Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Matematički odsjek

Miljenko Huzak

Matematička statistika

Predavanja

2020./2021.

Poglavlje 1

Slučajne varijable i vektori

Navodimo osnovne pojmove i rezultate teorije vjerojatnosti koji su velikim dijelom sadržani u kolegijima Vjerojatnost *i* Statistika *s Preddiplomskog studija matematike na Matematičkom odsjeku PMF-a u Zagrebu. Manji dio sadržaja koji se spominje obrađuje se u puno većem opsegu kroz predmete* Teorija vjerojatnosti 1 *i* 2 *s diplomske razine studija na istome fakultetu. Pretpostavlja se da je student upoznat s osnovnim pojmovima i rezultatima teorije mjere i integriranja obuhvaćenim sadržajem predmeta* Mjera i integral *s Preddiplomskog studija matematike na Matematičkom odsjeku PMF-a u Zagrebu.*

Neka je $(\Omega, \mathcal{F})$ bilo koji izmjeriv prostor, a $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ izmjeriv prostor sa σ-algebrom Borelovih skupova (vidjeti [9], poglavlje 8.1.) u euklidskom prostoru $\mathbb{R}^k$ za $k \geq 1$ prirodan broj. Nadalje, neka je $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$. Tada označimo sa

$$\langle -\infty, x] := \langle -\infty, x_1] \times \langle -\infty, x_2] \times \cdots \times \langle -\infty, x_k],$$

podskup od $\mathbb{R}^k$. Ukoliko je $k = 1$, $\langle -\infty, x]$ je odozdo neomeđen segment. Za $k \geq 2$ radi se o jednostrano neomeđenom (k-dimenzionalnom) pravokutniku. Nadalje, neka su $a = (a_1, a_2, \ldots, a_k)$, $b = (b_1, b_2, \ldots, b_k) \in \mathbb{R}^k$ i neka je $\leq$ parcijalni uređaj definiran na $\mathbb{R}^k$ sa

$$a \leq b \text{ ako je } (\forall i)\ a_i \leq b_i.$$

Relacija $\leq$ je relacija potpunog uređaja jedino za $k = 1$. Primijetimo da je

$$\{X \leq x\} = X^{-1}(\langle -\infty, x]) = \{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k\}.$$

Definicija 1.1. *Kažemo da je $X : \Omega \to \mathbb{R}^k$* **$k$-dimenzionalna slučajna veličina** *ukoliko je X izmjerivo preslikavanje u paru σ-algebri $(\mathcal{F}, \mathcal{B}(\mathbb{R}^k))$, tj. ako vrijedi da je*

$$(\forall B \in \mathcal{B}(\mathbb{R}^k))\ \{X \in B\} \in \mathcal{F}.$$

Ako je $k = 1$, X zovemo **slučajna varijabla***, a ako je $k \geq 2$, X zovemo* **slučajni vektor***.*

Prema teoremima 8.1 i 8.2 iz [9], str. 235. i 237., X je k-dimenzionalna slučajna veličina ako i samo ako vrijedi:

$$(\forall x \in \mathbb{R}^k)\ \{X \leq x\} \in \mathcal{F}.$$

Može se pokazati da je $X = (X_1, X_2, \ldots, X_k) : \Omega \to \mathbb{R}^k$ slučajni vektor ako i samo ako su sve komponente X_1, X_2,..., X_k slučajne varijable (vidjeti [9], str. 238., propozicija 8.4). Nadalje, ako su X_1, X_2,..., X_k slučajne varijable definirane na istome izmjerivom prostoru $(\Omega, \mathcal{F})$ i $g : \mathbb{R}^k \to \mathbb{R}$ je Borelova funkcija, tada je kompozicija slučajnog vektora $X =$

$(X_1, X_2, \dots, X_k)$ i funkcije g, u oznaci $g(X_1, X_2, \dots, X_k) =: g \circ X : \Omega \to \mathbb{R}$, slučajna variabla (vidjeti [9], str. 239., propozicija 8.7). Naime, kompozicija izmjerivih preslikavanja je izmjerivo preslikavanje. Podsjetimo se, $g : \mathbb{R}^k \to \mathbb{R}$ je *Borelova funkcija* ako je izmjeriva u paru σ-algebri $(\mathcal{B}(\mathbb{R}^k), \mathcal{B}(\mathbb{R}))$ (vidjeti [9], str. 236. i 239.).

Primjer 1.2. Neka je A događaj na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ i $X := \mathbb{1}_A$ neka je *indikatorska funkcija od A*. Indikatorska funkcija skupa A, $\mathbb{1}_A$, je takva funkcija da je $\mathbb{1}_A(\omega) = 1$ ako je $\omega \in A$ i $\mathbb{1}_A(\omega) = 0$ za $\omega \in A^c$. Tada je X slučajna varijabla. Naime, za proizvoljan Borelov skup B u $\mathbb{R}$,

$$\{X \in B\} = \begin{cases} \Omega & \text{ako } 0 \in B \text{ i } 1 \in B, \\ A & \text{ako } 0 \notin B \text{ i } 1 \in B, \\ A^c & \text{ako } 0 \in B \text{ i } 1 \notin B, \\ \emptyset & \text{ako } 0 \notin B \text{ i } 1 \notin B. \end{cases}$$

U sva četiri slučaja je $\{X \in B\} \in \mathcal{F}$, dakle, događaj. Budući da je B bio proizvoljan Borelov skup, pokazana tvrdnja vrijedi za svaki $B \in \mathcal{B}(\mathbb{R})$. Prema definiciji 1.1, X je slučajna varijabla. Općenitije, neka su A_1, A_2,..., A_n događaji i α_1, α_2,..., α_n realni brojevi. Tada je $X := \alpha_1 \mathbb{1}_{A_1} + \alpha_2 \mathbb{1}_{A_2} + \cdots + \alpha_n \mathbb{1}_{A_n}$ slučajna varijabla. Svaku slučajnu varijablu oblika X, dakle takvu koja se može prikazati kao (konačna) linearna kombinacija indikatorskih funkcija događaja, zovemo *jednostavna slučajna varijabla*. Pokažimo izmjerivost od X. Prvo, primijetimo da se X može prikazati u sljedećem obliku:

$$X = 0 \cdot \mathbb{1}_{H_0} + a_1 \mathbb{1}_{H_1} + a_2 \mathbb{1}_{H_2} + \cdots + a_m \mathbb{1}_{H_m},$$

gdje su $a_0 := 0$, a_1, a_2,..., a_m međusobno različiti realni brojevi, a H_0, H_1,..., H_m čine konačnu *particiju* od Ω i događaji su. Naime, za svaki i, a_i je jednak nekom od elemenata skupa $\{0, \alpha_1, \dots, \alpha_n\}$ ili zbroju nekih elemenata tog skupa, a budući da je $H_i = \{X = a_i\}$, jednak je nekoj konačnoj uniji međusobno disjunktnih presjeka događaja A_1, A_2,..., A_n ili njihovih komplemenata $(0 \le i \le m)$. Neka je sada B proizvoljan Borelov skup. Tada, ili B ne sadrži niti jedan broj iz skupa $S := \{0, a_1, \dots, a_m\}$, ili sadrži točno k $(k \ge 1)$ njih, recimo da su to a_{i_1}, a_{i_2},..., a_{i_k} $(0 \le i_1 < \cdots < i_k \le m)$. Tada je

$$\{X \in B\} = \begin{cases} H_{i_1} \cup H_{i_2} \cup \cdots \cup H_{i_k} & \text{ako je } B \cap S = \{a_{i_1}, a_{i_2}, \dots, a_{i_k}\}, \\ \emptyset & \text{ako je } B \cap S = \emptyset. \end{cases}$$

Jasno je da je u oba slučaja $\{X \in B\} \in \mathcal{F}$. Dakle, s istim argumentom kao prije zaključujemo da je X slučajna varijabla. $\square$

U teoriji vjerojatnosti i općenito u teoriji mjere, jednostavne slučajne varijable imaju značajnu ulogu jer se, na primjer, matematičko očekivanje jednostavne slučajne varijable jednostavno računa, a svaka se slučajna varijabla može prikazati kao limes jednostavnih. Naime, vrijedi sljedeći teorem (za dokaz vidjeti [9], teorem 8.5 i korolar 8.3, str. 242.-243.).

Teorem 1.3. *Neka je X slučajna varijabla definirana na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Tada postoji niz $(X_n)_{n \in \mathbb{N}}$ jednostavnih slučajnih varijabli takvih da je za sve $\omega \in \Omega$ i sve $n \in \mathbb{N}$, $|X_n(\omega)| \le |X(\omega)|$ i $\lim_n X_n(\omega) = X(\omega)$. Specijalno, ako je X nenegativna slučajna varijabla, tada postoji niz jednostavnih slučajnih varijabli $(X_n)_{n \in \mathbb{N}}$ takvih da je za sve $\omega \in \Omega$ i sve $n \in \mathbb{N}$, $0 \le X_n(\omega) \le X_{n+1}(\omega) \le X(\omega)$ i $\lim_n X_n(\omega) = X(\omega)$.*

Ta tvrdnja je osnova za tzv. *induktivnu metodu* u teoriji mjere i često se primijenjuje.

Zadatak 1.4. Neka su X_1, X_2,..., X_n slučajne varijable definirane na istome izmjerivom prostoru.Tada su $Y := \min_{1 \le i \le n} X_i$ i $Z := \max_{1 \le i \le n} X_i$ slučajne varijable. Dokažite.

1.1 Zakon razdiobe slučajnih varijabli i vektora

Pretpostavimo da je definirana vjerojatnost $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ na izmjerivom prostoru $(\Omega, \mathcal{F})$. Dakle, zadan je vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$.

Definicija 1.5. *Neka je $X : \Omega \to \mathbb{R}^k$ k-dimenzionalna slučajna veličina definirana na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Induciranu vjerojatnost $\mathbb{P}_X$, definiranu na izmjerivom prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ relacijom*

$$\mathbb{P}_X(B) := \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B)),\ B \in \mathcal{B}(\mathbb{R}^k),$$

zovemo **zakon razdiobe od** X.

Očito je $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_X)$ vjerojatnosni prostor kojega zovemo **inducirani vjerojatnosni prostor slučajnom veličinom** (**vektorom** ili **varijablom**) X. Nas će u statistici zanimati samo takav vjerojatnosni prostor.

Zakone razdioba slučajnih varijabli ili vektora možemo karakterizirati realnim funkcijama koje zovemo funkcijama razdioba ili funkcijama distribucija.

Definicija 1.6. *Neka je X k-dimenzionalna slučajna veličina sa zakonom razdiobe $\mathbb{P}_X$. Funkciju $F_X : \mathbb{R}^k \to \mathbb{R}$ definiranu sa*

$$F_X(x) := \mathbb{P}_X(\langle -\infty, x]),\ x \in \mathbb{R}^k,$$

zovemo **funkcija razdiobe (ili distribucije) od** X.

Primjetimo da iz definicija 1.6 i 1.5 slijedi da je za sve $x \in \mathbb{R}^k$,

$$F_X(x) = \mathbb{P}(X \leq x). \tag{1.1}$$

Vrijedi sljedeći teorem (za dokaz vidjeti [9], str. 251., teorem 9.1).

Teorem 1.7. *Funkcija razdiobe $F \equiv F_X$ slučajne varijable X je*

(1) *neprekidna zdesna,*
(2) *neopadajuća i*
(3) *takva da je $F(-\infty) \equiv \lim_{x \to -\infty} F(x) = 0$ i $F(+\infty) \equiv \lim_{x \to +\infty} F(x) = 1$.*

Iz teorema 1.7 slijedi da funkcija razdiobe F_X slučajne varijable X u svakoj točki $x \in \mathbb{R}$ ima limes slijeva:

$$F_X(x-) \equiv \lim_{y \to x-} F_X(y).$$

Primijetimo da zbog konačne aditivnosti vjerojatnosti, za svaki poluotvoreni interval $\langle a, b]$ (gdje su $a, b \in \mathbb{R}$ takvi da je $a < b$) vrijedi

$$\mathbb{P}_X(\langle a, b]) = \mathbb{P}_X(\langle -\infty, b] \setminus \langle -\infty, a]) = \mathbb{P}_X(\langle -\infty, b]) - \mathbb{P}_X(\langle -\infty, a]) = F_X(b) - F_X(a).$$

Zbog nenegativnosti vjerojatnosti, odavde odmah slijedi da je F_X neopadajuća funkcija. Naime, $a < b$ povlači

$$F_X(b) - F_X(a) = \mathbb{P}_X(\langle a, b]) \geq 0. \tag{1.2}$$

Slična svojstva ima i funkcija razdiobe slučajnog vektora. Da bismo ih u potpunosti mogli iskazati, uvodimo operator diferenciranja po komponentama argumenta funkcije. Neka je $g : \mathbb{R}^k \to \mathbb{R}$ funkcija, a, $b \in \mathbb{R}^k$ bilo koje točke te a_i i b_i njihove i-te komponente $(1 \leq i \leq k)$. Slijedeći oznake iz knjige [9], str. 269., neka je

$$\Delta_{b_i - a_i} g(a) := g(a_1, \ldots, a_{i-1}, b_i, a_{i+1} \ldots, a_k) - g(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_k)$$

i iterativno za $i \geq 2$,

$$\begin{aligned} & \Delta_{b_{i-1}-a_{i-1}}(\Delta_{b_i-a_i} g(a)) := \\ = \ & \Delta_{b_{i-1}-a_{i-1}} g(a_1, \ldots, a_{i-1}, b_i, a_{i+1} \ldots, a_k) - \Delta_{b_{i-1}-a_{i-1}} g(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_k). \end{aligned}$$

Sada definiramo broj

$$\Delta_{b-a} g(a) := \Delta_{b_1-a_1}(\Delta_{b_2-a_2}(\cdots \Delta_{b_k-a_k} g(a) \cdots)).$$

Na sličan način kako je definiran parcijalni uređaj $\leq$ u $\mathbb{R}^k$, definiramo relaciju $<$ u $\mathbb{R}^k$:

$$a < b \ \text{ ako je } \ (\forall i) \ a_i < b_i.$$

Teorem 1.8. *Funkcija razdiobe $F \equiv F_X$ k-dimenzionalnog slučajnog vektora X je*

(1) *neprekidna zdesna,*
(2) *za svaki $a, b \in \mathbb{R}^k$ takve da je $a < b$, vrijedi $\Delta_{b-a} F(a) \geq 0$, i*
(3) *takva da je za bilo koji $i \in \{1, 2, \ldots, k\}$, $\lim_{x_i \to -\infty} F(x) = 0$ i $\lim_{\forall i, \, x_i \to +\infty} F(x) = 1$.*

Prvo, primijetimo da je u slučaju dimenzije $k = 1$, za realne brojeve $a < b$, $\Delta_{b-a} F_X(a) = F_X(b) - F_X(a)$ što je nenegativno jer je prema svojstvu (2) teorema 1.7, funkcija razdiobe slučajne varijable neopadajuća funkcija. U tom smislu je svojstvo (2) iz teorema 1.7 zapravo svojstvo (2) iz teorema 1.8. Slično, svojstvo (3) iz teorema 1.7 je također svojstvo (3) iz teorema 1.8.

Relacija (1.2) nam daje interpretaciju izraza $\Delta_{b-a} F_X(a)$ za jednodimenzijski slučaj. Pogledajmo na primjeru kada je dimenzija $k = 2$, koja je interpretacija istoga izraza u višedimenzijskim slučajevima. Na taj način ćemo i dokazati svojstvo (2) iz prethodnog teorema. Neka su $a = (a_1, a_2)$, $b = (b_1, b_2) \in \mathbb{R}^2$ takve točke da je $a < b$. Tada je, koristeći se definicijama operatora diferenciranja i funkcije razdiobe, te konačnom aditivnošću vjerojatnosti,

$$\begin{aligned} \Delta_{b-a} F_X(a) &= \Delta_{b_1-a_1}(\Delta_{b_2-a_2} F_X(a)) = \\ &= \Delta_{b_1-a_1} F_X(a_1, b_2) - \Delta_{b_1-a_1} F_X(a_1, a_2) = \\ &= (F_X(b_1, b_2) - F_X(a_1, b_2)) - (F_X(b_1, a_2) - F_X(a_1, a_2)) = \\ &= (\mathbb{P}_X(\langle -\infty, b_1] \times \langle -\infty, b_2]) - \mathbb{P}_X(\langle -\infty, a_1] \times \langle -\infty, b_2])) - \\ & \quad -(\mathbb{P}_X(\langle -\infty, b_1] \times \langle -\infty, a_2]) - \mathbb{P}_X(\langle -\infty, a_1] \times \langle -\infty, a_2])) = \\ &= \mathbb{P}_X(\langle a_1, b_1] \times \langle -\infty, b_2]) - \mathbb{P}_X(\langle a_1, b_1] \times \langle -\infty, a_2]) = \\ &= \mathbb{P}_X(\langle a_1, b_1] \times \langle a_2, b_2]) = \mathbb{P}_X(\langle a, b]). \end{aligned}$$

Dakle,

$$\Delta_{b-a} F_X(a) = \mathbb{P}_X(\langle a, b]) \tag{1.3}$$

u slučajevima $1 \leq k \leq 2$, a na sličan način može se pokazati da (1.3) vrijedi i za $k > 2$, dakle, za svaki prirodan broj k. Budući da je vjerojatnost nenegativna mjera, $\Delta_{b-a} F_X(a) = \mathbb{P}_X(\langle a, b]) \geq 0$ pa slijedi svojstvo (2) iz teorema 1.8. Dakako, ovdje je općenito za $a = (a_1, \ldots, a_k) < b = (b_1, \ldots, b_k)$,

$$\langle a, b] = \langle a_1, b_1] \times \langle a_2, b_2] \times \cdots \times \langle a_k, b_k]$$

k-dimenzionalni poluotvoreni pravokutnik.

1.2 Neprekidne i diskretne slučajne varijable i vektori

Neka je $\lambda \equiv \lambda^k = \lambda^1 \times \cdots \times \lambda^1$ Lebesgueova mjera definirana na izmjerivom prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ (vidjeti [9], str. 257. i 271. i dodatak A).

Definicija 1.9. *Kažemo da je k-dimenzionalna slučajna veličina X* **apsolutno neprekidna** *ili, kraće,* **neprekidna** *ako postoji nenegativna Borelova funkcija f_X definirana na $\mathbb{R}^k$ takva da se funkcija razdiobe F_X može prikazati na sljedeći način:*

$$F_X(x) = \int\limits_{\langle -\infty, x]} f_X(y)\, d\lambda(y), \ x \in \mathbb{R}^k.$$

Funkciju f_X zovemo **funkcija gustoće razdiobe od** X *ili, kraće,* **gustoća od** X.

Iz definicije 1.9 i svojstava funkcija razdioba odmah slijedi da za gustoću f neprekidne slučajne veličine vrijedi relacija

$$\int\limits_{\mathbb{R}^k} f(x)\, d\lambda(x) = 1. \tag{1.4}$$

Nadalje, funkcija razdiobe neprekidne k-dimenzionalne slučajne veličine je neprekidna funkcija kao funkcija s euklidskog prostora $\mathbb{R}^k$ u euklidski prostor $\mathbb{R}$.

Primijetimo da za neprekidnu k-dimenzionalnu slučajnu veličinu X s gustoćom f_X vrijedi

$$(\forall B \in \mathcal{B}(\mathbb{R}^k)) \ \mathbb{P}_X(B) = \int\limits_B d\mathbb{P}_X = \int\limits_B f_X(x)\, d\lambda(x). \tag{1.5}$$

Odavde slijedi da, ako je B takav Borelov skup da je $\lambda(B) = 0$, tada je nužno i $\mathbb{P}_X(B) = 0$. Dakle, zakon razdiobe neprekidne slučajne varijable (ili vektora) je *apsolutno neprekidan u odnosu na Lebesgueovu mjeru.* Pri tome je Radon-Nikodymova derivacija $d\mathbb{P}_X/d\lambda$ zakona razdiobe $\mathbb{P}_X$ od X u odnosu na Lebesgueovu mjeru (vidjeti [9], str. 297.), jednaka gustoći. Naime, prema Radon-Nikodymovom teoremu (vidjeti [9], str. 297., teorem 10.5), gustoća od X je u odnosu na λ *skoro svuda*, ili kraće, λ-s.s., jedinstvena, tj.

$$\lambda(\{x \in \mathbb{R}^k \ : \ f_X(x) \neq \frac{d\mathbb{P}_X}{d\lambda}(x)\}) = 0.$$

Dakle, ako gustoći neprekidne slučajne varijable ili vektora X promijenimo vrijednost u, na primjer, prebrojivo mnogo točaka, modificirana funkcija će također biti gustoća od X.

Navest ćemo teorem (vidjeti [9], str. 273.-274. i [7] IV, §4.5, zadaci 49. i 50., str. 106.) koji se često koristi u primjenama. Prije iskaza, dogovorno ćemo reći da je neka funkcija $g : \mathbb{R}^k \to \mathbb{R}$ *skoro svuda* neprekidna *u odnosu na Lebesgueovu mjeru* ako je neprekidna u svakoj točki domene osim možda na podskupu koji je Lebesgueove mjere nula. Kraće kažemo da je g λ-s.s. neprekidna. Slično, kažemo da je g *skoro svuda* k-puta neprekidno diferencijabilna *u odnosu na Lebesgueovu mjeru* (ili λ-s.s. k-puta neprekidno diferencijabilna) ako je klase C^{k-1} na čitavoj domeni i klase C^k svuda na domeni osim na podskupu koji je Lebesgueove mjere nula.

Teorem 1.10. *Ako je gustoća $f = f_X$ neprekidne k-dimenzionalne slučajne veličine X skoro svuda neprekidna u odnosu na Lebesgueovu mjeru, tada je funkcija razdiobe od X, F_X, skoro svuda k-puta neprekidno diferencijabilna u odnosu na Lebesgueovu mjeru i vrijedi*

$$f(x_1, x_2, \ldots, x_k) = \frac{\partial^k F_X(x_1, x_2, \ldots, x_k)}{\partial x_1 \partial x_2 \cdots \partial x_k} \ \text{za } \lambda\text{-s.s. } (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k. \tag{1.6}$$

Obratno, ako je F_X λ-skoro svuda k-puta neprekidno diferencijabilna funkcija i (1.6) *vrijedi za neku nenegativnu Borelovu funkciju $f : \mathbb{R}^k \to \mathbb{R}$ takvu da je* (1.4) *ispunjeno, tada je X neprekidna k-dimenzionalna slučajna veličina s gustoćom f.*

Definicija 1.11. *Slučajna veličina X dimenzije k je* **diskretna** *ako postoji prebrojiv skup $D \subseteq \mathbb{R}^k$ takav da je $\mathbb{P}_X(D) = 1$.*

Na primjer, jednostavne slučajne varijable (vidjeti primjer 1.2) su diskretne. Naime, za skup D iz definicije 1.11 možemo uzeti konačan skup vrijednosti jednostavne slučajne varijable.

Općenito, za neku diskretnu slučajnu veličinu X dimenzije k neka je $D = \{a_1, a_2, \ldots\}$ *najmanji* prebrojiv skup takav da je $\mathbb{P}_X(D) = 1$. To znači da za (međusobno različite) elemente a_1, a_2,... od D vrijedi da je $p_i := \mathbb{P}_X(\{a_i\}) = \mathbb{P}(X = a_i) > 0$ za sve $i \geq 1$. Tada je $D \subseteq X(\Omega)$. Iako se D ne mora podudarati sa slikom od X, često kažemo da je D skup (bitnih) vrijednosti od X. Definirajmo mjeru μ_D na izmjerivom prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ na sljedeći način:

$$\mu_D(B) := \sum_{i \geq 1} \mathbb{1}_B(a_i) = |B \cap D|,$$

gdje je sa $|A|$ označen broj elemenata skupa A. Tu mjeru zovemo *brojeća mjera* u odnosu na D. Nadalje, primijetimo da je zakon razdiobe $\mathbb{P}_X$ od X *apsolutno neprekidan u odnosu na pripadnu brojeću mjeru μ_D*. Budući da je brojeća mjera σ-konačna, prema Radon-Nikodymovom teoremu postoji nenegativna izmjeriva funkcija $f_X : \mathbb{R}^k \to \mathbb{R}$ (Radon-Nikodymova derivacija od $\mathbb{P}_X$ u odnosu na μ_D), takva da je za sve $B \in \mathcal{B}(\mathbb{R}^k)$,

$$\mathbb{P}_X(B) = \int_B f_X(x) d\mu_D(x) = \sum_{a_i \in B} f_X(a_i).$$

S druge strane, zbog $\mathbb{P}_X(B) = \mathbb{P}_X(B \cap D)$ i σ-aditivnosti od $\mathbb{P}_X$ je

$$\mathbb{P}_X(B) = \sum_{a_i \in B} p_i. \tag{1.7}$$

Dakle, za $B = \{a_j\}$, izjednačavanjem prethodne dvije jednakosti slijedi

$$f_X(a_j) = p_j = \mathbb{P}(X = a_j).$$

Budući da je gustoća f_X od X μ_D-s.s. jedinstvena, te da je $\mathbb{P}(X = x) = 0$ za $x \notin D$, slijedi da možemo uzeti da

$$f_X(x) = \mathbb{P}(X = x) = \mathbb{P}_X(\{x\}) \tag{1.8}$$

vrijedi za svaki $x \in \mathbb{R}^k$. Funkciju $f_X : \mathbb{R}^k \to \mathbb{R}$ definiranu formulom (1.8) zovemo **funkcija gustoće razdiobe diskretne k-dimenzionalne slučajne veličine** X ili, kraće, **gustoća diskretne k-dimenzionalne slučajne veličine** X. Njenu restrikciju na D često zovemo i *funkcijom vjerojatnosti od X* ili *funkcijom mase od X*, i poistovjećujemo je sa zakonom razdiobe od X. Zapisujemo ju tabelarno na sljedeći način:

$$\begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}.$$

Primijetimo da za gustoću f_X diskretne k-dimenzionalne slučajne veličine X vrijedi da joj je *nosač*, tj. skup $\text{supp} f_X := \{x \in \mathbb{R}^k : f_X(x) \neq 0\}$, jednak najmanjem prebrojivom skupu D takvom da je $\mathbb{P}_X(D) = 1$.

Općenito, neka je $g : \mathbb{R}^k \to \mathbb{R}$ bilo koja funkcija čiji nosač je prebrojiv skup S. Ako je S konačan skup ili je red $\sum_{x \in S} |g(x)|$ konvergentan, tada je za svaki podskup B od $\mathbb{R}^k$ dobro definirana suma (reda ili konačna)

$$\sum_{x \in B} g(x) := \sum_{x \in B \cap S} g(x).$$

Uz oznake kao gore, iz definicije 1.11 i relacije (1.7), za gustoću $f = f_X$ diskretne k-dimenzionalne slučajne veličine X vrijedi

$$\sum_{x} f(x) = \sum_{x \in D} f(x) = \sum_{i \geq 1} f(a_i) = \sum_{i \geq 1} p_i = 1. \tag{1.9}$$

Nadalje, iz relacije (1.7) i definicije funkcije razdiobe slijedi da je funkcija razdiobe $F = F_X$ diskretne slučajne veličine X jednaka

$$F(x) = \sum_{y \leq x} f(y) = \sum_{a_i \leq x} f(a_i) = \sum_{a_i \leq x} p_i, \quad x \in \mathbb{R}^k. \tag{1.10}$$

Odavde odmah slijedi da je funkcija razdiobe diskretne slučajne veličine po dijelovima konstantna.

Zadatak 1.12. Neka je λ Lebesgueova mjera na $\mathbb{R}$ i neka je $\mathcal{B}([0,1])$ σ-algebra Borelovih skupova na intervalu $[0,1]$. Tada je $([0,1], \mathcal{B}([0,1]), \lambda)$ vjerojatnosni prostor. Pokažite da je funkcija $X : [0,1] \to \mathbb{R}$ definirana sa $X(\omega) := \omega$, $\omega \in [0,1]$, (funkcija identiteta), neprekidna slučajna varijabla s gustoćom jednakom $f(x) = \mathbb{1}_{[0,1]}(x)$, $x \in \mathbb{R}$. Za neprekidnu slučajnu varijablu kojoj je funkcija f gustoća, kažemo da je *uniformno distribuirana na segmentu* $[0,1]$.

1.3 Zadavanje slučajnih varijabli i vektora

Za zadavanje nekog k-dimenzionalnog slučajnog vektora (ili varijable ako je $k = 1$) dovoljno je da imamo zadanu realnu funkciju $F : \mathbb{R}^k \to \mathbb{R}$ sa svojstvima $(1-3)$ iz teorema 1.8. Takvu funkciju F zovemo **vjerojatnosna funkcija razdiobe**.

Što znači zadati slučajnu varijablu ili vektor? To znači da treba konstruirati vjerojatnosti prostor i na njemu definirati slučajnu varijablu ili vektor čija će funkcija razdiobe biti jednaka zadanoj vjerojatnosnoj funkciji razdiobe.

Neka je $F : \mathbb{R}^k \to \mathbb{R}$ vjerojatnosna funkcija razdiobe. Za proizvoljni pravokutnik $\langle a, b] \subseteq \mathbb{R}^k$ definirajmo mjeru $\mathbb{P}_F$ tog pravokutnika na sljedeći način:

$$\mathbb{P}_F(\langle a, b]) := \Delta_{b-a} F(a). \tag{1.11}$$

Zapravo, pokazuje se da je $\mathbb{P}_F$ mjera na množini $\mathcal{S} = \{\langle a, b] \ : \ a, b \in \mathbb{R}^k \text{ i } a < b\}$ svih pravokutnika koja se, pomoću svojstava funkcije F, može proširiti do vjerojatnosti (u istoj oznaci) na σ-algebri Borelovih skupova $\mathcal{B}(\mathbb{R}^k)$ (vidjeti [9], str. 252., teorem 9.2 i str. 270., teorem 9.3). Na taj način dolazimo do vjerojatnosnog prostora $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_F)$. Dakle, $\Omega = \mathbb{R}^k$, $\omega = x \in \Omega$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^k)$ and $\mathbb{P} = \mathbb{P}_F$. Definirajmo sada preslikavanje $X : \Omega \to \mathbb{R}^k$ kao preslikavanje identiteta, $X(\omega) := \omega$. Trivijalno za sve $x \in \mathbb{R}^k$ vrijedi da je

$$\{X \leq x\} = \{\omega \in \Omega = \mathbb{R}^k \ : \ X(\omega) = \omega \leq x\} = \langle -\infty, x] \in \mathcal{B}(\mathbb{R}^k),$$

pa je X k-dimenzionalna slučajna veličina. Nadalje, za proizvoljni $x \in \mathbb{R}^k$ je, prema (1.1), gornjoj jednakosti skupova i činjenici da je $\mathbb{P} = \mathbb{P}_F$, neprekidnosti vjetrojatnosti u odnosu na rastuće događaje, definiciji (1.11) i svojstvu (3) od F,

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}_F(\langle -\infty, x]) = \lim_{n\to\infty} \mathbb{P}_F(\langle -\mathbf{n}, x]) = \lim_{n\to\infty} \Delta_{x-(-\mathbf{n})} F(x) = F(x),$$

gdje je $\mathbf{n} = (n, \ldots, n) \in \mathbb{R}^k$. Dakle, zadana vjerojatnosna funkcija razdiobe F je funkcija razdiobe od X, a $\mathbb{P}_F$ je zakon razdiobe od X. Primijetimo da je u ovakvoj konstrukciji izmjeriv prostor $(\Omega, \mathcal{F})$ uvijek jednak izmjerivom prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ Borelovih skupova na $\mathbb{R}^k$, k je dimenzija slučajne veličine X koja je zadana u svojoj *kanonskoj formi*, kao preslikavanje identiteta $X(\omega) = \omega$ $(\omega \in \Omega = \mathbb{R}^k)$. Zadavanjem vjerojatnosti $\mathbb{P} = \mathbb{P}_F$ na $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ pomoću vjerojatnosne funkcije razdiobe zadajemo odmah i zakon razdiobe $\mathbb{P}_X = \mathbb{P}_F$ od X. Na taj način ista k-dimenzionalna slučajna veličina (kao funkcija) može imati razne zakone razdiobe.

Zadatak 1.13. Ako je $F : \mathbb{R} \to \mathbb{R}$ vjerojatnosna funkcija razdiobe, tada postoji slučajna varijabla X na vjerojatnosnom prostoru iz zadatka 1.12, kojoj je F funkcija razdiobe. (**Uputa:** Za $\omega \in [0,1]$, neka je $X(\omega) := \sup\{x : F(x) < \omega\}$. Primijetite, ako je F strogo rastuća funkcija, da je tada $X = F^{-1}$.)

Budući da ćemo se u nastavku baviti samo neprekidnim i diskretnim slučajnim varijablama i vektorima, opisat ćemo posebno kako se zadaju takve slučajne veličine.

Neka je f nenegativna realna Borelova funkcija definirana na $\mathbb{R}^k$ sa svojstvom (1.4). Takvu funkciju zovemo **vjerojatnosna gustoća neprekidne razdiobe**. Tada je relacijom

$$F(x) := \int_{\langle -\infty, x]} f(y)\, d\lambda(y),\ x \in \mathbb{R}^k, \tag{1.12}$$

definirana vjerojatnosna funkcija razdiobe (to se relativno lako pokaže). Dakle, postoji vjerojatnosni prostor i k-dimenzionalna slučajna veličina X kojoj je F funkcija razdiobe. Iz definicije od F trivijalno slijedi da je X apsolutno neprekidna slučajna veličina, pri čemu je f gustoća od X. Dakle, neprekidne slučajne varijable ili vektore možemo zadavati njihovim gustoćama.

Spomenimo ovdje i jedan tehnički detalj vezan uz zadavanje neprekidnih razdioba. Ukoliko je podintegralna funkcija nenegativna i Riemann-integrabilna na području integracije (koje može biti i neograničeni interval, a integral je tada nepravi Riemannov integral), tada je funkcija integrabilna i u odnosu na Lebesgueovu mjeru, i ta dva integrala se podudaraju (vidjeti, na primjer za slučaj jedne dimenzije, [9], str. 305., teorem 10.10). U primjenama se gotovo uvijek neprekidne distribucije zadaju vjerojatnosnim gustoćama koje su po dijelovima neprekidne (zbog toga su Borelove) i imaju (neprave) Riemannove integrale na $\mathbb{R}^k$ jednake 1.

U nastavku pokažimo kako zadajemo diskretne slučajne varijable ili vektore. Neka je zadan prebrojiv podskup $D = \{a_1, a_2, \ldots\}$ (konačan ili beskonačan) od $\mathbb{R}^k$ i preslikavanje $P : D \to \mathbb{R}$ takvo da za elemente $p_i := P(a_i)$ $(i \geq 1)$ slike od P vrijedi da su pozitivni brojevi manji od jedan, te da je $\sum_{i\geq 1} p_i = 1$. Drugim riječima, P je **vjerojatnosna funkcija mase**. Tada je formulom

$$f(x) := \sum_{i\geq 1} p_i \cdot \mathbb{1}_{\{a_i\}}(x),\quad x \in \mathbb{R}^k,$$

zadana gustoća neke k-dimenzionalne diskretne slučajne veličine. Naime, može se pokazati da je funkcija $F : \mathbb{R}^k \to \mathbb{R}$ definirana formulom

$$F(x) := \sum_{y\leq x} f(y),\quad x \in \mathbb{R}^k, \tag{1.13}$$

vjerojatnosna funkcija razdiobe. Primijetite da je desna strana u (1.13) dobro definirana jer je supp$f = D$, dakle, prebrojiv skup. Općenito, ukoliko je $f : \mathbb{R}^k \to \mathbb{R}$ nenegativna Borelova funkcija takva da joj je nosač prebrojiv skup i da vrijedi

$$\sum_x f(x) = 1,$$

tada je formulom (1.13) zadana vjerojatnosna funkcija razdiobe. Takvu funkciju f zovemo **vjerojatnosnom gustoćom diskretne razdiobe**. Budući da je tada F vjerojatnosna funkcija razdiobe, postoji vjerojatnosni prostor i diskretna k-dimenzionalna slučajna veličina kojoj je F funkcija razdiobe, a f gustoća.

Dakle, diskretne ili apsolutno neprekidne slučajne varijable ili vektore zadajemo njihovim gustoćama.

Zadatak 1.14. Neka su $a < b$ realni brojevi. Pokažite da je $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := \frac{1}{b-a} \cdot \mathbb{1}_{[a,b]}(x),\ x \in \mathbb{R},$$

vjerojatnosna gustoća neke neprekidne slučajne varijable. Odredite funkciju razdiobe te varijable i pokažite da je diferencijabilna u svim realnim točkama osim u a i b, te da je za $x \in \mathbb{R} \setminus \{a, b\}$, $F'(x) = f(x)$. Neprekidnu slučajnu varijablu X kojoj je funkcija f gustoća zovemo *uniformna slučajna varijabla na segmentu* $[a, b]$, i pišemo $X \sim U(a, b)$, a njen zakon razdiobe zovemo *uniformna razdioba na segmentu* $[a, b]$. Za primjer (specijalni slučaj), vidjeti zadatak 1.12.

Zadatak 1.15. Ako je X slučajna varijabla s neprekidnom funkcijom razdiobe $F : \mathbb{R} \to \mathbb{R}$, tada je kompozicija $F \circ X = F(X)$ uniformno distribuirana slučajna varijabla na segmentu $[0, 1]$.

Zadatak 1.16. Pokažite da je za bilo koje realne brojeve μ i $\sigma^2 > 0$, funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

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

vjerojatnosna gustoća neprekidne razdiobe. Neprekidnu slučajnu varijablu X kojoj je funkcija f gustoća zovemo *normalna slučajna varijabla s parametrima μ i σ^2*, i pišemo $X \sim N(\mu, \sigma^2)$, a njen zakon razdiobe zovemo *normalna razdioba s parametrima μ i σ^2*. Ukoliko su parametri od X $\mu = 0$ i $\sigma^2 = 1$, X zovemo *standardna* (ili *jedinična*) *normalna* slučajna varijabla.

Zadatak 1.17. Pokažite da je za bilo koje pozitivne realne brojeve $\alpha > 0$ i $\beta > 0$, funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := \frac{1}{\Gamma(\alpha)\beta^\alpha}\, x^{\alpha-1} e^{-\frac{x}{\beta}} \cdot \mathbb{1}_{\langle 0,+\infty\rangle}(x),\ x \in \mathbb{R}$$

vjerojatnosna gustoća neprekidne razdiobe. Ovdje je sa Γ označena *gama* funkcija $\Gamma : \langle 0, +\infty\rangle \to \mathbb{R}$ s pravilom pridruživanja

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt,\ x > 0.$$

Zakon razdiobe neprekidne slučajne varijable X kojoj je funkcija f gustoća zovemo *gama razdioba s parametrima α i β*, i pišemo $X \sim \Gamma(\alpha, \beta)$. Specijalno, ako je $\alpha = 1$ i $\beta = 1/\lambda$ za $\lambda > 0$, tada kažemo da X ima *eksponencijalnu razdiobu s parametrom* λ i pišemo $X \sim \mathrm{Exp}\,(\lambda)$.

Zadatak 1.18. Pokažite da je za bilo koje pozitivne realne brojeve $\alpha > 0$ i $\beta > 0$, funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \cdot \mathbb{1}_{\langle 0,1\rangle}(x),\ x \in \mathbb{R},$$

vjerojatnosna gustoća neprekidne razdiobe. Ovdje je sa B označena *beta* funkcija $B : \langle 0, +\infty\rangle^2 \to \mathbb{R}$ s pravilom pridruživanja

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt,\ x, y > 0.$$

Za funkciju B vrijedi relacija

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},\ \alpha, \beta > 0.$$

Zakon razdiobe neprekidne slučajne varijable X kojoj je funkcija f gustoća zovemo *beta razdioba s parametrima α i β*, i pišemo $X \sim B(\alpha, \beta)$.

Zadatak 1.19. Neka je $x_0 \in \mathbb{R}^k$ bilo koja realna točka u $\mathbb{R}^k$. Realna izmjeriva funkcija $f : \mathbb{R}^k \to \mathbb{R}$ definirana sa

$$f(x) := \mathbb{1}_{\{x_0\}}(x),\ x \in \mathbb{R}^k,$$

je vjerojatnosna gustoća k-dimenzionalnog diskretnog slučajnog vektora (ili varijable ukoliko je $k = 1$) X za čiji zakon razdiobe vrijedi

$$\mathbb{P}(X = x_0) = 1.$$

Takvu slučajnu veličinu zovemo *degeneriranom*. Još kažemo da X *ima degeneriranu razdiobu*, ili da je *gotovo sigurno jednaka konstanti* x_0. Kraće pišemo: $X = x_0$ g.s. Zakon razdiobe degenerirane veličine X označavamo sa δ_{x_0} (*delta mjera*).

Zadatak 1.20. Neka je n prirodan broj i $p \in \langle 0, 1\rangle$ realan broj (vjerojatnost). Pokažite da je funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := \binom{n}{x} p^x(1-p)^{n-x} \cdot \mathbb{1}_{\{0,1,\ldots,n\}}(x),\ x \in \mathbb{R},$$

vjerojatnosna gustoća diskretne razdiobe s nosačem jednakim $\{0, 1, \ldots, n\}$. Diskretnu slučajnu varijablu X s gustoćom jednakom funkciji f zovemo *binomna slučajna varijabla s parametrima n i p* i pišemo: $X \sim b(n, p)$. Zakon razdiobe od X zovemo *binomna razdioba s parametrima n i p*. U slučaju kada je $n = 1$, X zovemo *Bernoullijeva slučajna varijabla s parametrom p*, a njen zakon razdiobe *Bernoullijeva razdioba s parametrom p*. Parametar p se zove *vjerojatnost uspjeha*.

Zadatak 1.21. Označimo sa $\mathbb{N}$ skup svih prirodnih brojeva (tj. skup svih pozitivnih cijelih brojeva). Neka je $p \in \langle 0, 1\rangle$ vjerojatnost. Pokažite da je funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := p(1-p)^{x-1} \cdot \mathbb{1}_{\mathbb{N}}(x),\ x \in \mathbb{R},$$

vjerojatnosna gustoća diskretne razdiobe s nosačem jednakim $\mathbb{N}$. Diskretnu slučajnu varijablu X s gustoćom jednakom funkciji f zovemo *geometrijska slučajna varijabla s parametrom p* i pišemo: $X \sim G(p)$. Zakon razdiobe od X zovemo *geometrijska razdioba s parametrom p*. Kao i kod binomne razdiobe, parametar p je vjerojatnost uspjeha.

Zadatak 1.22. Označimo sa $\mathbb{N}_0$ skup svih prirodnih brojeva zajedno s nulom (tj. skup svih nenegativnih cijelih brojeva). Neka je $\lambda > 0$ bilo koji pozitivni realni broj. Pokažite da je funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) := \frac{\lambda^x}{x!} e^{-\lambda} \cdot \mathbb{1}_{\mathbb{N}_0}(x),\ x \in \mathbb{R},$$

vjerojatnosna gustoća diskretne razdiobe s nosačem jednakim $\mathbb{N}_0$. Diskretnu slučajnu varijablu X s gustoćom jednakom funkciji f zovemo *Poissonova slučajna varijabla s parametrom* λ i pišemo: $X \sim P(\lambda)$. Zakon razdiobe od X zovemo *Poissonova razdioba s parametrom* λ.

Zadatak 1.23. Neka su $n < M < N$ prirodni brojevi. Nadalje, neka je $j \in \mathbb{N}_0$ takav broj da je $j \leq n$. Pokažite da je tada formulom

$$P(j) := \frac{\binom{M}{j}\binom{N-M}{n-j}}{\binom{N}{n}}$$

zadana vjerojatnosna funkcija mase na skupu $\{0, 1, \ldots, n\}$. Diskretna slučajna varijabla X kojoj je zakon radiobe određen sa P zove se *hipergeometrijska slučajna varijabla s parametrima* (n, M, N), a njen zakon razdiobe *hipergeometrijska razdioba s parametrima* (n, M, N). Pišemo: $X \sim H(n, M, N)$.

1.4 Uvjetne i marginalne gustoće

Neka je $k \geq 2$ i neka je $f : \mathbb{R}^k \to \mathbb{R}$ vjerojatnosna gustoća neprekidne razdiobe. Dakle, f je gustoća nekog k-dimenzionalnog neprekidnog slučajnog vektora $X = (X_1, X_2, \ldots, X_k)$. Budući da je Lebesgueov integral od f duž $\mathbb{R}^k$ jednak 1, dakle konačan, iz Fubinijevog teorema (za Lebesguevu mjeru vidjeti [7], str. 120., §5.3, teorem 3 ili teorem A.1 u dodatku A) slijedi da postoji nenegativna i integrabilna (u odnosu na Lebesgueovu mjeru) Borelova funkcija $f_1 : \mathbb{R} \to \mathbb{R}$ takva da je

$$f_1(x) = \int\limits_{\mathbb{R}^{k-1}} f(x, x_2, \ldots, x_k)\, d\lambda^{k-1}(x_2, \ldots x_k) \text{ za } \lambda^1\text{-s.s. } x \in \mathbb{R}. \tag{1.14}$$

Također, primjenom Fubinijevog teorema slijedi da je

$$\begin{aligned}\int\limits_{\mathbb{R}} f_1(x)\, d\lambda^1(x) &= \int\limits_{\mathbb{R}} \Big(\int\limits_{\mathbb{R}^{k-1}} f(x, x_2, \ldots x_k)\, d\lambda^{k-1}(x_2, \ldots, x_k)\Big)\, d\lambda^1(x) \overset{\text{(Fubinijev tm.)}}{=} \\ &= \int\limits_{\mathbb{R}^k} f(x_1, x_2, \ldots, x_k)\, d\lambda^k(x_1, x_2, \ldots, x_k) = 1.\end{aligned}$$

Dakle, f_1 je vjerojatnosna gustoća neprekidne razdiobe. Gustoću f_1 zovemo **marginalna gustoća od f u odnosu na prvu komponentu**. Analogno, za bilo koju komponetu od f opisanu indeksom j $(1 \leq j \leq k)$, nenegativna Borelova funkcija $f_j : \mathbb{R} \to \mathbb{R}$ takva da je

$$f_j(x) = \int\limits_{\mathbb{R}^{k-1}} f(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k)\, d\lambda^{k-1}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k),\ \lambda^1\text{-s.s. } x \in \mathbb{R},$$

vjerojatnosna je gustoća neke neprekidne razdiobe, koju zovemo marginalna gustoća od f u odnosu na j-tu komponentu.

Koja je interpretacija marginalne gustoće? Promotrimo marginalnu gustoću f_1. Neka je $x \in \mathbb{R}$ i $B := \langle -\infty, x] \times \mathbb{R}^{k-1}$. Tada je, prema definiciji zakona razdiobe neprekidnog slučajnog vektora i prema Fubinijevom teoremu,

$$\begin{aligned}
\mathbb{P}_{X_1}(\langle -\infty, x]) &= \mathbb{P}(X_1 \leq x) = \mathbb{P}((X_1, X_2, \ldots, X_k) \in \langle -\infty, x] \times \mathbb{R}^{k-1}) = \mathbb{P}(X \in B) = \\
&= \mathbb{P}_X(B) = \int_B f(y)\, d\lambda^k(y) = \int_{\mathbb{R}^k} \mathbb{1}_B(y) \cdot f(y)\, d\lambda^k(y) \overset{\text{(Fubinijev tm.)}}{=} \\
&= \int_{\langle -\infty, x]} \left(\int_{\mathbb{R}^{k-1}} f(y_1, y_2, \ldots, y_k)\, d\lambda^{k-1}(y_2, \ldots, y_k) \right) d\lambda^1(y_1) = \\
&= \int_{\langle -\infty, x]} f_1(y_1)\, d\lambda^1(y_1).
\end{aligned}$$

Dakle, f_1 je gustoća slučajne varijable X_1, prve komponente vektora X.

Općenito, neka je $1 \leq \ell < k$ i neka su $1 \leq j_1 < \ldots < j_\ell \leq k$. Tada se analognom primjenom Fubinijevog teorema kao gore može pokazati da je nenegativna Borelova funkcija $f_{j_1,\ldots,j_\ell} : \mathbb{R}^\ell \to \mathbb{R}$ takva da je

$$\begin{aligned}
&f_{j_1,\ldots,j_\ell}(x_1, \ldots, x_\ell) = \\
&= \int_{\mathbb{R}^{k-\ell}} f(y_1, .., y_{j_1-1}, x_1, y_{j_1+1}, .., y_{j_\ell-1}, x_\ell, y_{j_\ell+1}, .., y_k) d\lambda^{k-1}(y_1, .., y_{j_1-1}, y_{j_1+1}, .., y_{j_\ell-1}, y_{j_\ell+1}, .., y_k),
\end{aligned}$$

za λ^ℓ-s.s. $(x_1, \ldots, x_\ell) \in \mathbb{R}^\ell$, vjerojatnosna gustoća ℓ-dimenzionalne apsolutno neprekidne razdiobe, a koja je, po interpretaciji, gustoća slučajnog vektora $(X_{j_1}, \ldots, X_{j_\ell})$, ℓ-dimenzionalnog podvektora od X. Dakle, funkciju $f_{j_1,\ldots,j_\ell}$ dobijemo tako da vjerojatnosnu gustoću f prointegriramo po svim varijablama osim po komponentama na mjestima j_1,..., j_ℓ. Funkciju $f_{j_1,\ldots,j_\ell}$ zovemo **marginalna gustoća od f u odnosu na komponente** $j_1, \ldots, j_\ell$.

Primijetite da smo upravo pokazali da su komponente apsolutno neprekidnih slučajnih vektora također apsolutno neprekidne slučajne veličine.

Od sada pa nadalje, radi pojednostavljenja oznaka, sve integrale neke funkcije $f : \mathbb{R}^k \to \mathbb{R}$ u odnosu na Lebesgueovu mjeru na $\mathbb{R}^k$ za $k \geq 1$ kraće ćemo označavati na sljedeći način:

$$\int_{\mathbb{R}^k} f(y_1, y_2, \ldots, y_k)\, d\lambda^k(y_1, y_2, \ldots, y_k) \equiv \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(y_1, y_2, \ldots, y_k)\, dy_1 dy_2 \cdots dy_k$$

ili u formi vektorskog zapisa varijable integracije:

$$\int_{\mathbb{R}^k} f(y)\, d\lambda^k(y) \equiv \int_{\mathbb{R}^k} f(y)\, dy.$$

Primijetite da je ta oznaka konzistentna s Fubinijevim teoremom i činjenicom da kadagod je nenegativna podintegralna funkcija Riemann-integrabilna (ima nepravi Riemannov integral na $\mathbb{R}^k$), da je tada i Lebesgue-integrabilna i da im se integrali podudaraju (vidjeti teorem 10.10 u [9], str. 305).

Sumirajmo ova razmatranja o neprekidnim slučajnim vektorima u sljedeći teorem.

Teorem 1.24. *Neka je $X = (X_1, X_2, \ldots, X_k)$ neprekidni slučajni vektor s gustoćom $f = f_X$. Tada je svaka komponenta $X' = (X_{j_1}, \ldots, X_{j_\ell})$ od X dimenzije ℓ ($1 \leq \ell < k$, $1 \leq j_1 <$*

$\dots < j_\ell \leq k$), *neprekidna ℓ-dimenzionalna slučajna veličina i za gustoću $f_{X'}$ od X' vrijedi*

$$f_{X'}(x_1,\dots,x_\ell) = \\ = \int_{\mathbb{R}^{k-\ell}} f(y_1,..,y_{j_1-1},x_1,y_{j_1+1},..,y_{j_\ell-1},x_\ell,y_{j_\ell+1},..,y_k)\, d(y_1,..,y_{j_1-1},y_{j_1+1},..,y_{j_\ell-1},y_{j_\ell+1},..,y_k),$$

za λ^ℓ-s.s. $(x_1,\dots,x_\ell) \in \mathbb{R}^\ell$.

Neka je sada $f : \mathbb{R}^k \to \mathbb{R}$ $(k \geq 2)$ vjerojatnosna gustoća diskretne razdiobe. Dakle, f je gustoća nekog k-dimenzionalnog diskretnog slučajnog vektora $X = (X_1, X_2, \dots, X_k)$ s vrijednostima iz diskretnog skupa $D = \operatorname{supp} f$. Analogno apsolutno neprekidnom slučaju može se pokazati da je

$$f_1(x) = \sum_{y\in\mathbb{R}^{k-1}} f(x,y),\ x \in \mathbb{R}, \tag{1.15}$$

gustoća diskretne razdiobe koju također zovemo marginalna gustoća od f u odnosu na prvu komponentu. f_1 je ujedno gustoća prve komponente X_1 od X. U dokazivanju te činjenice također se koristi Fubinijev teorem, ovoga puta primijenjen na integrale u odnosu na brojaću mjeru (vidjeti dodatak A i teorem A.1). Na u potpunosti analogni način apsolutno neprekidnom slučaju, zamjenom integrala sa sumama, može se definirati i marginalna gustoća od f u odnosu na bilo koju drugu komponentu, kao i u odnosu na bilo koji podskup komponenti, uz istu interpretaciju da su to gustoće odgovarajućih podvektora diskretnog slučajnog vektora X. U tom smislu vrijedi i analogon teorema 1.24 za komponente diskretnog slučajnog vektora.

Neka je sada za $k \geq 2$, $f : \mathbb{R}^k \to \mathbb{R}$ gustoća neprekidnog ili diskretnog slučajnog vektora $X = (X_1, X_2, \dots, X_k)$. Nadalje, neka je $f_{2,\dots,k} : \mathbb{R}^{k-1} \to \mathbb{R}$ marginalna gustoća od f, gustoća podvektora $Y = (X_2, \dots, X_k)$ od X, i neka je $y = (x_2, \dots, x_k) \in \mathbb{R}^{k-1}$ bilo koja točka iz nosača gustoće $f_Y = f_{2,\dots,k}$ od Y, dakle takva točka da je $f_Y(y) > 0$. Tada se može jednostavno pokazati da je funkcija $f_{X_1|Y}(\cdot|y) : \mathbb{R} \to \mathbb{R}$, definirana formulom

$$f_{X_1|Y}(x|y) := \frac{1}{f_Y(y)} \cdot f(x,y),\ x \in \mathbb{R}, \tag{1.16}$$

vjerojatnosna gustoća neke jednodimenzionalne razdiobe. Ukoliko je X apsolutno neprekidan vektor, ta razdioba je apsolutno neprakidna, a ukoliko je X diskretan, razdioba je diskretna. Tu razdiobu (u oznaci $\mathbb{P}_{X_1|Y}(\cdot|y)$) definiranu na izmjerivom prostoru $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, zovemo **uvjetni zakon razdiobe od X_1 uz dano $Y = y$**, a pripadnu funkciju $f_{X_1|Y}(\cdot|y)$ zovemo **uvjetna gustoća od X_1 uz dano $Y = y$**.

Općenito, neka je $1 \leq \ell < k$ i neka su $1 \leq j_1 < \dots < j_\ell \leq k$. Nadalje, neka je Y podvektor od X koji sadrži sve komponente od X koje nisu jednake komponentama X_{j_1},..., X_{j_ℓ}. Nadalje, neka je $y \in \mathbb{R}^{k-\ell}$ točka iz nosača gustoće f_Y od Y. Tada se analogno kao gore može pokazati da je funkcija $f_{X_{j_1},\dots,X_{j_\ell}|Y}(\cdot|y) : \mathbb{R}^\ell \to \mathbb{R}$, definirana formulom

$$f_{X_{j_1},\dots,X_{j_\ell}|Y}(x_1,\dots,x_\ell|y) := \frac{1}{f_Y(y)} \cdot f(y_1,\dots,y_{j_1-1},x_1,y_{j_1+1},\dots,y_k),\ (x_1,\dots,x_\ell) \in \mathbb{R}^\ell,$$

vjerojatnosna gustoća neke ℓ-dimenzionalne razdiobe koju zovemo **uvjetni zakon razdiobe od $(X_{j_1}, \dots, X_{j_\ell})$ uz dano $Y = y$** i označavamo sa $\mathbb{P}_{X_{j_1},\dots,X_{j_\ell}|Y}(\cdot|y)$, a pripadnu funkciju $f_{X_{j_1},\dots,X_{j_\ell}|Y}(\cdot|y)$ zovemo **uvjetna gustoća od $(X_{j_1}, \dots, X_{j_\ell})$ uz dano $Y = y$**. Odavde također slijedi da, ako je X apsolutno neprekidan vektor, da je uvjetna razdioba od $(X_{j_1}, \dots, X_{j_\ell})$ uz dano $Y = y$ apsolutno neprekidna, a ukoliko je X diskretan, ta razdioba je diskretna. Dakle, za $B \in \mathcal{B}(\mathbb{R}^\ell)$, vjerojatnost $\mathbb{P}_{X_{j_1},\dots,X_{j_\ell}|Y}(B|y)$ možemo interpretirati kao *uvjetnu vjerojatnost događaja* $\{(X_{j_1},\ldots,X_{j_\ell})\in B\}$ *uz dano* $Y=y$, dakle, na sljedeći način:

$$\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B\,|\,Y=y)\equiv \mathbb{P}_{X_{j_1},\ldots,X_{j_\ell}|Y}(B|y):=\int_B f_{X_{j_1},\ldots,X_{j_\ell}|Y}(x_1,\ldots,x_\ell|y)\,dx_1\cdots dx_\ell, \tag{1.17}$$

u apsolutno neprekidnom slučaju, odnosno izrazom

$$\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B\,|\,Y=y)\equiv \mathbb{P}_{X_{j_1},\ldots,X_{j_\ell}|Y}(B|y):=\sum_{(x_1,\ldots,x_\ell)\in B} f_{X_{j_1},\ldots,X_{j_\ell}|Y}(x_1,\ldots,x_\ell|y), \tag{1.18}$$

u diskretnom slučaju. Nadalje, primijetite da se lijeva strana jednakosti (1.18) (u diskretnom slučaju) interpretira kao uvjetna vjerojatnost događaja $\{(X_{j_1},\ldots,X_{j_\ell})\in B\}$ *uvjetno na* događaj $\{Y=y\}$ koji je pozitivne vjerojatnosti ($\mathbb{P}(Y=y)=f_Y(y)>0$):

$$\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B\,|\,Y=y)=\frac{\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B,\,Y=y)}{\mathbb{P}(Y=y)}, \tag{1.19}$$

a što je konzistentno s desnom stranom iste jednakosti jer je:

$$\sum_{(x_1,\ldots,x_\ell)\in B} f_{X_{j_1},\ldots,X_{j_\ell}|Y}(x_1,\ldots,x_\ell|y)=\sum_{(x_1,\ldots,x_\ell)\in B}\frac{\mathbb{P}(X_{j_1}=x_1,\ldots,X_{j_\ell}=x_\ell,Y=y)}{\mathbb{P}(Y=y)}=$$

$$=\frac{\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B,\,Y=y)}{\mathbb{P}(Y=y)}=\mathbb{P}((X_{j_1},\ldots,X_{j_\ell})\in B\,|\,Y=y).$$

Oprez, za lijevu stranu u (1.17), dakle **u neprekidnom slučaju, ne vrijedi relacija (1.19)**. Naime, desna strana te relacije nije definirana jer je $\mathbb{P}(Y=y)=0$.

Primjer 1.25. *Polinomna razdioba.* Neka su k i n prirodani brojevi i neka su p_0, p_1,..., p_k pozitivni realni brojevi (vjerojatnosti) takvi da je $p_0+p_1+\cdots+p_k=1$. Nadalje, neka je $H_n:=\{(x_0,x_1,\ldots,x_k)\in\mathbb{R}^{k+1}\ :\ x_0+x_1+\cdots+x_k=n\}$ hiperravnina u $\mathbb{R}^{k+1}$. Tada je $D_n:=(\mathbb{N}_0)^{k+1}\cap H_n$ konačan skup svih nenegativnih cjelobrojnih rješenja jednadžbe $x_0+x_1+\cdots+x_k=n$. Funkcija $g:\mathbb{R}^{k+1}\to\mathbb{R}$ definirana formulom

$$g(x_0,x_1,\ldots,x_k)=\frac{n!}{x_0!x_1!\cdots x_k!}\cdot p_0^{x_0}p_1^{x_1}\cdots p_k^{x_k}\cdot\mathbb{1}_{D_n}(x_0,x_1,\ldots,x_k),\ (x_0,x_1,\ldots,x_k)\in\mathbb{R}^{k+1},$$

je gustoća $k+1$-dimenzionalnog diskretnog slučajnog vektora (s vrijednostima u D_n) kojeg zovemo *polinomni slučajni vektor*, a njegov zakon razdiobe *polinomna razdioba*. Naime, g je nenegativna Borelova funkcija s nosačem jednakim D_n i za nju vrijedi (primjenom multinomnog teorema):

$$\sum_{(x_0,x_1,\ldots,x_k)} g(x_0,x_1,\ldots,x_k)=\sum_{(j_0,j_1,\ldots,j_k)\in D_n}\frac{n!}{j_0!j_1!\cdots j_k!}p_0^{j_0}p_1^{j_1}\cdots p_k^{j_k}=(p_0+p_1+\cdots+p_k)^n=1.$$

Označimo vektor kojemu je g gustoća sa $(X_0,X_1,\ldots,X_k)$. Pišemo: $(X_0,X_1,\ldots,X_k)\sim M(n;p_0,p_1,\ldots,p_k)$. Primijetite da s vjerojatnosti 1 vrijedi da je $X_0=n-(X_1+\cdots+X_k)$. Zbog toga se možemo ograničiti na slučajni vektor $X=(X_1,\ldots,X_k)$. Tada je funkcija $f:\mathbb{R}^k\to\mathbb{R}$, definirana kao (na prirodan način parametrizirana) restrikcija od g na H_n:

$$f(x_1,\ldots,x_k):=g(n-(x_1+\cdots+x_k),x_1,\ldots,x_k),\ (x_1,\ldots,x_k)\in\mathbb{R}^k,$$

gustoća od X. Ako je $(x_1, x_2, \dots, x_k) \in \mathbb{R}^k \cap (\mathbb{N}_0)^k$ tako da je $x_1 + x_2 + \cdots + x_k \le n$, tada je

$$f(x_1, x_2, \dots, x_k) = \binom{n}{x_1}\binom{n - x_1}{x_2} \cdots \binom{n - (x_1 + \cdots + x_{k-1})}{x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} p_0^{n-(x_1+\cdots+x_k)}.$$

Specijalno, u slučaju kada je $k = 1$, X je binomna slučajna varijabla $b(n, p_1)$.
Pokažimo da je razdioba bilo koje komponente X_j polinomnog slučajnog vektora binomna $b(n, p_j)$, $j = 0, 1, \dots, k$. Pokazat ćemo tu tvrdnju za komponentu X_1. Zbog simetrije, tvrdnja će onda vrijediti i za sve ostale komponente. Neka je $i \in \{0, 1, \dots, n\}$. Pomoću binomnog teorema računamo marginalnu gustoću od f u odnosu na prvu komponentu:

$$f_{X_1}(i) = \sum_{i_2=0}^{n-i} \cdots \sum_{i_k=0}^{n-(i_2+\cdots+i_{k-1})} g(n - (i + i_2 + \cdots + i_k), i, i_2, \dots, i_k) =$$

$$= \sum_{i_2=0}^{n-i} \cdots \sum_{i_k=0}^{n-(i_2+\cdots+i_{k-1})} f(i, i_2, \dots, i_k) =$$

$$= \binom{n}{i} p_1^i \sum_{i_2=0}^{n-i} \cdots \sum_{i_k=0}^{n-(i_2+\cdots+i_{k-1})} \binom{n - (i + i_2 \cdots + i_{k-1})}{i_k} p_k^{i_k} p_0^{n-(i+i_2+\cdots+i_k)} = (\text{binomni t.})$$

$$= \binom{n}{i} p_1^i \sum_{i_2=0}^{n-i} \cdots \sum_{i_{k-1}=0}^{n-(i_2+\cdots+i_{k-2})} \binom{n - (i + i_2 \cdots + i_{k-2})}{i_{k-1}} p_{k-1}^{i_{k-1}} (p_0 + p_k)^{n-(i+i_2+\cdots+i_{k-1})} =$$

$$= \cdots = \binom{n}{i} p_1^i (p_0 + p_2 + \cdots + p_k)^{n-i} = \binom{n}{i} p_1^i (1 - p_1)^{n-i}.$$

Dakle, $X_1 \sim b(n, p_1)$. Na sličan način može se pokazati da za marginalnu gustoću $f_{2,\dots,k}$ od f, dakle gustoću od $(X_2, \dots, X_k)$, vrijedi

$$\begin{aligned} f_{2,\dots,k}(i_2, \dots, i_k) &= \binom{n}{i_2} \cdots \binom{n - (i_2 + \cdots + i_{k-1})}{i_k} \cdot p_2^{i_1} \cdots p_k^{i_k} r^{n-(i_2+\cdots+i_k)} = \\ &= \frac{n!}{i_2! \cdots i_k!(n - i_2 - \cdots - i_k)!} p_2^{i_2} \cdots p_k^{i_k} r^{n-i_2-\cdots-i_k} \end{aligned}$$

gdje je $r = 1 - (p_2 + \cdots + p_k) = p_0 + p_1$, za $(i_2, \dots i_k) \in (\mathbb{N}_0)^{k-1}$ takve da je $i_2 + \cdots + i_k \le n$.
Dakle, $(n - (X_2 + \cdots + X_k), X_2, \dots, X_k) \sim M(n; r, p_2, \dots, p_k)$. Ovaj rezultat se može poopćiti na bilo koji podvektor od X.
Sada možemo odrediti uvjetnu razdiobu od X_1 uz dano $Y = (X_2, \dots, X_k) = (i_2, \dots, i_k)$, gdje su i_2,..., i_k nenegativni cijeli brojevi takvi da je $i_2 + \cdots + i_k \le n$. Za $i \in \{0, 1, \dots, n - (i_2 + \cdots + i_k)\}$, prema (1.15),

$$\begin{aligned} f_{X_1|Y}(i|i_2, \dots, i_k) &= \frac{f(i, i_2, \dots, i_k)}{f_{2,\dots,k}(i_2, \dots, i_k)} = \\ &= \frac{\frac{n!}{i! i_2! \cdots i_k!(n-i-i_2-\cdots-i_k)!} p_1^i p_2^{i_2} \cdots p_k^{i_k} p_0^{n-i-i_2-\cdots-i_k}}{\frac{n!}{i_2! \cdots i_k!(n-i_2-\cdots-i_k)!} p_2^{i_2} \cdots p_k^{i_k} r^{n-i_2-\cdots-i_k}} = \\ &= \binom{n - i_2 - \cdots - i_k}{i} \left(\frac{p_1}{p_0 + p_1}\right)^i \left(1 - \frac{p_1}{p_0 + p_1}\right)^{n-i_2-\cdots-i_k-i} \end{aligned}$$

Dakle, uvjetna razdioba od X_1 uz dano $(X_2, \dots, X_k) = (i_2, \dots, i_k)$ je binomna razdioba $b(n - i_2 - \cdots - i_k, p_1/(p_0 + p_1))$. Zbog simetrije, analogni rezultat vrijedi za uvjetne razdiobe ostalih komponenata od X. □

Primjer 1.26. *Bivarijatna normalna razdioba.* Kažemo da slučajni vektor (X, Y) ima *bivarijatnu normalnu razdiobu* ako je neprekidan s gustoćom $f : \mathbb{R}^2 \to \mathbb{R}$, danom formulom

$$f(x,y) =$$
$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\},$$

za $(x, y) \in \mathbb{R}^2$. Parametri razdiobe su realni brojevi μ_1, μ_2, $\sigma_1 > 0$, $\sigma_2 > 0$ i $-1 < \rho < 1$. Odredimo zakon razdiobe prve komponente X. Prvo, primijetimo da je

$$A^2 - 2\rho AB + B^2 = (1-\rho^2)A^2 + (B-\rho A)^2 = (A-\rho B)^2 + (1-\rho^2)B^2.$$

Ako stavimo da je $A = (x - \mu_1)/\sigma_1$ i $B = (y - \mu_2)/\sigma_2$, te uz primjenu zamjene varijabli $z = B - \rho A$, slijedi da je marginalna gustoća od f u odnosu na prvu komponentu, dakle gustoća f_X od X,

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f(x,y)\,dy = \frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-\frac{1}{2}A^2}\int_{\mathbb{R}}\frac{1}{\sqrt{1-\rho^2}\sqrt{2\pi}}e^{-\frac{z^2}{2(1-\rho^2)}}dz = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-\frac{1}{2}A^2}\cdot 1 = \frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2}, \end{aligned}$$

gdje smo u trećoj jednakosti iskoristili činjenicu da je integral vjerojatnosne gustoće (u ovome slučaju normalne razdiobe $N(0, 1-\rho^2)$) jednak 1. Dakle, razdioba prve komponente X je normalna $N(\mu_1, \sigma_1^2)$. Zbog simetrije slijedi da je razdioba druge komponente Y normalna $N(\mu_2, \sigma_2^2)$.
Sada možemo odrediti i uvjetnu razdiobu od X. Neka je $y \in \mathbb{R}$ proizvoljan. Tada je $f_Y(y) > 0$ gdje je f_Y gustoća od Y. Uz iste supstitucije A i B kao gore, uvjetna gustoća od X uz dano $Y = y$ je, prema (1.16),

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,e^{-\frac{1}{2(1-\rho^2)}[(A-\rho B)^2+(1-\rho^2)B^2]}}{\frac{1}{\sigma_2\sqrt{2\pi}}\,e^{-\frac{1}{2}B^2}} = \\ &= \frac{1}{\sigma_1\sqrt{1-\rho^2}\sqrt{2\pi}}\,e^{-\frac{1}{2(1-\rho^2)}(A-\rho B)^2} = \\ &= \frac{1}{\sigma_1\sqrt{1-\rho^2}\sqrt{2\pi}}\,e^{-\frac{1}{2\sigma_1^2(1-\rho^2)}(x-\mu_1-\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2))^2},\ x \in \mathbb{R}. \end{aligned}$$

Dakle, uvjetna razdioba od X uz dano $Y = y$ je normalna $N(\mu_1+\rho\sigma_1(y-\mu_2)/\sigma_2, \sigma_1^2(1-\rho^2))$ razdioba. Zbog simetrije slijedi da je za sve $x \in \mathbb{R}$, uvjetna razdioba od Y uz dano $X = x$ normalna $N(\mu_2 + \rho\sigma_2(x - \mu_1)/\sigma_1, \sigma_2^2(1 - \rho^2))$ razdioba.
Na kraju, provjerimo da je f zaista vjerojatnosna gustoća neprekidnog dvodimenzionalnog slučajnog vektora. Trivijalno je f pozitivna i Borelova funkcija (jer je neprekidna), a primjenom Fubinijevog teorema i činjenice da su marginalne i uvjetne gustoće od f gustoće normalnih razdioba, dakle vjerojatnosne, slijedi i da je

$$\int_{\mathbb{R}^2} f(x,y)\,dxdy = \int_{\mathbb{R}} f_X(x)\left(\int_{\mathbb{R}} f_{Y|X}(y|x)\,dy\right)dx = \int_{\mathbb{R}} f_X(x)\cdot 1\,dx = 1.$$

Dakle, f je vjerojatnosna gustoća. □

1.5 Nezavisnost slučajnih varijabli i vektora

Neka je $(X_n; n \in I)$ konačan ili (prebrojivo) beskonačan niz slučajnih varijabli ili vektora definiranih na istome vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Ovdje je I skup indeksa koji je prebrojiv, dakle možemo uzeti da je I konačan podskup skupa prirodnih brojeva $\mathbb{N}$ ili $I = \mathbb{N}$. Općenito, X_n je k_n-dimenzionalna slučajna veličina, za $n \in I$.

Definicija 1.27. *Niz $(X_n; n \in I)$ je niz* **nezavisnih** *slučajnih veličina (varijabli ili vektora) ako za svaki konačni podskup $I' = \{n_1, n_2, \dots n_\ell\}$ skupa indeksa I (ovdje je $\ell = |I'|$) i sve $B_1 \in \mathcal{B}(\mathbb{R}^{k_{n_1}})$, $B_2 \in \mathcal{B}(\mathbb{R}^{k_{n_2}})$,..., $B_\ell \in \mathcal{B}(\mathbb{R}^{k_{n_\ell}})$ vrijedi*

$$\mathbb{P}(X_{n_1} \in B_1, X_{n_2} \in B_2, \dots, X_{n_\ell} \in B_\ell) = \mathbb{P}(X_{n_1} \in B_1) \cdot \mathbb{P}(X_{n_2} \in B_2) \cdots \mathbb{P}(X_{n_\ell} \in B_\ell).$$

Kažemo da je $(X_n; n \in I)$ niz **zavisnih** *slučajnih veličina ako nije niz nezavisnih slučajnih veličina.*

Neka je $(X_n; n \in I)$ niz nezavisnih slučajnih varijabli ili vektora. Tada je moguće formirati nove nizove nezavisnih slučajnih varijabli ili vektora na sljedeći način. Neka je $1 < n_1 < n_2 < \cdots$ strogo rastući podniz elemenata iz skupa indeksa I kojim je definirana jedna particija elemenata niza $(X_n; n \in I)$:

$$(X_1, X_2, \dots, X_{n_1}),\ (X_{n_1+1}, X_{n_1+2}, \dots, X_{n_2}),\ (X_{n_2+1}, X_{n_2+2}, \dots, X_{n_3}), \dots$$

Na svaku od tih particija primijenimo izmjerive transformacije. Preciznije, neka su

$$g_1 : \mathbb{R}^{k_1+k_2+\cdots+k_{n_1}} \to \mathbb{R}^{m_1},\ g_2 : \mathbb{R}^{k_{n_1+1}+k_{n_1+2}+\cdots+k_{n_2}} \to \mathbb{R}^{m_2},\ g_3 : \mathbb{R}^{k_{n_2+1}+k_{n_2+2}+\cdots+k_{n_3}} \to \mathbb{R}^{m_3}, \dots$$

Borelove funkcije. Tada su slučajne veličine

$$Y_1 := g_1(X_1, X_2, \dots, X_{n_1}),\ Y_2 := g_2(X_{n_1+1}, X_{n_1+2}, \dots, X_{n_2}),\ Y_3 := g_3(X_{n_2+1}, X_{n_2+2}, \dots, X_{n_3}), \dots$$

također nezavisne (vidjeti [9], str. 356., teorem 11.3).

Moguće je konstruirati niz nezavisnih slučajnih varijabli ili vektora. Preciznije, za zadani niz vjerojatosnih funkcija razdioba $(F_n; n \in I)$ postoji vjerojatnosni prostor i niz nezavisnih slučajnih veličina $(X_n; n \in I)$ definiranih na tom prostoru, takvih da je za $n \in I$, F_n funkcija razdiobe od X_n (vidjeti [9], str. 358., teorem 11.7). Jedna takva realizacija je sljedeći *kanonski vjerojatnosni prostor* (opisan u dokazu citiranog teorema). Označimo sa Ω skup svih nizova $\omega = (\omega_1, \omega_2, \dots)$ čiji elementi su točke $\omega_j \in \mathbb{R}^{k_j}$, $j = 1, 2, \dots$. Sa $\mathcal{F}$ označimo σ-algebru generiranu *cilindrima* s konačno dimenzionalnim Borelovim skupovima za baze (za primjer konstrukcije vidjeti [9], str. 335.-336.). Na tom izmjerivom prostoru $(\Omega, \mathcal{F})$ definiramo koordinatne projekcije $(X_n; n \in I)$ sa $X_n(\omega) := \omega_n$, $n \in I$, koje su izmjeriva preslikavanja. Nadalje, neka je za svaki $n \in I$, $\mathbb{P}_n$ vjerojatnost na izmjerivom prostoru $(\mathbb{R}^{k_n}, \mathcal{B}(\mathbb{R}^{k_n}))$ inducirana vjerojatnosnom funkcijom razdiobe F_n. Tada je vjerojatnost na $(\Omega, \mathcal{F})$ *produktna vjerojatnost* $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2 \times \cdots$. Prema korolaru 10.9 iz [9], str. 338., takva vjerojatnost (i vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$) postoji. Uz tu vjerojatnost je *kanonski niz* $(X_n; n \in I)$ niz nezavisnih slučajnih veličina takav da je za svaki $n \in I$, zakon razdiobe od X_n jednak $\mathbb{P}_n$.

Za konačan niz slučajnih varijabli imamo jednostavnu karakterizaciju nezavisnosti (vidjeti [9], str. 354., teorem 11.1).

Teorem 1.28. *Neka je $X = (X_1, X_2, \dots, X_k)$ k-dimenzionalni slučajni vektor. Komponente X_1, X_2,..., X_k vektora X su nezavisne ako i samo ako za funkciju razdiobe F_X od X vrijedi*

$$(\forall (x_1, x_2, \dots, x_k) \in \mathbb{R}^k) \quad F_X(x_1, x_2, \dots, x_k) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdots F_{X_k}(x_k),$$

gdje su F_{X_1}, F_{X_2},..., F_{X_k} funkcije razdioba komponenti od X.

Ako slučajni vektor $X = (X_1, X_2, \ldots, X_k)$ ima gustoću f_X (kao apsolutno neprekidan ili diskretan slučajni vektor), tada je iz prethodnog potpoglavlja poznato da i komponente imaju gustoće, u oznaci f_{X_j}, $j = 1, 2, \ldots, k$. Sada se nezavisnost komponenti od X može karakterizirati pomoću njihovih gustoća.

Teorem 1.29. *Komponente X_1, X_2,..., X_k neprekidnog slučajnog vektora $X = (X_1$, X_2,..., $X_k)$ s gustoćom f_X su nezavisne ako i samo ako vrijedi*

$$f_X(x_1, x_2, \ldots, x_k) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_k}(x_k) \tag{1.20}$$

za sve $(x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ osim možda za točke iz skupa Lebesgueove mjere nula.

Za dokaz teorema 1.29 vidjeti [9], str. 355., teorem 11.2. Ako je X diskretan slučajni vektor, tada jednakost u (1.20) vrijedi za sve $(x_1, x_2, \ldots, x_k) \in D$, gdje je D, nosač od f_X u $\mathbb{R}^k$, prebrojiv skup vrijednosti od X (vidjeti [9], str. 92.-93., teorem 5.3).

S druge strane, ako imamo konačan niz nezavisnih neprekidnih slučajnih varijabli X_1, X_2,..., X_k s gustoćama f_{X_1}, f_{X_2},..., f_{X_k} tada se može pokazati da je slučajni vektor $X = (X_1, X_2, \ldots, X_k)$ kojemu su te varijable komponente, također apsolutno neprekidan. Naime, vrijedi sljedeći teorem (vidjeti [9], str. 356., korolar 11.1).

Teorem 1.30. *Neka su X_1, X_2,..., X_k nezavisne neprekidne slučajne varijable s gustoćama f_{X_1}, f_{X_2},..., f_{X_k}. Tada je slučajni vektor $X = (X_1, X_2, \ldots, X_k)$ neprekidan s gustoćom:*

$$f(x_1, \ldots, x_k) := f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_k}(x_k),\ (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k. \tag{1.21}$$

Ukoliko su te slučajne varijable zavisne, X općenito ne mora biti apsolutno neprekidan slučajni vektor (vidjeti primjer 1.37 u sljedećem potpoglavlju). U slučaju diskretnih slučajnih varijabli trivijalno je $X = (X_1, X_2, \ldots, X_k)$ diskretan slučajni vektor ako i samo ako su X_1, X_2,..., X_k diskretne slučajne varijable. Specijalno, ako su X_1, X_2,..., X_k nezavisne diskretne slučajne varijable s gustoćama f_{X_1}, f_{X_2},..., f_{X_k} i s vrijednostima u prebrojivim skupovima D_1, D_2,..., D_k, tada je gustoća diskretnog slučajnog vektora $X = (X_1, X_2, \ldots, X_k)$, s vrijednostima u skupu $D = D_1 \times D_2 \times \cdots \times D_k$, također dana formulom (1.21).

Zadatak 1.31. Neka je X_1, X_2,..., X_n niz nezavisnih jednako distribuiranih neprekidnih slučajnih varijabli s gustoćom f i funkcijom razdiobe F. Definiramo *uređajne statistike* $X_{(1)}$, $X_{(2)}$,..., $X_{(n)}$ kao niz dobiven iz polaznog niza sortiranjem vrijednosti uzlazno, tako da vrijedi

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}.$$

Preciznije, $X_{(1)} := \min_{1 \leq i \leq n} X_i$, a $X_{(j)} := \min\{X_i : X_i > X_{(j-1)}\}$ induktivno za $j \geq 2$. Primijetite da je $X_{(n)} = \max_{1 \leq i \leq n} X_i$, te da gornje stroge nejednakosti vrijede *gotovo sigurno*. (Zašto?)

(a) Pokažite da su $X_{(j)}$ za $1 \leq j \leq n$ slučajne varijable, odredite im funkcije razdioba i pokažite da su apsolutno neprekidne. Odredite im gustoće. (Pretpostavite da je $F'(x) = f(x)$ za sve osim možda za konačno mnogo $x \in \mathbb{R}$.)

(b) Pretpostavite da je F funkcija razdiobe uniformne distribucije $U(0, 1)$. Odredite zakone razdioba uređajnih statistika $X_{(j)}$ $(1 \leq j \leq n)$.

(c) Uz istu pretpostavku kao u (b), pokažite da je slučajni vektor $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ neprekidan i odredite mu gustoću.

Zadatak 1.32. Neka je (X, Y) dvodimenzionalan neprekidan slučajni vektor s gustoćom $f = f_{X,Y}$. Neka su redom f_X i f_Y (marginalne) gustoće od X i Y. Tada su X i Y

nezavisne slučajne varijable ako i samo ako je za sve $y \in \operatorname{supp} f_Y$, $f_{X|Y}(\cdot|y) = f_X$ λ-s.s. U tom slučju je i $f_{Y|X}(\cdot|x) = f_Y$ λ-s.s. za sve $x \in \operatorname{supp} f_X$. Dokažite. Pokažite da ista karakterizacija nezavisnosti pomoću marginalnih i uvjetnih gustoća vrijedi i u diskretnom slučaju. Odavde zaključite da su (u kontekstu ovoga zadatka) X i Y nezavisne slučajne varijable ako i samo ako je uvjetni zakon razdiobe jedne (prve) u odnosu na drugu varijablu jednak (bezuvjetnom) zakonu razdiobe prve varijable, dakle, neovisno o vrijednosti druge varijable.

1.6 Funkcije slučajnih varijabli i vektora

Na početku poglavlja smo naveli da, ako je X slučajna veličina (varijabla ili vektor) i g Borelova funkcija, da je tada $Y = g(X)$ slučajna veličina, dakako, uz uvjet da je navedena kompozicija funkcija dobro definirana. To znači da mora biti ispunjeno da je $\mathbb{P}(X \in D_g) = \mathbb{P}_X(D_g) = 1$, gdje je D_g domena funkcije g. Jasno je da, ako je X diskretna slučajna varijabla ili vektor, ili je g diskretna Borelova funkcija, da je tada i $Y = g(X)$ diskretna slučajna veličina. U ovom potpoglavlju probat ćemo odgovoriti na pitanje: ako je X *neprekidna* slučajna varijabla ili vektor s gustoćom f, uz koje uvjete na g je $Y = g(X)$ *neprekidna* slučajna veličina i kako se pomoću f i g računa njezina gustoća?

