Steinova metoda i normalna aproksimacija

Ivanović, Krunoslav

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Krunoslav Ivanović

STEINOVA METODA I NORMALNA APROKSIMACIJA

Diplomski rad

Voditelj rada:
prof. dr. sc. Bojan Basrak

Zagreb, 2024.

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

Uvod

Centralni granični teorem jedan je od povijesno najznačajnijih rezultata u teoriji vjerojatnosti. Klasični dokaz temelji se na Fourierovim metodama, odnosno, karakterističnim funkcijama pa ga nije nimalo jednostavno adaptirati u slučaju kada ne radimo sa sumom nezavisnih jednako distribuiranih slučajnih varijabli.

Sedamdesetih godina prošlog stoljeća, Charles Stein je uveo novu metodu dokazivanja graničnih teorema koja je danas poznata kao Steinova metoda. Ta metoda nam omogućava da dobijemo eksplicitne ograde na grešku prilikom aproksimacije čak i u slučajevima kada nemamo nezavisnost. Štoviše, metode inspirirane Steinovom su omogućile slične rezultate i u slučajevima kada granična distribucija nije normalna. Također, Steinova metoda nalazi primjene ne samo u vjerojatnosti, nego i statistici i strojnom učenju.

U prvom dijelu rada fokusiramo se na Steinovu metodu za normalne aproksimacije. Na početku, iznosimo generalnu teoriju koja je potrebna za daljnje razumijevanje, a nakon toga prelazimo na različite kontekste u kojima se Steinova metoda može koristiti. Prvo proučavamo klasičnu teoriju zbroja nezavisnih slučajnih varijabli pri čemu dokazujemo sve klasične centralne granične teoreme. Nakon toga promatramo metode koje se mogu koristiti u slučajevima kada slučajne varijable nisu nezavisne. U ovom dijelu slijedimo [7], [3] te [4].

U drugom dijelu rada, dokazujemo egzistenciju sparivanja između jednostavne simetrične slučajne šetnje u jednoj dimenziji te Brownovog gibanja, koje konstruiramo na istom vjerojatnosnom prostoru kao i šetnju. Metoda koju koristimo je inspirirana i povezana sa Steinovom lemom i fundamentalno je različita od originalnog dokaza. U ovom dijelu slijedimo [1].

Poglavlje 1

Steinova metoda

1.1 Osnovne definicije

Kako se u ovom radu bavimo konvergencijom nizova slučajnih varijabli te brzinom konvergencije, prije svega potrebno je definirati metrike koje ćemo promatrati i pogledati njihova osnovna svojstva.

Definicija 1.1 (Vjerojatnosne metrike)**.** *Za dvije vjerojatnosne mjere μ i ν i familiju testnih funkcija $\mathcal{H}$, definiramo*

$$d_{\mathcal{H}}(\mu, \nu) := \sup_{f \in \mathcal{H}} \left| \int h(x)\mu(\mathrm{dx}) - \int h(x)\nu(\mathrm{dx}) \right|.$$

Također, ako su X i Y slučajne varijable pisati ćemo $\mathrm{d}_{\mathcal{H}}(X, Y)$ kao pokratu za $\mathrm{d}_{\mathcal{H}}(\mathbb{P}_X, \mathbb{P}_Y)$, gdje $\mathbb{P}_X$ i $\mathbb{P}_Y$ označavaju vjerojatnosne mjere inducirane slučajnim varijablama X i Y, redom.

Ovisno o familiji $\mathcal{H}$ koju promatramo, dobivamo različite metrike. Tri najpoznatije metrike koje dobivamo na ovaj način su:

- Ako uzmemo
$$\mathcal{H} = \left\{ \mathbb{1}_{(-\infty, t]}(x) \colon t \in \mathbb{R} \right\}, \tag{1.1}$$
odnosno, skup indikatorskih funkcija intervala oblika $(-\infty, t]$, gdje je t realan broj, dobivamo Kolmogorovljevu metriku, koju označavamo s d_K. Direktno iz definicije (1.1), vidimo da vrijedi
$$\begin{aligned} \mathrm{d}_K(X, Y) &= \sup_{t \in \mathbb{R}} \left| \int \mathbb{1}_{(-\infty, t]}(x)\mu(\mathrm{dx}) - \int \mathbb{1}_{(-\infty, t]}(x)\nu(\mathrm{dx}) \right| \\ &= \sup_{t \in \mathbb{R}} |\mathbb{P}(X \leqslant t) - \mathbb{P}(Y \leqslant t)| = \|F_X - F_Y\|_\infty, \end{aligned} \tag{1.2}$$

gdje F_X i F_Y predstavljaju funkcije distribucije slučajnih varijabli X i Y.

- Ako uzmemo
$$\mathcal{H} = \left\{h\colon \mathbb{R} \to \mathbb{R}\colon \|h'\|_\infty \leqslant 1\right\},$$
skup svih Lipschitzovih funkcija s Lipschitzovom konstantom manjom ili jednakom jedan, dobivamo Wassersteinovu udaljenost koju označavamo s d_W.

- Ako uzmemo
$$\mathcal{H} = \{\mathbb{1}_A(x)\colon A \subseteq \mathbb{R} \text{ Borelov}\},$$
dobivamo udaljenost totalne varijacije koju označavamo sa d_{TV}.

Jasno, odmah se vidi da je $\mathrm{d}_K \leqslant \mathrm{d}_{TV}$ s obzirom da uzimamo supremum po većem skupu. Zanimljiviji je odnos u sljedećoj propoziciji.

Propozicija 1.2. *Ako slučajna varijabla Z ima gustoću koja je ograničena nekom konstantom $C > 0$, tada, za bilo koju slučajnu varijablu X vrijedi*

$$\mathrm{d}_K(X,Z) \leqslant \sqrt{2C\,\mathrm{d}_W(X,Z)}. \tag{1.3}$$

Dokaz. Označimo s $g_t(x) = \mathbb{1}_{(-\infty,t]}(x)$. Uvedimo

$$g_{t,\varepsilon}(x) = \begin{cases} 1 & x \leqslant t \\ 0 & x > t+\varepsilon \\ \text{linearno} & \text{između.} \end{cases}$$

Sada, imamo

$$\begin{aligned} \mathbb{E}[g_t(X)] - \mathbb{E}[g_t(Z)] &= \mathbb{E}[g_t(X)] - \mathbb{E}[g_{t,\varepsilon}(Z)] + \mathbb{E}[g_{t,\varepsilon}(Z)] - \mathbb{E}[g_t(Z)] \\ &\leqslant \mathbb{E}[g_{t,\varepsilon}(X)] - \mathbb{E}[g_{t,\varepsilon}(Z)] + C\varepsilon/2 \\ &\leqslant \mathrm{d}_W(X,Z)/\varepsilon + C\varepsilon/2, \end{aligned}$$

pri čemu prvi dio prve nejednakosti vrijedi jer je $g_{t,\varepsilon} \geqslant g_t$, a drugi dio vrijedi jer je

$$\mathbb{E}[g_{t,\varepsilon}(Z)] - \mathbb{E}[g_t(Z)] = \int_t^{t+\varepsilon} f_Z(w)(g_{t,\varepsilon}(w) - g_t(w))\mathrm{d}w \leqslant C\varepsilon/2.$$

Druga nejednakost vrijedi jer je $\varepsilon g_{t,\varepsilon}$ Lipschitzova s konstantom jedan. Uzimanjem $\varepsilon = \sqrt{2\,\mathrm{d}_W(X,Z)/C}$, uočavanjem da smo potpuno analogno mogli dobiti ogradu u drugom smjeru te uzimanjem supremuma po svim g_t, dobivamo tvrdnju leme.

□

lema koju smo upravo dokazali dopušta da, u slučaju kada je $Z \sim N(0,1)$, zaključimo da konvergencija u Wassersteinovoj metrici povlači i konvergenciju u Kolmogorovljevoj metrici. Iz (1.2), vidimo da će konvergencija u Kolmogorovljevoj metrici povlačiti konvergenciju u distribuciji. Zbog toga, ako želimo dokazati da neki niz $(X_n)_{n\in\mathbb{N}}$ konvergira u distribuciji prema jediničnoj normalnoj slučajnoj varijabli, dovoljno je dokazati konvergenciju u Wassersteinovoj metrici. No, ograde na Kolmogorovljevu metriku koje dobijemo iz (1.3) nisu optimalne. Za optimalne ograde bismo trebali ograničavati Kolmogorovljevu metriku direktno.

Naposljetku, prisjetimo se definicije apsolutno neprekidnih funkcija te nekih njihovih svojstava.

Definicija 1.3. *Funkcija $F\colon [a,b] \to \mathbb{R}$ je apsolutno neprekidna ako za sve $\varepsilon > 0$ postoji $\delta > 0$ takav da*

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| < \varepsilon \quad \textit{kad god} \quad \sum_{k=1}^{N} (b_k - a_k) < \delta,$$

gdje su proizvoljni intervali $(a_k, b_k), k = 1, \dots, N$ disjunktni. Funkcija $F\colon \mathbb{R} \to \mathbb{R}$ je apsolutno neprekidna ako je apsolutno neprekidna svaka njena restrikcija $F\big|_{[a,b]}$.

Ključno svojstvo apsolutno neprekidnih funkcija je u sljedećem teoremu, koji navodimo bez dokaza.

Teorem 1.4. *Neka je F apsolutno neprekidna funkcija na $[a,b]$. Tada, F' postoji gotovo svuda i integrabilna je. Štoviše,*

$$F(x) - F(a) = \int_a^x F'(y)\mathrm{d}y, \quad \textit{za sve } a \leqslant x \leqslant b.$$

Obratno, ako je f integrabilna na $[a,b]$, tada postoji apsolutno neprekidna funkcija F takva da je $F'(x) = f(x)$, gotovo svugdje. Možemo uzeti $F(x) = \int_a^x f(y)\mathrm{d}y$.

Također, bitna nam je sljedeća propozicija.

Propozicija 1.5. *Svaka Lipschitz neprekidna funkcija je apsolutno neprekidna.*

Dokaz. Neka je $f\colon \mathbb{R} \to \mathbb{R}$ Lipschitz neprekidna s konstantom L, odnosno, vrijedi

$$|f(x) - f(y)| \leqslant L|x - y|$$

za sve realne x i y. Tada za disjunktne intervale $(a_k, b_k), k = 1, \dots, N$, vrijedi

$$\sum_{k=1}^{N} |f(b_k) - f(a_k)| \leqslant \sum_{k=1}^{N} L|b_k - a_k|$$

pa za proizvoljan $\varepsilon > 0$ možemo uzeti $\delta = \varepsilon/L$ u definiciji apsolutne neprekidnosti. $\square$

1.2 Steinova lema

Na početku, objasnimo heuristički što želimo napraviti. Ključna činjenica koju koristimo je Steinova lema — za slučajnu varijablu Z vrijedi da je

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[f'(Z)] \tag{1.4}$$

za sve „lijepe” funkcije f ako i samo ako je $Z \sim N(0,1)$. Iz toga je razumno pretpostaviti da je neka slučajna varijabla W „približno normalna” ako je jednakost (1.4) približno zadovoljena.

Sada, za ograničenu i izmjerivu funkciju $g \colon \mathbb{R} \to \mathbb{R}$ promatramo f, ograničeno rješenje diferencijalne jednadžbe

$$f'(x) - xf(x) = g(x) - \mathbb{E}[g(Z)], \tag{1.5}$$

pri čemu je $Z \sim N(0,1)$. Naravno, potrebno je dokazati da takvo ograničeno rješenje postoji, no ako postoji, koristeći (1.5), dobivamo

$$\mathbb{E}[g(W)] - \mathbb{E}[g(Z)] = \mathbb{E}[g(W) - \mathbb{E}g(Z)] = \mathbb{E}[f'(W) - Wf(W)]. \tag{1.6}$$

Vidimo da supremum lijeve strane nad nekom klasom funkcija g možemo ocijeniti tako što gledamo supremum desne strane nad odgovarajućim funkcijama f induciranim od g. Pri tome, važno je istaknuti da se na desnoj strani unutar očekivanja nalazi samo jedna slučajna varijabla pa je za očekivati da je taj izraz lakši za ocijeniti nego lijevu stranu. Upravo u tome je ključ Steinove metode.

Kako bi ovo formalizirali, potrebno je prije svega precizno iskazati i dokazati Steinovu lemu te analizirati povezanost svojstava funkcije g i svojstava rješenja f diferencijalne jednadžbe (1.5).

Lema 1.6. *Za fiksan $z \in \mathbb{R}$ i $\Phi(z) := \mathbb{P}(Z \leqslant z)$, jedinstveno ograničeno rješenje $f(w) := f_z(w)$ jednadžbe*

$$f'(w) - wf(w) = \mathbb{1}_{(-\infty,z]}(w) - \Phi(w) \tag{1.7}$$

dano je s

$$f_z(w) = \begin{cases} \sqrt{2\pi}e^{w^2/2}\Phi(w)(1-\Phi(z)) & w \leqslant z \\ \sqrt{2\pi}e^{w^2/2}\Phi(z)(1-\Phi(w)) & w > z. \end{cases} \tag{1.8}$$

Dokaz. Množenje obje strane u (1.7) s $e^{-w^2/2}$ daje jednakost

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(e^{-w^2/2}f(w)\right) = e^{-w^2/2}\left(\mathbb{1}_{(-\infty,z]}(w) - \Phi(z)\right).$$

Integriranjem dobivamo

$$\begin{aligned}
f_z(w) &= e^{w^2/2}\int_{-\infty}^{w}(\mathbb{1}_{(-\infty,z]}(x) - \Phi(z))e^{-x^2/2}\mathrm{d}x \\
&= -e^{w^2/2}\int_{w}^{\infty}(\mathbb{1}_{(-\infty,z]}(x) - \Phi(z))e^{-x^2/2}\mathrm{d}x
\end{aligned}$$

što je upravo (1.8). Ograničenost će slijediti iz leme 1.8.

Opće rješenje diferencijalne jednadžbe (1.7) je dano kao zbroj $f_z(w)$ i rješenja odgovarajuće homogene jednadžbe, koje je oblika $ce^{w^2/2}$, za neku realnu konstantu c. Jasno, ako želimo da je rješenje ograničena funkcija, jedina opcija je $c = 0$.

□

Lema 1.7 (Steinova lema)**.** *Ako W ima jediničnu normalnu distribuciju, tada*

$$\mathbb{E}[f'(W)] = \mathbb{E}[Wf(W)] \tag{1.9}$$

za sve apsolutno neprekidne funkcije $f\colon \mathbb{R} \to \mathbb{R}$ za koje je $\mathbb{E}|f'(W)| < \infty$.

Obratno, ako je jednakost (1.9) zadovoljena za sve ograničene, neprekidne i po dijelovima glatko derivabilne funkcije $f\colon \mathbb{R} \to \mathbb{R}$ za koje je $\mathbb{E}|f'(Z)| < \infty$, gdje je $Z \sim N(0,1)$, tada W ima jediničnu normalnu distribuciju.

Dokaz. Neka je W jedinična normalna i $f\colon \mathbb{R} \to \mathbb{R}$ apsolutno neprekidna funkcija takva da je $\mathbb{E}|f'(W)| < \infty$. Tada

$$\begin{aligned}
\mathbb{E}[f'(W)] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f'(w)e^{-w^2/2}\mathrm{d}w = \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} f'(w)\left(\int_{-\infty}^{w} -xe^{-x^2/2}\mathrm{d}x\right)\mathrm{d}w \\
&+ \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty} f'(w)\left(\int_{w}^{\infty} xe^{-x^2/2}\mathrm{d}x\right)\mathrm{d}w.
\end{aligned}$$

Zbog uvjeta da je $\mathbb{E}|f'(W)| < \infty$, smijemo primijeniti Fubinijev teorem. Dobivamo

$$\begin{aligned}
\mathbb{E}[f'(W)] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}\left(\int_{x}^{0} f'(w)\mathrm{d}w\right)(-x)e^{-x^2/2}\mathrm{d}x \\
&+ \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\left(\int_{0}^{x} f'(w)\mathrm{d}w\right)xe^{-x^2/2}\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(f(x) - f(0))xe^{-x^2/2}\mathrm{d}x = \mathbb{E}[Wf(W)],
\end{aligned}$$

pri čemu zadnja jednakost vrijedi jer integral koji sadrži $f(0)$ iščezava (integriramo neparnu funkciju po cijelom $\mathbb{R}$).

Za drugi smjer, neka je W slučajna varijabla koja zadovoljava (1.9) za sve apsolutno neprekidne funkcije $f\colon \mathbb{R} \to \mathbb{R}$ za koje je $\mathbb{E}|f'(Z)| < \infty$. Ako za f uzmemo funkciju f_z definiranu u (1.8), iz leme 1.6 dobivamo ograničenost, dok je neprekidnost i po dijelovima glatka derivabilnost jasna iz definicije funkcije. Uvrštavanjem u (1.9) dobivamo

$$0 = \mathbb{E}[f_z'(W) - Wf_z(W)] = \mathbb{E}[(-\infty,](W) - \Phi(z)] = \mathbb{P}(W \leqslant z) - \Phi(z).$$

Kako ovo vrijedi za sve $z \in \mathbb{R}$, vidimo da je $W \sim N(0,1)$.

□

Naposljetku, ostaje lema koja govori o svojstvima rješenja f_z jednadžbe (1.7).

Lema 1.8. *Neka je $z \in \mathbb{R}$ i f_z funkcija dana u (1.8). Tada je $wf_z(w)$ rastuća funkcija u w. Štoviše, za sve realne u, v i w vrijedi*

$$|wf_z(w)| \leqslant 1, \qquad |wf_z(w) - uf_z(u)| \leqslant 1 \tag{1.10}$$

$$|f_z'(w)| \leqslant 1, \qquad |f_z'(w) - f_z'(u)| \leqslant 1 \tag{1.11}$$

$$0 < f_z(w) \leqslant \min(\sqrt{2\pi}/4, 1/|z|). \tag{1.12}$$

Također, vrijedi i

$$|(w+u)f_z(w+u) - (w+v)f_z(w+v)| \leqslant (|w| + \sqrt{2\pi}/4)(|u| + |v|). \tag{1.13}$$

Dokaz. Jer je $f_z(w) = f_{-z}(-w)$, dovoljno je tvrdnju dokazati za $z \geqslant 0$. Za $w > 0$,

$$\int_w^\infty e^{-x^2/2} \mathrm{d}x \leqslant \int_w^\infty \frac{x}{w} e^{-x^2/2} \mathrm{d}x = -\frac{e^{-x^2/2}}{w}\Bigg|_w^\infty = \frac{e^{-w^2/2}}{w}.$$

Također, jer je

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{xe^{-x^2/2}}{1+x^2}\right) = -\frac{(x^4 + 2x^2 - 1)e^{-x^2/2}}{(1+x^2)^2} \leqslant e^{-x^2/2}$$

integriranjem dobivamo

$$\int_w^\infty e^{-x^2/2} \mathrm{d}x \geqslant \frac{we^{-w^2/2}}{1+w^2}.$$

Spajanjem dobivenih nejednakosti, dobivamo

$$\frac{we^{-w^2/2}}{(1+w^2)\sqrt{2\pi}} \leqslant 1-\Phi(w) \leqslant \frac{e^{-w^2/2}}{w\sqrt{2\pi}}. \tag{1.14}$$

Direktnim računom iz leme 1.6 dobivamo da je

$$(wf_z(w))' = \begin{cases} \sqrt{2\pi}(1-\Phi(z))\left((1+w^2)e^{w^2/2}\Phi(w) + \frac{w}{\sqrt{2\pi}}\right) & \text{za } w < z, \\ \sqrt{2\pi}\Phi(z)\left((1+w^2)e^{w^2/2}(1-\Phi(w)) - \frac{w}{\sqrt{2\pi}}\right) & \text{za } w > z. \end{cases}$$

Koristeći donju ogradu iz (1.14), vidimo da je $(wf_z(w))' \geqslant 0$ za $w > z$. Kako je za predznak „bitni" dio izraza simetričan u oba slučaja (zamijenimo w s $-w$), analogno zaključujemo da je i $(wf_z(w))' \geqslant 0$ za $w < z$. Dakle, $(wf_z(w))' \geqslant 0$ za sve w, to jest, dokazali smo prvu tvrdnju leme.

Također, iz nejednakosti (1.14), za $w > 0$, vrijedi

$$\frac{w^2}{1+w^2} \leqslant (1-\Phi(w))we^{w^2/2}\sqrt{2\pi} \leqslant 1.$$

Puštanjem w u ∞, vidimo da je $\lim_{w\to\infty}(1-\Phi(w))we^{w^2/2}\sqrt{2\pi} = 1$. Uvrštavanjem $-w$ umjesto w, vidimo da je i $\lim_{w\to-\infty}\Phi(w)we^{w^2/2}\sqrt{2\pi} = -1$. Kombinirajući ove limese i jednakost (1.8), dobivamo

$$\lim_{w\to-\infty} wf_z(w) = \Phi(z) - 1, \quad \lim_{w\to\infty} wf_z(w) = \Phi(z) \tag{1.15}$$

iz čega jasno slijedi (1.10).

Prisjetimo se, iz definicije, odnosno, rješenja Steinove jednadžbe, imamo

$$\begin{aligned} f_z'(w) &= wf_z(w) + \mathbb{1}_{(-\infty,z]}(w) - \Phi(z) \\ &= \begin{cases} wf_z(w) + 1 - \Phi(z) & \text{za } w < z, \\ wf_z(w) - \Phi(z) & \text{za } w > z; \end{cases} \\ &= \begin{cases} (\sqrt{2\pi}we^{w^2/2}\Phi(w) + 1)(1-\Phi(z)) & \text{za } w < z, \\ (\sqrt{2\pi}we^{w^2/2}(1-\Phi(w)) - 1)\Phi(z) & \text{za } w > z. \end{cases} \end{aligned}$$

Jer je $wf_z(w)$ rastuća funkcija u w, koristeći (1.15), vrijedi

$$0 < f_z'(w) \leqslant zf_z(z) + 1 - \Phi(z) < 1 \quad \text{za } w < z$$

te

$$-1 < zf_z(z) - \Phi(z) \leqslant f_z'(w) < 0 \quad \text{za } w > z$$

čime smo dokazali prvu nejednakost u (1.11). Za drugu nejednakost, ako su $f_z'(u)$ i $f_z'(w)$ istog predznaka, iskoristimo prethodnu ogradu. Ako ne, kombiniranjem gornjih ocjena, imamo

$$|f_z'(w) - f_z'(u)| \leqslant zf_z(z) + 1 - \Phi(z) - (zf_z(z) - \Phi(z)) = 1.$$

Uočimo da zbog prethodnih ocjena na f_z', f_z postiže maksimum u z. Dakle,

$$0 < f_z(w) \leqslant f_z(z) = \sqrt{2\pi}e^{z^2/2}\Phi(z)(1 - \Phi(z)).$$

Iz (1.14), imamo da je $f_z(z) \leqslant \frac{1}{z}$. Definirajmo

$$g(z) := \Phi(z)(1 - \Phi(z)) - \frac{e^{-z^2/2}}{4}, \quad g_1(z) := \frac{1}{\sqrt{2\pi}} + \frac{z}{4} - \frac{2\Phi(z)}{\sqrt{2\pi}}.$$

Direktnim računom dobivamo da je $g'(z) = e^{-z^2/2}g_1(z)$ te

$$g_1(0) = 0, g_1'(0) < 0, g_1''(z) = \frac{z}{\pi}e^{-z^2/2}, \lim_{z\to\infty} g_1(z) = \infty.$$

Dakle, funkcija g_1 je konveksna na $[0, \infty)$ te postoji $z_1 > 0$ takav da je $g_1(z) < 0$ za $z < z_1$ i $g_1(z) > 0$ za $z > z_1$. Konkretno, na $[0, \infty)$, funkcija $g(z)$ pada za $z < z_1$ i raste za $z > z_1$, dakle, njen supremum se postiže ili u 0, ili u ∞, to jest,

$$g(z) \leqslant \max(g(0), g(\infty)) = 0, \forall z \in [0, \infty)$$

što je ekvivalentno s $f_z(z) \leqslant \sqrt{2\pi}/4$. Ovime smo dokazali (1.12).

Naposljetku, za posljednju nejednakost, napišimo

$$\begin{aligned}&(w+u)f_z(w+u) - (w+v)f_z(w+v)\\ &\quad = w(f_z(w+u) - f_z(w+v)) + uf_z(w+u) - vf_z(w+v)\end{aligned}$$

pa primijenimo teorem srednje vrijednosti i (1.11) na prvi član te nejednakost (1.12) na drugi član.

