УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Марија Д. Цупарић

ТЕСТОВИ САГЛАСНОСТИ ЗАСНОВАНИ НА L^2 и L^∞ РАСТОЈАЊИМА И ЊИХОВА АСИМПТОТСКА ЕФИКАСНОСТ

— Докторска дисертација —

Београд, 2021.

UNIVERSITY OF BELGRADE
FACULTY OF MATHEMATICS

Marija D. Cuparić

GOODNESS-OF-FIT TESTS BASED ON L^2 AND L^∞ DISTANCES AND THEIR ASYMPTOTIC EFFICIENCY

— Doctoral Dissertation —

Belgrade, 2021.

Ментор:

др Бојана Милошевић, доцент
Универзитет у Београду, Математички факултет

Чланови комисије:

проф. др Павле Младеновић, редовни професор
Универзитет у Београду, Математички факултета

проф. др Милан Меркле, редовни професор у пензији
Универзитет у Београду, Електротехнички факултета

др Марко Обрадовић, доцент
Универзитет у Београду, Математички факултета

др Бојана Милошевић, доцент
Универзитет у Београду, Математички факултет

Датум одбране:

Захвалница

Током рада на овој дисертацији имала сам срећу да будем окружена људима који су радо улагали своје време у моје истраживање, давали корисне савете и улагали своје знање у превазилажењу препрека на које сам наилазила.

Посебно бих желела да се захвалим ментору доц. др. Бојани Милошевић на несебичној помоћи, залагању и активном учествовању током израде дисертације. Такође, захваљујем се на томе што је личним примером показала са коликом преданошћу се треба бавити науком и за знање које сам стекла у раду са њом. Велику захвалност дугујем и доц. др Марку Обрадовићу на корисним саветима који су допринели да се квалитет дисертације побољша и који је увек био спреман да одговори на сва моја питања. Захваљујем се и проф. др Павлу Младеновићу и проф. др Милану Мерклеу који су детаљним читањем дисертације и сугестијама допринели њеном коначном облику. Захвалност дугујем и Благоју Ивановићу за несебичну помоћ у реализацији техничких решења која су била неопходна за добијање резултата приказаних у дисертацији.

Захваљујем се оцу, мајци и сестри на сталној подршци током живота. Захваљујем се пријатељима који су веровали у мене и били ми велика подршка током израде дисертације. Највећу захвалност дугујем супругу Тому за неизмерну љубав, разумевање и подршку коју ми је пружио и тиме ми помогао да превазиђем све тешкоће.

Марија Цупарић

Наслов дисертације: *Тестови сагласности засновани на* L^2 *и* L^∞ *растојањима и њихова асимптотска ефикасност*

Резиме: Циљ дисертације је конструкција нових тестова сагласности, анализа њихових својстава, као и добијање нових теоријских резултата везаних за граничне расподеле слабо дегенерисаних V-статистика са оцењеним параметрима. Нови тестови сагласности се базирају на карактеризацијама једнаке расподељености две функције узорка. Тест статистике су формиране као L^2 растојања V-емпиријских функција расподела статистика из карактеризације, односно као L^2 и L^∞ растојања V-емпиријских Лапласових трансформација тих статистика. У другом случају резултујуће статистике се могу посматрати као V-статистике са оцењеним параметром или функције таквих статистика.

До сада су били познати гранични резултати за недегенерисане V-статистике са оцењеним параметрима и слабо дегенерисане V-статистике са оцењеним параметрима у случају језгра реда два. У дисертацији су изведени гранични резултати за одговарајућу класу слабо дегенерисаних V-статистика са оцењеним параметром реда m, где је m паран број. Захваљујући њима, одређена су асимптотска својства предложених тестова. Да би се проверио квалитет ових тестова, одређене су емпиријске моћи помоћу Монте Карло симулација, као и приближна Бахадурова ефикасност. Приказани су нови резултати у вези са приближном Бахадуровом ефикасношћу у случају блиских алтернатива, која је примењива и када гранична расподела статистике при нултој хипотези није нормална. У наведеном смислу, урађено је и поређење великог броја тестова како класичних, тако и неких тестова развијених у последње време.

Сви претходно поменути резултати су добијени за потпуне узорке. Предложена је додатно и модификација тестова када су подаци случајно цензурисани. У том случају, за апроксимацију $p-$вредности предложен је нови бутстреп метод, чије је коришћење теоријски оправдано.

Кључне речи: $V-$статистике, $V-$емпиријски процеси, слабо дегенерисано језгро, карактеризације расподела, Лапласова трансформација, Бахадурова асимптотска ефикасност, цензурисани подаци, скалирање инверзом вероватноће цензурисања, бутстреп

Научна област: Математика

Ужа научна област: Вероватноћа и статистика

АМС 2010 класификација: 62E20, 62G10, 62G20, 62N01, 62N03

Dissertation title: *Goodness-of-fit tests based on L^2 and L^∞ distances and their asymptotic efficiency*

Abstract: The goal of this dissertation is the construction of new goodness-of-fit tests, analysis of their properties, as well as to obtain new theoretical findings regarding the limiting distributions of weakly degenerated $V-$statistics with estimated parameters. New goodness-of-fit tests are based on equidistributional type characterizations of two sample functions. Test statistics are formed as L^2 distances between $V-$empirical distribution functions of statistics from characterization, and also as L^2 and L^∞ distances between $V-$empirical Laplace transformations of those statistics. In the latter case, resulting test statistics can be observed as $V-$statistics with an estimated parameter or as functions of those statistics.

Until now, limiting results were known for non-degenerate $V-$statistics with estimated parameters, as well as for weakly degenerate $V-$statistics of degree two with estimated parameters. Limiting results for the appropriate class of weakly degenerate $V-$statistics with an estimated parameter of degree m, where m is even number, are derived in this dissertation. Owing to these results, asymptotic properties for presented tests are determined. To assess the quality of these tests, empirical powers were determined using Monte Carlo simulations, as well as approximate Bahadur efficiency. New results are presented regarding the approximate Bahadur efficiency in case of close alternatives, which is applicable also when the limiting distribution of statistics under the null hypothesis is not normal. In this sense, the comparison between many tests is performed, both classical tests and recently developed tests.

All previously mentioned results were obtained for complete samples. Additional, modification of previously introduced tests for randomly censored data was also proposed. In such a case, the new theoretically justified bootstrap method is proposed for the approximation of $p-$value.

Keywords: V-statistics, $V-$empirical processes, weakly degenerate kernel, characterization of distributions, Laplace transform, Bahadur asymptotic efficiency, censored data, inverse probability of censored weight, bootstrap

Scietific Area: Mathematics

Scietific Sub-area: Probability and Statistics

AMS 2010 Classification: 62E20, 62G10, 62G20, 62N01, 62N03

Садржај

Глава 1

УВОД

Од најранијих почетака статистике, статистичари су почињали своју анализу претпостављајући расподелу података и затим проверавали да ли је та претпоставка адекватна. Током времена је предложен велики број статистичких процедура за проверу сагласности са расподелом од којих највећи део чине статистички тестови сагласности.

У општем случају, тестира се нулта хипотеза H_0 да узорак прати задату расподелу са функцијом расподеле F. У наставку под узорком ће се подразумевати низ независних и једнако расподељених случајних величина. Нулта хипотеза може бити проста, када је функција расподеле F у потпуности одређена, или сложена, када је расподела одређена до на непознати параметар. Са друге стране, у већини ситуација алтернативна хипотеза је сложена и често само претпоставља да нулта хипотеза није тачна. С обзиром да је алтернативна хипотеза често сувише општа врло ретко се може рећи да је нека процедура тестирања сагласности супериорнија над другим процедурама. У случају да је алтернативна расподела на неки начин одређена, тестови се формирају тако да буду осетљиви на својства расподеле која се јавља у алтернативној хипотези.

Током времена настао је велики број тестова сагласности заснованих на различитим својствима. Први такав објављен тест је Пирсонов χ^2 тест (*Pearson (1900)*), који је и данас у широкој употреби. Иако је овај тест често мање моћан од других класа тестова, он се може применити и на дискретне и на непрекидне, као и једнодимензионалне и вишедимензионалне податке. Три најчешће коришћена теста сагласности су Колмогоров-Смирновљев тест, Крамер-фон Мизесов тест и Андерсон-Дарлингов тест (детаљно размотрени у *Anderson and Darling (1952)*). Ти тестови су засновани на емпиријској функцији расподеле и могу се прилагодити за тестирање различитих расподела. У последње време појавио се велики број тестова намењених за тестирање конкретне нулте хипотезе који користе својства претпостављене расподеле (видети на пример *Henze and Meintanis (2005), Obradović (2015a), Obradović et al. (2015), Nikitin (2017), Miloševic (2017), Torabi et al. (2018), Henze and Koch (2020), Jiménez-Gamero et al. (2020)*).

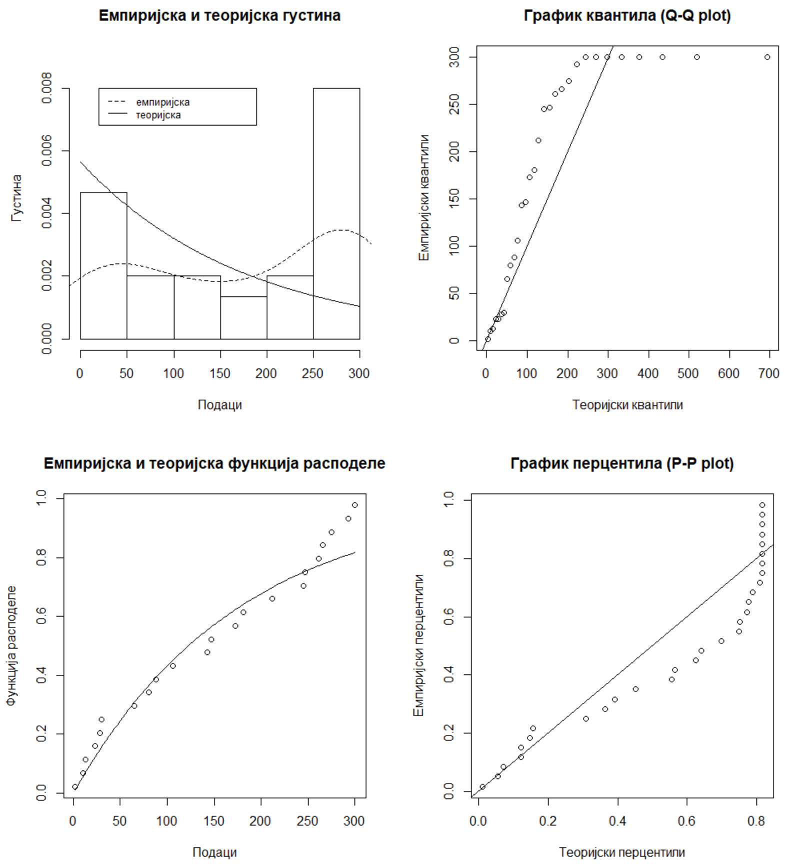
Када је доступан велики број тестова за тестирање одређене нулте хипотезе, потребно је наћи рационалан разлог за коришћење неког од њих. Значај избора одговарајућег теста од великог броја расположивих тестова може се видети на следећем примеру.

Пример 1.1 *Размотримо времена отказивања и времена рада, оних елемената који нису отказали, за узорак уређаја великог електричног система (Meeker and Escobar (2014)). У одређеном тренутку, 30 јединица је пуштено у рад при нормалним условима. За сваку од јединица бележена су времена отказа, ако је до отказа дошло, као и узрок отказа. Узроци отказа су подељени у две групе: прву где су кварови настали нагомилавањем случајно насталих оштећења услед високих напона што је узроковало отказ појединачне незаштићене компоненте; и другу где се откази појављују након 100000 циклуса коришћења, што је узроковано нормалним током производње. Наравно, постоје и оне јединице које су и даље оперативне након 300000 циклуса. У наставку, ми ћемо само разматрати времена отказа, односно времена рада, без посматрања узрока, па су у табели 1.1 приказани обједињени подаци.*

Табела 1.1: Времена отказивања и времена рада електричног система

275	13	147	23	181	30	65	10	300	173
106	300	300	212	300	300	300	2	261	293
88	247	28	143	300	23	300	80	245	266

Желимо да тестирамо да ли су подаци експоненцијално расподељени. Пре формалног тестирања података, можемо посматрати графички приказ.



Слика 1.1: Графици времена отказивања и времена рада електричног система

На слици 1.1 приказане су четири врсте графика података. Први график приказује хистограм и емпиријску густину података, као и теоријску густину расподеле из нулте хипотезе. Други график је график теоријских и емпиријских квантила. Трећи график даје емпиријску и теоријску функцију расподеле, док четврти график даје емпиријске и теоријске перцентиле. Оно што се може уочити са графика јесте да подаци не прате експоненцијалну расподелу.

Наведену хипотезу ћемо тестирати уз помоћ неколико опште познатих и често коришћених тестова експоненцијалности, њихове формуле као и референце на радове у којима су предложени биће приказане у поглављу 4. У табели 1.2 приказане су $p-$вредности тих тестова експоненцијалности.

Табела 1.2: $p-$вредности тестова

Статистика	EP	CO	G_n	MO	ω^2
p-вредност	0.0583	0.3086	0.015	0.0200	0.0059

Дакле, можемо закључити да у овом случају различити тестови дају различите закључке у вези са тим да ли прихватити или одбацити нулту хипотезу.

Са друге стране, ако посматрамо само истраживање, познато је да су машине посматране до тренутка отказа или док се није навршило 300000 циклуса рада. Дакле, нису познати тачни тренуци отказа свих 30 јединица. На тај начин дошло је до цензурисања типа I, односно цензурисања унапред познатом фиксираном вредношћу. То је један могући разлог за различите закључке тестова, али и показатељ да процедуре тестирања треба прилагодити таквом истраживању.

Као што смо видели у примеру, као додатак формалном тестирању хипотеза често се користе и графичке технике. Односно цртају се одговарајући графици на којима се може уочити подударање или одступање података од претпостављене расподеле као одређена карактеристика на графику.

У наставку ће бити представљене класе тест статистика засноване на L^2 и L^∞ растојањима између одговарајућих случајних величина. С обзиром да се ове тест статистике могу написати као функције $V-$статистика или $V-$емпиријских процеса, прво ће бити приказана теорија $V-$статистика из које ће следити асимптотска својства предложених тест статистика. За проверу квалитета ових тестова, поред емпријских моћи тестова, биће приказана и асимптотска ефикасност како нових тако и неких раније развијених тестова. Ова два критеријума доприносе избору одговарајућег теста када је велики број тестова на располагању.

Чак и када се изабере најбољи тест (по одређеном критеријуму), ако се примени на податке који су цензурисани, резултати таквог тестирања могу бити погрешно протумачени. Због тога, процедуре које се користе на потпуним подацима треба прилагодити цензурисаним подацима. Последњи део дисертације је посвећен прилагођавању нових тестова цензурисаним подацима, одређивању њихових асимптотских својстава и одређивању емпријских моћи тих тестова.

Глава 2

ТЕОРИЈА $V-$СТАТИСТИКА

Приликом формирања тест статистике теста сагласности са неком расподелом посебно је погодно ако тест статистика припада некој већој класи статистика чија су својства позната. Пример такве класе је класа $V-$статистика чија су асимптотска својства доста истражена. Заправо, испоставља се да велики број тест статистика тестова који се примењују за различите расподеле се може написати у облику $V-$статистике или неког функционала од $V-$емпиријског процеса. Први резултати који су били познати за ову класу статистика (видети на пример *Korolyuk and Borovskikh (1994)* или *Serfling (2009)*), омогућавали су познавање асимптотских својстава великог броја тест статистика за тестирање простих хипотеза.

Са друге стране, 80-тих година XX века долази се до првих својстава $V-$статистика са оцењеним параметром (видети *Randles (1982)*), које заправо одговарају тест статистикама којима се тестирају сложене хипотезе.

У наставку биће уведен појам статистичког функционала, чији специјалан случај представљају $V-$статистике. Затим ће бити приказано како су се јавиле прве $V-$статистике, као и асимптотска својства $V-$статистика у недегенерисаном и слабо дегенерисаном случају. У последњем делу поглавља дата је теорија $V-$емпиријских функција.

2.1 Статистички функционали

Размотримо функционал T дефинисан на скупу $\mathcal{F}$ функција расподела дефинисаних на $\mathbb{R}$, односно реално вредносне функције $T(F),\ F \in \mathcal{F}$. Претпоставимо да желимо да оценимо тај функционал на основу узорка $X_1, ..., X_n$ из расподеле са функцијом расподеле F.

Формулом

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{I}\{\mathrm{X_i} \leqslant \mathrm{x}\},$$

где је I индикатор случајна промењива, се дефинише емпиријска функција расподеле. Следећа теорема даје својства емпиријске функције расподеле. За више детаља видети *Wasserman (2006)*.

Теорема 2.1 *Нека је $X_1, ..., X_n$ узорак из расподеле F и нека је F_n емпиријска функција расподеле.*

1. *За било коју фиксирану вредност x, $E(F_n(x)) = F(x)$ и $D(F_n(x)) = \frac{F(x)(1-F(x))}{n}$. На основу тога, за средње квадратну грешку важи $E(F_n(x) - F(x))^2 = \frac{F(x)(1-F(x))}{n} \to 0$ и $F_n(x) \xrightarrow{P} F(x)$.*

2. *(Гливенко-Кантелијева теорема) $\sup\limits_x |F_n(x) - F(x)| \xrightarrow{ss} 0$.*

3. *(Дворецки-Кифер-Волфовицова неједнакост) За свако $\varepsilon > 0$*

$$P\Big\{\sup_x |F(x) - F_n(x)| > \varepsilon\Big\} \leqslant 2e^{-2n\varepsilon^2}.$$

Дакле, захваљујући својствима емпиријске функције расподеле из претходне теореме, као оцена раније поменутог функционала $T(F)$ може се размотрити функција у којој је теоријска расподела замењена емпиријском функцијом расподеле, односно $T(F_n)$. Ову оцену називамо статистичким функционалом. Поставља се питање да ли се таква оцена може сматрати добром у смислу да је (асимптотски) непристрасна и постојана оцена. Размотримо два примера.

Пример 2.1 *Нека је функционал $T(F) = \int x dF(x)$, односно математичко очекивање расподеле. Одговарајући статистички функционал је*

$$T(F_n) = \int x dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n.$$

Познато је да је ова оцена непристрасна и постојана оцена наведеног функционала. Такође, $\sqrt{n}(\overline{X} - T(F))$ конвергира у расподели ка случајној величини са нормалном расподелом са очекивањем 0 и дисперзијом $D(X_1)$.

Пример 2.2 *Нека је μ_k $k-$ти центрирани моменат од F, односно $\mu_k = E(X - E(X))^k$. Тада је*

$$\begin{aligned} T(F_n) = M_k = \int \Big(x - \int x dF_n(x)\Big)^k dF_n(x) &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^k \\ &= \frac{1}{n}\sum_{i=1}^{n} X_i^k - k\overline{X}_n \frac{1}{n}\sum_{i=1}^{n} X_i^{k-1} + \cdots + (-\overline{X}_n)^k. \end{aligned}$$

Баз губитка општости, пошто расподела од M_k не зависи од $E(X_1)$ (другоразредна статистика за тај параметар), можемо претпоставити да је $E(X_1) = 0$. Након рачуна, добија се да је

$$E(M_k) = \mu_k + \frac{1}{n}\left(\binom{k}{2}\mu_{k-2}\mu_2 - k\mu_k\right) + O_p\Big(\frac{1}{n^2}\Big),\ n \to \infty.$$

Слично, за дисперзију се добија да је

$$D(M_k) = \frac{1}{n}(\mu_{2k} - 2k\mu_{k-1}\mu_{k+1} - \mu_k^2 + k^2\mu_2\mu_{k-1}^2) + O_p\left(\frac{1}{n^2}\right),\ n \to \infty.$$

Дакле, може се закључити да је M_k асимптотски непристрасна и постојана оцена функционала μ_k.

Такође, може се показати да је гранична расподела од $\sqrt{n}(M_k - \mu_k)$ нормална расподела са очекивањем 0 и дисперзијом $\mu_{2k} - 2k\mu_{k-1}\mu_{k+1} - \mu_k^2 + k^2\mu_2\mu_{k-1}^2$ (за више детаља видети Lehmann (2004)).

Статистички функционали који се јављају у претходним примерима су (асимптотски) непристрасни и постојани. Такође, у оба примера гранична расподела је нормална. Међутим, с обзиром да је скуп свих статистичких функционала широк, ове особине се не могу лако уопштити на све статистичке функционале. Због тога ћемо се у наставку бавити одређеном класом функционала и њима одговарајућим статистичким функционалима.

2.2 $V-$статистике

Размотримо класу функционала који се могу представити као очекивање одређене функције Φ. Такви функционали се дефинишу на следећи начин.

Дефиниција 2.1 *Нека је $X_1, ..., X_n$ узорак из расподеле F. Пресликавање $T : F \to E(\Phi(X_1, ..., X_m))$ се назива функционал очекивања, где је $\Phi : \mathbb{R}^m \to \mathbb{R}$ мерљива функција симетрична по својим аргументима. Функција Φ се назива језгро функционала.*

Најмање m за које постоји пресликавање T се назива редом језгра Φ. Без губитка општости увек се може претпоставити да је језгро симетрична функција по својим аргументима. Уколико није, можемо је заменити симетричном функцијом $\Phi^\star(x_1, ..., x_m)$ дефинисаном са

$$\Phi^\star(x_1, ..., x_m) = \frac{1}{m!} \sum_{\pi \in \Pi(m)} \Phi(x_{\pi(1)}, ..., x_{\pi(m)}),$$

где је $\Pi(m)$ скуп свих пермутација скупа $\{1, 2, ..., m\}$. У наставку ће се претпостављати да је језгро симетрично.

Размотримо сада оцену функционала очекивања која се добија једноставном заменом функције F њеном оценом, односно одговарајући статистички функционал.

Дефиниција 2.2 *Нека је $X_1, ..., X_n$ узорак из расподеле F. V-статистика реда m са језгром Φ је статистички функционал облика*

$$V_n = \int \cdots \int \Phi(x_1, ..., x_m) dF_n(x_1) \cdots dF_n(x_m) = \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \Phi(X_{i_1}, ..., X_{i_m}), \tag{2.1}$$

где је $\Phi : \mathbb{R}^m \to \mathbb{R}$ мерљива функција.

Прве $V-$статистике (или фон Мизесови функционали) су предложене 1947. године у раду фон Мизеса (*von Mises (1947)*) као статистички функционали једноставнијег облика којима се уз помоћ Тејлоровог развоја могу апроксимирати диференцијабилни статистички функционали.

Асимптотска расподела диференцијабилног статистичког функционала $T(F_n)$ зависи највише од понашања функционала $T(\cdot)$ у тачки F. Односно, тип асимптотске расподеле диференцијабилног статистичког функционала $T(F_n)$ зависи од тога који је први члан различит од нуле у Тејлоровом развоју од $T(\cdot)$ у тачки F. Због тога, посматра се гранична расподела од $n^{\frac{r}{2}}(T(F_n)-T(F))$, где је r ред првог ненула елемента Тејлоровог развоја од $T(\cdot)$.