Primjer 1.33. Neka je X neprekidna slučajna varijabla s gustoćom $f = f_X$. Tada je $Y = X^2$ neprekidna slučajna varijabla s gustoćom

$$h(x) = \frac{1}{2\sqrt{x}} \left(f(-\sqrt{x}) + f(\sqrt{x}) \right) \cdot \mathbb{1}_{\langle 0,+\infty\rangle}(x).$$

Pokažimo to. Neka je $F(y) = F_Y(y) = \mathbb{P}(Y \le y)$ funkcija distribucije od Y. Ako je $y \le 0$, tada je

$$F(y) = \mathbb{P}(X^2 \le y) = 0 = \int\limits_{-\infty}^{y} h(s)\, ds.$$

Neka je sada $y > 0$. Računamo:

$$\begin{aligned} F(y) &= \mathbb{P}(X^2 \le y) = \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}) \overset{(1.5)}{=} \int\limits_{-\sqrt{y}}^{\sqrt{y}} f(t)\, dt = \\ &= \int\limits_{-\sqrt{y}}^{0} f(t)\, dt + \int\limits_{0}^{\sqrt{y}} f(t)\, dt \overset{\text{(tm.z.v.)}}{=} \int\limits_{y}^{0} f(-\sqrt{s}) \frac{1}{2\sqrt{s}}\, ds + \int\limits_{0}^{y} f(\sqrt{s}) \frac{1}{2\sqrt{s}}\, ds = \\ &= \int\limits_{0}^{y} (f(-\sqrt{s}) + f(\sqrt{s})) \frac{1}{2\sqrt{s}}\, ds = \int\limits_{-\infty}^{y} h(s)\, ds \end{aligned}$$

pri čemu je druga jednakost u drugom retku posljedica *teorema o zamjeni varijabli* (vidjeti Teorem 20.1, str. 180. u [10] i Teorem 10, IV., §5.4, str. 130 u [7]) primijenjenom na funkcije $t = g_1^{-1}(s) = -\sqrt{s}$ i $t = g_1^{-1}(s) = \sqrt{s}$ redom, dakle, u prvom integralu na bijekciju $g_1 = g|_{\langle -\sqrt{y},0\rangle} : \langle -\sqrt{y}, 0\rangle \to \langle 0, y\rangle$, a u drugom na bijekciju $g_2 = g|_{\langle 0,\sqrt{y}\rangle} : \langle 0, \sqrt{y}\rangle \to \langle 0, y\rangle$, gdje je $g(x) = x^2$. Primijetimo da su u oba slučaja funkcije zamjene varijabli, inverzi restrikcija g_1 i g_2 od g, *neprekidno diferencijabilne injekcije* na području integracije pa je moguće primijeniti navedeni teorem o zamjeni varijabli. Da je Y neprekidna slučajna varijabla s gustoćom h slijedi iz definicije 1.9. $\square$

Iz prethodnog primjera je jasno da je za neprekidnu slučajnu varijablu X s gustoćom $f = f_X$, slučajna varijabla $Y = g(X)$ neprekidna ukoliko je g takva Borelova funkcija da postoje međusobno disjunktni otvoreni intervali D_1, D_2,... (najviše prebrojivo mnogo njih) i Borelov skup N takvi da je $\mathbb{P}_X(N) = 0$, da za domenu D_g od g vrijedi $D_g = N \cup D_1 \cup D_2 \cup \cdots$, a da su restrikcije $g_k := g|_{D_k} : D_k \to S_k := g_k(D_k)$ $(k = 1, 2, \ldots)$ bijekcije (strogo monotone funkcije) takve da su njihovi inverzi $g_k^{-1} : S_k \to D_k$ $(k = 1, 2, \ldots)$ neprekidno diferencijabilne funkcije (tj. klase C^1). Naime, za svaki $y \in \mathbb{R}$ vrijedi ($I_y := \langle -\infty, y]$)

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(g(X) \le y) = \mathbb{P}(X \in g^{-1}(I_y)) \overset{(1.5)}{=} \int\limits_{g^{-1}(I_y)} f(t)\,dt = \int\limits_{g^{-1}(I_y)\cap D_g} f(t)\,dt = \\
&= \int\limits_{g^{-1}(I_y)\cap N} f(t)\,dt + \int\limits_{g^{-1}(I_y)\cap \bigcup_k D_k} f(t)\,dt \overset{\text{(LTMK)}}{=} 0 + \sum_k \int\limits_{g^{-1}(I_y)\cap D_k} f(t)\,dt \overset{\text{(tm.z.v.)}}{=} \\
&= \sum_k \int\limits_{g_k(g^{-1}(I_y)\cap D_k)} f(g_k^{-1}(s)) |\frac{d}{ds} g_k^{-1}(s)|\,ds = \sum_k \int\limits_{I_y \cap S_k} f(g_k^{-1}(s)) |\frac{d}{ds} g_k^{-1}(s)|\,ds \overset{\text{(LTMK)}}{=} \\
&= \int\limits_{I_y} \left(\sum_k f(g_k^{-1}(s)) |\frac{d}{ds} g_k^{-1}(s)| \cdot \mathbb{1}_{S_k}(s) \right) ds \overset{(1.22)}{=} \int\limits_{-\infty}^{y} f_Y(s)\,ds.
\end{aligned}$$

Druga jednakost u drugom retku i prva jednakost u zadnjem retku su posljedica primjene Lebesgueovog teorema o monotonoj konvergenciji (kraće, LTMK, vidjeti npr. [7], §3.4, teorem 10, str. 42.) ili, jednostavno, linearnosti integrala u slučaju konačno mnogo funkcija g_k, i činjenice da je $\mathbb{P}_X(N) = 0$ (u prvom slučaju), a prva jednakost u trećem retku je posljedica teorema o zamjeni varijabli primijenjenog na difeomorfizme g_k. Dakle, po definiciji 1.9, $Y = g(X)$ je neprekidna slučajna varijabla s gustoćom:

$$f_Y(y) = \sum_k f_X(g_k^{-1}(y)) |\frac{d}{dy} g_k^{-1}(y)| \cdot \mathbb{1}_{S_k}(y). \tag{1.22}$$

U višedimenzionalnom slučaju, kada je X d-dimenzionalni slučajni vektor, uz analogne uvjete na Borelovu funkciju g kao u jednodimenzionalnom slučaju, vrijedit će da je $Y = g(X)$ d-dimenzionalni neprekidni slučajni vektor.

Primjer 1.34. *Box-Müllerova metoda.* Neka su X_1, X_2 dvije nezavisne slučajne varijable s uniformnom razdiobom na segmentu $[0, 1]$. Definiramo dvije nove varijable:

$$Y_1 = \sqrt{-2\log X_1} \cdot \cos(2\pi X_2), \quad Y_2 = \sqrt{-2\log X_1} \cdot \sin(2\pi X_2).$$

Tada su Y_1 i Y_2 nezavisne slučajne varijable, obje s jediničnom normalnom razdiobom. Primijetite da je definicija od Y_1, Y_2 dobra jer je $\mathbb{P}(-2\log X_1 > 0, X_1 > 0) = \mathbb{P}(X_1 \in \langle 0, 1\rangle) = 1$. Nadalje, slučajni vektor (X_1, X_2) je neprekidan s gustoćom

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) = \mathbb{1}_{\langle 0,1\rangle}(x_1) \cdot \mathbb{1}_{\langle 0,1\rangle}(x_2) = \mathbb{1}_{\langle 0,1\rangle^2}(x_1, x_2), \quad x_1, x_2 \in \mathbb{R},$$

budući da su X_1, X_2 neprekidne i nezavisne slučajne varijable. S druge strane, vrijedi da je $(Y_1, Y_2) = g(X_1, X_2)$ za funkciju $g : \langle 0, 1\rangle \times \langle 0, 1\rangle \to \mathbb{R} \times \mathbb{R}$ definiranu formulom

$$g(x_1, x_2) = (\sqrt{-2\log x_1} \cdot \cos(2\pi x_2), \sqrt{-2\log x_1} \cdot \sin(2\pi x_2)).$$

Funkcija g nije injekcija na svojoj domeni, ali se može rastaviti na dvije bitne bijekcije:

$$\begin{aligned}
g_1 : &= g|_{\langle 0,1\rangle \times \langle 0,\frac{1}{2}\rangle} : \langle 0, 1\rangle \times \langle 0, \tfrac{1}{2}\rangle =: D_1 \to S_1 := \mathbb{R} \times \langle 0, +\infty\rangle \\
g_2 : &= g|_{\langle 0,1\rangle \times \langle \frac{1}{2},1\rangle} : \langle 0, 1\rangle \times \langle \tfrac{1}{2}, 1\rangle =: D_2 \to S_2 := \mathbb{R} \times \langle -\infty, 0\rangle
\end{aligned}$$

jer vrijedi da je $D_g = \langle 0,1\rangle \times \langle 0,1\rangle = N \cup D_1 \cup D_2$ gdje je $N = \langle 0,1\rangle \times \{\frac{1}{2}\}$ očito takav da je $\mathbb{P}_{X_1,X_2}(N) = \mathbb{P}(X_2 = \frac{1}{2}) = 0$. U tom slučaju su njihovi inverzi, $g_i^{-1} : S_i \to D_i$ $(i = 1,2)$, dani formulama:

$$\begin{aligned} g_1^{-1}(y_1,y_2) &= (e^{-\frac{1}{2}(y_1^2+y_2^2)}, \tfrac{1}{2\pi}\operatorname{arcctg}\tfrac{y_1}{y_2}) \\ g_2^{-1}(y_1,y_2) &= (e^{-\frac{1}{2}(y_1^2+y_2^2)}, \tfrac{1}{2\pi}\operatorname{arcctg}\tfrac{y_1}{y_2} + \tfrac{1}{2}), \end{aligned}$$

neprekidno diferencijabilne funkcije na svojim domenama. Primjetite da je u oba slučaja

$$|\det Dg_1^{-1}(y_1,y_2)| = |\det Dg_2^{-1}(y_1,y_2)| = \frac{1}{2\pi}e^{-\frac{1}{2}(y_1^2+y_2^2)}.$$

Sada primijenimo sličan argument kao u prethodnom primjeru. Neka je $z \in \mathbb{R}^2$ proizvoljan i $I_z := \langle -\infty, z]$ pripadni poluotvoreni odozdo neomeđeni pravokutnik. Tada je, primjenom teorema o zamjeni varijabli:

$$\begin{aligned} & F_{Y_1,Y_2}(z) = \mathbb{P}(g(X_1,X_2) \le z) = \mathbb{P}((X_1,X_2) \in g^{-1}(I_z)) \overset{(1.5)}{=} \\ = & \iint\limits_{g^{-1}(I_z)} \mathbb{1}_{\langle 0,1\rangle^2}(x_1,x_2)\, dx_1 dx_2 = \int_0^1\int_0^1 \mathbb{1}_{g^{-1}(I_z)}(x_1,x_2)\, dx_1 dx_2 = \\ = & \int_0^1 dx_1 \int_0^{\frac{1}{2}} dx_2\, \mathbb{1}_{g^{-1}(I_z)} + \int_0^1 dx_1 \int_{\frac{1}{2}}^1 dx_2\, \mathbb{1}_{g^{-1}(I_z)}(x_1,x_2) \overset{\text{(tm.z.v.)}}{=} \\ = & \int_{-\infty}^{+\infty} dy_1 \int_0^{+\infty} dy_2\, \mathbb{1}_{g^{-1}(I_z)}(g_1^{-1}(y_1,y_2))|\det Dg_1^{-1}(y_1,y_2)| + \\ & + \int_{-\infty}^{+\infty} dy_1 \int_{-\infty}^{0} dy_2\, \mathbb{1}_{g^{-1}(I_z)} g_2^{-1}(y_1,y_2))|\det Dg_2^{-1}(y_1,y_2)| = \\ = & \int_{-\infty}^{+\infty} dy_1 \int_0^{+\infty} dy_2\, \mathbb{1}_{I_z}(y_1,y_2)\frac{1}{2\pi}e^{-\frac{1}{2}(y_1^2+y_2^2)} + \int_{-\infty}^{+\infty} dy_1 \int_{-\infty}^{0} dy_2\, \mathbb{1}_{I_z}(y_1,y_2)\frac{1}{2\pi}e^{-\frac{1}{2}(y_1^2+y_2^2)} = \\ = & \iint\limits_{\langle -\infty, z]} \frac{1}{2\pi}e^{-\frac{1}{2}(y_1^2+y_2^2)}\, dy_1 dy_2. \end{aligned}$$

Dakle, prema definiciji 1.9, slučajni vektor (Y_1, Y_2) je neprekidan s gustoćom

$$f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{2\pi}e^{-\frac{1}{2}(y_1^2+y_2^2)} = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_1^2} \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_2^2} = \phi(y_1)\phi(y_2),$$

gdje je ϕ funkcija gustoće jedinične normalne razdiobe (vidjeti zadatak 1.16). Sada se lako vidi da su marginalne gustoće jednake:

$$f_{Y_1}(y) = \int_{-\infty}^{+\infty} f_{Y_1,Y_2}(y,v)\, dv = \int_{-\infty}^{+\infty} \phi(y)\phi(v)\, dv = \phi(y)\int_{-\infty}^{+\infty} \phi(v)\, dv = \phi(y)\cdot 1 = \phi(y),\ y \in \mathbb{R},$$

odnosno $f_{Y_2}(y) = \phi(y)$, $y \in \mathbb{R}$, zbog simetrije. Dakle, Y_1 i Y_2 imaju jediničnu normalnu distribuciju. Nadalje, budući da je za sve $y_1, y_2 \in \mathbb{R}$,

$$f_{Y_1,Y_2}(y_1,y_2) = \phi(y_1)\phi(y_2) = f_{Y_1}(y_1)\cdot f_{Y_2}(y_2)$$

prema teoremu 1.29 zaključujemo da su Y_1 i Y_2 nezavisne slučajne varijable. □

U skladu s prethodnim primjerima, vrijedi sljedeći teorem (za dokaz vidjeti [9], teorem 11.8 i napomenu iza njega na str. 363. - 354., te teorem 11.9 i napomenu iza njega na str. 266. - 367.).

Teorem 1.35. *Ako je X neprekidna d-dimenzionalna slučajna veličina s gustoćom f_X i $g : D_g \to \mathbb{R}^d$ ($D_g \subseteq \mathbb{R}^d$) Borelova funkcija takva da:*

(i) $\mathbb{P}_X(D_g) = 1$, tj, kompozicija $g \circ X$ je dobro definirana,
(ii) postoje međusobno disjunktni otvoreni skupovi D_1, D_2,... i Borelov skup N u $\mathbb{R}^d$ takvi da je $D_g = N \cup D_1 \cup D_2 \cup \cdots$ i $\mathbb{P}_X(N) = 0$, te da
(iii) restrikcije $g_k := g|_{D_k} : D_k \to S_k := g_k(D_k)$ ($k = 1, 2, \ldots$) od g su bijekcije sa svojstvom da su njihovi inverzi $g_k^{-1} : S_k \to D_k$ ($k = 1, 2, \ldots$) klase C^1,

tada je d-dimenzionalna slučajna veličina $Y = g(X)$ neprekidna s gustoćom:

$$f_Y(y) = \sum_k f_X(g_k^{-1}(y))|\det Dg_k^{-1}(y)| \cdot \mathbb{1}_{S_k}(y). \tag{1.23}$$

Teorem se može primijeniti i na slučajne varijable koje su dovoljno glatke funkcije slučajnih vektora.

Primjer 1.36. Neka su X i Y takve slučajne varijable da je slučajni vektor (X, Y) kojemu su to komponente, neprekidan s gustoćom $f = f_{X,Y}$. Tada su slučajne varijable $U = X+Y$ i $V = Y - X$ neprekidne s gustoćama

$$f_U(u) = \int_{-\infty}^{+\infty} f(x, u - x)\, dx,\ u \in \mathbb{R},\ \ f_V(v) = \int_{-\infty}^{+\infty} f(x, v + x)\, dx,\ v \in \mathbb{R}.$$

Specijalno, ako su X i Y neprekidne i nezavisne slučajne varijable s gustoćama f_X i f_Y redom, tada su i U i V neprekidne s gustoćama

$$f_U(u) = \int_{-\infty}^{+\infty} f_X(x)f_Y(u - x)\, dx,\ u \in \mathbb{R},\ \ f_V(v) = \int_{-\infty}^{+\infty} f_X(x)f_Y(v + x)\, dx,\ v \in \mathbb{R}.$$

Pokazat ćemo prvu tvrdnju. Druga tvrdnja je posljedica prve tvrdnje i činjenice da ako su X i Y neprekidne i nezavisne slučajne varijable, da je tada slučajni vektor (X, Y) neprekidan s gustoćom $f(x, y) = f_X(x)f_Y(y)$ ($x, y \in \mathbb{R}$) prema teoremu 1.30.
Neka je $g(x, y) := (x + y, y - x)$. Tada je $g : \mathbb{R}^2 \to \mathbb{R}^2$ bijekcija, $g^{-1}(u, v) = \frac{1}{2}(u - v, u + v)$ je klase C^1 i $|\det Dg^{-1}(u, v)| = \frac{1}{2}$. Prema teoremu 1.35, $(U, V) = g(X, Y)$ je neprekidan slučajni vektor s gustoćom

$$g(u, v) = f_{U,V}(u, v) = \frac{1}{2} f(\frac{1}{2}(u - v), \frac{1}{2}(u + v)),\ u, v \in \mathbb{R}.$$

Odavde slijedi da su U i V neprekidne slučajne varijable kao komponente neprekidnog slučajnog vektora s gustoćama jednakim odgovarajućim marginalnim gustoćama:

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{+\infty} g(u, v)\, dv = \int_{-\infty}^{+\infty} \frac{1}{2} f(\frac{1}{2}(u - v), \frac{1}{2}(u + v))\, dv = \int_{-\infty}^{+\infty} f(x, u - x)\, dx, \quad u \in \mathbb{R}, \\ f_V(v) &= \int_{-\infty}^{+\infty} g(u, v)\, du = \int_{-\infty}^{+\infty} \frac{1}{2} f(\frac{1}{2}(u - v), \frac{1}{2}(u + v))\, du = \int_{-\infty}^{+\infty} f(x, v + x)\, dx, \quad v \in \mathbb{R} \end{aligned}$$

pri čemu je zadnja jednakost u slučaju gustoće f_U posljedica zamjene varijable $v = 2x + u$, a u slučaju gustoće f_V je posljedica zamjene varijable $u = 2x + v$. □

Primjer 1.37. Pokažimo da slučajni vektor (X, Y) ne mora biti neprekidan ako su X i Y zavisne neprekidne slučajne varijable. Neka je X neka neprekidna slučajna varijabla. Tada je i $Y := -X$ neprekidna slučajna varijabla (jer funkcija $g : \mathbb{R} \to \mathbb{R}$, $g(x) = -x$, zadovoljava uvjete teorema 1.35, a $Y = g(X)$). Kada bi (X, Y) bio neprekidan slučajni vektor, tada bi, prema prethodnom primjeru 1.36, $X + Y$ bila neprekidna slučajna varijabla. Dakle, $\mathbb{P}(X + Y = 0) = 0$. S druge strane je očito da je $\mathbb{P}(X + Y = 0) = 1$ po definiciji od Y. Dakle, imamo kontradikciju pa je pretpostavka o neprekidnosti od (X, Y) netočna. □

Zadatak 1.38. Dokažite sljedeće tvrdnje. Neka su X i Y komponente neprekidnog slučajnog vektora (X, Y) s gustoćom $f = f_{X,Y}$. Tada su slučajne varijable $U = X \cdot Y$ i $V = Y/X$ neprekidne s gustoćama

$$f_U(u) = \int_{-\infty}^{+\infty} f(x, \frac{u}{x}) \frac{1}{|x|}\, dx, \ u \in \mathbb{R}, \ \ f_V(v) = \int_{-\infty}^{+\infty} f(x, vx)|x|\, dx, \ v \in \mathbb{R}.$$

Specijalno, ako su X i Y neprekidne i nezavisne slučajne varijable s gustoćama f_X i f_Y redom, tada su i U i V neprekidne s gustoćama

$$f_U(u) = \int_{-\infty}^{+\infty} f_X(x) f_Y(\frac{u}{x}) \frac{1}{|x|}\, dx, \ u \in \mathbb{R}, \ \ f_V(v) = \int_{-\infty}^{+\infty} f_X(x) f_Y(vx)|x|\, dx, \ v \in \mathbb{R}.$$

Zadatak 1.39. Neka je $X \sim N(\mu, \sigma^2)$ normalna slučajna varijabla (vidi zadatak 1.16), te a i $b \neq 0$ realni brojevi. Pokažite da je tada $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$.

Zadatak 1.40. Neka su $X \sim N(\mu, \sigma^2)$ i $Y \sim N(\nu, \tau^2)$ dvije nezavisne normalne slučajne varijable, te $a, b, c \in \mathbb{R}$ takvi da je $|a| + |b| > 0$. Dokažite da je $aX + bY + c \sim N(a\mu + b\nu + c, a^2\sigma^2 + b^2\tau^2)$.

Zadatak 1.41. Neka je X k-dimenzionalni neprekidni slučajni vektor s gustoćom f_X, A regularna matrica reda k i $a \in \mathbb{R}^k$ zadani vektor. Pokažite da je tada $Y = a + AX$ k-dimenzionalni neprekidni slučajni vektor s gustoćom:

$$f_Y(y) = f_X(A^{-1}(y - a)) \cdot \frac{1}{|\det A|}, \ \ y \in \mathbb{R}^k.$$

Zadatak 1.42. Neka su X_1 i X_2 nezavisne standardne normalno distribuirane slučajne varijable (vidi zadatak 1.16), te neka je $\rho \in \langle -1, 1 \rangle$ realan broj. Nadalje, neka je

$$Y = (Y_1, Y_2) := (X_1, \rho X_1 + \sqrt{1 - \rho^2} X_2)$$

dvodimenzionalan slučajni vektor. Dokažite da Y ima bivarijatnu normalnu razdiobu s parametrima $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$ i ρ (vidi primjer 1.26). Nadalje, pokažite da su Y_1 i Y_2 standardne normalne slučajne varijable, a ako je $\rho \neq 0$, da su tada Y_1 i Y_2 zavisne.

Zadatak 1.43. Neka su $\beta > 0$ i $\alpha_i > 0$ $(i = 1, 2, \ldots, n)$ realni brojevi i neka je $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. Dokažite da, ako su $X_i \sim \Gamma(\alpha_i, \beta)$ $(i = 1, 2, \ldots, n)$ nezavisne gama distribuirane slučajne varijable (vidi zadatak 1.17), da je tada

$$Y = X_1 + X_2 + \cdots + X_n \sim \Gamma(\alpha, \beta).$$

Zadatak 1.44. Neka su X_1, X_2,..., X_n nezavisne standardno normalno distribuirane slučajne varijable (vidi zadatak 1.16), te $Y = X_1^2 + X_2^2 \cdots + X_n^2$. Pokažite da je $Y \sim \Gamma(\frac{n}{2}, 2)$ (vidi zadatak 1.17). Za slučajnu varijablu X koja ima gama razdiobu $\Gamma(\frac{n}{2}, 2)$ za neki prirodan broj n, kažemo da ima *χ^2-razdiobu s n stupnjeva slobode* i pišemo $X \sim \chi^2(n)$.
(**Uputa:** Koristi se matematička indukcija po n. Za bazu indukcije, uočite da je $\Gamma(1/2) = \sqrt{\pi}$.)

Zadatak 1.45. Neka su $Z \sim N(0,1)$ i $V \sim \chi^2(n)$ dvije nezavisne slučajne varijable (vidi zadatke 1.16 i 1.44). Pokažite da je $T = Z/\sqrt{V/n}$ neprekidna slučajna varijabla i odredite joj gustoću $f = f_T$. Za neprekidnu slučajnu varijablu X koja ima dobivenu gustoću f kažemo da ima *Studentovu t-razdiobu s n stupnjeva slobode* i pišemo $X \sim t(n)$. Studentovu razdiobu s $n = 1$ stupnjeva slobode zovemo još *jediničnom Cauchyjevom razdiobom*.

Zadatak 1.46. Neka su $X_1 \sim \chi^2(m_1)$ i $X_2 \sim \chi^2(m_2)$ dvije nezavisne slučajne varijable (vidi zadatak 1.44). Pokažite da je

$$F = \frac{X_1/m_1}{X_2/m_2}$$

neprekidna slučajna varijabla i odredite joj gustoću $f = f_F$. Za neprekidnu slučajnu varijablu X koja ima dobivenu gustoću f kažemo da ima *F-razdiobu s uređenim parom stupnjeva slobode* (m_1, m_2) i pišemo $X \sim F(m_1, m_2)$.

Zadatak 1.47. Neka su X_1, X_2,..., X_n nezavisne jednako distribuirane Bernoullijeve slučajne varijable s istom vjerojatnosti uspjeha p (vidi zadatak 1.20), te $Y = X_1 + X_2 \cdots + X_n$.
(a) Pokažite da je $Y \sim b(n, p)$.
(b) Neka je $k \in \{1, 2, \ldots, n\text{–}1\}$. Tada je uvjetna razdioba slučajnog vektora $(X_1, X_2, \ldots, X_n)$ uz dano $Y = k$, diskretna uniformna na skupu svih uređenih n-torki elemenata 0 i 1 takvih da svaka n-torka ima točno k elemenata jednakih 1. Dokažite.

Zadatak 1.48. Neka su X_1, X_2,..., X_n nezavisne Poissonove slučajne varijable (vidi zadatak 1.22) takve da su $X_i \sim P(\lambda_i)$ za $\lambda_i > 0$ $(i = 1, 2, \ldots, n)$, te $Y = X_1 + X_2 \cdots + X_n$.
(a) Pokažite da je $Y \sim P(\lambda)$ za $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
(b) Neka je m prirodan broj i $p_i := \lambda_i/\lambda$ $(i = 1, 2, \ldots, n)$. Tada je uvjetna razdioba slučajnog vektora $(X_1, X_2, \ldots, X_n)$ uz dano $Y = m$, polinomna razdioba $M(m; p_1, p_2, \ldots, p_n)$ (vidi primjer 1.25). Dokažite.

Zadatak 1.49. Neka je X neprekidna slučajna varijabla s gustoćom f i nosačem te gustoće jednakim čitavom skupu $\mathbb{R}$. Ako je $Y = \operatorname{tg} X$, pokažite da je Y neprekidna slučajna varijabla i izračunajte joj gustoću.

Zadatak 1.50. Neka su $X \sim \Gamma(\alpha, \beta)$ i $Y \sim \Gamma(\gamma, \beta)$ nezavisne gama distribuirane slučajne varijable s parametrima $\alpha > 0$, $\beta > 0$ i $\gamma > 0$ (vidjeti zadatak 1.17). Odredite zakon razdiobe slučajne varijable $U = X/(X + Y)$.

1.7 Matematičko očekivanje

Neka je X slučajna varijabla definirana na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Definiramo $X^+ := \max\{0, X\}$ i $X^- := \max\{0, -X\}$. Očito je $X = X^+ - X^-$ i $|X| = X^+ + X^-$. Budući da su X^+ i X^- nenegativne izmjerive funkcije, *Lebesgue-Stieltjesov integrali od X^+ i X^- (u odnosu na mjeru $\mathbb{P}$)*, $\mathbb{E}X^+ := \int_\Omega X^+ \, d\mathbb{P}$ i $\mathbb{E}X^- := \int_\Omega X^+ \, d\mathbb{P}$, postoje i vrijednosti su im u skupu $\mathbb{R} \cup \{+\infty\}$ (vidjeti [9], §10, str. 289.).

Definicija 1.51. *Kažemo da slučajna varijabla X* **ima matematičko očekivanje** *ako je barem jedan od integrala $\mathbb{E}X^+$ i $\mathbb{E}X^-$ konačan. U tom slučaju je* matematičko očekivanje

$$\mathbb{E}X := \mathbb{E}X^+ - \mathbb{E}X^-.$$

Očito je da, ako slučajna varijabla X ima matematičko očekivanje, da je tada $\mathbb{E}X \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. Drugim riječima (vidjeti [9], §10, str. 290.), X ima matematičko očekivanje ako Lebesgue-Stieltjesov integral od X (u odnosu na mjeru $\mathbb{P}$) $\int_\Omega X \, d\mathbb{P}$ postoji, i u tom slučaju je

$$\mathbb{E}X = \int_\Omega X \, d\mathbb{P}. \tag{1.24}$$

Za $A \in \mathcal{F}$, dakle bilo koji događaj A, $\int_A X \, d\mathbb{P}$ je Lebesgue-Stieltjesov integral od restrikcije $X|_A : A \to \mathbb{R}$ na restringiranom prostoru mjere $(A, \mathcal{F} \cap A, \mathbb{P})$ ako postoji. U tom slučaju je

$$\mathbb{E}(X \cdot \mathbb{1}_A) = \int_A X \, d\mathbb{P}. \tag{1.25}$$

Za izračunati matematičko očekivanje slučajne varijable X dovoljno je znati njen zakon razdiobe budući da prema *teoremu o zamjeni varijabli* vrijedi da je

$$\mathbb{E}g(X) = \int_\Omega g(X) \, d\mathbb{P} = \int_\mathbb{R} g \, d\mathbb{P}_X \tag{1.26}$$

za neku Borelovu funkciju $g : \mathbb{R} \to \mathbb{R}$. Još više, oba navedena integrala postoje i jednaka su ako jedan od njih postoji (vidjeti teorem 10.11 u [9], str. 306.). Nadalje, ako je X neprekidna slučajna varijabla s gustoćom f_X, tada je izraz (1.26), u stvari jednak

$$\mathbb{E}g(X) = \int_\mathbb{R} g(x) f_X(x) \, dx, \tag{1.27}$$

a ako je X diskretna slučajna varijabla s gustoćom f_X, tada je (1.26) jednak

$$\mathbb{E}g(X) = \sum_x g(x) f_X(x). \tag{1.28}$$

Definicija 1.52. *Kažemo da slučajna varijabla X ima* **konačno** *matematičko očekivanje ako je $\mathbb{E}|X| < +\infty$.*

Budući da je $\mathbb{E}|X| = \mathbb{E}X^+ + \mathbb{E}X^-$, odmah slijedi da X ima konačno matematičko očekivanje u skladu s prethodnom definicijom ako i samo ako su $\mathbb{E}X^+ < +\infty$ i $\mathbb{E}X^- < +\infty$, a to je ekvivalentno tome da X ima matematičko očekivanje u skladu s definicijom 1.51 i da je $\mathbb{E}X \in \mathbb{R}$.

Primjer 1.53. Neka je $A \in \mathcal{F}$ i $x_0 \in \mathbb{R}$ zadani broj. Tada je $X = x_0 \cdot \mathbb{1}_A$ diskretna slučajna varijabla sa zakonom razdiobe

$$X \sim \begin{pmatrix} 0 & x_0 \\ 1 - \mathbb{P}(A) & \mathbb{P}(A) \end{pmatrix}.$$

$\mathbb{E}X$ postoji jer je, ili $X \geq 0$ (pa je $\mathbb{E}X^- = 0$) ili $-X \geq 0$ (pa je $\mathbb{E}X^+ = 0$) ovisno o tome je li $x_0 \geq 0$ ili $x_0 < 0$. Prema formuli (1.28),

$$\mathbb{E}X = 0 \cdot (1 - \mathbb{P}(A)) + x_0 \cdot \mathbb{P}(A) = x_0 \mathbb{P}(A).$$

Specijalno, ako je $X \equiv x_0$ za zadani broj $x_0 \in \mathbb{R}$, dakle konstanta ili je X degenerirana slučajna varijabla takva da je $\mathbb{P}(X = x_0) = 1$ (vidi zadatak 1.19), tada je

$$\mathbb{E}X = \mathbb{E}x_0 = x_0.$$

□

Primjer 1.54. Neka je X neprekidna slučajna varijabla s jediničnom Cauchyjevom distribucijom. Prema zadatku 1.45, gustoća od X je $f(x) = 1/(\pi(1+x^2))$, $x \in \mathbb{R}$. Izračunajmo

$$\mathbb{E}|X| = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)}\,dx = \frac{2}{\pi}\int_{0}^{+\infty} \frac{x}{1+x^2}\,dx = \frac{1}{\pi}\log(1+x^2)|_0^{+\infty} = +\infty.$$

Druga jednakost je posljedica parnosti podintegralne funkcije. Dakle, X nema konačno matematičko očekivanje. Budući da je

$$\mathbb{E}X^+ = \int_{-\infty}^{+\infty} \max\{0,x\}f(x)\,dx = \int_{-\infty}^{+\infty} \max\{0,-x\}f(-x)\,dx = \int_{-\infty}^{+\infty} \max\{0,-x\}f(x)\,dx = \mathbb{E}X^-$$

X nema matematičko očekivanje uopće. U gornjem nizu nejednakosti koristili smo zamjenu varijabli $x \mapsto -x$ i parnost gustoće. □

Definicija 1.55. *Neka je $X = (X_1, X_2, \ldots, X_k)$ slučajni vektor. Kažemo da X* **ima matematičko očekivanje** *ako svaka komponenta X_1, X_2,..., X_k tog vektora ima matematičko očekivanje, te u tom slučaju definiramo* **matematičko očekivanje od** *X kao vektor*

$$\mathbb{E}X := (\mathbb{E}X_1, \mathbb{E}X_2, \ldots, \mathbb{E}X_k).$$

Iz prethodnih definicija je jasno da će slučajni vektor X imati *konačno* matematičko očekivanje ako će mu svaka komponenta imati konačno matematičko očekivanje.

Označimo sa $\mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ skup svih slučajnih varijabli definiranih na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ koje imaju konačno matematičko očekivanje. Nadalje, podsjetimo sa da kada kažemo da neka tvrdnja vrijedi *gotovo sigurno* ili, kraće, *g.s.*, da to znači da ta ista tvrdnja vrijedi na događaju vjerojatnosti jedan. Na primjer, ako su X i Y slučajne varijable, tada je $X = Y$ g.s. ako je $\mathbb{P}(X = Y) = 1$, ili $X < Y$ g.s. ako je $\mathbb{P}(X \le Y) = 1$.

Vrijedi sljedeći teorem.

Teorem 1.56. *Neka su $X, Y, X_n \in \mathcal{L}^1$, $n \in \mathbb{N}$, te $\alpha, \beta \in \mathbb{R}$ proizvoljni. Tada vrijedi:*

(i) $\alpha X + \beta Y \in \mathcal{L}^1$ *i* $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}X + \beta\mathbb{E}Y$.
(ii) $X \le Y$ *g.s.* $\Rightarrow \mathbb{E}X \le \mathbb{E}Y$
(iii) $|\mathbb{E}X| \le \mathbb{E}|X|$
(iv) $(X \ge 0$ *g.s. i* $\mathbb{E}X = 0) \Rightarrow X = 0$ *g.s.*
(v) *Za realne brojeve* $1 \le r < s$ *vrijedi da je* $\mathbb{E}|X|^r \le 1 + \mathbb{E}|X|^s$.
(vi) *Ako su X i Y nezavisne slučajne varijable, tada je* $XY \in \mathcal{L}^1$ *i* $\mathbb{E}(XY) = \mathbb{E}X \cdot \mathbb{E}Y$.
(vii) (Jensenova nejednakost) *Ako je $\phi : \mathbb{R} \to \mathbb{R}$ konveksna funkcija takva da je $\phi(X) \in \mathcal{L}^1$, tada je $\phi(\mathbb{E}X) \le \mathbb{E}\phi(X)$.*

Svojstva $(i - iii)$ (vidjeti [9], teorem 10.1, str. 291.-292.) kažu da je $\mathcal{L}^1$ vektorski potprostor realnog vektorskog prostora svih slučajnih varijabli definiranih na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, pri čemu je operator matematičkog očekivanja monotoni linearni funkcional na $\mathcal{L}^1$. Isti prostor se može smatrati i (pseudo)normiranim prostorom s normom

$\|X\|_1 := \mathbb{E}|X|$. Naime, nejednakost trokuta trivijalno slijedi iz nejednakosti trokuta za apsolutnu vrijednost $|\cdot|$ i monotonosti očekivanja (ii). Nenegativna homogenost slijedi iz istog svojstva za apsolutnu vrijednost i linearnosti očekivanja (i). Iz primjera 1.53 i svojstva (iv) (vidjeti [9], propozicija 10.4, str. 294.) slijedi pseudopozitivna definitnost. Da bi dobili normu bez prefiksa "pseudo", potrebno je $\mathcal{L}^1$ zamijeniti (a zapravo reprezentirati) s kvocijentnim skupom u odnosu na relaciju ekvivalencije "biti jednak gotovo sigurno" (kraće, "= (g.s.)") u $\mathcal{L}^1$, u oznaci $L^1 = L^1(\Omega, \mathcal{F}, \mathbb{P})$ (vidjeti §6 u [7], str. 50.-51.). Za svojstvo (v) vidjeti dokaz propozicije 6.1 u [9], str. 137., a za svojstvo (vi) teorem 11.5 u [9], str. 357. Jensenova nejednakost (vii) je posljedica sljedeće činjenice o konveksnim funkcijama. Naime, za svaki $c \in \mathbb{R}$ postoji $m = m(c) \in \mathbb{R}$ takav da je $\phi(x) \geq m(x-c) + \phi(c)$ za svaki $x \in \mathbb{R}$. Budući da konveksna funkcija u svakoj točki c domene ima derivacije slijeva $D_-\phi(c)$ i zdesna $D_+\phi(c)$, za m je dovoljno uzeti bilo koji element segmenta $[D_-\phi(c), D_+\phi(c)]$ (vidjeti [11], 6.6., str. 61.). Ako za c uzmemo $c = \mathbb{E}X$, tada imamo da je $\phi(X) \geq m(X - \mathbb{E}X) + \phi(\mathbb{E}X)$ pa primjenom monotonosti (ii) i linearnosti (i) očekivanja slijedi tvrdnja (vii).

Vezano uz konveksne funkcije, u drugom poglavlju trebat će nam sljedeća napomena (vidjeti napomenu u [11], 6.6, str. 61.).

Napomena 1.57. Budući da za svaku konveksnu funkciju $\phi : \mathbb{R} \to \mathbb{R}$ očito vrijedi da je $\phi(x) = \sup_{q\in\mathbb{Q}}(D_-\phi(q)(x-q) + \phi(q))$ za svaki $x \in \mathbb{R}$ i neprekidna je, postoje nizovi realnih brojeva $(a_n, n \in \mathbb{N})$ i $(b_n, n \in \mathbb{N})$ takvi da je

$$(\forall x \in \mathbb{R})\ \phi(x) = \sup_n (a_n x + b_n).$$

Svojstvo (ii) iz prethodnog teorema 1.56 zovemo *monotonost matematičkog očekivanja*. Primijetimo da pomoću te tvrdnje i činjenice da, ako je $\mathbb{P}(N) = 0$, slijedi da je $\mathbb{E}(X\mathbb{1}_N) = 0$, možemo pokazati da, ako je $X = Y$ g.s. i $X \in \mathcal{L}^1$, da je $Y \in \mathcal{L}^1$ i $\mathbb{E}X = \mathbb{E}Y$. Obrati tih tvrdnji općenito ne vrijede, ali vrijedi sljedeći teorem.

Teorem 1.58. *Neka su $X, Y \in \mathcal{L}^1$ i u odnosu na σ-algebru $\mathcal{F}$ izmjerive slučajne varijable. Ako je za svaki $A \in \mathcal{F}$, $\mathbb{E}(X\mathbb{1}_A) \leq \mathbb{E}(Y\mathbb{1}_A)$, tada je $X \leq Y$ g.s. Specijalno, ako je za svaki $A \in \mathcal{F}$, $\mathbb{E}(X\mathbb{1}_A) = \mathbb{E}(Y\mathbb{1}_A)$, tada je $X = Y$ g.s.*

Teorem 1.58 ćemo dokazati u narednom poglavlju dokazujući g.s. jedinstvenost uvjetnog očekivanja.

Za razvoj statističke teorije posebno su važni rezultati o graničnom ponašanju matematičkog očekivanja iskazani u sljedećem teoremu. Naime, radi se o dobro poznatim Lebesgueovim teoremima o monotonoj i dominiranoj konvergenciji, te Fatouovoj lemi iz teorije Lebesgueovog integrala, iskazanima u terminima nizova slučajnih varijabli i matematičkog očekivanja (vidjeti [9], §10.2, teoremi 10.2, 10.3 i 10.6, str. 295.-297.).

Podsjetimo se, niz slučajnih veličina $(X_n)_{n\in\mathbb{N}}$ *konvergira gotovo sigurno* prema slučajnoj veličini X ako je $\lim_n X_n(\omega) = X(\omega)$ za $\omega \in E$, gdje je E događaj vjerojatnosti jedan. Kraće pišemo $\lim_n X_n = X$ g.s. Slično, $\underline{\lim}_n X_n = X$ g.s. znači da je $\underline{\lim}_n X_n(\omega) = X(\omega)$ za sve $\omega \in E$ gdje je E događaj vjerojatnosti jedan.

Iskažimo sada teoreme o konvergenciji slučajnih varijabli.

Teorem 1.59. *Neka su X, X_n, $n \in \mathbb{N}$, slučajne varijable. Tada vrijede sljedeće tvrdnje.*

(i) (Lebesgueov teorem o monotonoj konvergenciji) *Ako je za svaki n, $0 \leq X_n \leq X_{n+1}$ g.s. i $\lim_n X_n = X$ g.s., tada je $(\mathbb{E}X_n)_{n\in\mathbb{N}}$ rastući niz u $\overline{\mathbb{R}}$ i $\lim_n \mathbb{E}X_n = \mathbb{E}X$.*

(ii) (Fatouova lema) *Ako je za svaki n, $0 \leq X_n$ g.s. i $\underline{\lim}_n X_n = X$ g.s., tada je $\mathbb{E}X \leq \underline{\lim}_n \mathbb{E}X_n$.*

(*iii*) (Lebesgueov teorem o dominiranoj konvergenciji) *Ako je za svaki n, $|X_n| \leq V$ g.s. za neku slučajnu varijablu $V \in \mathcal{L}^1$ i $\lim_n X_n = X$ g.s., tada je $\lim_n \mathbb{E}X_n = \mathbb{E}X$.*

Primijetite da za tvrdnje (*i*), Lebesgueov teorem o monotonoj konvergenciji (kraće LTMK), i (*ii*), Fatouovu lemu, nije nužno pretpostaviti da je niz $(X_n, n \in \mathbb{N})$ iz $\mathcal{L}^1$, ali, budući da se radi o nenegativnim slučajnim varijablama, njihovo matematičko očekivanje postoji i postoji matematičko očekivanja njihovog limesa, odnosno gomilišta. U oba slučaja, nije nužno da su te granične vrijednosti u $\mathcal{L}^1$. S druge strane, u tvrdnji (*iii*), Lebesgueovom teoremu o dominiranoj konvergenciji (kraće LTDK), niz $(X_n, n \in \mathbb{N})$ je nužno u $\mathcal{L}^1$ i, uz uvjete tog teorema, g.s. po točkama limes tog niza je također nužno u $\mathcal{L}^1$.

Iz teorema 1.56, svojstvo (*v*), slijedi da, ako za slučajnu varijablu X vrijedi da ima konačan *drugi moment*, tj. da je $\mathbb{E}(X^2) < +\infty$, da tada X ima konačno matematičko očekivanje $\mu_X := \mathbb{E}X$ i da slučajna varijabla $(X - \mu_X)^2$ ima konačno matematičko očekivanje. Još više, ukoliko je Y slučajna varijabla takva da ima konačan drugi moment, tada primjenom monotonosti očekivaja (uz oznaku $\mu_Y := \mathbb{E}Y$) slijedi

$$\mathbb{E}|(X - \mu_X)(Y - \mu_Y)| \leq \frac{1}{2}(\mathbb{E}(X - \mu_X)^2 + \mathbb{E}(Y - \mu_Y)^2) < +\infty.$$

Dakle, sljedeća definicija je korektna.

Definicija 1.60. *Neka su X i Y slučajne varijable takva da je $\mathbb{E}(X^2) < +\infty$ i $\mathbb{E}(Y^2) < +\infty$. Označimo s $\mu_X = \mathbb{E}X$ i $\mu_Y = \mathbb{E}Y$.*

(*a*) **Varijanca od** X *je nenegativni broj*

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

(*b*) **Standardna devijacija od** X *jest nenegativan broj*

$$\mathrm{std}(X) := \sqrt{\mathbb{V}\mathrm{ar}X}.$$

(*c*) **Kovarijanca od** X **i** Y *je broj*

$$\mathrm{cov}(X, Y) := \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

(*d*) *Ako su $\sigma_X = \mathrm{std}(X) > 0$ i $\sigma_Y = \mathrm{std}(Y) > 0$, tada su dobro definirane* **standardizirane varijable od** X **i** Y *izrazima:*

$$Z_X := \frac{X - \mu_X}{\sigma_X}, \quad Z_Y := \frac{Y - \mu_Y}{\sigma_Y}.$$

U tom slučaju je **koeficijent korelacije od** X **i** Y *broj*

$$\mathrm{corr}(X, Y) := \mathrm{cov}(Z_X, Z_Y).$$

Primijetite da, ako je $\mathbb{E}(X^2) < +\infty$, $\mathbb{E}(Y^2) < +\infty$ i $a, b \in \mathbb{R}$, da je tada i

$$\mathbb{E}[(X + Y)^2] \leq 2(\mathbb{E}(X^2) + \mathbb{E}(Y^2)) < +\infty \ \text{ i } \ \mathbb{E}[(aX + b)^2] \leq 2(a^2\mathbb{E}(X^2) + b^2) < +\infty.$$

Dakle, ako postoje i konačne su varijance od X i Y, postoje i konačne su varijance od $X + Y$ i $aX + b$. Iz prethodne definicije i teorema 1.56 slijede sljedeća svojstva varijance i kovarijance.

Teorem 1.61. *Neka su X, Y i Z takve slučajne varijable da je $\mathbb{E}(X^2) < +\infty$, $\mathbb{E}(Y^2) < +\infty$, i $\mathbb{E}(Z^2) < +\infty$, te $a, b \in \mathbb{R}$ proizvoljni. Tada vrijedi:*

(i) $\mathbb{V}\mathrm{ar}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$
(ii) $\mathbb{V}\mathrm{ar}(aX+b) = a^2\mathbb{V}\mathrm{ar}X$
(iii) $\mathbb{V}\mathrm{ar}X = 0$ *ako i samo ako je X degenerirana slučajna varijabla.*
(iv) $\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}X + 2\,\mathrm{cov}(X,Y) + \mathbb{V}\mathrm{ar}Y$
(v) *Ako su X i Y nezavisne slučajne varijable, tada je* $\mathrm{cov}(X,Y) = 0$ *i* $\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}X + \mathbb{V}\mathrm{ar}Y$.
(vi) $\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X \cdot \mathbb{E}Y$
(vii) $\mathrm{cov}(aX+b,Y) = a\,\mathrm{cov}(X,Y)$
$(viii)$ $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$
(ix) $\mathrm{cov}(X+Y,Z) = \mathrm{cov}(X,Z) + \mathrm{cov}(Y,Z)$
(x) $\mathrm{cov}(X,X) = \mathbb{V}\mathrm{ar}X$

Primjer 1.62. Neka je X normalna slučajna varijabla $N(\mu,\sigma^2)$ (vidjeti zadatak 1.16). Tada je, prema zadatku 1.39, $X = \mu + \sigma Z$ gdje je Z jedinična normalna varijabla. Vrijedi:

$$\mathbb{E}|Z| = \int\limits_{-\infty}^{+\infty} |x|\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = \frac{2}{\sqrt{2\pi}}\int\limits_{0}^{+\infty} xe^{-\frac{1}{2}x^2}\,dx = -\frac{2}{\sqrt{2\pi}}\,e^{-\frac{1}{2}x^2}\Big|_0^{+\infty} = \frac{2}{\sqrt{2\pi}} < +\infty.$$

Dakle, Z ima konačno matematičko očekivanje i

$$\mathbb{E}Z = \int\limits_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = 0$$

jer je podintegralna funkcija neparna, a područje integracije je simetrično oko nule. Prema teoremu 1.56 (i) X ima konačno matematičko očekivanje i

$$\mathbb{E}X = \mu + \sigma\mathbb{E}Z = \mu.$$

Nadalje,

$$\begin{aligned}\mathbb{V}\mathrm{ar}Z &= \mathbb{E}(Z^2) = \int\limits_{-\infty}^{+\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = \frac{2}{\sqrt{2\pi}}\int\limits_{0}^{+\infty} x^2e^{-\frac{1}{2}x^2}\,dx = \text{(parcijalna integracija)}\\ &= \frac{2}{\sqrt{2\pi}}\Big(-xe^{-\frac{1}{2}x^2}\Big|_0^{+\infty} + \int\limits_{0}^{+\infty} e^{-\frac{1}{2}x^2}\,dx\Big) = \frac{2}{\sqrt{2\pi}}\int\limits_{0}^{+\infty} e^{-\frac{1}{2}x^2}\,dx = \int\limits_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = 1.\end{aligned}$$

Prema teoremu 1.61 (ii),

$$\mathbb{V}\mathrm{ar}X = \sigma^2\mathbb{V}\mathrm{ar}Z = \sigma^2.$$

Dakle, interpretacija parametara normalne varijable $N(\mu,\sigma^2)$ je da su to njeno matematičko očekivanje (μ) i varijanca (σ^2). Primijetite da je ovdje Z standardizirana varijabla od X. □

Primjer 1.63. Neka slučajni vektor (X,Y) ima bivarijatnu normalnu razdiobu s parametrima $\mu_1 \in \mathbb{R}$, $\mu_2 \in \mathbb{R}$, $\sigma_1 > 0$, $\sigma_2 > 0$ i $\rho \in \langle -1, 1\rangle$ (vidjeti primjer 1.26). U primjeru 1.26 je pokazano da su komponente X i Y normalno distribuirane slučajne varijable: $X \sim N(\mu_1,\sigma_1^2)$ i $Y \sim N(\mu_2,\sigma_2^2)$. Na osnovi prethodnog primjera 1.62, zaključujemo da su μ_1 i μ_2 njihova matematička očekivanja, a σ_1 i σ_2 njihove standardne devijacije. Dakle, komponente od (X,Y) imaju konačne varijance pa postoji i konačna je njihova kovarijanca $\mathrm{cov}(X,Y)$. Izračunajmo ju. U tu svrhu, primijetimo da, ako su U, V nezavisne jedinične normalne slučajne varijable, da slučajan vektor

$$(X',Y') := (\sigma_1 U + \mu_1, \sigma_2\rho U + \sigma_2\sqrt{1-\rho^2}\cdot V + \mu_2)$$

ima bivarijatnu normalnu razdiobu s istim parametrima kao vektor (X, Y). To se može pokazati korištenjem rezultata iz zadataka 1.41 i 1.42. Budući da su (X, Y) i (X', Y') jednako distribuirani, vrijedi da je

$$\mathrm{cov}(X, Y) = \int_{\mathbb{R}^2} (x - \mu_1)(y - \mu_2) f(x, y)\, dx\, dy = \mathrm{cov}(X', Y'),$$

gdje je f funkcija gustoće zajedničke bivarijatne normalne razdiobe. Korištenjem linearnosti matematičkog očekivanja, nezavisnosti od U i V i teorema 1.56 (vi), te činjenice da je $\mathbb{E}U = \mathbb{E}V = 0$, te $\mathbb{E}(U^2) = \mathbb{V}\mathrm{ar}U = 1$ i $\mathbb{E}(V^2) = \mathbb{V}\mathrm{ar}V = 1$ slijedi:

$$\begin{aligned}
\mathrm{cov}(X, Y) &= \mathrm{cov}(X', Y') = \mathbb{E}[(X' - \mu_1)(Y' - \mu_2)] = \\
&= \mathbb{E}[\sigma_1 U \cdot \sigma_2(\rho U + \sqrt{1-\rho^2} \cdot V)] = \\
&= \sigma_1\sigma_2(\rho \mathbb{E}(U^2) + \sqrt{1-\rho^2} \cdot \mathbb{E}(UV)) = \\
&= \sigma_1\sigma_2(\rho \mathbb{V}\mathrm{ar}(U^2) + \sqrt{1-\rho^2} \cdot \mathbb{E}U \cdot \mathbb{E}V)) = \\
&= \rho\sigma_1\sigma_2.
\end{aligned}$$

Dakle,

$$\rho = \frac{\mathrm{cov}(X, Y)}{\sigma_1\sigma_2} \overset{(\mathrm{tm.}1.61\ (vii))}{=} \mathrm{cov}\left(\frac{X - \mu_1}{\sigma_1}, \frac{Y - \mu_2}{\sigma_2}\right) = \mathrm{corr}(X, Y)$$

je koeficijent korelacije od X i Y. □

Primjer 1.64. Neka je X slučajna varijabla s gama distribucijom $\Gamma(\alpha, \beta)$ (vidjeti zadatak 1.17). Budući da joj je gustoća jednaka $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$ za $x > 0$ i $f(x) = 0$ za $x \le 0$, slijedi da je $X \ge 0$ g.s. pa X ima matematičko očekivanje. Pokažimo da je ono konačno tako da ga izračunamo. Slično za varijancu.

$$\begin{aligned}
\mathbb{E}X &= \int_0^{+\infty} x \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx = \frac{\Gamma(\alpha+1)\beta}{\Gamma(\alpha)} \int_0^{+\infty} \frac{1}{\Gamma(\alpha+1)\beta^{\alpha+1}} x^{(\alpha+1)-1} e^{-\frac{x}{\beta}}\, dx = \\
&= \frac{\Gamma(\alpha+1)\beta}{\Gamma(\alpha)} \cdot 1 = \alpha\beta
\end{aligned}$$

gdje smo u predzadnjoj jednakosti iskoristili činjenicu da je podintegralna funkcija gustoća $\Gamma(\alpha+1, \beta)$-razdiobe, a u zadnjoj jednakosti svojstvo Γ-funkcije da je $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Slično,

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_0^{+\infty} x^2 \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx = \frac{\Gamma(\alpha+2)\beta^2}{\Gamma(\alpha)} \int_0^{+\infty} \frac{1}{\Gamma(\alpha+2)\beta^{\alpha+2}} x^{(\alpha+1)-1} e^{-\frac{x}{\beta}}\, dx = \\
&= \frac{\Gamma(\alpha+2)\beta^2}{\Gamma(\alpha)} \cdot 1 = \alpha(\alpha+1)\beta^2
\end{aligned}$$

pa je, prema teoremu 1.61 (i),

$$\mathbb{V}\mathrm{ar}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2 = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2.$$

Specijalno, odavde slijedi da za eksponencijalno distribuiranu slučajnu varijablu $Y \sim \mathrm{Exp}(\lambda) \equiv \Gamma(1, \frac{1}{\lambda})$ vrijedi

$$\mathbb{E}Y = \frac{1}{\lambda}, \quad \mathbb{V}\mathrm{ar}X = \frac{1}{\lambda^2},$$

a za χ^2-distribuiranu slučajnu varijablu $V \sim \chi^2(n) \equiv \Gamma(\frac{n}{2}, 2)$ s n stupnjeva slobode,

$$\mathbb{E}V = n, \quad \mathbb{V}\mathrm{ar}V = 2n.$$

□

Primjer 1.65. Neka je $X \sim b(n,p)$ binomna slučajna varijabla (zadatak 1.20). Tada, prema zadatku 1.47, postoje nezavisne jednako distribuirane Bernoullijeve slučajne varijable X_1, X_2,..., X_n s vjerojatnosti uspjeha p takve da je $X_1 + X_2 + \cdots + X_n \sim b(n,p)$, dakle, takvih da je njihov zbroj po distribuciji jednak X. Budući da je $\mathbb{E}X_i = 0 \cdot (1-p) + 1 \cdot p = p$, zbog linearnosti očekivanja,

$$\mathbb{E}X = \mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}X_1 + \cdots + \mathbb{E}X_n = np.$$

Primijetite da je prva jednakost posljedica činjenice da matematičko očekivanje *ovisi samo o zakonu razdiobe od* X. Nadalje, budući da je

$$\mathbb{V}\mathrm{ar}X_i = \mathbb{E}(X_i^2) - (\mathbb{E}X_i)^2 = 0^2 \cdot (1-p) + 1^2 \cdot p - p^2 = p(1-p),$$

i X_i, $i = 1, 2, \ldots, n$, su nezavisne, prema teoremu 1.61 (v),

$$\mathbb{V}\mathrm{ar}X = \mathbb{V}\mathrm{ar}(X_1 + X_2 + \cdots + X_n) = \mathbb{V}\mathrm{ar}X_1 + \cdots + \mathbb{V}\mathrm{ar}X_n = np(1-p).$$

□

Primjer 1.66. Neka je $X \sim G(p)$ geometrijska slučajna varijabla s vjerojatnosti uspjeha $p \in \langle 0, 1\rangle$ (vidjeti zadatak 1.21). Tada je $X > 0$ g.s. pa postoji matematičko očekivanje. Prema formuli 1.28, uz pokratu $q := 1 - p$,

$$\mathbb{E}X = \sum_{k=1}^{+\infty} kpq^k = pq \sum_{k=1}^{+\infty} kq^{k-1} = pq \frac{d}{dy} \frac{1}{1-y}\bigg|_{y=q} = pq \frac{1}{(1-q)^2} = \frac{q}{p} = \frac{1-p}{p},$$

pri čemu je $\sum_{k=1}^{+\infty} kq^{k-1}$ jednako derivaciji geometrijskog reda $\sum_{k=0}^{+\infty} y^k = \frac{1}{1-y}$ kao apsolutno konvergentnog reda potencija definiranog na domeni $D = \langle -1, 1\rangle$ i uniformno konvergentnog ka funkciji $y \mapsto \frac{1}{1-y}$, u točki $y = q \in D$. Da bi pokazali da X ima konačnu varijancu i izračunali je, prvo izračunajmo

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{k=1}^{+\infty} k(k-1)pq^k = \sum_{k=2}^{+\infty} k(k-1)pq^k = pq^2 \sum_{k=2}^{+\infty} k(k-1)q^{k-2} = \\
&= pq^2 \frac{d^2}{dy^2} \frac{1}{1-y}\bigg|_{y=q} = pq^2 \frac{2}{(1-q)^3} = 2\frac{q^2}{p^2}.
\end{aligned}$$

Koristeći teorem 1.61 (i) i linearnost očekivanja,

$$\mathbb{V}\mathrm{ar}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}X - (\mathbb{E}X)^2 = 2\frac{q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q}{p^2} = \frac{1-p}{p^2}.$$

□

Zadatak 1.67. Neka je $X \sim P(\lambda)$ Poissonova slučajna varijabla s parametrom $\lambda > 0$ (vidjeti zadatak 1.22). Pokažite da X ima konačne matematičko očekivanje i varijancu, te da vrijedi

$$\lambda = \mathbb{E}X = \mathbb{V}\mathrm{ar}X.$$

Zadatak 1.68. Neka je $X \sim H(n, M, N)$ hipergeometrijska slučajna varijabla s parametrima $n < M < N$ $(n, M, N \in \mathbb{N}$, vidjeti zadatak 1.20). Pokažite da X ima konačne matematičko očekivanje i varijancu, te da je

$$\mathbb{E}X = np, \quad \mathbb{V}\mathrm{ar}X = np(1-p)\frac{1 - \frac{n}{N}}{1 - \frac{1}{N}},$$

gdje je $p := M/N$.

Zadatak 1.69. Neka je $X \sim U(a,b)$ uniformno distribuirana slučajna varijabla na segmentu $[a,b]$ (vidjeti zadatak 1.14). Budući da je $a < X < b$ g.s., X ima konačne matematičko očekivanje i varijancu, ali i momente svakoga reda. Pokažite da je za svaki prirodan broj n,

$$\mathbb{E}(X^n) = \frac{1}{n+1}(a^n + a^{n-1}b + \cdots + ab^{n-1} + b^n).$$

Broj $\mathbb{E}(X^n)$ zovemo *n-ti moment od X*. Koristeći dobivenu formulu, pokažite da je

$$\mathbb{E}X = \frac{a+b}{2}, \quad \mathbb{V}\mathrm{ar}X = \frac{1}{12}(b-a)^2.$$

Zadatak 1.70. Neka je $X \sim B(\alpha,\beta)$ beta distribuirana slučajna varijabla s parametrima $\alpha > 0$ i $\beta > 0$ (vidjeti zadatak 1.18). Budući da je $0 < X < 1$ g.s., X ima konačne sve momente, pa specijalno i matematičko očekivanje i varijancu. Vrijedi

$$\mathbb{E}X = \frac{\alpha}{\alpha+\beta}, \quad \mathbb{V}\mathrm{ar}X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Zadatak 1.71. Neka je $X \sim t(n)$ Studentova slučajna varijabla s n stupnjeva slobode ($n \in \mathbb{N}$, vidjeti zadatak 1.45). Ispitajte za koje n postoje očekivanje i varijanca od X i izračunajte ih.

Zadatak 1.72. Neka slučajna varijabla $X \sim F(m,n)$ ima F-razdiobu s (m,n) stupnjeva slobode ($m,n \in \mathbb{N}$, vidjeti zadatak 1.46). Ispitajte za koje m,n postoje očekivanje i varijanca od X i izračunajte ih.

Zadatak 1.73. Neka je X pozitivna slučajna varijabla takva da je $\mathbb{E}X < +\infty$ i $\mathbb{P}(X \neq \mathbb{E}X) > 0$. Dokažite da postoji $\mathbb{E}[\log X]$ i da vrijedi *stroga Jensenova nejednakost* za logaritamsku funkciju:

$$\mathbb{E}[\log X] < \log \mathbb{E}X.$$

Poglavlje 2

Uvjetno matematičko očekivanje

2.1 Prostor $\mathcal{L}^2$

Označimo sa $\mathcal{L}^2 \equiv \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ prostor svih slučajnih varijabli s konačnim drugim momentom definiranih na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Nije teško pokazati da je $\mathcal{L}^2$ (realan) vektorski prostor, potprostor od $\mathcal{L}^1 \equiv \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Primijetite da slučajna varijabla ima konačan drugi moment ako i samo ako ima konačnu varijancu (u tom slučaju implicitno slijedi da joj je i matematičko očekivanje konačno). Nadalje, ako su $X, Y \in \mathcal{L}^2$, tada je

$$\mathbb{E}(|XY|) \leq \frac{1}{2}(\mathbb{E}(X^2) + \mathbb{E}(Y^2)) < +\infty$$

pa je na $\mathcal{L}^2$ dobro definiran *pseudoskalarni produkt* $\langle \cdot, \cdot \rangle : \mathcal{L}^2 \times \mathcal{L}^2 \to \mathbb{R}$ formulom

$$\langle X, Y \rangle := \mathbb{E}[XY], \quad X, Y \in \mathcal{L}^2. \tag{2.1}$$

Naime, pomoću teorema 1.56 nije teško pokazati da je $\langle \cdot, \cdot \rangle$ u oba argumenta linearno preslikavanje (*bilinearnost*) takvo da je za sve $X \in \mathcal{L}^2$, $\langle X, X \rangle = E(X^2) \geq 0$ (*nenegativna definitnost*) i $\langle X, X \rangle = 0$ ako i samo ako je $X = 0$ g.s. (*pseudopozitivna definitnost*). Budući da je $\mathcal{L}^2$ vektorski prostor snabdjeven pseudoskalarnim produktom, kažemo da je $\mathcal{L}^2$ *pseudounitarni prostor*. Pseudoskalarni produkt inducira *pseudonormu* $\| \cdot \|_2 : \mathcal{L}^2 \to [0, +\infty\rangle$ na $\mathcal{L}^2$ formulom

$$\|X\|_2 := \sqrt{\langle X, X \rangle} = \sqrt{\mathbb{E}[X^2]}, \quad X \in \mathcal{L}^2. \tag{2.2}$$

Iz definicije pseudonorme (2.2) slijedi da i u pseudounitarnom prostoru vrijedi *relacija paralelograma*:

$$(\forall X, Y \in \mathcal{L}^2) \; \|X + Y\|_2^2 + \|X - Y\|_2^2 = 2\|X\|_2^2 + 2\|Y\|_2^2. \tag{2.3}$$

Definicija 2.1. *Neka je $(X_n)_{n \in \mathbb{N}}$ niz u $\mathcal{L}^2$.*

(a) *Kažemo da niz $(X_n)_{n \in \mathbb{N}}$* **konvergira u $\mathcal{L}^2$** *ako postoji $X \in \mathcal{L}^2$ takav da je*

$$\lim_n \|X_n - X\|_2 = 0.$$

X je limes danog niza, u oznaci $X = (L^2) \lim_n X_n$ ili $X_n \xrightarrow{L^2} X$.

(b) *Niz $(X_n)_{n \in \mathbb{N}}$ je* **Cauchyjev niz u $\mathcal{L}^2$** *ako vrijedi*

$$(\forall \varepsilon > 0)(\exists n_\varepsilon \in \mathbb{N})(\forall r, s \in \mathbb{N}) \; r, s \geq n_\varepsilon \Rightarrow \|X_r - X_s\|_2 < \varepsilon.$$

Za niz konvergentan u $\mathcal{L}^2$ još kažemo da je *konvergentan u srednjem reda dva* i pišemo $(m^2)\lim_n X_n = X$ ili $X_n \overset{m^2}{\to} X$ (vidjeti [9], str. 319.-320.). Tu konvergenciju treba razlikovati od drugih tipova konvergencije slučajnih varijabli (po vjerojatnosti, gotovo sigurno, po distribuciji). Nadalje, primijetimo da, ako niz konvergira u $\mathcal{L}^2$, da mu je tada limes *gotovo sigurno jedinstven* (kraće: *g.s. jedinstven*), tj. ako su X i X' dva limesa istoga niza u $\mathcal{L}^2$, tada je $\mathbb{P}(X = X') = 1$.

Na kvocijentnom prostoru L^2 klasa ekvivalencija elemenata od $\mathcal{L}^2$ definiranih u odnosu na relaciju ekvivalencije *biti jednak gotovo sigurno* ("= g.s.") prirodno se pomoću pseudonorme (2.2) u $\mathcal{L}^2$ definira norma u L^2, a pomoću relacije paralelograma definira se odgovarajući skalarni produkt. Naime, norma klase ekvivalencije je jednaka pseudonormi (bilo kojeg) predstavnika te klase, pa je i odgovarajući skalarni produkt dviju klasa ekvivalencije jednak pseudoskalarnom produktu (2.1) njihovih predstavnika. Pokazuje se da je unitarni prostor L^2 *potpun*, tj. da je svaki Cauchyjev niz u tom prostoru i konvergentan u L^2 (vidjeti Teorem 3 u [7], V., §3.1, str. 208.). Dakle, L^2 je *Hilbertov prostor*. Odavde slijedi da svaki Cauchyjev niz u $\mathcal{L}^2$ i konvergira u $\mathcal{L}^2$, pa je $\mathcal{L}^2$ potpun pseudounitarni prostor. Također ćemo reći da je $\mathcal{L}^2$ Hilbertov prostor.

Teorem 2.2. (Schwartz-Cauchyjeva nejednakost). *Neka su X, Y slučajne varijable iz $\mathcal{L}^2$. Tada je*

$$(\mathbb{E}(XY))^2 \le \mathbb{E}(X^2) \cdot \mathbb{E}(Y^2). \tag{2.4}$$

Jednakost u (2.4) vrijedi ako i samo ako postoje brojevi $\lambda, \mu \in \mathbb{R}$ koji nisu oba jednaka nuli i takvi da je $\lambda X + \mu Y = 0$ g.s.

Dokaz. Ako je $\mathbb{E}(X^2) = 0$, tada je $X = 0$ g.s. pa je $\mathbb{E}(X^2) = \mathbb{E}(XY) = 0$. Dakle, vrijedi (2.4) kao jednakost. Nadalje, $\mathbb{E}(X^2) = 0$ ako i samo ako je $X = 0$ g.s., odnosno, ako i samo ako je $\lambda X + 0 \cdot Y = 0$ g.s. za bilo koji $\lambda \in \mathbb{R} \setminus \{0\}$ i $\mu = 0$ pa vrijedi i druga tvrdnja. Pretpostavimo da je $\mathbb{E}(X^2) > 0$. Neka je $f : \mathbb{R} \to \mathbb{R}$ funkcija definirana sa

$$f(\lambda) := \|Y - \lambda X\|_2^2 = (\text{def. } 2.2) = \|X\|_2^2\lambda^2 - 2\langle X, Y\rangle\lambda + \|Y\|_2^2, \ \ \lambda \in \mathbb{R}.$$

Budući da je $\|X\|_2^2 = \mathbb{E}(X^2) > 0$, f je nenegativna konveksna kvadratna funkcija. To znači da je njena diskriminanta D nužno nepozitivna, pri čemu postiže vrijednost 0 ako i samo ako je $D = 0$. Dakle,

$$D = (-2\langle X, Y\rangle)^2 - 4\|X\|_2^2\|Y\|_2^2 \le 0 \Leftrightarrow (\langle X, Y\rangle)^2 \le \|X\|_2^2\|Y\|_2^2,$$

što je (2.4). Jednakost se postiže u (2.4), a što je ekvivalentno sa $D = 0$, ako i samo ako postoji realan broj λ takav da je $0 = f(\lambda) = \|Y - \lambda X\|_2^2$, dakle, ako i samo ako je $Y = \lambda X$ g.s. Primijetimo da je ta g.s. jednakost ekvivalentna sa $\lambda X + \mu Y = 0$ g.s. za $\mu = -1$. □

Nejednakost (2.4) je *Schwartz-Cauchyjeva nejednakost* u pseudounitarnom prostoru $\mathcal{L}^2$. Dokaz teorema 2.2 je isti za sve (pseudo)unitarne prostore.

Navedimo jednu važnu primjenu Schwartz-Cauchyjeve nejednakosti u statistici. Ako u (2.4) umjesto X i Y iz $\mathcal{L}^2$ stavimo $X - \mu_X$ i $Y - \mu_Y$, gdje su $\mu_X = \mathbb{E}X$ i $\mu_Y = \mathbb{E}Y$, te uzmemo u obzir definicije (vidjeti definiciju 1.60) kovarijance $\text{cov}(X, Y)$ od X i Y i njihovih standardnih devijacija $\sigma_X = \text{std}(X)$ i $\sigma_Y = \text{std}(Y)$, tada imamo sljedeću inačicu Schwartz-Cauchyjeve nejednakosti:

$$|\text{cov}(X, Y)| \le \sigma_X \cdot \sigma_Y. \tag{2.5}$$

U skladu s teoremom 2.2, jednakost u (2.5) vrijedi ako i samo ako postoje realni brojevi λ i μ ne oba jednaka nuli i takvi da je

$$\lambda(X - \mu_X) + \mu(Y - \mu_Y) = 0 \text{ g.s.} \ \Leftrightarrow \ \lambda X + \mu Y + \nu = 0 \text{ g.s. za } \nu \equiv -\lambda\mu_X - \mu\mu_Y.$$

Dakle, jednakost u (2.5) vrijedi ako i samo ako postoje realni brojevi λ, μ, ν takvi da je $|\lambda| + |\mu| \neq 0$ i $\lambda X + \mu Y + \nu = 0$ g.s.

Korolar 2.3. *Neka je (X,Y) slučajan vektor takav da su X, Y nedegenerirane slučajne varijable konačnih varijanci. Tada za koeficijent korelacije od X i Y, $\operatorname{corr}(X,Y)$, vrijedi*

(i) $-1 \leq \operatorname{corr}(X,Y) \leq 1$
(ii) $\operatorname{corr}(X,Y) = 1$ *ako i samo ako postoje $\alpha, \beta \in \mathbb{R}$ takvi da je $\beta > 0$ i $Y = \beta X + \alpha$ g.s.*
(iii) $\operatorname{corr}(X,Y) = -1$ *ako i samo ako postoje $\alpha, \beta \in \mathbb{R}$ takvi da je $\beta < 0$ i $Y = \beta X + \alpha$ g.s.*

Dokaz. Budući da su X i Y nedegenerirane slučajne varijable konačnih varijanci, vrijedi da je $\mathbb{V}\mathrm{ar}X > 0$, $\mathbb{V}\mathrm{ar}Y > 0$, pa su i standardne devijacije $\sigma_X = \operatorname{std}(X)$, $\sigma_Y = \operatorname{std}(Y)$ konačne i pozitivne. Dakle, realan broj $\rho_{X,Y} := \operatorname{cov}(X,Y)/(\sigma_X \cdot \sigma_Y)$ je dobro definiran i za njega vrijedi da je $|\rho_{X,Y}| \leq 1$ prema Schwartz-Cauchyjevoj nejednakosti (2.5). S druge strane i $\operatorname{corr}(X,Y)$ je dobro definiran broj i za njega, prema definiciji 1.60 i teoremu 1.61 (vii), vrijedi

$$\operatorname{corr}(X,Y) = \operatorname{cov}\left(\frac{X-\mu_X}{\sigma_X}, \frac{Y-\mu_Y}{\sigma_Y}\right) = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \rho_{X,Y},$$

odakle odmah slijede nejednakosti (i). Budući da jednakost u Schwartz-Cauchyjevoj nejednakosti (2.5) vrijedi ako i samo ako postoje realni brojevi λ, μ, ν takvi da je $\lambda X + \mu Y + \nu = 0$ g.s. i $\lambda \neq 0$ ili $\mu \neq 0$, imamo da je $|\operatorname{corr}(X,Y)| = 1$ ako i samo ako je $\lambda X + \mu Y + \nu = 0$ g.s. za neke $\lambda, \mu, \nu \in \mathbb{R}$ pri čemu su $\lambda \neq 0$ i $\mu \neq 0$. Naime, kada bi $\lambda = 0$ tada bi Y bila degenerirana slučajna varijabla, a u slučaju da je $\mu = 0$, X bi bila degenerirana varijabla. Dakle, $|\operatorname{corr}(X,Y)| = 1$ ako i samo ako postoje realni brojevi α i $\beta \neq 0$ takvi da je $Y = \beta X + \alpha$ g.s. U tom slučaju je (prema teoremu 1.61), $\sigma_Y = \sqrt{\mathbb{V}\mathrm{ar}Y} = \sqrt{\beta^2 \mathbb{V}\mathrm{ar}X} = |\beta|\sigma_X$ i $\operatorname{cov}(X,Y) = \beta \mathbb{V}\mathrm{ar}(X) = \beta\, \sigma_X^2$ pa je

$$\operatorname{corr}(X,Y) = \rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{\beta\, \sigma_X^2}{|\beta| \sigma_X^2} = \operatorname{sign}\beta.$$

Dakle, $\beta > 0$ ako i samo ako je $\operatorname{corr}(X,Y) = 1$, i $\beta < 0$ ako i samo ako je $\operatorname{corr}(X,Y) = -1$, čime je upotpunjen dokaz tvrdnji (ii) i (iii). □

Napomena 2.4. Primijetite da iz dokaza korolara 2.3 slijedi da vrijede sljedeće tvrdnje:
$(ii)'$ $\operatorname{corr}(X,Y) = 1$ ako i samo ako postoje $\gamma, \delta \in \mathbb{R}$ takvi da je $\delta > 0$ i $X = \delta Y + \gamma$ g.s.
$(iii)'$ $\operatorname{corr}(X,Y) = -1$ ako i samo ako postoje $\gamma, \delta \in \mathbb{R}$ takvi da je $\delta < 0$ i $X = \delta Y + \gamma$ g.s.

2.2 Teorem o ortogonalnoj projekciji

Neka su $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor i $\mathcal{L}^2 \equiv \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ prostor svih slučajnih varijabli na $(\Omega, \mathcal{F}, \mathbb{P})$ s konačnim drugim momentima, uveden u prethodnom potpoglavlju. Kažemo da je podskup $\mathcal{K}$ od $\mathcal{L}^2$ *potpun potprostor od* $\mathcal{L}^2$ ako je vektorski potprostor vektorskog prostora $\mathcal{L}^2$, pseudounitaran je s istim pseudoskalarnim produktom kao u $\mathcal{L}^2$ i potpun kao pseudounitarni prostor. Ekvivalentno, potprostor $\mathcal{K}$ od $\mathcal{L}^2$ je potpun ako i samo ako je *zatvoren* u $\mathcal{L}^2$.

Ako su $X, Y \in \mathcal{L}^2$ takvi da je $\langle X, Y \rangle = \mathbb{E}(XY) = 0$, tada kažemo da su elementi X i Y *ortogonalni* i pišemo: $X \perp Y$.

Teorem 2.5. (Teorem o ortogonalnoj projekciji). *Neka je $Y \in \mathcal{L}^2$ i neka je $\mathcal{K}$ potpun potprostor od $\mathcal{L}^2$. Tada postoji g.s. jedinstvena slučajna varijabla $\hat{Y} \in \mathcal{K}$ takva da vrijede sljedeće dvije ekvivalentne tvrdnje:*

(1) $\mathbb{E}(Y-\hat{Y})^2 = \min_{Z\in\mathcal{K}} \mathbb{E}(Y-Z)^2$;
(2) $(\forall Z \in \mathcal{K})\ \mathbb{E}(\hat{Y}Z) = \mathbb{E}(YZ)$.

Slučajnu varijablu $\hat{Y}$ koja, za zadanu varijablu $Y \in \mathcal{L}^2$ i potpun potprostor $\mathcal{K}$ od $\mathcal{L}^2$, zadovoljava uvjete (1) i (ili) (2), zovemo *verzija ortogonalne projekcije od Y na $\mathcal{K}$.*

Dokaz. Zapišimo tvrdnje (1) i (2) u terminima pseudonorme i pseudoskalarnog produkta, odnosno, interpretirajmo (2) pomoću pojma ortogonalnosti:

(1) $\|Y-\hat{Y}\|_2 = \min_{Z\in\mathcal{K}} \|Y-Z\|_2$;

(2) $(\forall Z \in \mathcal{K})\ Y-\hat{Y} \perp Z$.

Dokaz egzistencije $\hat{Y}$. Neka je $\Delta := \inf_{Z\in\mathcal{K}} \|Y-Z\|_2$. Budući da je konstantna funkcija 0 u $\mathcal{K}$, vrijedi da je $0 \le \Delta \le \|Y\|_2 < +\infty$. Nadalje, budući da je Δ najveća donja međa skupa $\{\|Y-Z\|_2 : Z \in \mathcal{K}\}$, postoji niz $(Y_n)_{n\in\mathbb{N}}$ u $\mathcal{K}$ takav da je $\lim_n \|Y-Y_n\|_2 = \Delta$. Za proizvoljne $r, s \in \mathbb{N}$, primijenimo relaciju paralelograma (2.3) na slučajne varijable $Y-(Y_r+Y_s)/2$ i $(Y_r-Y_s)/2$, a zatim, prvo, iskoristimo činjenicu da je $\|Y-(Y_r+Y_s)/2\|_2 \ge \Delta$ (jer je $(Y_r+Y_s)/2 \in \mathcal{K}$), a onda da je $\lim_n \|Y-Y_n\|_2 = \Delta$ uzimanjem prvo limesa po r, a zatim po s pri čemu dobijamo sljedeći niz implikacija:

$$\begin{array}{rrcl} & \|Y-Y_r\|_2^2 + \|Y-Y_s\|_2^2 & = & 2\|Y-(Y_r+Y_s)/2\|_2^2 + 2\|(Y_r-Y_s)/2\|_2^2 \\ \Rightarrow & \|Y-Y_r\|_2^2 + \|Y-Y_s\|_2^2 & \ge & 2\Delta^2 + \frac{1}{2}\|Y_r-Y_s\|_2^2 \\ \Rightarrow & \Delta^2 + \|Y-Y_s\|_2^2 & \ge & 2\Delta^2 + \frac{1}{2}\overline{\lim}_r\|Y_r-Y_s\|_2^2 \\ \Rightarrow & 2\Delta^2 & \ge & 2\Delta^2 + \frac{1}{2}\overline{\lim}_s\overline{\lim}_r\|Y_r-Y_s\|_2^2 \ge 2\Delta^2 \end{array}$$

Dakle, $\overline{\lim}_{r,s}\|Y_r-Y_s\|_2 = \lim_{r,s}\|Y_r-Y_s\|_2 = 0$ odakle slijedi da je $(Y_n)_{n\in\mathbb{N}}$ Cauchyjev niz u potpunom prostoru $\mathcal{K}$. Dakle, postoji $\hat{Y} \in \mathcal{K}$ takav da je $\hat{Y} = (L^2)\lim_n Y_n$. Preostaje pokazati da za $\hat{Y}$ vrijedi (1). Nakon toga ćemo pokazati da su (1) i (2) ekvivalentne tvrdnje, pa će za $\hat{Y}$ vrijediti i tvrdnja (2).
Budući da je za svaki $n \in \mathbb{N}$, prema nejednakosti trokuta,

$$\|Y-\hat{Y}\|_2 \le \|Y-Y_n\|_2 + \|Y_n-\hat{Y}\|_2,$$

slijedi da je, uzimanjem limesa po n, prema teoremu o sendviču, $\|Y-\hat{Y}\|_2 \le \Delta$. S druge strane, budući da je $\hat{Y} \in \mathcal{K}$, vrijedi da je $\Delta \le \|Y-\hat{Y}\|_2$. Dakle, $\|Y-\hat{Y}\|_2 = \Delta$ pa vrijedi tvrdnja (1).

Pretpostavimo da vrijedi (1). Ako je $Z = 0$ g.s., $Y-\hat{Y} \perp Z$ po definiciji ortogonalnosti. Neka je $Z \in \mathcal{K}$ bilo koji element takav da $Z \ne 0$ g.s. Za sve $\lambda \in \mathbb{R}$, $\hat{Y}+\lambda Z \in \mathcal{K}$ pa zbog (1) vrijedi:

$$\begin{array}{rl} & (\forall\lambda \in \mathbb{R})\ \ \|Y-\hat{Y}-\lambda Z\|_2^2 \ge \|Y-\hat{Y}\|_2^2 \\ \Leftrightarrow & (\forall\lambda \in \mathbb{R})\ \ \lambda^2\|Z\|_2^2 - 2\lambda\langle Z, Y-\hat{Y}\rangle \ge 0 \\ \Rightarrow & \langle Z, Y-\hat{Y}\rangle = 0. \end{array}$$

Naime, da bi konveksna kvadratna funkcija s realnim nultočkama bila nenegativna, nultočke se moraju podudarati. Budući da je jedna nultočka očito jednaka nuli, i druga nultočka, $\lambda_2 = 2\langle Z, Y-\hat{Y}\rangle/\|Z\|_2^2$, je nužno jednaka 0. Dakle, $Y-\hat{Y} \perp Z$. Budući da je $Z \in \mathcal{K}$ bio proizvoljan, vrijedi tvrdnja (2).

Obratno, pretpostavimo da vrijedi (2). Za proizvoljno $Z \in \mathcal{K}$ je $\hat{Y} - Z \in \mathcal{K}$ pa prema (2) vrijedi da je $\langle Y - \hat{Y}, \hat{Y} - Z\rangle = 0$. Prema tome,

$$\begin{aligned}\|Y - Z\|_2^2 &= \|(Y - \hat{Y}) + (\hat{Y} - Z)\|_2^2 = \|Y - \hat{Y}\|_2^2 + 2\langle Y - \hat{Y}, \hat{Y} - Z\rangle + \|\hat{Y} - Z\|_2^2 = \\ &= \|Y - \hat{Y}\|_2^2 + \|\hat{Y} - Z\|_2^2.\end{aligned}$$

Dakle, dokazali smo da vrijedi *Pitagorin poučak*:

$$\|Y - Z\|_2^2 = \|Y - \hat{Y}\|_2^2 + \|\hat{Y} - Z\|_2^2. \tag{2.6}$$

Sada imamo da je za sve $Z \in \mathcal{K}$,

$$\|Y - Z\|_2^2 = \|Y - \hat{Y}\|_2^2 + \|\hat{Y} - Z\|_2^2 \geq \|Y - \hat{Y}\|_2^2 \geq \Delta^2$$

odakle slijedi da je

$$\Delta = \inf_{Z \in \mathcal{K}} \|Y - Z\|_2 \geq \|Y - \hat{Y}\|_2 \geq \Delta \ \Rightarrow \Delta = \|Y - \hat{Y}\|_2$$

pa vrijedi tvrdnja (1).

Dokaz g.s. jedinstvenosti od $\hat{Y}$. Pretpostavimo da je $\tilde{Y} \in \mathcal{K}$ neki drugi element za koji vrijede (1) i (2). Tada primjenom (1) na $\tilde{Y}$ i $\hat{Y}$, te Pitagorinog poučka (2.6) za $Z = \tilde{Y}$, slijedi

$$\Delta^2 = \|Y - \tilde{Y}\|_2^2 = \|Y - \hat{Y}\|_2^2 + \|\hat{Y} - \tilde{Y}\|_2^2 = \Delta^2 + \|\hat{Y} - \tilde{Y}\|_2^2.$$

Dakle, $\|\hat{Y} - \tilde{Y}\|_2 = 0$ pa je $\tilde{Y} = \hat{Y}$ g.s. □

Primijetite da je iz dokaza očito da teorem o ortogonalnoj projekciji vrijedi u svakom (pseudo)unitarnom prostoru. Sustav jednakosti iz tvrdnje (2) zove se *normalne jednadžbe*. Pomoću njih se računaju ortogonalne projekcije.