$\square$

Sada vidimo osnovne ideje iza Steinove metode. Naime, ako želimo ograničiti udaljenost između W i Z u Kolmogorovljevoj metrici, koristeći (1.6), to možemo napraviti ograničavajući izraz $\mathbb{E}[f_z'(W) - Wf_z(W)]$. Naravno, pri tome koristimo svojstva funkcije f_z koja smo iskazali u lemi 1.8. Kako bi to učinili, potrebno je nekako iskoristiti strukturu slučajne varijable W.

Prije analiziranja različitih konteksta unutar kojih se slučajne varijable W mogu dobro ograničiti, proširiti ćemo dosad uvedenu metodu na ograničavanje u Wassersteinovoj metrici. Kao što ćemo vidjeti, funkcije f_h koje ćemo dobiti analizom u Wassersteinovoj metrici će imati ljepša svojstva nego funkcije f_z koje smo upravo uveli. Kako su dokazi sljedećih tvrdnji slični već napravljenim dokazima, dokaze, koji se mogu naći u [3] izostavljamo.

Za izmjerivu funkciju $h\colon \mathbb{R} \to \mathbb{R}$ za koju je $\mathbb{E}|h(Z)| < \infty$, promatramo Steinovu jednadžbu

$$f'(w) - wf(w) = h(w) - \mathbb{E}[h(Z)].$$

Ako rješenje te jednadžbe označimo s f_h, za proizvoljnu slučajnu varijablu W vrijedi

$$\mathbb{E}[h(W)] - \mathbb{E}[h(Z)] = \mathbb{E}[f_h'(W) - Wf_h(W)] \tag{1.16}$$

te opet lijevu stranu možemo ograničiti gledajući isključivo desnu. Kao i prije, važno pitanje je egzistencija i svojstva funkcije f_h.

Lema 1.9. *Neka je $h\colon \mathbb{R} \to \mathbb{R}$ izmjeriva te neka je $\mathbb{E}|h(Z)| < \infty$, gdje je $Z \sim N(0,1)$. Jedinstveno ograničeno rješenje diferencijalne jednadžbe*

$$f'(w) - wf(w) = h(w) - \mathbb{E}[h(Z)] \tag{1.17}$$

je dano s

$$\begin{aligned} f_h(w) &= e^{w^2/2} \int_{-\infty}^{w} (h(x) - \mathbb{E}[h(Z)])\, e^{-x^2/2} \mathrm{d}x \\ &= -e^{w^2/2} \int_{w}^{\infty} (h(x) - \mathbb{E}[h(Z)])\, e^{-x^2/2} \mathrm{d}x \end{aligned} \tag{1.18}$$

Lema 1.10. *Za danu izmjerivu funkciju $h\colon \mathbb{R} \to \mathbb{R}$, neka je f_h ograničeno rješenje Steinove jednadžbe (1.17) dano u (1.18). Ako je h ograničena, tada*

$$\|f_h\|_\infty \leqslant \sqrt{\pi/2}\|h(\cdot) - \mathbb{E}[h(Z)]\|_\infty, \qquad \|f_h'\|_\infty \leqslant 2\|h(\cdot) - \mathbb{E}[h(Z)]\|_\infty.$$

Ako je h apsolutno neprekidna, tada

$$\|f_h\|_\infty \leqslant 2\|h'\|_\infty, \quad \|f_h'\|_\infty \leqslant \sqrt{\pi/2}\|h'\|_\infty, \quad \|f_h''\|_\infty \leqslant 2\|h'\|_\infty. \tag{1.19}$$

Uočimo, u slučaju kada je h apsolutno neprekidna (što sigurno jest ako je h Lipschitzova) imamo ogradu na f_h te njene prve dvije derivacije, dok smo u slučaju Kolmogorovljeve metrike u lemi 1.8 imali samo ogradu na f_z i njenu prvu derivaciju. Zbog toga je u pravilu lakše raditi s funkcijama f_h.

1.3 Sume nezavisnih slučajnih varijabli

Nakon što smo napravili osnovnu analizu svojstava rješenja Steinove jednažbe, fokus prebacujemo na različite kontekste u kojima možemo dobro ograničiti izraz $\mathbb{E}[f_h'(W) - Wf_h(W)]$, gdje je f_h ograničeno rješenje Steinove jednadžbe za izmjerivu funkciju h. U ovom poglavlju gledamo slučaj kada je W dana kao zbroj n nezavisnih slučajnih varijabli. Dokazati ćemo Lindenbergov centralni granični teorem te verziju Berry-Esseenovog teorema.

Prije svega, definirajmo notaciju koju ćemo koristiti kroz ostatak poglavlja. Neka su $\xi_1, \xi_2, \ldots, \xi_n$ nezavisne slučajne varijable s očekivanjem nula i konačnim varijancama za koje vrijedi $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$. Označimo

$$W = \sum_{i=1}^n \xi_i, \quad W^{(i)} = W - \xi_i$$

te definirajmo

$$K_i(t) = \mathbb{E}[\xi_i(\mathbb{1}_{\{0 \leqslant t \leqslant \xi_i\}} - \mathbb{1}_{\{\xi_i \leqslant t < 0\}})].$$

Iz definicije je jasno da je $K_i(t) \geqslant 0$ za sve realne t. Štoviše, zbog nenegativnosti, primjenom Fubinijevog teorema, dobivamo

$$\int_{-\infty}^{\infty} K_i(t)\mathrm{d}t = \mathbb{E}[\xi_i^2], \quad \int_{-\infty}^{\infty} |t|K_i(t)\mathrm{d}t = \frac{1}{2}\mathbb{E}|\xi_i|^3. \tag{1.20}$$

Neka je $h\colon \mathbb{R} \to \mathbb{R}$ izmjeriva funkcija za koju je $\mathbb{E}|h(Z)| < \infty$ te označimo s $f := f_h$ ograničeno rješenje Steinove jednadžbe. Cilj nam je ograničiti desnu stranu jednakosti

$$\mathbb{E}[h(W)] - \mathbb{E}[h(Z)] = \mathbb{E}[f'(W) - Wf(W)].$$

Jer su ξ_i i $W^{(i)}$ nezavisni i jer je $\mathbb{E}[\xi_i] = 0$, za svaki $i = 1, \ldots, n$, slijedi

$$\mathbb{E}[Wf(W)] = \sum_{i=1}^n \mathbb{E}[\xi_i f(W)] = \sum_{i=1}^n \mathbb{E}[\xi_i(f(W) - f(W^{(i)}))].$$

Koristeći Newton-Leibnizovu formulu, dobivamo

$$\begin{aligned}
\mathbb{E}[Wf(W)] &= \sum_{i=1}^n \mathbb{E}\left[\xi_i \int_0^{\xi_i} f'(W^{(i)} + t)\mathrm{d}t\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\int_{-\infty}^{\infty} f'(W^{(i)} + t)\xi_i(\mathbb{1}_{\{0 \leqslant t \leqslant \xi_i\}} - \mathbb{1}_{\{\xi_i \leqslant t < 0\}})\mathrm{d}t\right] \\
&= \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[f'(W^{(i)} + t)]K_i(t)\mathrm{d}t, \qquad (1.21)
\end{aligned}$$

gdje zadnja jednakost vrijedi zbog Fubinijevog teorema, definicije funkcije K_i i nezavisnosti varijabli ξ_i i $W^{(i)}$. Također, jer je

$$\sum_{i=1}^{n}\int_{-\infty}^{\infty}K_i(t)\mathrm{d}t=\sum_{i=1}^{n}\mathbb{E}[\xi_i^2]=1,$$

imamo

$$\mathbb{E}[f'(W)]=\sum_{i=1}^{n}\int_{-\infty}^{\infty}\mathbb{E}[f'(W)]K_i(t)\mathrm{d}t.$$

Spajanjem dvaju dobivenih jednakosti dobivamo

$$\mathbb{E}[f'(W)-Wf(W)]=\sum_{i=1}^{n}\int_{-\infty}^{\infty}\mathbb{E}[f'(W)-f'(W^{(i)}+t)]K_i(t)\mathrm{d}t. \tag{1.22}$$

Jednakost (1.22) će nam biti ključna u ogradama.

Sada možemo preći na dokaz centralnog graničnog teorema. Prvo ćemo dokazati slabiju verziju u kojoj pretpostavljamo postojanje trećih momenata, a nakon toga dokazati teorem sa slabijim pretpostavkama.

Teorem 1.11. *Neka su ξ_1, ξ_2, ..., ξ_n nezavisne slučajne varijable s očekivanjem nula i konačnim trećim momentom. Također, neka vrijedi $\sum_{i=1}^n\mathbb{E}[\xi_i^2]=1$. Tada, ako s δ označimo $\delta=3\sum_{i=1}^n\mathbb{E}|\xi_i|^3$, vrijedi*

$$|\mathbb{E}[h(W)]-\mathbb{E}[h(Z)]|\leqslant\delta\|h'\|_\infty,$$

gdje je $h\colon\mathbb{R}\to\mathbb{R}$ proizvoljna Lipschitzova funkcija, a $Z\sim N(0,1)$.

Dokaz. Iz (1.19) i (1.22), koristeći teorem srednje vrijednosti, imamo

$$\begin{aligned}|\mathbb{E}[f_h'(W)-Wf_h(W)]| &\overset{(1.22)}{\leqslant}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\mathbb{E}|f_h'(W)-f_h'(W^{(i)}+t)|K_i(t)\mathrm{d}t\\ &\leqslant\sum_{i=1}^{n}\int_{-\infty}^{\infty}\mathbb{E}[\|f_h''\|_\infty|W-(W^{(i)}+t)|]K_i(t)\mathrm{d}t\\ &\overset{(1.19)}{\leqslant}2\|h'\|_\infty\sum_{i=1}^{n}\int_{-\infty}^{\infty}\mathbb{E}[|t|+|\xi_i|]K_i(t)\mathrm{d}t.\end{aligned}$$

Koristeći (1.20), dobivamo

$$\begin{aligned}|\mathbb{E}[f_h'(W)-Wf_h(W)]| &\leqslant 2\|h'\|_\infty\sum_{i=1}^{n}\Big(\mathbb{E}|\xi_i|^3/2+\mathbb{E}|\xi_i|\mathbb{E}[\xi_i^2]\Big)\\ &\leqslant 3\|h'\|_\infty\sum_{i=1}^{n}\mathbb{E}|\xi_i|^3\end{aligned}$$

što dokazuje teorem.

□

Pretpostavka o postojanju trećih momenata nije potrebna, kao što vidimo u sljedećem teoremu.

Teorem 1.12. *Neka su ξ_1, ξ_2, ..., ξ_n nezavisne slučajne varijable s očekivanjem nula takve da je $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$. Ako su*

$$\beta_2 = \sum_{i=1}^n \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>1\}}], \quad \beta_3 = \sum_{i=1}^n \mathbb{E}|\xi_i|^3 \mathbb{1}_{\{|\xi_i|\leqslant 1\}}$$

i $\delta = 4(4\beta_2 + 3\beta_3)$, tada vrijedi

$$|\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]| \leqslant \delta \|h'\|_\infty,$$

za sve Lipschitzove funkcije $h\colon \mathbb{R} \to \mathbb{R}$, gdje je $Z \sim N(0,1)$.

Dokaz. Zbog leme 1.10, vrijedi

$$|f_h'(W) + f_h'(W^{(i)} + t)| \leqslant 2\|f_h'\|_\infty \leqslant 8\|h'\|_\infty.$$

Također, koristeći istu lemu i teorem srednje vrijednosti, imamo

$$|f_h'(W) + f_h'(W^{(i)} + t)| \leqslant \|f_h''\|_\infty |W - W^{(i)} - t| \leqslant 2\|h'\|_\infty (|\xi_i| + |t|).$$

Kombinirajući ove dvije ograde, slijedi

$$|f_h'(W) + f_h'(W^{(i)} + t)| \leqslant \|h'\|_\infty \min(8, 2(|t| + |\xi_i|)) \leqslant 8\|h'\|_\infty (|t| \wedge 1 + |\xi_i| \wedge 1),$$

gdje $a \wedge b$ standardno označava $\min(a,b)$.

Sada,

$$\begin{aligned} |\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]| &= |\mathbb{E}[f_h'(W) - W f_h(W)]| \\ &\overset{(1.22)}{\leqslant} \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[|f_h'(W) - f_h'(W^{(i)} + t)|] K_i(t) \mathrm{d}t \\ &\leqslant 8\|h'\|_\infty \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[|t| \wedge 1 + |\xi_i| \wedge 1] K_i(t) \mathrm{d}t. \end{aligned}$$

Jer je

$$\int_{-\infty}^{\infty} (|t| \wedge 1) \left(\mathbb{1}_{[0,x]}(t) - \mathbb{1}_{[-x,0)}(t)\right) \mathrm{d}t = \begin{cases} \frac{1}{2}|x|t^2 + |x|(|x| - 1) & \text{za } |x| > 1, \\ \frac{1}{2}|x|^3 & \text{za } |x| \leqslant 1, \end{cases}$$

puštajući $|t| \wedge 1$ unutar integrala u $K_i(t)$, dobivamo

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathbb{E}[|t| \wedge 1 + |\xi_i| \wedge 1] K_i(t) \mathrm{d}t \\
&\quad = \mathbb{E}[|\xi_i|(|\xi_i| - 1)\mathbb{1}_{\{|\xi_i|>1\}}] + \frac{1}{2}\mathbb{E}[|\xi_i(|\xi_i| \wedge 1)^2] + \mathbb{E}[\xi_i^2 \mathbb{E}[|\xi_i| \wedge 1]] \\
&\quad = \mathbb{E}[\xi_i^2 \mathbb{1}_{\{|\xi_i|>1\}}] - \frac{1}{2}\mathbb{E}[|\xi_i \mathbb{1}_{\{|\xi_i|>1\}}] + \frac{1}{2}\mathbb{E}[|\xi_i|^3 \mathbb{1}_{\{|\xi_i|\leqslant 1\}}] + \mathbb{E}[\xi_i^2 \mathbb{E}[|\xi_i| \wedge 1]].
\end{aligned}$$

Zbrajanjem po svim i, dobivamo

$$|\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]| \leqslant 8\|h'\|_\infty \left(\beta_2 + \frac{1}{2}\beta_3 + \sum_{i=1}^{n} \mathbb{E}[\xi_i^2]\mathbb{E}[|\xi_i| \wedge 1]\right).$$

Kako su i x^2 i $x \wedge 1$ rastuće funkcije za $x \geqslant 0$, slijedi da za proizvoljnu slučajnu varijablu ξ,

$$\mathbb{E}[\xi^2]\mathbb{E}[|\xi| \wedge 1] \leqslant \mathbb{E}[\xi^2(|\xi| \wedge 1)] = \mathbb{E}[|\xi|^3 \mathbb{1}_{\{|\xi|\leqslant 1\}}] + \mathbb{E}[\xi^2 \mathbb{1}_{\{|\xi|>1\}}]$$

pa vrijedi

$$|\mathbb{E}[h(W) - \mathbb{E}[h(Z)]| \leqslant \|h'\|_\infty \cdot 8 \left(\beta_2 + \frac{1}{2}\beta_3 + (\beta_2 + \beta_3)\right)$$

što je tražena tvrdnja.

□

Naposljetku, dokažimo Lindenbergov centralni granični teorem.

Teorem 1.13 (Lindenberg)**.** *Neka su $X_1, \dots, X_n$ nezavisne slučajne varijable s očekivanjem nula i konačnom varijancom. Označimo*

$$S_n = \sum_{i=1}^{n} X_i, \quad B_n^2 = \sum_{i=1}^{n} \mathbb{E}[X_i^2].$$

Ako vrijedi Lindenbergov uvjet

$$\frac{1}{B_n^2} \sum_{i=1}^{n} \mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|>\varepsilon B_n\}}] \to 0, \textit{ kad } n \to \infty, \textit{ za sve } \varepsilon > 0,$$

tada $S_n / B_n \xrightarrow{D} N(0,1)$.

Dokaz. Definirajmo

$$\xi_i = X_i/B_n, \quad W = S_n/B_n$$

te β_2 i β_3 kao u teoremu 1.12. Za $0 < \varepsilon < 1$,

$$\begin{aligned}
\beta_2 + \beta_3 &= \frac{1}{B_n^2}\sum_{i=1}^{n}\mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|>B_n\}}] + \frac{1}{B_n^3}\sum_{i=1}^{n}\mathbb{E}[|X_i|^3 \mathbb{1}_{\{|X_i|\leqslant B_n\}}] \\
&\leqslant \frac{1}{B_n^2}\sum_{i=1}^{n}\mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|>B_n\}}] + \frac{1}{B_n^3}\sum_{i=1}^{n}B_n\mathbb{E}[|X_i|^2 \mathbb{1}_{\{\varepsilon B_n\leqslant|X_i|\leqslant B_n\}}] \\
&+ \frac{1}{B_n^3}\sum_{i=1}^{n}\varepsilon B_n\mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|<\varepsilon B_n\}}] \\
&\leqslant \frac{1}{B_n^2}\sum_{i=1}^{n}\mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|>\varepsilon B_n\}}] + \varepsilon.
\end{aligned}$$

Sada, zbog Lindenbergovog uvjeta, puštanjem n u ∞, dobivamo

$$\beta_2 + \beta_3 \leqslant \varepsilon.$$

Kako je $0 < \varepsilon < 1$ bio proizvoljan, puštanjem ε u nulu, dobivamo da $\beta_2 + \beta_3 \to 0$ kad $n \to \infty$.

Koristeći teorem 1.12 imamo da vrijedi

$$|\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]| \leqslant 8\|h'\|_\infty(\beta_2 + \beta_3),$$

za svaku Lipschitzovu funkciju h. Zbog toga,

$$\mathrm{d}_W(W, Z) \lesssim \beta_2 + \beta_3.$$

Iz propozicije 1.2 slijedi

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(S_n/B_n \leqslant z) - \Phi(z)| \lesssim (\beta_2 + \beta_3)^{1/2} \to 0 \text{ kad } n \to \infty.$$

□

Pogledajmo sada poseban slučaj kada je $(\xi_n)_{n\in\mathbb{N}}$ niz nezavisnih jednako distribuiranih slučajnih varijabli s zajedničkom distribucijom ξ. Bez smanjenja općenitosti, pretpostavljamo da ξ ima očekivanje nula i varijancu jedan. U kontekstu prethodnog teorema, gledamo $X_i := \xi_i/\sqrt{n}$. Tada je $S_n = \sum_{i=1}^{n} \xi_i/\sqrt{n}$ te $B_n^2 = 1$ te vrijedi $S_n \xrightarrow{D} N(0,1)$. Zanima nas brzina kojom se ta konvergencija događa, u ovisnosti o n.

Za to je dovoljno pogledati koliko iznosi $\beta_2 + \beta_3$. Iz definicije,

$$\begin{aligned}\beta_2 + \beta_3 &= \frac{1}{B_n^2}\sum_{i=1}^n \mathbb{E}[X_i^2 \mathbb{1}_{\{|X_i|>B_n\}}] + \frac{1}{B_n^3}\sum_{i=1}^n \mathbb{E}[|X_i|^3 \mathbb{1}_{\{|X_i|\leqslant B_n\}}] \\ &= n\mathbb{E}\left[\frac{\xi^2}{n}\mathbb{1}_{\{|\xi/\sqrt{n}|>1\}}\right] + n\mathbb{E}\left[\frac{|\xi|^3}{n\sqrt{n}}\mathbb{1}_{\{|\xi/\sqrt{n}|\leqslant 1\}}\right] \\ &= \mathbb{E}[\xi^2 \mathbb{1}_{\{|\xi|>\sqrt{n}\}}] + \frac{1}{\sqrt{n}}\mathbb{E}[|\xi|^3 \mathbb{1}_{\{|\xi|\leqslant\sqrt{n}\}}].\end{aligned}$$

Ako još dodamo pretpostavku da je ξ ograničena slučajna varijabla, vidimo da za dovoljno velike n, prvi sumand je jednak nuli, dok je drugi sumand $\mathbb{E}|\xi|^3/\sqrt{n}$.

Dakle, brzina konvergencije u Wassersteinovoj metrici je reda $n^{-1/2}$, a u Kolmogorovljevoj $n^{-1/4}$. Kao što smo naveli nakon propozicije 1.2, ovakva ocjena za Kolmogorovljevu metriku nije optimalna i treba nam drugačiji pristup.

Prvo ćemo dokazati Berry-Esseenov teorem u slučaju kada su slučajne varijable koje promatramo ograničene.

Teorem 1.14. *Neka su ξ_1, ξ_2, ..., ξ_n nezavisne slučajne varijable s očekivanjem nula za koje vrijedi $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$ te neka su*

$$W^{(i)} = W - \xi_i = \sum_{j=1}^n \xi_j - \xi_i, \qquad K_i(t) = \mathbb{E}[\xi_i(\mathbb{1}_{\{0\leqslant t\leqslant \xi_i\}} - \mathbb{1}_{\{\xi_i\leqslant t<0\}})].$$

Tada vrijedi

$$\left|\sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{P}(W^{(i)} + t \leqslant z)K_i(t)\mathrm{d}t - \Phi(z)\right| \leqslant 2.44\sum_{i=1}^n \mathbb{E}|\xi_i|^3. \tag{1.23}$$

Dodatno, ako još vrijedi i $|\xi_i| \leqslant \delta_0$ za sve $i = 1, \dots, n$ za neku konstantu $\delta_0 > 0$, slijedi

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leqslant z) - \Phi(z)| \leqslant 3.3\delta_0.$$

Dokaz. Za $z \in \mathbb{R}$, označimo s f ograničeno rješenje f_z Steinove jednadžbe. Koristeći (1.21), slijedi

$$\begin{aligned}\mathbb{E}[Wf(W)] &= \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[f'(W^{(i)} + t)]K_i(t)\mathrm{d}t \\ &= \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[(W^{(i)} + t)f(W^{(i)} + t) + \mathbb{1}_{\{W^{(i)}+t\leqslant z\}} - \Phi(z)]K_i(t)\mathrm{d}t\end{aligned}$$

pri čemu u prijelazu iz prvog u drugi red koristimo činjenicu da je f rješenje Steinove jednadžbe.

Sređivanjem izraza i korištenjem da je $\sum_{i=1}^n \int_{-\infty}^{\infty} K_i(t)\mathrm{d}t = 1$ dobivamo

$$\begin{aligned}\sum_{i=1}^n \int_{-\infty}^{\infty} &\mathbb{P}(W^{(i)} + t \leqslant z)K_i(t)\mathrm{d}t - \Phi(t) \\ &= \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[Wf(W) - (W^{(i)} + t)f(W^{(i)} + t)]K_i(t)\mathrm{d}t.\end{aligned}$$

Sada, koristeći ocjene iz leme 1.8, konkretno nejednakost (1.13), vrijedi

$$\begin{aligned}\sum_{i=1}^n \mathbb{E} \int_{-\infty}^{\infty} &\left| Wf(W) - (W^{(i)} + t)f(W^{(i)} + t) \right| K_i(t)\mathrm{d}t \\ &= \sum_{i=1}^n \mathbb{E} \int_{-\infty}^{\infty} \left| (W^{(i)} + \xi_i)f(W^{(i)} + \xi_i) - (W^{(i)} + t)f(W^{(i)} + t) \right| K_i(t)\mathrm{d}t \\ &\leqslant \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[(|W^{(i)}| + \sqrt{2\pi}/4)(|\xi_i| + |t|)]K_i(t)\mathrm{d}t \\ &\leqslant (1 + \sqrt{2\pi}/4) \sum_{i=1}^n \int_{-\infty}^{\infty} (\mathbb{E}|\xi_i| + |t|)K_i(t)\mathrm{d}t,\end{aligned}$$

pri čemu zadnja nejednakost vrijedi jer su $W^{(i)}$ i ξ_i nezavisni te je $\mathbb{E}[{W^{(i)}}^2] \leqslant 1$.

