

Primjena statističkih metoda u analizi Brownovog gibanja

Željeznjak, Ivan Vatroslav

Master's thesis / Diplomski rad

2023

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

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:599059>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-02-11**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

Ivan Vatroslav Željeznjak

Primjena statističkih metoda u analizi
Brownovog gibanja

Diplomski rad

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

INTEGRIRANI PREDDIPLOMSKI I DIPLOMSKI SVEUČILIŠNI STUDIJ
FIZIKA I INFORMATIKA; SMJER NASTAVNIČKI

Ivan Vatroslav Željeznjak

Diplomski rad

**Primjena statističkih metoda u analizi
Brownovog gibanja**

Voditelj diplomskog rada: doc. dr. sc. Petar Žugec

Ocjena diplomskog rada: _____

Povjerenstvo: 1. _____

2. _____

3. _____

Datum polaganja: _____

Zagreb, 2023.

Sažetak

Stohastički procesi su matematički objekti koji opisuju procese koji podliježu zakonima teorije vjerojatnosti a u sebi sadrže elemente slučajnosti ili neuređenosti te se primjenjuju kod matematičkog modeliranja fenomena u mnoštvu disciplina. Jedno od najvažnijih stohastičkih procesa je Brownovo gibanje čija je makroskopska realizacija difuzija, koju računalnim simulacijama i statističkim metodama iskazujemo kao graničnu raspodjelu slučajnih šetnji u slučaju infinitezimalnih koraka u konačnom intervalu.

Ključne riječi: teorija vjerojatnosti, stohastički procesi, Brownovo gibanje, slučajna šetnja, difuzija.

Application of statistical methods in analysis of Brownian motion

Abstract

Stochastic processes are mathematical objects that describe processes that are subject to the laws of probability theory and contain elements of randomness or disorder, and are used in the mathematical modeling of phenomena in many disciplines. One of the most important stochastic processes is Brownian motion, the macroscopic realization of which is diffusion, which we express with computer simulations and statistical methods as the limiting distribution of random walks in the case of infinitesimal steps in a finite interval.

Keywords: probability theory, stochastic processes, Brownian motion, random walk, diffusion.

Sadržaj

1	Uvod	2
1.1	Brownovo otkriće	2
1.2	Termodinamika i kinetičko-molekularna teorija	3
1.3	Einsteinovo objašnjenje	4
2	Slučajna šetnja	5
2.1	Različite okolnosti nasumične šetnje	5
2.2	Generiranje koraka nasumične šetnje	6
2.3	Simulacije nasumičnih šetnji u dvije dimenzije	8
3	Besselova korekcija	11
3.1	Populacije i uzorci	12
3.2	Statistički procjenitelj i njegova svojstva	12
3.2.1	Priistranost	13
3.3	Besselova korekcija - Izvod	13
3.4	Besselova korekcija - računalne simulacije	15
3.5	Podcjenjivanje pristranog procjenitelja	19
4	Centralni granični teorem	21
4.1	Zabune oko centralnog graničnog teorema	21
4.2	Kocke	22
4.3	Primjer CGT na nasumičnoj šetnji	24
5	Pólyjino brojanje	26
5.1	Pólyina šetnja	26
5.2	Pólyine konstante slučajne šetnje	26
6	Nasumična šetnja po sferi	28
6.1	Zarobljena šetnja	28
6.2	Određivanje sfernog koraka	28
6.3	Određivanje kutne varijable	30
6.4	Cikličnost	31
6.5	Analitički primjeri	32
6.6	Generiranje smjerova	33

6.7	Centralni granični teorem na sferi	36
7	Poveznica raspodjela	40
7.1	Fizikalna veza	40
7.2	Izvod intervalne raspodjele	40
8	Difuzija	44
9	Dodatak	47
9.1	Rješenje integrala 7.14 preko teorema o reziduumima	47

1 Uvod

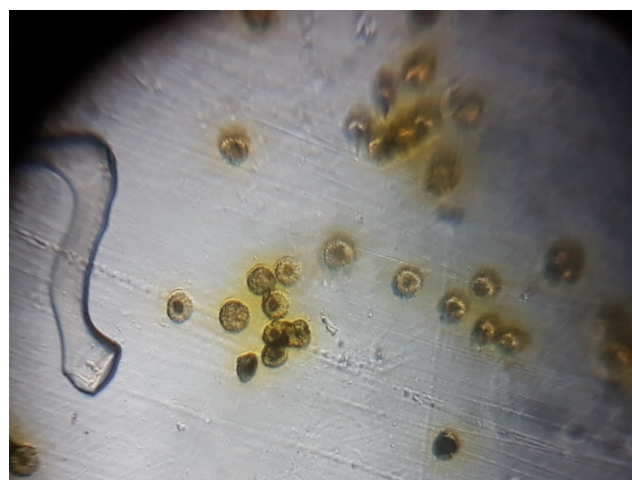
Ljude je od pamtivijeka zanimala struktura materije. Danas nam je dobro poznato da sva materija oko nas se sastoji od diskretnih građevnih jedinica koje nazivamo atomima. Makar su najstariji zapisi koncepta atoma stari preko dvije tisuće godina, teorija atoma je počela biti razmatrana tek u 19. stoljeću, razvojem kemije i kinetičko-molekularne teorije, dok je prvi direktni dokaz postojanja atoma bilo Brownovo gibanje.

1.1 Brownovo otkriće

Godine 1827., škotski botaničar Robert Brown proučavao je mikroskopom (slika 1.1) zrnca peludi suspendirana u vodi (slika 1.2) i zapazio pojavu koju zovemo Brownovo gibanje. Pojavu je slučajno zapazio, proučavajući na koji način peludna zrnca oploduju jajnu stanicu.



Slika 1.1: Brownov mikroskop [1]



Slika 1.2: Uvećani pogled zrnaca peludi pod mikroskopom [2]

To ga je navelo da provede niz istraživanja. Uočio je kako se zrnca ne gibaju jednako po fluidu u kojem se nalaze, već se njihov relativan položaj i sam oblik mijenja s vremenom. Takvo gibanje zrnaca ga je podsjetilo na gibanje mikroskopskih životinja, stoga je pretpostavio kako njihovo gibanje ne proizlazi od samog fluida u kojem se nalaze, već da je samo zrnce "oživjelo" i počelo se gibati. Postupak je ponovio sa zrcima drugih vrsta kao i sa stoljeće starim osušenim uzorcima,

gdje je također opazio jednako gibanje. Kako bi odstranio mogućnost da su zrnca peludi uistinu "oživjela", promatrao je razne usitnjene neorganske materijale i svaki put je opazio isto gibanje. Brown zaključuje kako se čestice sitne tvari koje su suspendirane u tekućini gibaju neovisno jedna o drugoj, bile one organske ili neorganske. Makar nije mogao objasniti ovaj fenomen, njegovo detaljno vođenje bilješki služi kao poučan primjer znanstvene metode. Prošlo je gotovo čitavo stoljeće od Brownovog otkrića do njegovog objašnjenja koje je rezultiralo revolucionarnim napretkom znanosti.

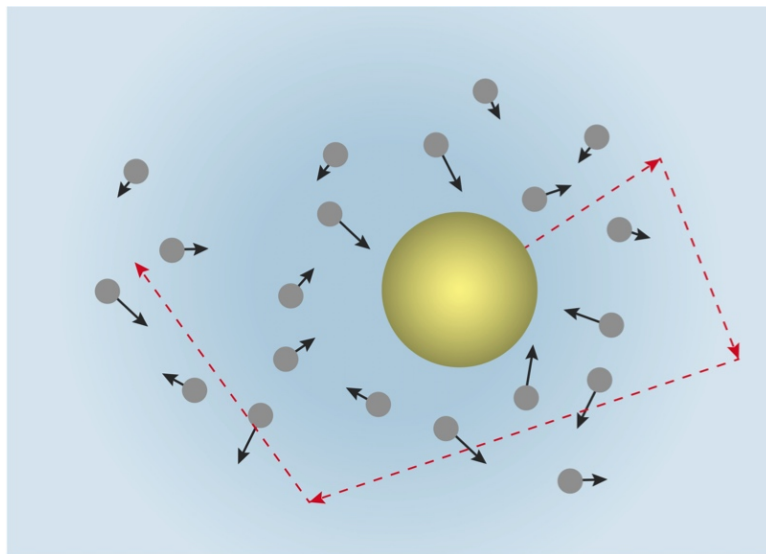
1.2 Termodinamika i kinetičko-molekularna teorija

Termodinamika kao teorija, bila je jedno od vrhunskih dostignuća znanosti 19. stoljeća. Koristeći samo zakone termodinamike, koncepte energije i entropije bilo je moguće opisati mnoštvo pojava. No nije postojalo obrazloženje koje bi opravdalo točnost tih zakona, pogotovo jer se činilo da je termodinamika kontradiktorna s dobro utemeljenom Newtonovom mehanikom. Naime, zbog Newtonovog determinizma svaki bi proces trebao biti potpuno reverzibilan, što nije u skladu s drugim zakonom termodinamike. U pokušaju razrješenja tog problema, Maxwell i Boltzmann su pokazali da se eksperimentalno dobiveni rezultati mogu objasniti statističkim¹ ponašanjem velike skupine čestica koje se neprestano i nasumično gibaju i međusobno sudaraju. No njihova teorija je zapravo dodatno naglasila paradoks o reverzibilnosti jer se temelji na Newtonovoj mehanici. Razrješenje paradoksa ponudio je Boltzmann tako da je postulirao da se svaka zasebna čestica ponaša deterministički, ali da je njihova nakupina podložna zakonima vjerojatnosti koji su asimetrični u vremenu, gdje prolaskom vremena sustav teži u vjerojatnija stanja što dovodi do drugog zakona termodinamike. Znanstvena zajednica je odbacila ovu "statističku mehaniku", koja je tražila da se žrtvuje determinizam za matematičku teoriju za koju nije postojala direktna indikacija da je adekvatna, već samo mehanička analogija dobro utemeljene teorije termodinamike.

¹Takav opis je prvi statistički zakon fizike iz koje se razvila molekularno-kinetička teorija.

1.3 Einsteinovo objašnjenje

Jedan od četiri rada Einsteinove famozne 1905. godine je upravo bio o Brownovom gibanju. Kao pobornik atoma je u tom radu konstruirao model u kojemu bi se moglo opaziti djelovanje atoma koji su po njegovoj procjeni (u to doba) bili presitni i prebrzi za opaziti. Njegova tvrdnja je bila da, nasumično sudaranje atoma s česticom koju možemo opaziti rezultira različitim tlakom na suprotnim stranama i da bi se čestica zbog toga nasumično gibala (prikazano na slici 1.3).



Slika 1.3: Ilustracija Brownovog gibanja [3]

Preko tog modela je izračunao efektivni radijus molekule šećera, a u kasnijem radu 1908. godine je predložio pokus koji bi potvrdilo njegova predviđanja. Iste te godine Jean Baptiste Perrin je taj pokus napravio i uspješno dokazao atomsku građu tvari, za što mu je 1926. godine dodjeljena Nobelova nagrada za fiziku. S ovim dokazom su i najveći protivnici prihvatili teoriju atoma.

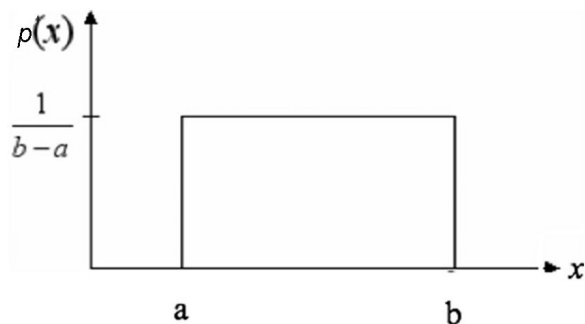
2 Slučajna šetnja

Slučajna šetnja je model šetnje u kojemu šetač pri svakome koraku nasumično mijenja pojedinosti uzastopnih koraka. Pogodna je jer je jednostavna za shvatiti, a može nam dati uvid u ponašanje kompleksnih nasumičnih procesa poput Brownovog gibanja.

Sama šetnja može se realizirati na više načina, ovisno o broju dimenzija, slobodi smjera koraka i samoj duljini koraka. U ovom poglavlju razmatramo kako različite okolnosti kod slučajnih šetnji utječu na njihov razvoj.

2.1 Različite okolnosti nasumične šetnje

Za okolnosti smjera koraka nasumičnog šetača uzimat ćemo nasumično generirane vrijednosti kontinuirane i diskretne uniformne raspodjele preko koje nam služe za usporedbu šetnje u proizvoljnom smjeru i šetnje po rešetci, a za okolnosti duljine koraka nasumičnog šetača uzimat ćemo nasumično generirane vrijednosti eksponencijalne raspodjele i jediničnu duljinu koraka kojima uspoređujemo utjecaj duljine koraka na razvoj šetnje. Navedene raspodjele definiramo na slijedeći način:

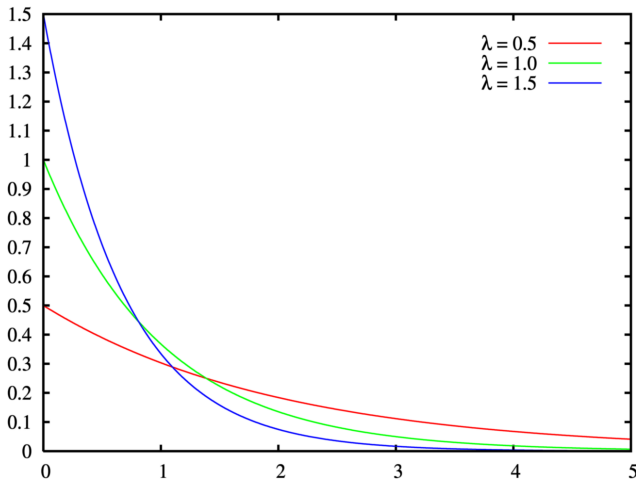


$$p(x) = \begin{cases} \frac{1}{b-a} & \text{za } a \leq x \leq b, \\ 0 & \text{inače.} \end{cases}$$

Slika 2.1: Prikaz kontinuirane uniformne raspodjele [8].

Kontinuirana² uniformna raspodjela (slika 2.1.), također poznata kao pravokutna raspodjela, gdje za kontinuiranu slučajnu varijablu x definiranu na intervalu $x \in [a, b]$, njena funkcija gustoće vjerojatnosti poprima konstantnu vrijednost obrnuto proporcionalnoj širini intervala $\frac{1}{b-a}$, dok je izvan intervala nula. Zgodno je svojstvo ovakve raspodjele da je vrlo jednostavna za uporabu i da svaka vrijednost na zadanom intervalu ima jednaku vjerojatnost, što u slučaju kada ne poznamo pravu "prirodu" raspodjele koju proučavamo razumno polazište (ali rijetko točno).

²Ukoliko se radi o diskretnoj slučajnoj varijabli, odnosno kada varijabla može poprimiti konačan i diskretan broj vrijednosti, dobivamo diskretnu uniformnu raspodjelu čija je konstantna vrijednost funkcije vjerojatnosti obrnuto proporcionalna broju vrijednosti koje raspodjela može poprimiti $\frac{1}{n}$.



Slika 2.2: Prikaz eksponencijalna distribucije [11]

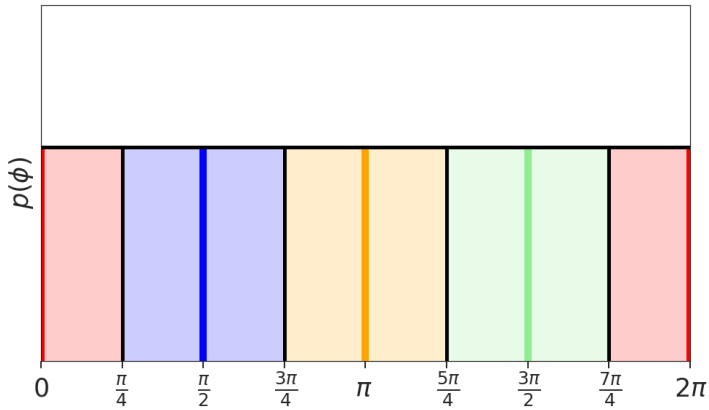
$$p(x; \lambda) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{za } x \geq 0, \\ 0 & \text{inače.} \end{cases}$$

Eksponencijalna raspodjela (prikazano na slici 2.2) je kontinuirana raspodjela koja je nesimetrična i monotono padajuća, gdje se manje vrijednosti češće ostvaruju od većih, opisuje proces čiji se događaji odvijaju nezavisno konstantnom prosječnom učestalošću. Parametar λ predstavlja tu prosječnu mjeru, a kako se eksponencijalna raspodjela često upotrebljava kod razmatranja vremenskih fenomena, često se zove srednje vrijeme raspada, gdje je srednja vrijednost eksponencijalne raspodjele jednaka upravo tom parametru. Jedinstveno svojstvo eksponencijalne raspodjele je svojstvo "zaboravljanja", odnosno da ako promatramo t_1 vremena događaj koji je eksponencijalno raspodjeljen te se on još nije ostvario, vjerojatnost čekanje daljnjeg t_2 vremena je neovisno o prvobitnom vremenu promatranja t_1 . To možemo zapisati:

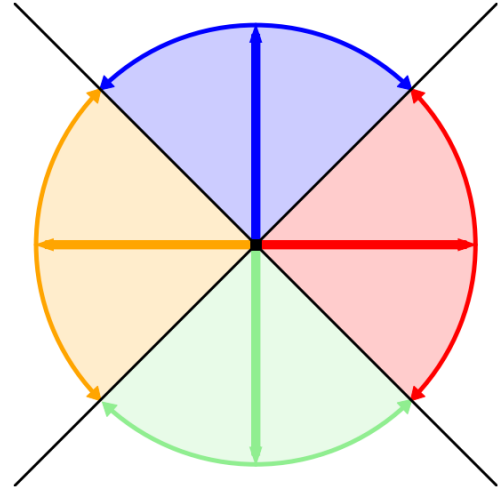
$$p(t > t_1 + t_2 | t_1) = \frac{p(t > t_1 + t_2)}{p(t > t_1)} = \frac{e^{-\frac{1}{\lambda}(t_1+t_2)}}{e^{-\frac{1}{\lambda}t_1}} = e^{-\frac{t_2}{\lambda}} = p(t > t_2). \quad (2.1)$$

2.2 Generiranje koraka nasumične šetnje

Kako bismo uopće mogli usporedno pratiti razvoj šetnji u različitim okolnostima, moramo osigurati da su raspodjele generiranih koraka usporedive, gdje mijenjamo određene okolnosti *osnovne* šetnje i potom pratiti kako izmjena utječe na razvoj šetnje. Za osnovnu šetnju uzimamo šetnju gdje šetač ima korak duljine jedan (dalje u tekstu *jedinični* korak) i slobodan je kretati se u proizvoljnome smjeru oko sebe, gdje svakim korakom odabire kut kontinuirane uniformne raspodjele $\phi \in [0, 2\pi)$ (dalje u tekstu *izotropni* smjer).



