Brownov most i Kolmogorov-Smirnovljeva statistika

Blažević, Nikolina

Master's thesis / Diplomski rad

2016

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:258241

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-09-11**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Nikolina Blažević

BROWNOV MOST I KOLMOGOROV-SMIRNOVLJEVA STATISTIKA

Diplomski rad

Voditelj rada:
doc. dr. sc. Ante Mimica

Zagreb, veljača 2016.

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

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

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Sadržaj **iii**

Uvod **1**

1 Izabrane definicije i rezultati iz teorije vjerojatnosti **2**
1.1 Vjerojatnosni prostor . . . 2
1.2 Svojstva slučajnih varijabli . . . 4
1.3 Primjeri slučajnih varijabli . . . 10
1.4 Empirijska funkcija distribucije . . . 16

2 Brownovo gibanje **20**
2.1 Gaussovski proces . . . 22
2.2 Jako Markovljevo svojstvo . . . 24
2.3 Lévyjev zakon trojke . . . 29
2.4 Princip invarijantnosti . . . 33
2.5 Brownov most . . . 36

3 Kolmogorov-Smirnovljeva statistika **41**
3.1 Testiranje statističkih hipoteza . . . 41
3.2 Kolmogorov-Smirnovljev test . . . 42
3.3 Asimptotski rezultat . . . 44

Bibliografija **49**

Uvod

Promatrate li kroz mikroskop čestice peludi u kapljici vode primijetit ćete neprekidno, nepravilno gibanje čestica koje ne ovisi o sastavu i gustoći čestica, o gibanju drugih čestica, o tokovima u tekućini, niti o isparavanju. Ovo su neke od činjenica koje je škotski botaničar Robert Brown prvi zabilježio u svojim znanstvenim radovima iz 1828. godine *On the existence of active molecules in organic and inorganic bodies* i *Additional remarks on active molecules* promatrajući ovaj fenomen.

Slijedeći Brownove opservacije pojavilo se nekoliko teorija među kojima je i ona Einsteinova iz 1905. godine koja je dala točno objašnjenje i smirila ondašnje diskusije o postojanju atoma. Jean Perrin je 1909. godine kombinirao Einsteinovu teoriju i eksperimentalne opservacije Brownovog gibanja kako bi potvrdio postojanje i odredio veličinu atoma. Nezavisno od Einsteina, M. von Smoluchowski došao je do jednake interpretacije Brownovog gibanja. Brownovo gibanje kao matematički objekt prvi puta rigozorno je definirao Wiener koji uvodi Wienerovu mjeru na prostoru $C[0, 1]$ gradeći na Einsteinovom i von Smoluchowskijevom radu. Zatim slijede konstrukcije Kolmogorova, Lévya, Ciesielskija i Donskera. Povijesno gledajući, Brownovo gibanje prvi je stohastički proces u neprekidnom vremenu i kao takav utjecao je primjerice na razvoj Gaussovskih procesa, martingala i Markovljevih procesa pa bismo stoga mogli nedvojbeno reći da je Brownovo gibanje jedan od najvažnijih stohastičkih procesa.

Ovaj matematički pojam našao je mnoge primjene u znanosti i financijskoj matematici, a jedna od primjena u statistici pa dalje posredno u svim aspektima primjene statistike opisana je u ovom radu koji je inspirian knjigom S. Resnicka ([9]). Cilj ovog rada jest izvod asipmtotske distribucije Kolmogorov-Smirnovljeve statistike koja služi za testiranje činjenice dolazi li slučajni uzorak iz neke populacije s neprekidnom razdiobom.

Prvo poglavlje daje podlogu o teoriji vjerojatnosti. Nakon toga uvodi se pojam Brownovog mosta i dokazuje jako Markovljevo svojstvo i Lévyjev zakon trojke. Iskazan je princip invarijantnosti, definiran Brownov most i dokazano svojstvo da je Brownov most aproksimativno jednak Brownovom gibanju uvjetovanom da ima vrijednost 0 u vremenu 1. U zadnjem poglavlju nalazi se pregled osnovnih pojmova o testiranju statističke hipoteze, Kolmogorov-Smirnovljeva statistika i pripadajući test i konačno izvod asimptotske distribucije Kolmogorov-Smirnovljeve statistike.

Poglavlje 1

Izabrane definicije i rezultati iz teorije vjerojatnosti

Kao što i sam naslov poglavlja sugerira, u prvom poglavlju ovog rada, pomalo i nepravedno stavljamo teoriju vjerojatnosti u okvir, odnosno uzet ćemo samo osnovne i potrebne rezultate i definicije. Smisao ovog poglavlja jest kratko ponavljanje ili uvođenje čitatelja u pojmove bez kojih je teško razumjeti ostatak rada. Nakon što uvedemo pojmove potrebne za definiranje vjerojatnosnog prostora, definiramo slučajnu varijablu i navodimo svojstva. U potpoglavlju 1.3 nalaze se primjeri slučajnih varijabli i vektora, te samo oni rezultati koji će nam koristiti u nastavku rada. Opširnije o ovoj temi možete pronaći u [11] odakle su preuzete definicije i rezultati iz ovog poglavlja. Ako nije drugačije navedeno, dokaze tvrdnji možete pronaći u toj knjizi. Dobro je znati da je najznačajniji dio ovog poglavlja za rad upravo potpoglavlje 1.4 gdje od definicije empirijske funkcije i raspisa njenih svojstava dolazimo do početne motivacije iz koje će uslijediti i konačni rezultat ovog rada.

1.1 Vjerojatnosni prostor

Neka je Ω neprazan skup i $\mathcal{F}$ familija podskupova od Ω. Kažemo da je $\mathcal{F}$ σ-algebra na Ω ako vrijedi

(i) $\emptyset \in \mathcal{F}$

(ii) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

(iii) $A_i \in \mathcal{F}, i \in \mathbb{N} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Dakle, $\mathcal{F}$ je σ-algebra ako je zatvorena na komplementiranje i prebrojive unije, a lako se pokaže da je zatvorena i na prebrojive presjeke i skupovne razlike:

$$A_i \in \mathcal{F}, i \in \mathbb{N} \implies \bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c \right)^c \in \mathcal{F}$$
$$A_1, A_2 \in \mathcal{F} \implies A_1 \backslash A_2 = A_1 \cap A_2^c \in \mathcal{F}.$$

Trivijalne σ-algebre skupa Ω su $\{\emptyset, \Omega\}$ i $\mathcal{P}(\Omega)$. $\mathcal{P}(\Omega)$ jest oznaka za pojam partitivnog skupa od Ω, što je po definiciji skup svih podskupova skupa Ω.

σ-algebru generiranu familijom svih otvorenih skupova na $\mathbb{R}$ nazivamo σ-algebra Borelovih skupova i označavamo s $\mathcal{B}$. Kažemo da je funkcija $f : \mathbb{R} \to \mathbb{R}$ Borelova, odnosno izmjeriva funkcija ako vrijedi da je $f^{-1}(B) \in \mathcal{B}$ za svaki $B \in \mathcal{B}$. Svaka neprekidna funkcija $f : \mathbb{R} \to \mathbb{R}$ je Borelova funkcija. Poopćeno, $\mathcal{B}^n$ je σ-algebra na $\mathbb{R}^n$ generirana familijom svih otvorenih skupova na $\mathbb{R}^n$ i analogno vrijedi tvrdnja za funkcije na $\mathbb{R}^n$.

Izmjeriv prostor je uređen par $(\Omega, \mathcal{F})$ gdje je $\mathcal{F}$ σ-algebra na Ω. Mjera na $(\Omega, \mathcal{F})$ jest funkcija $\mu : \mathcal{F} \to [0, +\infty]$ koja zadovoljava uvjete:

(i) $\mu(\emptyset) = 0$

(ii) Za niz međusobno disjunktnih skupova $(A_n)_{n\in\mathbb{N}}$ iz $\mathcal{F}$ vrijedi $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

Ako je $\mu(\Omega) = 1$, tada se mjera μ naziva vjerojatnosnom mjerom i označavat ćemo je s P. Navedimo i neka osnovna svojstva vjerojatnosti

- (monotonost) $A, B \in \mathcal{F}$ i $A \subset B \implies \mathrm{P}(A) \le \mathrm{P}(B)$
- (neprekidnost u odnosu na rastući/padajući niz događaja)
  $A_n \in \mathcal{F}, n \in \mathbb{N}, A_1 \subset A_2 \subset \dots$ i $A = \bigcup_{n=1}^{\infty} A_n \implies \mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n)$
  $A_n \in \mathcal{F}, n \in \mathbb{N}, A_1 \supset A_2 \supset \dots$ i $A = \bigcap_{n=1}^{\infty} A_n \implies \mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n)$
- (σ-poluaditivnost) $A_n \in \mathcal{F}, n \in \mathbb{N} \implies \mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} \mathrm{P}(A_n)$
- (vjerojatnost unije dvaju događaja) $A, B \in \mathcal{F} \implies \mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$
- (vjerojatnost suprotnog događaja) $A \in \mathcal{F} \implies \mathrm{P}(A^c) = 1 - \mathrm{P}(A)$.

Konačno, definirajmo osnovni objekt u teoriji vjerojatnosti. Uređena trojka $(\Omega, \mathcal{F}, \mathrm{P})$ gdje je Ω skup, $\mathcal{F}$ σ-algebra na Ω i P vjerojatnost na $(\Omega, \mathcal{F})$ naziva se vjerojatnosni prostor.

Elemente σ-algebre $\mathcal{F}$ nazivamo događajima, a broj P(A), za A $\in \mathcal{F}$ jest vjerojatnost događaja A.

Neka je $(\Omega, \mathcal{F}, \mathrm{P})$ vjerojatnosni prostor. Funkcija $X : \Omega \to \mathbb{R}$ jest slučajna varijabla na Ω ako za svaki Borelov skup $B \subset \mathbb{R}$ vrijedi $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$. Generalizirano, funkcija $X : \Omega \to \mathbb{R}^n$ jest slučajni vektor (n-dimenzionalni slučajni vektor) na Ω ako vrijedi $X^{-1}(B) \in \mathcal{F}$ za svako $B \in \mathcal{B}^n$. Ako je X slučajni vektor i $f : \mathbb{R}^n \to \mathbb{R}^m$ Borelova funkcija, tada je $f \circ X = f(X)$ m-dimenzionalni slučajni vektor.

1.2 Svojstva slučajnih varijabli

Konvergencija slučajnih varijabli

Niz slučajnih varijabli $(X_n)_{n\in\mathbb{N}}$ na Ω konvergira

- gotovo sigurno (g.s.) prema $X : \Omega \to \mathbb{R}$ ako je

$$\mathrm{P}\left(\{\omega \in \Omega : \lim X_n(\omega) = X(\omega)\}\right) = 1.$$

  Konvergenciju g.s od X_n prema X zapisujemo $\lim\limits_{n\to\infty} X_n = X$ g.s ili $X_n \xrightarrow{g.s.} X$.

- po vjerojatnosti prema $X : \Omega \to \mathbb{R}$ ako za svaki $\epsilon > 0$

$$\mathrm{P}(|X_n - X| > \epsilon) \to 0, \quad n \to \infty.$$

  Konvergenciju po vjerojatnosti od X_n prema X zapisujemo $X_n \xrightarrow{\mathrm{P}} X$.

- po distribuciji prema $X : \Omega \to \mathbb{R}$ ako vrijedi da je za svaki x u kojem je F neprekidna

$$\lim_{n\to\infty} F_n(x) = F(x).$$

  Konvergenciju po distribuciji od X_n prema X zapisujemo $X_n \implies X$. Slučajnu varijablu X nazivamo još i asimptotska distribucija.

Funkcija distribucije

Definicija 1.2.1. *Neka je X slučajna varijabla na Ω. Funkcija distribucije od X jest funkcija $F : \mathbb{R} \to [0,1]$ definirana sa*

$$F(x) = \mathrm{P}(X^{-1}(-\infty, x]) = \mathrm{P}\left(\{\omega \in \Omega : X(\omega) \leq x\}\right) = \mathrm{P}(X \leq x), \quad x \in \mathbb{R}.$$

Ako za slučajne varijable X i Y vrijedi $\mathrm{P}(X \leq z) = \mathrm{P}(Y \leq z)$, $z \in \mathbb{R}$, tada kažemo da su slučajne varijable X i Y jednake po distribuciji i pišemo $X \stackrel{d}{=} Y$.

Teorem 1.2.2. *Funkcija distribucije F slučajne varijable X zadovoljava sljedeća svojstva*

(i) F je rastuća funkcija

(ii) F je neprekidna zdesna na $\mathbb{R}$, odnosno $\lim_{y \to x+} F(y) = F(x)$

(iii) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \to +\infty} F(x) = 1$.

Definicija 1.2.3. *Neka je X slučajna varijabla s funkcijom distribucije F. Generalizirani inverz funkcije distribucije F definiramo sa*

$$F^{\leftarrow}(y) = \inf \{u : F(u) \geq y\} . \tag{1.1}$$

Propozicija 1.2.4. *Neka je F neprekidna funkcija distribucije i $F^{\leftarrow}$ pripadni generalizirani inverz. Tada vrijedi:*

(i) $F(F^{\leftarrow}(t)) \geq t$

(ii) $F^{\leftarrow}$ je strogo rastuća na $(0, 1)$

(iii) $F(F^{\leftarrow}(t)) = t$

(iv) $F^{\leftarrow}(y) \leq t \iff y \leq F(t)$.

Dokaz. Počnimo redom za svaku tvrdnju propozicije:

(i) Neka je (y_n) niz takav da je $F(y_n) \geq t$ i $y_n \downarrow F^{\leftarrow}(t)$. Budući da je F neprekidna slijedi $t \leq \lim_{n \to \infty} F(y_n) = F(F^{\leftarrow}(t))$.

(ii) Neka su $0 < a < b < 1$. Kako je F rastuća funkcija, tada vrijedi da je $\{u : F(u) \geq a\} \supset \{u : F(u) \geq b\}$, pa je $\inf\{u : F(u) \geq a\} \leq \inf\{u : F(u) \geq b\}$, odnosno $F^{\leftarrow}(a) \leq F^{\leftarrow}(b)$. Preostaje pokazati da vrijedi stroga nejednakost. Iz tvrdnje (i) slijedi $F(F^{\leftarrow}(b)) \geq b$. Budući da je F neprekidna, za $\epsilon = \frac{b-a}{2} > 0$ postoji $\delta > 0$ takav da je $F(u) \geq b - \epsilon > a$ za svaki $u \in (F^{\leftarrow}(b) - \delta, F^{\leftarrow}(b) + \delta)$. Tada zbog $F^{\leftarrow}(a) \leq F^{\leftarrow}(b) - \delta$ slijedi $F^{\leftarrow}(a) < F^{\leftarrow}(b)$.

(iii) Dokazat ćemo ovu tvrdnju tako da pokažemo da je $F(F^{\leftarrow}(t)) \leq t$ i $F(F^{\leftarrow}(t)) \geq t$. $F(F^{\leftarrow}(t)) \geq t$ vrijedi prema (i). Zbog neprekidnosti od F možemo pretpostaviti da je $F(a) = t$. Tada je $a \in \{u : F(u) \geq t\}$. Iz definicije infimuma slijedi da je $F^{\leftarrow}(t) \leq a$, a budući da je F rastuća i neprekidna slijedi $F(F^{\leftarrow}(t)) \leq F(a) = t$.

(iv) Pretpostavimo prvo da vrijedi $F^{\leftarrow}(y) \leq t$. Budući da je F rastuća i neprekidna imamo $F(F^{\leftarrow}(y)) \leq F(t)$ i kada iskoristimo tvrdnju pod (ii) slijedi $y \leq F(t)$. S druge strane, ako je $y \leq F(t)$, tada je $t \in \{u : F(u) \geq y\}$, pa je $F^{\leftarrow}(y) \leq t$ prema definiciji infimuma.

□

Diskretna slučajna varijabla

Neka je $(\Omega, \mathcal{F}, \mathrm{P})$ vjerojatnosni prostor, X slučajna varijabla na Ω i F pripadna funkcija distribucije. Slučajna varijabla X jest diskretna ukoliko postoji prebrojiv skup $D \subset \mathcal{B}(\mathbb{R})$ takav da je $\mathrm{P}(X \in D) = 1$. Funkcija gustoće diskretne slučajne varijable X jest funkcija $f : \mathbb{R} \to \mathbb{R}$ definirana sa

$$f(x) = \mathrm{P}(X = x)$$

i prema tome za funkciju distribucije definiranu u Definiciji 1.2.1 u diskretnom slučaju vrijedi

$$F(x) = \sum_{\{y \in D\,:\, y \leq x\}} f(y).$$

Diskretne slučajne varijable obično zadajemo tako da zadamo skup $D = \{x_1, x_2, \ldots\}$ i brojeve $p_n = \mathrm{P}(X = x_n)$, što zapisujemo u tablicu

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix} \tag{1.2}$$

koju nazivamo distribucija ili zakon razdiobe slučajne varijable X.

Neka su X_1 i X_2 diskretne slučajne varijable definirane na istom vjerojatnosnom prostoru $(\Omega,\mathcal{F},\mathrm{P})$. Zajednička funkcija gustoće jest funkcija $f_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}$ za koju vrijedi

$$f_{X_1,X_2}(x_1, x_2) = \mathrm{P}(X_1 = x_1, X_2 = x_2).$$

f_{X_1,X_2} naziva se i gustoćom slučajnog vektora (X_1, X_2). Za njihovu zajedničku funkciju distribucije (funkciju distribucije slučajnog vektora (X_1, X_2)) $F_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}$ vrijedi

$$F_{X_1,X_2}(x_1, x_2) = \mathrm{P}(X_1 \leq x_1, X_2 \leq x_2).$$

Za slučajni vektor (X_1, X_2) s gustoćom razdiobe f_{X_1,X_2} marginalna gustoća od X_1 jest

$$f_{X_1}(x_1) = \sum_{x_2 \in \mathrm{Im}\, X_2} f_{X_1,X_2}(X_1, X_2), \quad x_1 \in \mathbb{R}.$$

Ako je $f_{X_2} = \mathrm{P}(X_2 = x_2) > 0$, $x_2 \in \mathrm{Im}\, X_2$, tada je uvjetna gustoća od X_1 uz dano $X_2 = x_2$

$$f_{X_1|X_2} = \mathrm{P}(X_1 = x_1 | X_2 = x_2) = \frac{\mathrm{P}(X_1 = x_1, X_2 = x_2)}{\mathrm{P}(X_2 = x_2)} = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_2}(x_2)}, \quad x_1 \in \mathbb{R}.$$

Druga jednakost posljednjeg izraza jest definicija uvjetne vjerojatnosti.