Koristeći (1.20), slijedi

$$\begin{aligned}&\left| \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{P}(W^{(i)} + t \leqslant z)K_i(t)\mathrm{d}t - \Phi(z) \right| \\ &\qquad \leqslant (1 + \sqrt{2\pi}/4) \sum_{i=1}^n (\mathbb{E}|\xi_i|\mathbb{E}[\xi_i^2] + \frac{1}{2}\mathbb{E}|\xi_i|^3) \\ &\qquad \leqslant \frac{3}{2}(1 + \sqrt{2\pi}/4) \sum_{i=1}^n \mathbb{E}|\xi_i|^3 \leqslant 2.44 \sum_{i=1}^n \mathbb{E}|\xi_i|^3.\end{aligned}$$

Kada bi u gornjoj ocjeni mogli zamijeniti $\mathbb{P}(W^{(i)} + t \leqslant z)$ s $\mathbb{P}(W \leqslant z)$, uz (1.20), bili bi gotovi. Prvo, uočimo da je

$$\mathbb{P}(W \leqslant z) = \mathbb{P}(W^{(i)} + \xi_i \leqslant z) = \mathbb{P}(W^{(i)} \leqslant z - \xi_i).$$

Iz toga, direktno dobivamo

$$|\mathbb{P}(W^{(i)} + t \leqslant z) - \mathbb{P}(W \leqslant z)| \leqslant \mathbb{P}(z - \max\{\xi_i, t\} \leqslant W^{(i)} \leqslant z - \min\{\xi_i, t\}).$$

Ako su i $|t|$ i $|\xi_i|$ dovoljno mali, ocjena bi trebala biti dobra.

U našem slučaju, svi $|\xi_i|$ su orgraničeni s δ_0. Zbog toga, $K_i(t) = 0$ za $|t| > \delta_0$ pa je dovoljno promatrati gornju ogradu jedino u slučaju kada su i $|t|$ i $|\xi_i|$ ograničeni s δ_0. Posebno, imamo

$$\mathbb{P}(W^{(i)} + t \leqslant z) = \mathbb{P}(W - \xi_i + t \leqslant z) \begin{cases} \geqslant \mathbb{P}(W \leqslant z - 2\delta_0) \\ \leqslant \mathbb{P}(W \leqslant z + 2\delta_0). \end{cases}$$

Uvršatavajući $z + 2\delta_0$ umjesto z u gornju nejednakost te primjenom na prvu tvrdnju iz teorema, dobivamo

$$\begin{aligned} \mathbb{P}(W \leqslant z) - \Phi(z) &\leqslant \Phi(z + 2\delta) - \Phi(z) + \frac{3}{2}\left(1 + \sqrt{2\pi}/4)\right) \sum_{i=1}^{n} \mathbb{E}|\xi_i|^3 \\ &\leqslant \frac{2\delta_0}{\sqrt{2\pi}} + \frac{3}{2}(1 + \sqrt{2\pi}/4)\delta_0 \leqslant 3.3\delta_0. \end{aligned}$$

Potpuno analogno, uvrštavajući $z - 2\delta_0$ dobivamo drugi smjer tražene nejednakosti, čime smo dokazali teorem.

□

U prethodnom dokazu, ključno je bilo „zamijeniti" $\mathbb{P}(W^{(i)} + t \leqslant z)$ s $\mathbb{P}(W \leqslant z)$. To smo napravili koristeći činjenicu da je svaka od varijabli ξ_i uniformno ograničena s δ_0, no, Berry-Esseenov teorem će vrijediti i u generalnijem slučaju, što ćemo upravo i dokazati. Iako se dokaz može provesti i induktivno, koristiti ćemo koncentracijsku nejednakost koja će omogućiti spomenutu zamjenu.

Konstanta koju dobijemo ovim pristupom nije optimalna, no prednost je što se ovaj pristup može proširiti i na druge situacije, a ne samo na zbroj nezavisnih slučajnih varijabli.

Prije svega, iskažimo i dokažimo koncentracijsku nejednakost koju ćemo koristiti.

Lema 1.15. *Za sve realne $a < b$, i sve $1 \leqslant i \leqslant n$, vrijedi*

$$\mathbb{P}(a \leqslant W^{(i)} \leqslant b) \leqslant \sqrt{2}(b - a) + 2(\sqrt{2} + 1) \sum_{i=1}^{n} \mathbb{E}|\xi_i|^3.$$

Dokaz. Radi lakšeg zapisa, uvedimo oznaku $\gamma := \sum_{i=1}^{n} \mathbb{E}|\xi_i|^3$. Uzmimo $\delta = \gamma$ i definirajmo

$$f(w) = \begin{cases} -\frac{1}{2}(b - a) - \delta & \text{ako } w < a - \delta, \\ w - \frac{1}{2}(b + a) & \text{ako } a - \delta \leqslant w \leqslant b + \delta, \\ \frac{1}{2}(b - a) + \delta & \text{ako } w > b + \delta, \end{cases}$$

odnosno, neka je f takva da je $f' = \mathbb{1}_{[a-\delta,b+\delta]}$ te $\|f\| = \frac{1}{2}(b-a)+\delta$. Definirajmo i

$$\widehat{K}_j(t) := \xi_j\left(\mathbb{1}_{\{-\xi_j \leqslant t \leqslant 0\}} - \mathbb{1}_{\{0<t\leqslant -\xi_j\}}\right), \quad \widehat{K}(t) = \sum_{j=1}^{n} \widehat{K}_j(t).$$

Jer je ξ_j nezavisan od $W^{(i)} - \xi_j$ za sve $j \neq i$, ξ_i nezavisan od $W^{(i)}$ te $\mathbb{E}[\xi_j] = 0$, vrijedi

$$\begin{aligned}
\mathbb{E}[W^{(i)} f(W^{(i)})] - \mathbb{E}[\xi_i f(W^{(i)} - \xi_i)] &= \sum_{j=1}^{n} \mathbb{E}[\xi_j(f(W^{(i)}) - f(W^{(i)} - \xi_j))] \\
&= \sum_{j=1}^{n} \mathbb{E}\left[\xi_j \int_{-\xi_j}^{0} f'(W^{(i)}+t)\mathrm{d}t\right] \\
&= \sum_{j=1}^{n} \mathbb{E}\left[\int_{-\infty}^{\infty} f'(W^{(i)}+t)\widehat{K}_j(t)\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_{-\infty}^{\infty} f'(W^{(i)}+t)\widehat{K}(t)\mathrm{d}t\right].
\end{aligned}$$

Jer je $f' \geqslant 0$ te $\widehat{K} \geqslant 0$, očito vrijedi

$$\mathbb{E}\left[\int_{-\infty}^{\infty} f'(W^{(i)}+t)\widehat{K}(t)\mathrm{d}t\right] \geqslant \mathbb{E}\left[\int_{|t|\leqslant\delta} f'(W^{(i)}+t)\widehat{K}(t)\mathrm{d}t\right].$$

Iz definicije funkcije f, slijedi

$$\begin{aligned}
\mathbb{E}\left[\int_{|t|\leqslant\delta} f'(W^{(i)}+t)\widehat{K}(t)\mathrm{d}t\right] \geqslant \mathbb{E}\left[\mathbb{1}_{\{a\leqslant W^{(i)}\leqslant b\}} \int_{|t|\leqslant\delta} \widehat{K}(t)\mathrm{d}t\right] \\
\mathbb{E}\left[\mathbb{1}_{\{a\leqslant W^{(i)}\leqslant b\}} \sum_{j=1}^{n} |\xi_j| \min(\delta, |\xi_j|)\right] \geqslant H_{1,1} - H_{1,2},
\end{aligned}$$

uz oznake

$$\begin{aligned}
H_{1,1} &= \mathbb{P}(a \leqslant W^{(i)} \leqslant b) \sum_{j=1}^{n} \mathbb{E}[|\xi_j| \min(\delta, |\xi_j|)] \\
H_{1,2} &= \mathbb{E}\left|\sum_{j=1}^{n} \left(|\xi_j| \min(\delta, |\xi_j|) - \mathbb{E}|\xi_j| \min(\delta, |\xi_j|)\right)\right|.
\end{aligned}$$

Zadnja nejednakost vrijedi tako što dodamo i oduzmemo $\sum_{j=1}^{n} \mathbb{E}|\xi_j| \min(\delta, |\xi_j|)$, primijenimo da je $1 \geqslant \mathbb{1}_A$ za sve skupove A te iskoristimo nejednakost trokuta.

Sada, jer je

$$\min(x,y) \geqslant x - \frac{x^2}{4y}$$

za $x, y > 0$, slijedi

$$\sum_{j=1}^{n} \mathbb{E}|\xi_j| \min(\delta, |\xi_j|) \geqslant \sum_{j=1}^{n} \left(\mathbb{E}[\xi_j^2] - \frac{\mathbb{E}|\xi_j|^3}{4\delta} \right) = \frac{1}{2},$$

zbog odabira δ pa je

$$H_{1,1} \geqslant \frac{1}{2}\mathbb{P}(a \leqslant W^{(i)} \leqslant b).$$

Također, primjenom Cauchy-Schwarz nejednakosti, slijedi

$$\begin{aligned} H_{1,2} &\leqslant \left(\mathbb{E}\left[\left(\sum_{j=1}^{n} (|\xi_j| - \mathbb{E}|\xi_j|) \min(\delta, |\xi_j|) \right)^2 \right] \right)^{1/2} \\ &\leqslant \left(\mathbb{E}\left[\sum_{j=1}^{n} |\xi_j|^2 \min(\delta, |\xi_j|)^2 \right] \right)^{1/2} \\ &\leqslant \delta \left(\sum_{j=1}^{n} \mathbb{E}[\xi_j^2] \right)^{1/2} = \delta. \end{aligned}$$

Dakle, dokazali smo

$$\mathbb{E}\left[\int_{-\infty}^{\infty} f'(W^{(i)} + t)\widehat{K}(t)\mathrm{d}t \right] \geqslant \frac{1}{2}\mathbb{P}(a \leqslant W^{(i)} \leqslant b) - \delta.$$

U drugu ruku, jer je $\|f\| = \frac{1}{2}(b-a) + \delta$, vrijedi

$$\begin{aligned} \mathbb{E}[W^{(i)} f(W^{(i)}] - \mathbb{E}[\xi_i f(W^{(i)} - \xi_i)] &\leqslant \left(\frac{1}{2}(b-a) + \delta \right) \left(\mathbb{E}|W^{(i)}| + \mathbb{E}|\xi_i| \right) \\ &\leqslant \frac{1}{\sqrt{2}} \left((\mathbb{E}|W^{(i)}|)^2 + (\mathbb{E}|\xi_i|)^2 \right)^{1/2} (b - a + 2\delta) \\ &\leqslant \frac{1}{\sqrt{2}} \left(\mathbb{E}|W^{(i)}|^2 + \mathbb{E}|\xi_i|^2 \right)^{1/2} (b - a + 2\delta) \\ &\leqslant \frac{1}{\sqrt{2}} (b - a + 2\delta), \end{aligned}$$

pri čemu prva nejednakost vrijedi zbog vrijednosti $\|f\|$, druga zbog nejednakosti $2(a^2 + b^2) \geqslant (a+b)^2$, a treća zbog Cauchy-Schwarz nejednakosti.

Kombinirajući nejednakosti koje smo dokazali, dobivamo

$$\mathbb{P}(a \leqslant W^{(i)} \leqslant b) \leqslant \sqrt{2}(b-a) + 2(1+\sqrt{2})\delta = \sqrt{2}(b-a) + (1+\sqrt{2})\gamma,$$

što smo i željeli dokazati.

□

Nakon dokaza leme, dokaz Berry-Esseenovog teorema postaje skoro trivijalan.

Teorem 1.16 (Berry-Esseen)**.** *Neka su $\xi_1, \xi_2, \ldots, \xi_n$ nezavisne slučajne varijable s očekivanjem nula za koje vrijedi $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$. Tada za $W = \sum_{i=1}^n \xi_i$ vrijedi*

$$\sup_{z \in \mathbb{R}} |\mathbb{P}(W \leqslant z) - \Phi(z)| \leqslant 9.4\gamma, \quad \textit{gdje } \gamma = \sum_{i=1}^n \mathbb{E}|\xi_i|^3.$$

Dokaz. Koristeći iste oznake kao i prije u poglavlju, imamo

$$\begin{aligned}
&\left| \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{P}(W^{(i)} + t \leqslant z) K_i(t) \mathrm{d}t - \mathbb{P}(W \leqslant z) \right| \\
&\quad = \left| \sum_{i=1}^n \int_{-\infty}^{\infty} \Big(\mathbb{P}(W^{(i)} + t \leqslant z) - \mathbb{P}(W \leqslant z) \Big) K_i(t) \mathrm{d}t \right| \\
&\quad \leqslant \sum_{i=1}^n \int_{-\infty}^{\infty} |\mathbb{P}(W^{(i)} + t \leqslant z) - \mathbb{P}(W \leqslant z)| K_i(t) \mathrm{d}t \\
&\quad = \sum_{i=1}^n \int_{-\infty}^{\infty} |\mathbb{P}(W^{(i)} + t \leqslant z) - \mathbb{P}(W^{(i)} + \xi_i \leqslant z)| K_i(t) \mathrm{d}t \\
&\quad = \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[\mathbb{P}(z - (t \vee \xi_i) \leqslant W^{(i)} \leqslant z - (t \wedge \xi_i) \mid \xi_i)] K_i(t) \mathrm{d}t \\
&\quad \leqslant \sum_{i=1}^n \int_{-\infty}^{\infty} \mathbb{E}[\sqrt{2}(|t| + |\xi_i|) + 2(\sqrt{2}+1)\gamma] K_i(t) \mathrm{d}t \\
&\quad = \sqrt{2} \sum_{i=1}^n \left(\frac{1}{2} \mathbb{E}|\xi_i|^3 + \mathbb{E}|\xi_i| \mathbb{E}[\xi_i^2] \right) + 2(\sqrt{2}+1)\gamma \\
&\quad \leqslant (3.5\sqrt{2} + 2)\gamma \leqslant 6.95\gamma
\end{aligned}$$

pri čemu prvi redak slijedi iz (1.20), drugi iz nejednakosti trokuta, treći raspisivanjem, četvrti iz poznatog svojstva uvjetnog očekivanja, peti primjenom upravo dokazane leme, a zadnji opet iz (1.20).

U kombinaciji s ocjenom (1.23) iz teorema 1.14, koristeći nejednakost trokuta, slijedi

$$|\mathbb{P}(W \leqslant z) - \Phi(z)| \leqslant (6.95 + 2.44)\gamma \leqslant 9.4\gamma$$

čime smo dokazali teorem.

□

Opet, naglašavamo da dobivena konstanta 9.4 nije optimalna. Također, pretpostavka o postojanju trećih momenata nije nužna, nego se može dobiti ograda koja ovisi o konstantama β_2 i β_3 koje su definirane kao u Lindenbergovom centralnom graničnom teoremu.

Naposljetku, vratimo se opet na primjer sa sumom nezavisnih jednako distribuiranih slučajnih varijabli koji smo analizirali nakon dokaza Lindenbergovog centralnog graničnog teorema. U kontekstu upravo dokazanog Berry-Esseenovog teorema, gledamo $\xi_i = X_i/\sqrt{n}$, gdje su $X_1, \dots, X_n$ nezavisne jednako distribuirane slučajne varijable s očekivanjem nula i varijancom 1. Teorem nam govori

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(W \leqslant z) - \Phi(z)| \lesssim \sum_{i=1}^{n} \mathbb{E}\left|\frac{X_i}{\sqrt{n}}\right|^3 = \frac{\mathbb{E}|X_1|^3}{\sqrt{n}}.$$

Vidimo da je brzina konvergencije reda $n^{-1/2}$, što je bolje nego što smo prethodno dokazali.

1.4 Lokalna zavisnost

U prethodnom poglavlju, pogledali smo kako se Steinova metoda može primijeniti u proučavanju suma nezavisnih, ne nužno jednako distribuiranih, slučajnih varijabli te smo dokazali neke klasične rezultate. U ovom i sljedećih nekoliko poglavlja, promatrati ćemo druge, različite kontekste u kojima klasične metode teže prolaze, no Steinova metoda relativno jednostavno i elegantno vodi do eksplicitnih ograda udaljenosti od jedinične normalne distribucije.

Prije svega, definirajmo susjedstva zavisnosti slučajnih varijabli $X_1, \dots, X_n$.

Definicija 1.17. *Za slučajne varijable X_1, X_2, ..., X_n kažemo da imaju susjedstva zavisnosti $N_i \subseteq \{1, \dots, n\}$, $i = 1, 2, \dots, n$ ako je $i \in N_i$ te je X_i nezavisna od $\{X_j\}_{j \notin N_i}$.*

Prirodno je zamisliti graf s n vrhova, gdje vrh i predstavlja slučajnu varijablu X_i te postoji brid između vrhova i i j ako i samo ako su te dvije varijable zavisne, odnosno, ako je $j \in N_i$. Ključan rezultat je sljedeći.

Teorem 1.18. *Neka su X_1, X_2, ..., X_n slučajne varijable za koje vrijedi $\mathbb{E}[X_i^4] < \infty$, $\mathbb{E}[X_i] = 0$, za sve $i = 1, 2, \ldots n$. Definirajmo $\sigma^2 = \mathrm{Var}(\sum_i^n X_i)$ i $W = \sum_{i=1}^n X_i/\sigma$. Neka varijable $X_1, \ldots, X_n$ imaju susjedstva zavisnosti $N_i, i = 1, \ldots, n$ i neka je $D = \max_{1 \leqslant i \leqslant n} |N_i|$. Tada,*

$$d_W(W, Z) \leqslant \frac{D^2}{\sigma^3} \sum_{i=1}^n \mathbb{E}|X_i|^3 + \frac{\sqrt{28} D^{3/2}}{\sqrt{\pi}\sigma^2} \sqrt{\sum_{i=1}^n \mathbb{E}[X_i^4]}$$

gdje je $Z \sim N(0, 1)$.

Dokaz. Zbog (1.16) i leme 1.10, dovoljno je ocijeniti

$$|\mathbb{E}[f'(W) - Wf(W)]|, \quad \text{za } f \text{ takve da je } \|f\|_\infty, \|f''\|_\infty \leqslant 2, \|f'\|_\infty \leqslant \sqrt{2/\pi}.$$

Označimo $W_i = \sum_{j \notin N_i} X_i$ i uočimo da je X_i nezavisna od W_i. Uočimo da je

$$\begin{aligned} \mathbb{E}[f'(W) - Wf(W)] = \mathbb{E}\left[\frac{1}{\sigma}\sum_{i=1}^n X_i \left(f(W) - f(W_i) - (W - W_i)f'(W)\right)\right] \\ + \mathbb{E}\left[f'(W)\left(1 - \frac{1}{\sigma}\sum_{i=1}^n X_i(W - W_i)\right)\right] \end{aligned}$$

pri čemu smo samo dodali i oduzeli drugi član, a u prvom članu dodali nulu u obliku $\mathbb{E}[\sum_{i=1}^n X_i f(W_i)]$, jer su X_i i W_i nezavisne te $\mathbb{E}[X_i] = 0$. Primjenom nejednakosti trokuta, vidimo da je dovoljno ograničiti apsolutne vrijednosti članova na desnoj strani, što ćemo i napraviti.

Za ogradu prvog člana, prvo ćemo uočiti da izraz u zagradi možemo ocijeniti preko Taylorovog teorema, odnosno, vrijedi

$$f(W) - f(W_i) - (W - W_i)f'(W) = \frac{f''(W^*)}{2}(W - W_i)^2,$$

pri čemu je $W*$ između W i W_i. Zbog toga, nejednakosti trokuta, i prebacivanjem apsolutne vrijednosti unutar očekivanja, dobivamo da je gornja ograda za prvi izraz

$$\frac{\|f''\|_\infty}{2\sigma}\sum_{i=1}^n \mathbb{E}|X_i(W - W_i)^2| \leqslant \frac{1}{\sigma^3}\sum_{i=1}^n \mathbb{E}|X_i(\sum_{j \in N_i} X_j)^2| \leqslant \frac{1}{\sigma^3}\sum_{i=1}^n \sum_{j,k \in N_i} \mathbb{E}|X_i X_j X_k|.$$

Primjena nejednakosti između aritmetičke i geometrijske sredine, uz linearnost očekivanja, daje

$$\mathbb{E}|X_i X_j X_k| \leqslant \frac{1}{3}(\mathbb{E}|X_i|^3 + \mathbb{E}|X_j|^3 + \mathbb{E}|X_k|^3).$$

Naposljetku, jer je $|N_i| \leqslant D$ za sve i, svaki sumand $\mathbb{E}|X_i|^3$ se u sumi pojavi najviše $3D^2$ puta čime smo dobili ogradu

$$\mathbb{E}\left[\frac{1}{\sigma}\sum_{i=1}^{n} X_i\left(f(W) - f(W_i) - (W - W_i)f'(W)\right)\right] \leqslant \frac{D^2}{\sigma^3}\sum_{i=1}^{n}\mathbb{E}|X_i|^3$$

gdje na desnoj strani prepoznajemo prvi sumand ograde u teoremu.

Sada, ograničavamo drugi član. Opet, grubom ogradom na f', gornja ograda na drugi član je

$$\frac{\|f'\|_\infty}{\sigma^2}\mathbb{E}\left|\sigma^2 - \sum_{i=1}^{n} X_i \sum_{j\in N_i} X_j\right| \leqslant \frac{\sqrt{2}}{\sigma^2\sqrt{\pi}}\sqrt{\operatorname{Var}\left(\sum_{i=1}^{n}\sum_{j\in N_i} X_i X_j\right)}$$

gdje smo koristili da je

$$\sigma^2 = \mathbb{E}\left[\sum_{i=1}^{n} X_i \sum_{j\in N_i} X_j\right]$$

te smo primijenili Cauchy-Schwarzovu nejednakost. Naime, jednakost vrijedi jer je

$$\sigma^2 = \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i - \underbrace{\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]}_{=0}\right)^2\right]$$

te onda razmnožimo desnu stranu i iskoristimo da je $\mathbb{E}[X_i X_j] = 0$, za X_i i X_j nezavisne, jer je $\mathbb{E}[X_i] = 0$, za sve i.

Sada, ograničimo varijancu koja se nalazi unutar korijena. Raspišimo

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^{n}\sum_{j\in N_i} X_i X_j\right)^2\right] &= \sum_{i\neq j}\sum_{k\in N_i}\sum_{l\in N_j}\mathbb{E}[X_i X_j X_k X_l] \\ &\quad + \sum_{i=1}^{n}\sum_{j\in N_i}\mathbb{E}[X_i^2 X_j^2] + \sum_{i=1}^{n}\sum_{j\in N_i}\sum_{k\in N_i\setminus\{j\}}\mathbb{E}[X_i^2 X_j X_k].\end{aligned}$$

Opet, ograničavamo član po član. Primjenom nejednakosti između aritmetičke i geometrijske sredine, vrijedi

$$\sum_{i=1}^{n}\sum_{j\in N_i}\mathbb{E}[X_i^2 X_j^2] \leqslant \frac{1}{2}\sum_{i=1}^{n}\sum_{j\in N_i}\left(\mathbb{E}[X_i^4] + \mathbb{E}[X_j^4]\right) \leqslant D\sum_{i=1}^{n}\mathbb{E}[X_i^4].$$

Također, vrijedi i

$$\sum_{i=1}^{n}\sum_{j\in N_i}\sum_{k\in N_i\setminus\{j\}}\mathbb{E}[X_i^2X_jX_k] \leqslant \frac{1}{4}\sum_{i=1}^{n}\sum_{j\in N_i}\sum_{k\in N_i\setminus\{j\}}(2\mathbb{E}[X_i^4]+\mathbb{E}[X_j^4]+\mathbb{E}[X_k^4])$$
$$\leqslant D(D-1)\sum_{i=1}^{n}\mathbb{E}[X_i^4].$$

Naposljetku, rastavljamo sumu u dva dijela

$$\sum_{i\neq j}\sum_{k\in N_i}\sum_{l\in N_j}\mathbb{E}[X_iX_jX_kX_l] = \sum_{\{i,k\},\{j,l\}}\mathbb{E}[X_iX_k]\mathbb{E}[X_jX_l] + \sum_{\{i,j,k,l\}}\mathbb{E}[X_iX_jX_kX_l]$$

gdje prvi dio označava one indekse za koje je $\{X_i, X_k\}$ nezavisno od $\{X_j, X_l\}$, a drugi dio ostatak. Za drugi dio, opet primjenom nejednakosti između aritmetičke i geometrijske sredine, ograničavamo

$$\sum_{\{i,k,j,l\}}\mathbb{E}[X_iX_jX_kX_l] \leqslant 6D^3\sum_{i=1}^{n}\mathbb{E}[X_i^4]$$

s obzirom da u grafu u kojem je maksimalan stupanj D, postoji najviše $D\times 2D\times 3D$ četvorki koje zadovoljavaju zadana svojstva.