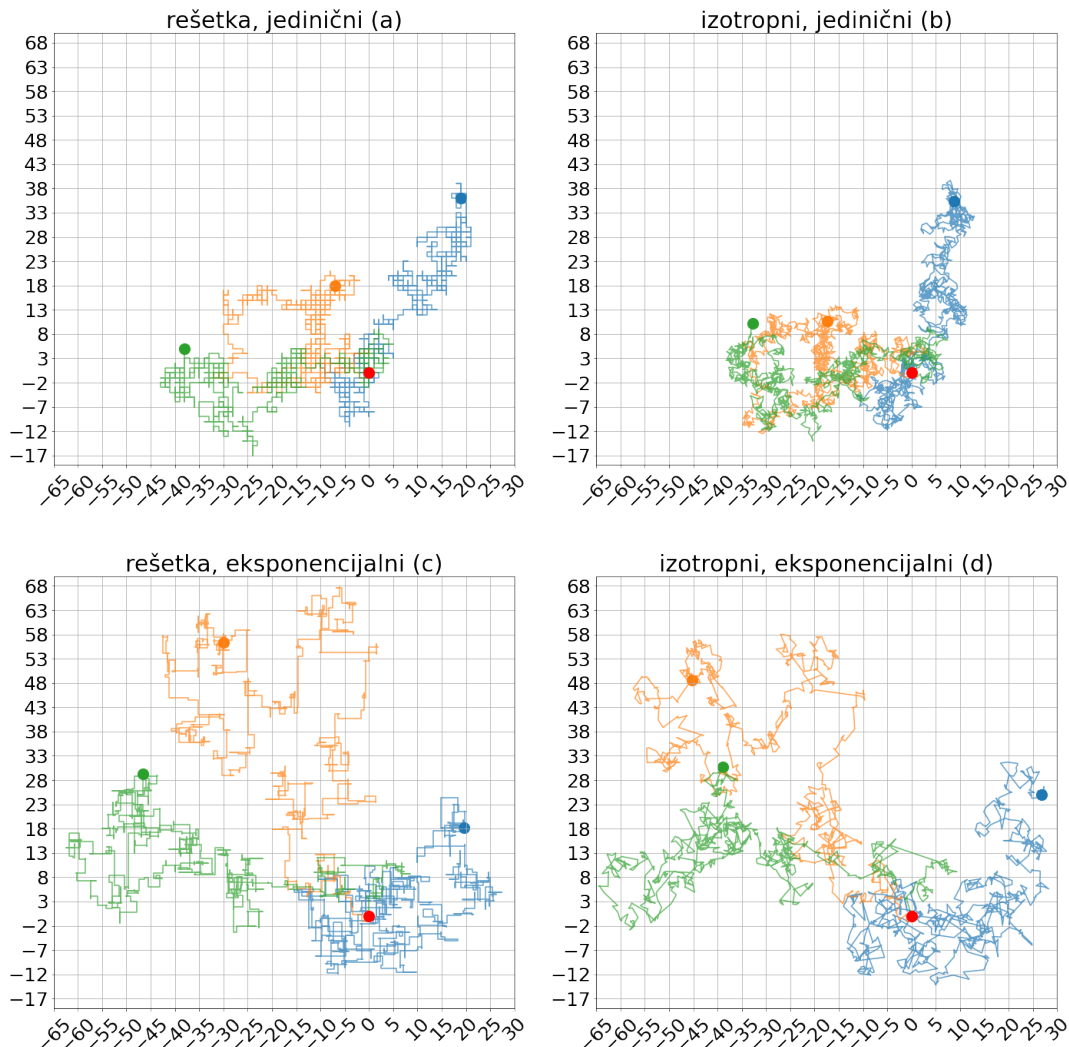
Slika 2.3: Ilustracija kontinuirane i diskretne uniformne raspodjele za odabir smjera koraka.



Slika 2.4: Ilustracija pretvorbe kontinuiranog smjera koraka u diskretni smjer koraka.

Kako bismo od prethodno navedene šetnje postigli šetača na rešetci (dalje u tekstu šetnja po *rešetci*), odnosno šetnju gdje se šetač kreće isključivo u kardinalnim smjerovima, pretvaramo kontinuiranu uniformnu raspodjelu u diskretnu (slika 2.3), gdje se pripadni interval svodi na svoj pripadni kardinalni smjer (slika 2.4). Također promatramo kako sama duljina koraka utječe na razvoj šetnje. U tu svrhu generiramo duljine koraka iz eksponencijalne raspodjele (dalje u tekstu *eksponencijalni* korak) s parametrom srednjeg slobodnog koraka $\lambda = 1$. Preko ovih kombinacija dviju raspodjela za duljinu koraka i dviju raspodjela za kutni otklon koraka možemo konstruirati četiri međusobno usporediva para šetnji.

2.3 Simulacije nasumičnih šetnji u dvije dimenzije

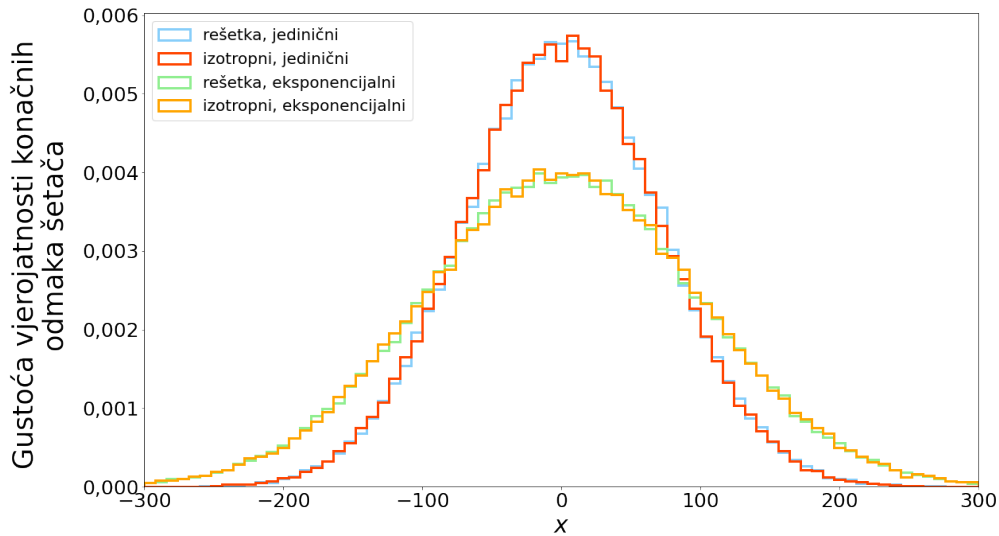


Slika 2.5: Simulacija tri šetača nakon $k = 1000$ koraka za okolnosti:

- (a) šetnja po rešetki s jediničnom duljinom koraka,
- (b) šetnja s izotropnim smjerom i jediničnom duljinom koraka,
- (c) šetnja po rešetki s nasumičnom duljinom koraka generiranom iz eksponencijalne distribucije s $\lambda = 1$,
- (d) šetnja s izotropnim smjerom i nasumičnom duljinom koraka generiranom iz eksponencijalne distribucije s $\lambda = 1$.

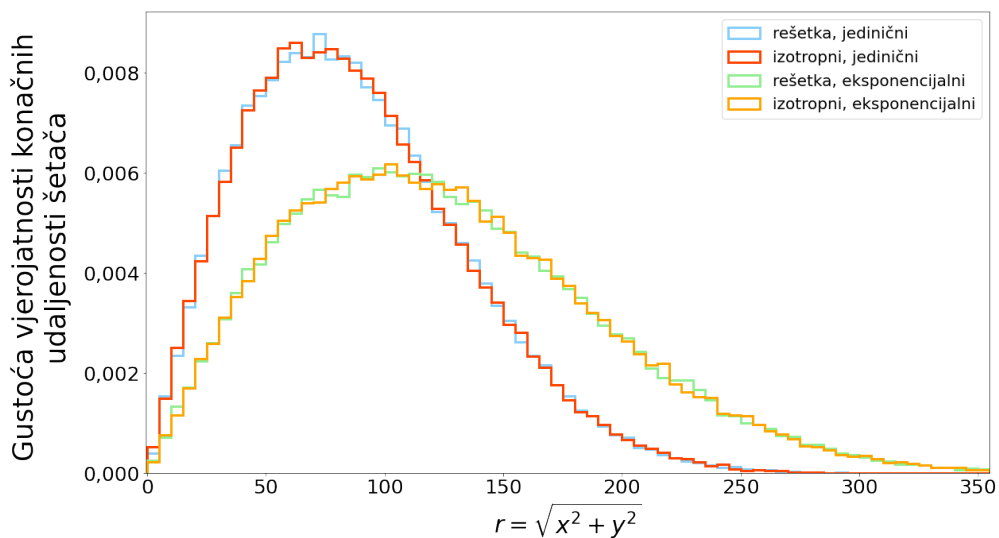
Na slici 2.5 su prikazane simulacije šetnji u prethodno navedenim okolnostima. Crvena točka predstavlja ishodište $(0,0)$ svake šetnje, a kako bi prikaz bio vjerniji osi svih grafova su jednake. Primjećujemo kako šetnje u različitim okolnostima poprimaju sličan oblik, gdje promjena okolnosti smjera naizgled slabo utječe na sami razvoj šetnje, dok promjena okolnosti duljine koraka naizgled dramatično utječu na razvoj šetnje. Kako bismo detaljnije razradili razliku utjecaja tih okolnosti, razmatramo raspodjele konačnih položaja.

U tu svrhu smo simulirali $n = 10^5$ šetača koji rade $k = 10^4$ koraka, za svaku navedenu šetnju.



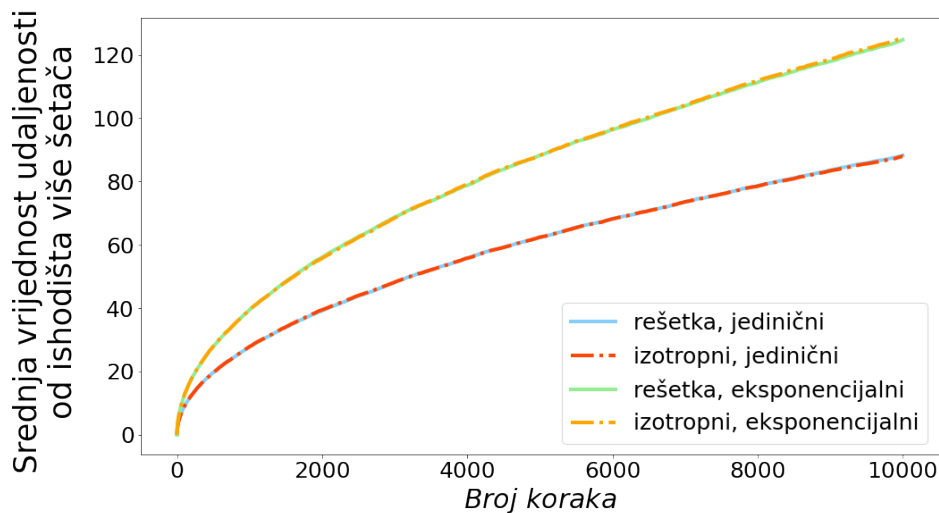
Slika 2.6: Raspodjele konačnih pomaka šetača od ishodišta u x-smjeru za $n = 10^5$ šetača u navedenim okolnostima šetnji nakon $k = 10^4$ koraka.

Na slici 2.6 prikazana je učestalost konačnih x-koordinata za navedene šetnje. Primjećujemo kako raspodjele za sve okolnosti poprimaju oblik normalne raspodjele (rezultat koji ćemo razmatrati u poglavlju 4), ali da nastaje razlika kod promjene okolnosti između jediničnog i eksponencijalnog koraka. Konačne x-koordinate svih šetnji se najviše koncentriraju oko ishodišta, odnosno gustoća vjerojatnosti konačnih x-koordinata šetača se smanjuje udaljavanjem od ishodišta x-osi, gdje su šetači s eksponencijalnim korakom na kraju više raspršeni od šetača s jediničnim korakom.



Slika 2.7: Raspodjele konačnih udaljenosti šetača od ishodišta za $n = 10^5$ šetača u navedenim okolnostima šetnji nakon $k = 10^4$ koraka.

Na slici 2.7 prikazane su raspodjele konačnih udaljenosti od ishodišta dobivene transformacijom varijabli $\sqrt{x^2 + y^2}$ konačnih x i y koordinata navedenih šetnji. Iz priloženog možemo naslutiti kako se se u prosjeku šetači svih vrsta šetnji odmiču, što je naizgled kontradiktorno s rezultatima koje smo dobili za pomak u x smjeru (a i sama raspodjela za pomake u y smjeru poprima isti oblik), gdje šetač "vrluda" oko ishodišta x -osi. Razlog tome je da vrlo rijetko da se šetač istovremeno slabo odmiče od obje osi, odnosno kako bi $\bar{r} \approx 0$, x i y koordinate moraju biti istovremeno jednaki nuli što bi značilo da su smjerovi korelirani, a znamo da su međusobno nezavisni stoga ne mogu biti ni na koji način korelirani. Da to potvrdimo razmatramo kako ovisi udaljenost šetača o broju koraka.



Slika 2.8: Srednje vrijednosti udaljenosti kod svakog koraka šetača iz ishodišta za 10^5 šetača.

Slika 2.8 prikazuje evoluciju aritmetičke sredine konačnih odmakova svih šetnji pripadnih okolnosti. Krivulje poprimaju (barem aproksimativno) oblik funkcije korijena, što nam ukazuje kako uistinu šetač udaljava od ishodišta, te da se šetači s eksponencijalnim korakom "brže" udaljavaju, makar je u prosjeku duljina koraka jednaka jediničnome. No, svejedno smo dobili jednak oblik raspodjela, čime zaključujemo kako makar su se promijenili detalji, statistika (za $k \gg 1$ koraka) je ostala ista.

3 Besselova korekcija

Kao što je navedeno u prošlom poglavlju, slučajna šetnja predstavlja slučajan proces, stoga se može razmatrati uporabom statistike. Jedan od načina statističke analize je preko metode statističkih momenata, gdje nam statistički momenti ukazuju karakteristična svojstva raspodjele promatrane veličine. Među najosnovnijim statističkim momentima su prvi statistički moment, odnosno očekivana vrijednost (μ) koja predstavlja vrijednost populacijske sredine koja se približava pravoj vrijednosti ako broj mjerenja teži prema beskonačnosti. Sama po sebi srednja vrijednost nam ne daje nikakav opis raspodjele, stoga nam je od izuzetne važnosti istovremeno promatrati i drugi statistički moment, odnosno varijancu (σ^2), koja predstavlja mjeru raspršenosti podataka oko srednje vrijednosti raspodjele. Varijanca populacije koja se sastoji od n elemenata, procjenjuje se kvadratom razlike pojedinog elementa uzorka i očekivane vrijednosti, podijeljena s ukupnim brojem elemenata n i možemo je koristiti isključivo kada nam je poznata očekivana vrijednost μ :

$$\hat{V}_0(X_i) = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.$$

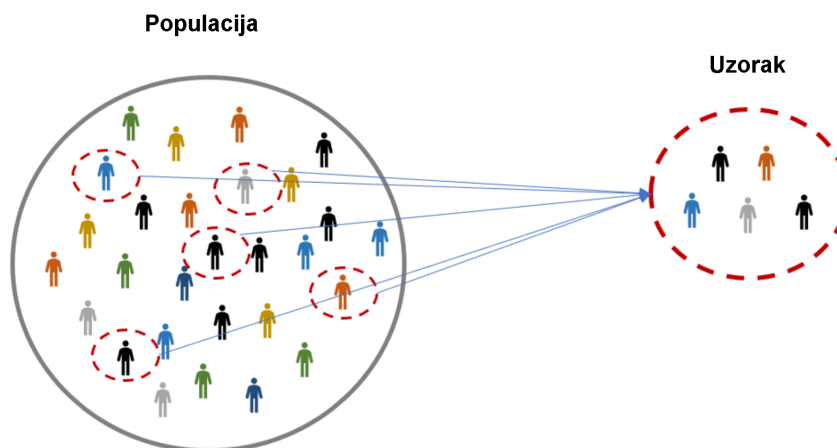
U slučaju kada su nam parametri populacije nepoznati, a sama populacija nepraktična za proučavanje, uzimamo uzorak populacije preko koje procjenjujemo parametre populacije. Varijancu uzorka (s^2) procjenjujemo na sličan način kao i varijancu populacije, gdje zbog nepoznavanja očekivane vrijednosti μ moramo izračunati srednju vrijednost uzorka \bar{X} i dijeliti sa $n - 1$ umjesto s n :

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

Zamjena srednje vrijednosti populacije srednjom vrijednosti uzorka $\mu \rightarrow \bar{X}$ je jasna, dok izmjena faktora $n \rightarrow n - 1$, inače poznata kao Besselova korekcija, dolazi upravo zbog primoranosti korištenja \bar{x} umjesto μ smo uveli pristranost u procjeni varijance uzorka. U ovom poglavlju prikazujemo izvod Besselove korekcije i ilustriramo je računalnim simulacijama, a zbog jasnoće se navode i definiraju potrebni statistički pojmovi.

3.1 Populacije i uzorci

Populacija je skup koji uključuje sve elemente s određenim zajedničkim svojstvima koje proučavamo, a sama populacija može biti konačna ili beskonačna.



Slika 3.1: Ilustracija pravilnog uzimanja uzorka populacije [4]

Proučavanje čitave populacije često je nepraktično ako ne i nemoguće, stoga se najčešće uzima samo uzorak, odnosno podskup populacije. Rezultate dobivene analizom uzorka generaliziramo na čitavu populaciju, odnosno od karakterističnih svojstava uzorka procjenjujemo svojstva populacije. Kako bi rezultati dobiveni uzorkovanjem bili reprezentativni za čitavu populaciju, od izuzetne je važnosti nepristrano i nasumično uzimanje samog uzorka kao što je prikazano na slici 3.1.

3.2 Statistički procjenitelj i njegova svojstva

Procjenitelj je statistika³ čije se vrijednosti interpretiraju kao procjene danoga populacijskog parametra P . Greška procjenitelja $\hat{P}(x)$ slučajne varijable x je:

$$g(x) = \hat{P}(x) - P. \quad (3.1)$$

³Statistika (jednina) ili statistika uzorka je svaka veličina izračunata iz vrijednosti uzorka. Statističke svrhe uključuju procjenu parametra populacije, opis uzorka ili procjenu hipoteze. Prosjek (poznat i kao srednja vrijednost) uzorka je primjer statistike. Pojam statistika koristi se i za funkciju i za vrijednost funkcije koja nam na danom uzorku govori kako odrediti procjenu na temelju parametara opažanja nekog uzorka. Kada se statistika koristi za određenu svrhu, na nju se može ukazati imenom koje označava njezinu svrhu. Postoje statistički kriteriji (nepristranost, dosljednost, efikasnost itd.) koji za danu veličinu mogu biti nekompatibilni, no za varijancu su svi zadovoljeni.