Да би се израчунали елементи Тејлоровог развоја, потребно је увести појам Гатоовог извода или, како се ређе назива, фон Мизесовог извода. Ова два извода, иако врло слична, у општем случају нису идентични, мада се у случају рада са функционалима могу изједначити (за више детаља видети *Fernholz (2012)*). Ако постоји гранична вредност

$$d_1T(F;F_n-F)=\frac{d}{d\lambda}T(F+\lambda(F_n-F))-T(F)|_{\lambda=0+}=\lim_{\lambda\to0+}\frac{T(F+\lambda(F_n-F))-T(F)}{\lambda},$$

она се назива Гатоов извод првог реда од T у тачки F у правцу F_n. У општем случају, Гатоов извод $k-$тог реда од T у тачки F у правцу F_n се дефинише са

$$d_kT(F;F_n-F)=\frac{d^k}{d\lambda^k}T(F+\lambda(F_n-F))-T(F)|_{\lambda=0+},$$

ако гранична вредност постоји.

Применом Тејлоровог развоја на $T(F_n)-T(F)$ следи

$$T(F_n)-T(F)=\sum_{j=1}^{r}\frac{1}{j!}d_jT(F;F_n-F)+R_{rn}=V_{rn}+R_{rn}, \tag{2.2}$$

за одређено r, где је V_{rn} $V-$статистика а R_{rn} остатак. За одређивање граничне расподеле од $n^{\frac{r}{2}}(T(F_n)-T(F))$, потребно је одредити ранг статистике V_{rn}, односно најмање r тако да V_{rn} није дегенерисана случајна величина, и показати да $n^{\frac{r}{2}}R_{rn}\xrightarrow{p}0$. Након тога се сва асимтотска својства статистичког функционала $T(F_n)$ добијају из асимптотских својстава $n^{\frac{r}{2}}V_{rn}$. У наставку ћемо разматрати понашање статистика V_{rn}, док се доказ да $n^{\frac{r}{2}}R_{rn}\xrightarrow{p}0$ може наћи у *von Mises (1947)* или *Serfling (2009)* (поглавље 6.2.2).

У општем случају, Гатоов извод $k-$тог реда $d_kT(F;F_n-F)$ је $k-$линеаран ако постоји функција $\Phi_F(x_1,..,x_k)$, $(x_1,...,x_k)\in\mathbb{R}^k$, таква да важи

$$\begin{aligned}d_kT(F;F_n-F)&=\int\cdots\int\Phi_F(x_1,...,x_k)\prod_{i=1}^{k}d(F_n(x_i)-F(x_i))=\int\cdots\int\widetilde{\Phi}_F(x_1,...,x_k)\prod_{i=1}^{k}dF_n(x_i)\\&=\frac{1}{n^k}\sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}\widetilde{\Phi}_F(X_{i_1},...,X_{i_k}),\end{aligned}$$

где је

$$\widetilde{\Phi}_F(x_1,...,x_k)=\Phi_F(x_1,..,x_k)-\sum_{i=1}^{k}\int\Phi_F(x_1,...,x_k)dF(x_i)+\sum_{1\leqslant i<j\leqslant k}\iint\Phi_F(x_1,...,x_k)dF(x_i)dF(x_j)$$
$$+\cdots+(-1)^k\int\cdots\int\Phi_F(x_1,...,x_k)\prod_{i=1}^{k}dF(x_k).$$

На основу тога, и из чињенице да је први члан суме различит од нуле у (2.2) $r-$ти, следи да је

$$V_{rn}=\sum_{j=1}^{r}\frac{1}{j!}d_jT(F;F_n-F)=\frac{1}{n^r}\sum_{i_1=1}^{n}\cdots\sum_{i_r=1}^{n}\widetilde{\Phi}_F(X_{i_1},...,X_{i_r}). \tag{2.3}$$

Можемо приметити да израз (2.3) има исти облик као (2.1). Ове статистике су назване $V-$статистикама јер их је фон Мизес (*von Mises*) први увео.

Најзначајнији случајеви су када је $r=1$, тада је у питању недегенерисана статистика, и када је $r=2$, тада је у питању слабо дегенерисана статистика. У претходном случају, који је разматрао фон Мизес, ранг $V-$статистике и њен ред се поклапају. Тада, асимптотска расподела од V_{1n} директно следи из централне граничне теореме, пошто је ова статистика у ствари узорачка средина. Асимптотску расподелу у случају дегенерисане $V-$статистике реда 2 је фон Мизес показао у свом раду (*von Mises (1947)*). У општем случају ред и ранг статистике не морају да се поклапају, па ће сва остала тврђења бити приказана у општем случају.

Годину дана након што је фон Мизес увео $V-$статистике, Хефдинг је објавио рад *Hoeffding (1948)* у којем је увео нови тип статистика, које су назване $U-$статистике, и одредио њихову асимптотску расподелу. За $V-$статистику дефинисану у (2.1), одговарајућа $U-$статистика је одређена са

$$U_n=\frac{1}{\binom{n}{m}}\sum_{1\leqslant i_1<...<i_m\leqslant n}\Phi(X_{i_1},...,X_{i_m}). \tag{2.4}$$

Ако је $m=1$, $U-$ и $V-$статистика се поклапају. Поред тога, у истом раду Хефдинг је дао везу између $U-$ и $V-$статистика. У општем случају, важи Хефдингова репрезентација $V-$статистика преко $U-$статистика

$$n^mV_n=m!\binom{n}{m}U_n+\sum_{(i_1,...,i_m)}\Phi(X_{i_1},...,X_{i_m}), \tag{2.5}$$

где је сума узета по свим $m-$торкама $(i_1,...,i_m)$ таквим да важи $i_k=i_l,\ k\neq l$, за бар један пар. Број елемената у тој суми је реда n^{m-1}, а веза (2.5) не зависи од ранга језгра. У општем случају, најједноставнији начин за одређивање асимптотске расподеле $V-$статистика је коришћење асимптотске расподеле $U-$статистика и наведене везе.

2.3 Асимптотска својства $V-$статистика

Ако је ред језгра $m = 1$, $V-$статистика је одређена са

$$V_n = \frac{1}{n}\sum_{i=1}^{n}\Phi(X_i).$$

У том случају је $E(V_n) = T(F)$, где је $T(F)$ одговарајући функционал. Ако је дисперзија језгра коначна, из централне граничне теореме ће следити асимптотска расподела ове статистике. Међутим, у случају када је $m > 1$, ситуација је компликованија, пре свега јер чланови суме више нису независни међусобно. Размотримо следећи пример.

Пример 2.3 *Нека је $m = 2$ и $\Phi(x_1, x_2) = |x_1 - x_2|$. Тада је одговарајућа $V-$статистика дефинисана са*

$$V_n = \frac{1}{n^2}\sum_{i,j=1}^{n}\Phi(X_i, X_j) = \frac{1}{n^2}\sum_{i,j=1}^{n}|X_i - X_j|.$$

Према томе, важи

$$E(V_n) = E\bigg(\frac{1}{n^2}\sum_{i,j=1}^{n}|X_i - X_j|\bigg) = E\bigg(\frac{1}{n^2}\sum_{i\neq j}|X_i - X_j| + \frac{1}{n^2}\sum_{i=j}|X_i - X_j|\bigg).$$

Пошто је $|X_i - X_j| = 0$ увек када је $i = j$, следи да је

$$E(V_n) = \frac{1}{n^2}n(n-1)E(|X_1 - X_2|) = \frac{n-1}{n}T(F),$$

чиме се добија да оцена није непристрасна, али јесте асимптотски непристрасна. $\triangle$

Дакле, у општем случају $V-$статистика није непристрасна оцена функционала. С обзиром да се пристрасност добија због елемената који се дуплирају у суми, једна могућност да би се редуковала пристрасност јесте да се сумира само по различитим индексима. То доводи до $U-$статистика, поменутих у претходном поглављу. За $V-$статистику дефинисану у (2.1), одговарајућа $U-$статистика је дата у (2.4). Да бисмо видели како су повезане $U-$ и $V-$статистике, размотримо следећи пример.

Пример 2.4 *Размотримо узорак $X_1, ..., X_n$ и функционал $T(F) = P\{X_1 + X_2 > 0\}$. Одговарајући статистички функционал, односно $V-$статистика, је дат са $V_n = \frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{I}\{X_i + X_j > 0\}$. Тада важи*

$$V_n = \frac{n(n-1)}{n^2}U_n + \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{I}\{X_i > 0\} = \frac{\binom{n}{2}}{n^2}U_n + O_p\Big(\frac{1}{n}\Big),\ n \to \infty.$$

Према томе, већина асимптотских својстава ове $V-$статистике се може добити на основу асимптотских својстава њој одговарајуће $U-$статистике. $\triangle$

Слично као у претходном примеру, и у општем случају ће асимптотска својства $V-$статистика, при одређеним условима, следити из асимптотских својстава $U-$статистика уз коришћење репрезентације (2.5).

Пре увођења асимптотских својстава ових статистика, увешћемо још неке појмове везане за $U-$ и $V-$статистике. Претпоставимо да језгро $\Phi(x_1, ..., x_m)$ задовољава услов $E\Phi^2(X_1, ..., X_n) < \infty$. Пројекција реда $c,\ c \leqslant m$, језгра Φ је дефинисана са

$$\varphi_c(x_1, ..., x_c) = E(\Phi(X_1, ..., X_c, X_{c+1}, ..., X_m)|X_1 = x_1, ..., X_c = x_c).$$

Ако је $c = m$, онда је $\varphi_m(x_1, ..., x_m) = \Phi(x_1, ..., x_m)$, односно одговарајућа пројекција је једнака самом језгру. Дефинишемо

$$\begin{aligned} g_1(x_1) &= \varphi_1(x_1) - \theta, \\ g_c(x_1, ..., x_c) &= \varphi_c(x_1, ..., x_c) - \theta - \sum_{i=1}^{c} g_1(x_i) - \sum_{1\leqslant i_1<i_2\leqslant c} g_2(x_{i_1}, x_{i_2}) - \cdots \\ &\quad - \sum_{1\leqslant i_1<\cdots<i_{c-1}\leqslant c} g_{c-1}(x_{i_1}, ..., x_{i_{c-1}}),\ c = 2, ..., m, \end{aligned} \tag{2.6}$$

где је $\theta = E(\Phi(X_1, ..., X_m))$. Функције g_c су симетричне по својим аргументима и важи да је $E(g_c(x_1, ..., x_{c-1}, X_c)) = 0,\ c = 1, ..., m$. Уз помоћ ових функција може се увести Хефдингова репрезентација $U-$статистика (*Hoeffding (1948)*) на следећи начин

$$U_n - \theta = \sum_{c=1}^{m} \frac{\binom{m}{c}}{\binom{n}{c}} S_{cn}, \tag{2.7}$$

где је $S_{cn} = \sum\limits_{1\leqslant i_1<\cdots<i_c\leqslant n} g_c(X_{i_1}, ..., X_{i_c}),\ c = 1, ..., m.$

Дефинишемо $\sigma_0^2 = 0$ и

$$\sigma_c^2 = D(\varphi_c(X_1, ..., X_c)),$$

за $1 \leqslant c \leqslant m$. Приметимо да, применом Јенсенове неједнакости и особина условног очекивања, следи

$$\begin{aligned} \sigma_{c+1}^2 &= D(\varphi_{c+1}(X_1, ..., X_{c+1})) \\ &= E(\varphi_{c+1}(X_1, ..., X_{c+1}) - E(\varphi_{c+1}(X_1, ..., X_{c+1})))^2 \\ &= E(E(\varphi_{c+1}(X_1, ..., X_{c+1}) - E(\varphi_{c+1}(X_1, ..., X_{c+1})))^2|(X_1, ..., X_c)) \\ &\geqslant E(E(\varphi_{c+1}(X_1, ..., X_{c+1}) - E(\varphi_{c+1}(X_1, ..., X_{c+1}))|(X_1, ..., X_c)))^2 \\ &= E(E(\varphi_{c+1}(X_1, ..., X_{c+1})|(X_1, ..., X_c)) - E(E(\varphi_{c+1}(X_1, ..., X_{c+1})|(X_1, ..., X_c)))^2 \\ &= D(E(\varphi_{c+1}(X_1, ..., X_{c+1})|(X_1, ..., X_c))) \\ &= D(\varphi_c(X_1, ..., X_c)) = \sigma_c^2. \end{aligned}$$

Дакле, важи да је $0 = \sigma_0^2 \leqslant \sigma_1^2 \leqslant \cdots \leqslant \sigma_m^2 = D(\Phi(X_1, ..., X_m)) < \infty.$

Дисперзија $U-$статистике је одређена изразом

$$D(U_n) = \frac{1}{\binom{n}{m}} \sum_{c=0}^{m} \binom{m}{c} \binom{n-m}{m-c} \sigma_c^2 \tag{2.8}$$

и може се показати да важи (*Serfling (1984)*)

$$\frac{m^2}{n}\sigma_1^2 \leqslant D(U_n) \leqslant \frac{m}{n}\sigma_m^2. \tag{2.9}$$

Ако важи $0 = \sigma_{c-1}^2 < \sigma_c^2$, $U-$статистика, а самим тим и $V-$статистика, је ранга c.

Слично као код $V-$статистика, једино у случају када је ред $U-$статистике једнак 1, чланови суме којом је дефинисана $U-$статистика су међусобно независни. У општем случају, за добијање граничне расподеле, могу се дефинисати пројекционе $U-$статистике на следећи начин

$$U_n^\star - \theta = \frac{\binom{m}{c}}{\binom{n}{c}} \sum_{1 \leqslant i_1 < \cdots < i_c \leqslant n} (\varphi_c(X_{i_1}, ..., X_{i_c}) - \theta). \tag{2.10}$$

Ове статистике могу представљати апроксимацију $U-$статистика у зависности од њеног ранга. Претпоставимо да је $0 = \sigma_{c-1}^2 < \sigma_c^2,\ c > 1$. Тада важи

$$U_n^\star - \theta = \frac{\binom{m}{c}}{\binom{n}{c}} S_{cn} = \frac{\binom{m}{c}}{\binom{n}{c}} \sum_{1 \leqslant i_1 < \cdots < i_c \leqslant n} g_c(X_{i_1}, ..., X_{i_c}). \tag{2.11}$$

Како је разлика $U_n - U_n^\star$ такође $U-$статистика са језгром

$$\phi(x_1, ..., x_m) = \Phi(x_1, ..., x_m) - \theta - \sum_{1 \leqslant i_1 < \cdots < i_c \leqslant m} (\varphi_c(x_{i_1}, ..., x_{i_c}) - \theta),$$

при чему је $E(\phi(X_1, ..., X_m)) = 0$ и све пројекције језгра ϕ до реда c су једнаке нули, из формуле за дисперзију $U-$статистике (2.8) следи да је $E(U_n - U_n^\star)^2 = O(n^{-(c+1)})$ и да

$$n^{\frac{c}{2}}(U_n - U_n^\star) \xrightarrow{P} 0. \tag{2.12}$$

Па се гранична расподела од $n^{\frac{c}{2}}(U_n - \theta)$ може добити одређивањем граничне расподеле од $n^{\frac{c}{2}}(U_n^\star - \theta)$.

У наставку следе теореме које дају асимптотска својства $U-$ и $V-$статистика. Прва теорема представља закон великих бројева за $U-$статистике, док следеће две теореме дају граничну расподелу недегенерисаних, односно слабо дегенерисаних, $U-$статистика. Након тога следе аналогне теореме у случају $V-$статистика. Докази тих теорема су засновани на претходно уведеним репрезентацијама и могу се наћи у *Korolyuk and Borovskikh (1994)* или *Serfling (2009)*.

Теорема 2.2 *Нека је $X_1, ..., X_n$ узорак из расподеле F и нека за језгро Φ $U-$статистике U_n важи $E|\Phi(X_1, ..., X_m)| < \infty$. Тада U_n скоро сигурно конвергира ка θ када $n \to \infty$.*

Теорема 2.3 *Нека је $X_1, ..., X_n$ узорак из расподеле F. Ако за језгро Φ $U-$статистике U_n важи $E(\Phi^2(X_1, ..., X_m)) < \infty$ и $\sigma_1^2 > 0$, тада*

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

У случају слабо дегенерисаних $U-$статистика, прва пројекција језгра је једнака θ, па је и $\sigma_1^2 = 0$. Због тога се у том случају мора применити другачији метод него у случају недегенерисаних статистика. У ту сврху користи се метод ортогоналног развоја. Нека је A линеарни оператор дефинисан на $L^2(\mathbb{R}, F)$ са

$$Aq(x) = \int \varphi_2(x, y)q(y)dF(y),\ x \in \mathbb{R},\ q \in L^2, \tag{2.13}$$

где је $\varphi_2(x, y)$ друга пројекција језгра $\Phi(\cdot)$. Нека су његове сопствене вредности дате са $\upsilon_k, k = 1, 2, ...$, а ортонормирани сопствени вектори са $e_k, k = 1, 2, ...$ Тада важи

$$\varphi_2(x, y) - \theta = \sum_{k=1}^{n} \upsilon_k e_k(x)e_k(y). \tag{2.14}$$

Користећи овај развој и (2.12) добија се асимптотска расподела слабо дегенерисане $U-$статистике.

Теорема 2.4 *Нека је $X_1, ..., X_n$ узорак из расподеле F. Ако за језгро Φ $U-$статистике U_n важи $E(\Phi^2(X_1, ..., X_m)) < \infty$ и $0 = \sigma_1^2 < \sigma_2^2$, тада*

$$n(U_n - \theta) \xrightarrow{D} \binom{m}{2} \sum_{j=1}^{\infty} \upsilon_j (Z_j^2 - 1),$$

где је $\{\upsilon_k\}, k = 1, 2, ...,$ низ сопствених вредности интегралног оператора A дефинисаног у (2.13), а $Z_k, k = 1, 2, ...,$ независне случајне величине са стандардном нормалном расподелом.

Већ је поменуто да при одговарајућим условима, $U-$ и $V-$статистике са истим језгром се понашају на сличан начин. Следећа лема даје потребне услове за то, а њен доказ је заснован на Хефдинговој репрезентацији (2.5) (за детаље видети *Serfling (2009)*).

Лема 2.1 *Нека је r позитиван цео број и нека за језгро Φ важи $E|\Phi(X_{i_1}, ..., X_{i_m})|^r < \infty$, за све $1 \leqslant i_1, ..., i_m \leqslant n$. Тада*

$$E|U_n - V_n|^r = O_p\Big(\frac{1}{n^r}\Big),\ n \to \infty.$$

Следеће теореме, које дају закон великих бројева и асимптотску расподелу $V-$статистика, редом, следе на основу теореме 2.2, односно теорема 2.3 и 2.4, и леме 2.1.

Теорема 2.5 *Нека је $X_1, ..., X_n$ узорак из расподеле F и нека за језгро Φ $V-$статистике V_n важи $E|\Phi(X_{i_1}, ..., X_{i_m})| < \infty$, за све $1 \leqslant i_1, ..., i_m \leqslant n$. Тада V_n скоро сигурно конвегира ка θ када $n \to \infty$.*

Теорема 2.6 *Нека је $X_1, ..., X_n$ узорак из расподеле F и нека за језгро Φ $V-$статистике V_n важи $E(\Phi^2(X_{i_1}, ..., X_{i_m})) < \infty$, за све $1 \leqslant i_1, ..., i_m \leqslant n$. Тада*

1. *ако је $\sigma_1^2 > 0$, онда*
$$\sqrt{n}(V_n - \theta) \xrightarrow{D} \mathcal{N}(0, m^2\sigma_1^2).$$

2. *ако је $0 = \sigma_1^2 < \sigma_2^2$, онда*
$$n(V_n - \theta) \xrightarrow{D} \binom{m}{2} \sum_{j=1}^{\infty} \upsilon_j Z_j^2,$$
где је $\{\upsilon_k\}, k = 1, 2, ...,$ низ сопствених вредности интегралног оператора A дефинисаног у (2.13), а Z_k, $k = 1, 2, ...,$ независне случајне величине са стандардном нормалном расподелом.

У случају дегенерисаних статистика, односно статистика ранга $r > 2$, одговарајућа тврђења се могу пронаћи у *Korolyuk and Borovskikh (1994)*.

2.4 Асимптотска својства $V-$статистика са оцењеним параметром

У претходном поглављу, $V-$статистике су од расподеле F из фамилије расподела $\mathcal{F}$ зависиле само кроз узорак $X_1, ..., X_n$ из те расподеле. Међутим, може се десити да $V-$статистике у свом запису садрже и неки битан фактор фамилије расподела $\mathcal{F}$ који је непознат и који се мора оценити на основу узорка. У том случају, статистика укључује узорак и кроз помоћну статистику којом је оцењен дати параметар. Такве статистике се називају $V-$статистикама са оцењеним параметром. Поставља се питање какав је утицај тог оцењеног параметра на граничну расподелу те статистике.

$V-$статистике са оцењеним параметром се могу записати у следећем облику

$$V_n(\widehat{\lambda}_n) = \frac{1}{n^m} \sum_{i_1, ..., i_m = 1}^{n} \Phi(X_{i_1}, ..., X_{i_m}; \widehat{\lambda}_n), \tag{2.15}$$

где је $\widehat{\lambda}_n$ постојана оцена параметра λ добијена на основу узорка $X_1, ..., X_n$. Ако бисмо уместо статистике $V_n(\widehat{\lambda}_n)$ посматрали статистику $V_n(\gamma)$, за неко фиксирано γ, асимптотски резултати би следили из претходног поглавља. Циљ је утврдити да ли су граничне расподеле ове две статистике исте, односно утврдити да ли оцена параметра $\widehat{\lambda}_n$ има утицаја на асимптотску расподелу статистике. Као и у претходном поглављу, од интереса су статистике ранга 1 и 2, односно недегенерисане и слабо дегенерисане статистике.

Нека је $U_n(\widehat{\lambda}_n)$ одговарајућа $U-$статистика реда m са оцењеним параметром и истим језгром као $V-$статистика (2.15), односно статистика облика

$$U_n(\widehat{\lambda}_n) = \frac{1}{\binom{n}{m}} \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} \Phi(X_{i_1}, ..., X_{i_m}; \widehat{\lambda}_n). \tag{2.16}$$

Нека је $\theta(\gamma) = E_\lambda(\Phi(X_1, ..., X_m; \gamma))$ очекивање језгра ове статистике, при чему је λ вредност параметра. Следећа теорема представља закон великих бројева $U-$статистика са оцењеним параметром. Доказ ове теореме се може наћи у *Iverson and Randles (1989)*.

Теорема 2.7 *Нека је $E|\Phi(X_1, ..., X_m; \lambda)| < \infty$ и нека постоји околина $K(\lambda)$ од λ, таква да ако је $D(\lambda, d)$ сфера са центром λ и полупречником d таква да $D(\lambda, d) \subset K(\lambda)$ онда*

$$\lim_{d \to 0} E(\sup_{\gamma \in D(\lambda, d)} |\Phi(X_1, ..., X_m; \gamma) - \Phi(X_1, ..., x_m; \lambda)|) = 0.$$

При тим условима

1. *ако $\widehat{\lambda}_n \xrightarrow{P} \lambda$, онда $U_n(\widehat{\lambda}_n) \xrightarrow{P} \theta(\lambda)$;*

2. *ако $\widehat{\lambda}_n \xrightarrow{ss} \lambda$, онда $U_n(\widehat{\lambda}_n) \xrightarrow{ss} \theta(\lambda)$.*

У раду *Randles (1982)* добијени су први резултати у вези са асимптотском расподелом недегенерисаних $U-$статистика са оцењеним параметром. Доказ следеће теореме се може наћи у наведеном раду.