Primjer 2.6. *Regresijski pravac i najbolji linearni predviditelj.* Neka su X i Y nedegenerirane slučajne varijable konačnih varijanci i neka su definirane na istome vjerojatnosnom prostoru. Cilj je naći afinu funkciju $x \mapsto \beta x + \alpha$ takvu da $\hat{Y} := \beta X + \alpha$ *najbolje aproksimira* Y u smislu da je *srednjekvadratna greška aproksimacije* najmanja, tj. da vrijedi

$$\mathbb{E}(Y - \beta X - \alpha)^2 = \min_{a,b \in \mathbb{R}} \mathbb{E}(Y - bX - a)^2. \tag{2.7}$$

Pravac $y = \beta x + \alpha$ zove se *regresijski pravac*, a slučajna varijabla $\hat{Y} = \beta X + \alpha$ *najbolji linearni predviditelj od* Y *pomoću* X. Grešku predviđanja $Y - \hat{Y}$ zovemo *rezidual*. Dakle, reziduali su slučajne varijable. Pomoću teorema o ortogonalnoj projekciji 2.5 možemo pokazati da takav predviditelj (a time i pravac) postoji, a rješavanjem normalnih jednadžbi i izračunati ga. Pokažimo kako. Definirajmo potprostor $\mathcal{K}$ od $\mathcal{L}^2$ kao linearnu ljusku konstantne varijable 1 i slučajne varijable X. Budući da je X nedegenerirana varijabla, 1 i X su linearno nezavisni elementi od $\mathcal{L}^2$ pa je $\mathcal{K}$ dvodimenzionalan potprostor. Dakle, konačne je dimenzije pa je i potpun. Primijetite da je uvjet na minimalnost pogreške (2.7), u stvari, uvjet (1) iz teorema 2.5. Dakle, prema teoremu o ortogonalnoj projekciji, $\hat{Y} = \beta X + \alpha$ postoji tako da zadovoljava uvjet (2.7), g.s. je jedinstven, a α, $\beta \in \mathbb{R}$ su takvi da vrijede normalne jednadžbe na elementima baze od $\mathcal{K}$, tj. za $Z \in \{1, X\}$ (vidjeti (2) u iskazu teorema 2.5):

$$\left.\begin{aligned}\mathbb{E}(1 \cdot \hat{Y}) &= \mathbb{E}(1 \cdot Y) \\ \mathbb{E}(X \cdot \hat{Y}) &= \mathbb{E}(X \cdot Y)\end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned}\beta\,\mathbb{E}X + \alpha &= \mathbb{E}Y \\ \beta\,\mathbb{E}(X^2) + \alpha\,\mathbb{E}X &= \mathbb{E}(XY).\end{aligned}\right.$$

Uvođenjem uobičajenih oznaka za očekivanja, standardne devijacije i koeficijent korelacije od X i Y (vidjeti dokaz korolara 2.3), te primjenom teorema 1.61 (i) i (vi), normalne jednadžbe prelaze u ekvivalentan oblik:

$$\left.\begin{array}{rcl} \beta\mu_X + \alpha & = & \mu_Y \\ \beta(\mathbb{V}\mathrm{ar}X + \mu_X^2) + \alpha\mu_X & = & \mathrm{cov}(X,Y) + \mu_X\mu_Y \end{array}\right\} \Leftrightarrow \left\{ \begin{array}{rcl} \alpha & = & \mu_Y - \beta\mu_X \\ \beta & = & \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}. \end{array}\right.$$

Time smo odredili regresijski pravac. Primijetite da je regresijski pravac jednoznačno određen budući da normalne jednadžbe čine Cramérov sustav linearan po nepoznanicama α i β. To je posljedica linearne nezavisnosti varijabli 1 i X.

Iskoristimo ovaj primjer i za interpretaciju koeficijenta korelacije. Korištenjem linearnosti matematičkog očekivanja, teorema 1.61 i upravo izračunane koeficijente regresijskog pravca, zaključujemo da su srednjekvadratna greška aproksimacije (ekvivalentno, kvadrat norme reziduala) i varijanca najboljeg lineranog predviditelja:

$$\mathbb{E}(Y-\hat{Y})^2 = \mathbb{E}(Y-\beta X-\alpha)^2 = (1-\rho_{X,Y}^2)\sigma_Y^2 = (1-\mathrm{corr}^2(X,Y))\mathbb{V}\mathrm{ar}Y \tag{2.8}$$

$$\mathbb{V}\mathrm{ar}\hat{Y} = \mathbb{V}\mathrm{ar}(\beta X+\alpha) = \beta^2\mathbb{V}\mathrm{ar}X = \left(\rho_{X,Y}\frac{\sigma_Y}{\sigma_X}\right)^2\sigma_X^2 = \mathrm{corr}^2(X,Y)\mathbb{V}\mathrm{ar}Y. \tag{2.9}$$

Primijetite da su dobivene relacije za srednjekvadratnu grešku i varijancu predviditelja konzistentne s korolarom 2.3: najmanja je vrijednost srednjekvadratne greške kada je koeficijent korelacije jednak -1 ili 1. Tada je njen iznos 0 odakle slijedi da je $Y = \beta X + \alpha$ g.s. Dakako, tada je $\mathbb{V}\mathrm{ar}\hat{Y} = \mathbb{V}\mathrm{ar}Y$. Obratno, ako je Y g.s. afina funkcija od X, tada je srednjekvadratna greška jednaka 0, pa je koeficijent korelacije jednak -1 ili 1, a afina funkcija se podudara s regresijskim pravcem. S druge strane, vrijednost srednjekvadratne greške je najveća kada je koeficijent korelacije jednak 0. Tada kažemo da su X i Y *nekorelirane* slučajne varijable. Primijetite da su, prema teoremu 1.61, nezavisne slučajne varijable ujedno i nekorelirane (obrat općenito ne vrijedi, vidjeti primjer iza definicije nekoreliranosti u [9], str. 139.-140.). U slučaju da su X i Y nekorelirane varijable, najbolji linearni predviditelj od Y pomoću X je degenerirana varijabla jednaka μ_Y, dakle, ne ovisi o X. To je jasno iz izraza za koeficijent smjera β i relacije (2.9). Nadalje, primijetite da što je apsolutna vrijednost koeficijenta korelacije veća, to se Y od svog lineranog predviditelja pomoću X manje razlikuje u smislu srednjekvadratne greške. Kažemo da su X i Y *negativno korelirane* varijable ako je $\mathrm{corr}(X,Y) < 0$ (što je ekvivalentno tome da je koeficijent smjera β regresijskog pravca negativan), a X i Y su *pozitivno korelirane* slučajne varijable ako je $\mathrm{corr}(X,Y) > 0$ (ekvivalentno, ako je $\beta > 0$). Zbog toga je koeficijent korelacije bezdimenzijska veličina kojom se mjeri koliko su i kako X i Y *linearno* korelirani. U skladu s napomenom 2.4, istu interpretaciju koeficijenta korelacije bismo imali ako bismo aproksimirali X pomoću afine funkcije od Y. To i očekujemo budući da je definicija koeficijenta korelacije simetrična u X i Y.

Na kraju, zbrajajući desne strane izraza za srednjekvadratnu grešku aproksimacije (2.8) i varijancu najboljeg lineranog predviditelja (2.9) slijedi *raščlamba varijance od Y* na varijancu najboljeg linearnog predviditelja $\hat{Y}$ i srednjekvadratnu grešku aproksimacije $\hat{Y}$ od Y:

$$\mathbb{V}\mathrm{ar}Y = \mathbb{V}\mathrm{ar}\hat{Y} + \mathbb{E}(Y-\hat{Y})^2. \tag{2.10}$$

Dobivena relacija (2.10) je osnova za konstrukciju i interpretaciju *tablice analize varijance* (kraće, tablice ANOVA) koja se koristi u regresijskoj analizi (vidjeti, npr. poglavlje XI.3 i jednakost (45) na str. 283 u [8]). Nadalje, pomoću nje se razumije *metoda redukcije varijance* koja se primijenjuje u povećanju učinkovitosti računanja određenih integrala po dijelovima neprekidnih funkcija metodom *Monte Carlo* kadagod je to moguće (vidjeti, npr. poglavlje 4. u [4]). □

Zadatak 2.7. Za potpun potprostor $\mathcal{K}$ od $\mathcal{L}^2$ i $Y \in \mathcal{L}^2$ označimo sa $P_{\mathcal{K}}Y$ jednu verziju ortogonalne projekcije od Y na $\mathcal{K}$. Dokažite sljedeće tvrdnje.
(i) Ako je $Y \in \mathcal{K}$, tada je $P_{\mathcal{K}}Y = Y$ g.s.
(ii) Ako je $1 \in \mathcal{K}$, tada je za svaki $\hat{Y} = P_{\mathcal{K}}Y$ g.s., $\mathbb{E}\hat{Y} = \mathbb{E}Y$.
(iii) Ako su $X, Y \in \mathcal{K}$ i $\alpha, \beta \in \mathbb{R}$, tada je $P_{\mathcal{K}}(\alpha X + \beta Y) = \alpha P_{\mathcal{K}}X + \beta P_{\mathcal{K}}Y$ g.s.
(iv) Ako je $\mathcal{M}$ takav potpun potprostor od $\mathcal{L}^2$ da je $\mathcal{M} \subset \mathcal{K}$, tada je $P_{\mathcal{K}}(P_{\mathcal{M}}Y) = P_{\mathcal{M}}(P_{\mathcal{K}}Y) = P_{\mathcal{M}}Y$ g.s. Specijalno, $P_{\mathcal{K}}(P_{\mathcal{K}}Y) = P_{\mathcal{K}}Y$ g.s.
(v) Vrijedi *Pitagorin poučak*: za svaki $Y \in \mathcal{L}^2$ je $\|Y\|_2^2 = \|P_{\mathcal{K}}Y\|_2^2 + \|Y - P_{\mathcal{K}}Y\|_2^2$.

2.3 Postojanje i g.s. jedinstvenost uvjetnog matematičkog očekivanja

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor. Označimo kao i prije sa $\mathcal{L}^1 \equiv \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ prostor svih slučajnih varijabli definiranih na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ i konačnog matematičkog očekivanja. $\mathcal{L}^1$ je potpuni pseudonormirani prostor s normom $\|X\|_1 \equiv \mathbb{E}(|X|)$ ($X \in \mathcal{L}^1$). Nadalje, neka je $\mathcal{L}^2 \equiv \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ Hilbertov prostor svih slučajnih varijabli definiranih na istom vjerojatnosnom prostoru i s konačnim drugim momentom. Na početku ovoga poglavlja smo spomenuli da je $\mathcal{L}^2$ vektorski potprostor od $\mathcal{L}^1$ (vidjeti također i prethodno poglavlje). Konvergentan i Cauchyjev niz u prostoru $\mathcal{L}^1$ se definiraju na isti način kao u definiciji 2.1 samo u odnosu na (pseudo)normu $\|\cdot\|_1$ umjesto $\|\cdot\|_2$. Ako je $X \in \mathcal{L}^2$, tada je primjenom Schwartz-Cauchyjeve nejednakosti (2.4) (uz $Y = 1$):

$$\|X\|_1 = \mathbb{E}(|X|) = \mathbb{E}(|X| \cdot 1) \leq \sqrt{\mathbb{E}(X^2) \cdot 1} = \|X\|_2.$$

Dakle, ako je niz $(X_n)_{n \in \mathbb{N}}$ iz $\mathcal{L}^2$ konvergentan ili je Cauchyjev niz u $\mathcal{L}^2$, tada je on takav niz i u $\mathcal{L}^1$. Prema tome, ako je $\mathcal{K}$ potpun potprostor od $\mathcal{L}^2$, tada je $\mathcal{K}$ *potpun pseudonormirani potprostor* od $\mathcal{L}^1$ u smislu da je $\mathcal{K}$ vektorski potprostor od $\mathcal{L}^1$ koji je pseudonormirani prostor obzirom na normu prostora $\mathcal{L}^1$ i potpun obzirom na tu normu. Kraće kažemo da je $\mathcal{K}$ (i) potpun potprostor od $\mathcal{L}^1$.

Nadalje, neka je $\mathcal{G}$ σ-algebra podskupova od Ω takva da je $\mathcal{G} \subset \mathcal{F}$. Kažemo da je $\mathcal{G}$ *σ-podalgebra od $\mathcal{F}$*. Nadalje, kažemo da je slučajna varijabla X, definirana na izmjerivom prostoru $(\Omega, \mathcal{F})$, *$\mathcal{G}$-izmjeriva* ako je

$$(\forall B \in \mathcal{B}(\mathbb{R})) \;\; X^{-1}(B) = \{X \in B\} \in \mathcal{G}.$$

Označimo sa $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ skup svih $\mathcal{G}$-izmjerivih slučajnih varijabli konačnog matematičkog očekivanja, a sa $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ skup svih $\mathcal{G}$-izmjerivih slučajnih varijabli konačnog drugog momenta. Jednostavno se vidi da je $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ potpuni potprostor od $\mathcal{L}^1$, a $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ potpun potprostor od $\mathcal{L}^2$ koji je ujedno i potpuni potprostor od $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$. Dakle, vrijede slijedeće relacije:

$$\begin{array}{lllllll} \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P}) \cap \mathcal{L}^2 & = & \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) & \subset & \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P}) & \subset & \\ & & & \subset & \mathcal{L}^2 & \subset & \mathcal{L}^1 \end{array}$$

pri čemu "$\subset$" označava relaciju "biti potpun potprostor od".

Neka je $Y \in \mathcal{L}^1$. Cilj je naći slučajnu varijablu $\hat{Y}$ takvu da je $\hat{Y} \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ i da za sve događaje $G \in \mathcal{G}$ vrijedi da je

$$\int_G \hat{Y} \, d\mathbb{P} = \int_G Y \, d\mathbb{P}. \tag{2.11}$$

Primijetite da su oba integrala u (2.11) konačna, te da je jednakost (2.11) ekvivalentna sa $\mathbb{E}[(Y - \hat{Y}) \cdot \mathbb{1}_G] = 0$. Ako bi Y bio iz prostora $\mathcal{L}^2$, tada bi $\hat{Y}$ postojao prema teoremu o ortogonalnoj projekciji 2.5 za potprostor $\mathcal{K} = \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ budući da tada iz normalnih jednadžbi primijenjenih na slučajne varijable $Z = \mathbb{1}_G$ slijedi (2.11). U tom slučaju je $\hat{Y} \in \mathcal{K} = \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$. Pitanje je da li slična "projekcija" postoji u prostoru $\mathcal{L}^1$ na potprostor $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$. Sljedeći teorem je odgovor na to pitanje.

Teorem 2.8. *Neka je Y slučajna varijabla definirana na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ takva da je $\mathbb{E}|Y| < +\infty$ i neka je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Tada postoji g.s. jedinstvena slučajna varijabla $\hat{Y}$ takva da vrijedi:*

(1) *$\hat{Y}$ je $\mathcal{G}$-izmjeriva,*
(2) *$\mathbb{E}|\hat{Y}| < +\infty$ i*
(3) *$(\forall G \in \mathcal{G}) \int_G \hat{Y}\, d\mathbb{P} = \int_G Y\, d\mathbb{P}$.*

Dokaz. Prvo ćemo pokazati da, ako postoji slučajna varijabla $\hat{Y}$ takva da vrijedi $(1-3)$, da je tada nužno g.s. jedinstvena. Zatim ćemo dokazati ekzistenciju od $\hat{Y}$ i to, prvo, u slučaju da je $Y \in \mathcal{L}^2$, zatim, drugo, kada je $Y \geq 0$, te, na kraju, u općem slučaju za bilo koji $Y \in \mathcal{L}^1$.

Dokaz g.s. jedinstvenosti. (To je ujedno i dokaz teorema 1.58.) Neka su $\hat{Y}$ i $\tilde{Y}$ dvije slučajne varijable takve da vrijedi $(1-3)$. Pretpostavimo da one nisu g.s. jedinstvene. Tada mora vrijediti da je $\mathbb{P}(\hat{Y} \neq \tilde{Y}) > 0$. Budući da je $\{\hat{Y} \neq \tilde{Y}\} = \{\hat{Y} < \tilde{Y}\} \cup \{\hat{Y} > \tilde{Y}\}$, vrijedi

$$0 < \mathbb{P}(\hat{Y} \neq \tilde{Y}) = \mathbb{P}(\hat{Y} < \tilde{Y}) + \mathbb{P}(\hat{Y} > \tilde{Y}) \;\Rightarrow\; \mathbb{P}(\hat{Y} < \tilde{Y}) > 0 \text{ ili } \mathbb{P}(\hat{Y} > \tilde{Y}) > 0.$$

Bez smanjenja općenitosti, pretpostavimo da je $\mathbb{P}(\hat{Y} > \tilde{Y}) > 0$. Budući da je

$$\{\hat{Y} > \tilde{Y}\} = \bigcup_{n=1}^{\infty} \left\{\hat{Y} \geq \tilde{Y} + \frac{1}{n}\right\} \quad \text{i} \quad \left\{\hat{Y} \geq \tilde{Y} + \frac{1}{n}\right\} \subseteq \left\{\hat{Y} \geq \tilde{Y} + \frac{1}{n+1}\right\},\ n \in \mathbb{N},$$

prema svojstvu *neprekidnosti vjerojatnosti u odnosu na rastuće događaje* (vidjeti teorem 2.1 (d) u [9], str. 14.) je

$$\lim_n \mathbb{P}\left(\hat{Y} \geq \tilde{Y} + \frac{1}{n}\right) = \mathbb{P}(\hat{Y} > \tilde{Y}) > 0$$

pa postoji $n \in \mathbb{N}$ takav da je

$$\mathbb{P}\left(\hat{Y} \geq \tilde{Y} + \frac{1}{n}\right) > \frac{1}{2}\mathbb{P}(\hat{Y} > \tilde{Y}) > 0.$$

Neka je $G := \{\hat{Y} \geq \tilde{Y} + \frac{1}{n}\}$. Budući da za obje slučajne varijable $\hat{Y}$ i $\tilde{Y}$ vrijedi (1), $G \in \mathcal{G}$ (vidjeti korolar 8.1. u [9], str. 240.). Dakle, korištenjem linearnosti matematičkog očekivanja (što je korektno jer vrijedi (2)) i tvrdnje (3),

$$\mathbb{E}[(\hat{Y} - \tilde{Y}) \cdot \mathbb{1}_G] \overset{\text{(lin.m.o.)}}{=} \mathbb{E}(\hat{Y} \cdot \mathbb{1}_G) - \mathbb{E}(\tilde{Y} \cdot \mathbb{1}_G) \overset{(3)}{=} \mathbb{E}(Y \cdot \mathbb{1}_G) - \mathbb{E}(Y \cdot \mathbb{1}_G) = 0.$$

S druge strane, budući da je $(\hat{Y} - \tilde{Y}) \cdot \mathbb{1}_G \geq \frac{1}{n} \cdot \mathbb{1}_G$, prema monotonosti matematičkog očekivanja, uzimajući u obzir da je $\mathbb{P}(G) > 0$,

$$\mathbb{E}[(\hat{Y} - \tilde{Y}) \cdot \mathbb{1}_G] \geq \mathbb{E}[\frac{1}{n}\mathbb{1}_G] = \frac{1}{n}\mathbb{P}(G) > 0.$$

Dakle, dobili smo kontradikciju. Budući da je ta kontradikcija dobivena na osnovi pretpostavke da $\hat{Y}$ i $\tilde{Y}$ nisu g.s. jedinstveni, slijedi da je ta tvrdnja lažna.

Dokaz egzistencije za $Y \in \mathcal{L}^2$. Ako je $Y \in \mathcal{L}^2$, tada, prema teoremu o ortogonalnoj projekciji 2.5, postoji $\hat{Y}$ u potpunom potprostoru $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ tako da za svaki $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ vrijedi da je $\mathbb{E}[(Y - \hat{Y})Z] = 0$. Specijalno, za svaki $G \in \mathcal{G}$ je $\mathbb{E}[(Y - \hat{Y}) \cdot \mathbb{1}_G] = 0$ budući da je $\mathbb{1}_G \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$. Dakle, vrijedi (3). Nadalje, budući da je $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \subset \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, $\hat{Y}$ zadovoljava i $(1-2)$.

Za nastavak dokaza teorema 2.8, treba nam sljedeća lema. Prije nego što ju iskažemo i dokažemo, podsjetimo se da je slučajna varijabla Y *ograničena* ako je ograničena jednoliko na Ω, preciznije, ako postoji konstanta $B > 0$ takva da je za sve $\omega \in \Omega$, $|Y(\omega)| \leq B$. U tom slučaju je nužno $Y \in \mathcal{L}^2$ pa, prema upravo dokazanom, postoji $\hat{Y}$ takva slučajna varijabla da za nju vrijede svojstva $(1-3)$ iz teorema.

Od sada pa nadalje sa $\hat{Y}$ označavamo bilo koju verziju slučajne varijable pridruženu varijabli $Y \in \mathcal{L}^2$ i σ-podalgebri $\mathcal{G}$ od $\mathcal{F}$ tako da je ispunjeno $(1-3)$ iz teorema.

Lema 2.9. *Neka su U i V slučajne varijable u $\mathcal{L}^2$ takve da je $0 \leq U \leq V$. Tada je $0 \leq \hat{U} \leq \hat{V}$ g.s.*

Dokaz leme. Prvo dokazujemo da je $\hat{U} \geq 0$ g.s. Pretpostavimo suprotno, da vrijedi $\mathbb{P}(\hat{U} < 0) > 0$. Budući da je

$$\{\hat{U} < 0\} = \bigcup_{n=1}^{\infty} \left\{\hat{U} \leq -\frac{1}{n}\right\} \quad \text{i} \quad \left\{\hat{U} \leq -\frac{1}{n}\right\} \subseteq \left\{\hat{U} \leq -\frac{1}{n+1}\right\},\ n \in \mathbb{N},$$

slično kao u dokazu g.s. jedinstvenosti možemo primijeniti svojstvo neprekidnosti vjerojatnosti u odnosu na rastuće događaje i zaključiti da postoji $n \in \mathbb{N}$ tako da je

$$\mathbb{P}\left(\hat{U} \leq -\frac{1}{n}\right) \geq \frac{1}{2}\mathbb{P}(\hat{U} < 0) > 0.$$

Stavimo da je $G := \{\hat{U} \leq -\frac{1}{n}\}$. Vrijedi $G \in \mathcal{G}$ jer je $\hat{U}$ $\mathcal{G}$-izmjeriva slučajna varijabla. Sada je, prema svojstvu (3) i monotonosti matematičkog očekivanja,

$$0 \leq \mathbb{E}[U \cdot \mathbb{1}_G] \overset{(3)}{=} \mathbb{E}[\hat{U} \cdot \mathbb{1}_G] \overset{\text{(mon.m.o.)}}{\leq} \mathbb{E}[-\frac{1}{n}\mathbb{1}_G] = -\frac{1}{n}\mathbb{P}(G) < 0,$$

dakle, kontradikcija. Slijedi da je $\hat{U} \geq 0$ g.s.
Neka je sada $W := V - U \geq 0$. Tada je, prema upravo pokazanom, $\hat{W} \geq 0$ g.s. Ako pokažemo da je $\hat{W} = \hat{U} - \hat{V}$ g.s., tada slijedi da je $\hat{U} \leq \hat{V}$ g.s. Neka je $G \in \mathcal{G}$ proizvoljan. Tada je, primjenom svojstva (3) u odnosu na odgovarajuću varijablu i linearnosti matematičkog očekivanja,

$$\mathbb{E}[(\hat{U} - \hat{V}) \cdot \mathbb{1}_G] \overset{\text{(lin.m.o.)}}{=} \mathbb{E}[\hat{U} \cdot \mathbb{1}_G] - \mathbb{E}[\hat{V} \cdot \mathbb{1}_G] \overset{(3)}{=} \mathbb{E}[U \cdot \mathbb{1}_G] - \mathbb{E}[V \cdot \mathbb{1}_G] \overset{\text{(lin.m.o.)}}{=} \mathbb{E}[W \cdot \mathbb{1}_G],$$

a budući da je $\hat{U} - \hat{V} \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, slijedi da je $\hat{U} - \hat{V} = \hat{W}$ g.s. □

Dokaz egzistencije za $Y \geq 0$. Neka je $Y \in \mathcal{L}^1$ i $Y \geq 0$. Za $n \in \mathbb{N}$ definirajmo $Y_n := Y \cdot \mathbb{1}_{\{Y \leq n\}}$. Tada je $(Y_n)_{n \in \mathbb{N}}$ rastući niz nenegativnih ograničenih varijabli takav da je $\lim_n Y_n = Y$. Za svaki $n \in \mathbb{N}$ zbog ograničenosti od Y_n postoji $\hat{Y}_n$. Prema lemi 2.9, $(\hat{Y}_n)_{n \in \mathbb{N}}$ je g.s. rastući niz g.s. nenegativnih slučajnih varijabli. Definirajmo $W(\omega) := \underline{\lim}_n \hat{Y}_n(\omega)$ $(\omega \in \Omega)$. Pokazat ćemo da W ima tražena svojstva $(1-3)$ iz teorema, tj. da je $\hat{Y} = W$ g.s. Iz definicije od W je jasno da je W $\mathcal{G}$-izmjeriva slučajna varijabla jer su $\hat{Y}_n$ $\mathcal{G}$-izmjerive za sve $n \in \mathbb{N}$ (vidjeti teorem 8.4. (i) u [9], str. 242.). Dakle, vrijedi (2). Neka je sada $G \in \mathcal{G}$ proizvoljno. Tada je $(Y_n \cdot \mathbb{1}_G)_{n\in\mathbb{N}}$ rastući niz nenegativnih brojeva takav da je $\lim_n Y_n \cdot \mathbb{1}_G = Y \cdot \mathbb{1}_G$. Prema Lebesguevom teoremu 1.59 (i) o monotonoj konvergenciji (LTMK) slijedi $\lim_n \mathbb{E}(Y_n \cdot \mathbb{1}_G) = \mathbb{E}(Y \cdot \mathbb{1}_G)$. S druge strane, $(\hat{Y}_n \cdot \mathbb{1}_G)_{n\in\mathbb{N}}$ je g.s. rastući niz g.s. nenegativnih slučajnih varijabli takvih da je $\lim_n \hat{Y}_n \cdot \mathbb{1}_G = W \cdot \mathbb{1}_G$ g.s. Prema LTMK, $\lim_n \mathbb{E}(\hat{Y}_n \cdot \mathbb{1}_G) = \mathbb{E}(W \cdot \mathbb{1}_G)$. Budući da je $\mathbb{E}(\hat{Y}_n \cdot \mathbb{1}_G) = \mathbb{E}(Y_n \cdot \mathbb{1}_G)$ za svaki $n \in \mathbb{N}$, limesi tih dvaju nizova realnih brojeva su jedinstveni: $\mathbb{E}(W \cdot \mathbb{1}_G) = \mathbb{E}(Y \cdot \mathbb{1}_G)$. Shematski:

$$\begin{array}{lccc} 0 \le \hat{Y}_n \cdot \mathbb{1}_G \uparrow W \text{ g.s.} & & \mathbb{E}(\hat{Y}_n \cdot \mathbb{1}_G) \longrightarrow & \mathbb{E}(W \cdot \mathbb{1}_G) \\ & \Rightarrow & \| & \| \\ 0 \le Y_n \cdot \mathbb{1}_G \uparrow Y & & \mathbb{E}(Y_n \cdot \mathbb{1}_G) \longrightarrow & \mathbb{E}(Y \cdot \mathbb{1}_G) \end{array}$$

Dakle, W zadovoljava (3). Ako u (3) uzmemo $G = \Omega$, tada je

$$\mathbb{E}|W| = \mathbb{E}W = \mathbb{E}(W \cdot \mathbb{1}_\Omega) \stackrel{(3)}{=} \mathbb{E}(Y \cdot \mathbb{1}_\Omega) = \mathbb{E}Y = \mathbb{E}|Y| < +\infty,$$

pa vrijedi i (2).

Dokaz egzistencije za $Y \in \mathcal{L}^1$. $Y = Y^+ - Y^-$ gdje su $Y^+ \ge 0$ i $Y^- \ge 0$, te $\mathbb{E}Y^+ < +\infty$ i $\mathbb{E}Y^- < +\infty$. Primijenimo prethodni rezultat na te varijable. Naime, neka su $\hat{Y}^+$ i $\hat{Y}^-$ pridružene varijablama Y^+, odnosno Y^-, tako da za njih vrijedi tvrdnja teorema 2.8. Pokažimo da $W := \hat{Y}^+ - \hat{Y}^-$ zadovoljava $(1-3)$ iz teorema (u odnosu na Y). Budući da su $Y^+, Y^- \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ slijedi da je i $W = \hat{Y}^+ - \hat{Y}^- \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ pa vrijedi $(1-2)$. Neka je sada $G \in \mathcal{G}$ proizvoljan. Tada je zbog linearnosti matematičkog očekivanja i pripadajućih svojstava (3),

$$\mathbb{E}(W \cdot \mathbb{1}_G) = \mathbb{E}(\hat{Y}^+ \cdot \mathbb{1}_G) - \mathbb{E}(\hat{Y}^- \cdot \mathbb{1}_G) \stackrel{(3)}{=} \mathbb{E}(Y^+ \cdot \mathbb{1}_G) - \mathbb{E}(Y^- \cdot \mathbb{1}_G) = \mathbb{E}(Y \cdot \mathbb{1}_G).$$

Dakle, zbog proizvoljnosti od G, vrijedi i (3) pa je $\hat{Y} = W$. □

Napomena 2.10. Alternativni dokaz teorema 2.8 bazira se na Radon-Nikodymovom teoremu (vidjeti teorem 10.5 u [9], str. 297.) na sljedeći način. Za nenegativnu slučajnu varijablu $Y \in \mathcal{L}^1$ vrijedi da je sa $Q(G) := \mathbb{E}[Y \cdot \mathbb{1}_G]$ definirana konačna mjera (vidjeti teorem 10.4 u [9], str. 296.) na izmjerivom prostoru $(\Omega, \mathcal{G})$ koja je apsolutno neprekidna u odnosu na vjerojatnost $\mathbb{P}$. Tada prema Radon-Nikodymovom teoremu postoji $\mathcal{G}$-izmjeriva slučajna varijabla $\hat{Y} = dQ/d\mathbb{P}$ (Radon-Nikodymova derivacija) takva da je $Q(G) = \mathbb{E}[\hat{Y} \cdot \mathbb{1}_G]$ za svaki $G \in \mathcal{G}$. Budući da je u općem slučaju $Y = Y^+ - Y^-$, varijabla $\hat{Y}$, takva da za nju vrijede tvrdnje $(1-3)$ iz teorema, se konstruira na uobičajeni način.

Definicija 2.11. *Neka je Y slučajna varijabla na $(\Omega, \mathcal{F}, \mathbb{P})$ takva da je $\mathbb{E}|Y| < +\infty$ i neka je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Slučajnu varijablu $\hat{Y}$ za koju vrijede tvrdnje:*

(1) $\hat{Y}$ *je* $\mathcal{G}$*-izmjeriva,*
(2) $\mathbb{E}|\hat{Y}| < +\infty$ *i*
(3) $(\forall G \in \mathcal{G}) \int\limits_G \hat{Y}\, d\mathbb{P} = \int\limits_G Y\, d\mathbb{P}.$

zovemo **verzijom uvjetnog matematičkog očekivanja od** Y **uz dano** $\mathcal{G}$ *i označavamo sa* $\mathbb{E}[Y|\mathcal{G}]$.

Prema teoremu 2.8 verzija uvjetnog matematičnog očekivanja pridružena zadanim elementima: varijabli s konačnim matematičkim očekivanjem i σ-podalgebri na koju se uvjetuje, postoji i g.s. je jedinstvena. Strogo formalno, u normiranom prostoru L^1, uvjetno matematičko očekivanje postoji i jedinstveno je (za zadane elemente) i interpretira se kao *klasa* svih verzija uvjetnog matematičkog očekivanja pridruženih zadanim elementima u skladu s prethodnom definicijom 2.11. Dakako, ta klasa je element potprostora $L^1(\Omega, \mathcal{G}, \mathbb{P})$ od L^1, koji je dobiven kao kvocijentni skup elemenata iz potprostora $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ u odnosu na relaciju *biti jednak g.s.* Primijetite da su elementi proizvoljne klase iz $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ $\mathcal{G}$-*izmjerive* slučajne varijable konačnog očekivanja koje su međusobno g.s. jednake.

Od sada pa nadalje, kada napišemo (ili kažemo): *neka je* $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$ ili *neka je* $\hat{Y}$ *uvjetno matematičko očekivanje od* Y *uz dano* $\mathcal{G}$, tada taj izraz ima ekvivalentno značenje: *neka je* $\hat{Y}$ *jedna verzija uvjetnog matematičkog očekivanja od* Y *uz dano* $\mathcal{G}$.

Prvo ćemo pokazati konstrukciju uvjetnog matematičkog očekivanja kada se σ-podalgebra $\mathcal{G}$ može generirati prebrojivim potpunim sustavom događaja. Kažemo da je prebrojiva (konačna ili beskonačna) množina $\mathcal{H} := \{H_1, H_2, \ldots\}$ *potpuni sustav događaja* na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ ako su događaji $\{H_1, H_2, \ldots\}$ (1.) pozitivne vjerojatnosti: $\mathbb{P}(H_i) > 0$ za svaki $i \in \{1, 2, \ldots\}$ i (2.) čine particiju od Ω: za svaki $i \neq j$ je $H_i \cap H_j = \emptyset$ i $\cup_{i \geq 1} H_i = \Omega$. Nadalje, označimo sa $\sigma(\mathcal{H})$ najmanju σ-algebru podskupova od Ω koja sadrži množinu $\mathcal{H}$. Kažemo da je $\sigma(\mathcal{H})$ σ-*algebra generirana množinom* $\mathcal{H}$.

Prije iskaza tvrdnje o uvjetnom očekivanju slučajne varijable $Y \in \mathcal{L}^1 \equiv \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ uvjetno na $\mathcal{G} = \sigma(\mathcal{H})$ za potpun sustav događaja $\mathcal{H}$ navedimo i dokažimo jednu lemu. Za nju nam treba pojam uvjetnog očekivanja u odnosu na dani događaj. Za događaj $B \in \mathcal{F}$ takav da je $\mathbb{P}(B) > 0$ i slučajnu varijablu Y s konačnim očekivanjem, broj $\int\limits_\Omega Y(\omega)\, d\mathbb{P}(\omega|B)$ je po definiciji matematičko očekivanje od Y u odnosu na uvjetnu vjerojatnost $\mathbb{P}(\cdot|B)$. Označimo ga s

$$\mathbb{E}[Y|B] := \int\limits_\Omega Y\, d\mathbb{P}(\cdot|B) \equiv \int\limits_\Omega Y(\omega)\, d\mathbb{P}(\omega|B). \tag{2.12}$$

Taj broj zovemo *uvjetno matematičko očekivanje od* Y *uz dani događaj* B.

Lema 2.12. *Za događaj* $B \in \mathcal{F}$ *takav da je* $\mathbb{P}(B) > 0$ *i* $Y \in \mathcal{L}^1$ *vrijedi*

$$\mathbb{E}[Y|B] = \frac{\mathbb{E}[Y \cdot \mathbb{1}_B]}{\mathbb{P}(B)}. \tag{2.13}$$

Dokaz. Neka je događaj B fiksiran i neka $\mathcal{I}$ označava množinu svih $Y \in \mathcal{L}^1$ takvih da vrijedi (2.13). Dokaz leme sprovodimo postepeno, od jednostavnih slučajnih varijabli do općeg slučaja. Prvo ćemo pokazati da je $Y = \mathbb{1}_A \in \mathcal{I}$ za bilo koji događaj A (vidjeti primjere 1.2 i 1.53):

$$\int\limits_B \mathbb{1}_A(\omega)\, d\mathbb{P}(\omega|B) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}[\mathbb{1}_{A \cap B}]}{\mathbb{P}(B)} = \frac{\mathbb{E}[\mathbb{1}_A \cdot \mathbb{1}_B]}{\mathbb{P}(B)}.$$

Dakle, $\mathbb{1}_A \in \mathcal{I}$. Nadalje, budući da je matematičko očekivanje linearni funkcional, jednostavne slučajne varijable (tj. linearne kombinacije karakterističnih funkcija događaja) su također u množini $\mathcal{I}$. Neka je sada $Y \geq 0$. Tada se Y može dobiti kao rastući limes jednostavnih nenegativnih slučajnih varijabli Y_n ($n \in \mathbb{N}$) (vidjeti teorem 1.3). Budući da su $0 \leq Y_n \uparrow Y$ i $0 \leq Y \cdot \mathbb{1}_B \uparrow Y \cdot \mathbb{1}_B$ kada $n \to +\infty$ i $Y_n \in \mathcal{I}$ za svaki $n \in \mathbb{N}$, primjenom Lebesgueova teorema 1.59 (i) o monotonoj konvergenciji (LTMK) i (2.13) dobijamo

$$\int\limits_\Omega Y(\omega)\, d\mathbb{P}(\omega|B) \overset{\text{(LTMK)}}{=} \lim_n \int\limits_\Omega Y_n(\omega)\, d\mathbb{P}(\omega|B) \overset{(2.13)}{=} \lim_n \frac{\mathbb{E}[Y_n \cdot \mathbb{1}_B]}{\mathbb{P}(B)} \overset{\text{(LTMK)}}{=} \frac{\mathbb{E}[Y \cdot \mathbb{1}_B]}{\mathbb{P}(B)}.$$

Dakle, $Y \in \mathcal{I}$. Na kraju, neka je $Y \in \mathcal{L}^1$. Tada je $Y = Y^+ - Y^-$. Budući da su $Y^+ \geq 0$ i $Y^- \geq 0$ u $\mathcal{I}$, zbog linearnosti matematičkog očekivanja slijedi da je i $Y \in \mathcal{I}$. $\square$

Teorem 2.13. *Neka je $\mathcal{H} = \{H_1, H_2, \ldots\}$ prebrojiv potpuni sustav događaja na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G} = \sigma(\mathcal{H})$ i $Y \in \mathcal{L}^1$. Tada je*

$$\mathbb{E}[Y|\mathcal{G}] = \sum_{k\geq 1} \mathbb{E}[Y|H_k] \cdot \mathbb{1}_{H_k}. \tag{2.14}$$

Dokaz. Nije teško pokazati da je

$$\sigma(\mathcal{H}) = \{\, H_{i_1} \cup H_{i_2} \cup \cdots \, : \, \{i_1, i_2, \ldots\} \subseteq \{1, 2, \ldots\}\, \}, \tag{2.15}$$

tj. $\sigma(\mathcal{H})$ je množina svih prebrojivih (konačnih i, ako je $\mathcal{H}$ beskonačna množina, beskonačnih) disjunktnih unija elemenata potpunog sustava događaja $\mathcal{H}$. Budući da je za svaki k, $\mathbb{P}(H_k) > 0$, dobro je definirana uvjetna vjerojatnost $\mathbb{P}(\cdot|H_k)$ na $(\Omega, \mathcal{F})$, a kako je i $Y \in \mathcal{L}^1$, dobro je definirano uvjetno matematičko očekivanje od Y uz dani događaj H_k i, prema (2.13), vrijedi

$$\mathbb{E}[Y|H_k] = \frac{\mathbb{E}(Y \cdot \mathbb{1}_{H_k})}{\mathbb{P}(H_k)}, \quad k \geq 1. \tag{2.16}$$

Označimo s $y_k := \mathbb{E}[Y|H_k]$ za sve $k \geq 1$ i s $\hat{Y} := \sum_{k\geq 1} y_k \cdot \mathbb{1}_{H_k}$. Neka je $B \in \mathcal{B}(\mathbb{R})$ proizvoljan Borelov skup. Tada taj skup može sadržavati (ali i ne mora) brojeve y_k za $k \geq 1$. Označimo sa $k_j(B)$ $(j \geq 1)$ sve indekse $k \geq 1$ takve da je $y_k \in B$, tj.

$$\mathcal{J}(B) \equiv \{k_1(B), k_2(B), \ldots\} := \{k \geq 1 : y_k \in B\}.$$

Ako B ne sadrži niti jednu vrijednost y_k, tada je $\mathcal{J}(B) = \emptyset$. Slijedi da je

$$\{\hat{Y} \in B\} = \begin{cases} H_{k_1(B)} \cup H_{k_2(B)} \cup \cdots & \text{ako je } \mathcal{J}(B) \neq \emptyset \\ \emptyset & \text{ako je } \mathcal{J}(B) = \emptyset. \end{cases}$$

Dakle, $\{\hat{Y} \in B\} \in \mathcal{G} \equiv \sigma(\mathcal{H})$ prema (2.15), pa je $\hat{Y}$ $\mathcal{G}$-izmjeriva slučajna varijabla. Dakle, vrijedi svojstvo (1) iz definicije 2.11. Nadalje, korištenjem (2.16) i monotonosti matematičkog očekivanja (vidjeti teorem 1.56 (*iii*)), slijedi

$$|\hat{Y}| = \sum_{k\geq 1} |y_k| \cdot \mathbb{1}_{H_k} = \sum_{k\geq 1} \frac{|\mathbb{E}(Y \cdot \mathbb{1}_{H_k})|}{\mathbb{P}(H_k)} \cdot \mathbb{1}_{H_k} \leq \sum_{k\geq 1} \frac{\mathbb{E}(|Y| \cdot \mathbb{1}_{H_k})}{\mathbb{P}(H_k)} \cdot \mathbb{1}_{H_k} = \sum_{k\geq 1} \mathbb{E}[|Y|\,|H_k] \cdot \mathbb{1}_{H_k}.$$

Još jednom upotrebom monotononosti matematičkog očekivanja slijedi

$$\mathbb{E}(|\hat{Y}|) \leq \mathbb{E}(\sum_{k\geq 1} \mathbb{E}[|Y|\,|H_k] \cdot \mathbb{1}_{H_k}).$$

Ako je suma zdesna u prethodnoj nejednakosti konačna, tada je zbog linearnosti matematičkog očekivanja i svojstva (2.16),

$$\begin{aligned} \mathbb{E}(\sum_{k\geq 1} \mathbb{E}[|Y|\,|H_k] \cdot \mathbb{1}_{H_k}) &\overset{\text{(lin.m.o.)}}{=} \sum_{k\geq 1} \mathbb{E}[|Y|\,|H_k]\mathbb{P}(H_k) \overset{(2.16)}{=} \\ = \sum_{k\geq 1} \mathbb{E}(|Y| \cdot \mathbb{1}_{H_k}) &\overset{\text{(lin.m.o.)}}{=} \mathbb{E}(|Y|) < +\infty, \end{aligned}$$

dakle, $\mathbb{E}(|\hat{Y}|) < +\infty$, pa vrijedi (2) iz definicije 2.11. U suprotnom, ako je suma zdesna beskonačna, primijetimo da su tada sume redova funkcija $\sum_{k\geq 1} \mathbb{E}[|Y|\,|H_k] \cdot \mathbb{1}_{H_k}$ i $|Y| =$

$\sum_{k\geq 1}|Y|\cdot \mathbb{1}_{H_k}$ rastući limesi njihovih (nenegativnih) parcijalnih suma, pa je, prema Lebesgueovom teoremu o monotonoj konvergenciji i linearnosti matematičkog očekivanja,

$$\begin{aligned}
&\mathbb{E}(\sum_{k\geq 1}\mathbb{E}[|Y|\,|H_k]\cdot \mathbb{1}_{H_k}) \stackrel{\text{(LTMK)}}{=} \lim_n \mathbb{E}(\sum_{k=1}^{n}\mathbb{E}[|Y|\,|H_k]\cdot \mathbb{1}_{H_k}) \stackrel{\text{(lin.m.o.)}}{=} \\
= \; & \lim_n \sum_{k=1}^{n}\mathbb{E}[|Y|\,|H_k]\mathbb{P}(H_k) \stackrel{(2.16)}{=} \lim_n \sum_{k=1}^{n}\mathbb{E}(|Y|\cdot \mathbb{1}_{H_k}) \stackrel{\text{(lin.m.o.)}}{=} \\
= \; & \lim_n \mathbb{E}(\sum_{k=1}^{n}|Y|\cdot \mathbb{1}_{H_k}) \stackrel{\text{(LTMK)}}{=} \mathbb{E}(|Y|) < +\infty,
\end{aligned}$$

dakle, i u tom slučaju je $\mathbb{E}(|\hat{Y}|) < +\infty$. Na kraju, neka je $G \in \mathcal{G}$ proizvoljno. Budući da je $\mathcal{G} = \sigma(\mathcal{H})$ i vrijedi (2.15), postoji podskup $\{i_1, i_2, \ldots\}$ skupa indeksa $k \geq 1$ takav da je $G = H_{i_1} \cup H_{i_2} \cup \cdots$. Slijedi

$$\mathbb{E}(\hat{Y}\cdot \mathbb{1}_G) = \mathbb{E}((\sum_{k\geq 1} y_k \cdot \mathbb{1}_{H_k})\cdot \mathbb{1}_{H_{i_1}\cup H_{i_2}\cup\cdots}) = \mathbb{E}(\sum_{k\geq 1} y_k \cdot \mathbb{1}_{H_k \cap (H_{i_1}\cup H_{i_2}\cup\cdots)}) = \mathbb{E}(\sum_{j\geq 1} y_{i_j}\cdot \mathbb{1}_{H_{i_j}}).$$

Slično kao i u prethodnom slučaju, ako je suma zdesna u prethodnoj jednakosti (po j) konačna, tada je zbog linearnosti matematičkog očekivanja i (2.16),

$$\begin{aligned}
&\mathbb{E}(\sum_{j\geq 1} y_{i_j}\cdot \mathbb{1}_{H_{i_j}}) \stackrel{\text{(lin.m.o.)}}{=} \sum_{j\geq 1} y_{i_j}\mathbb{P}(H_{i_j}) \stackrel{(2.16)}{=} \sum_{j\geq 1}\mathbb{E}(Y\cdot \mathbb{1}_{H_{i_j}}) \stackrel{\text{(lin.m.o.)}}{=} \\
= \; & \mathbb{E}(Y\cdot \mathbb{1}_{H_{i_1}\cup H_{i_2}\cup\cdots}) = \mathbb{E}(Y\cdot \mathbb{1}_G),
\end{aligned}$$

dakle, $\mathbb{E}(\hat{Y}\cdot \mathbb{1}_G) = \mathbb{E}(Y\cdot \mathbb{1}_G)$. U suprotnom, ako je suma zdesna beskonačna po j, tada su parcijalne sume reda $\sum_{j\geq 1} y_{i_j}\cdot \mathbb{1}_{H_{i_j}}$ i parcijalne sume reda $Y\cdot \mathbb{1}_G = \sum_{j\geq 1} Y\cdot \mathbb{1}_{H_{i_j}}$ uniformno dominirane slučajnim varijablama iz $\mathcal{L}^1$, tj. za svaki $n \in \mathbb{N}$ je

$$|\sum_{j=1}^{n} y_{i_j}\cdot \mathbb{1}_{H_{i_j}}| = \sum_{j=1}^{n}|y_{i_j}|\cdot \mathbb{1}_{H_{i_j}} \leq \sum_{k\geq 1}|y_k|\cdot \mathbb{1}_{H_k} = |\hat{Y}|$$

i

$$|\sum_{j=1}^{n} Y\cdot \mathbb{1}_{H_{i_j}}| = \sum_{j=1}^{n}|Y|\cdot \mathbb{1}_{H_{i_j}} \leq \sum_{k\geq 1}|Y|\cdot \mathbb{1}_{H_k} = |Y|.$$

Primjenom Lebesgueova teorema 1.59 (*iii*) o dominiranoj konvergenciji (LTDK), linearnosti matematičkog očekivanja i (2.16), slijedi

$$\begin{aligned}
&\mathbb{E}(\sum_{j\geq 1} y_{i_j}\cdot \mathbb{1}_{H_{i_j}}) \stackrel{\text{(LTDK)}}{=} \lim_n \mathbb{E}(\sum_{j=1}^{n} y_{i_j}\mathbb{1}_{H_{i_j}}) \stackrel{\text{(lin.m.o.)}}{=} \lim_n \sum_{j=1}^{n} y_{i_j}\mathbb{P}(H_{i_j}) \stackrel{(2.16)}{=} \\
= \; & \lim_n \sum_{j=1}^{n}\mathbb{E}(Y\cdot \mathbb{1}_{H_{i_j}}) \stackrel{\text{(lin.m.o.)}}{=} \lim_n \mathbb{E}(\sum_{j=1}^{n} Y\cdot \mathbb{1}_{H_{i_j}}) \stackrel{\text{(LTDK)}}{=} \mathbb{E}(Y\cdot \mathbb{1}_G),
\end{aligned}$$

dakle, $\mathbb{E}(\hat{Y}\cdot \mathbb{1}_G) = \mathbb{E}(Y\cdot \mathbb{1}_G)$. Zbog proizvoljnosti od $G \in \mathcal{G}$, vrijedi i (3) iz definicije 2.11. Prema tome, u skladu s definicijom 2.11, $\hat{Y}$ je verzija uvjetnog očekivanja od Y uz dano $\mathcal{G} \equiv \sigma(\mathcal{H})$, dakle, vrijedi (2.14). $\square$

Postoji veza između uvjetnog matematičkog očekivanja i matematičkog očekivanja u odnosu na uvjetni zakon razdiobe (kako je opisano u potpoglavlju 1.4. prethodnog poglavlja). Poznavanje te veze omogućit će nam da u važnom specijalnom slučaju izračunamo verziju uvjetnog matematičkog očekivanja. Opišimo taj slučaj.

Neka su X i Y slučajne veličine definirane na istome vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, pri čemu je X k-dimenzionalna slučajna veličina ($k \geq 1$), a Y slučajna varijabla s konačnim matematičkim očekivanjem. Označimo sa

$$\sigma(X) := \{X^{-1}(B) \,:\, B \in \mathcal{B}(\mathbb{R}^k)\} = \{\,\{X \in B\} \,:\, B \in \mathcal{B}(\mathbb{R}^k)\}$$

množinu događaja iz $\mathcal{F}$ koji su oblika $\{X \in B\}$ za neki Borelov skup B iz $\mathbb{R}^k$. Nije teško pokazati da je $\sigma(X)$ σ-podalgebra od $\mathcal{F}$ koju zovemo *σ-algebra generirana s X*. Tu σ-algebru interpretiramo kao *informacije* kojima raspolažemo *ako nam je poznat ishod od X*. Kraće označavamo s

$$\mathbb{E}[Y|X] \equiv \mathbb{E}[Y|\sigma(X)]$$

uvjetno matematičko očekivanje od Y uz dano $\sigma(X)$. $\mathbb{E}[Y|X]$ postoji prema teoremu 2.8. Slučajnu veličinu $\mathbb{E}[Y|X]$ interpretiramo kao (najbolju) *procjenu* ili *predviđanje* ishoda od *Y ako nam je X poznato*.

Navedimo sada važan rezultat koji će nas motivirati za konstrukciju $\mathbb{E}[Y|X]$.

Teorem 2.14. *Neka su k-dimenzionalna slučajna veličina X ($k \geq 1$) i slučajna varijabla Z definirane na istom izmjerivom prostoru. Tada je Z $\sigma(X)$-izmjeriva varijabla ako i samo postoji Borelova funkcija $g : \mathbb{R}^k \to \mathbb{R}$ takva da je $Z = g(X)$.*

Dokaz. Neka je $Z = g(X)$ za neku Borelovu funkciju $g : \mathbb{R}^k \to \mathbb{R}$. Tada za proizvoljan Borelov skup B iz $\mathbb{R}^k$ vrijedi da je $g^{-1}(B)$ Borelov skup u $\mathbb{R}$ (jer je g Borelova funkcija), pa je

$$Z^{-1}(B) = (g \circ X)^{-1}(B) = X^{-1}(g^{-1}(B)) \in \sigma(X).$$

Zbog proizvoljnosti od B slijedi da je Z $\sigma(X)$-izmjeriva varijabla.

Pokažimo sada da vrijedi obratna tvrdnja, tj. da za svaku $\sigma(X)$-izmjerivu varijablu Z postoji Borelova funkcija $g : \mathbb{R}^k \to \mathbb{R}$ takva da je $Z = g(X)$. Prvo pretpostavimo da je $Z = \mathbb{1}_{\{X \in B\}}$ za neki Borelov skup $B \in \mathcal{B}(\mathbb{R}^k)$. Ako stavimo da je $g := \mathbb{1}_B$, tada je $g(X) = \mathbb{1}_B(X) = \mathbb{1}_{\{X \in B\}} = Z$ pa imamo tvrdnju. Neka je sada Z jednostavna $\sigma(X)$-izmjeriva varijabla. Tada postoje Borelovi skupovi $B_1, B_2, \ldots, B_\ell \in \mathcal{B}(\mathbb{R}^k)$ i realni brojevi $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ takvi da je

$$Z = \alpha_1 \cdot \mathbb{1}_{\{X \in B_1\}} + \alpha_2 \cdot \mathbb{1}_{\{X \in B_2\}} + \cdots + \alpha_\ell \cdot \mathbb{1}_{\{X \in B_\ell\}}.$$

Ako stavimo da je $g := \alpha_1 \cdot \mathbb{1}_{B_1} + \alpha_2 \cdot \mathbb{1}_{B_2} + \cdots + \alpha_\ell \cdot \mathbb{1}_{B_\ell}$, tada je g jednostavna Borelova funkcija i očito je $Z = g(X)$. Dakle, i u tom slučaju vrijedi tvrdnja. Ako je Z proizvoljna $\sigma(X)$-izmjeriva varijabla, tada postoji niz $(Z_n)_{n \in \mathbb{N}}$ jednostavnih $\sigma(X)$-izmjerivih varijabli takvih da je $Z = \lim_n Z_n$ po točkama (vidjeti teorem 1.3). Prema prethodno pokazanom, za svaki $n \in \mathbb{N}$ postoje Borelove funkcije g_n takve da je $Z_n = g_n(X)$. Stavimo da je $g(x) := \lim_n g_n(x)$ za $x \in \mathbb{R}^k$ za koje limes niza brojeva $(g_n(x))_{n \in \mathbb{N}}$ postoji i konačan je, inače stavimo $g(x) := 0$. Tada je $g : \mathbb{R}^k \to \mathbb{R}$ Borelova funkcija (vidjeti, na primjer, propoziciju 8.9 u [9], str. 244.) i vrijedi

$$Z = \lim_n Z_n = \lim_n g_n(X) = g(X).$$

Dakle, tvrdnja vrijedi za sve $\sigma(X)$-izmjerive varijable. □

Vratimo se na naš slučaj. Budući da je slučajna varijabla $Z = \mathbb{E}[Y|X]$ $\sigma(X)$-izmjeriva, prema teoremu 2.14 postoji Borelova funkcija $\phi : \mathbb{R}^k \to \mathbb{R}$ takva da je $\mathbb{E}[Y|X] = \phi(X)$.

Koja je interpretacija funkcije ϕ? Interpretirat ćemo je u slučaju kada je $Y \in \mathcal{L}^2$. Tada je nužno i $\mathbb{E}[Y|X] \in \mathcal{L}^2$, iz čega slijedi da je $\phi \in \mathcal{L}^2(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_X)$. Taj prostor kraće označavamo $\mathcal{L}^2(\mathbb{P}_X)$. Prema teoremu o projekciji 2.5 i uzimajući u obzir konstrukciju uvjetnog očekivanja iz dokaza teorema 2.8, ϕ je $\mathbb{P}_X$-g.s. jedinstvena funkcija iz $\mathcal{L}^2(\mathbb{P}_X)$ takva da je

$$\mathbb{E}(Y - \phi(X))^2 = \inf_{g \in \mathcal{L}^2(\mathbb{P}_X)} \mathbb{E}(Y - g(X))^2. \tag{2.17}$$

Dakle, ϕ je Borelova funkcija iz prostora $\mathcal{L}^2(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_X)$ za koju je srednjekvadratna pogreška predviđanja Y pomoću $\phi(X) = \mathbb{E}[Y|X]$, najmanja. Uspoređujući kriterij (2.17) s kriterijem (2.7) i uzimajući u obzir terminologiju iz primjera 2.6, uvjetno matematičko očekivanje $\mathbb{E}[Y|X]$ mogli bismo zvati i *najboljim predviditeljem od Y pomoću X*, a funkciju $\phi \in \mathcal{L}^2(\mathbb{P}_X)$ takvu da vrijedi (2.17), *regresijskom funkcijom.* Grešku aproksimacije $Y - \mathbb{E}[Y|X]$ i ovdje zovemo *rezidual.* U nastavku ćemo precizno definirati pojam regresijske funkcije u općem slučaju kada je $Y \in \mathcal{L}^1$, ali za slučajne vektore (X, Y) koji su neprekidni ili takvi da je X diskretna slučajna veličina.

Kako odrediti (izračunati, konstruirati) funkciju ϕ? Izračunat ćemo ju uz dodatne i nešto općenitije pretpostavke na slučajni vektor (X, Y). Neka je X diskretna slučajna veličina dimenzije k ili je $k+\ell$-dimenzionalni slučajni vektor (X, Y) neprekidan s gustoćom $f \equiv f_{X,Y}$. Dakle, Y je općenito ℓ-dimenzionalna slučajna veličina ($k, \ell \geq 1$). U oba slučaja postoji marginalna gustoća $f_X : \mathbb{R}^k \to \mathbb{R}$, takva da je $f_X(x) = \mathbb{P}(X = x)$ ako je X diskretna veličina, a $f_X(x) = \int\limits_{\mathbb{R}} f(x, y)\, dy$ za λ^k-s.s. $x \in \mathbb{R}^k$, ako je (X, Y) neprekidan vektor. Neka je $x \in \mathbb{R}^k$ takav da je $f_X(x) > 0$, dakle, $x \in \operatorname{supp} f_X$. Tada je dobro definiran uvjetni zakon razdiobe od Y uz dano $X = x$ induciran vjerojatnosnom funkcijom razdiobe $F_{Y|X}(\cdot|x) : \mathbb{R}^\ell \to \mathbb{R}$, definiranom za $y \in \mathbb{R}^\ell$ s

$$F_{Y|X}(y|x) := \mathbb{P}(Y \leq y|X = x) = \begin{cases} \frac{\mathbb{P}(X=x, Y \leq y)}{\mathbb{P}(X=x)} & \text{ako je } X \text{ diskretna s.v.} \\ \int\limits_{-\infty}^{y} f_{Y|X}(z|x)\, dz & \text{ako je } (X, Y) \text{ neprekidan s.v.} \end{cases}$$

Ovdje je $f_{Y|X}(y|x) = f(x, y)/f_X(x)$ ($y \in \mathbb{R}^\ell$) uvjetna gustoća od Y uz dano $X = x$. Dakako, ako je slučajni vektor (X, Y) diskretan (tj. ako su X i Y obje diskretne slučajne veličine), tada (vidjeti potpoglavlje 1.4.) vrijedi da je

$$F_{Y|X}(y|x) = \mathbb{P}(Y \leq y|X = x) = \frac{\mathbb{P}(X = x,\, Y \leq y)}{\mathbb{P}(X = x)} = \sum_{z \leq y} f_{Y|X}(z|x), \ y \in \mathbb{R}^\ell,$$

gdje je $f_{Y|X}(y|x) = f_{X,Y}(x, y)/f_X(x) = \mathbb{P}(X = x, Y = y)/\mathbb{P}(X = x)$ ($y \in \mathbb{R}^\ell$) uvjetna gustoća diskretne razdiobe od Y uz dano $X = x$.
Označimo s $\mathbb{P}_{Y|X}(\cdot|x)$ uvjetni zakon razdiobe od Y uz dano $X = x$ za $x \in \operatorname{supp} f_X$. Ako je $g : \mathbb{R}^\ell \to \mathbb{R}$ Borelova funkcija, tada je g slučajna varijabla na vjerojatnosnom prostoru $(\mathbb{R}^\ell, \mathcal{B}(\mathbb{R}^\ell), \mathbb{P}_{Y|X}(\cdot|x))$. Pretpostavimo da je $g \in \mathcal{L}^1(\mathbb{R}^\ell, \mathcal{B}(\mathbb{R}^\ell), \mathbb{P}_{Y|X}(\cdot|x))$, tj. da matematičko očekivanje od g u odnosu na $\mathbb{P}_{Y|X}(\cdot|x)$ postoji i konačno je. Označimo ga s

$$\mathbb{E}[g(Y)|X = x] := \int\limits_{\mathbb{R}^\ell} g\, d\mathbb{P}_{Y|X}(\cdot|x) \equiv \int\limits_{\mathbb{R}^\ell} g(y)\, d\mathbb{P}_{Y|X}(y|x). \tag{2.18}$$

Ta oznaka je konzistentna sa svim do sada uvedenim istim oznakama. Naime, ako je X diskretna slučajna varijabla i $x \in \mathrm{supp} f_X$, tada je dobro definirana uvjetna vjerojatnost $\mathbb{P}(\cdot|\{X = x\}) \equiv \mathbb{P}(\cdot|X = x)$ na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. U tom slučaju je, prema teoremu o zamjeni varijabli (vidjeti teorem 10.11 u [9], str. 306.),

$$\int\limits_{\mathbb{R}^\ell} g(y)\, d\mathbb{P}_{Y|X}(y|x) \stackrel{\text{(tm.z.v.)}}{=} \int\limits_{\Omega} g(Y(\omega))\, d\mathbb{P}(\omega|X = x) \stackrel{(2.12)}{=}$$

$$= \quad \mathbb{E}[g(Y)|\{X = x\}] \stackrel{(2.13)}{=} \frac{\mathbb{E}(g(Y) \cdot \mathbb{1}_{\{X=x\}})}{\mathbb{P}(X = x)},$$

pri čemu lijeva strana postoji ako i samo ako postoji desna strana (prve) jednakosti. Ako je još i Y diskretna slučajna varijabla, tada je i

$$\int\limits_{\mathbb{R}^\ell} g(y)\, d\mathbb{P}_{Y|X}(y|x) = \sum_{y \in \mathbb{R}^\ell} g(y) f_{Y|X}(y|x). \tag{2.19}$$

Nadalje, ako je (X, Y) neprekidan slučajan vektor, tada je

$$\int\limits_{\mathbb{R}^\ell} g(y)\, d\mathbb{P}_{Y|X}(y|x) = \int\limits_{\mathbb{R}^\ell} g(y) f_{Y|X}(y|x)\, dy.$$

Dakle,

$$\mathbb{E}[g(Y)|X = x] = \begin{cases} \frac{\mathbb{E}(g(Y) \cdot \mathbb{1}_{\{X=x\}})}{\mathbb{P}(X=x)} & \text{ako je } X \text{ diskretna s.v.} \\ \int\limits_{\mathbb{R}^\ell} g(y) f_{Y|X}(y|x)\, dy & \text{ako je } (X, Y) \text{ neprekidna s.v.} \end{cases} \tag{2.20}$$

Definicija 2.15. *Neka su k-dimenzionalna slučajna veličina X i ℓ-dimenzionalna slučajna veličina Y $(k, \ell \geq 1)$ definirane na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ pri čemu je X diskretna slučajna veličina ili je (X, Y) apsolutno neprekidan slučajni vektor. Nadalje, neka je $g : \mathbb{R} \to \mathbb{R}$ takva Borelova funkcija da je za $\mathbb{P}_X$-g.s. x, takve da je $f_X(x) > 0$, $\mathbb{E}[|g(Y)|\,|X = x] < +\infty$. To znači da za Borelov skup*

$$N := \{x \in \mathrm{supp} f_X : \mathbb{E}[|g(Y)|\,|X = x] = +\infty\}$$

vrijedi da je $\mathbb{P}_X(N) = 0$. Tada funkciju $\phi : \mathbb{R}^k \to \mathbb{R}$ definiranu sa

$$\phi(x) := \begin{cases} \mathbb{E}[g(Y)|X = x], & f_X(x) > 0 \text{ i } x \notin N \\ 0, & f_X(x) = 0 \text{ ili } x \in N \end{cases}$$

zovemo **regresijska funkcija od $Z = g(Y)$ u odnosu na X.**

Sljedeći teorem kaže da u našem specijalnom slučaju regresijska funkcija postoji, da je integrabilna Borelova funkcija te da je kompozicija te funkcije i X jedna verzija od $\mathbb{E}[Z|X]$. Dakle, teorem je poveznica pojmova uvjetnog očekivanja i očekivanja u odnosu na uvjetnu distribuciju.

Kao i prije, označimo sa $\mathcal{L}^1(\mathbb{P}_X) \equiv \mathcal{L}^1(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_X)$ prostor svih realnih Borelovih funkcija definiranih na $\mathbb{R}^k$ koje su (konačno) integrabilne u odnosu na zakon razdiobe $\mathbb{P}_X$ od X.

Teorem 2.16. *Neka su k-dimenzionalna slučajna veličina X i ℓ-dimenzionalna slučajna veličina Y $(k, \ell \geq 1)$ definirane na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ pri čemu je X diskretna slučajna veličina ili je (X, Y) apsolutno neprekidan slučajni vektor. Nadalje, neka je $g : \mathbb{R} \to \mathbb{R}$ Borelova funkcija. Ako je $\mathbb{E}(|g(Y)|) < +\infty$, tada je*

(i) $\mathbb{E}[|g(Y)|\,|X=x] < +\infty$ *za* $\mathbb{P}_X$*-g.s.* x *takve da je* $f_X(x) > 0$,
(ii) *postoji regresijska funkcija* ϕ *od* $g(Y)$ *u odnosu na* X *i* $\phi \in \mathcal{L}^1(\mathbb{P}_X)$, *te*
(iii) $\phi(X) = \mathbb{E}[g(Y)|X]$ *g.s.*

Dokaz. Pretpostavimo prvo da je X diskretna slučajna veličina. Neka je nosač gustoće od X probrojiv (konačan ili beskonačan) skup $D = \text{supp} f_X = \{x_1, x_2, \ldots\}$. Pretpostavljamo da su elementi skupa D indeksirani na način da su međusobno različiti, tj. da vrijedi da kadagod je $k \neq \ell$ da je nužno $x_k \neq x_\ell$. Tada množina

$$\mathcal{H} = \{H_k := \{X = x_k\} \, : k \geq 1\}$$

čini potpun sustav događaja na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ i vrijedi da je $\sigma(\mathcal{H}) = \sigma(X)$. U skladu s primjerom 2.13 stavimo da je za $k \geq 1$, $y_k := \mathbb{E}[g(Y)|H_k] = \mathbb{E}[g(Y)|X = x_k]$. Tada je uvjetno matematičko očekivanje od $g(Y)$ uz dano $\sigma(X)$ dano izrazom (2.14):

$$\mathbb{E}[g(Y)|X] \overset{(2.14)}{=} \sum_{k\geq 1} y_k \cdot \mathbb{1}_{H_k} = \sum_{k\geq 1} y_k \cdot \mathbb{1}_{\{X=x_k\}} = \sum_{k\geq 1} y_k \cdot \mathbb{1}_{\{x_k\}}(X) = \psi(X)$$

gdje je $\psi := \sum_{k\geq 1} y_k \cdot \mathbb{1}_{\{x_k\}}$ očito Borelova funkcija. Nadalje, budući da je za $x \in \mathbb{R}^k$,

$$\psi(x) = \left\{ \begin{array}{ll} y_k = \mathbb{E}[g(Y)|X = x_k], & x = x_k \in D \\ 0, & x \notin D \end{array} \right\} = \phi(x),$$

gdje je $\phi : \mathbb{R}^k \to \mathbb{R}$ regresijska funkcija, slijede tvrdnje *(ii)* i *(iii)*. Naime, tvrdnja da je $\phi \in \mathcal{L}^1(\mathbb{P}_X)$ vrijedi ako je, uz izmjerivost od ϕ, ispunjeno

$$\int_{\mathbb{R}^k} |\phi(x)| d\mathbb{P}_X(x) = \mathbb{E}[|\phi(X)|] < +\infty,$$

a to je svojstvo (2) iz definicije uvjetnog očekivanja 2.11. Budući da je $\phi(X) = \mathbb{E}[g(Y)|X]$, vrijedi definicijsko svojstvo (2), dakle, $\phi \in \mathcal{L}^1(\mathbb{P}_X)$.

Tvrdnja *(i)* trivijalno slijedi iz formule (2.13) (ili (2.16)), monotonosti matematičkog očekivanja i pretpostavke da je $g(Y) \in \mathcal{L}^1$:

$$\mathbb{E}[|g(Y)|\,|X = x] \overset{(2.13)}{=} \frac{\mathbb{E}[|g(Y)| \cdot \mathbb{1}_{\{X=x\}}]}{\mathbb{P}(X = x)} \overset{\text{(mon.m.o)}}{\leq} \frac{\mathbb{E}[|g(Y)|]}{\mathbb{P}(X = x)} < +\infty$$

za sve $x \in \mathbb{R}^k$ takve da je $f_X(x) = \mathbb{P}(X = x) > 0$. Ovdje je zanemariv skup N iz definicije 2.15 prazan skup ($N = \emptyset$).

Neka je sada (X, Y) neprekidan slučajni vektor s gustoćom $f : \mathbb{R}^{k+\ell} \to \mathbb{R}$. Budući da su X i Y neprekidne slučajne veličine s (marginalnim) gustoćama $f_X(x) = \int_{\mathbb{R}} f(x, y)\, dy$ (za λ^k-s.s. $x \in \mathbb{R}^k$) i $f_Y(y) = \int_{\mathbb{R}} f(x, y)\, dx$ (za λ^ℓ-s.s. $y \in \mathbb{R}^\ell$) i budući da vrijedi da je $f(x, y) = f_{Y|X}(y|x) f_X(x)$ za sve $y \in \mathbb{R}^\ell$ i $x \in \text{supp} f_X$, prema Fubinijevom teoremu A.1 *(i)* i korištenjem linearnosti integrala,

$$\mathbb{E}(|g(Y)|) = \int\limits_{\mathbb{R}^\ell} |g(y)| f_Y(y) dy = \int\limits_{\mathbb{R}^\ell} |g(y)| (\int\limits_{\mathbb{R}^k} f(x,y) dx) dy \overset{\text{(lin.int.)}}{=}$$

$$= \int\limits_{\mathbb{R}^\ell} (\int\limits_{\mathbb{R}^k} |g(y)| f(x,y) dx) dy \overset{\text{(Fubini tm.)}}{=} \int\limits_{\mathbb{R}^k} (\int\limits_{\mathbb{R}^\ell} |g(y)| f(x,y) dy) dx =$$

$$= \int\limits_{\mathrm{supp} f_X} (\int\limits_{\mathbb{R}^\ell} |g(y)| f_{Y|X}(y|x) f_X(x) dy) dx \overset{\text{(lin.int.)}}{=} \int\limits_{\mathrm{supp} f_X} (\int\limits_{\mathbb{R}^\ell} |g(y)| f_{Y|X}(y|x)\, dy) f_X(x) dx \overset{(2.18)}{=}$$

$$= \int\limits_{\mathrm{supp} f_X} \mathbb{E}[|g(Y)|\,|X = x] f_X(x) dx = \int\limits_{\mathrm{supp} f_X} \mathbb{E}[|g(Y)|\,|X = x] d\mathbb{P}_X(x).$$

Po pretpostavci znamo da je $\mathbb{E}(|g(Y)|) < \infty$. Dakle, za $\mathbb{P}_X$-g.s. $x \in \mathrm{supp} f_X$ je $\mathbb{E}[|g(Y)|\,|X = x] < +\infty$ (vidjeti npr. propoziciju 10.4 (*iii*) iz [9], str. 294.), pa vrijedi tvrdnja (*i*). Prema tome, regresijska funkcija ϕ od $g(Y)$ u odnosu na X je dobro definirana. Neka je $N = \{x \in \mathrm{supp} f_X : \mathbb{E}[|g(Y)|\,|X = x] = +\infty\}$ Borelov skup u $\mathbb{R}^k$ takav da je $\mathbb{P}_X(N) = 0$. Budući da je $N \subset \mathrm{supp} f_X$, primijetite da je

$$\lambda^k(N) = \int\limits_N dx = \int\limits_N \frac{1}{f_X(x)} f_X(x)\, dx = \int\limits_N \frac{1}{f_X(x)}\, dP_X(x) = 0,$$

dakle, N je skup Lebesgueove mjere nula (na $\mathbb{R}^k$). Budući da je $g(Y) \in \mathcal{L}^1$, ponovo primjenom Fubinijevog teorema A.1 (*ii*) (jer je podintegrabilna funkcija inegrabilna) i linearnosti integrala, slijedi da je

$$\begin{aligned} \mathbb{E}g(Y) &= \int\limits_{\mathbb{R}^\ell} g(y) f_Y(y) dy = \int\limits_{\mathbb{R}^\ell} g(y) (\int\limits_{\mathbb{R}^k} f(x,y) dx) dy \overset{\text{(lin.int.)}}{=} \\ &= \int\limits_{\mathbb{R}^\ell} (\int\limits_{\mathbb{R}^k} g(y) f(x,y) dx) dy \overset{\text{(Fubini tm.)}}{=} \int\limits_{\mathbb{R}^k} (\int\limits_{\mathbb{R}^\ell} g(y) f(x,y) dy) dx. \end{aligned}$$

Iz Fubinijevog teorema A.1 (*ii*) zaključujemo da postoji Borelova funkcija $h : \mathbb{R}^k \to \mathbb{R}$ takva da je $h(x) = \int\limits_{\mathbb{R}^\ell} g(y) f(x,y) dy$ za $x \in \mathrm{supp} f_X \cap N^c$. Budući da je za $x \in \mathrm{supp} f_X \cap N^c$,

$$h(x) = \int\limits_{\mathbb{R}^\ell} g(y) f(x,y) dy = f_X(x) \int\limits_{\mathbb{R}^\ell} g(y) f_{Y|X}(y|x)\, dy = f_X(x) \mathbb{E}[g(Y)|X = x],$$

$\mathrm{supp} f_X$ i N su Borelovi skupovi, a f_X Borelova funkcija, te da je za sve $x \in \mathbb{R}^k$,

$$\phi(x) = \begin{cases} \frac{h(x)}{f_X(x)}, & x \in \mathrm{supp} f_X \cap N^c \\ 0, & x \in (\mathrm{supp} f_X)^c \cup N, \end{cases}$$

regresijska funkcija ϕ je Borelova kao izmjeriva kombinacija takvih funkcija. Da bi dovršili dokaz da je $\phi \in \mathcal{L}^1(\mathbb{P}_X)$, a time i dokazali tvrdnju (*ii*) u potpunosti, dovoljno je prvo pokazati (*iii*) (tada će ta tvrdnja slijediti iz definicijskog svojstva (2) uvjetnog očekivanja). Dakle, preostaje dokazati da je $\hat{Z} := \phi(X)$ verzija uvjetnog očekivanja od $Z = g(Y)$ uz dano $\sigma(X)$. To znači da treba pokazati da za $\hat{Z}$ vrijede svojstva $(1-3)$ iz definicije 2.11 uvjetnog očekivanja. Budući da je $\hat{Z} = \phi(X)$ i ϕ je Borelova funkcija, $\hat{Z}$ je, prema teoremu

2.14, $\sigma(X)$-izmjeriva varijabla. Dakle, vrijedi svojstvo (1). Dokažimo sada svojstvo (2), tj. da je $\hat{Z} \in \mathcal{L}^1$. Neka je $N \subset \mathrm{supp} f_X$ zanemariv skup na kojem matematičko očekivanje uvjetne distribucije od $|g(Y)|$ uz dano $X = x$ nije konačno. Stavimo $S := \mathrm{supp} f_X \cap N^c$. Tada, primjenom monotonosti integrala i Fubinijevog teorema (podintegralna funkcija je nenegativna), slijedi:

$$\begin{aligned}
\mathbb{E}(|\hat{Z}|) &= \int\limits_{\mathbb{R}^k} |\phi(x)| f_X(x)\,dx = \int\limits_S |\mathbb{E}[g(Y)|X=x]| f_X(x)\,dx \overset{(2.18)}{=} \\
&= \int\limits_S |\int\limits_{\mathbb{R}^\ell} g(y) f_{Y|X}(y|x)\,dy| f_X(x)\,dx \overset{\text{(mon.int.)}}{\leq} \int\limits_S (\int\limits_{\mathbb{R}^\ell} |g(y)| f_{Y|X}(y|x)\,dy) f_X(x)\,dx \overset{\text{(lin.int.)}}{=} \\
&= \int\limits_S (\int\limits_{\mathbb{R}^\ell} |g(y)| f_{Y|X}(y|x) f_X(x)\,dy)\,dx = \int\limits_S (\int\limits_{\mathbb{R}^\ell} |g(y)| f(x,y)\,dy)\,dx = \\
&= \int\limits_{\mathbb{R}^k} (\int\limits_{\mathbb{R}^\ell} |g(y)| f(x,y)\,dy)\,dx \overset{\text{(Fubini tm.)}}{=} \int\limits_{\mathbb{R}^\ell} (\int\limits_{\mathbb{R}^k} |g(y)| f(x,y)\,dx)\,dy \overset{\text{(lin.int.)}}{=} \\
&= \int\limits_{\mathbb{R}^\ell} |g(y)| (\int\limits_{\mathbb{R}^k} f(x,y)\,dx)\,dy = \int\limits_{\mathbb{R}^\ell} |g(y)| f_Y(y)\,dy = \mathbb{E}(|g(Y)|) < +\infty,
\end{aligned}$$

dakle, vrijedi (2). Neka je sada $G \in \sigma(X)$ proizvoljni događaj. Tada postoji Borelov skup $B \subseteq \mathbb{R}^k$ takav da je $G = \{X \in B\}$. Vrijedi da je

$$\begin{aligned}
\mathbb{E}(\hat{Z} \cdot \mathbb{1}_G) &= \mathbb{E}(\phi(X) \cdot \mathbb{1}_B(X)) = \int\limits_{\mathbb{R}^k} \phi(x) \mathbb{1}_B(x) f_X(x)\,dx \overset{(2.18)}{=} \\
&= \int\limits_S (\int\limits_{\mathbb{R}^\ell} g(y) f_{Y|X}(y|x)\,dy) \mathbb{1}_B(x) f_X(x)\,dx \overset{\text{(lin.int.)}}{=} \int\limits_S (\int\limits_{\mathbb{R}^\ell} g(y) \mathbb{1}_B(x) f(x,y)\,dy)\,dx = \\
&= \int\limits_{\mathbb{R}^k} (\int\limits_{\mathbb{R}^\ell} g(y) \mathbb{1}_B(x) f(x,y)\,dy)\,dx \overset{\text{(Fubini tm.)}}{=} \int\limits_{\mathbb{R}^k \times \mathbb{R}^\ell} g(y) \mathbb{1}_B(x) f(x,y)\,d(x,y) \overset{\text{(tm.z.v.)}}{=} \\
&= \mathbb{E}(g(Y) \cdot \mathbb{1}_B(X)) = \mathbb{E}(g(Y) \cdot \mathbb{1}_G),
\end{aligned}$$

primjenom Fubinijevog teorema (jer je $\mathbb{E}(|g(Y) \cdot \mathbb{1}_B(X)|) \leq \mathbb{E}(|g(Y)|) < +\infty$) i teorema o zamjeni varijabli. Budući da je $G \in \sigma(X)$ bio proizvoljan, vrijedi i svojstvo (3) pa je, po definiciji 2.11, $\phi(X) = \hat{Z} = \mathbb{E}[g(Y)|X]$. $\square$

Primjer 2.17. Neka su $X_1, X_2, \ldots, X_n$ nezavisne jednako distribuirane Bernoullijeve slučajne varijable s vjerojatnosti uspjeha $p \in \langle 0, 1\rangle$ $(n \geq 2)$. Tada je $T := X_1 + \cdots + X_n$ binomna $b(n,p)$ slučajna varijabla (vidjeti zadatak 1.47 (a)). T se interpretira kao ukupan broj uspjeha u nizu od n nezavisnih ponavljanja istovrsnih bernoullijevskih pokusa. Želimo naći $\mathbb{E}[X_1|T]$. Budući da je T diskretna slučajna varijabla, $\mathbb{E}[X_1|T] = \phi(T)$ prema teoremu 2.16 za regresijsku funkciju $\phi : \mathbb{R} \to \mathbb{R}$ od X_1 u odnosu na T. Kako je $T \sim b(n,p)$, dovoljno je izračunati $\phi(k)$ za $k \in \mathbb{N}_n := \{0, 1, \ldots, n\}$. Primijetimo da je X_1 Bernoullijeva slučajna varijabla i uz uvjetnu razdiobu od X_1 uz dano $T = k$. Dakle,

$$\begin{aligned}
\phi(k) &= \mathbb{E}[X_1|T=k] = \mathbb{P}(X_1 = 1|T = k) = \frac{\mathbb{P}(X_1 = 1, X_1 + X_2 + \cdots + X_n = k)}{\mathbb{P}(T = n)} = \\
&= \frac{\mathbb{P}(X_1 = 1, X_2 + \cdots + X_n = k-1)}{\mathbb{P}(T = n)} \overset{\text{(nez.)}}{=} \frac{\mathbb{P}(X_1 = 1)\mathbb{P}(X_2 + \cdots + X_n = k-1)}{\mathbb{P}(T = n)} = \\
&= \frac{p \cdot \binom{n-1}{k-1} p^{k-1}(1-p)^{n-k}}{\binom{n}{k} p^k (1-p)^{n-k}} = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\binom{n-1}{k-1}}{\frac{n}{k}\binom{n-1}{k-1}} = \frac{k}{n},
\end{aligned}$$

pri čemu smo u drugom retku koristili činjenicu da su X_1 i $X_2+\cdots+X_n$ nezavisne slučajne varijable, a pri prijelazu iz drugog u treći redak da je $X_2+\cdots+X_n \sim b(n-1,p)$ (zadatak 1.47 (a)). Prema tome, $\mathbb{E}[X_1|T]=\phi(T)=T/n$. Interpretacija: ako je poznat ukupan broj uspjeha, uvjetna vjerojatnost uspjeha u prvom pokusu je jednaka relativnoj frekvenciji uspjeha. Moglo bi se još reći da je najbolja procjena vjerojatnosti uspjeha u prvom pokusu ako je poznat ukupan broj uspjeha, jednaka relativnoj frekvenciji uspjeha. Primijetimo da bi isti rezultat dobili da smo računali $\mathbb{E}[X_j|T=k]$ za $j=2,\ldots,n$. Dakle,

$$\mathbb{E}[X_1|T]=\mathbb{E}[X_1|T]=\cdots=\mathbb{E}[X_n|T]=\frac{T}{n} \text{ g.s.}$$

Slijedi da su slučajne varijable $X_1,\ldots,X_n$ uvjetno jednako distribuirane Bernoullijeve s vjerojatnosti uspjeha jednakoj relativnoj frekvenciji uspjeha uvjetno na dani ukupan broj uspjeha.

Je li niz $X_1,\ldots,X_n$ i uvjetno nezavisan uz dano $T=k$? Pokažimo da je odgovor na to pitanje niječan. Prema zadatku 1.47 (b), uvjetna distribucija slučajnog vektora $(X_1,\ldots,X_n)$ uz dano $T=k$ je diskretna uniformna na skupu svih uređenih n-torki elemenata 0 i 1 takvih da svaka n-torka ima točno k elemenata jednakih 1. Dakle, ako je $(x_1,\ldots,x_n)$ jedna takva n-torka, tada je

$$\mathbb{P}(X_1=x_1,\ldots,X_n=x_n|T=k)=\frac{1}{\binom{n}{k}}=\frac{k\cdot(k-1)\cdots 1}{n\cdot(n-1)\cdots(n-k+1)}=(*).$$

S druge strane, kada bi $X_1,\ldots,X_n$ bile uvjetno nezavisne slučajne varijable uz dano $T=k$, tada bi za istu n-torku nula i jedinica bilo

$$\begin{aligned}\mathbb{P}(X_1=x_1,\ldots,X_n=x_n|T=k) &\overset{\text{(nez.)}}{=} \mathbb{P}(X_1=x_1|T=k)\cdots\mathbb{P}(X_n=x_n|T=k)=\\ &= \left(\frac{k}{n}\right)^k\left(1-\frac{k}{n}\right)^{n-k}=\frac{k^k\cdot\ (n-k)^{n-k}}{n^n}\end{aligned}$$

što nije jednako izrazu $(*)$. Dakle, slučajne varijable $X_1,\ldots,X_n$ nisu uvjetno nezavisne uz dano $T=k$. □

Primjer 2.18. Neka slučajni vektor (X,Y) ima bivarijatnu normalnu razdiobu (vidjeti primjer 1.26) s očekivanjima μ_1 i μ_2, varijancama $\sigma_1^2>0$ i $\sigma_2^2>0$ od X i Y redom, te koeficijentom korelacije $\rho\in\langle -1,1\rangle$. U primjeru 1.26 je pokazano da je za svaki $x\in\mathbb{R}$ uvjetna razdioba od Y uz dano $X=x$ normalna razdioba s očekivanjem $\mu_2+\rho\sigma_2(x-\mu_1)/\sigma_1$ i varijancom $\sigma_2^2(1-\rho^2)$. Dakle, regresijska funkcija $\phi:\mathbb{R}\to\mathbb{R}$ je jednaka $\phi(x)=\mathbb{E}[Y|X=x]=\mu_2+\rho\sigma_2(x-\mu_1)/\sigma_1$ za $x\in\mathbb{R}$. Prema teoremu 2.16,

$$\mathbb{E}[Y|X]=\mu_2+\rho\cdot\frac{\sigma_2}{\sigma_1}\cdot(X-\mu_1).$$

Zbog simetrije je

$$\mathbb{E}[X|Y]=\mu_1+\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot(Y-\mu_2).$$

Primijetite da je $\mathbb{E}[Y|X]$ jednako najboljem linearnom predviditelju od Y pomoću X (vidjeti primjer 2.6). Analogno je $\mathbb{E}[X|Y]$ jednako najboljem linearnom predviditelju od X pomoću Y. Općenito se najbolji linearni predviditelj i uvjetno matematičko očekivanje ne moraju podudarati (niti g.s.) (vidjeti zadatak 2.25 u nastavku). □

Zadatak 2.19. Pretpostavimo da za zadanu slučajnu varijablu X takvu da je $\mathbb{E}|X|<+\infty$ i zadanu σ-podalgebru $\mathcal{G}$ od $\mathcal{F}$ postoji slučajna varijabla Y na istom vjerojatnosnom prostoru $(\Omega,\mathcal{F},\mathbb{P})$ takva da je

(i) Y je $\mathcal{G}$-izmjeriva i
(ii) za svaki $G \in \mathcal{G}$, $\mathbb{E}[Y \cdot \mathbb{1}_G]$ postoji i $\mathbb{E}[Y \cdot \mathbb{1}_G] = \mathbb{E}[X \cdot \mathbb{1}_G]$.

