Aproksimacija očekivanja potencijskih sredina velikih uzoraka

Rebić Taučer, Borna

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 911

APROKSIMACIJA OČEKIVANJA POTENCIJSKIH SREDINA VELIKIH UZORAKA

Borna Rebić Taučer

Zagreb, lipanj 2023.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 911

APROKSIMACIJA OČEKIVANJA POTENCIJSKIH SREDINA VELIKIH UZORAKA

Borna Rebić Taučer

Zagreb, lipanj 2023.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Zagreb, 10. ožujka 2023.

ZAVRŠNI ZADATAK br. 911

Pristupnik: **Borna Rebić Taučer (0036526910)**

Studij: Elektrotehnika i informacijska tehnologija i Računarstvo

Modul: Računarstvo

Mentor: izv. prof. dr. sc. Tomislav Burić

Zadatak: **Aproksimacija očekivanja potencijskih sredina velikih uzoraka**

Opis zadatka:

Potencijske sredine su najvažnija klasa sredina u matematici i imaju široku primjenu u raznim područjima. Najčešće korišteni primjeri takvih sredina su aritmetička, geometrijska, harmonijska i kvadratna sredina, a ovaj rad se bavi njihovim određivanjem u statistici. Za uzorke s velikim podacima takve se sredine općenito teško računaju pa je cilj rada aproksimirati očekivanu vrijednost potencijskih sredina za takve velike uzorke. Aproksimacija očekivanja se temelji na asimptotskim razvojima potencijskih sredina pa je prvo potrebno implementirati algoritme kojima se određuju koeficijenti u takvim razvojima. Zatim treba usporediti dobiveno asimptotsko očekivanje s očekivanjem potencijskih sredina slučajnih uzoraka dobivenih Monte Carlo simulacijama. Rezultate treba testirati i analizirati na uzorcima iz najčešće korištenih statističkih razdioba: uniformne, normalne i hi-kvadrat razdiobe.

Rok za predaju rada: 9. lipnja 2023.

Htio bih izraziti par riječi zahvale svima koji su mi svojom bezuvjetnom podrškom omogućili da završim ovaj studij. Zahvaljujem svim profesorima i zaposlenicima Fakulteta elektrotehnike i računarstva, na posevećenom vremenu i prenesenom znanju, a posebno mentoru izv. prof. dr. sc. Tomislavu Buriću na svoj pomoći, sugestijama i stručnim savjetima da bi ovaj rad bio uspješno završen. Posebna zahvala ide mojoj obitelji i prijateljima, koji su mi bili neizreciva potpora tijekom cijelog studija i bez čije podrške i ljubavi ni jedan moj uspjeh, pa tako ni ovaj, ne bi bio moguć ni potpun.

SADRŽAJ

1. Uvod

Teorija vjerojatnosti jedno je od najvažnijih područja matematike koje se najviše primjenjuje u područjima matematičke statistike, kao i u raznim područjima fizike, biologije, ekonomije i drugdje. Područje teorije vjerojatnosti bavi se proučavanjem slučajnih varijabli, njihovih distribucija i karakteristika kao što su očekivanje i disperzija. Matematička statistika bavi se proučavanjem i obradom manjeg dijela populacije koju nazivamo uzorak. Podatci prikupljeni iz odabranog uzorka sortiraju se, analiziraju i interpretiraju metodama teorije vjerojatnosti.

Kako bismo bolje razumjeli očekivanje potencijalnih velikih uzoraka i njihovu aproksimaciju, prvo je važno upoznati se s osnovama teorije vjerojatnosti i matematičke statistike. U drugom poglavlju objašnjena je teorijska pozadina potrebna za razumijevanje osnovnih pojmova kao što su slučajna varijabla, očekivanje i uzorak, dok je u trećem poglavlju detaljnije opisan postupak aproksimacije očekivanja korištenjem asimptotskog razvoja.

U četvrtom poglavlju predstavit ćemo rezultate aproksimacije za očekivanja koja smo definirali u prethodnom poglavlju. Koristeći razvijenu aplikaciju, aproksimirat ćemo očekivanja za uzorke različitih duljina na slučajnim varijablama s jednolikom i normalnom distribucijom.

2. Teorijska podloga

U ovom poglavlju upoznat ćemo se s osnovnim pojmovima iz područja vjerojatnosti i statistike. Definirat ćemo slučajnu varijablu i bitne pojmove vezane uz njih kao što su očekivanje, ishodišni i centralni momenti i sl.

2.1. Slučajne varijable

Prilikom realizacije nekog pokusa ostvaruje se elementaran događaj $\omega \in \Omega$. Cilj pokusa je upravo mjerenje neke numeričke veličine koja ovisi o realizaciji tog elementarnog događaja. Jednostavan primjer je model bacanja novčića. Elementarni događaji su pismo i glava koje ćemo redom označiti s 0 i 1. Time smo definirali preslikavanje iz skupa Ω u skup S = $\{0, 1\}$. Upravo takvo preslikavanje nazivamo slučajna varijabla.

Razlikujemo dvije vrste slučajnih varijabli, diskretne i kontinuriane slučajne varijable. Gore spomenuti primjer predstavlja diskretnu slučajnu varijablu. Glavna razlika je što diskretne varijable poprimaju svoje vrijednosti unutar diskretnog skupa, dok kontinuirane mogu poprimiti vrijednost kao bilo koji realni broj unutar nekog intervala. Cilj ovog poglavlja je definirati detaljnije obje slučajne varijable i uočiti razlike i sličnosti.

Preslikavanje $X : \Omega \to S$ je diskretna slučajna varijabla ako je za svaki $x_k \in S$ skup $A_k := \{\omega \in \Omega : X(\omega) = x_k\}$ događaj. Označimo

$$p_k := P(A_k) = P(X = x_k). \tag{2.1}$$

Za ove brojeve vrijedi:
1) Vjerojatnost svakog događaja mora biti veća od nule, $p_k > 0$,
2) Zbroj svih vjerojatnosti mora iznositi 100% odnosno 1, $\sum p_k = 1$.

