УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Марија О. Живановић

ТЕСТОВИ ЕКСПОНЕНЦИЈАЛНОСТИ ЗАСНОВАНИ НА ЕМПИРИЈСКИМ ЛАПЛАСОВИМ ТРАНСФОРМАЦИЈАМА

мастер рад

Београд, 2019.

Ментор:

др Бојана Милошевић, доцент
Универзитет у Београду, Математички факултет

Чланови комисије:

др Марко Обрадовић, доцент
Универзитет у Београду, Математички факултет

др Милан Јовановић, доцент
Универзитет у Београду, Математички факултет

Датум одбране: 25. септембар 2019.

Брату

Наслов мастер рада: Тестови експоненцијалности засновани на емпиријским Лапласовим трансформацијама

Резиме: Тестирање сагласности са неком од расподела један је од најважнијих задатака математичке статистике. Велики број класичних тестова заснован је на разлици између претпостављене функције расподеле и њене постојане оцене - емпиријске функције расподеле. У скорије време веома популаран приступ је да се усместо поменуте разлике посматра разлика између Лапласове трансформације расподеле и њене постојане оцене. Поред оваквих тестова, актуелни су и тестови засновани на различитим типовима карактеризација на основу једнаке расподељености двеју статистика. Тада се за оцене Лапласове трансформације сваке од статистика користе U - емпиријске Лапласове трансформације које су природно уопштење емпиријских Лапласових трансформација.

У овом раду биће размотрено неколико класа тестова експоненцијалности: тестови засновани на растојању између емпиријске и теоријске Лапласове трансформације, тестови засновани на диференцијалној једначини коју задовољава теоријска Лапласова трансформација, као и тестови засновани на карактеризацијама једнаке расподељености две статистике узорка у којима се одговарајуће Лапласове трансформације оцењују U - емпиријским Лапласовим трансформацијама. Тестови ће бити упоређивани коришћењем емпиријске моћи.

Кључне речи: тестови сагласности, експоненцијална расподела, Лапласова трансформација, U - статистика, карактеризација, емпиријска моћ

Садржај

1 Увод **1**

1.1 Статистички функционали 2

1.2 U - статистике 3

1.3 V - статистике 10

2 Емпиријске Лапласове трансформације **12**

2.1 Лапласове трансформације случајних величина 12

2.2 Емпиријска Лапласова трансформација 13

2.3 U - емпиријска Лапласова трансформација 14

3 Тестови сагласности са експоненцијалном расподелом **16**

3.1 Карактеризације експоненцијалне расподеле 16

3.2 Тестови засновани на емпиријским Лапласовим трансформацијама 17

4 Тестови засновани на U - емпиријским Лапласовим трансфомацијама **20**

4.1 Асимптотска својства $J_{n,a}$ 21

4.2 Тест статистике засноване на карактеризацијама ескпоненцијалне расподеле 23

4.3 Утицај параметра a на понашање тест статистика 26

5 Емпиријска моћ **29**

5.1 Монте - Карло методе 29

5.2 Брзина конвергенције и поређење тестова 31

5.3 Закључак 35

Литература **36**

Глава 1

Увод

Пре него што се почнемо бавити изучавањем области из наслова овог рада најпре ћемо дефинисати неколико основних теорема из области вероватноће, које ће нам бити потребне за доказивање осталих теорема наведених у овом раду.

Теорема 1.0.1 (Чебишовљева неједнакост)**.** *Нека је X произвољна случајна величина, а ϵ и r позитивни бројеви. Тада важи неједнакост*

$$P\{|X| \geq \epsilon\} \leq \frac{\mathsf{E}\,|X|^r}{\epsilon^r}.$$

Специјално за $r = 2$ важи неједнакост

$$P\{|X - \mathsf{E}\,X| \geq \epsilon\} \leq \frac{D(X)}{\epsilon^2}.$$

Теорема 1.0.2 (Чебишовљев закон великих бројева)**.** *Нека је (X_n) низ независних случајних величина и $C > 0$ константа таква да за сваки природан број n важи $DX_n \leq C$. Ако је $S_n = \sum_{k=1}^{n} X_k$, онда важи*

$$\frac{S_n - \mathsf{E}\,S_n}{n} \xrightarrow{P} 0, \quad n \to \infty.$$

Теорема 1.0.3 (Хинчинов закон великих бројева)**.** *Нека је (X_n) низ независних случајних величина са истом расподелом и коначним математичким очекивањем m. Тада*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} m, \quad n \to \infty.$$

Теорема 1.0.4 (Централна гранична теорема)**.** *Нека је (X_n) низ независних случајних величина са истом расподелом, коначним математичким очекивањем m и коначном дисперзијом $\sigma^2 > 0$. Ако је $S_n = \sum_{k=1}^{n} X_k$, онда за свако $x \in \mathbb{R}$ при $n \to \infty$ важи*

$$P\Big\{\frac{S_n - nm}{\sigma\sqrt{n}} \leq x\Big\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

Теорема 1.0.5 (Теорема Слуцког)**.** *Ако низ случајних величина (X_n) у расподели конвергира ка X, $X_n \xrightarrow{D} X$, и низ случајних величина (Y_n) у расподели конвергира ка некој константи $a \in \mathbb{R}$, $Y_n \xrightarrow{D} a$, тада $X_n + Y_n \xrightarrow{D} X + a$ и $X_n Y_n \xrightarrow{D} aX$.*

1.1 Статистички функционали

Нека су (X_n) низ независних једнако расподељених случајних величина са функцијом расподеле $F(x) = P(X \leq x)$. Функцију расподеле $F(x)$ оцењујемо емпиријском функцијом расподеле.

Дефиниција 1.1.1. *Емпиријска функција расподеле $\widehat{F}_n$ је дефинисана са*

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$$

где је

$$I(X_i \leq x) = \begin{cases} 1, & \text{ако је } X_i \leq x \\ 0, & \text{ако је } X_i > x \end{cases}.$$

Дефиниција 1.1.2. *Ако уместо функције расподеле F хоћемо да оценимо $h(F)$, где је h нека функција, $h(F)$ се назива статистички функционал.*

Примери статистичких функционала:

1. Математичко очекивање је статистички функционал.

$$\mu = \mathsf{E}\, X = \int_{-\infty}^{+\infty} x dF(x)$$

2. Дисперзија је статистички функционал.

$$\sigma^2 = DX = \int_{-\infty}^{+\infty} (x - \mu)^2 dF(x) = \mathsf{E}\,(X - \mu)^2$$

3. Медијана је статистички функционал.

$$m = F^{-1}(1/2)$$

Уочавамо да математичко очекивање и дисперзију можемо да представимо у облику $\mathsf{E}\, g(X)$, где је $g(x) = x$ у случају математичког очекивања а у случају дисперзије $g(x) = (x-\mu)^2$, док медијану није могуће представити у том облику. Имамо следећу дефиницију.

Дефиниција 1.1.3. *Уколико $h(F)$ можемо да запишемо као $h(F) = \mathsf{E}\, g(X)$, онда такве функционале називамо функционали средње вредности.*

1.2 U - статистике

Параметар који оцењујемо $\Theta = h(F)$ је реална функција која може да буде очекивање, стандардна девијација, медијана, вероватноћа итд.. Постоје и друге могућности за Θ које захтевају више од једне случајне величине X за њено дефинисање. Примери оваквих функционала су $\mathsf{E}\,|X_2 - X_1|$, вероватноћа $p = P\big[(X_1, X_2) \in S\big]$, али и нека функција од p, p^2, $\sqrt{p}, 1/p$, итд..

Ово поглавље посветићемо изучавању функционала средње вредности. Очигледно је да $\mathsf{E}(X_1)$, $\mathsf{E}\,|X_2 - X_1|$ припадају тој групи, али и $P(X_1 < a)$ јер је вероватноћу могуће записати као $\mathsf{E}\,\phi(X_1)$, где је

$$\phi(x) = \begin{cases} 1, & \text{ако је } x < a \\ 0, & \text{иначе} \end{cases}.$$

Исто важи и за $P\big[(X_1, X_2) \in S\big]$. Супротно се примећује да ни медијана, ни стандардна девијација не припадају овој групи, као ниједна функција од p која није полиномна (на пример $\sqrt{p}, 1/p$).

Дакле, разматрамо Θ за које постоји природан број m и функција ϕ од m аргумената таква да је

$$\Theta = \mathsf{E}\left[\phi(X_1, \ldots, X_m)\right].$$

Из облика за Θ уочавамо да је $\phi(X_1, \ldots, X_m)$ једна непристрасна оцена параметра Θ. Без губитка општости можемо да претпоставио да је $\phi(X_1, \ldots, X_m)$ симетрична функција њених агрумената. Уколико то није случај симетризујемо, тј. уместо функције ϕ посматрамо функцију ϕ^* дефинисану као

$$\phi^*(X_1, X_2, \ldots, X_m) = \frac{1}{m!} \sum_{(i_1, \ldots, i_m)} \cdots \sum \phi(X_{i_1}, \ldots, X_{i_m})$$

где су $(i_1, \ldots, i_m)$ све m - пермутације. Примећујемо да како је $\phi(X_1, \ldots, X_m)$ непристрасна оцена за Θ, тако је и било која пермутација $\phi(X_{i_1}, \ldots, X_{i_m})$ њена непристрасна оцена, па самим тим линеарна комбинација $\phi^*(X_1, X_2, \ldots, X_m)$ је такође непристрасна оцена за Θ.

Дакле, без губитка општости претпоставимо да је $\phi(X_1, \ldots, X_m)$ симетрична функција по својим аргументима и дефинишемо

$$U = \frac{1}{\binom{n}{m}} \sum_{(i_1,\ldots,i_m)} \cdots \sum \phi(X_{i_1}, \ldots, X_{i_m}) \tag{1.1}$$

где су $(i_1, \ldots, i_m)$ све m - комбинације од n елемената. Примећујемо да је U такође непристрасна и симетрична по n - аргумената $X_1, \ldots, X_n$ оцена која се назива U - статистика. Шта више, U - статистика је једина непристрасна симетрична оцена за све F за које постоји $\Theta = h(F)$ и која има најмању дисперзију од свих других непристрасних оцена. Функција $\phi(x_1, \ldots, x_m)$ се назива језгро U - статистике које је реда m.

Примери U - статистика реда $m = 1$:

1. Нека је језгро U - статистике $\phi(x) = x$. Тада имамо

   $$U = \frac{1}{n} \sum_{i=1}^{n} \phi(X_i) = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}$$

   што је непристрасна оцена параметра $\Theta = E(X_1)$. Како је у питању сума независних једнако расподељених случајних величина на основу централне граничне теореме, уз услов да је $0 < D(X_1) < \infty$, имамо да важи $\sqrt{n}(U - \Theta) \to \mathcal{N}(0, D(X_1))$. У даљем раду показаћемо да варијација централне граничне теореме важи за U - статистике било ког реда $m \geq 1$ (видети 1.2.4).

2. Уколико је језгро U - статистике $\phi(x) = I\{x \leq t\}$ тада је U - статистика облика

   $$U = \frac{1}{n} \sum_{i=1}^{n} I\{X_1 \leq t\}.$$

   Овако дефинисаном U - статистиком смо оценили $\Theta = \mathsf{E}\, I\{X_1 \leq t\} = P\{X_1 \leq t\}$.