Теорема 2.8 *Претпоставимо да постоји околина $K(\lambda)$ од λ и константа $C > 0$, таква да ако $\gamma \in K(\lambda)$ и $D(\gamma, d)$ је сфера са центром γ и полупречником d таква да $D(\gamma, d) \subset K(\lambda)$ и важе услови*

$$E(\sup_{\gamma' \in D(\gamma, d)} |\Phi(X_1, ..., X_m; \gamma') - \Phi(X_1, ..., X_m; \gamma)|) \leqslant Cd$$

и

$$\lim_{d \to 0} E(\sup_{\gamma' \in D(\gamma, d)} |\Phi(X_1, ..., X_m; \gamma') - \Phi(X_1, ..., X_m; \gamma)|^2) = 0.$$

Додатно нека важи један од следећа два услова:

1. *Нека $\theta(\gamma)$ има извод једнак 0 у тачки $\gamma = \lambda$, нека је $\sqrt{n}(\widehat{\lambda}_n - \lambda) = O_p(1)$, $n \to \infty$, и нека је*
$$\sqrt{n}(U_n(\lambda) - \theta(\lambda)) \xrightarrow{D} \mathcal{N}(0, \tau^2),$$
где је
$$\tau^2 = D(E(\Phi(X_1, ..., X_m; \lambda)|X_1)) > 0. \tag{2.17}$$

2. *Нека $\theta(\gamma)$ има ненула извод у $\gamma = \lambda$ и додатно нека*
$$\sqrt{n}(U_n(\lambda) - \theta(\lambda), \widehat{\lambda}_n - \lambda) \xrightarrow{D} \mathcal{N}_{p+1}(0, \Sigma).$$

Тада

$$\sqrt{n}(U_n(\widehat{\lambda}_n) - \theta(\lambda)) \xrightarrow{D} \mathcal{N}(0, \tau^2),$$

ако је $\tau^2 > 0$, *где је* τ^2 *дато са* (2.17) *или са*

$$\tau^2 = D^T \Sigma D, \tag{2.18}$$

где је $D' = \left(1, \frac{\partial\theta(\cdot)}{\partial\gamma_1}, ..., \frac{\partial\theta(\cdot)}{\partial\gamma_p}\right)\Big|_{\gamma=\lambda}$.

Уз коришћење Хефдингове репрезентације (2.5), тврђење ће важити и за $V-$статистике са оцењеним параметром, односно статистике облика (2.15). Претходни резултат се може проширити и на диференцијабилне статистичке функционале. За више детаља видети *Randles (1982)*.

Размотримо сада слабо дегенерисане $V-$статистике са оцењеним параметром. Претпоставимо да за језгро $\Phi(x_1, ..., x_m; \lambda)$ важе следећи услови

$$\varphi_1(x;\lambda) = E(\Phi(X_1, X_2, ..., X_m; \lambda)|X_1 = x) = 0, \tag{2.19}$$

за свако x, и

$$\sigma_2^2 = E\varphi_2^2(X_1, X_2; \lambda) > 0, \tag{2.20}$$

где је $\varphi_2(x_1, x_2; \lambda) = E(\Phi(X_1, X_2, ..., X_m; \lambda)|X_1 = x_1, X_2 = x_2)$. Наведени услови представљају услове слабе дегенерисаности језгра Φ. Размотримо специјалан случај $V-$статистика реда $2m$ са оцењеним параметром облика

$$V_n(\widehat{\lambda}) = \int\limits_{-\infty}^{+\infty} \left(\frac{1}{n^m}\sum_{i_1,...,i_m=1}^{n} h(X_{i_1}, ..., X_{i_m}, t; \widehat{\lambda}_n)\right)^2 dM(t),$$

чије се језгро може написати у симетризованом облику

$$\Phi(X_1, ..., X_{2m}; \widehat{\lambda}_n) = \frac{1}{(2m)!}\sum_{\pi\in\Pi(2m)}\int\limits_{-\infty}^{+\infty} h(X_{\pi(1)}, ..., X_{\pi(m)}, t; \widehat{\lambda}_n) h(X_{\pi(m+1)}, ..., X_{\pi(2m)}, t; \widehat{\lambda}_n) dM(t), \tag{2.21}$$

где је $\Pi(2m)$ скуп свих пермутација скупа $\{1, 2, ..., 2m\}$, h нека функција, а M коначна позитивна мера. За овако дефинисано језгро и добијање асимптотских резултата није неопходно захтевати његову диференцијабилност. Испоставља се да $V-$статистике са овако дефинисаним језгром имају велику примену, што ће се видети кроз примере тест статистика у поглављу 3.

Де Вет и Рандлес су у свом раду (*de Wet and Randles (1987)*) показали да слабо дегенерисана V-статистика са оцењеним параметром реда два конвергира у расподели ка бесконачној линеарној комбинацији независних случајних величина са χ_1^2 расподелом. Ти резултати се могу уопштити тако да важе за V-статистике реда $2m$. Прво ћемо приказати услове који треба да важе да би се могли доказати главни резултати.

Услов 2.1 *Претпоставимо да* $\mu(t;\gamma) = E_\lambda(h(X_1,...,X_m,t;\gamma))$ *постоји за свако* γ *у околини* $\gamma=\lambda$ *и да задовољава* $\mu(t;\lambda)\equiv 0$ *за свако* t. *Додатно претпоставимо да за свако* $\varepsilon>0$ *постоји ограничена сфера* S *у* $\mathbb{R}^p$ *са центром* λ, *таква да за* $\gamma\in S$ *важи*

$$\frac{1}{||\gamma-\lambda||^2}\int\limits_{-\infty}^{+\infty}(\mu(t;\gamma)-d_1\mu(t;\lambda)^T(\gamma-\lambda))^2dM(t)<\varepsilon,$$

где је $d_1\mu(t;\lambda)$ *вектор парцијалних извода од* $\mu(t;\gamma) = E_\lambda(h(X_1,...,X_m,t;\gamma))$ *у* $\gamma=\lambda$, *који задовољава*

$$\int\limits_{-\infty}^{+\infty}(d_1\mu(t;\lambda)_r)^2dM(t)<\infty,$$

за $r=1,...,p$, *где је* $d_1\mu(\cdot)_r$ $r-$*та компонента вектора* $d_1\mu(\cdot)$.

Услов 2.2 *Нека је*

$$\widehat{\lambda}_n=\lambda+\frac{1}{n}\sum_{i=1}^{n}\alpha(X_i)+o_p\Big(\frac{1}{\sqrt{n}}\Big),\ n\to\infty,$$

где је $E(\alpha(X_i)_r)=0$ *и* $E(\alpha(X_i)_r\cdot\alpha(X_i)_{r'})<\infty$ *за свако* $1\leqslant r\leqslant r'\leqslant p$.

Услов 2.3 *За било коју варијацију* $(i_1,...,i_m)$ *индекса из скупа* $\{1,2,....,n\}$, *постоји околина* $K(\lambda)$ *од* λ *и константа* $C>0$, *таква да ако је* $\gamma\in K(\lambda)$ *и ако је* $D(\gamma,d)$ *сфера са центром* γ *и полупречником* d *таква да* $D(\gamma,d)\subset K(\lambda)$, *онда*

$$\int\limits_{-\infty}^{+\infty}(E(\sup_{\gamma'\in D(\gamma,d)}|h(X_{i_1},...,X_{i_m},t;\gamma')-h(X_{i_1},...,X_{i_m},t;\gamma)|))^2dM(t)\leqslant Cd^2.$$

Додатно, за неко $\varepsilon>0$ *постоји* $d^\star>0$ *тако да* $0<d<d^\star$, *онда*

$$\int\limits_{-\infty}^{+\infty}E(\sup_{\gamma'\in D(\gamma,d)}|h(X_{i_1},...,X_{i_m},t;\gamma')-h(X_{i_1},...,X_{i_m},t;\gamma)|)^4dM(t)<\varepsilon.$$

Дефинишимо нову $V-$статистику, која зависи од тачне вредности параметра, односно

$$\begin{aligned}V_n^\star(\lambda)=\frac{1}{n^{2m}}\sum_{i_1,...,i_{2m}=1}^{n}\int\limits_{-\infty}^{+\infty}&\Bigg(h(X_{i_1},...,X_{i_m},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{i_j\in\{i_1,...,i_m\}}\alpha(X_{i_j})\Bigg)\\&\times\Bigg(h(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{i_k\in\{i_{m+1},..,i_{2m}\}}\alpha(X_{i_k})\Bigg)dM(t),\end{aligned}\tag{2.22}$$

где је функција α дефинисана у услову 2.2. Означимо са $\Phi_\star(X_1,...,X_{2m};\lambda)$ симетризацију јегра $V-$статистике (2.22), односно

$$\Phi_\star(X_1,...,X_{2m};\widehat{\lambda}_n)=\frac{1}{(2m)!}\sum_{\pi\in\Pi(2m)}\int\limits_{-\infty}^{+\infty}\Big(h(X_{\pi(1)},...,X_{\pi(m)},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{j=1}^{m}\alpha(X_{\pi(j)})\Big)$$
$$\times\Big(h(X_{\pi(m+1)},...,X_{\pi(2m)},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{k=m+1}^{2m}\alpha(X_{\pi(k)})\Big)dM(t), \tag{2.23}$$

где је $\Pi(2m)$ скуп свих пермутација скупа $\{1,2,...,2m\}$. Приметимо да је

$$\begin{aligned}\varphi_1^\star(x;\lambda)&=E(\Phi_\star(x,X_2,...,X_{2m};\lambda))\\&=E\Big(\int\limits_{-\infty}^{+\infty}\Big(h(x,X_2,...,X_m,t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x)+\sum_{j=2}^{m}\alpha(X_j)\Big)\Big)\\&\qquad\times\Big(h(X_{m+1},...,X_{2m},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{k=m+1}^{2m}\alpha(X_k)\Big)dM(t)\Big)\\&=\int\limits_{-\infty}^{+\infty}E\Big(h(x,X_2,...,X_m,t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x)+\sum_{j=2}^{m}\alpha(X_j)\Big)\Big)\\&\qquad\times E\Big(h(X_{m+1},...,X_{2m},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{k=m+1}^{2m}\alpha(X_k)\Big)dM(t)\Big)=0\end{aligned}$$

што следи из услова 2.1 и 2.2. Такође,

$$\begin{aligned}&\varphi_2^\star(x_1,x_{m+1};\lambda)=E(\Phi_\star(x_1,X_2,...,X_m,x_{m+1},X_{m+2}...,X_{2m};\lambda))\\&=\frac{2m^2(2m-2)!}{(2m)!}E\Big(\int\limits_{-\infty}^{+\infty}\Big(h(x_1,X_2,...,X_m,t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x_1)+\sum_{j=2}^{m}\alpha(X_j)\Big)\Big)\\&\qquad\times\Big(h(x_{m+1},X_{m+2},...,X_{2m},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x_{m+1})+\sum_{k=m+2}^{2m}\alpha(X_k)\Big)\Big)dM(t)\Big)\\&=\frac{2m^2(2m-2)!}{(2m)!}\int\limits_{-\infty}^{+\infty}E\Big(h(x_1,X_2,...,X_m,t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x_1)+\sum_{j=2}^{m}\alpha(X_j)\Big)\Big)\\&\qquad\times E\Big(h(x_{m+1},X_{m+2},...,X_{2m},t;\lambda)+d_1\mu(t;\lambda)^T\frac{1}{m}\Big(\alpha(x_{m+1})+\sum_{k=m+2}^{2m}\alpha(X_k)\Big)\Big)dM(t)\\&=\frac{2m^2(2m-2)!}{(2m)!}\int\limits_{-\infty}^{+\infty}\Big(h_1(x_1,t;\lambda)+d_1\mu(t;\lambda)^T\frac{\alpha(x_1)}{m}\Big)\Big(h_1(x_{m+1},t;\lambda)+d_1\mu(t;\lambda)^T\frac{\alpha(x_{m+1})}{m}\Big)dM(t),\end{aligned}\tag{2.24}$$

где је $h_1(x,t;\lambda)=E_\lambda(h(x,X_2,...,X_m,t;\lambda))$.

Следећа теорема даје асимптотска својства слабо дегенерисане $V-$статистике $V_n(\widehat{\lambda}_n)$ са оцењеним параметром и доказана је у раду аутора (*Cuparić et al. (2020)*).

Теорема 2.9 (Цупарић, Милошевић, Обрадовић 2020) *Нека је $X_1, ..., X_n$ узорак из расподеле F. Нека важе услови 1, 2 и 3, и додатно $E(\Phi_\star^2(X_{i_1}, ..., X_{i_{2m}})) < +\infty$ и $E(\Phi_\star(X_{i_1}, ..., X_{i_{2m}})) < +\infty$, за све варијације са понављањем $(i_1, ..., i_{2m})$ скупа $\{1, ..., n\}$. Тада*

$$n(V_n(\widehat{\lambda}_n) - V_n^\star(\lambda)) \xrightarrow{P} 0.$$

Додатно, ако је $E(\varphi_2^\star(X_1, X_2; \lambda))^2 > 0$, онда

$$nV_n(\widehat{\lambda}_n) \xrightarrow{D} \binom{2m}{2} \sum_{k=1}^{+\infty} \upsilon_k Z_k^2,$$

где је $\{\upsilon_k\}, k = 1, 2, ...,$ низ сопствених вредности оператора $A^\star$ дефинисаног на $L^2(\mathbb{R}, F)$ са

$$A^\star q(x) = \int\limits_R \varphi_2^\star(x, y) q(y) dF(y),\ x \in \mathbb{R},\ q \in L^2, \tag{2.25}$$

при чему је $\varphi_2^\star(x, y)$ друга пројекција језгра $\Phi_\star(x_1, ..., x_{2m})$, а $Z_k, k = 1, 2, ...,$ су независне случајне величине са стандардном нормалном расподелом.

Доказ: Без губитка општости, претпоставимо да је $\int\limits_{-\infty}^{+\infty} dM(t) = 1$. Дефинишемо помоћну статистику Y_n, која на неки начин представља прелаз између $V_n^\star(\lambda)$ и $V_n(\widehat{\lambda}_n)$, са

$$Y_n = \int\limits_{-\infty}^{+\infty} \left(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} h(X_{i_1}, ..., X_{i_m}, t; \lambda) + \mu(t; \widehat{\lambda}_n) \right)^2 dM(t).$$

Показаћемо да $nY_n - nV_n^\star(\lambda) \xrightarrow{P} 0$ и $nV_n(\widehat{\lambda}_n) - nY_n \xrightarrow{P} 0$, а одатле ће следити тврђење теореме.

Посматрајмо прво разлику између Y_n и $V_n^\star(\lambda)$. Тада, користећи једнакост $a^2 - b^2 = (a-b)^2 + 2b(a-b)$, следи

$$\begin{aligned}
Y_n - V_n^\star(\lambda) &= \int\limits_{-\infty}^{+\infty} \left(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} h(X_{i_1}, ..., X_{i_m}, t; \lambda) + \mu(t; \widehat{\lambda}_n) \right)^2 dM(t) \\
&\quad - \int\limits_{-\infty}^{+\infty} \left(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} \Big(h(X_{i_1}, ..., X_{i_m}, t; \lambda) + d_1\mu(t; \lambda)^T \frac{1}{m} \sum_{i_k \in \{i_1, \ldots, i_m\}} \alpha(X_{i_k}) \Big) \right)^2 dM(t) \\
&= \int\limits_{-\infty}^{+\infty} \Bigg(\bigg(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} \Big(h(X_{i_1}, ..., X_{i_m}, t; \lambda) + \mu(t; \widehat{\lambda}_n) - h(X_{i_1}, ..., X_{i_m}, t; \lambda) - d_1\mu(t; \lambda)^T \frac{1}{m} \sum_{i_k \in \{i_1, \ldots, i_m\}} \alpha(X_{i_k}) \Big) \bigg)^2 \\
&\quad + 2 \bigg(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} \Big(h(X_{i_1}, ..., X_{i_m}, t; \lambda) + d_1\mu(t; \lambda)^T \frac{1}{m} \sum_{i_k \in \{i_1, \ldots, i_m\}} \alpha(X_{i_k}) \Big) \bigg) \bigg(\frac{1}{n^m} \sum_{i_1, \ldots, i_m} \Big(h(X_{i_1}, ..., X_{i_m}, t; \lambda) \\
&\quad + \mu(t; \widehat{\lambda}_n) - h(X_{i_1}, ..., X_{i_m}, t; \lambda) - d_1\mu(t; \lambda)^T \frac{1}{m} \sum_{i_k \in \{i_1, \ldots, i_m\}} \alpha(X_{i_k}) \Big) \bigg) \Bigg) dM(t)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty}\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t) + 2\int_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h(X_{i_1},\dots,X_{i_m},t;\lambda)\\
&\quad + d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{i_k\in\{i_1,\dots,i_m\}}\alpha(X_{i_k})\Big)\Big)\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)dM(t).
\end{aligned}
$$

Применом Коши-Шварцове неједнакости на други сабирак претходног израза следи

$$
\begin{aligned}
n(Y_n - V_n^{\star}(\lambda)) \leqslant{}& n\int_{-\infty}^{+\infty}\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t)\\
&+ 2\Big(\int_{-\infty}^{+\infty} n\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}h(X_{i_1},\dots,X_{i_m},t;\lambda) + d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t)\Big)^{\frac{1}{2}}\\
&\times\Big(\int_{-\infty}^{+\infty} n\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t)\Big)^{\frac{1}{2}}\\
={}& n\int_{-\infty}^{+\infty}\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t)\\
&+ 2\sqrt{nV_n^{\star}(\lambda)}\Big(\int_{-\infty}^{+\infty} n\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T\frac{1}{n}\sum_{k=1}^{n}\alpha(X_k)\Big)^2 dM(t)\Big)^{\frac{1}{2}}. \qquad (2.26)
\end{aligned}
$$

Из услова 2.1 следи да је

$$
n\int_{-\infty}^{+\infty}\Big(\mu(t;\widehat{\lambda}_n) - d_1\mu(t;\lambda)^T(\widehat{\lambda}_n - \lambda)\Big)^2 dM(t) < \varepsilon||\sqrt{n}(\widehat{\lambda}_n - \lambda)||^2. \qquad (2.27)
$$

Како из услова 2.2 важи да је $\widehat{\lambda}_n - \lambda$ ограничено у вероватноћи, то следи да први сабирак израза (2.26) конвергира у вероватноћи ка 0. С обиром да је $nV_n^{\star}(\lambda)$ ограничено у вероватноћи као $V-$статистика са језгром које задовољава услове теореме, користећи услов 2.1 и теорему Слуцког, следи да и други сабирак у (2.26) конвергира ка 0.

Размотримо сада разлику између $V_n(\widehat{\lambda}_n)$ и Y_n. Слично као раније следи

$$
\begin{aligned}
V_n(\widehat{\lambda}_n) - Y_n ={}& \int_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}h(X_{i_1},\dots,X_{i_m},t;\widehat{\lambda}_n)\Big)^2 dM(t) - \int_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}h(X_{i_1},\dots,X_{i_m},t;\lambda) + \mu(t;\widehat{\lambda}_n)\Big)^2 dM(t)\\
={}& \int_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h(X_{i_1},\dots,X_{i_m},t;\widehat{\lambda}_n) - h(X_{i_1},\dots,X_{i_m},t;\lambda)\Big) - \mu(t;\widehat{\lambda}_n)\Big)^2 dM(t)\\
&+2\int_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}h(X_{i_1},\dots,X_{i_m},t;\lambda) + \mu(t;\widehat{\lambda}_n)\Big)\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h(X_{i_1},\dots,X_{i_m},t;\widehat{\lambda}_n)
\end{aligned}
$$

$$-h(X_{i_1},...,X_{i_m},t;\lambda)\Big)-\mu(t;\widehat{\lambda}_n)\Big)dM(t)=R_1+2R_2.$$

Из Коши-Шварцове неједнакости следи да је

$$\begin{aligned}nR_2 &\leqslant \Big(\int\limits_{-\infty}^{+\infty} n\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m} h(X_{i_1},...,X_{i_m},t;\lambda)+\mu(t;\widehat{\lambda}_n)\Big)^2 dM(t)\Big)^{\frac{1}{2}}\\ &\times\Big(n\int\limits_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h(X_{i_1},...,X_{i_m},t;\widehat{\lambda}_n)-h(X_{i_1},...,X_{i_m},t;\lambda)\Big)-\mu(t;\widehat{\lambda}_n)\Big)^2 dM(t)\Big)^{\frac{1}{2}}\\ &\leqslant \sqrt{nY_n}\sqrt{nR_1}\end{aligned}$$

Како је nY_n ограничено у вероватноћи, ако се докаже да $nR_1 \xrightarrow{P} 0$, онда ће важити

$$nV_n(\widehat{\lambda}_n)-nY_n \xrightarrow{P} 0.$$

Нека је

$$Q_n(s)=\int\limits_{-\infty}^{+\infty}\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\Big)^2 dM(t).$$

Треба доказати да

$$nQ_n(\sqrt{n}(\widehat{\lambda}_n-\lambda)) \xrightarrow{P} 0.$$

Из услова 2.2 следи да постоји сфера S у $\mathbb{R}^p$ и $\delta>0$ тако да за свако n важи

$$P\{\sqrt{n}(\widehat{\lambda}_n-\lambda)\notin S\}\leqslant\frac{\delta}{2}.$$

За произвољно $\varepsilon'>0$ важи

$$\begin{aligned}P\{nQ_n(\sqrt{n}(\widehat{\lambda}_n-\lambda))>\varepsilon'\} &= P\{nQ_n(\sqrt{n}(\widehat{\lambda}_n-\lambda))>\varepsilon', \sqrt{n}(\widehat{\lambda}_n-\lambda)\in S\}\\ &+P\{nQ_n(\sqrt{n}(\widehat{\lambda}_n-\lambda))>\varepsilon', \sqrt{n}(\widehat{\lambda}_n-\lambda)\notin S\}\\ &\leqslant P\{\sup_{s\in S} nQ_n(s)>\varepsilon'\}+P\{\sqrt{n}(\widehat{\lambda}_n-\lambda)\notin S\}.\end{aligned}$$

Дакле, довољно је показати да $\sup\limits_{s\in S} nQ_n(s) \xrightarrow{P} 0$.