Dokažite da je tada nužno $\mathbb{E}|Y| < +\infty$. Primijetite da je tada svojstvo (2) u definiciji 2.11 uvjetnog očekivanja suvišno. Dakako, tada se u svojstvu (3) implicitno pretpostavlja i postojanje integrala s lijeve strane jednakosti.

Zadatak 2.20. Neka su $X_1, X_2, \ldots, X_n$ ($n \geq 2$) nezavisne slučajne varijable. Za Borelovu ograničenu funkciju $h : \mathbb{R}^n \to \mathbb{R}$ definiramo preslikavanje $\gamma_h : \mathbb{R} \to \mathbb{R}$ sa

$$\gamma_h(x) := \mathbb{E}[h(x, X_2, \ldots, X_n)], \quad x \in \mathbb{R}.$$

Dokažite da je $\gamma_h(X_1) = \mathbb{E}[h(X_1, X_2, \ldots, X_n)|X_1]$.
(**Uputa:** Zbog nezavisnosti komponenata, zakon razdiobe slučajnog vektora $X = (X_1, X_2, \ldots, X_n)$ je $\mathbb{P}_X = \mathbb{P}_1 \times \mathbb{P}_2 \times \cdots \times \mathbb{P}_n$ gdje su $\mathbb{P}_k \equiv \mathbb{P}_{X_k}$ ($1 \leq k \leq n$), vidjeti dodatak A. Primijenite Fubinijev teorem i za dokaz izmjerivosti funkcije γ_h, i za dokaz da je $\gamma_h(X_1) = \mathbb{E}[h(X_1, X_2, \ldots, X_n)|X_1]$.)

Zadatak 2.21. Neka su: (X, Y) slučajan vektor iz iskaza teorema 2.16, $g : \mathbb{R}^\ell \to \mathbb{R}$ Borelova funkcija i $p \geq 1$ realan broj. Kažemo da je slučajna varijabla Z u prostoru $\mathcal{L}^p \equiv \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ ako je $\mathbb{E}[|Z|^p] < +\infty$. Slično, sa $\mathcal{L}^p(\mathbb{P}_X) \equiv \mathcal{L}^p(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mathbb{P}_X)$ označimo prostor svih Borelovih funkcija $f : \mathbb{R}^k \to \mathbb{R}$ takvih da je $\int |f(x)|^p d\mathbb{P}_X(x) = \mathbb{E}[|f(X)|^p] < +\infty$. Dokažite da vrijedi sljedeća tvrdnja. Ako je $Z := g(Y) \in \mathcal{L}^p$ tada vrijedi

(i) $\mathbb{E}[|Z|^p\, |X = x] < +\infty$ za $\mathbb{P}_X$-g.s. x takve da je $f_X(x) > 0$,
(ii) regresijska funkcija ϕ od Z u odnosu na X postoji i $\phi \in \mathcal{L}^p(\mathbb{P}_X)$,
(iii) svaka verzija od $\mathbb{E}[Z|X]$ je u $\mathcal{L}^p$.

(**Uputa:** Dokazati da je $\mathbb{E}[|\hat{Z}|^p] < +\infty$, za $\hat{Z} = \phi(X)$, na isti način kao što je u dokazu teorema 2.16 pokazano da je $\mathbb{E}[|\hat{Z}|] < +\infty$, samo što se umjesto monotonosti matematičkog očekivanja treba primijeniti Jensenova nejednakost (teorem 1.56 (vii)) na funkciju $x \mapsto |x|^p$.)

Zadatak 2.22. Neka je (X, Y) neprekidan $k + \ell$-dimenzionalan slučajni vektor ili je X k-dimenzionalna diskretna slučajna veličina. Nadalje, neka je $h : \mathbb{R}^{k+\ell} \to \mathbb{R}$ ograničena Borelova funkcija. Ako sa ϕ_h definiramo preslikavanje takvo da je $\phi_h(x) := \mathbb{E}[h(x, Y)|X = x]$ ako je $f_X(x) > 0$ i $\phi_h(x) := 0$ ako je $f_X(x) = 0$, tada je $\phi_h \in \mathcal{L}^1(\mathbb{P}_X)$ i $\phi_h(X) = \mathbb{E}[h(X, Y)|X]$. Dokažite.

Zadatak 2.23. Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor t.d. je $\Omega = [-1, 1]$, $\mathcal{F} = \mathcal{B}([-1, 1])$ i $\mathbb{P} = \frac{1}{2}\lambda$ ($\lambda \equiv \lambda^1$ je Lebesgueova mjera na $\mathbb{R}$). Neka su X, Y slučajne varijable takve da je $\mathbb{E}[|Y|] < +\infty$ i $X(\omega) := |\omega|$, $\omega \in \Omega$. Nađite $\mathbb{E}[Y|X]$ ako postoji.
(**Uputa:** Pokažite da je $\sigma(X) = \{(-B) \cup B : B \in \mathcal{B}([0, 1])\}$, te da je svaka parna funkcija izmjeriva u odnosu na tu σ-algebru. Tada je parni dio od X jedna verzija od $\mathbb{E}[Y|X]$.)

Zadatak 2.24. Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor takav da je je $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$ i $\mathbb{P} = \lambda$ Lebesgueova mjera na $\mathbb{R}$. Neka su X, Y slučajne varijable definirane sa

$$X(\omega) := 2 \cdot \mathbb{1}_{[0, \frac{1}{3}\rangle}(\omega) + 5 \cdot \mathbb{1}_{[\frac{1}{3}, \frac{2}{3}]}(\omega), \quad Y(\omega) := \omega, \quad \omega \in \Omega.$$

Odredite $\mathbb{E}[Y|X]$ ako postoji.

Zadatak 2.25. Neka je (U, V) bivarijatni normalni slučajni vektor s oba parametra očekivanja jednaka 0, oba parametra varijance jednaka 1 i s koeficijentom korelacije od U i V jednakim $\rho \in \langle -1, 1\rangle$ (vidjeti primjer 1.26). Definiran je novi slučajni vektor $(X, Y) := (e^U, e^V)$. Odredite $\mathbb{E}[Y|X]$ i $\mathbb{E}[X|Y]$ ako postoje. Da li se ta uvjetna očekivanja g.s. podudaraju s odgovarajućim najboljim linearnim predviditeljima?

2.4 Svojstva uvjetnog matematičkog očekivanja

Navest ćemo osnovna svojstva uvjetnog matematičkog očekivanja u jednom teoremu. Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ zadani vjerojatnosni prostor, te neka je $\mathcal{L}^1$ pripadni prostor slučajnih varijabli konačnog matematičkog očekivanja, definiranih na tom prostoru. Nadalje, kažemo da su slučajna varijabla Y, definirana na tom vjerojatnosnom prostoru, i σ-podalgebra $\mathcal{G}$ od $\mathcal{F}$ *nezavisni* ako su za svaki događaj $G \in \mathcal{G}$ slučajne varijable Y i $\mathbb{1}_G$ nezavisne.

Teorem 2.26. *Neka su Y i Y_n ($n \in \mathbb{N}$) slučajne varijable iz prostora $\mathcal{L}^1$, te $\mathcal{H}$ i $\mathcal{G}$ σ-podalgebre od $\mathcal{F}$. Tada vrijedi sljedeće:*

(*i*) $\mathbb{E}(\mathbb{E}[Y|\mathcal{G}]) = \mathbb{E}Y$
(*ii*) *Ako je Y $\mathcal{G}$-izmjeriva varijabla, tada je $\mathbb{E}[Y|\mathcal{G}] = Y$ g.s.*
(*iii*) (Linearnost uvjetnog očekivanja) *Za svaki $\alpha_1, \alpha_2 \in \mathbb{R}$ je* $\mathbb{E}[\alpha_1 Y_1 + \alpha_2 Y_2|\mathcal{G}] = \alpha_1 \mathbb{E}[Y_1|\mathcal{G}] + \alpha_2 \mathbb{E}[Y_2|\mathcal{G}]$ *g.s.*
(*iv*) (Monotonost uvjetnog očekivanja) *Ako je $Y \geq 0$ g.s., tada je $\mathbb{E}[Y|\mathcal{G}] \geq 0$ g.s. Posljedično, ako je $Y_1 \leq Y_2$ g.s., tada je $\mathbb{E}[Y_1|\mathcal{G}] \leq \mathbb{E}[Y_1|\mathcal{G}]$ g.s.*
(*v*) (Teorem o uvjetnoj monotonoj konvergenciji) *Ako je za svaki n, $0 \leq Y_n \leq Y_{n+1}$ g.s. i $\lim_n \uparrow Y_n = Y$ g.s., tada je $\lim_n \uparrow \mathbb{E}[Y_n|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$ g.s.*
(*vi*) (Uvjetna Fatouova lema) *Ako je za svaki n, $0 \leq Y_n$ g.s. i $\underline{\lim}_n Y_n = Y$ g.s., tada je $\mathbb{E}[Y|\mathcal{G}] \leq \underline{\lim}_n \mathbb{E}[Y_n|\mathcal{G}]$ g.s.*
(*vii*) (Teorem o uvjetnoj dominiranoj konvergenciji) *Ako je za svaki n, $|Y_n| \leq V$ g.s. za neku slučajnu varijablu $V \in \mathcal{L}^1$ i $\lim_n Y_n = Y$ g.s., tada je $\lim_n \mathbb{E}[Y_n|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$ g.s.*
(*viii*) (Jensenova uvjetna nejednakost) *Ako je $\phi : \mathbb{R} \to \mathbb{R}$ konveksna funkcija takva da je $\phi(Y) \in \mathcal{L}^1$, tada je $\phi(\mathbb{E}[Y|\mathcal{G}]) \leq \mathbb{E}[\phi(Y)|\mathcal{G}]$ g.s.*
(*ix*) (Teleskopsko pravilo) *Ako je $\mathcal{H} \subset \mathcal{G}$, tada je $\mathbb{E}[\,\mathbb{E}[Y|\mathcal{G}]\,|\mathcal{H}] = \mathbb{E}[Y|\mathcal{H}]$ g.s.*
(*x*) *Ako je Z $\mathcal{G}$-izmjeriva varijabla takva da je $Z \cdot Y \in \mathcal{L}^1$, tada je $\mathbb{E}[ZY|\mathcal{G}] = Z \cdot \mathbb{E}[Y|\mathcal{G}]$ g.s.*
(*xi*) *Ako su Y i $\mathcal{G}$ nezavisni, tada je $\mathbb{E}[Y|\mathcal{G}] = \mathbb{E}Y$ g.s.*

Dokaz. Svojstva (*i*) i (*ii*) slijede odmah iz definicije 2.11 uvjetnog očekivanja: ako za G u (3) uzmeno $G = \Omega$, slijedi (*i*), a $\hat{Y} \equiv Y$ trivijalno zadovoljava definicijske uvjete $(1-3)$ uvjetnog očekivanja $\mathbb{E}[Y|\mathcal{G}]$, pa slijedi i (*ii*).

Dokažimo linearnost (*iii*). Neka su $\hat{Y}_i := \mathbb{E}[Y_i|\mathcal{G}]$, $i = 1, 2$, verzije uvjetnih očekivanja. Za proizvoljne realne brojeve α_i, $i = 1, 2$, neka je $Z := \alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2$. Tada je linearna kombinacija Z $\mathcal{G}$-izmjeriva varijabla konačnog matematičkog očekivanja jer je $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ vektorski potprostor od $\mathcal{L}^1$. Nadalje, neka je $G \in \mathcal{G}$ proizvoljan. Tada je zbog lineranosti očekivanja (teorem 1.56 (*i*)) i svojstva (3) definicije 2.11 primijenjenog na verzije $\hat{Y}_1$ i $\hat{Y}_2$,

$$\begin{aligned} \mathbb{E}(Z \cdot \mathbb{1}_G) &= \mathbb{E}((\alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2) \cdot \mathbb{1}_G) \stackrel{\text{(lin.m.o.)}}{=} \alpha_1 \mathbb{E}(\hat{Y}_1 \cdot \mathbb{1}_G) + \alpha_2 \mathbb{E}(\hat{Y}_2 \cdot \mathbb{1}_G) \stackrel{(3)}{=} \\ &= \alpha_1 \mathbb{E}(Y_1 \cdot \mathbb{1}_G) + \alpha_2 \mathbb{E}(Y_2 \cdot \mathbb{1}_G) \stackrel{\text{(lin.m.o.)}}{=} \mathbb{E}((\alpha_1 Y_1 + \alpha_2 Y_2) \cdot \mathbb{1}_G). \end{aligned}$$

Dakle, $Z = \alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2$ ima definicijska svojstva $(1-3)$ uvjetnog očekivanja $\mathbb{E}[\alpha_1 Y_1 + \alpha_2 Y_2|\mathcal{G}]$ pa vrijedi (*iii*).

Dokažimo monotonost uvjetnog očekivanja (*iv*). Prvo, neka je $Y \in \mathcal{L}^1$ takva da je $Y \geq 0$ g.s. Definirajmo nenegativnu slučajnu varijablu

$$Y'(\omega) := \begin{cases} Y(\omega), & \omega \in \Omega' \\ 0, & \omega \notin \Omega', \end{cases}$$

gdje je $\Omega' := \{Y \geq 0\}$ događaj vjerojatnosti jedan. Budući da je $Y' = Y$ g.s., Y' ima konačno očekivanje, a kako je i nenegativna, na nju možemo primijeniti lemu 2.9 (primijetite da se u dokazu te leme koristi samo da su varijable iz $\mathcal{L}^1$). Dakle, za svaku verziju $\hat{Y}'$ od $\mathbb{E}[Y'|\mathcal{G}]$ vrijedi da je $\hat{Y}' \geq 0$ g.s. Pokažimo da je $\mathbb{E}[Y'|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$ pa će time prva tvrdnja biti dokazana. Prvo, pokažimo da je bilo koja verzija $\hat{Y}'$ od $\mathbb{E}[Y'|\mathcal{G}]$ i verzija od $\mathbb{E}[Y|\mathcal{G}]$. Neka je $G \in \mathcal{G}$ proizvoljan. Tada, primjenom definicijskog svojstva (3) slijedi

$$\mathbb{E}(\hat{Y}' \cdot \mathbb{1}_G) \overset{(3)}{=} \mathbb{E}(Y' \cdot \mathbb{1}_G) = \mathbb{E}(Y' \cdot \mathbb{1}_G \cdot \mathbb{1}_{\Omega'}) = \mathbb{E}(Y \cdot \mathbb{1}_G \cdot \mathbb{1}_{\Omega'}) = \mathbb{E}(Y \cdot \mathbb{1}_G).$$

Ako u gornjem nizu jednakosti zamijenimo $\hat{Y}'$ s verzijom $\hat{Y}$ od $\mathbb{E}[Y|\mathcal{G}]$, te Y' zamijenimo s Y i obratno, vidimo da vrijedi i obratna tvrdnja: da je svaka verzija od $\mathbb{E}[Y|\mathcal{G}]$ ujedno i verzija od $\mathbb{E}[Y'|\mathcal{G}]$. Dakle, $\mathbb{E}[Y'|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$. Posebna tvrdnja je direktna posljedica prve tvrdnje primijenjene na varijablu $Y := Y_2 - Y_1$ (vidjeti dokaz druge tvrdnje leme 2.9).

Dokažimo teorem o uvjetnoj monotonoj konvergenciji (v). Zbog istoga razloga kao i u prethodnom slučaju, bez smanjenja općenitosti možemo pretpostaviti da je za svaki $\omega \in \Omega$ i svaki $n \in \mathbb{N}$, $0 \leq Y_n(\omega) \leq Y_{n+1}(\omega)$. Neka su $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$ i $\hat{Y}_n = \mathbb{E}[Y_n|\mathcal{G}]$, $n \in \mathbb{N}$, odgovarajuće verzije uvjetnih očekivanja. Prema lemi 2.9 vrijedi da je za svaki $n \in \mathbb{N}$, $0 \leq \hat{Y}_n \leq \hat{Y}_{n+1}$ g.s. Neka je $Z(\omega) := \overline{\lim}_n \hat{Y}_n(\omega)$, $\omega \in \Omega$. Z je $\mathcal{G}$-izmjeriva varijabla (možda takva da je $\mathbb{P}(Z = +\infty) > 0$) i vrijedi da je $Z = \lim_n \hat{Y}_n$ g.s. Tvrdnja (v) će biti dokazana ako pokažemo da je $Z = \hat{Y}$ g.s. Budući da je Z nenegativna varijabla, dovoljno je dokazati da Z zadovoljava definicijski uvjet (3) za $\mathbb{E}[Y|\mathcal{G}]$. U tom slučaju će nužno slijediti da je $\mathbb{E}Z = \mathbb{E}Y < +\infty$, tj. $Z \in \mathcal{L}^1$ (pa će nužno biti i $\mathbb{P}(Z < +\infty) = 1$). Neka je $G \in \mathcal{G}$ proizvoljan. Budući da je za svaki $n \in \mathbb{N}$, $0 \leq \hat{Y}_n \cdot \mathbb{1}_G \leq \hat{Y}_{n+1}$ g.s. i $Z \cdot \mathbb{1}_G = \lim_n \hat{Y}_n \cdot \mathbb{1}_G$ g.s., prema LTMK (teorem 1.59 (i)) je nužno

$$\mathbb{E}[Z \cdot \mathbb{1}_G] = \lim_n \mathbb{E}[\hat{Y}_n \cdot \mathbb{1}_G] \overset{(3)}{=} \lim_n \mathbb{E}[Y_n \cdot \mathbb{1}_G] = \mathbb{E}[Y \cdot \mathbb{1}_G].$$

Naime, i prva i zadnja jednakost su posljedice LTMK, pri čemu je u drugom slučaju LTMK primijenjen na rastući nenegativni niz $(Y_n \cdot \mathbb{1}_G; n \in \mathbb{N})$ čiji limes je $Y \cdot \mathbb{1}_G = \lim_n Y_n \cdot \mathbb{1}_G$, a druga jednakost je posljedica svojstva (3) iz definicije 2.11 primijenjenog na $\mathbb{E}[Y_n|\mathcal{G}]$, $n \in \mathbb{N}$. Dakle, $\lim_n \hat{Y}_n = Z = \mathbb{E}[Y|\mathcal{G}]$ g.s. pa vrijedi tvrdnja (v).

Uvjetna Fatouova lema (vi) i teorem o uvjetnoj dominiranoj konveregenciji (vii) sada se dokazuju pomoću teorema o uvjetnoj monotonoj konvergenciji (v) na način kako se Fatouova lema i teorem o dominiranoj kovergenciji dokazuju pomoću teorema o monotonoj kovergenciji, dakako za bezuvjetno matematičko očekivanje (vidjeti, na primjer, dokaze teorema 10.3 i 10.6 u [9], §10.2, str. 295.-297.).

Slijedi dokaz uvjetne Jensenove nejednakosti $(viii)$. Neka je $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$ jedna verzija, $\phi : \mathbb{R} \to \mathbb{R}$ konveksna funkcija i $\phi(Y) \in \mathcal{L}^1$. Budući da je za sve n, $\phi(Y) \geq a_n Y + b_n$ za neke nizove realnih brojeva $(a_n, n \in \mathbb{N})$ i $(b_n, n \in \mathbb{N})$ prema napomeni 1.57, zbog monotonosti (iv) i linearnosti (iii) uvjetnog očekivanja je $\mathbb{E}[\phi(Y)|\mathcal{G}] \geq a_n \hat{Y} + b_n$ g.s. za sve n. Prema istoj napomeni, $\mathbb{E}[\phi(Y)|\mathcal{G}] \geq \sup_n (a_n \hat{Y} + b_n) = \phi(\hat{Y}) = \phi(\mathbb{E}[Y|\mathcal{G}])$ g.s.

Dokaz teleskopskog pravila (ix). Neka su $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$ i $\tilde{Y} = \mathbb{E}[Y|\mathcal{H}]$. Budući da je $\tilde{Y} \in \mathcal{L}^1(\Omega, \mathcal{H}, \mathbb{P})$, za dokazati da je $\tilde{Y}$ verzija od $\mathbb{E}[\hat{Y}|\mathcal{H}]$ dovoljno je pokazati da vrijedi svojstvo (3) iz definicije 2.11 za $\mathbb{E}[\hat{Y}|\mathcal{H}]$. Neka je $H \in \mathcal{H}$ proizvoljan. Tada je

$$\mathbb{E}[\tilde{Y} \cdot \mathbb{1}_H] \overset{(3)}{=} \mathbb{E}[Y \cdot \mathbb{1}_H] \overset{(3)}{=} \mathbb{E}[\hat{Y} \cdot \mathbb{1}_H],$$

pri čemu se u prvoj jednakosti pozivamo na definicijsko svojstvo (3) za $\mathbb{E}[Y|\mathcal{H}]$, a u drugoj na isto svojstvo, ali za $\mathbb{E}[Y|\mathcal{G}]$ jer je $H \in \mathcal{H} \subset \mathcal{G}$. Zbog proizvoljnosti od $H \in \mathcal{H}$ slijedi tvrdnja (3) za $\mathbb{E}[\hat{Y}|\mathcal{H}]$. Dakle, $\tilde{Y} = \mathbb{E}[\hat{Y}|\mathcal{H}]$.

Dokažimo (x). Neka je $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$. Prvo, neka je $Z = \mathbb{1}_A$ za $A \in \mathcal{G}$. Tada je trivijalno $\mathbb{1}_A \cdot Y \in \mathcal{L}^1$ i za svaki $G \in \mathcal{G}$, budući da je tada i $A \cap G \in \mathcal{G}$, primjenom definicijskog svojstva (3) od $\mathbb{E}[Y|\mathcal{G}]$,

$$\mathbb{E}[\mathbb{1}_A\hat{Y}\cdot\mathbb{1}_G]=\mathbb{E}[\hat{Y}\cdot\mathbb{1}_{A\cap G}]\stackrel{(3)}{=}\mathbb{E}[Y\cdot\mathbb{1}_{A\cap G}]=\mathbb{E}[\mathbb{1}_AY\cdot\mathbb{1}_G].$$

Budući da je trivijalno $\mathbb{1}_A\hat{Y}\in\mathcal{L}^1(\Omega,\mathcal{G},\mathbb{P})$, vrijedi da je $\mathbb{1}_A\hat{Y}=\mathbb{E}[\mathbb{1}_AY|\mathcal{G}]$. Dakle, za svaki $A\in\mathcal{G}$,

$$\mathbb{E}[\mathbb{1}_AY|\mathcal{G}]=\mathbb{1}_A\mathbb{E}[Y|\mathcal{G}] \text{ g.s.} \tag{2.21}$$

Nadalje, neka je sada $Z=\sum_{i=1}^k\alpha_i\mathbb{1}_{A_i}$ jednostavna $\mathcal{G}$-izmjeriva varijabla. Dakle, A_1, A_2,..., $A_k\in\mathcal{G}$. Tada je zbog linearnosti $(\sum_{i=1}^k\alpha_i\mathbb{1}_{A_i})Y\in\mathcal{L}^1$, te linearnosti uvjetnog očekivanja (iii) i prethodnog slučaja

$$\mathbb{E}[\sum_{i=1}^k\alpha_i\mathbb{1}_{A_i}Y|\mathcal{G}]\stackrel{(iii)}{=}\sum_{i=1}^k\alpha_i\mathbb{E}[\mathbb{1}_{A_i}Y|\mathcal{G}]\stackrel{(2.21)}{=}\sum_{i=1}^k\alpha_i\mathbb{1}_{A_i}\mathbb{E}[Y|\mathcal{G}]=(\sum_{i=1}^k\alpha_i\mathbb{1}_{A_i})\mathbb{E}[Y|\mathcal{G}] \text{ g.s.} \tag{2.22}$$

Neka je Z nenegativna $\mathcal{G}$-izmjeriva slučajna varijabla takva da je $ZY\in\mathcal{L}^1$. Budući da je $Z\geq 0$ i $\mathcal{G}$-izmjeriva, postoji rastući niz jednostavnih nenegativnih $\mathcal{G}$-izmjerivih varijabli $(Z_n, n\in\mathbb{N})$ takvih da je $\lim_n Z_n=Z$ (vidjeti teorem 1.3). Pretpostavimo da je i Y nenegativna slučajna varijabla. Budući da je prema prethodnom za svaki n, $Z_nY\in\mathcal{L}^1$, te da je $(Z_nY, n\in\mathbb{N})$ rastući niz nenegativnih slučajnih varijabli, prema teoremu o uvjetnoj monotonoj konvergenciji (v) (kraće, TUMK),

$$\mathbb{E}[ZY|\mathcal{G}]\stackrel{(\text{TUMK})}{=}\lim_n\mathbb{E}[Z_nY|\mathcal{G}]\stackrel{(2.22)}{=}\lim_n Z_n\mathbb{E}[Y|\mathcal{G}]=Z\mathbb{E}[Y|\mathcal{G}] \text{ g.s.} \tag{2.23}$$

Konačno, neka je Z proizvoljna $\mathcal{G}$-izmjeriva varijabla takva da je $ZY\in\mathcal{L}^1$. Prikažimo Z i Y pomoću njihovih pozitivnih i negativnih dijelova: $Z=Z^+-Z^-$, $Y=Y^+-Z^-$. Tada su i Z^-Y^-, Z^-Y^+, Z^+Y^-, $Z^+Y^+\in\mathcal{L}^1$ i primjenom linearnosti uvjetnog očekivanja (iii), prema prethodnom je

$$\begin{aligned}\mathbb{E}[ZY|\mathcal{G}] &\stackrel{(iii)}{=} \mathbb{E}[Z^+Y^+|\mathcal{G}]+\mathbb{E}[Z^-Y^-|\mathcal{G}]-\mathbb{E}[Z^+Y^-|\mathcal{G}]-\mathbb{E}[Z^-Y^+|\mathcal{G}]\stackrel{(2.23)}{=}\\ &= Z^+\mathbb{E}[Y^+|\mathcal{G}]+Z^-\mathbb{E}[Y^-|\mathcal{G}]-Z^+\mathbb{E}[Y^-|\mathcal{G}]-Z^-\mathbb{E}[Y^+|\mathcal{G}]=\\ &= Z\mathbb{E}[Y|\mathcal{G}] \text{ g.s.}\end{aligned}$$

Na kraju, dokažimo (xi). Neka su Y i $\mathcal{G}$ nezavisni. Stavimo da je $\hat{Y}=\mathbb{E}Y$. Budući da je $\hat{Y}$ konstanta, trivijalno je u $\mathcal{L}^1(\Omega,\mathcal{G},\mathbb{P})$ pa za $\hat{Y}$ vrijede definicijske tvrdnje $(1-2)$ za $\mathbb{E}[Y|\mathcal{G}]$. Pokažimo da vrijedi i tvrdnja (3). Neka je $G\in\mathcal{G}$ proizvoljni događaj. Tada je zbog nezavisnosti slučajnih varijabli $\mathbb{1}_G$ i Y i linearnosti matematičkog očekivanja (teorem 1.56 (i)),

$$\mathbb{E}[Y\cdot\mathbb{1}_G]\stackrel{(\text{nez.})}{=}\mathbb{E}Y\cdot\mathbb{E}\mathbb{1}_G\stackrel{(\text{lin.m.o.})}{=}\mathbb{E}[(\mathbb{E}Y)\cdot\mathbb{1}_G]=\mathbb{E}(\hat{Y}\cdot\mathbb{1}_G).$$

Zbog proizvoljnosti od $G\in\mathcal{G}$, gornja jednakost vrijedi za svaki $G\in\mathcal{G}$, dakle, vrijedi definicijska tvrdnja (3) za $\mathbb{E}[Y|\mathcal{G}]$. □

Korolar 2.27. *Neka su $Y\in\mathcal{L}^1$ i $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Tada vrijedi:*

(i) $|\mathbb{E}[Y|\mathcal{G}]|\leq\mathbb{E}[|Y|\,|\mathcal{G}]$ *g.s.*
(ii) $\|\mathbb{E}[Y|\mathcal{G}]\|_1\leq\|Y\|_1$.

Dokaz. Tvrdnja (i) slijedi iz monotonosti uvjetnog očekivanja. Naime, budući da vrijedi $-|Y|\leq Y\leq|Y|$, primjenom teorema 2.26 (iv) i (iii) slijedi da je $-\mathbb{E}[|Y|\,|\mathcal{G}]\leq\mathbb{E}[Y|\mathcal{G}]\leq\mathbb{E}[|Y|\,|\mathcal{G}]$ g.s. što je ekvivalentno sa svojstvom (i) iz korolara. Ako na tvrdnju (i) primijenimo svojstvo monotonosti matematičkog očekivanja (teorem 1.56 (ii)) te teorem 2.26 (i), odmah slijedi i tvrdnja (ii) korolara. $\square$

Nejednakost (i) iz korolara 2.27 mogla se dokazati i direktno primjenom Jensenove uvjetne nejednakosti (teorem 2.26 $(viii)$) budući da je apsolutna vrijednost $|\cdot| : \mathbb{R} \to \mathbb{R}$ konveksno preslikavanje.

Interpretirajmo nejednakost (ii). Ako za fiksnu σ-podalgebru $\mathcal{G}$ od $\mathcal{F}$ uvjetno matematičko očekivanje uvjetno na $\mathcal{G}$ interpretiramo kao preslikavanje u L^1 koje klasi (obzirom na relaciju "= g.s.") reprezentiranoj sa $Y \in \mathcal{L}^1$, pridružuje klasu reprezentiranu verzijom od $\mathbb{E}[Y|\mathcal{G}]$, tada tvrdnja (ii) iz korolara 2.27 kaže da je to preslikavanje *kontrakcija* u L^1. Primijetite da je to preslikavanje dobro definirano (vidjeti početak dokaza monotonosti uvjetnog očekivanja).

Primjer 2.28. Poopćimo primjer 2.17. Neka su $X_1, X_2, \ldots, X_n$ nezavisne jednako distribuirane (kraće n.j.d.) slučajne varijable sa zakonom razdiobe $\mathbb{P}_X$ i s konačnim matematičkim očekivanjem. Za $T := X_1 + X_2 + \cdots + X_n$, odredimo $\mathbb{E}[X_i|T]$. Prvo, primijetimo da je za $\hat{X}_i := \mathbb{E}[X_i|T]$ $(0 \le i \le n)$ i proizvoljni $B \in \mathcal{B}(\mathbb{R})$, upotrebom svojstva (3) uvjetnog očekivanja i teorema o zamjeni varijabli, a zbog nezavisnosti, jednake distribuiranosti i simetrije,

$$\begin{aligned}
\mathbb{E}[\hat{X}_i \cdot \mathbb{1}_{\{T\in B\}}] &\overset{(3)}{=} \mathbb{E}[X_i \cdot \mathbb{1}_{\{T\in B\}}] = \int\limits_{T^{-1}(B)} X_i \, d\mathbb{P} \overset{\text{(tm.z.v. i n.j.d.)}}{=} \\
= \int\limits_{x_1+\cdots+x_n\in B} x_i \, d\mathbb{P}_X(x_1)\cdots d\mathbb{P}_X(x_n) &\overset{\text{(simet.)}}{=} \int\limits_{x_1+\cdots+x_n\in B} x_1 \, d\mathbb{P}_X(x_1)\cdots d\mathbb{P}_X(x_n) = \\
= \mathbb{E}[X_1 \cdot \mathbb{1}_{\{T\in B\}}] &\overset{(3)}{=} \mathbb{E}[\hat{X}_1 \cdot \mathbb{1}_{\{T\in B\}}].
\end{aligned}$$

Dakle, zbog proizvoljnosti od $B \in \mathcal{B}(\mathbb{R})$, prema teoremu 1.58 je $\hat{X}_1 = \hat{X}_2 = \cdots = \hat{X}_n$ g.s. Sada, prvo primjenom svojstva (ii), a zatim i svojstva (iii) (linearnosti) iz teorema 2.26, slijedi:

$$T \overset{(ii)}{=} \mathbb{E}[T|T] = \mathbb{E}[\sum_{i=1}^n X_i|T] \overset{(iii)}{=} \sum_{i=1}^n \mathbb{E}[X_i|T] = \sum_{i=1}^n \hat{X}_i = n\hat{X}_1 \text{ g.s.}$$

Dakle,

$$\mathbb{E}[X_1|T] = \mathbb{E}[X_2|T] = \cdots = \mathbb{E}[X_n|T] = \frac{T}{n} = \overline{X}_n \text{ g.s.}$$

Drugim riječima, najbolja procjena srednjeg ishoda svakog pojedinog opažanja (mjerenja) u n.j.d. pokusima ako je poznat sumarni ishod svih opažanja, jednaka je aritmetičkoj sredini ishoda svih opažanja. $\square$

Primjer 2.29. Neka su k-dimenzionalna slučajna veličina X i ℓ-dimenzionalna slučajna veličina Y $(k, \ell \ge 1)$ definirane na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ pri čemu je X diskretna slučajna veličina ili je (X, Y) apsolutno neprekidan slučajni vektor, kao u iskazu teorema 2.16. Pretpostavimo da su poznati uvjetni zakoni razdiobe od Y uz dano $X = x$ za $\mathbb{P}_X$-g.s. $x \in \mathbb{R}^k$. Označimo sa N skup svih $x \in \mathbb{R}^k$ za koje uvjetna razdioba od Y uz dano $X = x$ nije definirana, dakle $N = \mathbb{R}^k \setminus \text{supp} f_X$ pa je $\mathbb{P}_X(N) = 0$. Tada za svaki $y \in \mathbb{R}^\ell$, primjenom teorema 2.26 (i) i teorema 2.16, vrijedi

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \le y) = \mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)] \overset{\text{tm. } 2.26\,(i)}{=} \mathbb{E}(\mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)|X]) \overset{\text{tm. } 2.16}{=} \\
&= \int\limits_{\mathbb{R}^k \cap N^c} \mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)|X = x]\, d\mathbb{P}_X(x) = \int\limits_{\mathbb{R}^k \cap N^c} \mathbb{P}(Y \le y | X = x)\, d\mathbb{P}_X(x) = \\
&= \int\limits_{\mathbb{R}^k \cap N^c} F_{Y|X}(y|x)\, d\mathbb{P}_X(x). \qquad (2.24)
\end{aligned}$$

Dobivena formula je poopćena verzija *formule potpune vjerojatnosti* (vidjeti [9] teorem 4.1, str. 63.). Na sličan način, za svaki $(x, y) \in \mathbb{R}^{k+\ell}$ je

$$\begin{aligned}
F_{X,Y}(x, y) &= \mathbb{P}(X \le x, Y \le y) = \mathbb{E}[\mathbb{1}_{\langle -\infty, x] \times \langle -\infty, y]}(X, Y)] \overset{\text{tm. } 2.26\,(i,x)}{=} \\
&= \mathbb{E}(\mathbb{1}_{\langle -\infty, x]}(X)\mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)|X]) \overset{\text{tm. } 2.16}{=} \\
&= \int\limits_{\langle -\infty, x] \cap N^c} \mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)|X = s]\, d\mathbb{P}_X(s) = \int\limits_{\langle -\infty, x] \cap N^c} \mathbb{P}(Y \le y | X = s)\, d\mathbb{P}_X(s) = \\
&= \int\limits_{\langle -\infty, x] \cap N^c} F_{Y|X}(y|s)\, d\mathbb{P}_X(s). \qquad (2.25)
\end{aligned}$$

Primijetite da uz zadane elemente, vjerojatnosne funkcije razdiobe F_X (ekvivalentno $\mathbb{P}_X$) i $F_{Y|X}(\cdot|x)$ za $x \in \mathbb{R}^k \cap N^c$ (uz $\mathbb{P}_X(N) = 0$), su formulama (2.24) i (2.25) zadane vjerojatnosne funkcije razdiobe (neke) slučajne veličine Y dimenzije ℓ, odnosno (nekog) slučajnog vektora (X, Y) (u smislu da je moguće konstruirati vjerojatnosni prostor i na njemu slučajnu veličinu Y, odnosno slučajni vektor (X, Y), tako da su formulama (2.24), odnosno (2.25), dane njihove funkcije razdioba).
(a) *Ako je za $\mathbb{P}_X$-g.s. $x \in \mathbb{R}^k$ uvjetna razdioba od Y uz dano $X = x$ neprekidna s gustoćom $f_{Y|X}(\cdot|x)$, tada je i Y neprekidna slučajna veličina (neovisno o tome je li X neprekidna ili diskretna slučajna veličina).* Dokažimo to i odredimo gustoću od Y. Za proizvoljno $y \in \mathbb{R}^\ell$, primjenom jednakosti (2.24) i Fubinijevog teorema A.1 slijedi

$$\begin{aligned}
F_Y(y) &\overset{(2.24)}{=} \int\limits_{\mathbb{R}^k} F_{Y|X}(y|x)\, d\mathbb{P}_X(x) = \int\limits_{\mathbb{R}^k} (\int\limits_{\langle -\infty, y]} f_{Y|X}(t|x)\, dt)\, d\mathbb{P}_X(x) \overset{\text{(Fubinijev tm.)}}{=} \\
&= \int\limits_{\langle -\infty, y]} (\int\limits_{\mathbb{R}^k} f_{Y|X}(t|x)\, d\mathbb{P}_X(x))\, dt. \qquad (2.26)
\end{aligned}$$

Dakle, prema definiciji 1.9, Y je neprekidna slučajna veličina dimenzije ℓ s gustoćom

$$f_Y(y) = \int\limits_{\mathbb{R}^k} f_{Y|X}(y|x)\, d\mathbb{P}_X(x). \qquad (2.27)$$

(b) Pretpostavimo da je X diskretna slučajna varijabla sa zakonom razdiobe $w_k = \mathbb{P}(X = a_k)$ za $k = 1, 2, \dots, m$. Nadalje, označimo kraće sa $F_k \equiv F_{Y|X}(\cdot|a_k)$, $k = 1, 2, \dots, m$, niz vjerojatnosnih funkcija razdioba. Tada iz formule (2.24) slijedi da je funkcija razdiobe F_Y od Y jednaka

$$F_Y = \sum_{k=1}^{m} w_k F_k,$$

drugim riječima, funkcija razdiobe od Y je konveksna kombinacija vjerojatnosnih funkcija razdioba. Zakon razdiobe čija se funkcija razdiobe može prikazati kao konveksna kombinacija vjerojatnosnih funkcija razdioba zovemo *mješavinom razdioba* ili *mješavinom distribucija* (engl. *mixture of distributions*). Dakle, prema tvrdnji (a), mješavina neprekidnih razdioba je neprekidna razdioba.
(c) Pretpostavimo sada da je X neprekidna slučajna veličina s gustoćom $w = f_X$. Nadalje, kao i u (b), označimo sa $F(\cdot|x) \equiv F_{Y|X}(\cdot|x)$, $x \in \mathbb{R}^k$, množinu vjerojatnosnih funkcija razdioba. Tada je primjenom formule (2.24) funkcija razdiobe od Y jednaka

$$F_Y = \int\limits_{\mathbb{R}^k} F(\cdot|x)w(x)\,dx.$$

Zakon razdiobe čija se funkcija razdiobe može dobiti kao težinski integral vjerojatnosnih funkcija razdioba (s težinama ukupne mase 1), zovemo *složena razdioba* ili *složena distribucija* (engl. *compound distribution*). Dakle, prema tvrdnji (a), složena razdioba sastavljena od neprekidnih razdioba je neprekidna razdioba.
(d) *Ako su za $\mathbb{P}_X$-g.s. $x \in \mathbb{R}^k$ uvjetna razdioba od Y uz dano $X = x$ i marginalna razdioba od X neprekidne s gustoćama, redom $f_{Y|X}(\cdot|x)$ i f_X, tada je i (X, Y) neprekidan slučajan vektor s gustoćom*

$$f(x,y) = \begin{cases} f_{Y|X}(y|x)f_X(x), & x \in \mathbb{R}^k \cap N^c, \quad y \in \mathbb{R}^\ell \\ 0, & x \in N, \quad\quad\quad\; y \in \mathbb{R}^\ell, \end{cases} \tag{2.28}$$

gdje je N $\mathbb{P}_X$-zanemariv skup iz (2.25). Dokažimo i to. Za proizvoljne $(x, y) \in \mathbb{R}^{k+\ell}$, primjenom jednakosti (2.25) i Fubinijevog teorema A.1 slijedi

$$\begin{aligned} F_{X,Y}(x,y) &\overset{(2.25)}{=} \int\limits_{\langle -\infty,x]\cap N^c} F_{Y|X}(y|s)\,d\mathbb{P}_X(s) = \int\limits_{\langle -\infty,x]\cap N^c} \Big(\int\limits_{\langle -\infty,y]} f_{Y|X}(t|s)\,dt\Big) f_X(s)\,ds \overset{\text{(Fubini. tm.)}}{=} \\ &\overset{(2.28)}{=} \int\limits_{\langle -\infty,x]\times\langle -\infty,y]} f(s,t)\,ds\,dt. \end{aligned} \tag{2.29}$$

Dakle, prema definiciji 1.9, (X, Y) je neprekidan slučajan vektor s gustoćom (2.28). □

Napomena 2.30. Svojstva uvjetnog očekivanja iskazana teoremom 2.26 i teorem 2.14 omogućavaju nam poopćenje definicije uvjetne distribucije $\mathbb{P}_{Y|X}(\cdot|y)$ na slučaj kada se slučajni vektor (X, Y) dimenzije $k + \ell$ sastoji i od *neprekidne* slučajne veličine X dimenzije $k \geq 1$ pri čemu Y može biti neprekidna ili diskretna slučajna veličina dimenzije $\ell \geq 1$. Bitno je jedino da su X i Y definirane na istome vjerojatnosnom prostoru.
Neka je $y \in \mathbb{R}^\ell$ proizvoljan. Tada sličnom argumentacijom kao u primjeru 2.29 slijedi:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{E}[\mathbb{1}_{\langle -\infty,y]}(Y)] \overset{\text{tm. } 2.26\,(i)}{=} \mathbb{E}(\mathbb{E}[\mathbb{1}_{\langle -\infty,y]}(Y)|X]) \overset{\text{tm. } 2.14}{=} \mathbb{E}g_y(X)$$

gdje je $g_y : \mathbb{R}^k \to \mathbb{R}$ Borelova funkcija. Naime, budući da je svaka verzija od $\mathbb{E}[\mathbb{1}_{\langle -\infty,y]}(Y)|X]$ $\sigma(X)$-izmjeriva slučajna varijabla, prema teoremu 2.14 postoji Borelova funkcija g_y takva da je $g_y(X) = \mathbb{E}[\mathbb{1}_{\langle -\infty,y]}(Y)|X]$. Kako je, prema teoremu o zamjeni varijabli (teorem 10.11 u [9], str. 306.),

$$F_Y(y) = \mathbb{E}(\mathbb{E}[\mathbb{1}_{\langle -\infty,y]}(Y)|X]) = \mathbb{E}g_y(X) \overset{\text{tm. o zamj. v.}}{=} \int\limits_{\mathbb{R}^k} g_y(x)f_X(x)\,dx,$$

izraz (2.24) iz primjera 2.29 upućuje da je razumno definirati vrijednost $F_{Y|X}(y|x)$ uvjetne funkcije distribucije od Y uz dano $X = x$ u točki y pomoću $g_y(x)$ za sve $x \in \mathbb{R}^k$ za koji je to moguće učiniti. Tu vrijednost ćemo definirati za $x \in \mathrm{supp} f_X$ gdje je f_X gustoća od X. Uzmimo sada da je $y = r = (q_1, q_2, \ldots, q_\ell) \in \mathbb{Q}^\ell$ točka iz $\mathbb{R}^\ell$ kojoj su sve komponente racionalni brojevi i odaberimo pripadnu Borelovu funkciju q_r takvu da je $g_r(X)$ jedna verzija od $\mathbb{E}[\mathbb{1}_{\langle -\infty, y]}(Y)|X]$. Takvih točaka, a onda i odabranih funkcija, ima prebrojivo mnogo. Promotrimo sljedeće događaje iz σ-podalgebre $\sigma(X)$ opisane pomoću funkcija g_r za $r \in \mathbb{Q}^\ell$:

$$\begin{aligned}
A_0 &:= \bigcap_{r\in\mathbb{Q}^\ell} \{0 \le g_r(X) \le 1\} \\
A_1 &:= \bigcap_{r\in\mathbb{Q}^\ell} \{\lim_{\mathbb{Q}^\ell \ni s \to r+} g_s(X) = g_r(X)\} \\
A_2 &:= \bigcap_{r_1<r_2} \{\Delta_{r_2-r_1} g_{r_1}(X) \ge 0\} \\
A_3 &:= \bigcap_{j=1}^{\ell} \{\lim_{q_j\to-\infty} g_{(q_1,q_2,\ldots,q_\ell)}(X) = 0\} \cap \{\lim_{\forall j,\, q_j\to+\infty} g_{(q_1,q_2,\ldots,q_\ell)}(X) = 1\} \\
A &:= A_0 \cap A_1 \cap A_2 \cap A_3.
\end{aligned}$$

Sada ćemo neka od svojstva $(i - xi)$ uvjetnog očekivanja iz teorema 2.26 iskoristiti da dokažemo da su navedeni događaji vjerojatnosti jedan.
Budući da za indikatorsku slučajnu varijablu vrijedi da je $0 \le \mathbb{1}_{\langle -\infty, r]}(Y) \le 1$, iz svojstva (iv) monotonosti uvjetnog očekivanja je $0 \le g_r(X) \le 1$ g.s. za sve $r \in \mathbb{Q}^\ell$. Budući da je A_0 (prebrojiv) presjek događaja $\{0 \le q_r(X) \le 1\} \equiv \{X \in g_r^{-1}(\langle 0, 1\rangle)\}$ vjerojatnosti jedan, A_0 je događaj vjerojatnosti jedan i vrijedi da je

$$A_0 = \{X \in \bigcap_{r\in\mathbb{Q}^\ell} g_r^{-1}(\langle 0,1\rangle)\}.$$

Nadalje, neka su $r, s \in \mathbb{Q}^\ell$ takvi da je $r \le s$ gdje je '$\le$' relacija parcijalnog uređaja u $\mathbb{R}^\ell$ definirana kao na početku poglavlja 1. Tada je

$$(\forall s, r \in \mathbb{Q}^\ell)\ r \le s \ \Rightarrow\ 0 \le \mathbb{1}_{\langle -\infty, r]}(Y) \le \mathbb{1}_{\langle -\infty, s]}(Y) \le 1 \ \text{ i } \ \lim_{s\to r+} \mathbb{1}_{\langle -\infty, s]}(Y) = \mathbb{1}_{\langle -\infty, r]}(Y).$$

Prema teoremu o uvjetnoj dominiranoj konvergenciji (vii) slijedi da je nužno

$$\lim_{s\to r+} g_s(X) = g_r(X) \text{ g.s.},$$

dakle, događaj $\{\lim_{s\to r+} g_s(X) = g_r(X)\} \equiv \{X \in \{\lim_{s\to r+} g_s = g_r\}\}$ je vjerojatnosti jedan za sve $r \in \mathbb{Q}^\ell$. Budući da je A_1 prebrojiv presjek događaja vjerojatnosti jedan, i sam je događaj vjerojatnosti jedan i vrijedi da je

$$A_1 = \{X \in \bigcap_{r\in\mathbb{Q}^\ell} \{\lim_{s\to r+} g_s = g_r\}\}.$$

Neka su sada $r_1, r_2 \in \mathbb{Q}^\ell$ takvi da je $r_1 < r_2$ u smislu da odgovarajuće komponente tih vektora u toj relaciji (kao što je navedeno neposredno prije teorema 1.8). Koristeći definiciju operatora Δ (vidjeti također neposredno prije teorema 1.8), vrijedi da je

$$\Delta_{r_2-r_1} \mathbb{1}_{\langle -\infty, r_1]}(Y) = \mathbb{1}_{\langle r_1, r_2]}(Y) \ge 0.$$

Odavde slijedi korištenjem svojstava (iii) linearnosti i (iv) monotonosti uvjetnog očekivanja, da je

$$\Delta_{r_2-r_1} g_{r_1}(X) \overset{(iii)}{=} \mathbb{E}[\Delta_{r_2-r_1} \mathbb{1}_{\langle -\infty, r_1]}(Y)|X] = \mathbb{E}[\mathbb{1}_{\langle r_1, r_2]}(Y)|X] \overset{(iv)}{\geq} 0 \text{ g.s.}$$

Dakle, događaj $\{\Delta_{r_2-r_1} g_{r_1}(X) \geq 0\} \equiv \{X \in \{\Delta_{r_2-r_1} g_{r_1} \geq 0\}\}$ je vjerojatnosti jedan, pa je i događaj A_2 vjerojatnosti jedan kao prebrojiv produkt tih događaja i vrijedi:

$$A_2 = \{X \in \bigcap_{r_1 < r_2} \{\Delta_{r_2-r_1} g_{r_1} \geq 0\}\}.$$

Na sličan način, korištenjem činjenica da za bilo koju j-tu koordinatu q_j racionalne točke $r = (q_1, q_2, \ldots, q_\ell) \in \mathbb{Q}^\ell$ (za $j \in \{1, 2, \ldots, \ell\}$) vrijedi:

$$\lim_{q_j \to -\infty} \mathbb{1}_{\langle -\infty, r]}(Y) = 0$$

i da su indikatorske funkcije omeđene, preme teoremu o uvjetnoj dominiranoj konvergenciji $(viii)$ slijedi da je

$$\lim_{q_j \to -\infty} g_r(X) = \lim_{q_j \to -\infty} \mathbb{E}[\mathbb{1}_{\langle -\infty, r]}(Y)|X] = \mathbb{E}[0|X] = 0 \text{ g.s.}$$

Dakle, za sve j događaji $\{\lim_{q_j \to -\infty} g_{(q_1, q_2, \ldots, q_\ell)}(X) = 0\}$ su vjerojatnosti jedan. Analogno, budući da je

$$\lim_{\forall j,\, q_j \to +\infty} \mathbb{1}_{\langle -\infty, r]}(Y) = 1,$$

po istom teremu o uvjetnoj dominiranoj konvergenciji slijedi da je

$$\lim_{\forall j,\, q_j \to +\infty} g_r(X) = 1 \text{ g.s.}$$

Dakle, i događaj $\{\lim_{\forall j,\, q_j \to +\infty} g_{(q_1, q_2, \ldots, q_\ell)}(X) = 1\}$ je vjerojatnosti jedan. Prema tome je događaj A_3 kao konačan presjek događaja vjerojatnosti jedan, vjerojatnosti jedan i vrijedi:

$$A_3 = \{X \in \bigcap_{j=1}^{\ell} \{\lim_{q_j \to -\infty} g_{(q_1, q_2, \ldots, q_\ell)} = 0\} \cap \{\lim_{\forall j,\, q_j \to +\infty} g_{(q_1, q_2, \ldots, q_\ell)} = 1\}\}.$$

Na kraju, događaj A je vjerojatnosti jedan kao konačan presjek takvih i $A = \{X \in B\}$ za Borelov skup

$$\begin{aligned} B \;&=\; \bigcap_{r \in \mathbb{Q}^\ell} g_r^{-1}(\langle 0, 1\rangle) \cap \bigcap_{r \in \mathbb{Q}^\ell} \{\lim_{s \to r+} g_s = g_r\} \cap \bigcap_{r_1 < r_2} \{\Delta_{r_2-r_1} g_{r_1} \geq 0\} \cap \\ &\cap \bigcap_{j=1}^{\ell} \{\lim_{q_j \to -\infty} g_{(q_1, q_2, \ldots, q_\ell)} = 0\} \cap \{\lim_{\forall j,\, q_j \to +\infty} g_{(q_1, q_2, \ldots, q_\ell)} = 1\}. \end{aligned}$$

Neka je sada $N := B^c \cap \operatorname{supp} f_X$. Tada vrijedi da je N Borelov skup i da je Lebesguove mjere nula:

$$\begin{aligned} \lambda^k(N) \;&=\; \int_N dx = \int_{B^c \cap \operatorname{supp} f_X} \frac{1}{f_X(x)} f_X(x)\, dx) = \int_{B^c} \mathbb{1}_{\operatorname{supp} f_X}(x) \frac{1}{f_X(x)} d\mathbb{P}_X(x) \overset{\text{tm.o.zamj.v.}}{=} \\ &= \int_{A^c} \mathbb{1}_{\{f_X(X) > 0\}} \frac{1}{f_X(X)}\, d\mathbb{P} = \mathbb{E}[\mathbb{1}_{A^c} \frac{1}{f_X(X)}] \overset{\mathbb{P}(A^c)=0}{=} 0. \end{aligned}$$

Za $x \in \text{supp} f_X$ i $y \in \mathbb{R}^\ell$ definirajmo:

$$F_{Y|X}(y|x) := \begin{cases} \lim_{r\to y+} g_r(x), & x \in N^c \\ G(y), & x \in N \end{cases} \tag{2.30}$$

gdje je G bilo koja vjerojatnosna funkcija distribucije definirana na $\mathbb{R}^\ell$. Nije teško pokazati da je funkcija $F_{Y|X}(\cdot|x) : \mathbb{R}^\ell \to \mathbb{R}$ vjerojatnosna funkcija distribucije za svaki $x \in \text{supp} f_X$ (vidjeti sljedeći zadatak). Prema tome, $F_{Y|X}(\cdot|x)$ definira zakon razdiobe $\mathbb{P}_{Y|X}(\cdot|x)$ kojeg zovemo *uvjetni zakon razdiobe od Y uz dano $X = x$*. Oznaka:

$$\mathbb{P}(Y \in B|X = x) \equiv \mathbb{P}_{Y|X}(B|x), \quad B \in \mathcal{B}(\mathbb{R}^\ell). \tag{2.31}$$

Uzimajući u obzir definiciju matematičkog očekivanja u odnosu na uvjetnu distribuciju $\mathbb{P}_{Y|X}(\cdot|x)$ danog relacijom (2.18), tada je definicijom 2.15 definirana regresijska funkcija i u odnosu na upravo dobivenu uvjetnu distribuciju. □

Zadatak 2.31. (*a*) Dokažite da je funkcija $F_{Y|X}(\cdot|x) : \mathbb{R}^\ell \to \mathbb{R}$ iz napomene 2.30 dobro definirana sa (2.30) za $x \in N^c$ i da je vjerojatnosna funkcija distribucije.
(*b*) Dokažite da teorem 2.16 vrijedi za slučajne vektore (X, Y) takve da je X neprekidna slučajna veličina.

Zadatak 2.32. Za σ-algebre $\mathcal{F}$ i $\mathcal{G}$ definiramo sa $\mathcal{F} \vee \mathcal{G} := \sigma(\mathcal{F} \cup \mathcal{G})$ najmanju σ-algebru koja sadrži obje σ-algebre. Neka su $\mathcal{H}$ i $\mathcal{G}$ σ-podalgebre od $\mathcal{F}$ i $Y \in \mathcal{L}^1$. Ako su $\mathcal{H}$ i $\sigma(Y) \vee \mathcal{G}$ nezavisne σ-algebre, dokažite da je tada $\mathbb{E}[Y|\mathcal{H} \vee \mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$. Primijetite da je tada tvrdnja (*xi*) teorema 2.26 posljedica ove tvrdnje za $\mathcal{G} = \{\emptyset, \Omega\}$.
(**Uputa:** Primijetite da vrijedi: (1) $\mathcal{H}$ i $\sigma(Y) \vee \mathcal{G}$ su nezavisne ako i samo ako su za sve $H \in \mathcal{H}$ i $G \in \mathcal{G}$, $\mathbb{1}_H$ i $Y \cdot \mathbb{1}_G$ nezavisne slučajne varijable, (2) $\mathcal{H} \vee \mathcal{G} = \sigma(\{H \cap G : H \in \mathcal{H}, G \in \mathcal{G}\})$.)

Zadatak 2.33. Neka je (X, Y) dvodimenzionalni slučajni vektor takav da je za svaki $x \in \mathbb{R}$, uvjetna distribucija od Y uz dano $X = x$ normalna $N(x, \nu^2)$, a marginalna distribucija od X je normalna $N(\mu, \sigma^2)$. Koji je zakon razdiobe od Y, a koji od (X, Y)?

Zadatak 2.34. Neka je (R, S) dvodimenzionalni slučajni vektor takav da je $N = R + S$ Poissonova slučajna varijabla $P(\lambda)$, a uvjetna razdioba od (R, S) uvjetno na $N = n$ je polinomna razdioba $M(n; p, q)$ ($\lambda > 0$, $p, q \in \langle 0, 1\rangle$ takvi da je $p + q = 1$). Pokažite da su tada R i S nezavisne Poissonove varijable, $R \sim P(p\lambda)$ i $S \sim P(q\lambda)$ (usporedite sa zadatkom 1.48).

2.5 Uvjetno matematičko očekivanje u prostoru $\mathcal{L}^2$

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ zadani vjerojatnosni prostor, te neka je $\mathcal{L}^2$ pripadni prostor slučajnih varijabli konačne varijance koje su definirane na tom prostoru. Nadalje, neka je $Y \in \mathcal{L}^2$ i neka je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Iz dokaza teorema 2.8 slijedi da je uvjetno očekivanje $\mathbb{E}[Y|\mathcal{G}]$ (preciznije, svaka verzija uvjetnog očekivanja) ortogonalna projekcija od Y u potprostor $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ od $\mathcal{L}^2$. Dakle, $\mathbb{E}[Y|\mathcal{G}] \in \mathcal{L}^2$, tj. ako je Y konačne varijance, onda je i uvjetno očekivanje od Y uz dano $\mathcal{G}$ konačne varijance. Slijedi da je rezidual $Y - \mathbb{E}[Y|\mathcal{G}] \in \mathcal{L}^2$. Stoga je sljedeća definicija korektna.

Definicija 2.35. *Neka je $Y \in \mathcal{L}^2$ i $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Tada je* **uvjetna varijanca od Y uz dano $\mathcal{G}$** *slučajna varijabla koja je verzija od $\mathbb{E}[(X - \mathbb{E}[Y|\mathcal{G}])^2|\,\mathcal{G}]$. Označimo:*

$$\mathbb{V}\text{ar}[Y|\mathcal{G}] := \mathbb{E}[(X - \mathbb{E}[Y|\mathcal{G}])^2|\,\mathcal{G}].$$

Ako je i $Z \in \mathcal{L}^2$, tada je **uvjetna kovarijanca od** Y **i** Z **uz dano** $\mathcal{G}$ *slučajna varijabla koja je verzija od* $\mathbb{E}[(Y - \mathbb{E}[Y|\mathcal{G}]) \cdot (Z - \mathbb{E}[Z|\mathcal{G}]) \,|\, \mathcal{G}]$. *Označimo:*

$$\mathrm{cov}[Y, Z|\mathcal{G}] := \mathbb{E}[(Y - \mathbb{E}[Y|\mathcal{G}]) \cdot (Z - \mathbb{E}[Z|\mathcal{G}]) \,|\, \mathcal{G}].$$

Iz gornje definicije je jasno da je

$$\mathrm{cov}[Y, Y|\mathcal{G}] \equiv \mathbb{V}\mathrm{ar}[Y|\mathcal{G}], \tag{2.32}$$

te da su uvjetna varijanca i uvjetna kovarijanca g.s. jedinstvene. Nadalje, budući da je $(Y - \mathbb{E}[Y|\mathcal{G}])^2 \geq 0$, primjenom svojstva monotonosti uvjetnog očekivanja (teorem 2.26 (iv)) slijedi da je uvjetna varijanca g.s. nenegativna.

U sljedećim teoremima navedena su neka svojstva uvjetne varijance i uvjetne kovarijance.

Teorem 2.36. *Neka su $X, Y, Z \in \mathcal{L}^2$ i neka je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Tada vrijedi:*

(i) $\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] = \mathbb{E}[Y^2|\mathcal{G}] - (\mathbb{E}[Y|\mathcal{G}])^2$ *g.s.*
(ii) $\mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]] + \mathbb{V}\mathrm{ar}(\mathbb{E}[Y|\mathcal{G}]) = \mathbb{V}\mathrm{ar}Y$
(iii) $\mathbb{V}\mathrm{ar}(\mathbb{E}[Y|\mathcal{G}]) \leq \mathbb{V}\mathrm{ar}Y$
(iv) $\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] = 0$ *g.s. ako i samo ako je* $Y = \hat{Y}$ *g.s. za neku verziju* $\hat{Y}$ *od* $\mathbb{E}[Y|\mathcal{G}]$
(v) *Ako su U i W $\mathcal{G}$-izmjerive varijable, $W \in \mathcal{L}^2$ i U je ograničena, tada je* $\mathbb{V}\mathrm{ar}[UY + W \,|\, \mathcal{G}] = U^2 \mathbb{V}\mathrm{ar}[Y|\mathcal{G}]$ *g.s.*
(vi) $\mathrm{cov}[Y, Z|\mathcal{G}] = \mathbb{E}[Y \cdot Z|\mathcal{G}] - \mathbb{E}[Y|\mathcal{G}] \cdot \mathbb{E}[Z|\mathcal{G}]$ *g.s.*
(vii) $\mathrm{cov}[Y, Z|\mathcal{G}] = \mathrm{cov}[Z, Y|\mathcal{G}]$ *g.s.*
$(viii)$ $\mathbb{E}(\mathrm{cov}[Y, Z|\mathcal{G}]) + \mathrm{cov}(\mathbb{E}[Y|\mathcal{G}], \mathbb{E}[Z|\mathcal{G}]) = \mathrm{cov}(Y, Z)$
(ix) *Ako je Z $\mathcal{G}$-izmjeriva varijabla, tada je* $\mathrm{cov}[Y, Z|\mathcal{G}] = 0$ *g.s.*
(x) *Ako su U, V, W $\mathcal{G}$-izmjerive varijable, $W \in \mathcal{L}^2$ i U, V su ograničene, tada je* $\mathrm{cov}[UY + VZ + W, X \,|\, \mathcal{G}] = U \cdot \mathrm{cov}[Y, X|\mathcal{G}] + V \cdot \mathrm{cov}[Z, X|\mathcal{G}]$ *g.s.*

Dokaz. Neka je $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$. Prema definiciji uvjetne varijance 2.35, iz linearnosti uvjetnog očekivanja (iii) i svojstava (ii) i (x) iz teorema 2.26 slijedi

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] &\overset{\text{(def. 2.35)}}{=} \mathbb{E}[(Y - \hat{Y})^2|\mathcal{G}] = \mathbb{E}[Y^2 - 2\hat{Y}Y + \hat{Y}^2|\mathcal{G}] \overset{\text{tm. 2.26}\,(ii,iii,x)}{=} \\
&= \mathbb{E}[Y^2|\mathcal{G}] - 2\hat{Y}\mathbb{E}[Y|\mathcal{G}] + \hat{Y}^2 = \mathbb{E}[Y^2|\mathcal{G}] - \hat{Y}^2 = \mathbb{E}[Y^2|\mathcal{G}] - (\mathbb{E}[Y|\mathcal{G}])^2 \text{ g.s.}
\end{aligned}$$

što povlači tvrdnju (i).

Kao i prije, neka je $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$ i $\mu_Y = \mathbb{E}Y$. Korištenjem linearnosti matematičkog očekivanja (teorem 1.56 (i)), jednakosti (2.33), te činjenice da je $\mathbb{E}\hat{Y} = \mathbb{E}Y$ (prema teoremu 2.26 (i)), slijedi

$$\begin{aligned}
\mathbb{V}\mathrm{ar}Y &= \mathbb{E}[(Y - \mu_Y)^2] = \mathbb{E}[((Y - \hat{Y}) + (\hat{Y} - \mu_Y))^2] = \\
&= \mathbb{E}[(Y - \hat{Y})^2 + 2(\hat{Y} - \mu_Y)(Y - \hat{Y}) + (\hat{Y} - \mu_Y)^2] \overset{\text{(lin.m.o.)}}{=} \\
&= \mathbb{E}[(Y - \hat{Y})^2] + 2\mathbb{E}[(\hat{Y} - \mu_Y)(Y - \hat{Y})] + \mathbb{E}[(\hat{Y} - \mu_Y)^2].
\end{aligned}$$

Budući da je $\mathbb{V}\mathrm{ar}(\mathbb{E}[X|\mathcal{G}]) = \mathbb{V}\mathrm{ar}\hat{Y} = \mathbb{E}[(\hat{Y} - \mu_Y)^2]$, ako dokažemo da je srednji član u prethodno dobivenom izrazu jednak nuli, te da je $\mathbb{E}[(Y - \hat{Y})^2] = \mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]]$, onda smo dokazali tvrdnju (ii). Prvo,

$$\begin{aligned}
&\mathbb{E}[(\hat{Y} - \mu_Y)(Y - \hat{Y})] \overset{\text{tm. 2.26}\,(i)}{=} \mathbb{E}[\mathbb{E}[(\hat{Y} - \mu_Y)(Y - \hat{Y})|\mathcal{G}]] \overset{\text{tm. 2.26}\,(x)}{=} \\
= \; &\mathbb{E}[(\hat{Y} - \mu_Y)\mathbb{E}[(Y - \hat{Y})|\mathcal{G}]] \overset{\text{tm. 2.26}\,(ii-iii)}{=} \mathbb{E}[(\hat{Y} - \mu_Y)(\mathbb{E}[Y|\mathcal{G}] - \hat{Y})] = \\
= \; &\mathbb{E}[(\hat{Y} - \mu_Y) \cdot 0] = 0.
\end{aligned}$$

Nadalje, iz svojstva uvjetnog očekivanja (i), teorem 2.26, slijedi da je

$$\mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]] \overset{(\mathrm{def.}\,2.35)}{=} \mathbb{E}[\mathbb{E}[(Y-\hat{Y})^2|\mathcal{G}]] \overset{\mathrm{tm.}\,2.26\,(i)}{=} \mathbb{E}[(Y-\hat{Y})^2]. \tag{2.33}$$

Dakle, prema prethodnom, slijedi tvrdnja (ii). Odavde odmah slijedi i nejednakost (iii), budući da je $\mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]] \geq 0$.

Neka je sada $Y = \hat{Y}$ g.s. za neku verziju $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$. Tada je $Y^2 = \hat{Y}^2$ g.s., pa vrijedi da je $\hat{Y}^2$ jedna verzija od $\mathbb{E}[Y^2|\mathcal{G}]$ (prema (ii) iz teorema 2.26 i vidjeti početak dokaza svojstva monotonosti uvjetnog očekivanja (iv) iz istoga teorema). Sada iz (i) slijedi

$$\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] \overset{(i)}{=} \mathbb{E}[Y^2|\mathcal{G}] - (\mathbb{E}[Y|\mathcal{G}])^2 = \hat{Y}^2 - \hat{Y}^2 = 0 \text{ g.s.}$$

Obratno, neka je $\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] = 0$ g.s. i $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$. Budući da je $\mathbb{V}\mathrm{ar}[Y|\mathcal{G}] = 0$ g.s., jednakost (2.33) povlači da je $\mathbb{E}[(Y-\hat{Y})^2] = 0$ odakle slijedi da je $Y = \hat{Y}$ g.s. Prema tome, vrijedi (iv).

Neka su U i W $\mathcal{G}$-izmjerive varijable takve da je U ograničena i $W \in \mathcal{L}^2$. Zbog linearnosti uvjetnog očekivanja (iii) i svojstva (x) iz teorema 2.26 je $\mathbb{E}[UY + W\,|\,\mathcal{G}] = U\hat{Y} + W$ g.s., gdje je $\hat{Y} = \mathbb{E}[Y|\mathcal{G}]$. Dakle, gotovo sigurno je

$$\mathbb{V}\mathrm{ar}[UY + W\,|\,\mathcal{G}] \overset{(\mathrm{def.}\,2.35)}{=} \mathbb{E}[U^2(Y-\hat{Y})^2|\mathcal{G}] \overset{\mathrm{tm.}\,2.26\,(x)}{=} U^2\mathbb{E}[(Y-\hat{Y})^2|\mathcal{G}] \overset{(\mathrm{def.}\,2.35)}{=} U^2\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]$$

čime je dokazana tvrdnja (v).

Tvrdnje $(vi - x)$ za uvjetnu kovarijancu dokazuju se analogno. □

Za svako $Y \in \mathcal{L}^2$ vrijedi

$$\mathbb{E}(\mathbb{V}\mathrm{ar}[Y|\mathcal{G}]]) = \mathbb{E}[(Y - \mathbb{E}[Y|\mathcal{G}])^2] = \inf_{Z\in\mathcal{G}} \mathbb{E}[(Y-Z)^2]. \tag{2.34}$$

Prva jednakost slijedi iz dokaza teorema 2.36, a druga jednakost je uvjet ortogonalnosti (1) iz teorema o projekciji 2.5. Dakle, srednja vrijednost uvjetne varijance od Y jednaka je srednjekvadratnoj pogrešci predviđanja Y pomoću $\mathbb{E}[Y|\mathcal{G}]$ na osnovi informacija reprezentiranih s $\mathcal{G}$. Rezidual $Y - \mathbb{E}[Y|\mathcal{G}]$ je pogreška predviđanja. Slično, za svako $Y, Z \in \mathcal{L}^2$ vrijedi

$$\mathbb{E}(\mathrm{cov}[Y,Z|\mathcal{G}]) = \mathbb{E}[(Y - \mathbb{E}[Y|\mathcal{G}]) \cdot (Z - \mathbb{E}[Z|\mathcal{G}])] = \mathrm{cov}(Y - \mathbb{E}[Y|\mathcal{G}], Z - \mathbb{E}[Z|\mathcal{G}]). \tag{2.35}$$

Nadalje, prva jednakost u (2.34) kaže da je (bezuvjetno) matematičko očekivanje uvjetne varijance jednako kvadratu $\mathcal{L}^2$-pseudonorme reziduala, a prva jednakost u (2.35) kaže da je (bezuvjetno) matematičko očekivanje uvjetne kovarijance dviju varijabli $\mathcal{L}^2$-pseudoskalarni produkt njihovih reziduala. Druga jednakost u (2.35) je posljedica teorema 1.61 (vi) jer je $\mathbb{E}(Y - \mathbb{E}[Y|\mathcal{G}]) = \mathbb{E}(Z - \mathbb{E}[Z|\mathcal{G}]) = 0$, a što je posljedica teorema 2.26 (i). Dakle, ako fiksiramo σ-algebru na koju uvjetujemo, matematičko očekivanje uvjetne varijance neke slučajne varijable iz $\mathcal{L}^2$ interpretiramo kao varijancu reziduala koji nastaje kada predviđamo tu varijablu pomoću njenog uvjetnog očekivanja, a matematičko očekivanje uvjetne kovarijance dviju slučajnih varijabli iz $\mathcal{L}^2$, kao kovarijancu njihovih reziduala. Prema tome, sljedeći korolar je direktna posljedica prethodnog teorema 2.36.

Korolar 2.37. *Neka su $X, Y, Z \in \mathcal{L}^2$ i neka je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Tada vrijedi:*

(i) $\|Y - \mathbb{E}[Y|\mathcal{G}]\|_2^2 = \mathbb{V}\mathrm{ar}(Y - \mathbb{E}[Y|\mathcal{G}]) = \mathbb{V}\mathrm{ar}Y - \mathbb{V}\mathrm{ar}(\mathbb{E}[Y|\mathcal{G}])$
(ii) $\langle Y - \mathbb{E}[Y|\mathcal{G}], Z - \mathbb{E}[Z|\mathcal{G}]\rangle = \mathrm{cov}(Y - \mathbb{E}[Y|\mathcal{G}], Z - \mathbb{E}[Z|\mathcal{G}]) =$
$= \mathrm{cov}(Y,Z) - \mathrm{cov}(\mathbb{E}[Y|\mathcal{G}], \mathbb{E}[Z|\mathcal{G}])$.

Primijetite da je druga jednakost u (i) prethodnog korolara Pitagorin poučak (vidjeti zadatak 2.7 (v)) ili, u statistici uobičajeni naziv, *raščlamba varijance* slučajne varijable Y kada tu varijablu predviđamo pomoću $\mathbb{E}[Y|\mathcal{G}]$ na osnovi informacija sadržanih u σ-podalgebri $\mathcal{G}$.

Nejednakost (iii) iz teorema 2.36 ćemo koristiti u poglavlju 3 za konstrukciju optimalnog procjenitelja.

U sljedećem primjeru pokazat ćemo jednu primjenu tvrdnje (i) iz teorema 2.26 i tvrdnje (ii) iz teorema 2.36.

Primjer 2.38. Neka je (X, Y) bivarijatni slučajni vektor takav da je X diskretna slučajna varijabla čiji je zakan razdiobe zadan funkcijom mase $\mathbb{P}(X = k) = p_k$ za $k = 1, 2, \ldots, m$, a Y je slučajna varijabla sa svojstvom da je matematičko očekivanje uvjetne razdiobe od Y uz dano $X = k$ jednako μ_k, a varijanca iste uvjetne razdiobe je jednaka σ_k^2 za $k = 1, 2, \ldots, m$. Tada su bezuvjetno matematičko očekivanje i varijanca od Y, primjenom teorema 2.26 (i), teorema 2.36 (ii) i regresijske funkcije (teorema 2.16), jednaki:

$$\mathbb{E}Y \overset{\text{tm.2.26}(i)}{=} \mathbb{E}(\mathbb{E}[Y|X]) \overset{\text{(reg.fun.)}}{=} \sum_{k=1}^{m} \mathbb{E}[Y|X = k]\, p_k = \sum_{k=1}^{m} p_k \mu_k =: \overline{\mu} \tag{2.36}$$

$$\begin{aligned}
\mathbb{V}\mathrm{ar}Y &\overset{\text{tm.2.36}(ii)}{=} \mathbb{E}(\mathbb{V}\mathrm{ar}[Y|X]) + \mathbb{V}\mathrm{ar}(\mathbb{E}[Y|X]) \overset{\text{(reg.fun.)}}{=} \\
&= \sum_{k=1}^{m} \mathbb{V}\mathrm{ar}[Y|X = k]\, p_k + \sum_{k=1}^{m} (\mathbb{E}[Y|X = k] - \mathbb{E}(\mathbb{E}[Y|X]))^2 p_k \overset{(2.36)}{=} \\
&= \sum_{k=1}^{m} \sigma_k^2 p_k + \sum_{k=1}^{m} (\mu_k - \overline{\mu})^2 p_k.
\end{aligned} \tag{2.37}$$

Prvi član u raščlambi (bezuvjetne) varijance od Y (2.37), $\sum_{k=1}^{m} \sigma_k^2 p_k$, zovemo *srednja varijabilnost od* Y **unutar** *grupe*, a drugi član, $\sum_{k=1}^{m} (\mu_k - \overline{\mu})^2 p_k$, zovemo *varijabilnost srednjih vrijednosti od* Y **među** *grupama*, pri čemu pod *grupom* mislimo na statističke jedinice reprezentirane klasom (događajem) $\{X = k\}$ $(k = 1, 2, \ldots, m)$. U tom smislu, (bezuvjetno) matematičko očekivanje od Y (2.36) je srednja vrijednost matematičkih očekivanja od Y obzirom na grupe. □

Slijedeći teorem je direktna posljedica uvjetne Jensenove nejednakosti $(viii)$ iz teorema 2.26.

Teorem 2.39. *Neka su* $Y \in \mathcal{L}^2$ *i* $\mathcal{G}$ σ*-podalgebra od* $\mathcal{F}$*. Tada vrijedi:*

(i) $(\mathbb{E}[Y|\mathcal{G}])^2 \leq \mathbb{E}[Y^2|\mathcal{G}]$ *g.s.*
(ii) $\|\mathbb{E}[Y|\mathcal{G}]\|_2 \leq \|Y\|_2$

Dokaz. Neka je $\phi(x) = x^2$, $x \in \mathbb{R}$. $\phi : \mathbb{R} \to \mathbb{R}$ je konveksna funkcija i Y i $\phi(Y) = Y^2$ su konačnih matematičkih očekivanja. Prema uvjetnoj Jensenovoj jednakosti teorem 2.26 $(viii)$ slijedi tvrdnja (i). Tvrdnja (ii) slijedi iz monotonosti matematičkog očekivanja teorem 1.56 (ii). □

Interpretacija tvrdnje (ii) prethodnog teorema za prostor $\mathcal{L}^2$ je slična interpretaciji tvrdnje (ii) korolara 2.27 za prostor $\mathcal{L}^1$: preslikavanje koje svakoj klasi ekvivalencije iz L^2 pridružuje klasu uvjetnog očekivanja uz danu σ-podalgebru je kontrakcija u L^2.

Zadatak 2.40. Dokažite tvrdnje $(vi - x)$ teorema 2.36 za uvjetnu kovarijancu.

Zadatak 2.41. Neka je $p \geq 1$ realan broj i neka je $\mathcal{L}^p$ prostor svih slučajnih varijabli definiranih na istome vjerojatnosnom prostoru s konačnim p-tim momentom. Nadalje, neka je L^p kvocijentni skup od $\mathcal{L}^p$ obzirom na relaciju ekvivalenciju “biti jednak g.s.” Dokažite da je tada preslikavanje koje svakoj klasi ekvivalencije iz L^p pridružuje klasu uvjetnog očekivanja uz danu σ-podalgebru, kontrakcija u L^p.

Poglavlje 3

Osnovni pojmovi matematičke statistike

3.1 Statistički model

Poopćimo pojam vjerojatnosnog prostora.

Definicija 3.1. *Neka je* $(\Omega, \mathcal{F})$ *izmjeriv prostor i* $\mathcal{P}$ *množina vjerojatnosnih mjera definiranih na* $(\Omega, \mathcal{F})$*. Tada uređenu trojku* $(\Omega, \mathcal{F}, \mathcal{P})$ *zovemo* **statistička struktura**.

Primijetite da statističku strukturu $(\Omega, \mathcal{F}, \mathcal{P})$ s jednočlanom množinom $\mathcal{P} = \{\mathbb{P}\}$ možemo poistovjetiti s vjerojatnosnim prostorom $(\Omega, \mathcal{F}, \mathbb{P})$.

Množina $\mathcal{P}$ je često parametrizirana konačnodimenzijskim *parametrom* θ:

$$\mathcal{P} = \{\mathbb{P}_\theta \, : \, \theta \in \Theta\}.$$

Ovdje je Θ podskup od $\mathbb{R}^m$ $(m \geq 1)$ koji zovemo *parametarski prostor*. Intuitivno, množina $\mathcal{P}$ predstavlja sve moguće kandidate za pravu razdiobu (vjerojatnost) koju želimo odrediti na osnovi opažanja jedne ili više slučajnih veličina definiranih na izmjerivom prostoru $(\Omega, \mathcal{F})$. Ako je množina $\mathcal{P}$ parametrizirana konačnodimenzijskim parametrom, tada se zadaća određivanja prave razdiobe svodi na pronalaženje nepoznate vrijednosti parametra θ. Dakako, parametarski prostor Θ tada interpretiramo kao skup svih mogućih vrijednosti od θ. Da bi ta osnovna zadaća statistike bila moguća, parametar θ mora zadovoljavati uvjet **raspoznavanja** (engl. *identifiability*):

$$(\forall \theta_1, \theta_2 \in \Theta) \;\; \theta_1 \neq \theta_2 \; \Rightarrow \; \mathbb{P}_{\theta_1} \neq \mathbb{P}_{\theta_2}.$$

Od sada pa nadalje, *za sve paremetarske modele pretpostavljamo da je zadovoljen uvjet raspoznavanja.*

Primjer 3.2. Primjeri $(a - d)$ parametriziranih statističkih struktura $(\Omega, \mathcal{F}, \mathcal{P})$, gdje je $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$.
(a) Neka su $\Omega = \{0, 1\}$ dvočlani skup i $\mathcal{F} = \mathcal{P}(\Omega)$ partitivni skup od Ω. Za parametar $\theta \in \langle 0, 1\rangle = \Theta$, definiramo vjerojatnost $\mathbb{P}_\theta$ djelovanjem na elementarnim događajima: $\mathbb{P}_\theta(\{0\}) := 1 - \theta$, $\mathbb{P}_\theta(\{1\}) := \theta$.
(b) Neka su $\Omega = \{0, 1, 2, \ldots\} = \mathbb{N}_0$ skup nenegativnih cijelih brojeva i $\mathcal{F} = \mathcal{P}(\mathbb{N}_0)$. Označimo sa λ parametar s vrijednostima u $\langle 0, +\infty\rangle = \Theta$ i definiramo pripadnu vjerojatnost sa $\mathbb{P}_\lambda(\{k\})) := (\lambda^k) \exp(-\lambda)/k!$, $k \in \mathbb{N}_0$.

(c) Neka su $\Omega = \mathbb{R}$ i $\mathcal{F} = \mathcal{B}(\mathbb{R})$ σ-algebra Borelovih skupova u $\mathbb{R}$. Za parametar $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \langle 0, +\infty\rangle = \Theta$ definiramo vjerojatnost $\mathbb{P}_\theta$ sa

$$\mathbb{P}_\theta(B) := \int\limits_B \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2},\ B \in \mathcal{B}(\mathbb{R}).$$

(d) Općenito, neka je $(\Omega, \mathcal{F}) = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ Borelov izmjeriv prostor $(k \geq 1)$. Ako je $f(\cdot;\theta)$ gustoća neke k-dimenzionalne neprekidne ili diskretne slučajne veličine, parametrizirana nekim parametrom $\theta \in \Theta$, tada, za zadanu vrijednost parametra θ, možemo definirati vjerojatnost $\mathbb{P}_\theta$ sa:

$$\mathbb{P}_\theta(B) := \begin{cases} \sum_{x \in B} f(x;\theta) & \text{u diskretnom slučaju} \\ \int\limits_B f(x;\theta)\, dx & \text{u neprekidnom slučaju} \end{cases}$$

za $B \in \mathcal{B}(\mathbb{R}^k)$. $\square$

Bazični pojam u matematičkoj statistici je pojam *slučajnog uzorka* koji predstavlja model za opažanja ili mjerenja statističkih varijabli. Slučajan uzorak se može definirati općenitije od definicije koju ćemo mi dati, ali za potrebe ovoga izlaganja dovoljno je pretpostaviti da su mjerenja slučajnih veličina međusobno nezavisna, odnosno da se statističke jedinice biraju u uzorak na isti *jednostavan* način i *bez ponavljanja*.

Definicija 3.3. Slučajan uzorak duljine *n na statističkoj strukturi $(\Omega, \mathcal{F}, \mathcal{P})$ je niz $X_1, X_2, \ldots, X_n$ slučajnih veličina na $(\Omega, \mathcal{F})$ takvih da su nezavisne i jednako distribuirane u odnosu na svaku vjerojatnost $\mathbb{P} \in \mathcal{P}$.*

Često puta ćemo slučajni uzorak duljine n prikazivati u obliku n-torke $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ i kraće ćemo reći da je to *slučajan uzorak duljine n*. Ako su elementi te n-torke slučajne varijable, tada će se raditi o slučajnom vektoru dimenzije n kojemu su komponente nezavisne i jednako distribuirane (n.j.d.) slučajne varijable u odnosu na svaku vjerojatnost iz množine $\mathcal{P}$. U računima tako zapisani slučajni uzorak interpretiramo u transponiranom obliku:

$$(X_1, X_2, \ldots, X_n)^\tau = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}.$$

Dakle, slučajan uzorak varijabli duljine n je matrica tipa $(n, 1)$, a slučajni uzorak iste duljine n, ali d-dimenzionalnih vektora je matrica tipa (n, d). U oba slučaja su retci matrica n.j.d. slučajne veličine.

Postoje li netrivijalna statistička struktura i slučajni uzorak na toj strukturi proizvoljne duljine n? Postoje i za njegovu konstrukciju možemo koristiti uputu iz poglavlja 1 (potpoglavlje 1.5).

Definicija 3.4. Statistika *na statističkoj strukturi $(\Omega, \mathcal{F}, \mathcal{P})$ je svaka slučajna veličina koja je izmjeriva funkcija nekog slučajnog uzorka na toj statističkoj strukturi.*

Posebno, slučajnu veličinu T *konačne* dimenzije k $(k \geq 1)$ zovemo statistikom ako postoje brojevi $n \in \mathbb{N}$, $d \in \mathbb{N}$, izmjerivo preslikavanje $t : \mathbb{R}^{dn} \to \mathbb{R}^k$ i n-dimenzionalni slučajni uzorak $(X_1, X_2, \ldots, X_n)$ veličina dimenzije d $(d \geq 1)$, takvi da je

$$T = t(X_1, X_2, \ldots, X_n).$$

Primjer 3.5. Neka je $X_1, X_2, \ldots, X_n$ slučajan uzorak varijabli, duljine n. Tada su statistike:

$$\begin{aligned}
\overline{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i \ \ (\textit{aritmetička sredina}) \\
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \ \ (\textit{uzoračka varijanca}).
\end{aligned}$$

Obje navedene statistke su dimenzije 1. Za aritmetičku sredinu je $\overline{X}_n = t(X_1, X_2, \ldots, X_n)$ gdje je

$$t(x_1, x_2, \ldots, x_n) = \overline{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n), \ \ (x_1, x_2, \ldots, x_n) = \mathbf{x} \in \mathbb{R}^n,$$

izmjeriva funkcija $t : \mathbb{R}^n \to \mathbb{R}$, a za uzoračku varijancu je $S_n^2 = v(X_1, X_2, \ldots, X_n)$ gdje je

$$v(x_1, x_2, \ldots, x_n) = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, \ \ (x_1, x_2, \ldots, x_n) = \mathbf{x} \in \mathbb{R}^n,$$

izmjeriva funkcija $v : \mathbb{R}^n \to \mathbb{R}$. $\square$

Neka je $X : \Omega \to \mathbb{R}^d$ slučajna veličina dimenzije d ($d \geq 1$) definirana na parametriziranoj statističkoj strukturi $(\Omega, \mathcal{F}, \mathcal{P})$, $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$. Za $\theta \in \Theta$ stavimo

$$F(x;\theta) \equiv F_\theta(x) = \mathbb{P}_\theta(X \leq x), \ x \in \mathbb{R}^d.$$

Tada je $F(\cdot;\theta)$ funkcija distribucije od X uz vjerojatnost $\mathbb{P}_\theta$. Kažemo da X *pripada* **statističkom modelu** $\mathcal{P}' = \{F(\cdot;\theta) : \theta \in \Theta\}$.