3.2.1 Pristranost

Pristranost procjenitelja je očekivana vrijednost odstupanja procijenjene vrijednosti od prave vrijednosti:

$$\mathbb{E}[g(x)] = \mathbb{E}[\hat{P}(x) - P] = \mathbb{E}[\hat{P}(x)] - P. \quad (3.2)$$

Za procjenitelj kažemo da je nepristran ukoliko mu je očekivano odstupanje od prave vrijednosti jednako nuli:

$$\mathbb{E}[\hat{P}(x)] - P = 0. \quad (3.3)$$

3.3 Besselova korekcija - Izvod

Varijanca, koja predstavlja mjeru odstupanja elemenata skupa Y od njezine srednje vrijednosti, općenito se može zapisati kao očekivanje kvadrata razlike tog skupa i njene srednje vrijednosti:

$$\mathbb{V}(Y) = \mathbb{E}[(Y - \mathbb{E}(Y))^2], \quad (3.4)$$

koje se može proširiti u razliku očekivanja kvadrata i kvadrata očekivanja:

$$\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2, \quad (3.5)$$

a samo očekivanje kvadrata možemo zapisati:

$$\mathbb{E}(Y^2) = \mathbb{V}(Y) + \mathbb{E}(Y)^2. \quad (3.6)$$

Kod sume nekoleriranih⁴ varijabli Y_i varijanca glasi:

$$\mathbb{V}\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n \mathbb{V}(Y_i). \quad (3.7)$$

Također, ukoliko su sve vrijednosti skalirane za konstantu a , varijanca je skalirana za kvadrat te konstante:

$$\mathbb{V}(aY) = a^2\mathbb{V}(Y). \quad (3.8)$$

Neka je skup x uzorak od n elemenata nasumičnih i nezavisnih varijabli, gdje su očekivane

⁴Općeniti izraz glasi $\mathbb{V}(\sum_{i=1}^n Y_i) \stackrel{\text{def}}{=} \sum_{i,j=1}^n \text{Cov}(Y_i, Y_j) = \sum_{i=1}^n \mathbb{V}(Y_i) + \sum_{i \neq j} \text{Cov}(Y_i, Y_j)$.

vrijednosti srednje vrijednosti i varijance *populacije*:

$$\mathbb{E}(x_i) = \mu, \quad (3.9)$$

$$\mathbb{V}(x_i) = \sigma^2, \quad (3.10)$$

a očekivanje kvadrata:

$$\mathbb{E}(x_i^2) = \mathbb{V}(x_i) + \mathbb{E}(x_i)^2 = \sigma^2 + \mu^2. \quad (3.11)$$

Očekivana vrijednost i varijanca srednje vrijednosti *uzorka*:

$$\mathbb{E}(\bar{x}) = \mathbb{E}\left(\frac{\sum_{i=1}^n x_i}{n}\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n x_i\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(x_i) = \mu, \quad (3.12)$$

$$\mathbb{V}(\bar{x}) = \mathbb{V}\left(\frac{\sum_{i=1}^n x_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}(x_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}, \quad (3.13)$$

i pripadno očekivanje kvadrata:

$$\mathbb{E}(\bar{x}^2) = \mathbb{V}(\bar{x}) + \mathbb{E}(\bar{x})^2 = \frac{\sigma^2}{n} + \mu^2. \quad (3.14)$$

Neka je probni procjenitelj varijance uzorka $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}, \quad (3.15)$$

to možemo raspisati na sljedeći način:

$$\begin{aligned} \mathbb{E}\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}\right] &= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right] \\ &= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n x_i^2 - \sum_{i=1}^n 2x_i\bar{x} + \sum_{i=1}^n \bar{x}^2\right] = \\ &= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2\right] = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n x_i^2 - n\bar{x}^2\right] = \\ &= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n x_i^2\right) - \frac{1}{n}\mathbb{E}(n\bar{x}^2) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(x_i^2) - \mathbb{E}(\bar{x}^2). \end{aligned} \quad (3.16)$$

Prepoznavamo očekivanje kvadrata populacije (3.11 i uzorka 3.14) koje uvrštavamo ⁵:

$$\begin{aligned} & \frac{1}{n} \sum_i^n [\mathbb{V}(x_i) + \mathbb{E}(x_i)] - [\mathbb{V}(\bar{x}) + \mathbb{E}(\bar{x})] = \\ & = \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} + \mu^2 \right) = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2. \end{aligned} \quad (3.17)$$

Primjećujemo kako probni procjenitelj ne procjenjuje σ^2 , već ima faktor:

$$\frac{n-1}{n} < 1, \quad (3.18)$$

čime zaključujemo da je pristran (podcjenjuje σ^2) te kako bi bio nepristran moramo ga pomnožiti recipročnom vrijednosti od jednadžbe 3.18:

$$\frac{n}{n-1} \mathbb{E}[\hat{\sigma}^2] = \frac{n}{n-1} \mathbb{E} \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \right] = \mathbb{E} \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right] = \sigma^2. \quad (3.19)$$

Vidimo kako je izvod dugačak i netransparentan, te nam ne daje uvid u temeljni razlog potrebe Besselove korekcije, koja se može vrlo lako razumjeti. Stoga koristimo računalne simulacije kao laboratorij, gdje imamo izuzetnu prigodu poznavanja karakterističnih parametara raspodjele kao što su srednja vrijednost populacije μ i varijanca populacije σ^2 koje zadajemo računalu. Time smo u mogućnosti rekonstruirati osnovne ideje njihovom realizacijom i predočiti na intuitivan način nužnost Besselove korekcije.

3.4 Besselova korekcija - računalne simulacije

U ovom potpoglavlju konstruiramo "probne" procjenitelje. Razmatramo kako izmjena pojedinih parametara naših "probni" procjenitelja utječe na njihova svojstva, prvenstveno na njihovu pristranost kod procjene.

Koristimo procjenitelj varijance populacije:

$$\hat{\sigma}_{\mu,n}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2, \quad (3.20)$$

kao bazu kojoj mijenjamo parametar srednje vrijednosti populacije u srednju vrijednost uzorka

⁵Kod ovog člana se vidi odakle dolazi Besselova korekcija, a taj član proizlazi iz $\mathbb{V}(\bar{x})$ što je posljedica upotrebe \bar{x} umjesto μ .

($\mu \rightarrow \bar{x}$) da konstruiramo procjenitelja varijance:

$$\hat{\sigma}_{\bar{x},n}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (3.21)$$

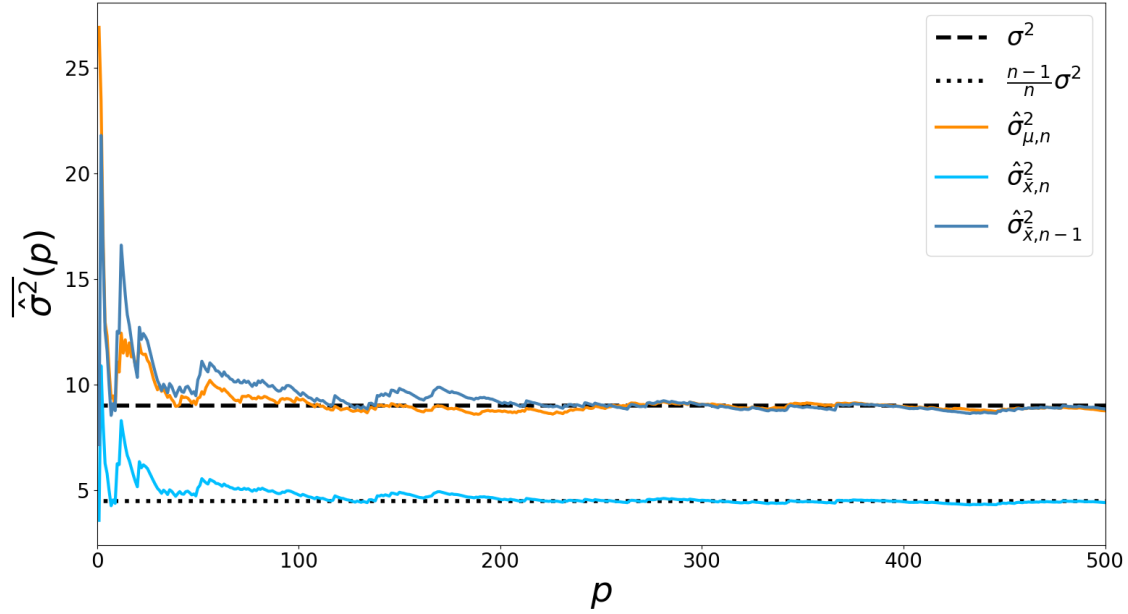
i procjenitelj kojem ujedno dodajemo i Besselovu korekciju ($\mu \rightarrow \bar{x}, n \rightarrow n - 1$):

$$\hat{\sigma}_{\bar{x},n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (3.22)$$

Generirali smo par⁶ nasumičnih vrijednosti (x_1, x_2) iz normalne distribucije $\mathcal{N}(10, 3)$. Preko kojih računamo srednju vrijednost uzorka \bar{x} i tri navedena procjenitelja. Postupak smo ponovili više puta, svaki put generirajući novi par vrijednosti, gdje smo za svaki p -ti generirani par, odnosno "ponovljeni" pokus, računali srednju vrijednost pojedinog procjenitelja $\overline{\hat{\sigma}^2}$ do p -tog para. *Ukoliko se radi o nepristranom procjenitelju*, očekivana vrijednost svakog i -tog izračunatog procjenitelja varijance $\hat{\sigma}_i^2$ je upravo zadana varijancom σ^2 tj. $\mathbb{E}(\hat{\sigma}_i^2) = \sigma^2$, a i također posljedično srednja vrijednost p parova procjenitelja varijance:

$$\mathbb{E} \left(\frac{1}{p} \sum_i^p \hat{\sigma}_i^2 \right) = \frac{1}{p} \sum_i^p \mathbb{E}(\hat{\sigma}_i^2) = \frac{1}{p} (\mathbb{E}(\hat{\sigma}_1^2) + \mathbb{E}(\hat{\sigma}_2^2) \dots + \mathbb{E}(\hat{\sigma}_p^2)) = \sigma^2. \quad (3.23)$$

⁶Odabir dvaju elemenata je iz razloga da je to najmanji mogući broj članova (jedan događaj nije dovoljan za statistiku) za račun varijance i tim odabirom se postiže najizraženiji prikaz pristranosti procjenitelja tj. korekcijski faktor $\frac{n}{n-1}$ najveći za $n = 2$, makar oni u asimptoti ($n \rightarrow \infty$) postizu isti rezultat.



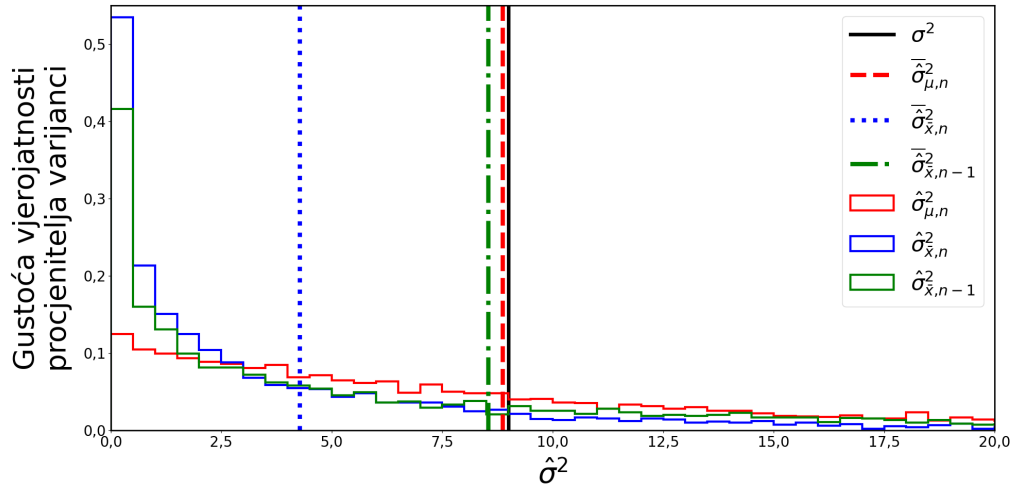
Slika 3.2: Ovisnost srednje vrijednosti procjenitelja varijance o broju generiranih parova iz normalne raspodjele s zadanim parametrima $\mu = 10$, $\sigma = 3$.

Rezultati na slici 3.2 prikazuju kako procjenitelji pri malom broju generiranih parova fluktuiraju, no kako se povećanjem ponavljanja stabiliziraju oko određene vrijednosti. Procjenitelji $\hat{\sigma}_{\mu,n}^2$ i $\hat{\sigma}_{\bar{x},n-1}^2$ stabiliziraju se upravo oko zadane vrijednosti "prave" varijance raspodjele σ^2 što ukazuje na njihovu nepristranost⁷, za razliku od procjenitelja $\hat{\sigma}_{\bar{x},n}^2$ kod kojeg se očituje pristranost. Možemo primijetiti kako krivulje procjenitelja $\hat{\sigma}_{\bar{x},n-1}^2$ i $\hat{\sigma}_{\bar{x},n}^2$ poprimaju isti oblik, no procjenitelj $\hat{\sigma}_{\bar{x},n}^2$ koji je skaliran za faktor $\frac{1}{2}$ što je upravo iznos predivđen preko Besselove korekcije.

⁷Makar je procjenitelj varijance uzorka zbog Besselove korekcija nepristran, uzimanje korijena nepristranog procjenitelja varijance čini procjenitelj standardne devijacije pristranim. Razlog tome je da funkcija korijena nije linearna i zbog toga ne komutira sa očekivanjem. Također funkcija korijena je konkavna, pa po Jensenovoj nejednakosti podcjenjuje očekivanu standardnu devijaciju uzorka:

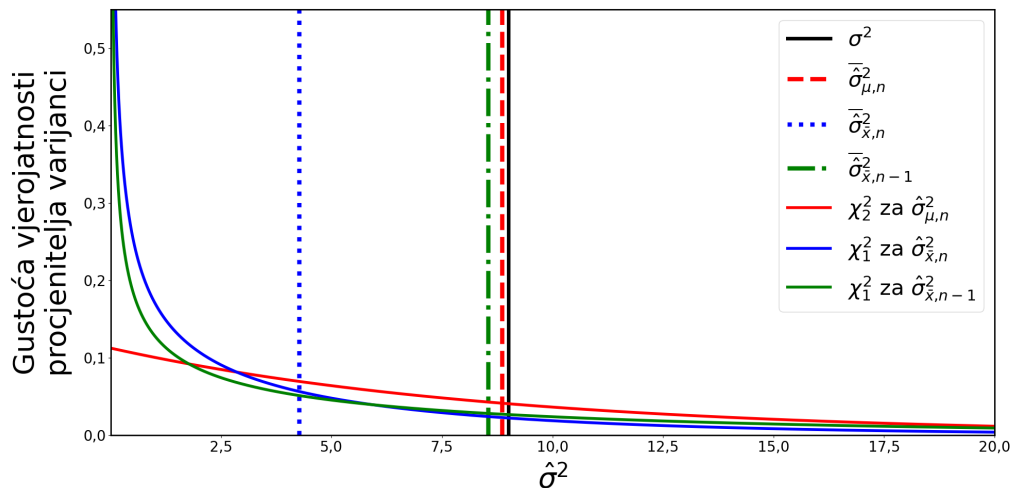
$$\mathbb{E}[\sqrt{\mathbb{V}(x)}] \leq \sqrt{\mathbb{E}[\mathbb{V}(x)]} = \sigma.$$

Za razliku od procjenitelja srednje vrijednosti i varijance, ne postoji nepristran procjenitelj standardne devijacije koji je neovisan o raspodjeli. Iz tog razloga se standardna devijacija definira kao korijen varijance.



Slika 3.3: Gustoća vjerojatnosti triju procjenitelja varijance dobivene iz $p = 5000$ generiranih parova iz normalne raspodjele s zadanim parametrima $\mu = 10, \sigma = 3$.

Također smo nanovo generirali novi skup parova, ovaj put računajući kvadrate navedenih procjenitelja (odnosno "probni" varijanci). Učestalost izračunatih vrijednosti svakog procjenitelja prikazujemo histogramom na slici 3.3. Aritmetička sredina raspodjele svakog procjenitelja prikazana je vertikalnim linijama, gdje vidimo veliko odstupanje od zadane vrijednosti σ^2 kod procjenitelja $\hat{\sigma}_{\bar{x},n}^2$ i slično kao u prošleme primjeru, umanjeno za faktor 2, kao i sličnost oblika raspodjela procjenitelja $\hat{\sigma}_{\bar{x},n}^2$ i $\hat{\sigma}_{\bar{x},n-1}^2$, gdje procjenitelj $\hat{\sigma}_{\mu,n}^2$ poprima drugačiji oblik.



Slika 3.4: χ^2 raspodjele triju procjenitelja varijance dobivene iz $p = 5000$ generiranih parova iz normalne raspodjele s zadanim parametrima $\mu = 10, \sigma = 3$, skalirana statističkim momentima generiranih raspodjela.

Te χ^2 raspodjele skalirane pripadnim momentima procjenitelja radi preglednosti prikazujemo na slici 3.4. Kako vidimo, raspodjela za procjenitelj $\hat{\sigma}_{\mu,n}^2$ poprima oblik χ^2 raspodjele sa parametrom stupnja slobode $k = 2$, odnosno da se radi o zbroju dviju nezavisnih varijabli normalne

raspodjele, dok je za raspodjele za procjenitelje $\sigma_{\bar{x},n}^2$ i $\sigma_{\bar{x},n-1}^2$ broj stupnjeva slobode $k = 1$. Kako su sama srednja vrijednost populacije \bar{x} usko vezana uz članove uzorka x_i , a srednja vrijednost populacije μ je potpuno neovisna veličina o raspodjeli uzorkovanja, slijedi da je za očekivati kako su članovi uzorka x_i "bliži" \bar{x} nego μ , da će za pojedini član vrijediti:

$$\mathbb{E}[(x_i - \bar{x})^2] \leq \mathbb{E}[(x_i - \mu)^2], \quad (3.24)$$

te kako bi poništili ovo podcjenjivanje, povećavamo izraz dijeleći ga manjim brojem.

3.5 Podcjenjivanje pristranog procjenitelja

Procjenitelj varijance koji koristi srednju vrijednost uzorka ima svojstvo da osigurava najmanju procijenjenu varijancu. Da bismo to pokazali konstruiramo procjenitelj varijance gdje promatramo njegovu ovisnost o probnom procjenitelju srednje vrijednosti μ^* :

$$\hat{\sigma}(\mu^*)^2 = \frac{\sum_{i=1}^n (x_i - \mu^*)^2}{n}. \quad (3.25)$$

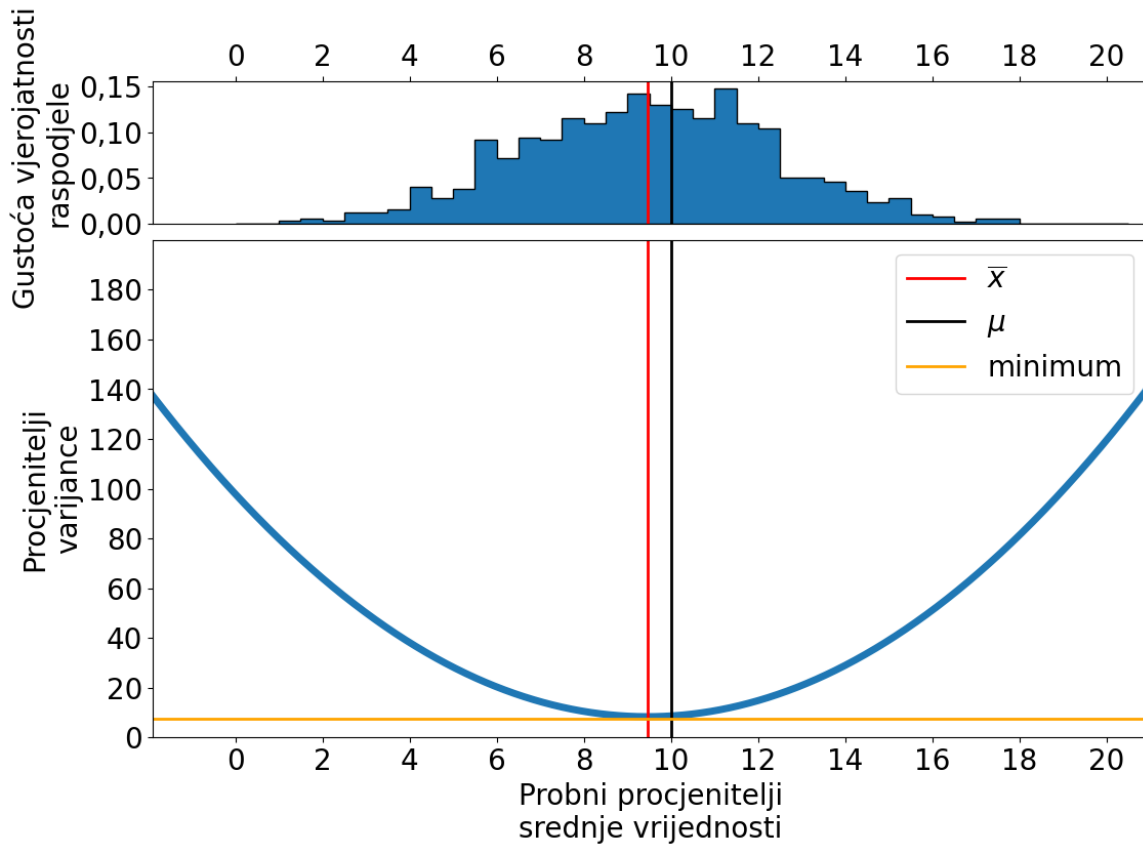
Da odredimo za koju vrijednost će izraz biti minimalan, potreban nam je lokalni ekstrem koji možemo dobiti određivanjem nultočke prve derivacije:

$$\begin{aligned} \frac{d}{d\mu^*} \left(\frac{\sum_{i=1}^n (x_i - \mu^*)^2}{n} \right) &= 0, \\ \frac{-2}{n} \sum_{i=1}^n (x_i - \mu^*) &= 0, \\ \sum_{i=1}^n x_i - \sum_{i=1}^n \mu^* &= 0. \end{aligned} \quad (3.26)$$

Dobijemo da je lokalni ekstrem $\mu^* = \bar{x}$. Možemo zaključiti kako se radi o lokalnom minimumu, jer u suprotnom bi značilo da bi uvrštavanjem srednje vrijednosti raspodjele dobili najveću moguću varijancu, tj. svi drugi procjenitelj bi davali manju. No svejedno vrijedi provjeriti analitički:

$$\frac{d^2}{d\mu^{*2}} \left(\frac{\sum_{i=1}^n (x_i - \mu^*)^2}{n} \right) \Bigg|_{\mu^* = \bar{x}} = 2 > 0. \quad (3.27)$$

Time smo dokazali da je za dani skup x , najmanja moguća varijanca upravo kada koristimo srednju vrijednost uzorka \bar{x} , odnosno u izuzetnom slučaju kada se iznosom pogodimo točno $\bar{x} = \mu$. To još dodatno ilustriramo računalnom simulacijom.