Примери U - статистика реда $m = 2$:

1. U - статистика чије је језгро $\phi(x_1, x_2) = |x_1 - x_2|$ је облика

$$U = \frac{1}{\binom{n}{2}} \sum \sum_{i<j} \phi(X_i, X_j) = \frac{1}{\binom{n}{2}} \sum \sum_{i<j} |X_i - X_j|.$$

   Уочавамо да је језгро U - статистике симетрично по својим аргументима и да је параметар који се оцењује $\Theta = \mathsf{E}\,|X_1 - X_2|$. Овако дефинисана U - статистика назива се Џинијева средња разлика.

2. Ако је језгро U - статистике $\phi(x_1, x_2) = I\{x_1 + x_2 > 0\}$ тада имамо

$$U = \frac{1}{\binom{n}{2}} \sum \sum_{i<j} \phi(X_i, X_j) = \frac{1}{\binom{n}{2}} \sum \sum_{i<j} I\{X_i + X_j > 0\}.$$

   И у овом случају је језгро симетрично, а параметар који оцењујемо је $\Theta = P(X_1 + X_2 > 0)$.

Да бисмо одредили асимптотско понашање U - статистика најпре ћемо дефинисати шта се подразумева под i - том пројекцијом U - статистике.

Дефиниција 1.2.1. *Нека је $\phi(X_1, \ldots, X_m)$ језгро U - статистике реда m. Њену i - ту пројекцију $1 \leq i \leq m$ дефинишемо са :*

$$\phi_i(x_1, ..x_i) = \mathsf{E}(\phi(x_1, .., x_i, X_{i+1}, .., X_m)). \tag{1.2}$$

У наредном делу доказаћемо два помоћна тврђења (Лему 1.2.1 и Теорему 1.2.1) чији ће нам резултат бити потребан за доказивање централинх теорема овог поглавља, закона великих бројева и централне граничне теореме за U - статистике.

Лема 1.2.1. *Очекивање i - те пројекције језгра U - статистике је једнако парамету који оцењујемо, тј.*

$$\mathsf{E}(\phi_i(X_1, \ldots, X_i)) = \Theta \text{ за свако } 1 \leq i \leq m \tag{1.3}$$

и ако важи

$$D(\phi_i(X_1, \ldots, X_i)) = {\sigma_i}^2 \text{ за свако } 1 \leq i \leq m \tag{1.4}$$

тада је

$$Cov(\phi(X_1, .., X_i, X_{i+1}, .., X_m), \phi(X_1, .., X_i, X_{i+1}^{'}, .., X_m^{'}) = \sigma_i^2 \tag{1.5}$$

где су $X_1, \ldots, X_i, X_{i+1}, \ldots, X_m, X_{i+1}^{'}, \ldots, X_m^{'}$ независне случајне величине са функцијом расподеле F.

Доказ. Доказаћемо случај $i = 1$ и $m = 2$. Како је

$$\phi_1(x_1) = \mathsf{E}(\phi(x_1, X_2)) = \int \phi(x_1, x_2) f(x_2) dx_2$$

и како сви интеграли постоје, па могуће је мењати поредак интеграције, имамо да је

$$\mathsf{E}(\phi_1(X_1)) = \int \Big(\int \phi(x_1, x_2) f(x_2) dx_2 \Big) f(x_1) dx_1$$

$$= \int \int \phi(x_1, x_2) f(x_1) f(x_2) dx_1 dx_2 = \mathsf{E}\, \phi(X_1, X_2) = \Theta$$

чиме смо показали први део леме.

Да бисмо показали други део леме најпре ћемо да распишемо чему је једнако очекивање квадрата прве пројекције.

$$\mathsf{E}\, \phi_1^2(X_1) = \int \Big(\int \phi(x_1, x_2) f(x_2) dx_2 \Big) \Big(\int \phi(x_1, x_3) f(x_3) dx_3 \Big) f(x_1) dx_1$$

$$= \int \int \int \phi(x_1, x_2) \phi(x_1, x_3) f(x_1) f(x_2) f(x_3) dx_1 dx_2 dx_3 = \mathsf{E}((\phi(X_1, X_2) \phi(X_1, X_3))$$

На основу претходне једнакости имамо да важи

$$cov(\phi(X_1, X_2), \phi(X_1, X_3)) = \mathsf{E}((\phi(X_1, X_2) \phi(X_1, X_3)) - \mathsf{E}((\phi(X_1, X_2))\, \mathsf{E}(\phi(X_1, X_3))$$

$$= \mathsf{E}\, \phi_1^2(X_1) - \Theta^2 = \mathsf{E}\, \phi_1^2(X_1) - \mathsf{E}(\phi_1(X_1))^2 = D(\phi_1(X_1)) = \sigma_1^2.$$

Уопштење важи за $m > 2$ и $i > 1$. ■

Теорема 1.2.1. *Дисперзија U - статистике једнака је*

$$DU = \frac{1}{\binom{n}{m}} \sum_{i=1}^{m} \binom{m}{i} \binom{n-m}{m-i} {\sigma_i}^2.$$

Уколико је ${\sigma_1}^2 > 0$ и ${\sigma_i}^2 < \infty, i = 1, \ldots, m$ тада важи

$$D(\sqrt{n} U) \xrightarrow[n \to \infty]{} m^2 {\sigma_1}^2.$$

Доказ. Први део теореме доказаћемо за $m = 2$

$$DU = D\Big(\frac{1}{\binom{n}{2}} \sum \sum_{i<j} \phi(X_i, X_j) \Big) = \frac{1}{\binom{n}{2}^2} \sum \sum_{i<j} \sum \sum_{k<l} cov(\phi(X_i, X_j), \phi(X_k, X_l))$$

$$= \frac{1}{\binom{n}{2}} \frac{2}{n(n-1)} \Big(\frac{n(n-1)}{2} \sigma_2^2 + \binom{n}{2} (n-2) \sigma_1^2 \Big) = \frac{1}{\binom{n}{2}} \Big(2(n-2)\sigma_1^2 + \sigma_2^2 \Big)$$

$$= \frac{1}{\binom{n}{2}} \sum_{i=1}^{2} \binom{2}{i} \binom{n-2}{2-i} {\sigma_i}^2.$$

Тврђење остаје на снази за било које $m \geq 2$.
Да бисмо показали да је $D(\sqrt{n}U) \xrightarrow[n\to\infty]{} m^2{\sigma_1}^2$ користићемо први део теореме. Уочимо да је $\binom{n-m}{k} = \frac{1}{k!}(n-m)\dots(n-m-k+1) \sim \frac{n^k}{k!}$, одакле добијамо

$$DU \sim \frac{1}{\binom{n}{m}} \sum_{i=1}^{m} \binom{m}{i} \frac{n^{m-i}}{(m-i)!} \sigma_i^2 \sim \frac{m!}{n^m} m \frac{n^{m-1}}{(m-1)!} \sigma_1^2 = \frac{1}{n} m^2 \sigma_1^2.$$

Друга еквивалентност у претходном запису важи због услова ${\sigma_i}^2 < \infty, i = 1, \dots, m$, јер у супротном израз $\frac{1}{n^m}\sigma_i^2$ није дефинисан, па имамо тврђење из теореме

$$D(\sqrt{n}U) \xrightarrow[n\to\infty]{} m^2\sigma_1^2.$$

■

Теорема 1.2.2 (Закон великих бројева за U - статистике)**.** *Уколико је* ${\sigma_1}^2 > 0$ *и* ${\sigma_i}^2 < \infty, i = 1, \dots, m$ *тада важи* $U \xrightarrow{P} \Theta$.

Доказ. У доказу теореме користићемо Чебишовљеву неједнакост примењену на $U - \Theta$. Како на основу теореме 1.2.1 имамо да $D(\sqrt{n}U) \xrightarrow[n\to\infty]{} m^2{\sigma_1}^2$, па следи

$$P(|U - \Theta| \geq \epsilon) = P(|U - \mathsf{E}\,U| \geq \epsilon) \leq \frac{DU}{\epsilon^2} = \frac{D(\sqrt{n}U)}{n\epsilon^2} \xrightarrow[n\to\infty]{} 0, \tag{1.6}$$

где је $\epsilon > 0$. На основу дефиниције конвергенције у вероватноћи важи тврђење из теореме $U \xrightarrow{P} \Theta$. ■

Теорема 1.2.3 (Јаки закон великих бројева за U - статистике)**.** *Када је* ${\sigma_1}^2 > 0$ *и* ${\sigma_i}^2 < \infty, i = 1, \dots, m$ *важи* $U \xrightarrow{C.C.} \Theta$.

За доказ теореме погледати [15].

Теорема 1.2.4 (Централна гранична теорема за U - статистике)**.** *Уколико је* $0 < {\sigma_i}^2 < \infty, i = 1, \dots, m$ *тада важи*

$$\sqrt{n}(U - \Theta) \xrightarrow{D} \mathcal{N}(0, m^2{\sigma_1}^2). \tag{1.7}$$

Доказ. Означимо са $\psi(X_i) = m(\phi_1(X_i) - \Theta), 1 \leq i \leq n$. Имамо

$$\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\psi(X_i) = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}m\big(\phi_1(X_i)-\Theta\big) = m\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\phi_1(X_i)-\Theta\big) \xrightarrow{D} \mathcal{N}(0, m^2\sigma_1^2) \tag{1.8}$$

Како важи (1.3) и (1.4) и при томе су $\phi_1(X_i), 1 \leq i \leq n$, низ једнако расподељених независних случајних величина према централној граничној теореми имамо да важи (1.8). Показаћемо да $\sqrt{n}(U-\Theta)$ можемо представити у следећем облику

$$\sqrt{n}(U - \Theta) \stackrel{D}{=} \sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\psi(X_i) + \sqrt{n}R_n \tag{1.9}$$

где је $\sqrt{n}R_n$ остатак за који важи $\sqrt{n}R_n \xrightarrow{D} 0$. Запишимо (1.9) у облику

$$\sqrt{n}R_n \stackrel{D}{=} \sqrt{n}(U - \Theta) - \sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i) - \Theta\big) \tag{1.10}$$

и покажимо да $\sqrt{n}(U-\Theta)-\sqrt{n}\frac{a}{n}\sum_{i=1}^{n}\big(\phi_1(X_i)-\Theta\big)$ тежи 0 у средње квадратном смислу. Тиме показујемо да $\sqrt{n}R_n$ у расподели тежи 0, па на основу теореме Слуцког примењене на (1.9), и имајући у виду (1.8), имамо да важи тврђење из теореме.

$$\begin{aligned}
&\mathsf{E}\Big(\sqrt{n}(U - \Theta) - \sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i) - \Theta\big)\Big)^2 \\
&= D\Big(\sqrt{n}(U - \Theta) - \sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i) - \Theta\big)\Big) \\
&= nD\Big((U - \Theta) - \frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i) - \Theta\big)\Big) \\
&= nD\Big(U - \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i) - (m-1)\Theta\Big) \\
&= nD\Big(U - \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big)
\end{aligned}$$

