemo oblik Beta(a,b) apriorne gustoće koji glasi:

$$g(\pi; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \pi^{a-1} (1-\pi)^{b-1}. \tag{3.3}$$

3.2 Korištenje Beta apriorne distribucije

Ukoliko za π upotrijebimo Beta(a,b) apriornu gustoću (3.3), dobivamo:

$$g(\pi|y) = \frac{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \pi^{a-1} (1-\pi)^{b-1} \binom{n}{y} \pi^y (1-\pi)^{n-y}}{\binom{n}{y} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+y)\Gamma(b+n-y)}{\Gamma(a+b+n)} \int_0^1 \frac{\Gamma(a+b+n)}{\Gamma(a+y)\Gamma(b+n-y)} \pi^{a+y-1} (1-\pi)^{b+n-y-1} d\pi}$$

$$= \frac{\Gamma(a+b+n)}{\Gamma(a+y)\,\Gamma(b+n-y)} \pi^{a+y-1}\ (1-\pi)^{b+n-y-1}. \tag{3.4}$$

Ako izostavimo konstante, vrijedi:

$$g(\pi|y) \propto \pi^{a+y-1}\ (1-\pi)^{b+n-y-1} \quad za\ \ 0 \le \pi \le 1. \tag{3.5}$$

Prepoznajemo iz (3.4) da se radi o Beta distribuciji sa parametrima $a' = a + y$ i $b' = b + n - y$. Primijetimo da zapravo dodajemo broj uspjeha parametru a, a broj neuspjeha parametru b. Ponovno smo aposteriornu distribuciju lako dobili bez integriranja.

Uočimo da je uniformna aposteriorna distribucija poseban slučaj Beta distribucije kada su parametri $a = 1$ i $b = 1$.

Familija Beta distribucija je konjugirana za binomnu raspodjelu

Definicije su preuzete iz [2]

Definicija 3.2.1. *Bayesovski statistički model slučajne varijable (ili slučajnog vektora) $X : \Omega \to \mathbb{R}^k$ je familija $\mathcal{P} = \{f(\cdot|\theta) : \theta \in \Theta\}$ (iz definicije 1.3.7), pri čemu je parametru θ pridružena njegova apriorna distribucija s gustoćom $f(\theta)$, $\theta \in \Theta$.*

Definicija 3.2.2. *Za familiju $\mathcal{F}$ vjerojatnosnih distribucija na Θ kažemo da je konjugirana za statističku strukturu $\mathcal{P}$ ako za svaki $f \in \mathcal{F}$ također vrijedi $f(\theta\,|\,x) \in \mathcal{F}$, za svaki x. Tada apriornu i aposteriornu distribuciju nazivamo konjugiranim distribucijama, a apriornu funkciju gustoće nazivamo konjugiranim apriorom za funkciju vjerodostojnosti danu pomoću modela $\mathcal{P}$.*

Ukoliko krenemo s Beta apriornom distribucijom, dobivamo Beta aposteriornu distribuciju, prateći jednostavno pravilo: dodaj broj uspjeha parametru a te broj neuspjeha parametru b. Beta distribucija je konjugirana familija za binomnu raspodjelu.

Jeffreyjeva apriorna distribucija

Beta $\left(\frac{1}{2}, \frac{1}{2}\right)$ apriorna distribucija poznata je kao Jeffreyjeva apriorna distribucija za binomnu raspodjelu. Za dani parametar π, dana je sa :

$$g(\pi) \propto \sqrt{I(\pi|y)},$$

gdje je $I(\pi|y)$ Fisherova informacija koja je dana formulom :

$$I(\pi|y) = -\mathbb{E}\left(\frac{\partial^2 log f(y|\pi)}{\partial \pi^2}\right).$$

Slijedi raspis u slučaju binomne raspodjele.
Pogledamo funkciju vjerodostojnosti iz formule (3.1) te ju logaritmiramo:

$$log\, f(y|\pi) = log\binom{n}{y} + ylog\pi + (n-y)\, log\,(1-\pi).$$

Nadalje,

$$\frac{\partial}{\partial \pi} log f(y|\pi) = \frac{y}{\pi} - \frac{n-y}{1-\pi} = y\pi^{-1} - (n-y)(1-\pi)^{-1}$$

$$\frac{\partial^2}{\partial \pi^2} log f(y|\pi) = -y\pi^{-2} - (n-y)(1-\pi)^{-2}.$$

Budući da vrijedi $\mathbb{E}[Y] = n\pi$, imamo :

$$-\mathbb{E}\left[\frac{\partial^2}{\partial\pi^2} log f\left(y|\pi\right)\right] = n\pi^{-1}\left(1-\pi\right)^{-1}.$$

Slijedi:

$$\sqrt{-\mathbb{E}\left[\frac{\partial^2}{\partial\pi^2} log f\left(y|\pi\right)\right]} = \sqrt{n}\pi^{-\frac{1}{2}}\left(1-\pi\right)^{-\frac{1}{2}}.$$

Dakle,

$$g\left(\pi\right) \propto \sqrt{n}\pi^{-\frac{1}{2}}\left(1-\pi\right)^{-\frac{1}{2}}.$$

Jednostavno slijedi:

$$g\left(\pi\right) = \frac{\Gamma\left(1\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\pi^{-\frac{1}{2}}\left(1-\pi\right)^{-\frac{1}{2}}$$

što je upravo funkcija gustoće Beta $\left(\frac{1}{2}, \frac{1}{2}\right)$ distribucije.

3.3 Odabir apriorne distribucije

Bayesov teorem daje nam metodu kojom možemo revidirati distribuciju parametra, uz dane podatke. U ovom dijelu predstavit ćemo neke metode koje nam mogu pomoći pri odabiru apriorne distribucije, kao i ono na što trebamo obratiti pažnju.

Odabir konjugirane apriorne distribucije ukoliko imamo nejasno prethodno znanje

Kada na početku imamo maglovito, nejasno znanje o apriornoj distribuciji, odabir jedne od Beta(a,b) distribucija će biti prikladan odabir. Recimo da jedino što na početku znamo o parametru π jest da je on jako malen. Tada će Beta $(0.5, 1)$, Beta $(0.5, 2)$, Beta $(0.5, 3)$, Beta $(1, 2)$ ili Beta $(1, 3)$ biti zadovoljavajuć izbor. Svaka od tih distribucija najviše vjerojatnosti daje upravo onim vrijednostima π koje su male. Primjere Beta distribucija možemo vidjeti u Dodatku na kraju rada (5). Nije jako bitno koju od tih distribucija ćemo izabrati; konačna distribucija će u svakom slučaju biti vrlo slična.

Odabir konjugirane apriorne distribucije ukoliko prethodno znamo mjere lokacije i raspršenosti

Beta (a, b) distribucija ima razne oblike, a bilo bi dobro da se naš odabir apriorne distribucije podudara s našim početnim uvjerenjima. Predlažemo odabir one Beta (a, b) distribucije koja se podudara s našim uvjerenjima o mjerama lokacije i raspršenosti. Neka je π_0 apriorno očekivanje, a σ_0 apriorna standardna devijacija. Srednja vrijednost Beta (a, b) raspodjele je $\frac{a}{a+b}$. To izjednačimo s našim očekivanjem srednje vrijednosti :

$$\pi_0 = \frac{a}{a+b}. \tag{3.6}$$

Nadalje, standardna devijacija Beta distribucije je $\sqrt{\frac{ab}{(a+b)^2(a+b+1)}}$. To izjednačimo s našom pretpostavkom o standardnoj devijaciji. Jer vrijedi $\frac{a}{a+b} = \pi_0$ i $\frac{b}{a+b} = 1 - \pi_0$, imamo :

$$\sigma_0 = \sqrt{\frac{\pi_0 (1 - \pi_0)}{a + b + 1}}. \tag{3.7}$$

Rješavanje tih dviju jednadžbi daje nam parametre a i b, odnosno Beta (a, b) apriornu distribuciju.

Na što treba paziti prije upotrebe odabrane apriorne distribucije?

1. Prvo moramo nacrtati odabranu Beta (a, b) apriornu distribuciju. Ukoliko je oblik funkcije prilično blizu prvotnim stavovima o raspodjeli parametra, možemo ju koristiti u daljnjem računu. U suprotnom, smijemo korigirati π_0 i σ_0 sve dok graf ne bude aproksimativno odgovarao prethodnim vjerovanjima.

2. Znamo da uzorak iz binomne raspodjele (n, π), čija je proporcija $\tilde{\pi} = \frac{y}{n}$, ima varijancu jednaku : $\frac{\pi(1-\pi)}{n}$. Izjednačimo li to (uvrstivši π_0) s apriornom varijancom, dobivamo:

$$\frac{\pi_0 (1 - \pi_0)}{n_{ek}} = \frac{ab}{(a+b)^2 \times (a+b+1)}.$$

   Budući da je $\pi_0 = \frac{a}{a+b}$ i $(1 - \pi_0) = \frac{b}{a+b}$, ekvivalentna veličina uzorka je $n_{ek} = a+b+1$. Ona kaže da je količina informacija o parametru zasnovana na našoj apriornoj distribuciji jednaka količini koju bismo dobili iz slučajnog uzorka te veličine. Uvijek moramo provjeriti je li ta veličina nerealistično velika. Trebamo se pitati: "Je li naše znanje o proporciji π zbilja jednako onomu koje bi dobili iz novog slučajnog uzorka veličine n_{ek}?" Ako je odgovor ne, trebamo povećati pretpostavljenu standardnu devijaciju.

Generalno o konstrukciji neprekidne apriorne distribucije

Ukoliko apriorna distribucija ne odgovara nekoj Beta distribuciji, možemo konstruirati diskretnu distribuciju koju zatim interpoliramo kako bismo dobili neprekidnu. Ipak, ovim pristupom moramo provesti integriranje u računu.To vidimo u sljedećem primjeru.

Primjer 4. *Tri studenta konstruiraju svoja apriorna uvjerenja o π, proporciji građana grada Hamiltona koji podržavaju izgradnju casina u njihovom gradu. Ana misli da je srednja vrijednost njezina apriorna odabira 0.2, a standardna devijacija 0.8. Beta (a, b) raspodjela koja zadovoljava njezin odabir nađe se tako da, korištenjem formule* (3.7), *izračunamo:*

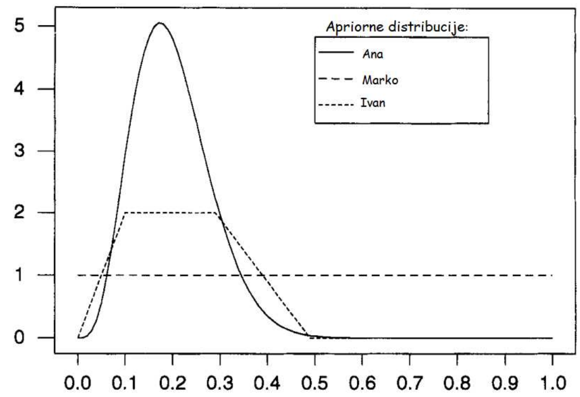
$$\frac{0.2 \times 0.8}{a + b + 1} = 0.008^2.$$

Stoga je njezina ekvivalentna veličina $a + b + 1 = 25$. Za Aninu apriornu distribuciju zato vrijedi da je $a = 4.8$ i $b = 19.2$. Budući da je Marko pridošlica u gradu, on nema nikakvu ideju o omjeru ljudi koji su za i koji su protiv gradnje casina. Zato se on odlučuje za uniformnu raspodjelu. Dakle, za njega vrijedi, $a = b = 1$. Veličina njegova uzorka je a+b+1=3. Ivan ne može pronaći Beta (a, b) distribuciju koja će odgovarati njegovim apriornim uvjerenjima. On stoga daje težine pojedinim vrijednostima, što vidimo u tablici ispod, te ih linearno interpolira kako bi dobio neprekidnu raspodjelu.