U primjenama izučavamo veličinu X i njenu *populacijsku* razdiobu. Stoga će nam od interesa biti ona množina $\mathcal{P}$ statističke strukture koja se sastoji od mogućih zakona razdioba od X indeksiranih parametrom θ s vrijednostima u Θ. Uz tu pretpostavku postoji jednoznačna veza između $\mathbb{P}_\theta \in \mathcal{P}$ i $F(\cdot;\theta) \in \mathcal{P}'$ pa možemo poistovjetiti $\mathcal{P}$ sa $\mathcal{P}'$. Dakle,

$$\mathcal{P} \equiv \{F(\cdot;\theta) : \theta \in \Theta\}.$$

U primjenama će često X biti neprekidna ili diskretna slučajna veličina (varijabla ili vektor) s gustoćom $f(\cdot;\theta) \equiv f_\theta : \mathbb{R}^d \to [0, +\infty\rangle$ u odnosu na vjerojatnost $\mathbb{P}_\theta$. Budući da je i veza funkcije distribucije i gustoće jednoznačna na način kako je to opisano u poglavlju 1 (potpoglavlje 1.3), možemo $\mathcal{P}$ poistovjetiti i s množinom gustoća:

$$\mathcal{P} \equiv \{f(\cdot;\theta) : \theta \in \Theta\}.$$

Od sada pa nadalje pretpostavljamo da je

$$\mathcal{P} = \{f(\cdot;\theta) \ : \ \theta \in \Theta\}.$$

Fraza:

Neka je $(X_1, X_2, \ldots, X_n)$ *slučajan uzorak iz modela* $\mathcal{P}$

znači da je $(X_1, X_2, \ldots, X_n)$ takav slučajan uzorak da zakon razdiobe od X_i (za sve i) ima gustoću $f(\cdot;\theta)$ za $\theta \in \Theta$.

3.2 Dovoljna statistika

Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$. Podsjetimo se da je cilj statističke analize zaključivati o pravoj (populacijskoj) vrijednosti parametra θ na osnovi informacije sadržene u uzorku. Pojam *dovoljnosti* odnosi se na pojednostavljenje forme kroz transformaciju uzorka tako da ukupna informacija o populacijskoj vrijednosti parametra ostane ista.

Definicija 3.6. *Statistika $T = t(X_1, X_2, \ldots, X_n)$ dimenzije k ($k \geq 1$) je* **dovoljna za** θ *ako za svako $y \in \mathbb{R}^k$ za koje postoji uvjetna razdioba slučajnog uzorka $(X_1, X_2, \ldots, X_n)$ uz uvjet $T = y$, ta uvjetna razdioba ne ovisi o parametru θ.*

Primijetimo da, ako postoji zajednička gustoća $f_{\mathbf{X},T}(\cdot,\cdot;\theta)$ slučajnog uzorka $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ i statistike $T = t(\mathbf{X})$, tada je T dovoljna statistika ako uvjetna gustoća

$$f_{\mathbf{X}|T}(\mathbf{x}|y) = \frac{f_{\mathbf{X},T}(\mathbf{x}, y;\theta)}{f_T(y;\theta)}, \quad \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn},$$

od $\mathbf{X}$ uz uvjet $T = y$, ne ovisi o θ za sve $y \in \text{supp} f_T$.

Primjer 3.7. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak iz Bernoullijevog modela $b(1,\theta)$, $\theta \in \langle 0, 1\rangle$. Pokažimo da je statistika $T = X_1 + X_2 + \cdots + X_n$ dovoljna za θ. T ima binomnu razdiobu $b(n,\theta)$. Budući da je model diskretan, $\mathbf{X}$ i T imaju zajedničku gustoću. Odredimo gustoću uvjetne razdiobe od $\mathbf{X}$ uz dano $T = k$ gdje je $k \in \{0, 1, \ldots, n\}$ bilo koji broj kojega T postiže (odnosno, binomna razdioba s parametrima n i θ) s pozitivnom vjerojatnosti. Tada je za $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ takvu n-torku da su svi brojevi $x_i \in \{0, 1\}$ sa svojstvom da je $x_1 + x_2 + \cdots + x_n = k$, uvjetna gustoća jednaka:

$$\begin{aligned}
f_{\mathbf{X}|T}(\mathbf{x}|k) &= \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | T = k) = \\
&= \frac{\mathbb{P}_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, T = k)}{\mathbb{P}_\theta(T = k)} = \\
&= \frac{\mathbb{P}_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)}{\mathbb{P}_\theta(T = k)} \overset{\text{nez. od } X_1,\ldots,X_n}{=} \\
&= \frac{\mathbb{P}_\theta(X_1 = x_1) \cdot \mathbb{P}_\theta(X_2 = x_2) \cdots \mathbb{P}_\theta(X_n = x_n)}{\mathbb{P}_\theta(T = k)} \overset{X_i \sim b(1,\theta), T \sim b(n,\theta)}{=} \\
&= \frac{\theta^{x_1+x_2+\cdots+x_n}(1-\theta)^{n-(x_1+x_2+\cdots+x_n)}}{\binom{n}{k}\theta^k(1-\theta)^{n-k}} = \frac{\theta^k(1-\theta)^{n-k}}{\binom{n}{k}\theta^k(1-\theta)^{n-k}} = \\
&= \frac{1}{\binom{n}{k}}.
\end{aligned}$$

Ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ takva n-torka da nisu svi njeni elementi jednaki 0 ili 1, ili $x_1 + x_2 + \cdots x_n \neq k$, tada je

$$f_{\mathbf{X}|T}(\mathbf{x}|k) = \frac{\mathbb{P}_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, T = k)}{\mathbb{P}_\theta(T = k)} = \frac{\mathbb{P}_\theta(\emptyset)}{\mathbb{P}_\theta(T = k)} = 0.$$

Dakle, za $k \in \{0, 1, \ldots, n\}$ je

$$f_{\mathbf{X}|T}(\mathbf{x}|k) = \begin{cases} \frac{1}{\binom{n}{k}}, & (\forall i)(x_i \in \{0,1\}) \text{ i } x_1 + x_2 + \cdots + x_n = k \\ 0, & (\exists i)(x_i \notin \{0,1\}) \text{ ili } x_1 + x_2 + \cdots + x_n \neq k \end{cases} \tag{3.1}$$

neovisno o θ. Prema tome, T je dovoljna statistika za θ. Primijetimo da iz (3.1) slijedi da je uvjetna razdioba slučajnog uzorka $\mathbf{X}$ uz dano $T = k$ diskretna uniformna na skupu svih uređenih n-torki $(x_1, x_2, \ldots, x_n)$ nula i jedinica takvih da je $x_1 + x_2 + \cdots + x_n = k$. □

Propozicija 3.8. *Statistika $T = t(X_1, X_2, \ldots, X_n)$ dimenzije k $(k \geq 1)$ je dovoljna za θ ako i samo ako za svaku statistiku $S = s(X_1, X_2, \ldots, X_n)$ vrijedi da uvjetna razdioba od S uz dano $T = y$ ne ovisi o θ, za svako $y \in \mathbb{R}^k$ za koje postoji ta uvjetna razdioba.*

Dokaz. Smjer '$\Leftarrow$': Ako za svaku statistiku S vrijedi da joj je uvjetna razdioba uz uvjet $T = y$ neovisna o θ, tada specijalno za S možemo uzeti slučajan uzorak $\mathbf{X}$, tj. $S = \mathbf{X} = (X_1, X_2, \ldots, X_n) = s(\mathbf{X})$ gdje je $s : \mathbb{R}^{dn} \to \mathbb{R}^{dn}$ funkcija identiteta. Po definiciji (3.6), T je dovaljna statistika za θ.

Smjer '$\Rightarrow$': Neka je T dovoljna statistika za θ i $S = s(\mathbf{X})$ proizvoljna statistika, recimo dimenzije ℓ $(\ell \geq 1)$. Tada je za proizvoljni $y \in \mathbb{R}^k$ takav da uvjetni zakon razdiobe $\mathbb{P}_{S|T}(\cdot|y;\theta)$ (u odnosu na $\mathbb{P}_\theta$) od S uvjetno na $T = y$ postoji i za proizvoljni Borelov skup $B \in \mathcal{B}(\mathbb{R}^\ell)$, vrijedi:

$$\begin{aligned}
\mathbb{P}_{S|T}(B|y;\theta) &\overset{(1.17-1.18,\,2.31)}{=} \mathbb{P}_\theta(S \in B | T = y) = \\
&= \mathbb{P}_\theta(s(\mathbf{X}) \in B | T = y) = \mathbb{P}_\theta(\mathbf{X} \in s^{-1}(B) | T = y) \overset{(1.17-1.18,\,2.31)}{=} \\
&= \mathbb{P}_{\mathbf{X}|T}(s^{-1}(B)|y)
\end{aligned}$$

što ne ovisi o θ. Naime, uvjetni zakon razdiobe $\mathbb{P}_{\mathbf{X}|T}(\cdot|y)$ od $\mathbf{X}$ uz dano $T = y$ ne ovisi o θ jer je T dovoljna statistika za θ. Primijetite da $\mathbb{P}_{\mathbf{X}|T}(s^{-1}(\cdot)|y)$ postoji za dani $y \in \mathbb{R}^k$ za koji postoji $\mathbb{P}_{S|T}(\cdot|y;\theta)$. Budući da su $B \in \mathcal{B}(\mathbb{R}^\ell)$ i y bili proizvoljni, slijedi da uvjetni zakon razdiobe $\mathbb{P}_{S|T}(\cdot|y;\theta)$ od S uz uvjet $T = y$ ne ovisi o θ. Zbog proizvoljnosti od S, slijedi tvrdnja. $\square$

Navedimo sada važnu karakterizaciju dovoljnosti. Podsjetimo se da smo pretpostavili da su svi statistički modeli koje izučavamo diskretni ili neprekidni, dakle, imaju gustoću. Prema tome, ako je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$, tada je gustoća slučajnog uzorka kao vektora dimenzije dn jednaka

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta), \quad \mathbf{x} = (x_1, x_2, \ldots x_n) \in \mathbb{R}^{dn}.$$

Teorem 3.9 (Neyman-Fisherov teorem o faktorizaciji)**.** *Statistika $T = t(\mathbf{X})$ dimenzije k $(k \geq 1)$ je dovoljna za θ ako i samo ako se gustoća slučajnog uzorka $\mathbf{X}$ može faktorizirati na sljedeći način:*

$$f_{\mathbf{X}}(\mathbf{x};\theta) = g_\theta(t(\mathbf{x})) \cdot h(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{dn}, \tag{3.2}$$

gdje su h, te g_θ za sve $\theta \in \Theta$, nenegativne izmjerive funkcije.

Dokaz. Dokaz sprovodimo samo za slučaj diskretnog modela. Za neprekidni slučaj pogledajte u [5].

Smjer '$\Rightarrow$': Neka je $T = t(\mathbf{X})$ dovoljna statistika za θ gdje je $t : \mathbb{R}^{dn} \to \mathbb{R}^k$ Borelova funkcija. Neka je $\theta \in \Theta$ proizvoljno te neka su $\mathbf{x} \in \mathbb{R}^{dn}$ i $y \in \mathbb{R}^k$ takvi da je $y = t(\mathbf{x})$ i $f_T(y;\theta) = \mathbb{P}_\theta(T = y) > 0$. Tada je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, T = y = t(\mathbf{x})) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = t(\mathbf{x})) \cdot \mathbb{P}_\theta(T = t(\mathbf{x})).$$

Budući da je T dovoljna za θ, $\mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = t(\mathbf{x})) =: h(\mathbf{x})$ ne ovisi o θ. Ako stavimo u desnu stranu gornje jednakosti da je $g_\theta(y) := \mathbb{P}_\theta(T = y)$, tada imamo da je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = t(\mathbf{x})) \cdot \mathbb{P}_\theta(T = t(\mathbf{x})) = h(\mathbf{x}) \cdot g_\theta(t(\mathbf{x})),$$

dakle, vrijedi (3.2).

Smjer '$\Leftarrow$': Pretpostavimo da vrijedi (3.2) za slučajni uzorak $\mathbf{X}$. Neka je $\mathbf{x} \in \mathbb{R}^{dn}$ proizvoljno. Tada je za $\theta \in \Theta$:

$$\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}; \theta) \overset{(3.2)}{=} g_\theta(t(\mathbf{x}))h(\mathbf{x}). \tag{3.3}$$

Neka je $y \in \mathbb{R}^k$ bilo koja vrijednost takva da je $f_T(y; \theta) = \mathbb{P}_\theta(T = y) > 0$. Tada je:

$$\mathbb{P}_\theta(T = y) = \sum_{t(\boldsymbol{\xi})=y} \mathbb{P}_\theta(\mathbf{X} = \boldsymbol{\xi}) \overset{(3.3)}{=} \sum_{t(\boldsymbol{\xi})=y} g_\theta(t(\boldsymbol{\xi}))h(\boldsymbol{\xi}) = g_\theta(y) \sum_{t(\boldsymbol{\xi})=y} h(\boldsymbol{\xi}). \tag{3.4}$$

U gornjem izrazu se sumira po svim $\boldsymbol{\xi} \in \mathbb{R}^{dn}$ takvima da je $t(\boldsymbol{\xi}) = y$. Ako je $\mathbf{x}$ takav da je $y = t(\mathbf{x})$, tada je:

$$\begin{aligned}
\mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = y) &= \frac{\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, T = y = t(\mathbf{x}))}{\mathbb{P}_\theta(T = y)} = \frac{\mathbb{P}_\theta(\mathbf{X} = \mathbf{x})}{\mathbb{P}_\theta(T = y)} \overset{(3.3-3.4)}{=} \\
&= \frac{g_\theta(t(\mathbf{x}))h(\mathbf{x})}{g_\theta(y) \sum_{t(\boldsymbol{\xi})=y} h(\boldsymbol{\xi})} \overset{y=t(\mathbf{x})}{=} \\
&= \frac{h(\mathbf{x})}{\sum_{t(\boldsymbol{\xi})=y} h(\boldsymbol{\xi})}.
\end{aligned}$$

Ako je $\mathbf{x}$ takav da je $y \neq t(\mathbf{x})$ tada je:

$$\mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = y) = \frac{\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, T = y)}{\mathbb{P}_\theta(T = y)} = \frac{\mathbb{P}_\theta(\emptyset)}{\mathbb{P}_\theta(T = y)} = 0.$$

Dakle,

$$f_{\mathbf{X}|T}(\mathbf{x}|y) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x} | T = y) = \begin{cases} \frac{h(\mathbf{x})}{\sum_{t(\boldsymbol{\xi})=y} h(\boldsymbol{\xi})}, & t(\mathbf{x}) = y \\ 0, & t(\mathbf{x}) \neq y \end{cases}$$

ne ovisi o θ pa je, po definiciji (3.6), T dovoljna statistika za θ. □

Primjer 3.10. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak iz normalnog modela $N(\mu, \sigma^2)$ s vektorskim parametrom $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \langle 0, +\infty\rangle$. Pomoću teorema o faktorizaciji 3.9 pokazat ćemo da je $T = (\overline{X}_n, S_n^2)$ dovoljna statistika za θ.

Za $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, neka su

$$t(\mathbf{x}) = \overline{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \text{ i } v(\mathbf{x}) = s^2 = \frac{1}{n-1}(\sum_{i=1}^n (x_i - \overline{x})^2)$$

aritmetička sredina i uzoračka varijanca realizacije $\mathbf{x}$ uzorka $\mathbf{X}$. Tada je

$$\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}; \theta) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2} = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2\right) = \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n ((x_i - \overline{x}) - (\overline{x} - \mu))^2\right) = \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n ((x_i - \overline{x})^2 - 2(\overline{x} - \mu)(x_i - \overline{x}) + (\overline{x} - \mu)^2)\right) = \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n (x_i - \overline{x})^2 + n(\overline{x} - \mu)^2\right)\right) = \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(n(\overline{x} - \mu)^2 + (n-1)s^2\right)\right) = \\
&= g_\theta(\overline{x}, s^2) \cdot h(\mathbf{x})
\end{aligned}$$

za

$$\begin{aligned} g_\theta(z_1,z_2) &= \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n}\cdot\exp\left(-\frac{1}{2\sigma^2}\left(n(z_1-\mu)^2+(n-1)z_2\right)\right), \quad z_1,z_2\in\mathbb{R} \\ h(\mathbf{x}) &\equiv 1, \quad \mathbf{x}\in\mathbb{R}^n. \end{aligned}$$

Prema teoremu o faktorizaciji 3.9, $T=(\overline{X}_n,S_n^2)$ je dovoljna statistika za θ. □

Je li funkcijska transformacija dovoljne statistike dovoljna statistika? Ne uvijek, ali *bijektivna* (i izmjeriva) transformacija dovoljne statistike jest dovoljna statistika. To je sadržaj prvog korolara Neyman-Fisherovog teorema o faktorizaciji 3.9. Prvo, uvedimo pojam *ekvivalentnih statistika.*

Definicija 3.11. *Za dvije statistike V i W dimenzije k ($k\geq 1$) kažemo da su* **ekvivalentne** *ako postoji bijekcija $\eta:\mathbb{R}^k\to\mathbb{R}^k$ takva da su η i η^{-1} Borelove funkcije i $\eta(V)=W$.*

Korolar 3.12. *Neka su V i W ekvivalentne statistike u statističkom modelu parametriziranom parametrom θ. Tada je W dovoljna statistika za θ ako i samo ako je V dovoljna statistika za θ.*

Dokaz. Neka je $V=v(\mathbf{X})$ dovoljna statistika za θ dimenzije k. Tada se gustoća slučajnog uzorka $\mathbf{X}$ iz modela parametriziranog sa θ, prema teoremu o faktorizaciji 3.9, može faktorizirati na sljedeći način:

$$f_{\mathbf{X}}(\mathbf{x};\theta)=g_\theta(v(\mathbf{x}))\cdot h(\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^{dn}. \tag{3.5}$$

Definirajmo funkciju $\tilde{g}_\theta:\mathbb{R}^k\to[0,+\infty\rangle$ sa

$$\tilde{g}_\theta(y):=g_\theta(\eta^{-1}(y)), \quad y\in\mathbb{R}^k, \tag{3.6}$$

gdje je $\eta:\mathbb{R}^k\to\mathbb{R}^k$ bijekcija takva da je $W=\eta(V)$. $\tilde{g}_\theta$ je izmjeriva funkcija jer je kompozicija izmjerivih funkcija. Neka je $W=w(\mathbf{X})$. Tada je za sve $\mathbf{x}\in\mathbb{R}^{dn}$,

$$\tilde{g}_\theta(w(\mathbf{x}))\stackrel{(3.6)}{=}g_\theta(\eta^{-1}(w(\mathbf{x})))\stackrel{w=\eta\circ v}{=}g_\theta(v(\mathbf{x})). \tag{3.7}$$

Dakle, prema prethodnom je:

$$f_{\mathbf{X}}(\mathbf{x};\theta)\stackrel{(3.5)}{=}g_\theta(v(\mathbf{x}))\cdot h(\mathbf{x})\stackrel{(3.7)}{=}\tilde{g}_\theta(w(\mathbf{x}))\cdot h(\mathbf{x}).$$

Primjenom obrata teorema o faktorizaciji 3.9, slijedi da je $W=w(\mathbf{X})$ dovoljna statistika za θ. Drugi smjer slijedi na isti način zbog simetrije. □

Drugi korolar Neyman-Fisherovog teorema o faktorizaciji je karakterizacija dovoljnosti za statističke modele koji se sastoje od gustoća čiji nosači ne ovise o parametru. Preciznije, neka je $\mathcal{P}=\{f(\cdot,\theta):\theta\in\Theta\}$ statistički model. Tada nosač $\mathrm{supp}f(\cdot;\theta)$ ne ovisi o θ ako je za sve $\theta,\theta'\in\Theta$ takve da je $\theta\neq\theta'$, $\mathrm{supp}f(\cdot;\theta)=\mathrm{supp}f(\cdot;\theta')$. Označimo sa:

$$S=\mathrm{supp}f(\cdot;\theta)=\{x\in\mathbb{R}^d:f(x;\theta)>0\} \text{ za svaki } \theta\in\Theta,$$

zajednički nosač gustoća iz $\mathcal{P}$. Neka je, kao i prije, sa $f_{\mathbf{X}}(\mathbf{x};\theta)$ označena funkcija gustoće slučajnog uzorka duljine n. Tada je nosač te gustoće

$$\mathrm{supp}f_{\mathbf{X}}(\cdot;\theta)=\overbrace{S\times S\times\cdots\times S}^{n}\equiv S^n \text{ za svaki } \theta\in\Theta.$$

Dakle, nosači gustoća slučajnog uzorka su također neovisni o θ i S^n je zajednički nosač tih gustoća.

Korolar 3.13. *Neka je* $\mathbf{X}$ *slučajan uzorak duljine* n *iz statističkog modela koji se sastoji od gustoća čiji nosači ne ovisi o parametru* θ *te neka je njihov zajednički nosač* S. *Tada je* $T = t(\mathbf{X})$ *dovoljna statistika za* θ *ako i samo ako je zadovoljen sljedeći uvjet:*

$$(\forall \theta, \theta' \in \Theta)\ \theta \neq \theta' \ \Rightarrow\ \frac{f_{\mathbf{X}}(\mathbf{x};\theta)}{f_{\mathbf{X}}(\mathbf{x};\theta')} = \psi_{\theta,\theta'}(t(\mathbf{x})),\ \mathbf{x} \in S^n,\ \textit{za neku funkciju}\ \psi_{\theta,\theta'}. \tag{3.8}$$

Dokaz. Smjer '$\Rightarrow$': Neka je $T = t(\mathbf{X})$ dovoljna statistika za θ. Tada je po teoremu o faktorizaciji 3.9:

$$f_{\mathbf{X}}(\mathbf{x};\theta) = g_\theta(t(\mathbf{x})) \cdot h(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^{dn},$$

pa za $\theta, \theta' \in \Theta$ takve da je $\theta \neq \theta'$ i $\mathbf{x} \in S^n$ je:

$$\frac{f_{\mathbf{X}}(\mathbf{x};\theta)}{f_{\mathbf{X}}(\mathbf{x};\theta')} = \frac{g_\theta(t(\mathbf{x}))}{g_{\theta'}(t(\mathbf{x}))} = \psi_{\theta,\theta'}(t(\mathbf{x}))$$

gdje je

$$\psi_{\theta,\theta'}(y) := \frac{g_\theta(y)}{g_{\theta'}(y)},\ \ y \in t(S^n),$$

funkcija definirana na slici restrikcije preslikavanja t na S^n.

Smjer '$\Leftarrow$': Pretpostavimo da je za preslikavanje t i slučajni uzorak $\mathbf{X}$ ispunjen uvjet (3.8). Fiksirajmo jednu vrijednost $\theta_0 \in \Theta$ parametra modela. Tada za $\theta \neq \theta_0$ i $\mathbf{x} \in S^n$ iz (3.8) vrijedi da je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \psi_{\theta,\theta_0}(t(\mathbf{x})) \cdot f_{\mathbf{X}}(\mathbf{x};\theta_0). \tag{3.9}$$

Stavimo da je:

$$\begin{aligned} g_\theta(y) &:= \begin{cases} \psi_{\theta,\theta_0}(y), & y \in t(S^n) \text{ i } \theta \neq \theta_0 \\ 1, & y \notin t(S^n) \text{ ili } \theta = \theta_0, \end{cases} \\ h(\mathbf{x}) &:= f_{\mathbf{X}}(\mathbf{x};\theta_0),\ \mathbf{x} \in \mathbb{R}^{nd}. \end{aligned}$$

Tada je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \left\{ \begin{array}{ll} \psi_{\theta,\theta_0}(t(\mathbf{x})) \cdot f_{\mathbf{X}}(\mathbf{x};\theta_0) & \text{za } \mathbf{x} \in S^n \text{ i } \theta \neq \theta_0 \text{ prema (3.9)} \\ f_{\mathbf{X}}(\mathbf{x};\theta_0) = 1 \cdot f_{\mathbf{X}}(\mathbf{x};\theta_0) & \text{za } \mathbf{x} \in S^n \text{ i } \theta = \theta_0 \\ 0 = g_\theta(t(\mathbf{x})) \cdot f_{\mathbf{X}}(\mathbf{x};\theta_0) & \text{za } \mathbf{x} \in \mathbb{R}^{nd} \setminus S^n \end{array} \right\} = g_\theta(t(\mathbf{x})) \cdot h(\mathbf{x})$$

za svaki $\mathbf{x} \in \mathbb{R}^d$ i $\theta \in \Theta$. Prema obratu teorema o faktorizaciji 3.9 slijedi da je $T = t(\mathbf{X})$ dovoljna statistika za θ. □

Među svim dovoljnim statistikama razumno je izabrati onu statistiku koja je, u nekom smislu, optimalne forme.

Primjer 3.14. Neka je $\mathbf{X} = (X_1, X_2, X_3)$ slučajni uzorak duljine 3 iz Bernoullijevog modela $b(1,\theta)$, $\theta \in \langle 0, 1\rangle$. Očito je sam slučajni uzorak $\mathbf{X}$ jedna dovoljna statistika za θ, a iz primjera 3.7 slijedi da je dovoljna statistika za θ i ukupan broj uspjeha $T = X_1 + X_2 + X_3$. Nije teško pokazati da je i $S = (X_1, X_2 + X_3)$ dovoljna statistika za θ. Osim što je T dovoljna statistika za θ najmanje dimenzije među tri navedene dovoljne statistike, ona se može dobiti i kao *funkcija* statistika S i $\mathbf{X}$. Tako je $T = t(\mathbf{X})$ za $t(x_1, x_2, x_3) = x_1 + x_2 + x_3$ i $T = \phi(S)$ za $\phi(z_1, z_2) = z_1 + z_2$. Može se pokazati i više, da je T *funkcija bilo koje druge dovoljne statistike za* θ. □

Za dovoljnu statistiku koja je funkcija svake druge dovoljne statistike kažemo da je *minimalna*. Navedimo preciznu definiciju.

Definicija 3.15. *Statistika T je* **minimalna dovoljna statistika za** θ *ako je dovoljna za θ i za svaku drugu dovoljnu statistiku S za θ postoji izmjeriva funkcija ϕ takva da je $T = \phi(S)$.*

Primjer 3.16. Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajni uzorak duljine n iz normalnog modela $N(\theta, 1)$ s parametrom očekivanja $\theta \in \mathbb{R}$. Dakle, gustoća od $\mathbf{X}$ je jednaka:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}; \theta) &= \prod_{i=1}^{n} f(x_i; \theta) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2} = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot \exp(-\frac{1}{2}\sum_{i=1}^{n} x_i^2) \cdot \exp(\theta \sum_{i=1}^{n} x_i - \frac{n}{2}\theta^2), \ \ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n. \end{aligned}$$

Stavimo li da je:

$$\begin{aligned} t(\mathbf{x}) &:= x_1 + x_2 + \cdots + x_n, \ \ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n, \\ h(\mathbf{x}) &:= \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2}|\mathbf{x}|^2}, \ \ \mathbf{x} \in \mathbb{R}^n, \\ g_\theta(y) &:= e^{\theta y - \frac{n}{2}\theta^2}, \ \ y \in t(\mathbb{R}^n), \end{aligned}$$

tada je

$$f_{\mathbf{X}}(\mathbf{x}; \theta) = h(\mathbf{x}) \cdot g_\theta(t(\mathbf{x})), \ \ \mathbf{x} \in \mathbb{R}^n.$$

Prema teoremu o faktorizaciji 3.9, $T = t(\mathbf{X}) = X_1 + X_2 + \cdots + X_n$ je dovoljna statistika.

Neka je sada $S = s(\mathbf{X})$ proizvoljna dovoljna statistika za θ, recimo dimenzije $\ell \geq 1$. Pokažimo da postoji izmjeriva funkcija $\phi : \mathbb{R}^\ell \to \mathbb{R}$ takva da je $T = \phi(S)$.

Budući da je $S = s(\mathbf{X})$ dovoljna statistika za θ, prema teoremu o faktorizaciji 3.9 vrijedi:

$$f_{\mathbf{X}}(\mathbf{x}; \theta) = \tilde{h}(\mathbf{x}) \cdot \tilde{g}_\theta(s(\mathbf{x})), \ \ \mathbf{x} \in \mathbb{R}^n. \tag{3.10}$$

Neka su sada $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ dva proizvoljna vektora takva da je $s(\mathbf{x}) = s(\mathbf{y})$. Prvo, za svaki $\theta \in \Theta$ vrijedi da je

$$\frac{f_{\mathbf{X}}(\mathbf{x}; \theta)}{f_{\mathbf{X}}(\mathbf{y}; \theta)} = e^{-\frac{1}{2}(|\mathbf{x}|^2 - |\mathbf{y}|^2)} \cdot e^{\theta(t(\mathbf{x}) - t(\mathbf{y}))}. \tag{3.11}$$

S druge strane je, zbog pretpostavke da je $s(\mathbf{x}) = s(\mathbf{y})$:

$$\frac{f_{\mathbf{X}}(\mathbf{x}; \theta)}{f_{\mathbf{X}}(\mathbf{y}; \theta)} \overset{(3.10)}{=} \frac{\tilde{h}(\mathbf{x}) \cdot \tilde{g}_\theta(s(\mathbf{x}))}{\tilde{h}(\mathbf{y}) \cdot \tilde{g}_\theta(s(\mathbf{y}))} \overset{s(\mathbf{x}) = s(\mathbf{y})}{=} \frac{\tilde{h}(\mathbf{x})}{\tilde{h}(\mathbf{y})}.$$

Zaključujemo da kvocijent $f_{\mathbf{X}}(\mathbf{x}; \theta)/f_{\mathbf{X}}(\mathbf{y}; \theta)$ ne ovisi o θ. Iz izraza (3.11) za taj kvocijent slijedi da je tada nužno $t(\mathbf{x}) = t(\mathbf{y})$. Dakle, pokazali smo da vrijedi sljedeća tvrdnja:

$$(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n) \ s(\mathbf{x}) = s(\mathbf{y}) \ \Rightarrow \ t(\mathbf{x}) = t(\mathbf{y}). \tag{3.12}$$

Definirajmo sada funkciju $\phi : \mathbb{R}^\ell \to \mathbb{R}$ na sljedeći način. Ako je $\mathbf{z} = (z_1, z_2, \dots, z_\ell) \in s(\mathbb{R}^n)$, dakle, u slici preslikavanja s, tada postoji $\mathbf{x} \in \mathbb{R}^n$ takav da je $\mathbf{z} = s(\mathbf{x})$. Stavimo tada da je $\phi(\mathbf{z}) := t(\mathbf{x})$. Broj $\phi(\mathbf{z})$ je jednoznačno definiran. Naime, ako je $\mathbf{y} \in \mathbb{R}^n$ neki drugi vektor takav da je $\mathbf{z} = s(\mathbf{y}) = s(\mathbf{x})$, tada je, zbog (3.12), nužno $t(\mathbf{y}) = t(\mathbf{x}) = \phi(\mathbf{z})$. S druge strane, ako je $\mathbf{z} \in \mathbb{R}^\ell \setminus s(\mathbb{R}^n)$, tada stavimo da je $\phi(\mathbf{z}) := 0$. Za tako definirano preslikavanje vrijedi

$$\phi(S) = \phi(s(\mathbf{X})) \overset{\text{po def.}}{=} t(\mathbf{X}) = T.$$

Dakle, T je dovoljna statistika za θ za koju vrijedi da se može prikazati kao funkcija svake druge dovoljne statistike za θ. Prema definiciji 3.15, T je minimalna dovoljna statistika za θ. $\square$

Statistički model je *konačan* ako se sastoji od konačno mnogo funkcija gustoće:

$$\mathcal{P} = \{f_0, f_1, \ldots, f_m\}.$$

U tom slučaju ulogu parametra modela ima indeks j gustoća $f_j \in \mathcal{P}$, a parametarski prostor je skup vrijednosti indeksa $J = \{0, 1, \ldots, m\}$. Za dovoljnu statistiku u modelu $\mathcal{P}$ u kojemu nije spacificiran parametar reći ćemo da je *dovoljna statistika za model* $\mathcal{P}$.

Ako su nosači svih gustoća iz konačnog statističkog modela jednaki, tada se minimalna dovoljna statistika za taj model može eksplicitno konstruirati.

Teorem 3.17. *Neka je* $\mathcal{P} = \{f_0, f_1, \ldots, f_m\}$ *konačan statistički model takav da su nosači svih gustoća iz modela jednaki, te neka je* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *slučajni uzorak duljine* n *iz tog modela. Tada je statistika*

$$T = t(\mathbf{X}) = \left(\prod_{i=1}^{n} \frac{f_1(X_i)}{f_0(X_i)}, \prod_{i=1}^{n} \frac{f_2(X_i)}{f_0(X_i)}, \ldots, \prod_{i=1}^{n} \frac{f_m(X_i)}{f_0(X_i)} \right)$$

minimalna dovoljna za $\mathcal{P}$.

Dokaz. Pokažimo prvo da je T dovoljna statistika. Budući da u modelu $\mathcal{P}$ nosači gustoća ne ovise o indeksu $j \in J = \{0, 1, \ldots, m\}$, u tu svrhu koristit ćemo karakterizaciju dovoljnosti iz korolara 3.13. Primijetite da je gustoća slučajnog uzorka tada također indeksirana sa j na sljedeći način:

$$f_{\mathbf{X}}(\mathbf{x}; j) = \prod_{i=1}^{n} f_j(x_i), \quad \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn}.$$

Neka je indeks $j \neq 0$. Tada je za $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ iz zajedničkog nosača S^n gustoća slučajnog uzorka $\mathbf{X}$:

$$\begin{aligned} \frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; 0)} &= \frac{\prod_{i=1}^{n} f_j(x_i)}{\prod_{i=1}^{n} f_0(X_i)} = \pi_j(t(\mathbf{x})) \\ \frac{f_{\mathbf{X}}(\mathbf{x}; 0)}{f_{\mathbf{X}}(\mathbf{x}; j)} &= \frac{\prod_{i=1}^{n} f_0(x_i)}{\prod_{i=1}^{n} f_j(X_i)} = \frac{1}{\pi_j(t(\mathbf{x}))} \equiv \frac{1}{\pi_j}(t(\mathbf{x})) \end{aligned}$$

gdje je $\pi_j : \mathbb{R}^m \to \mathbb{R}$ projekcija na j-tu koordinatu: $\pi_j(z_1, z_2, \ldots, z_m) = z_j$.
Neka su sada $j \neq k$ indeksi takvi da $j \neq 0$ i $k \neq 0$. Tada je za $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in S^n$:

$$\frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; k)} = \frac{\prod_{i=1}^{n} f_j(x_i)}{\prod_{i=1}^{n} f_k(x_i)} = \frac{\prod_{i=1}^{n} f_j(x_i) / \prod_{i=1}^{n} f_0(x_i)}{\prod_{i=1}^{n} f_k(x_i) / \prod_{i=1}^{n} f_0(x_i)} = \frac{\pi_j(t(\mathbf{x}))}{\pi_k(t(\mathbf{x}))} \equiv \frac{\pi_j}{\pi_k}(t(\mathbf{x}))$$

Pokazali smo da za indekse $j, k \in J$ takve da je $j \neq k$ postoji funkcija $\psi_{j,k} : \mathbb{R}^m \to \mathbb{R}$ takva da je

$$\frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; k)} = \psi_{j,k}(t(\mathbf{x})), \quad \mathbf{x} \in S^n.$$

Funkcija $\psi_{j,k}$ je definirana sa:

$$\psi_{j,k}(\mathbf{z}) := \begin{cases} \pi_j(\mathbf{z}), & j \neq 0 \text{ i } k = 0 \\ \frac{1}{\pi_k}(\mathbf{z}), & j = 0 \text{ i } k \neq 0 \\ \frac{\pi_j}{\pi_k}(\mathbf{z}), & j \neq 0 \text{ i } k \neq 0, \end{cases}$$

za $\mathbf{z} \in \mathbb{R}^m$. Prema korolaru 3.13, $T = t(\mathbf{X})$ je dovoljna statistika za $\mathcal{P}$.

Neka je $S = s(\mathbf{X})$ neka druga dovoljna statistka za $\mathcal{P}$, recimo dimenzije $\ell \geq 1$. Tada, po korolaru 3.13, za indeks $j \neq 0$ postoji funkcija $\phi_{j,0} : \mathbb{R}^\ell \to \mathbb{R}$ takva da je za $\mathbf{x} \in S^n$:

$$\phi_{j,0}(s(\mathbf{x})) = \frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; 0)}.$$

S druge strane, za isti $\mathbf{x}$ i indeks $j \neq 0$ je:

$$\frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; 0)} = \pi_j(t(\mathbf{x})).$$

Prema tome je:

$$\phi_{j,0}(s(\mathbf{x})) = \frac{f_{\mathbf{X}}(\mathbf{x}; j)}{f_{\mathbf{X}}(\mathbf{x}; 0)} = \pi_j(t(\mathbf{x})).$$

Dakle, ako stavmo da je $\phi : \mathbb{R}^\ell \to \mathbb{R}^m$ funkcija kojoj su komponente $\phi_{j,0}$ za $j = 1, 2, \ldots, m$, tj. $\phi := (\phi_{1,0}, \phi_{2,0}, \ldots, \phi_{m,0})$, tada je

$$\begin{aligned}
\phi(S) &= \phi(s(\mathbf{X})) = (\phi_{1,0}(s(\mathbf{X})), \phi_{2,0}(s(\mathbf{X})), \ldots, \phi_{m,0}(s(\mathbf{X}))) = \\
&= (\pi_1(t(\mathbf{X})), \pi_2(t(\mathbf{X})), \ldots, \pi_m(t(\mathbf{X}))) = t(\mathbf{X}) = \\
&= T.
\end{aligned}$$

Budući da je T dovoljna statistika za $\mathcal{P}$ i može se prikazati kao funkcija svake dovoljne statistike za isti model, T je minimalna dovoljna statistika za $\mathcal{P}$. □

3.3 Potpuna statistika

U statističkom zaključivanju važnu ulogu imaju *potpune statistike*.

Definicija 3.18. *Statistika T u statističkom modelu parametriziranom parametrom θ je* **potpuna statistika** *ako za svaku Borelovu realnu funkciju g takvu da je za sve $\theta \in \Theta$, $\mathbb{E}_\theta[g(T)] = 0$, nužno slijedi da je $g(T) = 0$ $\mathbb{P}_\theta$-g.s. za sve $\theta \in \Theta$.*

Važno svojstvo potpunosti je invarijantnost na izmjerive transformacije.

Propozicija 3.19. *Neka je T potpuna statistika dimenzije k ($k \geq 1$). Ako je $\phi : \mathbb{R}^k \to \mathbb{R}^\ell$ Borelova funkcija ($\ell \geq 1$), tada je $S = \phi(T)$ potpuna statistika.*

Dokaz. Neka je $g : \mathbb{R}^\ell \to \mathbb{R}$ Borelova funkcija takva da je za sve $\theta \in \Theta$, $\mathbb{E}_\theta[g(S)] = 0$. Tada vrijedi i

$$(\forall \theta \in \Theta)\ 0 = \mathbb{E}_\theta[g(S)] = \mathbb{E}_\theta[g(\phi(T))] = \mathbb{E}_\theta[(g \circ \phi)(T)]$$

za funkciju $g \circ \phi : \mathbb{R}^k \to \mathbb{R}$. Primijetite da je $g \circ \phi$ Borelova funkcija kao kompozicija Borelovih funkcija. Budući da je T potpuna statistika, iz definicije 3.18 slijedi da je nužno

$$(\forall \theta \in \Theta)\ 0 = (g \circ \phi)(T) = g(\phi(T)) = g(S)\ \mathbb{P}_\theta - \text{g.s.}.$$

Dakle, prema definiciji 3.18, S je potpuna statistika. □

Primjer 3.20. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine n iz Bernoullijevog modela $b(1, \theta)$ s parametrom $\theta \in \langle 0, 1 \rangle$. Nadalje, neka je $T = X_1 + X_2 + \cdots + X_n$. Pokazat ćemo da je T potpuna statistika. Znamo da je $T \sim b(n, \theta)$. Zbog toga je za svaku izmjerivu funkciju $g : \mathbb{R} \to \mathbb{R}$ i $\theta \in \langle 0, 1 \rangle$.

$$\mathbb{E}_\theta[g(T)] = \sum_{k=0}^{n} g(k) \binom{n}{k} \theta^k (1-\theta)^{n-k}.$$

Neka je sada g takva Borelova funkcija da je za sve $\theta \in \langle 0,1\rangle$, $\mathbb{E}_\theta[g(T)] = 0$, ekvivalentno:

$$
\begin{aligned}
& (\forall\theta \in \langle 0,1\rangle)\ 0 = \sum_{k=0}^{n} g(k)\binom{n}{k}\theta^k(1-\theta)^{n-k} \\
\Leftrightarrow\quad & (\forall\theta \in \langle 0,1\rangle)\ 0 = (1-\theta)^n \sum_{k=0}^{n} g(k)\binom{n}{k}\left(\frac{\theta}{1-\theta}\right)^k \\
\Leftrightarrow\quad & (\forall x > 0)\ 0 = \sum_{k=0}^{n} g(k)\binom{n}{k}x^k
\end{aligned}
$$

za $x = \theta/(1-\theta)$. Primijetite da preslikavanje $\theta \mapsto \theta/(1-\theta)$ bijektivno preslikava interval $\langle 0,1\rangle$ na interval $\langle 0,+\infty\rangle$. Po teoremu o jednakosti polinoma slijedi da je nužno

$$
(\forall k \in \{0,1,\dots,n\})\ \ g(k)\binom{n}{k} = 0 \ \ \Rightarrow\ \ (\forall k \in \{0,1,\dots,n\})\ \ g(k) = 0.
$$

Sada nije teško pokazati da je za svako $\theta \in \langle 0,1\rangle$:

$$
\mathbb{P}_\theta(g(T)=0) = \sum_{g(k)=0} \mathbb{P}_\theta(T=k) = \sum_{k=0}^{n}\binom{n}{k}\theta^k(1-\theta)^{n-k} = 1.
$$

Budući da je $g(T) = 0$ $\mathbb{P}_\theta$-g.s. za sve $\theta \in \langle 0,1\rangle$, po definiciji je T potpuna statistika. □

Statistika T iz prethodnog primjera 3.20 je ujedno i dovoljna statistika. Znači li to da je svaka dovoljna statistika ujedno i potpuna? Odgovor je da nije. Kontraprimjer je naveden u sljedećem primjeru.

Primjer 3.21. (Dovoljna statistika koja nije potpuna)
Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajni uzorak duljine n iz normalnog modela $N(\theta,\theta^2)$ s očekivanjem i standarnom devijacijom jednakim $\theta > 0$. Slično kao i u primjeru 3.10 pokaže se da je $T = (\overline{X}_n, S_n^2)$ dovoljna statistika za θ. Da bismo pokazali da T nije potpuna statistika, treba pronaći barem jednu izmjerivu funkciju $g : \mathbb{R}^2 \to \mathbb{R}$ takvu da je za sve $\theta > 0$, $\mathbb{E}_\theta[g(T)] = 0$, ali $\mathbb{P}_\theta(g(T) \neq 0) > 0$ za neki $\theta > 0$. Stavimo

$$
g(z_1,z_2) := \frac{n(n-1)}{2}z_1^2 - \frac{n^2-1}{2}z_2.
$$

Neka je $\theta > 0$ proizvoljno. Računamo:

$$
\begin{aligned}
\mathbb{E}_\theta[g(T)] &= \mathbb{E}_\theta[\frac{n(n-1)}{2}(\overline{X}_n)^2 - \frac{n^2-1}{2}S_n^2] \overset{\text{lin.m.o.}}{=} \\
&= \frac{n(n-1)}{2}\mathbb{E}_\theta[(\overline{X}_n)^2] - \frac{n^2-1}{2}\mathbb{E}_\theta[S_n^2] = \frac{n(n-1)}{2}\left(\frac{\theta^2}{n}+\theta^2\right) - \frac{n^2-1}{2}\theta^2 = \\
&= \frac{n^2-1}{2}\theta^2 - \frac{n^2-1}{2}\theta^2 = 0.
\end{aligned}
$$

S druge strane je, također za proizvoljno $\theta > 0$:

$$
\mathbb{P}_\theta(g(T)>0) = \mathbb{P}_\theta\left(\frac{n(n-1)}{2}(\overline{X}_n)^2 - \frac{n^2-1}{2}S_n^2 > 0\right) = \mathbb{P}_\theta\left(\left|\frac{\overline{X}_n}{S_n}\right| > \sqrt{\frac{n+1}{n}}\right) > 0
$$

jer statistika $\overline{X}_n/S_n$ ima gustoću (kao kvocijent dviju nezavisnih neprekidnih slučajnih varijabli). Prema tome, T nije potpuna statistika. □

Pokazuje se da, ako je statistika dovoljna i potpune, da je nužno i minimalna dovoljna. Sljedeći teorem navodimo bez dokaza. Za detalje pogledati u [5] i [6].

Teorem 3.22. *Neka je T potpuna i dovoljna statistika za θ. Tada je T minimalna dovoljna statistika za θ.*

3.4 Eksponencijalne familije

Eksponencijalne familije su široka klasa modela koje dopuštaju dovoljne i potpune statistike.

Označimo sa $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ za slučajnu veličinu dimenzije d ($d \geq 1$). Pretpostavimo da je parametarski prostor dimenzije m ($m \geq 1$).

Definicija 3.23. *Model $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ je* **k-parametarska eksponencijalna familija** *ako za gustoće vrijedi da se mogu prikazati u obliku:*

$$f(x;\theta) = C(\theta) \cdot h(x) \cdot \exp\left(\sum_{j=1}^{k} Q_j(\theta) t_j(x)\right) \tag{3.13}$$

gdje su: k prirodan broj, $C : \Theta \to \langle 0, +\infty\rangle$, $h : \mathbb{R}^d \to [0, +\infty\rangle$, $Q_j : \Theta \to \mathbb{R}$ i $t_j : \mathbb{R}^d \to \mathbb{R}$ ($j = 1, \ldots, k$) Borelove funkcije, pri čemu funkcije $\{t_1, \ldots, t_k\}$ nisu konstante i čine linearno nezavisan skup funkcija.

Postoji prirodna reparametrizacija u k-parametarskom eksponencijalnom modelu. Stavimo da je

$$\eta_j := Q_j(\theta),\ j = 1, 2, \ldots, k.$$

Tada je $\eta = (\eta_1, \eta_2, \ldots, \eta_k)$ novi parametar kojeg zovemo *prirodni parametar*. Zapis gustoće (3.13) k-parametarske familije u reparametriziranom obliku pomoću prirodnog parametra:

$$f(x;\eta) = C_0(\eta) \cdot h(x) \cdot \exp\left(\sum_{j=1}^{k} \eta_j t_j(x)\right) \tag{3.14}$$

zovemo *kanonskom formom.* Označimo sada sa $\Sigma \subseteq \mathbb{R}^k$ prostor vrijednosti prirodnog parametra za koji je izrazom (3.14) definirana vjerojatnosna funkcija gustoće. Dakle,

$$\eta \in \Sigma \quad \Leftrightarrow \quad \begin{cases} \int\limits_{\mathbb{R}^d} h(x) \cdot \exp\left(\sum_{j=1}^{k} \eta_j t_j(x)\right) dx < +\infty & \text{ako je model neprekidan} \\ \sum_{x \in \mathbb{R}^d} h(x) \cdot \exp\left(\sum_{j=1}^{k} \eta_j t_j(x)\right) < +\infty & \text{ako je model diskretan.} \end{cases}$$

Skup Σ zovemo *prirodni parametarski prostor.* Nije teško pokazati da je Σ konveksan skup u $\mathbb{R}^k$.

Zadatak 3.24. Neka je Σ neprazan prirodni parametarski prostor za k-parametarsku eksponencijalnu familiju $\mathcal{P}$ s gustoćama oblika (3.14). Dokažite da je Σ konveksan skup u $\mathbb{R}^k$. (**Uputa.** Za $\lambda \in \langle 0, 1\rangle$ postoje brojevi $p > 1$ i $q > 1$ takvi da su $\frac{1}{p} + \frac{1}{q} = 1$ te $\lambda = 1/p$ i $1 - \lambda = 1/q$. Primijeninite Hölderovu nejednakost na integrale (sume ako je model diskretan) iz definicije od Σ s tako odabranim p i q.)

Primijetite da je za $\eta \in \Sigma$ normalizirajuća konstanta $C_0(\eta)$ iz (3.14) jednaka:

$$C_0(\eta) = \left\{ \begin{array}{ll} \left(\int\limits_{\mathbb{R}^d} h(x) \cdot \exp\left(\sum_{j=1}^{k} \eta_j t_j(x)\right) dx\right)^{-1} & \text{ako je model neprekidan} \\ \left(\sum_{x \in \mathbb{R}^d} h(x) \cdot \exp\left(\sum_{j=1}^{k} \eta_j t_j(x)\right)\right)^{-1} & \text{ako je model diskretan} \end{array} \right\} \text{za } \eta \in \Sigma.$$

Označimo sa Q preslikavanje:

$$Q : \Theta \to \mathbb{R}^k,\ \ Q(\theta) := (Q_1(\theta), Q_2(\theta), \ldots, Q_k(\theta)),\ \ \theta \in \Theta, \tag{3.15}$$

gdje su Q_j $(j = 1, 2, \ldots, k)$ preslikavanja iz izraza za gustoću (3.13) k-parametarske eksponencijalne familije. Tada je nužno slika preslikavanja Q sadržana u prirodnom parametarskom prostoru, tj. $Q(\Theta) \subseteq \Sigma$. Nadalje, za konstantu $C(\theta)$ iz (3.13) za $\theta \in \Theta$ takav da je $\eta = Q(\theta)$ nužno vrijedi:

$$C(\theta) = C_0(\eta) > 0.$$

Zbog toga je nosač gustoće (3.13):

$$\mathrm{supp} f(\cdot;\theta) = \mathrm{supp}\, h$$

za sve $\theta \in \Theta$. Dakle, gustoće dane eksponencijalne familije $\mathcal{P}$ imaju jednake nosače (neovisne o θ). Nadalje, primijetite da dimenzije parametra θ i prirodnog parametra η ne moraju biti jednake.

Za nalaženje minimalnih dovoljnih i potpunih statistika u eksponencijalnim familijama, nužno je znati jesu li *punog ranga*.

Definicija 3.25. *Neka je $\mathcal{P}$ k-parametarska eksponencijalna familija zadana gustoćama (3.23) i neka je Q preslikavanje (3.15). Ako slika od Q, $Q(\Theta)$, sadrži neprazni otvoreni pravokutnik dimenzije k, tada je k-parametarska familija $\mathcal{P}$* **punog ranga**.

Primjer 3.26. Primjeri eksponencijalnih familija:

(a) Bernoullijev model $b(1,\theta)$, $\theta \in \langle 0, 1\rangle$. Gustoća je:

$$f(x;\theta) = \theta^x(1-\theta)^{1-x} \cdot \mathbb{1}_{\{0,1\}}(x) = (1-\theta)\cdot \mathbb{1}_{\{0,1\}}(x) \cdot \exp\left(x \log \frac{\theta}{1-\theta}\right),\ x \in \mathbb{R}.$$

Ovdje su $C(\theta) = 1-\theta > 0$, $h(x) = \mathbb{1}_{\{0,1\}}(x)$, $Q(\theta) = Q_1(\theta) = \log(\theta/(1-\theta)$ i $t(x) = t_1(x) = x$ je netrivijalna funkcija, pa je jednočlani skup $\{t\}$ linerano nezavisan. Dakle, radi se o 1-parametarskoj eksponencijalnoj familiji. Preslikavanje Q bijektivno preslikava $\Theta = \langle 0, 1\rangle$ na $\mathbb{R}$. Budući da je slika $Q(\Theta) = \mathbb{R} = \langle -\infty, +\infty\rangle$ neprazni otvoreni interval u $\mathbb{R}$, familija je punog ranga.

(b) Normalni model $N(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \langle 0, +\infty\rangle$. Gustoća je:

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) = \frac{e^{-\frac{\mu^2}{\sigma^2}}}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right),\ x \in \mathbb{R}.$$

Stavimo:

$$C(\theta) = \frac{e^{-\frac{\mu^2}{\sigma^2}}}{\sqrt{2\pi\sigma^2}},\ h(x) \equiv 1,\ Q_1(\theta) = \frac{\mu}{\sigma^2},\ Q_2(\theta) = -\frac{1}{2\sigma^2},\ t_1(x) = x.\ t_2(x) = x^2.$$

Skup nekonstantnih funkcija $\{t_1, t_2\}$ je linearno nezavisan jer se radi o elementarnim polinomima stupnjeva 1 i 2. Dakle, normalni model je 2-parametarska eksponencijalna familija. Preslikavanje Q definirano sa $Q(\theta) = (Q_1(\theta), Q_2(\theta))$ za $\theta \in \Theta = \mathbb{R} \times \langle 0, +\infty\rangle$ bijektivno preslikava Θ na $\mathbb{R} \times \langle -\infty, 0\rangle$. Budući da je slika od Q jedan neprazni otvoreni dvodimenzionalni pravokutnik u $\mathbb{R}^2$, normalni model je punog ranga.

(c) Polinomni model $M(n; p_0, p_1, \ldots, p_k)$, $\theta = (p_1, p_2, \ldots, p_k) \in \Theta$ gdje je

$$\Theta = \{(p_1, p_2, \ldots, p_k) \in \langle 0, 1\rangle^k : p_1 + p_2 + \cdots + p_k < 1\}$$

parametarski prostor. Za $\theta = (p_1, p_2, \ldots, p_k) \in \theta$, označimo sa $p_0 := 1 - (p_1 + \cdots + p_k)$. Gustoća je, za $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ i $x_0 := n - (x_1 + x_2 + \cdots + x_k)$,

$$\begin{aligned} f(x;\theta) &= \frac{n!}{x_0!x_1!\cdots x_k!} p_0^{x_0} p_1^{x_1} \cdots p_k^{x_k} \mathbb{1}_{S_n}(x) = \\ &= p_0^n \cdot \frac{n!}{x_0!x_1!\cdots x_k!} \cdot \mathbb{1}_{S_n}(x) \cdot \exp\left(x_1 \log \frac{p_1}{p_0} + \cdots + x_k \log \frac{p_k}{p_0}\right) \end{aligned}$$

gdje je $S_n = \{(x_1, x_2, \dots, x_k) \in \mathbb{N}_0^k : x_1 + x_2 + \cdots + x_k \leq n\}$. Ovdje su:

$$\begin{aligned}
C(\theta) &= p_0^n = (1 - (p_1 + p_2 + \cdots + p_k))^n, \\
h(x) &= \frac{n!}{x_0! x_1! \cdots x_k!} \cdot \mathbb{1}_{S_n}(x), \\
Q_j(\theta) &= \log \frac{p_j}{p_0} = \frac{p_j}{1 - (p_1 + p_2 + \cdots + p_k)}, \ j = 1, 2, \dots, k, \\
t_j(x) &= x_j = \pi_j(x_1, x_2, \dots, x_k), \ j = 1, 2, \dots, k.
\end{aligned}$$

Budući da su funkcije t_j zapravo koordinatne projekcije π_j ($j = 1, 2, \dots, k$), nekonstantne su i skup $\{t_1, t_2, \dots, t_k\}$ je linearno nezavisan. Dakle, polinomni model je jedna k-parametarska eksponencijalna familija. Preslikavanje $Q = (Q_1, Q_2, \dots, Q_k) : \Theta \to \mathbb{R}^k$ bijektivno preslikava Θ na $\mathbb{R}^k$, pa je model i punog ranga. □

Dovoljnu statistiku u eksponencijalnim familijama možemo jednostavno odrediti.

Teorem 3.27. *Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak duljine n iz k-parametarske familije čije gustoće su dane izrazom (3.13). Tada je*

$$T = \left(\sum_{i=1}^n t_1(X_i), \sum_{i=1}^n t_2(X_i), \dots, \sum_{i=1}^n t_k(X_i) \right) \tag{3.16}$$

dovoljna statistika za θ dimenzije k.

Dokaz. Gustoća slučajnog uzorka $\mathbf{X}$ je:

$$\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}; \theta) &= \prod_{i=1}^n f(x_i; \theta) \overset{(3.13)}{=} \\
&= C(\theta)^n \prod_{i=1}^n h(x_i) \exp\left(\sum_{i=1}^n \sum_{j=1}^k Q_j(\theta) t_j(x_i) \right) = \\
&= C(\theta)^n \cdot \exp\left(\sum_{j=1}^k Q_j(\theta) \left(\sum_{i=1}^n t_j(x_i) \right) \right) \cdot \prod_{i=1}^n h(x_i), \ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^{dn}.
\end{aligned}$$

Ako stavimo da je

$$\begin{aligned}
\tau(\mathbf{x}) &:= \left(\sum_{i=1}^n t_1(x_i), \sum_{i=1}^n t_2(x_i), \dots, \sum_{i=1}^n t_k(x_i) \right), \ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^{dn}, \\
\mathbf{h}(\mathbf{x}) &:= \prod_{i=1}^n h(x_i), \ \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^{dn}, \\
g_\theta(\mathbf{y}) &:= C(\theta)^n \cdot \exp\left(\sum_{j=1}^k Q_j(\theta) y_j \right) = C(\theta)^n e^{(Q(\theta), \mathbf{y})}, \ \mathbf{y} = (y_1, y_2, \dots, y_k) \in \mathbb{R}^k,
\end{aligned}$$

tada je

$$f_{\mathbf{X}}(\mathbf{x}; \theta) = \mathbf{h}(\mathbf{x}) \cdot g_\theta(\tau(\mathbf{x})), \ \mathbf{x} \in \mathbb{R}^{dn},$$

pa je, po teoremu o faktorizaciji 3.9, $T = \tau(\mathbf{X})$ dovoljna statistika za θ. □

Dovoljnu statistiku (3.16) iz teorema 3.27 zovemo **prirodna dovoljna statistika**.

Primjer 3.28. (Krivuljna eksponencijalna familija). Neka je statistički model normalan $N(\theta, \theta^2)$ s parametrom $\theta > 0$ (kao u primjeru 3.21). Gustoća je jednaka:

$$f(x;\theta) = \frac{1}{\theta\sqrt{2\pi}} \cdot e^{-\frac{1}{2\theta^2}(x-\theta)^2} = \frac{1}{\theta\sqrt{2\pi e}} \cdot \exp\left(\frac{1}{\theta}x - \frac{1}{2\theta^2}x^2\right), \ x \in \mathbb{R}.$$

Očito se radi o 2-parametarskoj eksponencijalnoj familiji. Je li punog ranga? Ako prikažemo sliku preslikavanja $Q = (Q_1, Q_2) : \langle 0, +\infty\rangle \to \mathbb{R}^2$, gdje su $Q_1(\theta) = 1/\theta$ i $Q_2(\theta) = -1/(2\theta^2)$, u prostoru prirodnih parametara $\eta = (\eta_1, \eta_2)$, vidimo da se radi o dijelu parabole. Naime, budući da je

$$\eta_1 = \frac{1}{\theta} > 0 \ \& \ \eta_2 = -\frac{1}{2\theta^2} \quad \Rightarrow \quad \eta_2 = -\frac{1}{2}\left(\frac{1}{\theta}\right)^2 = -\frac{1}{2}\eta_1^2 \ \& \ \eta_1 > 0$$

$Q(\langle 0, +\infty\rangle)$ je dio parabole $\eta_2 = -\frac{1}{2}\eta_1^2$ koji se nalazi u (otvorenom) IV. kvadrantu. Dakle, $Q(\Theta) = Q(\langle 0, +\infty\rangle)$ *ne sadrži* neprazni otvoreni pravokutnik u $\mathbb{R}^2$ pa taj model nije punog ranga. Ovo je ujedno i primjer kada je dimenzija parametarskog prostora manja od dimenzije prirodnog parametarskog prostora. □

Navedimo dovoljan uvjet kada će prirodna dovoljna statistika biti minimalna dovoljna. Prije glavnog teorema, dokažimo potrebnu lemu.

Lema 3.29. *Neka su $\mathcal{P}$ i $\mathcal{P}_0$ statistički modeli takvi da je $\mathcal{P}_0 \subset \mathcal{P}$. Ako je T minimalna dovoljna statistika za $\mathcal{P}_0$ i dovoljna statistika za $\mathcal{P}$, tada je T minimalna dovoljna statistika za $\mathcal{P}$.*

Dokaz. Budući da je T dovoljna statistika za $\mathcal{P}$, dovoljno je pokazati da je T funkcija svake druge dovoljne statistike za $\mathcal{P}$. Neka je S bilo koja dovoljna statistika za $\mathcal{P}$. Budući da je $\mathcal{P}_0$ podmnožina od $\mathcal{P}$, S je dovoljna statistika i za $\mathcal{P}_0$. T je minimalna dovoljna statistika za $\mathcal{P}_0$ pa postoji funcija ϕ takva da je $T = \phi(S)$. Dakle, T je minimalna dovoljna i za $\mathcal{P}$. □

Neka su gustoće k-parametarske eksponencijalne familije dane sa (3.13). Podsjetimo se preslikavanja $Q : \Theta \to \mathbb{R}^k$ takvog da je

$$Q(\theta) = (Q_1(\theta), Q_2(\theta), \ldots, Q_k(\theta)), \ \ \theta \in \Theta.$$

Teorem 3.30. *Neka je T prirodna dovoljna statistika k-parametarske eksponencijalne familije $\mathcal{P}$ parametrizirane s θ kojoj su gustoće dane sa (3.13). Tada je T minimalna dovoljna statistika za θ ako je ispunjen jedan od sljedećih uvjeta:*

(1) *Model $\mathcal{P}$ je punog ranga;*
(2) *Slika preslikavanja Q sadrži $k+1$ točku koje razapinju $\mathbb{R}^k$.*

Dokaz. Prvo primijetimo da (1) povlači (2). Naime, ako $Q(\Theta)$ sadrži neprazni otvoreni pravokutnik dimenzije k, tada je moguće pronaći $k+1$ točku:

$$\eta_j = Q(\theta_j), \ \ j = 0, 1, 2, \ldots, k, \tag{3.17}$$

takve da je skup vektora u $\mathbb{R}^k$, $L = \{\eta_1 - \eta_0, \eta_2 - \eta_0, \ldots, \eta_k - \eta_0\}$, linearno nezavisan. Tada je L i baza za $\mathbb{R}^k$, pa točke (3.17) *razapinju* $\mathbb{R}^k$. Dakle, pretpostavimo da vrijedi uvjet (2).

Neka je skup L linearno nezavisan za dane točke (3.17). Nadalje, neka je $\mathcal{P}_0$ konačna podmnožina od $\mathcal{P}$ definirana pomoću tih točaka na sljedeći način:

$$\mathcal{P}_0 = \{f(\cdot;\theta_0), f(\cdot;\theta_1), f(\cdot;\theta_2), \ldots, f(\cdot;\theta_k)\}.$$

Prema teoremu 3.17 (kojeg možemo primijeniti na konačan model $\mathcal{P}_0$ jer gustoće imaju jednake nosače), minimalna dovoljna statistika za $\mathcal{P}_0$ je:

$$T_0 = (T_{01}, T_{02}, \ldots, T_{0k}) = \left(\prod_{i=1}^{n} \frac{f(X_i; \theta_1)}{f(X_i; \theta_0)}, \prod_{i=1}^{n} \frac{f(X_i; \theta_2)}{f(X_i; \theta_0)}, \ldots, \prod_{i=1}^{n} \frac{f(X_i; \theta_k)}{f(X_i; \theta_0)} \right).$$

Sada je za svaki $j = 1, 2, \ldots, k$, uzimajući u obzir ekvivalentnost zapisa (3.14) i (3.13):

$$\begin{aligned} T_{0j} &= \left(\frac{C(\theta_j)}{C(\theta_0)} \right)^n \exp\left(\sum_{i=1}^{n} \sum_{\ell=1}^{k} (Q_\ell(\theta_j) - Q_\ell(\theta_0)) t_\ell(X_i) \right) = \\ &= \left(\frac{C_0(\eta_j)}{C_0(\eta_0)} \right)^n \exp\left(\sum_{\ell=1}^{k} (\eta_{\ell j} - \eta_{\ell 0}) T_\ell \right), \end{aligned}$$

gdje je $T = (T_1, T_2, \ldots, T_k)$ prirodna dovoljna statistika. Tada je to ekvivalentno sa

$$\sum_{\ell=1}^{k} (\eta_{\ell j} - \eta_{\ell 0}) T_\ell = \log T_{0j} - n \log \frac{C_0(\eta_j)}{C_0(\eta_0)}, \quad 1 \leq \ell \leq k. \tag{3.18}$$

Neka je sada A matrica čiji stupci su linearno nezavisni vektori iz skupa L:

$$A = [A_{\ell j}] = [\eta_{\ell j} - \eta_{\ell 0}],$$

a $\psi : \langle 0, +\infty \rangle^k \to \mathbb{R}^k$ funkcija definirana sa:

$$\begin{aligned} & \psi(y_1, y_2, \ldots, y_k) := \\ = & \left(\log y_1 - n \log \frac{C_0(\eta_1)}{C_0(\eta_0)}, \log y_2 - n \log \frac{C_0(\eta_2)}{C_0(\eta_0)}, \ldots, \log y_k - n \log \frac{C_0(\eta_k)}{C_0(\eta_0)} \right)^\tau, \end{aligned}$$

za $(y_1, y_2, \ldots, y_k) \in \mathbb{R}^k$. Primijetite da je ψ bijekcija. Tada se (3.18) može zapisati u matričnom obliku:

$$AT = \psi(T_0) \quad \Leftrightarrow \quad T_0 = \psi^{-1}(AT).$$

Zbog linearne nezavisnosti vektora u L, A je regularna matrica pa je:

$$T = A^{-1} \psi(T_0).$$

Odavde slijedi da su T i T_0 ekvivalentne statistike, pa, budući da je T_0 dovoljna statistika za $\mathcal{P}_0$, prema korolaru 3.12 T je dovoljna za $\mathcal{P}_0$. Neka je S bilo koja dovoljna statistika za $\mathcal{P}_0$. Budući da je T_0 minimalna dovoljna statistika za $\mathcal{P}_0$, postoji funkcija ϕ takva da je $T_0 = \phi(S)$. Dakle,

$$T = A^{-1} \psi(T_0) = A^{-1} \psi(\phi(S)) = (A^{-1} \circ \psi \circ \phi)(S).$$

Budući da je kompozicija izmjerivih funkcija izmjeriva, T je funkcija od S. Dakle, T je dovoljna statistika za $\mathcal{P}_0$ i funkcija je svake druge dovoljne statistika za $\mathcal{P}_0$. Slijedi da je T minimalna dovoljna statistika za $\mathcal{P}_0$. Budući da je, prema teoremu 3.27, T dovoljna statistika za $\mathcal{P}$ i minimalna je dovoljna statistika za $\mathcal{P}_0$, prema lemi 3.29 je minimalna dovoljna za $\mathcal{P}$. $\square$

Navedimo i dovoljan uvjet da je prirodna dovoljna statistika potpuna.

Teorem 3.31. *Prirodna dovoljna statistika k-parametarske eksponencijalne familije punog ranga je potpuna.*

Dokaz. Neka je $T = (T_1, T_2, \dots, T_k) = \tau(\mathbf{X})$ prirodna dovoljna statistika i neka je $g : \mathbb{R}^k \to \mathbb{R}$ proizvoljna izmjeriva funkcija takva da vrijedi

$$(\forall \theta \in \Theta) \ \ \mathbb{E}_\theta[g(T)] = 0. \tag{3.19}$$

Treba dokazati da je nužno $g(T) = 0$ $\mathbb{P}_\theta$-g.s. za sve $\theta \in \Theta$.

Pretpostavimo da je gustoća iz modela $\mathcal{P}$ dana u obliku (3.13) pri čemu je odgovarajući kanonski oblik (3.14). Neka je Θ parametarski prostor, a Σ prirodni parametarski prostor. Budući da je model punog ranga, postoji neprazni otvoreni pravokutnik I dimenzije k takav da je $I \subseteq Q(\Theta) \subset \Sigma$. Bez smanjenja općenitosti možemo pretpostaviti da je

$$I = \langle -a_1, a_1 \rangle \times \langle -a_2, a_2 \rangle \times \cdots \times \langle -a_k, a_k \rangle,$$

za $a_i > 0$ $(i = 1, \dots, k)$. Naime, translacijom funkcije Q za fiksni vektor $\mathbf{b} = (b_1, b_2, \dots, b_k)$ dobijemo ekvivalentnu formu gustoće:

$$f(x;\theta) = C(\theta)h(x)\exp\left(\sum_{j=1}^k Q_j(\theta)t_j(x)\right) = C(\theta)\tilde{h}(x)\exp\left(\sum_{j=1}^k \tilde{Q}_j(\theta)t_j(x)\right)$$

gdje su

$$\begin{aligned}
\tilde{Q}(\theta) &= Q(\theta) - \mathbf{b} = (Q_1(\theta) - b_1, Q_2(\theta) - b_2, \dots, Q_k(\theta) - b_k), \ \theta \in \Theta \\
\tilde{h}(x) &= h(x)\exp\left(\sum_{j=1}^k b_j t_j(x)\right), \ x \in \mathbb{R}^d.
\end{aligned}$$

Odgovarajućim odabirom vektora $\mathbf{b}$ moguće je postići da je $\mathbf{0} \in I$.

Prvo, pretpostavimo da je model $\mathcal{P}$ neprekidan. Budući da smo pretpostavili da vrijedi (3.19), zbog $I \subseteq Q(\Theta) \subseteq \Sigma$ specijalno vrijedi da je

$$\begin{aligned}
& (\forall \eta \in I) \ \ \mathbb{E}_\eta[g(T)] = 0 \\
\Leftrightarrow \ \ & (\forall \eta \in I) \ \ \int_{\mathbb{R}^{nd}} g(\tau(\mathbf{x}))\mathbf{h}(\mathbf{x})\exp\left(\sum_{j=1}^k \eta_j \tau_j(\mathbf{x})\right) d\mathbf{x} = 0 \\
\Leftrightarrow \ \ & (\forall \eta \in I) \ \ \int_{\mathbb{R}^k} g(\mathbf{s})\exp\left(\sum_{j=1}^k \eta_j s_j\right) d\mu(\mathbf{s}) = 0 \\
\Leftrightarrow \ \ & (\forall \eta \in I) \ \ \int_{\mathbb{R}^k} g^+(\mathbf{s})\exp\left(\sum_{j=1}^k \eta_j s_j\right) d\mu(\mathbf{s}) = \int_{\mathbb{R}^k} g^-(\mathbf{s})\exp\left(\sum_{j=1}^k \eta_j s_j\right) d\mu(\mathbf{s}) \qquad (3.20)
\end{aligned}$$

pri čemu smo u trećoj ekvivalentnoj izreci koristili zamjenu varijabli $(s_1, s_2, \dots, s_k) = \mathbf{s} = \tau(\mathbf{x})$, a u četvrtoj da se Borelova funkcija g može prikazati kao razlika $g = g^+ - g^-$ dviju nenegativnih izmjerivih funkcija g^+, g^-. Nova mjera po kojoj integriramo nakon zamjene varijabli je inducirana mjera

$$\mu(B) = \int_{\tau^{-1}(B)} \mathbf{h}(\mathbf{x})\, d\mathbf{x}, \ \ B \in \mathcal{B}(\mathbb{R}^k).$$

Ako u (3.20) stavimo $\eta = (\eta_1, \dots, \eta_k) = \mathbf{0}$, tada dobijamo zajednički broj

$$A := \int_{\mathbb{R}^k} g^+(\mathbf{s})\, d\mu(\mathbf{s}) = \int_{\mathbb{R}^k} g^-(\mathbf{s})\, d\mu(\mathbf{s}).$$

Ako je $A = 0$, tada su $g^+ = g^- = 0$ μ-s.s. Tada je očito i $g = 0$ μ-s.s., iz čega slijedi da je

$$\mathbb{P}_\eta(g(T) = 0) = \int\limits_{g^{-1}(0)} \exp\left(\sum_{j=1}^{k} \eta_j s_j\right) d\mu(\mathbf{s}) = \int\limits_{\mathbb{R}^k} \exp\left(\sum_{j=1}^{k} \eta_j s_j\right) d\mu(\mathbf{s}) = 1$$

za sve $\eta \in \Sigma$, pa specijalno i za sve $\theta \in \Theta$.
Neka je $A > 0$. Tada možemo normirati funkcije g^+ i g^- tako da vrijedi

$$\int\limits_{\mathbb{R}^k} g^+(\mathbf{s})\, d\mu(\mathbf{s}) = \int\limits_{\mathbb{R}^k} g^-(\mathbf{s})\, d\mu(\mathbf{s}) = 1.$$

U tom slučaju nenegativne izmjerive funkcije g^+ i g^- možemo interpretirati kao vjerojatnosne gustoće u odnosu na mjeru μ. Prema tome, svaka od tih dviju funkcija definira po jednu vjerojatnosnu mjeru na izmjerivom prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ formulom

$$\nu_f(B) := \int\limits_{B} f(\mathbf{s})\, d\mu(\mathbf{s}), \quad B \in \mathcal{B}(\mathbb{R}^k),$$

za $f = g^+$ ili $f = g^-$. Nadalje, može se pokazati da iz (3.20) slijedi (vidjeti Theorem 9 u §2, u [5]):

$$(\forall \xi_1, \xi_2, \ldots, \xi_k \in \mathbb{R}) \quad \int\limits_{\mathbb{R}^k} \exp(i \sum_{j=1}^{k} \xi_j s_j) g^+(\mathbf{s})\, d\mu(\mathbf{s}) = \int\limits_{\mathbb{R}^k} \exp(i \sum_{j=1}^{k} \xi_j s_j) g^-(\mathbf{s})\, d\mu(\mathbf{s}). \tag{3.21}$$

Ako za svaku vjerojatnost ν na $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ definiramo kompleksnu funkciju $\phi_\nu : \mathbb{R}^k \to \mathbf{C}$ na sljedeći način:

$$\phi_\nu(\xi_1, \xi_2, \ldots \xi_k) := \int\limits_{\mathbb{R}^k} \exp(i \sum_{j=1}^{k} \xi_j s_j)\, d\nu(\mathbf{s}), \tag{3.22}$$

tada je (3.21) ekvivalentno sa:

$$(\forall \xi_1, \xi_2, \ldots, \xi_k \in \mathbb{R}) \;\; \phi_{\nu_{g^+}}(\xi_1, \xi_2, \ldots, \xi_k) = \phi_{\nu_{g^-}}(\xi_1, \xi_2, \ldots, \xi_k). \tag{3.23}$$

Funkcija (3.22) zove se *karakteristična funkcija* od ν. Prema *teoremu jedinstvenosti* (vidjeti teorem 13.2 u [9], str. 445.), preslikavanje $\nu \mapsto \phi_\nu$ je injekcija. Dakle, prema tom teoremu, iz (3.23) slijedi da je $\nu_{g^+} = \nu_{g^-}$, tj. vrijedi da je

$$(\forall B \in \mathcal{B}(\mathbb{R}^k)) \quad \int\limits_{B} g^+(\mathbf{s})\, d\mu(\mathbf{s}) = \int\limits_{B} g^-(\mathbf{s})\, d\mu(\mathbf{s}).$$

Dakle, $g^+ = g^-$ μ-s.s., pa je $g(T) = 0$ $\mathbb{P}_\theta$-g.s. za sve $\theta \in \Theta$. Dakle, T je potpuna statistika.
Analogno se dokaže tvrdnja za slučaj diskretne familije. □

Sve eksponencijalne familije iz primjera 3.26 su punog ranga, pa su njihove prirodne dovoljne statistike ujedno i minimalne dovoljne i potpune. U slučaju krivuljne eksponencijalne familije iz primjera 3.28, prirodna dovoljna statistika je minimalna dovoljna, ali ne možemo zaključiti da je potpuna.

Za kraj ovoga potpoglavlja, navedimo primjer važne distribucije koja nije eksponencijalna familija.

Primjer 3.32. Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajni uzorak duljine n iz neprekidnog $U(0,\theta)$-modela s parametrom $\theta > 0$. Gustoća je:

$$f(x;\theta) = \frac{1}{\theta} \cdot \mathbb{1}_{[0,\theta]}(x), \quad x \in \mathbb{R}.$$

Uniformni model $U(0,\theta)$ nije eksponencijalna familija (*Zašto?*).

Nađimo jednu dovoljnu statistiku. Gustoća slučajnog uzorka je:

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot \mathbb{1}_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \cdot \mathbb{1}_{[x_{(n)},+\infty\rangle}(\theta), \quad \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n.$$

Ako stavimo da je $g_\theta(y) = (1/\theta^n)\mathbb{1}_{[y,+\infty\rangle}(\theta)$ i $h(\mathbf{x}) \equiv 1$, slijedi da je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = g_\theta(x_{(n)})h(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n.$$

Prema teoremu o faktorizaciji 3.9, dovoljna statistika je $T = X_{(n)} = \max_{1\le i\le n} X_i$. Potpunost te statistike ispitat ćemo po definiciji.

Prvo, odredimo zakon razdiobe od T. Za $\theta \in \Theta = \langle 0,+\infty\rangle$, funkcija distribucije od T je:

$$\begin{aligned} F_T(x;\theta) &= \mathbb{P}_\theta(X_{(n)} \le x) = \mathbb{P}_\theta(X_1 \le x, X_2 \le x, \dots, X_n \le x) \overset{\text{nez.}}{=} \\ &= \mathbb{P}_\theta(X_1 \le x)\mathbb{P}_\theta(X_2 \le x)\cdots\mathbb{P}_\theta(X_n \le x) \overset{\text{jed. dist.}}{=} \\ &= F(x;\theta)^n = \begin{cases} 0 & x \le 0 \\ \frac{x^n}{\theta^n}, & 0 < x < \theta \\ 1, & x \ge \theta. \end{cases} \end{aligned}$$

Ta funkcija je derivabilna svuda osim u točkama $x = 0$ i $x = \theta$. Budući da je

$$F_T'(x;\theta) = \begin{cases} 0, & x < 0 \text{ ili } x > \theta \\ \frac{n}{\theta^n}x^{n-1}, & 0 < x < \theta, \end{cases}$$

T je neprekidna statistika s gustoćom

$$f_T(x;\theta) = \begin{cases} 0, & x \le 0 \text{ ili } x \ge \theta \\ \frac{n}{\theta^n}x^{n-1}, & 0 < x < \theta, \end{cases}$$

Neka je sada $g : \mathbb{R} \to \mathbb{R}$ bilo koja Borelova funkcija takva da je:

$$(\forall\theta > 0) \;\; 0 = \mathbb{E}_\theta[g(T)] = \int_0^\theta g(x)\frac{n}{\theta^n}x^{n-1}dx.$$

Ako g prikažemo kao: $g = g^+ - q^-$, tada se gornji uvjet može napisati u ekvivalentnom obliku:

$$(\forall\theta > 0) \;\; \int_0^\theta g^+(x)x^{n-1}dx = \int_0^\theta g^-(x)x^{n-1}dx.$$

Fiksirajmo sada $\theta > 0$ i označimo sa $A(\theta)$ zajedničku nenegativnu konstantu:

$$A(\theta) = \int_0^\theta g^+(x)x^{n-1}dx = \int_0^\theta g^-(x)x^{n-1}dx.$$

Ako je $A(\theta) = 0$ tada je za λ-s.s. $x \in \langle 0, \theta]$,

$$x^{n-1}g^+(x) = x^{n-1}g^-(x) = 0 \quad \Rightarrow \quad g^+(x) = g^-(x) = 0 \quad \Rightarrow \quad g(x) = 0.$$

Odavde slijedi da je $\langle 0, \theta] \subseteq g^{-1}(0)$ λ-s.s., pa je:

$$\mathbb{P}_\theta(g(T) = 0) = \int\limits_{g^{-1}(0)} f_T(x;\theta)\,dx \geq \int\limits_0^\theta f_T(x;\theta)\,dx = 1.$$

Dakle, $\mathbb{P}_\theta(g(T) = 0) = 1$. Ako je $A(\theta) > 0$, tada su funkcije

$$f_+(x;\theta) := \begin{cases} 0, & x \leq 0 \text{ ili } x \geq \theta, \\ \frac{x^{n-1}}{A(\theta)} \cdot g^+(x), & 0 < x < \theta, \end{cases} \quad f_-(x;\theta) := \begin{cases} 0, & x \leq 0 \text{ ili } x \geq \theta, \\ \frac{x^{n-1}}{A(\theta)} \cdot g^-(x), & 0 < x < \theta \end{cases}$$

vjerojatnosne gustoće s nosačem jednakim $\langle 0, \theta\rangle$. Tada za funkcije distribucije tih dviju razdioba (označimo te razdiobe s μ_+ i μ_-) vrijedi:

$$(\forall x \in \langle 0, \theta]) \;\; F_{\mu_+}(x;\theta) = \int\limits_0^x f_+(y;\theta)\,dy = \int\limits_0^x f_-(y;\theta)\,dy = F_{\mu_-}(x;\theta).$$

Odavde slijedi da je $F_{\mu_+}(\cdot;\theta) \equiv F_{\mu_-}(\cdot;\theta)$, a odavde odmah zaključujemo da su za λ-s.s. $x \in \langle 0, \theta]$,

$$f_+(x;\theta) = f_-(x;\theta) \quad \Rightarrow \quad \frac{x^{n-1}}{A(\theta)} \cdot g^+(x) = \frac{x^{n-1}}{A(\theta)} \cdot q^-(x) \quad \Rightarrow \quad g^+(x) = q^-(x) \quad \Rightarrow \quad g(x) = 0.$$

Sada kao i u prethodnom slučaju zaključujemo da je $g(T) = 0$ $\mathbb{P}_\theta$-g.s. Budući da je $\theta > 0$ bio proizvoljan, slijedi da je za sve $\theta > 0$, $g(T) = 0$ $\mathbb{P}_\theta$-g.s. Dakle, po definiciji potpunosti, $T = \max_{1 \leq i \leq n} X_i$ je potpuna statistika. $\square$

Zadatak 3.33. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz Poissonovog modela $P(\lambda)$ za $\lambda > 0$ i $T = X_1 + X_2 + \cdots + X_n$ statistika.

(a) Odredite uvjetni zakon razdiobe slučajnog uzorka $\mathbf{X}$ uz uvjet da je $T = k$ za k takav da ta uvjetna razdioba postoji.
(b) Nađite jednu verziju od $\mathbb{E}[e^{X_1}|T]$.
(c) Pokažite da je T potpuna i dovoljna statistika za λ.

Zadatak 3.34. Neka je $\mathcal{P} = \{f(\cdot;\theta) : \theta > 0\}$ neprekidni statistički model s gustoćom:

$$f(x;\theta) = \frac{x}{\theta} \cdot e^{-\frac{x^2}{2\theta}} \cdot \mathbb{1}_{\langle 0, +\infty\rangle}(x), \quad x \in \mathbb{R}.$$

Nađite jednu potpunu i jednu minimalnu dovoljnu statistiku za θ.

Poglavlje 4

Statistička procjena

4.1 Uvod u problem optimalne procjene

Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n $(n \geq 1)$ iz parametarskog modela

$$\mathcal{P} = \{f(\cdot;\theta) \,:\, \theta \in \Theta\}$$

gdje su gustoće $f(\cdot;\theta)$ slučajnih veličina $X_1, X_2, \ldots, X_n$ dimenzije d $(d \geq 1)$ parametrizirane parametrom θ dimenzije m $(m \geq 1)$. Cilj je procijeniti vrijednost (populacijskog) parametra $\tau(\theta)$ gdje je $\tau : \Theta \to \mathbb{R}^k$ izmjeriva funkcija $(k \geq 1)$. Kažemo da je $\tau(\theta)$ funkcija od θ dimenzije k. Ako je $k = m$ i $\tau(\theta) = \theta$ $(\theta \in \Theta)$, tada kažemo da želimo procijeniti parametar θ.