Neprekidna slučajna varijabla

Neka je $(\Omega, \mathcal{F}, \mathrm{P})$ vjerojatnosni prostor, X slučajna varijabla na Ω i F pripadna funkcija distribucije. Slučajna varijabla X jest neprekidna slučajna varijabla ako postoji Borelova funkcija $f : \mathbb{R} \to \mathbb{R}_+$ takva da je

$$P(a \leq X \leq b) = \int_a^b f(y)\, dy.$$

Funkciju f nazivamo funkcijom gustoće slučajne varijable X. Prema tome, za funkciju distribucije definiranu u Definiciji 1.2.1, u neprekidnom slučaju vrijedi

$$F(x) = \int_{-\infty}^x f(y)\, dy. \tag{1.3}$$

Neka su X_1 i X_2 neprekidne slučajne varijable definirane na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathrm{P})$. Zajednička razdioba tih slučajnih varijabli jest funkcija $f_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}$ za koju vrijedi

$$P(a \leq X_1 \leq b, c \leq X_2 \leq d) = \int_a^b \int_c^d f_{X_1,X_2}\, dx_1\, dx_2.$$

f_{X_1,X_2} naziva se i gustoćom slučajnog vektora (X_1, X_2). Za njihovu zajedničku funkciju distribucije (funkciju distribucije slučajnog vektora (X_1, X_2)) $F_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}$ vrijedi

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(u, v)\, du\, dv.$$

Za slučajni vektor (X_1, X_2) s gustoćom razdiobe f_{X_1,X_2} marginalna gustoća od X_1 jest

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\, dx_2, \quad x_1 \in \mathbb{R}.$$

Ako je $f_{X_2} > 0$, $x_2 \in \operatorname{Im} X_2$, tada je uvjetna gustoća od X_1 uz dano $X_2 = x_2$

$$f_{X_1|X_2} = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_2}(x_2)}, \quad x_1 \in \mathbb{R}. \tag{1.4}$$

.

Matematičko očekivanje i varijanca

Definicija 1.2.5. *Definirajmo* $X^+ = \max\{X, 0\}$, $X^- = \max\{-X, 0\}$. *Kažemo da slučajna varijabla* $X = X^+ - X^-$ *ima matematičko očekivanje ukoliko je barem jedan od integrala*

$$\mathrm{E}X^+ := \int_\Omega X^+ dP, \quad \mathrm{E}X^- := \int_\Omega X^- dP$$

konačan. Tada matematičko očekivanje, u oznaci $\mathrm{E}X$*, definiramo kao* $\mathrm{E}X = \mathrm{E}X^+ - \mathrm{E}X^-$.

Teorem 1.2.6. *Neka su X i Y nenegativne slučajne varijable ili neka vrijedi* $\mathrm{E}|X| < \infty$, $\mathrm{E}|Y| < \infty$*. Tada vrijedi*

$$\mathrm{E}[aX + Y] = a\mathrm{E}X + \mathrm{E}Y.$$

Ovaj važan rezultat koji će biti korišten više puta nazivamo linearnost matematičkog očekivanja,odnosno linearnost očekivanja i u nastavku rada pozivat ćemo se na njega.

Ako je X diskretna slučajna varijabla, tada vrijedi

$$\mathrm{E}X = \sum_x x\mathrm{P}(X = x) = \sum_i x_i p_i, \tag{1.5}$$

$$\mathrm{E}[g(X)] = \sum_i g(x_i) p_i, \tag{1.6}$$

gdje su jednakosti napisane korištenjem oznaka iz (1.2).

Ako je X neprekidna slučajna varijabla s gustoćom f, tada je

$$\mathrm{E}X = \int_{-\infty}^{\infty} x f(x) dx, \tag{1.7}$$

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx, \tag{1.8}$$

gdje integriramo u odnosu na Lebesgueovu mjeru.

Definicija 1.2.7. *Varijanca slučajne varijable X definira se s*

$$VarX = \mathrm{E}[(X - \mathrm{E}X)^2] \in [0, \infty].$$

Ako je $VarX < \infty$*, onda kažemo da X ima konačnu varijancu.*

Korisno je vidjeti da zbog linearnosti matematičkog očekivanja vrijede i sljedeća dva zaključka:

$$\begin{aligned} VarX &= \mathrm{E}[(X - \mathrm{E}X)(X - \mathrm{E}X)] \\ &= \mathrm{E}[X^2 - X \cdot \mathrm{E}X - \mathrm{E}X \cdot X + (\mathrm{E}X)^2] \\ &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 - (\mathrm{E}X)^2 + (\mathrm{E}X)^2 \\ &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 \end{aligned} \tag{1.9}$$

$$\begin{aligned} Var[aX+b] &= \mathrm{E}[(aX+b)^2] - [\mathrm{E}(aX+b)]^2 \\ &= \mathrm{E}(a^2X^2 + 2abX + b^2) - (a\mathrm{E}X + b)^2 \\ &= a^2\mathrm{E}(X^2) + 2ab\mathrm{E}X + b^2 - a^2(\mathrm{E}X)^2 - 2ab\mathrm{E}X - b^2 \\ &= a^2[\mathrm{E}(X^2) - (\mathrm{E}X)^2] \\ &= a^2 VarX. \end{aligned} \tag{1.10}$$

Definicija 1.2.8. *Kovarijanca slučajnih varijabli X, Y se definira s*

$$Cov(X,Y) = \mathrm{E}[(X-\mathrm{E}X)(Y-\mathrm{E}Y)].$$

Za slučajne varijable X i Y kažemo da su nekorelirane ako vrijedi $Cov(X,Y) = 0$.

Analogno kao i kod izvoda (1.9) slijedi

$$Cov(X,Y) = \mathrm{E}(XY) - (\mathrm{E}X)(\mathrm{E}Y) \tag{1.11}$$

Nezavisnost slučajnih varijabli

Definicija 1.2.9. *Neka su* $X_1, X_2, \ldots, X_n$ *slučajne varijable. Slučajne varijable* $X_1, X_2, \ldots, X_n$ *su nezavisne ako za proizvoljne* $B_1, B_2, \ldots, B_n \in \mathcal{B}$ *vrijedi*

$$\mathrm{P}(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n) = \prod_{i=1}^{n} \mathrm{P}(X_i \in B_i).$$

Neka je I skup indeksa. Kažemo da je $\{X_i : i \in I\}$ *familija nezavisnih varijabli ako su slučajne varijable nezavisne za svaki podskup različitih indeksa iz I.*

Propozicija 1.2.10. *Neka je* $\{X_i : i \in I\}$ *familija nezavisnih slučajnih varijabli za neki skup indeksa I i neka su* $g_i : \mathbb{R} \to \mathbb{R}$, $i \in I$ *Borelove funkcije. Tada je* $\{g_i(X_i), i \in I\}$ *familija nezavisnih slučajnih varijabli.*

Teorem 1.2.11. *Neka su* $X_1, X_2, \ldots, X_n$ *nezavisne slučajne varijable takve da su sve nenegativne ili da sve imaju konačno očekivanje. Tada postoji* $\mathrm{E}\left(\prod_{i=1}^{n} X_i\right)$ *i vrijedi*

$$\mathrm{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathrm{E}X_i.$$

Teorem 1.2.12. *Neka su* $X_1, X_2, \ldots, X_n$ *nezavisne slučajne varijable čije varijance postoje. Tada postoji i varijanca od* $\sum_{i=1}^{n} X_i$ *i vrijedi*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} VarX_i.$$

Propozicija 1.2.13. *Neka su $X_1, \dots, X_n$ nezavisne slučajne varijable s gustoćom f_i, $1 \leq i \leq n$. Tada $(X_1, \dots, X_n)$ ima gustoću*

$$f(x_1, \dots, x_n) = f_1(x_1) \cdots f_n(x_n), \quad (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Definicija 1.2.14. *Karakteristična funkcija slučajne varijable X čija je funkcija distribucije F jest funkcija φ definirana sa*

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x), \quad t \in \mathbb{R}.$$

Za karakterističnu funkciju slučajne varijable X vrijediti će

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x) = \mathrm{E}e^{itX}, \quad t \in \mathbb{R}.$$

Analogno definiramo i karakterističnu funkciju slučajnog vektora $X = (X_1, \dots, X_n)$ i vrijedi

$$\varphi(t) = \mathrm{E}e^{i\langle t, X\rangle}, \quad t = (t_1, \dots, t_n) \in \mathbb{R}^n,$$

gdje je $\langle t, X\rangle = t_1X_1 + \cdots + t_nX_n$.

Teorem 1.2.15. *Slučajne varijable $X_1, \dots, X_n$ su nezavisne ako i samo ako vrijedi*

$$\varphi_{X_1,\dots,X_n}(t_1, \dots, t_n) = \prod_{k=1}^{n} \varphi_{X_k}(t_k), \quad t = (t_1, \dots, t_n) \in \mathbb{R}^n.$$

1.3 Primjeri slučajnih varijabli

Primjer 1.3.1. *Diskretna slučajna varijabla X čija je distribucija dana s*

$$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}, \quad 0 < p < 1$$

naziva se Bernoullijeva slučajna varijabla. Očito je prema (1.5) $\mathrm{E}X = 0\cdot(1-p) + 1\cdot p = p$, *a zbog $X^2 = X$ i $\mathrm{E}X^2 = \mathrm{E}X$, prema* (1.9) *slijedi da je $VarX = \mathrm{E}X^2 - (\mathrm{E}X)^2 = p - p^2 = p(1-p)$.*

Diskretna slučajna varijabla X čija je distribucija zadana pomoću

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n, \quad n \in \mathbb{N}$$

naziva se binomna slučajna varijabla i označava sa $X \sim B(n, p)$. *Budući da je binomna slučajna varijabla suma n nezavisnih Bernoulijevih slučajnih varijabli, zbog linearnosti očekivanja i Teorema 1.2.12 slijedi*

$$\mathrm{E}X = np$$

$$VarX = np(1-p).$$

Primjer 1.3.2. *Kažemo da neprekidna slučajna varijabla X ima normalnu distribuciju s parametrima* μ *i* σ^2 *i označavamo* $X \sim N(\mu, \sigma^2)$ *ako joj je funkcija gustoće f zadana sa*

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

Vrijedi da je $\mathrm{E}X = \mu$, $VarX = \sigma^2$.

Karakteristična funkcija normalne slučajne varijable $X \sim N(\mu, \sigma^2)$ *jest*

$$\varphi_X(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}.$$

Neka je $Y = aX + b$, $a \neq 0$, *Tada je* $Y \sim N(a\mu + b, a^2\sigma^2)$. *Nadalje, za* $X_1, \ldots, X_n$ *nezavisne slučajne varijable takve da su* $X_i \sim N(\mu_i, \sigma_i^2)$ *vrijedi da je*

$$\sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right).$$

Neka je $Z = (Z_1, Z_2, ..., Z_k)$, $k \in \mathbb{N}$ *slučajni vektor takav da su* $Z_i \sim N(0, 1)$ *nezavisne, jednako distribuirane slučajne varijable.* $X^T = (X_1, X_2, ..., X_n)$ *je normalni slučajni vektor ako postoje matrica* $A \in M_{nxk}$ *i* $\mu \in \mathbb{R}^n$ *takvi da je* $X = AZ + \mu$.

Lema 1.3.3. *Linearna transformacija normalnog slučajnog vektora je normalni slučajni vektor.*

Za dokaz tvrdnje vidjeti [3], 29. Limit Theorems in $\mathbb{R}^k$, Normal Distributions in $\mathbb{R}^k$.

Definicija 1.3.4. *Kažemo da dvije slučajne varijable X i Y imaju bivarijantnu normalnu razdiobu ako postoje dvije nezavisne normalno distribuirane slučajne varijable U, V i skalari a, b, c, d takvi da vrijedi:*

$$X = aU + bV,$$
$$Y = cU + dV.$$

Lema 1.3.5. *Neka su X i Y slučajne varijable koje imaju bivarijantnu normalnu distribuciju. Pretpostavimo da X i Y nisu korelirane. Tada su one nezavisne.*

Dokaz. Dovoljno je pokazati da tvrdnja vrijedi u slučaju kada X i Y imaju očekivanje 0. Tada vrijedi da je $\mathrm{E}U = \mathrm{E}V = 0$. Neka je $Z = t_1X + t_2Y = (t_1a + t_2c)U + (t_1b + t_2d)V$. Tada su zbog Propozicije 1.2.10 i Primjera 1.3.2 X, Y i Z normalne slučajne varijable. Označimo $X \sim N(0, \sigma_X^2)$, $Y \sim N(0, \sigma_Y^2)$. Budući da prema pretpostavki X i Y nisu korelirane, vrijedi da je $Cov(X, Y) = 0$, odnosno ekivalentno tome je $\mathrm{E}(XY) = \mathrm{E}X\mathrm{E}Y$ pa vrijedi

$$
\begin{aligned}
Var(Z) &= Var(t_1X + t_2Y)\\
&= \mathrm{E}(t_1X + t_2Y)^2 - [\mathrm{E}(t_1X + t_2Y)]^2\\
&= \mathrm{E}(t_1^2X^2 + 2t_1t_2XY + t_2^2Y^2) - [\mathrm{E}(t_1X) + \mathrm{E}(t_2Y)]^2\\
&= t_1^2\mathrm{E}X^2 + 2t_1t_2\mathrm{E}(XY) + t_2^2\mathrm{E}Y^2 - t_1^2(\mathrm{E}X)^2 - 2t_1t_2\mathrm{E}X\mathrm{E}Y - t_2^2(\mathrm{E}Y)^2\\
&= t_1^2\mathrm{E}X^2 - t_1^2(\mathrm{E}X)^2 + t_2^2\mathrm{E}Y^2 - t_2^2(\mathrm{E}Y)^2\\
&= t_1^2 VarX + t_2^2 VarY.
\end{aligned}
$$

Zbog linearnosti očekivanja je $Z \sim N(0, t_1^2\sigma_X^2 + t_2^2\sigma_Y^2)$. Prema Primjeru 1.3.2 i definiciji karakteristične funkcije slučajnog vektora vrijedi

$$
\varphi_{X,Y}(t_1, t_2) = Ee^{it_1X + it_2Y} = Ee^{itZ} = e^{-\frac{(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2)}{2}}.
$$

Pretpostavimo sada da su $\hat{X} \sim N(0, \sigma_X^2)$, $\hat{Y} \sim N(0, \sigma_Y^2)$ nezavisne slučajne varijable. Vrijedi Teorem 1.2.11 i može se pokazati da su one i nekorelirane, odnosno $Cov(\hat{X}, \hat{Y}) = \mathrm{E}(\hat{X}\hat{Y}) - \mathrm{E}\hat{X}\mathrm{E}\hat{Y} = \mathrm{E}\hat{X}\mathrm{E}\hat{Y} - \mathrm{E}\hat{X}\mathrm{E}\hat{Y} = 0$. Tada za karakterističnu funkciju $(\hat{X},\hat{Y})$ vrijedi isti argument kao i za karakterističnu funkciju (X,Y). Prema teoremu jedinstvenosti slijedi da ako su karakteristične funkcije $(\hat{X},\hat{Y})$ i (X,Y) jednake, tada su im jednake i distribucije. Budući da su $\hat{X}$ i $\hat{Y}$ nezavisne, jednako vrijedi i za X i Y. □

Primjer 1.3.6. *Kažemo da neprekidna slučajna varijabla X ima uniformnu distribuciju na segmentu [a,b], $a, b \in \mathbb{R}$, $a < b$ ako joj je funkcija gustoće f zadana s*

$$
f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b\\ 0, & x \notin [a, b]. \end{cases}
$$

Prema (1.7) *i* (1.8) *slijedi da je*

$$
\mathrm{E}X = \int_a^b \frac{x}{b-a}\,dx = \frac{a+b}{2} \qquad i \qquad \mathrm{E}X^2 = \int_a^b \frac{x^2}{b-a}\,dx = \frac{b^2 + ab + a^2}{3},
$$

pa vrijedi da je

$$
VarX = \frac{b^2 + ab + a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(a-b)^2}{12}.
$$

Propozicija 1.3.7. *Neka je X slučajna varijabla s neprekidnom funkcijom distribucije F. Tada je $F(X)$ slučajna varijabla s uniformnom distribucijom.*

Dokaz. Budući da je F neprekidna tada je i Borelova, pa je kompozicija slučajne varijable i Borelove funkcije ponovno slučajna varijabla. Dakle, $F(X)$ je slučajna varijabla. Preostaje nam pokazati da $F(X)$ ima uniformnu distribuciju. Neka je $0 \leq x \leq 1$. Prema Propoziciji 1.2.4 vrijedi:

$$F(X) \leq x \iff F^{\leftarrow}(F(X)) \leq F^{\leftarrow}(x) \iff X \leq F^{\leftarrow}(x).$$

Stoga slijedi

$$\mathrm{P}\left(F(X) \leq x\right) = \mathrm{P}\left(X \leq F^{\leftarrow}(x)\right) = F(F^{\leftarrow}(x)) = x.$$

Dakle, $F(X)$ ima uniformnu razdiobu, $F(X) \sim U(0,1)$. □

Pretpostavimo da su $X_1, \ldots, X_n$ nezavisne, jednako distribuirane slučajne varijable definirane na istom prostoru Ω sa zajedničkom funkcijom distribucije F. Ako pretpostavimo da je F neprekidna funkcija distribucije, tada se jednake realizacije među tim slučajnim varijablama događaju s vjerojatnosti 0 pa ih možemo zanemariti. Definirajmo sada slučajne varijable $X_{(1)}, \ldots, X_{(n)}$ na domeni Ω, za $\omega \in \Omega$:

$$\begin{aligned}
X_{(1)}(\omega) &= \min\{X_1(\omega), \ldots, X_n(\omega)\}\\
X_{(2)}(\omega) &= \text{ drugi najmanji od } \{X_1(\omega), \ldots, X_n(\omega)\}\\
&\vdots\\
X_{(n-1)}(\omega) &= \text{ drugi najveći od } \{X_1(\omega), \ldots, X_n(\omega)\}\\
X_{(1)}(\omega) &= \max\{X_1(\omega), \ldots, X_n(\omega)\},
\end{aligned}$$

pa imamo $X_{(1)} < \ldots < X_{(n)}$. Slučajne varijable $X_{(1)}, \ldots, X_{(n)}$ nazivamo uređene statistike.