Koristeći da je

$$\sigma^4 = \mathbb{E}\left[\sum_{i=1}^{n}X_i\sum_{j\in N_i}X_j\right]^2,$$

radeći istu dekompoziciju kao i prije, dobivamo

$$\sigma^4 = \sum_{i\neq j}\sum_{k\in N_i}\sum_{l\in N_j}\mathbb{E}[X_iX_k]\mathbb{E}[X_jX_l]$$
$$+\sum_{i=1}^{n}\sum_{j\in N_i}\mathbb{E}[X_iX_j]^2+\sum_{i=1}^{n}\sum_{j\in N_i}\sum_{k\in N_i\setminus\{j\}}\mathbb{E}[X_iX_j]\mathbb{E}[X_iX_k].$$

Prvi član rastavljamo na sumu po $\{i,k\}, \{j,l\}$ i $\{i,j,k,l\}$ kao i gore te uočimo da je drugi član u sumi uvijek nenegativan, iz čega dobivamo da je

$$\sigma^4 - \sum_{\{i,k,j,l\}}\mathbb{E}[X_iX_k]\mathbb{E}[X_jX_l] - \sum_{i=1}^{n}\sum_{j\in N_i}\sum_{k\in N_i\setminus\{j\}}\mathbb{E}[X_iX_j]\mathbb{E}[X_iX_k]$$

gornja ograda za $\sum_{\{i,k\},\{j,l\}} \mathbb{E}[X_i X_k]\mathbb{E}[X_j X_l]$. Naposljetku, uzastopnim primjenama nejednakosti između arimetičke i geometrijske sredine dobivamo

$$\begin{aligned}
-\mathbb{E}[X_i X_k]\mathbb{E}[X_j X_l] &\leqslant \frac{1}{2}(\mathbb{E}[X_i X_k]^2 + \mathbb{E}[X_j X_l]^2) \\
&\leqslant \frac{1}{2}(\mathbb{E}[X_i^2 X_k^2] + \mathbb{E}[X_j^2 X_l^2]) \\
&\leqslant \frac{1}{4}(\mathbb{E}[X_i^4] + \mathbb{E}[X_j^4] + \mathbb{E}[X_k^4] + \mathbb{E}[X_l^4]).
\end{aligned}$$

Spajajući zajedno sve ocjene koje smo upravo dokazali, dobili smo

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n}\sum_{j\in N_i} X_i X_j\right) &= \mathbb{E}\left[\left(\sum_{i=1}^{n}\sum_{j\in N_i} X_i X_j\right)^2\right] - \sigma^4 \\
&\leqslant (12D^3 + 2D^2)\sum_{i=1}^{n}\mathbb{E}[X_i^4] \leqslant 14D^3\sum_{i=1}^{n}\mathbb{E}[X_i^4]
\end{aligned}$$

što je upravo tražena ograda iz tvrdnje teorema.

□

Iako smo u dokazu teorema ograničavali sve izraze do kraja, u nekim primjenama je često dovoljno, a i bolje direktno ograničavati izraze koji se pojavljuju u međukoracima, na primjer, varijancu koju smo ograničavali u drugom dijelu dokaza. Uočimo da u slučaju kada su $X_1, \ldots, X_n$ nezavisne i jednako distribuirane, maksimalan stupanj u grafu zavisnosti može rasti s n te će centralni granični teorem još uvijek vrijediti.

Primjer 1.19 (Broj trokuta u Erdős–Rényijevom grafu). *Neka je $G = G(n, p)$, Erdős–Rényijev graf na n vrhova, odnosno, neka je G slučajan graf s n vrhova gdje je svaki par vrhova $i \neq j$ povezan bridom s vjerojatnosti p, nezavisno od drugih bridova. Neka je T broj trokuta u grafu G, odnosno, ako raspišemo preko indikatora,*

$$T = \sum_{i=1}^{N} Y_i, \text{ gdje } N = \binom{n}{3} \text{ i } Y_i = \mathbb{1}\{i - \text{ti trokut postoji}\}$$

gdje smo proizvoljno numerirali trokute.

Zbog nezavisnosti pri odabiru bridova, postoji zavisnost između slučajnih varijabli Y_i i Y_j ako i samo ako dijele neki zajednički brid. Zbog toga, $|N_i| = 3(n-3)+1$ za sve

$i = 1, \dots, n$ pa primijenimo teorem 1.18 uz $X_i = Y_i - \mathbb{E}[Y_i] = Y_i - p^3$ i $D = 3n - 8$. S obzirom da je $\mathbb{P}(Y_i = 1) = p^3$, vrijedi

$$\begin{aligned}\mathbb{E}|X_i|^k &= \mathbb{P}(Y_i = 1) \cdot (1 - p^3)^k + \mathbb{P}(Y_i = 0) \cdot (p^3)^k \\ &= p^3 \cdot (1 - p^3)^k + (1 - p^3) \cdot p^{3k} \\ &= p^3(1 - p^3)[(1 - p^3)^{k-1} + p^{3(k-1)}]\end{aligned}$$

za sve $k \geqslant 1$.

Dakle, jedino preostaje izračunati $\sigma^2 = \mathrm{Var}(T)$. Znamo da je

$$\begin{aligned}\mathrm{Var}(T) = \mathrm{Var}\left(\sum_{i=1}^{n} Y_i\right) &= \mathbb{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^2\right] - \mathbb{E}\left[\sum_{i=1}^{n} Y_i\right]^2 \\ &= \sum_{i,j=1}^{n} \mathbb{E}[Y_i Y_j] - \binom{n}{3}^2 p^6,\end{aligned}$$

gdje u zadnjem redu koristimo linearnost očekivanja. Naravno, ovisno o indeksima i i j, vrijednost $\mathbb{E}[Y_i Y_j]$ će biti različita. Pogledajmo slučajeve.

- *Ako $i = j$, vrijedi $\mathbb{E}[Y_i Y_j] = p^3$ i takvih članova u sumi ima $\binom{n}{3}$.*
- *Ako trokuti Y_i i Y_j dijele točno dva vrha, da bi oba trokuta postojala, potrebno je 5 bridova pa je $\mathbb{E}[Y_i Y_j] = p^5$, a takvih članova ima $\binom{n}{4} \cdot \binom{4}{2} \cdot \binom{2}{1} = 12\binom{n}{4}$.*
- *Ako trokuti Y_i i Y_j dijele točno jedan vrh, da bi oba trokuta postojala, potrebno je 6 bridova pa je $\mathbb{E}[Y_i Y_j] = p^6$, a takvih članova ima $\binom{n}{5} \cdot \binom{5}{1} \cdot \binom{4}{2} = 30\binom{n}{5}$.*
- *Ako trokuti Y_i i Y_j ne dijele niti jedan vrh, da bi oba trokuta postojala, potrebno je 6 bridova pa je $\mathbb{E}[Y_i Y_j] = p^6$, a takvih članova ima $\binom{n}{6} \cdot \binom{6}{3} \cdot \binom{3}{3} = 20\binom{n}{6}$.*

Dakle, dobili smo da je

$$\begin{aligned}\sigma^2 = \mathrm{Var}(T) &= \binom{n}{3} p^3 + 12\binom{n}{4} p^5 + 30\binom{n}{5} p^6 + 20\binom{n}{6} p^6 - \binom{n}{3}^2 p^6 \\ &= \binom{n}{3} p^3(1 + 8p^3 - 9p^2 - 3np^3 + 3np^2).\end{aligned}$$

Primjenom teorema 1.18 dobivamo da za $W = (T - \mathbb{E}[T])/\sigma$ *i* $Z \sim N(0,1)$ *vrijedi*

$$\begin{aligned}
\mathrm{d}_W(W,Z) &\leqslant \frac{(3n-8)^2}{\sigma^3}\sum_{i=1}^{N}\mathbb{E}|X_i|^3 + \frac{\sqrt{26}(3n-8)^{3/2}}{\sqrt{\pi}\sigma^2}\sqrt{\sum_{i=1}^{N}\mathbb{E}|X_i|^4} \\
&\leqslant \frac{(3n-8)^2}{\sigma^3}\binom{n}{3}p^3(1-p^3)[(1-p^3)^2+p^6] \\
&\quad + \frac{\sqrt{26}(3n-8)^{3/2}}{\sqrt{\pi}\sigma^2}\sqrt{\binom{n}{3}p^3(1-p^3)[(1-p^3)^3+p^9]}
\end{aligned}$$

gdje je σ *eksplicitno raspisan u prethodnoj jednakosti. Ocjena koju smo upravo dokazali vrijedi za sve* $n \geqslant 3$ *i sve* $0 \leqslant p \leqslant 1$*. Analiziranjem ocjene možemo naći odnos između* p *i* n *za koji će ta ograda ići prema nula kada* n *ide u beskonačnost, odnosno, slučajeve u kojima će vrijediti centralni granični teorem. Ispostavi se da smo ovim načinom dobili samo podskup vrijednosti za koje će vrijediti centralni granični teorem, no, drugačijom analizom i ocjenama (npr. eksplicitno računajući vrijednosti koje se pojavljuju u međukoracima dokaza teorema 1.18), možemo dobiti bolje ocjene.*

1.5 Izmjenjivi parovi

U ovom poglavlju, proučavamo još jedan kontekst u kojem Steinova metoda daje zadovoljavajuće rezultate.

Definicija 1.20. *Uređeni par* (W, W') *slučajnih varijabli je izmjenjivi par ako vrijedi* $(W, W') \stackrel{d}{=} (W', W)$*. Dodatno, ako za neki* $0 < a \leqslant 1$ *izmjenjivi par* (W, W') *zadovoljava i jednakost*

$$\mathbb{E}[W' \mid W] = (1-a)W,$$

tada uređeni par (W, W') *zovemo a-Steinovim parom.*

Dokažimo neka osnovna svojstva Steinovih parova.

Propozicija 1.21. *Neka je* (W, W') *izmjenjivi par.*

1. *Ako je* $F\colon \mathbb{R}^2 \to \mathbb{R}$ *antisimetrična funkcija, to jest,* $F(x,y) = -F(y,x)$*, tada je* $\mathbb{E}[F(W,W)'] = 0$*.*

Ako je (W, W') *a-Steinov par s* $\mathrm{Var}(W) = \sigma^2$*, tada je*

2. $\mathbb{E}[W] = 0$ *i* $\mathbb{E}[(W'-W)^2] = 2a\sigma^2$*.*

Dokaz. Prva točka slijedi iz sljedećeg niza jednakosti

$$\mathbb{E}[F(W,W')] = \mathbb{E}[F(W',W)] = -\mathbb{E}[F(W,W')]$$

pri čemu prva jednakost vrijedi zbog izmjenjivosti, a druga zbog antisimetričnosti funkcije F.

Za drugu tvrdnju, prvo, uočimo da je

$$\mathbb{E}[W] = \mathbb{E}[W'] = \mathbb{E}[\mathbb{E}[W' \mid W]] = \mathbb{E}[(1-a)W] = (1-a)\mathbb{E}[W]$$

pri čemu prva jednakost vrijedi zbog izmjenjivosti, zbog čega je $\mathbb{E}[W] = 0$. Nadalje,

$$\begin{aligned}
\mathbb{E}[(W'-W)^2] &= \mathbb{E}[W'^2] + \mathbb{E}[W^2] - 2\mathbb{E}[W'W] \\
&= 2\sigma^2 - 2\mathbb{E}[\mathbb{E}[W'W \mid W]] \\
&= 2\sigma^2 - 2\mathbb{E}[W\mathbb{E}[W' \mid W]] \\
&= 2\sigma^2 - 2(1-a)\mathbb{E}[W^2] \\
&= 2a\sigma^2.
\end{aligned}$$

□

Kao i prije, ključan će nam biti teorem koji daje ogradu na Wassersteinovu udaljenost između neke slučajne varijable od interesa i jedinične normalne slučajne varijable.

Teorem 1.22. *Ako je (W,W') a-Steinov par uz $\mathbb{E}[W^2] = 1$ te $Z \sim N(0,1)$, tada je*

$$\mathrm{d}_W(W,Z) \leqslant \frac{\sqrt{\mathrm{Var}(\mathbb{E}[(W'-W)^2 \mid W])}}{\sqrt{2\pi}a} + \frac{\mathbb{E}|W'-W|^3}{3a}.$$

Dokaz. Neka je f ograničena funkcija s ograničenom prvom i drugom derivacijom i neka je $F(w) := \int_0^w f(t)\mathrm{d}t$. Sada, zbog izmjenjivosti para (W,W') i Taylorovog razvoja, vrijedi

$$\begin{aligned}
0 &= \mathbb{E}[F(W') - F(W)] \\
&= \mathbb{E}\left[(W'-W)f(W) + \frac{1}{2}(W'-W)^2 f'(W) + \frac{1}{6}(W'-W)^3 f''(W^*)\right],
\end{aligned}$$

gdje je W^* slučajna veličina između W i W'. Jer je (W,W') a-Steinov par, vrijedi

$$\begin{aligned}
\mathbb{E}[(W'-W)f(W)] &= \mathbb{E}[\mathbb{E}[(W'-W)f(W) \mid W]] \\
&= \mathbb{E}[f(W)\mathbb{E}[(W'-W) \mid W]] \\
&= -a\mathbb{E}[Wf(W)].
\end{aligned}$$

Kombinirajući prethodne dvije jednakosti, dobivamo

$$\mathbb{E}[Wf(W)] = \mathbb{E}\left[\frac{(W'-W)^2 f'(W)}{2a} + \frac{(W'-W)^3 f''(W^*)}{6a}\right].$$

Iz zadnje jednakosti, slijedi da je

$$|\mathbb{E}[f'(W) - Wf(W)]| \leqslant \|f'\|_\infty \mathbb{E}\left|1 - \frac{\mathbb{E}[(W'-W)^2 \mid W]}{2a}\right| + \|f''\|_\infty \frac{\mathbb{E}|W'-W|^3}{6a}$$

gdje smo na uvjetnu vjerojatnost u prvom članu na desnoj strani prešli na isti način kao i prije.

Sada, uz pretpostavku da je $\mathrm{Var}(W) = 1$, koristeći drugu činjenicu iz prethodne propozicije, vrijedi $\mathbb{E}[\mathbb{E}[(W'-W)^2 \mid W]] = 2a$. Primjenom Cauchy-Schwarzove nejednakosti na taj član, uz to da promatramo funkcije f takve da je $\|f'\|_\infty \leqslant \sqrt{2/\pi}$ i $\|f''\|_\infty \leqslant 2$, dobivamo traženu ogradu iz tvrdnje teorema.

□

Generalna strategija je za neku slučajnu varijablu W pronaći slučajnu varijablu W' na istom vjerojatnosnom prostoru takvu da (W, W') bude a-Steinov par. Pronalazak takve slučajne varijable često nije jednostavan. Također, često je korisno koristiti činjenicu da je

$$\mathrm{Var}(\mathbb{E}[(W'-W)^2 \mid W]) \leqslant \mathrm{Var}(\mathbb{E}[(W'-W)^2 \mid \mathcal{F}]),$$

za sve σ-algebre $\mathcal{F}$ koje sadrže $\sigma(W)$. Ovo nam omogućava da uvjetujemo na veće σ-algebre, za koje je lakše računati uvjetno očekivanje. Dokaz ove činjenice nije pretjerano kompliciran. Raspisivanjem definicije varijance i korištenjem svojstva uvjetnog očekivanja, dobijemo da je ekvivalentno dokazati

$$\mathbb{E}[\mathbb{E}[(W'-W)^2 \mid W]^2] \leqslant \mathbb{E}[\mathbb{E}[(W'-W)^2 \mid \mathcal{F}]^2].$$

U svrhu lakšeg zapisa, označimo $X = (W'-W)^2$, $Y = \mathbb{E}[X \mid \mathcal{F}]$. Opet, iz svojstva uvjetnog očekivanja, znamo da je $\mathbb{E}[X \mid W] = \mathbb{E}[Y \mid W]$, jer je $\sigma(W) \subseteq \mathcal{F}$. Dakle, dokazujemo

$$\mathbb{E}[\mathbb{E}[Y \mid W]^2] \leqslant \mathbb{E}[Y^2],$$

no to je posljedica Jensenove nejednakosti i činjenice da je $\mathbb{E}[Y \mid W] = \mathbb{E}[Y]$.

Primjer 1.23 (Centralni granični teorem)**.** *Neka su X_1, ..., X_n nezavisne slučajne varijable koje zadovoljavaju $\mathbb{E}[X_i^4] < \infty$, $\mathbb{E}[X_i] = 0$ i $\mathrm{Var}(X_i) = 1$, za sve i. Definirajmo $W = n^{-1/2}\sum_{i=1}^n X_i$. Neka je I slučajno, uniformno odabran indeks iz $\{1,\dots,n\}$, $(X_1',\dots,X_n')$ nezavisna kopija vektora $(X_1,\dots,X_n)$ i definirajmo*

$$W' = W - \frac{X_I}{\sqrt{n}} + \frac{X_I'}{\sqrt{n}}.$$

Lagano se vidi da je (W, W') *izmjenjiv par. Dokažimo da je također i* $1/n$*-Steinov par. Računamo*

$$\begin{aligned}\mathbb{E}[W' - W \mid (X_1, \dots, X_n)] &= \frac{1}{\sqrt{n}} \mathbb{E}[X'_I - X_I \mid (X_1, \dots, X_n)] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{n} \mathbb{E}[X'_i - X_i \mid (X_1, \dots, X_n)] \\ &= -\frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{\sqrt{n}} = -\frac{W}{n},\end{aligned}$$

gdje u drugom redu uvjetujemo po I*, a u zadnjem redu koristimo nezavisnost od* X' *i* X *te činjenicu da je* $\mathbb{E}[X'_i] = 0$ *za sve* i*.*

Sada, jer je $\sigma(W) \subseteq \sigma(X_1, \dots, X_n)$*, koristeći svojstvo uvjetnog očekivanja, dobivamo* $\mathbb{E}[W' - W \mid W] = -W/n$*, kao što smo i htjeli dokazati.*

Sada možemo primjeniti teorem 1.22. Ocijenimo

$$\mathbb{E}|W' - W|^3 = \frac{1}{n^{3/2}} \sum_{i=1}^{n} \mathbb{E}|X_i - X'_i|^3 \leqslant \frac{8}{n^{3/2}} \sum_{i=1}^{n} \mathbb{E}|X_i|^3,$$

gdje smo koristili nejednakost između arimetičke i geometrijske sredine u obliku

$$(a + b)^3 = a^3 + 3a^2 b + 3ab^2 + b^3 \leqslant 4a^3 + 4b^3$$

i činjenicu da su X_i *i* X'_i *jednako distribuirani.*

Za drugi član, računamo

$$\begin{aligned}\mathbb{E}[(W' - W)^2 \mid (X_1, \dots, X_n)] &= \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{E}[(X'_i - X_i)^2 \mid (X_1, \dots, X_n)] \\ &= \frac{1}{n^2} \sum_{i=1}^{n} (1 + X_i^2)\end{aligned}$$

gdje smo u prvom jednakosti opet uvjetovali po I*, a u drugoj koristili nezavisnost* X' *i* X *te činjenicu da je* $\mathbb{E}[X_i'^2] = 1$*. Sada, uzimajući varijancu na obje strane, vrijedi*

$$\mathrm{Var}(\mathbb{E}[(W' - W)^2 \mid (X_1, \dots, X_n)]) \leqslant \frac{1}{n^4} \sum_{i=1}^{n} \mathbb{E}[X_i^4]$$

pa koristeći napomenu nakon dokaza teorema 1.22 smo dobili

$$\mathrm{Var}(\mathbb{E}[(W' - W)^2 \mid W]) \leqslant \frac{1}{n^4} \sum_{i=1}^{n} \mathbb{E}[X_i^4]$$

Kombinirajući upravo dokazane ograde, uz primjenu teorema 1.22, dobili smo

$$\mathrm{d}_W(W,Z) \leqslant \sqrt{\frac{2}{\pi}}\frac{\sqrt{\sum_{i=1}^{n}\mathbb{E}[X_i^4]}}{2n} + \frac{2}{3n}\sum_{i=1}^{n}\mathbb{E}|X_i|^3.$$

U slučaju da su X_i nezavisne i jednako distribuirane, vidimo da je ocjena reda veličine $n^{-1/2}$, što je upravo najbolje moguće. Također, pretpostavka o četvrtim momentima nije potrebna ako se odlučimo na ocjenjivanje izraza koji se pojavi kao međukorak u dokazu teorema.

1.6 Veličinom pristrano sparivanje

U ovom poglavlju promatramo kako ocijeniti Wassersteinovu udaljenosti između nenegativne slučajne varijable X i jedinične normalne slučajne varijable, koristeći veličinom pristran par.

Definicija 1.24. *Za nenegativnu slučajnu varijablu $X \geqslant 0$ s $\mathbb{E}[X] = \mu < \infty$, kažemo da X^s ima veličinom pristranu distribuciju s obzirom na X ako za sve funkcije $f\colon \mathbb{R} \to \mathbb{R}$ takve da je $\mathbb{E}|Xf(X)| < \infty$ vrijedi*

$$\mathbb{E}[Xf(X)] = \mu\mathbb{E}[f(X^s)].$$

Ono što se odmah vidi iz definicije, raspisujući oba očekivanja, jest da ako su F i F^s redom funkcije distribucije za X i X^s, vrijedi

$$\mathrm{d}F^s(x) = \frac{x}{\mu}\mathrm{d}F(x).$$

Posebno, ako X ima gustoću f s obzirom na Lebesgueovu mjeru, ima i X^s i ta gustoća je dana s $f^s(x) := \frac{x}{\mu}f(x)$. Također, ako je X diskretna slučajna varijabla s vrijednostima u $\mathbb{N}_0$, vrijedi

$$\mathbb{P}(X^s = k) = \frac{k}{\mu}\mathbb{P}(X = k).$$

Opet nam je ključan teorem koji daje ocjenu na Wassersteinovu udaljenost od jedinične normalne distribucije.

Teorem 1.25. *Neka je $X \geqslant 0$ nenegativna slučajna varijabla takva da je $\mathbb{E}[X] = \mu < \infty$ i $\mathrm{Var}(X) = \sigma^2 < \infty$. Neka je X^s slučajna varijabla na istom vjerojatnosnom prostoru kao i*

X i neka ima veličinom pristranu distribuciju s obzirom na X. Ako je $W = (X-\mu)/\sigma$ i $Z \sim N(0,1)$, vrijedi

$$d_W(W,Z) \leqslant \frac{\mu}{\sigma^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\mathbb{E}[X^s - X \mid X])} + \frac{\mu}{\sigma^3}\mathbb{E}[(X^s-X)^2].$$

Dokaz. Direktnim raspisivanjem iz definicija dobivamo

$$\mathbb{E}[Wf(W)] = \mathbb{E}\left[\frac{X-\mu}{\sigma} f\left(\frac{X-\mu}{\sigma}\right)\right] = \frac{\mu}{\sigma}\left(f\left(\frac{X^s-\mu}{\sigma}\right) - f\left(\frac{X-\mu}{\sigma}\right)\right).$$

Kao i inače, zanimaju nas ograničene funkcije f s ograničene prve dvije derivacije. Taylorov razvoj izraza na desnoj strani nam daje

$$\mathbb{E}[Wf(W)] = \frac{\mu}{\sigma}\mathbb{E}\left[\frac{X^s-X}{\sigma} f'\left(\frac{X-\mu}{\sigma}\right) + \frac{(X^s-X)^2}{2\sigma^2} f''\left(\frac{X^*-\mu}{\sigma}\right)\right],$$

gdje je X^* slučajna veličina između X i X^s. Iz nejednakosti trokuta slijedi

$$\begin{aligned} |\mathbb{E}[f'(W) - Wf(W)]| \leqslant & \left|\mathbb{E}\left[f'(W)\left(1 - \frac{\mu}{\sigma^2}(X^s - X)\right)\right]\right| + \\ & + \frac{\mu}{2\sigma^3}\left|\mathbb{E}\left[f''\left(\frac{X^*-\mu}{\sigma}\right)(X^s-X)^2\right]\right|. \end{aligned}$$

S obzirom da samo promatramo funkcije f takve da je $\|f'\|_\infty \leqslant \sqrt{2/\pi}$ i $\|f''\|_\infty \leqslant 2$, lagano se vidi da je drugi član na desnoj strani ograničen drugim sumandom iz tvrdnje teorema, a prvi član je ograničen sa

$$\sqrt{\frac{2}{\pi}}\mathbb{E}\left|1 - \frac{\mu}{\sigma^2}\mathbb{E}[X^s - X \mid X]\right| \leqslant \frac{\mu}{\sigma^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\mathbb{E}[X^s - X \mid X])},$$

pri čemu smo koristili Cauchy-Schwarz nejednakost i činjenicu da je $\mathbb{E}[X^s] = (\sigma^2 + \mu^2)/\mu$.