Vrijednost	*Težina*
0	*0*
0.05	*1*
0.1	*2*
0.3	*2*
0.4	*1*
0.5	*0*

$$g(\pi) = \begin{cases} 20\pi & za\ 0 \leq \pi \leq 0.10 \\ 0.2 & za\ 0.10 \leq \pi \leq 0.30 \\ 5 - 10\pi & za\ 0.30 \leq \pi \leq 0.50 \end{cases}$$

Navedene tri apriorne raspodjele vidimo na Slici 3.1.[1] *Primjetimo, površina ispod Ivanove krivulje nije jednaka 1, stoga ona nije uistinu funkcija gustoće. Međutim, to nam ne predstavlja problem jer relativne težine koje su nam dane tim oblikom funkcije su sve što je zbilja i potrebno, budući da se konstanta svakako pokrati u računu.*



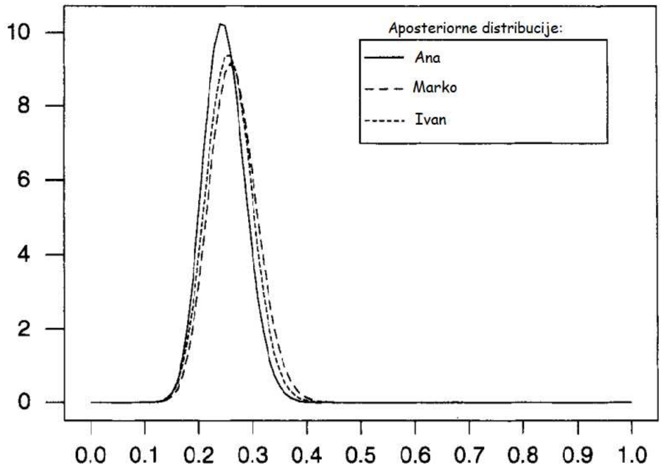
Slika 3.1: Apriorne distribucije

Efekti apriornih distribucija

U slučaju velikog broja podataka, dobit ćemo jako slične aposteriorne distribucije, makar krenuli s dosta različitim početnim distribucijama. Razlog tomu je taj što će podaci na neki način "preplaviti" sam odabir apriorne raspodjele. To vidimo u nastavku prethodnog primjera.

Primjer 5. *(nastavak Primjera 4.) Tri studenta uzela su nasumičan uzorak građana grada Hamiltona veličine n=100 i saznali njhova stajališta o izgradnji casina. Pokazalo se da njih y=26 podržava izgradnju. Znamo da je Anina aposteriorna distribucija Beta* $(4.8 + 26, 19.2 + 74)$*, a Markova Beta* $(1 + 26, 1 + 74)$*, dok Ivanovu nađemo numerički korištenjem odgovarajuće formule. Te tri distribucije prikazane su na Slici 3.2.*[2] *Vidimo da na kraju imamo vrlo slične aposteriorne distribucije, unatoč tomu što smo krenuli s poprilično različitim apriornim distribucijama.*

[1] Slika je preuzeta iz [1]
[2] Slika je preuzeta iz [1]



Slika 3.2: Aposteriorne distribucije

3.4 Zaključno o aposteriornim distribucijama

Graf aposterirone distribucije prikazuje nam sve što možemo znati o parametru u tom trenu. Mjere lokacije i mjere raspršenosti pomažu nam okarakterizirati novodobivenu distribuciju.

MJERE LOKACIJE

Mjera lokacije je jedna vrijednost koja pokušava opisati dani skup podataka tako da identificira središnji položaj unutar tog skupa podataka. Razlikujemo srednju vrijednost, mod i medijan.

Aposteriorna srednja vrijednost

Računamo je na sljedeći način:

$$m' = \int_0^1 \pi g\,(\pi|y)\ d\pi. \tag{3.8}$$

Kada je aposteriorna distribucija Beta (a', b'), srednja vrijednost iznosi:

$$m' = \frac{a'}{a' + b'}. \tag{3.9}$$

Aposteriorni medijan

Medijan je vrijednost iznad koje se nalazi 50% skupa (distribucije) podataka, a 50% ispod. Ako $g(\pi|y)$ prati Beta (a', b') raspodjelu, tada je medijan rješenje sljedeće formule:

$$\int_0^{medijan} g(\pi|y)\ d\pi = 0.5. \tag{3.10}$$

Jedini nedostatak je što dani integral moramo računati numerički.

Aposteriorni mod

Mod je najčešća vrijednost u skupu podataka. Ako je aposteriorna distribucija neprekidna, nalazimo ga tako da $g(\pi|y)$ deriviramo i izjednačimo s 0. U slučaju da se radi o Beta(a', b') distribuciji, derivacija je:

$$g'(y) = (a'-1)\,\pi^{a'-2} \times (1-\pi)^{b'-1} + \pi^{a'-1} \times (-1)\,(b'-1)\,(1-\pi)^{b'-2}.$$

Izjednačimo li to s 0, dobivamo:

$$mod \ = \ \frac{a'-1}{a'+b'-2}.$$

MJERE RASPRŠENOSTI

Mjere raspršenosti govore nam o varijabilnosti podataka, a najpoznatije su standardna devijacija, varijanca, percentili i interkvartili.

Varijanca

Varijanca aposteriorne distribucije je :

$$Var(\pi|y) = \int_0^1 (\pi - m')^2 g(\pi|y)\ d\pi. \tag{3.11}$$

U slučaju kada imamo Beta (a', b') aposteriornu distribuciju, varijanca je :

$$Var(\pi|y) = \frac{a' \times b'}{(a'+b')^2 \times (a'+b'+1)}. \tag{3.12}$$

Standardna devijacija

Budući da se varijanca izražava u kvadratnim jedinicama, jednostavnije je pogledati korijen varijance što nam daje standardnu devijaciju.

Percentili

K-ti percentil aposteriorne distribucije je vrijednost π_k, koja sadrži $k\%$ distribucije ispod. Numerički se nađe rješavanjem :

$$k = 100 \times \int_{-\infty}^{\pi_k} g(\pi|y)\ d\pi.$$

Interkvartilna razlika

To je razlika između gornjeg i donjeg kvartila te predstavlja raspon unutar kojeg se nalazi središnjih 50% podataka. Vrlo često se koristi, a glasi:

$$IQR = Q_3 - Q1.$$

Osoba	Aposteriorna distribucija	Prosjek	Medijan	Std.Dev.	IQR
Ana	Beta $(30.8, 93.2)$	0.248	0.247	0.039	0.053
Marko	Beta $(27, 75)$	0.265	0.263	0.043	0.059
Ivan	numerički	0.261	0.255	0.041	0.057

Tablica 3.1 : Mjere lokacija i raspršenosti aposteriornih distribucija

Primjer 6. *(nastavak Primjera 5.). Ana, Marko i Ivan izmjerili su neke mjere lokacije i mjere raspršenosti za svoje apriorne distribucije. Ana i Marko koristili su formule* (3.9) *i* (3.12)*, budući da su imali aposteriorne Beta raspodjele. Ivan je koristio formule* (3.8) *i* (3.11) *budući da on nije imao Beta distribuciju. Vrijednosti koje su dobili prikazani su u Tablici 3.1. Jasno vidimo da svi troje imaju slične rezultate finalnih statistika, iako su krenuli s različitim apriornim distribucijama.*

3.5 Bayesovski intervali vjerodostojnosti

Vrlo često želimo pronaći interval koji ima veliku vjerojatnost da sadrži parametar. Raspon vrijednosti, koji ima $1-\alpha$ vjerojatnost da sadrži parametar, uzevši u obzir dane podatke, odnosno dani slučajni uzorak, zovemo Bayesovski interval vjerodostojnosti. U narednom poglavlju vidjet ćemo da oni daju relevantniji odgovor od intervala pouzdanosti, budući da imaju direktniju vjerojatnosnu interpretaciju.

Bayesovski interval vjerodostojnosti za proporciju π

Ako smo koristili Beta (a, b) apriornu distribuciju, aposteriorna distribucija za $\pi \mid y$ je Beta (a', b'). Jednaka površina repova 95%-og intervala vjerodostojnosti može se dobiti pomoću 97.5 i 2.5 percentila. Naime, aproksimiramo Beta (a', b') aposteriornu distribuciju s normalnom distribucijom, uz istu srednju vrijednost i varijancu:

$$(\pi|y) \propto N\Big[m'; (s')^2\Big]$$

gdje je aposteriorna srednja vrijednost dana s:

$$m' = \frac{a'}{a'+b'},$$

a aposteriorna varijanca dana je s:

$$(s')^2 = \frac{a'b'}{(a'+b')^2\,(a'+b'+1)}. \tag{3.13}$$

Tada je $(1-\alpha)\times 100\%$ interval vjerodostojnosti za proporciju π aproksimativno dan s:

$$m' \pm z_{\frac{\alpha}{2}} \times s', \tag{3.14}$$

gdje je $z_{\frac{\alpha}{2}}$ vrijednost koju pronađemo u tablici funkcije distribucije standardne normalne razdiobe. Za 95% interval vjerodostojnosti, $z_{0.025} = 1.96$. Aproksimacija je jako dobra ukoliko je $a' \geq 10$ i $b' \geq 10$.