Прва у претходном низу једнакости важи јер је $\mathsf{E}(\sqrt{n}(U - \Theta)) = 0$ и $\mathsf{E}(\sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i) - \Theta\big)) = 0$ па је очекивање квадрата једнако дисперзији. Последња једнакост важи јер је дисперзија константе, у овом случају $(m-1)\Theta$, једнака 0. Даље имамо

$$nD\Big(U - \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big) = nDU + nD\Big(\frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big) - 2n\mathrm{Cov}\Big(U, \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big).$$

На основу теореме 1.2.1 имамо да први сабирак $nDU \underset{n\to\infty}{\to} m^2{\sigma_1}^2$, а на основу дефиниције за ${\sigma_1}^2$, имамо да је други сабирак, $nD\big(\frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i),\big)$ једнак $m^2{\sigma_1}^2$. Остаје да се покаже да је $n\mathrm{Cov}\big(U, \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\big)$ такође једнако $m^2{\sigma_1}^2$ и самим тим би важило $\mathsf{E}\left(\sqrt{n}(U-\Theta)-\sqrt{n}\frac{m}{n}\sum_{i=1}^{n}\big(\phi_1(X_i)-\Theta\big)\right)^2 \to 0$, кад $n\to\infty$, чиме би комплетирали доказ теореме.

$$
\begin{aligned}
&n\mathrm{Cov}\Big(U, \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big)\\
&= n\mathrm{Cov}\Big(\frac{1}{\binom{n}{m}}\sum_{(i_1,\dots,i_m)}\cdots\sum\phi(X_{i_1},\dots,X_{i_m}), \frac{m}{n}\sum_{i=1}^{n}\phi_1(X_i)\Big)\\
&= n\frac{1}{\binom{n}{m}}\frac{m}{n}\mathrm{Cov}\Big(\sum_{(i_1,\dots,i_m)}\cdots\sum\phi(X_{i_1},\dots,X_{i_m}), \sum_{i=1}^{n}\phi_1(X_i)\Big)\\
&= m\frac{1}{\binom{n}{m}}\binom{n}{m}m\mathrm{Cov}\big(\phi(X_1,\dots,X_m),\phi_1(X_1)\big)\\
&= m^2\,\mathsf{E}(\phi(X_1,\dots,X_m)\phi_1(X_1)) - \mathsf{E}(\phi(X_1,\dots,X_m))\,\mathsf{E}(\phi_1(X_1))\\
&= m^2(\mathsf{E}\,\phi_1^2(X_1) - (\mathsf{E}\,\phi_1(X_1))^2) = m^2D(\phi_1(X_1)) = m^2\sigma_1^2
\end{aligned}
$$

У претходним једнакостима смо користили својство $\mathsf{E}(\phi(X_1,\dots,X_m)\phi_1(X_1)) = \mathsf{E}\,\phi_1^2(X_1), \forall m \geq 1$, које ћемо ради једноставности записа показати да важи за $m=2$.

$$
\begin{aligned}
&\mathsf{E}(\phi(X_1,X_2)\phi_1(X_1))\\
&= \int\int\phi(x_1,x_2)\int\phi(x_1,x_2')f(x_2')dx_2'f(x_1)f(x_2)dx_1dx_2\\
&= \int\int\int\phi(x_1,x_2)\phi(x_1,x_2')f(x_1)f(x_2)f(x_2')dx_1dx_2dx_2'\\
&= \int\Big(\int\phi(x_1,x_2)f(x_2)dx_2\int\phi(x_1,x_2')f(x_2')dx_2'\Big)f(x_1)dx_1\\
&= \int\phi_1^2(x_1)f(x_1)dx_1 = \mathsf{E}\,\phi_1^2(X_1)
\end{aligned}
$$

■

1.3 V - статистике

Нека је $\Theta = h(F) = \int \phi(X_1, \ldots, X_m) dF$ функционал средње вредности. Као оцену параметра Θ методом замене добијамо статистику

$$V = h(\widehat{F}_n) = \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \phi(X_{i_1}, \ldots, X_{i_m}).$$

Предложена статистика назива се V - статистика.

Својства V - статистика:

- За $m = 1$ V - статистике су једнаке U - статистикама:

$$V = \frac{1}{n} \sum_{i=1}^{n} \phi(X_i) = U$$

- За $m = 2$ имамо да је V - статистика облика

$$V = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi(X_i, X_j) = \frac{1}{n^2} \sum_{i \neq j} \phi(X_i, X_j) + \frac{1}{n^2} \sum_{i=1}^{n} \phi(X_i, X_i).$$

  U - статистика реда $m = 2$ је

$$U = \frac{1}{\binom{n}{2}} \sum \sum_{i<j} \phi(X_i, X_j) = \frac{1}{n(n-1)} \sum \sum_{i<j} \phi(X_i, X_j).$$

  За U - статистику смо констатовали да је то непристрасна симетрична оцена параметра Θ, а за V - статистику имамо

$$\mathsf{E}\, V = \frac{n-1}{n} \Theta + \frac{1}{n} \mathsf{E}(X_1, X_1) = \Theta + \frac{1}{n} (\mathsf{E}(X_1, X_1) - \Theta)$$

  па је V - статистика асимптотски непристрасна оцена за Θ. Можемо приметити да су U и V статистике јако сличне.

- За $m = 2$, уколико је $0 < {\sigma_i}^2 < \infty, i = 1, m$, гранична расподела за V - статистике је иста као гранична расподела одговарајуће U - статистике, тј. важи

$$\sqrt{n}(V - \Theta) \xrightarrow{D} \mathcal{N}(0, 4{\sigma_1}^2).$$

Запишимо V - статистику у зависности од одговарајуће U - статистике.

$$V = \frac{1}{n^2}\sum_{i\neq j}\phi(X_i, X_j) + \frac{1}{n^2}\sum_{i=1}^{n}\phi(X_i, X_i)$$
$$= \frac{n-1}{n}\frac{1}{n(n-1)}\sum_{i\neq j}\phi(X_i, X_j) + \frac{1}{n^2}\sum_{i=1}^{n}\phi(X_i, X_i)$$
$$= \frac{n-1}{n}U + \frac{1}{n^2}\sum_{i=1}^{n}\phi(X_i, X_i)$$

Одузимањем Θ са леве и десне стране а затим множењем са $\sqrt{n}$ добијамо

$$\sqrt{n}(V-\Theta) = \sqrt{n}\Big(\frac{n-1}{n}U - \Theta + \frac{1}{n^2}\sum_{i=1}^{n}\phi(X_i, X_i)\Big)$$
$$= \frac{n-1}{n}\sqrt{n}(U-\Theta) + \frac{\sqrt{n}}{n}\Theta + \frac{\sqrt{n}}{n^2}\sum_{i=1}^{n}\phi(X_i, X_i)$$
$$= \frac{n-1}{n}\sqrt{n}(U-\Theta) + \frac{\sqrt{n}}{n^2}\sum_{i=1}^{n}\big(\phi(X_i, X_i) - \Theta\big).$$

Први сабирак претходног израза на основу Теореме 1.2.4 тежи у расподели ка $\mathcal{N}(0, 4{\sigma_1}^2)$ а други сабирак према закону великих бројева у вероватноћи тежи ка 0. На основу теореме Слуцког $\sqrt{n}(V-\Theta) \xrightarrow{D} \mathcal{N}(0, 4\sigma_1^2)$.

Важи уопштење претходне особине за било које $m > 2$, тј. гранична расподела за V - статистике је иста као гранична расподела одговарајуће U - статистике.

Теорема 1.3.1. *Ако важи* $0 < \sigma_i^2 < \infty$ $i = 1, \ldots, m$ *онда кад* $n \to \infty$ *важи*

$$\sqrt{n}(V-\Theta) \xrightarrow{D} \mathcal{N}(0, m^2\sigma_1^2).$$

Глава 2

Емпиријске Лапласове трансформације

2.1 Лапласове трансформације случајних величина

Нека је X ненегативна случајна величина чија је функција расподеле F.

Дефиниција 2.1.1. *Лапласова трансформација L случајне величине X је функција дефинисана на $\mathbb{R}_+$*

$$L(t) = \mathsf{E}(e^{-tX}) = \int_0^\infty e^{-tx} dF(x), t > 0.$$

Како је $e^{-tx} \leq 1$ и F је функција расподеле имамо да је $L(t) \leq 1$ за свако $t > 0$ па је Лапласова трансформација дефинисана за сваку ненегативну случајну величину.

Особине Лапласове трансфомације:

- Лапласова трансформација на јединствен начин одређује расподелу чија је она трансформација.

- Нека су X_1 и X_2 ненегативне, независне случајне величине са функцијама расподеле F_1 и F_2 и Лапласовим трансформацијама L_{X_1} и L_{X_2} редом. Тада важи

$$L_{X_1+X_2}(t) = L_{X_1}(t) L_{X_2}(t). \tag{2.1}$$

Користећи поменуту дефиницију Лапласове трансформације случајне величине и чињеницу да је очекивање производа независних случајних величина једнако производу њихових очекивања имамо (2.1)

$$L_{X_1+X_2}(t) = \mathsf{E}\, e^{-t(X_1+X_2)} = \mathsf{E}(e^{-tX_1}e^{-tX_2}) = \mathsf{E}\, e^{-tX_1}\, \mathsf{E}\, e^{-tX_2} = L_{X_1}(t)L_{X_2}(t).$$

- За $t > 0$, L има изводе свих редова $n \geq 1$ и сваки извод се добија диференцирањем под знаком интеграла

$$(-1)^n \frac{d^n}{dt^n} L(t) = \int_0^\infty e^{-tx} x^n dF(x), t > 0.$$

  Уочимо да је $\mathsf{E}\, X = -L^{'}(0)$ и $\mathsf{E}\, X^2 = L^{''}(0)$.

- За Лапласову трансформацију случајне величине важи

$$\int_0^\infty e^{-tx} F(x) dx = \frac{L(t)}{t}.$$

  Расписивањем израза са леве стране једнакости и коришћењем чињенице да су у питању мерљиве, ненегативне функције и да сви интеграли постоје, могуће је мењати поредак интеграције, закључујемо

$$\int_0^\infty e^{-tx} F(x) dx = \int_0^\infty e^{-tx} \Big(\int_0^x dF(y) \Big) dx$$

$$= \int_0^\infty \Big(\int_{x=y}^\infty e^{-tx} dx \Big) dF(y) = \int_0^\infty \frac{1}{t} e^{-ty} dF(y) = \frac{L(t)}{t}.$$

2.2 Емпиријска Лапласова трансформација

Нека је $X_1, X_2, X_3, \ldots, X_n$ прост случајан узорак случајне величине X са коначним математичким очекивањем и коначном дисперзијом. Имајући у виду дефиницију Лапласове трансформације случајне величине природно је да је њена емпиријска Лапласова трансформација облика

$$L_n(t) = \int_0^\infty e^{-tx} d\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n e^{-tX_i}, t \geq 0$$

где је $\widehat{F}_n$ емпиријска функција расподеле.

У наставку ћемо навести својства која важе за емпиријске Лапласове трансформације.

Лема 2.2.1. *За свако $t \geq 0$ важи*

$$L_n(t) \xrightarrow{P} L(t)$$

где је $L(t) = \mathsf{E}(e^{-Xt})$.