Procjenitelj od $\tau(\theta)$ je statistika $T = t(\mathbf{X})$ iste dimenzije kao i funkcija τ. Za takvu statistiku T još kažemo da je *točkovni procjenitelj* od $\tau(\theta)$. Dakle, procjenitelj može biti bilo koja statistika iste dimenzije kao funkcija parametra koju želimo procijeniti. Ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn}$ jedna realizacije slučajnog uzorka $\mathbf{X}$, tada je $t(\mathbf{x})$ **procjena** (ili *točkovna procjena*) od $\tau(\theta)$ na osnovi opaženog uzorka $\mathbf{x}$.

Često se procjenitelj T za $\tau(\theta)$ označava sa $\hat{\tau}(\theta)$ ili $\tilde{\tau}(\theta)$ i sl. Dakle, $T = \hat{\tau}(\theta)$ ili $T = \tilde{\tau}(\theta)$. Istu oznaku, na primjer $\hat{\tau}(\theta)$, možemo koristiti i za odgovarajuću procjenu od $\tau(\theta)$ kao realizaciju procjenitelja $T = t(\mathbf{X}) = \hat{\tau}(\theta)$ na osnovi opažene vrijednosti $\mathbf{x}$ slučajnog uzorka $\mathbf{X}$ kada je iz konteksta jasno da se radi o procjeni. Dakle, može biti i $\hat{\tau}(\theta) = t(\mathbf{x})$ ovisno o kontekstu.

Nećemo svaku statistiku dimenzije k koristiti za procjenu od $\tau(\theta)$ nego ćemo između raznih takvih statistika odabrati za procjenitelja onu statistiku $T = t(\mathbf{X})$ koja je, u nekom smislu, barem razuman izbor (vidjeti primjer 4.8), a u konačnici i *optimalan* izbor.

Razne procjenitelje za isti parametar $\tau(\theta)$ razumno je uspoređivati obzirom na preciznost procjene. Preciznost procjene možemo mjeriti na razne načine. U tu svrhu mi ćemo koristiti *kvadrat greške procjene*. Greška procjene $\tau(\theta)$ s $T = t(\mathbf{X})$ je broj $|t(\mathbf{x}) - \tau(\theta)|$ za danu realizaciju uzorka $\mathbf{x} \in \mathbb{R}^{dn}$. Ovdje je $|\cdot|$ oznaka za euklidsku normu. Tada je *kvadrat pogreške* jednak $|t(\mathbf{x}) - \tau(\theta)|^2$. Primijetite da je kvadrat euklidske norme $\mathbf{y} \mapsto |\mathbf{y}|^2$ funkcija lijepših svojstava od norme. Naime, norma nije derivabilna u svakoj točki dok kvadratna funkcija jest. Primijetite da nam je vrijednost $\tau(\theta)$ nepoznata, a da se svakim novim mjerenjem slučajnog uzorka $\mathbf{X}$ mijenja i vrijednost procjene $t(\mathbf{x})$. Zbog toga za *mjeru* greške procjene $\tau(\theta)$ pomoću procjenitelja $T = t(\mathbf{X})$ uzimamo *srednjekvadratnu pogrešku* od T (engl. *mean squared error*).

Definicija 4.1. *Neka je $T = t(\mathbf{X})$ procjenitelj za $\tau(\theta)$ kojemu komponente imaju konačnu varijancu za svaki $\theta \in \Theta$. Tada je* **srednjekvadratna pogreška procjene od** $\tau(\theta)$ **s** T

u odnosu na $\mathbb{P}_\theta$ *broj*

$$\mathrm{MSE}_\theta(T) := \mathbb{E}_\theta[|T - \tau(\theta)|^2]. \tag{4.1}$$

Kada je iz konteksta jasno da je cilj procijeniti funkciju $\tau(\theta)$ parametra θ, tada $\mathrm{MSE}_\theta(T)$ jednostavno zovemo *srednjekvadratna pogreška od* T budući da njome mjerimo koliko je T dobar procjenitelj za $\tau(\theta)$ ako je $\theta \in \Theta$ prava vrijednost parametra modela.

Budući da ne znamo koja je prava vrijednost od θ, u svrhu usporedbe dva procjenitelja iste funkcije $\tau(\theta)$ na njihove srednjekvadratne pogreške gledamo kao na funkcije od θ. U tom smislu ćemo reći da je T *najbolji procjenitelj za* $\tau(\theta)$ ako je za svaki drugi procjenitelj S, srednjekvadratna pogreška od T manja ili jednaka od srednjekvadratne pogreške od S *uniformno po* $\theta \in \Theta$, tj. ako za svaki drugi procjenitelj S vrijedi

$$(\forall \theta \in \Theta)\ \mathrm{MSE}_\theta(T) \leq \mathrm{MSE}_\theta(S). \tag{4.2}$$

Problem je da takav najbolji procjenitelj ne mora postojati. Pogledajmo jednostavni primjer.

Primjer 4.2. (Najbolji procjenitelj ne postoji). Model je zadan statističkom strukturom $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{P})$ gdje je $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ za

$$\mathbb{P}_1(B) = \int_B \frac{1}{2} e^{-|x|} dx, \quad \mathbb{P}_2(B) = \int_B e^{-2|x|} dx, \quad B \in \mathcal{B}(\mathbb{R}).$$

Nije teško provjeriti da su $\mathbb{P}_1$ i $\mathbb{P}_2$ vjerojatnosti. Te dvije vjerojatnosti su parametrizirane s $\theta \in \Theta = \{1, 2\}$ na način da je:

$$\mathbb{P}_\theta(B) = \int_B \frac{\theta}{2} e^{-\theta|x|} dx, \quad B \in \mathcal{B}(\mathbb{R}).$$

Neka je $\tau(\theta) = \theta$. Pretpostavimo da postoji najbolji procjenitelj T za θ u smislu da za svaki drugi procjenitelj S vrijedi (4.2). Neka su $S_1 \equiv 1$ i $S_2 \equiv 2$ dva procjenitelja za θ. Tada jednostavno slijedi da su njihove srednjekvadratne pogreške jednake

$$\mathrm{MSE}_\theta(S_1) = \mathbb{E}_\theta[(S_1 - \theta)^2] = \mathbb{E}_\theta[(1 - \theta)^2] = (1 - \theta)^2$$

i slično $\mathrm{MSE}_\theta(S_2) = (2 - \theta)^2$ za sve $\theta \in \Theta$. Ako primijenimo (4.2) na $S = S_1$, dobijemo da je

$$\begin{aligned} & (\forall \theta \in \Theta)\ 0 \leq \mathrm{MSE}_\theta(T) \leq \mathrm{MSE}_\theta(S_1) = (1 - \theta)^2 \\ \Rightarrow\quad & 0 \leq \mathrm{MSE}_1(T) \leq 0 \quad \Rightarrow \quad \mathrm{MSE}_1(T) = \mathbb{E}_1[(T - 1)^2] = 0 \\ \Rightarrow\quad & (T - 1)^2 = 0\ \mathbb{P}_1\text{ - g.s} \quad \Rightarrow \quad T = 1\ \mathbb{P}_1\text{ - g.s.} \end{aligned}$$

Sličnom primjenom (4.2) na $S = S_2$ dobijemo da je $T = 2$ $\mathbb{P}_2$ - g.s. Budući da je $\mathbb{1}_{\{T \neq 1\}}$ nenegativna funkcija, vrijedi sljedeća implikacija:

$$0 = \mathbb{P}_1(T \neq 1) = \int_{\{T \neq 1\}} e^{-|x|} dx = \int_{\mathbb{R}} \mathbb{1}_{\{T \neq 1\}} \cdot e^{-|x|} dx \quad \Rightarrow \quad \mathbb{1}_{\{T \neq 1\}} = 0\ \lambda\text{- s.s.}$$

Dakle,

$$0 = \int_{\mathbb{R}} \mathbb{1}_{\{T \neq 1\}} \cdot \frac{1}{2} e^{-2|x|} dx = \int_{\{T \neq 1\}} \frac{1}{2} e^{-2|x|} dx = \mathbb{P}_2(T \neq 1) \quad \Rightarrow \quad \mathbb{P}_2(T \neq 1) = 0,$$

tj. $T = 1$ $\mathbb{P}_2$ - g.s., a što je u kontradikciji sa zaključkom da je $T = 2$ $\mathbb{P}_2$ - g.s. Prema tome, za zadani model u klasi svih procjenitelja za θ, ne postoji najbolji u smislu uniformno najmanje srednjekvadratne pogreške. □

Prethodni primjer upućuje na to da treba smanjiti klasu procjenitelja u kojoj tražimo najboljeg. Pokazuje se da najbolji procjenitelj postoji u klasi *nepristranih* procjenitelja.

4.2 Nepristrani procjenitelji

Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine n ($n \geq 1$) iz statističkog modela parametriziranog parametrom $\theta \in \Theta$. Pretpostavimo da je $\tau : \Theta \to \mathbb{R}^k$ izmjeriva funkcija ($k \geq 1$) te da želimo procijeniti $\tau(\theta)$.

Definicija 4.3. *Procjenitelj $T = t(\mathbf{X})$ za $\tau(\theta)$ je* **nepristran** *ako vrijedi*

$$(\forall \theta \in \Theta) \;\; \mathbb{E}_\theta[T] = \tau(\theta). \tag{4.3}$$

Procjenitelj koji nije nepristran kažemo da je **pristran**.

Primjer 4.4. Primjeri nepristranih procjenitelja:

(a) U Bernoullijevom modelu $b(1, \theta)$, $\theta \in \Theta = [0, 1]$ procjenitelj $\overline{X}_n$, relativna frekvencija uspjeha, je nepristrani procjenitelj za vjerojatnost uspjeha θ:

$$\mathbb{E}_\theta[\overline{X}_n] \overset{\text{lin.mat.oč.}}{=} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_\theta X_1 \overset{\text{jedn.distr.}}{=} \frac{1}{n} n\theta = \theta$$

za sve $\theta \in [0, 1]$.

(b) Općenitije, neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak iz modela s konačnim matematičkim očekivanjem. To znači da je za svaki $\theta \in \Theta$, $\mathbb{E}_\theta[|X_1|] < +\infty$, pa je dobro definirana funkcija $\mu : \Theta \to \mathbb{R}$ sa $\mu(\theta) := \mathbb{E}_\theta[X_1]$, $\theta \in \Theta$. Dakle, $\mu(\theta)$ je matematičko očekivanje. Uzoračka aritmetička sredina $\overline{X}_n$ je nepristrani procjenitelj za $\mu(\theta)$.

(c) Uz iste pretpostavke na model kao u (b), pretpostavimo da model dopušta i konačnu varijancu. Dakle, za sve $\theta \in \Theta$, funkcija $\sigma^2 : \Theta \to \mathbb{R}$ je dobro definirana (i konačna) sa $\sigma^2(\theta) := \mathbb{V}\mathrm{ar}_\theta X_1$. Tada je uzoračka varijanca S_n^2 nepristrani procjenitelj za $\sigma^2(\theta)$ (vidjeti predavanja iz Statistike, *Statistička procjena*, Zadatak 1 na slajdu 9).

(d) Neka je $\mathcal{P} = \{f(\cdot; \theta) : \theta \in \Theta\}$ statistički model jednodimenzionalne razdiobe. Označimo sa $F(\cdot; \theta)$ ($\theta \in \Theta$) odgovarajuće funkcije distribucije. Ako je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak iz tog modela, tada je vrijednost empirijske funkcije distribucije

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}$$

u točki $x \in \mathbb{R}$ nepristrani procjenielj za vrijednost funkcije distribucije $F(x; \theta)$ u istoj točki x. Naime, za sve $\theta \in \Theta$,

$$\mathbb{E}_\theta[\hat{F}_n(x)] \overset{\text{lin.mat.oč.}}{=} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_\theta[\mathbb{1}_{\{X_i \leq x\}}] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}_\theta(X_1 \leq x) \overset{\text{jedn.distr.}}{=} \frac{1}{n} n F_\theta(x; \theta) = F_\theta(x; \theta).$$

$\square$

Nepristrani procjenitelj za danu funkciju ne mora uvijek postojati. Zbog toga sljedeća definicija ima smisla.

Definicija 4.5. *Za zadani statistički model parametriziran parametrom θ i dani slučajni uzorak $\mathbf{X}$ iz tog modela, kažemo da je funkcija $\tau(\theta)$* **procjenjiva** *ako postoji barem jedan nepristrani procjenitelj $T = t(\mathbf{X})$ od $\tau(\theta)$.*

Primjer 4.6. Primjeri funkcija koje nisu procjenjive:

(a) (Procjenjivost ovisi o duljini uzorka). Neka je $X = X_1$ slučajni uzorak duljine $n = 1$ iz Bernoullijevog modela $b(1,\theta)$, $\theta \in \langle 0, 1\rangle$. Tada ne postoji nepristrani procjenitelj $T = t(X_1)$ za $\tau(\theta) = \theta^2$. Pretpostavimo suprotno: da postoji nepristrani procjenitelj $T = t(X_1)$ za θ^2, tada bi za svaki $\theta \in \langle 0, 1\rangle$ trebalo vrijediti da je

$$\theta^2 = \mathbb{E}_\theta[t(X_1)] = t(0)\cdot(1-\theta) + t(1)\cdot\theta = (t(1) - t(0))\theta + t(0).$$

To je nemoguće jer netrivijalni polinom drugog stupnja nije jednak polinomu stupnja najviše 1 (po θ). Primijetite da, ako je duljina uzorka $n \geq 2$, da je tada θ^2 procjenjiva funkcija u istome modelu.
(b) (Procjenjivost ovisi o statističkom modelu). Neka je zadan Bernoullijev model kao u (a). Tada funkcija $\tau(\theta) = 1/\theta$ nije procjenjiva za slučajni uzorak bilo koje duljine n. Pretpostavimo da je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n ($n \geq 1$), te da postoji nepristrani procjenitelj $T = t(\mathbf{X})$ za $1/\theta$. Tada za svaki $\theta \in \langle 0, 1\rangle$ vrijedi

$$\begin{aligned}\frac{1}{\theta} &= \mathbb{E}_\theta[t(\mathbf{X})] = \sum_{x_1,\ldots,x_n \in \{0,1\}} t(x_1,\ldots,x_n)\theta^{x_1+\cdots+x_n}(1-\theta)^{n-(x_1+\cdots+x_n)} = \\ &= t(0,\ldots,0)(1-\theta)^n + \theta \cdot \sum_{x_1+\cdots+x_n \geq 1} t(x_1,\ldots,x_n)\theta^{x_1+\cdots+x_n-1}(1-\theta)^{n-(x_1+\cdots+x_n)}.\end{aligned}$$

Ako pogledamo granične vrijednosti lijeve i desne strane jednakosti kada $\theta \to 0+$, tada dobijamo da je $+\infty = t(0,\ldots,0) \in \mathbb{R}$ što je kontradikcija. □

Je li svojstvo nepristranosti invarijantno na funkcijske transformacije? Općenito nije. Na primjer, ako se radi o statističkom modelu koji dopušta nedegeneriranu konačnu varijancu kao u primjeru 4.4 $(b-c)$ u kojemu je $T = \overline{X}_n$ nepristrani procjenitelj za matematičko očekivanje $\mu(\theta)$, tada je $T^2 = \overline{X}_n^2$ *pristrani* procjenitelj za funkciju $\tau(\theta) = \mu(\theta)^2$. Naime, za sve $\theta \in \Theta$ je

$$\mathbb{E}_\theta[T^2] = \mathbb{E}_\theta[\overline{X}_n^2] = \mathbb{V}\mathrm{ar}_\theta \overline{X}_n + (\mathbb{E}_\theta[\overline{X}_n])^2 = \sigma^2(\theta) + \mu(\theta)^2 > \mu(\theta)^2.$$

Dakle, funkcijske transformacije općenito ne čuvaju svojstvo nepristranosti procjenitelja. Jedina klasa funkcija koja čuva nepristranost je klasa afinih funkcija.

Propozicija 4.7. *Neka je T nepristrani procjenitelj za funkciju $\tau(\theta)$ dimenzije k u nekom statističkom modelu parametriziranom sa θ i neka je $\eta : \mathbb{R}^k \to \mathbb{R}^\ell$ afina funkcija $\eta(y) = Ay + \mathbf{b}$ ($y \in \mathbb{R}^k$) (za zadane $A \in M_{k,\ell}, \mathbf{b} \in \mathbb{R}^\ell$, gdje su $k, \ell \in \mathbb{N}$). Tada je $\eta(T) = AT + \mathbf{b}$ nepristrani procjenitelj za $\eta(\tau(\theta)) = A\tau(\theta) + \mathbf{b}$.*

Dokaz. Nepristranost slijedi direktno iz linearnosti matematičkog očekivanja. □

Primjer 4.8. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz Poissonovog modela $P(\lambda)$, $\lambda > 0$. Za statistiku $T = (-1)^{X_1}$ i proizvoljan $\lambda > 0$ računamo:

$$\mathbb{E}_\lambda[T] = \mathbb{E}_\lambda[(-1)^{X_1}] = \sum_{k=0}^{\infty}(-1)^k \frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(-\lambda)^k}{k!} = e^{-2\lambda}.$$

Dakle, $T = (-1)^{X_1}$ je nepristrani procjenitelj za $\tau(\lambda) = e^{-2\lambda}$. Primijetite da T razumno nikada ne bismo odabrali za procjenjivanje funkcije $e^{-2\lambda}$ budući da je za sve $\lambda > 0$, $e^{-2\lambda} \in \langle 0, 1\rangle$, a $T = (-1)^{X_1}$ poprima vrijednosti u skupu $\{-1, 1\}$. Ipak, i takvi procjenitelji imat će ulogu u konstrukciji najboljih procjenitelja. □

4.3 Nepristrani procjenitelj uniformno minimalne varijance

Na kraju uvodnog potpoglavlja zaključili smo da se za egzistenciju najboljeg procjenitelja za danu funkciju parametra modela, u smislu uniformno najmanje srednjekvadratne pogreške, moramo ograničiti na manju klasu procjenitelja. Ta klasa će biti nepristrani procjenitelji. U tehničkom smislu ograničit ćemo se na nepristrane procjenitelje *konačne varijance*. Za statistiku T definiranu na statističkom modelu parametriziranom parametrom $\theta \in \Theta$ kažemo da je **konačne varijance** ako je za svaki $\theta \in \Theta$ varijanca $\mathbb{V}\mathrm{ar}_\theta T$ od T konačna.

Označimo sa $\mathcal{W}_\tau$ množinu svih nepristranih procjenitelja za $\tau(\theta)$ konačne varijance. Primijetite da je $\mathcal{W}_\tau \subseteq \mathcal{L}^2(\mathbb{P}_\theta)$ za sve $\theta \in \Theta$.

Definicija 4.9. *Neka je $\tau(\theta)$ procjenjiva funkcija. Statistika T je* **nepristrani procjenitelj uniformno minimalne varijance od** $\tau(\theta)$ *(engl.* Uniformly Minimum Variance Unbiased Estimator*, kraticom* UMVUE*) ako vrijedi da je $T \in \mathcal{W}_\tau$ i da je*

$$(\forall S \in \mathcal{W}_\tau)\,(\forall \theta \in \Theta)\;\; \mathbb{V}\mathrm{ar}_\theta T \leq \mathbb{V}\mathrm{ar}_\theta S. \tag{4.4}$$

Navedimo jednu karakterizaciju nepristranih procjenitelja uniformno minimalne varijance.

Teorem 4.10. *Neka je $\mathcal{N} \equiv \mathcal{W}_\eta$ množina nepristranuih procjenitelja za $\eta(\theta) \equiv 0$ konačne varijance. Tada je $T \in \mathcal{W}_\tau$ nepristrani procjenitelj uniformno minimalne varijance od $\tau(\theta)$ ako i samo vrijedi*

$$(\forall N \in \mathcal{N})\,(\forall \theta \in \Theta)\;\; \mathbb{E}_\theta[T \cdot N] = 0. \tag{4.5}$$

Dokaz. Neka je $\theta \in \Theta$ proizvoljan. Pokazat ćemo da su za $T \in \mathcal{W}_\tau$ i taj fiksni θ, nejednakost iz (4.4) i jednakost iz (4.5) ekvivalentni (uniformno po $S \in \mathcal{W}_\tau$ i $N \in \mathcal{N}$).

Prvo, primijetimo da, ako je $\mathbb{E}_\theta[T^2] = 0$, da je tada $T = 0$ $\mathbb{P}_\theta$-g.s., pa je za sve $S \in \mathcal{W}_\tau$, $\mathbb{V}\mathrm{ar}_\tau T = 0 \leq \mathbb{V}\mathrm{ar}_\tau S$ (dakle, vrijedi nejednakost iz (4.4) uniformno po $S \in \mathcal{W}_\tau$) i za sve $N \in \mathcal{N}$ je $\mathbb{E}_\theta[T \cdot N] = \mathbb{E}_\theta[0 \cdot N] = 0$ (pa vrijedi i jednakost iz (4.5) uniformno po $N \in \mathcal{N}$). Primijetite da ova diskusija ima smisla samo ako je $\tau(\theta) = 0$ za takav θ. Prema tome, pretpostavimo da je $\mathbb{E}_\theta[T^2] > 0$.

Smjer '$\Leftarrow$': Neka za θ vrijedi jednakost iz (4.5) za sve $N \in \mathcal{N}$, te neka je $S \in \mathcal{W}_\tau$ proizvoljan. Tada je $T - S \in \mathcal{N}$ jer je

$$\mathbb{E}_\theta[T - S] \overset{\text{lin.mat.oč.}}{=} \mathbb{E}_\theta T - \mathbb{E}_\theta S = \tau(\theta) - \tau(\theta) = 0$$

neovisno o vrijednosti θ i T i S su konačne varijance pa je i njihova razlika konačne varijance. Primjenom (4.5) na $N = T - S$ i linearnosti matematičkog očekivanja dobijamo da je

$$0 \overset{(4.5)}{=} \mathbb{E}_\theta[T \cdot (T - S)] \overset{\text{lin.mat.oč.}}{=} \mathbb{E}_\theta[T^2] - \mathbb{E}_\theta[TS] \;\;\Leftrightarrow\;\; \mathbb{E}_\theta[T^2] = \mathbb{E}_\theta[T \cdot S].$$

Primjenom Cauchy - Schwarzove nejednakosti (teorem 2.2) dobivamo:

$$\begin{aligned}
&\mathbb{E}_\theta[T^2] = \mathbb{E}_\theta[T \cdot S] \overset{\text{C-S}}{\leq} \sqrt{\mathbb{E}_\theta[T^2]} \cdot \sqrt{\mathbb{E}_\theta[S^2]} \overset{\mathbb{E}_\theta[T^2]>0}{\Rightarrow} \\
\Rightarrow\;\; &\sqrt{\mathbb{E}_\theta[T^2]} \leq \sqrt{\mathbb{E}_\theta[S^2]} \;\;\Leftrightarrow\;\; \mathbb{E}_\theta[T^2] \leq \mathbb{E}_\theta[S^2] \;\;\Leftrightarrow\;\; \mathbb{E}_\theta[T^2] - \tau(\theta)^2 \leq \mathbb{E}_\theta[S^2] - \tau(\theta)^2 \;\;\Leftrightarrow \\
\Leftrightarrow\;\; &\mathbb{V}\mathrm{ar}_\theta T \leq \mathbb{V}\mathrm{ar}_\theta S.
\end{aligned}$$

Budući da je $S \in \mathcal{W}_\tau$ bio proizvoljan, nejednakost iz (4.4) vrijedi za sve $S \in \mathcal{W}_\tau$.

Smjer '$\Rightarrow$': Neka je T UMVUE, dakle, nejednakost iz (4.4) vrijedi za zadani θ i sve $S \in \mathcal{W}_\tau$. Neka je $N \in \mathcal{N}$ proizvoljan. Tada je za svaki realan broj λ, statistika $S_\lambda := T + \lambda N \in \mathcal{W}_\tau$. Naime, neovisno o vrijednosti $\theta \in \Theta$ je

$$\mathbb{E}_\theta[S_\lambda] = \mathbb{E}_\theta[T + \lambda N] \overset{\text{lin.mat.oč.}}{=} \mathbb{E}_\theta[T] + \lambda \mathbb{E}_\theta[N] = \tau(\theta) + \lambda \cdot 0 = \tau(\theta),$$

pa je S_λ nepristran procjenitelj za $\tau(\theta)$, a kako je linearna kombinacija procjenitelja konačne varijance, konačne je varijance također, za svaki $\lambda \in \mathbb{R}$. Dakle, prema (4.4) i korištenjem svojstva varijance (iv) iz teorema 1.61 vrijedi da je za sve $\lambda \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{V}\mathrm{ar}_\theta T \overset{(4.4)}{\leq} \mathbb{V}\mathrm{ar}_\theta[S_\lambda] = \mathbb{V}\mathrm{ar}_\theta[T+\lambda N] \overset{(iv)\mathrm{tm.}1.61}{=} \mathbb{V}\mathrm{ar}_\theta T + 2\lambda \cdot \mathrm{cov}_\theta(T,N) + \lambda^2 \mathbb{V}\mathrm{ar}_\theta N \\ \Rightarrow \quad & f(\lambda) := 2\lambda \cdot \mathrm{cov}_\theta(T,N) + \lambda^2 \mathbb{V}\mathrm{ar}_\theta N \geq 0. \end{aligned}$$

Ako je $\mathbb{V}\mathrm{ar}_\theta N = 0$, tada je nužno da je $\mathrm{cov}_\theta(T,N) = 0$ da bi dobivena nejednakost vrijedila za sve $\lambda \in \mathbb{R}$. Odavde slijedi da je i

$$0 = \mathrm{cov}_\theta(T,N) \overset{(vi)\mathrm{tm.}1.61}{=} \mathbb{E}_\theta[T \cdot N] - \mathbb{E}_\theta T \cdot \mathbb{E}_\theta N = \mathbb{E}_\theta[T \cdot N] - \tau(\theta) \cdot 0 = \mathbb{E}_\theta[T \cdot N],$$

dakle, vrijedi jednakost u (4.5). U slučaju da je $\mathbb{V}\mathrm{ar}_\theta N > 0$, dobivena kvadratna funkcija $f(\lambda)$ po λ je konveksna i mora biti nenegativna na čitavoj domeni. U tom slučaju obje realne nul-točke te funkcije moraju biti jednake nuli:

$$\lambda_1 = 0 \ \& \ \lambda_2 = -\frac{2\,\mathrm{cov}_\theta(T,N)}{\mathbb{V}\mathrm{ar}_\theta N} = 0 \quad \Rightarrow \quad \mathbb{E}_\theta[T \cdot N] = \mathrm{cov}_\theta(T,N) = 0,$$

pa slijedi jednakost u (4.5). Budući da je $N \in \mathcal{N}$ bio proizvoljan, jednakost u (4.5) vrijedi za sve $N \in \mathcal{N}$.

Zbog proizvoljnosti od θ, slijedi da su tvrdnje (4.4) i (4.5) ekvivalentne. □

Pokažimo da je nepristrani procjenitelj uniformno minimalne varijance jedinstven.

Teorem 4.11. *Neka je $\tau(\theta)$ procjenjiva funkcija. Ako postoji, nepristrani procjenitelj uniformno minimalne varijance od $\tau(\theta)$ je $\mathbb{P}_\theta$-g.s. jedinstven, za sve $\theta \in \Theta$.*

Dokaz. Pretpostavimo da postoje barem dva UMVUE za $\tau(\theta)$. Neka su to T i U. Tada za svaki $\theta \in \Theta$ vrijedi da je $\mathbb{E}_\theta[T-U] = 0$ i da je $T-U$ konačne varijance, te da je

$$\mathbb{V}\mathrm{ar}_\theta T \leq \mathbb{V}\mathrm{ar}_\theta U \ \& \ \mathbb{V}\mathrm{ar}_\theta U \leq \mathbb{V}\mathrm{ar}_\theta T \quad \Rightarrow \quad \mathbb{V}\mathrm{ar}_\theta T = \mathbb{V}\mathrm{ar}_\theta U$$

jer su T i U najmanje varijance u množini $\mathcal{W}_\tau$. Budući da je $T - U \in \mathcal{N}$, prema karakterizaciji UMVUE danoj teoremom 4.10, slijedi da je za sve $\theta \in \Theta$,

$$\mathbb{E}_\theta[T(T-U)] = \mathbb{E}_\theta[U(T-U)] = 0 \quad \Rightarrow \quad \mathbb{E}_\theta[T \cdot U] = \mathbb{E}_\theta[T^2] = \mathbb{E}_\theta[U^2].$$

Dakle, za proizvoljno $\theta \in \Theta$ je

$$\mathrm{cov}_\theta(T,U) = \mathbb{E}_\theta[T \cdot U] - \tau(\theta)^2 = \mathbb{E}_\theta[T^2] - \tau(\theta)^2 = \mathbb{V}\mathrm{ar}_\theta T = \mathbb{V}\mathrm{ar}_\theta U.$$

Ako je $\mathbb{V}\mathrm{ar}_\theta T = \mathbb{V}\mathrm{ar}_\theta U = 0$, tada je za taj θ, $T = U = \tau(\theta)$ $\mathbb{P}_\theta$-g.s. U suprotnom, ako je $\mathbb{V}\mathrm{ar}_\theta T = \mathbb{V}\mathrm{ar}_\theta U > 0$, tada je

$$\mathrm{corr}_\theta(T,U) = \frac{\mathrm{cov}_\theta(T,U)}{\sqrt{\mathbb{V}\mathrm{ar}_\theta T \cdot \mathbb{V}\mathrm{ar}_\theta U}} = \frac{\mathbb{V}\mathrm{ar}_\theta T}{\mathbb{V}\mathrm{ar}_\theta T} = 1.$$

Prema tvrdnji (ii) korolara 2.3 slijedi da postoje brojevi $\alpha(\theta) > 0$ i $\beta(\theta)$ takvi da je $T = \alpha(\theta)U + \beta(\theta)$ $\mathbb{P}_\theta$-g.s. Budući da je $\alpha(\theta) > 0$, vrijedi:

$$\mathbb{V}\mathrm{ar}_\theta T = \mathbb{V}\mathrm{ar}_\theta U = \mathbb{V}\mathrm{ar}_\theta[\alpha(\theta)T + \beta(\theta)] \overset{(ii)\mathrm{tm.}1.61}{=} \alpha(\theta)^2 \mathbb{V}\mathrm{ar}_\theta T \quad \Rightarrow \quad \alpha(\theta) = 1.$$

Dakle, $T = U + \beta(\theta)$ $\mathbb{P}_\theta$-g.s. Zbog nepristranosti od T i U i primjenom linearnosti matematičkog očekivanja, slijedi da je

$$\tau(\theta) = \mathbb{E}_\theta T = \mathbb{E}_\theta U = \mathbb{E}_\theta[T + \beta(\theta)] = \mathbb{E}_\theta T + \beta(\theta) = \tau(\theta) + \beta(\theta) \quad \Rightarrow \quad \beta(\theta) = 0.$$

Dakle, $T = U$ $\mathbb{P}_\theta$-g.s. Zbog proizvoljnosti od $\theta \in \Theta$, slijedi tvrdnja. □

Kako pronaći nepristrani procjenitelj uniformno minimalne varijance? Sljedeći teorem nam kaže da se u potrazi za UMVUE možemo ograničiti na funkcije dovoljnih statistika.

Kao i prije, neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz statističkog modela parametriziranog parametrom θ.

Teorem 4.12 (Rao-Blackwell). *Neka je $\tau(\theta)$ procjenjiva funkcija dimenzije jedan i neka je $T = t(\mathbf{X})$ dovoljna statistika za θ. Nadalje, neka je $S = s(\mathbf{X})$ nepristrani procjenitelj za $\tau(\theta)$ konačne varijance, te neka je $u : \mathbb{R} \to \mathbb{R}$ izmjeriva regresijska funkcija od S u odnosu na T. Tada za $U = u(T)$ vrijedi:*

(i) U je statistika i funkcija od T;
(ii) U je nepristrani procjenitelj za $\tau(\theta)$;
(iii) Za svaki $\theta \in \Theta$ je $\mathbb{V}\mathrm{ar}_\theta U \leq \mathbb{V}\mathrm{ar}_\theta S$.

Dokaz. (*i*): Za regresijsku funkciju u od S u odnosu na T (vidjeti definiciju 2.15) vrijedi da je $u(x) = \mathbb{E}_\theta[S|T = x]$ za sve x za koje je desna strana te jednakosti definirana. Budući da je T dovoljna statistika, prema propoziciji 3.8 slijedi da $u(x)$ ne ovisi o θ (pa pišemo jednostavno da je $u(x) = \mathbb{E}[S|T = x]$, $x \in \mathrm{supp} f_T(\cdot, \theta)$). Nadalje, u je izmjeriva funkcija, a kako ne ovisi o θ na $\mathrm{supp} f_T(\cdot; \theta)$ gdje T poprima vrijednosti $\mathbb{P}_\theta$-g.s., $U = u(T)$ je izmjeriva funkcija *samo* od T, dakle, statistika.

(*ii*): Budući da S ima konačno matematičko očekivanje i $U = u(T)$ je kompozicija regresijske funkcije od S u odnosu na T i statistike T, prema teoremu 2.16 je $u(T) = \mathbb{E}[S|T]$ P_θ-g.s. za sve $\theta \in \Theta$. Primjenom svojstva (*i*) uvjetnog očekivanja iz teorema 2.26 za proizvoljno $\theta \in \Theta$ dobijamo da je

$$\mathbb{E}_\theta U = \mathbb{E}_\theta[u(T)] = \mathbb{E}_\theta[\mathbb{E}[S|T]] = \mathbb{E}_\theta S = \tau(\theta)$$

jer je S nepristrani procjenitelj za $\tau(\theta)$ po pretpostavci. Zbog proizvoljnosti od θ slijedi tvrdnja.

(*iii*): Primijetite da je za sve $\theta \in \Theta$,

$$\mathbb{V}\mathrm{ar}_\theta U \leq \mathbb{V}\mathrm{ar}_\theta S \quad \Leftrightarrow \quad \mathbb{E}_\theta[U^2] \leq \mathbb{E}_\theta[S^2]$$

jer su U i S nepristrani procjenitelji iste funkcije $\tau(\theta)$. Dakle, dovoljno je dokazati (desnu) ekvivalentnu tvrdnju za proizvoljni $\theta \in \Theta$. Korištenjem svojstva linearnosti matematičkog očekivanja (*i*) teorem 1.56, za proizvoljno $\theta \in \Theta$ slijedi:

$$\mathbb{E}_\theta[S^2] = \mathbb{E}_\theta[((S - U) + U)^2] \overset{\text{lin.m.o.}}{=} \mathbb{E}_\theta[(S - U)^2] + 2\,\mathbb{E}_\theta[U(S - U)] + \mathbb{E}_\theta[U^2].$$

Budući da je, primjenom svojstava uvjetnog očekivanja $(i - iii)$ i (x) (jer je $U = u(T)$ $\sigma(T)$-izmjerivo) teorema 2.26:

$$\begin{aligned}\mathbb{E}_\theta[U(S - U)] &\overset{(i)\,\text{tm.}2.26}{=} \mathbb{E}_\theta[\,\mathbb{E}[U(S - U)|T]\,] \overset{(x)\,\text{tm.}2.26}{=} \mathbb{E}_\theta[U \cdot \mathbb{E}[S - U|T]\,] \overset{(iii,ii)\,\text{tm.}2.26}{=} \\ &= \mathbb{E}_\theta[U \cdot (\mathbb{E}[S|T] - U)] = \mathbb{E}_\theta[U(U - U)] = \mathbb{E}_\theta 0 = 0\end{aligned}$$

i $\mathbb{E}_\theta[(S - U)^2] \geq 0$, iz gornje jednakosti slijedi da je $\mathbb{E}_\theta[U^2] \leq \mathbb{E}_\theta[S^2]$, a time i tvrdnja (*iii*). Primijetite da ta tvrdnja slijedi i iz teorema 2.36 (*iii*). □

Koristeći se sljedećim teoremom moguće je prepoznati UMVUE.

Teorem 4.13 (Lehmann-Scheffé)**.** *Neka je $\tau(\theta)$ procjenjiva funkcija dimenzije jedan i neka je $T = t(\mathbf{X})$ potpuna i dovoljna statistika za θ. Ako je statistika $U = u(T)$ funkcija potpune i dovoljne statistike za neku izmjerivu funkciju u i nepristrani je procjenitelj za $\tau(\theta)$ konačne varijance, tada je U nepristrani procjenitelj uniformno minimalne varijance od $\tau(\theta)$.*

Dokaz. Prvo ćemo pokazati g.s. jedinstvenost nepristranog procjenitelja koji je izmjeriva funkcija potpune statistike T. Pretpostavimo suprotno, tj. da uz nepristrani procjenitelj $U = u(T)$ za $\tau(\theta)$ postoji i drugi nepristrani procjenitelj V iste funkcije koji je izmjeriva funkcija od T, dakle, $V = v(T)$ i za sve $\theta \in \Theta$ je $\mathbb{E}_\theta U = \mathbb{E}_\theta V = \tau(\theta)$. Neka je $w := u - v$. Tada je w Borelova funkcija i vrijedi da je za sve $\theta \in \Theta$:

$$\mathbb{E}_\theta[w(T)] = \mathbb{E}_\theta[u(T) - v(T)] \overset{\text{lin.m.o.}}{=} \mathbb{E}_\theta U - \mathbb{E}_\theta V \overset{\text{neprist.}}{=} \tau(\theta) - \tau(\theta) = 0.$$

Budući da je T potpuna statistika, slijedi da je za sve $\theta \in \Theta$:

$$w(T) = 0 \ \mathbb{P}_\theta\text{- g.s.} \ \Leftrightarrow \ U = V \ \mathbb{P}_\theta\text{- g.s.}$$

čime je dokazana tvrdnja o g.s. jedinstvenosti nepristranog procjenitelja koji je funkcija potpune statistike.

Iz pretpostavke na U je očito da je $U \in \mathcal{W}_\tau$. Dokažimo sada da je U uniformno po θ najmanje varijance u množini $\mathcal{W}_\tau$ svih nepristranih procjenitelja za $\tau(\theta)$ konačne varijance. Neka je $S = s(\mathbf{X}) \in \mathcal{W}_\tau$ proizvoljni procjenitelj. Budući da je S konačne varijance, u skladu s teoremom 2.16 i zadatkom 2.31, postoji izmjeriva regresijska funkcija ϕ od S u odnosu na T. Stavimo da je $V = \phi(T)$. Budući da je T po pretpostavci i dovoljna statistika, prema Rao-Blackwellovom teoremu 4.12, V je nepristrani procjenitelj za $\tau(\theta)$ konačne varijance koji je manje varijance od varijance od S uniformno po θ. Budući da smo pokazali da postoji g.s. najviše jedan nepristrani procjenitelj koji je funkcija potpune statistike, slijedi da je $V = U$ $\mathbb{P}_\theta$-g.s. za sve θ. Dakle, za svaki $\theta \in \Theta$ je

$$\mathbb{V}\text{ar}_\theta U \overset{U=V \text{ g.s.}}{=} \mathbb{V}\text{ar}_\theta V \overset{(iii)\,\text{tm.4.12}}{\leq} \mathbb{V}\text{ar}_\theta S.$$

Budući da je $S \in \mathcal{W}_\tau$ bio proizvoljan, slijedi da je U UMVUE. □

Primjer 4.14. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz modela za eksponencijalnu distribuciju $\text{Exp}(\lambda)$, $\lambda > 0$. Nađimo UMVUE za $\tau(\lambda) = 1/\lambda$. Prvo, primijetimo da je $\tau(\lambda) = 1/\lambda$ matematičko očekivanje $\text{Exp}(\lambda)$-razdiobe. Budući da je uzoračka aritmetička sredina $\overline{X}_n$ nepristrani procjenitelj za matematičko očekivanje od $X_1 \sim \text{Exp}(\lambda)$ (vidjeti primjer 4.4 (b)), $\overline{X}_n$ je nepristrani procjenitelj za $\tau(\lambda) = 1/\lambda$ varijance:

$$\mathbb{V}\text{ar}_\lambda \overline{X}_n = \frac{1}{n}\mathbb{V}\text{ar}_\lambda X_1 = \frac{1}{n\lambda^2} < +\infty$$

za sve $\lambda > 0$, dakle, konačne varijance. *Je li $\overline{X}_n$ funkcija potpune i dovoljne statistike?* Ako je, onda je i UMVUE prema Lehmann-Schefféovom teoremu 4.13. Da bi to pokazali, primijetimo da je $\text{Exp}(\lambda)$ 1-parametarska eksponencijalna familija punog ranga kojoj je, prema teoremima 3.27 i 3.31, prirodna dovoljna statistika $T = X_1 + X_2 + \cdots + X_n$ dovoljna i potpuna. Naime,

$$f(x; \lambda) = \lambda \cdot e^{-\lambda x} \cdot \mathbb{1}_{\langle 0,+\infty\rangle}(x), \ \ x \in \mathbb{R},$$

gdje je $t(x) = t_1(x) = x$ pa je $\{t\}$ jednočlani linearno nezavisni skup funkcija, a $Q(\lambda) = Q_1(\lambda) = -\lambda$ bijekcija koja preslikava parametarski prostor $\langle 0, +\infty\rangle$ na otvoreni interval $\langle -\infty, 0\rangle$. Očito su $C(\lambda) = \lambda$ i $h(x) = \mathbb{1}_{\langle 0,+\infty\rangle}(x)$. Sada iz definicija 3.23 i 3.25 slijedi da je $\text{Exp}(\lambda)$ za $\lambda > 0$ 1-parametarska familija punog ranga.
Budući da je $\overline{X}_n = T/n$, dakle, funkcija potpune i dovoljne statistike za λ, uzoračka aritmetička sredina je UMVUE od $1/\lambda$. □

Primjer 4.15. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz Benoullijeva modela $b(1, \theta)$, $\theta \in \langle 0, 1\rangle$. U tom modelu je $T = X_1 + X_2 + \cdots + X_n$ potpuna i dovoljna statistika.
(a) Relativna frekvencija uspjeha $\overline{X}_n = T/n$ je nepristrani procjenitelj za θ konačne varijance, a kako je funkcija potpune i dovoljne statistike T, prema Lehmann-Schefféovom teoremu 4.13 je UMVUE od vjerojatnosti uspjeha θ.
(b) Neka je $\sigma^2(\theta) = \theta(1-\theta) = \mathbb{V}\mathrm{ar}_\theta X_1$. Uzoračka varijanca S_n^2 je nepristrani procjenitelj za $\sigma^2(\theta)$ konačne varijance u tom modelu. Budući da su $X_i \in \{0, 1\}$ za sve i, slijedi da je:

$$S_n^2 = \frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - n\overline{X}_n^2\right) \overset{{X_i}^2 = X_i}{=} \frac{1}{n-1}\left(\sum_{i=1}^n X_i - n\left(\frac{T}{n}\right)^2\right) = \frac{T}{n-1}\left(1 - \frac{T}{n}\right)$$

funkcija potpune i dovoljne statistike T. Dakle, prema Lehmann-Schefféovom teoremu 4.13 S_n^2 je UMVUE od varijance $\theta(1-\theta)$. $\square$

Primjer 4.16. (Konstrukcija UMVUE) Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz Poissonovog modela $P(\lambda)$ za $\lambda > 0$, te $\tau(\lambda) = \lambda e^{-\lambda}$. Konstrukcija UMVUE za funkciju $\tau(\lambda)$ bazira se na Rao-Blackwellovom teoremu 4.12. Odredimo prvo potpunu i dovoljnu statistiku za taj model. Gustoća iz tog modela je:

$$f(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \cdot \mathbb{1}_{\mathbb{N}_0}(x) = e^{-\lambda} \cdot e^{(\log\lambda)x} \cdot \frac{1}{x!}\mathbb{1}_{\mathbb{N}_0}(x).$$

Ako stavimo da je $C(\lambda) = e^{-\lambda}$, $h(x) = \frac{1}{x!}\mathbb{1}_{\mathbb{N}_0}(x)$, $Q(\lambda) = Q_1(\lambda) = \log\lambda$ i $t(x) = t_1(x) = x$, tada je $P(\lambda)$ (za $\lambda > 0$) 1-parametarska eksponencijalna familija (vidjeti definiciju 3.23). Funkcija Q preslikava bijektivno parametarski prostor $\langle 0, +\infty\rangle$ na otvoreni interval $\mathbb{R}$, pa je, u skladu s definicijom 3.25, $P(\lambda)$ i familija punog ranga. To znači da je prirodna dovoljna statistika $T = X_1 + X_2 + \cdots + X_n$, potpuna i dovoljna za λ (prema teoremima 3.27 i 3.31). U drugom koraku nađimo barem jedan nepristrani procjenitelj za $\tau(\theta) = \lambda e^{-\lambda}$ konačne varijance. Budući da $\lambda e^{-\lambda}$ možemo inetrpretirati kao vjerojatnost događaja da Poissonova varijabla $X_1 \sim P(\lambda)$ poprimi vrijednost 1, uzmimo da je $S = \mathbb{1}_{\{X_1 = 1\}}$. Ta statistika je Bernoullijeva slučajna varijabla pa je omeđena što povlači da ima sve konačne momente, posebno ima konačnu varijancu. Nadalje, za proizvoljni $\lambda > 0$ je

$$\mathbb{E}_\lambda S = \mathbb{E}_\lambda \mathbb{1}_{\{X_1=1\}} = \mathbb{P}_\theta(X_1 = 1) = \lambda e^{-\lambda} = \tau(\lambda).$$

Dakle, S je nepristrani procjenitelj za $\tau(\lambda)$ konačne varijance. Odredimo regresijsku funkciju u od S u odnosu na T. Budući da je $T = X_1 + X_2 + \cdots + X_n \sim P(n\lambda)$, dovoljno je računati vrijednost te funkcija za $k \in \mathbb{N}_0$. Naime, $f_T(k; \lambda) > 0$ ako i samo ako je $k \in \mathbb{N}_0$ neovisno o λ. Prvo, neka je $k \in \mathbb{N}$. Računamo:

$$\begin{aligned}
u(k) &= \mathbb{E}[S|T = k] = 0 \cdot f_{S|T}(0|k) + 1 \cdot f_{T|S}(1|k) = f_{T|S}(1|k) = \\
&= \frac{\mathbb{P}_\lambda(S = 1, T = k)}{\mathbb{P}_\lambda(T = k)} = \frac{\mathbb{P}_\lambda(X_1 = 1, X_1 + X_2 + \cdots + X_n = k)}{\mathbb{P}_\lambda(T = k)} = \\
&= \frac{\mathbb{P}_\lambda(X_1 = 1, X_2 + \cdots + X_n = k-1)}{\mathbb{P}_\lambda(T = k)} \overset{\text{nez.}}{=} \\
&= \frac{\mathbb{P}_\lambda(X_1 = 1) \cdot \mathbb{P}_\lambda(X_2 + \cdots + X_n = k-1)}{\mathbb{P}_\lambda(T = k)} = \left[\begin{array}{c} X_1 \sim P(\lambda), T \sim P(n\lambda) \\ X_2 + \cdots + X_n \sim P((n-1)\lambda) \end{array}\right] \\
&= \frac{\lambda e^{-\lambda} \cdot \frac{((n-1)\lambda)^{k-1}}{(k-1)!} e^{-(n-1)\lambda}}{\frac{(n\lambda)^k}{k!} e^{-n\lambda}} = \frac{k}{n}\left(1 - \frac{1}{n}\right)^{k-1}
\end{aligned}$$

Ako je $k = 0$, tada je $u(0) = \mathbb{P}(X_1 = 1|T = 0) = 0$, a kako je i $(k/n)(1 - 1/n)^{k-1} = 0$ za $k = 0$, vrijedi da je:

$$u(k) = \frac{k}{n}\left(1 - \frac{1}{n}\right)^{k-1}, \quad k \in \mathbb{N}_0.$$

Budući da je T dovoljna statistika, prema Rao-Blackwellovom teoremu 4.12, $U = u(T)$ je nepristrani procjenitelj za $\tau(\theta)$ konačne varijance, a kako je U i funkcija potpune statistike, prema Lehmann-Schefféovom teoremu 4.13, U je UMVUE od $\tau(\theta)$. □

4.4 Učinkovito procjenjivanje

U slučajevima kada ne postoje ili je teško naći potpune i dovoljne statistike, najbolje procjenitelje tražimo drugim metodama. Jedan od pristupa je da izučavamo problem optimalnog procjenjivanja u parametarskim statističkim modelima koji zadovoljavaju dovoljno *regularne* uvjete. Mi ćemo se ograničiti na izučavanje regularnih statističkih modela za jednodimenzionalne razdiobe koji su parametrizirani jednim (jednodimanzionalnim) parametrom.

Definicija 4.17. *Statistički model $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ za jednodimenzionalne razdiobe je* **regularan model** *ako su zadovoljeni sljedeći uvjeti:*

(R1) *Nosač* $\operatorname{supp} f(\cdot;\theta) = \{x \in \mathbb{R} : f(x;\theta) > 0\}$ *ne ovisi o θ.*
(R2) *Parametarski prostor Θ je otvoreni interval u $\mathbb{R}$.*
(R3) *Za sve $x \in \mathbb{R}$, preslikavanje $\theta \mapsto f(x;\theta)$ je neprekidno diferencijabilno na Θ.*
(R4) *Za* **Fisherovu informaciju***:*

$$I(\theta) := \begin{cases} \int_{\mathbb{R}} \left(\frac{\partial}{\partial\theta}\log f(x;\theta)\right)^2 f(x;\theta)\,dx & \text{ako je model neprekidan,} \\ \sum_{x\in\mathbb{R}} \left(\frac{\partial}{\partial\theta}\log f(x;\theta)\right)^2 f(x;\theta) & \text{ako je model, diskretan} \end{cases}$$

vrijedi da je $0 < I(\theta) < +\infty$ za sve $\theta \in \Theta$.
(R5) *Za svaki $\theta \in \Theta$ vrijedi:*

$$\begin{aligned} 0 &= \tfrac{d}{d\theta}\int_{\mathbb{R}} f(x;\theta)\,dx = \int_{\mathbb{R}} \tfrac{\partial}{\partial\theta} f(x;\theta)\,dx && \text{ako je model neprekidan,} \\ 0 &= \tfrac{d}{d\theta}\sum_{x\in\mathbb{R}} f(x;\theta) = \sum_{x\in\mathbb{R}} \tfrac{\partial}{\partial\theta} f(x;\theta) && \text{ako je model diskretan.} \end{aligned}$$

Napomena 4.18. Prva jednakost u uvjetu (R5) dolazi od činjenice da je $\int_{\mathbb{R}} f(x;\theta)\,dx = 1$ u neprekidnom slučaju, odn. $\sum_{x\in\mathbb{R}} f(x;\theta) = 1$ u diskretnom slučaju budući da se radi o vjerojatnosnim gustoćama. Druga jednakost je, u stvari, uvjet koji se formulira u (R5): da derivacija po θ i integral (suma) gustoće po x, komutiraju.
Uz koje uvjete na $f(\cdot;\theta)$ će derivacija po θ i integral po x komutirati? Očito je da to vrijedi u slučaju kada se radi o diskretnom regularnom modelu za kojeg je nosač $\operatorname{supp} f(\cdot;\theta)$ konačan skup. Naime, konačna suma i operator deriviranja komutiraju. U slučaju da se radi o neprekidnom regularnom modelu ili je nosač gustoće u diskretnom modelu prebrojivo beskonačan skup, da bi derivacija i integral komutirali, trebamo dodatne uvjete.
Neka je X slučajna varijabla iz regularnog modela, dakle, kojoj je gustoća $f(\cdot;\theta)$ za $\theta \in \Theta$. Budući da vrijede uvjeti (R1-3), dovoljno je pretpostaviti da *za svaki $\theta_0 \in \Theta$ postoji ε-okolina od θ_0 u Θ i nenegativna Borelova funkcija $K_{\theta_0} : \mathbb{R} \to \mathbb{R}$ takvi da je $\mathbb{E}_{\theta_0} K_{\theta_0}(X) < +\infty$ te da vrijedi:*

$$(\forall\theta \in \langle\theta_0 - \varepsilon, \theta_0 + \varepsilon\rangle)(\forall x \in \operatorname{supp} f(\cdot;\theta_0)) \quad \left|\frac{\partial}{\partial\theta} f(x;\theta)\right| \le K_{\theta_0}(x) f(x;\theta_0). \tag{4.6}$$

Tada vrijedi uvjet (R5). Pokažimo to. Neka su $x \in \mathrm{supp} f(\cdot;\theta_0)$ i $\theta \in \langle \theta_0 - \varepsilon, \theta_0 + \varepsilon \rangle \subseteq \Theta$ proizvoljni. Budući da vrijedi (R3), prema Lagrangeovu teremu srednje vrijednosti postoji $\vartheta_{\theta,x} \in \langle \theta_0, \theta \rangle \subset \langle \theta_0 - \varepsilon, \theta_0 + \varepsilon \rangle$ takva točka da je

$$f(x;\theta) - f(x;\theta_0) = \frac{\partial}{\partial \theta} f(x; \vartheta_{\theta,x})(\theta - \theta_0)$$

odakle slijedi da je

$$\left| \frac{f(x;\theta) - f(x;\theta_0)}{\theta - \theta_0} \right| = \left| \frac{\partial}{\partial \theta} f(x; \vartheta_{\theta,x}) \right| \overset{\text{pretp.}}{\leq} K_{\theta_0}(x) f(x;\theta_0).$$

Neka je sada (θ_n) niz u okolini $\langle \theta_0 - \varepsilon, \theta_0 + \varepsilon \rangle$ takav da je $\lim_n \theta_n = \theta_0$. Primijetite da je tada i $\lim_n \vartheta_{\theta_n,x} = \theta_0$ po teoremu o sandviču, a kako je po (R3) $\theta \mapsto f(x;\theta)$ diferencijabilna funkcija i klase je C^1 (tj. derivacija $\theta \mapsto \frac{\partial}{\partial \theta} f(x;\theta)$ je neprekidna funkcija) na Θ, slijedi da je

$$\lim_n \frac{f(x;\theta_n) - f(x;\theta_0)}{\theta_n - \theta_0} = \lim_n \frac{\partial}{\partial \theta} f(x; \vartheta_{\theta_n,x}) = \frac{\partial}{\partial \theta} f(x;\theta_0).$$

Budući da su ti nizovi dominirani sa integrabilnom funkcijom $K_{\theta_0} f(\cdot;\theta_0)$ po pretpostavci, tj.

$$\left\{ \begin{array}{ll} \int_{\mathbb{R}} K_{\theta_0}(x) f(x;\theta_0)\, dx & \text{ako je razdioba neprekidna} \\ \sum_{x \in \mathbb{R}} K_{\theta_0}(x) f(x;\theta_0) & \text{ako je razdioba diskretna} \end{array} \right\} = \mathbb{E}_{\theta_0}[K_{\theta_0}(X)] \overset{\text{pretp.}}{<} +\infty,$$

prema Lebesgueovom teoremu o dominiranoj konvergenciji (LTDK) slijedi da je (za slučaj neprekidne distribucije):

$$\begin{aligned} \frac{d}{d\theta} \int_{\mathbb{R}} f(x;\theta_0)\, dx &= \lim_n \frac{1}{\theta_n - \theta_0} \left(\int_{\mathbb{R}} f(x;\theta_n)\, dx - \int_{\mathbb{R}} f(x;\theta_0) \right) \overset{\text{lin.int.}}{=} \\ &= \lim_n \int_{\mathbb{R}} \left(\frac{f(x;\theta_n) - f(x;\theta_0)}{\theta_n - \theta_0} \right) dx \overset{\text{LTDK}}{=} \int_{\mathbb{R}} \frac{\partial}{\partial \theta} f(x;\theta_0)\, dx. \end{aligned}$$

Istim načinom zaključivanja tvrdnja (R5) vrijedi i za slučaj diskretnog modela s prebrojivo beskonačnim nosačem. □

Od sada pa nadalje, sve tvrdnje će se zapisivati kao da se radi o neprekidnom statističkom modelu iako će vrijediti i za diskretni model (osim ako se ne naglasi drugačije).

Navedimo neka svojstva Fisherove informacije navedene u uvjetu (R4). Primijetite da, ako je X slučajna varijabla kojoj je gustoća $f(\cdot;\theta)$, da je tada Fisherova informacija jednaka:

$$I(\theta) = \mathbb{E}_\theta \left[\left(\frac{\partial}{\partial \theta} \log f(X;\theta) \right)^2 \right].$$

Uz pretpostavku da vrijedi (R5), slijedi da je

$$\mathbb{E}_\theta \left[\frac{\partial}{\partial \theta} \log f(X;\theta) \right] = \int_{\mathrm{supp}\, f(\cdot;\theta)} \frac{\partial}{\partial \theta} \log f(x;\theta) \cdot f(x;\theta)\, dx = \int_{\mathbb{R}} \frac{\partial}{\partial \theta} f(x;\theta)\, dx \overset{\text{(R5)}}{=} 0. \tag{4.7}$$

Zbog toga se Fisherova informacija može interpretirati i kao varijanca:

$$I(\theta) = \mathbb{V}\mathrm{ar}_\theta \left[\frac{\partial}{\partial \theta} \log f(X;\theta) \right]. \tag{4.8}$$

Neka je sada $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz regularnog modela, te neka je

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta), \quad \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n,$$

funkcija gustoće od $\mathbf{X}$ parametrizirana s $\theta \in \Theta$. Definiramo **Fisherovu informaciju obzirom na uzorak X** sa

$$I_n(\theta) := \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right)^2\right]. \tag{4.9}$$

Budući da je slučajna varijabla

$$\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i;\theta), \tag{4.10}$$

slijedi da je, zbog linearnosti matematičkog očekivanja (i) teorem 1.56:

$$\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right] \overset{\text{lin.m.o}}{=} \sum_{i=1}^{n}\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f(X_i;\theta)\right] \overset{(4.7)}{=} 0. \tag{4.11}$$

Zbog toga se i Fisherova informacija obzirom na uzorak može interpretirati kao varijanca zbroja (4.10) nezavisnih i jednako distribuiranih slučajnih varijabli. Računamo:

$$\begin{aligned}
I_n(\theta) &= \mathbb{V}\mathrm{ar}_\theta\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right] \overset{(4.10)}{=} \mathbb{V}\mathrm{ar}_\theta\left[\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i;\theta)\right] \overset{\text{nez. i tm.1.61}\,(v)}{=} \\
&= \sum_{i=1}^{n}\mathbb{V}\mathrm{ar}_\theta\left[\frac{\partial}{\partial\theta}\log f(X_i;\theta)\right] \overset{\text{j.d.}}{=} n\mathbb{V}\mathrm{ar}_\theta\left[\frac{\partial}{\partial\theta}\log f(X_1;\theta)\right] \overset{(4.8)}{=} \\
&= nI(\theta).
\end{aligned} \tag{4.12}$$

Navedimo sada osnovni rezultat koji daje donju granicu za varijancu nepristranog procjenitelja u regularnom modelu.

Teorem 4.19 (Cramér-Rao)**.** *Neka je $\mathbf{X} = (X_1, X_2, \dots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz regularnog modela parametriziranog s θ. Neka su $\tau(\theta)$ procjenjiva derivabilna funkcija i $T = t(\mathbf{X})$ nepristrani procjenitelj za $\tau(\theta)$ konačne varijance, takvi da za sve $\theta \in \Theta$ vrijedi da je:*

$$\tau'(\theta) = \int_{\mathbb{R}^n} t(\mathbf{x})\frac{\partial}{\partial\theta}f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x}. \tag{4.13}$$

Tada je

$$(\forall\theta \in \Theta)\ \ \mathbb{V}\mathrm{ar}_\theta T \geq \frac{(\tau'(\theta))^2}{nI(\theta)}. \tag{4.14}$$

Donju granicu za varijancu iz (4.14) zovemo **donja Cramér-Raova granica**. Oznaka:

$$\mathrm{DCRG}(\tau) \equiv \frac{(\tau'(\theta))^2}{nI(\theta)}.$$

Dokaz. Pokazat ćemo da je nejednakost (4.14) direktna posljedica Schwarz-Cauchyjeve nejednakosti u prostoru $\mathcal{L}^2(\mathbb{P}_\theta)$ primijenjene na slučajne varijable $T - \tau(\theta)$ i $\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)$ budući da su obje varijable u tom prostoru: $T - \tau(\theta) \in \mathcal{L}^2(\mathbb{P}_\theta)$ po pretpostavci, a da je

$\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta) \in \mathcal{L}^2(\mathbb{P}_\theta)$ slijedi iz (4.12) i uvjeta (R4). Dakle, prema teoremu 2.2 vrijedi da je:

$$\left(\mathbb{E}_\theta\left[(T-\tau(\theta))\cdot\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right]\right)^2 \le \mathbb{E}_\theta[(T-\tau(\theta))^2]\cdot\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right)^2\right]. \quad (4.15)$$

Budući da je T nepristrani procjenitelj za $\tau(\theta)$, vrijedi da je

$$\mathbb{E}_\theta[(T-\tau(\theta))^2] = \mathbb{V}\mathrm{ar}_\theta T.$$

Prema (4.9) i (4.12) je:

$$\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right)^2\right] = nI(\theta).$$

Ostaje još izračunati:

$$\begin{aligned}
& \mathbb{E}_\theta\left[(T-\tau(\theta))\cdot\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right] \overset{\text{lin.m.o.}}{=} \\
= \;& \mathbb{E}_\theta\left[T\cdot\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right] - \tau(\theta)\cdot\underbrace{\mathbb{E}_\theta[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)]}_{\overset{(4.11)}{=}0} = \\
= \;& \mathbb{E}_\theta\left[t(\mathbf{X})\cdot\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X};\theta)\right] = \int\limits_{\operatorname{supp} f_{\mathbf{X}}(\cdot;\theta)} t(\mathbf{x})\cdot\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{x};\theta)\cdot f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x} = \\
= \;& \int\limits_{\operatorname{supp} f_{\mathbf{X}}(\cdot;\theta)} t(\mathbf{x})\cdot\frac{\frac{\partial}{\partial\theta}f_{\mathbf{X}}(\mathbf{x};\theta)}{f_{\mathbf{X}}(\mathbf{x};\theta)}\cdot f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x} = \int\limits_{\mathbb{R}^n} t(\mathbf{x})\cdot\frac{\partial}{\partial\theta}f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x} \overset{(4.13)}{=} \\
= \;& \tau'(\theta).
\end{aligned}$$

Dakle, (4.15) je ekvivalentno sa

$$(\tau'(\theta))^2 \le \mathbb{V}\mathrm{ar}_\theta T\cdot nI(\theta).$$

Budući da je $I(\theta) > 0$ prema (R4), te da je $\theta\in\Theta$ bio proizvoljan, slijedi (4.14). □

Napomena 4.20. Uvjet (4.13) iz iskaza teorema je zapravo uvjet da derivacija po θ i integral (suma u diskretnom slučaju) po x komutiraju, slično kao i uvjet regularnosti (R5). Dakle, (4.13) je ekvivalentno sa:

$$\frac{d}{d\theta}\int\limits_{\mathbb{R}^n} t(\mathbf{x})f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x} = \int\limits_{\mathbb{R}^n} t(\mathbf{x})\cdot\frac{\partial}{\partial\theta}f_{\mathbf{X}}(\mathbf{x};\theta)\,d\mathbf{x}.$$

Ako se radi o diskretnom modelu s konačnim nosačem gustoće, tada će taj uvjet biti zadovoljen za svaki nepristrani procjenitelj $T = t(\mathbf{X})$ derivabilne funkcije $\tau(\theta)$. U slučaju da je model neprekidan ili diskretan s prebrojivo beskonačnim nosačem gustoće, tada taj uvjet neće nužno uvijek biti zadovoljen. Ako za regularan statistički model, slično kao u napomeni 4.18, pretpostavimo sljedeće: *za svaki $\theta_0\in\Theta$ postoji ε-okolina od θ_0 u Θ i nenegativna Borelova funkcija $K_{n\theta_0}:\mathbb{R}^n\to\mathbb{R}$ takvi da je $\mathbb{E}_{\theta_0}K_{n\theta_0}(\mathbf{X}) < +\infty$ te da vrijedi:*

$$(\forall\theta\in\langle\theta_0-\varepsilon,\theta_0+\varepsilon\rangle)(\forall\mathbf{x}\in\operatorname{supp} f_{\mathbf{X}}(\cdot;\theta_0))\quad\left|t(\mathbf{x})\frac{\partial}{\partial\theta}f_{\mathbf{X}}(\mathbf{x};\theta)\right| \le K_{n\theta_0}(\mathbf{x})f_{\mathbf{X}}(\mathbf{x};\theta_0), \quad (4.16)$$

tada će uvjet (4.13) biti ispunjen. □

Napomena 4.21. Budući da je nejednakost (4.14) ekvivalentna sa Schwarz-Cauchyjevom nejednakošću, onda u njoj jednakost vrijedi ako i samo ako su za svaki $\theta \in \Theta$ slučajne varijable $T - \tau(\theta) \equiv t(\mathbf{X}) - \tau(\theta)$ i $\frac{\partial}{\partial\theta} \log f_{\mathbf{X}}(\mathbf{X}; \theta)$ $\mathbb{P}_\theta$-g.s. kolinearne (vidjeti teorem 2.2), tj. ako i samo ako postoji konstanta $k(\theta)$ takva da je

$$t(\mathbf{X}) - \tau(\theta) = k(\theta) \cdot \frac{\partial}{\partial\theta} \log f_{\mathbf{X}}(\mathbf{X}; \theta) \ \ \mathbb{P}_\theta\text{-g.s.}$$

Na primjer, ako za svaki $\theta \in \Theta$ postoji broj $k(\theta) \in \mathbb{R}$ takav da za sve $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \operatorname{supp} f_{\mathbf{X}}(\cdot; \theta)$ vrijedi da je

$$t(\mathbf{x}) = \tau(\theta) + k(\theta) \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i; \theta), \tag{4.17}$$

tada je uvjet g.s. kolinearnosti ispunjen, pa u (4.14) vrijedi jednakost. □

Nepristrani procjenitelji kojima varijanca dostiže donju Cramér-Raovu granicu su dobri procjenitelji.

Neka je $\mathbf{X}$ slučajni uzorak iz regularnog modela parametriziranog s θ ili modela u kojem je barem zadovoljen uvjet regularnosti (R4).

Definicija 4.22. *Procjenitelj $T = \tau(\mathbf{X})$ je* **učinkovit** *ili* **efikasan procjenitelj** *za procjenjivu i derivabilnu funkciju $\tau(\theta)$ ako vrijedi sljedeće:*

(i) T je nepristrani procjenitelj za $\tau(\theta)$ konačne varijance;
(ii) Za svaki $\theta \in \Theta$ vrijedi da je $\mathbb{V}\mathrm{ar}_\theta T = \frac{(\tau'(\theta))^2}{nI(\theta)}$.

Ako je u nekom regularnom modelu za svaki nepristrani procjenitelj neke derivabilne funkcije zadovoljen uvjet (4.13), tada je efikasan procjenitelj te funkcije ujedno i UMVUE. Na primjer, to vrijedi u diskretnim regularnim modelima s konačni nosačem gustoća. S druge strane, UMVUE nije nužno i efikasan procjenitelj iste funkcije.

Primjer 4.23. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine $n \geq 1$ iz Bernoullijevog modela $b(1, \theta)$, $\theta \in \langle 0, 1 \rangle$. Primijetimo da je taj model regularan: nosač svih gustoća

$$f(x; \theta) = \theta^x (1 - \theta)^{1-x} \cdot \mathbb{1}_{\{0,1\}}(x)$$

je jednak $\operatorname{supp} f(\cdot; \theta) = \{0, 1\}$ pa očito ne ovisi o θ. Dakle, zadovoljen je uvjet (R1). Nadalje, parametarski prostor je $\Theta = \langle 0, 1 \rangle$ što je otvoren interval, pa vrijedi i (R2). Za $x = 1$ je $\theta \mapsto f(1; \theta) = \theta$, za $x = 0$ je $\theta \mapsto f(0; \theta) = 1 - \theta$, te za $x \notin \{0, 1\}$ je $\theta \mapsto f(x; \theta) \equiv 0$ što su sve funkcije klase C^1 na intervalu Θ. Dakle, vrijedi (R3). Tada vrijedi i (R5) jer je model diskretan s konačnim nosačem (vidjeti napomenu 4.18). Preostaje provjeriti je li zadovoljen uvjet (R4). Za $x \in \{0, 1\}$ je

$$\frac{\partial}{\partial\theta} \log f(x; \theta) = \frac{\partial}{\partial\theta} (x \log\theta + (1 - x) \log(1 - \theta)) = \frac{x}{\theta} - \frac{1 - x}{1 - \theta} = \frac{x - \theta}{\theta(1 - \theta)}$$

pa je za slučajnu varijablu X s gustoćom $f(\cdot; \theta)$:

$$I(\theta) = \mathbb{E}_\theta \left[\left(\frac{X - \theta}{\theta(1 - \theta)} \right)^2 \right] \overset{\text{lin.m.o.}}{=} \frac{\mathbb{E}_\theta[(X - \theta)^2]}{\theta^2(1 - \theta)^2} = \frac{\mathbb{V}\mathrm{ar}_\theta X}{\theta^2(1 - \theta)^2} = \frac{1}{\theta(1 - \theta)}.$$

Budući da je $I(\theta) = 1/(\theta(1 - \theta)) \in \langle 0, +\infty \rangle$ za svaki $\theta \in \langle 0, 1 \rangle = \Theta$, vrijedi i (R4).
Budući da se radi o diskretnom regularnom modelu s konačnim nosačem gustoće, za svaki

nepristrani procjenitelj derivabilne funkcije od θ je zadovoljen uvjet (4.13) iz Cramér-Raovog teorema 4.19 (vidjeti napomenu 4.20) pa je efikasan procjenitelj ujedno i UMVUE za tu funkciju od θ. Na primjer, za funkciju $\tau(\theta) = \theta$ je jedan nepristrani procjenitelj relativna frekvencija uspjeha $T = \overline{X}_n$. Budući da je $\tau'(\theta) = 1$ i stoga

$$\mathbb{V}\mathrm{ar}_\theta \overline{X}_n = \frac{1}{n}\mathbb{V}\mathrm{ar}_\theta X_1 = \frac{\theta(1-\theta)}{n} = \frac{1^2}{n \cdot \frac{1}{\theta(1-\theta)}} = \frac{(\tau'(\theta))^2}{nI(\theta)},$$

relativna frekvencija uspjeha je efikasan procjenitelj, a time onda i UMVUE za θ. □

Primjer 4.24. (Procjenitelj koji je UMVUE, a nije efikasan) Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine n ($n \geq 2$) iz Poissonovog modela $P(\lambda)$, $\lambda > 0$. Želimo konstruirati UMVUE za funkciju $\tau(\lambda) = e^{-\lambda}$. Konstrukciju ćemo sprovesti na način kako je to ilustrirano u primjeru 4.16. U tom primjeru je pokazano da je $T = X_1 + X_2 + \cdots + X_n$ potpuna i dovoljna statistika za λ. Uzmimo statistiku $S = \mathbb{1}_{\{X_1=0\}}$. Tada je S nepristrani procjenitelj za $e^{-\lambda}$: za sve $\lambda > 0$ je

$$\mathbb{E}_\lambda S = \mathbb{P}_\lambda(X_1 = 0) \overset{X_1 \sim P(\lambda)}{=} e^{-\lambda},$$

a kako je ograničena slučajna varijabla, konačne je varijance. Nije teško pokazati da je regresijska funkcija od S u odnosu na T, za $k \in \mathbb{N}_0$, jednaka:

$$u(k) = \mathbb{E}[S|T = k] = f_{S|T}(1|k) = \frac{\mathbb{P}_\lambda(X_1 = 0, X_2 + \cdots + X_n = k)}{\mathbb{P}(T = k)} = \left(1 - \frac{1}{n}\right)^k.$$

Budući da je T dovoljna statistika, iz Rao-Blackwellovog teorema 4.12 slijedi da je $U = (1 - 1/n)^T$ nepristrani procjenitelj za $e^{-\lambda}$ konačne varijance koji je i funkcija potpune i dovoljne statistike T, pa, prema Lehmann-Schefféovom teoremu 4.13 slijedi da je U UMVUE od $e^{-\lambda}$.

Pokažimo da je Poissonov model regularan. (R1) je zadovoljen jer je nosač gustoće Poissonove razdiobe jednak $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ neovisno o vrijednosti parametra. Budući da je parametarski prostor otvoreni interval $\langle 0, +\infty \rangle$, vrijedi i (R2). Za $k \in \mathbb{N}_0$ je funkcija $\lambda \mapsto e^{-\lambda}\lambda^k/k!$ neprekidno diferencijabilna na $\langle 0, +\infty \rangle$, pa vrijedi i (R3). Provjerimo uvjet (R4). Neka je $\lambda > 0$ proizvoljan i $X \sim P(\lambda)$. Budući da je za $k \in \mathbb{N}_0$:

$$\frac{\partial}{\partial \lambda} \log f(k; \lambda) = \frac{\partial}{\partial \lambda} \log \left(\frac{\lambda^k}{k!} e^{-\lambda} \right) = \frac{\partial}{\partial \lambda}(k \log \lambda - \lambda - \log(k!)) = \frac{k}{\lambda} - 1 = \frac{1}{\lambda}(k - \lambda),$$

te da je $\mathbb{E}_\lambda X = \lambda = \mathbb{V}\mathrm{ar}_\lambda X$, slijedi da je

$$I(\lambda) = \mathbb{E}_\lambda \left[\frac{1}{\lambda^2}(X - \lambda)^2 \right] \overset{\text{lin.m.o.}}{=} \frac{1}{\lambda^2}\mathbb{E}_\lambda[(X - \lambda)^2] = \frac{1}{\lambda^2}\mathbb{V}\mathrm{ar}_\lambda X = \frac{1}{\lambda},$$

dakle, $I(\lambda) = 1/\lambda \in \langle 0, +\infty \rangle$. Time je pokazano da vrijedi (R4). Pokažimo da vrijedi uvjet (R5). Budući da se radi o diskretnom modelu s prebrojivo beskonačnim nosačem, moramo računati:

$$\sum_{k=0}^{+\infty} \frac{\partial}{\partial \lambda} f(k; \lambda) = \sum_{k=0}^{+\infty} \left(\frac{k\lambda^{k-1}}{k!} - \frac{\lambda^k}{k!} \right) e^{-\lambda} = \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} - \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1 - 1 = 0.$$

Primijetite da je račun dobar jer je

$$\sum_{k=0}^{+\infty} \left| \frac{\partial}{\partial \lambda} f(k; \lambda) \right| \leq \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} + \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1 + 1 = 2 < +\infty.$$

Dakle, vrijedi i (R5).
Za funkciju $\tau(\lambda) = e^{-\lambda}$ izračunajmo donju Cramér-Raovu granicu:

$$\mathrm{DCRG}(\tau) = \frac{(\tau'(\lambda))^2}{nI(\lambda)} = \frac{\lambda}{n}e^{-2\lambda}.$$

Da bismo provjerili je li UMVUE U efikasan procjenitelj za $e^{-\lambda}$, dovoljno je provjeriti je li mu varijanca jednaka donjoj Cramér-Raovoj granici (da je nepristrani procjenitelj za tu funkciju konačne varijance znamo jer je UMVUE). Prije nego što izračunamo varijancu od U, za konstantu $a > 0$ i proizvoljni $\lambda > 0$ izračunajmo:

$$\mathbb{E}_\lambda[a^T] \overset{T\sim P(n\lambda)}{=} \sum_{k=0}^{+\infty} a^k \cdot \frac{(n\lambda)^k}{k!}e^{-n\lambda} = e^{-n\lambda}\sum_{k=0}^{+\infty}\frac{(na\lambda)^k}{k!} = e^{-n\lambda}\cdot e^{na\lambda} = e^{-(1-a)n\lambda}. \tag{4.18}$$

Sada izračunajmo varijancu od U:

$$\begin{aligned}
\mathbb{V}\mathrm{ar}_\lambda U &\overset{\text{tm.1.61}(i)}{=} \mathbb{E}_\lambda[U^2] - (\mathbb{E}_\lambda U)^2 = \mathbb{E}_\lambda\left[\left(\left(1-\frac{1}{n}\right)^T\right)^2\right] - e^{-2\lambda} = \\
&= \mathbb{E}_\lambda\left[\left(\left(1-\frac{1}{n}\right)^2\right)^T\right] - e^{-2\lambda} \overset{(4.18)}{=} \\
&= \exp\left(-\left(1-\left(1-\frac{1}{n}\right)^2\right)n\lambda\right) - e^{-2\lambda} = e^{-2\lambda+\frac{\lambda}{n}} - e^{-2\lambda} = \\
&= \left(e^{\frac{\lambda}{n}}-1\right)e^{-2\lambda}.
\end{aligned}$$

Primijetite da smo u predzadnjoj jednakosti iskoristili identitet (4.18) prijenjen na $a = (1-1/n)^2$. Ispitujući tok eksponencijalne funkcije $x \mapsto e^x$ možemo zaključiti da je za sve realne brojeve $x \neq 0$, $e^x - 1 > x$. Ako tu nejednakost primijenimo na $x = \lambda/n > 0$, tada dobijemo da je

$$\mathbb{V}\mathrm{ar}_\lambda U = \left(e^{\frac{\lambda}{n}}-1\right)e^{-2\lambda} > \frac{\lambda}{n}e^{-2\lambda} = \mathrm{DCRG}(\tau).$$

Dakle, varijanca od UMVUE U je strogo veća od donje Cramér-Raove granice, pa $U = (1-1/n)^T$ nije efikasan procjenitelj za $e^{-\lambda}$ u Poissonovom modelu. □

4.5 Procjena metodom najveće vjerodostojnosti

Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajan uzorak duljine n ($n \geq 1$) iz statističkog modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$. U ovom potpoglavlju dozvoljavamo općenitu situaciju: da je $f(\cdot;\theta)$ iz množine $\mathcal{P}$ gustoća slučajne veličine dimenzije d ($d \geq 1$) parametrizirane parametrom θ dimenzije m ($m \geq 1$).

Ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn}$ jedna realizacija od $\mathbf{X}$, tada je **vjerodostojnost** (engl. *likelihood*) funkcija $L : \Theta \to \mathbb{R}$ definirana sa

$$L(\theta) \equiv L(\theta|\mathbf{x}) := f_{\mathbf{X}}(\mathbf{x};\theta) = \prod_{i=1}^n f(x_i;\theta), \quad \theta \in \Theta.$$

Za zadanu vrijednost $\theta \in \Theta$, broj $L(\theta) \equiv L(\theta|\mathbf{x})$ zovemo *vjerodostojnost vrijednosti θ parametra na osnovi opaženog uzorka* $\mathbf{x}$.

Definicija 4.25. *Statistika* $\hat{\theta} \equiv \hat{\theta}(\mathbf{X})$ *je* **procjenitelj najveće** *(ili* **maksimalne***)* **vjerodostojnosti** *(engl.* maximum likelihood estimator*, MLE) ako vrijedi*

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta|\mathbf{X}). \tag{4.19}$$

Ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn}$ opažena vrijednost slučajnog uzorka $\mathbf{X}$ i $\hat{\theta} \equiv \hat{\theta}(\mathbf{X})$ procjenitelj za θ metodom maksimalne vjerodostojnosti, tada je *procjena metodom maksimalne vjerodostojnosti od* θ *na osnovi opaženog uzorka* jednaka $\hat{\theta}(\mathbf{x})$. Ako je iz konteksta jasno, procjenu $\hat{\theta}(\mathbf{x})$ ćemo često (također kao i odgovarajući procjenitelj) jednostavno označavati s $\hat{\theta}$ i imenovati kraticom MLE (od engl. *maximum likelihood estimate*).

Primjer 4.26. Neka je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}_+^n$ realizacije slučajnog uzorka $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ duljine n $(n \geq 1)$ iz neprekidnog uniformnog modela $U(0, \theta)$, $\theta > 0$. Funkcija vjerodostojnosti je:

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot \mathbb{1}_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \cdot \mathbb{1}_{[x_{(n)}, +\infty\rangle}(\theta), \quad \theta > 0,$$

gdje je $x_{(n)} = \max_{1 \leq i \leq n} x_i$. Tada je MLE od θ jednak $\hat{\theta}(\mathbf{X}) = X_{(n)}$. Naime, budući da je $\theta \mapsto 1/\theta^n$ strogo padajuća funkcija na $\mathbb{R}_+ \equiv \langle 0, +\infty\rangle$, te da su $X_i > 0$ $\mathbb{P}_\theta$-g.s. za sve $\theta > 0$, vrijedi da je

$$L(X_{(n)}) = \frac{1}{(X_{(n)})^n} = \max_{\theta > 0} \frac{1}{\theta^n} \cdot \mathbb{1}_{[X_{(n)}, +\infty\rangle}(\theta) = \max_{\theta > 0} L(\theta|\mathbf{X}).$$

Primijetite da je MLE jedinstven, te da je jednak dovoljnoj statistici za θ (vidjeti primjer 3.32). $\square$

Primjer 4.27. Neka je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ realizacije slučajnog uzorka $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ duljine n $(n \geq 1)$ iz normalnog modela $N(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \langle 0, +\infty\rangle =: \Theta$. Funkcija vjerodostojnosti je:

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2} = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right), \quad \theta \in \Theta.$$

Nosač gustoće je jednak $\mathbb{R}$. Zbog toga je jednostavnije gledati *log-vjerodostojnost*:

$$\ell(\theta) \equiv \ell(\theta|\mathbf{x}) := \log L(\theta|\mathbf{x}) = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 - \frac{n}{2} \log \sigma^2 + C, \quad \theta = (\mu, \sigma^2) \in \Theta,$$

gdje je C konstanta koja ne ovisi o parametrima μ i σ^2. Naime, $\hat{\theta} \in \Theta$ je točka maksimuma od L na Θ ako i samo ako je to točka maksimuma od ℓ na istoj domeni, budući da je prirodni logaritam strogo rastuća funkcija. Nadalje, za svaki $x \in \mathbb{R}$, $\theta \mapsto f(x; \theta)$ je pozitivna funkcija klase C^2 na otvorenom i konveksnom skupu $\Theta = \mathbb{R} \times \langle 0, +\infty\rangle$. Prema tome je log-vjerodostojnost $\ell \in C^2(\Theta)$. Budući da je Θ otvoren i konveksan skup, ako postoji točka maksimuma od funkcije ℓ, tada je ona stacionarna točka te funkcije. Dakle, točku maksimuma log-vjerodostojnosti tražimo među stacionarnim točkama. Riješimo sustav jednadžbi (stavimo da je $v \equiv \sigma^2$):

$$\left.\begin{array}{l} \frac{\partial}{\partial \mu} \ell(\mu, v) = \frac{1}{v} \sum_{i=1}^{n} (x_i - \mu) = 0 \\ \frac{\partial}{\partial v} \ell(\mu, v) = \frac{1}{2v^2} \sum_{i=1}^{n} (x_i - \mu)^2 - \frac{n}{2v} = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x} \\ \hat{\sigma}^2 = \hat{v} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 = \frac{n-1}{n} s^2, \end{array}\right.$$

dakle, jedinstvena stacionarna točka je $(\overline{x}, \frac{n-1}{n}s^2)$. Je li to točka *globalnog* maksimuma od ℓ na domeni Θ? Budući da hessian (tj. matrica druge derivacije od ℓ) nije negativno definitna matrica na Θ, ne možemo zaključiti da je ℓ konkavna funkcija. Zbog toga se drugim metodama mora zaključiti da je ta jedinstvena stacionarna točka ujedno i točka globalnog maksimuma. Preciznije, može se pokazati da vrijedi:

$$(\forall \mu \in \mathbb{R})(\forall \sigma^2 > 0) \quad \ell(\mu, \sigma^2) \leq \ell(\hat{\mu}, \sigma^2) \leq \ell(\hat{\mu}, \hat{\sigma}^2).$$

Naime, za prvu nejednakost, fiksirajte vrijednost parametra $\sigma^2 \equiv v$ i na log-vjerodostojnost ℓ gledajte kao na funkciju $\mu \mapsto \ell(\mu, v)$, dakle kao na funkciju samo od μ na $\mathbb{R}$. Ta funkcija je konkavna kvadratna funkcija po μ pa ima jedinstvenu točku globalnog maksimuma. Pokazuje se da je ta točka $\hat{\mu} = \overline{x}$ neovisno o fiksiranoj vrijednosti od v. Za drugu nejednakost, uvrstite u izraz za log-vjerodostojnost dobivenu procjenu $\hat{\mu} = \overline{x}$ za μ. Pokazuje se da tako dobivena funkcija $v \mapsto \ell(\hat{\mu}, v)$ strogo raste na intervalu $\langle 0, \hat{v}]$, gdje je $\hat{v} = (n-1)s^2/n$, te strogo pada na intervalu $[\hat{v}, +\infty\rangle$. Dakle, $\hat{v} \equiv \hat{\sigma}^2$ je točka globalnog maksimuma te funkcije.
Prema tome, $\hat{\theta} = (\overline{X}_n, \frac{n-1}{n}S_n^2)$ je MLE za $\theta = (\mu, \sigma^2)$. Primijetite da je i u ovom primjeru MLE jedinstven, te da je funkcija dovoljne statistike $T = (\overline{X}_n, S_n^2)$ za θ (vidjeti primjer 3.10). □

Primijetimo da su u oba prethodna primjera, procjenitelji metodom maksimalne vjerodostojnosti jedinstveni, te da su *funkcije* dovoljnih statistika. To nije slučajno.