Slika 3.5: Ovisnost procjenitelja varijance o probnom procjenitelju srednje vrijednosti za $n = 1000$ generiranih članova iz normalne distribucije sa zadanim parametrima populacije $\mu = 10$, $\sigma = 3$.

Na slici 3.5 histogramom je prikazana generirana raspodjela, gdje vidimo kako srednje vrijednosti uzorka odstupaju od srednje vrijednosti populacije. Plava krivulja oblika parabole predstavlja funkcionalnu ovisnost varijance o probnom procjenitelju srednje vrijednosti, te vidimo kako uistinu varijanca postiže minimalan iznos upravo za $\mu^* = \bar{x}$.

4 Centralni granični teorem

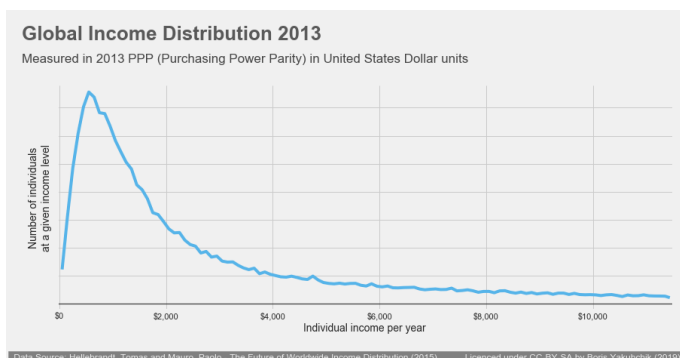
4.1 Zabune oko centralnog graničnog teorema

Centralni granični teorem (dalje u tekstu CGT⁸), jedan je od najvažnijih i najpoznatijih statističkih alata, ali često se neispravno spominje i koristi, te stvara krive predodžbe koje razmatramo u ovom poglavlju, kao i intuitivan primjer CGT-a i njegove primjene na nasumičnoj šetnji. Kako postoji više teorema koji su nazvani CGT, iskazat ćemo ga na jednostavan način: *Ukoliko uzimamo uzorke iz iste raspodjele, povećanjem veličine uzorka raspodjela srednje vrijednosti uzorka teži u normalnu (gausijansku) raspodjelu.*

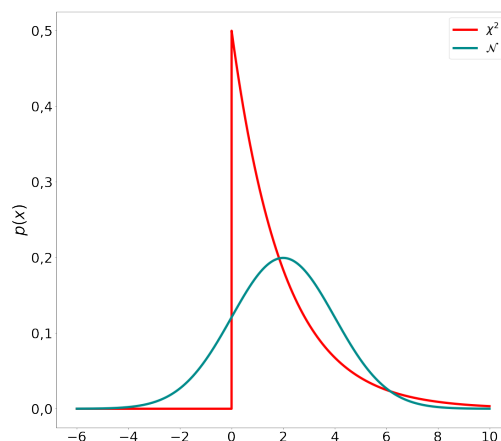
To nas dovodi do prve zabune.

1. Miješanje zakona velikih brojeva (dalje u tekstu ZVB) s CGT.

Poistovjećivanje CGT i ZVB je česta zabuna gdje se uzima da su sve raspodjele normalne, odnosno ukoliko je uzorak dovoljno velik, on mora biti normalno raspodjeljen.



Slika 4.1: Raspodjela globalnih godišnjih prihoda 2013 [6].



Slika 4.2: Usporedba normalne i χ^2 distribucije s jednakim srednjim vrijednostima i varijancama.

Kako možemo vidjeti na slici 4.1 (raspodjela koja je rađena na velikom uzorku), kao i na slici 4.2, gdje raspodjele imaju jednake parametre μ i σ , to jednostavno nije točno. Kako bismo istaknuli gdje nastaje zabuna, iskazujemo ZVB na jednostavan način:

Ukoliko uzimamo uzorke iz iste raspodjele, povećanjem veličine uzorka srednja vrijednost uzorka teži u očekivanu vrijednost populacije.

Kako vidimo, iskazi su vrlo slični, stoga je zabuna razumljiva. No, temeljna razlika ZVB i CGT

⁸Prvi spomen Centralnog graničnog teorema tim nazivom bio je 1920. od G. Pólya, da naglasi "centralnu" važnost tog teorema za teoriju vjerojatnosti.

je kako ZVB govori o samom **očekivanju** srednje vrijednosti \bar{x} , dok CGT govori o **raspodjeli** srednjih vrijednosti \bar{x} .

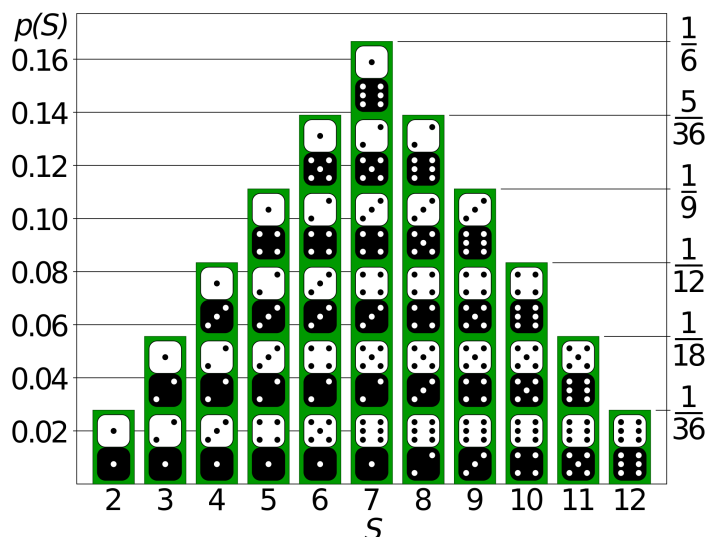
2. Miješanje veličine uzoraka n i broja uzorka z .

Vrlo jednostavna zabuna, no zamjena tih veličina $n \Leftrightarrow z$ nikako ne rezultira istom raspodjelom.

3. CGT vrijedi samo kod velikih uzorka, i to tek nakon $z \geq 30$ uzoraka.

Općenito, kako je CGT opisuje asimptotsko ponašanje raspodjele srednjih vrijednosti x . Stoga tvrdnje iz literature poput tih da CGT počinje vrijediti nakon $z = 30$ uzoraka, naravno, vrijede samo u aproksimativnom smislu. Specifična vrijednost z veličine uzorka nakon koje možemo smatrati da smo u režimu CGT-a ovisi o samoj raspodjeli koju promatramo.

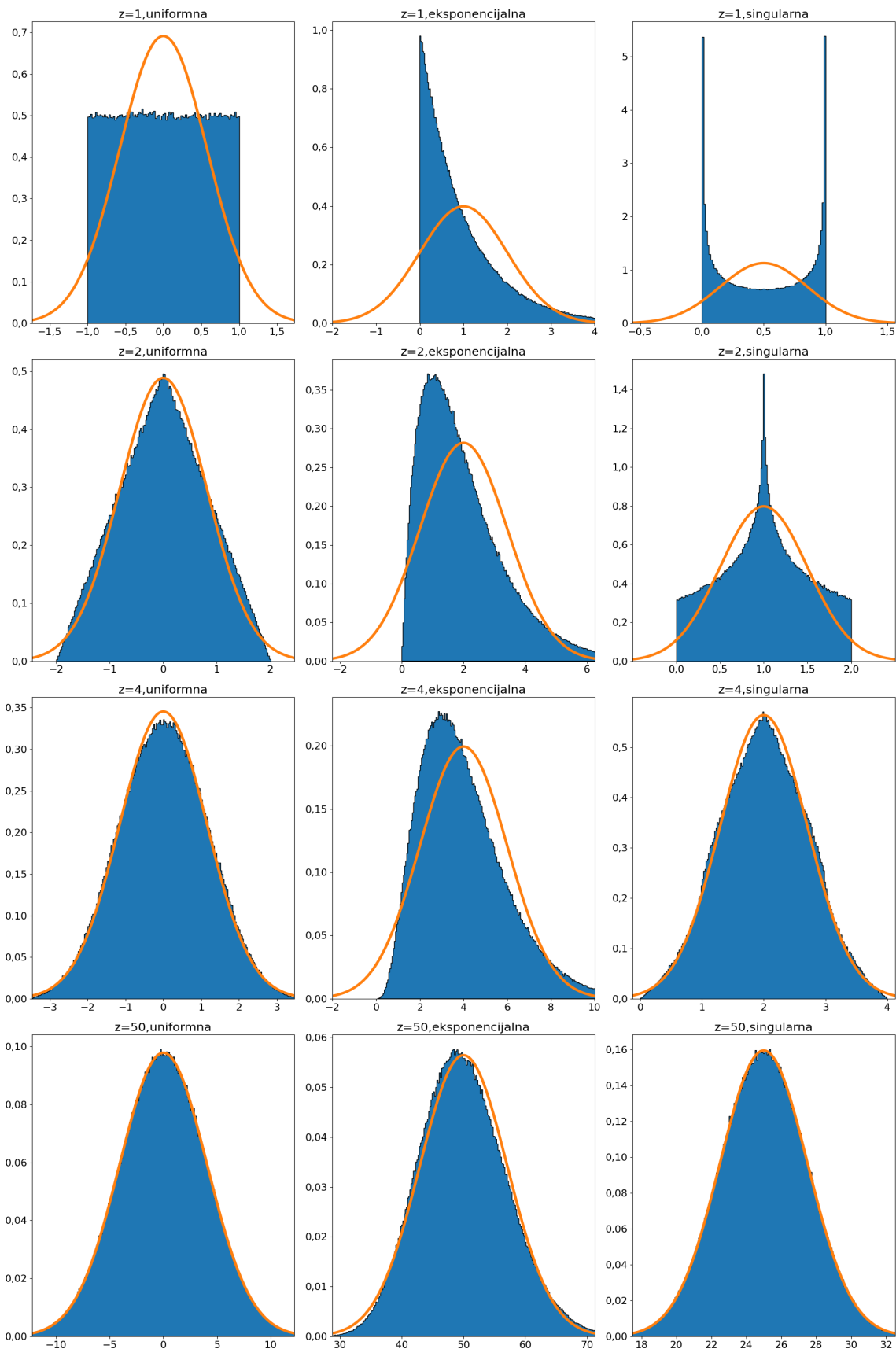
4.2 Kocke



Slika 4.3: Ilustracija svih mogućih kombinacija zbroja bacanja dviju kocki, preuzeto sa [5]

Vidimo kako kod bacanja dviju kocki, koja svaka zasebno ima diskretnu uniformnu distribuciju, zbroj njihovih bacanja daje distribuciju koja više nije uniformna. Neke konfiguracije mogu se realizirati na više načina, a stanja koja se mogu realizirati na više načina su stoga vjerojatnija. Ovim pristupom, koristeći računalnu simulaciju ilustrirat ćemo CGT za tri raspodjele i pratiti njihov razvoj pri dodavanju uzoraka istih raspodjela. Promatrat ćemo simetričnu raspodjelu (uniformnu raspodjelu), nesimetričnu (eksponencijalnu raspodjelu) i raspodjelu koja ima integrabilne singularitete. Rezultati su prikazani na slici 4.4. Preko vrijednosti na osima možemo vidjeti kako se sve raspodjele spuštaju i šire (zbog povećanja standardne devijacije ukupne raspodjele), i također da se svakim korakom odmiču za \bar{x} , osim uniformne raspodjele čiji je $\mathbb{E}(\bar{x} = 0)$. Također se vidi kako nesimetrična raspodjela čak i nakon pedeset koraka nije sasvim ušla u anvelopu normalne raspodjele, tj. kako $z \geq 30$ nije bio sasvim dovoljan.

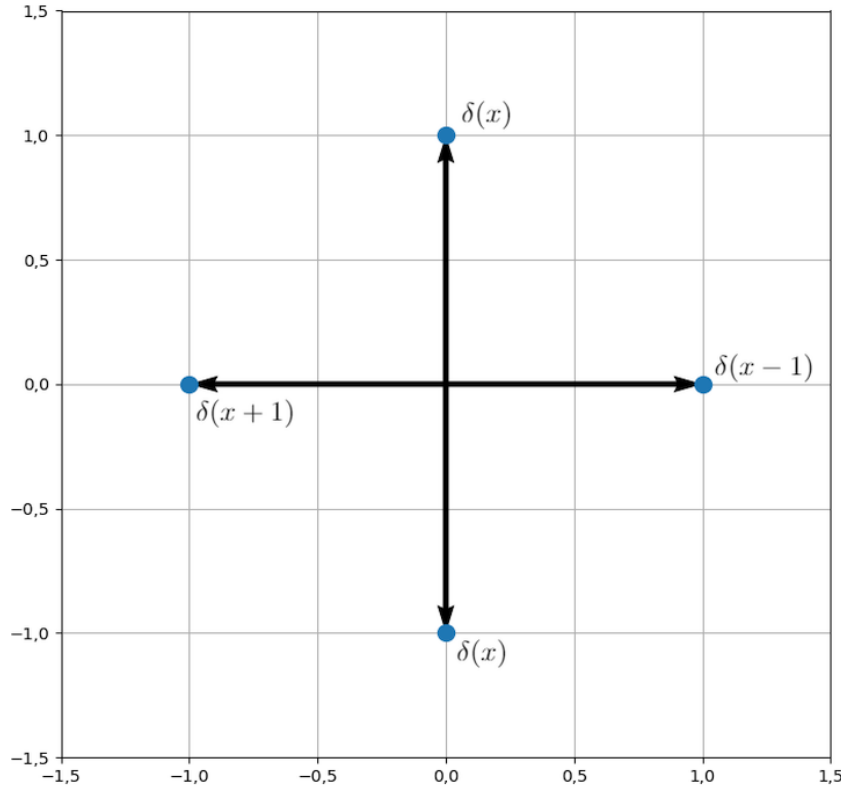
Gustoća vjerojatnosti raspodjela



Slika 4.4: Raspodjele dobivene zbrajanjem z skupova veličine $n = 10^6$ za uniformnu, eksponencijalnu i raspodjelu arkus sinusa, gdje narančasta krivulja predstavlja normalnu raspodjelu s pripadnim parametrima μ i σ za danu konvoluciju raspodjela. Napomena čitatelju kako su vrijednosti na osima različite kod svakog grafa.

4.3 Primjer CGT na nasumičnoj šetnji

Razmatramo slučaj nasumične šetnje po rešetki s jediničnim korakom. Nakon prvog koraka šetač se može nalaziti u jednom od kardinalnih smjerova kako je prikazano na slici 4.5.



Slika 4.5: Ilustracija mogućih položaja 2D šetača nakon prvog koraka iz ishodišta.

Ukoliko želimo koristiti kontinuirane alate s diskretnim varijablama, raspodjele iskazujemo preko Diracove delta funkcije. Kada na taj način iskazujemo raspodjelu diskretnih položaja u x-smjeru, dobivamo gustoću vjerojatnosti za položaje šetača:

$$g(x) = \frac{1}{4} [\delta(x-1) + \delta(x+1) + 2\delta(x)], \quad (4.1)$$

stoga su očekivani položaj i varijanca jednog koraka:

$$\mathbb{E}(x_1) = \bar{x}_1 = \int_{-\infty}^{\infty} xg(x)dx = 0, \quad (4.2)$$

$$\mathbb{V}(x_1) = \sigma_1^2 = \int_{-\infty}^{\infty} (x - \bar{x}_1)^2 g(x)dx = \frac{1}{2}. \quad (4.3)$$

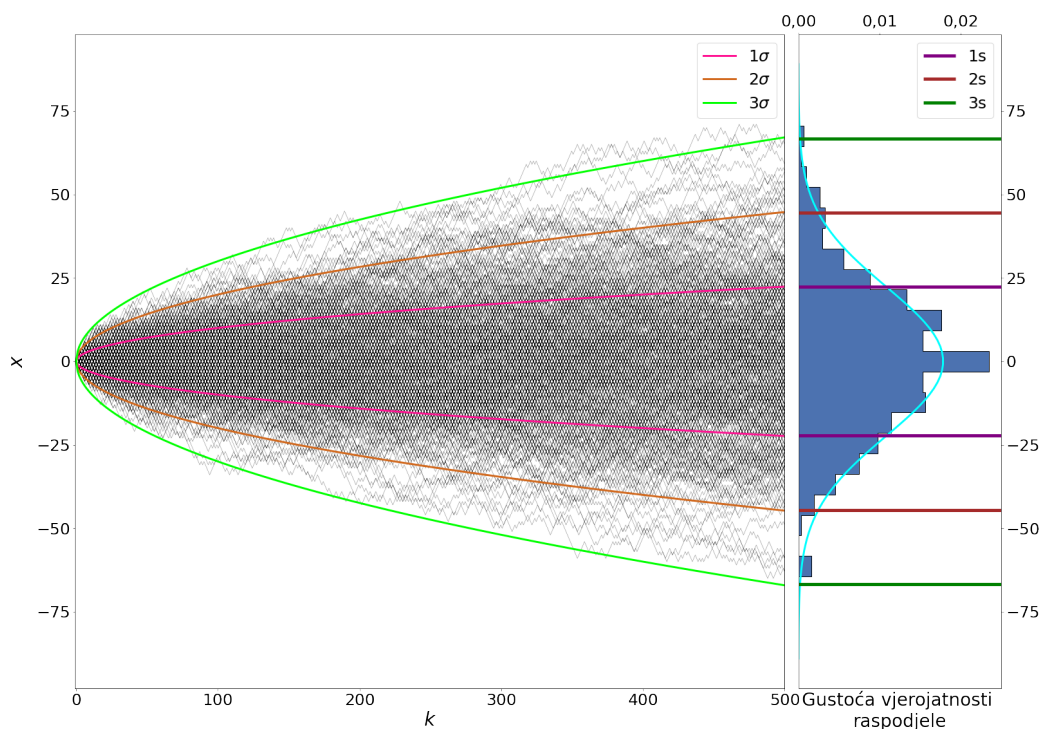
Kako smo već ranije ustvrdili da su koraci nasumične šetnje nezavisne i identične raspodijeljene

slučajne varijable, za zbroj očekivanje srednje vrijednosti varijance nakon n koraka vrijedi:

$$\mathbb{E}(x_{uk}) = \bar{x}_{uk} = \mathbb{E}\left(\sum_i^n x_i\right) = \sum_i^n \mathbb{E}(x_i) = \sum_i^n 0 = 0, \quad (4.4)$$

$$\mathbb{V}(x_{uk}) = \sigma_{uk}^2 = \mathbb{V}\left(\sum_i^n x_i\right) = \sum_i^n \mathbb{V}(x_i) = \sum_i^n \frac{1}{2} = \frac{n}{2}. \quad (4.5)$$

Zanimljiva posljedica 4.5 je da je očekivana standardna devijacija nasumične šetnje proporcionalna korijenu koraka, tj. rasipanje pomaka nasumičnog šetača u pojedinom koordinatom smjeru je proporcionalno korijenu za razliku od usmjerenog gibanja koje je linearno.