Доказ. Како су X_i, $1 \leq i \leq n$ независне случајне величине то важи и за e^{-tX_i}, $1 \leq i \leq n$ па на основу Чебишовљевог закона великих бројева имамо тврђење из леме. ■

Пре него што дефинишемо још једно својство емпиријске Лапласове трансформације дефинисаћемо шта се подразумева под Гаусовим процесом.

Дефиниција 2.2.1. *Стохастички процес $\{X_t\}_{t\geq 0}$ је Гаусов процес ако за свако n и за сваки избор $0 \leq t_1 < t_2 < \cdots < t_n$ вектор $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ има вишедимензиону нормалну расподелу.*

Лема 2.2.2. *За свако $t \geq 0$ важи*

$$\sqrt{n}(L_n(t) - L(t)) \xrightarrow{D} \eta(t)$$

где је $\eta(t)$ Гаусов процес.

2.3 U - емпиријска Лапласова трансформација

Нека су $X_1, \ldots, X_n$ прост случајни узорак из ненегативне непрекидне функције расподеле F која има коначно математичко очекивање и коначну дисперзију. Без губитка општости претпоставимо да је $\omega(X_1, \ldots, X_m)$, $m \leq n$, симетрична функција својих аргумената. У претходном поглављу смо дефинисали шта су U - статистике и емпиријске Лапласове трансформације, па узимајући у обзир те две дефиниције имамо да је U - емпиријска Лапласова трансформација следећег облика

$$L_n^U(t) = \frac{1}{\binom{n}{m}} \sum_{(i_1,\ldots,i_m)} \cdots \sum e^{-t\omega(X_{i_1},\ldots,X_{i_m})}.$$

Уочавамо да је језгро U - статистике облика $\phi(X_1, \ldots, X_a) = e^{-t\omega(X_1,\ldots,X_a)}$. Аналогно, V - емпиријска Лапласова трансформација је облика

$$L_n^V(t) = \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} e^{-t\omega(X_{i_1},\ldots,X_{i_m})}.$$

Лема 2.3.1. *За свако* $t \geq 0$ *важи* $L_n^U(t) \xrightarrow{P} L^U(t)$ *где је* $L^U(t) = \mathsf{E}\, e^{-t\omega(X_1,\ldots,X_a)}$.

Доказ. Када $t \geq 0$ фиксирамо, $L_n^U(t)$ можемо да посматрамо као U - статистику дефинисану у Поглављу 1. па за њу имамо да важи Теорема 1.2.2. ■

Глава 3

Тестови сагласности са експоненцијалном расподелом

3.1 Карактеризације експоненцијалне расподеле

Експоненцијална расподела је једна од најчешће коришћених расподела и има широку примену у моделовању података у теорији поузданости, за моделирање времена чекања између два узастопна догађаја и многим другим пољима. Њена одговарајућа функција расподеле има веома једноставан облик

$$F(x) = 1 - e^{-\lambda x}, x \geq 0, \lambda > 0.$$

Због тога, како и због њене широке употребе, постоји велики број карактеризација експоненцијалне расподеле које могу бити изражене у погодном облику. Дефинишимо шта се подразумева под карактеризацијом неке расподеле.

Дефиниција 3.1.1. *Неке за случајни вектор $\mathcal{X} = (X_1, X_2, X_3, \ldots, X_n)$ и функцију $\mathcal{L}$ постоји фамилија расподела $\mathcal{F}$ таква да $\mathcal{L}(\mathcal{X}) \in \mathcal{F}$ повлачи да $\mathcal{X}$ има неку особину $\mathcal{P}$. Ако важи и други смер овог тврђења, тј. ако из „$\mathcal{X}$ има особину $\mathcal{P}$" следи $\mathcal{L}(\mathcal{X}) \in \mathcal{F}$ тада имамо карактеризацију расподеле $\mathcal{F}$.*

Навешћемо неке карактеризације експоненцијалне расподеле које ће бити од користи при формирању тест статистика у овом раду.

Карактеризација 3.1.1. *(Десу (1971)) Нека је X случајна величина са функцијом расподеле F. Нека је $X_1, \ldots, X_n$ узорак из F и $W = min(X_1, \ldots, X_n)$. Ако је F недегенерисана функција расподеле тада за свако позитивно n, nW и*

X имају исту расподелу ако и само ако је $F(x) = 1 - e^{-\lambda x}$, за $x \geq 0$, где је λ позитивна константа.

Карактеризација 3.1.2. *(Пури и Рубин) Нека су X_1 и X_2 независне случајне величине са истом расподелом као случајна величина X која има функцију густине $f(x)$. Тада X и $|X_1 - X_2|$ имају исту расподелу ако и само ако за неко $\lambda > 0, f(x) = \lambda e^{-\lambda x}$, за $x \geq 0$.*

Карактеризација 3.1.3. *(Б. Милошевић и М. Обрадовић [10]) Нека су X_1 и X_2 независне једнако расподељене случајне величине са функцијом густине $f(x)$. Тада $\max(X_1, X_2) \stackrel{D}{=} \frac{1}{2}X_1 + X_2$ ако и само ако за неко $\lambda > 0$, $f(x) = \lambda e^{-\lambda x}$, за $x \geq 0$.*

Карактеризација 3.1.4. *(Л. Барингхаус и Н. Хензе [12]) Лапласова трансформација $\psi(t) = \mathsf{E}(e^{-Xt})$ случајне величине X задовољава диференцијалну једначину $(\lambda + t)\psi'(t) + \psi(t) = 0, t > 0$ ако и само ако случајна величина X има експоненцијалну расподелу са параметром λ, $X \in \varepsilon(\lambda)$.*

3.2 Тестови засновани на емпиријским Лапласовим трансформацијама

Нека је $X_1, \ldots, X_n$ прост случајни узорак из ненегативне непрекидне функције расподеле F. Тестирамо хипотезу

$$H_0 : F(x) = 1 - e^{-\lambda x}, \quad \lambda > 0,\ x \geq 0.$$

У претходном поглављу дефинисали смо емпиријску Лапласову трансформацију $X_1, \ldots, X_n$ која је облика

$$\psi_n(t) = \frac{1}{n} \sum_{j=1}^{n} e^{-tX_j}.$$

Уместо теста заснованог на $X_1, \ldots, X_n$ разматра се тест статистика заснована на узорку $Y_1, \ldots, Y_n$ где су $Y_i = X_i/\bar{X}_n$, $1 \leq i \leq n$ а $\bar{X}_n = n^{-1} \sum_{j=1}^{n} X_j$ узорачка средина. Означимо са

$$\varphi_n(t) = \frac{1}{n} \sum_{j=1}^{n} e^{-tY_j}$$

емпиријску Лапласову трансформацију узорка $Y_1, \ldots, Y_n$. У свом раду [14] Н. Хензе и С. Ментанис предлажу следећу тест статистику

$$T_{n,a} = n \int_0^\infty \big(\varphi_n(t) - \frac{1}{1+t}\big)^2 e^{-at} dt$$

где e^{-at} представља тежинску функцију која обезбеђују конвергенцију интеграла и a је позитивна константа. Уочавамо да при овако дефинисаним $Y_i, 1 \leq j \leq n$, при нултој хипотези, тест статистика $T_{n,a}$ не зависи од параметра експоненцијалне расподеле λ јер $\varphi_n(t)$ не зависи од параметра λ. Према леми 2.2.1 закључујемо да $\varphi_n(t)$ тежи ка Лапласовој трансформацији јединичне експоненцијалне расподеле, $\varphi(t) = \dfrac{1}{1+t}$, па ће велике вредности тест статистике $T_{n,a}$ бити разматране за одбацивање нулте хипотезе. Решавањем интеграла добијамо облик

$$T_{n,a} = \frac{1}{n} \sum_{j,k=1}^{n} \frac{1}{Y_j + Y_k + a} - 2 \sum_{j=1}^{n} e^{(Y_j+a)} \, \mathsf{E}_i(Y_j + a) + n(1 - ae^a \, \mathsf{E}_i(a))$$

где је

$$\mathsf{E}_i(z) = \int_z^\infty \frac{e^{-t}}{t} dt, z > 0$$

експоненцијални интеграл.

Варијација претходне тест - статистике је да уместо разлике $\varphi_n(t) - \dfrac{1}{1+t}$ посматрамо $(1+t)\varphi_n(t) - 1$, такође представљено у раду [14], што је блиско 0 при нултој хипотези па тест статистика постаје

$$L_{n,a} = n \int_0^\infty ((1-t)\varphi_n(t) - 1)^2 e^{-at} dt$$

коју можемо записати и на следећи начин

$$L_{n,a} = n \int_0^\infty \big(\varphi_n(t) - \frac{1}{1+t}\big)^2 (1-t)^2 e^{-at} dt.$$

Уочавамо да се $L_{n,a}$ разликује од $T_{n,a}$ само по тежинској функцији која је $(1-t)^2 e^{-at}$ уместо e^{-at}. Решавањем интеграла добијамо затворени облик за $L_{n,a}$ који не захтева израчунавање експоненцијалног интеграла

$$L_{n,a} = \frac{1}{n} \sum_{j,k=1}^{n} \frac{1 + (1 + a + Y_j + Y_k)^2}{(Y_j + Y_k + a)^3} - 2 \sum_{j=1}^{n} \frac{1 + Y_j + a}{(Y_j + a)^2} + \frac{n}{a}.$$

Статистике $L_{n,a}$ и $T_{n,a}$ можемо представити у заједничком облику

$$W_{n,a} = n \int_0^\infty (\varphi_n(t) - \frac{1}{1+t})^2 w(t) e^{-at} dt$$

где је $w(t) = 1$ за $T_{n,a}$ и $w(t) = (1-t)^2$ за $L_{n,a}$.

Теорема 3.2.1. *Претпоставимо да је $w : [0, \infty) \to \mathbb{R}$ непрекидна и ненегативна функција које има полиномијални раст, тј. постоји цели број r тако да је $w(t) = O(t^r)$, $t \to \infty$. Ако важи нулта хипотеза, $X_i \in \varepsilon(\lambda)$, $1 \leq i \leq n$, тада $W_{n,a}$ можемо представити у облику*

$$W_{n,a} = \sum_{i \geq 1} \lambda_i(w, a) N_i^2 \tag{3.1}$$

где су $N_1, N_2, \ldots$ независне стандардне нормалне случајне величине и $\lambda_i(w, a)$ функције које зависе од функције w и слободног параметра a.

Дакле $W_{n,a}$ се може представити као линеарна комбинација независних случајних величина са χ^2 расподелом. Доказ теореме је представљен у раду Н. Хензеа и С. Ментаниса, видети [14].

Узимајући у обзир Карактеризацију 3.1.4 за тестирање хипотезе о експоненцијалности Хензе и Ментанис предлажу још једну тест статистику која има веома погодан облик

$$D_{n,a} = n \int_0^\infty \left((\widehat{\lambda}_n + t)\psi_n'(t) + \psi_n(t) \right)^2 \bar{X}_n e^{-a\bar{X}_n t} dt$$

где је $\widehat{\lambda}_n = \bar{X}_n^{-1}$ оцена за λ методом максималне веродостојности и a позитивна константа. Хипотезу H_0 одбијамо за велике вредности тест статистика $D_{n,a}$ и $W_{n,a}$.