Teorem 4.28. *Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine $n \geq 1$ iz statističkog modela parametriziranog parametrom θ, te neka je $T = t(\mathbf{X})$ dovoljna statistika za θ. Ako procjenitelj metodom maksimalne vjerodostojnosti postoji i jedinstven je, tada je funkcija dovoljne statistike T.*

Dokaz. Budući da je $T = t(\mathbf{X})$ dovoljna statistika za θ, prema Neyman-Fisherovom teoremu o faktorizaciji 3.9 vrijedi da je za $\mathbf{x} \in \mathbb{R}^{dn}$ i $\theta \in \Theta$:

$$f_{\mathbf{X}}(\mathbf{x}; \theta) = h(\mathbf{x}) g_\theta(t(\mathbf{x})) \equiv L(\theta|\mathbf{x})$$

za neke nenegativne izmjerive funkcije h i g_θ. Po definiciji, za MLE $\hat{\theta}(\mathbf{x})$ vrijedi da je:

$$h(\mathbf{x}) g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{x})) = L(\hat{\theta}(\mathbf{x})|\mathbf{x}) = \max_{\theta \in \Theta} L(\theta|\mathbf{x}) = \max_{\theta \in \Theta} h(\mathbf{x}) g_\theta(t(\mathbf{x})) = h(\mathbf{x}) \max_{\theta \in \Theta} g_\theta(t(\mathbf{x})).$$

Budući da gornja jednakost ima smisla samo za $\mathbf{x} \in \operatorname{supp} f_{\mathbf{X}}(\cdot; \theta)$ za neki $\theta \in \Theta$, a što povlači da je nužno $h(\mathbf{x}) > 0$, slijedi da je

$$g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{x})) = \max_{\theta \in \Theta} g_\theta(t(\mathbf{x})) \tag{4.20}$$

kadagod je $h(\mathbf{x}) > 0$. Cilj nam je pokazati da je MLE funkcija od $t(\mathbf{x})$, tj. da $\hat{\theta}(\mathbf{x})$ ovisi o $\mathbf{x}$ preko $t(\mathbf{x})$. Neka su $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{dn}$ takvi da je $h(\mathbf{x}) > 0$ i $h(\mathbf{y}) > 0$, te da je $t(\mathbf{x}) = t(\mathbf{y})$. Ako pokažemo da je tada nužno i $\hat{\theta}(\mathbf{x}) = \hat{\theta}(\mathbf{y})$, tada je MLE $\hat{\theta}(\mathbf{x})$ funkcija od $t(\mathbf{x})$. Budući da su $\hat{\theta}(\mathbf{y}), \hat{\theta}(\mathbf{x}) \in \Theta$, slijedi da je

$$g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{x})) \overset{(4.20)}{\geq} g_{\hat{\theta}(\mathbf{y})}(t(\mathbf{x})) \overset{t(\mathbf{x})=t(\mathbf{y})}{=} g_{\hat{\theta}(\mathbf{y})}(t(\mathbf{y})) \overset{(4.20)}{\geq} g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{y})) \overset{t(\mathbf{y})=t(\mathbf{x})}{=} g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{x})).$$

Dakle,

$$g_{\hat{\theta}(\mathbf{x})}(t(\mathbf{x})) = g_{\hat{\theta}(\mathbf{y})}(t(\mathbf{x}))$$

odakle slijedi da je

$$L(\hat{\theta}(\mathbf{x})|\mathbf{x}) = L(\hat{\theta}(\mathbf{y})|\mathbf{x}) = \max_{\theta\in\Theta} L(\theta|\mathbf{x}).$$

Zbog pretpostavke da je MLE jedinstven, slijedi da je $\hat{\theta}(\mathbf{x}) = \hat{\theta}(\mathbf{y})$. □

Neka je sada $\tau : \Theta \to \mathbb{R}^p$ funkcija definirana na parametarskom prostoru Θ ($p \geq 1$). *Kako metodom maksimalne vjerodostojnosti procijeniti funkciju parametra $\tau(\theta)$?* Označimo sa $\Gamma := \tau(\Theta)$ sliku te funkcije. Tada je Γ prostor vrijednosti populacijskog parametra $\gamma = \tau(\theta)$ kojeg želimo procijeniti. Prvo, definirajmo *vjerodostojnost od γ*. Neka je $\mathbf{x}$ realizacija slučajnog uzorka $\mathbf{X}$ duljine n ($n \geq 1$) statističkog modela $\mathcal{P}$ parametriziranog sa θ. Označimo sa $\tau^{-1}(\gamma)$ prasliku od γ u prostoru Θ. Ako je $L(\theta) \equiv L(\theta|\mathbf{x})$ vjerodostojnost od θ, tada je **vjerodostojnost od** $\gamma = \tau(\theta)$ jednaka

$$B(\gamma) \equiv B(\gamma|\mathbf{x}) := \sup_{\theta\in\tau^{-1}(\gamma)} L(\theta|\mathbf{x}), \quad \gamma \in \Gamma. \tag{4.21}$$

Procjenitelj metodom maksimalne vjerodostojnosti od $\gamma = \tau(\theta)$ je statistika $\hat{\gamma} \equiv \hat{\gamma}(\mathbf{X})$ takva da je

$$B(\hat{\gamma}|\mathbf{X}) = \max_{\gamma\in\Gamma} B(\gamma|\mathbf{X}). \tag{4.22}$$

Pokazuje se da je procjenjivanje metodom maksimalne vjerodostojnosti (kao unarna operacija) invarijantno na funkcijske transformacije.

Teorem 4.29. *Neka je $\hat{\theta}$ procjenitelj metodom maksimalne vjerodostojnosti od θ na Θ i neka je $\tau(\theta)$ funkcija od θ. Tada je $\tau(\hat{\theta})$ procjenitelj metodom maksimalne vjerodostojnosti od $\tau(\theta)$ na $\tau(\Theta)$.*

Dokaz. Primijetite da množina $\{\tau^{-1}(\gamma) : \gamma \in \Gamma\}$ praslika od $\gamma \in \Gamma$ po τ čini jednu particiju od Θ. Zbog toga vrijedi da je

$$L(\hat{\theta}) = \max_{\theta\in\Theta} L(\theta) = \max_{\gamma\in\Gamma} \sup_{\theta\in\tau^{-1}(\gamma)} L(\theta) = \max_{\gamma\in\Gamma} B(\gamma). \tag{4.23}$$

Stavimo da je $\hat{\gamma} := \tau(\hat{\theta})$. Treba pokazati da tako definirani procjenitelj od γ zadovoljava definicijsku relaciju (4.22). Relacija $\hat{\gamma} = \tau(\hat{\theta})$ je ekvivalentna sa $\hat{\theta} \in \tau^{-1}(\hat{\gamma})$. Slijedi:

$$\sup_{\gamma\in\Gamma} B(\gamma) \overset{\hat{\gamma}\in\Gamma}{\geq} B(\hat{\gamma}) \overset{(4.21)}{=} \sup_{\theta\in\tau^{-1}(\hat{\gamma})} L(\theta) \overset{\hat{\theta}\in\tau^{-1}(\hat{\gamma})}{\geq} L(\hat{\theta}) \overset{(4.23)}{\geq} \max_{\gamma\in\Gamma} B(\gamma).$$

Budući da je supremum nekog skupa jednak maksimumu tog skupa (ako postoji), u prethodno dobivenom nizu nejednakosti imamo, zapravo, svuda jednakosti. Specijalno je

$$B(\hat{\gamma}) = \max_{\gamma\in\Gamma} B(\gamma),$$

a što je trebalo dokazati. □

4.6 Nizovi procjenitelja

Do sada su u razmatranjima slučajni uzorci bili fiksne duljine. U ovom potpoglavlju bavit ćemo se ponašanjem i svojstvima procjenitelja kada duljina uzorka teži u beskonačnost. Model će u početku biti dan u svoj svojoj općenitosti: $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ je statistički model za slučajnu veličinu dimenzije d ($d \geq 1$) zadan gustoćama $f(\cdot;\theta)$ koje su parametrizirane parametrom θ dimenzije m s vrijednostima u parametarskom prostoru $\Theta \subseteq \mathbb{R}^m$ ($m \geq 1$). Euklidsku normu u prostorima $\mathbb{R}^d$ ili $\mathbb{R}^m$ označavat ćemo sa $|\cdot|$.

Neka je sada $\tau(\theta)$ funkcija dimenzije k ($k \geq 1$) koju želimo procijeniti. Ta funkcija može (ali i ne mora) biti procjenjiva. Nadalje, neka je $(T_n;\, n \in \mathbb{N})$ *niz procjenitelja* za funkciju parametra $\tau(\theta)$. To znači da je za $n \in \mathbb{N}$, $T_n = t_n(X_1, X_2, \ldots, X_n)$ statistika s vrijednostima u skupu $\mathbb{R}^k$ gdje su $t_n : \mathbb{R}^{dn} \to \mathbb{R}^k$ Borelova funkcija i $(X_1, X_2, \ldots, X_n) =: \mathbf{X}_n$ slučajan uzorak iz modela $\mathcal{P}$ duljine n. Primijetite da smo u oznaci $\mathbf{X}_n$ za slučajni uzorak u matričnom obliku dodali indeks n kako bismo naglasili da se radi o uzorku duljine n.

Definicija 4.30. *Niz procjenitelja $(T_n;\, n \in \mathbb{N})$ za $\tau(\theta)$ je:*

(a) **jako konzistentan** *ako vrijedi da je*

$$(\forall \theta \in \Theta) \;\; \mathbb{P}_\theta(\lim_n T_n = \tau(\theta)) = 1 \tag{4.24}$$

(b) **slabo konzistentan** *(ili jednostavno* **konzistentan***) ako vrijedi da je*

$$(\forall \theta \in \Theta)\, (\forall \varepsilon > 0) \;\; \lim_n \mathbb{P}_\theta(|T_n - \tau(\theta)| \geq \varepsilon) = 0 \tag{4.25}$$

(c) **konzistentan u srednjem reda** r *(za $r \geq 1$) ako vrijedi da je*

$$(\forall \theta \in \Theta) \;\; \lim_n \mathbb{E}_\theta \left[|T_n - \tau(\theta)|^r\right] = 0. \tag{4.26}$$

Primijetite da je niz procjenitelja $(T_n;\, n \in \mathbb{N})$ jako konzistentan za $\tau(\theta)$ ako vrijedi da za sve $\theta \in \Theta$ niz $(T_n;\, n \in \mathbb{N})$ *gotovo sigurno konvergira k $\tau(\theta)$ u odnosu na vjerojatnost* $\mathbb{P}_\theta$, ili, kraće, $\mathbb{P}_\theta$-g.s. konvergira k $\tau(\theta)$ (vidjeti potpoglavlje 1.7). Tu konvergenciju još zapisujemo na sljedeće načine: za $\theta \in \Theta$,

$$\mathbb{P}_\theta\text{-g.s.} \lim_n T_n = \tau(\theta) \;\;\text{ ili }\;\; T_n \xrightarrow{\mathbb{P}_\theta\text{-g.s.}} \tau(\theta),\; n \to +\infty.$$

Konvergencija niza procjenitelja u definiciji slabe konzistentnosti je *konvergencija po vjerojatnosti* $\mathbb{P}_\theta$ (vidjeti [9], definicija na str. 145.) za $\theta \in \Theta$. Alternativnim zapisima, $(T_n;\, n \in \mathbb{N})$ je (slabo) konzistentan ako za sve $\theta \in \Theta$,

$$(\mathbb{P}_\theta) \lim_n T_n = \tau(\theta) \;\;\text{ ili }\;\; T_n \xrightarrow{\mathbb{P}_\theta} \tau(\theta),\; n \to +\infty.$$

Na kraju, konzistentnost u srednjem reda r za realan broj $r \geq 1$ je definirana pomoću konvergencije u Banachovom prostoru $\mathcal{L}^r(\mathbb{P}_\theta)$ za $\theta \in \Theta$ kako je to opisano u poglavlju 2. Naime, niz procjenitelja je konzistentan reda r ako za svaki $\theta \in \Theta$, niz slučajnih varijabli konvergira k $\tau(\theta)$ u Banachovom (a u slučaju $r = 2$ Hilbertovom) prostoru $\mathcal{L}^r(\mathbb{P}_\theta) \equiv \mathcal{L}^r(\Omega, \mathcal{F}, \mathbb{P}_\theta)$ u odnosu na pseudonormu

$$\|X\|_{\theta r} := \left(\mathbb{E}_\theta[|X|^r]\right)^{\frac{1}{r}}, \;\; X \in \mathcal{L}^r(\mathbb{P}_\theta),$$

tj. ako vrijedi da je za sve $\theta \in \Theta$, $\lim_n \|T_n - \tau(\theta)\|_{\theta r} = 0$ (vidjeti u [9] definiciju na str. 319.). Alternativni zapisi:

$$(m^r\text{-}\mathbb{P}_\theta) \lim_n T_n = \tau(\theta) \;\;\text{ ili }\;\; T_n \xrightarrow{m^r\text{-}\mathbb{P}_\theta} \tau(\theta),\; n \to +\infty \;\;\text{ ili }\;\; T_n \xrightarrow{L^r(\mathbb{P}_\theta)} \tau(\theta),\; n \to +\infty.$$

Budući da konvergencija gotovo sigurno i konvergencija u srednjem povlače konvergenciju po vjerojatnosti (teorem 10.12 u [9], str. 322.), jako konzistentan niz procjenitelja i konzistentan niz procjenitelja u srednjem su ujedno i (slabo) konzistentni nizovi procjenitelja.

Za dokazivanje konzistentnosti niza procjenitelja posebno je važan *jaki zakon velikih brojeva* (kraće: j.z.v.b.). Za zadani niz $(X_n;\, n \in \mathbb{N})$ slučajnih varijabli, označimo sa $(Z_n;\, n \in \mathbb{N})$ pripadni niz parcijalnih suma, tj. niz izvedenih slučajnih varijabli $Z_n := X_1+X_2+\cdots+X_n$ za $n \in \mathbb{N}$. Zbog toga i izvedeni niz aritmetičkih sredina $\overline{X}_n := Z_n/n$ $(n \in \mathbb{N})$ zovemo niz *parcijalnih* aritmetičkih sredina. Sljedeći teorem je dokazan u [9], teorem 12.15, str. 416.

Teorem 4.31 (Kolmogorov)**.** *Neka je $(X_n;\, n \in \mathbb{N})$ niz nezavisnih i jednako distribuiranih slučajnih varijabli. Tada niz parcijalnih aritmetičkih sredina $\overline{X}_n$, $n \in \mathbb{N}$, gotovo sigurno konvergira konačnom limesu ako i samo ako je $\mathbb{E}|X_1| < +\infty$. U tom slučaju je*

$$(\text{g.s.}) \lim_n \overline{X}_n = \mathbb{E}X_1. \tag{4.27}$$

Posebno je od koristi sljedeće poopćenje Kolmogorovljevog teorema (vidjeti [2], Theorem 2.4.5, str. 78.). Podsjetimo se da matematičko očekivanje slučajne varijable X *postoji* ako je $\mathbb{E}X^+ < +\infty$ ili $\mathbb{E}X^- < +\infty$ i u tom slučaju je $\mathbb{E}X = \mathbb{E}X^+ - \mathbb{E}X^- \in \overline{\mathbb{R}} \equiv \mathbb{R} \cup \{-\infty, +\infty\}$ (vidjeti potpoglavlje 1.7).

Korolar 4.32 (Jaki zakon velikih brojeva)**.** *Neka je $(X_n;\, n \in \mathbb{N})$ niz nezavisnih i jednako distribuiranih slučajnih varijabli. Ako postoji matematičko očekivanje $\mathbb{E}X_1 = \mu \in \overline{\mathbb{R}}$, tada za niz parcijalnih aritmetičkih sredina vrijedi da*

$$(\text{g.s.}) \lim_n \overline{X}_n = \mu. \tag{4.28}$$

Primijetite da, ako je $\mathbb{E}X^+ < +\infty$ i $\mathbb{E}X^- < +\infty$, da je tada $\mathbb{E}|X| = \mathbb{E}X^+ + \mathbb{E}X^- < +\infty$ pa Kolmogorovljev teorem 4.31 povlači da je (g.s.)$\lim_n X_n = \mathbb{E}X_1$, dakle, da niz $(X_n; n \in \mathbb{N})$ gotovo sigurno konvergira *konačnom* limesu koji je jednak broju $\mathbb{E}X_1$. Taj slučaj je obuhvaćen i korolarom 4.32. Općenitije, u slučaju kada je, na primjer, $\mathbb{E}X^+ < +\infty$ i $\mathbb{E}X^- = +\infty$, tada, prema korolaru 4.32, niz $(X_n; n \in \mathbb{N})$ također gotovo sigurno konvergira, ali prema *beskonačnom* limesu $\mathbb{E}X_1 = \mathbb{E}X^+ - \mathbb{E}X^- = -\infty$. Oba navedena slučaja trebat ćemo u sljedećem potpoglavlju.

Primjer 4.33. Pretpostavimo da za svaki $n \in \mathbb{N}$ imamo slučajan uzorak $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ duljine n iz statističkog modela zadanog vjerojatnosnim gustoćama $f(\cdot;\theta)$ i parametriziranog s $\theta \in \Theta$.

(a) Pretpostavimo da je za svako $\theta \in \Theta$, $\mathbb{E}_\theta|X_1| < +\infty$. Tada je dobro definirana realna funkcija $\mu(\theta) := \mathbb{E}_\theta(X_1)$, $\theta \in \Theta$. Iz jakog zakona velikih brojeva 4.32 slijedi da je niz parcijalnih aritmetičkih sredina $(\overline{X}_n;\, n \in \mathbb{N})$ jako konzistentan niz procjenitelja. Ako dodatno pretpostavimo da vrijedi:

$$(\forall \theta \in \Theta)\ \ \mathbb{E}_\theta(X_1^2) < +\infty, \tag{4.29}$$

tada je (zbog nepristranosti aritmetičke sredine za $\mu(\theta)$):

$$\mathbb{E}_\theta(|\overline{X}_n - \mu(\theta)|^2) = \mathbb{V}\mathrm{ar}_\theta \overline{X}_n \overset{\text{j.d. i tm.1.61}(ii)\&(v)}{=} \frac{1}{n}\mathbb{V}\mathrm{ar}_\theta X_1,$$

pa je $\lim_n \mathbb{E}_\theta(|\overline{X}_n - \mu(\theta)|^2) = 0$. Dakle, niz $(\overline{X}_n;\, n \in \mathbb{N})$ je konzistentan niz procjenitelja u srednjem reda 2 za $\mu(\theta)$. Odavde, a i iz jake konzistentnosti, slijedi da je taj niz procjenitelja i slabo konzistentan za očekivanje $\mu(\theta)$.

(b) Uz pretpostavku (4.29) na model, dobro je definirana realna funkcija $\sigma^2(\theta) := \mathbb{V}\mathrm{ar}_\theta X_1$, $\theta \in \Theta$. Tada je niz uzoračkih varijanci $(S_n^2;\, n \in \mathbb{N})$ jako konzistentan niz procjenitelja za varijancu $\sigma^2(\theta)$. Pokažimo to. Iz jakog zakona velikih brojeva 4.32 slijedi da je za proizvoljno $\theta \in \Theta$:

$$\text{(g.s.-}\mathbb{P}_\theta)\lim_n \overline{X}_n = \mu(\theta) \;\text{ i }\; \text{(g.s.-}\mathbb{P}_\theta)\lim_n \frac{1}{n}\sum_{i=1}^n X_i^2 = \mathbb{E}_\theta(X_1^2) \overset{\text{tm.1.61}(i)}{=} \sigma^2(\theta) + \mu(\theta)^2.$$

To znači da su događaji

$$\Omega_1 := \{\omega \in \Omega : \lim_n \overline{X}_n(\omega) = \mu(\theta)\} \;\text{ i }\; \Omega_2 := \{\omega \in \Omega : \lim_n \frac{1}{n}\sum_{i=1}^n X_i^2(\omega) = \sigma^2(\theta) + \mu(\theta)^2\}$$

vjerojatnosti 1. Tada je i njihov presjek $\Omega_1 \cap \Omega_2$ vjerojatnosti 1. Neka je $\omega \in \Omega_1 \cap \Omega_2$. Tada vrijedi:

$$\begin{aligned} S_n^2(\omega) &= \frac{1}{n-1}\sum_{i=1}^n (X_i(\omega) - \overline{X}_n(\omega))^2 = \frac{1}{n-1}(\sum_{i=1}^n X_i^2(\omega) - n(\overline{X}_n(\omega))^2) = \\ &= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^n X_i^2(\omega) - (\overline{X}_n(\omega))^2\right) \longrightarrow 1 \cdot (\sigma^2(\theta) + \mu(\theta)^2 - \mu(\theta)^2) = \sigma^2(\theta), \end{aligned}$$

kada $n \to +\infty$. Dakle,

$$\Omega_1 \cap \Omega_2 \subseteq \{\omega \in \Omega : \lim_n S_n^2(\omega) = \sigma^2\}.$$

Budući da je $\mathbb{P}_\theta(\Omega_1 \cap \Omega_2) = 1$, vrijedi da je i $\mathbb{P}_\theta(\lim_n S_n^2 = \sigma^2(\theta)) = 1$. Dakle, za sve $\theta \in \Theta$, (g.s.-$\mathbb{P}_\theta$)$\lim_n S_n^2 = \sigma^2(\theta)$ pa je niz $(S_n^2;\, n \in \mathbb{N})$ jako konzistentan niz procjenitelja za $\sigma^2(\theta)$.

(c) Neka je $x \in \mathbb{R}$ bilo koji zadani realni broj, te neka je $\hat{F}_n(x) = (\sum_{i=1}^n \mathbb{1}_{\{X_i \le x\}})/n$ vrijednost (parcijalne) empirijske funkcije distribucije u točki x, a na osnovi slučajnog uzorka $\mathbf{X}_n$, za $n \in \mathbb{N}$. Tada nam jaki zakon velikih brojeva 4.32 garantira da je $(\hat{F}_n(x);\, n \in \mathbb{N})$ jako konzistentan niz procjenitelja za vrijednost funkcije distribucije $F(x;\theta) = \mathbb{P}_\theta(X_1 \le x)$. Može se pokazati da je ta g.s. konvergencija *uniformna po* $x \in \mathbb{R}$. Naime, vrijedi da je

$$(\forall \theta \in \Theta)\;\; \mathbb{P}_\theta\left(\lim_n \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x;\theta)| = 0\right) = 1.$$

Taj rezultat je poznat kao *Glivenko-Cantellijev teorem* (teorem 12.17 u [9], str. 419.). Primijetite da se tim teoremom pokazuje da niz *funkcija* $\hat{F}_n$ (za $n \in \mathbb{N}$), procjenitelja funkcije distribucije $F(\cdot;\theta)$, gotovo sigurno u odnosu na $\mathbb{P}_\theta$ konvergira k funkciji $F(\cdot;\theta)$ u *sup-normi* na prostoru omeđenih realnih funkcija realne varijable. U tom smislu bismo mogli reći da je niz *funkcijskih procjenitelja* $(\hat{F}_n;\, n \in \mathbb{N})$ jako konzistentan za funkciju $F(\cdot;\theta)$. □

U analizi procjenitelja, važno je odrediti njihov *asimptotski zakon razdiobe*. Dobiveni asimptotski zakon razdiobe za dani procjenitelj ili statistiku općenito, predstavlja aproksimaciju *uzoračke razdiobe* te statistike u ovisnosti o parametru θ, za velike duljine uzorka n. U tom smislu je posebno važna normalna razdioba. Označimo s Φ funkciju distribucije jedinične (ili standardne) normalne slučajne varijable:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}y^2}, \;\; x \in \mathbb{R}.$$

Definicija 4.34. *Niz statistika* $(T_n;\, n \in \mathbb{N})$ *je* **asimptotski normalan** *ako za svaki* $\theta \in \Theta$ *postoje nizovi brojeva* $a_n(\theta) \in \mathbb{R}$, $b_n(\theta) > 0$ $(n \in \mathbb{N})$ *takvi da je*

$$(\forall x \in \mathbb{R}) \quad \lim_n \mathbb{P}_\theta \left(\frac{T_n - a_n(\theta)}{b_n(\theta)} \le x \right) = \Phi(x). \tag{4.30}$$

Pišemo: $T_n \sim AN(a_n(\theta), b_n^2(\theta))$, $n \to +\infty$.

Izrazom (4.30) se zahtjeva da niz slučajnih varijabli $((T_n - a_n(\theta))/b_n(\theta);\, n \in \mathbb{N})$ *konvergira po distribuciji* slučajnoj varijabli Z koja je jedinične normalne razdiobe, tj. $Z \sim N(0,1)$. To je novi, četvrti tip konvergencije slučajnih varijabli. Podsjetimo se (vidjeti definiciju na str. 320. u [9]): niz slučajnih varijabli $(X_n;\, n \in \mathbb{N})$ **konvergira po distribuciji** slučajnoj varijabli X ako za njihove funkcije distribucije F_X, F_{X_n} $(n \in \mathbb{N})$ vrijedi sljedeće:

$$(\forall x \in C(F_X)) \quad \lim_n F_{X_n}(x) = F_X(x). \tag{4.31}$$

Ovdje je sa $C(F_X)$ označen skup svih realnih brojeva x u kojima je funkcija distribucije F_X granične varijable neprekidna. Za konvergenciju niza $(X_n;\, n \in \mathbb{N})$ po distribuciji k X još koristimo oznake:

$$(D) \lim_n X_n = X \quad \text{ili} \quad X_n \xrightarrow{D} X, \; n \to +\infty.$$

Prema tome, definicijski uvjet (4.30) iz definicije asimptotski normalnog niza statistika mogli bismo ekvivalentno zapisati na sljedeće načine:

$$(D\text{-}\mathbb{P}_\theta) \lim_n \frac{T_n - a_n(\theta)}{b_n(\theta)} = Z \sim N(0,1) \quad \text{ili} \quad \frac{T_n - a_n(\theta)}{b_n(\theta)} \xrightarrow{D\text{-}\mathbb{P}_\theta} Z \sim N(0,1), \; n \to +\infty.$$

Primijetite da je $C(\Phi) = \mathbb{R}$.

Zadatak 4.35. Neka je $(X_n;\, n \in \mathbb{N})$ takav niz slučajnih varijabli da $(D) \lim_n X_n = Z \sim N(0,1)$. Dokažite da tada $(D) \lim_n X_n^2 = Z^2 \sim \chi^2(1)$.

Normalni zakon razdiobe se prirodno javlja kao granična distribucija u velikoj većini slučajeva u statistici. Naime, vrijedi sljedeći granični teorem (za dokaz vidjeti, na primjer Lévyjev centralni granični teorem 14.1 u [9], str. 507.).

Teorem 4.36 (Centralni granični teorem). *Neka je* $(X_n;\, n \in \mathbb{N})$ *niz nezavisnih i jednako distribuiranih slučajnih varijabli s konačnom varijancom* $\sigma^2 = \mathbb{V}\mathrm{ar} X_1$ *koja nije jednaka nuli. Ako označimo s* $\mu = \mathbb{E} X_1$ *očekivanje i sa* $\sigma = \sqrt{\mathbb{V}\mathrm{ar} X_1}$ *standardnu devijaciju, tada je*

$$\frac{\overline{X}_n - \mu}{\sigma} \cdot \sqrt{n} \xrightarrow{D} Z \sim N(0,1), \; n \to +\infty.$$

Primjer 4.37. Neka je za svaki $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz statističkog modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ s konačnom i ne-nula varijancom, tj. za svaki $\theta \in \Theta$ je $\sigma^2(\theta) = \mathbb{V}\mathrm{ar}_\theta X_1 \in \langle 0, +\infty \rangle$. Tada je dobro definirana realna funkcija $\mu(\theta) = \mathbb{E}_\theta X_1$, $\theta \in \Theta$. Prema centralnom graničnom teoremu 4.36, za niz parcijalnih aritmetičkih sredina $\overline{X}_n$ $(n \in \mathbb{N})$ vrijedi da je za sve $\theta \in \Theta$,

$$\frac{\overline{X}_n - \mu(\theta)}{\sigma(\theta)} \cdot \sqrt{n} \xrightarrow{D\text{-}\mathbb{P}_\theta} Z \sim N(0,1) \; n \to +\infty.$$

Dakle, niz procjenitelja $(\overline{X}_n;\, n \in \mathbb{N})$ za $\mu(\theta)$ je asimptotski normalan s asimptotskim očekivanjem $\mu(\theta)$ i asimptotskom varijancom $\sigma^2(\theta)/n$, odnosno asimptotskom standardnom devijacijom $\sigma(\theta)/\sqrt{n}$. Zaključujemo da je u modelima nedegeneriranih razdioba u $\mathbb{R}$ s konačnom varijancom aproksimativna uzoračka razdioba od $\overline{X}_n$ normalna s očekivanjem $\mu(\theta)$ i varijancom $\sigma^2(\theta)/n$ za veliki n. □

Primjer 4.38. Neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$ slučajni uzorak duljine n iz eksponencijalnog modela $\text{Exp}(\lambda)$, $\lambda > 0$. Pretpostavit ćemo da je n velik pa ćemo odrediti *aproksimativni* $(1-\alpha)\cdot 100\%$ pouzdani interval (p.i.) za parametar λ. Poznato je da je u tom modelu za sve $\lambda > 0$,

$$\mathbb{E}_\lambda X_1 = \mu(\lambda) = \frac{1}{\lambda}, \ \ \mathbb{V}\text{ar}_\lambda X_1 = \sigma^2(\lambda) = \frac{1}{\lambda^2} \in \langle 0, +\infty\rangle.$$

Prema prethodnom primjeru 4.37 je

$$(\lambda\overline{X}_n - 1)\sqrt{n} = \frac{\overline{X}_n - \frac{1}{\lambda}}{\frac{1}{\lambda}} \cdot \sqrt{n} \xrightarrow{D\text{-}\mathbb{P}_\theta} N(0,1), \ \ n \to +\infty.$$

Dakle, slučajna varijabla $Z = (\lambda\overline{X}_n - 1)\sqrt{n}$ za uzorak velike duljine n ima približno jediničnu normalnu razdiobu, a i strogo je monotona funkcija od λ. Prema tome, Z je *pivotna* slučajna veličina, tj.

(1.) Razdioba od Z (za veliki n) ne ovisi o parametru λ,
(2.) Z je kao funkcija od λ strogo monotona.

Zbog toga ju koristimo za konstrukciju pouzdanog intervala za λ. Neka je $\alpha \in \langle 0, 1\rangle$ zadan. Tada za $(1-\alpha/2)$-kvantil $z_{\frac{\alpha}{2}}$ jedinične normalne distribucije vrijedi da je za svaki $\lambda > 0$:

$$\begin{aligned} & \mathbb{P}_\lambda(-z_{\frac{\alpha}{2}} \le Z \le -z_{\frac{\alpha}{2}}) \approx 1-\alpha \\ \Leftrightarrow \ & \mathbb{P}_\lambda(-z_{\frac{\alpha}{2}} \le (\lambda\overline{X}_n - 1)\sqrt{n} \le -z_{\frac{\alpha}{2}}) \approx 1-\alpha \\ \Leftrightarrow \ & \mathbb{P}_\lambda\left(\frac{1}{\overline{X}_n}\left(1 - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}\right) \le \lambda \le \frac{1}{\overline{X}_n}\left(1 + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}\right)\right) \approx 1-\alpha. \end{aligned}$$

Dakle, približno $(1-\alpha)\cdot 100\%$ pouzdan interval za λ na osnovi velikog uzorka je slučajni interval:

$$\left[\frac{1}{\overline{X}_n}\left(1 - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}\right), \ \frac{1}{\overline{X}_n}\left(1 + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}\right)\right]. \quad \square$$

Asimptotska normalnost je invarijantna na glatke funkcijske transformacije. To je sadržaj sljedećeg, Cramérovog teorema koji je u statističkoj literaturu još poznat i kao Δ-*metoda*.

Teorem 4.39 (Cramér)**.** *Neka je $(X_n; n \in \mathbb{N})$ niz nezavisnih i jednako distribuiranih slučajnih varijabli s konačnom varijancom $\sigma^2 = \mathbb{V}\text{ar}X_1$ koja nije jednaka nuli i očekivanjem $\mu = \mathbb{E}X_1$. Ako je $g : \mathbb{R} \to \mathbb{R}$ Borelova funkcija sa svojstvom da je neprekidno diferencijabilna u otvorenoj okolini od μ i da je $g'(\mu) \neq 0$, tada je*

$$\frac{g(\overline{X}_n) - g(\mu)}{\sigma|g'(\mu)|} \cdot \sqrt{n} \xrightarrow{D} Z \sim N(0,1), \ n \to +\infty.$$

Za dokaz tog teorema treba nam sljedeća lema koju nećemo dokazivati (pogledajte zadatke 3.2.13 i 3.2.14 u [2], str. 125.), a koja spada u grupu tzv. *Slutskyjevih teorema* (vidjeti poglavlje 6 u [3]).

Lema 4.40. *Neka je X slučajna varijabla, y, z realni (neslučajni) brojevi i neka su $(X_n; n \in \mathbb{N})$, $(Y_n; n \in \mathbb{N})$ i $(Z_n; n \in \mathbb{N})$ nizovi slučajnih varijabli definirani na istome vjerojatnosnom prostoru takvi da*

$$X_n \xrightarrow{D} X, \ \ Y_n \xrightarrow{P} y, \ \ Z_n \xrightarrow{P} z, \ \ n \to +\infty.$$

Tada

$$Y_n \cdot X_n + Z_n \xrightarrow{D} yX + z, \ \ n \to +\infty.$$

Dokaz Cramérovog teorema. Neka je U otvoreni interval u $\mathbb{R}$ koji sadrži μ i takav da je $g|_U \in C^1(U)$. Budući da je, po j.z.v.b. 4.32, (g.s.)$\lim_n \overline{X}_n = \mu$, na događaju $\Omega_0 = \{\lim_n \overline{X}_n = \mu\}$ vjerojatnosti jedan postoji $n_0 \in \mathbb{N}$ takav da je za sve $n \geq n_0$, $\overline{X}_n \in U$. Drugim riječima, za $\omega \in \Omega_0$, postoji $n_0 \equiv n_0(\omega)$ takav da je za sve $n \geq n_0$, $\overline{X}_n(\omega) \in U$. U daljnjem izlaganju ću izostaviti pisanje 'malog' omega (ω). Tada, po Lagrangeovom teoremu srednje vrijednosti, postoji broj $\vartheta \equiv \vartheta(\overline{X}_n, \mu) \in \langle 0, 1\rangle$ takav da vrijedi

$$\begin{aligned} & g(\overline{X}_n) = g(\mu) + g'(\mu + \vartheta(\overline{X}_n - \mu)) \cdot (\overline{X}_n - \mu) \\ \Leftrightarrow \quad & \frac{g(\overline{X}_n) - g(\mu)}{\sigma|g'(\mu)|} \cdot \sqrt{n} = \frac{g'(\mu + \vartheta(\overline{X}_n - \mu))}{|g'(\mu)|} \cdot \frac{\overline{X}_n - \mu}{\sigma} \cdot \sqrt{n}. \end{aligned} \tag{4.32}$$

Zbog konveksnosti otvorenog intervala U je za sve $n \geq n_0$, $\mu + \vartheta(\overline{X}_n - \mu) \in U$, te

$$\min\{\mu, \overline{X}_n\} \leq \mu + \vartheta(\overline{X}_n - \mu) \leq \max\{\mu, \overline{X}_n\}.$$

Budući da smo na događaju Ω_0, vrijedi da je $\lim_n \overline{X}_n = \mu$, pa je po teoremu o sandviču i $\lim_n(\mu + \vartheta(\overline{X}_n - \mu)) = \mu$. Zbog neprekidnosti od g' na intervalu U je onda i $\lim_n g'(\mu + \vartheta(\overline{X}_n - \mu)) = g'(\mu)$. Budući da je po pretpostavci $g'(\mu) \neq 0$, vrijedi:

$$\lim_n \frac{g'(\mu + \vartheta(\overline{X}_n - \mu))}{|g'(\mu)|} = \frac{g'(\mu)}{|g'(\mu)|} = \begin{cases} 1, & g'(\mu) > 0 \\ -1, & g'(\mu) < 0. \end{cases}$$

Dakle, $\Omega_0 \subseteq \{\lim_n \frac{g'(\mu+\vartheta(\overline{X}_n-\mu))}{|g'(\mu)|} = \frac{g'(\mu)}{|g'(\mu)|}\}$ i $\mathbb{P}(\Omega_0) = 1$ pa je (g.s.)$\lim_n \frac{g'(\mu+\vartheta(\overline{X}_n-\mu))}{|g'(\mu)|} = \frac{g'(\mu)}{|g'(\mu)|}$, a budući da g.s. konvergencija povlači konvergenciju po vjerojatnosti, je i

$$(P)\lim_n \frac{g'(\mu + \vartheta(\overline{X}_n - \mu))}{|g'(\mu)|} = \frac{g'(\mu)}{|g'(\mu)|}.$$

Primijetite da je za $n \in \mathbb{N}$,

$$g'(\mu + \vartheta(\overline{X}_n - \mu)) \overset{(4.32)}{=} \frac{g(\overline{X}_n) - g(\mu)}{\overline{X}_n - \mu}$$

slučajna varijabla tako da je gore izvedena konvergencija po vjerojatnosti korektna. Nadalje, po centralnom graničnom teoremu 4.36 je

$$(D)\lim_n \frac{\overline{X}_n - \mu}{\sigma} \cdot \sqrt{n} = Z \sim N(0, 1).$$

Sada primjenom leme 4.40 na desnu stranu jednakosti (4.32) i nizove $\tilde{X}_n \equiv \frac{\overline{X}_n - \mu}{\sigma} \cdot \sqrt{n}$, $\tilde{Y}_n \equiv \frac{g'(\mu+\vartheta(\overline{X}_n-\mu))}{|g'(\mu)|}$, $\tilde{Z}_n \equiv 0$ $(n \in \mathbb{N})$, slijedi da je za $n \to +\infty$,

$$\frac{g(\overline{X}_n) - g(\mu)}{\sigma|g'(\mu)|} \cdot \sqrt{n} = \frac{g'(\mu + \vartheta(\overline{X}_n - \mu))}{|g'(\mu)|} \cdot \frac{\overline{X}_n - \mu}{\sigma} \cdot \sqrt{n} = \tilde{Y}_n \cdot \tilde{X}_n + \tilde{Z}_n \xrightarrow{D} \frac{g'(\mu)}{|g'(\mu)|} \cdot Z,$$

ali

$$\frac{g'(\mu)}{|g'(\mu)|} \cdot Z = \left\{ \begin{array}{rl} Z, & g'(\mu) > 0 \\ -Z, & g'(\mu) < 0 \end{array} \right\} \overset{D}{=} Z \sim N(0, 1)$$

što dokazuje teorem. Naime, jedinična normalna razdioba je simetrična oko nule pa je $-Z \overset{D}{=} Z$. $\square$

Napomena 4.41. Cramérov teorem vrijedi i za bilo koji asimptotski normalan niz slučajnih varijabli, a ne samo za parcijalne aritmetičke sredine. Naime, ako je $(T_n;\, n \in \mathbb{N})$ takav niz slučajnih varijabli da za $\mu \in \mathbb{R}$ i $\sigma > 0$ vrijedi da je

$$\frac{T_n - \mu}{\sigma} \cdot \sqrt{n} \xrightarrow{D} Z \sim N(0,1), \quad n \to +\infty,$$

te ako je g Borelova funkcija koja je neprekidno diferencijabilna u okolini od μ i $g'(\mu) \neq 0$, tada je

$$\frac{g(T_n) - g(\mu)}{\sigma |g'(\mu)|} \cdot \sqrt{n} \xrightarrow{D} Z \sim N(0,1), \quad n \to +\infty.$$

Za dokaz vidjeti u [3], Theorem 7, §7, str. 45. □

Primjer 4.42. (*Nastavak primjera 4.38.*) Neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz eksponencijalnog modela $\mathrm{Exp}(\lambda)$, $\lambda > 0$. Nađimo takvu netrivijalnu funkcijsku transformaciju g od $\overline{X}_n$ da asimptotska varijanca od $g(\overline{X}_n)$ ne ovisi o parametru λ. Prvo, reparametrizirajmo model tako da stavimo da je novi parametar $\mu = 1/\lambda > 0$, u interpretaciji, populacijska srednja vrijednost. Tada za populacijsku varijancu vrijedi da je $\mathbb{V}\mathrm{ar}_\lambda X_1 = 1/\lambda^2 = \mu^2 = \mathbb{V}\mathrm{ar}_\mu X_1$ u novom parametru. Prema centralnom graničnom teoremu 4.36 vrijedi da je za sve $\mu > 0$,

$$\frac{\overline{X}_n - \mu}{\mu} \cdot \sqrt{n} \xrightarrow{D\text{-}\mathbb{P}_\mu} Z \sim N(0,1), \quad n \to +\infty.$$

Budući da je u tom modelu $\mathbb{P}_\mu(\overline{X}_n > 0) = 1$, pretpostavimo da je $g : \langle 0, +\infty\rangle \to \mathbb{R}$ klase C^1 takva da je $g(1) = 0$ i $g'(\mu) > 0$ za $\mu > 0$. Primijenimo Cramérov teorem 4.39 na tu funkciju. Tada je

$$\frac{g(\overline{X}_n) - g(\mu)}{\mu g'(\mu)} \cdot \sqrt{n} \xrightarrow{D\text{-}\mathbb{P}_\mu} Z \sim N(0,1), \quad n \to +\infty,$$

pa je asimptotska varijanca od $g(\overline{X}_n)$ jednaka $(\mu g'(\mu))^2/n$. Budući da je zahtjev da ta varijanca ne ovisi o parametru modela, uzmimo da je $\mu g'(\mu) \equiv c$ za svaki $\mu > 0$ i neku konstantu $c > 0$ koja ne ovisi o μ ($\mu > 0$ i pretpostavili smo da je $g'(\mu) > 0$ pa je nužno $c > 0$). Uzmimo da je i $c = 1$ pa riješimo diferencijabilnu jednadžbu $\mu g'(\mu) = 1$ uz uvjet da je $g(1) = 0$. Rješenje te jednadžbe je:

$$g(\mu) = g(\mu) - g(1) = \int_1^\mu g'(\xi)\, d\xi = \int_1^\mu \frac{1}{\xi}\, d\xi = \log \mu - \log 1 = \log \mu, \quad \mu > 0.$$

Dakle, $g \equiv \log$ je jedna tražena transformacija. Za nju vrijedi: $\log \overline{X}_n \sim AN(-\log \lambda, \frac{1}{n})$ za velike n. Uzimajući razne druge vrijednosti za konstantu $c \in \mathbb{R}$, $c \neq 0$, i za početni uvjet $g(1) = b \in \mathbb{R}$, dobit ćemo i druge transformacije s traženim svojstvom. Primijetite da su sve takve transformacije oblika $g(x) = c \log x + b$.
Nalaženje takvih transformacija statistika koje daju asimptotske varijance neovisne o parametrima modela zovemo *stabilizacija varijance*. To je posebno važno u konstrukciji asimptotskih testova, odnosno testnih statistika čija je asimptotska varijanca neovisna o parametrima modela, a time i o hipotezama. □

Primjer 4.43. Neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n iz Bernoullijevog modela $b(1, \theta)$, $\theta \in \langle 0, 1\rangle =: \Theta$. Prema centralnom graničnom teoremu 4.36, za $\theta \in \Theta$,

$$\frac{\overline{X}_n - \theta}{\sqrt{\theta(1-\theta)}} \cdot \sqrt{n} \xrightarrow{D\text{-}\mathbb{P}_\theta} Z \sim N(0,1), \quad n \to +\infty.$$

Nije teško pokazati da je niz procjenitelja $T_n = \sqrt{\overline{X}_n(1-\overline{X}_n)}$ ($n \in \mathbb{N}$) za standardnu devijacije $\sigma(\theta) = \sqrt{\theta(1-\theta)}$ jako konzistentan, pa vrijedi da je i $(\mathbb{P}_\theta)\lim_n T_n = \sigma(\theta)$ što je ekvivalentno sa

$$(\mathbb{P}_\theta)\lim_n \frac{\sqrt{\theta(1-\theta)}}{\sqrt{\overline{X}_n(1-\overline{X}_n)}} = 1.$$

Primjenom leme 4.40 na nizove $\tilde{X}_n = \frac{\overline{X}_n-\theta}{\sqrt{\theta(1-\theta)}} \cdot \sqrt{n}$, $\tilde{Y}_n = \frac{\sqrt{\theta(1-\theta)}}{\sqrt{\overline{X}_n(1-\overline{X}_n)}}$ i $\tilde{Z} \equiv 0$ ($n \in \mathbb{N}$), slijedi da je, kada $n \to +\infty$,

$$\frac{\overline{X}_n-\theta}{\sqrt{\overline{X}_n(1-\overline{X}_n)}} \cdot \sqrt{n} = \frac{\sqrt{\theta(1-\theta)}}{\sqrt{\overline{X}_n(1-\overline{X}_n)}} \cdot \frac{\overline{X}_n-\theta}{\sqrt{\theta(1-\theta)}} \cdot \sqrt{n} = \tilde{Y}_n \cdot \tilde{X}_n + \tilde{Z}_n \overset{D\text{-}\mathbb{P}_\theta}{\longrightarrow} Z \sim N(0,1).$$

Dobivena slučajna veličina $(\overline{X}_n-\theta)\sqrt{n}/\sqrt{\overline{X}_n(1-\overline{X}_n)}$ je pivotna za konstrukciju pouzdanog intervala za θ.
Analizirajmo asimptotsku razdiobu procjenitelja varijance $\hat{\sigma}_n^2 = \overline{X}_n(1-\overline{X}_n)$. Uzmimo $g : \langle 0,1\rangle \to \mathbb{R}$, $g(x) := x(1-x)$, $x \in \langle 0,1\rangle$. Očito je $g \in C^1(\langle 0,1\rangle)$ i $g'(\theta) = 1-2\theta \neq 0$ ako i samo ako je $\theta \neq 1/2$. Prema Cramérovom teoremu 4.39, za sve $\theta \in \langle 0,1\rangle \setminus \{1/2\}$ je asimptotska razdioba od $\hat{\sigma}_n^2$ normalna:

$$\hat{\sigma}_n^2 \sim AN\left(\sigma^2(\theta), \frac{1}{n}(1-2\theta)^2\sigma(\theta)\right) \equiv AN\left(\theta(1-\theta), \frac{1}{n}(1-2\theta)^2\theta(1-\theta)\right), \ \ n \to +\infty.$$

Koja je asimptotska razdioba od $\hat{\sigma}_n^2$ za $\theta = 1/2$? Ideja je da provedemo potpuno isti izvod asimptotske razdiobe kao u dokazu Cramérovog teorema, samo što ćemo sada, zbog $g'(1/2) = 0$, Taylorov teorem srednje vrijednosti primijeniti na višu derivaciju od g. Budući da je $g''(x) \equiv -2 \neq 0$, dovoljno je ići do druge derivacije. Dakle, na događaju $\Omega_0 = \{\lim_n \overline{X}_n = 1/2\}$ vjerojatnosti $\mathbb{P}_{1/2}$ jednake 1, prema Taylorovom teoremu srednje vrijednosti postoji $\vartheta \equiv \vartheta(\overline{X}_n, 1/2) \in \langle 0,1\rangle$ takav broj da je

$$\begin{aligned} & g(\overline{X}_n) = g\left(\frac{1}{2}\right) + \frac{1}{2}g''\left(\frac{1}{2} + \vartheta\left(\overline{X}_n - \frac{1}{2}\right)\right)\left(\overline{X}_n - \frac{1}{2}\right)^2 \\ \Leftrightarrow \quad & \hat{\sigma}_n^2 = \frac{1}{4} + \frac{1}{2}(-2)\left(\overline{X}_n - \frac{1}{2}\right)^2 \\ \Leftrightarrow \quad & \left(\hat{\sigma}_n^2 - \frac{1}{4}\right) \cdot 4n = -\left(\frac{\overline{X}_n - \frac{1}{2}}{\frac{1}{2}} \cdot \sqrt{n}\right)^2. \end{aligned}$$

Sada primijenimo lemu 4.40 na nizove $\tilde{X}_n \equiv ((\overline{X}_n - 1/2)\sqrt{n}/(1/2))^2$, $\tilde{Y}_n \equiv -1$, $\tilde{Z}_n \equiv 0$ ($n \in \mathbb{N}$). Primijetite da po zadatku 4.35 slijedi da je $(D\text{-}\mathbb{P}_{1/2})\lim_n \tilde{X}_n = Z^2 \sim \chi^2(1)$ budući da je po centralnom graničnom teoremu $(D\text{-}\mathbb{P}_{1/2})\lim_n(\overline{X}_n - 1/2)\sqrt{n}/(1/2) = Z \sim N(0,1)$. Dakle, za $n \to +\infty$ je

$$\left(\hat{\sigma}_n^2 - \frac{1}{4}\right) \cdot 4n = -\left(\frac{\overline{X}_n - \frac{1}{2}}{\frac{1}{2}} \cdot \sqrt{n}\right)^2 = \tilde{Y}_n \cdot \tilde{X}_n + \tilde{Z}_n \overset{D\text{-}\mathbb{P}_{\frac{1}{2}}}{\longrightarrow} -Z^2.$$

Prema tome, asimptotska razdioba od $\hat{\sigma}_n$ za veliki n i u odnosu na $\mathbb{P}_{1/2}$ je, do na konstantne faktore, χ^2-razdioba. $\square$

Zadatak 4.44. Dokažite sljedeći teorem:
Neka je $(X_n;\, n \in \mathbb{N})$ niz nezavisnih i jednako distribuiranih slučajnih varijabli s konačnom varijancom $\sigma^2 = \mathbb{V}\mathrm{ar}X_1$ koja nije jednaka nuli i očekivanjem $\mu = \mathbb{E}X_1$. Ako je $g : \mathbb{R} \to \mathbb{R}$ Borelova funkcija sa svojstvom da je klase C^2 na otvorenoj okolini od μ i da je $g'(\mu) = 0$ i $g''(\mu) \neq 0$, tada je

$$\frac{g(\overline{X}_n) - g(\mu)}{\sigma^2 g''(\mu)} \cdot 2n \xrightarrow{D} Z^2 \sim \chi^2(1),\; n \to +\infty.$$

4.7 Konzistentnost i asimptotska normalnost procjenitelja maksimalne vjerodostojnosti

Neka je $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ regularan model parametriziran parametrom θ dimenzije jedan, te neka je, za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzotrak duljine n iz modela $\mathcal{P}$. Ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ opažena vrijednost uzorka $\mathbf{X}_n$ koja je iz zajedničkog nosača gustoće slučajnog uzorka $\mathbf{X}_n$ (model je regularan!), tada je dobro definirana log-vjerodostojnost $\ell_n : \Theta \to \mathbb{R}$:

$$\ell_n(\theta) \equiv \ell_n(\theta|\mathbf{x}) = \sum_{i=1}^{n} \log f(x_i;\theta).$$

Budući da je model regularan, vrijedi da je $\ell_n \in C^1(\Theta)$, odnosno da je log-vjerodostojnost neprekidno diferencijabilna funkcija na otvorenom intervalu Θ. Tada, ako postoji točka $\hat{\theta}_n \equiv \hat{\theta}_n(\mathbf{x}) \in \Theta$ (globalnog) maksimuma log-vjerodostojnosti, tj. ako $\hat{\theta}_n$ zadovoljava uvjet:

$$\ell_n(\hat{\theta}_n|\mathbf{x}) = \max_{\theta\in\Theta} \ell_n(\theta|\mathbf{x}),$$

tada je $\hat{\theta}_n$ nužno stacionarna točka log-vjerodostojnosti ℓ_n. Drugim riječima, $\hat{\theta}_n$ je rješenje *jednadžbe* log-*vjerodostojnosti*:

$$\ell_n'(\theta) \equiv \frac{\partial}{\partial\theta}\ell_n(\theta|\mathbf{x}) = 0.$$

Podsjetimo se da je $\hat{\theta}_n(\mathbf{x})$ *procjena*, a $\hat{\theta}_n(\mathbf{X}_n)$ *procjenitelj* metodom maksimalne vjerodostojnosti za θ.

Teorem 4.45. *Neka je $\theta_0 \in \Theta$ bilo koja vrijednost parametra regularnog modela, te neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak. Tada za svaku drugu vrijednost $\theta \in \Theta$ takvu da je $\theta \neq \theta_0$ vrijedi:*

$$\lim_n \mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)) = 1. \tag{4.33}$$

Dokaz. Primijetite da je

$$\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n) \quad \Leftrightarrow \quad \frac{1}{n}\sum_{i=1}^{n} \log \frac{f(X_i;\theta)}{f(X_i;\theta_0)} < 0. \tag{4.34}$$

Budući da je $\mathrm{supp}f(\cdot;\theta_0) = \mathrm{supp}f(\cdot;\theta)$ slijedi da je

$$\mathbb{E}_{\theta_0}\left(\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right) = \int \frac{f(x;\theta)}{f(x;\theta_0)} \cdot f(x;\theta_0)\,dx = \int\limits_{\mathrm{supp}f(\cdot;\theta_0)} f(x;\theta)\,dx = \int f(x;\theta)\,dx = 1 < +\infty,$$

a kako je

$$\mathbb{P}_{\theta_0}\left(\frac{f(X_1;\theta)}{f(X_1;\theta_0)} \neq 1\right) > 0$$

(jer bi u suprotnom bilo

$$\mathbb{P}_{\theta_0}(f(X_1;\theta_0) = f(X_1;\theta)) = 1 \quad \Rightarrow f(\cdot;\theta_0) = f(\cdot;\theta)\ \mathbb{P}_{\theta_0 X}\text{-g.s.},$$

a što je u kontradikciji s pretpostavljenim uvjetom raspoznavanja (vidjeti potpoglavlje 3.1) budući da je $\theta \neq \theta_0$), prema strogoj Jensenovoj nejednakosti za logaritamsku funkciju (vidjeti zadatak 1.73) slijedi da je

$$\mathbb{E}\left(\log\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right) < \log\mathbb{E}_{\theta_0}\left(\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right) = \log 1 = 0.$$

Budući da smo pokazali da postoji $\mathbb{E}\left(\log\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right)$ i da je strogo manje od nule (u $\overline{\mathbb{R}}$), možemo primijeniti jaki zakon velikih brojeva 4.32 na nezavisne i jednako distribuirane slučajne varijable $\log\frac{f(X_i;\theta)}{f(X_i;\theta_0)}$ $(i \geq 1)$. Zaključujemo da je događaj

$$\Omega_0 := \left\{\lim_n \frac{1}{n}\sum_{i=1}^{n}\log\frac{f(X_i;\theta)}{f(X_i;\theta_0)} = \mathbb{E}\left(\log\frac{f(X_1;\theta)}{f(X_1;\theta_0)}\right) < 0\right\}$$

$\mathbb{P}_{\theta_0}$-vjerojatnosti jedan. Primjenom neprekidnosti vjerojatnosti u odnosu na rastuće događaje i monotonosti vjerojatnosti, slijedi:

$$\begin{aligned}
1 &= \mathbb{P}_{\theta_0}(\Omega_0) \leq \mathbb{P}_{\theta_0}\left(\bigcup_{n_0=1}^{+\infty}\bigcap_{n=n_0}^{+\infty}\left\{\frac{1}{n}\sum_{i=1}^{n}\log\frac{f(X_i;\theta)}{f(X_i;\theta_0)} < 0\right\}\right) \overset{(4.34)}{=} \\
&= \mathbb{P}_{\theta_0}\left(\bigcup_{n_0=1}^{+\infty}\bigcap_{n=n_0}^{+\infty}\{\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)\}\right) \overset{\text{neprek. vjerojat.}}{=} \\
&= \lim_{n_0}\mathbb{P}_{\theta_0}\left(\bigcap_{n=n_0}^{+\infty}\{\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)\}\right) \overset{\text{monot. vjerojat.}}{\leq} \\
&\leq \underline{\lim}_n\mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)) \leq \overline{\lim}_n\mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)) \leq \\
&\leq 1.
\end{aligned}$$

Dakle, svugdje vrijede jednakosti pa je:

$$\begin{aligned}
1 &= \underline{\lim}_n\mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)) = \overline{\lim}_n\mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n)) = \\
&= \lim_n\mathbb{P}_{\theta_0}(\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta|\mathbf{X}_n))
\end{aligned}$$

što je trebalo dokazati. □

Prethodni teorem nam je trebao za dokaz (lokalne) egzistencije konzistentnog niza stacionarnih točaka log-vjerodostojnosti kao procjenitelja za θ.

Teorem 4.46. *Neka je $\theta_0 \in \Theta$ bilo koja vrijednost parametra regularnog modela, te neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak. Tada jednadžba log-vjerodostojnosti*

$$\frac{\partial}{\partial\theta}\ell_n(\theta|\mathbf{X}_n) = 0$$

na događaju čija $\mathbb{P}_{\theta_0}$-vjerojatnost teži ka jedan kada $n \to +\infty$, ima korijen $\hat{\theta}_n \equiv \hat{\theta}_n(\mathbf{X}_n)$ takav da

$$\hat{\theta}_n \xrightarrow{\mathbb{P}_{\theta_0}} \theta_0, \quad n \to +\infty.$$

Dokaz. Budući da je Θ otvoreni interval, postoji $\varepsilon > 0$ takav broj da je $\langle\theta_0-2\varepsilon, \theta_0+2\varepsilon\rangle \subseteq \Theta$. Tada su $\theta_0 - \varepsilon, \theta_0 + \varepsilon \in \Theta$ i nisu jednaki θ_0. Stavimo da je

$$B_n := \{\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta_0 - \varepsilon|\mathbf{X}_n)\}, \;\; C_n := \{\ell_n(\theta_0|\mathbf{X}_n) > \ell_n(\theta_0 + \varepsilon|\mathbf{X}_n)\}$$

i $A_n \equiv A_n(\varepsilon, \theta_0) := B_n \cap C_n$. Budući da, prema teoremu 4.45, $\lim_n \mathbb{P}_{\theta_0}(B_n) = 1$ i $\lim_n \mathbb{P}_{\theta_0}(C_n) = 1$, te da je zbog poluaditivnosti vjerojatnosti

$$A_n^c = B_n^c \cup C_n^c \;\; \Rightarrow \;\; 0 \le \mathbb{P}_{\theta_0}(A_n^c) \le \mathbb{P}_{\theta_0}(B_n^c) + \mathbb{P}_{\theta_0}(C_n^c),$$

po teoremu o sendviču slijedi da je $\lim_n \mathbb{P}_{\theta_0}(A_n^c) = 0$, a što je ekvivalentno tome da je $\lim_n \mathbb{P}_{\theta_0}(A_n) = 1$. Dakle, A_n je događaj čija $\mathbb{P}_{\theta_0}$-vjerojatnost teži k jedan kada $n \to +\infty$. Neka je $\omega \in A_n$ i $\mathbf{x} = \mathbf{X}_n(\omega)$ realizacija uzorka. Promotrimo funkciju $\ell_n(\theta) = \ell_n(\theta|\mathbf{x})$ na intervalu $[\theta_0-\varepsilon, \theta_0+\varepsilon]$. Budući da je (na tom događaju A_n) $\ell_n(\theta_0) > \max\{\ell_n(\theta_0-\varepsilon), \ell_n(\theta_0+\varepsilon)\}$, neprekidna funkcija ℓ_n restringirana na segment $[\theta_0-\varepsilon, \theta_0+\varepsilon]$ poprima svoj maksimum u nutrini tog intervala, dakle, u otvorenom intervalu $I \equiv I(\varepsilon, \theta_0) := \langle\theta_0 - \varepsilon, \theta_0 + \varepsilon\rangle$. Ta točka (općenito lokalnog) maksimuma je nužno stacionarna točka od diferencijabilne funkcije ℓ_n, dakle, korijen jednadžbe vjerodostojnosti. Kako se može dogoditi da postoji više stacionarnih točaka od ℓ_n na intervalu I, a skup stacionarnih točaka

$$\{\theta_s \in \Theta : \ell_n'(\theta_s) = 0\} = (\ell_n')^{-1}(\{0\})$$

je zatvoren skup ($\{0\}$ je očito zatvoren skup i ℓ_n' je neprekidno preslikavanje, a praslika zatvorenog skupa po neprekidnoj funkciji je zatvoren skup), označimo sa $\hat{\theta}_n$ onu stacionarnu točku od ℓ_n na tom intervalu I koja je najbliža θ_0. Ako postoje dvije najbliže stacionarne točke točki θ_0, onda za $\hat{\theta}_n$ odaberimo onu koja je manja. Time smo pokazali da je

$$A_n \subseteq \{|\hat{\theta}_n - \theta_0| < \varepsilon\} \;\; \Rightarrow \;\; 0 \le \mathbb{P}_{\theta_0}(|\hat{\theta}_n - \theta_0| \ge \varepsilon) \le \mathbb{P}_{\theta_0}(A_n^c)$$

pa po teoremu o sendviču slijedi da je $\lim_n \mathbb{P}_{\theta_0}(|\hat{\theta}_n-\theta_0| \ge \varepsilon) = 0$. Nadalje, ako je $0 < \varepsilon' < \varepsilon$, tada je jasno da će se stacionarna točka od ℓ_n odabrana gore opisanim postupkom na manjem intervalu $\langle\theta_0 - \varepsilon', \theta_0 + \varepsilon'\rangle$, podudarati s $\hat{\theta}_n$ odabranim istim postupkom na većem intervalu $\langle\theta_0 - \varepsilon, \theta_0 + \varepsilon\rangle$. Prema tome, $(\mathbb{P}_{\theta_0}) \lim_n \hat{\theta}_n = \theta_0$. □

Napomena 4.47. (a) Iz teorema 4.46 slijedi da, ako jednadžba vjerodostojnosti ima jedinstvenu stacionarnu točku $\hat{\theta}_n$ u Θ *za gotovo sve* $n \in \mathbb{N}$, da je tada nužno $(\hat{\theta}_n; n \in \mathbb{N})$ konzistentan niz procjenitelja za θ. Fraza 'za gotovo sve $n \in \mathbb{N}$' znači da za $\omega \in \Omega$, postoji $n_0(\omega) \in \mathbb{N}$ takav da za sve $n \in \mathbb{N}$ takve da je $n \ge n_0(\omega)$ jednadžba vjerodostojnosti $\frac{\partial}{\partial\theta}\ell_n(\theta|\mathbf{X}_n(\omega)) = 0$ ima jedinstvenu stacionarnu točku.
(b) Slično, ako je procjenitelj maksimalne vjerodostojnosti (MLE) za θ jedinstvena stacionarna točka log-vjerodostojnosti ℓ_n za gotovo sve $n \in \mathbb{N}$, tada je nužno niz MLE konzistentan niz procjenitelja za θ.
(c) Primijetite da teorem 4.46 ne tvrdi da procjenitelji maksimalne vjerodostojnosti postoje, a ako i postoje, ne tvrdi da su konzistentni za θ. Za dokaz egzistencije i jake konzistentnosti niza MLE za θ nužne su dodatne pretpostavke na model i parametarski prostor Θ (vidjeti, na primjer, u [3] Theorem 17 u §17, str. 114.). □

Za konzistentan niz $(\hat{\theta}_n; n \in \mathbb{N})$ procjenitelja za θ za koji vrijedi da za $\omega \in \Omega$ postoji $n_0(\omega) \in \mathbb{N}$ takav da je za svaki prirodan broj $n \ge n_0(\omega)$:

$$\frac{\partial}{\partial\theta}\ell_n(\hat{\theta}_n(\mathbf{X}_n(\omega))|\mathbf{X}_n(\omega)) = 0,$$

kraće kažemo da je **konzistentan niz korijena jednadžbe log-vjerodostojnosti po** θ.

U sljedećem teoremu navest ćemo dovoljne (dodatne) uvjete na regularan model da konzistentan niz korijena jednadžbe log-vjerodostojnosti za θ bude asimptotski normalan niz. Prije iskaza teorema navodimo jedan zadatak čiji rezultat se koristi u dokazu teorema.

Zadatak 4.48. Neka je $a > 0$ neslučajan broj, te neka su $(X_n;\, n \in \mathbb{N})$ i $(Y_n;\, n \in \mathbb{N})$ dva niza slučajnih varijabli takvih da je $(\mathbb{P}) \lim_n X_n = a$ i $(\mathbb{P}) \lim_n Y_n = 0$. Dokažite da je tada

$$(\mathbb{P}) \lim_n \frac{a}{X_n + Y_n} = 1.$$

Podsjetimo se da ćemo radi jednostavnosti sve tvrdnje i izvode vezane uz regularne modele iskazivati samo u terminima neprekidnih modela (integrala) iako će iste tvrdnje vrijediti i za diskretne regularne modele (sume), osim ako drugačije ne napomenemo.

Teorem 4.49. *Neka je za $n \in \mathbb{N}$, $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ slučajni uzorak iz regularnog modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ parametriziranog parametrom θ dimenzije jedan, za koji su ispunjeni dodatni uvjeti:*

(R3)' *Za sve $x \in \mathbb{R}$ preslikavanje $\theta \mapsto f(x;\theta)$ je klase C^3 na Θ.*
(R5)' *Za svaki $\theta \in \Theta$ i $k \in \{1, 2\}$ vrijedi:*

$$0 = \frac{d^k}{d\theta^k} \int_{\mathbb{R}} f(x;\theta)\, dx = \int_{\mathbb{R}} \frac{\partial^k}{\partial \theta^k} f(x;\theta).$$

(R6) *Za svaki $\theta_0 \in \Theta$ postoji nenegativna izmjeriva funkcija M_{θ_0} i otvorena okolina I_{θ_0} od θ_0 takvi da je $\mathbb{E}_{\theta_0}[M_{\theta_0}(X_1)] < +\infty$ i*

$$(\forall \theta \in I_{\theta_0})\, (\forall x \in \mathbb{R}) \quad \left| \frac{\partial^3}{\partial \theta^3} f(x;\theta) \right| \leq M_{\theta_0}(x).$$

Ako je $(\hat{\theta}_n;\, n \in \mathbb{N})$ konzistentan niz korijena jednadžbe log-vjerodostojnosti po θ, tada za svaki $\theta \in \Theta$ vrijedi:

$$(\hat{\theta}_n - \theta)\sqrt{nI(\theta)} \xrightarrow{D\text{-}\mathbb{P}_\theta} Z \sim N(0,1) \;\; n \to +\infty. \tag{4.35}$$

Dokaz. Neka je za $\omega \in \Omega$, $n_0(\omega)$ dovoljno velik prirodan broj takav da je za sve $n \geq n_0(\omega)$ i realizaciju uzorka $\mathbf{x} = \mathbf{X}_n(\omega)$, $\hat{\theta}_n \equiv \hat{\theta}_n(\mathbf{x}) \in \Theta$ stacionarna točka od $\ell_n \equiv \ell_n(\cdot|\mathbf{x})$. Nadalje, neka je $\theta_0 \in \Theta$ bilo koja vrijednost parametra. Budući da je zbog (R3)' ℓ_n' klase C^2 na otvorenom intervalu Θ, prema Taylorovom teoremu srednje vrijednosti postoji točka

$$\theta^* \equiv \theta^*(\theta_0, \hat{\theta}_n, \mathbf{x}) \in \langle \min\{\theta_0, \hat{\theta}_n\},\, \max\{\theta_0, \hat{\theta}_n\} \rangle \subseteq \Theta$$

takva da je

$$\ell_n'(\hat{\theta}_n) = \ell_n'(\theta_0) + \ell_n''(\theta_0)(\hat{\theta}_n - \theta_0) + \frac{1}{2}\ell_n'''(\theta^*)(\hat{\theta}_n - \theta_0)^2. \tag{4.36}$$

Budući da je $\ell_n'(\hat{\theta}_n) = 0$, jednakost (4.36) je ekvivalentna s izrazom

$$(\hat{\theta}_n - \theta_0)\sqrt{nI(\theta_0)} = \frac{I(\theta_0)}{\frac{1}{n}(-\ell_n''(\theta_0)) - \frac{1}{2n}\ell_n'''(\theta^*)(\hat{\theta}_n - \theta_0)} \cdot \frac{1}{\sqrt{nI(\theta_0)}}\ell_n'(\theta_0).$$

Dakle, za dovoljno veliki n i uz poistovjećivanje $\hat{\theta}_n \equiv \hat{\theta}_n(\mathbf{X}_n)$ i $\theta^* \equiv \theta^*(\theta_0, \hat{\theta}_n(\mathbf{X}_n), \mathbf{X}_n)$, vrijedi sljedeća jednakost među slučajnim varijablama:

$$(\hat{\theta}_n - \theta_0)\sqrt{nI(\theta_0)} = \frac{I(\theta_0)}{\frac{1}{n}(-\frac{\partial^2}{\partial\theta^2}\ell_n(\theta_0|\mathbf{X}_n)) - \frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n - \theta_0)} \cdot \frac{1}{\sqrt{nI(\theta_0)}}\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n). \tag{4.37}$$

Primijetite da je θ^* također statistika budući da je implicitno zadana algebarskim izrazom (4.36). Dovoljno je pokazati da vrijedi sljedeće:

(1) $\frac{1}{\sqrt{nI(\theta_0)}}\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n) \stackrel{D\text{-}\mathbb{P}_{\theta_0}}{\longrightarrow} Z \sim N(0,1)$, $n \to +\infty$,

(2) $\frac{1}{n}(-\frac{\partial^2}{\partial\theta^2}\ell_n(\theta_0|\mathbf{X}_n)) \stackrel{\mathbb{P}_{\theta_0}}{\longrightarrow} I(\theta_0)$, $n \to +\infty$,

(3) $\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n - \theta_0) \stackrel{\mathbb{P}_{\theta_0}}{\longrightarrow} 0$, $n \to +\infty$.

Naime, ako vrijedi (2) i (3), onda je i

$$\tilde{Y}_n \equiv \frac{I(\theta_0)}{\frac{1}{n}(-\frac{\partial^2}{\partial\theta^2}\ell_n(\theta_0|\mathbf{X}_n)) - \frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n - \theta_0)} \stackrel{\mathbb{P}_{\theta_0}}{\longrightarrow} \frac{I(\theta_0)}{I(\theta_0) - 0} = 1, \quad n \to +\infty,$$

jer se konvergencija po vjerojatnosti dobro ponaša na algebarske operacije na nizovima (vidjeti zadatak 4.48). Tada koristeći i (1) možemo primijeniti lemu 4.40 na nizove $\tilde{X}_n \equiv \frac{1}{\sqrt{nI(\theta_0)}}\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n)$, $\tilde{Y}_n$ i $\tilde{Z}_n \equiv 0$ ($n \in \mathbb{N}$). Slijedi da je

$$(\hat{\theta}_n - \theta_0)\sqrt{nI(\theta_0)} \stackrel{(4.37)}{=} \tilde{Y}_n \cdot \tilde{X}_n + \tilde{Z}_n \stackrel{D\text{-}\mathbb{P}_{\theta_0}}{\longrightarrow} Z \sim N(0,1), \quad n \to +\infty,$$

što je trebalo dokazati. Preostaje dokazati tvrdnje (1), (2) i (3).

Dokaz tvrdnje (1): Budući da je

$$\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i;\theta_0)$$

parcijalna suma niza $\mathbb{P}_{\theta_0}$-n.j.d. slučajnih varijabli $\frac{\partial}{\partial\theta}\log f(X_i;\theta_0)$ ($i \geq 1$), po (R4) s konačnom ne-nula varijancom

$$+\infty > \mathbb{V}\mathrm{ar}_{\theta_0}\left(\frac{\partial}{\partial\theta}\log(X_i;\theta_0)\right) \stackrel{(4.8)}{=} I(\theta_0) > 0$$

i matematičkim očekivanjem

$$\mathbb{E}_{\theta_0}\left(\frac{\partial}{\partial\theta}\log f(X_i;\theta_0)\right) \stackrel{(4.7)}{=} 0,$$

prema centralnom graničnom teoremu 4.36 vrijedi da je, kada $n \to +\infty$,

$$\frac{1}{\sqrt{nI(\theta_0)}}\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n) = \frac{\frac{1}{n}\frac{\partial}{\partial\theta}\ell_n(\theta_0|\mathbf{X}_n) - 0}{\sqrt{I(\theta_0)}} \cdot \sqrt{n} \stackrel{D\text{-}\mathbb{P}_{\theta_0}}{\longrightarrow} Z \sim N(0,1),$$

a što je tvrdnja (1).

Dokaz tvrdnje (2): Slično kao i u prethodnom slučaju, budući da je

$$\frac{1}{n}\left(-\frac{\partial^2}{\partial\theta^2}\ell_n(\theta_0|\mathbf{X}_n)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(-\frac{\partial^2}{\partial\theta^2}\log f(X_i;\theta_0)\right), \quad n \in \mathbb{N},$$

parcijalna suma niza $\mathbb{P}_{\theta_0}$-n.j.d. slučajnih varijabli $-\frac{\partial^2}{\partial\theta^2}\log f(X_i;\theta_0)$ ($i \geq 1$), prema jakom zakonu velikih brojeva 4.32 te će aritmetičke sredine $\mathbb{P}_{\theta_0}$-g.s. konvergirati prema matematičkom očekivanju tih varijabli ako ono postoji. Računamo:

$$\mathbb{E}_{\theta_0}\left(-\frac{\partial^2}{\partial\theta^2}\log f(X_1;\theta_0)\right) = \mathbb{E}_{\theta_0}\left(\left(\frac{\partial}{\partial\theta}\log f(X_1;\theta_0)\right)^2 - \frac{\frac{\partial^2}{\partial\theta^2}f(X_1;\theta_0)}{f(X_1;\theta_0)}\right). \tag{4.38}$$

Budući da očekivanja

$$\mathbb{E}_{\theta_0}\left(\left(\frac{\partial}{\partial\theta}\log f(X_1;\theta_0)\right)^2\right) \overset{\text{(R4)}}{=} I(\theta_0) \tag{4.39}$$

$$\mathbb{E}_{\theta_0}\left(\frac{\frac{\partial^2}{\partial\theta^2}f(X_1;\theta_0)}{f(X_1;\theta_0)}\right) = \int \frac{\frac{\partial^2}{\partial\theta^2}f(x;\theta_0)}{f(x;\theta_0)}f(x;\theta_0)\,dx = \int \frac{\partial^2}{\partial\theta^2}f(x;\theta_0)\,dx \overset{\text{(R5)'}}{=} 0 \tag{4.40}$$

postoje i konačna su, zbog linearnosti matematičkog očekivanja je:

$$\begin{aligned} &\mathbb{E}_{\theta_0}\left(-\frac{\partial^2}{\partial\theta^2}\log f(X_1;\theta_0)\right) \overset{\text{(4.38) \& lin.m.o.}}{=} \\ = \ & \mathbb{E}_{\theta_0}\left(\left(\frac{\partial}{\partial\theta}\log f(X_1;\theta_0)\right)^2\right) - \mathbb{E}_{\theta_0}\left(\frac{\frac{\partial^2}{\partial\theta^2}f(X_1;\theta_0)}{f(X_1;\theta_0)}\right) \overset{\text{(4.39–4.40)}}{=} \\ = \ & I(\theta_0). \end{aligned}$$

Dakle, gornje očekivanje postoji pa možemo primijeniti j.z.v.b.:

$$\frac{1}{n}\left(-\frac{\partial^2}{\partial\theta^2}\ell_n(\theta_0|\mathbf{X}_n)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(-\frac{\partial^2}{\partial\theta^2}\log f(X_i;\theta_0)\right) \overset{\text{g.s.-}\mathbb{P}_{\theta_0}}{\longrightarrow} I(\theta_0),\ \ n\to+\infty.$$

Budući da g.s. konvergencija povlači konvergenciju po vjerojatnosti, slijedi tvrdnja (2).

Dokaz tvrdnje (3): Prvo ćemo pokazati da je niz varijabli $\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)$ $(n\in\mathbb{N})$ *omeđen po vjerojatnosti* $\mathbb{P}_{\theta_0}$. To znači da vrijedi:

(4) $\lim_{K\to+\infty}\overline{\lim}_n\,\mathbb{P}_{\theta_0}\left(\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right|\geq K\right)=0.$

Na nenegativnu funkciju M_{θ_0} iz uvjeta (R6) i izvedeni niz n.j.d. slučajnih varijabli $M_{\theta_0}(X_i)$ $(i\geq 1)$ gdje su X_i elementi slučajnog uzorka $\mathbf{X}_n=(X_1,X_2,\ldots,X_n)$, možemo primijeniti jaki zakon velikih brojeva 4.32 budući da po pretpostavci (R6) vrijedi da je $\mathbb{E}_{\theta_0}(M_{\theta_0}(X_1))<+\infty$. Označimo sa $\Omega_0=\{\lim_n\frac{1}{n}\sum_{i=1}^n M_{\theta_0}(X_i)=\mathbb{E}_{\theta_0}(M_{\theta_0}(X_1))\}$ događaj $\mathbb{P}_{\theta_0}$-vjerojatnosti jedan. Tada za $\omega\in\Omega_0$ vrijedi da je niz parcijalnih aritmetičkih sredina od $M_{\theta_0}(X_i(\omega))$ $(i\geq 1)$ ograničen. To znači da postoji $K\equiv K(\omega)>0$ takav prirodni broj da je za sve $n\in\mathbb{N}$, $\frac{1}{n}\sum_{i=1}^n M_{\theta_0}(X_i(\omega))<K$. Dakle,

$$\Omega_0\subseteq\bigcup_{K=1}^{+\infty}\bigcap_{n=1}^{+\infty}\left\{\frac{1}{n}\sum_{i=1}^{n}M_{\theta_0}(X_i)<K\right\}.$$

Budući da je $\mathbb{P}_{\theta_0}(\Omega_0)=1$, slijedi da je i

$$\mathbb{P}_{\theta_0}\left(\bigcup_{K=1}^{+\infty}\bigcap_{n=1}^{+\infty}\left\{\frac{1}{n}\sum_{i=1}^{n}M_{\theta_0}(X_i)<K\right\}\right)=1,$$

pa je primjenom neprekidnosti vjerojatnosti u odnosu na padajuće događaje:

$$0=\mathbb{P}_{\theta_0}\left(\bigcap_{K=1}^{+\infty}\bigcup_{n=1}^{+\infty}\left\{\frac{1}{n}\sum_{i=1}^{n}M_{\theta_0}(X_i)\geq K\right\}\right)=\lim_{K\to+\infty}\mathbb{P}_{\theta_0}\left(\bigcup_{n=1}^{+\infty}\left\{\frac{1}{n}\sum_{i=1}^{n}M_{\theta_0}(X_i)\geq K\right\}\right). \tag{4.41}$$

Primijetimo da je za svako $n\in\mathbb{N}$:

$$\left\{\frac{1}{n}\sum_{i=1}^{n}M_{\theta_0}(X_i)\geq K\right\}\subseteq\bigcup_{m=1}^{+\infty}\left\{\frac{1}{m}\sum_{i=1}^{m}M_{\theta_0}(X_i)\geq K\right\},$$

pa je, prvo zbog monotonosti vjerojatnosti, a onda zbog relacije (4.41) primjenom teorema o sendviču:

$$
\begin{aligned}
&(\forall n \in \mathbb{N})\ \mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right) \leq \mathbb{P}_{\theta_0}\left(\bigcup_{m=1}^{+\infty}\left\{\frac{1}{m}\sum_{i=1}^{m} M_{\theta_0}(X_i) \geq K\right\}\right)\\
\Rightarrow\quad & 0 \leq \overline{\lim_{n}}\,\mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right) \leq \mathbb{P}_{\theta_0}\left(\bigcup_{m=1}^{+\infty}\left\{\frac{1}{m}\sum_{i=1}^{m} M_{\theta_0}(X_i) \geq K\right\}\right) \overset{(4.41)}{\Rightarrow}\\
\Rightarrow\quad & \lim_{K\to+\infty}\overline{\lim_{n}}\,\mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right) = 0. \qquad (4.42)
\end{aligned}
$$

Neka je I_{θ_0} otvorena okolina od θ_0 iz pretpostavke (R6), te neka je $\varepsilon > 0$ takav broj da je $\langle\theta_0 - \varepsilon, \theta_0 + \varepsilon\rangle \subseteq I_{\theta_0}$. Tada za $\omega \in \{|\hat{\theta}_n - \theta_0| < \varepsilon\}$ vrijedi da je $\hat{\theta}_n(\mathbf{X}_n(\omega)) \in I_{\theta_0}$. U tom slučaju je i $\theta^* \equiv \theta^*(\theta_0, \hat{\theta}(\mathbf{X}_n(\omega)), \mathbf{X}_n(\omega)) \in I_{\theta_0}$. Ako je

$$
\omega \in \{|\hat{\theta}_n - \theta_0| < \varepsilon\} \cap \left\{\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) < K\right\},
$$

tada je, zbog $\theta^* \in I_{\theta_0}$ i (R6):

$$
\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n(\omega))\right| \leq \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\partial^3}{\partial\theta^3}\log f(X_i(\omega);\theta^*)\right| \overset{(R6)}{\leq} \frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i(\omega)) < K
$$

pa je

$$
\begin{aligned}
&\{|\hat{\theta}_n - \theta_0| < \varepsilon\} \cap \left\{\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) < K\right\} \subseteq \left\{\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right| < K\right\}\\
\Leftrightarrow\quad & \left\{\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right| \geq K\right\} \subseteq \left\{\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right\} \cup \{|\hat{\theta}_n - \theta_0| < \varepsilon\}.
\end{aligned}
$$

Sada opet koristimo monotonost i poluaditivnost vjerojatnosti te konzistentnost od $\hat{\theta}_n$:

$$
\begin{aligned}
&\mathbb{P}_{\theta_0}\left(\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right| \geq K\right) \leq \mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right) + \mathbb{P}_{\theta_0}(|\hat{\theta}_n - \theta_0| \geq \varepsilon)\\
\Rightarrow\quad & 0 \leq \overline{\lim_{n}}\,\mathbb{P}_{\theta_0}\left(\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right| \geq K\right) \leq \underline{\lim_{n}}\,\mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right) + 0 \leq\\
&\leq \overline{\lim_{n}}\,\mathbb{P}_{\theta_0}\left(\frac{1}{n}\sum_{i=1}^{n} M_{\theta_0}(X_i) \geq K\right).
\end{aligned}
$$

Uzmino limes po $K \to +\infty$. Budući da vrijedi (4.42), primjenom teorema o sendviču slijedi tvrdnja (4). Preostaje dovršiti dokaz tvrdnje (3).

Neka su $K > 0$ i $\varepsilon > 0$ proizvoljni. Pretpostavimo da je

$$
\omega \in \left\{\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right| < K\right\} \cap \left\{|\hat{\theta}_n - \theta_0| < \frac{2\varepsilon}{K}\right\}.
$$

Tada vrijedi da je

$$
\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n(\omega))(\hat{\theta}_n(\omega) - \theta_0)\right| = \frac{1}{2n}\left|\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n(\omega))\right| \cdot |\hat{\theta}_n(\omega) - \theta_0| < \frac{K}{2}\cdot\frac{2\varepsilon}{K} = \varepsilon
$$

pa je

$$\left\{\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right|<K\right\}\cap\left\{|\hat{\theta}_n-\theta_0|<\frac{2\varepsilon}{K}\right\}\subseteq\left\{\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|<\varepsilon\right\}$$
$$\Leftrightarrow\left\{\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right\}\subseteq\left\{\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right|\geq K\right\}\cup\left\{|\hat{\theta}_n-\theta_0|\geq\frac{2\varepsilon}{K}\right\}.$$

Sada iz monotonosti i poluaditivnosti vjerojatnosti slijedi da je:

$$\begin{aligned}0 &\leq \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)\leq\\ &\leq \mathbb{P}_{\theta_0}\left(\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right|\geq K\right)+\mathbb{P}_{\theta_0}\left(|\hat{\theta}_n-\theta_0|\geq\frac{2\varepsilon}{K}\right).\end{aligned}$$

Uzimajući limes po n, a kako je $\lim_n \mathbb{P}_{\theta_0}\left(|\hat{\theta}_n-\theta_0|\geq\frac{2\varepsilon}{K}\right)=0$, iz gornje nejednakosti slijedi da je

$$\begin{aligned}0 &\leq \varliminf_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)\leq\\ &\leq \varlimsup_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)\leq\\ &\leq \varlimsup_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)\right|\geq K\right).\end{aligned}$$

Uzimamo limes kada $K\to+\infty$. Budući da vrijedi (4), tada po teoremu o sendviču slijedi da je:

$$\varliminf_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)=\varlimsup_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)=0.$$

Dakle,

$$\lim_n \mathbb{P}_{\theta_0}\left(\left|\frac{1}{2n}\frac{\partial^3}{\partial\theta^3}\ell_n(\theta^*|\mathbf{X}_n)(\hat{\theta}_n-\theta_0)\right|\geq\varepsilon\right)=0.$$

Budući da je $\varepsilon>0$ proizvoljan, dobivena konvergencija vrijedi za svaki $\varepsilon>0$ pa vrijedi (3).

