Bayesovsko zaključivanje za očekivanje normalne distribucije

Dunković, Dora

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Dora Dunković

BAYESOVSKO ZAKLJUČIVANJE ZA OČEKIVANJE NORMALNE DISTRIBUCIJE

Diplomski rad

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

Zagreb, studeni 2021.

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. ____________________________

Roditeljima, za beskonačnu ljubav i podršku.
Prijateljima, za sve pozitivne misli i motivaciju.
I desetogodišnjoj sebi, koja je s divljenjem gledala zgradu sa zlatnim slovima.

Sadržaj

Uvod

Statističko zaključivanje je proces kojim se, na temelju danih podataka, donose zaključci o čitavoj populaciji kojoj podaci pripadaju. Dva su osnovna pristupa statističkom zaključivanju: frekvencionistički i Bayesovski. Bez obzira na pristup koji se koristi, statističko zaključivanje zahtjeva ekstrapolaciju iz poznatog manjeg uzorka na veću nepoznatu populaciju stoga nikada ne možemo biti sigurni da su dobiveni rezultati u potpunosti točni. Ta nesigurnost se izražava kroz uporabu vjerojatnosti, bilo da se radi o intervalima pouzdanosti ili vjerodostojnosti, apriornim ili aposterirornim distribucijama.

Odabir pristupa koji koristimo za donošenje zaključaka bazira se na načinu na koji interpretiramo vjerojatnosti. Frekvencionistički ili klasični pristup vjerojatnosti doživljava kao frekvencije ponavljajućeg pokusa. Nasuprot tome, Bayesovski pristup frekvencije interpretira kao stupanj uvjerenja o događaju, temeljen na dostupnim informacijama. Prednost takvog subjektivnog shvaćanja vjerojatnosti jest mogućnost njihovog korištenja i u slučaju događaja koji se ne mogu ponavljati ili se događaju iznimno rijetko.

Temelje Bayesovske statistike postavio je Thomas Bayes u svom radu *An Essay Towards Solving Problem in the Doctrine of Chances*, koji je 1763. godine, dvije godine nakon autorove smrti, objavio njegov prijatelj Richard Price. U svom radu, Bayes je predstavio matematičko pravilo, tj. teorem za računanje uvjetnih vjerojatnosti koji danas nosi njegovo ime. Neovisno o Bayesu, francuski matematičar i astronom Pierre-Simon Laplace dolazi do istog rezultata u općenitijem obliku te ga objavljuje 1774. godine. Slijede godine intenzivnijeg korištenja Bayesovih metoda, sve do početka 20. stoljeća kada se javlja uvjerenje kako se znanosti ne mogu temeljiti na subjektivnim pretpostavkama. Do sredine 20. stoljeća interes za Bayesovskom statistikom je ponovno porastao te Bayesov teorem i danas nalazi širok spektar primjene koji uključuje genetiku, lingvistiku, kozmologiju, epidemiologiju, psihologiju, forenziku, evoluciju, ekologiju pa čak i rad fiktivnog detektiva Sherlocka Holmesa.

Neka od najvažnijih dostignuća koja koriste Bayesov teorem su dekodiranje Njemačke *Enigme* za vrijeme Drugog Svjetskog Rata, istraživanje koje dokazuje da cigarete uistinu uzrokuju rak pluća te filteri elektroničke pošte koji filtriraju neželjenu poštu.

Cilj ovog rada je opisati zaključivanje o parametru očekivanja normalne distribucije kada je parametar očekivanja slučajna varijabla te pokazati osnovne razlike između Bayesovskog pristupa i standardnog frekvencionističkog pristupa.

U prvom poglavlju uvodimo osnovne pojmove i rezultate koji su nam potrebni za razumijevanje Bayesovog teorema te njegovih posljedica u svrhu određivanja nepoznatog parametra očekivanja normalne slučajne varijable.

U drugom poglavlju upoznajemo se s diskretnom i neprekidnom verzijom Bayesovog teorema i njegovom osnovnom primjenom.

U trećem poglavlju uvodimo Bayesovske metode temeljene na Bayesovom teoremu kako bismo odredili vrijednost parametra očekivanja u diskretnom i neprekidnom slučaju. Također uvodimo različite apriorne distribucije i na primjerima ilustriramo njihov utjecaj na aposteriornu distribuciju.

U četvrtom poglavlju objašnjavamo razlike frekvencionističkog i Bayesovskog pristupa na točkovnoj i intervalnoj procjeni te statističkim testovima.

Konačno, u petom poglavlju primjenjujemo uvedenu teoriju i metode na konkretnim primjerima koje rješavamo u programskom jeziku R.

Poglavlje 1

Osnovni pojmovi

Kako bismo razumjeli Bayesovsko zaključivanje temeljeno na Bayesovom teoremu potrebno je definirati osnovne pojmove vjerojatnosti i statistike na kojima se ono temelji. U ovom poglavlju navodimo definicije i rezultate koji će nam biti potrebni za iskazivanje i razumijevanje daljnje teorije. Definicije su uglavnom preuzete iz [8].

1.1 Vjerojatnosni prostor

Neka je Ω proizvoljan neprazan skup. Označimo s $\mathcal{P}(\Omega)$ partitivni skup od Ω.

Definicija 1.1.1. *Familiju podskupova $\mathcal{F}$ od Ω $(\mathcal{F} \subset \mathcal{P}(\Omega))$ nazivamo* $\sigma -$ algebra *ako vrijede sljedeća tri svojstva:*
(i) $\mathcal{F} \in \Omega$;
(ii) (zatvorenost na komplement) Ako je $A \in \mathcal{F}$ onda je i $A^c \in \mathcal{F}$;
(iii) (zatvorenost na prebrojive unije) Ako su $A_j \in \mathcal{F},\ j \in \mathbb{N}$ onda je i $\bigcup_{j=1}^{\infty} A_j \in \mathcal{F}$.
Uređen par $(\Omega, \mathcal{F})$ zovemo izmjeriv prostor.

U idućoj propoziciji navodimo osnovna svojstva σ-algebra. Dokaz propozicije može se pronaći u [8].

Propozicija 1.1.2. *Neka je $\mathcal{F}$ σ-algebra na nepraznom skupu Ω. Tada vrijedi:*
(i) $\emptyset \in \mathcal{F}$;
(ii) (zatvorenost na prebrojive presjeke) Ako su $A_j \in \mathcal{F},\ j \in \mathbb{N}$ onda je i $\bigcap_{j=1}^{\infty} A_j \in \mathcal{F}$;
(iii) (zatvorenost na konačne unije) Za svaki $n \in \mathbb{N}$, ako su $A_1, A_2, ..., A_n \in \mathcal{F}$ onda je i $\bigcup_{j=1}^{n} A_j \in \mathcal{F}$.

Sada smo spremni navesti formalnu definiciju vjerojatnosti.

Definicija 1.1.3. *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 svojstva:*
(i) (nenegativnost) Za sve $A \in \mathcal{F}$ $\mathbb{P}(A) \geq 0$*;*
(ii) (normiranost) $\mathbb{P}(\Omega) = 1$*;*
(iii) (σ*-aditivnost) Za svaki niz* $(A_j)_{j \in \mathbb{N}}$ *po parovima disjunktnih događaja* $A_j \in \mathcal{F}$
$(A_i \cap A_j = \emptyset$ *za* $i \neq j)$ *vrijedi*

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

Uređenu trojku $(\Omega, \mathcal{F}, \mathbb{P})$ *zovemo* vjerojatnosni prostor.

U idućoj propoziciji navodimo neka svojstva vjerojatnosti. Dokaz propozicije može se pronaći u [8].

Propozicija 1.1.4. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor. Tada vrijedi:*
(i) $\mathbb{P}(\emptyset) = 0$*;*
(ii) (konačna aditivnost) Za svaki $n \in \mathbb{N}$ *i svaki konačan niz* $A_1, A_2, ..., A_n \in \mathcal{F}$ *po parovima disjunktnih događaja iz* $\mathcal{F}$ *vrijedi*

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

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

U Bayesovskoj statistici važnu ulogu igra pojam uvjetne vjerojatnosti koji uvodimo u idućoj definiciji.

Definicija 1.1.5. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor te* $B \in \mathcal{F}$ *takav da je* $\mathbb{P}(B) > 0$*.* Uvjetnu vjerojatnost A uz dano B *definiramo formulom*

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

Definicija 1.1.6. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor. Događaji A i B su* nezavisni *ako vrijedi*

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

Napomena 1.1.7. *Za nezavisne događaje A i B, uvjetna vjerojatnost od A uz dano B jednaka je vjerojatnosti događaja A:*

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

1.2 Slučajne varijable

Predmet statističkih istraživanja su često pokusi, tj. mjerenja nad kojima onda provodimo razne analize. Stoga su nam od iznimne važnosti funkcije koje elementarnim događajima, tj.elementima skupa Ω pridružuju realne brojeve. Takve funkcije nazivamo slučajnim varijablama. Njihovu formalnu definiciju te pojmove vezane uz njih navodimo u nastavku.

Definicija 1.2.1. *Neka je* $(\Omega, \mathcal{F}, \mathbb{P})$ *vjerojatnosni prostor.* Slučajna varijabla *na* $(\Omega, \mathcal{F}, \mathbb{P})$ *je svaka funkcija* $X : \Omega \to \mathbb{R}$ *takva da vrijedi*

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}, \quad x \in \mathbb{R}.$$

Definicija 1.2.2. Funkcija distribucije *slučajne varijable X je funkcija* $F : \mathbb{R} \to [0, 1]$ *definirana formulom*

$$F(x) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}.$$

Ukoliko je F neprekidna u $x \in \mathbb{R}$, tada je $\mathbb{P}(X = x) = 0$ za sve $x \in \mathbb{R}$.

Definicija 1.2.3. *Slučajna varijabla* $X : \Omega \to \mathbb{R}$ *je* apsolutno neprekidna *ako postoji* $f : \mathbb{R} \to [0, \infty\rangle$ *takva da za sve* $x \in \mathbb{R}$ *vrijedi*

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

Funkciju f nazivamo funkcija gustoće *slučajne varijable X.*

Napomena 1.2.4. *Ako je f funkcija gustoće neke slučajne varijable, tada vrijedi*
(i) $f(x) \geq 0$
(ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1.$

Primjer 1.2.5. *Odredimo konstantu* $c > 0$ *takvu da funkcija*

$$f(x) = ce^{-\frac{x^2}{2}}, \quad x \in \mathbb{R},$$

bude funkcija gustoće neke slučajne varijable.

Mora vrijediti $\int_{-\infty}^{\infty} f(x)\, dx = 1$ *pa slijedi*

$$1 = \int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} ce^{-\frac{x^2}{2}}\, dx = c \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx. \tag{1.1}$$

Označimo $I := \int_0^\infty e^{-\frac{x^2}{2}}\,dx$. *Tada slijedi*

$$\begin{aligned} I^2 &= \left(\int_0^\infty e^{-\frac{x^2}{2}}\,dx\right)\left(\int_0^\infty e^{-\frac{y^2}{2}}\,dy\right) = \int_0^\infty\int_0^\infty e^{-\frac{x^2+y^2}{2}}\,dx\,dy \\ &= \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-\frac{r^2}{2}}\,r\,dr\,d\theta \\ &= \frac{\pi}{2}\int_0^\infty e^{-\frac{r^2}{2}}\,r\,dr = \frac{\pi}{2}\int_0^\infty e^{-t}\,dt = \frac{\pi}{2} \end{aligned}$$

gdje smo u drugom koraku koristili Fubinijev teorem, u trećem prelazak na polarne koordinate te u petom supstituciju.
Slijedi $I = \sqrt{\frac{\pi}{2}}$. *Nadalje, pošto je podintegralna funkcija u* (1.1) *parna funkcija, vrijedi*

$$\int_{-\infty}^\infty e^{-\frac{x^2}{2}}\,dx = 2\int_0^\infty e^{-\frac{x^2}{2}}\,dx = 2\,I = 2\sqrt{\frac{\pi}{2}}$$

pa konačno slijedi

$$1 = c\,2\sqrt{\frac{\pi}{2}} \quad\Rightarrow\quad c = \frac{1}{\sqrt{2\,\pi}}.$$

Dobivamo da je f oblika

$$f(x) = \frac{1}{\sqrt{2\,\pi}}e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Slučajnu varijablu X s funkcijom gustoće f zovemo standardna normalna slučajna varijabla *i označavamo s* $X \sim N(0,1)$.

Općenito, slučajna varijabla X s funkcijom gustoće oblika

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

naziva se *normalna slučajna varijabla s parametrima* μ *i* σ^2. Oznaka je $X \sim N(\mu, \sigma^2)$.

Sada teoriju proširujemo na slučajne vektore kako bismo mogli definirati pojam uvjetne funkcije gustoće koji se koristi u neprekidnoj verziji Bayesovog teorema.

Definicija 1.2.6. *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 1.2.7. *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:*

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

Definicija 1.2.8. *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*

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

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

Definicija 1.2.9. *Neka je* (X, Y) *slučajan vektor s funkcijom gustoće* $f = f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$. Uvjetna funkcija gustoće *slučajne varijable X uz dano* $Y = y$ *definira se kao*

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

za sve $y \in \mathbb{R}$ *za koje vrijedi* $f_Y(y) > 0$.

Za kraj ovog poglavlja navodimo definicije matematičkog očekivanja te varijance.

Definicija 1.2.10. *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*

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

Navedimo i jedan teorem koji će koristiti kod računanja. Njegov dokaz se može pronaći u [8].

Teorem 1.2.11. *Neka je X apsolutno neprekidna slučajna varijaba s funkcijom gustoće* f_X *i neka je* $g : \mathbb{R} \to \mathbb{R}$ *strogo rastuća i diferencijabilna funkcija. Definirajmo* $Y := g \circ X = g(X)$. *Ako* $\int_{-\infty}^{\infty} |g(x)|\, f_X(x)\, dx < \infty$, *onda Y ima matematičko očekivanje i vrijedi*

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx.$$

Definicija 1.2.12. *Neka je X slučajna varijabla s očekivanjem* $\mathbb{E}(X)$.Varijanca *od X se definira kao*

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}(X))^2].$$

Standardna devijacija *od X je definirana kao* $\sigma(X) := \sqrt{\mathrm{Var}(X)}$.

Uočimo da vrijedi $0 \leq \mathrm{Var}(X) \leq \infty$ i $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Navedena formula je ponekad jednostavnija za računanje od definicije.

Primjer 1.2.13. *Izračunajmo matematičko očekivanje i varijancu normalne slučajne varijable X s parametrima μ i σ^2.*

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\ &= \left[\, y = \frac{x-\mu}{\sigma} \quad dy = \frac{1}{\sigma} dx \,\right] = \int_{-\infty}^{\infty} \frac{\sigma y + \mu}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \\ &= \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy}_{1} + \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} y\, e^{-\frac{y^2}{2}}\, dy}_{0} = \mu.\end{aligned}$$

U predzadnjem koraku smo za prvi integral iskoristili činjenicu da se radi o funkciji gustoće standardne normalne razdiobe, stoga je njen integral 1, a podintegralna funkcija pod drugim integralom je neparna funkcija pa je njen integral na simetričnoj domeni jednak 0. Dakle, očekivanje normalne slučajne varijable s parametrima μ i σ^2 je upravo μ.

Da bismo izračunali varijancu koristit ćemo $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. *Izračunajmo prvo* $\mathbb{E}(X^2)$.

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\ &= \left[\, y = \frac{x-\mu}{\sigma} \quad dy = \frac{1}{\sigma} dx \,\right] = \int_{-\infty}^{\infty} \frac{(\sigma y + \mu)^2}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \\ &= \int_{-\infty}^{\infty} \frac{\sigma^2 y^2}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy + \underbrace{\int_{-\infty}^{\infty} \frac{2\sigma\mu}{\sqrt{2\pi}} y\, e^{-\frac{y^2}{2}}\, dy}_{0} + \mu^2 \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy}_{1} \\ &= \begin{bmatrix} u = -y & dv = -e^{-\frac{y^2}{2}}\, dy \\ du = -dy & v = e^{\frac{y^2}{2}} \end{bmatrix} = -\frac{\sigma^2}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} y\, e^{-\frac{y^2}{2}}\, dy}_{0} + \sigma^2 \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy}_{1} + \mu^2 = \sigma^2 + \mu^2.\end{aligned}$$

Ovdje smo ponovno koristili svojstvo da je integral neparne funkcije na simetričnoj domeni jednak 0 te svojstvo da je integral funkcije gustoće neke neprekidne slučajne varijable jednak 1.

Sada slijedi

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

Dakle, varijanca normalne slučajne varijable s parametrima μ i σ^2 je σ^2.

Ovime smo dobili interpretaciju parametara normalne slučajne varijable.

Poglavlje 2

Bayesov teorem

U vjerojatnosti i statistici, Bayesov teorem ili Bayesovo pravilo jest matematička formula pomoću koje računamo uvjetne vjerojatnosti događaja. Ona nam omogućava da izmijenimo svoja predviđanja nekog događaja na temelju novih saznanja o tom događaju.

U suštini, pomoću Bayesovog teorema spajamo prijašnja saznanja, u obliku tzv. apriorne vjerojatnosti, s uočenim podacima kako bismo protumačili rezultate u svjetlu novih saznanja, u obliku tzv. aposteriorne vjerojatnosti.

Da bismo razumjeli način na koji zagovornici Bayesovske statistike ažuriraju predviđanja potrebno je razumjeti osnovne alate kojima se služe, stoga u nastavku uvodimo Bayesov teorem na kojem se čitava teorija zasniva. Počinjemo s diskretnom verzijom teorema koju ćemo proširiti na neprekidnu. Krajem poglavlja predstavit ćemo osnovne ideje Bayesovskog zaključivanja te uvesti i objasniti pojmove apriorne i aposteriorne vjerojatnosti. Kao izvor poslužili su [1], [4] i [7].

2.1 Diskretna verzija Bayesovog teorema

Pretpostavimo da je Ω konačan ili prebrojiv skup.
U prethodnom poglavlju definirali smo uvjetnu vjerojatnost događaja A uz dano B kao

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

Događaj B možemo zapisati kao B = $(A \cap B) \cup (A^c \cap B)$. Pošto su $(A \cap B)$ i $(A^c \cap B)$ disjunktni događaji vrijedi

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

Uvrstimo li to u definiciju uvjetne vjerojatnosti dobivamo

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)}. \tag{2.1}$$

Analogno možemo definirati uvjetnu vjerojatnost događaja B uz dano A

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

Izrazimo li sada $\mathbb{P}(A \cap B)$ dobivamo

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A). \tag{2.2}$$

Na sličan način možemo napisati $\mathbb{P}(A^c \cap B)$ te sve to uvrstiti u (2.1):

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

Ovime smo dobili Bayesovu fomulu za jedan događaj.