Primjer 7. *(nastavak na Primjer 6.) Ana, Marko i Ivan su izračunali* 95% *intervale vjerodostojnosti za proporciju π, uz jednake veličine repova. Koristili su dva načina: koristeći egzaktnu Beta funkciju gustoće i koristeći normalnu aproksimaciju. Ana, Marko i Ivan imaju djelomično različite intervale jer su krenuli s drugačijim pretpostavkama, ali ipak njihova početna uvjerenja su nadjačali podaci, stoga razlike nisu velike. Vidimo da se egzaktni intervali i oni dobiveni normalnom aproksimacijom neznatno razlikuju.*

Osoba	Aposteriorna distribucija	Interval vjerodostojnosti egzaktni		Interval vjerodostojnosti normalna aproksimacija	
		donja granica	gornja granica	donja granica	gornja granica
Ana	Beta $(30.8, 93.2)$	0.177	0.328	0.172	0.324
Marko	Beta $(27, 75)$	0.184	0.354	0.183	0.355
Ivan	numerički	0.181	0.340	0.181	0.341

Tablica 3.2 : Egzaktni i aproksimativni intervali vjerodostojnosti

Poglavlje 4

Usporedba frekvencionističkog i Bayesovskog zaključivanja za proporciju

Kada koristimo Bayesov pristup, aposteriorna distribucija uz opažene podatke daje konačan zaključak o parametru. S druge strane, frekvencionistički pristup statistici ima nekoliko različitih načina pomoću kojih donosi zaključke o parametru. To su točkovna procjena, procjena intervala te testiranje hipoteza. Naravno, sve to možemo izvesti i Bayesovskim pristupom.

4.1 Frekvencionistička interpretacija vjerojatnosti i parametara

Uzimamo slučajan uzorak iz nekog skupa podataka, odnosno populacije, koja ima nepoznat parametar. Pretpostavljamo da je on fiksna, ali nepoznata konstanta. Dakle, nemamo nikakvih vjerojatnosnih distribucija u vezi parametra, ali zato razmatramo distribuciju slučajnog uzorka veličine n.

Distribucija statistike nakon ponovno uzorkovanja

Uzmimo da su $Y_1, \dots, Y_n$ slučajni uzorci iz populacije čija distribucija ovisi o parametru θ. Neka je S statistika izračunata za slučajan uzorak. Vidimo da je S slučajna varijabla. Izračunamo S za svaki slučajan uzorak $Y_1, \dots, Y_n$ gdje su svi veličine n. Distribucija tih vrijednosti zove se distribucija statistike ponovnog uzorkovanja (sampling distribucija). Ta distribucija također ovisi o nepoznatoj vrijednosti parametra θ:

$$f(s \mid \theta).$$

U Bayesovu pristupu čitav zaključak temelji se na aposteriornoj distribuciji parametra, uzimajući u obzir točno opažene podatke:

$$g\left(\theta \,|\, podaci\right).$$

Također, svi ostali Bayesovski zaključci o procjenitelju ili intervalima vjerodostojnosti računaju se iz aposteriorne distribucije. Dakle, oni ovise o točno opaženim podacima.

4.2 Točkovna procjena

Definicija 4.2.1. *Neka je $X_1, \ldots, X_n$ slučajni uzorak iz distribucije F te θ nepoznati parametar te distribucije. Točkovni procjenitelj T od θ je funkcija slučajnog uzorka $T = t(X_1, \ldots, X_n)$.*

Budući da ne znamo pravu vrijednost parametra, ne možemo donijeti sud o preciznosti procjene, stoga ćemo promotriti neke kriterije uzoračke distribucije procjenitelja.

Nepristran procjenitelj

Očekivana vrijednost procjenitelja opisuje središte distribucije. To je srednja vrijednost koju bi procjenitelj imao da smo ga izračunali za sve moguće slučajne uzorke.

Kažemo da je procjenitelj nepristran ako je srednja vrijednost sampling distribucije jednaka pravoj vrijednosti parametra. Dakle, procjenitelj $\tilde{\theta}$ je nepristan ako:

$$\mathbb{E}\left(\tilde{\theta}\right) = \int \tilde{\theta} f\left(\tilde{\theta} \,|\, \theta\right) \, d\tilde{\theta} = \theta,$$

gdje je $f\left(\tilde{\theta} \,|\, \theta\right)$ sampling distribucija procjenitelja $\tilde{\theta}$ za parametar θ.
Pristranost procjenitelja $\tilde{\theta}$ definiramo kao:

$$bias\left(\tilde{\theta}\right) = \mathbb{E}\left(\tilde{\theta}\right) - \theta.$$

Očigledno, nepristranim procjeniteljima pristranost je jednaka 0. Bayesovski pristup većinom vodi do pristranih procjenitelja, dok frekventisti stavljaju naglasak na nepristrane procjenitelje.

Srednjekvadratna pogreška procjenitelja

Srednjekvadratna pogreška (frekvencionistička) procjenitelja $\tilde{\theta}$ je :

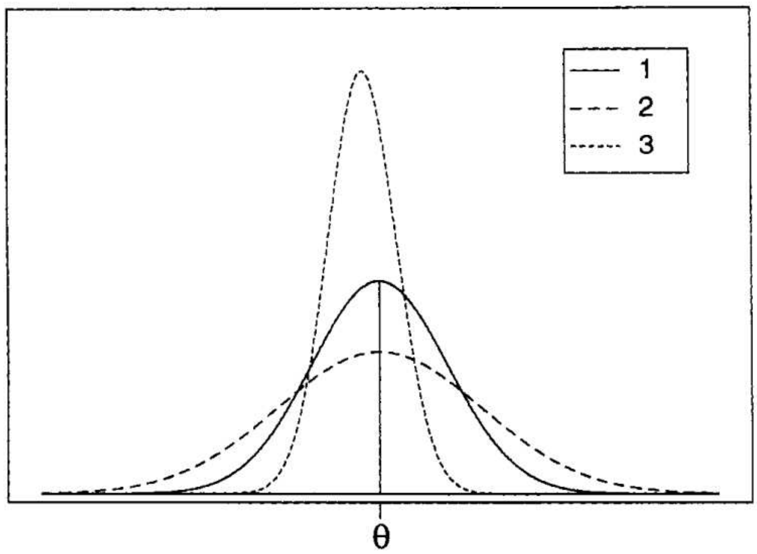
$$MSE\left(\tilde{\theta}\right) = \mathbb{E}\left(\tilde{\theta} - \theta\right)^2 = \int \left(\tilde{\theta} - \theta\right)^2 f\left(\tilde{\theta}|\theta\right) d\tilde{\theta}.$$

Pokazuje se da vrijedi:

$$MSE\left(\tilde{\theta}\right) = bias\left(\tilde{\theta}\right)^2 + Var\left(\tilde{\theta}\right).$$

Minimalna varijanca nepristranog procjenitelja

Kažemo da nepristran procjenitelj ima minimalnu varijancu ako niti jedan drugi nepristran procjenitelj nema manju varijancu. Slika 4.1 [1]prikazuje uzoračke distribucije tri moguća procjenitelja za θ. Procjenitelj 1 i 2 su nepristrani, a procjenitelj 3 je pristran. Procjenitelj 1 je najbolji budući da ima najmanju varijancu. Procjenitelj 3 ima manju varijancu nego procjenitelj 1. Trebali bismo moći nekako uspoređivati procjenitelje, a za to trebamo mjeriti njihovu efikanost. Mjera za efikasnost procjenitelja je srednjekvadratna pogreška.



Slika 4.1: Uzoračke distribucije za tri procjenitelja

[1]Slika je preuzeta iz [1]

4.3 Usporedba procjenitelja za proporciju

Bayesovski procjenitelji često imaju manje srednjekvadratne pogreške nego frekvencionistički procjenitelji. Frekvencionistički procjenitelj za π je:

$$\tilde{\pi}_f = \frac{y}{n},$$

gdje y, broj uspjeha u n nezavisnih pokusa, ima binomnu (n, π) distribuciju. $\tilde{\pi}_f$ je nepristran procjenitelj te vrijedi $Var\left(\tilde{\pi}_f\right) = \frac{\pi\times(1-\pi)}{n}$. Stoga je srednjekvadratna pogreška za $\tilde{\pi}_f$:

$$\begin{aligned} MSE\left(\tilde{\pi}_f\right) &= 0^2 + Var\left(\tilde{\pi}_f\right) \\ &= \frac{\pi\times(1-\pi)}{n}. \end{aligned}$$

Pretpostavimo da koristimo srednju vrijednost aposteriorne distribucije kao Bayesovsku procjenu za π pri čemu smo koristili Beta(1,1) kao apriornu distribuciju (uniformna apriorna distribucija), stoga imamo:

$$\tilde{\pi}_B = m' = \frac{a'}{a'+b'},$$

gdje je $a' = 1 + y$ i $b' = 1 + n - y$. Možemo to drugačije zapisati, kao linearnu funkciju u ovisnosti o y:

$$\tilde{\pi}_B = \frac{y+1}{n+2} = \frac{y}{n+2} + \frac{1}{n+2}.$$

Zato je srednja vrijednost ove uzoračke distribucije:

$$\frac{n\pi}{n+2} + \frac{1}{n+2},$$

a varijanca:

$$\left[\frac{1}{n+2}\right]^2 \times n\pi\,(1-\pi)\,.$$

Sada slijedi srednjekvadratna pogreška:

$$\begin{aligned} MSE\left(\tilde{\pi}_B\right) &= \left[\frac{n\pi}{n+2} \times \frac{1}{n+2} - \pi\right]^2 + \left[\frac{1}{n+2}\right]^2 \times n\pi\,(1-\pi) \\ &= \left[\frac{1-2\pi}{n+2}\right]^2 + \left[\frac{1}{n+2}\right]^2 \times n\pi\,(1-\pi)\,. \end{aligned}$$

Primjerice, pretpostavimo da je $\pi = 0.4$, a da je veličina uzorka n=10. Tada imamo:

$$MSE\left(\tilde{\pi}_f\right) = \frac{0.4 \times 0.6}{10} = 0.024$$

i

$$MSE\left(\tilde{\pi}_B\right) = \left[\frac{1 - 2 \times 0.4}{12}\right]^2 + \left[\frac{1}{12}\right]^2 \times 10 \times 0.4 \times 0.6 = 0.0169.$$

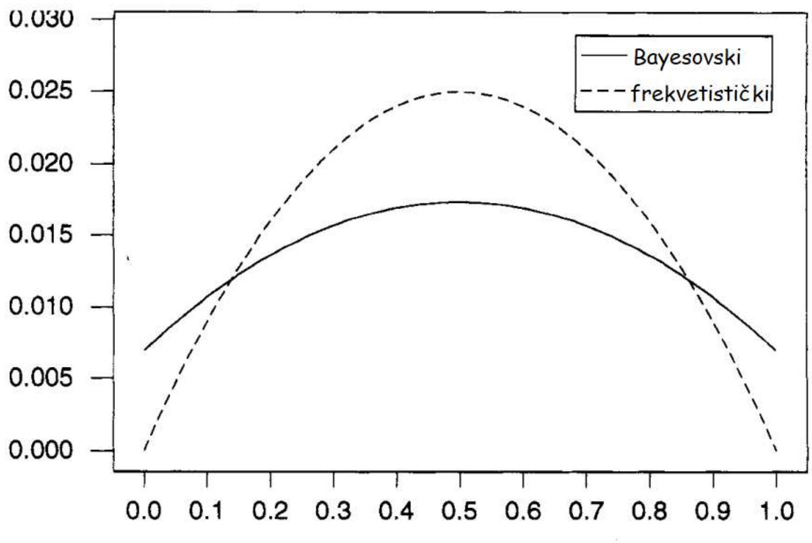
Nadalje, pretpostavimo da je $\pi = 0.5$ i n=10. Tada imamo:

$$MSE\left(\tilde{\pi}_f\right) = \frac{0.5 \times 0.5}{10} = 0.025$$

i

$$MSE\left(\tilde{\pi}_B\right) = \left[\frac{1 - 2 \times 0.5}{12}\right]^2 + \left[\frac{1}{12}\right]^2 \times 10 \times 0.5 \times 0.5 = 0.01736.$$

Vidimo da je Bayesovski procjenitelj u prosjeku bolji nego frekvencionistički (za te dvije vrijednosti proporcije π). Slika 4.2 [2] prikazuje srednjekvadratnu pogrešku za oba procjenitelja te vidimo da je u većini slučajeva Bayesovski procjenitelj (koristeći uniformnu apriornu distribuciju) bolji od frekvencionističkog.