□

Preostaje pitanje kako za slučajnu varijablu $X = \sum_{i=1}^n X_i$ konstruirati odgovarajuću veličinom pristranu slučajnu varijablu X^s. Pretpostavljamo da su X_i nenegativne slučajne varijable s konačnim očekivanjem $\mu_i = \mathbb{E}[X_i]$. Konstrukciju radimo kroz sljedeće korake.

1. Za svaki $i = 1, 2, \ldots, n$, neka je X_i^s veličinom pristrana slučajna varijabla s obzirom na X_i, nezavisna od $(X_j)_{j \neq i}$ i $(X_j^s)_{j \neq i}$. Definiramo slučajni vektor $(X_j^{(i)})_{j \neq i}$ tako da je njegova distribucija, uvjetno na $X_i^s = x$, jednaka distribuciji vektora $(X_j)_{j \neq i}$ uz uvjet $X_i = x$.

2. Odaberemo slučajan indeks $I \in \{1, \dots, n\}$, uz distribuciju $\mathbb{P}(I = i) = \mu_i/\mu$ gdje je $\mu = \mathbb{E}[X]$.

3. Definiramo $X^s = \sum_{j \neq I} X_j^{(I)} + X_I^s$.

U slučaju kada su X_i^s neprekidne slučajne varijable, događaj $X_i^s = x$ je vjerojatnosti nula pa uvjetujemo na događaje vjerojatnosti nula. Da bi to mogli napraviti, u slučaju kada slučajne varijable nemaju gustoću, koristimo regularne uvjetne distribucije, za koje znamo da postoje jer radimo na potpunom separabilnom metričkom prostoru, zajedno s Borel izmjerivim skupovima. Za više detalja, pogledati 5. poglavlje u [6].

Propozicija 1.26. *Neka je $X = \sum_{i=1}^n X_i$, pri čemu su $X_i \geqslant 0$, $\mathbb{E}[X_i] = \mu_i < \infty$ te $\mu = \mathbb{E}[X] = \sum_{i=1}^n \mu_i$. Ako je X^s konstruirana kao gore, tada X^s ima veličinom pristranu distribuciju s obzirom na X.*

Dokaz. Neka je $\mathbf{X} = (X_1, \dots, X_n)$ i za $i = 1, \dots, n$, neka je $\mathbf{X}^i$ vektor kojem je j-ta koordinata jednaka $X_j^{(i)}$ za $j \neq i$, a i-ta koordinata jednaka X_i^s. Da bi dokazali propoziciju, dovoljno je dokazati

$$\mathbb{E}[Xf(\mathbf{X})] = \mu \mathbb{E}[f(\mathbf{X}^I)]$$

za $f\colon \mathbb{R}^n \to \mathbb{R}$ za koje je $\mathbb{E}|Xf(\mathbf{X})| < \infty$. Ako na desnoj strani uvjetujemo na I i raspišemo, uočavamo da je dovoljno dokazati da je

$$\mathbb{E}[X_i f(\mathbf{X})] = \mu_i \mathbb{E}[f(\mathbf{X}^i)], \quad i = 1, \dots, n.$$

Neka je $h(X_i) := \mathbb{E}[f(\mathbf{X}) \mid X_i]$. Tada

$$\mathbb{E}[X_i f(\mathbf{X})] = \mathbb{E}[X_i h(X_i)] = \mu_i \mathbb{E}[h(X_i^s)],$$

što je upravo ono što smo htjeli dokazati.

□

Dvije jednostavne posljedice prethodne propozicije su sljedeće.

- Ako su slučajne varijable $X_1, \dots, X_n$ nezavisne, tada $X^s = X - X_I + X_I^s$ ima veličinom pristranu distribuciju s obzirom na X.

- Ako $X_1, \dots, X_n$ poprimaju vrijednosti nula ili jedan, tada $X^s = \sum_{j \neq I} X_j^{(I)} + 1$ ima veličinom pristranu distribuciju s obzirom na X.

Prva činjenica slijedi jer zbog nezavisnosti kao slučajni vektor $(X_j^{(i)})_{j\neq i}$ možemo koristiti upravo $(X_j)_{j\neq i}$. Druga činjenica slijedi jer je za bilo koju Bernoullijevu slučajnu varijablu X, $X^s = 1$.

Koristeći prvu činjenicu, na još jedan način možemo ocijeniti Wassersteinovu udaljenost od jedinične normalne distribucije za sumu nezavisnih slučajnih varijabli. Zanimljiviji primjer je sljedeći.

Primjer 1.27 (Broj izoliranih vrhova u Erdős–Rényijevom grafu)**.** *Neka je dan $G = G(n, p)$ Erdős–Rényijev graf i neka je X_i indikator da je vrh v_i izoliran, odnosno, stupnja nula. Tada je $X = \sum_{i=1}^n X_i$ ukupan broj izoliranih vrhova grafa G. Koristimo teorem 1.25 da bi ocijenili Wassersteinovu udaljenost između jedinične normalne distribucije i slučajne varijable $W = (X - \mu)/\sigma$, gdje je $\mu = \mathbb{E}[X]$ i $\sigma^2 = \operatorname{Var}(X)$.*

Potrebno je konstruirati slučajnu varijablu X^s, koja će imati veličinom pristranu distribuciju s obzirom na X. Prethodna propozicija i diskusija nakon nje nam govori da to radimo na sljedeći način. Prvo uniformno odaberemo indeks $I \in \{1, \dots, n\}$, onda stavimo $X_I = 1$ te prilagodimo ostale sumande uvjetujući na $X_I = 1$. S obzirom da $X_I = 1$ znači da je vrh v_I izoliran, to dobivamo tako što obrišemo sve bridove iz v_I. Posebno, G uz uvjet da je $X_I = 1$ je samo Erdős–Rényijev graf na preostalih $n - 1$ vrhova. Dakle, X^s je broj izoliranih vrhova u G nakon što smo obrisali sve bridove iz v_I.

Kako bi primijenili teorem 1.25, potrebno je izračunati $\mathbb{E}[X]$, $\operatorname{Var}(X)$, $\operatorname{Var}(\mathbb{E}[X^s - X \mid X])$ te $\mathbb{E}[(X^s - X)^2]$. Računamo redom. S obzirom da je vjerojatnost da je neki dani vrh izoliran jednaka $(1 - p)^{n-1}$, vrijedi

$$\mu = \mathbb{E}[X] = n \cdot (1 - p)^{n-1}.$$

Također, sve $X_1, \dots, X_n$ su jednako distribuirane pa vrijedi

$$\begin{aligned} \sigma^2 = \operatorname{Var}(X) &= \mu(1 - (1 - p)^{n-1}) + n(n - 1) \operatorname{Cov}(X_1, X_2) \\ &= \mu(1 + (np - 1)(1 - p)^{n-2}), \end{aligned}$$

jer je $\mathbb{E}[X_1 X_2] = (1 - p)^{2n-3}$. Označimo s d_i stupanj vrha v_i u grafu G i neka je D_i broj vrhova povezanih s v_i kojima je stupanj točno jedan. Tada je

$$X^s - X = D_I + \mathbb{1}[d_I > 0]$$

pa vrijedi

$$\begin{aligned} \operatorname{Var}(\mathbb{E}[X^s - X \mid G]) &= \frac{1}{n^2} \operatorname{Var}\left(\sum_{i=1}^n (D_i + \mathbb{1}[d_i > 0])\right) \\ &\leqslant \frac{2}{n^2}\left[\operatorname{Var}\left(\sum_{i=1}^n D_i\right) + \operatorname{Var}\left(\sum_{i=1}^n \mathbb{1}[d_i > 0]\right)\right] \end{aligned}$$

gdje prvu jednakost dobivamo tako što smo uvjetovali na I. Jer je $\sum_{i=1}^{n} \mathbb{1}[d_i > 0] = n - X$, drugi član na desnoj strani nejednakosti je jednak $\mathrm{Var}(X)$. *Uočimo da je $\sum_{i=1}^{n} D_i$ broj vrhova u G koji imaju stupanj jedan, što se može napisati i kao $\sum_{i=1}^{n} Y_i$, gdje je Y_i indikator da je stupanj vrha v_i u G jednak jedan. Slijedi*

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} D_i\right) &= n(n-1)p(1-p)^{n-2}(1-(n-1)p(1-p)^{n-2}) \\ &+ n(n-1)\,\mathrm{Cov}(Y_1, Y_2) \\ &= n(n-1)p(1-p)^{n-2}[1-(n-1)p(1-p)^{n-2} \\ &+ (1-p)^{n-2} + (n-1)^2p^2(1-p)^{n-3}],\end{aligned}$$

jer je $\mathbb{E}[Y_1Y_2] = p(1-p)^{2n-4} + (n-1)^2p^2(1-p)^{2n-5}$.

Naposljetku, računamo

$$\begin{aligned}\mathbb{E}[(X^s - X)^2] = \mathbb{E}[\mathbb{E}[(X^s - X)^2 \mid X]] &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[(D_i + \mathbb{1}[d_i > 0])^2] \\ &\leqslant \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[(D_i + 1)^2] = \mathbb{E}[D_1^2] + 2\mathbb{E}[D_1] + 1.\end{aligned}$$

Ako D_1 napišemo kao sumu indikatora, pokaže se da je

$$\mathbb{E}[D_1^2] = (n-1)p(1-p)^{n-2} + (n-1)(n-2)p^2(1-p)^{2n-5}.$$

Jer je $D_1 \leqslant D_1^2$ gotovo sigurno, dobili smo sve tražene ograde i možemo primijeniti teorem 1.25. Konkretno, dokazali smo sljedeći teorem.

Teorem 1.28. *Ako je X broj izoliranih vrhova u Erdős–Rényijevom grafu $G(n, p)$, $W = (X - \mu)/\sigma$ i za neki $1 \leqslant \alpha < 2$ vrijedi $\lim_{n\to\infty} n^\alpha p = c \in (0, \infty)$, tada*

$$\mathrm{d}_W(W, Z) \leqslant \frac{C}{\sigma},$$

za neku konstantu C.

1.7 Nula pristrano sparivanje

U prethodnom poglavlju promatrali smo veličinom pristrane distribucije, koje su bile definirane samo za nenegativne slučajne varijable. Sada promatramo nula pristrane distribucije koje su definirane za sve slučajne varijable kojima je očekivanje nula.

Definicija 1.29. *Neka je W slučajna varijabla za koju je $\mathbb{E}[W] = 0$ i $\mathrm{Var}(W) = \sigma^2 < \infty$. Kažemo da slučajna varijabla W^z ima nula pristranu distribuciju s obzirom na W ako za sve apsolutno neprekidne funkcije f takve da je $\mathbb{E}|Wf(W)| < \infty$ vrijedi*

$$\mathbb{E}[Wf(W)] = \sigma^2 \mathbb{E}[f'(W^z)].$$

Ako promatramo preslikavanje koje slučajnoj varijabli s očekivanjem nula i konačnom varijancom pridruži njenu odgovarajuću nula pristranu slučajnu varijablu, vidimo da je fiksna točka tog preslikavanja upravo normalna distribucija. Zbog toga, heuristički očekujemo da „skoro fiksne točke" tog preslikavanja budu približno normalne. Formalno je to opisano u sljedećem teoremu.

Teorem 1.30. *Neka je W slučajna varijabla s očekivanjem nula i varijancom jedan. Neka je W^z njena odgovarajuća nula pristrana slučajna varijabla definirana na istom vjerojatnosnom prostru. Ako je $Z \sim N(0,1)$, tada*

$$\mathrm{d}_W(W, Z) \leqslant 2\mathbb{E}|W^z - W|.$$

Dokaz. Neka je $\mathcal{F}$ familija svih funkcija $f \colon \mathbb{R} \to \mathbb{R}$ takvih da je $\|f'\|_\infty \leqslant \sqrt{2/\pi}$ i $\|f\|_\infty, \|f''\|_\infty \leqslant 2$. Tada

$$\begin{aligned}
\mathrm{d}_W(W, Z) &\leqslant \sup_{f \in \mathcal{F}} |\mathbb{E}[f'(W) - Wf(W)]| \\
&= \sup_{f \in \mathcal{F}} |\mathbb{E}[f'(W) - f'(W^z)]| \\
&= \sup_{f \in \mathcal{F}} \|f''\|_\infty \mathbb{E}|W - W^z|,
\end{aligned}$$

što je upravo tražena tvrdnja.

□

Dokažimo neka osnovna svojstva nula pristrane distribucije.

Propozicija 1.31. *Neka je W slučajna varijabla takva da je $\mathbb{E}[W] = 0$ i $\mathrm{Var}(W) = \sigma^2 < \infty$.*

1. *Postoji jedinstvena vjerojatnosna distribucija i slučajna varijabla W^z s tom distribucijom za koju vrijedi*
$$\mathbb{E}[Wf(W)] = \sigma^2 \mathbb{E}[f'(W^z)]$$
za sve apsolutno neprekidne funkcije f takve da je $\mathbb{E}|Wf(W)| < \infty$.

2. *Distribucija slučajne varijable W^z definirana u prethodnoj točki je apsolutno neprekidna s obzirom na Lebesgueovu mjeru s gustoćom*

$$p^z(w) = \sigma^{-2}\mathbb{E}[W\mathbb{1}[W > w]] = -\sigma^{-2}\mathbb{E}[W\mathbb{1}[W \leqslant w]].$$

Dokaz. Pretpostavimo da je $\sigma^2 = 1$. Generalni slučaj se rješava na sličan način. Dokazati ćemo obje točke istovremeno, dokazujući da je p^z uistinu funkcija gustoće koja definira distribuciju koja zadovoljava prvu točku.

Prvo promatramo $f(x) = \int_0^x g(t)\mathrm{d}t$ za neku nenegativnu funkciju g integrabilnu na kompaktima. Tada

$$\begin{aligned}\int_0^\infty f'(u)\mathbb{E}[W\mathbb{1}[W > u]]\mathrm{d}u &= \int_0^\infty g(u)\mathbb{E}[W\mathbb{1}[W > u]]\mathrm{d}u \\ &= \mathbb{E}\left[W\int_0^{\max\{0,W\}} g(u)\mathrm{d}u\right] = \mathbb{E}[Wf(W)\mathbb{1}[W \geqslant 0]].\end{aligned}$$

Koristeći isti argument, vrijedi $\int_{-\infty}^0 f'(u)p^z(u)\mathrm{d}u = \mathbb{E}[Wf(W)\mathbb{1}[W \leqslant 0]]$. Upravo dokazane jednakosti daju

$$\int_{\mathbb{R}} f'(u)p^z(u)\mathrm{d}u = \mathbb{E}[Wf(W)]$$

za sve funkcije f koje imaju gore opisanu reprezentaciju. No, kako je f apsolutno neprekidna funkcija, sigurno će postojati funkcija g takva da je $f(x) = \int_0^x g(t)\mathrm{d}t$ te funkciju g možemo rastaviti na pozitivni i negativni dio te se pozvati na upravo dokazano.

Dokažimo da je p^z vjerojatnosna distribucija. Iz same definicije vidimo da je funkcija nenegativna. Također, dvije reprezentacije iz definicije su jednake jer je $\mathbb{E}[W] = 0$. Vrijedi i

$$\int_0^\infty p^z(u)\mathrm{d}u = \mathbb{E}[W^2\mathbb{1}[W > 0]] \quad \text{i} \quad \int_{-\infty}^0 p^z(u)\mathrm{d}u = \mathbb{E}[W^2\mathbb{1}[W < 0]]$$

pa je $\int_{\mathbb{R}} p^z(u)\mathrm{d}u = \mathbb{E}[W^2] = 1$.

Naposljetku, jedinstvenost vrijedi jer za dvije slučajne varijable X i Y takve da je $\mathbb{E}[f'(X)] = \mathbb{E}[f'(Y)]$ za sve funkcije f koje su neprekidno derivabilne s kompaktnim nosačem vrijedi $X \overset{d}{=} Y$.

□

Uočimo, ako je W slučajna varijabla s očekivanjem nula i konačnom varijancom, iz definicije se jednostavno provjeri da je $(aW)^z \overset{d}{=} aW^z$, zbog čega nismo ništa izgubili u prethodnom dokazu pretpostavljajući da je $\mathrm{Var}(W) = 1$.

Naposljetku, opet nas zanima kako konstruirati nula pristranu distribuciju u slučaju kada je W zbroj nezavisnih slučajnih varijabli. Neka su $X_1, \dots, X_n$ nezavisne slučajne varijable s očekivanjem nula, $\mathrm{Var}(X_i) = \sigma_i^2$ te $\sum_{i=1}^n \sigma_i^2 = 1$. Definirajmo $W = \sum_{i=1}^n X_i$. Konstrukcija je sljedeća.

1. Za svaki $i = 1, 2, \dots, n$, neka X_i^z ima nula pristranu distribuciju s obzirom na X_i te neka je X_i^z nezavisna od $(X_j)_{j \neq i}$ i $(X_j^z)_{j \neq i}$.
2. Odaberimo nasumično $I \in \{1, \dots, n\}$ s vjerojatnosti $P(I = i) = \sigma_i^2$.
3. Definiramo $W^z = \sum_{j \neq I} X_j + X_I^z$.

Propozicija 1.32. *Neka je $W = \sum_{i=1}^n X_i$ definirana kao gore. Ako je W^z konstruirana slijedeći korake* 1-3, *tada W^z ima nula pristranu distribuciju s obzirom na W.*

Dokaz. Za sve apsolutno neprekidne funkcije f vrijedi

$$\begin{aligned}
\mathbb{E}[Wf(W)] &= \sum_{i=1}^n \mathbb{E}[X_i f(W - X_i + X_i)] \\
&= \sum_{i=1}^n \sigma_i^2 \mathbb{E}[f(W - X_i + X_i^z)] \\
&= \sum_{i=1}^n \mathbb{E}[f(W - X_i + X_i^z) \mathbb{1}[I = i]] \\
&= \mathbb{E}[f'(W - X_I + X_I^z)]
\end{aligned}$$

gdje smo u drugoj jednakosti koristili nezavisnost slučajnih varijabli X_i i $W - X_i$ te osnovno svojstvo nula pristrane distribucije, a u trećoj jednakosti koristimo nezavisnost varijable I od ostalih slučajnih varijabli.

$\square$

Poglavlje 2

Jaka ulaganja

Neka su $\varepsilon_1, \varepsilon_2, \ldots$ nezavisne i jednako distribuirane slučajne varijable očekivanja nula i $\mathbb{E}[\varepsilon_i^2] = 1$. Definirajmo $S_k = \sum_{i=1}^{k} \varepsilon_i$. Donskerov teorem nam govori da ako definiramo S^* kao linearnu interpolaciju procesa $(S_n)_{n\in\mathbb{N}}$, to jest,

$$S^*(t) = \begin{cases} S_k & \text{ako } t = k \in \mathbb{N} \\ \text{linearno na } [k, k+1], \end{cases}$$

vrijedi da $(S^*(nt)/\sqrt{n})_{t\geqslant 0}$ konvergira slabo prema standardnom Brownovom gibanju kada n ide u beskonačno, na prostoru $C[0,1]$.

Naš cilj u ovom poglavlju je konstruirati Brownovo gibanje $(B_t)_{t\geqslant 0}$ na istom vjerojatnosnom prostoru na kojem su i $(S_k)_{k\in\mathbb{N}}$ takvo da

$$\max_{1\leqslant k\leqslant n} |S_k - B_k|$$

raste sporo. Optimalnu ogradu su dali Komlós, Major i Tusnády u [5] gdje su dokazali da u slučaju kada ε_1 ima konačnu funkciju izvodnicu momenata na nekoj okolini nule, vrijedi

$$\max_{k\leqslant n} |S_k - B_k| = O(\log n).$$

Konkretno, dokazali su sljedeći teorem.

Teorem 2.1 (Komlós-Major-Tusnády)**.** *Neka su $\varepsilon_1, \varepsilon_2, \ldots$ nezavisne jednako distribuirane slučajne varijable za koje vrijedi $\mathbb{E}[\varepsilon_1] = 0, \mathbb{E}[\varepsilon_1^2] = 1$ i $\mathbb{E}[\exp(\theta|\varepsilon_1|)] < \infty$ za neki $\theta > 0$. Za $k \in \mathbb{N}$, definiramo $S_k = \sum_{i=1}^{k} \varepsilon_i$. Tada je za svaki n moguće konstruirati verziju procesa $(S_k)_{0\leqslant k\leqslant n}$ i standardno Brownovo gibanje $(B_t)_{0\leqslant t\leqslant n}$ na istom vjerojatnosnom prostoru takve da za svaki $x \geqslant 0$ vrijedi,*

$$\mathbb{P}(\max_{k\leqslant n} |S_k - B_k| \geqslant C\log n + x) \leqslant Ke^{-\lambda x},$$

za neke konstante C, K i λ koje ne ovise o n.

Mi ćemo dokazati slabiju varijantu teorema, u kojem su ε_i Rademacherove slučajne varijable (primaju vrijednosti ± 1, svaku s vjerojatnosti $1/2$), no za razliku od upravo iskazanog teorema, moći ćemo spariti cijeli proces $(S_k)_{k \in \mathbb{N}}$ s Brownovim gibanjem $(B_t)_{t \geqslant 0}$ tako da tvrdnja vrijedi za sve n, simultano. Iako je to samo poseban slučaj generalnog teorema, dovoljan je za analiziranje jednostavnih simetričnih slučajnih šetnji (SRW). Dokaz koji ćemo pokazati je fundamentalno različit od originalnog dokaza te tehnički manje zahtjevan.

Dokaz će se temeljiti na konceptu Steinovog koeficijenta koji će nam omogućiti da proizvoljnu slučajnu varijablu W sparimo s normalnom slučajnom varijablom Z tako da $W - Z$ ima eksponencijalno padajuće repove. Formalno, vrijedi sljedeći teorem.

Teorem 2.2. *Neka je W slučajna varijabla očekivanja nula i konačne varijance. Neka je T druga slučajna varijabla, definirana na istom vjerojatnosnom prostoru kao i W takva da kad god je φ Lipschitzova funkcija i φ' njena derivacija gotovo svugdje vrijedi*

$$\mathbb{E}[W\varphi(W)] = \mathbb{E}[\varphi'(W)T]. \tag{2.1}$$

Pretpostavimo da je $|T|$ gotovo sigurno ograničena konstantom. Tada, za svaki $\sigma^2 > 0$, možemo konstruirati $Z \sim N(0, \sigma^2)$ na istom vjerojatnosnom prostoru tako da za sve $\theta \in \mathbb{R}$ vrijedi

$$\mathbb{E}[\exp(\theta|W - Z|)] \leqslant 2\mathbb{E}\left[\exp\left(\frac{2\theta^2(T - \sigma^2)^2}{\sigma^2}\right)\right].$$

U slučaju da W i T zadovoljavaju prvi uvjet teorema, kažemo da je T Steinov koeficijent za W. Uočimo, ako je $T \simeq \sigma^2$ s velikom vjerojatnosti, W će biti blizu $N(0, \sigma^2)$, što zaključujemo iz (2.1) i Steinove leme.

Ključni koraci u provedbi dokaza KMT teorema za SRW će biti pronalazak Steinovih koeficijenata za odgovarajuće slučajne varijable, korištenje upravo iskazanog teorema te induktivni korak. Kao međukorak, dokazat ćemo i sljedeći teorem.

Teorem 2.3. *Postoje univerzalne konstante C, K i λ_0 takve da vrijedi sljedeće: neka je $n \geqslant 2$ proizvoljan. Pretpostavimo da su $\varepsilon_1, \ldots, \varepsilon_n$ izmjenjive slučajne varijable koje primaju vrijednosti ± 1. Za $k = 0, 1, \ldots, n$ neka je $S_k = \sum_{i=1}^{k} \varepsilon_i$ i neka je $W_k = S_k - \frac{k}{n} S_n$. Tada je moguće konstruirati verziju slučajnih varijabli $W_0, \ldots, W_n$ i standardan Brownov most $(\widetilde{B}_t)_{0 \leqslant t \leqslant 1}$ na istom vjerojatnosnom prostoru takve da za sve $0 < \lambda < \lambda_0$ vrijedi*

$$\mathbb{E}[\exp(\lambda \max_{k \leqslant n} |W_k - \sqrt{n}\widetilde{B}_{k/n}|)] \leqslant \exp(C \log n)\mathbb{E}\left[\exp\left(\frac{K\lambda^2 S_n^2}{n}\right)\right].$$

Podsjećamo, Brownov most na $[0,1]$ je slučajni proces definiran kao

$$W_t = B_t - tB_1,$$

pri čemu je B standardno Brownovo gibanje. Intuitivno, proces W možemo shvatiti kao da smo Brownovo gibanje uvjetovali na događaj $B_1 = 0$. Brownov most ima kovarijacijsku funkciju $\gamma(s,t) = s \wedge t - st$.