Zakon razdiobe slučajne varijable X sastoji se od područja vrijednosti koje ona poprima i odgovarajućih vjerojatnosti

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}. \tag{2.2}$$

Funkcija razdiobe kontinuirane slučajne varijable X je funkcija F: R -> [0,1] definirana formulom

$$F(x) := P(\{X < x\}). \tag{2.3}$$

Poznavanjem funkcije razdiobe, moguće je izračunati vjerojatnost pojave slučajne varijable na intervalu [a, b], a < b

$$\begin{aligned} F(b) &= P(\{X < b\}) \\ &= P(\{X < a\} \cup \{a \leq X < b\}) \\ &= P(\{X < a\}) + P(\{a \leq X < b\}) \\ &= F(a) + P(\{a \leq X < b\}) \end{aligned} \tag{2.4}$$

$$P(\{a \leq X < b\}) = F(b) - F(a).$$

Kod kontinuiranih slučajnih varijabli uvodimo novi pojam koji nazivamo funkcija gustoće. Za slučajnu varijablu X kažemo da je kontinuirana ukoliko postoji nenegativna funkcija gustoće $f : \mathbb{R} \to \mathbb{R}$

$$F(x) = \int_{-\infty}^{x} f(t)dt. \tag{2.5}$$

U točkama neprekinutosti od funkcije gustoće vrijedi

$$f(x) = \frac{dF(x)}{dx}. \tag{2.6}$$

Nezavisnost slučajnih varijabli je primjenjiva i na diskretne i na kontinuirane slučajne varijable. Slučajne varijable X i Y su nezavisne ukoliko za sve intervale A i B, gdje $A, B \in \mathbb{R}$ vrijedi

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B). \tag{2.7}$$

U nekom stohastičkom pokusu najčešće promatramo više slučajnih varijabli. Svaka od tih varijabli ima svoju razdiobu koja često, ali ne i nužno ovisi o drugim slučajnim varijablama. Takav sustav višedimenzionalnih varijabli nazivamo slučajjni vektori.

Slučajni vektor jest uređena n-torka slučajnih varijabli $X = (X_1, X_2, ..., X_n) : \Omega \to \mathbb{R}^n$. Takav vektor nazivamo n-dimenzonalan vektor. Funkciju razdiobe slučajnog vektora definiramo kao

$$F(x_1, \ldots, x_n) := P(X_1 < x_1, \ldots, X_n < x_n), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n, \tag{2.8}$$

a gustoću razdiobe slučajnog vektora kao

$$f(x_1, \ldots, x_n) = \frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n}. \tag{2.9}$$

2.2. Ishodišni i centralni moment

Slučajne varijable se opisuju i interesantne su nam zbog svojih numeričkih karakteristika. Neka je n prirodni broj. Ishodišni moment reda n definiramo formulom

$$E(X^n) = \sum_k x_k^n \cdot p_k. \tag{2.10}$$

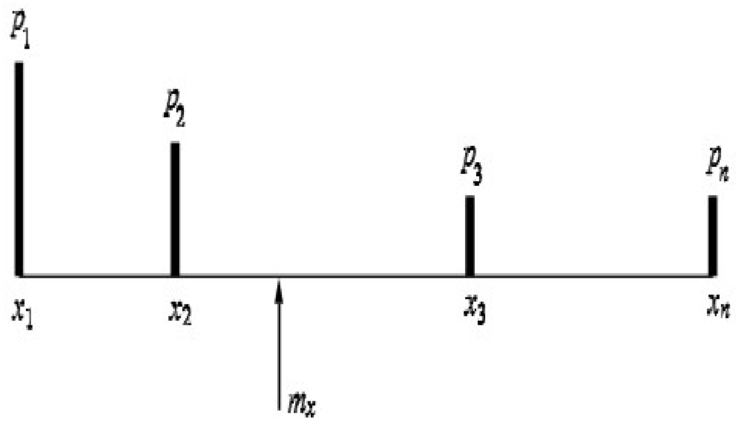
Centralni moment μ_n reda n za očekivanje m_X od X, definiramo

$$\mu_n := E[(X - m_X)^n] = \sum_k (x_k - m_X)^n \cdot p_k. \tag{2.11}$$

Očekivanje ili centralni moment reda 1 je najvažnija karakteristika slučajnih varijabli. Očekivanje diskretne slučajne varijable je definirano kao suma svih vrijednosti pomnoženih s vjerojatnosti pojavljivanja te vrijednosti

$$E(X) := \sum_k x_k p_k. \tag{2.12}$$

Očekivanje zapravo predstavlja težište slučajne varijable.



Slika 1. Očekivanje m_x diskretne slučajne varijable je težinska sredina njezinih realizacija.

Dok je za kontinuiranu slučajnu varijablu definirano kao

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx. \tag{2.13}$$

Očekivanje ima svojstvo linearnosti tj. za sve brojeve s i t, gdje $s, t \in \mathbb{R}$ vrijedi

$$E(sX + tY) = sE(X) + tE(Y). \tag{2.14}$$

Ukoliko su X i Y nezavisne, vrijedi

$$E(XY) = E(X)E(Y). \tag{2.15}$$

Disperzija je još jedna interesantna numerička karakteristika koja opisuje rasipanje slučajne varijable. Definiramo ju kao centralni moment drugog reda (n=2)

$$D(X) = E[(X-\mu)^2] = E(X^2) - \mu^2, \tag{2.16}$$

tj. za kontinuiranu slučajnu varijablu vrijedi

$$D(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2. \tag{2.17}$$

Svojstva disperzije su

$$\begin{aligned} D(X+Y) &= D(X) + D(Y), \\ D(sX) &= s^2 D(X). \end{aligned} \tag{2.18}$$

Veličina koja prikazuje odstupanje varijable X nazivamo standardna devijacija σ_X te vrijedi $\sigma_X := \sqrt{D(X)}$.

2.3. Primjeri razdioba slučajnih varijabli

U ovom poglavlju upoznat ćemo se s osnovnim pojmovima iz područja vjerojatnosti i statistike. Definirat ćemo slučajnu varijablu i bitne pojmove vezane uz njih kao što su očekivanje, ishodišni i centralni momenti i sl.

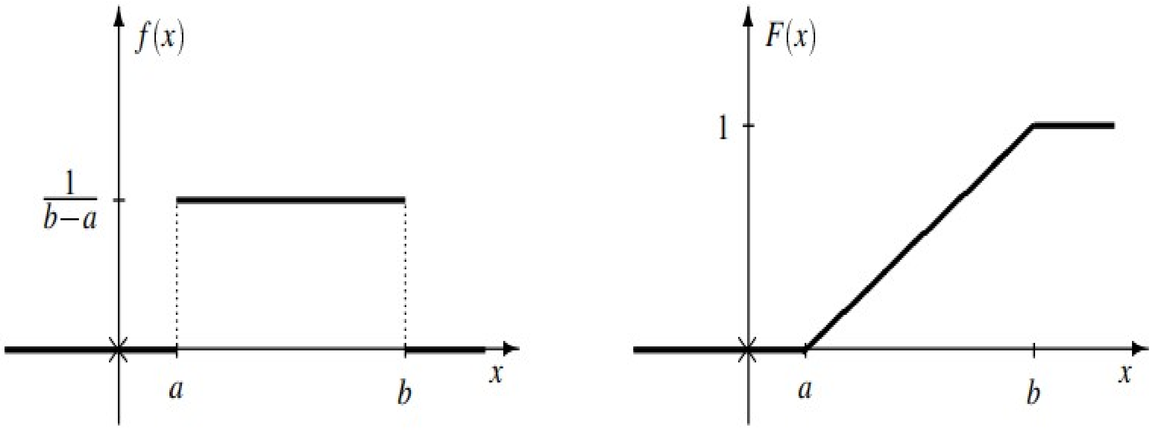
2.3.1. Jednolika razdioba

Slučajna varijabla X koja nasumično uzima broj unutar intervala $[a, b]$ ima jednoliku razdiobu na intervalu $[a, b]$. Kažemo da je varijabla *X* jednoliko ili uniformno distribuirana na intervalu $[a, b]$, ukoliko je zadana funkcijom razdiobe

$$F(x) = \frac{x-a}{b-a}, \quad a \leq x \leq b, \tag{2.19}$$

odnosno funkcijom gustoće

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b. \tag{2.20}$$



Slika 1. Grafovi funkcije gustoće i funkcije razdiobe jednolike razdiobe,na intervalu [a,b]