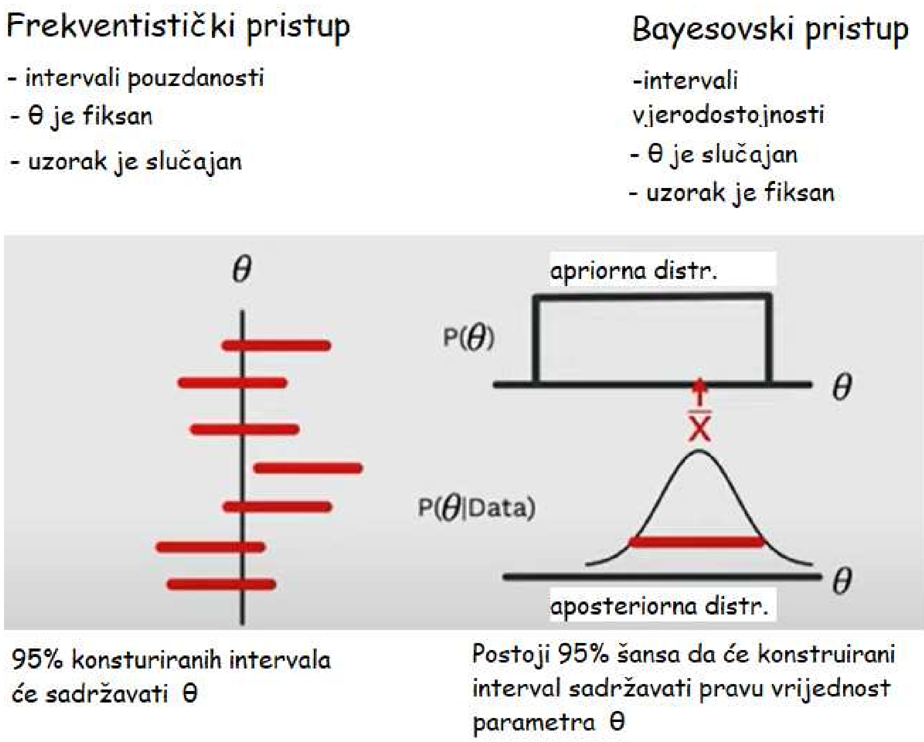


Slika 4.2: Srednjekvadratna greška

[2] Slika je preuzeta iz [1]

4.4 Intervalna procjena

Drugi tip zaključivanja koji promatramo je intervalna procjena. Želimo pronaći interval (l, u) koji ima unaprijed određenu vjerojatnost sadržavanja parametra. Znamo da će unaprijed određena proporcija intervala, izračunatih na raznim slučajnim uzorcima, sigurno sadržavati pravu vrijednost parametra. S druge strane, kako smo i vidjeli u prethodnom poglavlju, Bayesovski interval vjerodostojnosti za propoprciju π ima upravo željenu vjerojatnost. Upravo tu leži razlika intervala pouzdanosti i Bayesovskih intervala vjerodostojnosti. Lakše ćemo shvatiti pomoću slike 4.3[3].



Slika 4.3: Usporedba intervala pouzdanosti i vjerodostojnosti

[3] Slika je preuzeta s web adrese: https://www.youtube.com/watch?v=EJe3jiZNwUU

Intervali pouzdanosti

Intervali pouzdanosti su način na koji frekvencionistički statističari pokušavaju pronaći interval koji ima veliku vjerojatnost da sadrži pravu vrijednost parametra θ. $(1-\alpha)\times 100\%$-pouzdani interval za θ je slučajni interval (l, u) takav da je

$$\mathbb{P}(l \leq \theta \leq u) = 1 - \alpha.$$

Često je uzoračka distribucija procjenitelja aproksimativno normalna sa srednjom vrijednosti jednakom pravoj vrijednosti parametra. U tom slučaju, interval pouzdanosti ima oblik:

$$\textit{procjenitelj} \pm \textit{kritična vrijednost} \times \textit{standardna devijacija procjenitelja},$$

gdje kritičnu vrijednost dobijemo iz tablice standardne normalne razdiobe. Primjerice, ako je n jako velik, proporcija uzorka

$$\tilde{\pi}_f = \frac{y}{n}$$

aproksimativno prati normalnu raspodjelu sa srednjom vrijednosti π i standardnom devijacijom $\sqrt{\frac{\pi_f(1-\pi_f)}{n}}$.
To nam daje aproksimativni $(1-\alpha)\times 100\%$-pouzdani interval za proporciju π:

$$\tilde{\pi}_f \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{\tilde{\pi}_f\left(1-\tilde{\pi}_f\right)}{n}}.$$

Usporedba intervala pouzdanosti i intervala vjerodostojnosti za proporciju π

Bayesovski intervali su dosta korisniji statističarima jer se ne moraju brinuti ili vršiti računanje na događajima koji su se mogli dogoditi, ali nisu.

Primjer 8. *(nastavak Primjera 7. iz Poglavlja 3.) Sjetimo se da je u slučajnom uzorku od n=100 građana grada Hamiltona njih y=26 podržavalo izgradnju casina. Frekvencionistički 95%-interval pouzdanosti za proporciju π je :*

$$0.26 \pm 1.96\sqrt{\frac{0.26\times 0.74}{100}}$$
$$= (0.174, 0.346).$$

Usporedimo li to s intervalima vjerodostojnosti u Tablici 3.2., vidimo da se frekvencionistički interval najviše podudara s Aninim intervalom, a sličan je i preostalim.

4.5 Testiranje hipoteza

Treći tip statističkog zaključivanja koje promatramo je testiranje hipoteza. Statistička hipoteza je bilo koja pretpostavka o populacijskoj razdiobi X. Ona pruža pouzdan okvir za donošenje odluka o podacima koji nas zanimaju. Budući da je ovo jako zastupljena metoda u znanosti, pokazat ćemo i kako je izvesti Bayesovskim pristupom. Pojmovi su preuzeti iz [3] i [4].

Primjerice, pretpostavimo da nas zanima je li točna hipoteza:

$$H_0 : \mu \leq 50.$$

Zapravo, želimo na temelju realizacije slučajnog uzorka za X odlučiti hoćemo li odbaciti ili ne odbaciti hipotezu. Postupak donošenja odluke zove se testiranje statističke hipoteze. Primjetimo, postoji i njoj alternativna hipoteza:

$$H_1 : \mu > 50.$$

Neka je $X_1, \ldots, X_n$ slučajni uzorak za X,

$$X = (X_1, \ldots, X_n)\,.$$

Tada su realizacije $x = (x_1, x_2, \ldots, x_n)$ tog uzorka elementi od $\mathbb{R}^n$.

Definicija 4.5.1. *Test (hipoteze H_0 u odnosu na alternativu H_1) je preslikavanje $\tau : \mathbb{R}^n \to \{0, 1\}$.*

Interpretacija: Ako je za realizaciju x uzorka X, $\tau(x) = 1$, tada odbacujemo H_0 u korist H_1, a ako je $\tau(x) = 0$, tada ne odbacujemo H_0 u korist H_1.

Kritično područje za test τ, odnosno područje realizacije uzorka za koje se H_0 odbacuje u korist H_1, definiramo kao:

$$C = \tau^{-1}(1) = \{x \in \mathbb{R}^n : \tau(x) = 1\}.$$

Nadalje, neka X prati razdiobu zadanu funkcijom gustoće f, $X \sim f(x|\theta), \theta \in \Theta$. Funkcija vjerodostojnosti od θ:

$$L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta).$$

Jakost testa τ je preslikavanje $\gamma : \Theta \to [0, 1]$ definirano s:

$$\gamma(\theta) = \mathbb{E}[\tau(X)] = \mathbb{P}_\theta(X \in C) = \int_C L(\theta|x)dx\,.$$

Interpretacija: Ukoliko je θ_1 vrijednost parametra, za koju je H_1 istinito, jakost testa $\gamma(\theta_1)$ je sposobnost testa da odbaci H_0 ako je H_0 neistinita hipoteza.

Neka su:

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta_1.$$

Preslikavanje $\alpha : \Theta_0 \to [0, 1]$ definirano s:

$$\alpha(\theta) = \gamma(\theta) = \mathbb{P}_\theta(X \in C)$$

je vjerojatnost pogreške 1. vrste.

$$\alpha_\tau = sup_{\theta \in \Theta_0} \alpha(\theta)$$

je značajnost testa τ.
Kažemo da test ima razinu značajnosti α ako mu je značajnost manja ili jednaka α.

U daljnjem radu proučavat ćemo kako tretman utječe na ishod.

4.6 Jednostrane hipoteze

U slučaju parametarskih modela za populacijske razdiobe pretpostavka koju testiramo bit će bilo koja izjava o vrijednostima parametara.