Lema 1.3.8. *Neka su $U_1, \ldots, U_n$ nezavisne, jednako distribuirane, uniformne slučajne varijable na [0,t] i neka vrijedi $U_{(1)} < U_{(2)} < \ldots < U_{(n)}$. Tada je zajednička gustoća*

$$f_{U_{(1)},\ldots,U_{(n)}}(u_1, \ldots, u_n) = \begin{cases} \frac{n!}{t^n}, & 0 < u_1 < \ldots < u_n < t \\ 0, & \textit{inače} \end{cases}$$

Dokaz. Neka je Π skup svih permutacija skupa $\{1, \ldots, n\}$. Tada za $\pi \in \Pi$, $(U_{(1)}, \ldots, U_{(n)}) = (U_{\pi(1)}, \ldots, U_{\pi(n)})$ na skupu $[U_{\pi(1)} < \ldots < U_{\pi(n)}]$. Stoga, za bilo koju omeđenu funkciju g imamo

$$\mathrm{E}g(U_{(1)}, \ldots, U_{(n)}) = \sum_{\pi \in \Pi} \mathrm{E}g(U_{\pi(1)}, \ldots, U_{\pi(n)}) \mathbb{1}_{[U_{\pi(1)} < \ldots < U_{\pi(n)}]}.$$

Budući da vrijedi $f_{U_{\pi(1)},\dots,U_{\pi(n)}}(u_1,\dots,u_n) = f_{U_1,\dots,U_n}(u_1,\dots,u_n)$ i zbog nezavisnosti slučajnih varijabli $U_1,\dots,U_n$ slijedi

$$f_{U_{\pi(1)},\dots,U_{\pi(n)}}(u_1,\dots,u_n) = \begin{cases} t^{-n}, & (u_1,\dots,u_n) \in [0,t]^n \\ 0, & \text{inače.} \end{cases}$$

Stoga imamo

$$\begin{aligned} \mathrm{E}g(U_{(1)},\dots,U_{(n)}) &= \sum_{\pi\in\Pi} \int\limits_{[0<u_1<\ldots<u_n<t]} g(u_1,\dots,u_n)t^{-n}\,du_1\dots du_n \\ &= \int\limits_{[0,t]^n} g(u_1,\dots,u_n)\left(\frac{n!}{t^n}\mathbb{1}_{[u_1<\ldots<u_n]}(u_1,\dots,u_n)\right)du_1\dots du_n \end{aligned}$$

iz čega slijedi tvrdnja leme. □

Primjer 1.3.9. *Kažemo da slučajna varijabla X ima gama razdiobu s parametrima $\alpha > 0$ i $\lambda > 0$ ako je strogo pozitivna i ako je njena funkcija gustoće*

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

gdje je $\Gamma(\alpha) = \int\limits_0^\infty t^{\alpha-1}e^{-t}\,dt$ Γ*-funkcija.*

Slučajnu varijablu s gama razdiobom označavamo $X \sim \Gamma\left(\alpha, \frac{1}{\lambda}\right)$*. Prema* (1.7) *i* (1.8) *integriranjem dobijemo* $\mathrm{E}X = \frac{\alpha}{\lambda}$, $VarX = \frac{\alpha}{\lambda^2}$*. Karakteristična funkcija slučajne varijable* $X \sim \Gamma\left(\alpha, \frac{1}{\lambda}\right)$ *jest*

$$\varphi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^\alpha, \quad t \in \mathbb{R}.$$

U slučaju kada je $\alpha = 1$ kažemo da X ima eksponencijalnu razdiobu. Tada pišemo $X \sim Exp(\lambda)$*. Slijedi da su funkcija gustoće i funkcija distribucije ekponencijalne slučajne varijable*

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

Matematičko očekivanje i varijanca eksponencijalne razdiobe su tada $\mathrm{E}X = \frac{1}{\lambda}$, $VarX = \frac{1}{\lambda^2}$. *Karakteristična funkcija slučajne varijable* $X \sim Exp(\lambda)$ *jest*

$$\varphi_X(t) = \frac{\lambda}{\lambda - it}, \quad t \in \mathbb{R}.$$

Lema 1.3.10. *Neka su* $X_1, \ldots, X_n$ *nezavisne slučajne varijable takve da je* $X_i \sim Exp(\lambda)$, *za* $1 \leq i \leq n$. *Tada vrijedi* $X_1 + \cdots + X_n \sim \Gamma\left(n, \frac{1}{\lambda}\right)$.

Dokaz. Iz karakterističnih funkcija slučajnih varijabli $X_i \sim Exp(\lambda)$ i $X_1 + \cdots + X_n \sim \Gamma\left(n, \frac{1}{\lambda}\right)$ danih u Primjeru 1.3.9 tvrdnja slijedi direktno prema Teoremu 1.2.15. □

Propozicija 1.3.11. *Pretpostavimo da su* $\{E_n, n \geq 1\}$ *nezavisne, jednako distribuirane slučajne varijable,* $E_n \sim Exp(1)$. *Definirajmo* $\Gamma_n = E_1 + E_2 + \cdots + E_n$. *Tada je zajednička gustoća od* $\Gamma_1, \ldots, \Gamma_n$ *uz uvjet da je* $\Gamma_{n+1} = t$ *jednaka gustoći uređenih statistika od n nezavisnih, jednako distribuiranih slučajnih varijabli koje su uniformno distribuirane na* $[0, t]$.

Dokaz. Za početak pogledajmo zajedničku gustoću slučajnih varijabli $E_1, \ldots, E_{n+1}$. Budući da su to nezavisne, eksponencijalno distribuirane slučajne varijable, iz Primjera 1.3.9 i zbog Propozicije 1.2.13 zajednička gustoća je jednaka

$$f_{E_1,\ldots,E_{n+1}}(x_1, \ldots, x_{n+1}) = \prod_{i=1}^{n+1} e^{-x_i} = e^{-\sum\limits_{i=1}^{n+1} x_i}, \tag{1.12}$$

za $x_i > 0$, $1 \leq i \leq n+1$. Iz gustoće dane sa (1.12) možemo dobiti gustoću od $(\Gamma_1, \ldots, \Gamma_{n+1})$ zamjenom varijabli. Za $i = 1, \ldots, n+1$ definirajmo $s_i = \sum\limits_{j=1}^{i} x_j$. Vidimo stoga da je inverzna transformacija $x_i = s_i - s_{i-1}$, za $1 \leq i \leq n+1$ gdje je $s_0 = 0$. Pripadni Jakobijan inverzne transformacije jest

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{vmatrix} = 1.$$

Stoga vrijedi da je

$$f_{\Gamma_1,\ldots,\Gamma_{n+1}}(s_1, \ldots, s_{n+1}) = e^{-s_{n+1}}, \tag{1.13}$$

gdje je $0 \leq s_1 < s_1 < \ldots < s_{n+1}$. Kako prema Lemi 1.3.10 vrijedi da Γ_{n+1} ima gama razdiobu, iz (1.4) i (1.13) slijedi da je uvjetna zajednička gustoća jednaka

$$f_{\Gamma_1,\ldots,\Gamma_n \mid \Gamma_{n+1}=t}(s_1, \ldots, s_n) = \frac{f_{\Gamma_1,\ldots,\Gamma_{n+1}}(s_1, \ldots, s_n, t)}{f_{\Gamma_{n+1}}(t)} = \frac{e^{-t}}{\frac{e^{-t}t^n}{n!}} = \frac{n!}{t^n}, \tag{1.14}$$

za $0 < s_1 < \ldots < s_n < t$ što je prema Lemi 1.3.8 upravo gustoća uređenih statistika od n nezavisnih, jednako distribuiranih slučajnih varijabli koje su uniformno distribuirane na intervalu [0,t]. □

1.4 Empirijska funkcija distribucije

Definicija 1.4.1. *Neka je $X_1, X_2, \ldots, X_n$ slučajan uzorak duljine n iz distribucije F. Empirijska funkcija distribucije je slučajna funkcija $\hat{F}_n(\cdot)(x) : \mathbb{R} \to \mathbb{R}, \omega \in \Omega$ definirana s*

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]}, x \in \mathbb{R} \tag{1.15}$$

Prema tome, za svaki fiksni $x \in \mathbb{R}$, $\hat{F}_n(x)$ jest proporcija uzorka koji je manji ili jednak x. Prije nego nastavimo sa svojstvima empirijske funkcije distribucije koja će nam biti od koristi, pogledajmo nekolicinu poznatih teorema iz opće teorije vjerojatnosti koji će biti potrebni u dokazivanju svojstava empirijske funkcije distribucije.

Teorem 1.4.2. ***(Jaki zakon velikih brojeva)*** *Pretpostavimo da su $X_1, X_2, \ldots, X_n$ u parovima nezavisne, jednako distribuirane slučajne varijable takve da je $\mathrm{E}|X_i| < \infty$. Neka je $\mathrm{E}X_i = \mu$ i $S_n = X_1 + \cdots + X_n$. Tada $\frac{S_n}{n} \xrightarrow{g.s.} \mu$ kada $n \to \infty$.*

Dokaz možete pronaći u [5], 2.4 Strong Law of Large Numbers, dokaz Teorema 2.4.1.

Teorem 1.4.3. ***(Centralni granični teorem)*** *Neka su $X_1, X_2, \ldots$ nezavisne, jednako distribuirane slučajne varijable takve da je $\mathrm{E}X_i = \mu$ i $VarX_i = \sigma^2 \in (0, \infty)$. Ako je $S_n = X_1 + X_2 + \cdots + X_n$ tada $\frac{S_n - n\mu}{\sigma\sqrt{n}} \implies \chi$ gdje χ ima standardnu normalnu razdiobu.*

Dokaz se nalazi primjerice u [5], 3.4 Central Limit Theorems, dokaz Teorema 3.4.1.

Teorem 1.4.4. ***(Glivenko-Cantelli)*** *Neka su $X_1, X_2, \ldots$ nezavisne, jednako distribuirane slučajne varijable sa zajedničkom funkcijom distribucije F i neka je $\hat{F}_n$ pripadna empirijska funkcija distribucije. Tada $\hat{F}_n$ konvergira uniformno prema F kada $n \to \infty$, odnosno vrijedi*

$$\sup_x |\hat{F}_n(x) - F(x)| \xrightarrow{g.s.} 0.$$

Dokaz. Uzmimo fiksni x i neka je $Y_n = \mathbb{1}_{[X_n \leq x]}$. Kao što ćemo vidjeti detaljno u potpoglavlju Svojstva empirijske funkcije distribucije, Y_n su nezavisne, jednako distribuirane slučajne varijable i vrijedi $\mathrm{E}Y_n = \mathrm{P}(X_n \leq x) = F(x)$. Jaki zakon velikih brojeva, Teorem 1.4.2 povlači

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]} = \frac{1}{n} \sum_{i=1}^{n} Y_i \xrightarrow{g.s.} F(x).$$

Ako je $\hat{F}_n$ niz rastućih funkcija koje konvergiraju prema ograničenoj i neprekidnoj funkciji F, tada vrijedi tvrdnja teorema. Međutim, funkcija distribucije F može imati skokove. Ponovno, fiksirajmo x i neka je $Z_n = \mathbb{1}_{[X_n<x]}$. Budući da su Z_n nezavisne, jednako distribuirane slučajne varijable takve da je $\mathrm{E}Z_n = \mathrm{P}(X_n < x) = F(x-) = \lim\limits_{y\to x-} F(y)$, Teorem 1.4.2 implicira

$$\hat{F}_n(x-) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[X_i<x]} = \frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow{g.s.} F(x-).$$

Neka je za $1 \le j \le k-1$, $x_{j,k} = \inf\{y : F(y) \ge \frac{j}{k}\}$. Zbog konvergencije po točkama $\hat{F}_n(x)$ i $\hat{F}_n(x-)$, možemo uzeti N takav da za $n \ge N$ vrijedi

$$|\hat{F}_n(x_{j,k}) - F(x_{j,k})| < k^{-1} \qquad \text{i} \qquad |\hat{F}_n(x_{j,k}-) - F(x_{j,k}-)| < k^{-1},$$

za $1 \le j \le k-1$. Ako dodamo $x_{0,k} = -\infty$ i $x_{k,k} = \infty$, tada zadnje dvije nejednakosti vrijede i u slučaju $j = 0$ i $j = k$. Neka je $x \in (x_{j-1,k}, x_{j,k})$, $1 \le j \le k$ i $n \ge N$. Tada zbog monotonosti od $\hat{F}_n$, F i zbog $F(x_{j,k}-) - F(x_{j-1,k}) \le k^{-1}$ imamo

$$\hat{F}_n(x) \le \hat{F}_n(x_{j,k}-) \le F(x_{j,k}-) + k^{-1} \le F(x_{j-1,k}) + 2k^{-1} \le F(x) + 2k^{-1},$$

$$\hat{F}_n(x) \ge \hat{F}_n(x_{j-1,k}) \ge F(x_{j-1,k}) - k^{-1} \ge F(x_{j,k}-) - 2k^{-1} \ge F(x) - 2k^{-1}.$$

Vrijedi dakle $\sup\limits_x |\hat{F}_n(x) - F(x)| \le 2k^{-1}$, čime je tvrdnja dokazana. □

Svojstva empirijske funkcije distribucije

Neka su $X_1, X_2, \ldots, X_n$ nezavisne, jednako distribuirane slučajne varijable na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathrm{P})$. Označimo im funkciju distribucije s F. Indikatorske funkcije tih slučajnih varijabli jednake su kao i u definiciji empirijske funkcije distribucije

$$\mathbb{1}_{[X_i\le x]} = \begin{cases} 1, & X_i \le x \\ 0, & X_i > x. \end{cases}$$

Budući da je prema definiciji funkcije distribucije $F(x) = \mathrm{P}(X_i \le x)$, tada slijedi

$$\mathbb{1}_{[X_i\le x]} \sim \begin{pmatrix} 0 & 1 \\ 1 - F(x) & F(x) \end{pmatrix}.$$

Drugim riječima, slučajne varijable $\mathbb{1}_{[X_i\le x]}$ imaju Bernoulijevu razdiobu. Nadalje, zbog nezavisnosti slučajnih varijabli $X_1, X_2, \ldots, X_n$ i Propozicije 1.2.10 slijedi da su i nezavisne. Dakle, $\mathbb{1}_{[X_1\le x]}, \mathbb{1}_{[X_2\le x]}, \ldots, \mathbb{1}_{[X_n\le x]}$ su nezavisne Bernoulijeve slučajne varijable, pa za njihovu sumu vrijedi

$$\sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]} = n\hat{F}_n \sim B(n, F(x)).$$

Očekivanje i varijanca binomne varijable su nam poznati iz Primjera 1.3.1, pa imamo $\mathrm{E}(n\hat{F}_n(x)) = nF(x)$, $Var(n\hat{F}_n(x)) = nF(x)(1-F(x))$. Iz linearnosti matematičkog očekivanja i svojstva varijance (1.10) slijedi zaključak

$$\mathrm{E}(\hat{F}_n(x)) = F(x)$$

$$Var\hat{F}_n(x) = \frac{1}{n}F(x)(1 - F(x)).$$

Prema jakom zakonu velikih brojeva, Teoremu 1.4.2, slijedi da za svaki fiksni $x \in \mathbb{R}$

$$\lim_{n\to\infty} \hat{F}_n(x) = F(x) \text{ g.s.}$$

Uniformna konvergencija vrijedi prema Glivenko-Cantellijevom Teoremu 1.4.4.

Za fiksni $x \in \mathbb{R}$, centralni granični teorem, Teorem 1.4.3 primijenjen na sumu nezavisnih, jednako distribuiranih slučajnih varijabli $\mathbb{1}_{[X \leq x]}$ daje nam

$$\frac{\sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]} - nF(x)}{\sqrt{n}\sqrt{F(x)(1 - F(x))}} = \frac{n\left(\hat{F}_n(x) - F(x)\right)}{\sqrt{n}\sqrt{F(x)(1 - F(x))}} \implies N(0, 1)$$

$$\sqrt{n}(\hat{F}_n(x) - F(x)) \implies N(0, F(x)(1 - F(x))) \tag{1.16}$$

Ovaj rezultat može se proširiti u više dimenzija. Promotrimo stoga centralni granični teorem o slučajnim vektorima čiji dokaz se nalazi u [3], 29. Limit Theorems in $\mathbb{R}^k$, dokaz Teorema 29.5.

Teorem 1.4.5. *Neka su $X_n = (X_{n1}, X_{n2}, \ldots, X_{nk})$ nezavisni, jednako distribuirani slučajni vektori. Pretpostavimo da je $\mathrm{E}[X_{nu}^2] < \infty$. Neka je $c = (c_1, c_2, \ldots, c_k)$ vektor očekivanja gdje je $c_u = \mathrm{E}[X_{nu}]$ i $\Sigma = [\sigma_{ul}]$ kovarijacijska matrica gdje je $\sigma_{ul} = \mathrm{E}\left[(X_{nu} - c_u)(X_{nl} - c_l)\right]$. Neka je $S_n = X_1 + X_2 + \cdots + X_n$. Tada vektor $\frac{(S_n - nc)}{\sqrt{n}}$ konvergira po distribuciji k standardnoj normalnoj distribuciji s kovarijacijskom matricom Σ.*

Primijetimo prvo da za $x_1 < x_2$ vrijedi

$$\begin{aligned} Cov(\mathbb{1}_{[X_i \leq x_1]}, \mathbb{1}_{[X_i \leq x_2]}) &= E\mathbb{1}_{[X_i \leq x_1]}\mathbb{1}_{[X_i \leq x_2]} - E\mathbb{1}_{[X_i \leq x_1]}E\mathbb{1}_{[X_i \leq x_2]} \\ &= F(x_1 \wedge x_2) - F(x_1)F(x_2). \end{aligned}$$

Stoga prema Teoremu 1.4.5 slijedi

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\begin{pmatrix}\mathbb{1}_{[X_i\leq x_1]}\\ \mathbb{1}_{[X_i\leq x_2]}\end{pmatrix} \;-\; \begin{pmatrix}F(x_1)\\ F(x_2)\end{pmatrix}\right) = \sqrt{n}\left(\begin{pmatrix}\hat{F}_n(x_1)\\ \hat{F}_n(x_2)\end{pmatrix} \;-\; \begin{pmatrix}F(x_1)\\ F(x_2)\end{pmatrix}\right)$$

$$\Longrightarrow\; N\left(0 \quad , \quad \begin{pmatrix}F(x_1)(1-F(x_1)) & F(x_1)(1-F(x_2))\\ F(x_1)(1-F(x_2)) & F(x_2)(1-F(x_2)))\end{pmatrix}\right) \text{ u } \mathbb{R}^2.$$

Poglavlje 2

Brownovo gibanje

Kao što smo vidjeli u Uvodu, Brownovo gibanje jest jedan od značajnijih slučajnih procesa koji je utjecao na razvoj mnogih studija i koji je našao svoju primjenu u nekoliko sfera. Nakon definicije Brownovog gibanja glavna zadaća bit će pokazati da je Brownovo gibanje Gaussovski proces i dokazati da Brownovo gibanje zadovoljava jako Markovljevo svojstvo. Kao rezultat jakog Markovljevog svojstva vidjet ćemo princip refleksije, a potom dokazati i Lévyjev zakon trojke. Uvodimo pojam principa invarijantnosti poznatijeg kao Donskerov teorem. Brownov most definiran je u zadnjem potpoglavlju. Prikazana su svojstva Brownovog mosta, među kojima je najznačajnija Propozicija 2.5.3 koja će igrati ključnu ulogu u zadnjem poglavlju rada i koja nam govori da je Brownov most aproksimativno jednak Brownovom gibanju uvjetovanom da ima vrijednost 0 u vremenu 1. Na kraju poglavlja vidjet ćemo izraz koji upućuje na vezu Brownovog mosta i empirijske funkcije distrbucije.

Neka je $(\Omega, \mathcal{F}, \mathrm{P})$ vjerojatnosni prostor. Slučajna varijabla na tom vjerojatnosnom prostoru neovisna je o vremenu, no mnogi procesi koji se događaju u vremenu će ipak zahtijevati vremensku komponentu. Dolazimo do pojma slučajnog ili stohastičkog procesa gdje za svako vrijeme $t \geq 0$ imamo slučajnu varijablu koju ćemo označavati s X_t ili $X(t)$.