Sada kad znamo kako glasi funkcija gustoće jednolike razdiobe i kako se računa očekivanje neprekinutih slučajnih varijabli, moguće je izraziti očekivanje slučajne varijable *X* s jednolikom razdiobom na intervalu $[a, b]$

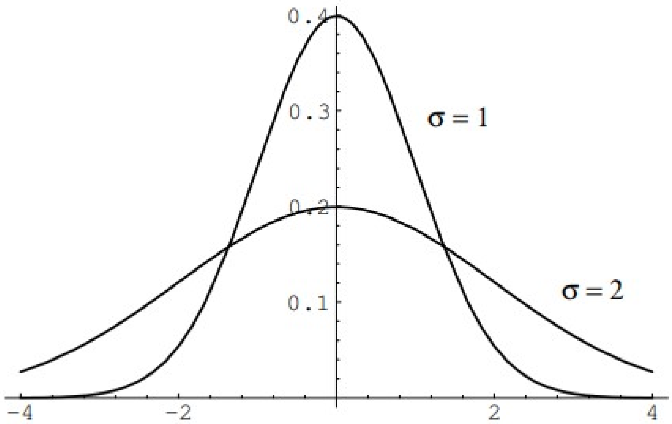
$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x)\, dx \\ &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{b+a}{2}. \end{aligned} \tag{2.21}$$

Na sličan način možemo odrediti disperziju te slučajne varijable

$$\begin{aligned} D(X) &= E[(X - m_X)^2] \\ &= E(X^2) - E(X)^2 \\ &= \int_a^b x^2 f(x) dx - E(X)^2 \\ &= \frac{(b-a)^2}{12}. \end{aligned} \tag{2.22}$$

2.3.2. Normalna razdioba

Normalna razdioba je jedna od najvažnijih razdioba upravo zbog svoje učestalosti. Slučajnu varijablu X s normalnom razdiobom označavamo kao $X \sim \mathcal{N}(a, \sigma^2)$. Graf ove funkcije je zvonolika krivulja koju još nazivamo i Gaussova krivulja.



Slika 1. Graf gustoće normalne razdiobe, za vrijednost parametra $a = 0$

Slučajna varijabla X ima normalnu razdiobu s parametrima $a \in \mathbb{R}$ i $\sigma^2 > 0$ ako je X neprekinuta slučajna varijabla s razdiobom

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \tag{2.23}$$

Izaberemo li parametre $a = 0$ i $\sigma = 1$, dobivamo slučajnu varijablu $\mathcal{N}(0, 1)$ koja se zove jedinična normalna razdioba. Njezinu gustoću označavamo sa

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}. \tag{2.24}$$

Znamo da je pomoću funkcije gustoće moguće odrediti funkciju razdiobe

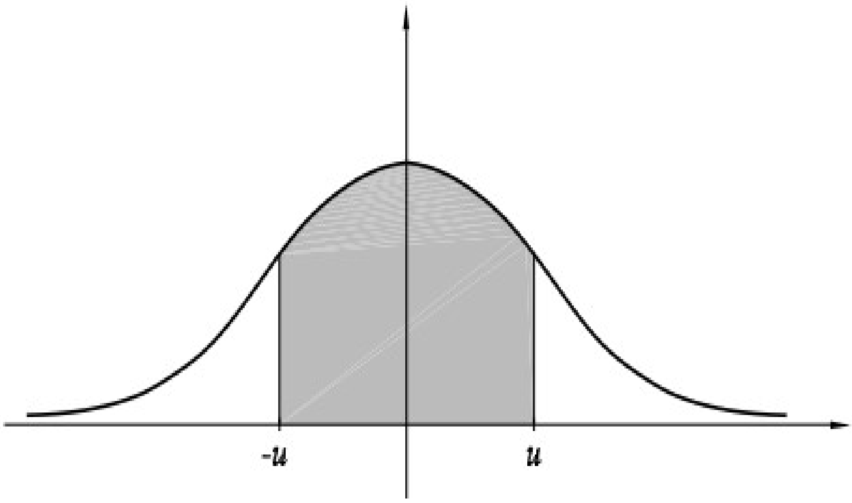
$$\Phi(u) = \int_{-\infty}^{u} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}t^2} dt. \tag{2.25}$$

Vidimo da je funkcija ϕ parna pa možemo napisati

$$\begin{aligned} \Phi(u) &= \int_{-\infty}^{u} \phi(t)dt \\ &= \int_{-\infty}^{0} \phi(t)dt + \int_{0}^{u} \phi(t)dt \\ &= \frac{1}{2} + \frac{1}{2} \int_{-u}^{u} \phi(t)dt \\ &= \frac{1}{2}\big[1 + \Phi^*(u)\big]. \end{aligned} \tag{2.26}$$

Funkciju razdiobe Φ možemo zapisati pomoću funkcije

$$\Phi^*(u) = \frac{1}{\sqrt{2\pi}} \int_{-u}^{u} e^{-\frac{1}{2}t^2} dt. \tag{2.27}$$



Slika 2. Vrijednost funkcije Φ^* jednaka je površini ispod grafa, unutar intervala [-u,u]

Očekivanje i disperziju normalne razdiobe moguće je izračunati koristeći karakterističnu funkciju. Za $X \sim \mathcal{N}(a, \sigma^2)$ vrijedi

$$\begin{gathered}
\vartheta(t) = e^{ita - \frac{1}{2}\sigma^2 t^2} \\
\vartheta'(t) = (ia - \sigma^2 t)e^{ita - \frac{1}{2}\sigma^2 t^2} \\
\vartheta'(0) = ia \\
\vartheta''(t) = \left[-\sigma^2 + (ia - \sigma^2 t)^2\right] e^{ita - \frac{1}{2}\sigma^2 t^2} \\
\vartheta''(0) = -\sigma^2 - a^2.
\end{gathered}$$

Sada je moguće izraziti očekivanje i disperziju

$$\begin{gathered}
E(X) = -i\vartheta'(0) = a \\
D(X) = -\vartheta''(0) + \vartheta'(0)^2 = \sigma^2.
\end{gathered}$$

2.4. Osnovni elementi statistike

Statistika se bavi proučavanjem nekog masovnog skupa, kojeg nazivamo populacija. Podatak koji proučavamo se naziva obilježje te se u praksi često proučava više obilježja. Populaciju čini skup Ω, a obilježje opisujemo pomoću slučajne varijable X. Glavni problem statistike je upravo određivanje numeričkih karakteristika i razdiobe slučajne varijable X.