Dobivena formula može se poopćiti na proizvoljnu familiju disjunktnih događaja.

Definicija 2.1.1. *Neka je $(H_i)_{i\in I}$ konačna ili prebrojiva familija događaja iz $\mathcal{F}$ takva da je $\mathbb{P}(H_i) > 0$ za sve $i \in I$, $H_i \cap H_j = \emptyset$ za $i \neq j$ te $\bigcup_{i\in I} H_i = \Omega$. Familiju $(H_i)_{i\in I}$ nazivamo* potpun sustav događaja.

Propozicija 2.1.2. *(Formula potpune vjerojatnosti) Neka je $(H_i)_{i\in I}$ potpun sustav događaja. Tada za svaki $A \in \mathcal{F}$ vrijedi*

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(H_i)\,\mathbb{P}(A \mid H_i).$$

Dokaz.

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) = \mathbb{P}\left(A \cap \left(\bigcup_{i\in I} H_i\right)\right) = \mathbb{P}\left(\bigcup_{i\in I}(A \cap H_i)\right) = \\ &= [\sigma - aditivnost] = \sum_{i\in I} \mathbb{P}(A \cap H_i) = \\ &\overset{(2.2)}{=} \sum_{i\in I} \mathbb{P}(H_i)\,\mathbb{P}(A \mid H_i).\end{aligned}$$

□

Teorem 2.1.3. *(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*

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

Dokaz. Po definiciji uvjetne vjerojatnosti slijedi

$$\begin{aligned}\mathbb{P}(H_j \,|\, A) &= \frac{\mathbb{P}(H_j \cap A)}{\mathbb{P}(A)} \overset{(2.2)}{=} \frac{\mathbb{P}(H_j)\,\mathbb{P}(A \,|\, H_j)}{\mathbb{P}(A)} \\ &\overset{2.1.2}{=} \frac{\mathbb{P}(H_j)\,\mathbb{P}(A \,|\, H_j)}{\sum_{i\in I} \mathbb{P}(H_i)\,\mathbb{P}(A \,|\, H_i)}.\end{aligned}$$

□

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

Primjer 2.1.4. *Pretpostavimo da se ujutro probudimo s povišenom tjelesnom temperaturom te odlučimo napraviti test na COVID-19. Naš test ispada negativan no mi smo pomalo skeptični pa se pitamo koja je vjerojatnost da smo ipak zaraženi, iako je test negativan. Pretpostavimo nadalje da znamo da je vjerojatnost zaraze 0.2, vjerojatnost da test općenito ispada negativan 0.9, a vjerojatnost da test ispadne negativan za zaraženu osobu 0.16.*

Po Bayesovoj formuli slijedi

$$\mathbb{P}(Z \,|\, -) = \frac{\mathbb{P}(- \,|\, Z)\,\mathbb{P}(Z)}{\mathbb{P}(-)}$$

gdje je $\mathbb{P}(Z)$ vjerojatnost zaraze, $\mathbb{P}(-)$ vjerojatnost da je test negativan, $\mathbb{P}(- \,|\, Z)$ vjerojatnost da test ispadne negativan za zaraženu osobu te $\mathbb{P}(Z \,|\, -)$ vjerojatnost da je osoba koja je dobila negativan rezultat testa ipak zaražena.

Uvrštavanjem u formulu dobivamo da je tražena vjerojatnost

$$\mathbb{P}(Z \,|\, -) = \frac{0.16 \cdot 0.2}{0.9} = 0.035.$$

Više primjera može se pronaći u [9].

2.2 Neprekidna verzija Bayesovog teorema

Prije nego što Bayesov teorem proširimo na neprekidne slučajne varijable uvodimo dodatne pojmove poput slučajnog uzorka i funkcije vjerodostojnosti koji će nam kasnije biti ključni za primjenu teorema. Definicije su preuzete iz [2].

Definicija 2.2.1. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor te X slučajna varijabla s funkcijom distribucije F. Slučajne varijable $X_1, X_2, ..., X_n$ čine* slučajan uzorak *duljine n iz distribucije F ako su $X_1, X_2, ..., X_n$ nezavisne, jednako distribuirane slučajne varijable s funkcijom distribucije F.*

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz razdiobe F te neka je f pripadna funkcija gustoće. S $f(x \,|\, \theta) = f(x_1, x_2, ..., x_n \,|\, \theta)$ označavamo zajedničku funkciju gustoće slučajnog uzorka $X_1, X_2, ..., X_n$ s parametrom $\theta \in \Theta$, gdje je Θ skup parametara.

Definicija 2.2.2. *Neka je $x = (x_1, x_2, .., x_n)$ opaženi uzorak za varijablu X s gustoćom $f(x \,|\, \theta)$ gdje je $\theta \in \Theta$ nepoznati parametar distribucije. Tada je* funkcija vjerodostojnosti *parametra θ funkcija $L : \Theta \to \mathbb{R}$ definirana s*

$$L(\theta) := f(x_1 \,|\, \theta) \cdots f(x_n \,|\, \theta), \quad \theta \in \Theta.$$

Primjer 2.2.3. *Odredimo funkciju vjerodostojnosti normalne slučajne varijable s parametrima μ i σ^2.*

Neka je $x_1, x_2, .., x_n$ opaženi uzorak iz $N(\mu, \sigma^2)$. Tada je

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x_i \,|\, \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(\sqrt{\sigma^2}\sqrt{2\pi})^n} e^{-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}.$$

Pošto se Bayesov teorem temelji na formuli potpune vjerojatnosti, u nastavku uvodimo njenu neprekidnu inačicu te neprekidnu verziju Bayesovog teorema. Teoremi, zajedno s primjerima, mogu se pronaći u [6].

Propozicija 2.2.4. *(Neprekidna verzija formule potpune vjerojatnosti) Neka su X i Y slučajne varijable na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ s funkcijama gustoće $f_X(x)$ i $f_Y(y)$ te neka je A proizvoljan Borelov skup. Vrijedi*

$$\mathbb{P}(X \in A) = \int_{\mathbb{R}} \mathbb{P}(X \in A \,|\, Y = y)\, f_Y(y)\, dy.$$

Dokaz.

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in A, Y \in \mathbb{R}) = \int_A \int_{\mathbb{R}} f_{XY}(x, y)\, dy\, dx = \int_A \int_{\mathbb{R}} f_{X|Y}(x|y)\, f_Y(y)\, dy\, dx =$$

$$= \int_{\mathbb{R}} \Big(\int_A f_{X|Y}(x|y)\, dx \Big) f_Y(y)\, dy = \int_{\mathbb{R}} \mathbb{P}(X \in A \,|\, Y = y)\, f_Y(y)\, dy.$$

□

Teorem 2.2.5. *(Neprekidna verzija Bayesovog teorema) Neka su X i Y neprekidne slučajne varijable td.* $f_Y(y) > 0$. *Tada vrijedi*

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

Dokaz. Po definiciji uvjetne funkcije gustoće slijedi

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

Sada slijedi

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

Druga jednakost slijedi iz propozicije 2.2.4. □

U skladu s terminologijom za diskretnu verziju teorema $f_X(x)$ nazivamo *apriorna funkcija gustoće* hipoteze, a $f_{X|Y}(x\,|\,y)$ *aposteriorna funkcija gustoće* hipoteze za dane podatke.

S terminologijom vezanom uz član $f_{Y|X}(y\,|\,x)$ pozabavit ćemo se u nastavku poglavlja.

2.3 Bayesovsko zaključivanje

U prethodna dva potpoglavlja vidjeli smo dvije verzije Bayesovog teorema koje se zajednički mogu napisati kao

$$\mathbb{P}\,(hipoteza\,|\,podaci) = \frac{\mathbb{P}\,(podaci\,|\,hipoteza)\,\mathbb{P}(hipoteza)}{\mathbb{P}(podaci)}.$$

Ako su podaci fiksni, nazivnik $\mathbb{P}(podaci)$ je konstantan te služi kao normalizacijska konstanta kako bi suma aposteriornih vjerojatnosti bila 1. Stoga Bayesov teorem možemo izraziti u obliku proporcionalnosti

$$\mathbb{P}\,(hipoteza\,|\,podaci) \;\propto\; \mathbb{P}\,(podaci\,|\,hipoteza)\,\mathbb{P}(hipoteza).$$

Jedna od glavnih primjena Bayesovog teorema jest opisivanje nepoznatog parametra na temelju podataka iz neke poznate distribucije.

Neka je $x = (x_1, x_2, ..., x_n)$ realizacija slučajnog uzorka $X = (X_1, X_2, ..., X_n)$ s funkcijom gustoće $f(x\,|\,\theta)$ gdje je $\theta \in \Theta$ nepoznat parametar. Uz dane podatke x, $f(x\,|\,\theta)$ možemo gledati kao funkciju parametra θ umjesto kao funkciju uzorka x. U tom slučaju $f(x\,|\,\theta)$

zovemo vjerodostojnost parametra θ za dani uzorak x i pišemo $L(\theta\,|\,x)$. Bayesova formula je sada oblika

$$f(\theta\,|\,x) = \frac{L(\theta\,|\,x)\,f(\theta)}{f(x)}.$$

Dakle, možemo pisati

$$\text{aposteriorna gustoća} \propto \text{vjerodostojnost} \times \text{apriorna gustoća}$$

tj.

$$f(\theta\,|\,x) \propto L(\theta\,|\,x)\,f(\theta).$$

Funkcija vjerodostojnosti $L(\theta\,|\,x)$ je funkcija kroz koju dani podaci izmjenjuju naše apriorne pretpostavke o parametru θ, stoga ona predstavlja informacije o parametru θ koje dobivamo iz uzorka.

Srž Bayesovskog zaključivanja je u tumačenju gornje formule. Član $f(\theta)$ predstavlja stupanj znanja (ili nezanja) o parametru θ prije skupljanja podataka, tj. naša apriori saznanja o nepoznatom parametru. Uz danu apriornu gustoću, vjerodostojnost te podatke, moguće je izračunati gustoću $f(\theta\,|\,x)$ parametra θ uz dani uzorak x. Na taj način ažuriramo svoja saznanja o nepoznatom parametru na temelju danog uzorka.

Kako je opisano u [5], prednost Bayesovog teorema je što nam omogućava da ažuriramo informacije o parametru θ kako naš uzorak raste.
Pretpostavimo da imamo početni uzorak x_1 iz distribucije X. Po Bayesovoj formuli slijedi

$$f(\theta\,|\,x_1) \propto L(\theta\,|\,x_1)\,f(\theta). \tag{2.3}$$

Pretpostavimo sada da imamo drugi uzorak x_2 iz iste distribucije X, nezavisan od prvog uzorka x_1. Tada

$$f(\theta\,|\,x_1, x_2) \;\propto\; L(\theta\,|\,x_1)\,L(\theta\,|\,x_2)\,f(\theta) \;\propto\; L(\theta\,|\,x_2)\,f(\theta\,|\,x_1). \tag{2.4}$$

Uočimo kako je izraz (2.4) istog oblika kao izraz (2.3) no u izrazu (2.3) $f(\theta\,|\,x_1)$ predstavlja aposteriornu gustoću dok u izrazu (2.4) ista funkcija predstavlja apriornu gustoću.

Ovaj proces možemo ponoviti proizvoljan broj puta. Općenito, ako imamo n nezavisnih slučajnih uzroraka $x_1, x_2, ..., x_n$ iz distribucije X, aposteriornu gustoću možemo računati dodavajući svaki uzorak posebno tako da u m-tom koraku vjerodostojnost m-tog uzorka množimo s aposteriornom gustoćom dobivenom u (m-1)-om koraku i na taj način dobivamo novu aposteriornu gustoću

$$f(\theta\,|\,x_1, .., x_m) \;\propto\; L(\theta\,|\,x_m)\,f(\theta\,|\,x_1, ..., x_{m-1}), \quad m = 2, ..., n$$

gdje je

$$f(\theta\,|\,x_1) \propto L(\theta\,|\,x_1)\,f(\theta).$$

Poglavlje 3

Bayesovsko zaključivanje za očekivanje normalne distribucije

Postoji razlog zašto se normalna distribucija naziva normalnom. Možemo ju pronaći u podatkovnim znanostima, strojnom učenju te mnogim različitim prirodnim pojavama. Od distribucije visine i težine, do rezultata ispita i IQ-a, čini se da je normalna distribucija prisutna svugdje. Razlog tomu leži u centralnom graničnom teoremu. Slučajna varijabla koja se može zapisati kao suma velikog broja nezavisnih, jednako distribuiranih slučajnih varijabli biti će aproksimativno normalna. Stoga takvu slučajnu varijablu, tj. slučajni uzorak iz te distribucije možemo aproksimirati normalnom distribucijom.

Najšire korištene statističke metode su upravo one koje koriste pretpostavku o normalnosti danih podataka. Tu pretpostavku koristimo u ovom poglavlju te pokazujemo kako se koristi Bayesov teorem, tj. Bayesovsko zaključivanje za slučajan uzorak iz normalne distribucije. Kao izvor glavnih ideja ovog poglavlja poslužio je [1].

3.1 Diskretni Bayesov teorem za očekivanje normalne distribucije

U ovom potpoglavlju pokazujemo kako odrediti nepoznati parametar očekivanja koristeći diskretnu verziju Bayesovog teorema. Prvo promatramo slučaj kada je dana jedna opažena vrijednost, a zatim dva načina za određivanje nepoznatog parametra u slučaju kada je dan slučajan uzorak opaženih vrijednosti.

Jedna opažena vrijednost

Neka je dana jedna opažena vrijednost x za slučajnu varijablu X s uvjetnom gustoćom $f(x \mid \theta)$ za koju znamo da je normalna s poznatom varijancom σ^2. Pretpostavimo da imamo

m mogućih vrijednosti $\mu_1, \mu_2, ..., \mu_m$ za parametar očekivanja. Za te vrijednosti biramo diskretnu razdiobu koja opisuje naše apriorno vjerovanje o nepoznatom parametru μ. Ako nemamo nikakve informacije o nepoznatom parametru μ, svim vrijednostima dajemo jednaku vjerojatnost. Kod Bayesovskog zaključivanja fiksiramo opažene vrijednosti, a variramo parametar μ po svim vrijednostima $\mu_1, \mu_2, ..., \mu_m$ koje smatramo mogućima. Po Bayesovom teoremu, aposteriorna vjerojatnost je oblika

$$\mathbb{P}(\mu \mid x) = \frac{vjerodostojnost \times apriorna\ vjerojatnost}{\sum vjerodostojnost \times apriorna\ vjerojatnost}$$

pošto će se bilo koja multiplikativna konstanta u apriornoj vjerojatnosti ili vjerodostojnosti pokratiti.

Vjerodostojnost parametra μ uz jednu danu opaženu vrijednost x jednaka je funkciji gustoće kojoj pripada ta opažena vrijednost. Pošto pretpostavljamo da je opažena vrijednost iz normalne distribucije s nepoznatim parametrom očekivanja μ i poznatim parametrom varijance σ^2, funkcija vjerodostojnosti je dana s

$$L(\mu \mid x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Dio funkcije vjerodostojnosti koji ne ovisi o parametru μ tretiramo kao konstantu, tj. pišemo

$$L(\mu \mid x) \propto e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{3.1}$$

Postoje dvije metode za računanje funkcije vjerodostojnosti. U prvoj metodi pronalazimo vjerodostojnost pomoću tablice distribucije standardne normalne distribucije. Definiramo

$$Z = \frac{X - \mu}{\sigma}$$

gdje je X slučajna varijabla koju promatramo, σ parametar standardne devijacije distribucije kojoj pripada X, a μ parametar očekivanja koji variramo.
Tada Z ima standardnu normalnu distribuciju $Z \sim N(0, 1)$ stoga vjerodostojnost gledamo iz tablice distribucije standardne normalne distribucije dane u dodatku.

Napomena 3.1.1. *Funkcija gustoće standardne normalne distribucije f simetrična je oko 0 stoga vrijedi $f(x) = f(-x)$.*

Druga metoda jest korištenje izraza (3.1). Opaženi uzorak fiksiramo, a nepoznati parametar variramo po svim mogućim vrijednostima.

Primjenimo obje metode na idućem primjeru.

Primjer 3.1.2. [1] *Pretpostavimo da je X normalna slučajna varijabla s nepoznatim očekivanjem μ i poznatom varijancom $\sigma^2 = 225$. Nadalje, pretpostavimo da imamo 5 mogućih vrijednosti za parametar μ : $180, 190, 200, 210, 220$ te da opažena vrijednost iznosi $x = 193$.*

Pošto nemamo dodatne informacije o nepoznatom parametru, pretpostavljamo da su sve vrijednosti jednako vjerojatne. Stoga je apriorna vjerojatnost za svaku vrijednost μ jednaka i iznosi 0.2.
Definiramo

$$z = \frac{x - \mu}{\sigma}$$

za svaku moguću vrijednost μ. Vrijednosti funkcije vjerodostojnosti $f(z)$ čitamo iz tablice distribucije standardne normalne distribucije dane u dodatku. Dobivene vrijednosti dane su u sljedećoj tablici.

μ	apriori	z	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2	0.87	0.2732	0.05464	0.2084
190	0.2	0.2	0.3910	0.0782	0.2982
200	0.2	-0.47	0.3572	0.07144	0.2724
210	0.2	-1.13	0.2107	0.04214	0.1607
220	0.2	-1.8	0.079	0.0158	0.0603
				0.26222	1

Ako vjerodostojnost računamo koristeći izraz (3.1)*, dobivamo iduću tablicu.*

μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2	0.6869	0.1374	0.2089
190	0.2	0.9802	0.196	0.2981
200	0.2	0.8968	0.1794	0.2728
210	0.2	0.5261	0.1052	0.16
220	0.2	0.1979	0.0396	0.0602
			0.6576	1

Vidimo da dobivamo vrlo slične vrijednosti, eventualne razlike nastale su zbog zaokruživanja.