Definicija 2.0.6. *Familija slučajnih varijabli* $X = \{X_t : t \geq 0\}$ *na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathrm{P})$ *naziva se slučajni ili stohastički proces. Primjetimo da slučajni proces možemo shvatiti kao funkciju dviju varijabli*

$$X : \mathbb{R}_+ \times \Omega \to \mathbb{R}.$$

Ako fiksiramo neki $\omega \in \Omega$ preslikavanje $t \mapsto X(t, \omega)$ nazivamo trajektorijom i to su realizacije procesa X tijekom vremena. Analogno, ako fiksiramo neki $t \geq 0$ preslikavanje $\omega \mapsto X(t, \omega)$ daje nam realizacije procesa u trenutku t.

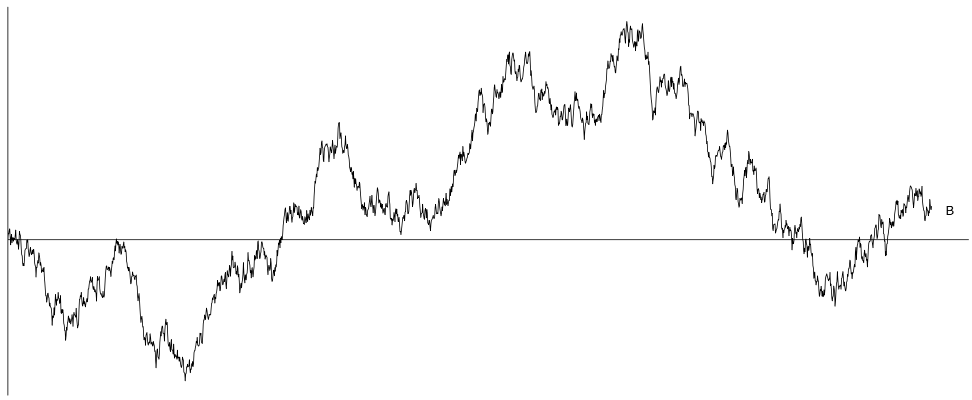
Definicija 2.0.7. *Neka je* $(\Omega, \mathcal{F}, \mathrm{P})$ *vjerojatnosni prostor. Slučajni proces* $B = \{B_t : t \geq 0\}$ *jest Brownovo gibanje ako vrijedi:*

(B1) $B_0 = 0$ P-*g.s.*

(B2) Za sve $m \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_m$ *su prirasti* $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_m} - B_{t_{m-1}}$ *nezavisne slučajne varijable.*

(B3) Za sve $0 \leq s < t$ *vrijedi* $B_t - B_s \sim N(0, t-s)$.

(B4) Za P-*g.s* $\omega \in \Omega$ *su putevi* $t \mapsto B_t(\omega) \in \mathbb{R}$, $t \in [0, \infty\rangle$ *neprekidne funkcije.*



Slika 2.1: Brownovo gibanje

Navedimo neka svojstva Brownovog gibanja koja slijede direktno iz definiciije. Budući da za $0 \leq s < t$ vrijedi $B_t - B_s \sim N(0, t-s)$, vidimo da za svaki t vrijedi $B_t \sim N(0, t)$, pa je

$$\mathrm{E}B_t = 0. \tag{2.1}$$

Nadalje slijedi da je $VarB_t = t$, pa je zbog (2.1)

$$t = VarB_t = \mathrm{E}B_t^2 - (\mathrm{E}B_t)^2 = \mathrm{E}B_t^2. \tag{2.2}$$

Osvrnimo se još i na slučajni vektor $(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_m} - B_{t_{m-1}})$, gdje je $0 = t_0 < t_1 <, \ldots, < t_m$. Vidjeli smo iz definicije da za svaki $0 \leq s < t$ vrijedi $B_t - B_s \sim N(0, t-s)$, odnosno vrijedi $B_t - B_s \sim \sqrt{t-s}N(0, 1)$. Budući da su prirasti Brownovog gibanja nezavisni, prema Primjeru 1.3.2 vidimo da je vektor prirasta Brownovog gibanja zapravo normalni slučajni vektor.

2.1 Gaussovski proces

Definicija 2.1.1. *Kovarijacijska funkcija slučajnog procesa $X = \{X_t : t \geq 0\}$ je definirana na sljedeći način*

$$Cov(X_s, X_t) = \mathrm{E}\left[(X_s - EX_s)(X_t - EX_t)\right].$$

Iz definicije kovarijacijske funkcije očito je da vrijedi

$$Cov(X_s, X_t) = \mathrm{E}\,(X_s X_t) - EX_s EX_t.$$

Definicija 2.1.2. *Slučajni proces $X = \{X_t : t \geq 0\}$ je Gaussovski ako je za sve $m \in \mathbb{N}$ i $0 < t_1 < t_2 < \ldots < t_m$ slučajni vektor $(X_{t_1}, X_{t_2}, \ldots, X_{t_m})$ normalni slučajni vektor.*

Neka je $X = (X_1, \ldots, X_n)$ Gaussovski proces, $m = (m_1, \ldots, m_n) \in \mathbb{R}^n$, gdje je $m_j = \mathrm{E}X_j$ i $\Sigma = (\sigma_{jk})_{j,k=1,\ldots,n} \in \mathbb{R}^{n\times n}$ kovarijacijska matrica slučajnog procesa, gdje je $\sigma_{jk} = Cov(X_j, X_k)$. Karakteristična funkcija Gaussovskog procesa jest

$$\mathrm{E}e^{i\langle t,X\rangle} = e^{i\langle t,m\rangle - \frac{1}{2}\langle t,\sigma t\rangle}. \tag{2.3}$$

Pretpostavimo sada da je $G(x)$, $x \in \mathbb{R}$ Gaussov proces sa očekivanjem 0 i funkcijom kovarijacije

$$Cov(G(x_1), G(x_2)) = F(x_1)(1 - F(x_2)), x_1 < x_2. \tag{2.4}$$

Tada zaključak (1.16) možemo zapisati kao

$$\sqrt{n}(\hat{F}_n(x) - F(x)) \implies G(x) \tag{2.5}$$

u smislu konvergencije konačno dimenzionalnih distribucija.

Propozicija 2.1.3. *Neka je $X = \{X_t : t \geq 0\}$ Gaussovski proces s vektorom očekivanja nula i matricom kovarijacije $C = (\min\{t_j, t_k\})_{j,k=1,\ldots,n}$. Ako $X = \{X_t : t \geq 0\}$ ima neprekidne trajektorije, tada je $X = \{X_t : t \geq 0\}$ Brownovo gibanje.*

Dokaz. Primijetimo prvo da je $X_0 \sim N(0, 0)$, pa će prema tome svojstva definicije Brownovog gibanja (B1) i (B4) kao pretpostavka vrijediti direktno.

Neka je $\Delta = (X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}})^T$, $\Gamma = (X_{t_1}, \ldots, X_{t_n})^T$. Tada vrijedi

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} X_{t_1} - X_{t_0} \\ X_{t_2} - X_{t_1} \\ \vdots \\ X_{t_n} - X_{t_{n-1}} \end{pmatrix} = \begin{pmatrix} X_{t_1} \\ X_{t_2} \\ \vdots \\ X_{t_n} \end{pmatrix},$$

odnosno ako donjetrokutastu matricu označimo slovom M, slijedi $M\Delta = \Gamma$. Nadalje vrijedi i $\Delta = M^{-1}\Gamma$, gdje je $M^{-1} = \begin{pmatrix} 1 & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}$. Budući da je Γ Gaussovski proces i zbog (2.3) slijedi

$$\mathrm{E}e^{i\langle\epsilon,\Delta\rangle} = \mathrm{E}e^{i\langle\epsilon,M^{-1}\Gamma\rangle} = \mathrm{E}e^{i\langle(M^{-1})^T\epsilon,\Gamma\rangle} = e^{-\frac{1}{2}\langle(M^{-1})^T\epsilon,C(M^{-1})^T\epsilon\rangle} = e^{-\frac{1}{2}\langle\epsilon,M^{-1}C(M^{-1})^T\epsilon\rangle}.$$

Jednostavnim matričnim množenjem dobijemo da je

$$M^{-1}C(M^{-1})^T = \begin{pmatrix} t_1 - t_0 & & & \\ & t_2 - t_1 & & \\ & & \ddots & \\ & & & t_n - t_{n-1} \end{pmatrix}.$$

Stoga je Γ Gaussovski slučajni vektor s nekoreliranim, stoga nezavisnim komponentama za koje vrijedi da su normalno distribuirane $\sim N(0, t_j - t_{j-1})$. To dokazuje (B2) i (B3). □

Propozicija 2.1.4. *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje Tada je B Gaussovski proces s vektorom očekivanja nula i kovarijacijskom funkcijom*

$$Cov(B_s, B_t) = \min\{s, t\}. \tag{2.6}$$

Dokaz. Neka je B Brownovo gibanje. Uzmimo proizvoljne $0 < t_1 < t_2 < \ldots < t_k$. Kao što smo vidjeli, slučajni vektor $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}})$ je normalni slučajni vektor, a slučajni vektor $(B_{t_1}, B_{t_2}, \ldots, B_{t_k})$ možemo prikazati na sljedeći način

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \begin{pmatrix} B_{t_1} \\ B_{t_2} - B_{t_1} \\ B_{t_3} - B_{t_2} \\ \vdots \\ B_{t_k} - B_{t_{k-1}} \end{pmatrix} = \begin{pmatrix} B_{t_1} \\ B_{t_2} \\ B_{t_3} \\ \vdots \\ B_{t_k} \end{pmatrix}$$

pa je prema Lemi 1.3.3 i $(B_{t_1}, B_{t_2}, \ldots, B_{t_k})$ normalni slučajni vektor. Dakle, prema Definiciji 2.1.2, B je Gausovski proces. Preostaje nam pokazati da vrijedi tvdnja o kovarijacijskoj funkciji. Neka je $0 < s \leq t$. Prema (1.11) i uz (2.2) slijedi

$$\begin{aligned}
Cov(B_s, B_t) &= \mathrm{E}(B_s B_t) - \mathrm{E}B_s \mathrm{E}B_t \\
&= \mathrm{E}[B_s(B_t - B_s)] + \mathrm{E}(B_s)^2 - 0 \\
&= \mathrm{E}B_s \mathrm{E}(B_t - B_s) + s \\
&= 0 + s \\
&= \min\{s, t\},
\end{aligned}$$

gdje druga jednakost vrijedi zbog linearnosti očekivanja, a treća zbog nezavisnosti prirasta Brownovog gibanja i Teorema 1.2.11. □

Lema 2.1.5. *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje. Tada je slučajni proces $\tilde{B} = \{\tilde{B}_t : t \geq 0\}$ gdje je $\tilde{B}_t = -B_t$ Brownovo gibanje.*

Dokaz. B je Brownovo gibanje pa je prema Propoziciji 2.1.4 $(B_{t_1}, B_{t_2}, \dots, B_{t_k})$ normalni slučajni vektor i $Cov(B_s, B_t) = s \wedge t$. Neka je $0 < t_1 < \dots < t_m$. Tada zbog Leme 1.3.3 i

$$(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_m})^T = \begin{pmatrix} -1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & -1 \end{pmatrix} \begin{pmatrix} B_{t_1} \\ \vdots \\ B_{t_m} \end{pmatrix}$$

slijedi da je $(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_m})$ normalni slučajni vektor, odnosno $\tilde{B}$ Gaussovski proces. Nadalje, $\mathrm{E}\tilde{B}_t = \mathrm{E}(-B_t) = -\mathrm{E}B_t = 0$ i zbog linearnosti očekivanja vrijedi

$$\begin{aligned}
Cov(\tilde{B}_s, \tilde{B}_t) &= Cov(-B_s, -B_t) \\
&= \mathrm{E}[(-B_s - \mathrm{E}(-B_s))(-B_t - \mathrm{E}(-B_t))] \\
&= \mathrm{E}[(B_s - \mathrm{E}B_s)(B_t - \mathrm{E}B_t)] \\
&= Cov(B_s, B_t) \\
&= s \wedge t.
\end{aligned}$$

Iz Propozicije 2.1.3 slijedi da je $\tilde{B}$ Brownovo gibanje. □

2.2 Jako Markovljevo svojstvo

Definicija 2.2.1. *Kažemo da su dva stohastička procesa $\{X_t : t \geq 0\}$ i $\{Y_t : t \geq 0\}$ nezavisna ako su za proizvoljne $t_1, t_2, \dots, t_n \geq 0$ i $s_1, s_2, \dots, s_m \geq 0$ slučajni vektori $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ i $(Y_{s_1}, Y_{s_2}, \dots, Y_{s_m})$ nezavisni.*

Definicija 2.2.2. *Filtracija na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, \mathrm{P})$ *je familija* $\{\mathcal{F}_t : t \geq 0\}$ σ*-algebri takva da vrijedi* $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ *za svaki* $s < t$*. Stohastički proces* $\{X_t : t \geq 0\}$ *definiran na vjerojatnosnom prostoru s filtracijom naziva se adaptiran ako je* X_t $\mathcal{F}_t$ *izmjeriva za proizvoljan* $t \geq 0$*. Za slučajnu varijablu* X *kažemo da je* $\mathcal{F}_t$ *izmjeriva ako za svaki* $B \in \mathcal{B}$ *vrijedi* $X^{-1}(B) \in \mathcal{F}_t$*.*

Propozicija 2.2.3. *Neka je* $B = \{B_t : t \geq 0\}$ *Brownovo gibanje. Tada je za* $s > 0$ *proces* $W = \{W_t : t \geq 0\}$*,* $W_t := B_{t+s} - B_s$ *za* $t \geq 0$ *Brownovo gibanje nezavisno od procesa* $\{B_t : 0 \leq t \leq s\}$*.*

Dokaz. Provjerimo da definirani proces zadovoljava Definiciju 2.0.7. $W_0 = B_{0+s} - B_s = 0$, dakle zadovoljen je uvjet (B1). Za $0 < t_1 < t_2 < \dots t_m$ vidimo da su $W_{t_1} = B_{t_1+s} - B_s, W_{t_2} - W_{t_1} = B_{t_2+s} - B_{t_1+s}, \dots W_{t_m} - W_{t_{m-1}} = B_{t_m+s} - B_{t_{m-1}+s}$ nezavisne slučajne varijable pa je zadovoljeno i svojstvo (B2). Neka su $0 \leq k < t$. Tada je $W_t - W_k = B_{t+s} - B_{k+s} \sim N(0, t+s-k-s) \sim N(0, t-k)$, čime je zadovoljeno svojstvo (B3). Preostaje nam pokazati da su trajektorije slučajnog procesa W neprekidne, odnosno da je zadovoljeno svojstvo (B4). Vidimo da vrijedi $t \mapsto W_t = B_{t+s} - B_s$, što je g.s neprekidno zbog neprekidnosti Brownovog gibanja B. Tvrdnja o nezavisnosti slijedit će direktno iz nezavisnosti prirasta Brownovog gibanja. □

Ako želimo znati kada Brownovo gibanje napušta ili ulazi u neki skup prvi put, kada dostiže svoj maksimum ili kada se vraća u 0, moramo koristiti slučajna vremena. Slučajna varijabla $\tau : \Omega \to [0, \infty]$ je vrijeme zaustavljanja u odnosu na filtraciju $\{\mathcal{F}_t : t \geq 0\}$ ako je $\{\tau \leq t\} \in \mathcal{F}_t$, za svaki $t \geq 0$. Tipičan primjer vremena zaustavljanja je vrijeme ulaska slučajnog procesa $X = \{X_t : t \geq 0\}$ u skup A. Prirodna filtracija $\mathcal{F}_t^X := \sigma(X_s : s \leq t)$ stohastičkog procesa $X = \{X_t : t \geq 0\}$ je relativno mala i stoga za mnoga zanimljiva vremena zaustavljanja moramo promatrati nešto veću filtraciju

$$\mathcal{F}_{t+}^X := \bigcap_{u>t} \mathcal{F}_u^X = \bigcap_{n \geq 1} \mathcal{F}_{t+\frac{1}{n}}^X.$$

Neka je τ vrijeme zaustavljanja u odnosu na filtraciju $\{\mathcal{F}_t : t \geq 0\}$. Definirajmo familije skupova

$$\mathcal{F}_\tau := \left\{ A \in \mathcal{F}_\infty := \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t \right) : A \cap \{\tau \leq t\} \in \mathcal{F}_t \, \forall t \geq 0 \right\}$$

$$\mathcal{F}_{\tau+} := \left\{ A \in \mathcal{F}_\infty := \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t \right) : A \cap \{\tau \leq t\} \in \mathcal{F}_{t+} \, \forall t \geq 0 \right\}.$$

Napomena 2.2.4. *Pretpostavimo da imamo Brownovo gibanje* $B = \{B_t : t \geq 0\}$ *i prirodnu filtraciju* $\mathcal{F}_t^B$*. Tada je Brownovo gibanje adaptirano na tu filtraciju i prema Propoziciji*

2.2.3 proces W iz te propozicije jest nezavisan od $\mathcal{F}_t^B$. Promotrimo sada i proširenu σ-algebru $\mathcal{F}_{s+}^B$ koja je ponovno filtracija. Zbog neprekidnosti vrijedi $B_{t+s} - B_s = \lim_{n\to\infty} B_{s_n+t} - B_{s_n}$, za strogo opadajući niz $\{s_n : n \in \mathbb{N}\}$ koji konvergira prema s. Prema Propoziciji 2.2.3 slijedi da je za proizvoljne $t_1, t_2, \ldots, t_m \geq 0$ vektor $(B_{t_1+s} - B_s, B_{t_2+s} - B_s, \ldots, B_{t_m+s} - B_s) = \lim_{j\uparrow\infty}(B_{t_1+s_j} - B_{s_j}, B_{t_2+s_j} - B_{s_j}, \ldots, B_{t_m+s_j} - B_{s_j})$ nezavisan od $\mathcal{F}_{s+}^B$, pa vrijedi i za proces W. Konačno, možemo proširiti Propoziciju 2.2.3 i reći da je za svaki $s \geq 0$ slučajni proces W nezavisan od σ-algebre $\mathcal{F}_{s+}^B$.

Teorem 2.2.5. ***(Jako Markovljevo svojstvo)*** *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje i T vrijeme zaustavljanja takvo da je $T < \infty$ g.s. Tada je proces*

$$\{B_{T+t} - B_T : t \geq 0\}$$

Brownovo gibanje koje je nezavisno od $\mathcal{F}_{T+}$.