Vrlo često je populacija prevelika i nepotrebno je statistički obraditi čitavu populaciju. U tom slučaju uzimamo manji skup te populacije koji nazivamo uzorak. Bitno je dobro odrediti uzorak kako bi on mogao predstavljati čitavu populaciju.

Neka je X slučajna varijabla s razdiobom F. Slučajne varijable $X_1, \ldots, X_n$ nazivamo nezavisnim kopijama slučajne varijable X ukoliko vrijedi da su:

1. međusobno nezavisne,

2. imaju razdiobu identičnu razdiobi slučajne varijable X.

Takvu n-torku slučajnih varijabli nazivamo uzorak. Ukoliko je x_1 realizacija slučajne varijable X_1, tj. x_n realizacija slučajne varijable X_n, tada $(x_1, \ldots, x_n)$ nazivamo realizacija ili vrijednost uzorka $(X_1, \ldots, X_n)$. Broj n predstavlja veličinu tj. dimenziju uzorka.
Za nepoznato očekivanje a populacije X odabiremo statistiku

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{2.28}$$

koju nazivamo sredina uzorka.

Označimo nepoznato očekivanje

$$E(\bar{X}) = a. \tag{2.29}$$

Varijabla X je slučajna te prema svojstvima očekivanja vrijedi

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) \\ &= \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots + E(X_n)\right) \\ &= a. \end{aligned}$$

Isto tako označimo disperziju

$$D(\bar{X}) = \sigma^2. \tag{2.30}$$

Znamo da su varijable $X_1, X_2, ..., X_n$ nezavisne pa vrijedi

$$\begin{aligned} D(\bar{X}) &= D\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) \\ &= \frac{\sigma^2}{n}. \end{aligned} \tag{2.31}$$

Gdje je σ^2 varijanca tj. disperzija populacije.

Vidimo da je disperzija statistike $\bar{X}$ obrnuto proporcionalna veličini uzorka. Iz čega možemo zaključiti da što je veći uzorak to će više vrijednosti biti koncentrirane oko srednje vrijednosti $E(\bar{X}) = a$. Što znači da će $\bar{X}$ biti dobra procjena za a.

Nepoznato očekivanje a populacije X procjenjujemo pomoću sredine uzorka

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{2.32}$$

Za tu slučajnu varijablu vrijedi

$$E(\bar{X}) = a, \tag{2.33}$$

$$D(\bar{X}) = \frac{\sigma^2}{n}. \tag{2.34}$$

3. Asimptotski razvoj

3.1. Općenito o asimtptoskim razvojima

Cilj ovog poglavlja je upoznati se s asimptotskim redovima, odnosno asimptotskim razvojem. Za početak je potrebno definirati O-notaciju. Neka su x i x_0 točke u području R kompleksne ravnine. Promatramo dvije funkcije $f(x)$ i $g(x)$, $g(x) \neq 0$.

Ako postoji broj A, takav da je $|f(x)/g(x)| \leq A$ za sve x iz R, tada pišemo

$$f(x) = O(g(x)), \quad x \to x_0 \quad u \quad \mathbb{R}. \tag{3.1}$$

Ako vrijedi $f(x)/g(x) \to 0$ kada $x \to x_0$, tada pišemo

$$f(x) = o(g(x)), \quad x \to x_0 \quad u \quad \mathbb{R}. \tag{3.2}$$

DEFINICIJA 1. *Kažemo da je red*

$$\textstyle\sum_{k=0}^{\infty} a_k \rho_k(x)$$

asimptotski razvoj funkcije f(x) kada x $\to x_0$ u području $\mathbb{R}$, ako vrijedi

$$f(x) = \sum_{k=0}^{n} a_k \rho_k(x) + o(\rho_n(x)), \quad x \to x_0 \quad u \quad \mathbb{R}. \tag{3.3}$$

Često pišemo

$$f(x) = \sum_{k=0}^{n} a_k \phi_k(x), \quad x \to x_0. \tag{3.4}$$

Mnogo funkcija može imati isti asimptotski razvoj i za te funkcije kažemo da su ekvivalentne. Funkcija $f(x)$ može imati drugačiji razvoj na nekom drugom području $\mathbb{R}$ i za neki drugi x_0.

DEFINICIJA 2. *Asimptotski razvoj*

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \frac{a_3}{x^3} + \ldots, \quad x \to \infty \tag{3.5}$$

zovemo asimptotski red potencija.

Prisjetimo se što znači kada konvergentan red potencija reprezentira funkciju. Ako vrijedi

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad |x| < r,$$
$$S_n(x) = \sum_{k=0}^{n} a_k x^k,$$

tada red potencija reprezentira $f(x)$ za svaki fiksni x, $|x| < r$, u smislu da je

$$\lim_{n\to\infty}(f(x) - S_n(x)) = 0.$$

Za dovoljno velik n i fiksni x, možemo aproksimirati funkciju $f(x)$ koristeći sumu $S_n(x)$.

Usporedimo to s asimptotskim redom potencija. Slijedi da za asimptotski red

$$f(x) = \sum_{k=0}^{\infty} a_k x^{-k}, \quad |x| < r,$$
$$S_n(x) = \sum_{k=0}^{n} a_k x^{-k},$$

i fiksni n vrijedi

$$\lim_{x\to\infty} x^n(f(x) - S_n(x)) = 0.$$

Odnosno za dovoljno velik x u R možemo aproksimirati funkciju $f(x)$ koristeći sumu $S_n(x)$.

Za asimptotski razvoj potencija koristit ćemo zapis

$$f(x) \sim \sum_{k=0}^{\infty} a_k x^{-k}. \tag{3.6}$$

Asimptotski red potencija ćemo kraće zvati asimptotski red, odnosno asimptotski razvoj potencija ćemo kraće zvati asimptotki razvoj.