Као што је већ изложено за тежинску функцију је изабрана $\bar{X}_n e^{-a\bar{X}_n t}$ из неколико разлога. Први разлог је њен пожељан облик добијен након решавања интеграла

$$D_{n,a} = \frac{1}{n} \sum_{j,k=1}^{n} \left[\frac{(1 - Y_j)(1 - Y_k)}{Y_j + Y_k + a} - \frac{(Y_k + Y_j)}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right]$$

где је $Y_j = X_j / \bar{X}_n$, $1 \leq j \leq n$. Имајући у виду претходни израз за $D_{n,a}$, наша статистика зависи од $X_1, \ldots, X_n$ преко $Y_1, \ldots, Y_n$ и добијамо да при нултој хипотези не зависи од параметра λ експоненцијалне расподеле.

Одабир слободног параметра a повећава моћ теста при одређеним хипотезама. Нпр. бирајући вредности за a блиске нули, а тиме допуштајући да тежинска функција опада споро, даје значајнију моћ теста алтернативама чија је густина расподеле концентрисана око нуле. Супротно, бирајући велике вредности слободног параметра a имамо да тежинска функција опада брзо ка нули и самим тим већу моћ теста против алтернатива које имају значајније разлике у погледу репа у односу на експоненцијалну расподелу.

Глава 4

Тестови засновани на U - емпиријским Лапласовим трансфомацијама

У наредном поглављу размотрићемо тестове експоненцијалности засноване на U - емпиријским Лапласовим трансформацијама. Размотримо карактеризацију експоненцијалне расподеле у форми

$$\omega_1(X_1, X_2, \ldots, X_m) \stackrel{D}{=} \omega_2(X_1, X_2, \ldots, X_m) \tag{4.1}$$

где су ω_1 и ω_2 ненегативне хомогене функције независних једнако расподељених случајних величина $X_1, X_2, X_3, \ldots, X_n$, $m \leq n$, за које важи

$$\omega_i(cX_1, cX_2, \ldots, cX_m) = c\omega_i(X_1, X_2, \ldots, X_m), \forall c \geq 0, i = 1, 2. \tag{4.2}$$

Додатно претпоставимо да је $X_1, X_2, X_3, \ldots, X_n$ узорак из ненегативне непрекидне функције расподеле F. У свом раду Б. Милошевић и М. Обрадовић [9] за тестирање нулте хипотезе $H_0 : F(x) = 1 - e^{-\lambda x}, \lambda > 0$ предлажу следећу тест статистику интегралног типа

$$J_{n,a} = \int_0^\infty (L_n^{(1)}(t) - L_n^{(2)}(t))\bar{X}e^{-a\bar{X}t}dt \tag{4.3}$$

где је $\bar{X}$ узорачка средина, број a позитивна константа и

$$L_n^{(k)}(t) = \frac{1}{n^{[m]}} \sum_{1 \leq i_1 < \cdots < i_m \leq n, \pi \in \Pi(m)} \cdots \sum e^{-t\omega_k\left(X_{i_{\pi(1)}}, \ldots, X_{i_{\pi(m)}}\right)}, k = 1, 2 \tag{4.4}$$

су U- емпиријске Лапласове трансформације, при чему $n^{[m]} = m!\binom{n}{m}$ и $\Pi(m)$ представља скуп свих пермутација бројева $\{1, \ldots, m\}$. Експоненцијална тежинска функција обезбеђује конвергенцију интеграла а улога узорачке средине је да статистика буде слободна од параметра λ под нултом хипотезом. Позитивна константа a се одређује да би повећали моћ теста против одређених алтернатива.

Велике позитивне и негативне вредности тест статистике $J_{n,a}$ биће значајне и тест ће бити посматран против свих алтернатива где теоријска вредност $J_{n,a}$ није нула, који укључује расподеле од практичног интереса.

Пре него што пређемо на израчунавање асимптотских својстава статистике $J_{n,a}$ уочимо да сменом $y = \bar{X}t$ у интегралу (4.3) тест статистика добија облик

$$J_{n,a} = \int_0^\infty (L_n^{(1)}(y) - L_n^{(2)}(y))e^{-ay}dy$$

где U - емпиријске Лапласове трансформације су облика

$$L_n^{(k)}(y) = \frac{1}{n^{[m]}} \sum_{1 \le i_1 < \cdots < i_m \le n, \pi \in \Pi(m)} \cdots \sum e^{-y\omega_k\left(Y_{i_{\pi(1)}}, \ldots, Y_{i_{\pi(m)}}\right)}, k = 1, 2$$

при чему је $Y_i = X_i/\bar{X}$ и користили смо својство (4.2).

Аналогно тест статистикама заснованим на U - емпиријским Лапласовим трансформацијама дефинишемо тест статистике засноване на V - емпиријским Лапласовим трансформацијама

$$K_{n,a} = \int_0^\infty (L_n^{(1)}(y) - L_n^{(2)}(y))e^{-ay}dy$$

где су

$$L_n^{(1)}(y) = \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} e^{-y\omega_k\left(Y_{i_1}, \ldots, Y_{i_m}\right)}, k = 1, 2.$$

4.1 Асимптотска својства $J_{n,a}$

Након интеграције израз (4.3) постаје

$$J_{n,a} = \frac{\bar{X}}{n^{[m]}} \sum_{1 \le i_1 < \cdots < i_m \le n, \pi \in \Pi(m)} \cdots \sum \frac{1}{a\bar{X} + \omega_1(X_{i_{\pi(1)}}, \ldots, X_{i_{\pi(m)}})} - \frac{1}{a\bar{X} + \omega_2(X_{i_{\pi(1)}}, \ldots, X_{i_{\pi(m)}})}. \tag{4.5}$$

Да бисмо нашли асимптотску расподелу за $\sqrt{n}J_{n,a}$ под H_0 посматрајмо помоћну функцију

$$J_{n,a}(\mu) = \frac{\mu}{n^{[m]}} \sum_{1\le i_1<\cdots<i_m\le n,\pi\in\Pi(m)}\cdots\sum \frac{1}{a\mu+\omega_1(X_{i_{\pi(1)}},\ldots,X_{i_{\pi(m)}})} - \frac{1}{a\mu+\omega_2(X_{i_{\pi(1)}},\ldots,X_{i_{\pi(m)}})} \tag{4.6}$$

гдје је $\mu = \lambda^{-1}$ јер $J_{n,a}$ зависи од оцењеног параметра $\widehat{\mu} = \bar{X}$. Прецизније записано $J_{n,a} = J_{n,a}(\widehat{\mu})$. За свако фиксирано $\mu > 0$ имајући у виду (4.2) закључујемо да $J_{n,a}(\mu)$ је U - статистика која не зависи од μ. Дакле, ради поједностављивања нека је $\mu = 1$. Уочавамо да по дефиницији наша U - статистика $J_{n,a}(1)$ има симетрично језгро

$$\Phi(X_1,\ldots,X_m;a) = \frac{1}{m!}\sum_{\pi\in\Pi(m)} \frac{1}{a+\omega_1(X_{\pi(1)},\ldots,X_{\pi(m)})} - \frac{1}{a+\omega_2(X_{\pi(1)},\ldots,X_{\pi(m)})}. \tag{4.7}$$

Ако је језгро недегенерисано можемо применити Теорему 1.2.4 и добити асимптотску расподелу за $\sqrt{n}J_{n,a}(1)$ која је $\mathcal{N}(0, m^2\sigma_1^2(a))$ гдје је $\sigma_1^2(a)$ дисперзија прве пројекције језгра, тј.

$$\sigma_1^2(a) = \mathsf{E}(\phi_1^2(X_1,a))$$

$$\phi_1(s,a) = \mathsf{E}(\Phi(X_1,\ldots,X_m;a)|X_1 = s).$$

Теорема 4.1.1. *Асимптотске расподеле за $\sqrt{n}J_{n,a}(\mu)$ и $\sqrt{n}J_{n,a} = \sqrt{n}J_{n,a}(\widehat{\mu})$ се поклапају.*

Доказ. Како је $\widehat{\mu} = \bar{X}$ оцена за μ развијамо $\sqrt{n}J_{n,a}(\widehat{\mu})$ у околини μ

$$\sqrt{n}J_{n,a}(\widehat{\mu}) = \sqrt{n}J_{n,a}(\mu) + \frac{\partial J_{n,a}(\mu)}{\partial\mu}\sqrt{n}(\widehat{\mu}-\mu) + o(\mu). \tag{4.8}$$

Означимо са $\omega_1 = \omega_1(X_{i_{\pi(1)}}\ldots,X_{i_{\pi(m)}})$ и $\omega_2 = \omega_2(X_{i_{\pi(1)}}\ldots,X_{i_{\pi(m)}})$.

$$\begin{aligned}\frac{\partial J_{n,a}(\mu)}{\partial\mu} &= \frac{1}{n^{[m]}}\sum_{1\le i_1<\cdots<i_m\le n,\pi\in\Pi(m)}\cdots\sum \frac{1}{(a+\frac{\omega_1}{\mu})^2}\frac{\omega_1}{\mu^2} - \frac{1}{(a+\frac{\omega_2}{\mu})^2}\frac{\omega_2}{\mu^2}\\ &= \frac{1}{n^{[m]}}\sum_{1\le i_1<\cdots<i_m\le n,\pi\in\Pi(m)}\cdots\sum \frac{\omega_1}{(a\mu+\omega_1)^2} - \frac{\omega_2}{(a\mu+\omega_2)^2}\end{aligned}$$

Уочавамо да је и $\dfrac{\partial J_{n,a}(\mu)}{\partial\mu}$ облика U - статистике па на основу теореме 1.2.2 важи

$$\frac{\partial J_{n,a}(\mu)}{\partial\mu} \xrightarrow{P} \mathsf{E}\Big(\frac{\partial J_{n,a}(\mu)}{\partial\mu}\Big).$$

С обзиром да важи (4.1) имамо да је $\mathsf{E}\Big(\dfrac{\partial J_{n,a}(\mu)}{\partial \mu}\Big) = 0$. Како на основу централне граничне теореме важи $\sqrt{n}(\widehat{\mu}-\mu) \xrightarrow{D} \mathcal{N}(0,\mu^2)$ применом теореме Слуцког имамо да производ

$$\frac{\partial J_{n,a}(\mu)}{\partial \mu}\sqrt{n}(\widehat{\mu}-\mu) \xrightarrow{D} 0$$

Ако узмемо у обзир последње поменуто асимптотско понашање и да важи $\sqrt{n}J_{n,a}(\mu) \xrightarrow{D} \mathcal{N}(0, m^2\sigma_1^2(a))$ поновном применом теореме Слуцког на (4.8) имамо асимптотску једнакост

$$\sqrt{n}J_{n,a}(\widehat{\mu}) \stackrel{D}{=} \sqrt{n}J_{n,a}(\mu)$$

тј. $\sqrt{n}J_{n,a}(\widehat{\mu}) \xrightarrow{D} \mathcal{N}(0, m^2\sigma_1^2(a))$. ■

Дакле, нулту хипотезу са нивом значајности α одбацићемо ако је

$$\frac{\sqrt{n}}{m\sigma_1(a)}|J_{n,a}| \geq u_{1-\alpha/2}$$

где $u_{1-\alpha/2}$ означава $1-\alpha/2$ квантил нормалне расподеле.