2.1 Teorem o sparivanju

Prvo ćemo dokazati teorem 2.2. To radimo kroz niz lema. Dokaz prve od njih je tehnički zahtjevan, a tehnike korištene u dokazu nisu potrebne za druge dokaze pa ga izostavljamo.

Lema 2.4. *Neka je n prirodan broj i neka je A neprekidna funkcija iz $\mathbb{R}^n$ u skup svih $n \times n$ pozitivno semidefinitnih matrica. Pretpostavimo da postoji konstanta $b \geqslant 0$ takva da za sve $x \in \mathbb{R}^n$,*

$$\|A(x)\| \leqslant b.$$

Tada postoji vjerojatnosna mjera μ na $\mathbb{R}^n$ takva da ako je X slučajni vektor s razdiobom μ, tada

$$\mathbb{E}[\exp(\langle \theta, X\rangle)] \leqslant \exp(b\|\theta\|^2)$$

za sve $\theta \in \mathbb{R}^n$ te

$$\mathbb{E}[\langle X, \nabla f(X)\rangle] = \mathbb{E}[\mathrm{Tr}(A(X)\,\mathrm{Hess}\, f(X))]$$

za sve $f \in C^2(\mathbb{R}^n)$ takve da su očekivanja

$$\mathbb{E}|f(X)^2|, \mathbb{E}\|\nabla f(X)\|^2 \text{ i } \mathbb{E}|\,\mathrm{Tr}(A(X)\,\mathrm{Hess}\, f(X))|$$

konačna.

Lema 2.5. *Neka su A i X kao u lemi 2.4. Neka su $1 \leqslant i < j \leqslant n$ proizvoljni. Definiramo*

$$v_{ij}(x) = a_{ii}(x) + a_{jj}(x) - 2a_{ij}(x),$$

gdje a_{ij} označava element na mjestu (i,j) matrice A. Tada, za sve $\theta \in \mathbb{R}$, vrijedi

$$\mathbb{E}[\exp(\theta|X_i - X_j|)] \leqslant 2\mathbb{E}[\exp(2\theta^2 v_{ij}(X))].$$

Dokaz. Neka je k prirodan broj. Definiramo $f \colon \mathbb{R}^n \to \mathbb{R}$

$$f(x) = (x_i - x_j)^{2k}.$$

Direktnim računom dobivamo

$$\langle x, \nabla f(x) \rangle = 2k(x_i - x_j)^{2k}$$

te

$$\operatorname{Tr}(A(x) \operatorname{Hess} f(x)) = 2k(2k-1)(x_i - x_j)^{2k-2} v_{ij}(x).$$

Jer je matrica A pozitivno definitna, funkcija v_{ij} je nenegativna. Primjenom Hölderove nejednakosti dobivamo

$$\mathbb{E}|\operatorname{Tr}(A(X) \operatorname{Hess} f(X))| \leqslant 2k(2k-1)(\mathbb{E}[(X_i - X_j)^{2k}])^{\frac{k-1}{k}}(\mathbb{E}[v_{ij}(X)^k])^{\frac{1}{k}}.$$

Iz drugog identiteta u lemi 2.4 možemo zaključiti

$$\mathbb{E}[(X_i - X_j)^{2k}] \leqslant (2k-1)(\mathbb{E}[(X_i - X_j)^{2k}])^{\frac{k-1}{k}}(\mathbb{E}[v_{ij}(X)^k])^{\frac{1}{k}}$$

što povlači

$$\mathbb{E}[(X_i - X_j)^{2k}] \leqslant (2k-1)^k \mathbb{E}[v_{ij}(X)^k].$$

Vrijedi

$$\begin{aligned}
\mathbb{E}[\exp(\theta|X_i - X_j|)] &\leqslant 2\mathbb{E}[\cosh(\theta(X_i - X_j))] \\
&= 2\sum_{k=0}^{\infty} \frac{\theta^{2k}\mathbb{E}[(X_i - X_j)^{2k}]}{(2k)!} \\
&\leqslant 2 + 2\sum_{k=1}^{\infty} \frac{(2k-1)^k \theta^{2k} \mathbb{E}[v_{ij}(X)^k]}{(2k)!}
\end{aligned}$$

pa koristeći nejednakost

$$\frac{(2k-1)^k}{(2k)!} \leqslant \frac{2^k}{k!}$$

dobivamo traženu tvrdnju.

□

Lema 2.6. *Neka je ρ vjerojatnosna funkcija gustoće na $\mathbb{R}$, pozitivna na intervalu (ograničenom ili neograničenom), a nula van tog intervala. Pretpostavimo $\int_{-\infty}^{\infty} x\rho(x)\mathrm{d}x = 0$. Za svaki x u nosaču funkcije ρ, definiramo*

$$h(x) = \frac{\int_x^{\infty} y\rho(y)\mathrm{d}y}{\rho(x)}.$$

Van nosača, neka je $h \equiv 0$. Neka je X slučajna varijabla s gustoćom ρ i konačnim drugim momentom. Tada vrijedi

$$\mathbb{E}[X\varphi(X)] = \mathbb{E}[h(X)\varphi'(X)] \tag{2.2}$$

za sve apsolutno neprekidne funkcije ρ takve da su obje strane jednakosti dobro definirane i $\mathbb{E}|h(X)\varphi(X)| < \infty$. Štoviše, ako je h_1 neka druga funkcija koja zadovoljava (2.2) za sve Lipschitzove φ, tada je $h_1 = h$ gotovo svugdje na nosaču funkcije ρ.

Obratno, ako je Y slučajna varijabla takva da vrijedi (2.2), gdje umjesto X stoji Y, za sve φ takve su $|\varphi(x)|$, $|x\varphi(x)|$ i $|h(x)\varphi'(x)|$ uniformno ograničene, tada Y ima gustoću ρ.

Dokaz. Neka je $u(x) = h(x)\rho(x)$. To je neprekidna funkcija, pozitivna na nosaču funkcije ρ te vrijedi $\lim_{x\to-\infty} u(x) = \lim_{x\to\infty} u(x) = 0$ jer

$$u(x) = \int_x^\infty y\rho(y)\mathrm{d}y = -\int_{-\infty}^x y\rho(y)\mathrm{d}y.$$

Jer je $\mathbb{E}[X^2] < \infty$, primjenom Fubinijevog teorema dobivamo

$$\mathbb{E}[h(X)] = \int_{-\infty}^\infty u(x)\mathrm{d}x = \mathbb{E}[X^2] < \infty.$$

Kada je φ ograničena Lipschitzova funkcija, jednakost (2.2) slijedi iz parcijalne integracije,

$$\begin{aligned}\int_{-\infty}^\infty x\varphi(x)\rho(x)\mathrm{d}x &= \begin{bmatrix} p = \varphi(x) & p' = \varphi'(x) \\ q' = x\varphi(x) & q = -u(x) \end{bmatrix} \\ &= \varphi(x)u(x)\Big|_{-\infty}^\infty + \int_{-\infty}^\infty \varphi'(x)u(x)\mathrm{d}x \quad = \int_{-\infty}^\infty \varphi'(x)u(x)\mathrm{d}x.\end{aligned}$$

Sada, neka je φ proizvoljna apsolutno neprekidna funkcija, i neka je $g\colon \mathbb{R} \to [0,1]$ C^∞ funkcija takva da je $g(x) = 1$ na $[-1,1]$ i $g(x) = 0$ izvan $[-2,2]$. Za svaki $a > 1$, neka je

$$\varphi_a(x) = \varphi(x)g(x/a).$$

Vrijedi

$$\varphi_a'(x) = \varphi'(x)g(x/a) + a^{-1}\varphi(x)g'(x/a).$$

Lagano se vidi da su funkcije φ_a i φ_a' ograničene i konvergiraju po točkama k φ i φ' kada $a \to \infty$. Štoviše, $|x\varphi_a(x)| \leqslant |x\varphi(x)|$ te

$$|h(x)\varphi_a'(x)| \leqslant |h(x)\varphi'(x)| + a^{-1}\|g'\|_\infty |h(x)\varphi(x)|.$$

Jer smo pretpostavili da su $\mathbb{E}|X\varphi(X)|$, $\mathbb{E}|h(X)\varphi'(X)|$ i $\mathbb{E}|h(X)\varphi(X)|$ konačni, možemo primijeniti teorem o dominiranoj konvergenciji i zaključujemo da (2.2) vrijedi za funkciju φ.

Sada, pretpostavimo da je h_1 neka druga funkcija koja zadovoljava (2.2) za sve Lipschitzove φ i $\mathbb{E}[X^2] < \infty$. Neka je $\varphi(x)$ Lipschitzova funkcija za koju je $\varphi'(x) = \operatorname{sign}(h_1(x) - h(x))$. Tada

$$0 = \mathbb{E}[\varphi'(X)(h_1(X) - h(X))] = \mathbb{E}|h_1(X) - h(X)|$$

pa je $h_1 = h$ gotovo svugdje na nosaču gustoće ρ.

Za obrat, neka X ima gustoću ρ i neka je $v\colon \mathbb{R} \to \mathbb{R}$ proizvoljna ograničena funkcija. Označimo $m = \mathbb{E}[v(X)]$ i definiramo

$$\varphi(x) = \frac{1}{u(x)} \int_{-\infty}^{x} \rho(y)(v(y) - m)\mathrm{d}y = \frac{-1}{u(x)} \int_{x}^{\infty} \rho(y)(v(y) - m)\mathrm{d}y$$

na nosaču funkcije ρ. Jer je u nenegativna i apsolutno neprekidna na nosaču funkcije ρ, funkcija φ je dobro definirana i apsolutno neprekidna. Dokažimo da je $|x\varphi(x)|$ uniformno ograničena. Ako je $x \geqslant 0$, vrijedi

$$\begin{aligned} |x\varphi(x)| &= \left| \frac{x}{u(x)} \int_{x}^{\infty} \rho(y)(v(y) - m)\mathrm{d}y \right| \\ &\leqslant \frac{2\|v\|_\infty}{|u(x)|} \int_{x}^{\infty} y\rho(y)\mathrm{d}y = 2\|v\|_\infty. \end{aligned}$$

Slično dokažemo da ista ograda vrijedi i za $x < 0$. Direktnom provjerom dokažemo da je

$$h(x)\varphi'(x) - x\varphi(x) = v(x) - m$$

zbog čega je $|h(x)\varphi'(x)|$ uniformno ograničena. Naposljetku, jer je φ neprekidna,

$$|\varphi(x)| \leqslant \sup_{|t| \leqslant 1} |\varphi(t)| + |x\varphi(x)|,$$

funkcija φ je također uniformno ograničena.

Dakle, ako je Y slučajna varijabla takva da (2.2) vrijedi za Y umjesto X i sve φ za koje su $|\varphi(x)|, |x\varphi(x)|$ i $|h(x)\varphi'(x)|$ uniformno ograničene, tada

$$\mathbb{E}[v(Y)] - \mathbb{E}[v(X)] = \mathbb{E}[v(Y) - m] = \mathbb{E}[h(Y)\varphi'(Y) - Y\varphi(Y)] = 0,$$

dakle X i Y su jednako distribuirane.

□

Naposljetku, dokažimo teorem 2.2

Dokaz. Pretpostavimo da W ima gustoću ρ s obzirom na Lebesgueovu mjeru koja je pozitivna i neprekidna svugdje. Definirajmo h kao u lemi 2.6. Tada, koristeći svojstva uvjetnog očekivanja, funkcija $w \mapsto \mathbb{E}[T \mid W = w]$ zadovoljava ista svojstva kao i funkcija h pa iz dokazane jedinstvenosti u lemi 2.6, vrijedi

$$h(w) = \mathbb{E}[T \mid W = w], \text{ g.s.}$$

Funkcija h je nenegativna po definiciji pa možemo definirati funkciju A iz $\mathbb{R}^2$ u prostor svih 2×2 pozitivno semidefinitnih matrica kao

$$A(x_1, x_2) = \begin{bmatrix} h(x_1) & \sigma\sqrt{h(x_1)} \\ \sigma\sqrt{h(x_1)} & \sigma^2 \end{bmatrix}.$$

Uočimo, funkcija A ne ovisi o x_2 i matrica je uistinu pozitivno semidefinitna. Također, jer smo pretpostavili da je ρ neprekidna, onda su također i h i A. Jer je T ograničena konstantom, onda je i h.

Neka je $X = (X_1, X_2)$ slučajan vektor koji zadovoljava svojstva iz leme 2.4 uz upravo definirani A. Neka je $\varphi \colon \mathbb{R} \to \mathbb{R}$ apsolutno neprekidna funkcija takva da su $|\varphi(x)|$, $|x\varphi(x)|$ i $|h(x)\varphi'(x)|$ uniformno ograničene. Neka je Φ takva da je $\Phi' = \varphi$. Možemo pretpostaviti da je $\Phi(0) = 0$. Definiramo $f \colon \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = \Phi(x_1)$. Tada, za neku konstantu C, za sve x_1, x_2 vrijedi

$$|f(x_1, x_2)| \leqslant C|x_1|, \|\nabla f(x_1, x_2)\| \leqslant C, |\operatorname{Tr}(A(x_1, x_2) \operatorname{Hess} f(x_1, x_2))| \leqslant C$$

gdje prva nejednakost slijedi primjenom teorema srednje vrijednosti. Sada, zbog leme 2.4, možemo zaključiti

$$\mathbb{E}[\langle X, \nabla f(X) \rangle] = \mathbb{E}[\operatorname{Tr}(A(X) \operatorname{Hess} f(X))],$$

što je zapravo

$$\mathbb{E}[X_1\varphi(X_1)] = \mathbb{E}[h(X_1)\varphi'(X_1)].$$

S obzirom da upravo dokazana jednakost vrijedi za sve φ za koje su $|\varphi(x)|$, $|x\varphi(x)|$ i $|h(x)\varphi'(x)|$ uniformno ograničeni, po 2.6 zaključujemo da su X_1 i W jednako distribuirani.

Na sličan način, uzimajući proizvoljnu φ takvu da su $|\varphi(x)|$, $|x\varphi(x)|$ i $|\varphi'(x)|$ uniformno ograničene, uz definiranje Φ preko $\Phi'(x) = \varphi(x)$ i $f(x_1, x_2) = \Phi(x_2)$, možemo zaključiti

$$\mathbb{E}[X_2\varphi(X_2)] = \sigma^2\mathbb{E}[\varphi'(X_2)]$$

što povlači da je $X_2 \sim N(0, \sigma^2)$. Sada, primjenjujemo lemu 2.5 na par (X_1, X_2). Vrijedi

$$v_{12}(x_1, x_2) = h(x_1) + \sigma^2 - 2\sigma\sqrt{h(x_1)} = \left(\sqrt{h(x_1)} - \sigma\right)^2.$$

Jer je $h(x_1) \geqslant 0$, vrijedi

$$\left(\sqrt{h(x_1)} - \sigma\right)^2 = \frac{(h(x_1) - \sigma^2)^2}{(\sqrt{h(x_1)} + \sigma)^2} \leqslant \frac{(h(x_1) - \sigma^2)^2}{\sigma^2}.$$

Kako su $h(X_1)$ i $h(W)$ jednako distribuirani te $h(W) = \mathbb{E}[T \mid W]$, tvrdnja vrijedi primjenom leme 2.5 i Jensenove nejednakosti.

Dakle, dokazali smo slučaj kada W ima gustoću ρ s obzirom na Lebesgueovu mjeru koja je pozitivna i neprekidna svugdje. Sada, za svaki $\varepsilon > 0$ definiramo $W_\varepsilon = W + \varepsilon Y$, gdje je $Y \sim N(0, 1)$ nezavisna od W. Ako ν označava distribuciju slučajne varijable W, tada W_ε ima gustoću

$$\rho_\varepsilon(x) = \int_{-\infty}^{\infty} \frac{e^{-(x-y)^2/2\varepsilon^2}}{\sqrt{2\pi}\varepsilon} \mathrm{d}\nu(y)$$

s obzirom na Lebesgueovu mjeru. Funkcija ρ_ε je pozitivna i neprekidna svugdje. Za svaku Lipschitzovu funkciju φ vrijedi

$$\begin{aligned}
\mathbb{E}[W_\varepsilon \varphi(W_\varepsilon)] &= \mathbb{E}[W\varphi(W + \varepsilon Y)] + \varepsilon \mathbb{E}[Y\varphi(W + \varepsilon Y)] \\
&= \mathbb{E}[T\varphi'(W + \varepsilon Y)] + \varepsilon^2 \mathbb{E}[\varphi'(W + \varepsilon Y)] \\
&= \mathbb{E}[(T + \varepsilon^2)\varphi'(W_\varepsilon)].
\end{aligned}$$

Dakle, po prethodno dokazanom, možemo konstruirati verziju slučajne varijable W_ε i $Z_\varepsilon \sim N(0, \sigma^2 + \varepsilon^2)$ na istom vjerojatnosnom prostoru takve da za sve θ,

$$\mathbb{E}[\exp(\theta|W_\varepsilon - Z_\varepsilon|)] \leqslant 2\mathbb{E}\left[\exp\left(\frac{2\theta^2(T - \sigma^2)^2}{\sigma^2 + \varepsilon^2}\right)\right].$$

Neka je μ_ε distribucija slučajnog para $(W_\varepsilon, Z_\varepsilon)$. Intuitivno je jasno da je familija mjera $\{\mu_\varepsilon\}_{0<\varepsilon<1}$ napeta. Naime, kako se ε smanjuje prema nuli, distribucija slučajne varijable W_ε je sve bliža fiksnoj distribuciji W, a slučajna varijabla Z_ε je po upravo dokazanom blisko sparena s W_ε.

Posebno, zbog napetosti, postojati će gomilište μ_0 kada $\varepsilon \to 0$. Neka $(W_0, Z_0) \sim \mu_0$. Jer $\varepsilon \to 0$, vrijedi $W_0 \sim W$ i $Z_0 \sim N(0, \sigma^2)$. Naposljetku, po Skorohodovom teoremu o reprezentaciji, možemo konstruirati odgovarajuće slučajne parove

$(W_\varepsilon, Z_\varepsilon)$ na istom vjerojatnosnom prostoru pa zbog Fatouove leme i teorema o monotonoj konvergenciji vrijedi

$$\mathbb{E}[\exp(\theta|W_0 - Z_0|)] \leqslant \liminf_{\varepsilon \to 0} \mathbb{E}[\exp(\theta|W_\varepsilon - Z_\varepsilon|)] \leqslant 2\mathbb{E}\left[\exp\left(\frac{2\theta^2(T-\sigma^2)^2}{\sigma^2}\right)\right].$$

□

2.2 Konstrukcije Steinovih koeficijenata

U prethodnom poglavlju dokazali smo teorem koji nam omogućava da slučajnu varijablu za koji postoji Steinov koeficijent sparimo s normalno distribuiranom slučajnom varijablom. Sada se fokusiramo na konstrukciju Steinovih koeficijenata za jednostavnu simetričnu slučajnu šetnju te sparivanje iste s odgovarajućom normalnom slučajnom varijablom. Dokazati ćemo dva teorema, od kojih drugi možemo shvatiti kao uvjetnu verziju prvog.

Teorem 2.7. *Postoje univerzalne konstante κ i $\theta_0 > 0$ takve da vrijedi sljedeće: neka je n prirodan broj i neka su $\varepsilon_1, \dots, \varepsilon_n$ nezavisne jednako distribuirane uniformne slučajne varijable na $\{-1, 1\}$. Neka je $S_n = \sum_{i=1}^n \varepsilon_i$. Tada je moguće konstruirati verziju slučajne varijable S_n i slučajnu varijablu $Z_n \sim N(0, n)$ na istom vjerojatnosnom prostoru tako da vrijedi*

$$\mathbb{E}[\exp(\theta_0|S_n - Z_n|)] \leqslant \kappa.$$

Teorem 2.8. *Neka su $\varepsilon_1, \dots, \varepsilon_n$ proizvoljni elementi skupa $\{-1, 1\}$. Neka je π uniformno slučajno odabrana permutacija skupa $\{1, \dots, n\}$. Za svaki $1 \leqslant k \leqslant n$, neka je $S_k = \sum_{\ell=1}^k \varepsilon_{\pi(\ell)}$ i neka je*

$$W_k = S_k - \frac{kS_n}{n}.$$

Tada postoje univerzalne konstante $c > 1$ i $\theta_0 > 0$ koje zadovoljavaju sljedeće. Neka je $n \geqslant 3$ proizvoljan, S_n proizvoljan i $n/3 \leqslant k \leqslant 2n/3$. Tada je moguće konstruirati verziju slučajne varijable W_k i normalno distribuiranu slučajnu varijablu Z_k s očekivanjem nula i varijancom $k(n-k)/n$ na istom vjerojatnosnom prostoru takve da za sve $\theta \leqslant \theta_0$,

$$\mathbb{E}[\exp(\theta|W_k - Z_k|)] \leqslant \exp\left(1 + \frac{c\theta^2 S_n^2}{n}\right).$$

Oba teorema ćemo dokazati koristeći teorem 2.2. Kao i prije, teoreme ćemo dokazati kroz niz lema.

Lema 2.9. *Neka su X i Y dvije nezavisne slučajne varijable, pri čemu je X uniformno distribuirana na $\{-1, 1\}$, a Y uniformno distribuirana na $[-1, 1]$. Tada za bilo koju Lipschitzovu funkciju φ vrijedi*

$$\mathbb{E}[X\varphi(X+Y)] = \mathbb{E}[(1-XY)\varphi'(X+Y)]$$

te

$$\mathbb{E}[Y\varphi(X+Y)] = \frac{1}{2}\mathbb{E}[(1-Y^2)\varphi'(X+Y)].$$

Dokaz. Vrijedi

$$\mathbb{E}[(1-XY)\varphi'(X+Y)] = \frac{1}{4}\int_{-1}^{1}(1+y)\varphi'(-1+y)\mathrm{d}y + \frac{1}{4}\int_{-1}^{1}(1-y)\varphi'(1+y)\mathrm{d}y.$$

Parcijalnom integracijom, slijedi

$$\int_{-1}^{1}(1+y)\varphi'(-1+y)\mathrm{d}y = 2\varphi(0) - \int_{-1}^{1}\varphi(-1+y)\mathrm{d}y$$

te

$$\int_{-1}^{1}(1-y)\varphi'(1+y)\mathrm{d}y = -2\varphi(0) + \int_{-1}^{1}\varphi(1+y)\mathrm{d}y.$$

Zbrajajući, dobivamo

$$\begin{aligned}\mathbb{E}[(1-XY)\varphi'(X+Y)] &= \frac{1}{4}\int_{-1}^{1}\varphi(1+y)\mathrm{d}y - \frac{1}{4}\int_{-1}^{1}\varphi(-1+y)\mathrm{d}y \\ &= \mathbb{E}[X\varphi(X+Y)].\end{aligned}$$

Za drugu tvrdnju, dovoljno je uočiti da za bilo koji x, parcijalna integracija daje

$$\frac{1}{2}\int_{-1}^{1} y\varphi(x+y)\mathrm{d}y = \frac{1}{2}\int_{-1}^{1}\frac{1-y^2}{2}\varphi'(x+y)\mathrm{d}y.$$

□

Sada smo spremni dokazati teorem 2.7

Dokaz. Zbog jednostavnosti, označimo S_n sa S. Neka je Y slučajna varijabla nezavisna od $\varepsilon_1, \ldots, \varepsilon_n$ i uniformno distribuirana na $[-1, 1]$. Pretpostavimo da su nam dane vrijednosti $\varepsilon_1, \ldots, \varepsilon_{n-1}$ i neka je $\mathbb{E}^-$ odgovarajuće uvjetno očekivanje. Neka su

$$S^- = \sum_{i=1}^{n-1}\varepsilon_i, X = \varepsilon_n.$$

Tada, po lemi 2.9, vrijedi

$$\mathbb{E}^-[X\varphi(S^- + X + Y)] = \mathbb{E}^-[(1 - XY)\varphi'(S + Y)] = \mathbb{E}^-[(1 - \varepsilon_n Y)\varphi'(S + Y)].$$

Uzimajući očekivanje na obje strane, dobivamo

$$\mathbb{E}[\varepsilon_n \varphi(S + Y)] = \mathbb{E}[(1 - \varepsilon_n Y)\varphi'(S + Y)].$$

Zbog simetrije, to daje

$$\mathbb{E}[S\varphi(S + Y)] = \mathbb{E}[(n - SY)\varphi'(S + Y)].$$

Opet, po lemi 2.9, vrijedi

$$\mathbb{E}[Y\varphi(S + Y)] = \frac{1}{2}\mathbb{E}[(1 - Y^2)\varphi'(S + Y)].$$

Ako označimo $\widetilde{S} = S + Y$ i

$$T = n - SY + \frac{1 - Y^2}{2},$$

zbrajanjem dobivamo

$$\mathbb{E}[\widetilde{S}\varphi(\tilde{S})] = \mathbb{E}[T\varphi'(\widetilde{S})].$$

Označimo $\sigma^2 = n$. Tada elementaran račun daje

$$\frac{(T - \sigma^2)^2}{\sigma^2} \leqslant \frac{2S^2 + \frac{1}{2}}{n}.$$

Sada, uočimo da je $\mathbb{E}[\widetilde{S}] = 0$, $\mathbb{E}[\widetilde{S}^2] < \infty$. Također, T je gotovo sigurno ograničena i vrijedi $\mathbb{E}[\widetilde{S}\varphi(\tilde{S})] = \mathbb{E}[T\varphi'(\widetilde{S})]$ zbog čega možemo primijeniti teorem 2.2. Po njemu, moguće je konstruirati verziju slučajne varijable $\widetilde{S}$ i slučajnu varijablu $Z \sim N(0, \sigma^2)$ na istom vjerojatnosnom prostoru takve da za sve θ,

$$\mathbb{E}[\exp(\theta|\widetilde{S} - Z|)] \leqslant 2\mathbb{E}[\exp(2\theta^2\sigma^{-2}(T - \sigma^2)^2].$$

Sada, zbog veze između S i $\widetilde{S}$, na istom prostoru možemo i konstruirati verziju od S takvu da je $|S - \widetilde{S}| \leqslant 1$. Slijedi

$$\mathbb{E}[\exp(\theta|S - Z|)] \leqslant 2\mathbb{E}[\exp(|\theta| + 2\theta^2\sigma^{-2}(T - \sigma^2)^2)].$$

Koristeći gore dobivenu ogradu na $(T - \sigma^2)^2/\sigma^2$, dobivamo

$$\mathbb{E}[\exp(\theta|S - Z|)] \leqslant 2\exp(|\theta| + \theta^2/n)\mathbb{E}[\exp(4\theta^2 S^2/n)].$$

Sada, ako je $V \sim N(0,1)$ nezavisna od S, imamo

$$\begin{aligned}\mathbb{E}[\exp(4\theta^2 S^2/n)] &= \mathbb{E}[\exp(\sqrt{8}\theta V S/\sqrt{n})] \\ &= \mathbb{E}[\mathbb{E}[\exp(\sqrt{8}\theta V \varepsilon_1/\sqrt{n} \mid V]^n] \\ &= \mathbb{E}[\cosh(\sqrt{8}\theta V/\sqrt{n})^n],\end{aligned}$$

gdje prva jednakost slijedi iz Fubinijevog teorema, druga iz svojstava uvjetnog očekivanja i nezavisnosti, a zadnja jer je ε_1 uniformna na $\{-1,1\}$. Naposljetku, koristeći jednostavnu nejednakost $\cosh x \leqslant \exp(x^2)$, dobivamo

$$\mathbb{E}[\exp(4\theta^2 S^2/n) \leqslant \mathbb{E}[\exp(8\theta^2 V^2)] = \frac{1}{\sqrt{1-16\theta^2}}, \text{za } 16\theta^2 < 1. \tag{2.3}$$

□

Prelazimo na dokaz drugog teorema.