3.2. Asimtptoski razvoji potencijskih sredina

Uzmimo sad slučajan vektor $\vec{e} = (1, 1, ..., 1)$ duljine n. Tada n-torku $(x_1, x_2, ..., x_n)$ možemo zapisati kao

$$(x_1, x_2, ..., x_n) = (x + a_1, x + a_2, ..., x + a_n) = x\vec{e}+\vec{a},$$

gdje x predstavlja srednju vrijednost n-torke $(x_1, x_2, ..., x_n)$. Definirajmo

$$m_k = m_k(\vec{a}, 0) = \frac{a_1^k + a_2^k + \ldots + a_n^k}{n} = M_k(\vec{a})^k, \quad m_0 := 1. \tag{3.7}$$

TEOREM 1. Opća formula za potencijsko očekivanje ima sljedeći asimptotski razvoj

$$M_r(x\vec{e} + \vec{a}) = x \cdot \sum_{k=0}^{\infty} c_k(r, \vec{a}) \cdot x^{-k}, \tag{3.8}$$

gdje je c

$$c_k(r, \vec{a}) = \frac{1}{k} \sum_{j=1}^{k} \left[j(1 + \frac{1}{r}) - k \right] \binom{r}{j} m_j c_{k-j}(r, \vec{a}), \quad k \in \mathbb{N}. \tag{3.9}$$

Gdje je prvih par koeficijenata

$$\begin{aligned}
c_0(r, \vec{a}) &= 1, \\
c_1(r, \vec{a}) &= m_1, \\
c_2(r, \vec{a}) &= -\frac{1}{2}(r-1)(m_1^2 - m_2), \\
c_3(r, \vec{a}) &= \frac{1}{6}(r-1)\left((2r-1)m_1^3 - 3(r-1)m_1m_2 + (r-2)m_3\right), \\
c_4(r, \vec{a}) &= -\frac{1}{24}(r-1)\Big((3r-1)(2r-1)m_1^4 - 6(r-1)(2r-1)m_1^2m_2 + 3(r-1)^2m_2^2 \\
&\quad + 4(r-2)(r-1)m_1m_3 - (r-3)(r-2)m_4\Big).
\end{aligned}$$

Za izračun točnih vrijednosti koeficijenata potrebno je supstituirati r s vrijednosti željenog reda.

Za sve X vrijedi

$$A(x\vec{e} + \vec{a}) = x + A(\vec{a}).$$

KOROLAR 1. Geometrijska srednja vrijednost ima sljedeći asimptotski razvoj

$$G(x\vec{e} + \vec{a}) = x \cdot \sum_{k=0}^{\infty} c_k x^{-k}, \tag{3.10}$$

gdje je $c_0 = 1$ i c_k

$$c_k = \frac{1}{k} \sum_{j=1}^{k} (-1)^{j-1} m_j c_{k-j}, \quad k \geq 1. \tag{3.11}$$

Srednja geometrijska vrijednost s prvih par koeficijenata glasi

$$\begin{aligned} G(xe+a) =& x + m_1 + \frac{1}{2}(m_1^2 - m_2)x^{-1} \\ &+ \frac{1}{6}(m_1^3 - 3m_1m_2 + 2m_3)x^{-2} \\ &+ \frac{1}{24}(m_1^4 - 6m_1^2m_2 + 3m_2^2 + 8m_1m_3 - 6m_4)x^{-3} \\ &+ \dots \end{aligned} \tag{3.12}$$

KOROLAR 2. Harmonijska srednja vrijednost ima sljedeći asimptotski razvoj

$$H(x\vec{e} + \vec{a}) = x\sum_{k=0}^{\infty} c_k x^{-k+1}, \tag{3.13}$$

gdje je $c_0 = 1$ i c_k

$$c_k = \sum_{j=1}^{k} (-1)^{j-1} m_j c_{k-j}, \quad k \geq 1. \tag{3.14}$$

Srednja harmonijska vrijednost s prvih par koeficijenata glasi

$$\begin{aligned} H(x\vec{e} + \vec{a}) =& x + m_1 + (m_1^2 - m_2)x^{-1} \\ &+ (m_1^3 - 2m_1m_2 + m_3)x^{-2} \\ &+ (m_1^4 - 3m_1^2m_2 + m_2^2 + 2m_1m_3 - m_4)x^{-3} \\ &+ \dots \end{aligned} \tag{3.15}$$

KOROLAR 3. Kvadratna srednja vrijednost ima sljedeći asimptotski razvoj

$$Q(x\vec{e} + \vec{a}) = x \cdot \sum_{k=0}^{\infty} c_k x^{-k}, \tag{3.16}$$

gdje je $c_0 = 1, c_1 = m_1$ i c_k

$$c_k = \left(\frac{3}{k} - 2\right) m_1 c_{k-1} + \left(\frac{3}{k} - 1\right) m_2 c_{k-2}, \quad k \geq 2. \tag{3.17}$$

Srednja kvadratna vrijednost s prvih par koeficijenata glasi

$$\begin{aligned} Q(x\vec{e} + \vec{a}) =& x + m_1 + \frac{1}{2}(-m_1^2 + m_2)x^{-1} \\ &+ \frac{1}{2}m_1(m_1^2 - m_2)x^{-2} \\ &+ \frac{1}{8}(-5m_1^4 + 6m_1^2m_2 - m_2^2)x^{-3} \\ &+ \dots \end{aligned} \tag{3.18}$$

TEOREM 2. Koeficijenti $c_k, k \in \mathbb{N}_0$ su homogeni polinomi u varijablama $(a_1, a_2, ..., a_n)$ i vrijedi

$$c_k(r, \vec{a}) = \sum_{\substack{\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0 \\ \alpha_1 + 2\alpha_2 + \cdots + k\alpha_k = k}} q_{\alpha_1, \ldots, \alpha_k}(r) m_1(\vec{a})^{\alpha_1} \cdots m_k(\vec{a})^{\alpha_k}, \tag{3.19}$$

gdje

$$\sum_{\substack{\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0 \\ \alpha_1 + 2\alpha_2 + \cdots + k\alpha_k = k}} q_{\alpha_1, \ldots, \alpha_k}(r) = 0, \quad k \geq 2. \tag{3.20}$$

Neka su $X_1.X_2, ..., X_n$ nezavisne kopije slučajne varijable X koje zadovoljavaju sljedeće uvjete:

1. Poznata je funkcija razdiobe varijable X i ona je simetrična oko neke točke μ tj. oko svoje srednje vrijednosti
2. μ je dovoljno velik, tako da vrijedi $P(X < 0) \to 0$ kada $\mu \to \infty$
3. $X \in L^k, k \in \mathbb{N}$.

Definirajmo sada centriranu slučajnu varijablu

$$A_i := X_i - \mu \tag{3.21}$$

i slučajan uzorak

$$A = (A_1, A_2, ..., A_n). \tag{3.22}$$

Sada ćemo za aproksimaciju očekivanja $E[M_r(X)]$, primjeniti asimptotsko proširenje

$$E[M_r(\vec{X})] = E[M_r(x\vec{e} + \vec{a})] = x \cdot \sum_{k=0}^{\infty} E[c_k(r, \vec{A})]x^{-k}, \quad x \to \infty. \tag{3.23}$$

LEMA 1. Za svaki neparni k vrijedi $E[C_k(r, \vec{A})] = 0$.