Из Чебишовљеве неједнакости следи да је

$$P\Big\{\sup_{s\in S} nQ_n(s)>\varepsilon'\Big\}\leqslant\frac{1}{\varepsilon'^2}E\Big(\sup_{s\in S} nQ_n(s)\Big)^2. \tag{2.28}$$

Приметимо да

$$
\begin{aligned}
nQ_n(s) &= n\int_{-\infty}^{+\infty}\left(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\right)^2 dM(t)\\
&= \frac{n}{n^{2m}}\int_{-\infty}^{+\infty}\sum_{i_1,\dots,i_{2m}}\Big(h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\\
&\quad\times\Big(h\Big(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)dM(t)\\
&= nQ'_n(s)+nQ''_n(s),
\end{aligned}
$$

где је

$$
\begin{aligned}
Q'_n(s) = \frac{1}{n^{2m}}\sum_{\{i_1,\dots,i_{2m}\}\in\mathcal{I}'}\int_{-\infty}^{+\infty}&\Big(h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\\
\times&\Big(h\Big(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)dM(t)
\end{aligned}
$$

и

$$
\begin{aligned}
Q''_n(s) = \frac{1}{n^{2m}}\sum_{\{i_1,\dots,i_{2m}\}\in\mathcal{I}''}\int_{-\infty}^{+\infty}&\Big(h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\\
\times&\Big(h\Big(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)dM(t),
\end{aligned}
$$

а $\mathcal{I}'$ скуп $2m$-торки индекса који се не понављају, односно $i_1 \neq \cdots \neq i_{2m}$, док је $\mathcal{I}''$ комплемент од $\mathcal{I}'$, односно скуп $2m-$торки индекса са бар једним понављањем. Према томе

$$
\sup_{s\in S} nQ_n(s) \leqslant \sup_{s\in S} nQ'_n(s) + \sup_{s\in S} nQ''_n(s). \tag{2.29}
$$

Размотримо први сабирак са десне стране израза (2.29). Ако је

$$
\begin{aligned}
H(x_1,...,x_{2m}) = \sup_{s\in S}\int_{-\infty}^{+\infty}&\Big|\Big(h\Big(x_1,...,x_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(x_1,...,x_m,t;\lambda)\Big)\\
\times&\Big(h\Big(x_{m+1},...,x_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(x_{m+1},...,x_{2m},t;\lambda)\Big)\Big|dM(t),
\end{aligned}
\tag{2.30}
$$

онда

$$
\sup_{s\in S} nQ'_n(s) \leqslant \frac{n}{n^{2m}}\sum_{\{i_1,\dots,i_{2m}\}\in\mathcal{I}'} H(X_{i_1},...,X_{i_{2m}}).
$$

Приметимо да

$$E\Big(\frac{n}{n^{2m}}\sum_{\{i_1,...,i_{2m}\}\in\mathcal{I}'} H(X_{i_1},...,X_{i_{2m}})\Big)^2 = \frac{n^2}{n^{4m}}\sum_{\substack{\{i_1,...,i_{2m}\}\in\mathcal{I}'\\ \{i_{2m+1},...,i_{4m}\}\in\mathcal{I}'}} E\big(H(X_{i_1},...,X_{i_{2m}})H(X_{i_{2m+1}},...,X_{i_{4m}})\big)$$
$$= \frac{n^2}{n^{4m}}\sum_{c=0}^{2m}\binom{n}{c}\binom{n-c}{4m-2c}\Psi_n(c),$$

где тачно c индекса из скупа $\{i_1,...,i_{2m}\}$ се поклапа са тачно c индекса из скупа $\{i_{2m+1},...,i_{4m}\}$, док су сви остали различити. Према томе, сваки члан претходне суме се асимптотски понаша као $\frac{1}{n^{c-2}}\Psi_n(c)$.

Ако је $c=0$, онда

$$\Psi_n(0) = (4m)!E(H(X_1,...,X_{2m})H(X_{2m+1},...,X_{4m}))) = (4m)!(E(H(X_1,...,X_{2m})))^2.$$

Заменом функције H у претходном изразу добија се

$$\Psi_n(0) = (4m)!\Bigg(E\Big[\sup_{s\in S}\int_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big)$$
$$\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)\Big]\Bigg)^2$$
$$\leqslant (4m)!\Bigg(E\Big[\int_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big)$$
$$\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)\Big]\Bigg)^2$$
$$\leqslant (4m)!\Bigg(E\Big[\int_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big|$$
$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big|dM(t)\Big]\Bigg)^2$$
$$= (4m)!\Bigg(\int_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big|$$
$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big|\Big]dM(t)\Bigg)^2$$
$$= (4m)!\Bigg(\int_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big|\Big]\Bigg)^2dM(t)\Bigg)^2.$$

Применом неједнакости $(a+b)^2 \leqslant 2a^2 + 2b^2$ и услова 2.3, добија се

$$\begin{aligned}\Psi_n(0) &\leqslant (4m)!\Big(2\int\limits_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big|\Big]\Big)^2 dM(t)\\ &\quad + 2\int\limits_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t)\Big)^2\\ &\leqslant 4(4m)!\Big(C\frac{d^2}{n}+\int\limits_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|d_1\mu\Big(t;\lambda\Big)^T\frac{s}{\sqrt{n}}\Big|\Big)^2 dM(t)\Big)^2\\ &\leqslant 4(4m)!\frac{d^4}{n^2}(C+K)^2,\end{aligned}$$

где је

$$K = \int\limits_{-\infty}^{+\infty}\sum_{r=1}^{p}(d_1\mu(t;\lambda)_r)^2 dM(t), \tag{2.31}$$

што је коначно из услова 2.1.

Ако је $c=1$, односно ако се тачно један индекс из скупа $\{i_1,...,i_{2m}\}$ поклапа са тачно једним индексом из скупа $\{i_{2m+1},...,i_{4m}\}$ док су сви остали индекси различити, следи

$$\Psi_n(1) = 4m^2(4m-2)!E(H(X_1,X_2,...,X_{2m})H(X_1,X_{2m+1},...,X_{4m-1}))).$$

Заменом функције H у претходном изразу добија се

$$\begin{aligned}\Psi_n(1) &= 4m^2(4m-2)!E\Big[\sup_{s\in S}\int\limits_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big)\\ &\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)\\ &\times\sup_{s\in S}\int\limits_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big)\\ &\times\Big(h\Big(X_{3m},...,X_{4m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t;\lambda)\Big)\Big|dM(t)\Big]\\ &\leqslant 4m^2(4m-2)!E\Big[\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big)\\ &\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)\\ &\times\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big)\end{aligned}$$

$$\times\Big(h\Big(X_{3m},...,X_{4m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$=4m^2(4m-2)!E\Big[\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big)$$

$$\times\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)$$

$$\times\Big(h\Big(X_{3m},...,X_{4m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big)\Big|dM(t_1)dM(t_2)\Big]$$

$$\leqslant 4m^2(4m-2)!E\Big[\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{3m},...,X_{4m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big|dM(t_1)dM(t_2)\Big]$$

$$=4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{3m},...,X_{4m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big|\Big]dM(t_1)dM(t_2).$$

Из независности случајних величина и претходног следи да је

$$\Psi_n(1)\leqslant 4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{3m},...,X_{4m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big|\Big]dM(t_1)dM(t_2).$$

Применом Коши-Шварцове неједнакости добија се да је

$$\Psi_n(1)\leqslant 4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|\Big]^2\Bigg)^{\frac{1}{2}}$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{3m},...,X_{4m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big|\Big]dM(t_1)dM(t_2)$$

$$=4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2\Bigg)^{\frac{1}{2}}$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]dM(t_1)$$

$$\times\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_{3m},...,X_{4m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m},...,X_{4m-1},t_2;\lambda)\Big|\Big]$$

$$\times\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)$$

$$-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|\Big]^2\Bigg)^{\frac{1}{2}}dM(t_2)$$

$$=4m^2(4m-2)!\Bigg(\int\limits_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2\Bigg)^{\frac{1}{2}}$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]dM(t_1)\Bigg)^2. \tag{2.32}$$

Поново применом Коши-Шварцове неједнакости на последњи израз добија се

$$\Psi_n(1)\leqslant 4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big]^2dM(t_1)$$

$$\times\int\limits_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big]\Bigg)^2dM(t_1)$$

$$\leqslant 4m^2(4m-2)!\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|+\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big]^2dM(t_1)$$

$$\times\int\limits_{-\infty}^{+\infty}\Bigg(E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]+\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Bigg)^2dM(t_1).$$

Применом неједнакости $(a+b)^2 \leqslant 2a^2+2b^2$ на оба члана претходног производа, а затим Коши-Шварцове неједнакости, следи

$$\Psi_n(1) \leqslant 4m^2(4m-2)!\Bigg(2\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2 dM(t)$$

$$+2\int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t_1)\Bigg)\Bigg(2\int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t_1)$$

$$+2\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]\Big)^2 dM(t_1)\Bigg)$$

$$\leqslant 4m^2(4m-2)!\Bigg(2\Big(\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-g(X_1,...,X_m,t_1;\lambda)\Big|\Big]^4 dM(t)\Big)^{\frac{1}{2}}$$

$$+2\int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t_1)\Bigg)\Bigg(2\int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t_1)$$

$$+2\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]\Big)^2 dM(t_1)\Bigg)$$

$$\leqslant 4m^2(4m-2)!\Big(2\varepsilon^{\frac{1}{2}}+2\frac{d^2}{n}K\Big)\Big(2\frac{d^2}{n}(C+K)\Big)\sim \varepsilon^{\frac{1}{2}}\frac{d^2}{n}+o_p\Big(\frac{1}{n}\Big),\ n\to\infty,$$

где је K дефинисано у (2.31).

Када је $c=2$, тада је

$$\Psi_n(2)=\Bigg(16\binom{m}{2}^2\varpi_1+4m^4\varpi_2+16\binom{m}{2}m^2\varpi_3\Bigg)(4m-4)!,$$

где је

$$\begin{aligned}\varpi_1&=E(H(X_1,X_2,X_3...,X_{2m})H(X_1,X_2,X_{2m+1},...,X_{4m-2}))),\\ \varpi_2&=E(H(X_1,X_2,...,X_{2m})H(X_1,X_{2m+1},...,X_{3m-1},X_{m+1},X_{3m},...,X_{4m-2}))),\\ \varpi_3&=E(H(X_1,X_2,...,X_{2m})H(X_1,X_{2m+1},...,X_{3m-1},X_2,X_{3m}...,X_{4m-2}))).\end{aligned}$$

Размотримо прво функцију ϖ_1. Тада

$$\varpi_1=E(H(X_1,X_2,X_3...,X_{2m})H(X_1,X_2,X_{2m+1},...,X_{4m-2})))$$

$$=E\Bigg[\sup_{s\in S}\int_{-\infty}^{+\infty}\Bigg|\Big(h\Big(X_1,X_2,X_3,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t;\lambda)\Big)$$

$$\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Bigg|dM(t)$$

$$\times\sup_{s\in S}\int\limits_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t;\lambda)\Big)$$

$$\times\Big(h\Big(X_{3m-1},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$\leqslant E\Big[\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)$$

$$\times\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t;\lambda)\Big)$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$=E\Big[\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda)\Big)$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t_2;\lambda)\Big)\Big|dM(t_1)dM(t_2)\Big]$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda)\Big)$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2)$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda)\Big)\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2).$$

Применом Коши-Шварцове неједнакости на претходни израз следи да је

$$
\begin{aligned}
\varpi_1 \leqslant & \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big] \\
& \times\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\Big|\Big]^2 \\
& \times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}} \\
& \times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2) \\
= & \int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_3,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,X_3,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}} \\
& \times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1) \\
& \times\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{3m-1},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{3m-1},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big] \\
& \times\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big) \\
& -h(X_1,X_2,X_{2m+1},...,X_{3m-2},t_2;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}dM(t_2) \\
= & \Big(\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}} \\
& \times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)\Big)^2.
\end{aligned}
$$

Последњи израз је исти као у (2.32), па слично као раније следи

$$
\varpi_1 \leqslant \varepsilon^{\frac{1}{2}}\frac{d^2}{n}+o_p\Big(\frac{1}{n}\Big),\ n\to\infty.
$$

Размотримо сада ϖ_2. Следи да је

$$
\begin{aligned}
\varpi_2 = & E(H(X_1,X_2,...,X_{2m})H(X_1,X_{2m+1},...,X_{3m-1},X_{m+1},X_{3m},...,X_{4m-2}))) \\
= & E\Big[\sup_{s\in S}\int_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big) \\
& \times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t) \\
& \times\sup_{s\in S}\int_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big)
\end{aligned}
$$

$$\times\Big(h\Big(X_{m+1},X_{3m},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},X_{3m},...,X_{4m-2},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$\leqslant E\Big[\int_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big|dM(t)$$

$$\times\int_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},X_{3m},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},X_{3m},...,X_{4m-2},t;\lambda)\Big|dM(t)\Big]$$

$$=E\Big[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|dM(t_1)dM(t_2)\Big]$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]dM(t_1)dM(t_2).$$

Из независности случајних величина следи

$$\varpi_2\leqslant\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|h\Big(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},X_{3m},...,X_{4m-2},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big|\Big]dM(t_1)dM(t_2)$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]\Big)^2dM(t_1)dM(t_2).$$

Применом Коши-Шварцове неједнакости добија се

$$
\begin{aligned}
\varpi_2 \leqslant & \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big|\Big]^2 \\
& \times E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2 dM(t_1)dM(t_2) \\
= & \Big(\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2 dM(t_1)\Big)^2 \\
\leqslant & \int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2\Big)^2 dM(t_1) \\
\leqslant & \int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|+\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big]^2\Big)^2 dM(t_1).
\end{aligned}
$$

Слично као раније, уз коришћење услова 2.3, добија се

$$
\begin{aligned}
\varpi_2 \leqslant & \int_{-\infty}^{+\infty}\Big(2E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2 \\
& + 2\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2\Big)^2 dM(t_1) \\
& \leqslant 8\Big(\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,...,X_m,t_1;\lambda)\Big|\Big]^2\Big)^2 dM(t_1) \\
& + \int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\Big|^4 dM(t_1)\Big) \\
& \leqslant 8\Big(\int_{-\infty}^{+\infty} E\Big[\sup_{s\in S}\Big|h\Big(X_1,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-g(X_1,...,X_m,t_1;\lambda)\Big|\Big]^4 dM(t_1) \\
& + \int_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big)\Big|^4 dM(t_1)\Big) \\
& < 8\Big(\varepsilon+\frac{d^4}{n^2}\int_{-\infty}^{+\infty}\Big(\Big|d_1\mu\Big(t;\lambda\Big)^T\Big|^2\Big)dM(t)\Big) = 8\varepsilon+o_p\Big(\frac{1}{n}\Big),\ n\to\infty.
\end{aligned}
$$

Размотримо сада ϖ_3. Следи да је

$$
\begin{aligned}
\varpi_3 = & E(H(X_1,X_2,...,X_{2m})H(X_1,X_{2m+1},...,X_{3m-1},X_2,X_{3m}...,X_{4m-2}))) \\
= & E\Big[\sup_{s\in S}\int_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,X_2,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t;\lambda)\Big)
\end{aligned}
$$

$$\times\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)$$

$$\times\sup_{s\in S}\int\limits_{-\infty}^{+\infty}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big)$$

$$\times\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$\leqslant E\Big[\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t;\lambda)\Big)\Big|dM(t)$$

$$\times\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t;\lambda)\Big)\Big|dM(t)\Big]$$

$$=E\Big[\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|dM(t_1)dM(t_2)\Big]$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2)$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2).$$

Применом Коши-Шварцове неједнакости добија се

$$\varpi_3 \leqslant \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1}...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|$$

$$\times\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)dM(t_2)$$

$$=\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]dM(t_1)$$

$$\times\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|\Big]^2$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}dM(t_2).$$

Поновном применом Коши-Шварцове неједнакости, слично као раније следи да је

$$\varpi_3 \leqslant \int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}$$

$$\times\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_{m+1},...,X_{2m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{m+1},...,X_{2m},t_1;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}dM(t_1)$$

$$\times\int_{-\infty}^{+\infty}\Big(E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_{2m+1},...,X_{3m-1},t_2;\lambda)\Big)\Big|\Big]^2$$

$$\times E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_2,X_{3m},...,X_{4m-2},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_2,X_{3m},...,X_{4m-2},t_2;\lambda)\Big)\Big|\Big]^2\Big)^{\frac{1}{2}}dM(t_2)$$

$$=\Big(\int_{-\infty}^{+\infty}E\Big[\sup_{s\in S}\Big|\Big(h\Big(X_1,X_2,...,X_m,t;\lambda+\frac{s}{\sqrt{n}}\Big)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_1,X_2,...,X_m,t;\lambda)\Big)\Big|\Big]^2dM(t)\Big)^2$$

$$\leqslant 8\varepsilon+o_p\Big(\frac{1}{n}\Big),\ n\to\infty.$$

Последњи корак следи исто као и у случају функције ϖ_2. Дакле,

$$\Psi_n(2)=\Bigg(16\binom{m}{2}^2\varepsilon^{\frac{1}{2}}\frac{d^2}{n}+32m^4\varepsilon+128\binom{m}{2}m^2\varepsilon+o_p\Big(\frac{1}{n}\Big)\Bigg)(4m-4)!\sim\varepsilon+o_p\Big(\frac{1}{n}\Big),\ n\to\infty.$$

За $c > 2$, $\frac{1}{n^{c-2}}\Psi_n(c) = o_p\Big(\frac{1}{n}\Big)$, $n \to \infty$, јер је очекивање одговарајуће функције ограничено. Дакле, узимајући у обзир да за свако $\varepsilon > 0$, d се може изабрати довољно мало, можемо закључити да

$$E(\sup_{s\in S} nQ'_n(s))^2 < const \cdot \varepsilon + o_p(1),\ n \to \infty.$$

Размотримо сада други сабирак са десне стране израза (2.29). Може се приметити да важи

$$\sup_{s\in S} nQ''_n(s) \leqslant \frac{n}{n^{2m}} \sum_{\{i_1,..,i_{2m}\}\in\mathcal{I}''} H(X_{i_1}, ..., X_{i_{2m}}),$$

где је $\mathcal{I}''$ скуп $2m-$торки индекса са бар једним понављањем и H дефинисано у (2.30). Тада

$$E\Bigg(\frac{n}{n^{2m}} \sum_{\{i_1,...i_{2m}\}\in\mathcal{I}''} H(X_{i_1}, ..., X_{i_{2m}})\Bigg)^2 = \frac{n^2}{n^{4m}} \sum_{\substack{\{i_1,...i_{2m}\}\in\mathcal{I}'' \\ \{i_{2m+1},...i_{4m}\}\in\mathcal{I}''}} E\Big(H(X_{i_1}, ..., X_{i_{2m}})H(X_{i_{2m+1}}, ..., X_{i_{4m}})\Big).$$

Размотримо општи члан ове суме

$$\begin{aligned}
&E\Big(H(X_{i_1}, ..., X_{i_{2m}})H(X_{i_{2m+1}}, ..., X_{i_{4m}})\Big)\\
&= E\Bigg(\sup_{s\in S}\int_{-\infty}^{+\infty}\Bigg|\Bigg(h\Big(X_{i_1}, ..., X_{i_m}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\\
&\qquad\times\Bigg(h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\Bigg|dM(t)\\
&\times\sup_{s\in S}\int_{-\infty}^{+\infty}\Bigg|\Bigg(h\Big(X_{i_{2m+1}}, ..., X_{i_{3m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}}, ..., X_{i_{3m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\\
&\qquad\times\Bigg(h\Big(X_{i_{3m+1}}, ..., X_{i_{4m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}}, ..., X_{i_{4m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\Bigg|dM(t)\Bigg)\\
&\leqslant E\Bigg(\int_{-\infty}^{+\infty}\sup_{s\in S}\Bigg|\Bigg(h\Big(X_{i_1}, ..., X_{i_m}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\\
&\qquad\times\Bigg(h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\Bigg|dM(t)\\
&\times\int_{-\infty}^{+\infty}\sup_{s\in S}\Bigg|\Bigg(h\Big(X_{i_{2m+1}}, ..., X_{i_{3m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}}, ..., X_{i_{3m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\\
&\qquad\times\Bigg(h\Big(X_{i_{3m+1}}, ..., X_{i_{4m}}, t; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}}, ..., X_{i_{4m}}, t; \lambda)-\mu\Big(t; \lambda+\frac{s}{\sqrt{n}}\Big)\Bigg)\Bigg|dM(t)\Bigg)
\end{aligned}$$

$$\leqslant E\Big(\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|dM(t)$$

$$\times\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}},...,X_{i_{3m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}},...,X_{i_{3m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}},...,X_{i_{4m}},t;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}},...,X_{i_{4m}},t;\lambda)-\mu\Big(t;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|dM(t)\Big)$$

$$=E\Big(\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|dM(t_1)dM(t_2)\Big)$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)dM(t_1)dM(t_2).$$

Дакле,

$$E\Big(\frac{n}{n^{2m}}\sum_{\{i_1,...,i_{2m}\}\in\mathcal{I}''}H(X_{i_1},...,X_{i_{2m}})\Big)^2$$

$$\leqslant\frac{n^2}{n^{4m}}\sum_{\substack{\{i_1,...,i_{2m}\}\in\mathcal{I}''\\ \{i_{2m+1},...,i_{4m}\}\in\mathcal{I}''}}\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}},...,X_{i_{2m}},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}},...,X_{i_{3m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|$$

$$\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}},...,X_{i_{4m}},t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)dM(t_1)dM(t_2). \tag{2.33}$$

Сабирци највећег реда претходног израза су они код којих се понављају тачно два индекса из скупа $\{i_1, ..., i_{2m}\}$ и тачно два индекса из скупа $\{i_{2m+1}, ..., i_{4m}\}$ и нема преклапања индекса из скупа $\{i_1, ..., i_{2m}\}$ са индексима из скупа $\{i_{2m+1}, ..., i_{4m}\}$. Таквих сабирака у суми има $\binom{n}{2}\binom{n-2}{4m-4}\binom{2m}{2}^2(4m-4)!2!$, односно реда n^{4m-2}. Дакле, сума тих чланова се асимптотски понаша као

$$\begin{aligned}\Upsilon_n = &\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1}, ..., X_{i_m}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}}, ..., X_{i_{3m}}, t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}}, ..., X_{i_{3m}}, t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}}, ..., X_{i_{4m}}, t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}}, ..., X_{i_{4m}}, t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)dM(t_1)dM(t_2).\end{aligned}$$

Из независности случајних величина следи

$$\begin{aligned}\Upsilon_n = &\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1}, ..., X_{i_m}, t_1; \lambda + \frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t_1; \lambda)-\mu\Big(t_1; \lambda + \frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda + \frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda)-\mu\Big(t_1; \lambda + \frac{s}{\sqrt{n}}\Big)\Big|\Big)\\
&\times E\Big(\sup_{s\in S}\Big|h\Big(X_{i_{2m+1}}, ..., X_{i_{3m}}, t_2; \lambda + \frac{s}{\sqrt{n}}\Big)-h(X_{i_{2m+1}}, ..., X_{i_{3m}}, t_2; \lambda)-\mu\Big(t_2; \lambda + \frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{3m+1}}, ..., X_{i_{4m}}, t_2;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{3m+1}}, ..., X_{i_{4m}}, t_2;\lambda)-\mu\Big(t_2;\lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)dM(t_1)dM(t_2)\\
= &\Big(\int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1}, ..., X_{i_m}, t_1; \lambda + \frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t_1; \lambda)-\mu\Big(t_1; \lambda + \frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda + \frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda)-\mu\Big(t_1; \lambda + \frac{s}{\sqrt{n}}\Big)\Big|\Big)dM(t_1)\Big)^2.\end{aligned}$$

Двоструком применом Коши-Шварцове неједнакости добија се

$$\begin{aligned}\Upsilon_n \leqslant &\int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1}, ..., X_{i_m}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big)\Big|\\
&\times\sup_{s\in S}\Big|h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^2 dM(t_1)\\
\leqslant &\int\limits_{-\infty}^{+\infty}\Big(E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1}, ..., X_{i_m}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1}, ..., X_{i_m}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big|\Big)^4\\
&\times E\Big(\sup_{s\in S}\Big|h\Big(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_{m+1}}, ..., X_{i_{2m}}, t_1; \lambda)-\mu\Big(t_1; \lambda+\frac{s}{\sqrt{n}}\Big)\Big|\Big)^4\Big)^{\frac{1}{2}} dM(t_1)\end{aligned}$$

$$
\begin{aligned}
&= \int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t_1;\lambda)-\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big|\Big)^4 dM(t_1)\\
&\leqslant \int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|h\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-h(X_{i_1},...,X_{i_m},t_1;\lambda)\Big|+\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big|\Big)^4 dM(t_1).
\end{aligned}
$$

Применом неједнакости $(a+b)^4 \leqslant 8a^4+8b^4$, слично као раније, добија се

$$
\begin{aligned}
\Upsilon_n &\leqslant 8\int\limits_{-\infty}^{+\infty} E\Big(\sup_{s\in S}\Big|g\Big(X_{i_1},...,X_{i_m},t_1;\lambda+\frac{s}{\sqrt{n}}\Big)-g(X_{i_1},...,X_{i_m},t_1;\lambda)\Big|\Big)^4 dM(t_1)\\
&+8\int\limits_{-\infty}^{+\infty}\Big(\sup_{s\in S}\Big|\mu\Big(t_1;\lambda+\frac{s}{\sqrt{n}}\Big|\Big)^4 dM(t_1)\\
&<8\Big(\varepsilon+\frac{d^4}{n^2}\int\limits_{-\infty}^{+\infty}\Big(\Big|d_1\mu\Big(t;\lambda\Big)^T\Big|^2\Big)dM(t)\Big)\\
&=8\varepsilon+o_p\Big(\frac{1}{n}\Big),\ n\to\infty.
\end{aligned}
$$

Сви остали чланови суме (2.33) се понашају као $o_p(n^{-1}),\ n\to\infty$, јер је очекивање одговарајуће функције ограничено из услова теореме.