Slika 4.6: Položaji $n = 1000$ 1D šetača tokom izvođenja $k = 500$ koraka.

Na slici 4.6 je prikazan razvoj šetnji s označenim standardnom devijacijom koja je $\sigma = \sqrt{k}$ i raspodjela konačnih položaja s označenom standardnom populacijom uzorka s .

5 Pólyjino brojanje

5.1 Pólyina šetnja

Pijanac će naći put do kuće, ali pijana ptica može se zauvijek izgubiti.

Shizuo Kakutani

George Pólya (1887.-1985.) bio je mađarski matematičar, koji je doprinjeo u mnogo različitih matematičkih područja, kao što su redovi, teorija brojeva, kombinatorika, i vjerojatnost. Pólya je volio šetnje, toliko da nije položio vozački ispit. Dok je šetao parkom ili šumom, znao je više puta susresti jedan te isti zaručeni par. To se počelo događati toliko često da se Pólya zabrinuo kako bi par mogao zaključiti da ih on namjerno prati. Kako bi se opravdao da je to puka slučajnost, zapitao se koliko je vjerojatno da će nasumično hodajući stazama naići na druge šetače. Može se pokazati kako je razlika između njihovih šetnji također primjer jednostavne slučajne šetnje, pa se pitanje o susretu može preformulirati u; koja je vjerojatnost da nasumični šetač posjeti neku proizvoljnu točku (u proizvoljnosti odabira točke, najjednostavnije je odabrati ishodište). Dokazao je zanimljiv rezultat opisan u sljedećem potpoglavlju.

5.2 Pólyine konstante slučajne šetnje

Promatramo li šetnju koja počinje iz ishodišta u d dimenzija gdje šetač čini beskonačno jediničnih koraka po d -dimenzionoj rešetci. Tada vjerojatnost jednog povratka u ishodište p_d zadovoljava:

$$1 + p_d + p_d^2 + \dots = m_d. \quad (5.1)$$

Kako se radi o geometrijskom redu s konstantim omjerom p_d vrijedi:

$$m_d = \frac{1}{1 - p_d}. \quad (5.2)$$

Vjerojatnost povratka u ishodište stoga glasi:

$$p_d = 1 - \frac{1}{m_d} \leq 1, \quad (5.3)$$

gdje je m_d :

$$m_d = \frac{d}{(2\pi)^d} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \left(d - \sum_{k=1}^d \cos(\theta_k) \right)^{-1} d\theta_1 d\theta_2 \dots d\theta_d. \quad (5.4)$$

Pólya je dokazao 1927. godine da m_d za $d=1$ i $d=2$ divergira, odnosno za jednu i dvije dimenzije šetač beskonačno puta posjećuje ishodište, pa je vjerojatnost posjećivanja ishodišta gotovo zagarantirana, $p_1 = p_2 = 1$. Također je dokazao da za $d > 2$ m_d konvergira, a iz toga slijedi da je $p_d < 1$. Račun za tri dimenzije glasi:

$$\begin{aligned} m_3 &= \frac{3}{(2\pi)^3} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{dx dy dz}{3 - \cos x - \cos y - \cos z} \\ &= \frac{\sqrt{6}}{32\pi^3} \Gamma\left(\frac{1}{24}\right) \Gamma\left(\frac{5}{24}\right) \Gamma\left(\frac{7}{24}\right) \Gamma\left(\frac{11}{24}\right) \\ &= 1,516386\dots \end{aligned} \quad (5.5)$$

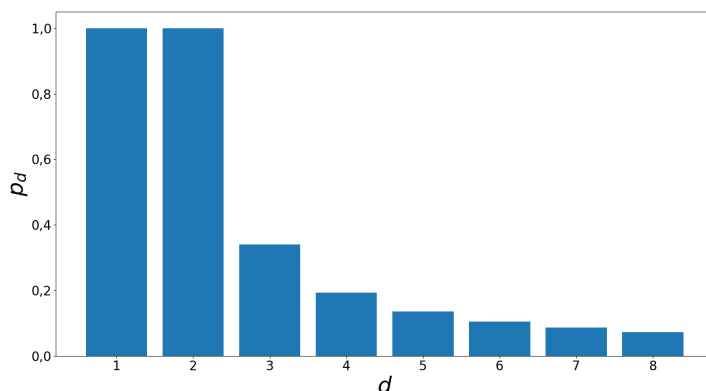
Iz čega proizlazi vjerojatnost posjećivanja ishodišta u tri dimenzije:

$$p_3 = 1 - \frac{1}{m_3} = 0,340537\dots < 1. \quad (5.6)$$

Zaključujemo da šetaču u tri dimenzije, za razliku od jedne i dvije, više nije garantiran povratak u ishodište. Pólyine konstante za više dimenzije prikazani su na tablici 5.1 i slici 5.1.

d	m_d	p_d
4	1.2394671218...	0.1932016732...
5	1.1563081248...	0.1351786098...
6	1.1169633732...	0.1047154956...
7	1.0939063155...	0.0858449341...
8	1.0786470120...	0.0729126499...

Tablica 5.1: Tablica Pólyinih konstanti nasumičnog hoda [7]



Slika 5.1: Grafički prikaz ovisnosti vjerojatnosti posjeta proizvoljne točke o dimenziji šetača.

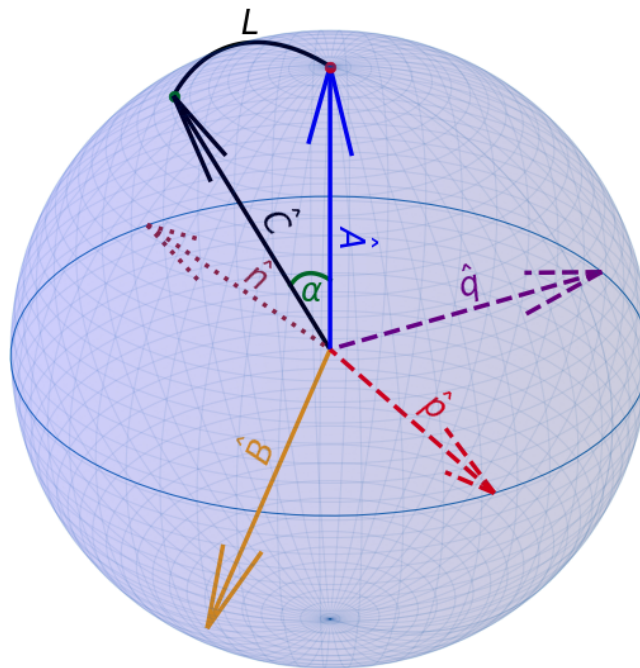
6 Nasumična šetnja po sferi

6.1 Zarobljena šetnja

U prethodnim smo poglavljima razmatrali šetnje gdje je prešutno pretpostavljeno kako je šetač slobodan kretati se po cijeloj ravnini. U ovom poglavlju razmatramo kako se nasumični šetač ponaša kada ga "zarobimo" u ograničen prostor, na površini sfere, te kakav utjecaj to ima na razvoj šetnje od tipičnih okolnosti šetača koji nije zarobljen.

6.2 Određivanje sfernog koraka

Određivanje novog koraka kod šetnje po jediničnoj sferi iz proizvoljne nije točke banalno. Razlog tome je da kada bismo za proizvoljnu točku zadanu u sfernom koordinatnom sustavu ($R = 1, \theta_0, \phi_0$) uzimali uniformno generirane kutne pomake $\Delta\theta$ i $\Delta\phi$, ne bismo dobili šetnju po sferi s fiksiranom duljinom koraka, već samo neujednačene skokove po sferi. Stoga se moramo pobrinuti da na ispravan način generiramo nasumičan smjer koraka, gdje onda šetač napravi korak zadane duljine L .



Slika 6.1: Ilustracija generiranja nasumičnog koraka po sferi.

Ilustracija jednog od mogućih načina prikazana je na slici 6.1. Počevši od početne točke koja se nalazi u sjevernom polu, koju tretiramo kao ishodište sferne šetnje i označavamo crvenom bojom, u koju pokazuje vektor \hat{A} . Kako bismo odredili smjer koraka, potrebna nam je ortonormirana baza, koju postizemo generiranjem nasumičnog vektora $\hat{B} \neq \pm\hat{A}$ i uzimanjem

vektorskih produkata:

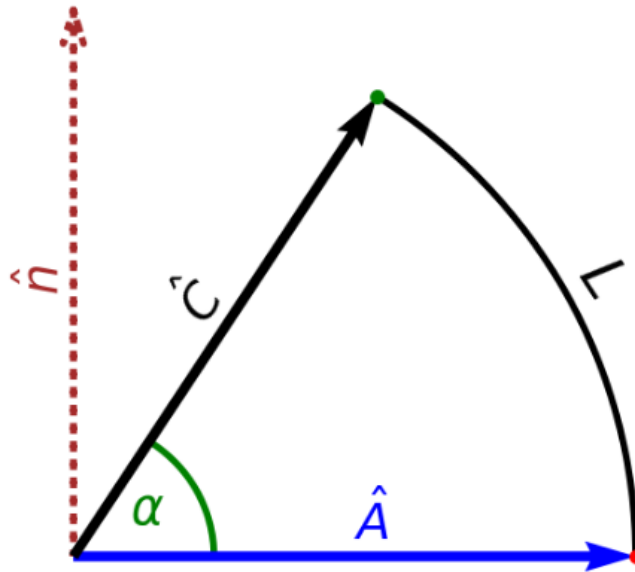
$$\hat{p} = \frac{\hat{A} \times \hat{B}}{\|\hat{A} \times \hat{B}\|}, \text{ i } \hat{q} = \frac{\hat{A} \times \hat{p}}{\|\hat{A} \times \hat{p}\|}. \quad (6.7)$$

Sada kada smo konstruirali ortonormiranu bazu, generiramo kut iz uniformne raspodjele $\phi \in [0, 2\pi)$ na ravnini koju razapinju vektori \hat{p} i \hat{q} , preko kojih zapisujemo vektor \hat{n} u polarnom obliku, koji pokazuje nasumični smjer koraka šetača:

$$\hat{n} = \cos \phi \hat{p} + \sin \phi \hat{q}. \quad (6.8)$$

Nakon određenog smjera koraka, željenu duljinu koraka L postizemo preko jednadžbe duljine kružnog luka, sa zadanim fiksnim pomakom α , zadanim u radijanima:

$$L = R\alpha \implies \alpha = \frac{L}{R}. \quad (6.9)$$



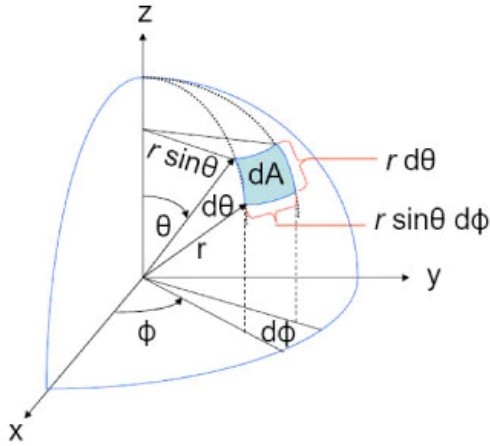
Slika 6.2: Ilustracija konačnog pomaka nasumičnog šetača po sferi nakon određivanja nasumičnog smjera.

Konačni položaj \hat{C} (prikazan na slici 6.2) zapisujemo:

$$\hat{C} = \cos \alpha \hat{A} + \sin \alpha \hat{n}. \quad (6.10)$$

Ovaj proces ponavljamo kod svakog koraka, gdje kod prethodnog koraka konačni vektor \hat{C} predstavlja novi početni vektor \hat{A} .

6.3 Određivanje kutne varijable



Slika 6.3: Površina elementa sferne ljuske [12].



Slika 6.4: Bočni prikaz sferne rešetke [13].

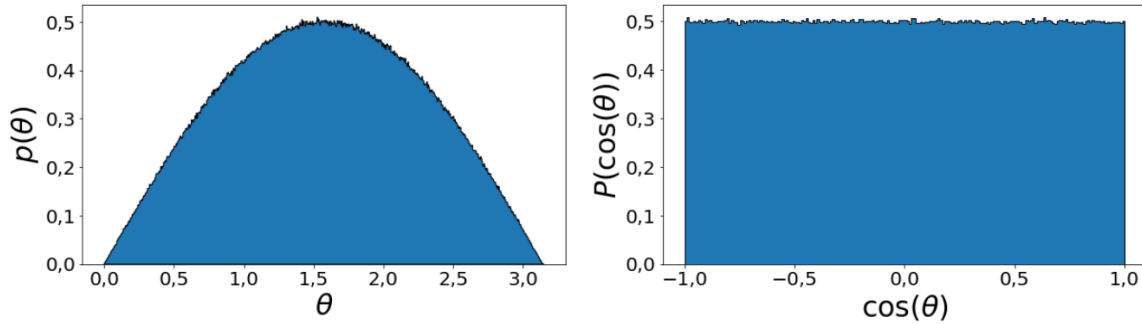
Nakon što smo odredili kako generirati same korake šetnje, također trebamo odrediti i koju varijablu proučavamo. Kako za ishodište uzimamo sjeverni pol, intuitivno bi bilo promatrati sferni kut θ koji predstavlja kutni pomak od sjevernog pola. Površina segmenta sferne ljuske je $dA = r^2 d\Omega$ (prikazano na slici 6.3) gdje je:

$$d\Omega(\theta, \phi) = \sin \theta d\theta d\phi, \quad (6.11)$$

te možemo zamijetiti kako površina segmenta sferne ljuske nije linearno zavisna o kutu θ , već zavisi ujedno i o sinusu tog kuta. Grafički prikazano na slici 6.4 vidimo kako se površina smanjuje udaljavanjem od ekvatora. Stoga je neuputno promatrati raspodjelu kutnih pomaka $\Delta\theta$ jer ovisi o položaju na kojem se šetač nalazi. Kako bi se uklonila ta ovisnost, uvodimo supstituciju $u = \cos \theta$ čiji je diferencijal $du = -\sin \theta d\theta$, kojom postizemo linearan površinski element:

$$d\Omega = \sin \theta d\theta d\phi = -d(\cos \theta) d\phi = -du d\phi. \quad (6.12)$$

Računalnom simulacijom sferne šetnje radimo provjeru raspodjele kuta θ i predložene supstitucije $u = \cos \theta$.

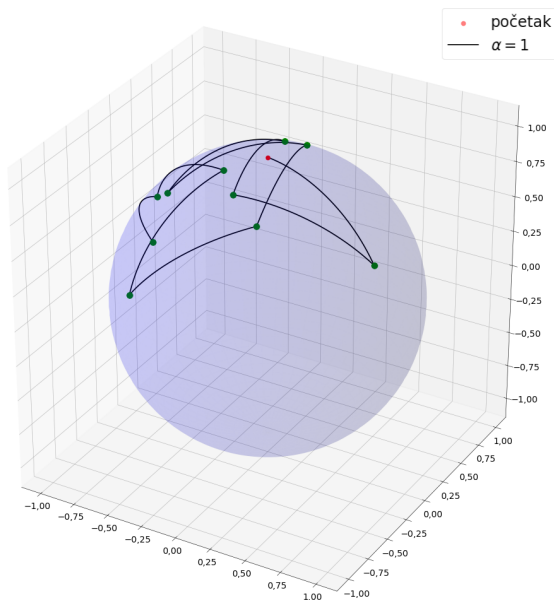


Slika 6.5: Kutne raspodjele simulacije sfernog šetača nakon $k = 10^7$ koraka s fiksnom duljinom kutnog pomaka $\alpha = 1$.

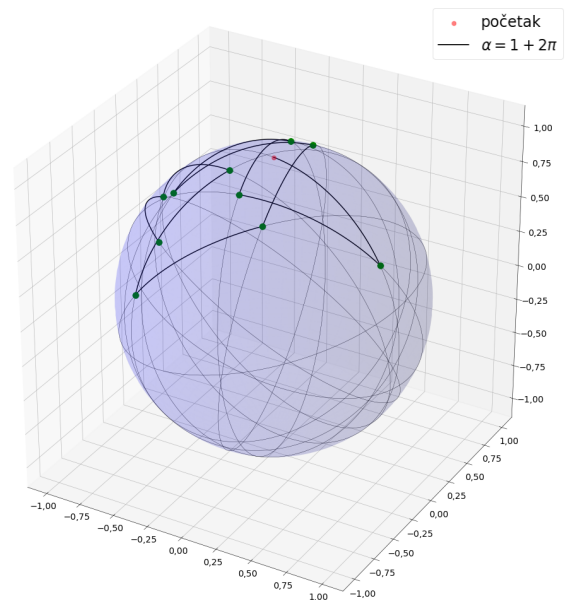
Očekivano, iz simulacije prikazanoj na slici 6.5, jasno se vidi da raspodjela kuta θ nije uniformna, te kako poprima oblik gornje polovice funkcije sinusa (zbog domene kuta θ), dok je raspodjela kosinusa kuta θ uniformna. Zaključujemo da je zato kosinus "prirodna" kutna varijabla čiju raspodjelu razmatramo u ostatku poglavlja.

6.4 Cikličnost

Kako smo prethodno naveli, temeljna razlika kod sfernog šetača je njegova zarobljenost, jer mu je prostor dostupan za šetanje konačan.



Slika 6.6: Prikaz sfernog šetača nakon $k = 10$ koraka s duljinom koraka $\alpha = 1$.



Slika 6.7: Prikaz sfernog šetača nakon $k = 10$ koraka s duljinom koraka $\alpha = 1 + 2\pi$.