Dokaz. Zbog Teorema 2, znamo da vrijedi

$$E[m_1^{\alpha_1} \ldots m_k^{\alpha_k}] = 0$$

tj.

$$E[A_1^{\beta_1} A_2^{\beta_2} \ldots A_n^{\beta_n}] = E[A_1^{\beta_1}] E[A_2^{\beta_2}] \ldots E[A_n^{\beta_n}] = 0.$$

LEMA 2. Vrijedi

$$
\begin{aligned}
E[m_1^2] &= \frac{\mu_2}{n}, \\
E[m_1^4] &= \frac{\mu_4}{n^3} + \frac{3(n-1)\mu_2^2}{n^3}, \\
E[m_1^2 m_2] &= \frac{\mu_4}{n^2} + \frac{(n-1)\mu_2^2}{n^2}, \\
E[m_2^2] &= \frac{\mu_4}{n} + \frac{(n-1)\mu_2^2}{n}, \\
E[m_1 m_3] &= \frac{\mu_4}{n}.
\end{aligned} \tag{3.24}
$$

Dokaz.

$$
E[m_1^2] = \frac{1}{n^2} E\left[\sum_{i=1}^{n} A_i\right]^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} E[A_i^2] + \sum_{i \neq j} E[A_i]E[A_j]\right) = \frac{\mu_2}{n}, \tag{3.25}
$$

$$
E[m_1^4] = \frac{1}{n^4}\left(\sum_{i=1}^{n} E[A_i^4] + 3\sum_{i \neq j} E[A_i^2]E[A_j^2]\right) = \frac{\mu_4}{n^3} + \frac{3(n-1)\mu_2^2}{n^3}. \tag{3.26}
$$

Koristeći ove dvije leme možemo procjeniti očekivanje potencijskih sredina na sljedeći način.

TEOREM 3. Očekivanje potencijske sredine ima sljedeće proširenje:

$$
E[Mr(\vec{X})] = \mu + \frac{d_2}{\mu} + \frac{d_4}{\mu^3} + O\left(\frac{1}{\mu^5}\right), \quad \mu \to \infty, \tag{3.27}
$$

gdje je μ srednja vrijednost i

$$
\begin{aligned}
d_2 &= \frac{r-1}{2} \cdot \frac{n-1}{n} \cdot \mu_2, \\
d_4 &= -\frac{r-1}{24} \cdot \frac{n-1}{n}\left[3\mu_2^2\left(\frac{(3r-1)(2r-1)}{n^2} - \frac{2(r-1)(2r-1)}{n} + (r-1)^2\right)\right. \\
&\qquad \left. -\mu_4\left(\frac{(3r-1)(2r-1)}{n^2} - \frac{(2r-1)(3r-5)}{n} + (r-2)(r-3)\right)\right].
\end{aligned}
$$

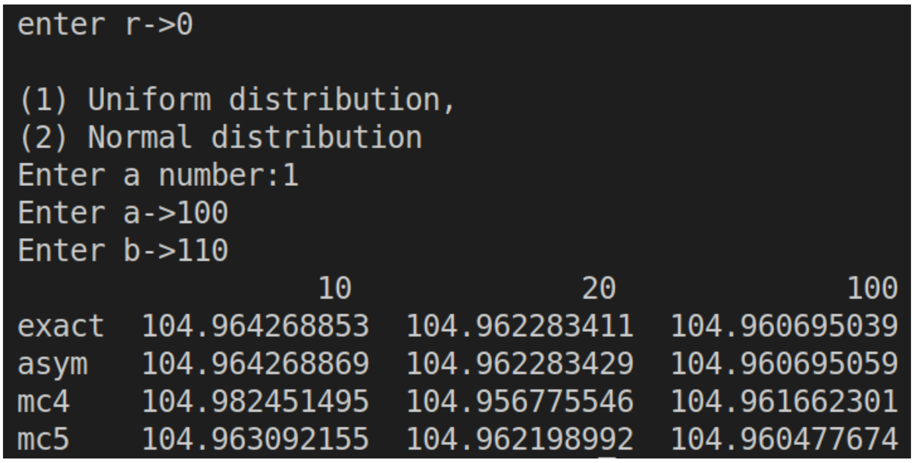
4. Implementacija i rezultati

U ovom poglavlju ćemo proći kroz implementaciju i rezultate aproksimacije za geometrijska, harmonijska i kvadratna očekivanja na uzorcima različitih duljina za slučajne varijable s jednolikom i normalnom razdiobom.

Za implementaciju zadatka razvili smo jednostavnu aplikaciju koristeći programski jezik Python, verziju 3.10.6. Za implementaciju uzoraka jednolike i normalne razdiobe korišten je paket *Numpy*, a za izračun egzaktne vrijednosti geometrijskog očekivanja normalne razdiobe bilo je potrebno implementirati gama funkciju i konfluentnu hipergeometrijsku funkciju za što je korišten paket *Scipy.special*.

Korisnik u terminalu odabire red očekivanja koje želi izračunati. Zatim odabire razdiobu te upisuje podatke koji se odnose na tu razdiobu, granice intervala a i b kada je u pitanju jednolika razdioba, odnosno μ i σ kada je u pitanju normalna razdioba. Aplikacija kao rezultat vraća tablicu s izračunom očekivanja dobivenih asimptotskim razvojem i Monte Carlo metodom s ponavljanjima od 10^4 i 10^5. U slučaju kada je odabran red očekivanja 0 tj. geometrijsko očekivanje, tada se računa i egzaktna vrijednost očekivanja.

Pokrenut ćemo aplikaciju za obje razdiobe, za geometrijsko, harmonijsko i kvadratno očekivanje, a za svaku razdiobu ćemo uzeti dvije različite sredine. Navedeni izračuni bit će provedeni na uzorcima duljine 10, 20 i 100. To će nam omogućiti da vidimo kako sredina i duljina uzorka utječu na samu aproksimaciju. Podatci dobiveni Monte Carlo metodom poslužit će nam za usporedbu rezultata.



```
enter r->0

(1) Uniform distribution,
(2) Normal distribution
Enter a number:1
Enter a->100
Enter b->110
                 10              20             100
exact  104.964268853  104.962283411  104.960695039
asym   104.964268869  104.962283429  104.960695059
mc4    104.982451495  104.956775546  104.961662301
mc5    104.963092155  104.962198992  104.960477674
```

Slika 10. Primjer korištenja aplikacije

4.1. Rezultati za jednoliku razdiobu

Pretpostavimo da X ima jednoliku razdiobu na intervalu $[a, b]$, gdje vrijedi 0<a<b. Tada

$$x = \mu = E(X) = \frac{a+b}{2} \tag{4.1}$$

slučajne varijable $A_1, ..., A_n$ su jednoliko rasprostranjene na $\left[-\frac{b-a}{2}, \frac{b-a}{2}\right] =: [-t, t]$ i tada vrijedi

$$\begin{aligned} \mu_2 &= E[A_1^2] = \frac{(b-a)^2}{12} = \frac{1}{3}t^2, \\ \mu_4 &= E[A_1^4] = \frac{(b-a)^4}{80} = \frac{1}{5}t^4. \end{aligned} \tag{4.2}$$