Дакле, можемо закључити да

$$E(\sup_{s\in S} nQ''_n(s))^2 < 8\varepsilon+o_p(n^{-1}),\ n\to\infty.$$

Применом неједнакости $(a+b)^2\leqslant 2a^2+2b^2$ и свега добијеног у (2.29) следи

$$E\Big(\sup_{s\in S} nQ_n(s)\Big)^2 \leqslant 2E\Big(\sup_{s\in S} nQ'_n(s)\Big)^2+2E\Big(\sup_{s\in S} nQ''_n(s)\Big)^2 \propto \varepsilon+o_p(1),\ n\to\infty.$$

Из претходног следи да

$$E\Big(\sup_{s\in S} nQ_n(s)\Big)^2\to 0, n\to\infty.$$

Заменом у (2.28) следи први део теореме.

Како је $V_n^\star(\lambda)$ слабо дегенерисана V-статистика реда $2m$, то из теореме 2.6 следи да

$$nV_n^\star(\lambda)\xrightarrow{D}\binom{2m}{2}\sum_{k=1}^{+\infty}\upsilon_k Z_k^2,$$

где је $\{\upsilon_k\},\ k=1,2,...$, низ сопствених вредности оператора $A^\star$ дефинисаног у (2.25) и $Z_k,\ k=1,2,...$, низ независних случајних величина са стандардном нормалном расподелом. Према томе, уз претходно доказано, следи други део тврђења теореме. ■

Може се приметити да утицај оцене параметра се огледа у члану $d_1\mu(t;\lambda)$. У случају када је тај члан једнак 0 за свако t, статистике $nV_n(\widehat{\lambda})$ и $nV_n(\lambda)$ имају исту граничну расподелу. Односно гранична расподела од $nV_n(\widehat{\lambda})$ је иста као када је λ познато. У супротном, постоји утицај оцене параметра на граничну расподелу.

Сличан резултат следи и за $U-$статистике са оцењеним параметрима. Дефинишемо

$$U_n(\widehat{\lambda}_n)=\frac{1}{\binom{n}{2m}}\sum_{1\leqslant i_1<\ldots<i_{2m}\leqslant n}\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n),$$

где је $\Phi(x_1,...,x_{2m};\widehat{\lambda}_n)$ дефинисано у (2.21) и

$$U_n(\lambda)=\frac{1}{\binom{n}{2m}}\sum_{1\leqslant i_1<\ldots<i_{2m}\leqslant n}\Phi_\star(X_{i_1},...,X_{i_{2m}};\lambda),$$

где је $\Phi_\star(x_1,...,x_{2m};\lambda)$ дефинисано у (2.23).

Теорема 2.10 (Цупарић, Милошевић, Обрадовић 2020) *Нека је $X_1,...,X_n$ узорак из расӣоделе F. Нека важе услови ӣреӣходне ӣеореме и доdаӣно*

$$E\Big(\int_{-\infty}^{+\infty}|h_1(X_1,t;\lambda)|^2dM(t)\Big)<\infty,$$

ӣdе је $h_1(x,t;\lambda)=E_\lambda(h(x,X_2,...,X_m,t;\lambda))$. Тада

$$nU_n(\widehat{\lambda}_n)\overset{D}{\to}\frac{2m(2m-1)}{2}\sum_{i=1}^{\infty}(\upsilon_kZ_k^2-\upsilon_k'),$$

ӣdе је $\{\upsilon_k\},k=1,2,...,$ низ соӣсӣвених вредносӣи оӣераӣора $A^\star$ дефинисаноӣ у (2.25), $\{\upsilon_k'\},k=1,2,...,$ је низ соӣсӣвених вредносӣи оӣераӣора A дефинисаноӣ у (2.13) и $Z_k,k=1,2,...,$ су независне случајне величине са сӣандардном нормалном расӣоделом.

Доказ: Из теореме 2.4 следи да

$$nU_n^\star(\lambda)\overset{D}{\to}\frac{2m(2m-1)}{2}\sum_{k=1}^{\infty}\upsilon_k(Z_k^2-1).$$

Према томе, довољно је доказати да

$$n(U_n(\widehat{\lambda}_n)-U_n^\star(\lambda))\overset{P}{\to}\binom{2m}{2}\sum_{k=1}^{\infty}(\upsilon_k-\upsilon_k'). \tag{2.34}$$

Размотримо разлику

$$\begin{aligned}n(U_n(\widehat{\lambda}_n)-U_n^\star(\lambda))&=\frac{n}{n(n-1)\cdots(n-2m+1)}\sum_{\{i_1,\cdots,i_{2m}\}\in\mathcal{I}'}\Big(\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n)-\Phi_\star(X_{i_1},...,X_{i_{2m}};\lambda)\Big)\\&=\frac{n}{n(n-1)\cdots(n-2m+1)}\Big(n^{2m}(V_n(\widehat{\lambda}_n)-V_n^\star(\lambda))-\sum_{\{i_1,...,i_{2m}\}\in\mathcal{I}''}\Big(\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n)-\Phi_\star(X_{i_1},...,X_{i_{2m}};\lambda)\Big)\Big)\end{aligned} \tag{2.35}$$

где је $\mathcal{I}'$ скуп $2m$-торки индекса који се не понављају, односно $i_1 \neq \cdots \neq i_{2m}$, док је $\mathcal{I}''$ комплемент од $\mathcal{I}'$, односно скуп $2m-$торки индекса са бар једним понављањем. Из претходне теореме следи да $n(V_n(\widehat{\lambda}_n) - V_n(\lambda)) \xrightarrow{P} 0$. Код другог члана претходног израза, односно израза

$$\frac{1}{(n-1)\cdots(n-2m+1)} \sum_{\substack{1\leqslant i_1,\dots,i_{2m}\leqslant n\\ \exists j\neq k\ i_j=i_k}} \Big(\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n) - \Phi^*(X_{i_1},...,X_{i_{2m}};\lambda)\Big), \tag{2.36}$$

разликујемо случајеве:

1. ако се тачно два индекса поклапају ($i_j = i_k \neq i_l$, за $l \in \{1,...,n\}\backslash\{k,j\}$ и $j \neq k$): У том случају одговарајући чланови суме (2.36) су једнаки

$$\begin{aligned}
\Xi_n =& \frac{n(n-1)\cdots(n-2m+2)}{(n-1)\cdots(n-2m+1)}(U_{1n}(\widehat{\lambda}_n) - U^{\star}_{1n}(\lambda)) = \frac{n}{n-2m+1}(U_{1n}(\widehat{\lambda}_n) - U^{\star}_{1n}(\lambda))\\
=& \frac{n}{n-2m+1}\frac{\binom{2m}{2}}{\binom{n}{2m-1}} \sum_{i_1\neq i_2\neq\cdots\neq i_{2m-1}} \Phi(X_{i_1},X_{i_1},X_{i_2},...,X_{i_{2m-1}};\widehat{\lambda}_n)\\
&- \frac{n}{n-2m+1}\frac{\binom{2m}{2}}{\binom{n}{2m-1}} \sum_{i_1\neq i_2\neq\cdots\neq i_{2m-1}} \Phi_{\star}(X_{i_1},X_{i_1},X_{i_2},...,X_{i_{2m-1}};\lambda)\\
=& \frac{n}{n-2m+1}\frac{1}{2m-1}\frac{\binom{2m}{2}}{\binom{n}{2m-1}} \sum_{i_1\neq i_2\neq\cdots\neq i_{2m-1}} \phi(X_{i_1},X_{i_2},...,X_{i_{2m-1}};\widehat{\lambda}_n)\\
&- \frac{n}{n-2m+1}\frac{1}{2m-1}\frac{\binom{2m}{2}}{\binom{n}{2m-1}} \sum_{i_1\neq i_2\neq\cdots\neq i_{2m-1}} \phi_{\star}(X_{i_1},X_{i_2},...,X_{i_{2m-1}};\lambda),
\end{aligned}$$

где су $\phi(X_{i_1},...,X_{i_{2m-1}};\widehat{\lambda}_n)$ и $\phi_{\star}(X_{i_1},...,X_{i_{2m-1}};\lambda)$ симетризације функција $\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n)$ и $\Phi_{\star}(X_{i_1},...,X_{i_{2m}};\lambda)$ у којима се тачно два индекса поклапају. Тада користећи закон великих бројева за $U-$статистике (теорема 2.2) и закон великих бројева за $U-$статистике са оцењеним параметром (теорема 2.7),

$$\Xi_n \xrightarrow{P} \binom{2m}{2}\frac{1}{2m-1}\big(E\phi(X_{i_1},...,X_{i_{2m-1}};\widehat{\lambda}_n) - E\phi_*(X_{i_1},X_{i_2},...,X_{i_{2m-1}};\lambda)\big).$$

Приметимо да је

$$\begin{aligned}
&\binom{2m}{2}\frac{1}{2m-1}E\phi_{\star}(X_{i_1},X_{i_2},...,X_{i_{2m-1}};\lambda)\\
&= \binom{2m}{2}\frac{m^2}{2m(2m-1)}E\Bigg(\int\limits_{-\infty}^{+\infty}\Bigg(h(X_1,X_2,...,X_{i_m},t;\lambda) + d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{j\in\{1,...,m\}}\alpha(X_j)\Bigg)\\
&\qquad\times\Bigg(h(X_1,X_{m+1},...,X_{2m-1},t;\lambda) + d_1\mu(t;\lambda)^T\frac{1}{m}\sum_{k\in\{1,m+1,..,2m-1\}}\alpha(X_k)\Bigg)dM(t)\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= \binom{2m}{2} \frac{m^2}{2m(2m-1)} \int\limits_{-\infty}^{+\infty} E\Bigg(\Bigg(h(X_1, X_2, ..., X_{i_m}, t; \lambda) + d_1\mu(t;\lambda)^T \frac{1}{m} \sum_{j\in\{1,...,m\}} \alpha(X_j)\Bigg) \\
&\quad \times \Bigg(h(X_1, X_{m+1}, ..., X_{2m-1}, t; \lambda) + d_1\mu(t;\lambda)^T \frac{1}{m} \sum_{k\in\{1,m+1,..,2m-1\}} \alpha(X_k)\Bigg)\Bigg) dM(t) \\
&= \binom{2m}{2} \frac{m^2}{2m(2m-1)} \int\limits_{-\infty}^{+\infty} E\Bigg(h_1(X_1, t; \lambda) + d_1\mu(t;\lambda)^T \frac{1}{m}\alpha(X_1)\Bigg)^2 dM(t) \\
&= \binom{2m}{2} E\varphi_2(X_1, X_1) = \frac{2m(2m-1)}{2} \sum_{k=1}^{\infty} \upsilon_k,
\end{aligned}
$$

ако је оператор A^* нуклеаран, односно ако је $\sum\limits_{k=1}^{\infty} |\upsilon_k| < \infty$. То ће важити ако је $E|\varphi_2^*(X_1, X_1)| < \infty$ (за више детаља видети *Korolyuk and Borovskikh (1994)*). Приметимо да

$$
\begin{aligned}
E|\varphi_2^*(X_1, X_1)| &= E\Bigg|\int\limits_{-\infty}^{+\infty} \Bigg(h_1(X_1, t; \lambda) + d_1\mu(t;\lambda)^T \frac{1}{m}\alpha(X_1)\Bigg)^2 dM(t)\Bigg| \\
&\leqslant E\int\limits_{-\infty}^{+\infty} \Bigg|h_1(X_1, t; \lambda) + d_1\mu(t;\lambda)^T \frac{1}{m}\alpha(X_1)\Bigg|^2 dM(t) \\
&\leqslant 2E\int\limits_{-\infty}^{+\infty} \Bigg|h_1(X_1, t; \lambda)\Bigg|^2 dM(t) + 2E\int\limits_{-\infty}^{+\infty} \Bigg|d_1\mu(t;\lambda)^T \frac{1}{m}\alpha(X_1)\Bigg|^2 dM(t) \\
&\leqslant 2E\int\limits_{-\infty}^{+\infty} \Bigg|h_1(X_1, t; \lambda)\Bigg|^2 dM(t) + \frac{2}{m^2} E\Bigg|\alpha(X_1)\Bigg|^2 \cdot \int\limits_{-\infty}^{+\infty} \Bigg|d_1\mu(t;\lambda)^T\Bigg|^2 dM(t).
\end{aligned}
$$

Први члан претходног израза је коначан из услова теореме, а из услова 2.1 и 2.2 други члан је коначан.

Слично као у претходном случају следи да $E\phi(X_{i_1}, ..., X_{i_{2m-1}}; \widehat{\lambda}) = (2m-1)\sum\limits_{k=1}^{\infty} \upsilon'_k$. Дакле, можемо закључити да

$$
\Xi_n \xrightarrow{P} \binom{2m}{2} \sum_{k=1}^{\infty} (\upsilon'_k - \upsilon_k).
$$

2. Ако се тачно $r, 1 < r < 2m$, индекса поклапа: У том случају одговарајући чланови суме (2.36) су пропорционални са

$$
\begin{aligned}
&\frac{n(n-1)\cdots(n-2m+r+1)}{(n-1)\cdots(n-2m+1)} (U_{2n}(\widehat{\lambda}_n) + U^{\star}_{2n}(\lambda)) \\
&= \frac{n}{(n-2m+1)\cdots(n-2m+r)} (U_{2n}(\widehat{\lambda}_n) + U^{\star}_{2n}(\lambda)),
\end{aligned}
$$

где су $U_{2n}(\widehat{\lambda}_n)$ и $U^{\star}_{2n}(\lambda)$ $U-$статистике са језгрима $\phi(X_{i_1},...,X_{i_{2m-r}};\widehat{\lambda}_n)$ и $\phi^{\star}(X_{i_1},...,X_{i_{2m-r}};\lambda)$, редом, која представљају одговарајуће симетризације функција $\Phi(X_{i_1},...,X_{i_{2m}};\widehat{\lambda}_n)$ и $\Phi_{\star}(X_{i_1},...,X_{i_{2m}};\lambda)$ у којима се тачно r индекса поклапа. Дакле, користећи закон великих бројева за $U-$статистике, следи да ти чланови суме конвергирају у вероватноћи ка 0.

Комбиновањем ових резултата у (2.35) добија се (2.34). Користећи теорему Слуцког доказ теореме је комплетиран. ■

Сличан резултат уз коришћење другачијег метода за $U-$статистике је добијен у раду *Arcones (2007)*.

2.5 $V-$емпиријске функције

$V-$емпиријске функције или $V-$емпиријски процеси су фамилије $V-$статистика посматране као случајни процеси.

Дефиниција 2.3 *Нека је $X_1,...,X_n$ узорак из расподеле F и нека је $\mathcal{H}$ фамилија реалних функција од m аргумената. $V-$емпиријска функција је колекција $V-$статистика за одговарајуће језгро Φ из $\mathcal{H}$, односно*

$$V_n(t)=\frac{1}{n^m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t).$$

Ако се фамилија $\mathcal{H}$ састоји из само једне функције, односно ако је t фиксирано, добија се стандардна $V-$статистика са језгром Φ. Специјалан случај фамилије $\mathcal{H}$ је фамилија индикатора повезаних са неком функцијом h, односно $\{\mathrm{I}\{h(\cdot,...,\cdot)\leqslant t\}, t\in\mathbb{R}\}$. На тај начин се добија $V-$емпиријска функција расподеле дефинисана са

$$H_n(t)=\frac{1}{n^m}\sum_{i_1,...,i_m=1}^{n}\mathrm{I}\{h(X_{i_1},...,X_{i_m})\leqslant t\}.$$

За фиксирано t, $H_n(t)$ је $V-$статистика, при чему је очекивање језгра једнако

$$H(t)=P\{h(X_1,...,X_m)\leqslant t\}.$$

Ако је $m=1$, а $h(x)=x$, H_n је уобичајна емпиријска функција расподеле.

Овако дефинисане $V-$емпиријске функције расподела, као и одговарајуће $U-$емпиријске функције расподела, су често проучаване у литератури. Тако на пример у раду *Silverman (1983)* добијени су резултати везани за слабу конвергенцију тих процеса, а затим је у раду *Helmers et al. (1988)* показано да важи Гливенко-Кантелијева теорема. Такође, ови процеси су проучавани и кроз практичне примене у раду *Serfling (1984)*, где је показана асимптотска нормалност одређених статистичких функционала који се могу представити као функције $V-$емпиријске функције расподеле, као и у радовима у којима су тест статистике представљене као функције $V-$емпиријске функције расподеле, као на пример *Jovanović et al. (2015)* и *Milošević and Obradović (2016b).*

Друга често коришћена фамилија је повезана са Лапласовом трансформацијом неке функције h. Нека је $X_1, ..., X_n$ узорак из ненегативне расподеле F. Нека је $h(X_1, ..., X_m)$ нека функција случајних величина. Лапласова трансформација те функције је дефинисана са

$$\mathcal{L}(t) = \int_0^{+\infty} \cdots \int_0^{+\infty} e^{-th(x_1,...,x_m)} \prod_{i=1}^{m} dF(x_i). \tag{2.37}$$

Тада је са

$$\mathcal{L}_n(t) = \int_0^{+\infty} \cdots \int_0^{+\infty} e^{-th(x_1,...,x_m)} \prod_{i=1}^{m} dF_n(x_i) = \frac{1}{n^m} \sum_{i_1,...,i_m=1}^{n} e^{-th(X_{i_1},...,X_{i_m})} \tag{2.38}$$

дефинисана V-емпиријска Лапласова трансформација функције h.

Као и код $V-$статистика, пожељно је знати асимптотска својства $V-$емпиријских процеса. У наставку ће бити приказане теореме везане за слабу конвергенцију процеса, које се могу применити за добијање граничних процеса $V-$емпиријских функција дефинисаних на том простору. Више детаља у случају функција дефинисаних на простору $C[0, 1]$ се могу наћи у књизи *Billingsley (1968)*, док се у случају простора $C[0, \infty)$ могу наћи у *Karatzas and Shreve (1991).*

На почетку уводимо неке основне појмове неопходне за разумевање теорема. Више детаља се може наћи у *Karatzas and Shreve (1991).*

Дефиниција 2.4 *Нека је (S, ρ) метрички простор са Бореловом $\sigma-$алгебром $\mathcal{B}(S)$. Нека је $\{P_n\}$ низ вероватносних мера на $(S, \mathcal{B}(S))$ и нека је P нека друга мера на том простору. Тада $\{P_n\}$ конвергира слабо ка P ако и само ако*

$$\lim_{n\to\infty} \int_S f(s) dP_n(s) = \int_S f(s) dP(s),$$

за сваку ограничену, непрекидну реалну функцију f на S.

Специјално, следи да је P вероватносна мера и да је јединствена.

Дефиниција 2.5 *Нека је $(\Omega_n, \mathcal{A}_n, P_n)$ низ простора вероватноћа и на сваком од њих размотримо случајну величину X_n са вредностима у метричком простору (S, ρ). Нека је $(\Omega, \mathcal{A}, P)$ други простор вероватноћа на којем је дефинисана случајна величина X која узима вредности у (S, ρ). Тада $\{X_n\}$ конвергира ка X у расподели ако низ вероватносних мера $\{P_n X_n^{-1}\}$ конвергира слабо ка мери PX^{-1}.*

Да би се ближе одредила слаба конвергенција у простору непрекидних функција, треба користити својства тог простора и према томе уводимо и следећу дефиницију.

Дефиниција 2.6 *Нека је (S,ρ) метрички простор и Π фамилија вероватносних мера на $(S,\mathcal{B}(S))$, где је $\mathcal{B}(S)$ Борелова $\sigma-$алгебра. Кажемо да је Π релативно компактан ако сваки низ елемената из Π садржи слабо конвергентан подниз. Кажемо да је Π густ ако за свако $\varepsilon > 0$, постоји компактни скуп $K \subset S$ такав да $P(K) \geqslant 1-\varepsilon$, за свако $P \in \Pi$.*

Следећа теорема је позната као Прохоровљева теорема и њен доказ се може наћи рецимо у *Billingsley (1968).*

Теорема 2.11 *Нека је Π фамилија вероватносних мера на комплетном, сепарабилном метричком простору S. Фамилија је релативно компактна ако и само ако је густа.*

Размотримо простор $C[0,\infty)$ непрекидних функција на $[0,\infty)$ са метриком у којој је растојање између две тачке x и y (односно две функције x и y од $t \in [0,\infty)$) дефинисано са

$$\rho(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\{1, \max_{0\leqslant t\leqslant n} |x(t)-y(t)|\}.$$

С обзиром да је простор $C[0,\infty)$ сепарабилан и комплетан, из Прохоровљеве теореме 2.11 следи да релативна компактност фамилије вероватносних мера на $(C[0,\infty),\mathcal{B}(C[0,\infty)))$ је еквивалентна томе да је фамилија густа. Одатле произилази следеће тврђење.