Dokaz. Pokažimo prvo tvrdnju za vremena zaustavljanja T_n koja diskretno aproksimiraju T odozgo. Definirajmo $T_n = \frac{m+1}{2^n}$ ako je $\frac{m}{2^n} \leq T < \frac{m+1}{2^n}$. Provjerimo da to jesu vremena zaustavljanja

$$\left\{T_n = \frac{k}{2^n}\right\} = \left\{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\right\} = \underbrace{\left\{T < \frac{k}{2^n}\right\}}_{\in\mathcal{F}(\frac{k}{2^n})} \setminus \underbrace{\left\{T < \frac{k-1}{2^n}\right\}}_{\in\mathcal{F}(\frac{k-1}{2^n})\subset\mathcal{F}(\frac{k}{2^n})}.$$

Označimo s $B^k = \{B_t^k : t \geq 0\}$ slučajni proces definiran s $B_t^k = B_{t+\frac{k}{2^n}} - B_{\frac{k}{2^n}}$. Prema Propoziciji 2.2.3 to je Brownovo gibanje. Nadalje, definirajmo proces $B^* = \{B_t^* : t \geq 0\}$ s $B_t^* = B_{t+T_n} - B_{T_n}$. Pretpostavimo da je $E \in \mathcal{F}_{T_n+}$. Tada za svaki događaj $\{B^* \in A\}$ imamo

$$\begin{aligned}
\mathrm{P}(\{B^* \in A\} \cap E) &= \sum_{k=0}^{\infty} \mathrm{P}(\{B^k \in A\} \cap E \cap \{T_n = \frac{k}{2^n}\}) \\
&= \sum_{k=0}^{\infty} \mathrm{P}(\{B^k \in A\})\mathrm{P}(E \cap \{T_n = \frac{k}{2^n}\}) \\
&= \mathrm{P}(\{B \in A\}) \sum_{k=0}^{\infty} \mathrm{P}(E \cap \{T_n = \frac{k}{2^n}\}) \\
&= \mathrm{P}(\{B \in A\})P(E).
\end{aligned}$$

Druga jednakost vrijedi jer je $\{B^k \in A\}$ nezavisan od $E \cap \{T_n = \frac{k}{2^n}\} \in \mathcal{F}^+(\frac{k}{2^n})$ prema Napomeni 2.2.4. Nadalje, prema Propoziciji 2.2.3 $P(B^k \in A) = P(B \in A)$ ne ovisi o k pa vrijedi treća nejednakost. Dakle, vrijedi da je B^* Brownovo gibanje nezavisno od E, pa stoga i od $\mathcal{F}_{T_n+}$. Kako $T_n \downarrow T$ vrijedi da je $\{B_{s+T_n} - B_{T_n} : s \geq 0\}$ Brownovo gibanje nezavisno od $\mathcal{F}_{T_n+} \supset \mathcal{F}_{T+}$. Stoga su prirasti

$$B_{s+t+T} - B_{t+T} = \lim_{n\to\infty} B_{s+t+T_n} - B_{t+T_n}$$

procesa $\{B_{T+t} - B_T : t \geq 0\}$ nezavisni, normalno distribuirani s očekivanjem nula i varijancom s. Kako je proces g.s. neprekidan, to je Brownovo gibanje. Štoviše, svi prirasti, a stoga i sam proces nezavisni su od $\mathcal{F}_{T+}$. □

Propozicija 2.2.6. ***(Princip refleksije)*** *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje i neka je $T_a = \inf\{t : B_t = a\}$ prvo vrijeme kada Brownovo gibanje pogodi nivo a, za neko $a > 0$. Definirajmo proces B^**

$$B_t^* = \begin{cases} B_t, & t \leq T_a \\ 2a - B_t, & t > T_a. \end{cases}$$

Tada je B^* Brownovo gibanje.

Dokaz. Definirajmo sa $C[0,\infty)$ skup neprekidnih funkcija na $[0,\infty)$. Neka je $C_0 \subset C[0,\infty)$ takav da za svaku funkciju $g \in C_0$ vrijedi $g(0) = 0$. Neka su $f \in C[0,\infty)$, $g \in C_0$, $t_0 \geq 0$. Definirajmo $\psi : C[0,\infty) \times [0,\infty) \times C_0 \mapsto C[0,\infty)$ sa

$$\psi(f, t_0, g) = \begin{cases} f(t), & t \leq t_0 \\ f(t_0) + g(t - t_0), & t \geq t_0. \end{cases}$$

Neka je sada

$$f(t,\omega) = B_{t\wedge T_a(\omega)} = \begin{cases} B_t(\omega), & t \leq T_a(\omega) \\ a, & t \geq T_a(\omega), \end{cases}$$

i definirajmo

$$g(t,\omega) = B_{T_a(\omega)+t} - B_{T_a(\omega)}.$$

Tada su slučajne varijable T_a i f određene Brownovim gibanjem do vremena T_a i nezavisne od g ili $-g$. Budući da je $T_a < \infty$ g.s. ([12], Primjer 5.9), prema Teoremu 2.2.5 i Lemi 2.1.5 vrijedi

$$g \stackrel{d}{=} -g \stackrel{d}{=} B.$$

Stoga su $(f, T_a, g) \stackrel{d}{=} (f, T_a, -g)$ kao slučajni elementi iz $C[0,\infty) \times [0,\infty) \times C_0$ i primjenom preslikavanja ψ dobijemo

$$\psi(f, T_a, g) \stackrel{d}{=} \psi(f, T_a, -g).$$

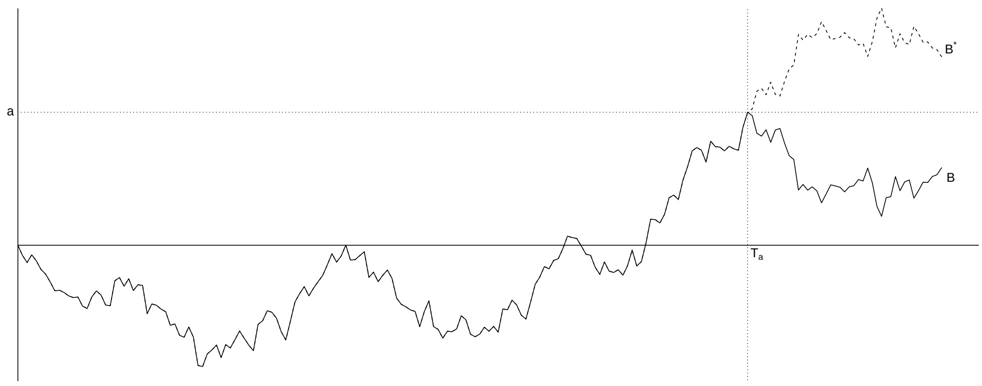
Na lijevoj strani jednakosti imamo

$$\begin{aligned}\psi(f, T_a, g) &= \begin{cases} f(t), & t \le T_a \\ f(T_a) + g(t - T_a), & t > T_a \end{cases} \\ &= \begin{cases} B_t, & t \le T_a \\ B_{T_a} + B_{T_a+t-T_a} - B_{T_a}, & t > T_a \end{cases} \\ &= B_t,\end{aligned}$$

a na desnoj strani jednakosti

$$\begin{aligned}\psi(f, T_a, -g) &= \begin{cases} f(t), & t \le T_a \\ f(T_a) - g(t - T_a), & t > T_a \end{cases} \\ &= \begin{cases} B_t, & t \le T_a \\ a - (B_{T_a+t-T_a} - B_{T_a}), & t > T_a \end{cases} \\ &= B_t^*.\end{aligned}$$

□



Slika 2.2: Princip refleksije

2.3 Lévyjev zakon trojke

Neka je $(\Omega,\mathcal{F},\mathrm{P})$ vjerojatnosni prostor i X slučajna varijabla takva da je $\mathrm{E}|X| < \infty$. Pretpostavimo da je $\mathcal{G}$ σ-podalgebra od $\mathcal{F}$. Slučajna varijabla Y koja je $\mathcal{G}$ izmjeriva, $\mathrm{E}|Y| < \infty$ i takva da za svaki $G \in \mathcal{G}$ vrijedi $\mathrm{E}(Y\mathbb{1}_G) = \mathrm{E}(X\mathbb{1}_G)$ naziva se verzija uvjetnog očekivanja X od $\mathcal{G}$ i pišemo $Y = \mathrm{E}[X\,|\,\mathcal{G}]$. Kada pišemo $\mathrm{E}[X\,|\,Y]$ to je zapravo oznaka za $\mathrm{E}[X\,|\,\sigma(Y)]$, gdje je $\sigma(Y)$ σ-algebra generirana s Y.

Definicija 2.3.1. *Neka je* $(\Omega,\mathcal{F},\mathrm{P})$ *vjerojatnosni prostor i* $\mathcal{F} = \{\mathcal{F}_t : t \geq 0\}$ *filtracija. Slučajni proces* $X = \{X_t : t \geq 0\}$ *je martingal ako vrijedi*

(i) X_t *je* $\mathcal{F}_t$ *izmjeriva za svaki* $t \geq 0$

(ii) $\mathrm{E}|X_t| < \infty$ *za svaki* $t \geq 0$

(iii) $\mathrm{E}[X_t\,|\,\mathcal{F}_s] = X_s$ *g.s za sve* $0 \leq s \leq t$.

Provjerimo da je Brownovo gibanje martingal u odnosu na prirodnu filtraciju $F = \{\mathcal{F}_t^B : t \geq 0\}$. Očito je prvo svojstvo iz definicije zadovoljeno. Nadalje, zbog $B_t \sim N(0,t)$ slijedi da je $\mathrm{E}|B_t| < \infty$. Konačno, pogledajmo zašto je zadovoljeno zadnje svojstvo:

$$\begin{aligned}
\mathrm{E}[B_t\,|\,\mathcal{F}_s] &= \mathrm{E}[B_t - B_s + B_s\,|\,\mathcal{F}_s] \\
&= \mathrm{E}[B_t - B_s\,|\,\mathcal{F}_s] + \mathrm{E}[B_s\,|\,\mathcal{F}_s] \\
&= \mathrm{E}(B_t - B_s) + B_s \\
&= B_s.
\end{aligned}$$

Posljednja jednakost slijedi zbog osnovnih svojstava uvjetnog očekivanja. Naime, vrijedi $\mathrm{E}[B_t - B_s\,|\,\mathcal{F}_s] = \mathrm{E}(B_t - B_s)$ zato što je $B_t - B_s$ nezavisno od $\mathcal{F}_s$ i $\mathrm{E}[B_s\,|\,\mathcal{F}_s] = B_s$ jer je B_s $\mathcal{F}_s$ izmjeriva.

Lema 2.3.2. *Neka je* $B = \{B_t : t \geq 0\}$ *Brownovo gibanje i* $\tau = \inf\{t \geq 0 : B_t \notin (-a,b)\}$ *vrijeme prvog ulaska u skup* $(-a,b)^c$*,* $a,b \geq 0$*. Tada vrijedi*

$$\mathrm{P}(B_\tau = -a) = \tfrac{b}{a+b}, \quad \mathrm{P}(B_\tau = b) = \tfrac{a}{a+b} \quad i \quad \mathrm{E}(\tau) = ab.$$

Dokaz. Budući da je B martingal i τ vrijeme zaustavljanja, prema teoremu o opcionalnom zaustavljanju slijedi da je tada zaustavljeni proces $B^\tau = \{B_{t\wedge t} : t \geq 0\}$ također martingal pa iz svojstva uvjetnog očekivanja vrijedi

$$\mathrm{E}B_0^\tau = \mathrm{E}B_t^\tau. \tag{2.7}$$

Vidimo da je $\mathrm{E}B_0^\tau = \mathrm{E}B_0 = 0$ i $\mathrm{E}B_t^\tau = \mathrm{E}B_{\tau\wedge t} \to \mathrm{E}B_\tau$ kada $t \to \infty$ prema Lebesgueovom teoremu o dominiranoj konvergenciji. Sada iz (2.7) slijedi

$$0 = \mathrm{E}B_\tau = -a\mathrm{P}(B_\tau = -a) + b\mathrm{P}(B_\tau = b).$$

Znamo da vrijedi i

$$1 = \mathrm{P}(B_\tau = -a) + \mathrm{P}(B_\tau = b).$$

Rješenje sustava dvije jednadžbe daje tvrdnju leme. Može se prema definiciji martingala, slično kao za Brownovo gibanje, provjeriti da je $W_t = B_t^2 - t$ martingal pa će tada i W^τ biti martingal prema teoremu o opcionalnom zaustavljanju. Tada je $0 = \mathrm{E}W_0^\tau = \mathrm{E}W_t^\tau = \mathrm{E}B_{t\wedge\tau}^2 - \mathrm{E}(t\wedge\tau)$ odakle zbog Lebesgueovog teorema o dominiranoj konvergenciji za $t \to \infty$ slijedi $\mathrm{E}\tau = \mathrm{E}B_\tau^2 = (-a)^2\mathrm{P}(B_\tau = -a) + b^2\mathrm{P}(B_\tau = b) = ab$. □

Uvedimo još neke oznake prije idućeg teorema: $\mathrm{P}^x(B \in A) := \mathrm{P}(B + x \in A)$, $\mathrm{E}^x(B) := \mathrm{E}(B + x)$. To znači da $\mathrm{P}^x(B \in A)$ označava vjerojatnost da Brownovo gibanje započinje u trenutku $t = 0$ u točki x. Očito je $\mathrm{P}^0 = \mathrm{P}$ i $\mathrm{E}^0 = \mathrm{E}$.

Teorem 2.3.3. *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje, τ vrijeme zaustavljanja u odnosu na filtraciju $\mathcal{F}_t^B$, $\eta \geq \tau$, gdje je η vrijeme zaustavljanja u odnosu na filtraciju $\mathcal{F}_{\tau+}^B$. Tada za svaki $\omega \in \{\eta < \infty\}$ i za svaku omeđenu, izmjerivu funkciju u na $\mathbb{R}$ vrijedi*

$$\begin{aligned}\mathrm{E}[u(B_\eta)|\mathcal{F}_{\tau+}^B](\omega) &= \mathrm{E}^{B_\tau(\omega)}[u(B_{\eta(\omega)-\tau(\omega)}(\cdot))]\\ &= \int u(B_{\eta(\omega)-\tau(\omega}(\omega'))\mathrm{P}^{B_\tau(\omega)}(d\omega').\end{aligned}$$

Dokaz. Zamjenom $u(B_\eta)$ s $u(B_\eta)\mathbb{1}_{[\eta<\infty]}$ možemo pretpostaviti da je $\eta < \infty$. Definirajmo

$$\sigma_j := \frac{(\lfloor 2^j(\eta - \tau)\rfloor + 1)}{2^j}.$$

Tada vrijedi da je $\inf\limits_{j\geq 1} \sigma_j = \eta - \tau$ i σ_j je očito $\mathcal{F}_{\tau+}^B$ izmjeriva budući da su η i τ $\mathcal{F}_{\tau+}^B$ izmjerive. Za sve $F \in \mathcal{F}_{\tau+}^B$ znamo da je $F \cap \left\{\sigma_j = \frac{k}{2^j}\right\} \in \mathcal{F}_{\tau+}^B$ što znači

$$\begin{aligned}\int_F u(B_{\tau+\sigma_j})d\mathrm{P} &= \sum_{k=1}^{\infty} \int_{F\cap\left\{\sigma_j=\frac{k}{2^j}\right\}} u(B_{\tau+\frac{k}{2^j}})d\mathrm{P}\\ &= \sum_{k=1}^{\infty} \int_{F\cap\left\{\sigma_j=\frac{k}{2^j}\right\}} \mathrm{E}[u(B_{\tau+\frac{k}{2^j}})\,|\,\mathcal{F}_{\tau+}^B]d\mathrm{P}\\ &= \sum_{k=1}^{\infty} \int_{F\cap\left\{\sigma_j=\frac{k}{2^j}\right\}} \mathrm{E}^{B_\tau(\omega)}u(B_{\frac{k}{2^j}})\mathrm{P}(d\omega)\\ &= \int_F \mathrm{E}^{B_\tau(\omega)}u(B_{\sigma_j(\omega)})\mathrm{P}(d\omega),\end{aligned}$$

gdje predzadnja nejednakost slijedi zbog jakog Markovljevog svojstva. Budući da je $t \mapsto B_t$ neprekidno $\lim_{j\to\infty} B_{\sigma_j} = B_{\eta-\tau}$ pa nam slijedi tvrdnja teorema za neprekidne, omeđene funkcije koristeći dominiranu konvergenciju. Za svaki kompaktni skup $K \subset \mathbb{R}$ možemo aproksimirati $\mathbb{1}_K$ s nizom neprekidnih funkcija. Stoga

$$\int_F \mathbb{1}_K(B_\eta)\, d\mathrm{P} = \int_F \mathrm{E}^{B_{\tau}(\omega)} \mathbb{1}_K(B_{\eta(\omega)-\tau(\omega)})\, \mathrm{P}(d\omega).$$

Kako kompaktni skupovi generiraju Borelovu σ-algebru, možemo koristiti teorem o jedinstvenosti za mjere da bismo proširili ovu jednakost na sve funkcije $\mathbb{1}_K$, gdje je $K \in \mathcal{B}(\mathbb{R})$ i odakle će zbog Beppo-Levijevog teorema slijediti tvrdnja teorema. □

Teorem 2.3.4. ***(Lévyjev zakon trojke)*** *Neka je $B = \{B_t : t \geq 0\}$ Brownovo gibanje i označimo sa $m_t = \inf_{s\leq t} B_s$ i $M_t = \sup_{s\leq t} B_s$. Tada za sve $t > 0$ i $a < 0 < b$ vrijedi*

$$\mathrm{P}(m_t > a, M_t < b, B_t \in dx) = \frac{dx}{\sqrt{2\pi t}} \sum_{n=-\infty}^{\infty} \left(e^{-\frac{(x+2n(b-a))^2}{2t}} - e^{-\frac{(x-2a-2n(b-a))^2}{2t}} \right).$$

Dokaz. Neka je $a < 0 < b$ i označimo s $\tau = \inf\{t \geq 0 : B_t \notin (a, b)\}$ prvo vrijeme izlaska iz intervala (a, b). Iz Leme 2.3.2 slijedi da je τ konačno vrijeme zaustavljanja. Označimo još $I = [c, d] \subset (a, b)$. Budući da je $\{\tau > t\} = \{m_t > a, M_t < b\}$ vidimo da vrijedi

$$\begin{aligned}
\mathrm{P}(m_t > a, M_t < b, B_t \in I) &= \mathrm{P}(\tau > t, B_t \in I) \\
&= \mathrm{P}(B_t \in I) - \mathrm{P}(\tau \leq t, B_t \in I) \\
&= \mathrm{P}(B_t \in I) - \mathrm{P}(B_\tau = a, \tau \leq t, B_t \in I) - \mathrm{P}(B_\tau = b, \tau \leq t, B_t \in I),
\end{aligned}$$

gdje zadnji korak slijedi iz činjenice da su $\{B_\tau = a\}$ i $\{B_\tau = b\}$, do na prazan skup $\{\tau = \infty\}$, disjunktna particija skupa Ω. Želimo uzastopno koristiti princip refleksije. Definirajmo prvo $r_a x := 2a - x$, $r_b x := 2b - x$, refleksije u (x, t)-ravnini u odnosu na $x = a$ i $x = b$. Primjenom Teorema 2.3.3 slijedi

$$\mathrm{P}(B_\tau = b, \tau \leq t, B_t \in I \,|\, \mathcal{F}_\tau)(\omega) = \mathbb{1}_{\{B_\tau = b\}\cap\{\tau\leq t\}}(\omega) \mathrm{P}^{B_\tau(\omega)}(B_{t-\tau(\omega)} \in I).$$