POSLJEDICA 1. Ako je slučajna varijabla X definirana s jednolikom razdiobom na intervalu $[a, b] \in \mathbb{R}$. Tada razvoj očekivanja glasi:

$$E[M_r(X)] = \mu + \frac{d^2}{\mu} + \frac{d^4}{\mu^3} + O\left(\frac{1}{\mu^5}\right), \quad \mu \to \infty, \tag{4.3}$$

gdje vrijedi

$$d^2 = \frac{r-1}{6} \cdot \frac{n-1}{n} t^2, \tag{4.4}$$

$$d^4 = -\frac{r-1}{360} \cdot \frac{n-1}{n} \left[2(3r-1)(2r-1) - (2r-1)(r+5) + (2r^2+5r+13)\right] t^4. \tag{4.5}$$

Kao što smo spomenuli, izrazit ćemo očekivanja reda 0, -1, 2 tj. geometrijsko, harmonijsko i kvadratno očekivanje, kada $\mu \to \infty$

$$E[G(X)] = \mu - \frac{1}{6} \cdot \frac{n-1}{n} t^2 \mu^{-1} + \frac{1}{360} \cdot \frac{n-1}{n} \left(2n^2 + 5n + 13\right) t^4 \mu^{-3} + O\left(\mu^{-5}\right), \tag{4.6}$$

$$E[H(X)] = \mu - \frac{1}{3} \cdot \frac{n-1}{n} t^2 \mu^{-1} + \frac{1}{90} \cdot \frac{n-1}{n} \left(12n^2 + 6n + 5\right) t^4 \mu^{-3} + O\left(\mu^{-5}\right), \quad (4.7)$$

$$E[Q(X)] = \mu + \frac{1}{8} \cdot \frac{n-1}{n} t^2 \mu^{-1} - \frac{1}{360} \cdot \frac{n-1}{n} \left(30n^2 - 21n + 31\right) t^4 \mu^{-3} + O\left(\mu^{-5}\right). \quad (4.8)$$

Za geometrijsko očekivanje također imamo egzaktnu formulu

$$E[G(X)] = \left(\frac{n}{n+1}\right) \cdot \frac{b^{1+1/n} - a^{1+1/n}}{b-a} \cdot \frac{1}{n}. \quad (4.9)$$

Pokrenut ćemo aplikaciju za geometrijsko, harmonijsko i kvadratno očekivanje slučajnih varijabli $X \sim U(100, 110)$ i $X \sim U(1000, 1010)$

Tablica 1. Geometrijsko očekivanje slučajne varijable $X \sim U(100, 110)$

	10	20	100
exact	104.964268853	104.962283411	104.960695039
asym	104.964268869	104.962283429	104.960695059
mc4	104.966142610	104.952547552	104.965890324
mc5	104.964576998	104.961135186	104.959533965

Tablica 2. Geometrijsko očekivanje slučajne varijable $X \sim U(1000, 1010)$

	10	20	100
exact	1004.996268637	1004.996061339	1004.995895500
asym	1004.996268638	1004.996061339	1004.995895500
mc4	1005.005131223	1004.996839054	1004.997816840
mc5	1004.996674217	1004.991519210	1004.995871544

Tablica 3. Harmonijsko očekivanje slučajne varijable $X \sim U(100, 110)$

	10	20	100
asym	104.928532124	104.924559464	104.921381424
mc4	104.921187778	104.930589204	104.923074674
mc5	104.929830912	104.926841952	104.921656867

Tablica 4. Harmonijsko očekivanje slučajne varijable $X \sim U(1000, 1010)$

	10	20	100
asym	1004.992537269	1004.992122670	1004.991790991
mc4	1004.976638724	1004.976736185	1004.992100022
mc5	1004.994636070	1004.994426483	1004.991407379

Tablica 5. Kvadratno očekivanje slučajne varijable $X \sim U(100, 110)$

	10	20	100
asym	105.035709967	105.037692678	105.039278598
mc4	105.040711607	105.035887337	105.039773041
mc5	105.038808954	105.041016210	105.038758074

Tablica 6. Kvadratno očekivanje slučajne varijable $X \sim U(1000, 1010)$

	10	20	100
asym	1005.003731338	1005.003938634	1005.004104469
mc4	1005.001810766	1005.011879189	1004.999756582
mc5	1005.002450066	1005.004197058	1005.004381436

4.2. Rezultati za normalnu raziobu

Primjena slučajne varijable X s normalnom razdiobom s očekivanjem $\mu > 0$i standardnom devijacijom σ.

Posljedica 2. Neka je slučajna varijabla X definirana normalnom razdiobom s očekivanjem $\mu > 0$ i standardnom devijacijom σ. Tada razvoj očekivanja glasi:

$$E[M_r(X)] = \mu + \frac{d^2}{\mu} + \frac{d^4}{\mu^3} + O\left(\frac{1}{\mu^5}\right), \quad \mu \to \infty, \tag{4.10}$$

gdje vrijedi

$$d_2 = \frac{r-1}{2} \cdot \frac{n-1}{n} \cdot \sigma^2, \tag{4.11}$$

$$d_4 = -\frac{r-1}{8} \cdot \frac{n-1}{n} \cdot \left[\frac{(2r-1)\cdot(r-3)}{n} + (3r-5)\right] \cdot \sigma^4. \tag{4.12}$$

Opet ćemo izraziti očekivanja reda 0, -1, 2 tj. geometrijsko, harmonijsko i kvadratno očekivanje, kada $\mu \to \infty$

$$E[G(X)] = \mu - \frac{n-1}{2n} \cdot \frac{\sigma^2}{\mu} - \frac{(n-1)(5n-3)}{8n^2} \cdot \frac{\sigma^4}{\mu^3} + O(\mu^{-5}), \tag{4.13}$$

$$E[H(X)] = \mu - \frac{n-1}{n} \cdot \frac{\sigma^2}{\mu} - \frac{(n-1)(2n-3)}{n^2} \cdot \frac{\sigma^4}{\mu^3} + O(\mu^{-5}), \tag{4.14}$$

$$E[Q(X)] = \mu - \frac{n-1}{2n} \cdot \frac{\sigma^2}{\mu} - \frac{(n-1)(n-3)}{8n^2} \cdot \frac{\sigma^4}{\mu^3} + O(\mu^{-5}). \tag{4.15}$$

Ovdje bi trebalo napomenuti kako je slučajna varijabla X definirana na skupu $\mathbb{R}$, a $M_r(X)$ je definiran samo za pozitivne brojeve. Kako gledamo slučajeve gdje $\mu \to \infty$, vjerojatnost $\mathcal{P}(X < 0)$ je zanemariva. Prisjetimo se funkcije gustoće jedinične normalne varijable

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, x \in \mathbb{R}. \tag{4.16}$$

Tada vrijedi

$$P(Z > x) = \int_x^{\infty} \phi(u)\, du \leq \int_x^{\infty} \frac{u}{x}\phi(u)\, du = \frac{1}{x}\phi(x). \tag{4.17}$$

Za opću slučajnu varijablu X vrijedi

$$P(X < 0) = P(Z < -\frac{\mu}{\sigma}) = P(Z > \frac{\mu}{\sigma}) \leq \frac{1}{\sqrt{2\pi}} \frac{\mu}{\sigma} e^{-\frac{\mu^2}{2\sigma^2}} \to 0, \quad \mu \to \infty. \tag{4.18}$$