Frekvencionistički jednostrani testovi

Neka μ_0 označava neki unaprijed zadani broj. Testiramo:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

ili

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0.$$

Primjer 9. *Pretpostavimo da želimo saznati je li novi lijek bolji od standardnog lijeka kojeg primaju pacijenti za neku određenu bolest. Ako to vrjedi, značilo bi da je π, proporcija pacijenata kojima se novi lijek pokazao bolji, veća od π_0, proporcije pacijenata kojima bolje djeluje stari lijek. Iz starijih istraživanja zna se da je $\pi_0 = 0.6$. Slučajno odabranoj grupi od 10 pacijenata dan je novi lijek. Y, broj pacijenata kojima je bolji novi lijek, prati binomnu (n, π) raspodjelu. Opazili smo y=8. To je bolje nego što bismo očekivali uz $\pi_0 = 0.6$. Ali, je li to dovoljno veća proporcija da donesemo generalan zaključak na razini značajnosti od 10%?*

Koraci :

1. *Postavimo nultu hipotezu o parametru (fiksan, ali nepoznat). Primjerice, $H_0 : \pi \leq 0.6$ (proporcija pacijenata kojima više pomaže novi lijek je manja ili jednaka od proporcije pacijenata kojima više pomaže stari lijek). Uključujemo sve π vrijednosti koje su manje ili jednake od 0.6 jer želimo odrediti je li novi lijek bolji. Nije nam u interesu odrediti da je novi lijek gori. Naime, nećemo preporučiti novi lijek ukoliko nije dokazivo da je bolji od staroga.*

2. *Postavimo alternativnu hipotezu: $H_1 : \pi > 0.6$.*

3. *Uzoračka distribucija testne statistike, uz pretpostavku da vrijedi nulta hipoteza, prati binomnu $(n, 0.6)$ raspodjelu gdje je n=10 broj pacijenata koji su primili novi lijek.*

4. *Biramo razinu značajnosti testa tako da bude što bliža $\alpha = 5\%$. Budući da y ima diskretnu raziobu, a α može poprimiti tek neke vrijednosti, morat ćemo odabrati razinu značajnosti koja je ili malo ispod ili malo iznad zadanih 5%.*

5. *Kritično područje biramo tako da ima vjerojatnost α u uvjetima nulte hipoteze. Ako odaberemo kritično područje $y \geq 9$, onda je $\alpha = 0.0463$, što vidimo u Tablici 4.1.*

6. *Ako vrijednost testne statistike na danom uzorku upada u kritično područje, onda odbacujemo nultu hipotezu na razini značajnosti α. U suprotnom, ne odbacujemo nultu hipotezu. U našem slučaju y=8, što ne upada u kritično područje.*

7. *Vrijednost koja mjeri jačinu dokaza protiv nulte hipoteze je p-vrijednost, a u ovom primjeru iznosi 0.1672.*

8. *Ako je p-vrijednost $< \alpha$, testna statistika upada u kritično područje i obratno. Stoga je ekvivalentan način zaključivanja: odbacujemo H_0 ako je p-vrijednost $< \alpha$. Dakle, u našem primjeru p-vrijednost > 0.05, stoga ne možemo odbaciti $H_0 : \pi \leq 0.6$. Dokazi nisu dovoljno jaki da zaključimo da vrijedi alternativna hipoteza.*

U binomnom slučaju p-vrijednost iznosi:

$$p - vrijednost = \sum_{y_{ops}}^{n} f(y|\pi_0),$$

gdje je y_{ops} opažena vrijednost od y.

Vrijednost	$f(y \mid \pi = 0.6)$	kritično područje
0	0.0001	ne odbacujemo
1	0.0016	ne odbacujemo
2	0.0106	ne odbacujemo
3	0.0425	ne odbacujemo
4	0.1115	ne odbacujemo
5	0.2007	ne odbacujemo
6	0.2508	ne odbacujemo
7	0.2150	ne odbacujemo
8	0.1209	ne odbacujemo
9	0.0403	odbacujemo
10	0.0060	odbacujemo

Tablica 4.1 : Distribucija od Y uz nultu hipotezu za jednostran test

Bayesovski jednostrani testovi

Želimo testirati:

$$H_0 : \pi \leq \pi_0$$

$$H_1 : \pi > \pi_0$$

na razini značajnosti α koristeći Bayesovsku metodu. Na sljedeći način možemo izračunati aposteriornu vjerojatnost uz istinitu nultu hipotezu:

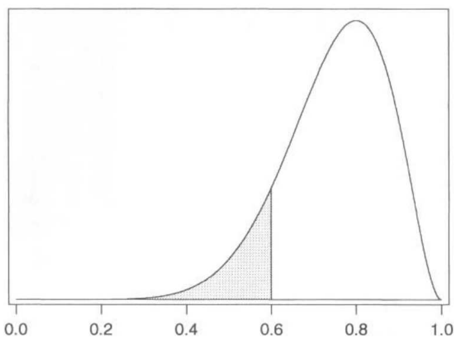
$$\mathbb{P}(H_0 : \pi \leq \pi_0|y) = \int_0^{\pi_0} g(\pi|y)\, d\pi.$$

Odbacujemo nultu hipotezu ako je aposteriorna vjerojatnost manja od razine značajnosti α. Prema tome, Bayesovki jednostrani test je zapravo test "male vjerojatnosti". Ovdje testiramo hipotezu o parametru koristeći aposterioronu distribuciju za parametar.

Primjer 10. *(Nastavak na prethodni Primjer 9.) Pretpostavimo da smo koristili Beta* $(1, 1)$ *apriornu disribuciju za* π*. Tada je, uz y=8, aposteriorna distribucija Beta* $(9, 3)$*. Aposteriorna vjerojatnost, uz uvjet nulte hipoteze, glasi:*

$$\mathbb{P}(\pi \leq 0.6|y = 8) = \int_0^{0.6} \frac{\Gamma(12)}{\Gamma(3)\Gamma(9)} \pi^2 (1-\pi)^8 \, d\pi$$
$$= 0.1189,$$

što nije manje od 0.05, stoga ne odbacujemo nultu hipotezu na razini značajnosti 5%. Na Slici 4.4. [4] *osjenčani dio predstavlja vjerojatnost nulte hipoteze.*



Slika 4.4: Aposteriorna vjerojatnost nulte hipoteze $H_0 : \pi \leq 0.6$

4.7 Dvostrane hipoteze

Dvostrani testovi, kao što smo prije spomenuli, utvrđuju postoji li promjena parametra u oba smjera.

Testiramo:

$$H_0 : \pi = \pi_0$$
$$H_1 : \pi \neq \pi_0$$

[4] Slika je preuzeta iz [1]

Frekvencionistički dvostrani testovi

Primjer 11. *Bacamo novčić 15 puta, a mogući ishodi su pismo (P) ili glava (G). Zabilježili smo 10 G. Je li 10 G u 15 bacanja novčića dovoljno da zaključimo da je novčić simetričan (fer)? Drugim riječima, je li proporcija π, vjerojatnost da padne G, ipak različita od $\frac{1}{2}$?*

1. *Postavimo nultu hipotezu o fiksnom, ali nepoznatom parametru π. Dakle, $H_0 : \pi = 0.5$.*

2. *Alternativna hipoteza je $H_1 : \pi \neq 0.5$. Želimo utvrditi postoji li promjena proporcije u bilo kojem smjeru, stoga imamo dvostrani test kao i dvostrano kritično područje.*

3. *Nulta distribucija je uzoračka distribucija slučajne varijable Y uz uvjet istinite nulte hipoteze. Radi se o binomnoj raspodjeli s parametrima 15 i 0.5.*

4. *Budući da je Y diskretna slučajna varijabla, biramo razinu značajnosti tako da ona bude što je moguće bliže 5%.*

5. *Kritično područje biramo tako da pod uvjetima nulte hipoteze ima vjerojatnost upravo α. Ako odaberemo kritično područje: $Y \leq 3 \cup Y \geq 12$, onda je $\alpha = 0.0352$. U Tablici 4.2. vidimo distribuciju uz H_0 i kritično područje.*

6. *Ako vrijednost testne statistike upada u kritično područje, odbacujemo nultu hipotezu na razini značajnosti α. Ako pak ne upada, ne odbacujemo H_0. U našem slučaju opazili smo y=10 što ne leži u kritičnom području. Dakle, ne odbacujemo nultu hipotezu.*

7. *Budući da se radi o dvostranom testu, p-vrijednost je $\mathbb{P}(Y \geq 10) + \mathbb{P}(Y \leq 5) = 0.302$, što je veće od α, stoga i na ovaj način zaključujemo da ne odbacujemo nultu hipotezu, odnosno da nemamo razloga sumnjati u istinitost nulte hipoteze.*

Vrijednost	$f(y \mid \pi = 0.5)$	kritično područje
0	0.0000	odbacujemo
1	0.0005	odbacujemo
2	0.0032	odbacujemo
3	0.0139	odbacujemo
4	0.0417	ne odbacujemo
5	0.0916	ne odbacujemo
6	0.1527	ne odbacujemo
7	0.1964	ne odbacujemo
8	0.1964	ne odbacujemo
9	0.1527	ne odbacujemo
10	0.0916	ne odbacujemo
11	0.0417	ne odbacujemo
12	0.0139	odbacujemo
13	0.0032	odbacujemo
14	0.0005	odbacujemo
15	0.0000	odbacujemo

Tablica 4.2 : Distribucija od Y uz nultu hipotezu za dvostrani test

Veza između dvostranih testova i intervala pouzdanosti

Kada testiramo dvostranu hipotezu na razini značajnosti α, njoj odgovara $(1-\alpha) \times 100\%$ interval pouzdanosti za parametar. Ako nultu hipotezu

$$H_O : \pi = \pi_0$$

odbacujemo, tada vrijednost π_0 leži van intervala pouzdanosti i obratno. Ako prihvaćamo nultu hipotezu, odnosno ne odbacujemo, tada π_0 leži unutar intervala pozdanosti i obratno. Interval pouzdanosti na neki način objedinjuje sve moguće nulte hipoteze koje ne bi bile odbačene da su bile testirane.

Bayesovski dvostrani testovi

U slučaju Bayesovskih dvostranih testova koristimo sličan način zaključivanja kao kod dvostranih testova i intervala pouzdanosti, samo što ovdje koristimo intervale vjerodostojnosti. Naime, izračunamo $(1-\alpha) \times 100\%$ interval vjerodostojnosti za π. Ako π_0 upada u taj interval, prihvaćamo nultu hipotezu $H_0 : \pi = \pi_0$, a ako ne upada, odbacujemo nultu hipotezu.

Primjer 12. *(nastavak na prethodni Primjer 11.) Ako koristimo uniformnu apriornu distribuciju, aposteriorna je Beta*$(10+1, 5+1)$. *95% Bayesovski interval pouzdanosti za* *proporciju π, uz korištenje normalne aproksimacije:*

$$\frac{11}{17} + 1.96 + \sqrt{\frac{11 \times 6}{(11+6)^2 \times (11+6+1)}}$$

$$= 0.647 \pm 0.221 = (0.426, 0.868)\,.$$

Vrijednost iz nulte hipoteze, $\pi = 0.5$, nalazi se u intervalu vjerodostojnosti, zato ne odbacujemo nultu hipotezu.

Poglavlje 5

Primjeri

U ovom poglavlju predstavit ćemo primjenu obrađene teorije kroz još par zadataka preuzetih iz [1]. Zadatke rješavamo pomoću programskog jezika R, a popratne kodove navodimo u dodatku.

Primjer 5.0.1. *Pretpostavimo da imamo 8 nezavisnih pokusa, a rezultat svakog je uspjeh ili neuspjeh. Vjerojatnost da je pokus uspješan konstantna je za svaki pokus. Neka slučajna varijabla Y broji uspješne pokuse. Dakle, $Y\,|\,\pi$ ima binomnu distribuciju s parametrima n=8 i π. Pretpostavimo da su moguće vrijednosti za π = 0, 0.2, 0.4, 0.6, 0.8 i 1. Radi se o diskretnoj distribuciji parametra π, a svaka moguća vrijednost je jednako vjerojatna.*

π	$g(\pi)$
0	0.1667
0.2	0.1667
0.4	0.1667
0.6	0.1667
0.8	0.1667
1	0.1667

Pretpostavimo da smo opazili 3 uspjeha. Nađite aposteriornu distribuciju od $g(\pi|y)$.

a) Kako izgleda matrica uvjetnih vjerojatnosti?