4.2 Тест статистике засноване на карактеризацијама ескпоненцијалне расподеле

У овој глави навешћемо неколико тест статистика заснованих на претходно дефинисаним карактеризацијама експоненцијалне расподеле и користећи претходну главу извести конкретне статистике $J_{n,a}$ на основу њих.

Узимајући у обзир Карактеризацију 3.1.1 нека је $\omega_1 = X_1$ и $\omega_2 = 2\min(X_1, X_2)$. Тада заменом у (4.5) добијамо тест статистику

$$J_{n,a}^D = \frac{\bar{X}}{n(n-1)} \sum\sum_{1\leq i_1<i_2\leq n, \pi\in\Pi(2)} \frac{1}{a\bar{X}+X_{i_{\pi(1)}}} - \frac{1}{a\bar{X}+2\min(X_{i_{\pi(1)}}, X_{i_{\pi(2)}})}. \tag{4.9}$$

Прва пројекција симетризованог језгра

$$\Phi^D(X_1, X_2; a) = \frac{1}{2}\Big(\frac{1}{a\bar{X}+X_1} + \frac{1}{a\bar{X}+X_2} - \frac{2}{a\bar{X}+2\min(X_1,X_2)}\Big)$$

на X_1 под нултом хипотезом је једнака

$$\phi_1^D(s;a) = \mathsf{E}(\Phi^D(X_1,X_2;a)|X_1=s) = \frac{1}{2(a+s)} + \frac{1}{2}e^a\,\mathsf{E}_i(-a)$$
$$- \frac{1}{2(a+2s)}e^{-s}\Big(2 + ae^{a/2+s}\Gamma(0,\tfrac{a}{2}) - ae^{a/2+s}\Gamma(0,\tfrac{a}{2}+s)$$
$$+ 2se^{a/2+s}\Gamma(0,\tfrac{a}{2}) - 2se^{a/2+s}\Gamma(0,\tfrac{a}{2}+s)\Big)$$

где је $\mathsf{E}_i(z) = \int_{-z}^{\infty} u^{-1}e^{-u}du$ експоненцијални интеграл и $\Gamma(a,z) = \int_{-z}^{\infty} t^{a-1}e^{-t}dt$ гама функција.

Користећи Карактеризацију 3.1.2 разматрана тест статистика $J_{n,a}$ је облика

$$J_{n,a}^P = \frac{\bar{X}}{n(n-1)} \sum\sum_{1\le i_1<i_2\le n,\pi\in\Pi(2)} \frac{1}{a\bar{X}+X_{i_{\pi(1)}}} - \frac{1}{a\bar{X}+|X_{i_{\pi(1)}}-X_{i_{\pi(2)}}|} \tag{4.10}$$

а прва пројекција језгра

$$\Phi^P(X_1,X_2;a) = \frac{1}{2}\Big(\frac{1}{a\bar{X}+X_1} + \frac{1}{a\bar{X}+X_2} - \frac{2}{a\bar{X}+|X_1-X_2|}\Big)$$

на X_1 под H_0 је

$$\phi_1^P(s;a) = \mathsf{E}(\Phi^P(X_1,X_2;a)|X_1=s) = \frac{1}{2(a+s)} + \frac{1}{2}e^a\,\mathsf{E}_i(-a)$$
$$- e^{-a-s}(e^{2a}\,\mathsf{E}_i(-a) + \mathsf{E}_i(a) - \mathsf{E}_i(a+s)).$$

Претходно поменуте статистике $J_{n,a}^D$ и $J_{n,a}^P$ су представљене у раду [9]. Поред тих статистика представљамо статистику засновану на Карактеризацији 3.1.3, па је статистика $J_{n,a}$ следећег облика

$$J_{n,a}^S = \frac{\bar{X}}{n(n-1)} \sum\sum_{1\le i_1<i_2\le n,\pi\in\Pi(2)} \frac{1}{a\bar{X}+\max(X_{i_{\pi(1)}},X_{i_{\pi(2)}})} - \frac{1}{a\bar{X}+0.5X_{i_{\pi(1)}}+X_{i_{\pi(2)}}} \tag{4.11}$$

а њена прва пројекција језгра

$$\Phi^S(X_1,X_2;a) = \frac{1}{2}\Big(\frac{2}{a\bar{X}+\max(X_1,X_2)} - \frac{1}{a\bar{X}+0.5X_1+X_2} - \frac{1}{a\bar{X}+X_1+0.5X_2}\Big)$$

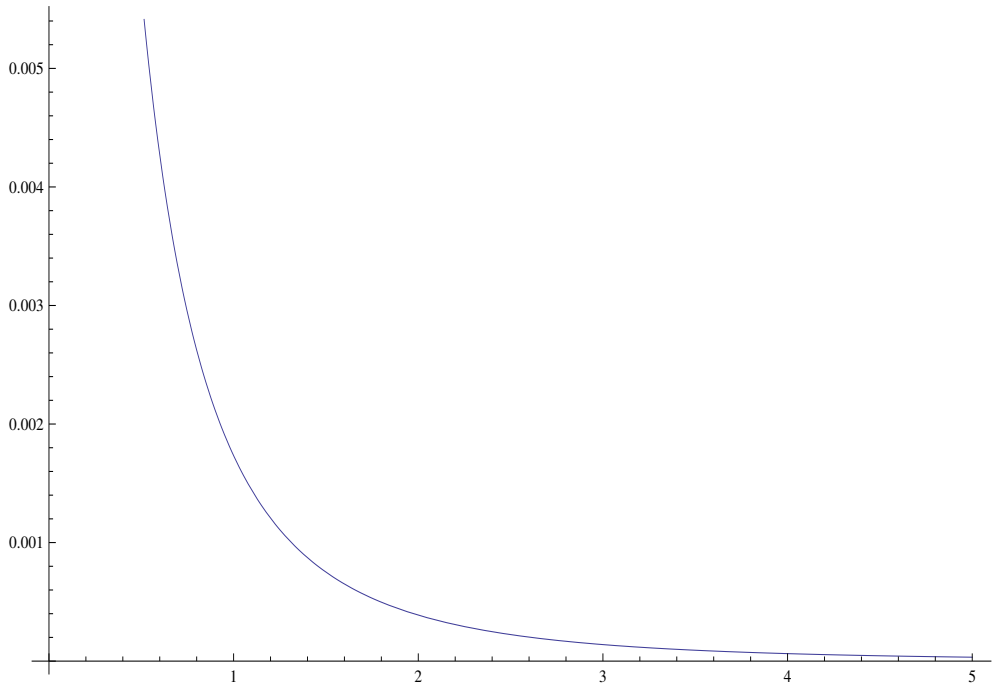
на X_1 под нултом хипотезом је

$$\phi_1^S(s;a) = \mathsf{E}(\Phi^S(X_1,X_2;a)|X_1=s) = \frac{1}{a+s}(1-e^s) + e^a\,\mathsf{E}_i(-a-s)$$
$$- \frac{1}{2}e^{a+s/2}\,\mathsf{E}_i(-a-\tfrac{s}{2}) - e^{2(a+s)}\,\mathsf{E}_i(-2(a+s)).$$

Табела 4.1: Вредности дисперзија првих пројекција $\sigma_D^2(a), \sigma_P^2(a), \sigma_S^2(a)$

a	1	2	5
$\sigma_D^2(a)$	3.52×10^{-3}	6.12×10^{-4}	4.17×10^{-5}
$\sigma_P^2(a)$	6.64×10^{-3}	9.89×10^{-4}	5.61×10^{-5}
$\sigma_S^2(a)$	4.33×10^{-4}	9.69×10^{-5}	8.25×10^{-6}

Уочавамо да су прве пројекције прилично компликованог облика, па није изненађујућа чињеница да дисперзије првих пројекција није могуће приказати у општем облику, али оне ипак могу бити компјутерски израчунате за свако конкретно a. Вредности свих дисперзија за одређено a приказани су у табели 4.1.



Слика 4.1: Функција $4\sigma_S^2(s;a)$ у зависности од слободног параметра a

Пре него што пређемо на практични део и поређење тестова из овог и поглавља 3, уочићемо једну лепу особину наше статистике.

4.3 Утицај параметра a на понашање тест статистика

Пре него што дефинишемо поменуте особине навешћемо теорему која ће нам бити потребна за њихово доказивање.

Теорема 4.3.1 (Абелова теорема за Лапласове трансформације)**.** *Претпоставимо да је $g : [0, \infty) \to \mathbb{R}$ мерљива функција која је интеграбилна на компактном интервалу. Додатно, претпоставимо да је $\int_0^\infty g(t)e^{-at}dt$ коначан за свако $a > 0$. Ако за неко γ и неку реалну константу A важи*

$$\lim_{s\to 0} \Gamma(\gamma + 1)s^{-\gamma} \int_0^s g(t)dt = A \tag{4.12}$$

тада

$$\lim_{a\to\infty} a^\gamma \int_0^\infty g(t)e^{-at}dt = A. \tag{4.13}$$

Имамо да (4.12) важи ако

$$\lim_{s\to 0} \Gamma(\gamma)s^{-(\gamma-1)}g(s) = A. \tag{4.14}$$

Доказ теореме се може наћи у књизи Лапласове трансформације - поглавље 5.1, видети [2].

Теорема 4.3.2. *За фиксирано n, имамо*

$$\lim_{a\to\infty} a^2 \int_0^\infty (L_n^{(1)}(t) - L_n^{(2)}(t))e^{-at}dt = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n \max(Y_i, Y_j) - \frac{3}{2}\bar{Y_n} \tag{4.15}$$

где је $Y_i = X_i/\bar{X}_n$, $1 \le i \le n$ и

$$L_n^{(1)}(t) = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n e^{-t(Y_i/2+Y_j)}$$

$$L_n^{(2)}(t) = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n e^{-t\max(Y_i,Y_j)}.$$

Доказ. Најпре уочимо да су $L_n^{(k)}(t), k = 1, 2$ V - статистике засноване на Карактеризацији 3.1.3. У поглављу 2 смо показали да су V и U - статистике веома сличне, па ради једноставности записа доказаћемо својство за V - статистике. Тврђење важи и за одговарајућу U - статистику.