Slučajan uzorak opaženih vrijednosti

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2 gdje je σ^2 poznat. Tada je funkcija vjerodostojnosti nepoznatog parametra μ dana s

$$L(\mu \,|\, x_1, x_2, ..., x_n) = f(x_1 \,|\, \mu) \cdots f(x_n \,|\, \mu).$$

[1] Više sličnih primjera može se pronaći u [1].

Sada po Bayesovom teoremu slijedi

$$\mathbb{P}(\mu \,|\, x_1, ..., x_n) = \frac{\mathbb{P}(\mu) \cdot f(x_1 \,|\, \mu) \cdots f(x_n \,|\, \mu)}{\sum_{\mu} \mathbb{P}(\mu) \cdot f(x_1 \,|\, \mu) \cdots f(x_n \,|\, \mu)}.$$

Aposteriorne vjerojatnosti možemo tražiti na dva načina. Prvi način sastoji se od analiziranja svake uočene vrijednost posebno na način aposteriorna vjerojatnost jedne vrijednosti postane apriorna vjerojatnost sljedeće vrijednosti, kako je opisano na kraju prethodnog poglavlja.

Primjer 3.1.3. *Pretpostavimo da je X normalna slučajna varijabla s nepoznatim očekivanjem* μ *i poznatom varijancom* $\sigma^2 = 225$ *te da imamo 5 mogućih vrijednosti za parametar* $\mu : 180, 190, 200, 210, 220.$ *Pretpostavimo da je sada dan slučajni uzorak opaženih vrijednosti* $193, 190, 215, 198, 184.$

Dobivamo sljedeću tablicu (vjerodostojnost računamo koristeći izraz (3.1)*).*

		x=193		
μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2	0.6869	0.1374	0.2089
190	0.2	0.9802	0.196	0.2981
200	0.2	0.8968	0.1794	0.2728
210	0.2	0.5261	0.1052	0.16
220	0.2	0.1979	0.0396	0.0602
			0.6576	1
		x=190		
μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2089	0.8	0.1617	0.2206
190	0.2981	1	0.2981	0.3936
200	0.2728	0.8	0.2182	0.288
210	0.16	0.41	0.0656	0.0866
220	0.0602	0.14	0.0084	0.011
			0.7574	1
		x=215		
μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2206	0.0657	0.0145	0.0382
190	0.3936	0.2494	0.0981	0.2587
200	0.288	0.6065	0.1747	0.4607
210	0.0866	0.946	0.0819	0.216
220	0.011	0.946	0.01	0.0264
			0.3792	1

x=198				
μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.0382	0.4868	0.0186	0.0216
190	0.2587	0.8674	0.2244	0.2602
200	0.4607	0.9912	0.4566	0.5295
210	0.216	0.7121	0.1538	0.1783
220	0.0264	0.3411	0.009	0.0104
			0.8624	1
x=184				
μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.0216	0.965	0.0208	0.0346
190	0.2602	0.9231	0.2402	0.3996
200	0.5295	0.5662	0.2998	0.4988
210	0.1783	0.2226	0.0397	0.067
220	0.0104	0.0561	0.0006	0.0009
			0.6011	1

Drugi način je da analiziramo sve opažene vrijednosti odjednom. Pošto je uzorak iz normalne distribucije s poznatim parametrom varijance σ^2, član $\frac{1}{\sqrt{2\pi}\sigma}$ je konstanta pa ga možemo zanemariti jer će se pokratiti u izrazu iz Bayesovog teorema. Funkcija vjerodostojnosti nepoznatog parametra μ je stoga oblika

$$\begin{aligned} L(\mu \,|\, x_1, x_2, ..., x_n) &\propto e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdots e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} \\ &\propto e^{-\frac{1}{2\sigma^2}\left[(x_1-\mu)^2+(x_2-\mu)^2+...+(x_n-\mu)^2\right]}. \end{aligned}$$

Raspišimo izraz u zagradi:

$$\begin{aligned} \left[(x_1-\mu)^2 + (x_2-\mu)^2 + ... + (x_n-\mu)^2\right] &= {x_1}^2 - 2x_1\mu + \mu^2 + ... + {x_n}^2 - 2x_n\mu + \mu^2 = \\ &= ({x_1}^2 + ... + {x_n}^2) - 2\mu\,(x_1 + ... + x_n) + n\mu^2 \\ &= ({x_1}^2 + ... + {x_n}^2) - 2n\mu\,\bar{x} + n\mu^2. \end{aligned}$$

Sada slijedi

$$\begin{aligned} L(\mu \,|\, x_1, x_2, ..., x_n) &\propto e^{-\frac{n}{2\sigma^2}\left[\frac{{x_1}^2+...+{x_n}^2}{n} - 2\mu\,\bar{x} + \mu^2\right]} \\ &\propto e^{-\frac{n}{2\sigma^2}\left[\frac{{x_1}^2+...+{x_n}^2}{n} - 2\mu\,\bar{x} + \mu^2 + \bar{x}^2 - \bar{x}^2\right]} \\ &\propto e^{-\frac{n}{2\sigma^2}\left[\mu^2 - 2\mu\,\bar{x} + \bar{x}^2\right]} \cdot e^{-\frac{n}{2\sigma^2}\left[\frac{{x_1}^2+...+{x_n}^2}{n} - \bar{x}^2\right]}. \end{aligned}$$

Drugi dio u posljednjem izrazu ne ovisi o nepoznatom parametru te ga možemo smatrati proporcionalnom konstantom pa pišemo

$$L(\mu \,|\, x_1, x_2, ..., x_n) \propto e^{-\frac{n(\bar{x}-\mu)^2}{2\sigma^2}} \propto e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}}.$$

Za aritmetičku sredinu $\bar{X}$ slučajnog uzorka $X_1, X_2, ..., X_n$ iz normalne distribucije s parametrima μ i σ^2 znamo da je normalno distribuirana s parametrima μ i $\frac{\sigma^2}{n}$. Dakle, funkcija vjerodostojnosti parametra μ uz dani slučajan uzorak $X_1, X_2, ..., X_n$ proporcionalan je vjerodostojnosti parametra μ uz dano uzoračko očekivanje $\bar{X}$.

Primjer 3.1.4. *Riješimo ponovno prethodni primjer, sada koristeći vjerodostojnost aritmetičke sredine uzorka.*

Koristimo

$$L(\mu \,|\, x_1, x_2, ..., x_n) \;\propto\; e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}}$$

gdje je

$$\bar{x} = \frac{193 + 190 + 215 + 198 + 184}{5} = 196 \quad i \quad \frac{\sigma^2}{n} = \frac{225}{5} = 45.$$

Ponovno nemamo dodatne informacije o nepoznatom parametru pa pretpostavljamo da su sve vrijednosti jednako vjerojatne. Dakle, apriorna vjerojatnost je ponovno jednaka za svaku vrijednost μ i iznosi 0.2.

μ	apriori	vjerodostojnost	vjerodostojnost × apriori	aposteriori
180	0.2	0.0582	0.0116	0.0345
190	0.2	0.6703	0.1341	0.399
200	0.2	0.8371	0.1674	0.4981
210	0.2	0.1133	0.0227	0.0675
220	0.2	0.0017	0.0003	0.0009
			0.3361	1

Dobivamo vrlo slične vrijednosti kao u prethodnom primjeru, no postupak je puno kraći.

3.2 Neprekidni Bayesov teorem za očekivanje normalne distribucije

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2, gdje je parametar σ^2 poznat. Pretpostavljamo da nepoznati parametar μ može biti bilo koja vrijednost unutar nekog intervala, umjesto da se nalazi u diskretnom skupu vrijednosti, stoga koristimo neprekidnu apriori distribuciju. Ponovno slijedi

$$f(\mu \,|\, x_1, ..., x_n) \propto f(\mu) \cdot L(\mu \,|\, x_1, x_2, ..., x_n) \propto f(\mu) \cdot f(x_1 \,|\, \mu) \cdots f(x_n \,|\, \mu)$$

no sada dopuštamo da $f(\mu)$ bude neprekidna funkcija gustoće.

U prošlom odjeljku raspisali smo funkciju vjerodostojnosti nepoznatog parametra μ uz dani slučajan uzorak $X_1, X_2, ..., X_n$, stoga neprekidna verzija Bayesove formule za normalnu distribuciju poprima sljedeći oblik

$$f(\mu \,|\, x_1, ..., x_n) = \frac{f(\mu) \cdot e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}}}{\int f(\mu) \cdot e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}} \, d\mu}. \tag{3.2}$$

Ovo možemo napraviti za bilo koju neprekidnu razdiobu $f(\mu)$, no u nazivniku se pojavljuje integral koji je nekada moguće izračunati jedino numerički.

Generalno razlikujemo informativne i neinformativne apriorne distribucije, ovisno o količini informacija o nepoznatom parametru koju sadrže. Neinformativne su one koje predstavljaju najniži nivo apriornog mišljenja o nepoznatom parametru, kao na primjer uniformna distribucija. Informativne apriorne distribucije sadrže određenu količinu informacija o nepoznatom parametru. Na primjer, informativna apriorna distribucija može biti aposteriorna distribucija prijašnjeg istraživanja. Najinformativnije apriorne distribucije su one koje nekim vrijednostima ili čitavim intervalima mogućih vrijednosti nepoznatog parametra dodjeljuju vjerojatnost 0.

Kod odabira apirorne distribucije važno je voditi računa o sljedećem:

- apriorna distribucija treba odražavati što više informacija o nepoznatom parametru
- apriorna distribucija treba biti što jednostavnijeg oblika za računanje.

Drugo navedeno svojstvo je posebno važno upravo radi integrala koji se pojavljuje u izrazu (3.2). Pogledajmo nekoliko posebnih slučajeva apriornih distribucija u kojima nije potrebno integrirati.

Jeffreyjeva apriorna distribucija

Ranije smo komentirali kako se multiplikativne konstante koje pripadaju apriornoj distribuciji u izrazu za aposteriornu distribuciju krate. Stoga nam nisu bitne same vrijednosti koje apriorna distribucija pridružuje svakoj mogućoj vrijednosti nepoznatog parametra već su bitne samo težine.

Jeffreyjeva apriorna distribucija svakoj mogućoj vrijednosti od μ daje jednaku težinu, $f(\mu) = 1$. Ona pripada skupini tzv. nepravilnih apriornih distribucija pošto integral funkcije gustoće nije 1. No, aposteriorna funkcija gustoće koja se dobije korištenjem Jeffreyjeve apriorne distribucije integrirana daje 1 pa ona jest funkcija gustoće.

Korištenje Jeffreyjeve apriorne distribucije sugerira kako nemamo dodatnih informacija o nepoznatom parametru, već sve moguće vrijednosti smatramo jednako vjerojatnima.

Neka je x opažena vrijednost za normalnu slučajnu varijablu X s očekivanjem μ i varijancom σ^2. Ako zanemarimo konstantu, funkcija vjerodostojnosti dana je s

$$L(\mu \,|\, x) \propto e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Pošto je apriorna gustoća uvijek jednaka 1, aposteriorna gustoća je proporcionalna funkciji vjerodostojnosti

$$f(\mu \,|\, x) \propto e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Dakle, aposteriorna gustoća jedne opažene vrijednosti je proporcionalna funkciji gustoće normalne slučajne varijable s parametrima μ i σ^2.

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2. U prethodnom odjeljku pokazali smo da je funkcija vjerodostojnosti slučajnog uzorka proporcionalna funkciji vjerodostojnosti aritmetičke sredine danog uzorka. Također znamo da je aritmetička sredina normalno distribuirana s parametrima μ i $\frac{\sigma^2}{n}$ stoga je vjerodostojnost opisana s

$$L(\mu \,|\, x_1, ..., x_n) \propto e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}}.$$

Ponovno, apriorna gustoća je uvijek jednaka 1 pa je stoga aposteriorna gustoća proporiconalna funkciji vjerodostojnosti

$$f(\mu \,|\, x_1, ..., x_n) \propto e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2/n}}.$$

Dakle, aposteriorna gustoća slučajnog uzorka je proporcionalna funkciji gustoće normalne slučajne varijable s parametrima μ i $\frac{\sigma^2}{n}$.

Normalna apriorna distribucija

Neka je dana jedna opažena vrijednost x za normalnu slučajnu varijablu X s očekivanjem μ i varijancom σ^2 te apriorna distribucija koja je normalna s parametrima m i s^2. Vrijedi

$$\begin{aligned}
f(\mu) \cdot L(\mu \,|x) &\propto e^{-\frac{(\mu-m)^2}{2s^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \propto e^{-\frac{1}{2}\left[\frac{(\mu-m)^2}{s^2}+\frac{(x-\mu)^2}{\sigma^2}\right]} \\
&\propto e^{-\frac{1}{2}\left[\frac{\sigma^2\mu^2-2\mu m\sigma^2+m^2\sigma^2+x^2s^2-2x\mu s^2+\mu^2 s^2}{s^2\sigma^2}\right]} \\
&\propto e^{-\frac{\sigma^2+s^2}{2\sigma^2 s^2}\left[\mu^2-2\frac{\sigma^2 m+s^2 x}{\sigma^2+s^2}\mu+\frac{m^2\sigma^2+x^2s^2}{\sigma^2+s^2}\right]} \\
&\propto e^{-\frac{\sigma^2+s^2}{2\sigma^2 s^2}\left[\mu^2-2\frac{\sigma^2 m+s^2 x}{\sigma^2+s^2}\mu+\left(\frac{\sigma^2 m+s^2 x}{\sigma^2+s^2}\right)^2\right]} \\
&\propto e^{-\frac{\sigma^2+s^2}{2\sigma^2 s^2}\left[\mu-\left(\frac{\sigma^2 m+s^2 x}{\sigma^2+s^2}\right)^2\right]}.
\end{aligned}$$

Dakle, aposteriorna distribucija je normalna s očekivanjem $m' = \frac{\sigma^2 m + s^2 x}{\sigma^2 + s^2}$ i varijancom $(s')^2 = \frac{\sigma^2 s^2}{\sigma^2 + s^2}$. Iz ovog zaključujemo da, počevši od normalne apriorne distribucije, dobivamo normalnu aposteriornu distribuciju. Radi toga nije potrebno integrirati da bismo dobili aposteriornu distribuciju već je dovoljno odrediti pravilo po kojem se mijenjaju parametri očekivanja i varijance.

Promotrimo recipročnu vrijednost varijance, još zvanu *preciznost* distribucije.

$$\frac{1}{(s')^2} = \frac{\sigma^2 + s^2}{\sigma^2 s^2} = \frac{1}{s^2} + \frac{1}{\sigma^2}.$$

Preciznost aposteriorne distribucije jednaka je zbroju preciznosti aprorne distribucije i preciznosti distribucije kojoj pripada opažena vrijednost. Očekivanje aposteriorne distribucije se sada može zapisati kao

$$m' = \frac{\sigma^2 m + s^2 x}{\sigma^2 + s^2} = \frac{\sigma^2}{\sigma^2 + s^2} \cdot m + \frac{s^2}{\sigma^2 + s^2} \cdot x = \frac{1/s^2}{1/\sigma^2 + 1/s^2} \cdot m + \frac{1/\sigma^2}{1/\sigma^2 + 1/s^2} \cdot x.$$

Dakle, očekivanje aposteriorne distribucije se može zapisati kao težinska sredina očekivanja apriorne distribucije i očekivanja distribucije kojoj pripada opažena vrijednost, gdje su težine zadane pomoću preciznosti.

Ovo pravilo funkcionira i za Jeffreyjevu apriornu distribuciju. Njena varijanca je ∞, stoga je preciznost jednaka 0. Aposteriorna preciznost je tada jednaka preciznosti distribucije kojoj pripada opažena vrijednost, tj. preciznosti funkcije vjerodostojnosti nepoznatog parametra

$$\frac{1}{(s')^2} = 0 + \frac{1}{\sigma^2} = \frac{1}{\sigma^2}$$

pa je varijanca aposterione distribucije jednaka varijanci funkcije vjerodostojnosti nepoznatog parametra.

Jeffreyjeva apriorna distribucija nema dobro definirano očekivanje, tj. ono može biti bilo što. No,

$$m' = \frac{0}{1/\sigma^2} \cdot \text{bilo što} + \frac{1/\sigma^2}{1/\sigma^2} \cdot x = x$$

pa je aposteriorno očekivanje jednako uočenoj vrijednosti x.

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2 te apriorna distribucija koja je normalna s parametrima m i s^2. Koristimo funkciju vjerodostojnosti aritmetičke sredine uzorka $\bar{X}$ koja je normalno distribuirana s parametrima μ i $\frac{\sigma^2}{n}$. Preciznost vjerodostojnosti aritmetičke sredine je $\frac{n}{\sigma^2}$.
Problem se svodi na određivanje aposteriorne distribucije za jednu vrijednost, aritmetičku

sredinu uzorka $\bar{X}$. Preciznost aposteriorne distribucije jednaka je zbroju preciznosti apriorne distribucije i preciznosti distribucije od $\bar{X}$

$$\frac{1}{(s')^2} = \frac{1}{s^2} + \frac{n}{\sigma^2} = \frac{\sigma^2 + ns^2}{\sigma^2 s^2}.$$

Sada je varijanca aposteriorne distribucije recipročna vrijednost preciznosti.
Očekivanje aposteriorne distribucije je težinska sredina očekivanja apriorne distribucije i očekivanja distribucije od $\bar{X}$ gdje su težine jednake pribrojnicima iz rastava preciznosti.

$$m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{x}.$$

Kao i ranije, u slučaju Jeffreyjeve apriorne distribucije aposteriorna preciznost jednaka je preciznosti funkcije vjerodostojnosti nepoznatog parametra. Ona je pak je uz dani slučajni uzorak $X_1, X_2, ..., X_n$ jednaka vjerodostojnosti aritmetičke sredine $\bar{X}$

$$\frac{1}{(s')^2} = 0 + \frac{n}{\sigma^2} = \frac{n}{\sigma^2}.$$

Očekivanje aposteriorne distribucije uz dani slučajni uzorak $x_1, x_2, ..., x_n$ jednako je $\bar{x}$

$$m' = \frac{0}{n/\sigma^2} \cdot \textit{bilo što} + \frac{n/\sigma^2}{n/\sigma^2} \cdot \bar{x} = \bar{x}.$$

Odabir normalne apriorne distribucije

Apriorna distribucija koju odaberemo trebala bi odražavati naša apriorna uvjerenja o nepoznatom parametru. Ako pretpostavimo da je apriorna distribucija normalna s parametrima m i s^2, $N(m, s^2)$, na temelju svojih uvjerenja trebamo odrediti parametre m i s^2.