Rj. Koristeći formulu

$$f(y\,|\,\pi) = \binom{n}{y}\pi^y\,(1-\pi)^{n-y}$$

računamo uvjetne vjerojatnosti. Primjerice, ako je y=1, a $\pi = 0.2$:

$$f(1\,|\,0.2) = \binom{8}{1}0.2^1\,(1-0.2)^{8-1} = 0.3355.$$

Matrica uvjetnih vjerojatnosti izračunatih u R-u:

	0	1	2	3	4	5	6	7	8
0	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.2	0.1678	0.3355	0.2936	0.1468	0.0459	0.0092	0.0011	0.0001	0.0000
0.4	0.0168	0.0896	0.2090	0.2787	0.2322	0.1239	0.0413	0.0079	0.0007
0.6	0.0007	0.0079	0.0413	0.1239	0.2322	0.2787	0.2090	0.0896	0.0168
0.8	0.0000	0.0001	0.0011	0.0092	0.0459	0.1468	0.2936	0.3355	0.1678
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

b) Izdvojite stupac u kojem je izračunata funkcija vjerodostojnosti uz opaženu vrijednost y=3.

Rj. Korištenjem formule (3.1)*, uvrstivši y=3, dobivamo:*

3
0.0000
0.1468
0.2787
0.1239
0.0092
0.0000

c) Nađite zajedničke vjerojatnosti.

Rj. Nalazimo ih korištenjem formule:

$$g(\pi) \times f(y \mid \pi),$$

odnosno tako da svaki član prethodne tablice pomnožimo s odgovarajućom vrijednosti apriorne distribucije, u našem slučaju sa $\frac{1}{6}$*. Rezultat dobiven programskim kodom:*

	0	1	2	3	4	5	6	7	8
[1,]	0.1667	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
[2,]	0.0280	0.0559	0.0489	0.0245	0.0076	0.0015	0.0002	0.0000	0.0000
[3,]	0.0028	0.0149	0.0348	0.0464	0.0387	0.0206	0.0069	0.0013	0.0001
[4,]	0.0001	0.0013	0.0069	0.0206	0.0387	0.0464	0.0348	0.0149	0.0028
[5,]	0.0000	0.0000	0.0002	0.0015	0.0076	0.0245	0.0489	0.0559	0.0280
[6,]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1667

d) Odredite marginalne vjerojatnosti od Y.

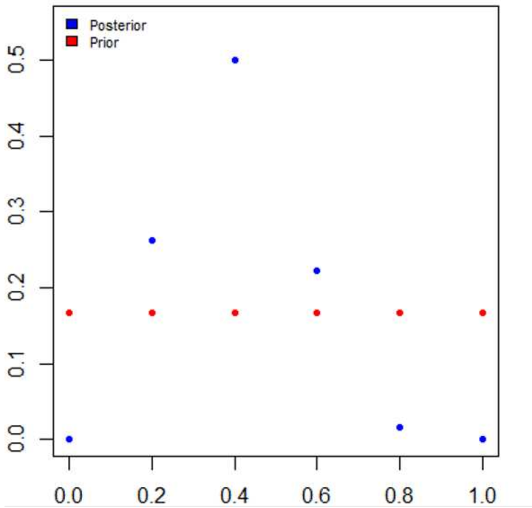
Rj. Marginalne vjerojatnosti dobijemo tako da sumiramo svaki stupac zasebno:

0	1	2	3	4	5	6	7	8
0.1975	0.0722	0.0908	0.0931	0.0927	0.0931	0.0908	0.0722	0.1975

e) Izračunajte aposteriornu distribuciju.

Rj. Aposteriornu distribuciju nalazimo tako da svaki član stupca koji odgovara y=3 podijelimo sa sumom tog stupca. Dobivamo tablicu koja sumira sve tražene podatke:

	Prior	Likelihood	Posterior
0	0.1666667	0.00000000	0.0000000
0.2	0.1666667	0.14680064	0.2628337
0.4	0.1666667	0.27869184	0.4989733
0.6	0.1666667	0.12386304	0.2217659
0.8	0.1666667	0.00917504	0.0164271
1	0.1666667	0.00000000	0.0000000



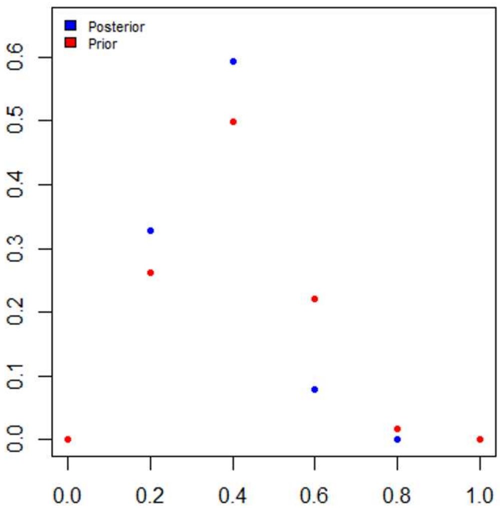
Slika 5.1: Apriorna i aposteriorna distribucija

Primjer 5.0.2. *Pretpostavimo da smo napravili dodatnih 7 pokusa u kojima smo zabilježili 2 uspjeha.*

a) Upotrijebite aposteriornu distribuciju iz prethodnog primjera kao apriornu distribuciju u ovome primjeru. Nađite novu aposteriornu distribuciju za π.

Rj. Na isti način kao u prethodnom primjeru, korištenjem programskog jezika R, dobivamo:

	Prior	Likelihood	Posterior
0	0.0000000	0.0000000	0.0000000000
0.2	0.2628337	0.2752512	0.3289134253
0.4	0.4989733	0.2612736	0.5927126725
0.6	0.2217659	0.0774144	0.0780526976
0.8	0.0164271	0.0043008	0.0003212045
1	0.0000000	0.0000000	0.0000000000

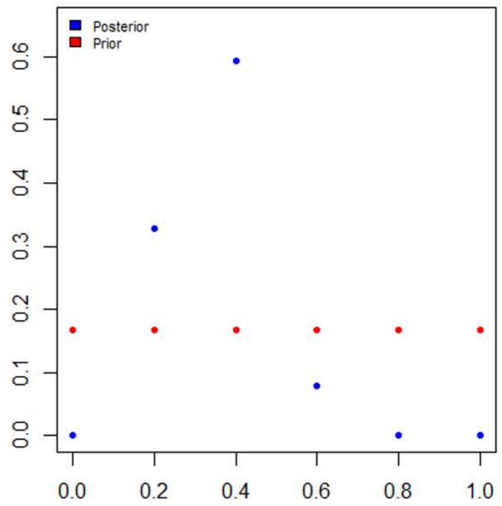


Slika 5.2: Apriorna i aposteriorna distribucija

b) Ukupno imamo 15 pokusa te 5 uspjeha. Koristeći apriornu distribuciju kao u prethodnom primjeru, izračunajte aposteriornu distribuciju.

Rj. Vidimo da smo dobili iste aposteriorne distribucije, a krenuli smo s drugačijim apriornim distribucijama. Razlog tome je što analiza opažanja u jednom koraku i sekvencijalno vode do istih rezultata, kako je i opisano u 2.1.

	Prior	Likelihood	Posterior
0	0.1666667	0.000000000	0.0000000000
0.2	0.1666667	0.103182294	0.3289134253
0.4	0.1666667	0.185937845	0.5927126725
0.6	0.1666667	0.024485642	0.0780526976
0.8	0.1666667	0.000100764	0.0003212045
1	0.1666667	0.000000000	0.0000000000



Slika 5.3: Apriorna i aposteriorna distribucija

Primjer 5.0.3. *Koristit ćemo funkcije u R-u kako bismo izračunali aposteriornu distribuciju u slučaju da $Y \mid \pi$ ima binomnu distribuciju s parametrima n i π, a apriorna distribucija za π je Beta(a, b). Znamo da je beta familija konjugirana za binomnu distribuciju. To znači da ako koristimo apriornu distribuciju iz beta familije, dobivamo aposteriornu distribuciju iz te iste familije (detaljnije u 3.2).*

Pretpostavimo da smo proveli 15 pokusa, svaki rezultira uspjehom ili neuspjehom te da smo opazili 6 uspjeha. Neka je Beta(1,1) apriorna distribucija.

a) Koliko iznose aposteriorna srednja vrijednost i standardna devijacija?

Rj. U slučaju beta apriorne distribucije vrlo lako računamo srednju vrijednost i standardnu devijaciju korištenjem formula (3.9) *i* (3.13)*, pri čemu je* $a' = 1 + 6$ *i* $b' = 1 + 15 - 6$:

$$m' = \frac{7}{17} = 0.412,$$

$$s' = \sqrt{\frac{7 \times 10}{(7+10)^2\,(7+10+1)}} = 0.116.$$

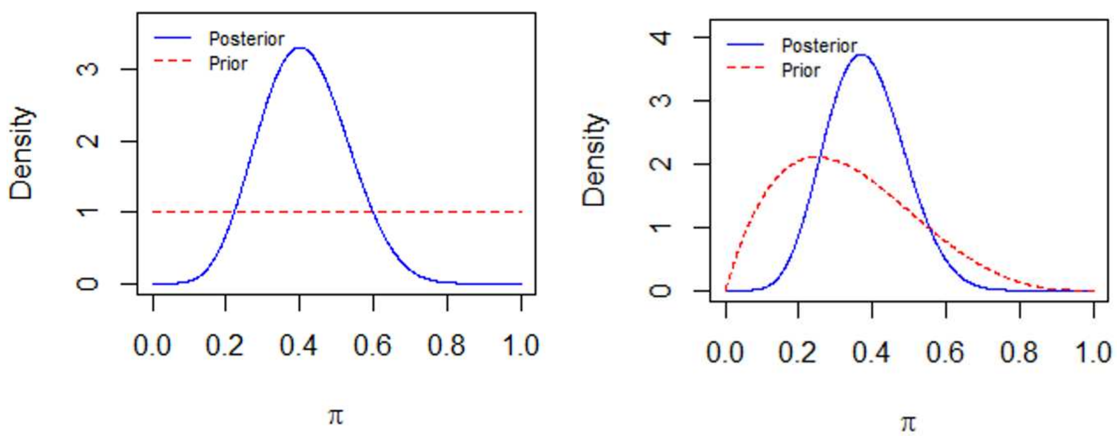
b) Nađite 95% *interval vjerodostojnosti za* π.