Теорема 2.12 *Нека су P_n и P вероватносне мере на $(C[0,\infty),\mathcal{B}(C[0,\infty)))$. Ако коначнодимензионалне расподеле од P_n конвергирају коначнодимензионалним расподелама од P и ако је $\{P_n\}$ густ, онда P_n слабо конвергира ка P.*

Следећа теорема даје услове који обезбеђују да је низ вероватносних мера густ и њен доказ се може пронаћи у *Karatzas and Shreve (1991).*

Теорема 2.13 *Низ вероватносних мера $\{P_n\}$ на $(C[0,\infty),\mathcal{B}(C[0,\infty)))$ је густ ако и само ако*

$$\lim_{\varepsilon\to\infty} \sup_{n\geqslant 1} P_n\{x : |x(0)| > \varepsilon\} = 0, \tag{2.39}$$

$$\lim_{\delta\to 0} \sup_{n\geqslant 1} P_n\{x : \max_{\substack{|s-t|\leqslant\delta \\ 0\leqslant s,t\leqslant T}} |x(s)-x(t)| > \varepsilon\} = 0,\ \forall T > 0,\ \varepsilon > 0. \tag{2.40}$$

Нека је $\{X_n\}$ низ случајних функција дефинисаних на простору $(\Omega_n, \mathcal{A}_n, P_n)$ које узимају вредности на скупу $C[0,\infty)$. Тај низ је густ ако је низ одговарајућих расподела густ. Следећа последица даје критеријум за утврђивање да је низ случајних величина $\{X_n\}$ густ на простору $C[0,\infty)$ (за више детаља видети *Karatzas and Shreve (1991)*).

Последица 2.1 *Нека је $\{X_n\}$ низ непрекидних случајних процеса $X_n = \{X_n(t), t \geqslant 0\}$ на $(\Omega,\mathcal{A},P)$, који задовољавају услове*

1. $\sup_{n\geqslant 1} E|X_n(0)|^{\nu} = M < \infty$,

2. $\sup_{n\geqslant 1} E|X_n(s) - X_n(t)|^{\alpha} \leqslant \Upsilon(T)|t-s|^{1+\beta}$, $\forall T > 0$ *и* $0 \leqslant s, t \leqslant T$

за неке позитивне константе α, β, ν *и функцију* $\Upsilon(T)$ *која зависи од* $T > 0$. *Тада је одговарајући низ веровантносних мера* P_n *на* $(C[0,\infty), \mathcal{B}(C[0,\infty)))$ *густ.*

Ситуација је нешто компликованија уколико уместо простора непрекидних функција C имамо простор D функција x које су непрекидне здесна и имају граничну вредност са леве стране, односно

1. за $0 \leqslant t < 1$, $x(t+) = \lim_{s\to t} x(s)$ постоји и $x(t+) = x(t)$;
2. за $0 < t \leqslant 1$, $x(t-) = \lim_{s\to t} x(s)$ постоји.

За функцију $x \in D$ се каже да има прекид прве врсте у t ако $x(t-)$ и $x(t+)$ постоје, али се разликују, и $x(t)$ се налази између њих. Сви прекиди у простору D су прве врсте, а захтев $x(t) = x(t+)$ представља нормализацију.

За претходне теореме које се односе на процесе дефинисане на простору $C[0,\infty)$, аналогне теореме важе и за процесе дефинисане на простору $D[0,1]$ са Скороходовом метриком. С обзиром да ћемо у наставку користити само процесе који су непрекидне функције, те теореме нећемо наводити али се оне могу пронаћи у *Billingsley (1968).*

Размотримо два примера уведена раније. У случају $V-$емпиријских функција расподеле, у питању је случај функција са прекидима прве врсте. Са друге стране, у случају $V-$емпиријских Лапласових трансформација, у питању су непрекидне функције па за доказивање слабе конвергенције треба проверити услове теореме 2.12, односно 2.13.

На крају овог поглавља наводимо још једну теорему, познату као теорема Слуцког за случајне процесе, која се често користи приликом одређивања граничне расподеле неке тест статистике. Њен доказ се заснива на Прохоровљевој теореми и теореми о непрекидном пресликавању и може се пронаћи у *Kosorok (2008).*

Теорема 2.14 *Нека су* $\{X_n\}$ *и* $\{Y_n\}$ *низови случајних процеса,* $X_n = \{X_n(t), t \geqslant 0\}$ *и* $Y_n = \{Y_n(t), t \geqslant 0\}$. *Нека* $X_n \xrightarrow{D} X$ *и* $Y_n \xrightarrow{P} c$, *где је* X *сепарабилан а* c *фиксирана константа.*

1. *Тада следи да* $(X_n, Y_n) \xrightarrow{D} (X, c)$.
2. *Ако су* X_n *и* Y_n *дефинисани у истом метричком простору, онда* $X_n + Y_n \xrightarrow{D} X + c$.
3. *Претпоставимо додатно да је* Y_n *скалар. Онда, кад год је* $c \in \mathbb{R}$, $Y_n X_n \xrightarrow{D} cX$. *Такође, ако је* $c \neq 0$, *онда* $\frac{X_n}{Y_n} \xrightarrow{D} \frac{X}{c}$.

Глава 3

ТЕСТОВИ САГЛАСНОСТИ ЗАСНОВАНИ НА L^2 И L^∞ РАСТОЈАЊИМА

Тестирање сагласности је једна од најзначајнијих области непараметарске статистике. Начини за формирање тест статистика тестова сагласности су различити. Када се тест формира за посебну класу расподела, он у своју тест статистику укључује неку особину претпостављене расподеле. Велики број тестова груписаних по особинама које се користе за њихово формирање, али груписаних и по самим расподелама за чије се тестирање користе, може се наћи на пример у *D'Agostino and Stephens (1986)*.

У последње време често коришћен и посебно погодан начин за формирање тестова сагласности је коришћење карактеризација расподеле. Карактеризације расподела имају једноставан и јасан облик када се адекватно интерпретирају уз помоћ елементарних концепата математичке анализе и једноставних функционалних једначина. Један од начина да се карактеризациона теорема укључи у тест статистику је кроз различите $V-$емпиријске трансформације. Једноставном заменом карактеризације овако формирани тестови се могу прилагодити за тестирање различитих расподела.

Приликом формирања теста пожељно је да он буде слободан од параметра расподеле која се тестира у нултој хипотези. Та особина омогућава да се тестира сложена хипотеза без коришћења компликованих метода реузорковања за апроксимацију расподеле статистике. У случају када сама тест статистика није слободна од параметра, често је могуће извршити одређену трансформацију узорка тако да се статистика ослободи од параметра расподеле из нулте хипотезе.

У наставку ћемо размотрити тестове сагласности са експоненцијалном расподелом, мада се наведени тестови могу прилагодити и за тестирање других расподела, уз коришћење одговарајућих карактеризација, и сви резултати ће важити уз мање модификације.

Експоненцијална расподела се појавила релативно касно у математичкој статистици. Кондо 1930. године у свом раду (*Kondo (1930)*) указује на експоненцијалну расподелу као Пирсонову расподелу типа X, док је Сухатме вероватно први који је у свом раду из 1937. (*Sukhatme (1937)*) предложио експоненцијалну расподелу, названу χ^2 са два степена слободе, као алтернативу нормалној расподели у ситуацијама где популација доста одступа од нормалне расподеле. У том раду је одређена расподела разлике између статистика поретка случајних величина из експоненцијалне расподеле, одакле су касније проистекла многа битна својства експоненцијалне расподеле. Међутим, експоненцијална расподела се у другим областима, као што су актуарство, биологија и инжењерске науке, користила више од 20 година пре него што се појавила у самој статистици. У статистици, експоненцијална расподела је прихваћена тек 50-тих година XX века великим бројем значајних радова који показују како њену примену тако и њена својства.

Иако је касно почела да се користи, експоненцијална расподела је данас, поред нормалне, једна од најшире коришћених расподела. Због тога постоји мноштво различитих тестова експоненцијалности који се базирају на различитим својствима. Стандардне процедуре за проверу експоненцијалности расподеле су тестови Колмогоров-Смирнова и Крамер-фон Мизеса, који користе емпиријску функцију расподеле. Међутим, развијени су и многи тестови специјално за експоненцијалну расподелу који се базирају на различитим карактеризацијама. Први тестови експоненцијалности су засновани на особини одсуства меморије (на пример, *Angus (1982), Ahmad and Alwasel (1999), Nikitin (1996)*). Још неке особине на којима се базирају тестови су максимална корелација (на пример, *Fortiana and Grané (2002)*), ентропија (на пример, *Choi et al. (2004), Ebrahimi et al. (1992)*), различите интегралне трансформације (на пример, *Jovanović et al. (2015), Milošević and Obradović (2016a)*) и друге.

У наставку ће прво бити приказане карактеризације експоненцијалне расподеле које ће затим бити коришћене за формирање тест статистика. Тестови експоненцијалности који су развијени представљају L^2 и L^∞ растојања између одређених трансформација случајних величина које се јављају у карактеризацијама. Одређена су њихова својства, а код тестова који зависе од параметра подешавања приказан је метод за одређивање оптималне вредности тог параметра. У последњем делу овог поглавља, као мера квалитета формираних тестова ће бити размотрена емпиријска моћ тестова, која се често користи у случају малих узорака и узорака умерене величине.

3.1 Карактеризације

Нека је $X_1, ..., X_n$ узорак. Претпоставимо да се тестира хипотеза $H_0(F \in \mathcal{F}_0)$ против алтернативе $H_1(F \in \mathcal{F}' - \mathcal{F}_0)$, где је $\mathcal{F}'$ нека велика класа расподела а $\mathcal{F}_0$ параметарска фамилија. Ако је процедура тестирања заснована на карактеризационим својствима од $\mathcal{F}_0$, онда је познато да: статистика $T(X_1, ..., X_n)$ при H_0 има расподелу Q, где је Q јединствено одређена функција расподеле, ако и само ако $F \in \mathcal{F}_0$.

За формирање тест статистика посебно су погодне карактеризације једнаке расподељености. Нека је дат узорак $X_1, .., X_n$ из расподеле F. Претпоставимо да је из особина расподеле F познато да за неке две хомогене функције ψ_1 и ψ_2 које зависе од m случајних величина важи једнакост у расподели. Тада карактеризација даје обрнути смер, односно ако су $\psi_1(X_1, ..., X_m)$ и $\psi_2(X_1, ..., X_m)$ једнаке у расподели, онда узорак потиче из расподеле F.

С обзиром да је и сама експоненцијална расподела релативно касно почела да се користи у статистици, није чудно што су се и карактеризације те расподеле релативно касно појавиле. Галамбош и Коц у својој књизи (*Galambos and Kotz (1978)*) претпостављају да се прва карактеризација експоненцијалне расподеле јавила у XIX веку у литератури повезаној са физиком, тачније везаној за теорију атома. Међутим, прве званично наведене карактеризације су повезане са теоријом екстремних вредности (на пример, *Fisher and Tippett (1928)*).

У наставку, за формирање тест статистика за тестирање експоненцијалне расподеле биће коришћене следеће две карактеризације, представљене у радовима *Puri and Rubin (1970)* и *Desu (1971)*, редом.

Карактеризација 3.1 *Нека је X, X_1, X_2 узорак из расподеле са густином f. Тада X и $|X_1 - X_2|$ имају исту расподелу ако и само ако $f(x) = \lambda e^{-\lambda x}$, $x \geqslant 0$, за неко $\lambda > 0$.*

Карактеризација 3.2 *Нека је $X, X_1, X_2, ..., X_m$ узорак обима $m + 1$ из расподеле са функцијом расподеле F. Тада за свако m, X и $m \min\{X_1, ..., X_m\}$ имају исту расподелу ако и само ако $F(x) = 1 - e^{-\lambda x}$, $x \geqslant 0$, за неко $\lambda > 0$.*

У наставку ћемо у случају Десуове карактеризације користити специјалан случај $m = 2$. Често се заправо тај специјалан случај карактеризације назива Десуовом карактеризацијом (иако једнакост само за $m = 2$ заправо не карактерише експоненцијалну расподелу).

3.2 Тестови L^2–типа засновани на V–емпиријској функцији расподеле

Нека је $X_1, X_2, ..., X_n$ узорак из непознате апсолутно непрекидне расподеле F. За тестирање нулте хипотезе $H_0 : F(x) = 1 - e^{-\lambda x}, x \geqslant 0, \lambda > 0$, у раду аутора (*Cuparić et al. (2020)*) предложена је следећа фамилија тест статистика:

$$W_n = \int\limits_0^{+\infty} \left(H_n^{(1)}(t) - H_n^{(2)}(t)\right)^2 dF_n(t), \tag{3.1}$$

где су

$$H_n^{(1)}(t) = \frac{1}{n^m} \sum_{i_1,...,i_m=1}^{n} \mathrm{I}\{\psi_1(X_{i_1}, ..., X_{i_m}) < t\} \tag{3.2}$$

и

$$H_n^{(2)}(t) = \frac{1}{n^m} \sum_{i_1,...,i_m=1}^{n} \mathrm{I}\{\psi_2(X_{i_1}, ..., X_{i_m}) < t\} \tag{3.3}$$

$V-$емпиријске функције расподеле случајних величина ψ_1 и ψ_2 из карактеризације. Ради добијања асимптотске расподеле, прво приметимо да се након интеграције добија следећи облик тест статистике

$$\begin{aligned} W_n &= \int_0^{+\infty} \Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\mathrm{I}\{\psi_1(X_{i_1},...,X_{i_m})<t\} - \frac{1}{n^m}\sum_{i_1,\dots,i_m}\mathrm{I}\{\psi_2(X_{i_1},...,X_{i_m})<t\}\Big)^2 dF_n(t) \\ &= \frac{1}{n^{2m}}\sum_{i_1,\dots,i_{2m}}\int_0^{+\infty}\Big(\mathrm{I}\{\psi_1(X_{i_1},...,X_{i_m})<t\}-\mathrm{I}\{\psi_2(X_{i_1},...,X_{i_m})<t\}\Big) \\ &\qquad\times\Big(\mathrm{I}\{\psi_1(X_{i_{m+1}},...,X_{i_{2m}})<t\}-\mathrm{I}\{\psi_2(X_{i_{m+1}},...,X_{i_{2m}})<t\}\Big)dF_n(t) \\ &= \frac{1}{n^{2m}}\sum_{i_1,\dots,i_{2m}}\int_0^{+\infty} h(X_{i_1},...,X_{i_m},t)h(X_{i_{m+1}},...,X_{i_{2m}},t)dF_n(t) \\ &= \frac{1}{n^{2m+1}}\sum_{i_1,\dots,i_{2m+1}}\Phi_W(X_{i_1},X_{i_2},...,X_{i_{2m+1}}), \end{aligned}$$

где је

$$\begin{aligned} \Phi_W(X_1,...,X_{2m+1}) = \frac{1}{(2m+1)!}\sum_{\pi\in\Pi(2m+1)} &\Big(\mathrm{I}\{\psi_1(X_{\pi(1)},...,X_{\pi(m)})<X_{\pi(2m+1)}\} \\ &-\mathrm{I}\{\psi_2(X_{\pi(1)},...,X_{\pi(m)})<X_{\pi(2m+1)}\}\Big)\Big(\mathrm{I}\{\psi_1(X_{\pi(m+1)},...,X_{\pi(2m)})<X_{\pi(2m+1)}\} \\ &-\mathrm{I}\{\psi_2(X_{\pi(m+1)},...,X_{\pi(2m)})<X_{\pi(2m+1)}\}\Big) \end{aligned}$$

и $\Pi(2m+1)$ скуп свих $(2m+1)!$ пермутација скупа $\{1,2,...,2m+1\}$.

Односно, статистика W_n се може написати као $V-$статистика реда $2m+1$. При нултој хипотези ове тест статистике су слободне од параметра нулте расподеле, што следи из тога да ако је узорак $X_1,...,X_n$ из експоненцијалне расподеле са параметром λ онда је $Y_1,...,Y_n$, где је $Y_i=\lambda X_i,\ i=1,...,n$, узорак из експоненцијалне расподеле са параметром 1. Тада из хомогености функције ψ_1 следи да је

$$\mathrm{I}\{\psi_1(X_1,...,X_m)<X_{2m+1}\}=\mathrm{I}\Big\{\psi_1\Big(\frac{Y_1}{\lambda},...,\frac{Y_m}{\lambda}\Big)<\frac{Y_{2m+1}}{\lambda}\Big\}=\mathrm{I}\{\psi_1(Y_1,...,Y_m)<Y_{2m+1}\}.$$

Аналогно важи и за све остале индикаторе које се јављају у језгру. То својство омогућава тестирање сложених хипотеза. Нулта хипотеза се одбацује за велике вредности тест статистике, односно критична област је облика $\{W_n \geqslant q_\alpha\}$, где је q_α $1-\alpha$ квантил одговарајуће расподеле.

Приметимо да због начина на који је формирана тест статистика, односно карakтеризација једнаке расподељености, при нултој хипотези важи

$$\begin{aligned}
\varphi_1(x_1) &= E(\Phi_W(X_1, X_2..., X_{2m+1})|X_1 = x_1) \\
&= \frac{(2m)!}{(2m+1)!} P\{\psi_1(x_1, X_2, ..., X_m) < X_{2m+1}, \psi_1(X_{m+1}, ..., X_{2m}) < X_{2m+1}\} \\
&+ \frac{1}{(2m+1)!} P\{\psi_1(X_2, ..., X_{m+1}) < x_1, \psi_1(X_{m+2}, ..., X_{2m+1}) < x_1\} \\
&+ \frac{(2m)!}{(2m+1)!} P\{\psi_2(x_1, X_2..., X_m) < X_{2m+1}, \psi_2(X_{m+1}, ..., X_{2m}) < X_{2m+1}\} \\
&+ \frac{1}{(2m+1)!} P\{\psi_1(X_2, ..., X_{m+1}) < x_1, \psi_1(X_{m+2}, ..., X_{2m+1}) < x_1\} \\
&- \frac{2m \cdot (2m)!}{(2m+1)!} P\{\psi_1(x_1, X_2..., X_m) < X_{2m+1}, \psi_2(X_{m+1}, ..., X_{2m}) < X_{2m+1}\} \\
&- \frac{2}{(2m+1)!} P\{\psi_1(X_2, ..., X_{m+1}) < x_1, \psi_1(X_{m+2}, ..., X_{2m+1}) < x_1\} \\
&- \frac{2m \cdot (2m)!}{(2m+1)!} P\{\psi_2(x_1, X_2..., X_m) < X_{2m+1}, \psi_1(X_{m+1}, ..., X_{2m}) < X_{2m+1}\} = 0.
\end{aligned}$$

Дакле, тест статистика је дегенерисана $V-$статистика реда $2m+1$. С друге стране, друга пројекција језгра је различита од нуле и може се одредити за конкретне ψ_1 и ψ_2. Из другог дела теореме 2.6 следи да, при H_0,

$$nW_n \overset{D}{\to} 10 \sum_{k=1}^{\infty} \upsilon_k Z_k^2, \tag{3.4}$$

где је $\{\upsilon_k\}, k = 1, 2, ...,$ низ сопствених вредности оператора A дефинисаног у (2.13), и $Z_k,\ k = 1, 2, ...,$ независне случајне величине са стандардном нормалном расподелом. Сопствене вредности $\{\upsilon_k\}, k = 1, 2, ...,$ се не могу експлицитно одредити. У поглављу 4.3 ће бити приказан нумерички метод којим се наведени низ сопствених вредности може одредити.

Са друге стране, можемо размотрити понашање статистике при фиксираној алтернативној хипотези (не зависи од n) која има очекивање μ. У том случају прва пројекција статистике неће бити једнака нули, односно статистика W_n је недегенерисана статистика и њена асимптотска расподела следи из првог дела теореме 2.6. Тада важи да

$$\sqrt{n}(W_n - \theta) \overset{D}{\to} \mathcal{N}(0, (2m+1)^2\sigma_1^2), \tag{3.5}$$

при чему θ зависи од параметра μ, а σ_1^2 је дисперзија одговарајуће прве пројекције.

Ако претходно применимо на конкретне случајне величине из карактеризација 3.1 и 3.2, добијају се тест статистике:

$$W_n^{\mathcal{P}} = \int_0^{\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i < t\} - \frac{1}{n^2} \sum_{i,j=1}^{n} \mathrm{I}\{|X_i - X_j| < t\} \right)^2 dF_n(t), \tag{3.6}$$

$$W_n^{\mathcal{D}} = \int\limits_0^{\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i < t\} - \frac{1}{n^2} \sum_{i,j=1}^{n} \mathrm{I}\{2\min\{X_i, X_j\} < t\} \right)^2 dF_n(t). \tag{3.7}$$

Након квадрирања и интеграције ове тест статистике се могу написати као слабо дегенерисане $V-$статистике реда 5 са језгрима

$$\begin{aligned}\Phi_W^{\mathcal{P}}(x_1,x_2,x_3,x_4,x_5) = \frac{1}{5!}\sum_{\pi\in\Pi(5)} \Big(& \mathrm{I}\{|x_{\pi(1)}-x_{\pi(2)}| < x_{\pi(5)}\}\mathrm{I}\{|x_{\pi(3)}-x_{\pi(4)}| < x_{\pi(5)}\} \\ & + \mathrm{I}\{x_{\pi(1)} < x_{\pi(5)}\}\mathrm{I}\{x_{\pi(2)} < x_{\pi(5)}\} - 2\mathrm{I}\{x_{\pi(1)} < x_{\pi(5)}\}\mathrm{I}\{|x_{\pi(2)}-x_{\pi(3)}| < x_{\pi(5)}\}\Big),\end{aligned}$$

$$\begin{aligned}\Phi_W^{\mathcal{D}}(x_1,x_2,x_3,x_4,x_5) = \frac{1}{5!}\sum_{\pi\in\Pi(5)} \Big(& \mathrm{I}\{2\min(X_{\pi(1)},X_{\pi(2)}) < x_{\pi(5)}\}\mathrm{I}\{2\min(X_{\pi(3)},X_{\pi(4)}) < x_{\pi(5)}\} \\ & + \mathrm{I}\{x_{\pi(1)} < x_{\pi(5)}\}\mathrm{I}\{x_{\pi(2)} < x_{\pi(5)}\} - 2\mathrm{I}\{x_{\pi(1)} < x_{\pi(5)}\}\mathrm{I}\{2\min(X_{\pi(2)},x_{\pi(3)}) < x_{\pi(5)}\}\Big),\end{aligned}$$

где је $\Pi(5)$ скуп свих $5!$ пермутација елемената скупа $\{1,2,3,4,5\}$.