To je prikazano primjerom na slici 6.6 gdje je fiksni kutni pomak $\alpha = 1$ i na primjeru slika 6.7, gdje šetač kutni pomak uvećan za 2π . Deblja linija predstavlja šetačev pomak, a tanja

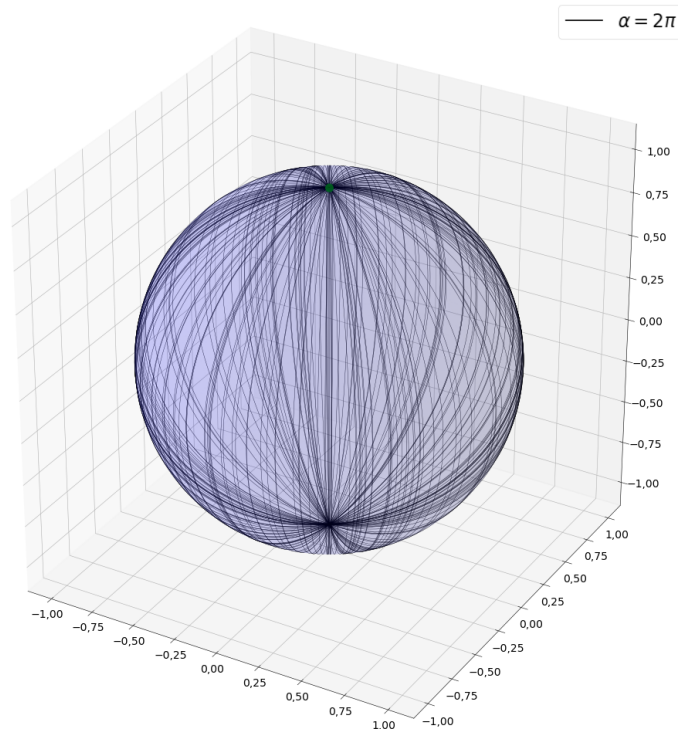
predstavlja šetačev put. Vidimo kako makar šetač čiji je korak uvećan za 2π prelazi veći put, njegov ukupni pomak ostaje jednak. Takav rezultat nije nimalo začuđujuć, ali svejedno predstavlja jako bitno svojstvo šetnje po sferi, gdje nam svojstvo cikličnosti ukazuje kako će raspodjela šetnji uvećana za svaki cjelobrojni višekratnik od 2π rezultirati istom raspodjelom:

$$p(\alpha) = p(\alpha + 2n\pi). \quad (6.13)$$

6.5 Analitički primjeri

Kod sferne šetnje možemo analitički iskazati malen broj šetnji. Primjer takve šetnje prikazan je na slici 6.8. Šetač započinje šetnju iz sjevernog pola s kutnim pomakom $\alpha = 2\pi$ i svakim pomakom se vraća u sjeverni pol, tj. on ostaje zarobljen na tom mjestu, što je ujedno i primjer cikličnosti. Kada šetač umjesto toga radio pomak $\alpha = \pi$, on bi skakao svakim korakom između sjevernog i južnog pola sfere. Ovakve šetnje možemo općenito iskazati na sljedeći način:

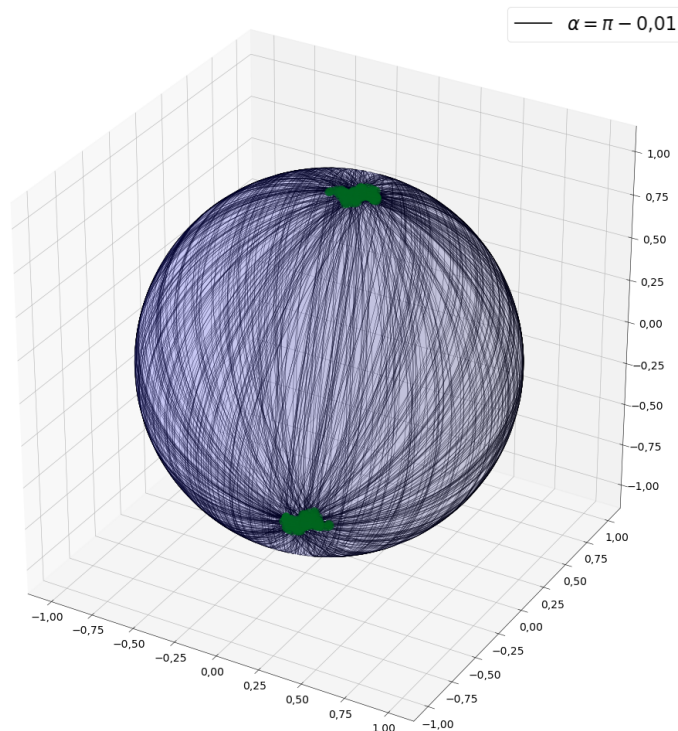
$$p_k(\theta; \alpha = n\pi) = \begin{cases} \delta(\theta), & \text{n ili k je paran,} \\ \delta(\theta - \pi), & \text{n i k neparni.} \end{cases} \quad (6.14)$$



Slika 6.8: Prikaz sfernog šetača nakon $k = 10$ koraka s fiksnim kutnim pomakom $\alpha = 2\pi$.

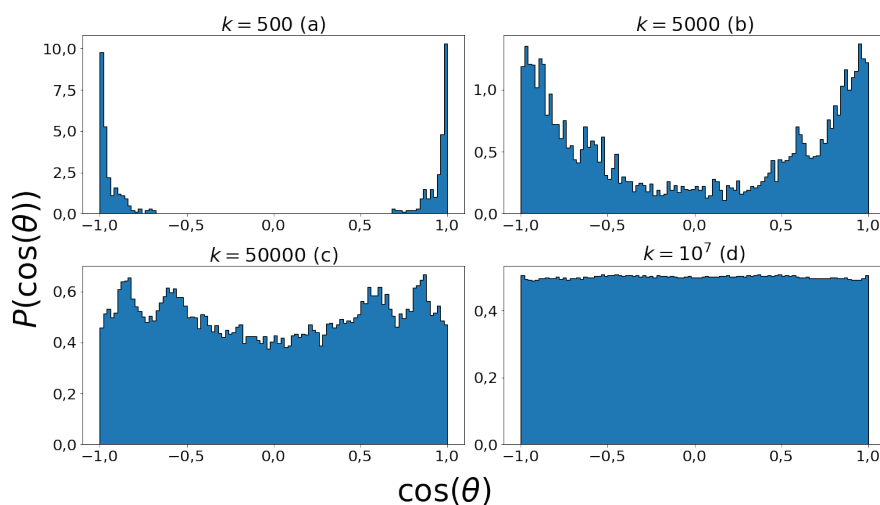
6.6 Generiranje smjerova

Uzimajući prethodni primjer gdje je šetač bio zarobljen na polovima, vrijedi razmatrati što će se dogoditi s raspodjelom položaja za malu izmjenu duljine koraka.



Slika 6.9: Prikaz sfernog šetača nakon $k = 100$ koraka s duljinom koraka $\alpha = \pi - 0,01$.

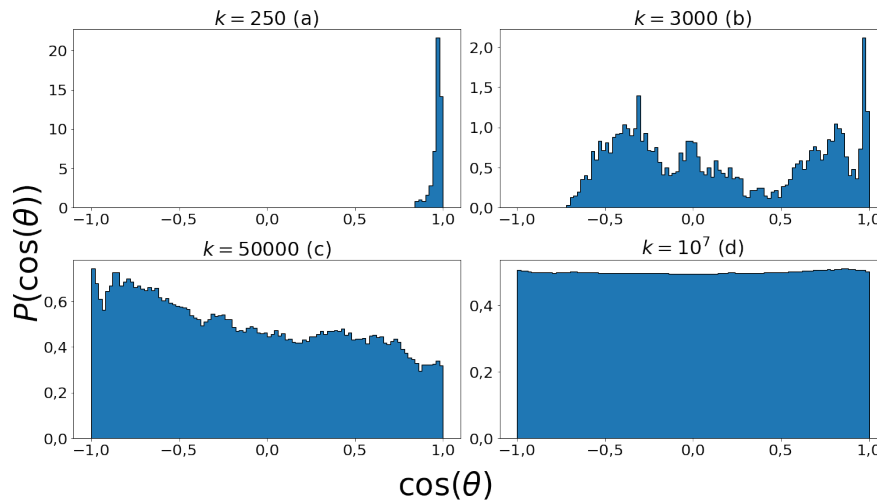
Kako je prikazano na slici 6.9, šetač se za malo odstupanje od cjelobrojnog višekratnika π "oslobodio" od polova, te se polako odmiče od njih. Stoga vrijedi razmatrati daljni razvoj raspodjele simuliranjem više koraka.



Slika 6.10: Kutne raspodjele sfernog šetača s fiksnim kutnim pomakom $\alpha = \pi - 0,01$ za (a) $k = 500$ koraka, (b) $k = 5000$ koraka, (c) $k = 50000$ koraka, (d) $k = 10^7$ koraka.

Slika 6.10 prikazuje razvoj raspodjele do određenog koraka. Vidimo kako se u početku simetrično nakuplja na polovima, no kako se broj koraka povećava, nalikuje na dva vala koja se šire jedan prema drugome dok se konačno ne postigne uniformna raspodjela.

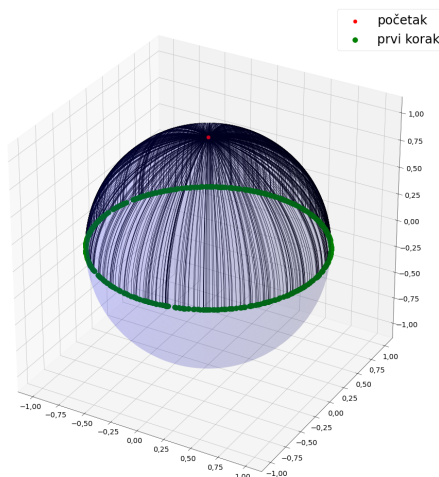
Također promatramo razvoj šetnje gdje šetač radi korake s malim odstupanjem od 2π prikazano na slici 6.11.



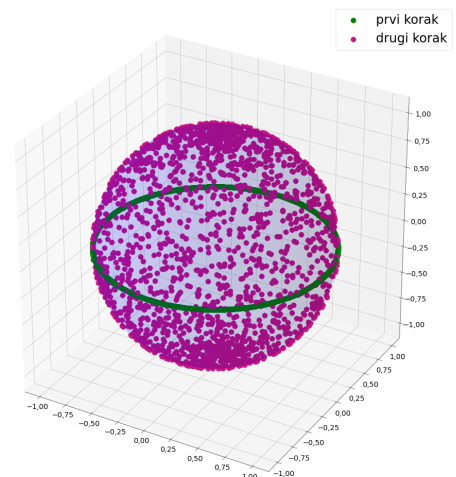
Slika 6.11: Kutne raspodjele sfernog šetača s fiksnim kutnim pomakom $\alpha = 2\pi - 0,01$ za (a) $n = 250$ koraka, (b) $n = 3000$, (c) $n = 50000$ koraka, (d) $n = 10^7$ koraka.

Ovaj put razvoj raspodjele nalikuje na val koji se širi s jedne strane dok ne dođe na zid od kojeg se odbija, sve dok na kraju ne postane uniformna raspodjela.

U sljedećem primjeru promatramo raspodjelu više istovremenih šetnji s fiksnim kutnim pomakom $\alpha = \frac{\pi}{2}$. Kako svaki šetač kreće s sjevernog pola može imati samo silaznu putanju, nakon prvog koraka svi šetači se isključivo nalaze na ekvatoru, kao što je prikazano na slici 6.12.



Slika 6.12: Simulacija položaja $n = 2000$ sfernih šetača nakon prvog koraka s fiksnim kutnim pomakom $\alpha = \frac{\pi}{2}$.

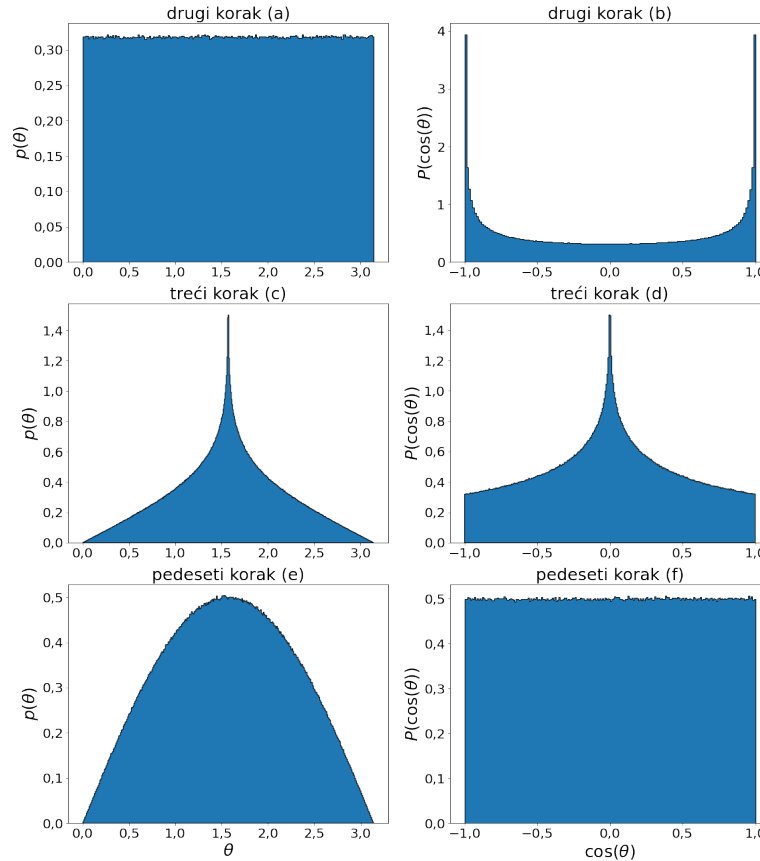


Slika 6.13: Simulacija položaja $n = 2000$ sfernih šetača nakon drugog koraka s fiksnim kutnim pomakom $\alpha = \frac{\pi}{2}$.

Raspodjele prvog koraka se daju analitički zapisati:

$$p_1(\theta; \alpha = \pi/2) = \delta\left(\theta - \frac{\pi}{2}\right), \quad P_1(\cos(\theta); \alpha = \pi/2) = \delta(\cos(\theta)), \quad (6.15)$$

ali već kod drugog koraka (slika 6.13) čini se kako se šetači nakupljaju oko polova. Stoga u ovom slučaju ujedno razmatramo i raspodjele kutova θ .



Slika 6.14: Kutne raspodjele $n = 10^7$ sfernih šetača s fiksnim kutnim pomakom $\alpha = \frac{\pi}{2}$ za drugi korak (a) i (b), treći korak (c) i (d) i pedeseti korak (e) i (f).

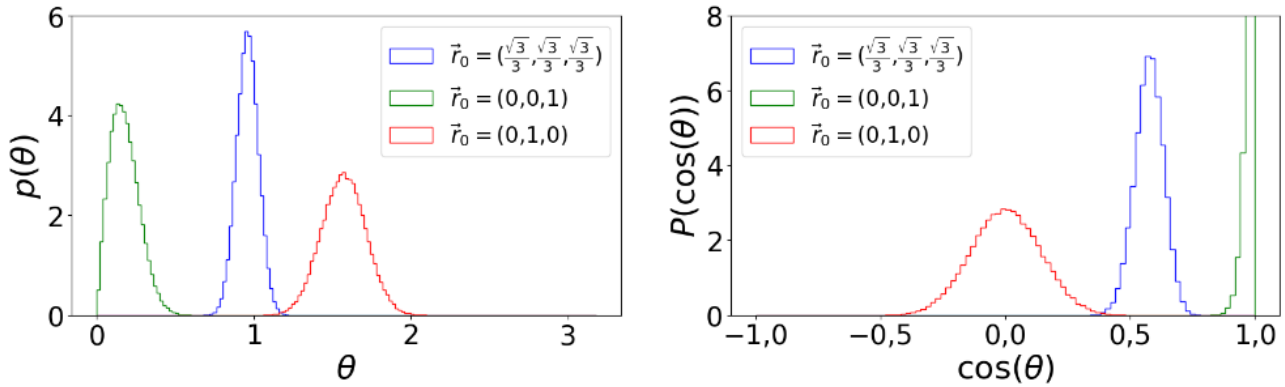
Vidimo na slici 6.14 zanimljivu pojavu da kod drugog koraka raspodjela $p_2(\theta; \alpha = \pi/2)$ postaje uniformnu raspodjelu, a raspodjela $P_2(\cos(\theta))$ ima integrabilne singularitete poput one na slici 4.4⁹, gdje je najveća vrijednost raspodjele oko polova. Nakon pedesetog koraka dobivamo isti rezultat kao na slici 6.5, ali i začuđujuć rezultat kod promatranja kutne varijable. Naime u poglavlju 4 smo proučavali, među ostalime, Centralni granični teorem na raspodjeli arkus sinusa, gdje za drugi i treći korak dobivamo oblikom slične raspodjele, no konačna raspodjela ne poprima oblik gausijanske krivulje, već uniformne.

⁹Analički zapis navedenih raspodjela:

$$p_2(\theta; \alpha = \pi/2) = \begin{cases} \frac{1}{\pi}, & \text{za } 0 \leq \theta \leq \pi, \\ 0, & \text{inače.} \end{cases} \quad P_2(\cos(\theta); \alpha = \pi/2) = \begin{cases} \frac{1}{\pi\sqrt{1-\cos(\theta)^2}}, & \text{za } -1 < \cos(\theta) < 1, \\ 0, & \text{inače.} \end{cases}$$

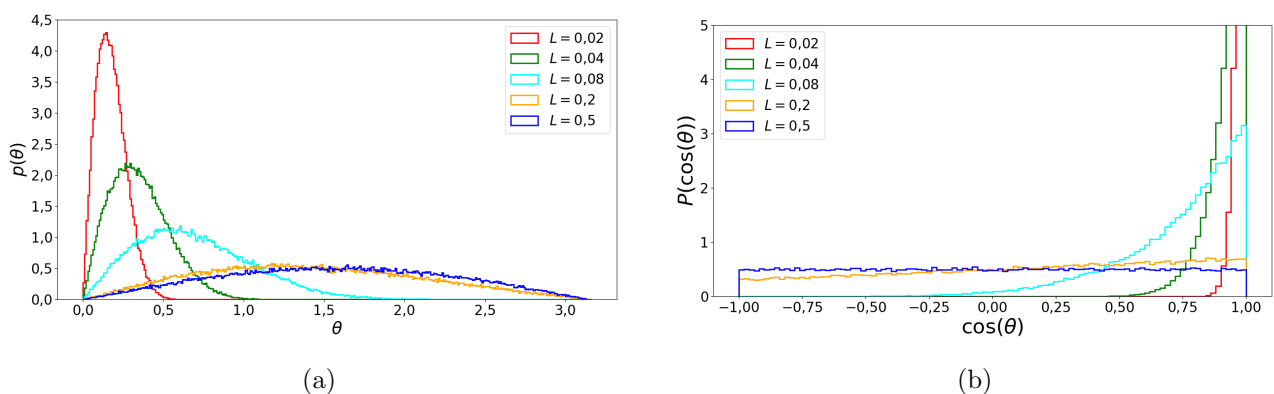
6.7 Centralni granični teorem na sferi

Kako CGT naizgled ne vrijedi za kutne raspodjele šetnje po sferi, razmatramo razvoj kutnih raspodjela malog broja kratkih koraka, za koje ne nastupa svojstvo cikličnosti.



Slika 6.15: Kutne raspodjele konačnih položaja $n = 10^5$ sfernih šetača nakon $k = 100$ koraka s fiksnim kutnim pomakom $\alpha = 0,01$ za različita ishodišta \vec{r}_0 .