Rj. Korištenjem formule (3.14) *računamo traženi interval te dobivamo:*

$$\left[\frac{7}{17} - 1.96 \times 0.116, \frac{7}{17} + 1.96 \times 0.116\right]$$
$$= [0.184, 0.639]\,.$$

c) Ponovite a) i b) za Beta(2,4) distribuciju. Nacrtajte obe aposteriorne distribucije te usporedite rezultate dobivene za aposteriorne srednje vrijednosti, standardnu devijaciju i interval vjerodostojnosti.

Rj. Prikažimo apriorne i aposteriorne distribucije za Beta(1,1) i Beta(2,4) dobivene pomoću programskog koda iz Dodatka s kraja rada.



Slika 5.4: Usporedba apriornih i aposteriornih distribucija za Beta(1,1) i Beta(2,4)

Uočavamo da su aposteriorne distribucije poprilično slične iako smo krenuli s različitim apriornim distribucijama. Razlog tomu je beta familija distribucija koja je konjugirana za binomnu distribuciju.

Za Beta(2,4) distribuciju dobivamo:

$$m' = \frac{8}{8+13} = 0.381,$$

$$s' = \sqrt{\frac{8 \times 13}{(8+13)^2\,(8+13+1)}} = 0.104,$$

$$\textit{te interval vjerodostojnosti } [0.177, 0.585].$$

Uočavamo da su srednje vrijednosti i standardne devijacije dosta slične u oba slučaja. Interval vjerodostojnosti je malo uži, odnosno precizniji u slučaju Beta(2,4) budući da ta apriorna distribucija daje više vjerojatnosti manjim vrijednostima parametra π*, za razliku od Beta(1,1) apriorne distribucije koja svim vrijednostima daje jednaku vjerojatnost.*

Za računanje 95% *intervala vjerodostojnosti u R-u koristimo 0.025 i 0.975 kvantile. Priloženi su rezulati dobiveni pomoću funkcije* binobp *u R-u. Vidimo da intervali pouzdanosti ne odgovaraju u potpunosti računski dobivenim intervalima, a to je zato što je aproksimacija jako dobro za* $a' \geq 10$ *i* $b' \geq 10$*, što u našem slučaju nije vrijedilo. Srednje vrijednosti i varijanca se podudaraju.*

```
> binobp(6,15,1,1)
Posterior Mean           :  0.4117647
Posterior Variance       :  0.0134564
Posterior Std. Deviation :  0.1160016

Prob.   Quantile
------  ---------
0.005   0.1471041
0.010   0.1664597
0.025   0.1975341
0.050   0.2266916
0.500   0.4082265
0.950   0.6089884
0.975   0.6456539
0.990   0.6865909
0.995   0.7132290
> binobp(6,15,2,4)
Posterior Mean           :  0.3809524
Posterior Variance       :  0.0107194
Posterior Std. Deviation :  0.1035347

Prob.   Quantile
------  ---------
0.005   0.1459844
0.010   0.1634200
0.025   0.1911901
0.050   0.2170686
0.500   0.3771052
0.950   0.5580345
0.975   0.5921885
0.990   0.6309392
0.995   0.6565686
```

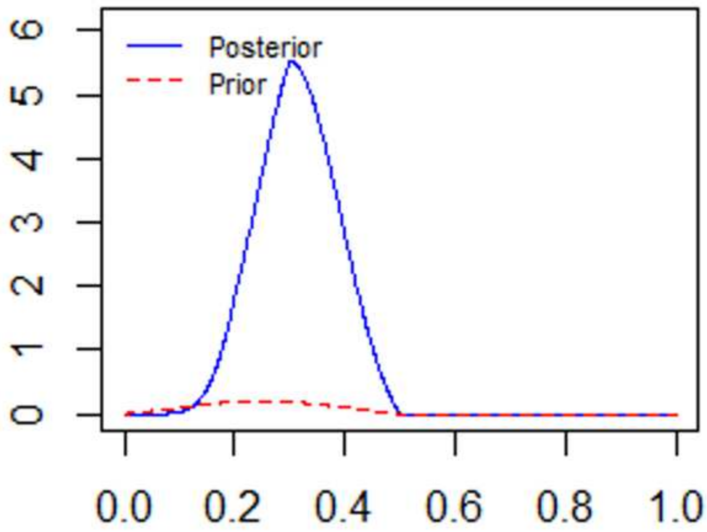
Primjer 5.0.4. *Pretpostavimo da $Y\,|\,\pi$ ima binomnu distribuciju s parametrima n i π, gdje je apriorna distribucija parametra π dana s:*

$$g(\pi) = \begin{cases} \pi & za\ \pi \le 0.2 \\ 0.2 & za\ 0.2 < \pi \le 0.3 \\ 0.5-\pi & za\ 0.3 < \pi \le 0.5 \\ 0 & za\ 0.5 < \pi \end{cases}$$

Pretpostavimo da smo u n=20 nezavisnih pokusa opazili y=7 uspjeha.

a) Odredite aposteriornu distribuciju od $g(\pi\,|\,y)$.

Rj. Programskim kodom dobijemo:



Slika 5.5: Aposteriorna distribucija $g(\pi\,|\,y)$

b) Pronađite aposteriornu srednju vrijednost i standardnu devijaciju.

Rj. Nalazimo ih računajući (3.8) *i* (3.11) *(korjenujemo da dobijemo standardnu devijaciju). Integrale računamo pomoću programskog koda koristeći funkciju* sintegral *te dobivamo:*

$$m' = 0.312,$$

$$s' = 0.071.$$

c) Nađite 95% interval vjerodostojnosti.

Rj. Pomoću R-a dobivamo interval:

$$[0.174, 0.446].$$

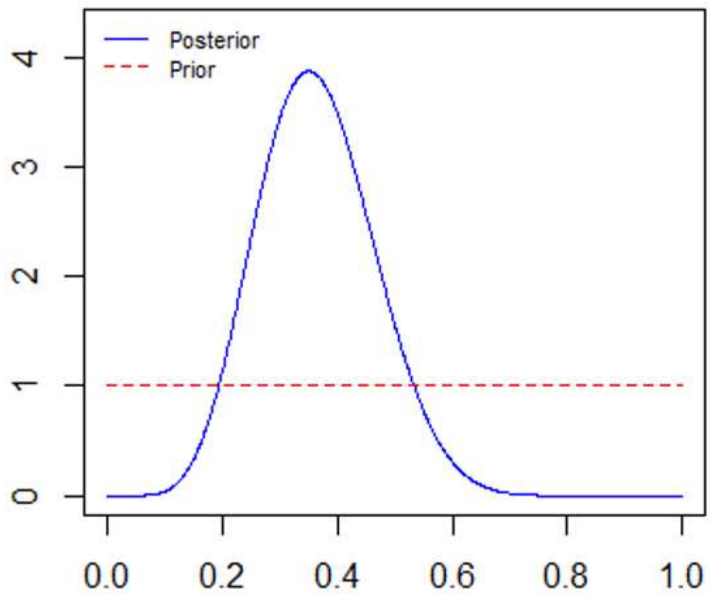
d) Ponovite sve za uniformnu distribuciju parametra π.

Rj. Znamo da je uniformna distribucija poseban slučaj Beta distribucije s parametrima a=b=1. Provedemo li sličan kod kao u prethodnim podzadacima, dobijemo:

$$m' = 0.365,$$

$$s' = 0.100,$$

$$\textit{interval vjerodostojnosti } [0.180, 0.569].$$



Slika 5.6: Aposteriorna distribucija $g(\pi \mid y)$

Dodatak

R kod za Primjer 5.0.1

```r
library('Bolstad')

pi=c (0,0.2,0.4,0.6,0.8,1)
pi.prior=rep(1/6,6)

results=binodp(3,8,uniform=TRUE,pi=pi, pi.prior=pi.prior,ret=TRUE)
```

R kod za Primjer 5.0.2

```r
library('Bolstad')

pi=c (0,0.2,0.4,0.6,0.8,1)
pi.prior=rep(1/6,6)

results=binodp(3,8,uniform=TRUE,pi=pi, pi.prior=pi.prior,ret=TRUE)

new_posterior_distribution=results$posterior

# analiza sekvencijalno
new_results=binodp(2,7,uniform=FALSE,pi=pi,pi.prior = new_posterior_distribution,ret=TRUE)

# analiza u jednom koraku
new_results_2=binodp(5,15,uniform=TRUE,pi=pi, pi.prior=pi.prior,ret=TRUE)
```

R kod za Primjer 5.0.3

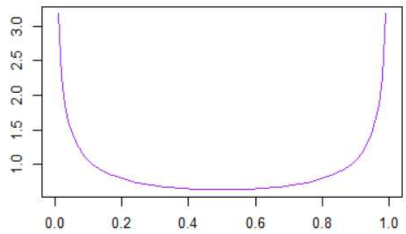
```
library('Bolstad')

#zadatak vrlo jednostavno rjesavamo koristenjem funkcije binobp

#a)
binobp (6,15,1,1)

#c)
binobp (6,15,2,4)
```

R kod za Primjer 5.0.4

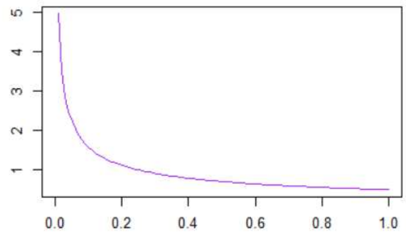
```
library('Bolstad')
pi=seq(0,1,by=0.001)
pi.prior=rep(0,length(pi))

pi.prior[pi<=0.2]=pi[pi<=0.2]
pi.prior[0.2<=pi & pi<=0.3]
pi.prior[0.2<pi & pi<=0.3]=0.2
pi.prior[0.3<pi & pi<=0.5]=0.5-pi[0.3<pi & pi<=0.5]
pi.prior[0.5<pi]=0

results=binogcp (7,20,"user",pi=pi,pi.prior=pi.prior,ret=TRUE)

dens=pi*results$posterior
post.mean=sintegral(pi,dens)
dens=(pi-post.mean$int)^2*results$posterior
post.var=sintegral(pi,dens)
post.sd=sqrt(post.var$int)

post.mean$int
post.sd

cdf=sintegral(pi,results$posterior,n.pts=length(pi))

lb=cdf$x[with(cdf,which.max(x[y<=0.025]))]
ub=cdf$x[with(cdf,which.max(x[y<=0.975]))]
lb
ub

cat(paste("Aproksimativni 95% interval vjerodostojnosti : [",round(lb,3)
    ,",", round (ub ,3),"]\n",sep=""))
```

```


#d) Ponovimo sve za uniformnu distribuciju parametra

pi=seq(0,1,by=0.001)
results=binobp(7,20,1,1)

dens=pi*results$posterior
post.mean=sintegral(pi,dens)
dens=(pi-post.mean$int)^2*results$posterior
post.var=sintegral(pi,dens)
post.sd=sqrt(post.var$int)

post.mean$int
post.sd

cdf=sintegral(pi,results$posterior,n.pts=length(pi))

lb=cdf$x[with(cdf,which.max(x[y<=0.025]))]
ub=cdf$x[with(cdf,which.max(x[y<=0.975]))]
lb
ub

cat(paste("Aproksimativni 95% interval vjerodostojnosti: [",round(lb,3)
    ,",", round (ub ,3),"]\n",sep=""))
```