Opet ćemo izraziti egzaktnu formulu za geometrijsko očekivanje

$$E[X^p] = E[|X|^p] + O(\sigma\mu e^{2\sigma^2}), \tag{4.19}$$

$$E[|X|_p] = \frac{1}{\sqrt{\pi}} \sigma^p 2^{\frac{p}{2}} \Gamma\left(\frac{p+1}{2}\right) {}_1F_1\left(-\frac{p}{2}, \frac{1}{2}, -\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2\right). \tag{4.20}$$

Za veliki μ dobivamo:

$$E[G(X)] \approx \sigma\sqrt{\frac{2}{\pi^2}} \left(\Gamma\left(\frac{n+1}{2n}\right) {}_1F_1\left(-\frac{1}{2n}, \frac{1}{2}, -\frac{1}{2}(\frac{\mu}{\sigma})^2\right)\right)^n. \tag{4.21}$$

Gama funkcija definirana je nepravim integralom

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\, dx, \alpha > 0, \tag{4.22}$$

gdje ${}_1F_1$ predstavlja konfluentnu hipergeometrijsku funkciju.

Pokrenut ćemo aplikaciju za geometrijsko, harmonijsko i kvadratno očekivanje slučajnih varijabli $X \sim N(105, \frac{25}{3})$ i $X \sim N(1005, \frac{25}{3})$

Tablica 7. Geometrijsko očekivanje slučajne varijable $X \sim N(105, \frac{25}{3})$

	10	20	100
exact	104.964253916	104.962266950	104.960677295
asym	104.964253995	104.962267038	104.960677390
mc4	104.948252201	104.957692886	104.958917892
mc5	104.962312346	104.960399022	104.959476926

Tablica 8. Geometrijsko očekivanje slučajne varijable $X \sim N(1005, \frac{25}{3})$

	10	20	100
exact	1004.996268621	1004.996061320	1004.995895480
asym	1004.996268621	1004.996061320	1004.995895480
mc4	1005.009870348	1004.999327175	1005.002280088
mc5	1004.994254630	1004.997333184	1004.995671373

Tablica 9. Harmonijsko očekivanje slučajne varijable $X \sim N(105, \frac{25}{3})$

	10	20	100
asym	104.928479646	104.924497744	104.921311575
mc4	104.933651330	104.929233251	104.919867742
mc5	104.932764150	104.927274622	104.920664082

Tablica 10. Harmonijsko očekivanje slučajne varijable $X \sim N(1005, \frac{25}{3})$

	10	20	100
asym	1004.992537209	1004.992122599	1004.991790911
mc4	1005.010104554	1004.989272224	1004.991521990
mc5	1004.992829325	1004.991463628	1004.991470616

Tablica 11. Kvadratno očekivanje slučajne varijable $X \sim N(105, \frac{25}{3})$

	10	20	100
asym	105.035709562	105.037692358	105.039278513
mc4	105.048329193	105.047526362	105.041051358
mc5	105.033863297	105.038022872	105.040111183

Tablica 12. Kvadratno očekivanje slučajne varijable $X \sim N(1005, \frac{25}{3})$

	10	20	100
asym	1005.003731338	1005.003938633	1005.004104469
mc4	1004.998257034	1005.007185298	1005.009569815
mc5	1004.999332095	1005.003048111	1005.003000987

5. Zaključak

Glavni cilj ovog rada bio je pokazati kako aproksimirati očekivanja potencijskih sredina velikih uzoraka. Za aproksimaciju očekivanja koristili smo asimptotski razvoj sredina. Trebalo je osmisliti pravilnu strukturu koda kojom bi implementirali izračune asimptotskog razvoja, očekivanja simulacija dobivenih Monte Carlo metodom i egzaktne formule za slučaj kada je u pitanju geometrijsko očekivanje.

Završni rad je ostvaren kroz tri poglavlja. U prva dva poglavlja upoznajemo osnovne pojmove iz vjerojatnosti i statistike, asimptotske razvoje potencijskih sredina, a u trećem prolazimo kroz implementaciju i rezultate samog zadatka.

Iz dobivenih podataka vidljivo je kako je aproksimacija bolja što je veće očekivanje, što se slaže s našom pretpostavkom. Korisno je spomenuti da je implementacija zadatka bila vrlo jednostavna te da je za pokretanje i izvođenje programa potrebno vrlo malo vremena. Potencijske sredine su najvažnija klasa sredina u matematici, ali problem je što ih je teško izračunati. Upravo zato nam je bio cilj pokazati točnost aproksimacije i brzinu kojom se dolazi do rješenja.

U rezultatima smo prikazali tri najčešće korištene sredine uz aritmetičku, a to su geometrijska, harmonijska i kvadratna. Iako je implementacija primjenjiva na potencijske sredine bilo kojeg reda.

LITERATURA

[1] Burić, Tomislav, Asimptotski razvoji gama funkcije i kvocijenta gama funkcija, doktorska disertacija. Zagreb : Prirodoslovno-matematicki fakultet, Matematički odsjek, 2011.

[2] Burić, Tomislav; Elezović, Neven; Mihoković, Lenka, Expectations of large data means, Journal of Mathematical Inequalities, 17 (2023), 1; 403-418

[3] Elezović, Neven, Vjerojatnost i statistika, Element, Zagreb, 2018.

[4] Pauše, Željko, Uvod u matematičku statistiku. Školska knjiga, Zagreb, 1993.

[5] Sarapa, Nikola, Teorija vjerojatnosti. Školska knjiga, Zagreb, 2002.

Aproksimacija očekivanja potencijskih sredina velikih uzoraka

Sažetak

Potencijske sredine su jedne od najvažnijih klasa sredina i imaju široku primjenu u raznim područjima. Cilj ovog završnog zadatka bio je aproksimirati ta očekivanja pomoću asimptotskog razvoja potencijskih sredina za velike uzorke. Rad prikazuje izvode matematičkih formula koje su nam bile potrebne za implementaciju zadatka. Zadatak smo implementirali u obliku aplikacije naredbenog retka koristeći programski jezik Python. Proveli smo testove za jednoliku i normalnu razdiobu, na uzorcima različitih dimenzija i različitih sredina. Također smo izvrtili testove za najčešće korištene sredine tj. geometrijsku, harmonijsku i kvadratnu.

Ključne riječi: potencijska očekivanja; asimptotski razvoj

Approximation of Expectation of Large Data Power Means

Abstract

Power means are one of the most important classes of means and have a wide range of applications in various fields. The goal of this final task was to approximate these expectations using the asymptotic expansion of power means for large samples. The paper presents the derivatives of mathematical formulas that were necessary for implementing the task. We implemented the task in the form of a command-line application using the Python programming language. We conducted tests for uniform and normal distributions, on samples of different dimensions and different means. We also conducted tests for the most commonly used means, namely geometric, harmonic, and quadratic means.

Keywords: power means; asymptotic expansion