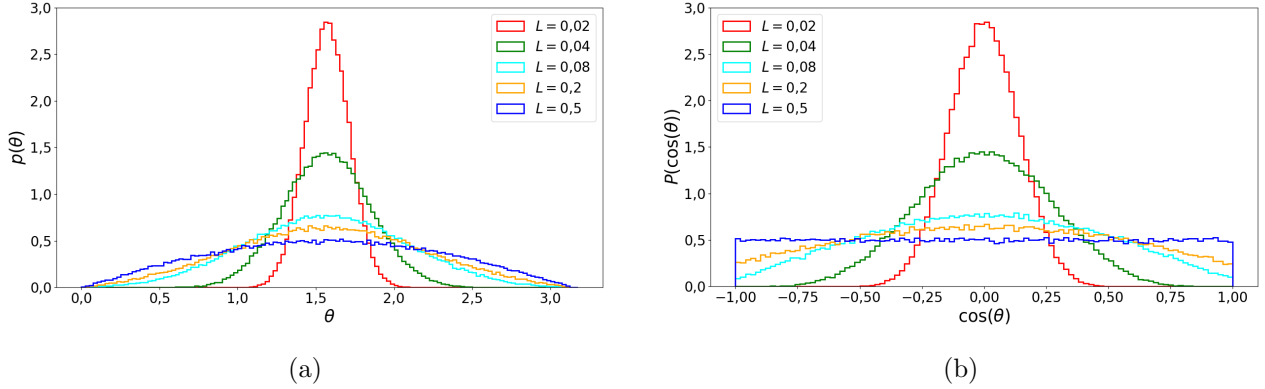
Kako je prikazano na slici 6.15, raspodjela sa ishodištem na ekvatoru ($\vec{r}_0 = (0, 1, 0)$) naizgled poprma oblik gausijanske krivulje, gdje se razvoj šetnje propagira jednako vjerojatno u svakom smjeru, dok raspodjela s ishodištem u sjevernom polu ($\vec{r}_0 = (0, 0, 1)$) ne. Razlog tome je da kako se šetač na sjevernom polu nalazi na vrhu sfere on se iz tog položaja jedino može spuštati prema dolje, što se iskazuje isključivim rastom sfernog kuta θ , odnosno isključivim padom kosinusa tog kuta. Time zaključujemo kako kutne raspodjele ovise o tome gdje šetnja započinje, te dalje razmatramo kako raspodjele izgledaju kada nastupi svojstvo cikličnosti.



Slika 6.16: Kutne raspodjele konačnih položaja $n = 10^5$ sfernih šetača nakon $k = 100$ koraka za fiksne kutne pomake α s ishodištem \vec{r}_0 na sjevernom polu.

Na slikama 6.16 i 6.17 vidimo kako raspodjele zadržavaju početni oblik, sve dok ne naiđu na "rub" svoje domene, nakon kojega teže u raspodjelu oblika $p(\theta) = \frac{1}{2} \sin(\theta)$ u slučaju 6.16a i 6.17a, tj. $P(\cos(\theta)) = \frac{1}{2}$ za 6.16b i 6.17b. Zaključujemo kako je normalna raspodjela tek

dobra aproksimacija za malen broj kratkih koraka, te nam u ovom slučaju CGT ne daje dobru reprezentaciju raspodjela.



Slika 6.17: Kutne raspodjele konačnih položaja $n = 10^5$ sfernih šetača nakon $k = 100$ koraka za fiksne kutne pomake α s ishodištem \vec{r}_0 na ekvatoru.

Raspodjela koja ispravno opisuje sfernu šetnju poznata je kao von Mises-Fisher (skraćeno vMF) raspodjela koja predstavlja obitelj raspodjela na površini hipersfera u d dimenzija, te joj je funkcija gustoće vjerojatnosti:

$$p_d(\vec{R}; \hat{\mu}, \kappa) = C_d(\kappa) e^{\kappa \hat{\mu}^T \cdot \vec{R}}, \quad (6.16)$$

gdje parametar $\hat{\mu}$ predstavlja srednji smjer raspodjele (ishodište), a parametar κ mjeru koncentracije raspodjele¹⁰, gdje je konstanta normalizacije za d dimenzija:

$$C_d(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)}. \quad (6.17)$$

U slučaju tri dimenzije jednadžbe možemo zapisati u zavisnosti o sfernim kutovima, gdje je položaj srednjeg smjera u sfernim koordinatama $\hat{\mu} = (\sin \vartheta \cos \varphi, \sin \vartheta \sin \varphi, \cos \vartheta)$:

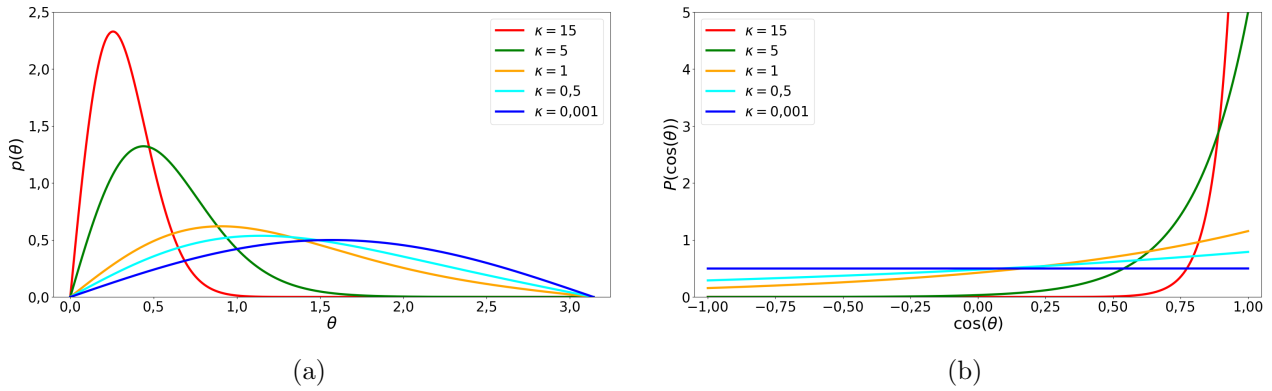
$$p(\theta, \phi; \kappa) = C_3(\kappa) \sin(\theta) e^{\kappa [\sin(\theta) \sin(\vartheta) \cos(\phi - \varphi) + \cos(\theta) \cos(\vartheta)]}. \quad (6.18)$$

Funkciju gustoće vjerojatnosti za raspodjelu kosinusa kuta (koristeći supstituciju $u = \cos(\theta)$) zapisujemo:

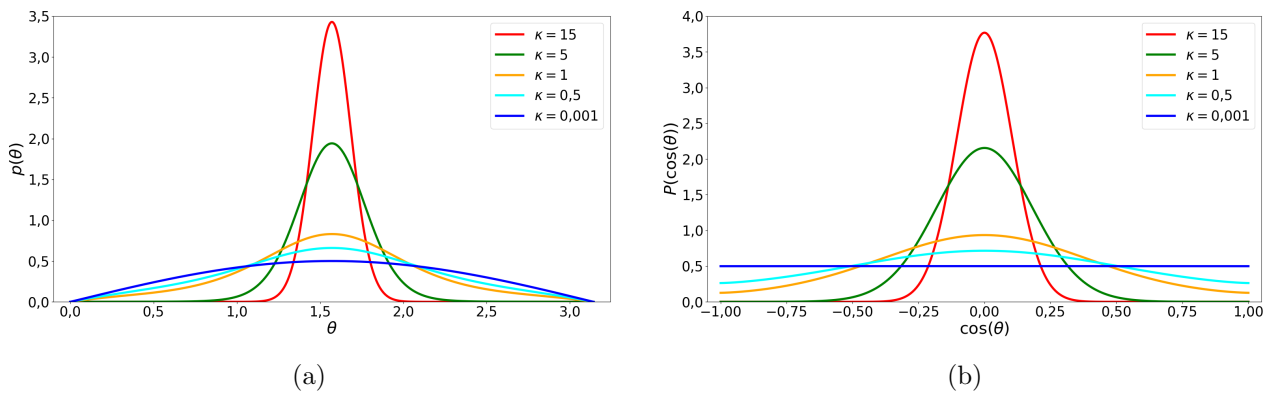
$$P(\cos(\theta), \phi; \kappa) = C_3(\kappa) e^{\kappa [\sqrt{(1-u^2)} \cos(\phi - \varphi) \sin(\vartheta) + \cos(\vartheta)u + \sqrt{(1-u^2)}]}. \quad (6.19)$$

Iz grafičkih prikaza na slikama 6.18 i 6.19 vidimo kako je u skladu s rezultatima 6.16 i 6.17.

¹⁰Parametri $\hat{\mu}$ i $\frac{1}{\kappa}$ su analogni parametrima μ i σ^2 kod normalne raspodjele.



Slika 6.18: Grafički prikaz von Mises-Fisher raspodjele s parametrom $\hat{\mu}$ na sjevernom polu za različite koncentracije κ .



Slika 6.19: Grafički prikaz von Mises-Fisher raspodjele s parametrom $\hat{\mu}$ na ekvatoru za različite koncentracije κ .

Kada je parametar κ malen, on predstavlja raspodjelu koja je jako raspršena (slabo koncentrirana) raspodjela teži u uniformnu raspodjelu, čiji limes možemo predstaviti kao:

$$\lim_{\kappa \rightarrow 0} \frac{\kappa e^{\kappa}}{2\pi(e^{\kappa} - e^{-\kappa})} = \frac{1}{4\pi}, \quad (6.20)$$

te su stoga raspodjele:

$$p(\theta, \phi; \kappa \rightarrow \infty) = \frac{\sin(\theta)}{4\pi}, \quad P(\cos(\theta), \phi; \kappa \rightarrow \infty) = \frac{1}{4\pi}, \quad (6.21)$$

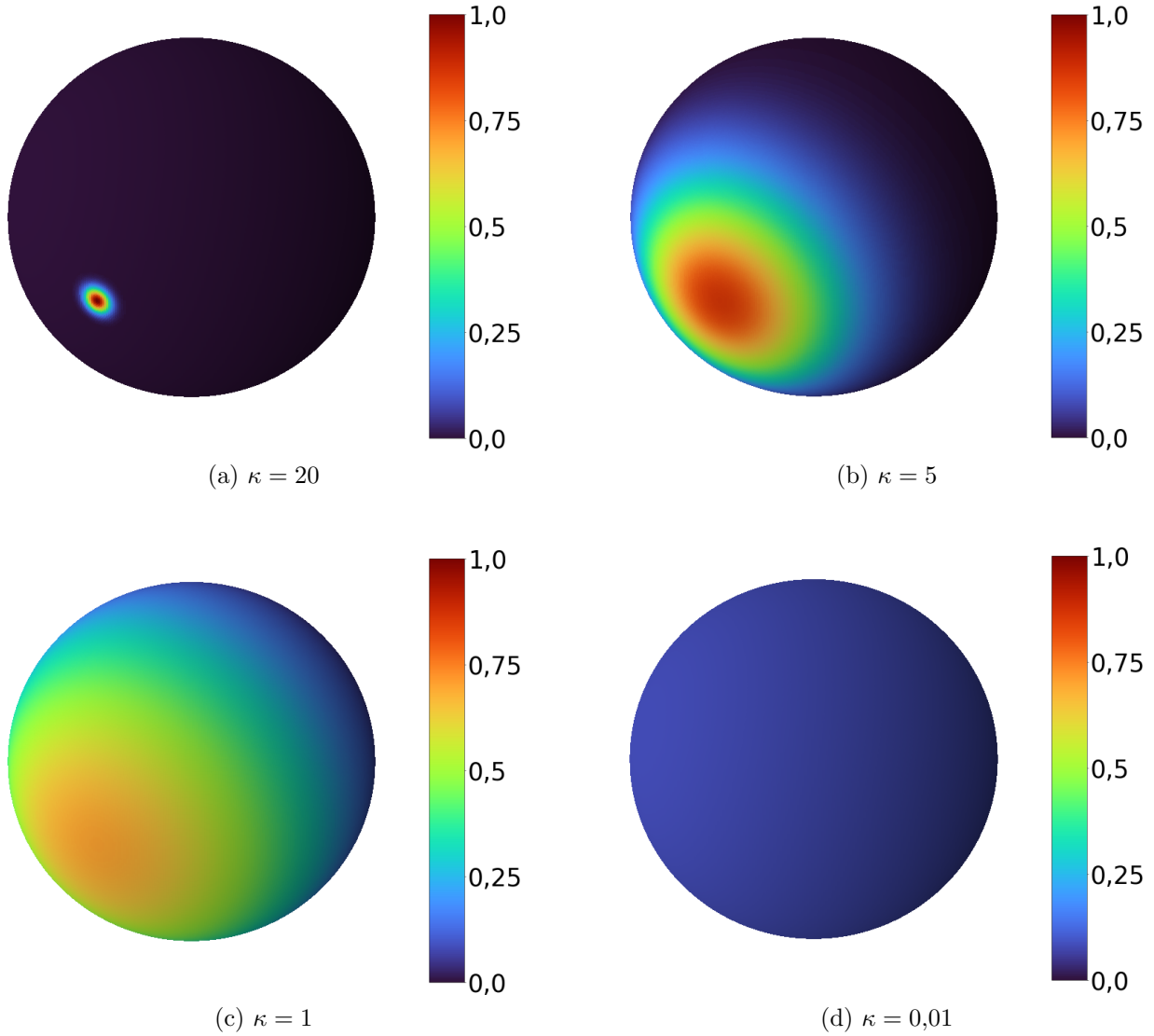
neovisne o parametru $\hat{\mu}$. Također, ukoliko parametar κ teži u beskonačnost, sva vjerojatnost je lokalizirana u jednu točku te dobivamo degeneriranu raspodjelu koja je opisana Diracovom delta funkcijom:

$$\lim_{\kappa \rightarrow \infty} \frac{\kappa e^{\kappa}}{2\pi(e^{\kappa} - e^{-\kappa})} = \infty. \quad (6.22)$$

U tom slučaju raspodjele poprimaju oblike:

$$p(\theta, \phi; \kappa \rightarrow \infty) = \delta(\theta - \vartheta)\delta(\phi - \varphi), \quad (6.23)$$

$$P(\cos(\theta), \phi; \kappa \rightarrow \infty) = \delta(\cos(\theta) - \cos(\vartheta))\delta(\phi - \varphi). \quad (6.24)$$



Slika 6.20: Prikaz von Mises–Fisher raspodjela na površini sfere s različitim parametrima κ i ishodištem raspodjele u $\hat{\mu} = (0, -1, 0)$ u Kartezijevom koordinatom sustavu.

Kada prikažemo vMF raspodjelu na sferi, može nam služiti kao ilustracija difuzije na površini sfere. Iz lokalizirane raspodjele (slika 6.20a), vidimo kako smanjenjem parametra koncentracije κ se raspodjela širi (slika 6.20b i 6.20c) te na kraju postaje uniformna (slika 6.20d). Zaključujemo kako se CGT na sfernoj šetnji realizira na drugačiji način, gdje je vMF raspodjela analogna normalnoj.

7 Poveznica raspodjela

7.1 Fizikalna veza

U prethodnim poglavljima smo razmatrali različite distribucije i statističke alate. U izradi matematičkih modela kojima prikazujemo fizikalne procese smo zato radili određene pretpostavke kako bismo pojednostavili modeliranje, stoga u ovom poglavlju povezujemo nasumični proces čiji događaji su diskretni (poput sudara čestica kod Brownovog gibanja) a događaju se u kontinuiranom vremenskom intervalu, raspodjelu koja opisuje vjerojatnosti opažanja tih događaja, gdje ćemo ujedno bez ikakvih nametanja izvesti i eksponencijalnu raspodjelu koja daje opis vremenskog razmaka između tih događaja.

7.2 Izvod intervalne raspodjele

Pretpostavke intervalne raspodjele:

1. Neka je $k \in \mathbb{Z}^+$ broj ostvarenih događaja unutar određenog intervala.
2. Pojava jednog događaja ne utječe na vjerojatnost drugog događaja, tj. događaji su nezavisni.
3. Vjerojatnost za svaki pojedini događaj je konstantna.
4. Dva događaja se ne mogu dogoditi u istom trenutku, tj. kako pratimo diskretne događaje s kontinuiranim prolaskom vremena, vjerojatnost dvaju (ili više) događaja je zanemariva, stoga se u svakom trenutku ostvari najviše jedan događaj.

Zapišemo li preko tih pravila vremensku ovisnost raspodjele, dobivamo:

$$\frac{\partial P_k(t)}{\partial t} = a [P_{k-1}(t) - P_k(t)]. \quad (7.1)$$

U slučaju kad tek počinjemo pratiti proces koji je opisan ovom raspodjelom, odnosno za početni događaj $k = 0$, ne postoji prethodni događaj (u ovome slučaju $k = -1$). Te se izraz pojednostavnjuje u:

$$\frac{\partial P_0(t)}{\partial t} = -aP_0(t). \quad (7.2)$$

Ovu diferencijalnu jednadžbu možemo riješiti metodom separacije varijabli:

$$\int_{P_0}^{P_0(t)} \frac{dP'_0}{P'_0} = -a \int_0^t dt', \quad (7.3)$$

čije je rješenje:

$$P_0(t) = e^{-at}. \quad (7.4)$$

Kako bismo odredili općeniti izraz, prikazujemo Fourierovom transformacijom.

$$P_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_k(\omega) e^{i\omega t} d\omega. \quad (7.5)$$

Vrijeme koje prethodi početak promatranja nema utjecaja te to svojstvo ostvarujemo množenjem step funkcijom $\Theta(t)$:

$$P_0(t) = e^{-at} \cdot \Theta(t), \quad \Theta(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0. \end{cases} \quad (7.6)$$

Ako parcijalno deriviramo po vremenu jednadžbu 7.5 dobivamo izraz:

$$\begin{aligned} \frac{\partial P_k(t)}{\partial t} &= \frac{1}{2\pi} \frac{\partial}{\partial t} \int_{-\infty}^{\infty} F_k(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial t} (F_k(\omega) e^{i\omega t}) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} i\omega F_k(\omega) e^{i\omega t} d\omega, \end{aligned} \quad (7.7)$$

a drugi izraz dobijamo uvrštavanjem jednadžbe 7.1:

$$\begin{aligned} \frac{\partial P_k(t)}{\partial t} &= \mathcal{F}^{-1} \left[\mathcal{F} \left\{ \frac{\partial P_k(t)}{\partial t} \right\} \right] = \mathcal{F}^{-1} [\mathcal{F} \{a [P_{k-1}(t) - P_k(t)]\}] \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} a [F_{k-1}(\omega) - F_k(\omega)] d\omega. \end{aligned} \quad (7.8)$$

Izraze 7.7 i 7.8 možemo izjednačiti:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} a [F_{k-1}(\omega) - F_k(\omega)] d\omega &= \frac{1}{2\pi} \int_{-\infty}^{\infty} i\omega F_k(\omega) e^{i\omega t} d\omega \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} \underbrace{\left(a [F_{k-1}(\omega) - F_k(\omega)] - i\omega F_k(\omega) \right)}_{=0} e^{i\omega t} d\omega &= 0. \end{aligned} \quad (7.9)$$

Rješenje integrala može biti jednako nuli, samo ako je izraz u zagradi jednak nuli, zbog toga dobivamo:

$$a [F_{k-1}(\omega) - F_k(\omega)] - i\omega F_k(\omega) = 0, \quad (7.10)$$

koji kada sredimo dobivamo rekurzivnu relaciju:

$$F_k(\omega) = \frac{a}{a + i\omega} F_{k-1}(\omega) = \frac{a^2}{(a + i\omega)^2} F_{k-2}(\omega) = \dots = \frac{a^k}{(a + i\omega)^k} F_0(\omega). \quad (7.11)$$