Prvo odabiremo vrijednost parametra očekivanja m apriorne distribucije. Parametar m predstavlja vrijednost oko koje centriramo svoja uvjerenja.

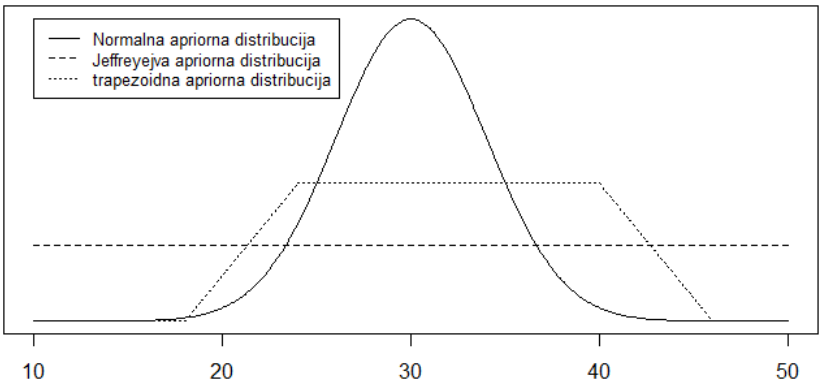
Nakon toga odabiremo vrijednost parametra standardne devijacije s apriorne distribucije. Prema [1], jedna od metoda za određivanje vrijednosti parametra s je razmotriti najveću i najmanju vrijednost koju bi mogla poprimiti slučajna varijabla koju promatramo, a zatim interval od odabrane najmanje do odabrane najveće vrijednosti podijeliti sa 6.

Primjer 3.2.1. [2] *Mjerenjem mase 10 istovrsnih kolutova sira (u kilogramima) dobiveni su sljedeći rezultati:*

38.7	*40.4*	*37.2*	*36.6*	*35.9*	*34.7*	*37.6*	*35.1*	*37.5*	*35.6*

Pretpostavimo da je masa koluta sira normalno distribuirana s poznatom standardnom devijacijom $\sigma = 3$*. Odredimo očekivanu težinu koluta sira koristeći*
(a) normalnu $N(30, 10^2)$ *apriornu distribuciju.*
(b) Jeffreyjevu apriornu distribuciju.
(c) apriornu distribuciju trapezoidnog oblika koja je jednaka nuli do 18kg, jedinici od 24 do 40kg i nuli od 46kg nadalje, a između zadanih vrijednosti je linearna.

Prikažimo najprije oblik sve tri apriorne distribucije na grafu.



Slika 3.1: Oblik apriornih distribucija

Napomena 3.2.2. *Za potrebe grafičkog prikaza Jeffreyjeva i trapezoidna apriorna distribucija su reskalirane t.d. su njihovi integrali na intervalu* $[10, 50]$ *jednaki 1. Račun je proveden za funkcije kako su zadane.*

U slučaju normalne i Jeffreyjeve apriorne distribucije, aposteriorne distribucije možemo izračunati pomoću ranije navedenog ažurirajućeg pravila. Za normalnu apriornu distri-

[2] Slični primjeri mogu se pronaći u [1].

buciju vrijedi

$$\frac{1}{(s')^2} = \frac{1}{s^2} + \frac{n}{\sigma^2} = \frac{1}{10^2} + \frac{10}{3^2}.$$

Odavdje slijedi $s' = 0.94$.

Nadalje

$$m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{x}$$

gdje $\bar{x} = 36.93$ *dobivamo iz danog uzorka.*

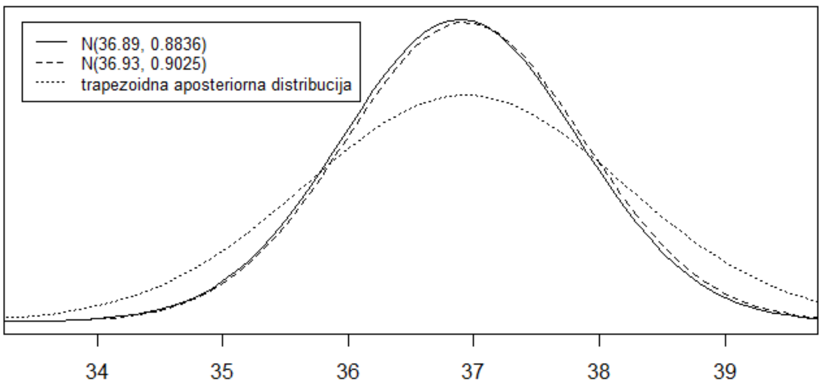
Slijedi $m' = 32.6$ *stoga je aposteriorna distribucija u prvom slučaju normalna s očekivanjem* $m' = 36.89$ *i varijancom* $\sigma^2 = 0.94^2 = 0.8836$*,* $N(36.89, 0.8836)$.

Za Jeffreyjevu apriornu distribuciju vrijedi

$$\frac{1}{(s')^2} = \frac{n}{\sigma^2} \quad i \quad m' = \bar{x}$$

iz čega slijedi $s' = 0.95$. *i* $m' = 36.93$. *Dakle, aposteriorna distribucija u ovom slučaju je normalna s očekivanjem* $m' = 36.93$ *i varijancom* $\sigma^2 = 0.95^2 = 0.9025$*,* $N(36.93, 0.9025)$.

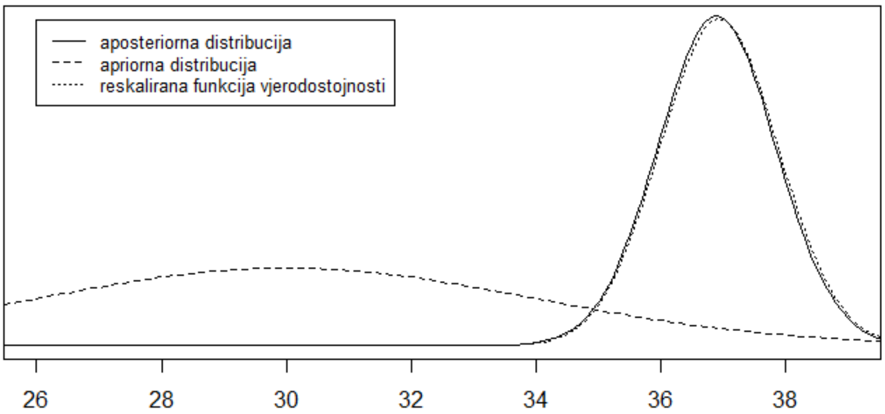
U slučaju trapezoidne apriorne distribucije, aposteriornu distribuciju računamo numerički, koristeći izraz (3.2)*. Prikažimo dobivene aposteriorne distribucije na grafu.*



Slika 3.2: Aposteriorne distribucije

Dobivamo relativno slične aposteriorne distribucije iako smo krenuli od jako različitih apriornih distribucija. To se događa upravo zato što za dobivanje aposteriorne distribucije koristimo (iste) dane podatke koji onda mijenjaju distribuciju, tj. naše zaključivanje o distribuciji očekivanja kroz funkciju vjerodostojnosti.

Odnos apriorne i aposteriorne distribucije te (reskalirane) funkcije vjerodostojnosti prikazan je na idućem grafu. Radi bolje vidljivosti, umjesto funkcije vjerodostojnosti parametra μ *uz dani slučajni uzorak* $x_1, ..., x_n$ *nacrtana je funkcija vjerodostojnosti parametra* μ *uz dano uzoračko očekivanje* $\bar{x}$ *pošto su one proporcionalne.*



Slika 3.3: Odnos apriorne i aposteriorne distribucije te (reskalirane) funkcije vjerodostojnosti

S grafa možemo vidjeti koliki utjecaj ima dani uzorak, kroz funkciju vjerodostojnosti, na zaključivanje o distribuciji nepoznatog parametra. Krenuli smo od normalne distribucije s očekivanjem 30 i standardnom devijacijom 10 i došli do normalne distribucije s očekivanjem 36.89 i standardnom devijacijom 0.8836.
Mogli bismo reći da nove informacije koje dobivamo iz uzorka "pomiću" distribuciju nepoznatog parametra prema očekivanju funkcije vjerodostojnosti nepoznatog parametra.

3.3 Bayesov interval vjerodostojnosti za očekivanje

Aposteriorna distribucija $f(\mu|x_1, x_2, ..., x_n)$ predstavlja naše zaključivanje o parametru μ na temelju slučajnog uzorka $X_1, X_2, ..., X_n$. Ona obuhvaća sve informacije i vjerovanja koje imamo o nepoznatom parametru μ uz dani slučajni uzorak. Pomoću aposteriorne distribucije možemo izračunati interval mogućih vrijednosti za nepoznati parametar koji smatramo da ne možemo odbaciti na određenoj razini značajnosti.

Bayesovski interval vjerodostojnosti je analogon frekvencionističkog intervala pouzdanosti, glavna razlika je u interpretaciji. Kod frekvencionističkog pristupa, 95% pouzdani interval za neki parametar θ označava interval koji u 95% slučajeva sadrži pravu vrijednost parametra θ, dok kod Bayesovskog pristupa 95%-tni interval vjerodostojnosti parametra θ označava interval za koji je uvjetna vjerojatnost da interval sadrži stvarnu vrijednost parametra θ jednaka 95%, uz dani slučajni uzorak.

Definicija 3.3.1. *Neka je $X_1, X_2, ...X_n$ slučajan uzorak iz distribucije s nepoznatim parametrom θ. Neka je nadalje $\alpha \in [0, 1]$ te a i b realni brojevi t.d.*

$$\int_{-\infty}^{a} f(\theta \,|\, x_1, x_2, ...x_n)\, d\theta = \int_{b}^{\infty} f(\theta \,|\, x_1, x_2, ...x_n)\, d\theta = \frac{\alpha}{2}.$$

Tada vrijedi

$$\mathbb{P}\left(\theta \in [a, b] \mid x_1, x_2...x_n\right) = \int_{a}^{b} f(\theta \,|\, x_1, x_2, ...x_n)\, d\theta = 1 - \alpha.$$

Interval $[a, b]$ nazivamo $(1 - \alpha) \cdot 100\%$ Bayesov interval vjerodostojnosti *za nepoznati parametar θ.*

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2, gdje je varijanca σ^2 poznata. Tada je uzoračko očekivanje $\bar{X}$ normalno distribuirano s parametrima μ i $\frac{\sigma^2}{n}$. Koristeći Jeffreyjevu ili normalnu apriornu distribuciju dobit ćemo normalnu aposteriornu distribuciju $N(m', (s')^2)$ gdje parametre m' i $(s')^2$ određujemo prema ranije navedenom ažurirajućem pravilu. Stoga slijedi

$$-z_{\frac{\alpha}{2}} \leq \frac{\mu - m'}{s'} \leq z_{\frac{\alpha}{2}}$$

gdje je $z_{\frac{\alpha}{2}}$ $\frac{\alpha}{2}$-kvantil standardne normalne distribucije. Raspišemo li gornji izraz dobivamo

$$m' - z_{\frac{\alpha}{2}} \cdot s' \leq \mu \leq m' + z_{\frac{\alpha}{2}} \cdot s'$$

stoga je $(1 - \alpha) \cdot 100\%$ Bayesov interval vjerodostojnosti za μ dan s

$$\left[m' - z_{\frac{\alpha}{2}} \cdot s', m' + z_{\frac{\alpha}{2}} \cdot s'\right].$$

Aposteriorna vjerojatnost da stvarna vrijednost parametra μ leži izvan dobivenog intervala pouzdanosti jest α.

Ako ne znamo varijancu σ^2, ne znamo odrediti preciznost pa ne možemo primjeniti ažurirajuće pravilo. U tom slučaju računamo uzoračku varijancu na temelju danog uzorka

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Zatim primjenimo pravilo za pronalazak $(s')^2$ i m' gdje umjesto nepoznate varijance σ^2 koristimo uzoračku varijancu $\hat{\sigma}^2$. Sada $\frac{\mu - m'}{\hat{\sigma}}$ ima Studentovu t-distribuciju s $n-1$ stupnjem slobode pa vrijedi

$$-t_{\frac{\alpha}{2}} \leq \frac{\mu - m'}{\hat{\sigma}} \leq t_{\frac{\alpha}{2}}$$

gdje je $t_{\frac{\alpha}{2}}$ $\frac{\alpha}{2}$-kvantil Studentove t-distribucije s $n-1$ stupnjem slobode gdje je n veličina danog uzorka. Stoga je $(1-\alpha)\cdot 100\%$ Bayesov interval vjerodostojnosti za μ dan s

$$\left[m' - t_{\frac{\alpha}{2}} \cdot s', m' + t_{\frac{\alpha}{2}} \cdot s'\right]$$

gdje $t_{\frac{\alpha}{2}}$ čitamo iz tablice Studentove t-distribucije.

U slučaju da počinjemo s apriornom distribucijom koja nije Jeffreyjeva ni normalna, aposteriornu distribuciju računamo pomoću Bayesovog teorema. Bayesov $(1-\alpha)\cdot 100\%$ interval vjerodostojnosti za μ tada određujemo prema definiciji. Tražimo donju vrijednost μ_l i gornju vrijednost μ_u td.

$$\int_{\mu_l}^{\mu_u} f(\mu \,|\, x_1, \ldots x_n)\, d\mu = 1 - \alpha. \tag{3.3}$$

Postoji mnogo takvih vrijednosti μ_l i μ_u, no najbolje je odabrati one koje čine najmanji interval $[\,\mu_l, \mu_u]$.

Primjer 3.3.2. *Izračunajmo 95% interval vjerodostojnosti za parametar očekivanja za težinu koluta sira iz prethodnog primjera.*

U slučaju normalne aposteriorne distribucije, interval je oblika

$$\left[m' - z_{\frac{\alpha}{2}} \cdot s' \ , \ m' + z_{\frac{\alpha}{2}} \cdot s'\right].$$

Za $\alpha = 0.05$ $\;z_{\frac{\alpha}{2}} = z_{0.025}$ *određujemo iz tablice distribucije standardne normalne distribucije i on iznosi* 1.96.
U prvom slučaju, za $N(36.89, 0.8836)$*, dobivamo interval*

$$[36.89 - 1.96 \cdot 0.94 \ , \ 36.89 + 1.96 \cdot 0.94] = [35.0476, 38.7324] \, .$$

U drugom slučaju, za $N(36.93, 0.9025)$*, dobivamo interval*

$$[36.93 - 1.96 \cdot 0.95\ ,\ 36.93 + 1.96 \cdot 0.95] = [35.068, 38.792]\,.$$

U trećem slučaju smo numerički računali aposteriornu distribuciju pa bismo numerički trebali računati i interval vjerodostojnosti koristeći (3.3).

Iako smo krenuli od poprilično različitih apriornih distribucija, dobivamo slične intervale vjerodostojnosti. Kao i kod računanja aposteriorne distribucije, utjecaj podataka je mnogo jači od utjecaja apriorne distribucije stoga dobivamo slične intervale.

Poglavlje 4

Razlika Bayesovskog i frekvencionističkog pristupa

Donošenje zaključaka o očekivanju za dani slučajni uzorak iz normalne distribucije jedna je od glavnih primjena statistike. S Bayesovske strane gledišta, aposteriorna distribucija sažima sve dostupne informacije o nepoznatom parametru, uz dani slučajni uzorak. No, s frekvencionističke strane, postoji više različitih metoda za donošenje zaključaka: točkovna procjena, intervalna procjena i testiranje hipoteza. Svaka od ovih metoda može se provesti i u Bayesovskom smislu.

U ovom poglavlju razmatramo frekvencionističke metode za točkovnu procjenu, intervalnu procjenu i testiranje hipoteza te ih uspoređujemo s odgovarajućim Bayesovskim metodama. Kao izvor korišteni su [1] i [3].

4.1 Točkovna procjena

Da bismo razumjeli način na koji se frekvencionistički određuje točkovni procjenitelj potrebno je prvo savladati osnovne pojmove vezane uz statističku procjenu.

Definicija 4.1.1. *Neka je $X_1, ..., X_n$ slučajan uzorak iz razdiobe F te θ nepoznati parametar te distribucije.* Točkovni procjenitelj *T od θ je funkcija slučajnog uzorka $T = t(X_1, ..., X_n)$.*

Točkovnih procjenitelja za neki parametar ima mnogo stoga je pitanje koji procjenitelj odabrati. Jedan od često korištenih kriterija jest nepristranost.

Definicija 4.1.2. *Procjenitelj T od θ je* nepristran *procjenitelj za θ ako vrijedi*

$$\mathbb{E}(T) = \theta.$$

Procjenitelj koji nije nepristran nazivamo *pristran* procjenitelj za θ, a razliku $\mathbb{E}(T) - \theta$ nazivamo *pristranost*.

Drugi kriterij je uzimanje procjenitelja s najmanjom varijancom obzirom na ostale procjenitelje za isti parametar. Takav procjenitelj se naziva *procjenitelj minimalne varijance* i generalno se koristi u frekvencionističkoj statistici.

Treći kriterij za uspoređivanje procjenitelja je srednjekvadratna greška.

Definicija 4.1.3. *Neka je T procjenitelj za θ.* Srednjekvadratna greška *od T je, ako postoji, broj*

$$\mathrm{MSE}(T) = \mathbb{E}(T - \theta)^2.$$

Srednjekvadratna greška se može rastaviti na zbroj varijance procjenitelja i kvadrata pristranosti:

$$\begin{aligned}\mathbb{E}(T-\theta)^2 &= \mathbb{E}(T - \mathbb{E}T)^2 + (\mathbb{E}T - \theta)^2 \\ &= Var(T) + pristranost^2.\end{aligned}$$

Neka je $X_1, X_2, ..., X_n$ slučajan uzorak iz normalne distribucije s parametrima μ i σ^2. Frekvencionistički točkovni procjenitelj za μ jest uzoračko očekivanje $\bar{X}$, koje je normalno distribuirano s parametrima μ i $\frac{\sigma^2}{n}$.

Kao Bayesovski procjenitelj za μ uzimamo očekivanje aposteriorne distribucije. Uz pretpostavku o normalnosti apriorne distribucije, aposteriorna distribucija je ponovno normalna gdje parametre m' i $(s')^2$ određujemo po ažurirajućem pravilu. Stoga vrijedi

$$\hat{\mu}_B = m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{X}.$$

Cilj je usporediti Bayesovski procjenitelj $\hat{\mu}_B$ s frekvencionističkim procjeniteljem $\hat{\mu}_F = \bar{X}$, uz pretpostavku da je parametar μ fiksirana nepoznata vrijednost.