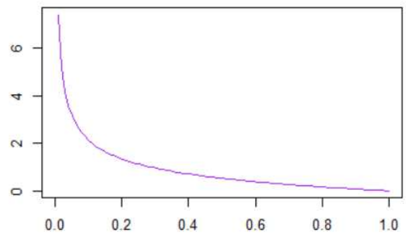
Beta distribucije



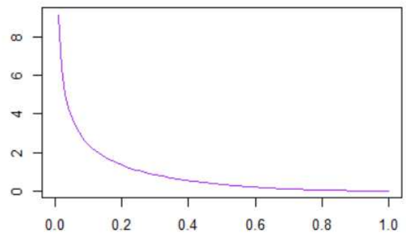
beta(0.5,0.5)



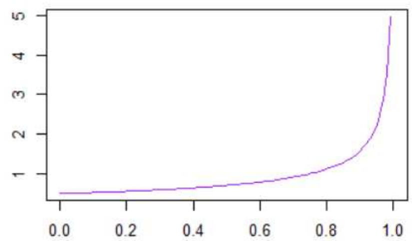
beta(0.5,1)



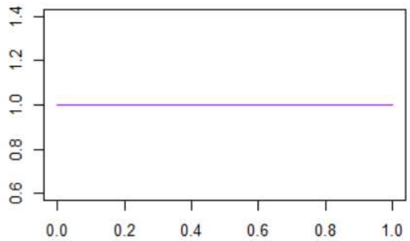
beta(0.5,2)



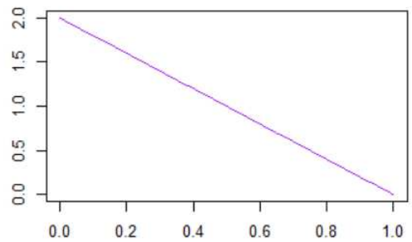
beta(0.5,3)



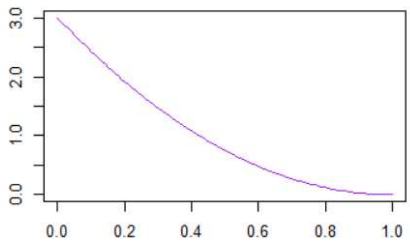
beta(1,0.5)



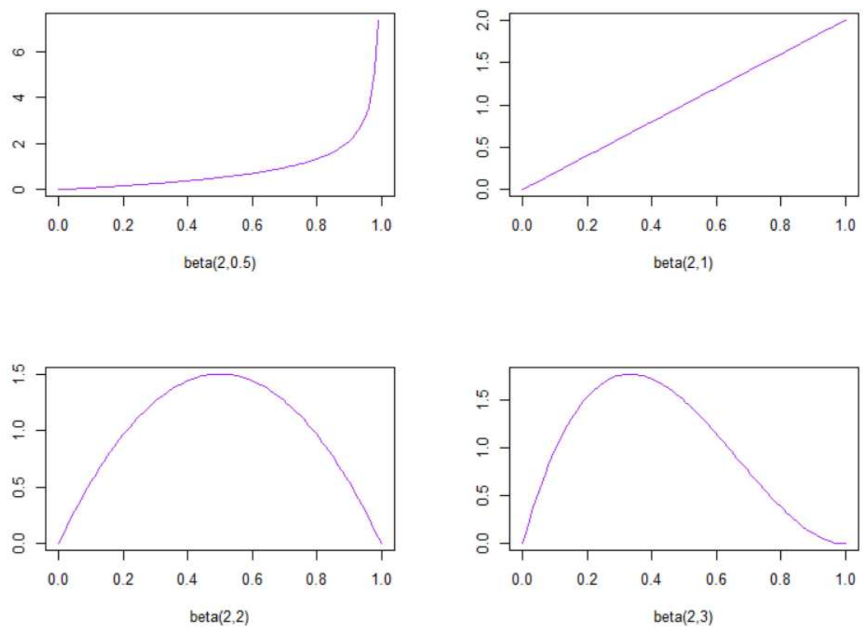
beta(1,1)

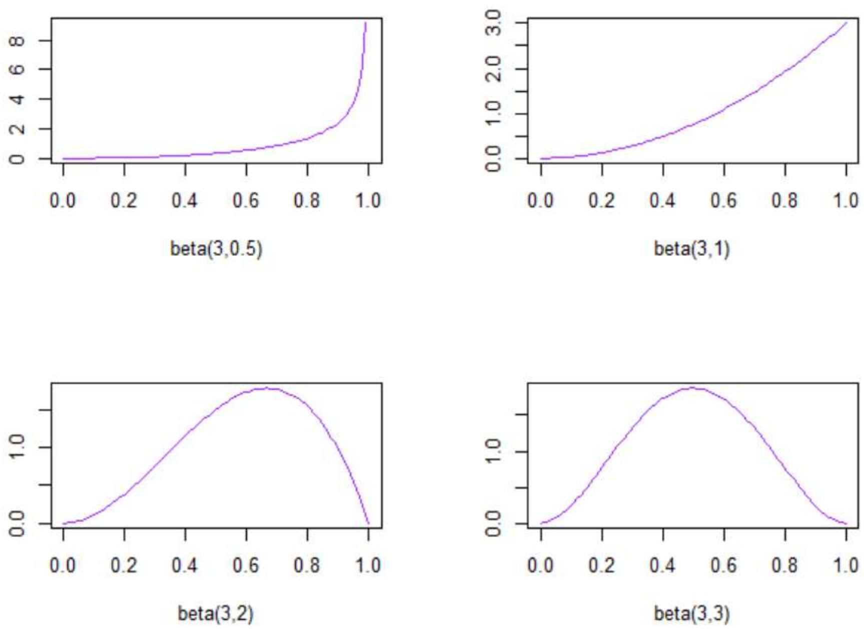


beta(1,2)



beta(1,3)


beta(2,0.5)
beta(2,1)
beta(2,2)
beta(2,3)


beta(3,0.5)
beta(3,1)
beta(3,2)
beta(3,3)

Bibliografija

[1] WM Bolstad, *Introduction to Bayesian Statistics. America: A John Wiley & Sons*, 2007.

[2] Lawrence D Brown, *Fundamentals of statistical exponential families: with applications in statistical decision theory*, Ims, 1986.

[3] Miljenko Huzak, *Vjerojatnost i matematicka statistika, predavanja*, PMF-MO, Zagreb (2006).

[4] ———, *Matematicka statistika*, PMF-MO predavanja (2012).

[5] N Sandrić i Z Vondraček, *Vjerojatnost–predavanja. 2019*.

[6] Nikola Sarapa, *Teorija vjerojatnosti*, Školska knjiga, 1987.

Sažetak

U ovom radu opisuje se Bayesovsko zaključivanje za proporciju binomne razdiobe. U uvodnom poglavlju uvedeni su osnovni pojmovi iz vjerojatnosti i statistike koji su potrebni za izgradnju teorije na kojoj se temelji Bayesov teorem. Također, u sklopu prvog poglavlja iskazana je diskretna i neprekidna verzija Bayesova teorema te su navedene distribucije koje smo koristili. Uvode se i ključni pojmovi vezani za teorem: apriorna i aposteriorna distribucija.

U drugom poglavlju opisuje se Bayesovsko zaključivanje za diskretne slučajne varijable- binomnu i Poissonovu. Prikazani primjeri predstavljaju primjenu teorema i vode do zaključka da podatke možemo analizirati istovremeno ili sekvencijalno te u oba slučaja doći do istih aposteriornih distribucija.

Treće poglavlje predstavlja okosnicu ovog rada, Bayesovsko zaključivanje za proporciju binomne razdiobe. Opisani su načini odabira apriornih distribucija, a obrađene su uniformna i Beta apriorna distribucija. Uvodi se bitan pojam konjugirane familije distribucija te se pokazuje da je Beta familija konjugirana za podatke koji dolaze iz binomne raspodjele. Objašnjeno je kako donijeti zaključke o aposteriornim distribucijama. Također, uvode se Bayesovski intervali vjerodostojnosti.

U četvrtom poglavlju uspoređujemo Bayesovske i frekvencionističke zaključke o parametru do kojih dolazimo točkovnom procjenom, procjenom intervala i testiranjem hipoteza. Razmatra se razlika u interpretaciji parametra, definiraju se procjenitelji, uspoređuju se intervali vjerodostojnosti i pouzdanosti te se promatraju jednostrane i dvostrane hipoteze s oba stajališta.

U posljednjem poglavlju nalaze se zadaci na kojima su primjenjene proučene metode. Zadaci su riješeni uz pomoć programskog jezika R, a kodovi su priloženi na kraju rada.

Summary

This paper describes Bayesian inference for a proportion of the binomial distribution. In the introductory chapter we introduce the basic concepts from probability and statistics that are needed to build the theory on which Bayes theorem is based. Also, in the first chapter, discrete and continuous versions of Bayes theorem are presented, along with the distributions we used. Key concepts related to the theorem are also introduced: a prior and a posterior distributions.

The second chapter describes Bayesian inference for discrete random variables, binomial and Poisson. The presented examples represent the application of the theorem and lead to the conclusion that we can analyze the data simultaneously or sequentially and in both cases arrive at the same posterior distributions.

The third chapter presents the backbone of this paper, Bayesian inference for a proportion of the binomial distribution. This chapter describes ways of selecting a prior distributions, and the uniform and Beta prior distributions are processed. The important notion of the conjugate family of distributions is introduced and it is shown that the Beta family is conjugate for data coming from the binomial distribution. It is explained how to draw conclusions about posterior distributions. Also, Bayesian credibility intervals are introduced.

In the fourth chapter, we compare Bayesian and frequentist conclusions about the parameter that we reach by point estimation, interval estimation and hypothesis testing. The difference in parameter interpretation is considered, estimators are defined, credibility and confidence intervals are compared, and one-side and and two-side hypotheses are observed from both points of view.

In the last chapter, there are problems to which the studied methods were applied. Problems were solved using the programming language R, and the codes are attached at the end of the paper.

Životopis

Rođena sam 10.7.1997. u Šibeniku, gdje upisujem Osnovnu školu Petra Krešimira IV. Srednjoškolsko obrazovanje stječem u Gimnaziji Antuna Vrančića, prirodoslovno-matematički smjer. Nakon toga selim u Zagreb te 2016. upisujem preddiplomski studij Matematika na Prirodoslovno-matematičkom fakultetu Zagrebačkog sveučilišta. Potom 2020. upisujem diplomski sveučilišni studij Matematička statistika.