Kako bismo odredili $F_0(\omega)$ potrebno je napraviti Fourierovu transformaciju od $P_0(t)$:

$$\begin{aligned} F_0(\omega) &= \mathcal{F}\{P_0(t)\} = \int_{-\infty}^{\infty} P_0(t) e^{-i\omega t} dt = \int_{-\infty}^{\infty} e^{-at} \Theta(t) e^{-i\omega t} dt = \\ &= \int_0^{\infty} e^{-(a+i\omega)t} dt = \frac{1}{a + i\omega}, \end{aligned} \quad (7.12)$$

koji kada uvrstimo u jednadžbu 7.11 dobijemo:

$$F_k(\omega) = \frac{a^k}{(a + i\omega)^{k+1}}, \quad (7.13)$$

a opći član raspodjele sada glasi:

$$P_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{a^k}{(a + i\omega)^{k+1}} e^{i\omega t} d\omega \quad (7.14)$$

Ovaj integral rješavamo metodom parcijalne integracije gdje su produkti funkcija:

$$f'(\omega) = \frac{1}{(a + i\omega)^{k+1}}, \quad g(\omega) = e^{i\omega t}. \quad (7.15)$$

Prvom integracijom dobivamo:

$$P_k(t) = \frac{a^k i}{2\pi k} (a + i\omega)^{-k} e^{i\omega t} \Big|_{-\infty}^{\infty} - \frac{i^2 t a^k}{k 2\pi} \int_{-\infty}^{\infty} (a + i\omega)^{-k} e^{i\omega t} dt. \quad (7.16)$$

Ako kompleksnu funkciju $h(\omega) = a + i\omega$ prebacimo u polarni oblik:

$$a + i\omega = \sqrt{a^2 + \omega^2} e^{i\theta}, \quad \tan(\theta) = \frac{\omega}{a}, \quad (7.17)$$

gdje je izraz čiji limes trebamo odrediti oblika $e^{ix}/|x|$. Zbog $|e^{ix}| = 1$ imamo:

$$\lim_{x \rightarrow \pm\infty} \frac{e^{ix}}{|x|^k} = 0, \quad (7.18)$$

gdje sa zadanim granicama integracije teži u nulu:

$$\begin{aligned} \lim_{\omega \rightarrow +\infty} (a + i\omega)^{-k} e^{i\omega t} &= \lim_{\omega \rightarrow +\infty} (a^2 + \omega^2)^{-\frac{1}{2}k} e^{i(\omega t - k\theta)} = 0, \\ \lim_{\omega \rightarrow -\infty} (a + i\omega)^{-k} e^{i\omega t} &= \lim_{\omega \rightarrow -\infty} (a^2 + \omega^2)^{-\frac{1}{2}k} e^{i(\omega t - k\theta)} = 0. \end{aligned} \quad (7.19)$$

Stoga je čitav član jednak nuli.

$$\frac{i}{2\pi k} (a + i\omega)^{-k} e^{i\omega t} \Big|_{-\infty}^{\infty} = \frac{i}{2\pi k} \left(\lim_{\omega \rightarrow +\infty} \frac{e^{i(\omega t - k\theta)}}{(a^2 + \omega^2)^{\frac{1}{2}k}} - \lim_{\omega \rightarrow -\infty} \frac{e^{i(\omega t - k\theta)}}{(a^2 + \omega^2)^{\frac{1}{2}k}} \right) = 0. \quad (7.20)$$

Ponovimo li isti postupak k puta, dobivamo:

$$P_k(t) = -\frac{a^k i^2 t}{k} \frac{1}{2\pi} \int_{-\infty}^{\infty} (a + i\omega)^{-k} e^{i\omega t} dt = \frac{a^k (-1)^k (i^2)^k t^k}{k!} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{(a + i\omega)} e^{i\omega t} dt. \quad (7.21)$$

Prepoznamo član koji je inverzna Fourierova transformacija od 7.12 te kako vrijedi:

$$P_0(t) \stackrel{\mathcal{F}}{=} F_0(\omega), \quad (7.22)$$

dobivamo:

$$P_k(t) = \frac{a^k (-1)^{2k} t^k}{k!} \mathcal{F}^{-1}(F_0(\omega)) = \frac{a^k t^k}{k!} P_0(t). \quad (7.23)$$

Uvrštavanjem jednadžbe 7.6 dobivamo konačno rješenje:

$$P_k(t) = \frac{a^k t^k e^{-at} \Theta(t)}{k!}. \quad (7.24)$$

Ako nas zanima vjerojatnost potrebnog vremenskog intervala između događaja, zapisujemo vjerojatnost:

$$P_1(T \leq t) = 1 - P_0(T \leq t) = 1 - e^{-at}, \quad (7.25)$$

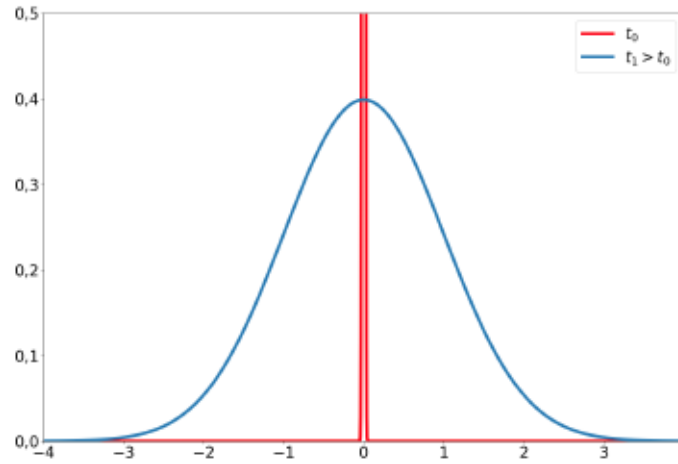
čime potvrđujemo poveznicu eksponencijalne i intervalne raspodjele, gdje eksponencijalna raspodjela opisuje vjerojatnost vremenskog trenutka događaja, a intervalna raspodjela opisuje vjerojatnost ostvarenja određenog broja događaja unutar intervala.

8 Difuzija

Jednadžba difuzije je parcijalna diferencijalna jednadžba koja opisuje promjenu gustoće na položaju $\vec{r} = (x, y, z)$ u vremenu t s konstantom difuzije D , te je zapisujemo na slijedeći način:

$$\frac{\partial c(\vec{r}, t)}{\partial t} = \nabla \cdot (D(\vec{r}, t) \nabla c(\vec{r}, t)). \quad (8.1)$$

Kako je jednadžba difuzije diferencijalna, ona opisuje Brownovo gibanje kod kojeg su koraci kontinuirani, tj. duljina njihovog koraka infinitezimalna. To znači da nakon prolaska nekog konačnog vremena od početka difuzije, jednadžba difuzije opisuje šetnju šetača koji je napravio beskonačno koraka, te se nalazi u limesu Centralnog graničnog teorema što možemo pokazati njezinim rješenjem. Razmatramo li česticu koja je savršeno lokalizirana u početnom vremenu



Slika 8.1: Ilustracija difuzije u trenutku $t_0 = 0$.

$t_0 = 0$ njenu početnu raspodjelu vjerojatnosti nalaženja opisujemo Diracovom delta funkcijom¹¹ (slika 8.1). Ukoliko konstanta difuzije ne ovisi o položaju, te se jednadžba 8.1 pojednostavi u linearnu diferencijalnu jednadžbu:

¹¹Jedna od (beskonačno) mogućih reprezentacija Diracove delta funkcije je limes gausijanske funkcije:

$$\int_{-\infty}^{\infty} f(x) \delta(x) dx = \lim_{a \rightarrow 0} \frac{1}{|a| \sqrt{\pi}} \int_{-\infty}^{\infty} f(x) e^{-\frac{x^2}{a^2}} dx = f(0),$$

gdje se skraćeno zapisuje:

$$\lim_{a \rightarrow 0} \frac{1}{|a| \sqrt{\pi}} e^{-\frac{x^2}{a^2}} = \delta(x),$$

te u ovom kontekstu Diracova delta funkcija predstavlja raspodjelu (s dimenzijom obrnuto proporcionalnom varijabli integracije).

$$\frac{\partial c(\vec{r}, t)}{\partial t} = D \nabla^2 c(\vec{r}, t), \quad (8.2)$$

te ukoliko promatramo difuziju u jednoj dimenziji se dodatno pojednostavni u:

$$\frac{\partial c(x, t)}{\partial t} = D \frac{\partial^2 c(x, t)}{\partial x^2} \quad (8.3)$$

Fourierova transformacija preko prostorne frekvencije $k = \frac{2\pi}{x}$ je:

$$C(k, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} c(x, t) e^{-ikx} dx, \quad (8.4)$$

i onda izraz 8.3 postaje:

$$\frac{\partial C(k, t)}{\partial t} = D \frac{\partial^2 C(k, t)}{\partial k^2}, \quad (8.5)$$

gdje desnu stranu izraza možemo raspisati:

$$\begin{aligned} D \frac{\partial^2 C(k, t)}{\partial k^2} &= D \frac{\partial^2}{\partial k^2} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} c(x, t) e^{-ikx} dx \right) \\ &= D \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d^2}{dk^2} \left(c(x, t) e^{-ikx} dx \right) \\ &= -k^2 D \frac{1}{2\pi} \int_{-\infty}^{\infty} c(x, t) e^{-ikx} dx \\ &= -k^2 DC(k, t), \end{aligned} \quad (8.6)$$

te jednačba difuzije postaje:

$$\frac{\partial C(k, t)}{\partial t} = -k^2 DC(k, t), \quad (8.7)$$

koje kada integriramo:

$$\int \frac{dC}{C(k, t)} = -k^2 D \int dt, \quad (8.8)$$

dobivamo rješenje:

$$C(k, t) = Ae^{-k^2 Dt}. \quad (8.9)$$

Kako nam je čestica lokalizirana u $t_0 = 0$, imamo početni uvjet $c(x, 0) = \delta(x)$, čija Fourierova transformacija glasi:

$$C(k, t = 0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(x) e^{-ikx} dx = \frac{1}{2\pi}, \quad (8.10)$$

gdje zaključujemo da je $A = \frac{1}{2\pi}$. Inverznom fourierovom transformacijom dobivamo:

$$c(x, t) = \mathcal{F}^{-1}(C(k, t)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-k^2 Dt} e^{ikx} dk. \quad (8.11)$$

Sređivanjem kvadrata u eksponentu:

$$\begin{aligned} c(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-k^2 Dt + ikx} dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(Dtk^2 - ikx)} dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(Dtk^2 - ikx - \frac{x^2}{4Dt} + \frac{x^2}{4Dt})} dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(\sqrt{Dt}k - \frac{ix}{2\sqrt{Dt}})^2} \cdot e^{-\frac{x^2}{4Dt}} dk \\ &= \frac{1}{2\pi} e^{-\frac{x^2}{4Dt}} \int_{-\infty}^{\infty} e^{-(\sqrt{Dt}k - \frac{ix}{2\sqrt{Dt}})^2} dk. \end{aligned} \quad (8.12)$$

Integral iz 8.12 rješavamo uvođenjem supstitucije $u = \sqrt{Dt}k - \frac{ix}{2\sqrt{Dt}}$, gdje je $du = \sqrt{Dt}dk$, gdje integral postaje gausijan:

$$\int_{-\infty}^{\infty} e^{-(\sqrt{Dt}k - \frac{ix}{2\sqrt{Dt}})^2} dk = \frac{1}{\sqrt{Dt}} \int_{-\infty}^{\infty} e^{-u^2} du, \quad (8.13)$$

gdje je rješenje integrala jednako $\sqrt{\pi}$ te rješenje jednadžbe difuzije glasi:

$$c(x, t) = \frac{1}{2\pi} e^{-\frac{x^2}{4Dt}} \frac{\sqrt{\pi}}{\sqrt{Dt}} = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}, \quad (8.14)$$

koja je kako vidimo gausijanska funkcija što je konzistentno s opažanjem kako smo rješavanjem jednadžbe difuzije odmah u režimu Centralnog graničnog teorema.

9 Dodatak

9.1 Rješenje integrala 7.14 preko teorema o reziduumima

U ovom poglavlju rješavanjem integrala

$$P_k(t) = \frac{a^k}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{(a + i\omega)^{k+1}} d\omega,$$

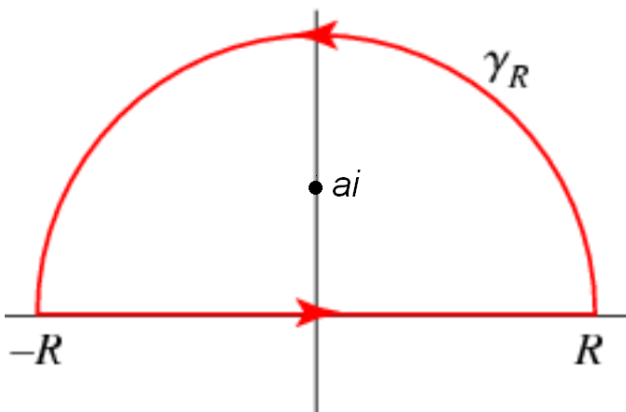
upotrebom alata kompleksne analize opravdamo uvođenje step funkcije $\Theta(t)$. Po teoremu o reziduumima:

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{z_0} \text{Res}(f(z), z_0), \quad (9.1)$$

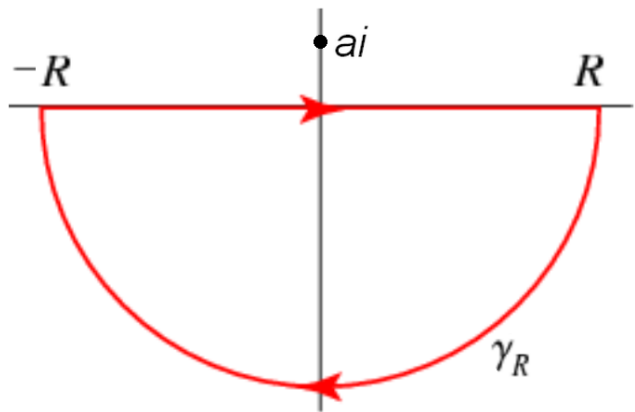
integral kompleksne funkcije $f(z)$ po pozitivno orijentiranoj zatvorenoj krivulji γ jednak sumi reziduuma funkcije $f(z)$ pomnožen s $2\pi i$. U tu svrhu prelazimo na integraciju kompleksne varijable z :

$$f(\omega) = \frac{e^{i\omega t}}{(a + i\omega)^{k+1}} \implies f(z) = \frac{e^{izt}}{(a + iz)^{k+1}}, \quad (9.2)$$

gdje je singularitet, odnosno pol $z_0 = ai$ reda $k + 1$.



Slika 9.1: Kontura integracije γ za $t > 0$ [15].



Slika 9.2: Kontura integracije γ za $t < 0$ [15].

Krivulju γ možemo rastaviti na "ravni" dio i kružni luk γ_R , koji se u slučaju $t > 0$ nalazi na gornjoj polu-ravnini (slika 9.1) i na donjoj polu-ravnini u u slučaju $t < 0$ (slika 9.2, te integral možemo zapisati:

$$\oint_{\gamma} f(z) dz = \int_{-R}^R f(z) dz + \int_{\gamma_R} f(z) dz. \quad (9.3)$$

Kako su granice integracije polaznog integrala ∞ i $-\infty$, po Jordanovoj lemi integracija po kružnom luku γ_R je za $t > 0$:

$$\lim_{R \rightarrow +\infty} \int_{\gamma_R} f(z) dz = \lim_{R \rightarrow +\infty} \int_0^\pi \frac{iR}{(a + iRe^{i\theta})^{k+1}} e^{iRe^{i\theta}} e^{i\theta} d\theta = 0, \quad (9.4)$$

te za $t < 0$:

$$\lim_{R \rightarrow +\infty} \int_{\gamma_R} f(z) dz = \lim_{R \rightarrow +\infty} \int_0^{-\pi} \frac{iR}{(a + iRe^{i\theta})^{k+1}} e^{-iRe^{i\theta}|t|} e^{i\theta} d\theta = 0, \quad (9.5)$$

koje reducira integral na:

$$\oint_{\gamma} f(z) dz = \int_{-R}^R f(z) dz = \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{(a + i\omega)^{k+1}} d\omega. \quad (9.6)$$

Kako se jedini pol nalazi na gornjoj polu-ravnini, u slučaju $t < 0$ kontura γ ne obuhvaća nijedan singularitet te je integral jednak nuli:

$$\oint_{\gamma} f(z) dz = \begin{cases} 2\pi i \sum_{z_0} \text{Res}(f(z), z_0), & \text{za } t > 0, \\ 0, & \text{za } t < 0, \end{cases} \quad (9.7)$$

što je jednako množenju step funkcijom $\Theta(t)$:

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{z_0} \text{Res}(f(z), z_0) \cdot \Theta(t). \quad (9.8)$$

Za određivanje reziduuma koristimo Rodriguesovu formulu za pol z_0 reda m :

$$\text{Res}(f(z), z_0) = \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} ((z - z_0)^m f(z)) \Big|_{z=z_0}, \quad (9.9)$$

kojim dobivamo:

$$\begin{aligned} \text{Res}(f(z), z_0) &= \frac{1}{k!} \frac{d^k}{dz^k} \left(\frac{(z - ai)^{k+1}}{(a + iz)^{k+1}} e^{izt} \right) \Big|_{z=ai} = \frac{1}{k!} \frac{d^k}{dz^k} \left(\frac{1}{i^{k+1}} e^{izt} \right) \Big|_{z=ai} = \\ &= \frac{1}{k!} \left(\frac{1}{i^{k+1}} i^k t^k e^{izt} \right) \Big|_{z=ai} = \frac{-it^k e^{-at}}{k!}. \end{aligned} \quad (9.10)$$

Uvrštavanjem dobivamo konačni rezultat:

$$P_k(t) = \frac{a^k}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{(a + i\omega)^{k+1}} d\omega = \frac{a^k}{2\pi} 2\pi i \frac{(-it^k e^{-at})}{k!} \Theta(t) = \frac{a^k t^k e^{-at}}{k!} \Theta(t). \quad (9.11)$$

Literatura

- [1] <https://www.anbg.gov.au/botanical-history/microscope-robert-brown.html>
- [2] <https://microcosmos.foldscope.com/?p=124663>
- [3] <https://scottbembenek.com/tag/brownian-motion/>
- [4] <https://s3-eu-west-1.amazonaws.com/blog.omniconvert.com-media/blog/wp-content/uploads/2019/10/21150245/sample-size-definition.png>
- [5] <https://math.stackexchange.com/questions/1204396/why-is-the-sum-of-the-rolls-of-two-dices-a-binomial-distribution-what-is-defined>
- [6] <https://income-inequality.info/images/3.png>
- [7] <http://yaroslavvb.com/papers/finch-mathematical-5.pdf>
- [8] https://www.researchgate.net/figure/Continuous-uniform-distribution-Source-14_fig8_332236648
- [9] https://upload.wikimedia.org/wikipedia/commons/thumb/d/db/Arcsin_density.svg/1920px-Arcsin_density.svg.png
- [10] https://www.researchgate.net/publication/319932351_Optimal_Capacitor_Placement_in_Wind_Farms_by_Considering_Harmonics_Using_Discrete_Lightning_Search_Algorithm/figures?lo=1
- [11] https://commons.wikimedia.org/wiki/File:Exponential_distribution_pdf.png
- [12] https://cdn.shopify.com/s/files/1/1026/4509/files/bde19dcca58d83e88aeb45491a ECB908_large.jpg?11184986474562496780
- [13] <https://www.vecteezy.com/png/1198025-grid-globe>
- [14] Gurland, John; Tripathi, Ram C. (1971), "A Simple Approximation for Unbiased Estimation of the Standard Deviation", *The American Statistician*, 25 (4): 30–32
- [15] https://www-thphys.physics.ox.ac.uk/people/FrancescoHautmann/ComplexVariable/s1_12_sl11p.pdf
- [16] <http://physics.ucsc.edu/~drip/5D/brown/brown.pdf>