Očekivanje aposteriorne distribucije $\hat{\mu}_B$ je linerna funkcija slučajne varijable $\bar{X}$ stoga je njeno očekivanje

$$\begin{aligned}\mathbb{E}(\hat{\mu}_B) &= \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \mathbb{E}(\bar{X}) \\ &= \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \mu.\end{aligned}$$

Sada je pristranost aposteriornog očekivanja $\hat{\mu}_B$ jednaka

$$\mathbb{E}(\hat{\mu}_B) - \mu = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \mu - \mu = \frac{\sigma^2}{ns^2 + \sigma^2} \cdot (m - \mu)$$

gdje su m i s^2 parametri očekivanja i varijance apriorne distribucije.

Dakle, $\hat{\mu}_B$ je pristran procjenitelj za μ. Pristranost bi bila 0 jedino kada bi apriorno očekivanje m bilo jednako stvarnoj vrijednosti nepoznatog parametra μ, no vjerojatnost da se to dogodi je 0.

S druge strane, frekvencionistički točkovni procjenitelj za μ, $\hat{\mu}_F$, jest nepristran procjenitelj za μ:

$$\mathbb{E}\,(\hat{\mu}_F) = \mathbb{E}\,(\bar{X}) = \mu \quad \Rightarrow \quad \mathbb{E}\,(\hat{\mu}_F) - \mu = 0.$$

Nadalje, varijanca aposteriornog očekivanja $\hat{\mu}_B$ je

$$\begin{aligned} Var\,(\hat{\mu}_B) &= Var\left(\frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{X}\right) \\ &= \left(\frac{n/\sigma^2}{n/\sigma^2 + 1/s^2}\right)^2 \cdot Var(\bar{X}) \\ &= \left(\frac{n/\sigma^2}{n/\sigma^2 + 1/s^2}\right)^2 \cdot \frac{\sigma^2}{n} \\ &= \left(\frac{ns^2}{ns^2 + \sigma^2}\right)^2 \cdot \frac{\sigma^2}{n}. \end{aligned}$$

To je očito manje od $\frac{\sigma^2}{n}$, što je varijanca frekvencionističkog procjenitelja $\hat{\mu}_F$.

Izračunajmo još srednjekvadratnu grešku oba procjenitelja.

$$\begin{aligned} \mathrm{MSE}\,(\hat{\mu}_B) &= Var\,(\hat{\mu}_B) + pristranost^2 \\ &= \left(\frac{ns^2}{ns^2 + \sigma^2}\right)^2 \cdot \frac{\sigma^2}{n} \;+\; \frac{\sigma^2}{ns^2 + \sigma^2} \cdot (m - \mu). \end{aligned}$$

S druge strane, $\hat{\mu}_F$ je nepristran procjenitelj za μ pa je

$$\mathrm{MSE}\,(\hat{\mu}_F) = Var\,(\hat{\mu}_F) = \frac{\sigma^2}{n}.$$

U slučaju kada imamo apriori informacije o nepoznatom parametru, apriorno očekivanje m će biti blizu μ pa Bayesovski procjenitelj $\hat{\mu}_B$ ima manju srednjekvadratnu grešku od frekvencionističkog $\hat{\mu}_F$.

Primjer 4.1.4. [1] *Bacamo 10 loptica za golf s visine od jednog metra i mjerimo visinu (u centimetrima) na koju se loptica podiže kada odskoči od tla. Izmjerene visine dane su u idućoj tablici.*

76.4	78.9	80.5	77.4	78.0	77.7	79.1	75.9	74.2	76.7

Pretpostavimo da je visina na koju loptica odskoči normalno distribuirana sa standardnom devijacijom $\sigma = 2$. Procijenimo parametar očekivanja koristeći:
(a) normalnu apriornu distribuciju $N(m, s^2)$ s parametrima $m = 75$ i $s = 10$
(b) normalnu apriornu distribuciju $N(m, s^2)$ s parametrima $m = 78$ i $s = 7$
(c) Jeffreyjevu apriornu distribuciju.

Bayesovski točkovni procjenitelj za parametar očekivanja bit će očekivanje aposteriorne distribucije. U sva tri slučaja procjenitelj možemo dobiti pomoću ažurirajućeg pravila

$$\hat{\mu}_B = m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{x}.$$

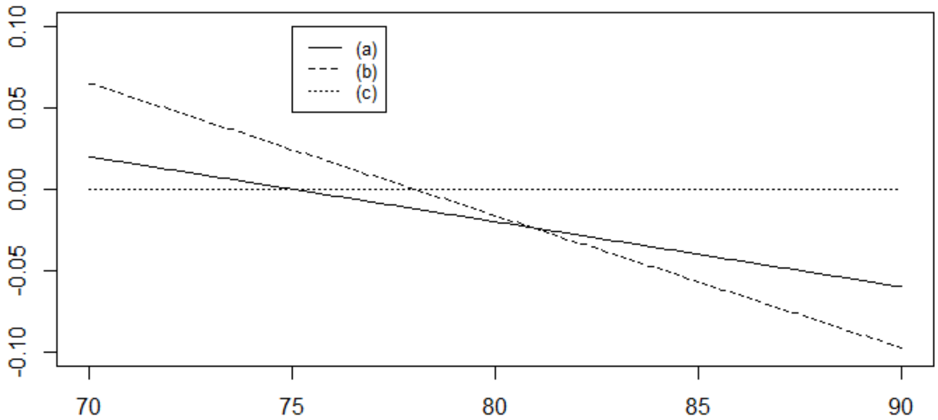
Označimo s $\hat{\mu}_B^a$, $\hat{\mu}_B^b$ te $\hat{\mu}_B^c$ Bayesovske procjenitelje koje dobivamo redom iz podzadataka (a), (b) i (c). Aritmetičku sredinu $\bar{x}$ dobivamo iz danog uzorka i ona iznosi $\bar{x} = 78.96$. Slijedi

$$\hat{\mu}_B^a = 78.94 \quad \hat{\mu}_B^b = 78.63 \quad \hat{\mu}_B^c = 78.96.$$

Izračunajmo još pristranost, varijancu i srednjekvadratnu grešku svakog od procjenitelja za različite stvarne vrijednosti parametra μ.
Pristranost možemo izraziti kao funkciju nepoznatog parametra na sljedeći način

$$bias\,(\hat{\mu}_B) = \mathbb{E}\,(\hat{\mu}_B) - \mu = \frac{\sigma^2}{ns^2 + \sigma^2} \cdot (m - \mu).$$

[1] Primjer preuzet iz [1].



Slika 4.1: Pristranosti procjenitelja u ovisnosti o stvarnoj vrijednosti parametra μ

S grafa možemo uočiti kako je jedino posljedni procjenitelj nepristran. On upravo odgovara frekvencionističkom točkovnom procjenitelju za μ, pošto je korištena Jeffreyjeva apriorna distribucija pa dobivamo $\hat{\mu}_B^c = \bar{x} = \hat{\mu}_F$.

Varijancu dobivamo koristeći izraz

$$Var\left(\hat{\mu}_B\right) = \left(\frac{ns^2}{ns^2 + \sigma^2}\right)^2 \cdot \frac{\sigma^2}{n}.$$

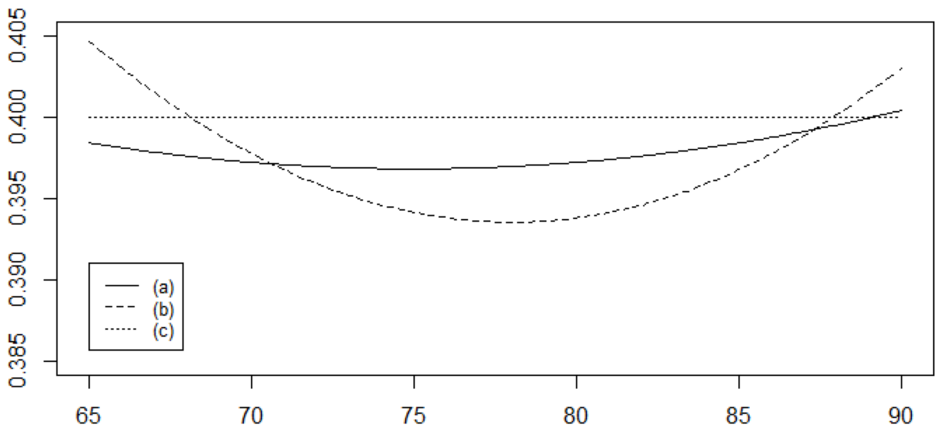
Dobivamo

$$Var\left(\hat{\mu}_B^a\right) = 0.397 \quad Var\left(\hat{\mu}_B^a\right) = 0.393 \quad Var\left(\hat{\mu}_B^a\right) = 0.4.$$

Dakle, za prave Bayesovske procjenitelje dobivamo varijancu manju od varijance frekvencionističkog procjenitelja.

Srednjekvadratnu grešku ponovno možemo napisati kao funkciju nepoznatog parametra te je prikazati na grafu za različite stvarne vrijednosti parametra μ:

$$\text{MSE}\left(\hat{\mu}_B\right) = Var\left(\hat{\mu}_B\right) + bias^2.$$



Slika 4.2: Srednjekvadratna greška procjenitelja u ovisnosti o stvarnoj vrijednosti parametra μ

Pošto se u posljednjem slučaju radi o frekvencionističkom procjenitelju, njegova pristranost jednaka je 0 stoga je srednjekvadratna greška jednaka varijanci. S grafa možemo uočiti da će srednjekvadratne greške preostala dva procjenitelja biti manje od srednjekvadratne greške frekvencionističkog procjenitelja kada je Bayesovski procjenitelj dovoljno blizu pravoj vrijednosti nepoznatog parametra, kako smo i ranije zaključili.

4.2 Intervalna procjena

U prošlom poglavlju upoznali smo pojam Bayesovskog intervala vjerodostojnosti i spomenuli kako je on analogon frekvencionstičkog intervala pouzdanosti. U ovom odjeljku uspoređujemo te dvije vrste intervala.

Definicija 4.2.1. *Neka su $L_n = l_n(X_1, ..., X_n)$ i $D_n = d_n(X_1, ..., X_n)$ funkcije slučajnog uzorka $X_1, ..., X_n$. Za $[L_n, D_n]$ kažemo da je $(1-\alpha)\cdot 100\%$* interval pouzdanosti *za parametar θ ako vrijedi*

$$\mathbb{P}(L_n \leq \theta \leq D_n) \geq 1 - \alpha, \quad \alpha \in \langle 0, 1 \rangle.$$

Frekvencionistički gledano, pouzdani interval za μ je interval za koji je velika vjerojatnost da sadrži pravu vrijednost nepoznatog parametra. Interval tražimo pomoću distribucije procjenitelja, u ovom slučaju to je uzoračko očekivanje $\bar{X}$. Distribucija od $\bar{X}$ je normalna s parametrima μ i $\frac{\sigma^2}{n}$ stoga slijedi

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1).$$

Iz toga slijedi

$$\mathbb{P}\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

gdje je $z_{\frac{\alpha}{2}}$ $\frac{\alpha}{2}$-kvantil standrardne normalne distribucije.

Zapišemo li gornju vjerojatnost t.d. je μ u sredini, dobivamo

$$\mathbb{P}\left(\bar{X} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Aritmetička sredina $\bar{X}$ je slučajna varijabla pa su krajevi intervala slučajni, a μ je fiksirana nepoznata vrijednost.

Interval zapisujemo u obliku

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Ovo se razlikuje od Bayesovskog intervala vjerodostojnosti koji smo dobili u prošlom poglavlju. Razlog tomu je što se Bayesovski interval vjerodostojnosti temelji na aposteriornoj distribuciji, a ona je dobivena iz subjektivno odabrane apriorne distribucije. Odabirom drugačije apriorne distribucije dobili bismo malo drugačiji interval.

U slučaju Jeffreyjeve apriori distribucije, apriorno očekivanje je $m' = \bar{X}$, a apriorna varijanca $(s')^2 = \frac{\sigma^2}{n}$. Stoga i Bayesov i frekvencionstički interval pouzdanost imaju sljedeći oblik

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

no imaju različitu interpretaciju.

U frekvencionističkom slučaju, μ je fiksirana vrijednost, a krajeve intervala računamo pomoću distribucije uzoračkog očekivanja $\bar{X}$. Uz dani uzorak dobivamo konkretan interval. Nivo značajnosti $(1 - \alpha) \cdot 100\%$ znači da $(1 - \alpha) \cdot 100\%$ intervala računatih na ovaj način sadrže pravu vrijednost nepoznatog parametra μ.

U Bayesovskom slučaju, μ je slučajna varijabla, a pouzdanost $(1-\alpha)\cdot 100\%$ znači da je uvjetna vjerojatnost da izračunati interval sadrži pravu vrijednost nepoznatog parametra μ, uz dane podatke, jednaka upravo $(1-\alpha)$.

Primjer 4.2.2. *Odredimo 95%-tni Bayesovski interval vjerodostojnosti za parametar očekivanja iz prethodnog primjera. U sva tri slučaja interval će biti oblika*

$$\left[m' - z_{\frac{\alpha}{2}} \cdot s' \ , \ m' + z_{\frac{\alpha}{2}} \cdot s'\right]$$

gdje su m' i s' očekivanje i standardna devijacija aposteriornih disribucija.
Očekivanja aposteriornih distribucija smo izračunali u prethodnom primjeru i ona iznose

$$m'_a = 78.94 \quad m'_b = 78.63 \quad m'_c = 78.96.$$

Standardne devijacije računamo pomoću ažurirajućeg pravila

$$s'_a = 0.6312 \quad s'_b = 0.6299 \quad s'_c = 0.6325.$$

Za $\alpha = 0.05$ $\ z_{\frac{\alpha}{2}} = z_{0.025}$ određujemo iz tablice distribucije standardne normalne distribucije i on iznosi 1.96.
Konačno, dobivamo intervale:

$$[77.7, 80.18]$$
$$[77.4, 79.86]$$
$$[77.7, 80.19]\,.$$

Intervali su jako slični, no ipak postoje male razlike upravo zbog subjektivnog izbora apriornih distribucija. U posljednjem slučaju korištena je Jeffreyjeva apriorna distribucija, stoga je taj interval jednak 95% frekvencionističkom pouzdanom intervalu za μ.
Frekvencionistički gledano, kada bismo ponavljali pokus, tj. kada bismo ponovno bacali loptice i mjerili visinu odskoka, interval [77.7, 80.19] *bi u 95% slučajeva trebao sadržavati pravu vrijednost parametra očekivanja μ.*
Bayesovski gledano, zanima nas samo provedeni pokus i tvrdimo da, uz podatke iz provedenog pokusa, vjerojatnost da interval [77.7, 80.19] *sadrži pravu vrijednost parametra μ iznosi 95%.*

4.3 Testiranje hipoteza o parametru μ

Prije no što promotrimo razlike Bayesovskog i frekvencionističkog pristupa kod testiranja hipoteza valja se upoznati s osnovnim pojmovima koji se koriste kod statističkih testova.

Statistička hipoteza jest pretpostavka o populacijskoj razdiobi promatrane varijable. Osnovnu hipotezu koju testiramo nazivamo *nulhipoteza* i označavamo s H_0. Uz nulhipotezu, postavlja se i njoj *alternativna hipoteza* koju označavamo s H_1.
Statistički test je pravilo podjele prostora vrijednosti uzoraka na dva podskupa: na područje vrijednosti uzoraka koji su konzistentni s H_0, i na njegov komplement u kojemu se nalaze vrijednosti nekonzistentne s H_0.
Odluka o odbacivanju ili neodbacivanju nulthipoteze donosi se na temelju *testne statistike*. Područje u kojem testna statistika poprima vrijednosti dijeli se na područje koje je konzistentno s H_0 i na područje nekonzistentno s H_0. Područje testne statistike koje je nekozistentno s H_0 naziva se *kritično područje*. Dakle, ako se opažena vrijednost testne statistike nalazi u kritičnom području, odbacujemo H_0 (u korist H_1). *Razina značajnosti testa* α je vjerojatnost odbacivanja H_0 ako je H_0 istinita hipoteza.

Jednostrane hipoteze

Frekvencionistički testovi o parametru temelje se na uzoračkoj distribuciji procjenitelja za taj parametar. Ako želimo provesti testove o parametru očekivanja μ, za procjenitelj uzimamo aritmetičku sredinu uzorka $\bar{X}$. Ako je dan slučajan uzorak $X_1, X_2, ..., X_n$ iz normalne distribucije $N(\mu, \sigma^2)$, tada je aritmetička sredina uzorka normalno distribuirana s parametrima μ i $\frac{\sigma^2}{n}$, $N(\mu, \frac{\sigma^2}{n})$.

Za početak postavljamo hipoteze

$$H_0 : \mu \leq \mu_0$$

$$H_1 : \mu > \mu_0$$

gdje je μ_0 unaprijed zadana vrijednost. U alternativnu hipotezu stavljamo ono što bismo željeli provjeriti pošto je cilj odbaciti nultu hipotezu. Uz pretpostavku da je H_0 istinita, aritmetička sredina $\bar{X}$ je normalno distribuirana s parametrima μ_0 i $\frac{\sigma^2}{n}$. Stoga varijabla

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

ima standardnu normalnu distribuciju. Slučajnu varijablu Z nazivamo *testna statistika*. Za unaprijed odabranu razinu značajnosti α računamo kritično područje. Zatim, na temelju danog uzorka, određujemo realizaciju testne statistike z. Ako vrijednost upadne u kritično područje, odbacujemo H_0 u korist H_1 na razini značajnosti α. U suprotnom ne možemo odbaciti H_0.