Zbog simetrije Brownovog gibanja, Lema 2.1.5 slijedi

$$\begin{aligned}
\mathrm{P}^b(B_{t-\tau(\omega)} \in I) &= \mathrm{P}(B_{t-\tau(\omega)} \in I - b) \\
&= \mathrm{P}(B_{t-\tau(\omega)} \in b - I) \\
&= \mathrm{P}^b(B_{t-\tau(\omega)} \in r_b I),
\end{aligned}$$

pa stoga vidimo da vrijedi

$$\begin{aligned}\mathrm{P}(B_\tau = b, \tau \le t, B_t \in I \,|\, \mathcal{F}_\tau)(\omega) &= \mathbb{1}_{\{B_\tau = b\}\cap\{\tau\le t\}}(\omega)\mathrm{P}^{B_\tau(\omega)}(B_{t-\tau(\omega)} \in I)\\ &= \mathrm{P}(B_\tau = b, \tau \le t, B_t \in r_b I \,|\, \mathcal{F}_\tau)(\omega).\end{aligned}$$

To nam daje da vrijedi

$$\mathrm{P}(B_\tau = b, \tau \le t, B_t \in I) = \mathrm{P}(B_\tau = b, \tau \le t, B_t \in r_b I) = \mathrm{P}(B_\tau = b, B_t \in r_b I).$$

Daljnje primjene principa refleksije daju

$$\begin{aligned}\mathrm{P}(B_\tau = b, B_t \in r_b I) &= \mathrm{P}(B_t \in r_b I) - \mathrm{P}(B_\tau = a, B_t \in r_b I)\\ &= \mathrm{P}(B_t \in r_b I) - \mathrm{P}(B_\tau = a, B_t \in r_a r_b I)\\ &= \mathrm{P}(B_t \in r_b I) - \mathrm{P}(B_t \in r_a r_b I) + \mathrm{P}(B_\tau = b, B_t \in r_a r_b I),\end{aligned}$$

odnosno

$$\mathrm{P}(B_\tau = b, \tau \le t, B_t \in I) = \mathrm{P}(B_t \in r_b I) - \mathrm{P}(B_t \in r_a r_b I) + \mathrm{P}(B_t \in r_b r_a r_b I) - + \dots$$

Konačno možemo zapisati

$$\begin{aligned}&\mathrm{P}(m_t > a, M_t < b, B_t \in I)\\ &\quad= \mathrm{P}(B_t \in I) - \mathrm{P}(B_t \in r_b I) + \mathrm{P}(B_t \in r_a r_b I) - \mathrm{P}(B_t \in r_b r_a r_b I) + - \dots\\ &\qquad\quad - \mathrm{P}(B_t \in r_a I) + \mathrm{P}(B_t \in r_b r_a I) - \mathrm{P}(B_t \in r_a r_b r_a I) + - \dots\\ &\quad= \mathrm{P}(B_t \in I) - \mathrm{P}(r_a B_t \in r_a r_b I) + \mathrm{P}(B_t \in r_a r_b I) - \mathrm{P}(r_a B_t \in (r_a r_b)^2 I) + - \dots\\ &\qquad\quad - \mathrm{P}(r_a B_t \in I) + \mathrm{P}(B_t \in r_b r_a I) - \mathrm{P}(r_a B_t \in r_b r_a I) + - \dots\end{aligned} \tag{2.8}$$

gdje smo u zadnjem koraku koristili činjenicu da je $r_a r_a = id$. Sve vjerojatnosti su vjerojatnosti međusobno disjunktnih događaja što znači da red konvergira apsolutno. Pogledajmo kako bismo to mogli zapisati. Vrijedi $r_a I = 2a - I$ i $r_b r_a I = 2b - (2a - I) = 2(b - a) + I$, pa tako za svaki $n \ge 0$

$$\begin{aligned}(r_b r_a)^n I &= 2n(b-a) + I\\ (r_a r_b)^n I &= 2n(a-b) + I = 2(-n)(b-a) + I.\end{aligned}$$

Stoga,

$$\begin{aligned}&\mathrm{P}(m_t > a, M_t < b, B_t \in I)\\ &\quad= \sum_{n=-\infty}^{\infty} \mathrm{P}(B_t \in 2n(a-b) + I) - \sum_{n=-\infty}^{\infty} \mathrm{P}(2a - B_t \in 2n(a-b) + I)\\ &\quad= \sum_{n=-\infty}^{\infty} \int_I \left(e^{-\frac{(x-2n(a-b))^2}{2t}} - e^{-\frac{(2a-x-2n(b-a))^2}{2t}}\right)\frac{dx}{\sqrt{2\pi t}},\end{aligned}$$

što je jednako tvrdnji teorema. □

2.4 Princip invarijantnosti

Neka je $\{X_n, n \geq 0\}$ niz nezavisnih, jednako distribuiranih slučajnih varijabli takvih da vrijedi da je $\mathrm{E}X_n = 0$ i $VarX_n = 1$. Definirajmo sada slučajnu šetnju s $S_0 = 0$ i $S_n = X_1 + \cdots X_n, \quad n \geq 1$.

Definirajmo i proces

$$B_t^n = \frac{S_{[nt]}}{\sqrt{n}}, \quad t \geq 0, \tag{2.9}$$

gdje je $[t]$ cijeli dio broja t, odnosno najveći cijeli broj koji je manji ili jednak od t. Tada vrijedi

$$B^n \implies B \tag{2.10}$$

gdje " $\implies$ " označava konvergenciju po distribuciji, a B Brownovo gibanje. Zadržimo se trenutno na konvergenciji konačno dimenzionalnih distribucija, odnosno za proizvoljan k, $0 \leq t_1 < t_2 < \cdots < t_k$ i realne brojeve $x_1, x_2, \ldots, x_k$ vrijedi

$$\lim_{n\to\infty} \mathrm{P}(B_{t_i}^n \leq x_1, i = 1, \ldots, k) = \mathrm{P}(B_{t_i} \leq x_i, i = 1, \ldots, k).$$

Provjerimo sada da tvrdnja (2.10) zaista vrijedi.

Dokaz relacije (2.10). Neka je $s < t$. Prema definiciji slučajnog procesa B^n slijedi

$$B_t^n - B_s^n = \frac{\sum\limits_{j=[ns]+1}^{[nt]} X_j}{\sqrt{n}},$$

a budući da su X_j nezavisne, jednako distribuirane slučajne varijable slijedi

$$\begin{aligned} B_t^n - B_s^n &= \frac{\sum\limits_{j=[ns]+1}^{[nt]} X_j}{\sqrt{n}} \overset{d}{=} \frac{\sum\limits_{j=1}^{[nt]-[ns]} X_j}{\sqrt{n}} = \frac{\sum\limits_{j=1}^{[nt]-[ns]} X_j}{\sqrt{[nt]-[ns]}} \frac{\sqrt{[nt]-[ns]}}{\sqrt{n}} \\ &= \frac{S_{[nt]-[ns]}}{\sqrt{[nt]-[ns]}} \frac{\sqrt{[nt]-[ns]}}{\sqrt{n}}, \end{aligned}$$

odnosno

$$\lim_{n\to\infty} \mathrm{P}(B_t^n - B_t^n \leq x) = \lim_{n\to\infty} \mathrm{P}\left(\frac{\sum\limits_{j=1}^{[nt]-[ns]} X_j}{\sqrt{[nt]-[ns]}} \frac{\sqrt{[nt]-[ns]}}{\sqrt{n}} \leq x \right). \tag{2.11}$$

Zbog

$$\frac{nt-ns-1}{n} \leq \frac{[nt]-[ns]}{n} \leq \frac{nt-ns+1}{n}$$

$$t - s - \frac{1}{n} \leq \frac{[nt] - [ns]}{n} \leq t - s + \frac{1}{n}$$

vrijedi $\lim_{n \to \infty} \frac{[nt]-[ns]}{n} = t - s$, pa je za $\epsilon \in (0, \sqrt{t-s})$

$$\sqrt{t-s} - \epsilon \leq \frac{\sqrt{[nt] - [ns]}}{\sqrt{n}} \leq \sqrt{t-s} + \epsilon \tag{2.12}$$

Konačno, iz (2.11) i (2.12) slijedi

$$\begin{aligned} \lim_{n\to\infty} \mathrm{P}\left(\frac{\sum_{j=1}^{[nt]-[ns]} X_j}{\sqrt{[nt]-[ns]}} \leq \frac{x}{\sqrt{t-s}+\epsilon} \right) &\leq \lim_{n\to\infty} \mathrm{P}(B_t^n - B_s^n \leq x) \\ &\leq \lim_{n\to\infty} \mathrm{P}\left(\frac{\sum_{j=1}^{[nt]-[ns]} X_j}{\sqrt{[nt]-[ns]}} \leq \frac{x}{\sqrt{t-s}-\epsilon} \right), \end{aligned}$$

odakle iz centralnog graničnog teorema slijedi

$$\int_{-\infty}^{\frac{x}{\sqrt{t-s}+\epsilon}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \leq \lim_{n\to\infty} \mathrm{P}(B_t^n - B_s^n) \leq \int_{-\infty}^{\frac{x}{\sqrt{t-s}-\epsilon}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy.$$

Za $\epsilon \to 0$ slijedi

$$\lim_{n\to\infty} \mathrm{P}(B_t^n - B_s^n) = \int_{-\infty}^{\frac{x}{\sqrt{t-s}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy. \tag{2.13}$$

Zamjenom varijabli lako vidimo da je izraz na desnoj strani jednakosti zapravo funkcija distribucije normalne slučajne varijable $\sim N(0, t-s)$, dakle vrijedi

$$B_t^n - B_s^n \implies N(0, t-s) \stackrel{d}{=} B_t - B_s.$$

Neka je $0 = t_0 < t_1 < t_2 < \ldots < t_k$. Varijable $B_{t_1}^n, B_{t_2}^n - B_{t_1}^n, \ldots, B_{t_k}^n - B_{t_{k-1}}^n$ su nezavisne budući da su disjunktne sume nezavisnih slučajnih varijabli. Stoga vrijedi prema Definiciji 1.2.9

$$\begin{aligned} &\mathrm{P}(B_{t_1}^n \leq x_1, B_{t_2}^n - B_{t_1}^n \leq x_2, \ldots, B_{t_k}^n - B_{t_{k-1}}^n \leq x_k) = \\ &\mathrm{P}(B_{t_1}^n \leq x_1) \cdot \mathrm{P}(B_{t_2}^n - B_{t_1}^n \leq x_2) \cdots \mathrm{P}(B_{t_k}^n - B_{t_{k-1}}^n \leq x_k) \end{aligned}$$

i budući da su prirasti Brownovog gibanja nezavisni ponovno prema definiciji vrijedi

$$\begin{aligned}&\mathrm{P}(B_{t_1} \leq x_1, B_{t_2} - B_{t_1} \leq x_2, \dots, B_{t_k} - B_{t_{k-1}} \leq x_k) = \\ &\mathrm{P}(B_{t_1} \leq x_1) \cdot \mathrm{P}(B_{t_2} - B_{t_1} \leq x_2) \cdots \mathrm{P}(B_{t_k} - B_{t_{k-1}} \leq x_k).\end{aligned}$$

Stoga vrijedi

$$\begin{aligned}&\lim_{n\to\infty} \mathrm{P}(B^n_{t_1} \leq x_1, B^n_{t_2} - B^n_{t_1} \leq x_2, \dots, B^n_{t_k} - B^n_{t_{k-1}} \leq x_k) \\ &= \lim_{n\to\infty} \mathrm{P}(B^n_{t_1} \leq x_1) \cdot \mathrm{P}(B^n_{t_2} - B^n_{t_1} \leq x_2) \cdots \mathrm{P}(B^n_{t_k} - B^n_{t_{k-1}} \leq x_k) \\ &= \mathrm{P}(B_{t_1} \leq x_1) \cdot \mathrm{P}(B_{t_2} - B_{t_1} \leq x_2) \cdots \mathrm{P}(B_{t_k} - B_{t_{k-1}} \leq x_k) \\ &= \mathrm{P}(B_{t_1} \leq x_1, B_{t_2} - B_{t_1} \leq x_2, \dots, B_{t_k} - B_{t_{k-1}} \leq x_k),\end{aligned}$$

odnosno

$$(B^n_{t_1}, B^n_{t_2} - B^n_{t_1}, \dots, B^n_{t_k} - B^n_{t_{k-1}}) \implies (B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}})$$

pa primjenom preslikavanja

$$(b_1, b_2, \dots, b_k) \mapsto (b_1, b_1 + b_2, \dots, b_1 + b_2 + \cdots + b_k)$$

slijedi tvrdnja koju je trebalo pokazati. □

Princip invarijantnosti ili kako se još naziva u literaturi funkcionalni centralni granični teorem, Donskerov teorem daje nam rezultat koji je značajniji od (2.10), konvergencije konačnodimenzionalnih distribucija. Načinimo od skupa $C[0,\infty)$,gdje $C[0,\infty)$ označava prostor svih realnih neprekidnih funkcija definiranih na domeni $[0,\infty)$, metrički prostor, gdje je metrika definirana na način da je konvergencija u tom prostoru jednaka kao lokalna uniformna konvergencija. Metriku deifniramo za $f, g \in C[0,\infty)$

$$\rho(f,g) = \sum_{n=1}^{\infty} \frac{1 \wedge \sup\limits_{0\leq t\leq n} |f(t) - g(t)|}{2^n}.$$

Stoga za $f_n \in C[0,\infty)$, $n \geq 0$,

$$\rho(f_n, f_0) \to 0 \iff \lim_{n\to\infty} \sup_{0\leq t\leq k} |f_n(t) - f_0(t)| = 0,$$

za proizvoljan k. Kako proces definiran u (2.9) nije neprekidan, prilagodit ćemo ga tako da mu trajektorije budu u skupu $C[0,\infty)$. Definirajmo

$$M^n_t = \frac{S_{[nt]}}{\sqrt{n}} + (nt - [nt])\frac{X_{[nt]+1}}{\sqrt{n}}, \quad t \geq 0.$$

Proširenje tvrdnje (2.10) jest da za proizvoljnu realnu funkciju $T : C[0,\infty) \mapsto \mathbb{R}$ vrijedi

$$T(M^n) \implies T(B), \tag{2.14}$$

gdje je $C[0,\infty)$ definirani metrički prostor i funkcija T je neprekidna sa svoje domene u $\mathbb{R}$. Primjerice, ako je $T(f) = \sup\limits_{0\le t\le 1} f(t)$, tada je $T(M^n) = \sup\limits_{0\le t\le 1}\left\{\frac{S_{[nt]}}{\sqrt{n}} + (nt-[nt])\frac{X_{[nt]+1}}{\sqrt{n}}\right\}$, pa za $n\to\infty$ imamo

$$T(M^n) = \sup_{0\le t\le 1}\frac{S_{[nt]}}{\sqrt{n}} = \frac{\sup\limits_{0\le j\le n} S_j}{\sqrt{n}} \implies \sup_{0\le t\le 1} B_t.$$

Ono što princip invarijantnosti čini značajnim jest teorem neprekinutog preslikavanja: Ako je $\psi : C[0,\infty) \to \chi$ preslikavanje iz $C[0,\infty)$ u potpun i separabilan metrički prostor χ i zadovoljava

$$\mathrm{P}(B \in \{f \in C[0,\infty) : \psi \text{ je neprekidna u f }\}) = 1 \tag{2.15}$$

tada vrijedi

$$\psi(M^n) \implies \psi(B).$$

Ako je ψ neprekidna u svim $f \in C[0,\infty)$ tada je (2.15) automatski zadovoljen. Ali, uvjet (2.15) govori nam da ψ ne mora biti neprekidna svuda nego samo na trajektorijama Brownovog gibanja. Ovdje vidimo važnost Brownovog gibanja za velike uzorke. Procjenitelj koji je funkcija parcijalnih suma slučajnog uzorka konvergirat će prema distribuciji funkcije Brownovog gibanja. To i jest glavni interes ovoga rada, a taj rezultat ćemo vidjeti u Teoremu 3.3.1.

2.5 Brownov most

Definicija 2.5.1. *Neka je $\{B_t, t \ge 0\}$ Brownovo gibanje. Definirajmo slučajni proces $\{B_t^{(m)}, 0 \le t \le 1\}$ na sljedeći način:*

$$B_t^{(m)} = B_t - tB_1, \quad 0 \le t \le 1. \tag{2.16}$$

Tako definirani slučajni proces nazivamo Brownov most.

Pogledajmo koja svojstva i zašto zadovoljava ovaj slučajni proces:

1. $B_0^{(m)} = B_1^{(m)} = 0.$
   Naime, budući da je prema definiciji Brownovog gibanja 2.0.7 $B_0 = 0$, $\mathrm{P} - g.s.$

$$B_0^{(m)} = B_0 - 0\cdot B_1 = B_0 = 0,$$
$$B_1^{(m)} = B_1 - 1\cdot B_1 = 0$$

odakle slijedi tvrdnja.

2. $\mathrm{E}B_t^{(m)} = 0$.
   Budući da je prema (2.1) $EB_t = 0$ vrijedi

$$\mathrm{E}B_t^{(m)} = \mathrm{E}[B_t - tB_1] = \mathrm{E}B_t - tEB_1 = 0 - t \cdot 0 = 0. \tag{2.17}$$

   Tvrdnja slijedi iz linearnosti očekivanja.

3. Za $0 \leq t_1 < t_2 \leq 1$, $Cov(B_{t_1}^{(m)}, B_{t_2}^{(m)}) = t_1(1 - t_2)$.
   Prema Propoziciji 2.1.4 slijedi da je $Cov(B_s, B_t) = \min\{s, t\}$. Uz svojstva nezavisnosti prirasta Brownovog gibanja iz Definicije 2.0.7, (2.1), (2.2), (2.17) i linearnosti matematičkog očekivanja spremni smo za izvod tvrdnje:

$$\begin{aligned}
Cov(B_{t_1}^{(m)}, B_{t_2}^{(m)}) &= \mathrm{E}[B_{t_1}^{(m)} B_{t_2}^{(m)}] - \mathrm{E}B_{t_1}^{(m)}\mathrm{E}B_{t_2}^{(m)} \\
&= \mathrm{E}[(B_{t_1} - t_1 B_1)(B_{t_2} - t_2 B_1)] - 0 \\
&= \mathrm{E}[B_{t_1} B_{t_2}] - t_2\mathrm{E}[B_1 B_{t_1}] - t_1\mathrm{E}[B_1 B_{t_2}] + t_1 t_2 \mathrm{E}B_1^2 \\
&= \mathrm{E}[B_{t_1}(B_{t_2} - B_{t_1})] + \mathrm{E}B_{t_1}^2 - t_2\mathrm{E}[B_{t_1}(B_1 - B_{t_1})] - t_2\mathrm{E}B_{t_1}^2 \\
&\quad - t_1\mathrm{E}[B_{t_2}(B_1 - B_{t_2})] - t_1\mathrm{E}B_{t_2}^2 + t_1 t_2 \mathrm{E}B_1^2 \\
&= \mathrm{E}B_{t_1}\mathrm{E}[B_{t_2} - B_{t_1}] + \mathrm{E}B_{t_1}^2 - t_2\mathrm{E}B_{t_1}\mathrm{E}[B_1 - B_{t_1}] - t_2\mathrm{E}B_{t_1}^2 \\
&\quad - t_1\mathrm{E}B_{t_2}\mathrm{E}[B_1 - B_{t_2}] - t_1\mathrm{E}B_{t_2}^2 + t_1 t_2 \mathrm{E}B_1^2 \\
&= 0 + t_1 - 0 - t_1 t_2 - 0 - t_1 t_2 + t_1 t_2 \\
&= t_1(1 - t_2)
\end{aligned}$$

   Treća nejednakost slijedi iz linearnosti očekivanja, a peta nejednakost iz nezavisnosti prirasta i Teorema 1.2.11.