Lema 2.10. *U notaciji teorema 2.8, za sve $\theta \in \mathbb{R}$ i sve $1 \leqslant k \leqslant n$, vrijedi*

$$\mathbb{E}[\exp(\theta W_k/\sqrt{k})] \leqslant \exp(\theta^2).$$

Dokaz. Fiksirajmo k i označimo $m(\theta) = \mathbb{E}[\exp(\theta W_k/\sqrt{k})]$. Jer je W_k ograničena slučajna varijabla, na standardan način se dokaže da je m diferencijabilna funkcija za koju vrijedi

$$m'(\theta) = \frac{1}{\sqrt{k}}\mathbb{E}[W_k \exp(\theta W_k/\sqrt{k})].$$

Sada, računamo

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{k}\sum_{j=k+1}^{n}(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}) &= \frac{(n-k)\sum_{i=1}^{k}\varepsilon_{\pi(i)} - k\sum_{j=k+1}^{n}\varepsilon_{\pi(j)}}{n} \\ &= \frac{(n-k)\sum_{i=1}^{k}\varepsilon_{\pi(i)} - k(S_n - \sum_{i=1}^{k}\varepsilon_{\pi(i)})}{n} \\ &= \sum_{i=1}^{k}\varepsilon_{\pi(i)} - \frac{kS_n}{n} = W_k.\end{aligned}$$

Dakle, dokazali smo

$$m'(\theta) = \frac{1}{n\sqrt{k}}\sum_{i=1}^{k}\sum_{j=k+1}^{n}\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})\exp(\theta W_k/\sqrt{k})].$$

Fiksirajmo $i \leqslant k < j$. Neka je $\pi' = \pi \circ (ij)$, dakle, $\pi'(i) = \pi(j), \pi'(j) = \pi(i)$, a jednake su na ostalim elementima. Opet, π' je uniformno distribuirana na skupu svih permutacija skupa $\{1, 2, \dots, n\}$. Štoviše, par (π, π') je izmjenjiv. Neka je

$$W'_k = \sum_{\ell=1}^{k} \varepsilon_{\pi'(\ell)} - \frac{kS_n}{n}.$$

Vrijedi

$$\begin{aligned}\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}) \exp(\theta W_k/\sqrt{k})] &= \mathbb{E}[(\varepsilon_{\pi'(i)} - \varepsilon_{\pi'(j)}) \exp(\theta W'_k/\sqrt{k})] \\ &= \mathbb{E}[(\varepsilon_{\pi(j)} - \varepsilon_{\pi(i)}) \exp(\theta W'_k/\sqrt{k})].\end{aligned}$$

Iz toga slijedi

$$\begin{aligned}&\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}) \exp(\theta W_k/\sqrt{k})] = \\ &\qquad = \frac{1}{2}\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})(\exp(\theta W_k/\sqrt{k}) - \exp(\theta W'_k/\sqrt{k}))].\end{aligned}$$

Koristeći nejednakost

$$|e^x - e^y| \leqslant \frac{1}{2}|x - y|(e^x + e^y)$$

i jer je $W_k - W'_k = \varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}$ slijedi

$$\begin{aligned}&|\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}) \exp(\theta W_k/\sqrt{k})]| \\ &\qquad \leqslant \frac{|\theta|}{4\sqrt{k}}\mathbb{E}[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})^2(\exp(\theta W_k/\sqrt{k}) + \exp(\theta W'_k/\sqrt{k}))] \\ &\qquad \leqslant \frac{|\theta|}{\sqrt{k}}\mathbb{E}[\exp(\theta W_k/\sqrt{k}) + \exp(\theta W'_k/\sqrt{k})] \\ &\qquad = \frac{2|\theta|}{\sqrt{k}}\mathbb{E}[\exp(\theta W_k/\sqrt{k})] = \frac{2|\theta|}{\sqrt{k}} m(\theta).\end{aligned}$$

Sada, ako iskoristimo ovu ocjenu u reprezentaciji $m'(\theta)$ koju smo dobili na početku dokaza, dobivamo

$$|m'(\theta)| \leqslant \frac{2|\theta|}{nk} \sum_{i=1}^{k} \sum_{j=k+1}^{n} m(\theta) \leqslant 2|\theta| m(\theta).$$

Uz činjenicu da je $m(0) = 1$, ocjena iz teorema jednostavno slijedi.

□

Lema 2.11. *Nastavljamo s notacijom teorema 2.8. Postoji univerzalna konstanta $\alpha_0 > 0$ takva da za sve n, sve moguće vrijednosti S_n, sve k takve da je $k \leqslant 2n/3$ i sve $\alpha \leqslant \alpha_0$ vrijedi*

$$\mathbb{E}[\exp(\alpha S_k^2/k)] \leqslant \exp\left(1 + \frac{3\alpha S_n^2}{4n}\right).$$

Dokaz. Neka je $Z \sim N(0,1)$ nezavisna. Tada

$$\mathbb{E}[\exp(\alpha S_k^2/k) = \mathbb{E}[\exp(ZS_k\sqrt{2\alpha/k})] = \mathbb{E}\left[\exp\left(ZW_k\sqrt{2\alpha/k} + Z\frac{kS_n}{n}\sqrt{2\alpha/k}\right)\right].$$

Po prethodnoj lemi,

$$\mathbb{E}\left[\exp\left(\sqrt{\frac{2\alpha}{k}}ZW_k\right) \middle| Z\right] \leqslant \exp(2\alpha Z^2).$$

Dakle, vrijedi

$$\mathbb{E}[\exp(\alpha S_k^2/k)] \leqslant \mathbb{E}\left[\exp\left(2\alpha Z^2 + \sqrt{\frac{2\alpha}{k}}\frac{kS_n}{n}Z\right)\right].$$

Jer je S_n deterministička vrijednost, desna strana je samo očekivanje funkcije $Z \sim N(0,1)$ što se lako izračuna. Nakon računa, dobivamo, za $0 < \alpha < 1/4$,

$$\mathbb{E}[\exp(\alpha S_k^2/k)] \leqslant \frac{1}{\sqrt{1-4\alpha}}\exp\left(\frac{\alpha k S_n^2}{(1-4\alpha)n^2}\right).$$

Ako k ograničimo s $2n/3$ i odaberemo dovoljno malen α_0, dokazali smo lemu. $\square$

Možemo preći na dokaz teorema 2.8.

Dokaz. Označimo $W = W_k$, $S = S_n$. Neka je Y slučajna varijabla uniformno distribuirana na $[-1,1]$ nezavisna od π. Fiksiramo $1 \leqslant i \leqslant k$ i $k < j \leqslant n$. Pretpostavimo da su nam dane vrijednosti $\{\pi(\ell), \ell \neq i, j\}$. Označimo s $\mathbb{E}^-$ uvjetnu vjerojatnost uz te informacije. Označimo

$$S^- = \sum_{\ell \neq i,j} \varepsilon_{\pi(\ell)}, W^- = \sum_{\ell \leqslant k, \ell \neq i} \varepsilon_{\pi(\ell)} - \frac{kS}{n}.$$

Ako je $S \neq S^-$, tada je $\varepsilon_{\pi(i)} = \varepsilon_{\pi(j)}$ i u tom slučaju

$$\mathbb{E}^-[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})\varphi(W+Y)] = 0.$$

U slučaju kada je $S = S^-$, uvjetna distribucija varijable $\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}$ je simetrična nad $\{-2, 2\}$. Neka je

$$X = \frac{\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}}{2} = \varepsilon_{\pi(i)}$$

i uočimo da vrijedi

$$W = W^- + X.$$

Dakle, na događaju $S = S^-$, po lemi 2.9 za sve Lipschitzove φ vrijedi

$$\begin{aligned} \mathbb{E}^-[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})\varphi(W+Y)] &= 2\mathbb{E}^-[X\varphi(W^- + X + Y)] \\ &= 2\mathbb{E}^-[(1 - XY)\varphi'(W+Y)] \\ &= \mathbb{E}^-[(2 - (\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})Y)\varphi'(W+Y)]. \end{aligned}$$

Uvedimo

$$a_{ij} = 1 - \varepsilon_{\pi(i)}\varepsilon_{\pi(j)} - (\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})Y.$$

Direktnim računom imamo

$$a_{ij} = \begin{cases} 2 - (\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})Y & \text{ako } \varepsilon_{\pi(i)} \neq \varepsilon_{\pi(j)} \\ 0 & \text{ako } \varepsilon_{\pi(i)} = \varepsilon_{\pi(j)}. \end{cases}$$

Dakle, neovisno o tome je li $S = S^-$ ili $S \neq S^-$, vrijedi

$$\mathbb{E}^-[(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)})\varphi(W+Y)] = \mathbb{E}^-[a_{ij}\varphi'(W+Y)].$$

Uzimanjem očekivanja, možemo zamijeniti $\mathbb{E}^-$ s $\mathbb{E}$ u gornjoj jednakosti. Sada, kao u dokazu leme 2.10, uočimo da je

$$W = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=k+1}^{n}(\varepsilon_{\pi(i)} - \varepsilon_{\pi(j)}).$$

Iz zadnje dvije dokazane jednakosti, slijedi

$$\mathbb{E}[W\varphi(W+Y)] = \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=k+1}^{n}a_{ij}\right)\varphi'(W+Y)\right].$$

Opet, po lemi 2.9, vrijedi

$$\mathbb{E}[Y\varphi(W+Y)] = \frac{1}{2}\mathbb{E}[(1 - Y^2)\varphi'(W+Y)].$$

Ako označimo $\widetilde{W} = W + Y$ i

$$T = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=k+1}^{n} a_{ij} + \frac{1 - Y^2}{2},$$

dobili smo

$$\mathbb{E}[\widetilde{W}\varphi(\widetilde{W})] = \mathbb{E}[T\varphi'(\widetilde{W})].$$

Direktan račun daje

$$\frac{1}{n} \sum_{i=1}^{k} \sum_{j=k+1}^{n} a_{ij} = \frac{k(n-k)}{n} - \frac{(\sum_{i=1}^{k} \varepsilon_{\pi(i)})(\sum_{j=k+1}^{n} \varepsilon_{\pi(j)})}{n} - WY.$$

Označimo $\sigma^2 = k(n-k)/n$. Jer je $n/3 \leqslant k \leqslant 2n/3$ i $|W| \leqslant |S_k| + \frac{2}{3}|S|$, vrijedi

$$\begin{aligned} \frac{(T - \sigma^2)^2}{\sigma^2} &\leqslant \frac{n}{k(n-k)}(|S_k| + |W| + 1/2)^2 \\ &\leqslant C\left(\frac{S_k^2}{k} + \frac{S^2}{n} + 1\right) \end{aligned}$$

gdje je C univerzalna konstanta.

Sada, vidimo da su zadovoljeni svi uvjeti da primijenimo teorem 2.2 na slučajnu varijablu $\widetilde{W}$, odnosno, možemo konstruirati verziju slučajne varijable $\widetilde{W}$ i $Z \sim N(0, \sigma^2)$ na istom vjerojatnosnom prostoru takve da za sve θ,

$$\mathbb{E}[\exp(\theta|\widetilde{W} - Z|)] \leqslant 2\mathbb{E}[\exp(2\theta^2\sigma^{-2}(T - \sigma^2)^2].$$

Jer vrijednost $\widetilde{W}$ određuje vrijednost W, možemo konstruirati verziju slučajne varijable W na istom prostoru takvu da je $|W - \widetilde{W}| \leqslant 1$. Slijedi

$$\mathbb{E}[\exp(\theta|W - Z|)] \leqslant 2\mathbb{E}[\exp(|\theta| + 2\theta^2\sigma^{-2}(T - \sigma^2)^2)].$$

Koristeći gore dokazanu ogradu na $(T - \sigma^2)^2/\sigma^2$, dobivamo

$$\mathbb{E}[\exp(\theta|W - Z|)] \leqslant 2\exp(|\theta| + C\theta^2 S^2/n + C\theta^2)\mathbb{E}[\exp(C\theta^2 S_k^2/k)],$$

gdje je C univerzalna konstanta. Sada smo gotovi primjenom leme 2.11.

□

2.3 Sparivanje s Brownovim mostom

Prelazimo na dokaz teorema 2.3, koji sparuje uvjetovanu slučajnu šetnju s Brownovim mostom. Teorem se može iskazati i na sljedeći način.

Teorem 2.12. *Koristimo notaciju teorema 2.8. Postoje pozitivne univerzalne konstante C, K i λ_0 takve da vrijedi sljedeće: za sve $n \geqslant 2$ i bilo koju moguću vrijednost S_n, moguće je konstruirati verzije slučajnih varijabli $W_0, W_1, \ldots, W_n$ i normalne slučajne varijable $Z_0, \ldots, Z_n$ s očekivanjem nula i*

$$\mathrm{Cov}(Z_i, Z_j) = \frac{(i \wedge j)(n - (i \vee j)}{n}$$

na istom vjerojatnosnom prostoru takve za sve $0 < \lambda < \lambda_0$, vrijedi

$$\mathbb{E}[\exp(\lambda \max_{i \leqslant n} |W_i - Z_i|)] \leqslant \exp\left(C \log n + \frac{K\lambda^2 S_n^2}{n}\right).$$

Prije dokaza teorema, heuristički ćemo objasniti ideju dokaza. Ideja je koristiti indukciju. Naime, fiksirati ćemo a, moguću vrijednost S_n, te ćemo fiksirati n i $n/3 \leqslant k \leqslant 2n/3$. Nakon toga, uzorkovati ćemo vrijednosti Z_k i S_k te onda slijeva i zdesna generirati odgovarajuću slučajnu šetnju, odnosno, normalni slučajni vektor. Onda ćemo te šetnje spojiti na odgovarajući način, u smislu da će dobivena šetnja, odnosno, dobivene normalne slučajne varijable imati traženu distribuciju i svojstva. To ćemo moći zbog ocjena iz prethodno dokazanih lema.

Također, treba naglasiti da na prvu nije očito da je ova formulacija teorema ekvivalentna s prethodnom formulacijom u teoremu 2.3. No, uočimo, ako varijable $Z_1, \ldots, Z_n$ skaliramo s $\sqrt{n}$ te ih složimo na $[0, 1]$, uniformno, dobijemo niz na $[0, 1]$ koji ima kovarijacijsku strukturu kao i Brownov most. Sada jednako kao u Lévyjevoj konstrukciji Brownovog gibanja, možemo nastaviti s dijadskim konstrukcijama i linearnim interpolacijama te se pokaže da dobijemo proces koji je upravo Brownov most.

Dokaz. Uzmimo univerzalnu konstantu α_0 iz leme 2.11 i konstante c i θ_0 iz teorema 2.8. Uzmimo

$$K = 8c, \ \lambda_0 \leqslant \sqrt{\frac{\alpha_0}{16c}} \wedge \frac{\theta_0}{2}, \ C \geqslant \frac{1 + \log 2}{\log(3/2)}.$$

Dokaz provodimo indukcijom po n. Za svaki n i svaku moguću vrijednost a od S_n, neka je $f_a^n(\mathbf{s})$ diskretna vjerojatnosna gustoća slučajnog vektora $(S_0, S_1, \ldots, S_n)$. Uočimo, taj vektor je uniformno distribuiran nad $\mathcal{A}_a^n$ gdje je

$$\mathcal{A}_a^n = \{\mathbf{s} \in \mathbb{Z}^{n+1} \colon s_0 = 0, s_n = a, \text{ i } |s_i - s_{i-1}| = 1 \text{ za sve } i\}.$$

Dakle, za svaki $\mathbf{s} \in \mathcal{A}_a^n$,

$$f_a^n(\mathbf{s}) = \frac{1}{|\mathcal{A}_a^n|}.$$

Neka je $\phi^n(\mathbf{z})$ gustoća normalnog slučajnog vektora $(Z_0, \ldots, Z_n)$ s očekivanjem nula i kovarijacijskom matricom iz tvrdnje teorema.

Želimo dokazati da za svaki n i svaku moguću vrijednost a od S_n možemo konstruirati zajedničku gustoću $\rho_a^n(\mathbf{s}, \mathbf{z})$ na $\mathbb{Z}^{n+1} \times \mathbb{R}^{n+1}$ takvu da je

$$\int \rho_a^n(\mathbf{s}, \mathbf{z}) \mathrm{d}\mathbf{z} = f_a^n(\mathbf{s}), \ \int \rho_a^n(\mathbf{s}, \mathbf{z}) \mathrm{d}\mathbf{s} = \phi^n(\mathbf{z})$$

te za sve $\lambda < \lambda_0$ vrijedi

$$\int \exp\left(\lambda \max_{i \leqslant n} \left| s_i - \frac{ia}{n} - z_i \right|\right) \rho_a^n(\mathbf{s}, \mathbf{z}) \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{z} \leqslant \exp\left(C \log n + \frac{K\lambda^2 a^2}{n}\right).$$

Pretpostavimo da se ρ_a^k može konstruirati za sve $k = 1, \ldots, n-1$ i sve odgovarajuće a. Sada konstruiramo ρ_a^n kada je a moguća vrijednost od S_n.

Prvo, fiksiramo a, vrijednost od S_n i indeks k takav da je $n/3 \leqslant k \leqslant 2n/3$, npr. $k = [n/2]$. Uz dano $S_n = a$, neka je $g_a^{n,k}(s)$ uvjetna gustoća od S_k. Brojanjem, dobivamo

$$g_a^{n,k}(s) = \frac{|\mathcal{A}_s^k||\mathcal{A}_{a-s}^{n-k}|}{|\mathcal{A}_a^n|}.$$

Neka je $h^{n,k}(z)$ gustoća normalno distribuirane slučajne varijable s očekivanjem 0 i varijancom $k(n-k)/n$. Po teoremu 2.8 i nejednakosti $\exp|x| \leqslant \exp(x) + \exp(-x)$, vidimo da postoji gustoća $\psi_a^{n,k}(s, z)$ na $\mathbb{Z} \times \mathbb{R}$ takva da je

$$\int \psi_a^{n,k}(s, z) \mathrm{d}z = g_a^{n,k}(s), \int \psi_a^{n,k}(s, z) \mathrm{d}s = h^{n,k}(z)$$

i za sve $0 < \theta \leqslant \theta_0$,

$$\int \exp\left(\theta \left| s - \frac{ka}{n} - z \right|\right) \psi_a^{n,k}(s, z) \mathrm{d}s\mathrm{d}z \leqslant \exp\left(1 + \frac{c\theta^2 a^2}{n}\right). \tag{2.4}$$

Sada definiramo $\gamma_a^n \colon \mathbb{Z} \times \mathbb{R} \times \mathbb{Z}^{k+1} \times \mathbb{R}^{k+1} \times \mathbb{Z}^{n-k+1} \times \mathbb{R}^{n-k+1} \to \mathbb{R}$ formulom

$$\gamma_a^n(s, z, \mathbf{s}, \mathbf{z}, \mathbf{s}', \mathbf{z}') = \psi_a^{n,k}(s, z) \rho_s^k(\mathbf{s}, \mathbf{z}) \rho_{a-s}^{n-k}(\mathbf{s}', \mathbf{z}').$$

Integrirajući po $\mathbf{s}', \mathbf{z}'$ pa po $\mathbf{s}, \mathbf{z}$ te naposljetku po s i z, lako se uvjerimo da je γ_a^n uistinu vjerojatnosna gustoća. Neka je $(S, Z, \mathbf{S}, \mathbf{Z}, \mathbf{S}', \mathbf{Z}')$ slučajan vektor s gustoćom

γ_a^n. Možemo shvatiti kao da prvo generiramo (S, Z) prema $\psi_a^{n,k}$, a nakon toga uz dano $S = s, Z = z$ nezavisno generiramo parove $(\mathbf{S}, \mathbf{Z})$ i $(\mathbf{S}', \mathbf{Z}')$ prema gustoćama ρ_s^k i ρ_{a-s}^{n-k}, redom.

Sada definiramo slučajne vektore $\mathbf{Y} \in \mathbb{R}^{n+1}$ i $\mathbf{U} \in \mathbb{Z}^{n+1}$. Za $i \leqslant k$, neka je

$$Y_i = Z_i + \frac{i}{k} Z,$$

a za $i \geqslant k$, neka je

$$Y_i = Z'_{i-k} + \frac{n-i}{n-k} Z.$$

Uočimo, definicije se poklapaju na $i = k$ jer je $Z_k = Z'_0 = 0$. Nadalje, definiramo $U_i = S_i$ za $i \leqslant k$ te $U_i = S + S'_{i-k}$ za $i \geqslant k$. Opet, definicije se poklapaju na $i = k$ jer je $S_k = S$ i $S'_0 = 0$. Tvrdimo da je gustoća vektora $(\mathbf{U}, \mathbf{Y})$ upravo ρ_a^n.

Prvo, promatramo marginalnu gustoću varijable $\mathbf{U}$. Direktnim računom dobivamo

$$\int \gamma_a^n(s, z, \mathbf{s}, \mathbf{z}, \mathbf{s}', \mathbf{z}') \mathrm{d}\mathbf{z} \mathrm{d}\mathbf{z}' \mathrm{d}z = g_a^{n,k}(s) f_s^k(\mathbf{s}) f_{a-s}^{n-k}(\mathbf{s}').$$

Uvrštavanjem, slijedi

$$g_a^{n,k}(s) f_s^k(\mathbf{s}) f_{a-s}^{n-k}(\mathbf{s}') = \frac{|\mathcal{A}_s^k||\mathcal{A}_{a-s}^{n-k}|}{|\mathcal{A}_a^n|} \frac{1}{|\mathcal{A}_s^k|} \frac{1}{|\mathcal{A}_{a-s}^{n-k}|} = \frac{1}{|\mathcal{A}_a^n|}.$$

Jer postoji bijekcija između $(S, \mathbf{S}, \mathbf{S}')$ te $\mathbf{U}$ može poprimiti bilo koju vrijednost u $\mathcal{A}_a^n$, dokazali smo da $\mathbf{U}$ ima marginalnu gustoću f_a^n.