Drugi način za donošenje zaključka je određivanje *p-vrijednosti* testa, tj. vjerojatnosti da testna statistika Z poprimi vrijednosti koje su, uz pretpostavku da je H_0 istinita, manje ili jednako vjerojatne od opažene vrijednosti testne statistike. U slučajnu jednostranog testa zadanog s

$$H_0 : \mu \geq \mu_0, \quad H_1 : \mu < \mu_0$$

p-vrijednost jednaka je

$$p - vrijednost = \mathbb{P}\left(Z < z \,|\, H_0\right)$$

gdje je Z testna statistika, a z realizacija testne statistike. U slučajnu jednostranog testa zadanog s

$$H_0 : \mu \leq \mu_0, \quad H_1 : \mu > \mu_0$$

p-vrijednost jednaka je

$$p - vrijednost = \mathbb{P}\left(Z > z \,|\, H_0\right).$$

Kod Bayesovskih testova, vrijednost μ_0 obično dolazi iz prethodnih zapažanja. Ako statističkim testom ne možemo odbaciti H_0, tj. ne možemo zaključiti da je parametar μ, uz dane nove podatke, različit od μ_0, dodavanjem novih podataka nismo dobili nikakve nove informacije.

Testiranje jednostranih hipoteza u Bayesovskoj statistici se provodi tako da se izračuna aposteriorna vjerojatnost nulte hipoteze

$$\mathbb{P}\left(H_0 : \mu \leq \mu_0 \,|\, x_1, ..., x_n\right) = \mathbb{P}\left(\mu \leq \mu_0 \,|\, x_1, ..., x_n\right) = \int_{-\infty}^{\mu_0} f(\mu \,|\, x_1, ..., x_n)\, d\mu.$$

U slučaju kada je aposteriorna distribucija $f(\mu \,|\, x_1, ..., x_n)$ normalna s parametrima m' i $(s')^2$ traženu vjerojatnost lako možemo dobiti iz tablice distribucije standardne normalne distribucije

$$\mathbb{P}\left(\mu \leq \mu_0 \,|\, x_1, ..., x_n\right) = \mathbb{P}\left(\frac{\mu - m'}{s'} \leq \frac{\mu_0 - m'}{s'}\right) = \mathbb{P}\left(Z \leq \frac{\mu_0 - m'}{s'}\right)$$

gdje je $Z = \frac{\mu - m'}{s'}$ slučajna varijabla sa standardnom normalnom distribucijom.

Ako je vjerojatnost manja od α odbacujemo H_0 i možemo zaključiti $\mu > \mu_0$.

Primjer 4.3.1. *Testirajmo sljedeće hipoteze za parametar očekivanja iz primjera 4.1.4 na razini značajnosti od 5% koristeći frekvencionistički i Bayesovski statistički test.*

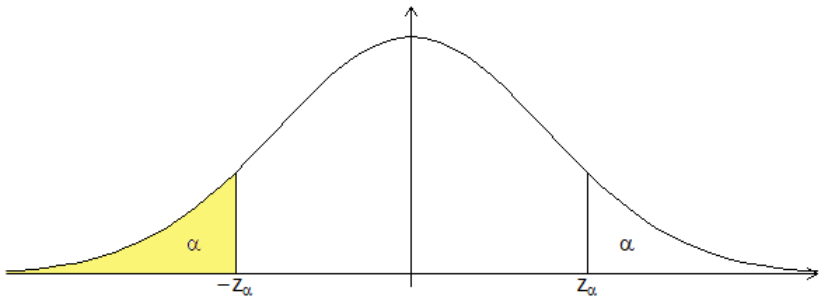
$$H_0 : \mu \geq 81$$

$$H_1 : \mu < 81$$

Frekvencionistički test provodimo na sljedeći način. Uz pretpostavku da je H_0 *istinita, frekvencionistički točkovni procjenitelj,* $\bar{X}$*, je normalno distribuirana slučajna varijabla s parametrima* $\mu_0 = 81$ *i* $\frac{\sigma^2}{n} = \frac{4}{10} = \frac{2}{5}$*. Artmetičku sredinu danog uzorka smo izračunali ranije i ona iznosi* $\bar{x} = 78.96$ *stoga je realizacija testne statistike dana s*

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{78.96 - 81}{2/\sqrt{10}} = -3.23.$$

Za danu razinu značajnosti $\alpha = 0.05$ *iz tablice distribucije standardne normalne distribucije određujemo* z_α *takav da* $\mathbb{P}(Z \leq -z_\alpha) = \mathbb{P}(Z \geq z_\alpha) = 0.05$ *i dobivamo* $z_\alpha = 1.64$.



Slika 4.3: Određivanje kritičnog područja u slučaju jednostranih hipoteza

Kritično pordučje je sada $\langle -\infty, -1.64]$*. S obzirom da realizacija testne statistike upada u kritično područje, odbacujemo nultu hipotezu na razini značajnosti* 5%*.*
Izračunajmo još p-vrijednost ovog testa.

$$\mathbb{P}(Z < z \,|\, H_0) = \mathbb{P}(Z < -3.23) = 0.0006.$$

Dakle, teško bismo mogli zaključiti da je vrijednost parametra μ *veća od 81.*

Provedimo još Bayesovski test. Podsjetimo se, u sva tri slučaja dobivamo normalnu aposteriornu distribuciju. U prethodna dva primjera izračunali smo parametre očekivanja i standardne devijacije aposteriornih distribucija i oni iznose

$$m'_a = 78.94 \quad m'_b = 78.63 \quad m'_c = 78.96,$$

$$s'_a = 0.6312 \quad s'_b = 0.6299 \quad s'_c = 0.6325.$$

Sada aposteriornu vjerojatnost nulte hipoteze možemo izračunati pomoću tablice distribucije standardne normalne distribucije. U prvom slučaju dobivamo sljedeće

$$\mathbb{P}(\mu \geq \mu_0 \,|\, x_1, ..., x_n) = \mathbb{P}\left(Z \geq \frac{\mu_0 - m'_a}{s'_a}\right) = \mathbb{P}(Z \geq 3.26) = 1 - \mathbb{P}(Z \leq 3.26) = 0.0006.$$

U preostala dva slučajeva dobivamo

$$\mathbb{P}\left(Z \geq \frac{\mu_0 - m'_b}{s'_b}\right) = \mathbb{P}\left(Z \geq 3.76\right) = 0.0001.$$

$$\mathbb{P}\left(Z \geq \frac{\mu_0 - m'_c}{s'_c}\right) = \mathbb{P}\left(Z \geq 3.23\right) = 0.0006.$$

Dakle, u sva tri slučaja je vjerojatnost manja od zadane razine značajnosti od 5% *stoga u sva tri slučajeva ponovno odbacujemo nulhipotezu u korist alternativne.*

Dvostrane hipoteze

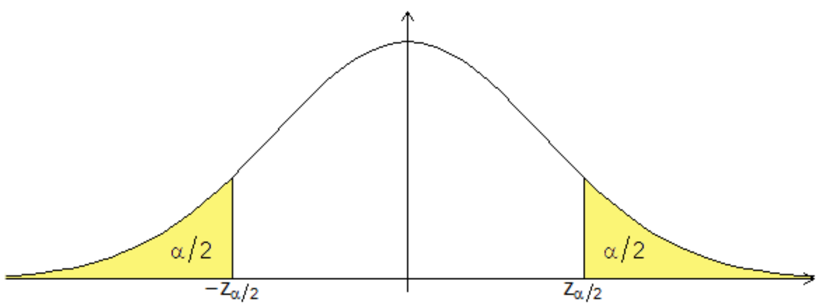
U slučaju dvostranog testa, hipoteze su sljedeće

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Procedura frekvencionističkog testa ista je kao kod jednostranog testa, no u ovom slučaju je kritično područje dvostrano. Ono je sada oblika

$$\left\langle -\infty, -z_{\frac{\alpha}{2}} \right] \bigcup \left[z_{\frac{\alpha}{2}}, \infty \right\rangle.$$



Slika 4.4: Određivanje kritičnog područja u slučaju dvostranih hipoteza

Također, p-vrijednost u ovom slučaju obuhvaća vjerojatnosti oba repa

$$p - vrijednost = \mathbb{P}\left(|Z| \ \geq \ |z| \,|\, H_0 \right).$$

Ponovno, ako je p-vrijednost $\leq \alpha$, odbacujemo H_0.

Nulhipotezu također odbacujemo ako realizacija testne statistike upadne u kritično područje, tj.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \in \left\langle -\infty, -z_{\frac{\alpha}{2}} \right] \bigcup \left[z_{\frac{\alpha}{2}}, \infty \right\rangle.$$

Izrazimo li μ_0 iz gornjeg izraza dobivamo

$$\mu_0 \geq \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad \text{i} \quad \mu_0 \leq \bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

Prisjetimo se, procjena frekvencionističkog interval pouzdanosti je oblika

$$\left[\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \ , \ \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right].$$

Dakle, ako odbacimo H_0 na razini značajnosti α onda μ_0 ne leži u $(1-\alpha)\cdot 100\%$ intervalu pouzdanosti za μ. Drugim rječima, $(1-\alpha)\cdot 100\%$ interval pouzadnosti sadrži sve vrjednosti μ_0 za koje ne bismo odbacili nulhipotezu, na razini značajnosti α.

Ako želimo testirati dvostranu hipotezu Bayesovskim testom, a imamo neprekidnu apriornu distribuciju, ne možemo samo izračunati aposteriornu vjerojatnost nulte hipoteze jer će ona biti 0. Umjesto toga računamo $(1-\alpha)\cdot 100\%$ interval vjerodostojnosti za μ koristeći aposteriornu distribuciju. Ako μ_0 leži u intervalu, zaključujemo da je moguće da μ poprimi vrijednost μ_0 pa u tom slučaju ne odbacujemo H_0. U suprotnom odbacujemo H_0. Ovakav način testiranja dvostranih hipoteza opravdan je upravo interpretacijom Bayesovskog intervala vjerodostojnosti. Ako je μ_0 prava vrijednost nepoznatog parametra, $(1-\alpha)\cdot 100\%$ interval vjerodostojnosti će ju sadržavati s vjerojatnošću $1-\alpha$.

Primjer 4.3.2. [2] *Uzet je slučajan uzorak od 25 računa za električnu energiju u rasponu od 3 mjeseca za kućanstva s Novog Zelanda. Izmjereni podaci dani su u sljedećoj tablici:*

514	*536*	*345*	*440*	*427*	*443*	*386*	*418*	*364*	*483*
506	*385*	*410*	*561*	*275*	*306*	*294*	*402*	*350*	*343*
480	*334*	*324*	*414*	*296*					

[2]Primjer preuzet iz [1].

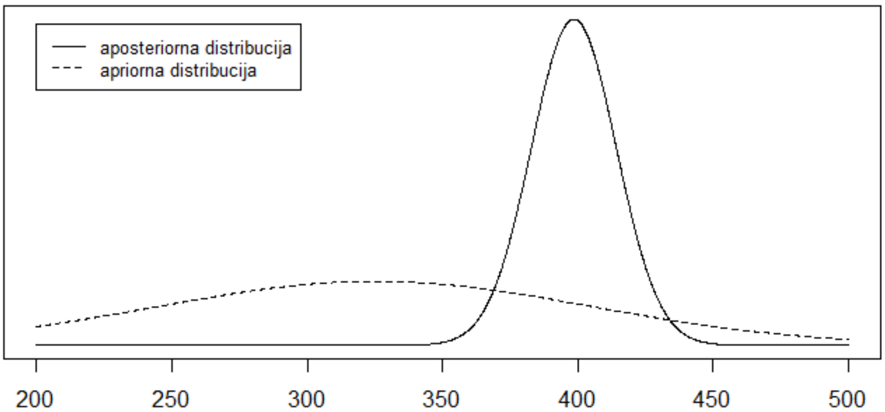
Pretpostavimo da je količina iskorištene električne energije normalno distribuirana s poznatom standardnom devijacijom $\sigma = 80$.

Uzmimo normalnu apriornu distribuciju $N(325, 80^2)$ *i izračunajmo aposteriornu distribuciju za nepoznati parametar očekivanja* μ.
Pošto je apriorna distribucija normalna, aposteriorna distribucija će također biti normalna $N(m', s'^2)$*, a parametre možemo izračunati pomoću ažurirajućeg pravila.*

$$\frac{1}{(s')^2} = \frac{1}{s^2} + \frac{n}{\sigma^2} = \frac{1}{80^2} + \frac{25}{80^2} \qquad s' = 15.69$$

$$m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{x} = \frac{1/80^2}{25/80^2 + 1/80^2} \cdot 325 + \frac{25/80^2}{25/80^2 + 1/80^2} \cdot 401.44 = 398.5$$

Dakle, aposteriorna distribucija je $N(398.5, 15.69^2)$.



Slika 4.5: Odnos apriorne i aposteriorne distribucije

Odredimo sada procjenu 95%*-tnog Bayesovskog intervala vjerodostojnosti za* μ. *Bayesovski* $(1-\alpha)\cdot 100\%$ *interval vjerodostojnosti je oblika*

$$\left[m' - z_{\frac{\alpha}{2}} \cdot s' \ , \ m' + z_{\frac{\alpha}{2}} \cdot s'\right]$$

gdje su m' i s' očekivanje i standardna devijacija aposteriorne disribucije stoga je procjena 95%-tnog Bayesovskog intervala vjerodostojnosti

$$[398.5 - 1.96 \cdot 15.69\ ,\ 398.5 + 1.96 \cdot 15.69] = [367.75, 429.25]\,.$$

Testirajmo još sljedeće hipoteze

$$H_0 : \mu = 350$$

$$H_1 : \mu \neq 350$$

na razini značajnosti od 5%.

Pošto se radi o testiranju dvostranih hipoteza, testiranje provodimo pomoću intervala vjerodostojnosti. Zadana je razina značajnosti od 5% što znači da koristimo 95%-tni interval vjerodostojnosti. Vrijednost $\mu_0 = 350$ ne leži u 95%-tnom intervalu vjerodostojnosti što znači da odbacujemo H_0 na razini značajnosti od 5%.

Poglavlje 5

Razni primjeri

U ovom poglavlju primjenjujemo uvedenu i obrađenu teoriju te metode za zaključivanje o nepoznatom parametru očekivanja na odabranim zadacima. Zadaci su preuzeti iz [1].

Primjer 5.0.1. *Neka je dan slučajan uzorak duljine $n = 10$ iz normalne distribucije $N(\mu, \sigma^2)$ gdje je $\sigma^2 = 4$ poznat.*

3.07	*7.51*	*5.95*	*6.83*	*8.80*	*4.19*	*7.44*	*7.06*	*9.67*	*6.89*

Pretpostavimo da su jedine moguće vrijedosti za nepoznati parametar

$$4.0,\ 4.5,\ 5.0,\ 5.5,\ 6.0,\ 6.5,\ 7.0,\ 7.5,\ 8.0,\ 8.5,\ 9.0 \ i\ 9.5$$

te da nemamo dodatnih informacija o nepoznatom parametru. Dakle, uzimamo da su sve moguće vrijednosti jednako vjerojatne:

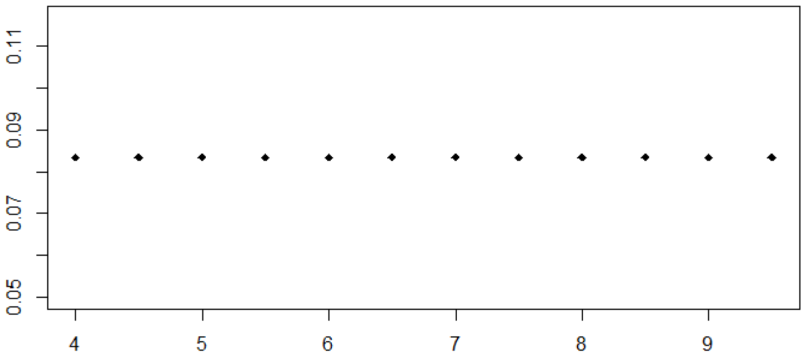
$$\mathbb{P}(\mu = \mu_0) = \frac{1}{12} \quad \forall \mu_0 \in \{4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5\}.$$

Pronađimo aposteriornu distribuciju za nepoznati parametar očekivanja μ.

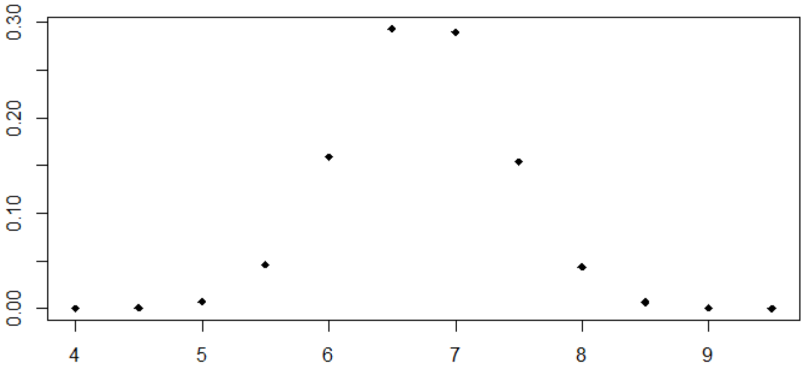
Zadana je diskretna apriorna distribucija stoga aposteriornu distribuciju računamo po

$$\mathbb{P}(\mu \,|\, x) = \frac{vjerodostojnost \times apriorna\ vjerojatnost}{\sum vjerodostojnost \times apriorna\ vjerojatnost}$$

gdje je apriorna vjerojatnost jednaka $\frac{1}{12}$ za svaku moguću vrijednost za μ, a vjerodostojnost računamo koristeći (3.1). *Račun provodimo u programskom jeziku R. Kod je dan u dodatku, a ovdje navodimo samo rezultate.*
U nastavku je dana tablica s dobivenom aposteriornom distribucijom te grafovi apriorne i aposteriorne distribucije.



Slika 5.1: Apriorna distribucija



Slika 5.2: Aposterirona distribucija

μ_0	*4.0*	*4.5*	*5.0*	*5.5*	*6.0*	*6.5*
apost	$2.6317 \cdot 10^{-5}$	$5.9227 \cdot 10^{-4}$	$7.1347 \cdot 10^{-3}$	$4.6003 \cdot 10^{-2}$	$1.5877 \cdot 10^{-1}$	$2.933 \cdot 10^{-1}$

μ_0	*7.0*	*7.5*	*8.0*	*8.5*	*9.0*	*9.5*
apost	$2.9002 \cdot 10^{-1}$	$1.535 \cdot 10^{-1}$	$4.3487 \cdot 10^{-2}$	$6.5943 \cdot 10^{-3}$	$5.3524 \cdot 10^{-4}$	$2.3253 \cdot 10^{-5}$

S grafova uočavamo kako se zaključivanje o nepoznatom parametru mijenja. Također, vrijednosti $\mu_0 = 6.5$ *i* $\mu_0 = 7.0$ *su vjerojatnije od ostalih što upućuje da bismo na temelju danog uzorka zaljučili da je vrijednost nepoznatog parametra očekivanja jednaka* 6.5 *ili* 7.