4. $B^{(m)}$ je Gaussovski proces.
   Neka je $0 < t_1 < t_2 < \ldots < t_m$, $m \in \mathbb{N}$ proizvoljan. $(B_{t_1}, B_{t_2}, \ldots, B_{t_m})$ jest normalni slučajni vektor, pa je i

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & -t_1 \\ 0 & 1 & \cdots & 0 & -t_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t_m \end{pmatrix} \begin{pmatrix} B_{t_1} \\ B_{t_2} \\ \vdots \\ B_{t_m} \\ B_1 \end{pmatrix} = \begin{pmatrix} B_{t_1}^{(m)} \\ B_{t_2}^{(m)} \\ \vdots \\ B_{t_m}^{(m)} \end{pmatrix}$$

   normalni slučajni vektor. Dakle, prema definiciji, $B^{(m)}$ je Gaussovski proces.

Lema 2.5.2. *Neka je $B = \{B_t, t \geq 0\}$ Brownovo gibanje i $B_t^{(m)} = B_t - tB_1$ pripadajući Brownov most. Tada je B_1 nezavisan od Brownovog mosta $\{B_t^{(m)}, 0 \leq t \leq 1\}$.*

Dokaz. Dovoljno je pokazati da za su neki fiksni broj t, $B_t^{(m)}$ i B_1 nezavisni. Vidimo da je $(B_t^{(m)}, B_1)$ (bivarijantni) normalni slučajni vektor kao linearna transformacija normalnog slučajnog vektora (B_t, B_1), pa nam zbog Leme 1.3.5 preostaje pokazati da su $B_t^{(m)}$ i B_1 nekorelirane.

$$\mathrm{E}B_t^{(m)}B_1 = \mathrm{E}[(B_t - tB_1)B_1] = \mathrm{E}B_tB_1 - t\mathrm{E}B_1^2 = (*)$$

Budući da je prema Propoziciji 2.1.4 za $s \geq 0, t \geq 0$, $Cov(B_s, B_t) = \mathrm{E}B_sB_t - \mathrm{E}B_s\mathrm{E}B_t = \mathrm{E}B_sB_t = \min\{s, t\}$, i iz (2.2) slijedi

$$(*) = \min\{t, 1\} - t \cdot 1 = t - t = 0.$$

Konačno zbog (1.11) i budući da su očekivanja Brownovog gibanja (2.1) i Brownovog mosta (2.17) jedaka nuli, imamo

$$Cov(B_t^{(m)}, B_1) = \mathrm{E}[B_t^{(m)}B_1] - \mathrm{E}B_t^{(m)}EB_1 = 0 - 0 = 0.$$

□

Propozicija 2.5.3. *Neka je $B = \{B_t, t \geq 0\}$ Brownovo gibanje i neka je $B^{(\epsilon)} = \{B_t^{(\epsilon)}, 0 \leq t \leq 1\}$ slučajni proces s neprekidnim trajektorijama koji za proizvoljan k, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k \leq 1$ zadovoljava*

$$\mathrm{P}[B_{t_1}^{(\epsilon)} \leq x_1, \ldots, B_{t_k}^{(\epsilon)} \leq x_k] = \mathrm{P}[B_{t_1} \leq x_1, \ldots, B_{t_k} \leq x_k \,|\, 0 \leq B_1 \leq \epsilon].$$

Tada za $\epsilon \to 0$ vrijedi

$$B^{(\epsilon)} \implies B^{(m)} \tag{2.18}$$

u smislu konvergencije na konačno dimenzionalnim distribucijama. Štoviše, vrijedi

$$\max_{0 \leq s \leq 1} |B_s^{(\epsilon)}| \implies \max_{0 \leq s \leq 1} |B_s^{(m)}|. \tag{2.19}$$

Dokaz. Dokazat ćemo konačnodimenzionalnu konvergenciju u (2.18), a za dokaz principa invarijantnosti vidi [2]. Jednom kada vidimo slučaj k=2, biti će jasno kako to primijeniti na neki općeniti broj k. Neka su $0 \leq t_1 \leq t_2 \leq 1$

$$\begin{aligned}
&\mathrm{P}(B_{t_1}^{(\epsilon)} \in A_1, B_{t_2}^{(\epsilon)} \in A_2) = \mathrm{P}(B_{t_1} \in A_1, B_{t_2} \in A_2 \,|\, 0 \leq B_1 \leq \epsilon) \\
&= \frac{\mathrm{P}(B_{t_1} \in A_1, B_{t_2} \in A_2, 0 \leq B_1 \leq \epsilon)}{\mathrm{P}(0 \leq B_1 \leq \epsilon)} \\
&= \frac{\mathrm{P}(B_{t_1} - t_1B_1 \in A_1 - t_1B_1, B_{t_2} - t_2B_1 \in A_2 - t_2B_1, 0 \leq B_1 \leq \epsilon)}{\mathrm{P}(0 \leq B_1 \leq \epsilon)}
\end{aligned} \tag{2.20}$$

Pogledajmo sada čemu je jednak brojnik iz razlomka danog u (2.20). Budući da je prema Lemi 2.5.2 Brownov most $B^{(m)}$ nezavisan od B_1 i koristeći Fubinijev teorem vrijedi

$$\begin{aligned}
&\mathrm{P}(B_{t_1} - t_1 B_1 \in A_1 - t_1 B_1, B_{t_2} - t_2 B_1 \in A_2 - t_2 B_1, 0 \leq B_1 \leq \epsilon) \\
&= \iiint\limits_{\substack{x_1 - t_1 y \in A_1 - t_1 y \\ x_2 - t_2 y \in A_2 - t_2 y \\ y \in [0,\epsilon]}} d\mathrm{P}_{(B_{t_1} - t_1 B_1, B_{t_2} - t_2 B_1, B_1)}(dx_1, dx_2, dy) \\
&= \int\limits_0^\epsilon \underbrace{\iint\limits_{\substack{x_1 - t_1 y \in A_1 - t_1 y \\ x_2 - t_2 y \in A_2 - t_2 y}} d\mathrm{P}_{(B_{t_1} - t_1 B_1, B_{t_2} - t_2 B_1)}(dx_1, dx_2)}_{\mathrm{P}(B_{t_1}^{(m)} \in A_1 - t_1 y, B_{t_2}^{(m)} \in A_2 - t_2 y)} \, d\mathrm{P}_{B_1}(dy)
\end{aligned}$$

Stoga prema (2.20) vidimo da je

$$\begin{aligned}
\mathrm{P}(B_{t_1}^{(\epsilon)} \in A_1, B_{t_2}^{(\epsilon)} \in A_2) &= \frac{\int\limits_0^\epsilon \mathrm{P}(B_{t_1}^{(m)} \in A_1 - t_1 y, B_{t_2}^{(m)} \in A_2 - t_2 y)\, d\mathrm{P}_{B_1}(dy)}{\mathrm{P}(0 \leq B_1 \leq \epsilon)} \\
&\xrightarrow{\epsilon \to 0} \frac{\mathrm{P}(B_{t_1}^{(m)} \in A_1, B_{t_2}^{(m)} \in A_2) n(0)}{n(0)} \\
&= \mathrm{P}(B_{t_1}^{(m)} \in A_1, B_{t_2}^{(m)} \in A_2),
\end{aligned}$$

gdje je n gustoća standardne normalne distribucije, $N(0, 1)$. □

Ova propozicija govori nam da o Brownovom mostu možemo razmišljati kao o Brownovom gibanju koje je uvjetovano da bude u okolini nule u vremenu 1. Također, ova propozicija nam daje vrlo bitan rezultat, način za računanje limesa po distribuciji Kolmogorov-Smirnovljeve statistike, što ćemo vidjeti u poglavlju 3.3 Asimptotski rezultat.

Brownov most i empirijska funkcija distribucije

Prisjetimo se sada već spomenutog Gaussovog procesa $G(x)$, $x \in \mathbb{R}$ s vektorom očekivanja nula i funkcijom kovarijacije zadane s (2.4)

$$Cov(G(x_1), G(x_2)) = F(x_1)(1 - F(x_2)), \quad x_1 < x_2.$$

Ako je funkcija distribucije F, funkcija distribucije uniformne razdiobe, $F(x) = U(x) = x$, za $0 \leq x \leq 1$, tada funkcija kovarijacije procesa G glasi:

$$Cov(G(x_1), G(x_2)) = F(x_1)(1 - F(x_2)) = U(x_1)(1 - U(x_2)) = x_1(1 - x_2), \quad x_1 < x_2.$$

Budući da ako dva Gaussova procesa imaju isti vektor očekivanja i funkciju kovarijacije tada imaju istu razdiobu, možemo reći da je G Brownov most. Definirajmo sada novi proces G^* s

$$G^*(x) = B^{(m)}(F(x)), \quad x \in \mathbb{R} \tag{2.21}$$

gdje je $B^{(m)}$ Brownov most. Pogledajmo kako izgleda funkcija kovarijacije definiranog procesa:

$$Cov(G^*(x_1), G^*(x_2)) = Cov(B^{(m)}(F(x_1)), B^{(m)}(F(x_2))) = F(x_1)(1 - F(x_2)),$$

gdje prva jednakost slijedi iz definicije slučajnog procesa G^*, a druga zbog trećeg navedenog svojstva Brownovog mosta o funkciji kovarijacije. Prisjetimo se izraza (2.4). Vidimo da su G i G^* dva Gaussova procesa s jednakom funkcijom kovarijacije i stoga u smislu jednakosti konačnodimenzionalnih distribucija možemo reći da vrijedi $G^* \stackrel{d}{=} G$. Promotrimo sada vezu između Brownovog mosta i empirijske funkcije distribucije. Budući da vrijedi $G^* \stackrel{d}{=} G$, izraz (2.5) možemo napisati kao

$$\sqrt{n}(\hat{F}_n(x) - F(x)) \implies B^m(F(x)),$$

gdje smo limes iz (2.5) napisali kao funkciju Brownovog mosta, što je veza empirijske funkcije distribucije i Brownovog mosta.

Poglavlje 3

Kolmogorov-Smirnovljeva statistika

Posljednje poglavlje ovog rada rezervirano je za, kao što naslov rada govori, Brownov most i Kolmogorov-Smirnovljevu statistiku. Najprije imamo kratak uvod o testiranju statističkih hipoteza, zatim uvodimo pojam Kolmogorov-Smirnovljeve statistike definirane preko empirijske funkcije distribucije i prikazujemo statistički test u kojemu ju koristimo. Ukratko, Kolmogorov-Smirnovljevu statistiku koristimo kao testnu statistiku prilikom statističkog testiranja dolazi li dani uzorak iz određene distribucije F. Vidjet ćemo u Teoremu 3.2.1 da testna statistika neće ovisiti o funkciji distribucije F ukoliko je ona neprekidna. Konačno, u posljednjem potpoglavlju dolazimo do željenog rezultata, a to jest da Kolmogorov-Smirnovljeva statistika ima asimptotsku distribuciju i upravo Brownov most, pa posljedično Brownovo gibanje daju nam način da izračunamo tu distribuciju.

3.1 Testiranje statističkih hipoteza

Neka je X slučajna varijabla, odnosno statističko obilježje koje nam je od interesa. Slučajni uzorak duljine n za tu slučajnu varijablu jest niz $X_1, X_2, \ldots, X_n$ nezavisnih, jednako distribuiranih slučajnih varijabli koje imaju razdiobu kao X i čiju realizaciju nazivamo uzorkom. Konačno, slučajna varijabla koja je funkcija slučajnog uzorka naziva se statistika.

Statistička hipoteza je bilo koja pretpostavka o razdiobi neke slučajne varijable X. Osnovna hipoteza koja se testira naziva se nulhipotezom i ona može biti neka pretpostavka o parametrima modela ili neka izjava koja ne ovisi o parametrima modela. Recimo, ako je za dani uzorak karakteristično da dolazi iz distribucije s matematičkim očekivanjem μ i varijancom σ^2, tada su μ i σ^2 parametri tog modela i statistička hipoteza može biti neka pretpostavka o parametrima tog modela. S druge strane, mogli bismo pretpostaviti u nulhipotezi da dani uzorak dolazi iz neke konkretne distribucije F što će i biti naš slučaj. Nulhipotezu označavamo s H_0. Ukoliko ta hipoteza jednoznačno određuje distribuciju od X naziva se jednostavnom, a u suprotnom kažemo da je složena. Nasuprot nulhipoteze zadajemo i njoj alternativnu hipotezu i označavamo sa H_1.

Realizacija slučajnog uzorka služi nam da odbacimo ili ne odbacimo nulhipotezu koju smo zadali i sam proces odlučivanja hoćemo li odbaciti nulhipotezu ili ne naziva se testiranje statističke hipoteze. Odluka se temelji na vrijednosti testne statistike. Detaljno o testiranju statističkih hipoteza možete pronaći u [7], a mi ćemo se sada usredotočiti na Kolmogorov-Smirnovljevu statistiku i pripadajući test.

3.2 Kolmogorov-Smirnovljev test

Neka je $X_1, X_2, \ldots, X_n$ slučajan uzorak duljine n. Pretpostavimo da želimo testirati dolazi li taj slučajni uzorak, odnosno slučajna varijabla X iz neke distribucije koju ćemo označiti s F. To je naša statistička hipoteza, odnosno naša pretpostavka o distribuciji X. Ova statistička hipoteza je jednostavna jer jednoznačno određuje razdiobu od X. Dakle, želimo testirati nulhipotezu H_0 koja tvrdi da uzorak dolazi iz distribucije F nasuprot alternativnoj hipotezi koja tvrdi da je H_0 pogrešna. To zapisujemo na sljedeći način

$$\begin{aligned} H_0 &: X \sim F \\ H_1 &: \neg H_0. \end{aligned}$$

Da bismo donijeli odluku o odbacivanju ili ne odbacivanju statističke hipoteze, odnosno da bismo izvršili statističko testiranje računamo empirijsku funkciju distribucije definiranu prije u ovome radu. Empirijska distribucija služi nam kao aproksimacija stvarnoj funkciji distribucije neke populacije. Prisjetimo se stoga Definicije 1.15

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[X_i \leq x]}, x \in \mathbb{R}.$$

Statističko testiranje hipoteze H_0 možemo provesti mjerenjem odstupanja $\hat{F}_n(x)$ od $F(x)$, gdje očito zaključujemo da pri velikim odstupanjima odbacujemo hipotezu H_0. Mjerenje odstupanja moguće je upravo pomoću Kolmogorov-Smirnovljeve statistike D_n koja je definirana na sljedeći način

$$D_n = \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)|.$$

D_n koristimo kao testnu statistiku i odbacujemo hipotezu kada je D_n velik, odnosno kada je $D_n > k_{n,1-\alpha}$, gdje je $k_{n,1-\alpha}$ $1-\alpha$ kvantil distribucije od D_n. Korištenje ove statistike je ostvarivo jer na skupu neprekidnih funkcija distribucija vrijedi vrlo značajan rezultat koji nam govori da testna statistika D_n ne ovisi o distribuciji F. Upravo ta činjenica omogućava tabeliranje kvantila distribucije od D_n a time i jednostavnu primjenu. Kritične vrijednosti

$k_{n,1-\alpha}$ koje zadovoljavaju $\mathrm{P}(D_n \geq k_{n,1-\alpha}) = \alpha$ su dane u mnogim tekstovima (npr. [4]) za različite vrijednosti α i određene vrijednosti n, uglavnom za $n \leq 80$. Pogledajmo konačno iskaz i dokaz ovog značajnog rezultata o distribuciji statistike D_n.

Teorem 3.2.1. *Pretpostavimo da su $X_1, X_2, ..., X_n$ nezavisne, jednako distribuirane slučajne varijable s neprekidnom funkcijom distribucije F i $U_1, U_2, ..., U_n$ nezavisne, uniformno distribuirane slučajne varijable s jednakom funkcijom distribucije $U(x) = x$, $0 \leq x \leq 1$ i empirijskom funkcijom distribucije $\hat{U}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[U_i \leq x]}$. Tada vrijedi*

$$D_n := \sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| = \sup_{x\in\mathbb{R}} |\hat{U}_n(x) - U(x)|.$$

Dokaz. Neka je S zatvarač od F, odnosno S je najmanji skup koji sadrži sve vrijednosti od F. Stoga je F(S)=1 i S^c je otvoren u $\mathbb{R}$. Bilo koji otvoreni skup može se prikazati kao unija otvorenih intervala pa stoga vrijedi $S^c = \bigcup_n I_n$, gdje su $I_n = (a_n, b_n)$ otvoreni intervali takvi da je $F(I_n) = 0$. To znači da je $F(a_n) = F(b_n)$ i zbog $0 = F(I_n) = P[X_i \in I_n]$ je

$$\hat{F}_n(I_n) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[X_i \in I_n]} = 0,$$

odakle slijedi da je $\hat{F}_n(b_n) = \hat{F}_n(a_n)$. Stoga se niti F, niti $\hat{F}_n$ ne mijenja na intervalima I_n pa možemo reći

$$D_n = \sup_{x\in\mathbb{R}\setminus\bigcup_n I_n} |\hat{F}_n(x) - F(x)| = \sup_{x\in S} |\hat{F}_n(x) - F(x)|.$$

Prisjetimo se definicije generaliziranog inverza (1.1) i svojstava prikazanih u Propozicji 1.2.4.

$$\begin{aligned}\hat{F}_n(F^{\leftarrow}(u)) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[X_i \leq F^{\leftarrow}(u)]} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[F(X_i) \leq F(F^{\leftarrow}(u))]} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[F(X_i) \leq u]} \stackrel{d}{=} \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[U_i \leq u]} \\ &= \hat{U}_n(u).\end{aligned}$$

Prva jednakost jest definicija empirijske funkcije distribucije, druga slijedi zbog svojstva da je funkcija distribucije rastuća funkcija i treća iz Propozicije 1.2.4. Pretposljednja jednakost slijedi prema Propozicji 1.3.7 i posljednja jest ponovno definicija emiprijske funkcije distribucije. Dakle vrijedi

$$\hat{F}_n(F^{\leftarrow}(u)) \stackrel{d}{=} \hat{U}_n(u). \tag{3.1}$$

U Propoziciji 1.2.4 smo vidjeli da je $F^{\leftarrow}$ stogo rastuća na $(0,1)$ i tada je $F^{\leftarrow}(u) \in S$. Zbog toga vrijedi

$$\begin{aligned} D_n &= \sup_{x\in S} |\hat{F}_n(x) - F(x)| \\ &= \sup_{u\in(0,1)} |\hat{F}_n(F^{\leftarrow}(u)) - F(F^{\leftarrow}(u))| \\ &= \sup_{u\in(0,1)} |\hat{F}_n(F^{\leftarrow}(u)) - u| \\ &= \sup_{u\in(0,1)} |\hat{U}_n(u) - u| \\ &= \sup_{u\in(0,1)} |\hat{U}_n(u) - U(u)|, \end{aligned}$$

gdje treća jednakost vrijedi ponovno zbog Propozicije 1.2.4, četvrta zbog (3.1) i peta zbog definicije uniformne distribucije prikazane u Primjeru 1.3.6. Time smo dokazali teorem. □

U praksi se D_n ponekad i računa na sljedeći način. Neka je $X_1, X_2, \ldots, X_n$ realizacija slučajne varijable i $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ pripadne uređene statistike. Budući da je $\hat{F}_n(X_{(i)}) = \frac{i}{n}$ i kako $\hat{F}_n(x)$ mijenja vrijednosti samo na uređenim statistikama imamo

$$D_n = \max_{1\le i\le n} |\frac{i}{n} - F(X_{(i)})| \vee |F(X_{(i)}) - \frac{i-1}{n}|. \tag{3.2}$$

3.3 Asimptotski rezultat

U prethodnom poglavlju vidjeli smo kako koristimo Kolmogorov-Smirnovljevu statistiku za male vrijednosti n, recimo $n \le 80$. Za velike n oslanjamo se na činjenicu da D_n, kada je pravilo normaliziran, ima limes po distribuciji. Krenimo odmah s iskazom i dokazom te tvrdnje, a kasnije ćemo vidjeti da zahvaljujući tom limesu po distribuciji možemo numerički izračunati vrijednosti Kolmogorov-Smirnovljeve testne statistike i za velike n.