Drugo, promatramo marginalnu gustoću varijable $\mathbf{Y}$. Tvrdimo da su Z, $\mathbf{Z}$ i $\mathbf{Z}'$ nezavisne s gustoćama $h^{n,k}, \phi^k$ i ϕ^n, redom. Opet, direktnim računom slijedi

$$\begin{aligned}
\int \gamma_a^n(s, z, \mathbf{s}, \mathbf{z}, \mathbf{s}', \mathbf{z}') \mathrm{d}\mathbf{s}' \mathrm{d}\mathbf{s} \mathrm{d}s &= \int \psi_a^{n,k}(s, z) \rho_s^k(\mathbf{s}, \mathbf{z}) \rho_{a-s}^{n-k}(\mathbf{s}', \mathbf{z}') \mathrm{d}\mathbf{s}' \mathrm{d}\mathbf{s} \mathrm{d}s \\
&= \phi^{n-k}(\mathbf{z}') \int \psi_a^n(s, z) \rho_s^k(\mathbf{s}, \mathbf{z}) \mathrm{d}\mathbf{s} \mathrm{d}s \\
&= \psi^{n-k}(\mathbf{z}') \phi^k(\mathbf{z}) \int \psi_a^n(s, z) \mathrm{d}s \\
&= \psi^{n-k}(\mathbf{z}') \psi^k(\mathbf{z}) h^{n,k}(z).
\end{aligned}$$

Dakle, Y je normalan slučajni vektor s očekivanjem nula. Jedino preostaje za izračunati $\mathrm{Cov}(Y_i, Y_j)$ što se lagano i napravi. Dakle, $\mathbf{Y} \sim \phi^n$.

Naposljetku, preostaje za dokazati eksponencijalnu ocjenu. Za $0 \leqslant i \leqslant n$, neka je

$$W_i = U_i - \frac{ia}{n}.$$

Trebamo dokazati da za $0 < \lambda < \lambda_0$,

$$\mathbb{E}[\exp(\lambda \max_{i \leqslant n} |W_i - Y_i|)] \leqslant \exp\left(C \log n + \frac{K\lambda^2 a}{n}\right),$$

gdje su C, K i λ_0 definirane na početku dokaza. Neka je

$$T_L = \max_{i \leqslant k} \left| S_i - \frac{iS}{k} - Z_i \right|, T_R = \max_{i \geqslant k} \left| S'_{i-k} - \frac{i-k}{n-k}(a-S) - Z'_{i-k} \right|$$

te

$$T = \left| S - \frac{ka}{n} - Z \right|.$$

Tvrdimo da je

$$\max_{i \leqslant n} |W_i - Y_i| \leqslant \max\{T_L, T_R\} + T.$$

Naime, za $i \leqslant k$, vrijedi

$$\begin{aligned}
|W_i - Y_i| &= \left| S_i - \frac{ia}{n} - \left(Z_i + \frac{iZ}{k} \right) \right| \\
&\leqslant \left| S_i - \frac{iS}{k} - Z_i \right| + \left| \frac{iS}{k} - \frac{ia}{n} - \frac{iZ}{k} \right| \\
&\leqslant T_L + \frac{i}{k} T \leqslant T_L + T.
\end{aligned}$$

Slično, za $i \geqslant k$,

$$\begin{aligned}
|W_i - Y_i| &= \left| S + S'_{i-k} - \frac{ia}{n} - \left(Z'_{i-k} + \frac{n-i}{n-k} Z \right) \right| \\
&\leqslant \left| S'_{i-k} - \frac{i-k}{n-k}(a-S) - Z'_{i-k} \right| + \left| S + \frac{i-k}{n-k}(a-S) - \frac{ia}{n} - \frac{n-i}{n-k} Z \right| \\
&= \left| S'_{i-k} - \frac{i-k}{n-k}(a-S) - Z'_{i-k} \right| + \frac{n-i}{n-k} \left| S - \frac{ka}{n} - Z \right| \\
&\leqslant T_R + T.
\end{aligned}$$

Sada, fiksiramo $\lambda < \lambda_0$. Koristeći $\exp(x \vee y) \leqslant \exp x + \exp y$, dobivamo

$$\exp(\lambda \max_{i \leqslant n} |W_i - Y_i|) \leqslant \exp(\lambda T_L + \lambda T) + \exp(\lambda T_R + \lambda T).$$

Po konstrukciji, lagano se provjeri da uvjetno na $S = s, Z = z$, $(\mathbf{S}, \mathbf{Z})$ ima uvjetnu gustoću ρ_s^k. Po pretpostavci indukcije, vrijedi

$$\mathbb{E}[\exp(\lambda T_L) \mid S, Z] \leqslant \exp\left(C \log k + \frac{K\lambda^2 S^2}{k}\right).$$

Jer su funkcije izvodnice momenata varijabli T_L i T konačne svugdje (doprinos od S možemo ograničiti konstantom, maksimum možemo ograničiti sumom, a Z je multivarijatna normalna), možemo primijeniti Cauchy-Schwarzovu nejednakost koja daje

$$\begin{aligned}\mathbb{E}[\exp(\lambda T_L)\exp(\lambda T)] &\leqslant \Big(\mathbb{E}[\mathbb{E}[\exp(\lambda T_L) \mid S, Z]^2]\mathbb{E}[\exp(2\lambda T)]\Big)^{1/2} \\ &\leqslant \exp(C\log k)\left(\mathbb{E}\left[\exp\left(\frac{2K\lambda^2 S^2}{k}\right)\right]\mathbb{E}[\exp(2\lambda T)]\right)^{1/2},\end{aligned}$$

pri čemu smo u prvoj nejednakosti koristili i Jensenovu nejednakost i svojstva uvjetnog očekivanja. Želimo primijeniti lemu 2.11 da bi ograničili prvi izraz unutar zagrade. Po definiciji,

$$2K\lambda^2 \leqslant 16c\frac{\alpha_0}{16c} = \alpha_0$$

te $n/3 \leqslant k \leqslant 2n/3$. Zbog toga, uistinu možemo primijeniti lemu 2.11 da bi dobili

$$\mathbb{E}\left[\exp\left(\frac{2K\lambda^2 S^2}{k}\right)\right] \leqslant \exp\left(1 + \frac{3K\lambda^2 a^2}{2n}\right).$$

Nadalje, opet po definiciji imamo $2\lambda \leqslant \theta_0$ pa koristeći nejednakost (2.4) s $\theta = 2\lambda$ dobivamo

$$\mathbb{E}[\exp(2\lambda T)] \leqslant \exp\left(1 + \frac{4c\lambda^2 a^2}{n}\right).$$

Spajajući što smo dobili u zadnja tri koraka, imamo

$$\mathbb{E}[\exp(\lambda T_L + \lambda T)] \leqslant \exp\left(C\log k + 1 + \frac{(3K+8c)\lambda^2 a^2}{4n}\right).$$

Iz definicije, $3K + 8c = 4K$. Jer je $n/3 \leqslant k \leqslant 2n/3$, imamo

$$\log k = \log n - \log(n/k) \leqslant \log n - \log(3/2)$$

dakle

$$\mathbb{E}[\exp(\lambda T_L + \lambda T)] \leqslant 2^{1/2}\exp\left(C\log n - C\log(3/2) + 1 + \frac{K\lambda^2 a^2}{n}\right).$$

Po simetriji, potpuno analognu ocjenu možemo dobiti za $\mathbb{E}[\exp(\lambda T_R + \lambda T)]$. Kombinirajući to s prethodno dokazanom ocjenom za $\mathbb{E}[\exp(\lambda \max_{i\leqslant n}|W_i - Y_i|)]$, dobili smo

$$\mathbb{E}[\exp(\lambda \max_{i\leqslant n}|W_i - Y_i|)] \leqslant 2\exp\left(C\log n - C\log(3/2) + 1 + \frac{K\lambda^2 a^2}{n}\right).$$

Naposljetku, zbog definicije konstante C, imamo

$$-C\log(3/2)+1+\log 2 \leqslant 0$$

čime smo dovršili korak indukcije.

Preostaje samo komentirati bazu indukcije. Za slučaj $n=2$, uzmimo da su (W_0, W_1, W_2) te (Z_0, Z_1, Z_2) nezavisni te uzmimo da je C dovoljno velik, a λ_0 dovoljno malen da svejedno vrijedi ograda.

□

2.4 Dovršetak dokaza

Preostaje za dokazati KMT teorem za jednostavnu simetričnu slučajnu šetnju. Prvo ćemo kao lemu dokazati verziju za „konačan n”, a onda i generalni teorem.

Lema 2.13. *Postoje univerzalne konstante $B>1$ i $\lambda>0$ takve da vrijedi sljedeće: neka je n prirodan broj i neka su $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ nezavisne jednako distribuirane slučajne varijable, uniformne na $\{-1,1\}$. Neka je $S_k=\sum_{i=1}^k \varepsilon_i$, za $k=0,1,\dots,n$. Moguće je konstruirati verziju niza $(S_k)_{k\leqslant n}$ i normalno distribuirane slučajne varijable $(Z_k)_{k\leqslant n}$ s očekivanjem 0 i $\mathrm{Cov}(Z_i, Z_j)=i\wedge j$ na istom vjerojatnosnom prostoru takve da je $\mathbb{E}[\exp(\lambda|S_n-Z_n|)]\leqslant B$ te*

$$\mathbb{E}[\exp(\lambda \max_{k\leqslant n}|S_k-Z_k|)]\leqslant B\exp(B\log n).$$

Dokaz. Uzmimo univerzalne konstante θ_0 i κ iz teorema 2.7 i C, K i λ_0 iz teorema 2.12. Odaberimo λ dovoljno malen takav da je

$$\lambda<\frac{\theta_0\wedge\lambda_0}{2} \text{ i } 16K\lambda^2<1.$$

Neka su f_a^n, ρ_a^n i ϕ^n gustoće kao u dokazu teorema 2.12. Neka su g^n i h^n gustoće slučajnih varijabli S_n i Z_n, redom. Po teoremu 2.7 i odabiru λ, postoji zajednička gustoća ψ^n na $\mathbb{Z}\times\mathbb{R}$ takva da je

$$\int \psi^n(s,z)\mathrm{d}z=g^n(s), \quad \int \psi^n(s,z)\mathrm{d}s=h^n(z),$$

te

$$\int \exp(2\lambda|s-z|)\psi^n(s,z)\mathrm{d}s\mathrm{d}z\leqslant\kappa. \tag{2.5}$$

Definiramo $\gamma^n\colon \mathbb{Z}\times\mathbb{R}\times\mathbb{Z}^{n+1}\times\mathbb{R}^{n+1}\to\mathbb{R}$ kao

$$\gamma^n(s,z,\mathbf{s},\mathbf{z})=\psi^n(s,z)\rho_s^n(\mathbf{s},\mathbf{z}).$$

Lagano se provjeri da je riječ o funkciji gustoće. Neka je $(S, Z, \mathbf{S}, \mathbf{Z})$ slučajan vektor s gustoćom γ^n. Kao i u dokazu teorema 2.12, pokaže se da je zajednička gustoća vektora $(Z, \mathbf{Z})$ jednaka $h^n(z)\phi^n(\mathbf{z})$. Definirajmo slučajni vektor $\mathbf{Y} = (Y_0, \ldots, Y_n)$ kao

$$Y_i = Z_i + \frac{i}{n} Z.$$

Zbog nezavisnosti među Z i $\mathbf{Z}$ i njihovih distribucija, $\mathbf{Y}$ je normalno distribuiran vektor s očekivanjem nula te vrijedi $\mathrm{Cov}(Y_i, Y_j) = i \wedge j$.

Integrirajući z i $\mathbf{z}$ vani, vidimo da je zajednička gustoća vektora $(S, \mathbf{S})$ jednaka $g^n(s) f_s^n(\mathbf{s})$. Jasno vidimo da je marginalna distribucija vektora $\mathbf{S}$ jednaka distribucije jednostavne simetrične slučajne šetnje do vremena n.

Dokažimo sada da par $(\mathbf{S}, \mathbf{Y})$ zadovoljava uvjete teorema. Neka je $W_i = S_i - iS/n$. Za bilo koji $i \leqslant n$, vrijedi

$$\begin{aligned} |S_i - Y_i| &= \left| S_i - \left(Z_i + \frac{i}{n} Z \right) \right| \\ &\leqslant |W_i - Z_i| + \frac{i}{n} |S - Z|. \end{aligned}$$

Uočimo da je uvjetna gustoća vektora $(\mathbf{S}, \mathbf{Z})$ uz dano $(S, Z) = (s, z)$ jednaka ρ_s^n. Jer je $\lambda < \lambda_0$, iz konstrukcije ρ_s^n, vrijedi

$$\mathbb{E}[\exp(\lambda \max_{i \leqslant n} |W_i - Z_i|) \mid S, Z] \leqslant \exp\left(C \log n + \frac{K\lambda^2 S^2}{n} \right).$$

Sada, koristeći Cauchy-Schwarzovu nejednakost i (2.5), dobivamo

$$\begin{aligned} &\mathbb{E}[\exp(\lambda \max_{i \leqslant n} |S_i - Y_i|)] \\ &\quad \leqslant [\mathbb{E}[\mathbb{E}[\exp(\lambda \max_{i \leqslant n}) |W_i - Z_i| \mid S, Z]^2] \mathbb{E}[\exp(2\lambda |S - Z|)]]^{1/2} \\ &\quad \leqslant \exp(C \log n) [\kappa \mathbb{E}[\exp(2K\lambda^2 S^2 / n)]]^{1/2}. \end{aligned}$$

Iz nejednakosti (2.3) i odabira λ, dokazali smo glavni dio teorema. Za drugu tvrdnju, dovoljno je uočiti da vrijedi (2.5) te $Y_n = Z$, jer je $Z_n = 0$.

□

Sada možemo preći na dokaz glavnog teorema.

Teorem 2.14 (KMT za SRW). *Neka su $\varepsilon_1, \varepsilon_2, \ldots$ nezavisne jednako distribuirane slučajne varijable uniformne na $\{-1, 1\}$. Za svaki k, neka je $S_k = \sum_{i=1}^k \varepsilon_i$. Moguće je konstruirati* *verziju niza $(S_k)_{k\geqslant 0}$ i standardno Brownovo gibanje $(B_t)_{t\geqslant 0}$ na istom vjerojatnosnom prostoru takve da za sve n i sve $x \geqslant 0$,*

$$\mathbb{P}(\max_{k\leqslant n} |S_k - B_k| \geqslant C\log n + x) \leqslant Ke^{-\lambda x},$$

pri čemu C, K i λ ne ovise o n.

Dokaz. Za $r = 1, 2, \ldots$, neka je $m_r = 2^{2^r}$ i $n_r = m_r - m_{r-1}$. Za svaki r, neka je $(S_k^{(r)}, Z_k^{(r)})_{0\leqslant k\leqslant n_r}$ slučajan vektor koji zadovoljava zaključak leme 2.13 i pretpostavimo da su ti vektori nezavisni. Induktivno, definiramo beskonačan niz $(S_k, Z_k)_{k\geqslant 0}$ na sljedeći način. Neka je $S_k = S_k^{(1)}$ i $Z_k = Z_k^{(1)}$ za $k \leqslant m_1$. Nakon definiranja $(S_k, Z_k)_{k\leqslant m_{r-1}}$, definiramo $(S_k, Z_k)_{m_{r-1}<k\leqslant m_r}$ kao

$$S_k = S^{(r)}_{k-m_{r-1}} + S_{m_{r-1}}, Z_k = Z^{(r)}_{k-m_{r-1}} + Z_{m_{r-1}}.$$

Jer su prirasti nezavisni, S_k i Z_k su slučajne šetnje s binarnim, odnosno, normalnim prirastima, redom.

Sada, uzmimo konstante B i λ iz leme 2.13. Za svaki r, iz leme 2.13 i nezavisnosti, vrijedi

$$\begin{aligned}\mathbb{E}[\exp(\lambda|S_{m_r} - Z_{m_r}|)] &\leqslant \mathbb{E}\left[\exp\left(\lambda \prod_{\ell=1}^{r} |S^{(\ell)}_{n_\ell} - Z^{(\ell)}_{n_\ell}|\right)\right] \\ &= \prod_{\ell=1}^{r} \mathbb{E}[\exp(\lambda|S^{(\ell)}_{n_\ell} - Z^{(\ell)}_{n_\ell}|)] \leqslant B^r. \end{aligned} \tag{2.6}$$

Neka je

$$C = \frac{1}{1 - \frac{\exp(-\frac{1}{2}B\log 4)}{B}}.$$

Dokazati ćemo indukcijom da za sve r, vrijedi

$$\mathbb{E}[\exp(\lambda \max_{k\leqslant m_r} |S_k - Z_k|)] \leqslant CB^r \exp(B\log m_r). \tag{2.7}$$

Po lemi 2.13 i jer je $B > 1$ i $C > 1$, tvrdnja vrijedi za $r = 1$. Pretpostavimo da tvrdnja vrijedi za $r - 1$. Iz nejednakost $\exp(x \vee y) \leqslant \exp x + \exp y$, imamo

$$\begin{aligned}\mathbb{E}[\exp(\lambda \max_{k\leqslant m_r} |S_k - Z_k|)] &\leqslant \mathbb{E}[\exp(\lambda \max_{m_{r-1}\leqslant k\leqslant m_r} |S_k - Z_k|)] \\ &\quad + \mathbb{E}[\exp(\lambda \max_{k\leqslant m_{r-1}} |S_k - Z_k|)]. \end{aligned} \tag{2.8}$$

Promotrimo prvi član na desnoj strani. Vrijedi

$$\max_{m_{r-1} \leqslant k \leqslant m_r} |S_k - Z_k| \leqslant \max_{1 \leqslant j \leqslant n_r} |S_j^{(r)} - Z_j^{(r)}| + |S_{m_{r-1}} - Z_{m_{r-1}}|.$$

Dakle, zbog nezavisnosti, leme 2.13 i nejednakosti (2.6), vrijedi

$$\mathbb{E}[\exp(\lambda \max_{m_{r-1} \leqslant k \leqslant m_r} |S_k - Z_k|)] \leqslant B^r \exp(B \log m_r).$$

Po pretpostavci indukcije i jer je $m_r = m_{r-1}^2$, drugi član u (2.8) je ograđen kao

$$\begin{aligned} \mathbb{E}[\exp(\lambda \max_{k \leqslant m_{r-1}} |S_k - Z_k|)] &\leqslant CB^{r-1} \exp(B \log m_{r-1}) \\ &= CB^{r-1} \exp\left(\frac{B \log m_r}{2}\right). \end{aligned}$$

Spajajući obje ocjene, dobivamo

$$\mathbb{E}[\exp(\lambda \max_{k \leqslant m_r} |S_k - Z_k|)] \leqslant B^r \exp(B \log m_r) \left(1 + \frac{C}{B} \exp\left(-\frac{B \log m_r}{2}\right)\right).$$

Iz definicije konstante C, nije teško provjeriti da je izraz u zagradama ograničen odozgo s C, čime dovršavamo korak indukcije.

Dakle, dokazali smo (2.7). Jer je $r \lesssim \log m_r$, slijedi da postoji konstanta K takva da za sve r,

$$\mathbb{E}[\exp(\lambda \max_{k \leqslant m_r} |S_k - Z_k|)] \leqslant K \exp(K \log m_r).$$

Preostaje dokazati istu nejednakost za proizvoljan n umjesto m_r. Neka je $n \geqslant 2$ proizvoljan te neka je r takav da je $m_{r-1} \leqslant n \leqslant m_r$. Tada $m_r = m_{r-1}^2 \leqslant n^2$. Dakle,

$$\begin{aligned} \mathbb{E}[\exp(\lambda \max_{k \leqslant n} |S_k - Z_k|)] &\leqslant \mathbb{E}[\exp(\lambda \max_{k \leqslant m_r} |S_k - Z_k|)] \\ &\leqslant K \exp(K \log m_r) \leqslant K \exp(2K \log n). \end{aligned}$$

Ocjena na vjerojatnost iz tvrdnje teorema se sada lagano dobije primjenom Markovljeve nejednakosti.

□

Bibliografija

[1] S. Chatterjee, *A new approach to strong embeddings*, Probability Theory and Related Fields **152** (2012), 231–264.

[2] R. Durrett, *Probability: Theory and Examples*, Cambridge University Press, 2019.

[3] L. H. Y. Chen i L. Goldstein i Q.-M. Shao, *Normal Approximation by Stein's Method*, Springer Berlin Heidelberg, 2011.

[4] A. D. Barbour i L. H. Y. Chen, *An Introduction to Stein's Method*, World Scientific Publishing, Singapore University Press, 2005.

[5] Komlós J. i Major P. i Tusnády G., *An approximation of partial sums of independent RV'-s, and the sample DF. I*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **32** (1975), br. 1–2, 111–131.

[6] O. Kallenberg, *Foundations of Modern Probability*, Springer International Publishing, 2021.

[7] N. Ross, *Fundamentals of Stein's method*, Probability Surveys **8** (2011), 210 – 293.

Sažetak

Ovaj rad proučava normalnu aproksimaciju Steinovom metodom te dokaz Komlós-Major-Tusnády (KMT) teorema za jednostavnu simetričnu slučajnu šetnju.

Rad inicijalno uvodi nekoliko metrika na prostoru vjerojatnosnih distribucija i dokazuje neke osnovne tvrdnje o njima. Rad posebno dokazuje Steinovu lemu koja karakterizira normalnu distribuciju i omogućava sve aproksimacije koje su korištene u ostatku rada.

Nakon toga, u Poglavlju 1 daje dokaze nekoliko graničnih teorema. Dokazani su klasični granični teoremi u slučaju zbroja nezavisnih slučajnih varijabli, a prikazana su i četiri različita konteksta u kojima se može koristiti Steinova metoda. Snaga ove metode je ilustrirana dokazom centralnog graničnog teorem za broj trokuta te broj izoliranih vrhova u Erdős–Rényijevom grafu. U drugom dijelu rada, kroz niz lema i pomoćnih teorema, dokazan je poznati KMT teorem za jednostavnu simetričnu slučajnu šetnju.

Summary

This thesis studies the normal approximation by Stein's method and the proof of the Komlós-Major-Tusnády (KMT) for the simple symmetric random walk.

The thesis initially introduces several metrics on the space of probability measures and proves some basic facts about them. Specially, this thesis proves Stein's lemma which characterizes the normal distribution and is the basis of all the approximations used in the remainder of the thesis.

After that, in Chapter 1, it gives proofs of several limit theorems. Classical limit theorems in the case of sum of independent random variables are proved, and four different contexts in which Stein's method can be used are presented. The power of the method is illustrated by the proof of the central limit theorem for the number of triangles and isolated vertices in the Erdős–Rényi graph. In the second part of the thesis, through a series of lemmas and auxiliary theorems, the well-known KMT theorem for the simple symmetric random walk is proved.

Životopis

Rođen sam 15. ožujka 2001. godine u Šibeniku. Nakon završetka Osnovne škole Petra Krešimira IV. u Šibeniku upisujem XV. gimnaziju u Zagrebu. Tijekom osnovne i srednje škole, sudjelovao sam na državnim i međunarodnim natjecanjima iz matematike, među kojima ističem Srednjoeuropsku matematičku olimpijadu 2018. i Romanian Master of Mathematics 2019.

Prijediplomski sveučilišni studij Matematika na Matematičkom odsjeku Prirodoslovno-matematičkog fakulteta u Zagrebu upisujem 2019. i završavam 2022. nakon čega upisujem diplomski studij Matematičke statistike na istom fakultetu. Tijekom studija bio sam demonstrator iz kolegija Linearna algebra 1 i 2, Diskretna matematika, Vjerojatnost, Integrali funkcija više varijabli, Algebarske strukture, Kompleksna analiza te Statistika.

Također, tijekom studija sam sudjelovao u radu udruge Mladi nadareni matematičari „Marin Getadlić" te sam držao pripreme za natjecanja u XV. gimnaziji. Za izniman uspjeh tijekom studija dodijeljene su mi dvije nagrade Matematičkog odsjeka, Dekanova nagrada te Stipendija Grada Zagreba za izvrsnost.