Pretpostavimo sada da smo naknadno dobili još jedan slučajan uzorak duljine $n = 6$ *iz iste distribucije* $N(\mu, 4)$.

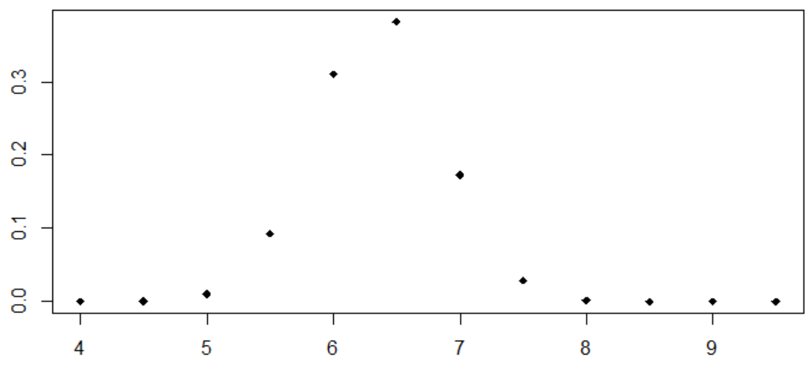
6.22	*3.99*	*3.67*	*6.35*	*7.89*	*6.13*

Izračunajmo aposteriornu distribuciju nepoznatog parametra uz dani novi uzorak.

Pošto sada već imamo neke informacije o nepoznatom parametru, ta saznanja uključujemo u apriornu distribuciju na način da aposteriorna distribucija iz prethodne analize postaje trenutna apriorna distribucija. Vrijednosti i graf aposteriorne distribucije dani su u nastavku.

μ_0	*4.0*	*4.5*	*5.0*	*5.5*	*6.0*	*6.5*
apost	$6.149 \cdot 10^{-6}$	$4.1314 \cdot 10^{-4}$	$1.0212 \cdot 10^{-2}$	$9.2855 \cdot 10^{-2}$	$3.1061 \cdot 10^{-1}$	$3.8224 \cdot 10^{-1}$

μ_0	*7.0*	*7.5*	*8.0*	*8.5*	*9.0*	*9.5*
apost	$1.7304 \cdot 10^{-1}$	$2.8819 \cdot 10^{-2}$	$1.7657 \cdot 10^{-3}$	$3.9797 \cdot 10^{-5}$	$3.2998 \cdot 10^{-7}$	$1.0066 \cdot 10^{-9}$



Slika 5.3: Aposteriorna distribucija dodanog uzorka

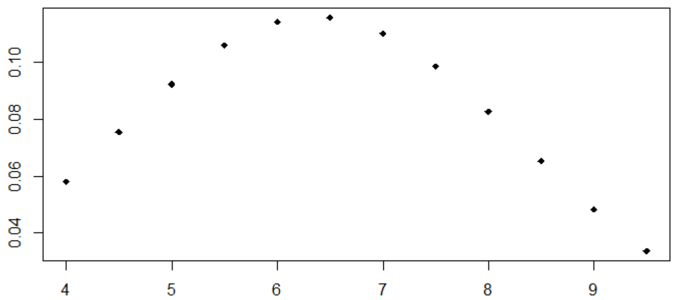
Ponovno uočavamo kako se zaključivanje o nepoznatom parametru mijenja uz dane nove podatke. Također, uz dani dodatni uzorak sada bismo zaključili da je vrijednost nepoznatog parametra očekivanja jednaka 6 *ili* 6.5.

Gledamo li dana dva uzorka kao jedan uzorak duljine $n = 16$ *dobili bismo iste rezultate pošto za apriornu distribuciju uz danu opaženu vrijednost uzimamo aposteriornu distribuciju dobivenu uz danu prethodnu opaženu vrijednost. Minimalne razlike bismo dobili kada bismo funkcije vjerodostojnosti za oba uzorka računali koristeći aritmetičke sredine uzoraka. Kod je dan u prilogu, a rezultat izostavljamo pošto razlike s obzirom na prvi način računanja nisu vidljive u rezultatima koje R ispisuje.*

Nadalje, umjesto da promatramo čitav uzorak duljine $n = 16$, *mogli bismo promatrati aritmetičku sredinu tog uzorka kao jednu danu vrijednost i na temelju nje izračunati aposteriornu distribuciju nepoznatog parametra.*

Budući da je slučajan uzorak dan iz normalne distribucije $N(\mu, \sigma^2)$, *aritmetička sredina* $\bar{X}$ *također je normalno distribuirana s parametrima* μ *i* $\frac{\sigma^2}{n}$. *Uzoračka aritmetička sredina iznosi* $\bar{x} = 6.3537$. *Tu vrijednost sada koristimo za računanje funkcije vjerodostojnosti i aposteriorne distribucije. Dobivena distribucija dana je u idućoj tablici.*

μ_0	*4.0*	*4.5*	*5.0*	*5.5*	*6.0*	*6.5*
apost	*0.058*	*0.0755*	*0.0923*	*0.1059*	*0.1142*	*0.1157*
μ_0	*7.0*	*7.5*	*8.0*	*8.5*	*9.0*	*9.5*
apost	*0.1101*	*0.0984*	*0.0826*	*0.0652*	*0.0483*	*0.0336*



Slika 5.4: Aposteriorna distribucija dobivena koristeći aritmetičku sredinu

Dobivamo drugačije vjerojatnosti, no distribucija je slična. Ponovno bismo zaključili da je vrijednost nepoznatog parametra očekivanja jednaka 6 *ili* 6.5.

Iz ovog primjera možemo zaključiti da je svejedno gledamo li sve uočene vrijednosti kao jedan veći uzorak ili odvajamo u više manjih uzoraka. Također, svejedno je na koji način računamo funkciju vjerodostojnosti. Konačno još možemo zaključiti da preciznije rezultate dobivamo analizirajući čitav uzorak nego jednu vrijednost, no uočavamo da je u suštini zaključak o vrijednosti nepoznatog parametra isti. U kontekstu ovog primjera, u oba slučaja dobivamo vrijednosti 6 *i* 6.5 *kao vjerojatnije vrijednosti nepoznatog parametra no u prvom slučaju, gdje analiziramo čitav uzorak, lakše bismo se odlučili za vrijednost* 6.5.

Primjer 5.0.2. *Neka je dan slučajan uzorak duljine $n = 15$ iz normalne distribucije $N(\mu, \sigma^2)$ gdje je $\sigma^2 = 4$ poznat te normalna apriorna distribucija $N(m, s^2)$ s parametrima $m = 20$ i $s^2 = 25$.*

26.8	26.3	28.03	28.5	26.3
31.9	28.5	27.2	20.9	27.5
28.0	18.6	22.3	25.0	31.5

Odredimo aposteriornu distribuciju te 95%-tni Bayesovski interval vjerodostojnosti.

Pošto je apriorna distribucija normalna, aposteriorna distribucija je također normalna, a njene parametre možemo pronaći pomoću ažurirajućeg pravila:

$$\frac{1}{(s')^2} = \frac{1}{s^2} + \frac{n}{\sigma^2}$$

$$m' = \frac{1/s^2}{n/\sigma^2 + 1/s^2} \cdot m + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2} \cdot \bar{x} = \frac{1/s^2}{1/s'^2} \cdot m + \frac{n/\sigma^2}{1/s'^2} \cdot \bar{x}$$

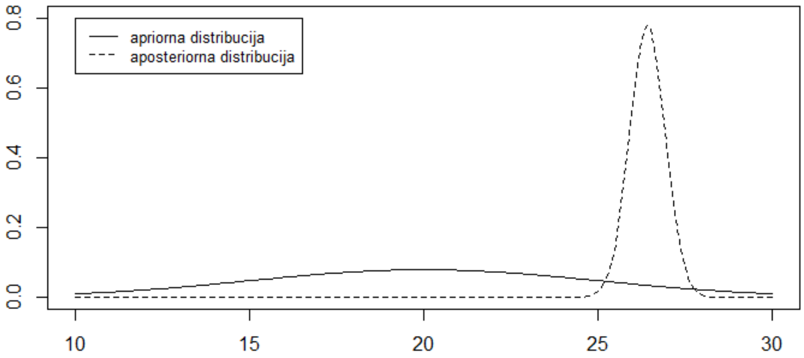
gdje su m' i s'^2 parametri aposteriorne distribucije, m i s^2 parametri apriorne distribucije, n duljina uzorka, σ^2 parametar varijance danog uzorka te $\bar{x}$ uzoračka aritmetička sredina.

Zadatak rješavamo u R-u, a ovdje navodimo dobivene rezultate. Kod iz R-a može se pronaći u dodatku.

Za parametre aposteriorne distribucije dobivamo

$$(s')^2 = 0.5137^2 = 0.2639 \quad m' = 26.4202$$

stoga je aposteriorna distribucija normalna $N(26.4, 0.3)$.



Slika 5.5: Odnos apriorne $N(20, 25)$ i aposteriorne $N(26.4, 0.3)$ distribucije

Još računamo 95%*-tni Bayesovski interval vjerodostojnosti. Pošto je aposteriorna distribucija normalna, interval je oblika*

$$\left[m' - z_{\frac{\alpha}{2}} \cdot s', m' + z_{\frac{\alpha}{2}} \cdot s'\right]$$

gdje je $z_{\frac{\alpha}{2}}$ $\frac{\alpha}{2}$*-kvantil standardne normalne distribucije. Za* $\alpha = 0.05$ *on iznosi* $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$*. stoga je traženi inteval dan s*

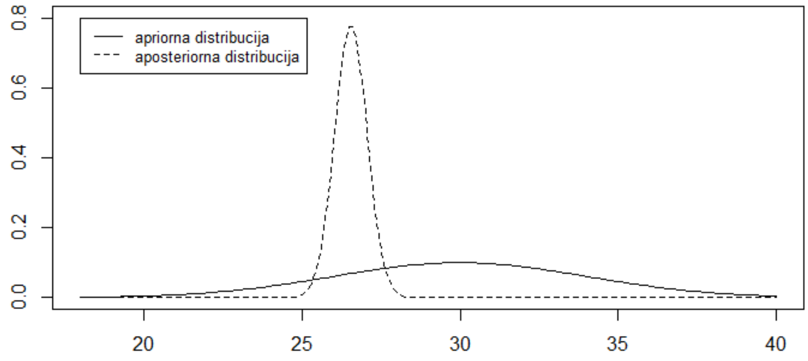
$$[26.4202 - 1.96 \cdot 0.5137\,,\ 26.4202 + 1.96 \cdot 0.5137] = [25.413, 27.427]\,.$$

Neka je sada apriorna distribucija normalna $N(m, s^2)$ *s parametrima* $m = 30$ *i* $s^2 = 16$*. Ponovno računamo aposteriornu distribuciju i* 95%*-tni Bayesovski interval vjerodostojnosti.*

Ponovno, apriorna distribucija je normalna pa će i aposteriorna distribucija biti normalna, a parametre ponovno određujemo po ažurirajućem pravilu. Dobivamo

$$(s')^2 = 0.5121^2 = 0.2623 \quad m' = 26.5462$$

stoga je aposteriorna distribucija normalna $N(26.5, 0.3)$*.*

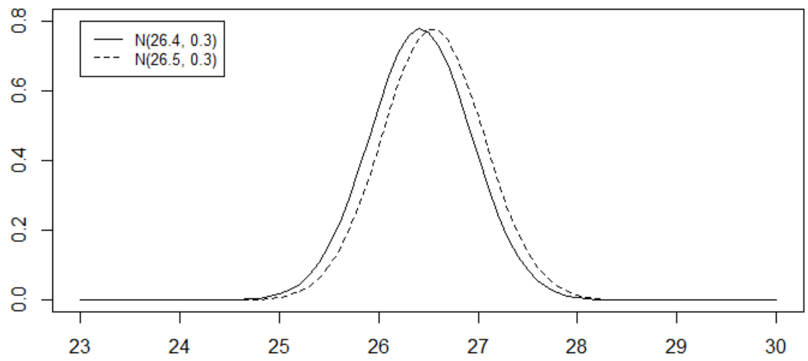


Slika 5.6: Odnos apriorne $N(30, 16)$ i aposteriorne $N(26.5, 0.3)$ distribucije

U ovom slučaju, 95%-tni Bayesovski interval vjerodostojnosti je

$$[26.5462 - 1.96 \cdot 0.5121\,,\ 26.5462 + 1.96 \cdot 0.5121] = [25.542, 27.55]\,.$$

Usporedimo sada dobivene aposteriorne distribucije.



Slika 5.7: Odnos aposteriornih distribucija

Primjećujemo kako dobivamo slične aposteriorne distribucije, a krenuli smo od poprilično različitih apriornih distribucija. Razlog tome leži u formuli za izračun očekivanja aposteriorne distribucije. Naime, koeficijent kojim se množi aritmetička sredina ima veću težinu od koeficijenta kojim se množi vrijednost očekivanja apriorne distribucije stoga dani uzorak ima veći utjecaj na parametar aposteriornog očekivanja nego apriorna distribucija. Što je dani uzorak veći, to je veći njegov utjecaj na aposteriornu distribuciju.

Ovo objašnjenje je u skladu i s intuitivnim objašnjenjem: što je dano više informacija o nepoznatom parametru, to je moguće točnije odrediti vrijednost nepoznatog parametra.

Vrijednosti parametara aposterinih distribucija razlikuju se minimalno, a posljedično tome, razlike između dobivenih 95%*-tnih intervala vjerodostojnosti su također male.*

Primjer 5.0.3. *Neka je dan slučajan uzorak duljine $n = 16$ iz normalne distribucije $N(\mu, \sigma^2)$ gdje je $\sigma^2 = 4$ poznat.*

4.09	4.68	1.87	2.62	5.58	8.68	4.07	4.78
4.79	4.49	5.85	5.9	2.4	6.27	6.3	4.47

Neka je apriorna distribucija zadana s

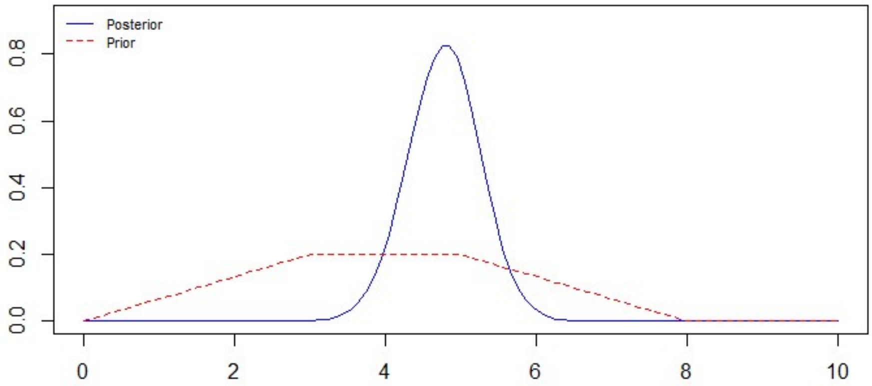
$$f(\mu) = \begin{cases} \mu, & 0 < \mu \leq 3 \\ 3, & 3 < \mu \leq 5 \\ 8 - \mu, & 5 < \mu \leq 8 \\ 0, & \mu \geq 8. \end{cases}$$

Odredimo aposteriornu distribuciju.

Za pronalazak aposteriorne distribucije u slučaju apriorne distribucije koja nije normalna koristimo funkciju `normgcp` *u programskom jeziku R. Kod je dan u prilogu, a ovdje prezentiramo samo rezultate.*

Napomena 5.0.4. *Apriorna distribucija je reskalirana kako bi vjerojatnosti u sumi dale 1.*

Po obliku krivulje aposteriorne distribucije mogli bismo zaključiti da se radi o normalnoj distribuciji, no parametri su nam nepoznati.



Slika 5.8: Apriorna i aposteriorna distribucija

Izračunajmo parametre aposteriornog očekivanja i aposteriorne standardne devijacije. Po definicijama matematičkog očekivanja i varijance parametre možemo računati na sljedeći način:

$$m' = \int_{-\infty}^{\infty} t\, f(t \,|\, x_1, x_2, ..., x_n)\, dt$$

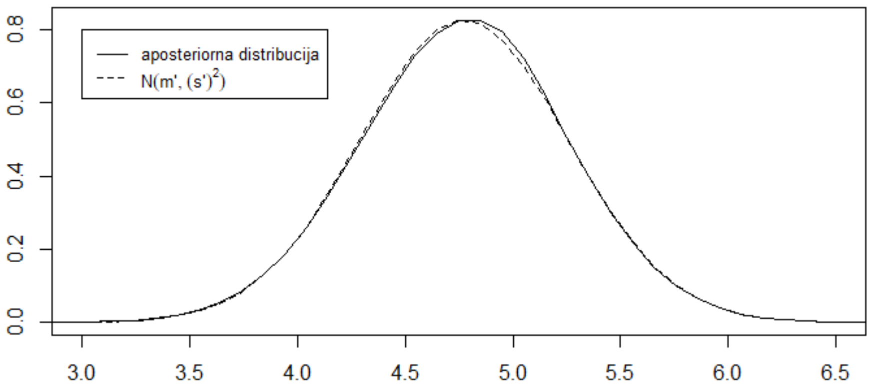
$$(s')^2 = \int_{-\infty}^{\infty} (t - m')^2\, f(t \,|\, x_1, x_2, ..., x_n)\, dt$$

gdje je $x_1, x_2, ..., x_n$ *dani uzorak pa je* $f(t \,|\, x_1, x_2, ..., x_n)$ *aposteriorna distribucija nepoznatog parametra uz dani uzorak.*

Za računanje aposteriornog očekivanja i aposteriorne standardne devijacije koristimo funkciju `integrate.xy` *u R-u koja računa integral numerički zadane funkcije.*

Dobivamo parametre $m' = 4.772$ *i* $s' = 0.4828$.

Usporedimo li dobivenu aposteriornu distribuciju s normalnom distribucijom s parametrima aposteriornog očekivanja i varijance, vidimo da su distribucije poprilično slične. Nedostatak ovako dobivene aposteriorne distribucije jest to što je ona numerički zadana pa s njom ne možemo dalje računati kako bismo mogli s nekom neprekidnom funkcijom. Zato je korisno, ako je moguće, numerički dobivenu aposteriornu distribuciju procijeniti nekom neprekidnom funkcijom.