С обзиром да је статистика слободна од параметра расподеле, можемо претпоставити без губитка општости да је $\lambda = 1$. Тада друга пројекција језгра $\Phi_W^{\mathcal{P}}$, односно $\Phi_W^{\mathcal{D}}$, је

$$\begin{aligned}\varphi_2^{\mathcal{P}}(x,y) &= \frac{1}{30} + \frac{3}{10}(e^{-2x-y}+e^{-x-2y}) - \frac{3}{20}(e^{-2y}+e^{-2x}) - \frac{16}{15}e^{-x-y} \\ &+ \frac{1}{15}e^{-\min(x,y)}(2-3\min(x,y)) + \frac{1}{30}e^{-\max(x,y)}(19-6\min(x,y)),\end{aligned} \tag{3.8}$$

$$\begin{aligned}\varphi_2^{\mathcal{D}}(x,y) &= \frac{8}{75} + \frac{2}{25}(e^{-5x}+e^{-5y}) + \frac{2}{15}(e^{-3x}+e^{-3y}) - \frac{1}{20}(e^{-2x}+e^{-2y}) \\ &+ \frac{2}{15}(e^{-\frac{3}{2}y} - e^{-3x})\mathrm{I}\{y \leqslant 2x\} - \frac{1}{10}\Big((1-e^{-x})\mathrm{I}\{y \leqslant x\} + (1-e^{-y})\mathrm{I}\{x \leqslant y\}\Big) \\ &+ \frac{2}{15}(e^{-3\frac{x}{2}} - e^{-3y})\mathrm{I}\{x \leqslant 2y\} - \frac{1}{10}(e^{-x}+e^{-y}) - \frac{1}{5}e^{-4\min(x,y)}.\end{aligned} \tag{3.9}$$

3.3 Тестови L^2–типа засновани на V–емпиријској Лапласовој трансформацији

Нека је $X_1, X_2, ..., X_n$ узорак из непознате апсолутно непрекидне расподеле F. За тестирање нулте хипотезе $H_0 : F(x) = 1 - e^{-\lambda x}, x \geqslant 0, \lambda > 0$, у раду аутора (*Cuparić et al. (2019b)*) предложена је следећа фамилија тест статистика:

$$M_{n,a}(\widehat{\lambda}_n) = \int\limits_0^{+\infty} \left(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t) \right)^2 e^{-at} dt, \tag{3.10}$$

где су

$$\mathcal{L}_n^{(1)}(t) = \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_1(Y_{i_1},\ldots,Y_{i_m})} \tag{3.11}$$

и

$$\mathcal{L}_n^{(2)}(t) = \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_2(Y_{i_1},\ldots,Y_{i_m})} \tag{3.12}$$

$V-$емпиријске Лапласове трансформације случајних величина из карактеризације и a позитивна константа. Уместо полазног узорка разматра се скалирани узорак $Y_i = \widehat{\lambda}_n X_i, i = 1, 2, ..., n,$ где је $\widehat{\lambda}_n = \overline{X}_n^{-1}$ постојана оцена параметра λ, како би статистика била слободна од параметра расподеле при нултој хипотези. Односно видимо да, из хомогености функција ψ_1 и ψ_2, следи

$$\begin{aligned}
M_{n,a}(\widehat{\lambda}_n) &= \int_0^{+\infty} \left(\frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_1(X_{i_1},\ldots,X_{i_m})\frac{1}{\overline{X}_n}} - \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_2(X_{i_1},\ldots,X_{i_m})\frac{1}{\overline{X}_n}} \right)^2 e^{-at} dt \\
&= \int_0^{+\infty} \left(\frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_1\left(\frac{Y_{i_1}}{\lambda},\ldots,\frac{Y_{i_m}}{\lambda}\right)\frac{\lambda}{\overline{Y}_n}} - \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_2\left(\frac{Y_{i_1}}{\lambda},\ldots,\frac{Y_{i_m}}{\lambda}\right)\frac{\lambda}{\overline{Y}_n}} \right)^2 e^{-at} dt \\
&= \int_0^{+\infty} \left(\frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_1(Y_{i_1},\ldots,Y_{i_m})\frac{1}{\overline{Y}_n}} - \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_2(Y_{i_1},\ldots,Y_{i_m})\frac{1}{\overline{Y}_n}} \right)^2 e^{-at} dt,
\end{aligned}$$

при чему случајне величине $Y_1, ..., Y_n$ имају експоненцијалну расподелу са параметром 1, ако полазни узорак $X_1, ..., X_n$ има експоненцијалну расподелу са параметром λ. Ово својство омогућава тестирање сложених хипотеза. Начин избора тежинске функције, односно функције e^{-at}, као и параметра a биће размотрен у поглављу 3.5. Нулта хипотеза се одбацује за велике вредности тест статистике.

Како би се испитала асимптотска својства тест статистике можемо приметити да

$$\begin{aligned}
M_{n,a}(\widehat{\lambda}_n) &= \int_0^{+\infty} \left(\frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_1(X_{i_1},\ldots,X_{i_m})\widehat{\lambda}_n} - \frac{1}{n^m} \sum_{i_1,\ldots,i_m=1}^{n} e^{-t\psi_2(X_{i_1},\ldots,X_{i_m})\widehat{\lambda}_n} \right)^2 e^{-at} dt \\
&= \frac{1}{n^{2m}} \int_0^{+\infty} \sum_{i_1,\ldots,i_{2m}=1}^{n} \left(e^{-t\psi_1(X_{i_1},\ldots,X_{i_m})\widehat{\lambda}_n} - e^{-t\psi_2(X_{i_1},\ldots,X_{i_m})} \right) \\
&\qquad \times \left(e^{-t\psi_1(X_{i_{m+1}},\ldots,X_{i_{2m}})\widehat{\lambda}_n} - e^{-t\psi_2(X_{i_{m+1}},\ldots,X_{i_{2m}})} \right) e^{-at} dt \\
&= \frac{1}{n^{2m}} \sum_{i_1,\ldots,i_{2m}} \int_0^{+\infty} h(X_{i_1}, ..., X_{i_m}, t; \widehat{\lambda}_n) h(X_{i_{m+1}}, ..., X_{i_{2m}}, t; \widehat{\lambda}_n) e^{-at} dt \\
&= \frac{1}{n^{2m}} \sum_{i_1,\ldots,i_{2m}} \Phi_M(X_{i_1}, ..., X_{i_{2m}}, a; \widehat{\lambda}_n),
\end{aligned}$$

где је

$$\Phi_M(X_1,...,X_{2m},a;\widehat{\lambda}_n)=\frac{1}{(2m)!}\sum_{\pi\in\Pi(2m)}\int\limits_0^{+\infty}h(X_{\pi(1)},...,X_{\pi(m)},t;\widehat{\lambda}_n)h(X_{\pi(m+1)},X_{\pi(2m)},t;\widehat{\lambda}_n)e^{-at}dt,$$

а $\Pi(2m)$ скуп свих $(2m)!$ пермутација скупа $\{1,2,...,2m\}$.

Слично као и код претходног теста, испоставља се да је прва пројекција језгра статистике $M_{n,a}(\lambda)$ једнака нули (директно следи применом Фубинијеве теореме), док је друга пројекција различита од нуле и може се одредити за конкретне карактеризације. Према томе, можемо закључити да је $M_{n,a}(\lambda)$ слабо дегенерисана $V-$статистика реда $2m$, а $M_{n,a}(\widehat{\lambda})$ слабо дегенерисана $V-$статистика са оцењеним параметром реда $2m$.

Следећа теорема даје асимптотску расподелу статистике $M_{n,a}(\widehat{\lambda}_n)$ при нултој хипотези. Испоставља се да статистика $M_{n,a}(\widehat{\lambda}_n)$ има исту граничну расподелу као статистика $M_{n,a}(\lambda)$. Теорема представља специјалан случај теореме 2.9. Због начина формирања тест статистике доказ ове теореме се може извести нешто једноставније него код теореме 2.9, што је урађено у радовима аутора (*Cuparić et al. (2019a)* и *Cuparić et al. (2019b)*), па ће тај начин бити приказан.

Теорема 3.1 (Цупарић, Милошевић, Обрадовић 2019) *Нека је $X_1,...,X_n$ узорак из експоненцијалне расподеле са параметром λ. Тада*

$$nM_{n,a}(\widehat{\lambda}_n)\overset{D}{\to}\binom{2m}{2}\sum_{k=1}^{\infty}\upsilon_k Z_k^2,$$

где $\upsilon_k, k=1,2,...,$ представља низ сопствених вредности оператора A дефинисаног на $L^2(\mathbb{R}^+,F)$ са

$$Aq(x)=\int\limits_0^{+\infty}\varphi_2(x,y,a)q(y)dF(y),\ x\in\mathbb{R}^+,\ q\in L^2, \tag{3.13}$$

где је $\varphi_2(x,y,a)$ друга пројекција језгра Φ_M, а $Z_k, k=1,2,...,$ независне случајне величине са стандардном нормалном расподелом.

Доказ: Статистика $M_{n,a}(\widehat{\lambda}_n)$ се може написати на следећи начин

$$M_{n,a}(\widehat{\lambda}_n)=\int\limits_0^{+\infty}\left(\frac{1}{n^2}\sum_{i_1,...,i_m=1}^{n}h(X_{i_1},...,X_{i_m},t,a;\widehat{\lambda}_n)\right)^2e^{-at}dt=\int\limits_0^{+\infty}V_n^2(\widehat{\lambda}_n)e^{-at}dt,$$

где је $V_n(\widehat{\lambda}_n)$ $V-$статистика реда m са оцењеним параметром и језгром $h(x_1,...,x_m,t,a;\widehat{\lambda}_n)$. Како је $h(x_1,...,x_m,t,a;\gamma)$ непрекидно диференцијабилна функција у односу на γ у тачки $\gamma=\lambda$, из теореме о средњој вредности добија се

$$V_n(\widehat{\lambda}_n)=V_n(\lambda)+(\widehat{\lambda}_n-\lambda)\frac{\partial V_n(\gamma)}{\partial\gamma}\Big|_{\gamma=\lambda^*},$$

за неко λ^* између λ и $\widehat{\lambda}_n$.

Из закона великих бројева за $V-$статистике (теорема 2.5), парцијални извод $\frac{\partial V_n(\gamma)}{\partial\gamma}$ конвергира ка

$$\begin{aligned}\frac{\partial}{\partial\gamma}E(h(X_1,...,X_m))|_{\gamma=\lambda} &= E\Big(\frac{\partial h(X_1,...,X_m)}{\partial\gamma}\Big)\Big|_{\gamma=\lambda}\\ &= E\Big(t\psi_2(X_1,...,X_m)e^{-t\psi_2(X_1,...,X_m)\gamma} - t\psi_1(X_1,...,X_m)e^{-t\psi_1(X_1,...,X_m)\gamma}\Big)\Big|_{\gamma=\lambda}\\ &= 0.\end{aligned}$$

Последњи корак следи из карактеризације. С обзиром да $\sqrt{n}(\widehat{\lambda}_n-\lambda)$ конвергира ка нормалној расподели, то је и ограничено у вероватноћи, па следи да су статистике $\sqrt{n}V_n(\widehat{\lambda}_n)$ и $\sqrt{n}V_n(1)$ асимптотски једнако расподељене. Дакле, оцена параметра нема утицаја на асимптотску расподелу статистике, па следи да и статистике $M_{n,a}(\lambda)$ и $M_{n,a}(\widehat{\lambda})$ имају исту асимптотску расподелу, која следи из теореме 2.6. ■

Уместо посматрања асимптотског понашања статистике при нултој хипотези, можемо посматрати њено понашање при алтернативној хипотези када је алтернатива фиксирана (не зависи од n). Претпоставимо да алтернативна расподела има очекивање μ. У том случају прва пројекција језгра $\Phi_M(X_1,...,X_{2m};\lambda)$ није једнака 0, па можемо закључити да статистика која се добија у том случају није дегенерисана $V-$статистика и због тога се гранична расподела у овом случају разликује од оне добијене у теореми 3.1.

Теорема 3.2 (Цупарић, Милошевић, Обрадовић 2019) *Нека је $X_1,...,X_n$ узорак из алтернативне расподеле са функцијом расподеле G. Тада*

$$\sqrt{n}(M_{n,a}(\mu)-\Delta(\mu)) \xrightarrow{D} \mathcal{N}(0,\Sigma),$$

где је $\Delta(\mu)=E(M_{n,a}(\mu))$ и

$$\begin{aligned}\Sigma = 4m^2D(\varphi_1(X_1,a)) &+ \Big(\int\limits_{(R^+)^{2m}}\Phi'_M(\boldsymbol{x},a;\mu)d\boldsymbol{G}(\boldsymbol{x})\Big)^2 D(X_1) + 4m\Big(\int\limits_{(R^+)^{2m}}x_1\Phi_M(\boldsymbol{x},a;\mu)d\boldsymbol{G}(\boldsymbol{x})\\ &- \int\limits_{R^+}x_1dG(x_1)\int\limits_{(R^+)^{2m}}\Phi_M(\boldsymbol{x},a;\mu)d\boldsymbol{G}(\boldsymbol{x})\Big)\Big(\int\limits_{(R^+)^{2m}}\Phi'_M(\boldsymbol{x},a;\mu)d\boldsymbol{G}(\boldsymbol{x})\Big),\end{aligned} \tag{3.14}$$

где је $\boldsymbol{x}=(x_1,...,x_{2m})$, $\boldsymbol{G}(\boldsymbol{x})=\prod\limits_{i=1}^{2m}G(x_i)$ и $\Phi'_M(\boldsymbol{x},a;\mu)=\frac{\partial\Phi_M(\boldsymbol{x},a;\gamma)}{\partial\gamma}\Big|_{\gamma=\mu}$.

Доказ: С обзиром да је језгро $\Phi_M(x_1,...,x_{2m},a;\mu)$ недегенерисано, из теореме 2.6 следи

$$\sqrt{n}(M_{n,a}(\mu)-\Delta) \xrightarrow{D} \mathcal{N}(0,(2m)^2D(\varphi_1(X_1,a))).$$

Како је $\Phi_M(x_1,..,x_{2m},a;\gamma)$ непрекидно диференцијабилна функција у односу на γ у тачки $\gamma=\mu$, из теореме о средњој вредности добија се

$$\sqrt{n}(M_{n,a}(\widehat{\mu})-\Delta(\mu))=\sqrt{n}(M_{n,a}(\mu)-\Delta(\mu))+\sqrt{n}(\widehat{\mu}-\mu)\frac{\partial M_{n,a}(\gamma)}{\partial\gamma}|_{\gamma=\mu^*},$$

за неко μ^* између μ и $\widehat{\mu}$. С обзиром да су $M_{n,a}(\mu)-\Delta(\mu)$ и $\widehat{\mu}-\mu$ две недегенерисане $V-$статистике, следи да је гранична расподела од $\sqrt{n}(M_{n,a}(\mu)-\Delta(\mu),\widehat{\mu}-\mu)$ дводимензионална нормална. Користећи закон великих бројева и теорему Слуцког, следи да $\sqrt{n}(M_{n,a}(\widehat{\mu})-\Delta(\mu))$ конвергира у расподели ка случајној величини са нормалном расподелом са очекивање 0 и дисперзијом једнаком

$$\begin{aligned}&(2m)^2D(\varphi_1(X_1,a))+\lim_{n\to\infty}E\bigg(\frac{\partial M_{n,a}(\gamma)}{\partial\gamma}\Big|_{\gamma=\mu}\bigg)^2D(\sqrt{n}\widehat{\mu})\\&+2\lim_{n\to\infty}\operatorname{cov}(\sqrt{n}M_{n,a}(\mu),\sqrt{n}\mu)E\bigg(\frac{\partial M_{n,a}(\gamma)}{\partial\gamma}|_{\gamma=\mu}\bigg).\end{aligned}$$

Одређивањем граничне вредности добија се (3.14). ■

Ако у тест статистику $M_{n,a}(\widehat{\lambda})$ укључимо карактеризације 3.1 и 3.2, добијају се две конкретне класе тестова

$$M_{n,a}^{\mathcal{P}}(\widehat{\lambda}_n)=\int\limits_0^{+\infty}\bigg(\frac{1}{n}\sum_{i=1}^n e^{-t\widehat{\lambda}_nX_i}-\frac{1}{n^2}\sum_{i,j=1}^n e^{-t\widehat{\lambda}_n|X_i-X_j|}\bigg)^2e^{-at}dt, \tag{3.15}$$

$$M_{n,a}^{\mathcal{D}}(\widehat{\lambda}_n)=\int\limits_0^{+\infty}\bigg(\frac{1}{n}\sum_{i=1}^n e^{-t\widehat{\lambda}_nX_i}-\frac{1}{n^2}\sum_{i,j=1}^n e^{-t2\widehat{\lambda}_n\min\{X_i,X_j\}}\bigg)^2e^{-at}dt. \tag{3.16}$$

Након квадрирања и интеграције добија се да ове тест статистике се могу представити као слабо дегенерисане $V-$статистике реда 4 са језгрима, редом,

$$\begin{aligned}\Phi_M^{\mathcal{P}}(x_1,x_2,x_3,x_4,a;\lambda)=&\frac{1}{4!}\sum_{\pi\in\Pi(4)}\Big(\frac{1}{a+\lambda x_{\pi(1)}+\lambda x_{\pi(3)}}-\frac{1}{a+\lambda x_{\pi(1)}+\lambda|x_{\pi(3)}-x_{\pi(4)}|}\\&-\frac{1}{\lambda x_{\pi(3)}+\lambda|x_{\pi(1)}-x_{\pi(2)}|}+\frac{1}{a+\lambda|x_{\pi(1)}-x_{\pi(2)}|+\lambda|x_{\pi(3)}-x_{\pi(4)}|}\Big),\end{aligned} \tag{3.17}$$

$$\begin{aligned}\Phi_M^{\mathcal{D}}(x_1,x_2,x_3,x_4,a;\lambda)=&\frac{1}{4!}\sum_{\pi\in\Pi(4)}\Big(\frac{1}{a+\lambda x_{\pi(1)}+\lambda x_{\pi(3)}}-\frac{1}{a+\lambda x_{\pi(1)}+2\lambda\min\{x_{\pi(3)},x_{\pi(4)}\}}\\&-\frac{1}{\lambda x_{\pi(3)}+2\lambda\min\{x_{\pi(1)},x_{\pi(2)}\}}+\frac{1}{a+2\lambda\min\{x_{\pi(1)},x_{\pi(2)}\}+2\lambda\min\{x_{\pi(3)},x_{\pi(4)}\}}\Big),\end{aligned} \tag{3.18}$$

где је $\Pi(4)$ скуп свих 4! пермутација скупа $\{1,2,3,4\}$. Прве пројекције језгара $\Phi_M^{\mathcal{I}}(x_1,x_2,x_3,x_4,a;\lambda),\mathcal{I}\in\{\mathcal{P},\mathcal{D}\}$, су једнаке нули. С обзиром да расподела тест статистика $M_{n,a}^{\mathcal{I}}(\lambda),\mathcal{I}\in\{\mathcal{P},\mathcal{D}\}$, не зависи од параметра λ, можемо претпоставити без губитка општости да је $\lambda = 1$. Друге пројекције језгара $\Phi_M(x_1, x_2, x_3, x_4, a; 1)$ су, редом,

$$\begin{aligned}\varphi_2^{\mathcal{P}}(x,y,a) &= -\frac{1}{2}+\frac{1}{3}(e^{-x}+e^{-y})+\frac{1}{6}e^{a-x-y}Ei(-a)\Big(a(e^x-2)(e^y-2)-e^x-e^y+4\Big)\\ &+\frac{1}{6}e^{-a-x-y}\Big(Ei(a)(4a+e^x+e^y-4)-(Ei(a+x)(4(a+x-1)+e^y)\\ &+Ei(a+y)(4(a+y-1)+e^x)-4(a+x+y-1)Ei(a+x+y))\Big)+\frac{1}{6(a+x+y)},\end{aligned} \tag{3.19}$$

$$\begin{aligned}\varphi_2^{\mathcal{D}}(x,y,a) &= \frac{1}{6}\Bigg(3+\frac{1}{a+x+y}-\frac{2e^{-x}}{a+2x+y}-\frac{2e^{-y}}{a+x+2y}-(4-a)e^aEi(-a)\\ &+e^{\frac{a+y}{2}}\Big(Ei\big(-\frac{a+y}{2}\big)-Ei\big(-\frac{a+2x+y}{2}\big)\Big)+e^{a+x}\Big(4Ei(-a-2x)-Ei(-a-x)\Big)\\ &+e^{\frac{a+x}{2}}\Big(Ei\big(-\frac{a+x}{2}\big)-Ei\big(-\frac{a+x+2y}{2}\big)\Big)+e^{a+y}\Big(4Ei(-a-2y)-Ei(-a-y)\Big)\\ &+\frac{e^{-x-y}}{a+2(x+y)}(2a+4(1+x+y))-2(e^{-x}+e^{-y})+e^{\frac{a}{2}}\Big(-(4+a+2x)Ei(-\frac{a}{2}-x)\\ &+(a+4)Ei(-\frac{a}{2})+(a+2(2+x+y))Ei(-\frac{a}{2}-x-y)-(4+a+2y)Ei(-\frac{a}{2}-y)\Big)\Bigg),\end{aligned} \tag{3.20}$$

где је $Ei(z) = -\int\limits_{-z}^{+\infty}\frac{e^{-t}}{t}dt$ експоненцијални интеграл.

3.4 Тестови L^∞–типа засновани на V–емпиријској Лапласовој трансформацији

Нека је $X_1, X_2, ..., X_n$ узорак из непознате апсолутно непрекидне расподеле F. За тестирање нулте хипотезе $H_0 : F(x) = 1 - e^{-\lambda x}, x \geqslant 0, \lambda > 0$, предложена је, у радовима аутора (*Cuparić et al. (2019a)* и *Cuparic et al. (2019)*), следећа фамилија тест статистика:

$$L_{n,a}(\widehat{\lambda}_n) = \sup_{t>0}|(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t))e^{-at}|, \tag{3.21}$$

где су $\mathcal{L}_n^{(1)}$ и $\mathcal{L}_n^{(2)}$ V–емпиријске Лапласове трансформације случајних величина из карактеризације дефинисане у (3.11) и (3.12) и a позитивна константа. Уместо полазног узорка разматра се скалирани узорак $Y_i = \widehat{\lambda}_n X_i,\ i = 1, 2, ..., n$, где је $\widehat{\lambda}_n = \overline{X}_n^{-1}$ постојана оцена параметра λ, како би статистика била слободна од параметра при нултој хипотези. Ово својство се показује на исти начин као код теста из поглавља 3.3. Тежинска функција e^{-at} код ове тест статистике има улогу да скалира вредност тест статистике. Може се узети и да је $a = 0$, односно да нема тежинске функције. Код проучавања карактеристика овог теста биће размотрена оба случаја.

Тест статистика $L_{n,a}(\widehat{\lambda}_n)$ се може представити на следећи начин

$$L_{n,a}(\widehat{\lambda}_n) = \sup_{t>0}|V_n(t, a; \widehat{\lambda}_n)| = \sup_{t>0}\Big|\frac{1}{n^m}\sum_{i_1,..,i_m=1}^{n}\Phi_L(X_{i_1}, ..., X_{i_m}, t, a; \widehat{\lambda}_n)\Big|, \tag{3.22}$$

где је $\Phi_L(X_1,...,X_m,t,a;\widehat{\lambda}_n)) = \left(e^{-t\widehat{\lambda}_n\psi_1(X_1,...,X_m)} - e^{-t\widehat{\lambda}_n\psi_2(X_1,...,X_m)}\right)e^{-at}$ симетрична функција по својим аргументима (или, ако није, посматра се одговарајућа симетризација). Дакле, за фиксирано t, $V_n(t,a;\widehat{\lambda}_n)$ је $V-$статистика реда m са оцењеним параметром, односно тест статистика је супремум апсолутне вредности $V-$емпиријског процеса.