Означимо са $g(t) = L_n^{(1)}(t) - L_n^{(2)}(t)$ и развијмо функцију $g(t)$ у околини тачке $t = 0$.

$$g(t) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(1-t(\frac{1}{2}Y_i+Y_j)+t^2(\frac{1}{2}Y_i+Y_j)^2-1+t\max(Y_i,Y_j)-t^2\max(Y_i,Y_j)^2+o(t^2)\big)$$

$$= \frac{t}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big((\max(Y_i,Y_j)-(\frac{1}{2}Y_i+Y_j)\big)+\frac{t^2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big((\frac{1}{2}Y_i+Y_j)^2-\max(Y_i,Y_j)^2)+o(t^2)\big)$$

Узимајући у обзир претходни развој функције имамо да важи (4.14)

$$\lim_{s\to 0}\Gamma(2)s^{-(2-1)}g(s) = \lim_{s\to 0}\Gamma(2)s^{-(2-1)}\Big(\frac{s}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(\max(Y_i,Y_j)-(\frac{1}{2}Y_i+Y_j)\big)$$

$$+\frac{s^2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\big((\frac{1}{2}Y_i+Y_j)^2-\max(Y_i,Y_j)^2\big)+o(s^2)\Big) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(Y_i,Y_j)-\frac{3}{2}\bar{Y_n}.$$

На основу Абелове теореме важи

$$\lim_{a\to\infty} a^2\int_0^\infty (L_n^{(1)}(t)-L_n^{(2)}(t))e^{-at}dt = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(Y_i,Y_j)-\frac{3}{2}\bar{Y_n}.$$

■

Теорема 4.3.3. *За фиксирано n важи*

$$\lim_{a\to\infty} a^2\int_0^\infty (L_n^{(1)}(t)-L_n^{(2)}(t))e^{-at}dt = \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\min(Y_i,Y_j)-\bar{Y_n} \tag{4.16}$$

где је $Y_i = X_i/\bar{X}_n$, $1 \le i \le n$ и

$$L_n^{(1)}(t) = \frac{1}{n}\sum_{i=1}^{n}e^{-tY_i}$$

$$L_n^{(2)}(t) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}e^{-2t\min(Y_i,Y_j)}.$$

Теорема 4.3.4. *За фиксирано n, имамо*

$$\lim_{a\to\infty} a^2\int_0^\infty (L_n^{(1)}(t)-L_n^{(2)}(t))e^{-at}dt = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}|Y_i-Y_j|-\bar{Y_n} \tag{4.17}$$

где је $Y_i = \dfrac{X_i}{\bar{X_n}}$, $1 \le i \le n$ и

$$L_n^{(1)}(t) = \frac{1}{n}\sum_{i=1}^{n}e^{-tY_i}$$

$$L_n^{(2)}(t) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} e^{-t|Y_i - Y_j|}.$$

Теореме 4.3.3 и 4.3.4 се доказују на исти начин као теорема 4.3.2, развојем функције $g(t) = L_n^{(1)}(t) - L_n^{(2)}(t)$ у околини тачке нула, и с обзиром да се до тврђења из теорема долази прилично праволинијски у овом раду их нећемо навести.

Како ће нам наведене граничне статистике бити потребне за рачунање емпиријске моћи при различитим алтернативама за њих ћемо увести следеће погодне ознаке:

$$K_n^D = \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \min(Y_i, Y_j) - \bar{Y}_n$$

$$K_n^P = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} |Y_i - Y_j| - \bar{Y}_n$$

$$K_n^S = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \max(Y_i, Y_j) - \frac{3}{2}\bar{Y}_n.$$

Глава 5

Емпиријска моћ

5.1 Монте - Карло методе

Монте - Карло је скуп алгоритама који користе принцип случајног узорковања да би се одредиле нумеричке оцене непознатих параметара. Тиме су омогућена моделовања комплексних ситуација где је велики број променљивих укључен, као и оцењивање потенцијалних ризика. Употреба Монте - Карло методе је веома распрострањена и довела је до запањујућих открића у области физике, теорије игара, финансија и другим гранама науке. Постоји велики спектар Монте - Карло метода, али све оне деле заједничко својство - засноване су на генерисању случајних бројева да би се решили детерминистички проблеми. У овом поглављу навешћемо неколико примера и основне принципе Монте - Карло метода и објаснити на који начин су битне за овај рад.

Концепт Монте - Карло медоте је увео Станислав Улам, пољски математичар и физичар, који је играјући карташку игру „Пасијанс" желио да сазна вероватноћу победе у поменутој игри. Како је рачунање свих могућих добитних комбинација користећи чисту комбинаторику било исувише компликовано Улам се запитао да ли би можда било једноставније одиграти „велики" број партија и на тај начин одредити фреквенцију победе. То математички можемо представити на следећи начин

$$X_i = \begin{cases} 1, & \text{победа} \\ 0, & \text{губитак} \end{cases}$$

где X_i представља i - ту одиграну партију $1 \leq i \leq n$. Нека је $p = P\{X_i = 1\}$ вероватноћа победе, и $1 - p = P\{X_i = 0\}$ вероватноћа губитка. Да бисмо одре-

дили вероватноћу победе у поменутој игри довољно је пребројати све добијене партије и поделити са укупним бројем одиграних партија, прецизније записано $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Применом Чебишовљевог закона великих бројева на S_n имамо да $S_n \xrightarrow{P} \mathsf{E}(X_1) = p$, што је Уламова фреквенција победе.

Пример 5.1.1. *Монте - Карло оцена броја π*
Да би смо одредили вредност броја π (са одређеним степеном поузданости) размотримо игру пикада. Ради једноставности посматрајмо први квадрант јединичног круга. Нека су (A_i, B_i) координате i - тог поготка и дефинишимо X_i, $1 \leq i \leq n$, на следећи начин

$$X_i = \begin{cases} 1, & \sqrt{A_i^2 + B_i^2} \leq 1 \\ 0, & \sqrt{A_i^2 + B_i^2} > 1 \end{cases}$$

Уочавамо да X_i узима вредност 1 уколико смо погодили у површину круга, а 0 у супротном. Како је површина четвртине јединичног круга $\pi/4$ природно је, користећи претходно дефинисано, број π оценити као $\widehat{\pi} = \dfrac{4}{n}\sum_{i=1}^{n} X_i$. Прецизност оцене, већи број тачних цифара, зависи од n, броја покушаја.

Пример 5.1.2. *Монте - Карло интеграција*
Претпоставимо да имамо дефинисану функцију $g(x)$ на ограниченом интервалу (a, b) коју није могуће аналитички једноставно интегралити на (a, b). Такође, додатно претпоставимо да је $0 < g(x) < c$. Да би одредили вредност интеграла $I = \int_a^b g(x)dx$, посматрајмо део координатног система на коме интегралимо функцију. Тај део представља један правоугаоник димензија $c \times (b - a)$. И у овом случају ћемо применити идеју погодака и промашаја. Претпоставимо да бирамо тачку из поменутог правоугаоника, (X_i, Y_i), и дефинишимо Z_i, $1 \leq i \leq n$ тако да узима вредност 1 ако је тачка (X_i, Y_i) испод графика функције $g(x)$, а у супротном 0.

$$Z_i = \begin{cases} 1, & Y_i \leq g(X_i) \\ 0, & Y_i > g(X_i) \end{cases}$$

Дакле, вредност интеграла I можемо да одредимо из пропорције $\frac{I}{c(b-a)} = \dfrac{1}{n}\sum_{i=1}^{n} Z_i$. Добијамо да је оцена површине интеграла Монте - Карло методом, погодака и промашаја, је $\widehat{I} = \dfrac{c(b-a)}{n}\sum_{i=1}^{n} Z_i$.

Сличну идеју Монте - Карло методу погодака и промашаја применићемо на наш пример одређивања моћи теста коју ћемо објаснити у наредној секцији.

5.2 Брзина конвергенције и поређење тестова

У Поглављу 4. у Теореми 4.1.1 смо доказали да асимптотска расподела за $\sqrt{n}J_{n,a}$ је $\mathcal{N}(0, m^2\sigma_1^2)$. Занимало нас је за који цео број n се постиже та асимптотска расподела. Поступак је следећи:

1. За узорак обима n и параметар a одредимо вредност тест статистике $J_{n,a}$ под нултом хипотезом.

2. Поновимо корак 1. N пута.

3. За добијени низ $\sqrt{n}J_{n,a}^1, \sqrt{n}J_{n,a}^2, \ldots, \sqrt{n}J_{n,a}^N$ Колмогоров - Смирновљевим тестом проверимо припадност $\mathcal{N}(0, m^2\sigma_1^2)$.

4. Уколико смо одбацили хипотезу о нормалности повећамо n и поновимо редом Кораке 1, 2 и 3.

5. Поступак понављамо док за неко n не прихватимо хипотезу о нормалности.

Описаним поступком, за $N = 10000$ и са почетном вредношћу $n = 20$, закључујемо да за $n > 300$ важи асимптотско понашање тест статистике $J_{n,a}$.

Како асимптотску расподелу не можемо да користимо за мале вредност n користићемо Монте - Карло методу за рачунање емпиријске моћи и помоћу њих поредити тест статистике. Принцип за одређивање емпиријске моћи теста је следећи:

1. За узорак обима n и параметар a одредимо вредност тест статистике $J_{n,a}$ под нултом хипотезом.

2. Поновимо корак 1. N пута.

3. С обзиром да су од значаја и велике позитивне и негативне вредности и на основу теореме 4.1.1 знамо да је асимптотска расподела за $J_{n,a}$ нормална са очекивањем 0 (расподела је центрирана око нуле), уместо да посматрамо двострану критичну област за добијени низ $J_{n,a}^1, J_{n,a}^2, J_{n,a}^3, \ldots, J_{n,a}^N$, одредићемо за ниво значајности α и низ $|J_{n,a}^1|, |J_{n,a}^2|, |J_{n,a}^3|, \ldots, |J_{n,a}^N|$ десну једностану критичну област, тј. одредићемо $\mu_{1-\alpha}$. (На основу теореме 4.1.1 $\mu_{1-\alpha}$ представља $1-\alpha$ квантил нормалне расподеле.)

4. Поновимо корак 1. и корак 2. под алтернативном хипотезом и добијени низ означимо $J_{n,a}^{1*}, J_{n,a}^{2*}, J_{n,a}^{3*}, \ldots, J_{n,a}^{N*}$.

5. Применом Монте - Карло методе, погодака и промашаја, за свако $J_{n,a}^{i*}, 1 \leq i \leq N$ одредимо $Z_i, 1 \leq i \leq N$ које је дефинисано са

$$Z_i = \begin{cases} 1, & |J_{n,a}^{i*}| > \mu_{1-\alpha} \\ 0, & \text{иначе} \end{cases}$$

6. Емпиријска моћ теста под разматраном алтернативном хипотезом је дефинисана као $\widehat{p}_{moc} = \sum_{i=1}^{N} Z_i/N$.

Претходно описани поступак одређивања емпиријских моћи применимо и за граничне тест стаистике K_n^D, K_n^P и K_n^S дефинисане у Теоремама 4.3.2, 4.3.3 и 4.3.4.

Сличан поступак примењујемо и за тест статистике дефинисане у Поглављу 3, $D_{n,a}, T_{n,a}, L_{n,a}$, са разликом што у том случају не посматрамо апсолутну вредност добијеног низа, јер су већ све вредности добијеног низа позитивне (по теореми 3.2.1 овако дефинисане тест статистике имају χ^2 расподелу). За ниво значајности α одредимо десну једностану критичну област тј. $\mu_{1-\alpha}$, и за $D_{n,a} > \mu_{1-\alpha}, T_{n,a} > \mu_{1-\alpha}, L_{n,a} > \mu_{1-\alpha}$ одбацујемо нулту хипотезу.