Slika 5.9: Odnos aposteriorne distribucije i $N(m', (s')^2)$

Želimo li odrediti 95%*-tni interval vjerodostojnosti za nepoznati parametar, trebali bismo koristiti izraz 3.3 pošto je dobivena aposteriorna distribucije numerički dobivena. No, ako aposteriornu distribuciju aproksimiramo normalnom distribucijom s parametrima aposteriornog očekivanja i varijance, mogli bismo početne vrijednosti gornje i donje granice intervala procijeniti pomoću izraza za* 95%*-tni interval vjerodostojnosti za normalnu aposteriornu distribuciju, a onda eventualno malo pomaknuti granice kako bismo dobili odgovarajući interval za numerički dobivenu aposteriornu distribuciju.*

U slučaju normalne aposteriorne distribucije s parametrima aposteriornog očekivanja i varijance $N(4.772, 0.233)$*, dobili bismo interval*

$$[3.8262,\ 5.7188]\,.$$

Pošto vrijedi

$$\int_{3.8262}^{5.7188} f(t\,|\,x_1, x_2, ..., x_n)\,dt = 0.9494$$

mogli bismo kao 95%*-tni interval vjerodostojnosti uzeti gore dobiveni interval.*

Možemo pokušati malo promijeniti granice kako bismo dobili bližu procjenu. Krećemo od inicijalnih granica intervala, 3.8262 i 5.7188. Od donje granice oduzimamo dovoljno malu vrijednost ϵ, a gornjoj granici pridodajemo tu istu vrijednost ϵ te ponovno računamo integral aposteriorne funkcije gustoće, sada s promjenjenim granicama:

$$\int_{3.8262-\epsilon}^{5.7188+\epsilon} f(t\,|\,x_1, x_2, ..., x_n)\,dt.$$

Ako je novo dobivena vrijednost integrala i dalje manja od 0.95 ponavljamo postupak, tj.od nove donje granice ponovno oduzimamo neku vrijednost ϵ_2, a novoj gornjoj granici dodajemo ϵ_2. Nova vrijednost ϵ_2 može biti jednaka ϵ, a može biti i manja ukoliko mislimo da bismo uzimajući ϵ dobili preveliku vrijednost integrala. Postupak ponavljamo sve dok ne dobijemo vrijednost dovoljno blizu 0.95.

U ovom slučaju, za ϵ uzimamo vrijednost 0.001 te nakon tri iteracije (sve s istom vrijednosti ϵ) dobivamo sljedeći integral

$$\int_{3.8232}^{5.7218} f(t\,|\,x_1, x_2, ..., x_n)\,dt = 0.9501.$$

Zaključujemo da je dobivena vrijednost dovoljno blizu 0.95 stoga kao 95%*-tni interval vjerodostojnosti uzimamo*

$$[3.8232,\ 5.7218]\,.$$

Zaključujemo kako i uz naizgled kompliciraniju aporiornu distribuciju dobivamo relativno jednostavnu aposteriornu distribuciju koja se, iako dobivena numerički, može aproksimirati normalnom distribucijom te se tako olakšava daljnja analiza.

Bibliografija

[1] W. M. Bolstad, *Introduction to Bayesian Statistics*, A John Wiley & Sons, 2007.

[2] M. Huzak, *Vjerojatnost i matematička statistika*, 2006.

[3] ______, *Matematička statistika , predavanja*.

[4] D. Kelly i C. Smith, *Bayesian Inference for Probabilistic Risk Assessment*, Springer, 2011.

[5] G. E. P. Box i G. C. Tjao, *Bayesian Inference in Statistical Analysis*, Addison Wesley Publishing Company, 1973.

[6] J.Orloff i J. Bloom, *Bayesian Updating with Continuous Priors, Class 13*, MIT predavanja, 2014.

[7] T. M. Donovan i R. M. Mickey, *Bayesian Statistics for Beginners*, Oxford University Press, 2019.

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

[9] J. V. Stone, *Bayes' Rule: A Tutorial Instruction to Bayesian Analysis*, Sebtel Press, 2013.

Dodatak

Tablica vrijednosti funkcije distribucije standardne normalne razdiobe

	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

R kod za Primjer 5.0.1

```
uzorak <- c(3.07, 7.51, 5.95, 6.83, 8.80, 4.19, 7.44, 7.06, 9.67, 6.89)
n <- length(uzorak)
sigma <- 2

#moguce vrijednosti za nepoznati parametar
mu_0 <- seq(4, 9.5, by= 0.5)

#funkcija za racunanje funkcije vjerodostojnosti
vjerodostojnost <- function(x, mu, sigma)
{
  return(exp(-(x-mu)^2/(2*sigma^2)))
}

apriori <- rep(1/12, 12)
#apriori vjerojatnosti jednake su za svaku vrijednost mu_0

#graf apriorne distribucije
plot(mu_0, apriori, pch = 18)

aposteriori <- c()
for(i in 1:length(uzorak))
{
  vjer_x_apriori <- c()
  for(j in 1:length(mu_0))
  {
    vjer_x_apriori[j] <- vjerodostojnost(uzorak[i], mu_0[j], sigma)*
    apriori[j]
  }
  suma <- sum(vjer_x_apriori)
  for(j in 1:length(mu_0))
  {
    aposteriori[j] <- vjer_x_apriori[j]/suma
  }
  apriori <- aposteriori
}

#graf aposteriorne distribucije
plot(mu_0, aposteriori, pch = 18)


uzorak2 <- c(6.22, 3.99, 3.67, 6.35, 7.89, 6.13)
n2 <- length(uzorak2)

apriori <- aposteriori
for(i in 1:length(uzorak2))
```

```
{
  vjer_x_apriori <- c()
  for(j in 1:length(mu_0))
  {
    vjer_x_apriori[j] <- vjerodostojnost(uzorak2[i], mu_0[j], sigma)*
   apriori[j]
  }
  suma <- sum(vjer_x_apriori)
  for(j in 1:length(mu_0))
  {
    aposteriori[j] <- vjer_x_apriori[j]/suma
  }
  apriori <- aposteriori
}

#graf aposteriorne distribucije nakon dodanog drugog uzorka
plot(mu_0, aposteriori, pch = 18)

#_______________________

#racunanje vjerodostojnosti pomocu aritmeticke sredine uzorka
uzorak <- c(3.07, 7.51, 5.95, 6.83, 8.80, 4.19, 7.44, 7.06, 9.67, 6.89)
n <- length(uzorak)
sigma <- 2
mean <- mean(uzorak)

mu_0 <- seq(4, 9.5, by= 0.5)

vjerodostojnost2 <- function(mu, sigma, n)
{
  return(exp(-(mean-mu)^2/(2*sigma^2/n)))
}

apriori <- rep(1/12, 12)

vjer_x_apriori <- c()
for(j in 1:length(mu_0))
{
  vjer_x_apriori[j] <- vjerodostojnost2(mu_0[j], sigma, n)*apriori[j]
}
suma <- sum(vjer_x_apriori)
aposteriori <- c()
for(j in 1:length(mu_0))
{
  aposteriori[j] <- vjer_x_apriori[j]/suma
}
plot(mu_0, aposteriori, pch = 18)
```

```r


uzorak2 <- c(6.22, 3.99, 3.67, 6.35, 7.89, 6.13)
n2 <- length(uzorak2)
apriori <-aposteriori
mean <- mean(uzorak2)

vjer_x_apriori <- c()
for(j in 1:length(mu_0))
{
  vjer_x_apriori[j] <- vjerodostojnost2(mu_0[j], sigma, n2)*apriori[j]
}
suma <- sum(vjer_x_apriori)
aposteriori <- c()
for(j in 1:length(mu_0))
{
  aposteriori[j] <- vjer_x_apriori[j]/suma
}
plot(mu_0, aposteriori, pch = 18)

#__________________________________________
#oba uzorka kao jedan, vjerodostojnost pomocu aritmeticke sredine uzorka

uzorak <- c(3.07, 7.51, 5.95, 6.83, 8.80, 4.19, 7.44, 7.06, 9.67, 6.89, 6.22, 3.99, 3.67, 6.35, 7.89, 6.13)
n <- length(uzorak)
sigma <- 2
mean <- mean(uzorak)
mu_0 <- seq(4, 9.5, by= 0.5)

vjerodostojnost2 <- function(mu, sigma, n)
{
  return(exp(-(mean-mu)^2/(2*sigma^2/n)))
}

apriori <- rep(1/length(uzorak), length(uzorak))

vjer_x_apriori <- c()
for(j in 1:length(mu_0))
{
  vjer_x_apriori[j] <- vjerodostojnost2(mu_0[j], sigma, n)*apriori[j]
}
suma <- sum(vjer_x_apriori)
aposteriori <- c()
for(j in 1:length(mu_0))
{
  aposteriori[j] <- vjer_x_apriori[j]/suma
```

```
}
plot(mu_0, aposteriori, pch = 18)

#_______________________________________________
uzorak <- c(3.07, 7.51, 5.95, 6.83, 8.80, 4.19, 7.44, 7.06, 9.67, 6.89,
    6.22, 3.99, 3.67, 6.35, 7.89, 6.13)
x <- mean(uzorak)

sigma <- 2


mu_0 <- seq(4, 9.5, by= 0.5)

apriori <- rep(1/12, 12)

vjerodostojnost <- function(x, mu, sigma)
{
  return(exp(-(x-mu)^2/(2*sigma^2)))
}

vjer_x_apriori <- c()
for(j in 1:length(mu_0))
{
  vjer_x_apriori[j] <- vjerodostojnost(x, mu_0[j], sigma)*apriori[j]
}
suma <- sum(vjer_x_apriori)
aposteriori <- c()
for(j in 1:length(mu_0))
{
  aposteriori[j] <- vjer_x_apriori[j]/suma
}
plot(mu_0, aposteriori, pch = 18)
```

R kod za Primjer 5.0.2

```r
sigma <- 2
m <- 20
s <- 5

uzorak <- c( 26.8, 26.3, 28.03, 28.5, 26.3, 31.9, 28.5, 27.2, 20.9,
    27.5, 28.0, 18.6, 22.3, 25.0, 31.5)
mean <- mean(uzorak)
n <- length(uzorak)

s_apost <- function(s, n, sigma) sqrt(1/(1/s^2 + n/sigma^2))

m_apost <- function(s, n, sigma, m, mean) (1/s^2)/(n/sigma^2+1/s^2)*m +
    (n/sigma^2)/(n/sigma^2+1/s^2)*mean

s_apost(s, n, sigma)
m_apost(s, n, sigma, m, mean)

#parametri aposterirone distribucije su sada m_apost i (s_apost)^2

t <- seq(10, 30, by=0.1)
plot(t, dnorm(t, 20, 5), type="l", ann=FALSE, ylim=c(0,0.8))
lines(t, dnorm(t, m_apost(s, n, sigma, m, mean), s_apost(s, n, sigma)),
    lty="dashed")
legend(10, 0.8, legend=c("apriorna distribucija", "aposteriorna
    distribucija"), lty=1:2, cex=0.8)

#__________________

sigma <- 2
m <- 30
s <- 4

uzorak <- c( 26.8, 26.3, 28.03, 28.5, 26.3, 31.9, 28.5, 27.2, 20.9,
    27.5, 28.0, 18.6, 22.3, 25.0, 31.5)
mean <- mean(uzorak)
n <- length(uzorak)

s_apost <- function(s, n, sigma) sqrt(1/(1/s^2 + n/sigma^2))

m_apost <- function(s, n, sigma, m, mean) (1/s^2)/(n/sigma^2+1/s^2)*m +
    (n/sigma^2)/(n/sigma^2+1/s^2)*mean

s_apost(s, n, sigma)
m_apost(s, n, sigma, m, mean)
```

```
#parametri aposterirone distribucije su sada m_apost i (s_apost)^2

t <- seq(18, 40, by=0.1)
plot(t, dnorm(t, 30, 4), type="l", ann=FALSE, ylim=c(0,0.8))
lines(t, dnorm(t, m_apost(s, n, sigma, m, mean), s_apost(s, n, sigma)),
    lty="dashed")
legend(18, 0.8, legend=c("apriorna distribucija", "aposteriorna
    distribucija"), lty=1:2, cex=0.8)

#__________________

t <- seq(23,30, by=0.1)
plot(t, dnorm(t, m_apost(5, n, sigma, 20, mean), s_apost(5, n, sigma)),
    type="l", ann=FALSE, ylim=c(0,0.8))
lines(t, dnorm(t, m_apost(4, n, sigma, 30, mean), s_apost(4, n, sigma)),
     lty="dashed")
legend(23, 0.8, legend=c("N(26.4, 0.3)", "N(26.5, 0.3)"), lty=1:2, cex
    =0.8)
```

R kod za Primjer 5.0.3

```
library('Bolstad')
library('sfsmisc')

uzorak <- c(4.09, 4.68, 1.87, 2.62, 5.58, 8.68, 4.07, 4.78, 4.79, 4.49,
    5.85, 5.9, 2.4, 6.27, 6.3, 4.47)
n <- length(uzorak)
sigma <- 2

mu_0 <- seq(0, 10, length.out = 100)

apriori <- c(0)
for(i in 1:length(mu_0))
{
  if(mu_0[i] > 0 & mu_0[i] <=3) apriori[i] <- mu_0[i]
  if(mu_0[i] > 3 & mu_0[i] <=5) apriori[i] <- 3
  if(mu_0[i] > 5 & mu_0[i] <=8) apriori[i] <- 8 - mu_0[i]
  if(mu_0[i] > 8) apriori[i] <- 0
}
plot(mu_0, apriori)

aposteriori <- normgcp( uzorak, sigma.x = 2, density = "user", n.mu =
    100, mu = mu_0, mu.prior = apriori)
#u varijabli aposteriori se nalaze vrijednosti funkcije vjerodostojnosi,
     aposteriorne distribucije te zadane moguce vrijedosti parametra mu
    kao i apriorna distribucija

y <- aposteriori$posterior

m <- integrate.xy(mu_0, mu_0*y, 0, 10, use.spline=TRUE, xtol=2e-08)
s2 <- integrate.xy(mu_0, (mu_0-m)^2*y, 0, 10, use.spline=TRUE, xtol=2e
    -08)

plot(mu_0, y, type="l", xlim=c(3,6.5), ann=FALSE)
lines(mu_0, dnorm(mu_0, m, sqrt(s2)), lty="dashed")
legend(3, 0.8, legend=c("aposteriorna distribucija", expression(N(m, (s)
    ^2))), lty=1:2, cex=0.8)

interval <- c(m - 1.96*sqrt(s2) , m + 1.96*sqrt(s2))

integrate.xy(mu_0, y, interval[1], interval[2], use.spline=TRUE, xtol=2e
    -08)
integrate.xy(mu_0, y, interval[1]-0.003, interval[2]+0.003, use.spline=
    TRUE, xtol=2e-08)
```

Sažetak

U ovom radu opisuju se metode za zaključivanje o vrijednosti parametra očekivanja normalne distribucije. Na početku se navode osnovni pojmovi teorije vjerojatnosti i statistike kako bi se izgradila teorija potrebna za uvođenje Bayesovog teorema i osnova Bayesovskog zaključivanja temeljenih na teoremu. Uvode se diskretna i neprekidna verzija Bayesovog teorema, zajedno s dokazima, te pojmovi usko vezani uz primjenu teorema poput apriorne i aposteriorne distribucije. Za kraj uvodnog dijela objašnjava se ideja Bayesovskog zaključivanja.

U nastavku se obrađuju metode zaključivanja, temeljene na diskretnoj verziji Bayesovog teorema, u slučaju kada je skup mogućih vrijednosti nepoznatog parametra diskretan. Zatim se obrađuju i metode temeljene na neprekidnoj verziji Bayesovog teorema za slučajeve kada pretpostavljamo da nepoznati parametar može biti jednak bilo kojoj vrijednosti unutar nekog intervala. Uvodi se i pojam Bayesovog intervala vjerodostojnosti za parametar očekivanja, koji je analogon standardnog frekvencionističkog intervala pouzdanosti, te njegova interpretacija u Bayesovskoj statistici te metode računanja intervala.

Nadalje, opisuju se razlike Bayesovskog i frekvencionističkog pristupa. Razmatraju se standardne metode za točkovnu procjenu, intervalnu procjenu i testiranje hipoteza te se uspoređuju s odgovarajućim Bayesovskim metodama.

Konačno, navedene metode se primjenjuju na odabranim zadacima koji se rješavaju u programskom jeziku R.

Summary

In this paper, we describe methods of Bayesian inference for mean of normal distribution. First, we introduce basic concepts and definitions of probability theory and statistics for the purpose of building theory needed to introduce and prove Bayes theorem and the basics of Bayesian inference. We introduce discrete and continuous version of Bayes theorem, along with their proofs, as well as concepts regarding theorem's use such as prior and posterior distribution. At the end of the introductory part, we explain the idea of Bayesian inference that is based on the theorem.

We continue with methods of inference, based on the discrete version of Bayes theorem, that is used in cases when the set of possible values for the mean is discrete. Next, we introduce the methods of inference based on the continuous version of Bayes theorem, used under the assumption that the mean can be equal to any value in an interval. We also introduce the concept of Bayesian credible interval for normal mean as well as its interpretation in Bayesian statistics and methods for its computation.

Furthermore, we explain the diffrences between Bayesian and frequentist inference. We compare frequentist and Bayesian point esitmators, intervals and testing hypothesis.

Finally, we use the explained methods on selected problems and solve the problems in programming language R.

Životopis

Rođena sam 19. veljače 1997. u Zagrebu, gdje sam odrasla i trenutno živim. Nakon završene Osnovne škole Matije Gupca, pohađam X. gimnaziju "Ivan Supek" u Zagrebu. 2015. godine školovanje nastavljam upisujući preddiplomski sveučilišni studij Matematika na Matematičkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu, a 2019. godine upisujem diplomski sveučilišni studij Financijska i poslovna matematika.

U sklopu osnovnoškolskog i srednjoškolskog obrazovanja učim engleski, njemački i francuski jezik, a izvan škole pohađam i sate talijanskog jezika. Također, od 2013. godine volontiram u Odredu izviđača "Plamen", te u sklopu toga pohadam tečajeve za dodatnu edukaciju.

Od ljeta 2021. godine radim kao student u Privrednoj banci Zagreb gdje se bavim razvojem internih kreditnih modela.