Time je dovršen dokaz teorema. □

Napomena 4.50. (a) Primijetite da teorem 4.49 kaže da za konzistentan niz $(\hat{\theta}_n; n\in\mathbb{N})$ korijena jednadžbe log-vjerodostojnosti po parametru θ u regularnom modelu koji zadovoljava dodatne uvjete regularnosti (R3)', (R5)' i (R6), vrijedi da je *asimptotski efikasan* (ili *asimptotski učinkovit*) niz procjenitelja za θ. Naime, *asimptotsko očekivanje* od $\hat{\theta}_n$ je jednako θ, a *asimptotska varijanca* $1/(nI(\theta))$ je jednaka donjoj Cramér-Raovoj granici za $\tau(\theta)=\theta$ (vidjeti definiciju 4.22).

(b) Ako regularni model iz iskaza teorema 4.49 dopušta jedinstveni niz stacionarnih točaka log-vjerodostojnosti koji je ujedno i niz MLE za θ, tada je niz $(\hat{\theta}_n;\ n\in\mathbb{N})$ procjenitelja maksimalne vjerodostojnosti za θ asimptotski normalan i asimptotski efikasan za θ. Ako je $\tau(\theta)$ takva diferencijabilna funkcija parametra da je $\tau'(\theta)\neq 0$, tada je i niz $(\tau(\hat{\theta}_n); n\in\mathbb{N})$ procjenitelja maksimalne vjerodostojnosti za $\tau(\theta)$ asimptotski normalan i asimptotski efikasan za $\tau(\theta)$. To je posljedica teorema 4.29, teorema 4.49 i Cramérovog teorema 4.39.

□

Dobivene rezultate u ovome potpoglavlju primijenit ćemo na dokaz asimptotskih svojstava procjenitelja maksimalne vjerodostojnosti parametara važnih klasa statističkih modela: jednoparametarskih eksponencijalnih familija. Radi jednostavnosti, promatrat ćemo te modele u kanonskom obliku i parametrizirane prirodnim parametrom s vrijednostima u nutrini prirodnog parametarskog prostora.

Neka je $\mathcal{P} = \{f(\cdot;\eta) : \eta \in \Theta \equiv \operatorname{Int}\Sigma\}$ jednoparametarska eksponencijalna familija, drugim riječima, statistički model pri čemu su vjerojatnosne gustoće oblika (za usporedbu vidjeti potpoglavlje 3.4 i (3.14)):

$$f(x;\eta) = \exp(\eta \cdot t(x) - A(\eta)) \cdot h(x), \quad x \in \mathbb{R}, \tag{4.43}$$

za neku realnu izmjerivu funkciju $t : \mathbb{R} \to \mathbb{R}$ koja nije konstanta i za nenegativnu izmjerivu funkciju $h : \mathbb{R} \to [0, +\infty\rangle$ koja ima neprazan nosač. To da t nije konstanta, znači da niti za jedan realan broj c i niti jedan $\eta \in \Theta$ nije $t(X_1) = c$ $\mathbb{P}_\eta$-g.s. gdje je X_1 slučajna varijabla s gustoćom $f(\cdot;\eta)$. Podsjetimo se da je nosač gustoće

$$\operatorname{supp} f(\cdot;\eta) = \operatorname{supp} h$$

neovisno o η i nužno je neprazan skup. Primijetite da je pozitivna normalizirajuća konstanta $C_0(\eta)$ iz izraza (3.14) zapisana u, za izlaganje u ovome poglavlju, pogodnijem obliku: $C_0(\eta) \equiv \exp(-A(\eta))$. Znamo da je prirodni parametarski prostor

$$\Sigma = \left\{ \eta \in \mathbb{R} : \int\limits_{\mathbb{R}} \exp(\eta \cdot t(x)) \cdot h(x)\, dx < +\infty \right\} \tag{4.44}$$

konveksan skup (vidjeti zadatak 3.24), pa je $\Theta = \operatorname{Int}\Sigma$ otvoren interval. Nadalje,

$$A(\eta) = \log \int\limits_{\mathbb{R}} \exp(\eta t(x)) \cdot h(x)\, dx, \quad \eta \in \Theta. \tag{4.45}$$

Podsjetimo se da i za model $\mathcal{P}$ dopuštamo da je diskretan s gustoćom oblika (4.43), te da će svi u nastavku izvedeni i dokazani rezultati vrijediti i za diskretne jednoparametarske eksponencijalne familije iako će zapis i sprovedeni računi biti uz korištenje integrala (a ne suma) kao da se radi samo o neprekidnim modelima. Dakako, ta se napomena odnosi i na definiciju prirodnog parametarskog prostora (4.44) i na izraz (4.45) za $A(\eta)$ gdje bi se u diskretnom slučaju integrali zamijenili sa sumama.

Propozicija 4.51. *Jednoparametarska eksponencijalna familija parametrizirana prirodnim parametrom $\eta \in \operatorname{Int}\Sigma$ je regularan model za koji vrijede i tvrdnje* (R3)', (R5)' *i* (R6).

Neka su t, h funkcije koje se pojavljuju u (4.43), te Σ prirodni parametarski prostor (4.44). Za dokaz ove propozicije, trebamo sljedeću lemu.

Lema 4.52. *Neka je $\phi : \mathbb{R} \to \mathbb{R}$ takva funkcija da je za svaki kompleksan broj $\xi = \eta + i\nu$ takav da je $\eta \in \Sigma$ i $\nu \in \mathbb{R}$ ispunjeno:*

$$\int\limits_{\mathbb{R}} |\phi(x) \exp(\xi \cdot t(x))| h(x)\, dx < +\infty. \tag{4.46}$$

Tada je funkcija

$$F(\xi) := \int\limits_{\mathbb{R}} \phi(x) \exp(\xi \cdot t(x)) h(x)\, dx \tag{4.47}$$

dobro definirana i analitička na području u $\mathbf{C}$ *koje sadrži skup* $\mathrm{Int}\Sigma + i\mathbb{R} \equiv \mathrm{Int}\Sigma \times \mathbb{R}$*. Nadalje, za svaki* $k \in \mathbb{N}$ *vrijedi da je*

$$\frac{d^k}{d\xi^k}F(\xi) = \int\limits_{\mathbb{R}} \phi(x)t(x)^k \exp(\xi \cdot t(x))h(x)\,dx. \tag{4.48}$$

Dokaz leme. Neka su $\eta_0 \in \mathrm{Int}\Sigma$ i $\nu_0 \in \mathbb{R}$ proizvoljni, te neka je $\xi_0 = \eta_0 + i\nu_0 \in \mathbf{C}$. Tada je, prema pretpostavci (4.46), dobro definiran broj

$$F(\xi_0) := \int\limits_{\mathbb{R}} \phi(x)\exp(\xi_0 \cdot t(x))h(x)\,dx.$$

Budući da je $\eta_0 \in \mathrm{Int}\Sigma$, postoji realan broj $\delta > 0$ takav da je, zbog (4.46), $F(\xi)$ dan izrazom (4.47), dobro definirano za sve $\xi = \eta + i\nu$ takve da je $|\eta - \eta_0| \le \delta$ i $\nu \in \mathbb{R}$. Tada je za $\xi \ne \xi_0$:

$$\frac{F(\xi) - F(\xi_0)}{\xi - \xi_0} \overset{(4.47)\text{ i lin.int.}}{=} \int\limits_{\mathbb{R}} \phi(x)\frac{\exp(\xi \cdot t(x)) - \exp(\xi_0 \cdot t(x))}{\xi - \xi_0}h(x)\,dx. \tag{4.49}$$

Budući da vrijedi ocjena:

$$(\forall z \in \mathbf{C})\quad 0 < |z| \le \delta \quad\Rightarrow\quad \left|\frac{e^{az} - 1}{z}\right| \le \frac{e^{\delta|a|}}{\delta},$$

tada je za $0 < |\xi_0 - \xi| \le \delta$:

$$\begin{aligned} \left|\frac{\exp(\xi \cdot t(x)) - \exp(\xi_0 \cdot t(x))}{\xi - \xi_0}\right| &= \left|e^{(\eta_0 + i\nu_0)t(x)}\frac{e^{(\xi - \xi_0)t(x)} - 1}{\xi - \xi_0}\right| \le \frac{1}{\delta}\left|e^{\eta_0 t(x) + \delta|t(x)|}\right| \le \\ &\le \frac{1}{\delta}(|\exp((\eta_0 - \delta)t(x))| + |\exp((\eta_0 + \delta)t(x))|). \end{aligned}$$

Dakle, za svaki $x \in \mathbb{R}$ je:

$$\begin{aligned} &\left|\phi(x)\frac{\exp(\xi \cdot t(x)) - \exp(\xi_0 \cdot t(x))}{\xi - \xi_0}\right|h(x) \le \\ &\le \frac{1}{\delta}(|\phi(x)\exp((\eta_0 - \delta)t(x))|h(x) + |\phi(x)\exp((\eta_0 + \delta)t(x))|h(x)). \end{aligned} \tag{4.50}$$

Kako su $\eta_0 - \delta, \eta_0 + \delta \in \Sigma$, iz (4.46) slijedi da je funkcija zdesne strane ocjene (4.50) integrabilna. Dakle, podintegralne funkcije u (4.49) su dominirane integrabilnom funkcijom. Budući da je eksponencijalna funkcija analitička u $\mathbf{C}$, slijedi da je za svaki niz $(\xi_n;\ n \in \mathbb{N})$ u $\mathbf{C}$ takav da je $\lim_n \xi_n = \xi_0$:

$$\lim_n \frac{\exp(\xi_n \cdot t(x)) - \exp(\xi_0 \cdot t(x))}{\xi_n - \xi_0} = t(x)\exp(\xi_0 \cdot t(x)),\ \ x \in \mathbb{R}.$$

Prema Lebesgueovu teoremu o dominiranoj konvergenciji (LTDK) slijedi da je

$$\begin{aligned} \lim_n \frac{F(\xi_n) - F(\xi_0)}{\xi_n - \xi_0} &\overset{(4.49)}{=} \lim_n \int\limits_{\mathbb{R}} \phi(x)\frac{\exp(\xi_n \cdot t(x)) - \exp(\xi_0 \cdot t(x))}{\xi_n - \xi_0}h(x)\,dx \overset{(\mathrm{LTDK})}{=} \\ &= \int\limits_{\mathbb{R}} \phi(x)t(x)\exp(\xi_0 \cdot t(x))h(x)\,dx. \end{aligned}$$

Budući da je $(\xi_n;\, n \in \mathbb{N})$ bio proizvoljan niz koji konvergira k ξ_0, slijedi da je kompleksna funkcija F derivabilna u proizvoljnoj točki $\xi_0 \in \mathrm{Int}\Sigma + i\mathbb{R}$, pa je analitička funkcija na području u $\mathbf{C}$ koje sadrži $\mathrm{Int}\Sigma + i\mathbb{R}$. Time je ujedno dokazana i formula (4.48) za $k = 1$. Ista formula za derivacije od F svakog reda $k \in \mathbb{N}$ slijedi matematičkom indukcijom. $\square$

Dokaz propozicije. U skladu s definicijom 4.17 regularnog modela, pokažimo da model zadovoljava uvjete regularnosti (R1-5). U svakoj eksponencijalnoj familiji nosač gustoća ne ovisi o parametru, pa vrijedi uvjet regularnosti (R1). Budući da je $\Theta = \mathrm{Int}\Sigma$ nutrina konveksnog skupa u $\mathbb{R}$, otvoreni je interval pa je ispunjen i uvjet (R2). Ako u lemu 4.52 stavimo da je $\phi \equiv 1$, tada je za $\eta \in \Sigma$ ispunjen uvjet (4.46) pa je, po toj lemi, funkcija

$$F(\eta + i\nu) = \int\limits_{\mathbb{R}} \exp((\eta + i\nu)t(x))h(x)\,dx$$

analitička kompleksna funkcija na području koje sadrži $\Theta + i\mathbb{R}$ i vrijedi (4.48). Tada restrikcija te funkcije na $\Theta \subseteq \mathbb{R}$ (u istoj oznaci $F \equiv F|_{\Theta}$) ima derivacije svakog reda i za njih vrijedi da je

$$\frac{d^k}{d\eta^k}F(\eta) = \int\limits_{\mathbb{R}} t(x)^k \exp(\eta t(x))h(x)\,dx, \quad \eta \in \Theta, \quad k \in \mathbb{N}. \tag{4.51}$$

Zbog toga je i funkcija

$$A(\eta) \overset{(4.45)}{=} \log F(\eta) \quad \Leftrightarrow \quad e^{-A(\eta)} = \frac{1}{F(\eta)}$$

klase $C^\infty(\Theta)$ kao kompozicija dviju neprekidno derivabilnih funkcija svakoga reda. Zbog toga je za svaki $\eta \in \Theta$, funkcija

$$\eta \mapsto f(x;\eta) \overset{(4.43)}{=} \exp(\eta t(x) - A(\eta))h(x)$$

klase (barem) $C^3(\Theta)$ pa vrijedi (R3)' (što povlači i (R3)). Neka je sada X slučajna varijabla kojoj je razdioba dana gustoćom $f(x;\eta)$. Za $\eta \in \Theta$ izračunajmo:

$$\begin{aligned} A'(\eta) &= \frac{F'(\eta)}{F(\eta)} \overset{(4.51)}{=} \int\limits_{\mathbb{R}} t(x)\exp(\eta t(x) - A(\eta))h(x)\,dx = \int\limits_{\mathbb{R}} t(x)f(x;\eta)\,dx = \\ &= \mathbb{E}_\eta(t(X)) \end{aligned} \tag{4.52}$$

$$\begin{aligned} A''(\eta) &= \frac{F''(\eta)}{F(\eta)} - \left(\frac{F'(\eta)}{F(\eta)}\right)^2 \overset{(4.51)}{=} \int\limits_{\mathbb{R}} t(x)^2 f(x;\eta)\,dx - \left(\int\limits_{\mathbb{R}} t(x)f(x;\eta)\,dx\right)^2 = \\ &= \mathbb{E}_\eta[t(X)^2] - (\mathbb{E}_\eta t(X))^2 = \mathbb{V}\mathrm{ar}_\eta t(X). \end{aligned} \tag{4.53}$$

Nadalje, za $x \in \operatorname{supp} h$ i $\eta \in \Theta$ računamo:

$$\frac{\partial}{\partial\eta}\log f(x;\eta) = \frac{\partial}{\partial\eta}(\eta t(x) - A(\eta) + \log h(x)) = t(x) - A'(\eta) \tag{4.54}$$

pa je Fišerova informacija:

$$I(\eta) = \mathbb{E}_\eta\left[\left(\frac{\partial}{\partial\eta}\log f(X;\eta)\right)^2\right] = \mathbb{E}_\eta[(t(X) - A'(\eta))^2] \overset{(4.52-4.53)}{=} \mathbb{V}\mathrm{ar}_\eta t(X) \in \langle 0, +\infty\rangle.$$

Naime, varijanca $\mathbb{V}\mathrm{ar}_\eta t(X)$ je jednaka $A''(\eta)$ po (4.53), pa je konačna, a budući da $t(X)$ nije degenerirana slučajna varijabla (jer t nije konstantna funkcija), $\mathbb{V}\mathrm{ar}_\eta t(X) > 0$. Dakle, vrijedi (R4). Nadalje, za $x \in \operatorname{supp} h$ i $\eta \in \Theta$ izračunajmo:

$$
\begin{aligned}
\frac{\partial}{\partial \eta} f(x;\eta) &= \frac{\partial}{\partial \eta}\left(\exp(\eta t(x) - A(\eta))h(x)\right) = \\
&= t(x)\exp(\eta t(x) - A(\eta))h(x) - A'(\eta)\exp(\eta t(x) - A(\eta))h(x) = \\
&= (t(x) - A'(\eta))f(x;\eta) \qquad (4.55)\\
\frac{\partial^2}{\partial \eta^2} f(x;\eta) &= \frac{\partial}{\partial \eta}\big((t(x) - A'(\eta))f(x;\eta)\big) = -A''(\eta)f(x;\eta) + (t(x) - A'(\eta))\frac{\partial}{\partial \eta}f(x;\eta) = \\
&= -A''(\eta)f(x;\eta) + (t(x) - A'(\eta))^2 f(x;\eta). \qquad (4.56)
\end{aligned}
$$

Odavde slijedi:

$$
\begin{aligned}
\int_{\mathbb{R}} \frac{\partial}{\partial \eta} f(x;\eta)\,dx &\overset{(4.55)}{=} \int_{\mathbb{R}} (t(x) - A'(\eta))f(x;\eta)\,dx = \mathbb{E}_\eta[t(X) - A'(\eta)] = \\
&= \mathbb{E}_\eta t(X) - A'(\eta) \overset{(4.52)}{=} \mathbb{E}_\eta t(X) - \mathbb{E}_\eta t(X) = 0 \\
\int_{\mathbb{R}} \frac{\partial^2}{\partial \eta^2} f(x;\eta)\,dx &\overset{(4.56)}{=} -A''(\eta)\int_{\mathbb{R}} f(x;\eta)\,dx + \int_{\mathbb{R}} (t(x) - A'(\eta))^2 f(x;\eta)\,dx = \\
&= -A''(\eta)\cdot 1 + \mathbb{E}_\eta[t(X) - A'(\eta)] \overset{(4.52-4.52)}{=} \\
&= -\mathbb{V}\mathrm{ar}_\eta t(X) + \mathbb{V}\mathrm{ar}_\eta t(X) = 0.
\end{aligned}
$$

Dakle, vrijedi (R5)' što povlači i da vrijedi (R5). Preostaje još pokazati da vrijedi (R6). Neka je $\eta_0 \in \Theta$ proizvoljan, te neka je $\varepsilon > 0$ takav broj da je $\langle \eta_0 - 2\varepsilon, \eta_0 + 2\varepsilon\rangle \subseteq \Theta$. Stavimo da je

$$I_{\eta_0} := \langle \eta_0 - \varepsilon, \eta_0 + \varepsilon\rangle.$$

Budući da je $\eta \mapsto |A'''(\eta)|$ neprekidna funkcija, na segmentu $\mathrm{Cl}I_{\eta_0} = [\eta_0 - \varepsilon, \eta_0 + \varepsilon]$ poprima maksimalnu vrijednost:

$$M_{\eta_0} := \max_{\eta \in \mathrm{Cl}I_{\eta_0}} |A'''(\eta)|.$$

Budući da je

$$\frac{\partial^3}{\partial \eta^3}\log f(x;\eta) \overset{(4.54)}{=} \frac{\partial^2}{\partial \eta^2}(t(x) - A'(\eta)) = -A'''(\eta),$$

za sve $x \in \mathbb{R}$ i sve $\eta \in I_{\eta_0}$ slijedi da je

$$\left|\frac{\partial^3}{\partial \eta^3}\log f(x;\eta)\right| = |A'''(\eta)| \le M_{\eta_0}.$$

M_{η_0} je konstantna nenegativna funkcija pa ima konačno matematičko očekivanje. Time je pokazano i svojstvo (R6). □

Teorem 4.53. *U jednoparametarskoj eksponencijalnoj familiji parametriziranoj prirodnim parametrom $\eta \in \mathrm{Int}\Sigma$ procjenitelji maksimalne vjerodostojnosti za η postoje, konzistentni su, asimptotski su normalno distribuirani i asimptotski su efikasni.*

Dokaz. Prema propoziciji 4.51 jednoparametarska familija je regularan model koji zadovoljava i dodatne uvjete (R3)', (R5)' i (R6). Zbog toga je, prema napomeni 4.50 (b), dovoljno dokazati da za gotovo sve duljine n slučajnih uzoraka $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ postoji procjenitelj maksimalne vjerodostojnosti $\hat{\eta}_n$ za η koji je ujedno jedinstvena stacionarna točka log-vjerodostojnosti:

$$\ell_n(\eta) = \sum_{i=1}^{n} \log f(X_i; \eta) = \eta \sum_{i=i}^{n} t(X_i) - nA(\eta) + \sum_{i=1}^{n} \log h(X_i).$$

U tom slučaju će niz MLE biti konzistentan niz procjenitelja prema teoremu 4.46 (vidjeti napomenu 4.47), a prema teoremu 4.49 će biti asimptotski normalan s asimptotskim očekivanje jednakim μ i asimptotskom varijancom jednakom $1/(nI(\eta)) = 1/(nA''(\eta))$, dakle, i asimptotski je efikasan.

Nađimo sve stacionarne točke od ℓ_n:

$$\ell_n'(\eta) = T_n - nA'(\eta) = 0 \quad \Leftrightarrow \quad A'(\eta) = \frac{T_n}{n}$$

gdje je $T_n = \sum_{i=1}^{n} t(X_i)$ prirodna dovoljna statistika. Tada su stacionarne točke rješenja jednadžbe: $A'(\eta) = T_n/n$. Neka je sada $\eta_0 \in \Theta = \mathrm{Int}\Sigma$ bilo koja vrijednost parametra. Prirodna dovoljna statitika T_n normirana sa n je parcijalna aritmetička sredina n.j.d., u odnosu na vjerojatnost $\mathbb{P}_{\eta_0}$, slučajnih varijabli $t(X_i)$ $(i \geq 1)$ koje imaju konačno matematičko očekivanje $\mathbb{E}_{\eta_0} t(X_1) = A'(\eta_0)$ prema (4.52). Zbog toga na taj niz možemo primijeniti jaki zakon velikih brojeva 4.31. Dakle, vrijedi:

$$\frac{T_n}{n} \overset{\text{g.s.-}\mathbb{P}_{\eta_0}}{\longrightarrow} A'(\eta_0), \;\; n \to +\infty.$$

S druge strane, neprekidna funkcija $\eta \mapsto A'(\eta)$ na otvorenom intervalu Θ strogo raste na tom intervalu jer je njena derivacija za svaki $\eta \in \Theta$

$$A''(\eta) \overset{(4.53)}{=} \mathbb{V}\mathrm{ar}_\eta t(X_1) > 0$$

budući da $t(X_1)$ nije g.s. konstanta. Slika strogo monotone neprekidne funkcije definirane na otvorenom intervalu je otvoreni interval $I := A'(\Theta) \subseteq \mathbb{R}$ i I sadrži limes $A'(\eta_0)$ od T_n/n na događaju $\mathbb{P}_{\eta_0}$-vjerojatnosti jedan. Tada na tom istome događaju postoji $n_0 \in \mathbb{N}$ takav da je za sve $n \geq n_0$, $T_n/n \in I$. Za svaki takav n jednadžba $T_n/n = A'(\eta)$ ima jedinstveno rješenje $\hat{\eta}_n \in \Theta$ i to je, dakle, jedinstvena stacionarna točka log-vjerodostojnosti. Budući da je

$$(\forall \eta \in \Theta) \;\; \ell_n''(\eta) = -nA''(\eta) < 0,$$

funkcija log-vjerodostojnosti je konkavna, pa je njena jedinstvena stacionarna točka $\hat{\eta}_n$ točka globalnog maksimuma od ℓ_n, dakle, MLE. Time je teorem dokazan. □

Zadatak 4.54. Neka je $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ za $n \in \mathbb{N}$ slučajan uzorak iz eksponencijalnog modela $\mathrm{Exp}(\lambda)$, $\lambda \in \langle 0, +\infty \rangle$. Nađite MLE i UMVUE za $\tau(\lambda) = e^{-\lambda}$. Pokažite da je MLE asimptotski efikasan, a UMVUE konzistentan procjenitelj za $\tau(\lambda)$. Ispitajte asimptotsku efikasnost od UMVUE za $\tau(\lambda)$.

Poglavlje 5

Testiranje statističkih hipoteza

Opisno, statistički test je postupak u kojemu donosimo odluku da li odbacujemo jednu statističku hipotezu u korist druge, njoj komplementarne statističke hipoteze, ili ju ne odbacujemo na osnovi opažene vrijednosti slučajnog uzorka i naprijed zadane razine značajnosti.

Statistička hipoteza je bilo koja tvrdnja o populacijskoj razdiobi statističkog obilježja kojeg izučavamo. U kontekstu parametarskih statističkih modela, statistička hipoteza je tvrdnja o populacijskim vrijednostima parametara. Razlikujemo jednostavne i složene statističke hipoteze. *Jednostavna* statistička hipoteza je statistička hipoteza kojom je (populacijski) zakon razdiobe statističkog obilježja od interesa, jedinstveno određen. Ako se radi o hipotezi kojom se određuje pripadnost populacijske razdiobe tog obilježja nekoj višečlanoj podklasi svih mogućih razdioba, kažemo da je statistička hipoteza *složena*.

Klasična matematička definicija testa je da je to funkcija koja realizacijama slučajnog uzorka pridružuje dvije moguće vrijednosti: jedna (1) koja se interpretira kao odluka o odbacivanju *osnovne* statističke hipoteze u korist *alternativne* i druga (0) čija je interpretacija da se ne odbacuje osnovna hipoteza. Dakako, odluka o neodbacivanju neke hipoteze ne znači da je ta hipoteza istinita, već samo da primijenjeni statistički test nije uspio detektirati odstupanje od nje.

U ovome poglavlju proširit ćemo klasičnu definiciju statističkog testa na tzv. *randomizirajući* test. Takvo određenje testa omogućit će nam jednostavniju konstrukciju i analizu njegove optimalnosti. Veličina koja kontrolira vjerojatnost odbacivanja osnovne hipoteze ako je ona istinita, zove se *značajnost* ili *signifikantnost* testa.

5.1 Randomizirajući test

Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n ($n \in \mathbb{N}$) iz parametarskog modela $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ zadanog vjerojatnosnim gustoćama. Označimo sa $f_{\mathbf{X}}(\cdot;\theta)$ ($\theta \in \Theta$) funkciju gustoće od $\mathbf{X}$. Podsjetimo se, ako je $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{dn}$ opažena vrijednost od $\mathbf{X}$, tada je vjerodostojnost od θ jednaka:

$$L(\theta|\mathbf{x}) := f_{\mathbf{X}}(\mathbf{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta), \quad \theta \in \Theta.$$

Neka je sada $\{\Theta_0, \Theta_1\}$ jedna particija parametarskog prostora Θ. To znači da su $\Theta_0, \Theta_1 \subset \Theta$ takvi podskupovi da su neprazni, *disjuntni* (tj. $\Theta_0 \cap \Theta_1 = \emptyset$) i da je $\Theta_0 \cup \Theta_1 = \Theta$. Pretpostavimo da želimo testirati statističke hipoteze:

$$\begin{aligned} H_0 &: \quad \theta \in \Theta_0 \\ H_1 &: \quad \theta \in \Theta_1. \end{aligned} \tag{5.1}$$

Hipotezu H_0 zovemo *osnovnom* ili *nul-hipotezom*, a H_1 *alternativnom hipotezom*.

Definicija 5.1. **Randomizirajući test** *za testiranje statističkih hipoteza* H_0 *u odnosu na* H_1 *je izmjerivo preslikavanje* $\tau : \mathbb{R}^{dn} \to [0,1]$.

Od sada pa nadalje, randomizirajući test ćemo kraće zvati *test*. Interpretacija testa τ je sljedeća. Za opaženu vrijednost $\mathbf{x} \in \mathbb{R}^{dn}$ slučajnog uzorka $\mathbf{X}$, imamo sljedeće. Ako je $\tau(\mathbf{x}) = 1$, tada je odluka da *odbacujemo osnovnu hipotezu u korist alternativne*. Ako je $\tau(\mathbf{x}) = 0$, tada *ne odbacujemo osnovnu hipotezu*. Za razliku od klasičnog testa, $\tau(\mathbf{x})$ može biti i broj iz intervala $\langle 0,1 \rangle$. Ako je to slučaj, dakle, ako je $\tau(\mathbf{x}) \in \langle 0,1 \rangle$, tada sprovedemo jedan dodatni bernoullijevski pokus neovisan od pokusa odabira uzorka, s vjerojatnosti uspjeha $\tau(\mathbf{x})$. Na primjer, bacamo novčić kojemu je vjerojatnost da padne pismo jednaka $\tau(\mathbf{x})$. Ako padne pismo, tada odbacujemo osnovnu hipotezu u korist alternativne. U suprotnom, ne odbacujemo osnovnu hipotezu. Dakle, vrijednost $\tau(\mathbf{x})$ testa τ interpretira se kao *vjerojatnost da ćemo odbaciti osnovnu hipotezu u korist alternativne hipoteze*.

Posebnu ulogu ima podskup $C \subseteq \mathbb{R}^{dn}$ koji je skup svih realizacija danog slučajnog uzorka za koji ćemo *sigurno* donijeti odluku o odbacivanju nul-hipoteze u korist alternativne. Dakle, $C \equiv C_\tau := \tau^{-1}(1)$. Skup C zovemo *kritično područje testa* τ.

Važna karakteristika test je njegova *jakost*. Opisno, jakost nekog testa je mjera sposobnosti toga testa da odbaci nul-hipotezu ako je alternativna hipoteza točna. Smatra se mjerom sposobnosti testa da detektira odstupanje empirijske razdiobe statističkog obilježja od interesa (na osnovi opaženog uzorka) od modelskih razdioba sadržanih u nul-hipotezi.

Definicija 5.2. **Jakost testa** τ *je funkcija* $\gamma_\tau : \Theta \to [0,1]$ *definirana sa*

$$\gamma_\tau(\theta) := \mathbb{E}_\theta[\tau(\mathbf{X})], \quad \theta \in \Theta.$$

Na primjer, ako je τ klasičan test, tj. slika od τ je dvočlani skup $\{0,1\}$, tada je $\gamma_\tau(\theta) = \mathbb{P}_\theta(\mathbf{X} \in C_\tau)$ za sve $\theta \in \Theta$. Tada je interpretacija broja $\gamma_\tau(\theta)$ da je to *vjerojatnost odbacivanja nul-hipoteze u korist alternativne ako je* θ *'prava' vrijednost parametra*. Na isti način ćemo interpretirati broj $\gamma_\tau(\theta)$ u slučaju randomizirajućeg testa τ.

Testiranje statističkih hipoteza, kao svaki statistički postupak, ne daje 100% pouzdane zaključke. Zbog toga kod donošenja odluke da li odbacujemo nul-hipotezu u korist alternativne hipoteze ili ne, možemo pogriješiti u smislu da:

(1.) odbacimo nul-hipotezu, a istinita je - u tom slučaju smo napravili *pogrešku prve vrste*, ili

(2.) ne odbacimo nul-hipotezu, a alternativa je istinita - to je slučaj *pogreške druge vrste*.

Budući da je nemoguće kontrolirati vjerojatnosti obiju pogrešaka, standardno se kontrolira samo vjerojatnost pogreške prve vrste. Za test hipoteza (5.1), vjerojatnosti pogreške prve vrste opisane su sa restrikcijom $\gamma_\tau|_{\Theta_0} : \Theta_0 \to [0,1]$ funkcije jakosti testa. Tada možemo definirati **značajnost testa** τ kao broj:

$$\alpha_\tau := \sup_{\theta \in \Theta_0} \gamma_\tau(\theta).$$

Kod konstrukcije testa obično si zadajemo broj $\alpha \in \langle 0,1 \rangle$ takav da je zahtjev na test τ kojeg želimo konstruitati: $\alpha_\tau \leq \alpha$. Broj α koji je gornja ograda vjerojatnosti pogrešaka prve vrste nekog testa zovemo **razina značajnosti** (ili **razina signifikantnosti**) tog testa. Obično se za α uzimaju brojevi 0.01, 0.05 ili 0.10. U tom slučaju, ako za test τ vrijedi da je $\alpha_\tau \leq \alpha$, obično se kaže da je test s razinom značajnosti $\alpha \cdot 100\%$.

5.2 Uniformno najjači testovi

Spomenuli smo da kod statističkih testova nije moguće istovremeno kontrolirati obje vrste pogrešaka: prve i druge vrste. Dogovorno se kontrolira pogreška prve vrste, a nastoji se konstruirati (ili odabrati) onaj test među svim testovima iste razine značajnosti koji ima *najmanju* vjerojatnost pogreške druge vrste. Takve testove ćemo zvati *uniformno najjačima* na zadanoj razini značajnosti.

Definicija 5.3. *Za statistiški test τ^* nul-hipoteze $H_0 : \theta \in \Theta_0$ naprema alternativi $H_1 : \theta \in \Theta_1$ i zadani broj $\alpha \in \langle 0, 1\rangle$, kažemo da je* **uniformno najjači na razini značajnosti** *α ako je $\alpha_{\tau^*} \leq \alpha$ i za svaki drugi test τ istih hipoteza vrijedi:*

$$\alpha_\tau \leq \alpha \quad \Rightarrow \quad (\forall \theta \in \Theta_1)\ \gamma_\tau(\theta) \leq \gamma_{\tau^*}(\theta).$$

Kraće ćemo reći da je τ^* iz definicije 5.3, UMP-test (engl. *Uniformly the Most Powerfull*) na razini značajnosti α.

Pokažimo da je UMP-test optimalan test u smislu da među svim testovima iste razine značajnosti ima najmanju vjerojatnost pogreške druge vrste. Za test τ hipoteza (5.1) vjerojatnosti pogrešaka druge vrste opisane su *funkcijom vjerojatnosti pogrešaka druge vrste testa* τ. To je funkcija $\beta_\tau : \Theta_1 \to [0, 1]$ takva da je $\beta_\tau = 1 - \gamma_\tau|_{\Theta_1}$. Ako je τ^* UMP-test na razini značajnosti α, prema definiciji 5.3 to znači da je za svaki drugi test τ takav da je $\alpha_\tau \leq \alpha$:

$$(\forall \theta \in \Theta_1)\ \gamma_\tau(\theta) \leq \gamma_{\tau^*}(\theta) \quad \Leftrightarrow \quad \beta_{\tau^*}(\theta) = 1 - \gamma_{\tau^*}(\theta) \leq 1 - \gamma_\tau(\theta) = \beta_\tau(\theta).$$

Dakle, UMP-test ima *uniformno* po θ najmanju vjerojatnost pogreške druge vrste među svim testovima koji su iste razine značajnosti.

Važan rezultat za konstrukciju UMP-testa je teorem koji je poznat pod tradicionalnim imenom Neyman-Pearsonova lema (vidjeti, na primjer, u [5]). Iako se odnosi samo na testiranje jednostavnih hipoteza u okviru najjednostavnijih netrivijalnih statističkih modela, dvočlanih, taj teorem je baza mnogih važnih rezultata. Neke od njih ćemo navesti i dokazati u ovome poglavlju.

Prije iskaza leme, navedimo neke pretpostavke i oznake koje ćemo u nastavku izlaganja koristiti. Pretpostavit ćemo da su nosači svih gustoća u modelu $\mathcal{P}$ jednaki, neovisni o parametru θ. Nadalje, model može biti neprekidan ili diskretan. Označimo s istom oznakom μ Lebesguevu mjeru λ ako je model neprekidan, odnosno odgovarajuću brojaću mjeru μ_D (za zajednički prebrojiv nosač D gustoća) ako je model diskretan. Drugim riječima, μ je σ-konačna mjera u odnosu na koju su zakoni razdiobe opisani sa $\mathcal{P}$ apsolutno neprekidni.

Teorem 5.4 (Neyman-Pearsonova lema)**.** *Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n ($n \in \mathbb{N}$) iz dvočlanog statističkog modela $\mathcal{P} = \{f(\cdot;\theta_0), f(\cdot;\theta_1)\}$ pri čemu za vrijednosti parametara vrijedi da su $\theta_0 \neq \theta_1$, a za vjerojatnosne gustoće da imaju zajednički nosač. Tada:*

(*i*) (Egzistencija) *Za zadani broj $\alpha \in \langle 0, 1\rangle$, postoje nenegativna realna konstanta k i test τ za testiranje hipoteza $H_0 : \theta = \theta_0$ nasuprot $H_1 : \theta = \theta_1$, takvi da vrijedi:*

$$\mathbb{E}_{\theta_0}[\tau(\mathbf{X})] = \alpha \tag{5.2}$$

$$\tau(\mathbf{x}) = \begin{cases} 1 & \text{za } L(\theta_1|\mathbf{x}) > kL(\theta_0|\mathbf{x}) \\ 0 & \text{za } L(\theta_1|\mathbf{x}) < kL(\theta_0|\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{dn}. \end{cases} \tag{5.3}$$

(*ii*) (Dovoljan uvjet za UMP) *Ako neki test τ zadovoljava uvjete* (5.2) *i* (5.3) *za dane $\alpha \in \langle 0, 1\rangle$ i konstantu k, tada je τ uniformno najjači test na razini značajnosti α za testiranje hipoteza H_0 vs. H_1 iz* (*i*).

(*iii*) (Nužan uvjet za UMP) *Ako je τ^* uniformno najjači test na razini značajnosti α za testiranje hipoteza H_0 vs. H_1 iz (i), tada postoji nenegativna konstanta k takva da τ^* zadovoljava* (5.3) *μ-s.s. Tada će test τ^* zadovoljavati i* (5.2) *ako ne postoji test $\tilde{\tau}$ takav da je $\alpha_{\tilde{\tau}} = \gamma_{\tilde{\tau}}(\theta_0) < \alpha$ i $\gamma_{\tilde{\tau}}(\theta_1) = 1$.*

Dokaz.

(*i*): Konstruirat ćemo test koji zadovoljava uvjete (5.2-5.3). Takav test ne mora biti jedinstven. Definirajmo funkciju $a : [0, +\infty\rangle \to [0, 1]$ sa

$$a(c) := \mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) > cL(\theta_0|\mathbf{X})), \quad c \geq 0.$$

Označimo sa S zajednički nosač gustoća slučajnog uzorka $f_{\mathbf{X}}(\cdot;\theta_0)$ i $f_{\mathbf{X}}(\cdot;\theta_1)$. Budući da je $\mathbb{P}_{\theta_0}(\mathbf{X} \in S) = \mathbb{P}_{\theta_1}(\mathbf{X} \in S) = 1$, slijedi da je $\mathbb{P}_{\theta_0}(L(\theta_0|\mathbf{X}) > 0) = \mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) > 0) = 1$ pa je, za $c \geq 0$:

$$a(c) = \mathbb{P}_{\theta_0}\left(\frac{L(\theta_1|\mathbf{X})}{L(\theta_0|\mathbf{X})} > c\right) = 1 - \mathbb{P}_{\theta_0}\left(\frac{L(\theta_1|\mathbf{X})}{L(\theta_0|\mathbf{X})} \leq c\right) = 1 - F(c;\theta_0),$$

gdje je $F(\cdot;\theta_0)$ funkcija distribucije slučajne varijable $L(\theta_1|\mathbf{X})/L(\theta_0|\mathbf{X})$ u odnosu na vjerojatnost $\mathbb{P}_{\theta_0}$. Zbog toga je funkcija a nerastuća, neprekidna je zdesna i vrijedi da su limesi $a(0+) \equiv \lim_{c\to 0+} a(c) = 1$ i $a(+\infty) \equiv \lim_{c\to+\infty} a(c) = 0$ (teorem 1.7). Nadalje, vrijedi da je

$$\mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) = cL(\theta_0|\mathbf{X})) = a(c-) - a(c).$$

Iz ove jednakosti slijedi: ako je a neprekidna u $c > 0$, tada je $\mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) = cL(\theta_0|\mathbf{X})) = 0$. Prema tome, za dani $\alpha \in \langle 0, 1\rangle$ sigurno postoji broj $c_0 > 0$ takav da je

$$a(c_0) \leq \alpha \leq a(c_0-).$$

Ako je $a(c_0-) - a(c_0) > 0$, za $\mathbf{x} \in \mathbb{R}^{dn}$ definiramo:

$$\tau(\mathbf{x}) := \begin{cases} 1, & L(\theta_1|\mathbf{x}) > c_0 L(\theta_0|\mathbf{x}) \\ \frac{\alpha - a(c_0)}{a(c_0-) - a(c_0)}, & L(\theta_1|\mathbf{x}) = c_0 L(\theta_0|\mathbf{x}) \\ 0, & L(\theta_1|\mathbf{x}) < c_0 L(\theta_0|\mathbf{x}). \end{cases}$$

Tada je slika preslikavanja τ sadržana u $[0, 1]$ i vrijedi:

$$\begin{aligned}
\mathbb{E}_{\theta_0}[\tau(\mathbf{X})] &= \mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) > c_0 L(\theta_0|\mathbf{X})) + \frac{\alpha - a(c_0)}{a(c_0-) - a(c_0)} \mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) = c_0 L(\theta_0|\mathbf{X})) = \\
&= a(c_0) + \frac{\alpha - a(c_0)}{a(c_0-) - a(c_0)} \cdot (a(c_0-) - a(c_0)) = \\
&= \alpha.
\end{aligned}$$

Dakle, τ je test koji, za $k := c_0$, zadovoljava (5.2-5.3).
Ako je $a(c_0-) - a(c_0) = 0$, tada je a neprekidno u točki c_0 i vrijedi da je $\alpha = a(c_0) = a(c_0-)$. Za $\mathbf{x} \in \mathbb{R}^{dn}$ definiramo:

$$\tau(\mathbf{x}) := \begin{cases} 1, & L(\theta_1|\mathbf{x}) > c_0 L(\theta_0|\mathbf{x}) \\ 0, & L(\theta_1|\mathbf{x}) \leq c_0 L(\theta_0|\mathbf{x}). \end{cases}$$

Tada je τ test i vrijedi:

$$\mathbb{E}_{\theta_0}[\tau(\mathbf{X})] = \mathbb{P}_{\theta_0}(L(\theta_1|\mathbf{X}) > c_0 L(\theta_0|\mathbf{X})) = a(c_0) = \alpha.$$

Dakle, i u ovome slučaju je τ test koji, za $k := c_0$, zadovoljava (5.2-5.3).

(ii): Neka je τ test koji zadovoljava (5.2-5.3) za dane $\alpha \in \langle 0,1\rangle$ i $k \geq 0$. Primijetite da je, prema (5.2), $\alpha = \alpha_\tau = \gamma_\tau(\theta_0)$. Neka je τ' bilo koji test hipoteza H_0 vs. H_1 koji zadovoljava uvjet da je

$$\gamma_{\tau'}(\theta_0) = \alpha_{\tau'} \leq \alpha. \tag{5.4}$$

Dovoljno je pokazati da vrijedi

$$\gamma_{\tau'}(\theta_1) \leq \gamma_\tau(\theta_1).$$

Naime, tada će, zbog proizvoljnosti od τ', slijediti da je τ UMP-test na razini značajnosti α po definiciji 5.3. Označimo:

$$\begin{array}{rcl} S^+ & := & \{\mathbf{x} \in \mathbb{R}^{dn} : \tau(\mathbf{x}) - \tau'(\mathbf{x}) > 0\} \equiv \{\tau - \tau' > 0\} \\ S^- & := & \{\mathbf{x} \in \mathbb{R}^{dn} : \tau(\mathbf{x}) - \tau'(\mathbf{x}) < 0\} \equiv \{\tau - \tau' < 0\}. \end{array}$$

Primijetite da je

$$S^- \cup S^+ = \{\mathbf{x} \in \mathbb{R}^{dn} : \tau(\mathbf{x}) - \tau'(\mathbf{x}) \neq 0\} \equiv \{\tau - \tau' \neq 0\}.$$

Nadalje, vrijedi:

$$\mathbf{x} \in S^+ \;\Rightarrow\; \tau(\mathbf{x}) > \tau'(\mathbf{x}) \geq 0 \;\Rightarrow\; \tau(\mathbf{x}) > 0 \;\overset{(5.3)}{\Rightarrow}\; L(\theta_1|\mathbf{x}) \geq kL(\theta_0|\mathbf{x}).$$

Slično:

$$\mathbf{x} \in S^- \;\Rightarrow\; \tau(\mathbf{x}) < \tau'(\mathbf{x}) \leq 1 \;\Rightarrow\; \tau(\mathbf{x}) < 1 \;\overset{(5.3)}{\Rightarrow}\; L(\theta_1|\mathbf{x}) \leq kL(\theta_0|\mathbf{x}).$$

Dakle,

$$S^- \cup S^+ \subseteq \{\mathbf{x} \in \mathbb{R}^{dn} : (\tau(\mathbf{x}) - \tau'(\mathbf{x}))(L(\theta_1|\mathbf{x}) - kL(\theta_0|\mathbf{x})) \geq 0\}. \tag{5.5}$$

Prema tome, koristeći linearnost integrala, slijedi:

$$\begin{aligned} & (\gamma_\tau(\theta_1) - \gamma_{\tau'}(\theta_1)) - k(\gamma_\tau(\theta_0) - \gamma_{\tau'}(\theta_0)) \overset{\text{def.5.2}}{=} \\ = & \int_{\mathbb{R}^{dn}} (\tau(\mathbf{x}) - \tau'(\mathbf{x}))(f_{\mathbf{X}}(\mathbf{x}, \theta_1) - kf_{\mathbf{X}}(\mathbf{x}, \theta_0))\, d\mu(x) = \\ = & \int_{\mathbb{R}^{dn}} (\tau(\mathbf{x}) - \tau'(\mathbf{x}))(L(\theta_1|\mathbf{x}) - kL(\theta_0|\mathbf{x}))\, d\mu(x) = \\ = & \int_{S^- \cup S^+} \underbrace{(\tau(\mathbf{x}) - \tau'(\mathbf{x}))(L(\theta_1|\mathbf{x}) - kL(\theta_0|\mathbf{x}))}_{\overset{(5.5)}{\geq} 0}\, d\mu(x) \geq 0. \end{aligned}$$

Dakle,

$$\gamma_\tau(\theta_1) - \gamma_{\tau'}(\theta_1) \geq k(\gamma_\tau(\theta_0) - \gamma_{\tau'}(\theta_0)) \overset{(5.2-5.4)}{\geq} k(\alpha - \alpha) = 0 \;\Rightarrow\; \gamma_{\tau'}(\theta_1) \leq \gamma_\tau(\theta_1).$$

(iii): Neka je τ^* UMP-test na razini značajnosti α za testiranje H_0 vs. H_1, te neka je τ test iz (i) za koji vrijede (5.2-5.3). Slično kao u dokazu tvrdnje (ii), neka su

$$S^- = \{\tau < \tau^*\},\; S^+ = \{\tau > \tau^*\} \;\Rightarrow\; S^- \cup S^+ = \{\tau^* \neq \tau\},$$

te stavimo da je

$$E := (S^- \cup S^+) \cap \{L(\theta_1|\cdot) \neq kL(\theta_0|\cdot)\} \;\Rightarrow\; E \subseteq \{(\tau - \tau^*)(L(\theta_1|\cdot) - kL(\theta_0|\cdot)) > 0\}.$$

Cilj nam je pokazati da je $\tau^* = \tau$ μ-s.s. na skupu $\{L(\theta_1|\cdot) \neq kL(\theta_0|\cdot)\}$. To znači da treba pokazati da je:

$$\mu(\{\tau^* \neq \tau\} \cap \{L(\theta_1|\cdot) \neq kL(\theta_0|\cdot)\}) = \mu(E) = 0.$$

Pretpostavimo suprotno, dakle, da je $\mu(E) > 0$. Tada je, zbog pozitivnosti podintegralne funkcije na E, i sljedeći integral pozitivan:

$$\begin{aligned} 0 &< \int_E (\tau(\mathbf{x}) - \tau^*(\mathbf{x}))(L(\theta_1|\mathbf{x}) - kL(\theta_0|\mathbf{x}))\, d\mu(\mathbf{x}) = \\ &= \int_{\mathbb{R}^{dn}} (\tau(\mathbf{x}) - \tau^*(\mathbf{x}))(L(\theta_1|\mathbf{x}) - kL(\theta_0|\mathbf{x}))\, d\mu(\mathbf{x}) = \\ &= (\gamma_\tau(\theta_1) - \gamma_{\tau^*}(\theta_1)) - k(\gamma_\tau(\theta_0) - \gamma_{\tau^*}(\theta_0)). \end{aligned}$$

Budući da je $\gamma_\tau(\theta_0) = \alpha$ i $\gamma_{\tau^*}(\theta_0) \leq \alpha$, na isti način kao u dokazu tvrdnje (ii) slijedi da je

$$\gamma_{\tau^*}(\theta_1) < \gamma_\tau(\theta_1),$$

a što je u kontradikciji s pretpostavkom da je τ^* UMP-test na razini značajnosti α. Dakle, $\mu(E) = 0$ pa τ^* zadovoljava uvjet (5.3) μ-s.s. Ako taj test ne zadovoljava (5.2), tada mu se može modificirati (povećati) vrijednost u točkama $\mathbf{x} \in \mathbb{R}^{dn}$ takvima da je $L(\theta_1|\mathbf{x}) = kL(\theta_0|\mathbf{x})$ tako da modificirani test ima vrijednost značajnosti $\gamma_{\tau^*}(\theta_0) = \alpha$. U tom slučaju će se povećati i jakost u točki θ_1 i time će modificirani test ostati UMP-test na razini značajnosti α i zadovoljavat će uvjet (5.3) μ-s.s. Kako i dalje taj test zadovoljava uvjet da je $\alpha_{\tau^*} = \alpha \leq \alpha$, a test τ iz (i) je prema (ii) UMP-test na razini značajnosti α, vrijedi da je uvećana vrijednost $\gamma_{\tau^*}(\theta_1) \leq \gamma_\tau(\theta_1)$. S druge strane, ako postoji neki test $\tilde{\tau}$ takav da je $\gamma_{\tilde{\tau}}(\theta_0) < \alpha$ i $\gamma_{\tilde{\tau}}(\theta_1) = 1$, tada je taj test UMP-test na razini značajnosti α, pa svaki UMP-test na istoj razini značajnosti ima jakost u točki θ_1 jednaku 1. To znači da, ako τ^* ima značajnost strogo manju od α, tada se njegova vrijednost ne može više uvećavati za $\mathbf{x} \in \mathbb{R}^{dn}$ takve da je $L(\theta_1|\mathbf{x}) = kL(\theta_0|\mathbf{x})$. Time je dokaz teorema gotov. □

Korolar 5.5. *Neka je τ uniformno najjači test na razini značajnosti $\alpha \in \langle 0, 1\rangle$ za testiranje hipoteza $H_0 : \theta = \theta_0$ nasuprot $H_1 : \theta = \theta_1$ za $\theta_0 \neq \theta_1$. Tada je*

$$\gamma_\tau(\theta_0) < \gamma_\tau(\theta_1).$$

Dokaz. Neka je $\tau' \equiv \alpha$ neki drugi test, za $\alpha = \gamma_\tau(\theta_0)$. Tada on zadovoljava uvjet $\alpha_{\tau'} = \alpha \leq \alpha$ pa je nužno

$$\alpha = \gamma_{\tau'}(\theta_1) \leq \gamma_\tau(\theta_1)$$

jer je τ UMP-test na razini značajnosti α po pretpostavci. Kada bi bilo da je $\gamma_{\tau'}(\theta_1) = \gamma_\tau(\theta_1) = \alpha < 1$, tada bi i test τ' bio UMP-test na razini značajnosti α. Tada bi, prema teoremu 5.4 (iii), morao zadovoljavati (5.3) μ-s.s., a to je moguće jedino ako je

$$\mu(\{\mathbf{x} \in \mathbb{R}^{dn} : L(\theta_1|\mathbf{x}) \neq kL(\theta_1)\}) = 0.$$

Odavde slijedi da je

$$f_{\mathbf{X}}(\mathbf{x}; \theta_1) = L(\theta_1|\mathbf{x}) = kL(\theta_0|\mathbf{x}) = kf_{\mathbf{X}}(\mathbf{x}; \theta_0) \text{ za } \mu\text{-s.s. } \mathbf{x} \in \mathbb{R}^{dn}.$$

Budući da su $f_{\mathbf{X}}(\cdot; \theta_1)$ i $f_{\mathbf{X}}(\cdot; \theta_0)$ vjerojatnosne funkcije gustoća, nužno je $k = 1$, a onda je $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_0}$ što je nemoguće zbog $\theta_0 \neq \theta_1$ i svojstva raspoznavanja. Dakle, $\gamma_\tau(\theta_0) < \gamma_\tau(\theta_1)$ što je trebalo dokazati. □

U nastavku ćemo primijeniti Neyman-Pearsonovu lemu 5.4 na netrivijalne jednoparametarske modele za testiranje složenijih hipoteza. Jednu takvu široku klasu modela gdje je moguće primijeniti Neyman-Pearsonovu lemu, čine modeli koji imaju svojsvo *monotonog omjera vjerodostojnosti*.

Od sada pa nadalje, neka je $\mathcal{P} = \{f(\cdot;\theta) : \theta \in \Theta\}$ jednoparametarski statistički model, dakle, $\Theta \subseteq \mathbb{R}$. Nadalje, neka je $\mathbf{x} \in \mathbb{R}^{dn}$ realizacija slučajnog uzorka $\mathbf{X}$ duljine n ($n \in \mathbb{N}$) iz tog modela.

Definicija 5.6. *Statistički model ima* **monotoni omjer vjerodostojnosti** *ako postoji funkcija $\xi(\mathbf{x})$ po $\mathbf{x} \in \mathbb{R}^{dn}$ takva da je, za sve $\theta', \theta'' \in \Theta$ takve da je $\theta' < \theta''$, omjer vjerodostojnosti $L(\theta''|\mathbf{x})/L(\theta'|\mathbf{x})$ neopadajuća funkcija po $\xi(\mathbf{x})$.*

Fraza da je 'omjer vjerodostojnosti $L(\theta''|\mathbf{x})/L(\theta'|\mathbf{x})$ neopadajuća funkcija po $\xi(\mathbf{x})$' znači da za $\theta' < \theta''$ postoji neopadajuća realna funkcija realne varijable $y \mapsto g_{\theta',\theta''}(y)$ takva da je za sve $\mathbf{x} \in \mathbb{R}^{dn}$:

$$\frac{L(\theta''|\mathbf{x})}{L(\theta'|\mathbf{x})} = g_{\theta',\theta''}(\xi(\mathbf{x})). \tag{5.6}$$

Teorem 5.7. *Neka je $\mathcal{P}$ jednoparametarski statistički model s monotonim omjerom vjerodostojnosti u funkciji $\xi(\mathbf{x})$ od $\mathbf{x} \in \mathbb{R}^{nd}$ gdje je $\mathbf{x}$ realizacija slučajnog uzorka $\mathbf{X}$ duljine n. Tada:*

(*i*) *Za bilo koje $\theta_0 \in \Theta$ i $\alpha \in \langle 0, 1\rangle$ postoji uniformno najjači test na razini značajnosti α za testiranje hipoteza*

$$H_0: \ \ \theta \leq \theta_0 \ \ \textit{nasuprot} \ \ H_1: \ \ \theta > \theta_0$$

definiran sa

$$\tau(\mathbf{x}) = \begin{cases} 1, & \xi(\mathbf{x}) > c \\ \beta, & \xi(\mathbf{x}) = c \\ 0, & \xi(\mathbf{x}) < c, \quad \mathbf{x} \in \mathbb{R}^{dn}, \end{cases} \tag{5.7}$$

pri čemu su konstante $\beta \in [0, 1\rangle$ i $c \in \mathbb{R}$ određene relacijom:

$$\mathbb{E}_{\theta_0}[\tau(\mathbf{X})] = \alpha. \tag{5.8}$$

(*ii*) *Jakost γ_τ testa iz (i) je strogo rastuća funkcija na skupu*

$$\{\theta \in \Theta : 0 < \gamma_\tau(\theta) < 1\}.$$

(*iii*) *Za svaki $\theta' \in \Theta$, test τ određen sa (5.7-5.8) je uniformno najjači test za testiranje hipoteza $H_0' : \theta \leq \theta'$ vs. $H_1' : \theta > \theta'$ na razini značajnosti $\alpha' = \gamma_\tau(\theta')$.*

(*iv*) *Za svaki $\theta < \theta_0$ test τ iz (i) minimizira vjerojatnost pogreške prve vrste $\gamma_{\tau'}(\theta)$ među svim testovima τ' koji zadovoljavaju uvjet $\mathbb{E}_{\theta_0}[\tau'(\mathbf{X})] = \alpha$.*

Dokaz.

(*i*) & (*ii*): Neka je θ_1 proizvoljna vrijednost parametra $\theta \in \Theta$ takva da zadovoljava hipotezu H_1, tj. takva da je $\theta_1 > \theta_0$. Promotrimo jednostavne hipoteze:

$$H_0^0: \ \ \theta = \theta_0 \ \ \text{nasuprot} \ \ H_1^0: \ \ \theta = \theta_1.$$

Prema Neyman-Pearsonovoj lemi 5.4, UMP-test za testiranje H_0^0 vs. H_1^0 je μ-s.s. jedinstveno određen s (5.2-5.3) za zadanu značajnost α i neku konstantu k. Pokazat ćemo da postoji k takav da za test određen relacijama (5.7-5.8) vrijede i relacije (5.2-5.3) iz Neyman-Pearsonove leme. U tom slučaju, Neyman-Pearsonovu lemu primijenjujemo na podmodel $\mathcal{P}_0 = \{f(\cdot;\theta_0), f(\cdot;\theta_1)\}$ od $\mathcal{P}$. Budući da, po pretpostavci, model $\mathcal{P}$ ima svojstvo monotonog omjera vjerodostojnosti u funkciji $\xi(\mathbf{x})$ za $\mathbf{x} \in \mathbb{R}^{dn}$, neka je, za $\theta' = \theta_0 < \theta'' = \theta_1$, g_{θ_0,θ_1} neopadajuća funkcija takva da vrijedi relacija (5.6). Stavimo $k := g_{\theta_0,\theta_1}(c)$ za c iz definicijske relacije (5.7) za test τ. Budući da je g_{θ_0,θ_1} neopadajuća funkcija, tada, koristeći i obrat po kontrapoziciji, za $\mathbf{x} \in \mathbb{R}^{dn}$ vrijedi:

$$\left(\xi(\mathbf{x}) \le c \;\Rightarrow\; \frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} = g_{\theta_0,\theta_1}(\xi(\mathbf{x})) \le g_{\theta_0,\theta_1}(c) = k\right) \;\Leftrightarrow\; \left(\frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} > k \;\Rightarrow\; \xi(\mathbf{x}) > c\right)$$
$$\left(\xi(\mathbf{x}) \ge c \;\Rightarrow\; \frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} = g_{\theta_0,\theta_1}(\xi(\mathbf{x})) \ge g_{\theta_0,\theta_1}(c) = k\right) \;\Leftrightarrow\; \left(\frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})} < k \;\Rightarrow\; \xi(\mathbf{x}) < c\right).$$

Dakle, ako je $\mathbf{x} \in \mathbb{R}^{dn}$ takav da je $L(\theta_1|\mathbf{x})/L(\theta_0|\mathbf{x}) > k$, tada je nužno i $\xi(\mathbf{x}) > c$ pa je $\tau(\mathbf{x}) = 1$. S druge strane, ako je $L(\theta_1|\mathbf{x})/L(\theta_0|\mathbf{x}) < k$, tada je $\xi(\mathbf{x}) < c$ pa je $\tau(\mathbf{x}) = 0$. Prema tome, test τ zadovoljava uvjet (5.3), a kako je (5.8) isto što i (5.2), slijedi da test τ zadovoljava oba uvjeta (5.2-5.3) iz Neyman-Pearsnove leme pa je, po dijelu (ii) tog teorema, UMP-test na razini značajnosti α za jednostavne hipoteze H_0^0 vs. H_1^0.

Uzmimo sada da su $\theta', \theta'' \in \Theta$ dvije točke takve da je $\theta' < \theta''$ i $\alpha' := \gamma_\tau(\theta') \in \langle 0, 1\rangle$. Tada se na potpuno isti način kao u slučaju vrijednosti $\theta_0 < \theta_1$, pokaže da je test τ UMP-test za testiranje hipoteza $H_0^{0'} : \theta = \theta'$ vs. $H_1^{0'} : \theta = \theta''$ na razini značajnosti α'. Primjenom korolara 5.5 slijedi da je nužno $\gamma_\tau(\theta') < \gamma_\tau(\theta'')$, dakle, γ_τ je strogo rastuća funkcija na skupu $\{\theta \in \Theta : 0 < \gamma_\tau(\theta) < 1\}$. Time smo pokazali tvrdnju (ii).

Preostaje dovršiti dokaz tvrdnje (i). Zbog rasta funkcije jakosti od τ, slijedi da je značajnost:

$$\alpha_\tau = \sup_{\theta \le \theta_0} \gamma_\tau(\theta) = \gamma_\tau(\theta_0) \overset{(5.8)}{=} \alpha.$$

Dakle, relacijom (5.8) zadana je značajnost testa τ kao testa hipoteza H_0 vs. H_1. Budući da je test τ UMP-test za testiranje jednostavnih hipoteza $H_0^0 : \theta = \theta_0$ vs. $H_1^0 : \theta = \theta_1$ na razini značajnosti α, za svaki drugi test τ' koji je razine značajnosti α za testiranje početnih složenih hipoteza $H_0 : \theta \le \theta_0$ vs. $H_1 : \theta > \theta_0$ vrijedi da je nužno $\gamma_{\tau'}(\theta_1) \le \gamma_\tau(\theta_1)$. Naime, testu τ' je tim više značajnost za jednostavnu nul-hipotezu H_0' također dominirana s α. Kako je θ_1 takav da je $\theta_1 > \theta_0$ proizvoljan, slijedi da je za sve $\theta > \theta_0$, $\gamma_{\tau'}(\theta) \le \gamma_\tau(\theta)$. Dakle, po definiciji 5.3, test τ je UMP-test za testiranje složenih hipoteza H_0 vs. H_1 na razini značajnosti α. Time je dovršen dokaz tvrdnje (i).

(iii): Tvrdnja (iii) se dokazuje na potpuno isti način kao tvrdnja (i).

(iv): Ako Neyman-Pearsonovu lemu 5.4 primijenimo na vrijednosti $\theta_1 \equiv \theta < \theta_0$, tada je test $\tilde{\tau} := 1 - \tau$ UMP-test za testiranje odgovarajućih jednostavnih hipoteza, na razini značajnosti $1 - \alpha$. To znači da za svaki drugi test τ' sa značajnosti α vrijedi da je:

$$1 - \gamma_{\tau'}(\theta_1) = \mathbb{E}_{\theta_1}[(1-\tau')(\mathbf{X})] = \gamma_{1-\tau'}(\theta_1) \le \gamma_{1-\tau}(\theta_1) = \mathbb{E}_{\theta_1}[(1-\tau)(\mathbf{X})] = 1 - \gamma_\tau(\theta_1)$$

što je ekvivalentno tome da je

$$\gamma_\tau(\theta) \equiv \gamma_\tau(\theta_1) \le \gamma_{\tau'}(\theta_1) \equiv \gamma_{\tau'}(\theta).$$

Time je dokazana tvrdnja (iv) i teorem. □

Napomena 5.8. Ako u teoremu 5.7 umjesto zadanih hipoteza stavimo da je

$$H_0 : \;\; \theta \ge \theta_0 \;\; \text{nasuprot} \;\; H_1 : \;\; \theta < \theta_0$$

tada će teorem vrijediti s definicijskom relacijom (5.8) za τ u kojoj je znak za uređaj '>' zamijenjen s '<' i obratno, i s tvrdnjom (*ii*) u kojoj je riječ 'rastuća' zamijenjena s riječi 'padajuća'. $\square$

Primjer 5.9. Neka je X slučajni uzorak duljine 1 iz modela s hipergeometrijskom razdiobom $H(n, M, N)$ pri čemu su nam n i veličina populacije N poznati, a M je (nepoznati) parametar distribucije. Pretpostavljamo da vrijedi da je $n < M < N$ pa je parametarski prostor $\Theta = \{n+1, n+2, \ldots, N-1\}$. Funkcija gustoće je tada

$$f(x; M) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \cdot \mathbb{1}_{\{0,1,\ldots,n\}}(x), \quad x \in \mathbb{R}.$$

Primijetite da je tada funkcija gustoće uzorka X duljine 1 jednaka $f(\cdot; M)$. Neka je $M_0 \in \Theta$ zadana vrijednost. Želimo testitrati hipoteze:

$$H_0: \; M \leq M_0 \;\text{ nasuprot }\; H_1: \; M > M_0.$$

Budući da je omjer vjerostojnosti za $x \in \{0, 1, \ldots, n\}$ i $n < M < N-1$ jednak:

$$\frac{L(M+1; x)}{L(M; x)} = \frac{f(x; N+1)}{f(x; N)} = \frac{\frac{\binom{M+1}{x}\binom{N-M-1}{n-x}}{\binom{N}{n}}}{\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}} = \frac{M+1}{N-M} \cdot \frac{N-M-n+x}{M+1-x},$$

vidimo da je strogo rastući u $\xi(x) = x$. Budući da model ima svojstvo monotonog omjera vjerodostojnosti u $\xi(x) = x$, prema teoremu 5.7, UMP-test za testiranje zadanih hipoteza na zadanoj razini značajnosti $\alpha \in \langle 0, 1\rangle$ je test:

$$\tau(x) = \begin{cases} 1, & x > c \\ \beta, & x = c \\ 0, & x < c \end{cases}$$

takav da za $c \in \{0, 1, \ldots, n\}$ i $\beta \in [0, 1\rangle$ vrijedi:

$$\begin{aligned} \alpha &= \mathbb{E}_{M_0}[\tau(X)] = \mathbb{P}_{M_0}(X > c) + \beta \cdot \mathbb{P}_{M_0}(X = c) = \\ &= 1 - F(c; M_0) + \beta(F(c; M_0) - F(c-1; M_0)). \end{aligned}$$

Ovdje je $F(\cdot; M)$ funkcija distribucije od X. Da bismo odredili c i β, sprovodimo sljedeći algoritam. Za $c = 1, 2, \ldots, n-1$ računamo:

$$\beta = \frac{\alpha + F(c; M_0) - 1}{F(c; M_0) - F(c-1; M_0)}.$$

Prvi puta kada je $\beta \in [0, 1\rangle$ zaustavimo se. Vrijednosti za β i c kod kojih smo zaustavili algoritam su tražene vrijednosti koje ulaze u definiciju od τ. Primijetimo da smo vrijednosti od c tražili u nosaču gustoće statistike $\xi(X) = X$. $\square$

U nastavku ćemo primijeniti teorem 5.7 na jednoparametarske eksponencijalne familije $\mathcal{P} = \{f(\cdot; \theta) : \theta \in \Theta\}$, neprekidne ili diskretne. Dakle, gustoće $f(\cdot; \theta) \in \mathcal{P}$ su oblika:

$$f(x; \theta) = C(\theta) \cdot e^{Q(\theta)t(x)} \cdot h(x), \quad x \in \mathbb{R}, \tag{5.9}$$

za $C(\theta) > 0$, preslikavanje $Q : \Theta \to \mathbb{R}$ i izmjeriva preslikavanja t i $h \geq 0$.

Korolar 5.10. *Neka je $\mathcal{P}$ jednoparametarska eksponencijalna familija za koju vrijedi da je preslikavanje $Q : \Theta \to \mathbb{R}$ iz izraza za gustoće* (5.9)*, strogo monotona funkcija. Tada postoji uniformno najjači test τ na zadanoj razini značajnosti $\alpha \in \langle 0, 1\rangle$ za testiranje hipoteza*

$$H_0 : \ \ \theta \le \theta_0 \ \ \textit{nasuprot} \ \ H_1 : \ \ \theta > \theta_0.$$

Ako je Q strogo rastuća funkcija, tada je

$$\tau(\mathbf{x}) = \begin{cases} 1, & \xi(\mathbf{x}) > c \\ \beta, & \xi(\mathbf{x}) = c \\ 0, & \xi(\mathbf{x}) < c, \ \ \mathbf{x} \in \mathbb{R}^n, \end{cases}$$

gdje je $\xi(\mathbf{x}) = t(x_1) + t(x_2) + \cdots + t(x_n)$ za $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ i za t funkciju iz izraza za gustoće (5.9)*. Konstante $\beta \in [0, 1\rangle$ i $c \in \mathbb{R}$ su određene relacijom:*

$$\mathbb{E}_{\theta_0}[\tau(\mathbf{X})] = \alpha$$

gdje je $\mathbf{X}$ slučajni uzorak duljine n ($n \in \mathbb{N}$) iz modela $\mathcal{P}$.
Ako je Q strogo padajuća funkcija, tada su nejednakosti u izrazu za τ obratne.

Dokaz. Dovoljno je provjeriti da jednoparametarski eksponencijalni model sa strogo monotonom funkcijom Q ima svojstvo monotonog omjera vjerodostojnosti u funkciji $\xi(\mathbf{x}) = \sum_{i=1}^n t(x_i)$ ako je Q strogo rastuća, odnosno u funkciji $-\xi(\mathbf{x})$ ako je Q strogo padajuća funkcija.

Prvo, neka je Q strogo rastuća funkcija. Neka su $\theta', \theta'' \in \Theta$ takve vrijedosnosti da je $\theta' < \theta''$. Tada je omjer vjerodostojnosti:

$$\frac{L(\theta''|\mathbf{x})}{L(\theta'|\mathbf{x})} = \prod_{i=1}^n \frac{f(x_i;\theta'')}{f(x_i;\theta')} \overset{(5.9)}{=} \left(\frac{C(\theta'')}{C(\theta')}\right) \cdot \exp\left((Q(\theta'') - Q(\theta')) \cdot \sum_{i=1}^n t(x_i)\right).$$

Budući da je Q strogo rastuća funkcija i $\theta' < \theta''$, slijedi da je $Q(\theta'') - Q(\theta') > 0$, pa je omjer vjerodostojnosti strogo rastuća (eksponencijalna) funkcija po $\xi(\mathbf{x}) = \sum_{i=1}^n t(x_i)$.

Ako je Q strogo padajuća funkcija, onda je:

$$\frac{L(\theta''|\mathbf{x})}{L(\theta'|\mathbf{x})} = \left(\frac{C(\theta'')}{C(\theta')}\right) \cdot \exp\left((Q(\theta') - Q(\theta'')) \cdot \left(-\sum_{i=1}^n t(x_i)\right)\right).$$

Sada je $Q(\theta') - Q(\theta'') > 0$ za $\theta' < \theta''$, pa je omjer vjerodostojnosti strogo rastuća funkcija u $-\xi(\mathbf{x}) = -\sum_{i=1}^n t(x_i)$.

Dakle, prema teoremu 5.7 slijede tvrdnje. □

Napomena 5.11. Ako u korolaru 5.10 umjesto zadanih hipoteza stavimo da je

$$H_0 : \ \ \theta \ge \theta_0 \ \ \text{nasuprot} \ \ H_1 : \ \ \theta < \theta_0$$

tada će korolar vrijediti ako u definicijskom izrazu za τ znak za uređaj '$>$' zamijenimo sa znakom '$<$' i obratno. To je direktna posljedica teorema 5.7 i slijedi iz napomene 5.8. □

Napomena 5.12. Primijetite da je $T = \xi(\mathbf{X})$ prirodna dovoljna statistika. Ako je T neprekidna slučajna varijabla, tada je $\mathbb{P}_{\theta_0}(T = c) = \mathbb{P}(\xi(\mathbf{X}) = c) = 0$ za bilo koji $c \in \mathbb{R}$ pa u definiciji testa τ iz korolara 5.10 možemo uzeti da je konstanta β jednaka 0. □

Primjer 5.13. Neka je $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak duljine n ($n \in \mathbb{N}$) iz normalnog modela $N(\mu, \sigma_0^2)$ s poznatom varijancom $\sigma_0^2 > 0$ i parametrom očekivanja $\mu \in \mathbb{R}$. Neka je μ_0 zadani realni broj.

(a) Pretpostavimo, prvo, da želimo testitrati hipotezu $H_0 : \mu \leq \mu_0$ nasuprot $H_1 : \mu > \mu_0$. Budući da je, za $\mu' < \mu''$, omjer vjerostojnosti:

$$\frac{L(\mu''|\mathbf{x})}{L(\mu'|\mathbf{x})} = \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}((x_i - \mu'')^2 - (x_i - \mu'')^2)\right) = e^{\frac{n}{2\sigma_0^2}(\mu''^2 - \mu'^2)} \cdot e^{\frac{1}{\sigma_0^2}(\mu'' - \mu')\xi(\mathbf{x})}$$

monotono rastuća funkcija od $\xi(\mathbf{x}) = x_1 + x_2 + \cdots + x_n$ (za realizaciju $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ uzorka $\mathbf{X}$), prema korolaru 5.10 i napomeni 5.12, UMP-test na razini značajnosti α za testiranje navedenih hipoteza je

$$\tau_1(\mathbf{x}) = \mathbb{1}_{\{\xi(\mathbf{x}) > c\}} = \mathbb{1}_{\{n\overline{x} > c\}} = \mathbb{1}_{\left\{\frac{\overline{x} - \mu_0}{\sigma_0}\sqrt{n} > c'\right\}} = \mathbb{1}_{\left\{\frac{\overline{x} - \mu_0}{\sigma_0}\sqrt{n} > z_\alpha\right\}}, \quad \mathbf{x} \in \mathbb{R}^n.$$

Naime, konstante c, odnosno c' odrede se iz uvjeta da je

$$\alpha = \gamma_{\tau_1}(\mu_0) = \mathbb{P}_{\mu_0}\left(\frac{\overline{X}_n - \mu_0}{\sigma_0}\sqrt{n} > c'\right) = 1 - \Phi(c').$$

Dakle, $c' = z_\alpha$, $(1 - \alpha)$-kvantilu jedinične normalne razdiobe. Test hipoteza o očekivanju normalne populacije kojemu je testna statistika

$$Z = \frac{\overline{X}_n - \mu_0}{\sigma_0}\sqrt{n}$$

zovemo *z-test*. Podsjetimo se da je z-test *test omjera vjerodostojnosti* za testiranje hipoteza o parametru očekivanja u normalnom modelu s poznatom populacijskom varijancom.

(b) Analogno, UMP-test za testiranje hipoteza $H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$ na razini značajnosti α je

$$\tau_2(\mathbf{x}) = \mathbb{1}_{\left\{\frac{\overline{x} - \mu_0}{\sigma_0}\sqrt{n} < -z_\alpha\right\}}, \quad \mathbf{x} \in \mathbb{R}^n,$$

dakle, radi se također o z-testu.

(c) *Postoji li uniformno najjači test za testiranje jednostavne nul-hipoteze $H_0 : \ \mu = \mu_0$ nasuprot dvostrane alternativne hipoteze $H_1 : \ \mu \neq \mu_0$?* Kada bi postojao takav UMP-test τ^* na razini značajnosti $\alpha \in \langle 0, 1\rangle$, tada bi on bio i UMP-test na istoj razini značajnosti za testiranje parova jednostavnih hipoteza $H_0' : \mu = \mu_0$ vs. $H_1' : \mu = \mu_1 > \mu_0$ i $H_0' : \mu = \mu_0$ vs. $H_1'' : \mu = \mu_2 < \mu_0$ za μ_1, μ_2 proizvoljne takve brojeve. Prema teoremu 5.4 (iii), tada bi nužno bilo da je $\tau^* = \tau_1$ λ-s.s. i $\tau^* = \tau_1$ λ-s.s., gdje su τ_1 i τ_2 z-testovi iz dijelova (a) i (b) ovoga primjera, a to je nemoguće. Ipak, za testiranje navedenih dvostranih hipoteza koristi se z-test i za te hipoteze iako nije UMP-test jer, kao test omjera vjerodostojnosti, ima dobra asimptotska svojstva (ali kojima se ovdje nećemo baviti). □

Zadatak 5.14. Neka je $\mathbf{X} = (X_1, \ldots, X_n)$ slučajni uzorak iz Bernoullijevog modela $b(1, \theta)$, $\theta \in \langle 0, 1\rangle$. Nađite UMP-test za testiranje $H_0 : \theta \geq 1/2$ vs. $H_1 : \theta < 1/2$ na razini značajnosti α. Ako za opaženu vrijednost $\mathbf{x}$ uzorka $\mathbf{X}$ duljine $n = 10$ vrijedi $\sum_{i=1}^{n} x_i = k = 3$, sprovedite dobiveni test na razini značajnosti od $\alpha = 5\%$.

Zadatak 5.15. Neka je $\mathbf{X} = (X_1, \ldots, X_n)$ slučajni uzorak iz eksponencijalnog modela $\text{Exp}\,(\lambda)$, $\lambda \in \langle 0, +\infty\rangle$. Nađite UMP-test za testiranje $H_0 : \lambda \leq 1$ vs. $H_1 : \lambda > 1$ na razini značajnosti α. Pretpostavite da je uzorak velike duljine.

Dodatak A

Fubinijev teorem

Neka su $(X, \mathcal{F}, \mu)$ i $(Y, \mathcal{G}, \nu)$ dva prostora sa σ-konačnim mjerama μ i ν. Označimo sa $\mathcal{F} \times \mathcal{G}$ najmanju σ-algebru generiranu množinom

$$\mathcal{S} = \{F \times G \,:\, F \in \mathcal{F},\, G \in \mathcal{G}\}.$$

Tada postoji jedinstvena mjera, u oznaci $\mu \times \nu$, definirana na izmjerivom prostoru $(X \times Y, \mathcal{F} \times \mathcal{G})$ takva da je (vidjeti [9], teorem 10.15, str. 328., ili [11], §8.2, str. 77.-78.)

$$(\forall F \times G \in \mathcal{S}) \;\; (\mu \times \nu)(F \times G) = \mu(F) \cdot \nu(G).$$

Mjeru $\mu \times \nu$ zovemo *produktna mjera* od μ i ν, a prostor mjere $(X \times Y, \mathcal{F} \times \mathcal{G}, \mu \times \nu)$ *produkt prostora mjera* $(X, \mathcal{F}, \mu)$ *i* $(Y, \mathcal{G}, \nu)$. Produktna mjera je također σ-konačna.

Neka $\mathcal{L}^1(Z, \mathcal{H}, \gamma)$ označava množinu svih $\mathcal{H}$-izmjerivih funkcija $g : Z \to \mathbb{R}$ definiranih na prostoru mjere $(Z, \mathcal{H}, \gamma)$ takvih da je $\int\limits_Z |g| \, d\gamma < +\infty$.

Teorem A.1. (Fubini) *Neka su $(X, \mathcal{F}, \mu)$ i $(Y, \mathcal{G}, \nu)$ prostori σ-konačnih mjera, te neka je $g : X \times Y \to \mathbb{R}$ funkcija.*

(i) Ako je g nenegativna $\mathcal{F} \times \mathcal{G}$-izmjeriva funkcija, tada postoje $\mathcal{F}$-izmjeriva nenegativna funkcija $g_X : X \to \overline{\mathbb{R}}$ i $\mathcal{G}$-izmjeriva nenegativna funkcija $g_Y : Y \to \overline{\mathbb{R}}$ takve da je $g_X(x) = \int\limits_Y g(x,y) d\nu(y)$ za $x \in X$ te $g_Y(y) = \int\limits_X g(x,y) d\mu(x)$ za $y \in Y$ i da vrijedi

$$\int\limits_X g_X(x)\, d\mu(x) = \int\limits_Y g_Y(y)\, d\nu(y) = \int\limits_{X \times Y} g(x,y)\, d(\mu \times \nu)(x,y). \tag{A.1}$$

(ii) Ako je $g \in \mathcal{L}^1(X \times Y, \mathcal{F} \times \mathcal{G}, \mu \times \nu)$, tada postoje funkcije $g_X \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ i $g_Y \in \mathcal{L}^1(Y, \mathcal{G}, \nu)$ takve da je $g_X(x) = \int\limits_Y g(x,y) d\nu(y)$ za μ-s.s. $x \in X$ te $g_Y(y) = \int\limits_X g(x,y) d\mu(x)$ za ν-s.s. $y \in Y$ i da vrijedi relacija (A.1).

(iii) Ako je g $\mathcal{F} \times \mathcal{G}$-izmjeriva funkcija i vrijedi da je

$$\int\limits_X \Big(\int\limits_Y |g(x,y)|\, d\nu(y) \Big)\, d\mu(x) < +\infty \;\; \textit{ili} \;\; \int\limits_Y \Big(\int\limits_X |g(x,y)|\, d\mu(x) \Big)\, d\nu(y) < +\infty, \tag{A.2}$$

tada je $g \in \mathcal{L}^1(X \times Y, \mathcal{F} \times \mathcal{G}, \mu \times \nu)$ pa vrijedi zaključak od (ii).

Relacija (A.1) se obično zapisuje u obliku:

$$\int\limits_X (\int\limits_Y g(x,y)\,d\nu(y))\,d\mu(x) = \int\limits_Y (\int\limits_X g(x,y)\,d\mu(x))\,d\nu(y) = \int\limits_{X\times Y} g(x,y)\,d(\mu\times\nu)(x,y). \quad (A.3)$$

Za dokaz teorema vidjeti [7], str. 115., §5.2, teorem 2 i [9], str. 330.-332., teorem 10.16 ili [11], str. 77.-78., §8.2 (za tvrdnje $(i-ii)$), te [9], str. 332., korolar 10.17 (za tvrdnju (iii)).

Na osnovi prethodnog teorema, za niz prostora $(X_i, \mathcal{F}_i, \mu_i)$, $0 \le i \le n$ $(n \ge 3)$, sa σ-konačnim mjerama induktivno se može definirati njihov produkt relacijom:

$$\Pi_{i=1}^n \mu_i := (\Pi_{i=1}^{n-1}\mu_i) \times \mu_n$$

i analogno produkt njihovih prostora mjera (vidjeti [9], §10.6, str. 327.-328.). Ako se radi o istim prostorima σ-konačne mjere, tj. $(X_i, \mathcal{F}_i, \mu_i) \equiv (X, \mathcal{F}, \mu)$ za svaki i, tada je dobro definirana *n-ta potencija mjere* μ, u oznaci μ^n, sa

$$\mu^n := \Pi_{i=1}^n \mu.$$

U primijeni Fubinijevog teorema sljedeće činjenice treba imati na umu.

1.) Budući da su zakoni razdioba slučajnih varijabli ili vektora definirani na Borelovim izmjerivim prostorima, Fubinijev teorem ćemo najčešće primijenjivati u tom kontekstu. Primijetimo (vidjeti, na primjer, [9], §8., zadatak 12. (a), str. 247.) da za σ-algebre Borelovih skupova vrijedi da je $(m, n \in \mathbb{N})$

$$\mathcal{B}(\mathbb{R}^m) \times \mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}^{m+n}).$$

2.) Nadalje, ako je X m-dimenzionalna, a Y n-dimenzionalna slučajna veličina, te ako su nezavisne, tada je zakon razdiobe slučajnog vektora (X, Y) dimenzije $m+n$ jednak produktu zakona razdioba od X i Y (prema definiciji 1.27 nezavisnosti slučajnih veličina i produkta mjera), tj.

$$\mathbb{P}_{X,Y} \overset{\text{(nez.)}}{=} \mathbb{P}_X \times \mathbb{P}_Y.$$

3.) Zakoni razdioba neprekidnih ili diskretnih slučajnih veličina su integrali po Lebesgueovoj, odnosno odgovarajućoj brojećoj mjeri. Obje mjere su σ-konačne. Pri tome primijetimo da je Lebesgueova mjera na k-dimenzionalnom prostoru $\mathbb{R}^k$, u oznaci λ^k, jednaka k-toj potenciji Lebesgueove mjere λ^1 na $\mathbb{R}$. Dakle,

$$\lambda^k = (\lambda^1)^k = \overbrace{\lambda^1 \times \cdots \times \lambda^1}^{k}.$$

Nadalje, neka je μ_D brojeća mjera na $\mathbb{R}^k$ obzirom na prebrojiv skup D. Ako sa π_i označimo *projekciju na i-tu koordinatu*, dakle, preslikavanje $\pi_i : \mathbb{R}^k \to \mathbb{R}$, $\pi_i(x_1, x_2, \ldots, x_k) := x_i$, za $1 \le i \le k$, te sa $D_i := \pi_i(D) \subset \mathbb{R}$, $1 \le i \le k$, projekcije skupa D na odgovarajuće koordinatne osi, tada je

$$\mu_D(B) = \int\limits_B \mathbb{1}_D(x)\,d\mu_{\Pi_{i=1}^k D_i}(x) = \int\limits_B \mathbb{1}_D(x)\,d\Pi_{i=1}^k \mu_{D_i}(x) = \sum_{x\in\Pi_{i=1}^k D_i} \mathbb{1}_{D\cap B}(x),$$

gdje je $\mu_{\Pi_{i=1}^k D_i}$ brojeća mjera u odnosu na Kartezijev produkt $\Pi_{i=1}^k D_i = D_1 \times D_2 \times \cdots \times D_k$, te vrijedi da je

$$\mu_{\Pi_{i=1}^k D_i} = \Pi_{i=1}^k \mu_{D_i}.$$

Drugim riječima, brojeća mjera μ_D je apsolutno neprekidna u odnosu na brojeću mjeru $\mu_{\Pi_{i=1}^k D_i}$ (s Radon-Nikodymovom derivacijom $\mathbb{1}_D$), a koja je produkt brojećih mjera. Zbog toga Fubinijev teorem možemo primijeniti na integrale u odnosu na Lebeguesovu i bilo koju brojeću mjeru.