Алтернативне расподеле које ћемо размотрити су:

1. Вејбулова расподела са густином

$$g(x, \Theta) = e^{-x^{1+\Theta}}(1+\Theta)x^{\Theta}, \Theta > 0, \quad x \geq 0$$

2. Гама расподела са густином

$$g(x, \Theta) = \frac{x^{\Theta}}{\Gamma(\Theta+1)} e^{-x}, \Theta > 0, \quad x \geq 0$$

3. Полунормална са густином

$$g(x, \Theta) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-x^2/(2\sigma^2)} \quad x > 0$$

4. Расподела линеарне стопе отказа са густином

$$g(x, a, b) = (a + bx)e^{-(ax+(bx^2)/2)}, \quad x \geq 0,\ a > 0,\ b > 0.$$

Табела 5.1: Емпиријске моћи за ниво значајности $\alpha = 0.05$ и $n = 20$

n=20	$\varepsilon(5)$	$\Gamma(2)$	$W(0.8)$	$W(1.4)$	HN	LF(2)	LF(4)
$J^D_{20,0.5}$	5	46	22	26	11	15	23
$J^D_{20,1}$	5	54	21	34	16	20	33
$J^D_{20,2}$	5	56	20	37	19	26	37
$J^D_{20,5}$	5	55	21	40	22	30	44
$J^D_{20,7}$	5	54	22	39	23	31	44
$J^D_{20,9}$	5	55	23	41	23	31	47
$J^P_{20,0.5}$	5	58	17	39	22	29	42
$J^P_{20,1}$	5	57	16	42	24	32	47
$J^P_{20,2}$	5	56	16	41	25	33	47
$J^P_{20,5}$	5	53	19	40	24	33	47
$J^P_{20,7}$	5	53	21	40	25	34	47
$J^P_{20,9}$	5	54	23	42	25	34	49
$J^S_{20,0.5}$	5	58	16	34	17	21	33
$J^S_{20,1}$	5	58	14	39	19	25	39
$J^S_{20,2}$	5	57	14	40	22	30	43
$J^S_{20,5}$	5	56	17	42	23	33	46
$J^S_{20,7}$	5	56	20	42	25	34	47
$J^S_{20,9}$	5	57	22	43	26	34	49
$L_{20,0.5}$	5	53	27	35	16	21	31
$L_{20,1}$	5	54	24	38	18	26	38
$L_{20,2}$	5	54	24	40	22	31	45
$L_{20,5}$	5	45	24	34	20	28	42
$L_{20,7}$	5	43	23	33	20	28	41
$L_{20,9}$	5	41	23	31	18	26	40
$D_{20,0.5}$	5	56	26	38	19	26	38
$D_{20,1}$	5	51	25	39	19	27	39
$D_{20,2}$	5	53	25	40	23	32	46
$D_{20,5}$	5	44	24	33	20	27	41
$D_{20,7}$	5	41	23	31	18	26	39
$D_{20,9}$	5	39	23	30	17	25	39
$T_{20,0.5}$	5	57	26	40	20	27	40
$T_{20,1}$	5	51	25	38	19	27	39
$T_{20,2}$	5	53	25	40	23	32	46
$T_{20,5}$	5	44	24	33	20	27	41
$T_{20,7}$	5	41	23	31	19	26	40
$T_{20,9}$	5	39	23	30	17	25	40
K^D_{20}	5	60	12	48	32	40	56
K^P_{20}	5	60	12	48	32	40	56
K^S_{20}	5	60	12	48	32	40	56

Табела 5.2: Емпиријске моћи за ниво значајности $\alpha = 0.05$ и $n = 50$

n=50	$\varepsilon(5)$	$\Gamma(2)$	$W(0.8)$	$W(1.4)$	HN	LF(2)	LF(4)
$J^D_{50,0.5}$	5	92	42	65	21	32	50
$J^D_{50,1}$	5	94	43	74	31	45	66
$J^D_{50,2}$	5	95	44	78	39	56	76
$J^D_{50,5}$	5	95	48	82	50	66	85
$J^D_{50,7}$	5	94	49	83	53	69	86
$J^D_{50,9}$	5	93	48	81	51	67	87
$J^P_{50,0.5}$	5	95	42	80	43	58	79
$J^P_{50,1}$	5	95	41	80	48	64	84
$J^P_{50,2}$	5	94	43	82	51	68	86
$J^P_{50,5}$	5	93	46	83	55	70	88
$J^P_{50,7}$	5	93	48	83	58	72	89
$J^P_{50,9}$	5	92	47	82	55	71	89
$J^S_{50,0.5}$	5	93	36	72	28	42	63
$J^S_{50,1}$	5	95	38	76	54	52	73
$J^S_{50,2}$	5	95	40	80	44	61	80
$J^S_{50,5}$	5	95	45	83	53	68	86
$J^S_{50,7}$	5	94	47	83	55	70	87
$J^S_{50,9}$	5	93	46	81	53	69	88
$L_{50,0.5}$	5	95	52	78	34	50	70
$L_{50,1}$	5	100	51	79	42	61	81
$L_{50,2}$	5	94	50	81	52	69	86
$L_{50,5}$	5	89	48	79	56	71	88
$L_{50,7}$	5	86	46	76	53	69	87
$L_{50,9}$	5	83	45	73	51	67	87
$D_{50,0.5}$	5	95	52	80	43	60	80
$D_{50,1}$	5	100	51	80	47	67	85
$D_{50,2}$	5	92	49	80	53	70	87
$D_{50,5}$	5	87	48	78	55	70	88
$D_{50,7}$	5	84	46	74	52	68	87
$D_{50,9}$	5	81	45	72	50	66	86
$T_{50,0.5}$	5	95	52	81	45	62	81
$T_{50,1}$	5	100	50	79	47	67	85
$T_{50,2}$	5	92	49	80	54	70	87
$T_{50,5}$	5	87	48	78	55	70	88
$T_{50,7}$	5	84	46	74	52	69	87
$T_{50,9}$	5	82	45	72	50	67	86
K^D_{50}	5	94	37	86	62	78	91
K^P_{50}	5	94	37	86	62	78	91
K^S_{50}	5	94	37	86	62	78	91

5.3 Закључак

У табелама 5.1 и 5.2 приказане су емпиријске моћи за различите алтернативе за ниво значајности $\alpha = 0.05$ и број Монте - Карло симулација $N = 10000$. У табели 5.1 је посматран узорак мањег обима $n = 20$, док је у табели 5.2 узорак обима $n = 50$.

За узорак обима $n = 20$ важи да са повећањем параметра a емпиријске моћи тест статистика дефинисаних у Поглављу 4. су значајно веће у односу на емпиријске моћи тест статистика дефинисаних у Поглављу 3. Такође за мање вредности параметра a тест статистике $L_{n,a}, D_{n,a}$, $T_{n,a}$, $J^D_{n,a}$, $J^P_{n,a}$ и $J^S_{n,a}$ су упоредиве, осим у случају Вејбулове расподеле $W(0.8)$ када су емпиријске моћи тест статистика $L_{n,a}$, $D_{n,a}$, $T_{n,a}$ значајно веће.

Што се тиче емпиријских моћи за узорак обима $n = 50$ за мање вредности параметра a емпиријске моћи тест статистика дефинисаних са $L_{n,a}$, $D_{n,a}$, $T_{n,a}$ су веће од емпиријских моћи тест статистика $J^D_{n,a}$, $J^P_{n,a}$ и $J^S_{n,a}$. Са повећањем параметра a емпиријске моћи $J^D_{n,a}$, $J^P_{n,a}$ и $J^S_{n,a}$ брже расту од емпиријских моћи тест статистика $L_{n,a}$, $D_{n,a}$, $T_{n,a}$ (важи да за $L_{n,a}$, $D_{n,a}$, $T_{n,a}$ у случају расподела $\Gamma(2)$, $W(0.8)$, $W(1.4)$ емпиријске моћи опадају са порастом a).

Уочавамо да у обе табеле, 5.1 и 5.2, емпиријске моћи граничних статистика дефинисах са K^D_n, K^P_n, K^S_n су значајно веће у односу на емпиријске моћи свих других тест статистика (осим у случају $W(0.8)$ расподеле).

На основу представљених резултата закључујемо да су све поменуте статистике упоредиве, и да њихова емпиријска моћ зависи од обима узорка n као и од слободног параметра a. У случају мањег узорка предложене тест статистике $J^D_{n,a}$, $J^P_{n,a}$ и $J^S_{n,a}$ имају већу моћ са порастом параметра a, док за узорак обима $n = 50$ моћ зависи од конкретне алтернативе и избора параметра. Ипак, тест статистике K^D_n, K^P_n, K^S_n, које су слободне од параметра a, су се показале као најбоље ако као меру узимамо њихову емпиријску моћ.

Литература

[1] B. Milošević. Asimptotska svojstva neparametarskih testova zasnovanih na U - statistikama i V - statistikama sa nedegenerisanim i slabo degenerisanim jezgrom. *Diss. Univerzitet u Beogradu - Matematički fakultet*, 2016.

[2] D. V. Widder. *The Laplace Transform.* Princeton university press, 1946.

[3] E. L. Lehmann. *Elements of Large Sample Theory.* Springer, 1999.

[4] J. L. Teugels S. Csorgo. Empirical Laplace Transform and Approximation of Compound Distributions. *Applied Probability Trust*, 27(1):88–101, 1990.

[5] L. Glavaš S. Janković. *Stohastički modeli u operacionim istraživanjima* . Univerzitet u Beogradu - Matematički fakultet, 2016.

[6] L. Wasserman. *All of Nonparametric Statistics.* Springer, 2005.

[7] M. Cuparić B. Milošević, M. Obradović. New L^2 - type exponentiality tests. *SORT: Statistics and Operations Research Transcations*, 43(1):25–50, 2019.

[8] M. Obradović. Karakterizacija nekih raspodela i Bahadurova efikasnost testova saglasnosti. *Diss. Univerzitet u Beogradu - Matematički fakultet*, 2015.

[9] M. Obradović B. Milošević. New class of exponentiality tests based on U - empirical Laplace transform. *Statistical Papers*, 57(4):977–990, 2016.

[10] M. Obradović B. Milošević. Some characterizations of the exponential distribution based on order statsitics. *AADN: Applicable Analysis and Discrete Mathematics*, 10:394–407, 2016.

[11] N. Gurtler L. Baringhaus and N. Henze. Weighted integral test statistics and components of smooth tests of fit. *Aust. N. Z. J. Statist.*, 42:179–192, 2000.

[12] N. Henze L. Baringhaus. A class of consistent tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Math.*, 43(3):551–564, 1990.

[13] P. Mladenović. *Verovanoća i Statistika.* Univerzitet u Beogradu - Matematički fakultet, 2008.

[14] S. G. Meintanis N.Henze. Tests of it for Exponentiality based on the empirical Laplace Transform. *Statistics: A Journal of Theoretical and Applied Statistics*, 36(2):147–161, 2010.

[15] W. Hoeffding. The strong law of large number for U - statistics. *North Carolina State University. Dept. of Statistics*, 1961.

Биографија аутора

Марија, Обрен, Живановић (*Фоча, 02. август 1993.*) је мастер студент Математичког факултета у Београду. Године 2016. дипломирала је на Математичком факултету, Универзитета у Београду, са просечном оценом 9.21 од 10. Претходно је основно образовање стекла у О.Ш. „Веселин Маслеша" Фоча и општу гимназију завршила у Фочи, као одличан ђак. Током средњошколског образовања била је члан ликовне групе „Ars in medias res" . Тренутно запослена у „Беолинк" на позицији софтвер инжињера.