Teorem 3.3.1. *Pretpostavimo da su $\{X_n, n \ge 1\}$ nezavisne, jednako distribuirane slučajne varijable s neprekidnom distribucijom F. Neka je D_n Kolmogorov-Smirnovljeva testna statistika, $D_n = \sup_x |\hat{F}_n(x) - F(x)|$ i definirajmo $D := \sup_{0\le x\le 1} |B^{(m)}(x)|$. Tada vrijedi*

$$\sqrt{n} D_n \implies D. \tag{3.3}$$

Dokaz. Prisjetimo se na trenutak iskaza Teorema 3.2.1 i uočimo da se nalazimo u uvjetima teorema, odnosno F jest neprekidna funkcija distribucije. Budući da D_n ima istu distribuciju kao i kada je F uniformna distribucija, možemo $\{X_n\}$ zamijeniti s $\{U_n\}$, gdje su

$U_n \sim U(0,1)$, $n \geq 1$. Odabrat ćemo uniformne slučajne varijable na poseban način. Propozicija 1.3.11 nam tvrdi da ako su $\{E_n\}$ nezavisne, jednako distribuirane eksponencijalne slučajne varijable, i ako je $\Gamma_n = E_1 + \cdots E_n$, tada u $\mathbb{R}^n$ vrijedi

$$(U_{(1)}, \dots, U_{(n)}) \stackrel{d}{=} \left(\frac{\Gamma_1}{\Gamma_{n+1}}, \dots, \frac{\Gamma_n}{\Gamma_{n+1}} \right),$$

gdje su $(U_{(1)}, \dots, U_{(n)})$ uređene statistike duljine n iz uniformne razdiobe. Iz Teorema 3.2.1 slijedi da je

$$D_n = \sup_{x \in \mathbb{R}} |\hat{U}_n(x) - U(x)|,$$

pa zbog (3.2) imamo

$$\begin{aligned}
\sqrt{n} D_n &= \sqrt{n} \sup_{x \in \mathbb{R}} \left| \hat{U}_n(x) - U(x) \right| \\
&= \sqrt{n} \max_{1 \leq i \leq n} \left| U_{(i)} - \frac{i}{n} \right| \vee \left| U_{(i)} - \frac{i-1}{n} \right| \\
&\stackrel{d}{=} \sqrt{n} \max_{1 \leq i \leq n} \left| \frac{\Gamma_i}{\Gamma_{n+1}} - \frac{i}{n} \right| \vee \left| \frac{\Gamma_i}{\Gamma_{n+1}} - \frac{i-1}{n} \right| \\
&= \sqrt{n} \max_{1 \leq i \leq n} \left| \frac{\Gamma_i}{\Gamma_{n+1}} - \frac{i}{n} \right| + O(\frac{1}{\sqrt{n}}) \\
&= \frac{n}{\Gamma_{n+1}} \max_{1 \leq i \leq n} \left| \frac{\Gamma_i - i}{\sqrt{n}} - \frac{i}{n} \frac{\Gamma_{n+1} - n}{\sqrt{n}} \right| + O(\frac{1}{\sqrt{n}}) \\
&= \frac{n}{\Gamma_{n+1}} \max_{1 \leq i \leq n} \left| \frac{\Gamma_i - i}{\sqrt{n}} - \frac{i}{n} \frac{\Gamma_n - n}{\sqrt{n}} \right| + o_p(1).
\end{aligned}$$

$o_p(1)$ je reda $\frac{\Gamma_{n+1} - \Gamma_n}{\sqrt{n}}$, što je prema definiciji jednako $\frac{E_{n+1}}{\sqrt{n}}$. Zbog definicije distribucije eksponencijalne slučajne varijable vidimo da $o_p \implies 0$ kada $n \to \infty$. Definirajmo sada proces M^n koji će imati neprekidne trajektorije s

$$\frac{\Gamma_{[nt]} - [nt]}{\sqrt{n}}, \quad 0 \leq t \leq 1.$$

Princip invarijantnosti je sada zadovoljen i vrijedi $M^n \implies B$. Budući da je preslikavanje sa skupa neprekidnih realnih funkcija u $\mathbb{R}$

$$f(\cdot) \to \sup_{0 \leq t \leq 1} |f(t) - t f(1)|$$

neprekidno i $\frac{n}{\Gamma_{n+1}} \to 1$ prema jakom zakonu velikih brojeva (Teorem 1.4.2) slijedi

$$\begin{aligned}\sqrt{n}D_n &\overset{d}{=} \frac{n}{\Gamma_{n+1}} \max_{1\leq i\leq n} \left|M^n_{\frac{i}{n}} - \frac{i}{n} M^n_1\right| + o_p(1) \\ &= \frac{n}{\Gamma_{n+1}} \sup_{1\leq t\leq n} \left|M^n_t - tM^n_1\right| + o_p(1) \\ &\Longrightarrow \sup_{1\leq t\leq n} |B_t - tB_1| \\ &= \sup_{1\leq t\leq n} \left|B^{(m)}_t\right|.\end{aligned}$$

□

Vidimo da je asimptotska distribucija od D_n određena sa supremumom apsolutne vrijednosti Brownovog mosta. Sada je vrijeme da iskoristimo Propoziciju 2.5.3, koja nam daje vezu Brownovog mosta i Brownovog gibanja, te iskažemo Korolar koji problem računanja distribucije D svodi na problem računanja distribucije Brownovog gibanja.

Korolar 3.3.2. *Vrijedi*

$$\begin{aligned}\mathrm{P}(D > d) &= \mathrm{P}(\max_{0\leq s\leq 1} |B^{(m)}_s| > d) \\ &= \lim_{\epsilon\to 0} \mathrm{P}(\max_{0\leq s\leq 1} |B^{(\epsilon)}_s| > d) \\ &= \lim_{\epsilon\to 0} \mathrm{P}(\max_{0\leq s\leq 1} |B_s| > d | 0 \leq B_1 \leq \epsilon).\end{aligned}$$

Korolar je izravna posljedica Teorema 3.3.1 i Propozicije 2.5.3.

Prisjetimo se sada izraza (2.8). Koristeći iste oznake iz dokaza mogli smo ga raspisati i na ovaj način:

$$\begin{aligned}&\mathrm{P}(m_t > a, M_t < b, B_t \in I) \\ &\quad= \mathrm{P}(B_t \in I) - \mathrm{P}(B_t \in r_b I) + \mathrm{P}(B_t \in r_a r_b I) - \mathrm{P}(B_t \in r_b r_a r_b I) + - \ldots \\ &\qquad\quad - \mathrm{P}(B_t \in r_a I) + \mathrm{P}(B_t \in r_b r_a I) - \mathrm{P}(B_t \in r_a r_b r_a I) + - \ldots \\ &\quad= \mathrm{P}(B_t \in I) - \mathrm{P}(r_b B_t \in I) + \mathrm{P}(B_t \in r_a r_b I) - \mathrm{P}(r_b B_t \in r_a r_b I) + - \ldots \\ &\qquad\quad - \mathrm{P}(r_b B_t \in r_b r_a I) + \mathrm{P}(B_t \in r_b r_a I) - \mathrm{P}(r_b B_t \in (r_b r_a)^2 I) + - \ldots\end{aligned}$$

iz čega bi slijedilo

$$\mathrm{P}(m_t > a, M_t < b, B_t \in I) = \sum_{n=-\infty}^{\infty} \mathrm{P}(B_t \in 2n(b-a)+I) - \sum_{n=-\infty}^{\infty} \mathrm{P}(2b - B_t \in 2n(a-b)+I). \quad (3.4)$$

Promotrimo sada situaciju u kojoj je $I = [c, d]$ i $t = 1$. Uz oznaku distribucije standardne normalne razdiobe N izraz (3.4) postaje

$$\begin{aligned}\mathrm{P}(a < \min_{0\le s\le 1} < \max_{0\le s\le 1} < b, B_1 \in [c, d]) = &\sum_{k=-\infty}^{\infty} N(c + 2k(b-a), d + 2k(b-a)) \\ &- \sum_{k=-\infty}^{\infty} N(2b - d + 2k(b-a), 2b - c + 2k(b-a)).\end{aligned} \tag{3.5}$$

Pokažimo za kraj da vrijedi sljedeći asimptotski rezultat

$$\mathrm{P}(D > v) = 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2v^2}, \tag{3.6}$$

odnosno dovoljno će biti pokazati da vrijedi

$$\mathrm{P}(D \le v) = 1 - \mathrm{P}(D > v) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2v^2} = 1 + 2\sum_{k=1}^{\infty}(-1)^{k}e^{-2k^2v^2}. \tag{3.7}$$

Ekvivalentan izraz (3.7) dokazat ćemo koristeći (3.5) uz zadane $a = -v$, $b = v$, $c = 0$, $d = \epsilon$.

$$\begin{aligned}
\mathrm{P}(D \le v) &= \lim_{\epsilon\to 0} \mathrm{P}(\max_{0\le s\le 1} |B_s| < v | 0 \le B_1 \le \epsilon)) \\
&= \lim_{\epsilon\to 0} \frac{\mathrm{P}(-v < \min_{0\le s\le 1} B_s < \max_{0\le s\le 1} B_s < v, 0 \le B_1 \le \epsilon)}{\mathrm{P}(0 \le B_1 \le \epsilon)} \\
&= \lim_{\epsilon\to 0} \frac{\sum_{k=-\infty}^{\infty} N(4kv, \epsilon + 4kv) - \sum_{k=-\infty}^{\infty} N(2v - \epsilon + 4kv, 2v + 4kv)}{N((0, \epsilon])} \\
&= \frac{\sum_{k=-\infty}^{\infty} n(4kv) - \sum_{k=-\infty}^{\infty} n((4k+2)v)}{n(0)} \\
&= \frac{\sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(4kv)^2}{2}} - \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{((4k+2)v)^2}{2}}}{\frac{1}{\sqrt{2\pi}}} \\
&= \frac{2\sum_{k=1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(4kv)^2}{2}} + \frac{1}{\sqrt{2\pi}} - 2\sum_{k=0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{((4k+2)v)^2}{2}}}{\frac{1}{\sqrt{2\pi}}} \\
&= 1 + 2\left(\sum_{k=1}^{\infty} e^{-8k^2v^2} - \sum_{k=0}^{\infty} e^{-\frac{((4k+2)v)^2}{2}}\right).
\end{aligned}$$

Potrebno je još pokazati da su posljednji izraz i izraz (3.7) ekvivalentni, što će slijediti zbog

$$\begin{aligned}
\sum_{k=1}^{\infty} e^{-8k^2v^2} - \sum_{k=0}^{\infty} e^{-\frac{((4k+2)v)^2}{2}} &= \sum_{k=1}^{\infty} e^{-8k^2v^2} - \sum_{k=0}^{\infty} e^{-\frac{((2(2k+1))v)^2}{2}} \\
&= \sum_{k=1}^{\infty} e^{-8k^2v^2} - \sum_{k=0}^{\infty} e^{-\frac{4((2k+1)v)^2}{2}} \\
&= \sum_{k=1}^{\infty} e^{-2(2k)^2v^2} - \sum_{k=0}^{\infty} e^{-2(2k+1)^2v^2} \\
&= \sum_{k=1}^{\infty} (-1)^k e^{-2k^2v^2}.
\end{aligned}$$

Bibliografija

[1] D. P. Bertsekas i J. N. Tsitsiklis, *Introduction to Probability, Lecture Notes*, Massachusetts Institute of Technology, 2000.

[2] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[3] ______, *Probability and Measure (Third Edition)*, John Wiley & Sons, New York, 1995.

[4] Z. W. Birnbaum, *Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size*, Journal of the American Statistical Association **47** (1952), 425–441.

[5] R. Durrett, *Probability: Theory and Examples (Edition 4.1)*, Cambridge University Press, Cambridge, 2013.

[6] M. Huzak, *Vjerojatnost i matematička statistika, Predavanja*, PMF-Matematički odjel, 2006, `http://aktuari.math.pmf.unizg.hr/docs/vms.pdf`.

[7] E. L. Lehmann i J. P. Romano, *Testing Statistical Hypotheses*, Springer, New York, 2005.

[8] P. Mörters i Y. Peres, *Brownian motion*, Cambridge University Press, New York, 2010.

[9] Sidney I. Resnick, *Adventures in Stochastic Processes*, Birkhäuser, Boston/Basel/Berlin, 1992.

[10] G. G. Roussas, *A Course in Mathematical Statistics (Second Edition)*, Academic Press, 1997.

[11] N. Sarapa, *Teorija vjerojatnosti*, Školska knjiga, Zagreb, 2002.

[12] R.L. Schilling i L. Partzsch, *Brownian motion, An Introduction to Stochastic Processes (2nd Edition)*, De Gruyter, Berlin/Boston, 2014.

[13] Z. Vondraček, *Financijsko modeliranje 2, Skripta*, 2013, https://web.math.pmf.unizg.hr/~vondra/2fm13-predavanja.html.

Sažetak

U ovome radu uveden je pojam Brownovog mosta pomoću kojeg je izvedena asimptotska distribucija Kolmogorov-Smirnovljeve statistike koja služi za testiranje činjenice dolazi li slučajni uzorak iz neke populacije s neprekidnom razdiobom.

Prvo poglavlje je kratko ponavljanje pojmova iz teorije vjerojatnosti bez kojih je teško razumjeti ostatak rada. Uvodimo pojam vjerojatnosnog prostora, slučajne varijable i nekih njenih svojstava te primjere poznatih distribucija koji će nam biti od koristi u nastavku. Definiranjem empirijske funkcije distribucije i raspisa njenih svojstava dolazimo do motivacije za nastavak rada.

U drugom poglavlju definiramo Brownovo gibanje, jedan od najznačajnih stohastičkih procesa i dokazujemo teoreme koji će poistovjetiti Brownovo gibanje s Gaussovskim procesom, pokazati da zadovoljava jako Markovljevo svojstvo i kao posljedicu tog svojstva dokazujemo prinip refleksije. Slijedi dokaz Lévyevog zakona trojke i iskaz principa invarijantnosti, poznatijeg kao Donskerov teorem. Na koncu je definiran Brownov most, prikazana su njegova svojstva i dokazan rezultat koji će nam biti ključan u zadnjem poglavlju, a koji nam govori da je Brownov most aproksimativno jednak Brownovom gibanju privezanom da bude u okolini nule u vremenu 1. Na kraju poglavlja imamo vezu Brownovog mosta i empirijske funkcije distrbucije.

Posljednje poglavlje ovog rada uvodi pojmove vezane uz testiranje statističkih hipoteza, zatim govori o Kolmogorov-Smirnovljevoj statistici definiranoj preko empirijske funkcije distribucije i prikazuje statistički test u kojem se koristi kao testna statistika. Konačno dolazimo do krajnjeg rezultata koji nam govori da Kolmogorov-Smirnovljeva statistika ima asimptotsku distribuciju i da upravo Brownov most, pa posljedično i Brownovo gibanje nam daju način kako da ju izračunamo.

Summary

This thesis introduces the concept of Brownian bridge with which is derived the asymptotic distribution of Kolmogorov-Smirnov statistic used for testing whether a random sample comes from a particular population with countinuous distribution.

The first chapter is a brief introduction to the concepts of probability theory without which is difficult to understand the subject matter of this paper. Probability space, random variables along with some of its properties and examples of well known distributions that will be useful later on, are specified in this chapter. Defining the empirical distribution function and its properties brings us the motivation for the main task of this paper.

Further, the second chapter defines one of the most important stochastic processes, Brownian motion. Theorems to identify Brownian motion with Gaussian process and to demonstrate that it meets strong Markov property are proved. As a result of strong Markov property, the reflection principle is shown and then applied repeatedly to obtain the proof of Lévy's triple law. The invariance principle known as Donsker's theorem is stated. Finally, Brownian bridge and its properties are defined, and it is proved that Brownian bridge is approximately Brownian motion conditioned to be in the neighborhood of the zero at time 1. At the end of this chapter we link Brownian bridge with empirical distribution function.

The last chapter of this thesis introduces concepts related to testing of statistical hypotheses, then talks about Kolmogorov-Smirnov statistics defined by the empirical distribution function and displays statistical test in which it is used as a test statistic. Finally we come to the main result which tells us that the Kolmogorov-Smirnov statistic has an asymptotic distribution and that Brownian bridge, and thus Brownian motion give us a way to compute it.

Životopis

Rođena sam 02.11.1991. godine u Beču. Obrazovanje sam započela u Područnoj školi Rakitje i nastavila u Osnovnoj školi Sveta Nedelja. Nakon osnovnog obrazovanja, 2006. godine upisala sam prirodoslovno-matematičku gimnaziju, V. gimnaziju u Zagrebu. Neposredno nakon završetka srednje škole, 2010. godine upisala sam Prirodoslovno-matematički fakultet, preddiplomski studij Matematika gdje sam 2013. godine stekla akademski naziv sveučilišne prvostupnice matematike. Svoje obrazovanje nastavila sam na istom fakultetu, na diplomskom studiju Financijska i poslovna matematika koji završavam ovim diplomskim radom. Tijekom studija dobila sam Priznanje za iznimne rezultate u izvannastavnim aktivnostima u akademskoj godini 2012./2013. Prirodoslovno-matematičkog fakulteta, sudjelovala sam u radu studentske udruge, usavršavala znanje stranih jezika.