Следећа теорема даје слабу конвергенцију процеса $V_n(t,a;\widehat{\lambda}_n)$, који је основа тест статистике $L_{n,a}(\widehat{\lambda}_n)$, а њена последица даје конвергенцију у расподели саме тест статистике при нултој хипотези. Овај резултат је добијен у раду аутора (*Cuparić et al. (2019a)*).

Теорема 3.3 (Цупарић, Милошевић, Обрадовић 2019) *Нека је $X_1,...,X_n$ узорак из експоненцијалне расподеле са параметром λ. Тада процес $\{V_n(t,a;\widehat{\lambda}_n), t\geqslant 0\}$ конвергира слабо у $C[0,\infty)$ ка центрираном Гаусовом процесу $\{\eta(t), t\geqslant 0\}$ са коваријационом функцијом*

$$K(s,t) = m^2E(\Phi_L(X_1,...,X_m,t,a;1)\Phi_L(X_1,...,X_m,s,a;1)). \tag{3.23}$$

Доказ: Слично као у доказу теореме 3.1 може се показати да, за фиксирано t, статистика $\sqrt{n}V_n(t,a;\widehat{\lambda}_n)$, дефинисана у (3.22), и статитика $\sqrt{n}V_n(t,a;\lambda)$ имају исту асимптотску расподелу, односно да оцена параметра λ нема утицаја на граничну расподелу статистике. Такође, с обзиром да је статистика слободна од параметра расподеле, без губитка општости можемо претпоставити да је $\lambda=1$.

Како је функција Φ_L непрекидна функција дефинисана на $[0,\infty)$, према теореми 2.12, да би се показала конвергенција процеса, довољно је доказати конвергенцију коначнодимензионалних расподела и да је скуп коначнодимензионалних расподела густ. С обзиром да је за фиксирано t, $V_n(t,a;1)$ недегенерисана V-статистика следи из теореме 2.6 да $\sqrt{n}V_n(t,a;1)$ има граничну нормалну расподелу. Исто важи и за све коначнодимензионалне расподеле. Још треба доказати да је одговарајући скуп густ. За то ћемо користити последицу 2.1. Пошто је $V_n(0,a;1)=0$, одакле следи да је $E|V_n(0,a;1)|^\nu=0$ чиме је испуњен први услов наведене последице. Да бисмо показали други услов, показаћемо да постоје α,β и функција Υ такви да

$$\sup_{n\geqslant 1} E|\sqrt{n}V_n(s,a;1)-\sqrt{n}V_n(t,a;1)|^\alpha \leqslant \Upsilon(T)|t-s|^{1+\beta},\ \forall T>0, 0\leqslant s,t\leqslant T.$$

Претпоставимо да је $s=t+\zeta,\ \zeta\in[0,T]$. Приметимо да је

$$\begin{aligned}
&E(|\sqrt{n}V_n(t+\zeta,a;1)-\sqrt{n}V_n(t,a;1)|^2)\\
&= nE\Bigg(\frac{1}{n^m}\sum_{i_1,...,i_m}\Phi_L(X_{i_1},...,X_{i_m},t+\zeta,a;1) - \frac{1}{n^m}\sum_{i_1,...,i_m}\Phi_L(X_{i_1},...,X_{i_m},t,a;1)\Bigg)^2\\
&= \frac{n}{n^{2m}}\sum_{i_1,...,i_{2m}}E\Bigg(\Phi_L(X_{i_1},...,X_{i_m},t+\zeta,a;1)-\Phi_L(X_{i_1},...,X_{i_m},t,a;1)\Bigg)\\
&\qquad\times\Bigg(\Phi_L(X_{i_{m+1}},...,X_{i_{2m}},t+\zeta,a;1)-\Phi_L(X_{i_{m+1}},...,X_{i_{2m}},t,a;1)\Bigg).
\end{aligned}$$

У зависности од тога колико индекса из скупа $\{i_1,...,i_{2m}\}$ се међусобно поклапа, разликујемо случајеве:

- ако су сви индекси међусобно различити из независности случајних величина и карактеризације следи да су ти чланови горње суме једнаки 0 и њих у суми има $\binom{n}{2m}(2m)!$;

- ако се тачно два индекса поклапају разликујемо два случаја

  - поклапају се два индекса у скупу $\{i_1, ..., i_m\}$ или се поклапају два индекса у скупу $\{i_{m+1}, ..., i_{2m}\}$, онда је општи члан суме једнак 0 због независности случајних величина и карактеризације, а таквих елемената у горњој суми има $n\binom{n-1}{2m-2}\binom{m}{2}(2m-2)!2!$;
  - један индекс из скупа $\{i_1, ..., i_m\}$ је једнак једном индексу из скупа $\{i_{m+1}, ..., i_{2m}\}$, онда је општи члан суме једнак

$$\begin{aligned}
&E((\Phi_L(X_1, X_2, ..., X_m, t+\zeta, a; 1) - \Phi_L(X_1, X_2, ..., X_m, t, a; 1))\\
&\times (\Phi_L(X_1, X_{m+1}, ..., X_{2m-1}, t+\zeta, a; 1) - \Phi_L(X_1, X_{m+1}, ..., X_{2m-1}, t, a; 1)))\\
&= E(\varphi_1(X_1, t+\zeta, a; 1) - \varphi_1(X_1, t, a; 1))^2,
\end{aligned}$$

  где је $\varphi_1(x, t, a; 1) = E(\Phi_L(X_1, X_2, ..., X_m, t, a; 1)|X_1 = x)$ прва пројекција језгра Φ_L, а таквих елемената има $n\binom{n-1}{2m-2}m^2(2m-2)!$ у горњој суми;

- ако се тачно c индекса поклапа ($c > 2$), применом Коши-Шварцове неједнакости и неједнакости $(a-b)^2 \leqslant 2a^2 + 2b^2$ следи да је

$$\begin{aligned}
&E(\Phi_L(X_{i_1}, ..., X_{i_m}, t+\zeta, a; 1) - \Phi_L(X_{i_1}, ..., X_{i_m}, t; 1))(\Phi_L(X_{i_{m+1}}, ..., X_{i_{2m}}, t+\zeta, a; 1)\\
&- \Phi_L(X_{i_{m+1}}, ..., X_{i_{2m}}, t, a; 1))\\
&\leqslant (E(\Phi_L(X_{i_1}, ..., X_{i_m}, t+\zeta, a; 1) - \Phi_L(X_{i_1}, ..., X_{i_m}, t, a; 1))^2\\
&\times E(\Phi_L(X_{i_{m+1}}, ..., X_{i_{2m}}, t+\zeta, a; 1) - \Phi_L(X_{i_{m+1}}, ..., X_{i_{2m}}, t, a; 1))^2)^{\frac{1}{2}}\\
&= E(\Phi_L(X_{i_1}, ..., X_{i_m}, t+\zeta, a; 1) - \Phi_L(X_{i_1}, ..., X_{i_m}, t, a; 1))^2\\
&\leqslant 2E(\Phi_L(X_{i_1}, ..., X_{i_m}, t+\zeta, a; 1))^2 + 2E(\Phi_L(X_{i_1}, ..., X_{i_m}, t, a; 1))^2 < \infty,
\end{aligned}$$

  а таквих чланова има $n^{2m} - \binom{n}{2m}(2m)! - n\binom{n-1}{2m-2}(m(m-1)+m^2)(2m-2)!$.

Комбиновањем свих претходних случајева добија се да је

$$\begin{aligned}
&E(|\sqrt{n}V_n(t+\zeta, a; 1) - \sqrt{n}V_n(t, a; 1)|^2)\\
&= \frac{1}{n^{2m-1}}n\binom{n-1}{2m-2}(2m)^2(2m-2)!E(\varphi_1(X_1, t+\zeta, a; 1) - \varphi_1(X_1, t, a; 1))^2 + o_p(1)\\
&\propto E(\varphi_1(X_1, t+\zeta, a; 1) - \varphi_1(X_1, t, a; 1))^2 + o_p(1),\ n \to \infty.
\end{aligned}$$

Применом теореме о средњој вредности добија се да је

$$E(\varphi_1(X_1, t+\zeta, a; 1) - \varphi_1(X_1, t, a; 1))^2 \leqslant |\zeta|^2 E\left(\frac{d}{dt}\varphi_1(X_1, t, a; 1)|_{t=\xi}\right)^2,$$

где је $\xi \in [t, t+\zeta]$. С обзиром да непрекидна функција на компактном скупу $[0, T]$ достиже свој максимум, следи да функција

$$\Upsilon(T) = \max_{\xi\in[0,T]} E\left(\frac{d}{dt}\varphi_1(X_1, \xi, a; 1)\right)^2$$

постоји и коначна је. Дакле, можемо закључити да

$$E(|\sqrt{n}V_n(s, a; 1) - \sqrt{n}V_n(t, a; 1)|^2) \leqslant |t-s|^2\Upsilon(T),$$

одакле следи други услов последице 2.1 (за $\alpha = 2$ и $\beta = 1$). Тиме је показана слаба конвергенција процеса $\sqrt{n}V_n(t, a; 1)$ у $C[0, \infty)$ ка Гаусовом процесу $\eta(t)$ са коваријационом функцијом (3.23). ■

Последица 3.1 (Цупарић, Милошевић, Обрадовић 2019) *Нека је $X_1, ..., X_n$ узорак из експоненцијалне расподеле са параметром λ. Тада*

$$\sqrt{n}L_{n,a}(\widehat{\lambda}_n) \xrightarrow{D} \sup_{t>0} |\eta(t)|,$$

где је $\eta(t)$ центрирани Гаусов процес са коваријационом функцијом дефинисаном у (3.23).

Доказ: Основу статистике $L_{n,a}(\widehat{\lambda}_n)$ чини процес $\{V_n(t, a; \widehat{\lambda}_n), t \geqslant 0\}$, што се може видети из (3.22). Увођењем смене $s = e^{-t} \in [0, 1]$, језгро процеса $\{V_n(s, a; \widehat{\lambda}_n), s \in [0, 1]\}$ има облик

$$\widetilde{\Phi}_L(x_1, ..., x_m, s, a; \widehat{\lambda}_n) = (s^{\psi_1(x_1,...,x_m)} - s^{\psi_2(x_1,...,x_m)})s^a.$$

Дакле, наведени процес представља процес дефинисан на простору $C[0, 1]$ са супремум нормом. Користећи сличне аргументе као у доказу претходне теореме, може се показати да процес $\{V_n(s, a; \widehat{\lambda}_n), s \in [0, 1]\}$ конвергира слабо у $C[0, 1]$ ка центрираном Гаусовом процесу $\{\eta(s), s \in [0, 1]\}$ са коваријационом функцијом која се добија из (3.23) увођењем исте смене. С обзиром да је супремум непрекидна функција на простору $C[0, 1]$, из теореме о непрекидном пресликавању и коришћењем трансформације $t = -\ln s$, следи тврђење последице. ■

Иако је позната гранична случајна величина на основу претходне теореме, експлицитни облик расподеле те случајне величине није позната. Због тога се критичне вредности морају одређивати симулацијама.

Ако у тест статистику $L_{n,a}(\widehat{\lambda})$ укључимо карактеризације 3.1 и 3.2, добијају се две конкретне класе тестова са тест статистикама

$$L_{n,a}^{\mathcal{P}}(\widehat{\lambda}_n) = \sup_{t>0} \left| \left(\frac{1}{n} \sum_{i=1}^{n} e^{-tY_i} - \frac{1}{n^2} \sum_{i,j=1}^{n} e^{-t|Y_i - Y_j|} \right) e^{-at} \right|, \tag{3.24}$$

$$L_{n,a}^{\mathcal{D}}(\widehat{\lambda}_n) = \sup_{t>0} \left| \left(\frac{1}{n} \sum_{i=1}^{n} e^{-tY_i} - \frac{1}{n^2} \sum_{i,j=1}^{n} e^{-t2\min\{Y_i, Y_j\}} \right)^2 e^{-at} \right|. \tag{3.25}$$

Коваријационе функције граничних процеса су, редом,

$$K^{\mathcal{P}}(s,t) = \frac{e^{-a(s+t)} st(4 + 4s + 4t + 3st)}{3(1+s)(2+s)(1+t)(2+t)(1+s+t)}, \tag{3.26}$$

$$K^{\mathcal{D}}(s,t) = \frac{e^{-a(s+t)} st(4 + 8s + 4s^2 + 8t + 15st + 6s^2t + 4t^2 + 6st^2)}{(1+s)(1+t)(1+s+t)(2+2s+t)(2+s+2t)(3+2s+2t)}. \tag{3.27}$$

3.5 Одређивање оптималне вредности параметра подешавања

Често се у тест статистикама појављује такозвани параметар подешавања као параметар у тежинској функцији. На пример, тестови представљени у поглављу 3.3 се у општем случају могу представити на следећи начин

$$M_{n,a}(\widehat{\lambda}_n) = \int_0^{+\infty} \left(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t) \right)^2 \omega(t) dt,$$

где је $\omega(t)$ нека тежинска функција. Улога тежинске функције у овом тесту је формирање интеграбилне подинтегралне функције. За тежинску функцију је изабрано $\omega(t) = e^{-at}$, где је a параметар подешавања, из два разлога. Први разлог је једноставност у рачуну, што се огледа у томе да језгра $\Phi_M^{\mathcal{P}}$, дато у (3.17), и $\Phi_M^{\mathcal{D}}$, дато у (3.18), добијена након интеграције, а и сам процес интеграције, нису компликовани. Други разлог потиче од Тауберове теореме везане за Лапласове трансформације. Испоставља се да при веома општим условима понашање Лапласове трансформације у близини нуле јединствено одређује асимптотско понашање расподеле дефинисане на $[0, \infty)$ када $x \to \infty$, односно репа дате расподеле, и обрнуто (за више детаља видети *Feller (1968)*). Због тога, ако се изабере мала вредност за параметар подешавања a, тежинска функција опада полако, што би као резултат требало да да високе моћи против алтернатива које имају бесконачну густину у нули. Са друге стране, велике вредности параметра подешавања a дају да тежинска функција брзо одлази у нулу, што би као резултат требало да да високе моћи против алтернатива које се значајно разликују од експоненцијалне расподеле у репу расподеле. Дакле, кроз избор параметра a може се повећати моћ теста у неком одређеном смеру кроз скуп алтернативних расподела.

Тест представљен у поглављу 3.4 се такође у општем случају може приказати коришћењем неке тежинске функције $\omega(t)$, односно

$$L_{n,a}(\widehat{\lambda}_n) = \sup_{t>0} |(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t))\omega(t)|.$$

У овом случају, као што је раније наведено, улога тежинске функције је да скалира вредности функције чији се супремум тражи.

У практичним применама, да би се применио тест потребно је узети конкретну вредност за a. С обзиром да би тест требало да буде што бољи против великог броја алтернатива, поставља се питање како изабрати вредност параметра подешавања a. Најчешће коришћен приступ у литератури (на пример, *Epps and Pulley (1983), Baringhaus and Henze (1991), Henze and Meintanis (2002b)*) је да се одреде емпиријске моћи теста за различите алтернативе и неки низ вредности параметра a, а да се као препорука за коришћење у пракси узме оно a које је дало највећу моћ за највећи број алтернатива. Међутим, на тај начин, користећи фиксирану вредност параметра a, може се добити тест који за неке алтернативе има врло ниску моћ. Због тога постоји потреба за избором параметра a на основу самих података. Први помаци у том правцу су били у виду вишеструког тестирања, односно комбиновање више процедура повезаних са различитим вредностима параметра у једну процедуру што резултује добрим перформансама теста против великог броја алтернатива (*Klar (2001), Fromont and Laurent (2006), Tenreiro (2011)*). У том случају, нулта хипотеза се одбацује ако једна од статистика покаже тенденцију ка одбацивању нулте хипотезе са нивоом значајности који представља калибрирану вредност тако да вишеструко тестирање има ниво значајности највише α.

У последње време су развијена два метода којима се на основу података одређује оптимална вредност параметра подешавања. Да би се ови методи применили фамилија тест статистика мора бити инваријантна у односу на параметре расподеле при нултој хипотези, односно слободна од тог параметра. Ту особину задовољавају обе класе тестова приказане у претходним поглављима које у свом запису имају параметар подешавања.

Први метод, развијен од стране Алисона и Сантане (*Allison and Santana (2015)*), у свом алгоритму претпоставља случајну природу избора параметра. Док други метод за избор оптималне вредности параметра подешавања, развијен од стране Тенреира (*Tenreiro (2019)*), превазилази ту случајну природу избора параметра. С обзиром на то, биће приказан само други метод одређивања оптималне вредности параметра подешавања и примењен на конкретне тест статистике.

Поступак добијања оптималне вредности параметра a се може описати следећим алгоритмом:

1. задати низ позитивних вредности параметра подешавања, $a \in \{a_1, ..., a_k\}$;
2. узети бутстреп узорак $(X_1^*, ..., X_n^*)$ из емпиријске функције расподеле од $(X_1, ..., X_n)$;
3. одредити вредност тест статистике за бутстреп узорак, $T_{n,a_i}(X_1^*, ..., X_n^*)$, $i = 1, ..., k$;

4. поновити кораке 1. и 2. велики број пута (B пута), при чему се добија низ вредности тест статистике $T^*_{1,a_i}, ..., T^*_{B,a_i}$, $i = 1, ..., k$;

5. одредити бутстреповану моћ $\widehat{P}_{a_i} = \frac{1}{B}\sum\limits_{j=1}^{B} I\{T^*_{j,a_i} \geqslant \widehat{q}_{n,a_i}(u)\}$, $i = 1, ..., k$, где је u калибрирано тако да тест има ниво значајности највише α;

6. оптимална вредност параметра a је она за коју важи $\widehat{a} = \underset{a\in\{a_1,...,a_k\}}{\operatorname{argmax}} \widehat{P}_a$.

Избор низа вредности у кораку 1. зависи од саме тест статистике и улоге параметра a у њој. Након добијања оптималне вредности параметра a на основу података, даље се Монте Карло симулацијама може одредити моћ или $p-$вредност теста на стандардан начин.

Избор калибрираног нивоа u је од великог значаја, јер не само да се њиме обезбеђује да је мера теста мања од номиналног нивоа α већ је тај ниво близак α што је више могуће. Приликом калибрације бира се највеће u из скупа $G_p = \{u_j, j \in I_p\}$ на интервалу $(0, 1)$, при чему је $u_1 = p$ и $u_{j+1} = u_j + p$, за неко $0 < p \leqslant \frac{\alpha}{k}$, где је k кардиналност скупа из корака 1 претходног алгоритма, такво да је $P_{\widehat{a}}(u) = P_{F_0}\{T_{n,\widehat{a}_u} \geqslant q_{n,\widehat{a}_u}(u)\} \leqslant \alpha$, односно

$$u^{\widehat{a}}_{n,\alpha,p} = \max_{u\in G_p}\{P_{\widehat{a}}(u) \leqslant \alpha\}. \tag{3.28}$$

Ова калибрација обезбеђује да наведени тест има меру што ближу α када p тежи 0. Могуће је урадити и други ниво калибрације, који осигурава да добијена вредност заиста јесте најближа могућа номиналном нивоу значајности. Међутим, у практичним случајевима испоставља се да тестови са једним нивом калибрације и тестови са два нивоа калибрације дају сличне резултате, а с обзиром да су временски значајно захтевнији тестови са две калибрације, предлаже се коришћење тестова са једном калибрацијом (*Tenreiro (2019)*).

Резултати примене овог метода за избор оптималне вредности параметра подешавања a у тестовима представљеним у поглављима 3.3 и 3.4 су приказани у таблема 3.1 и 3.2.

У емпиријским проучавањима код већине тестова се испоставља да оптимална вредност параметра a је између 0 и 5. Међутим, код тестова L^2-типа може бити значајно да се испита шта се дешава ако параметар подешавања тежи бесконачности, односно ако се посматра гранична вредност када $a \to \infty$. У ту сврху ћемо користити следећи став који представља последицу Абелове теореме за Лапласове трансформације (за више детаља видети *Widder (1946)*).

Став 3.1 *Нека је $\rho : [0, \infty) \to \mathbb{R}$ мерљива функција интеграбилна на компактним интервалима. Такође, претпоставимо да је $\int_0^\infty \rho(t)e^{-at}dt$ коначно за свако $a > 0$. Ако за неко $\gamma \geqslant 0$ и неку релну константу B*

$$\lim_{s\to 0}\Gamma(\gamma + 1)s^{-\gamma}\int\limits_0^s \rho(t)dt = B, \tag{3.29}$$

онда

$$\lim_{a\to\infty} a^{\gamma}\int\limits_0^{\infty}\rho(t)e^{-at}dt = B.$$

Сада (3.29) *важи ако*

$$\lim_{s\to 0}\Gamma(\gamma)s^{-(\gamma-1)}\rho(s) = B.$$

У следећој теореми је дато понашање статистике $M_{n,a}(\widehat{\lambda}_n)$ из поглавља 3.3 када $a\to\infty$, што је показано у раду аутора (*Cuparić et al. (2019b)*).

Теорема 3.4 (Цупарић, Милошевић, Обрадовић 2019) *За фиксирано n, следи*

$$\lim_{a\to\infty} a^3 M_{n,a}(\widehat{\lambda}_n) = 2\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_1(Y_{i_1},...,Y_{i_m}) - \psi_2(Y_{i_1},...,Y_{i_m})\Big)\bigg)^2,$$

īge je $Y_i = \widehat{\lambda}_n X_i,\ i=1,2,...,n,$ и $\widehat{\lambda}_n = \overline{X}^{-1}$.

Доказ: Означимо са $\rho(t) = \Big(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t)\Big)^2$. Тада се тест статистика може написати као $M_{n,a}(\widehat{\lambda}_n) = \int\limits_0^{\infty}\rho(t)e^{-at}dt$. Применом Маклореновог развоја на $\rho(t)$ добија се

$$\begin{aligned}
\rho(t) &= \bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(e^{-t\psi_1(Y_{i_1},...,Y_{i_m})} - e^{-t\psi_2(Y_{i_1},...,Y_{i_m})}\Big)\bigg)^2\\
&=\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(1-t\psi_1(Y_{i_1},...,Y_{i_m})+\frac{t^2\psi_1^2(Y_{i_1},...,Y_{i_m})}{2}-1+t\psi_2(Y_{i_1},...,Y_{i_m})-\frac{t^2\psi_2^2(Y_{i_1},...,Y_{i_m})}{2}+o_p(t^2)\Big)\bigg)^2\\
&=t^2\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_1(Y_{i_1},...,Y_{i_m})-\psi_2(Y_{i_1},...,Y_{i_m})\Big)\bigg)^2+\frac{t^3}{2}\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_1(Y_{i_1},...,Y_{i_m})-\psi_2(Y_{i_1},...,Y_{i_m})\Big)\bigg)\\
&\times\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_2^2(Y_{i_1},...,Y_{i_m})-\psi_1^2(Y_{i_1},...,Y_{i_m})\Big)\bigg)+\frac{t^4}{4}\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_2^2(Y_{i_1},...,Y_{i_m})-\psi_1^2(Y_{i_1},...,Y_{i_m})\Big)\bigg)+o_p(t^4),
\end{aligned}$$

кад $t\to 0$. Користећи став 3.1 и чињеницу да је

$$\lim_{s\to 0}\Gamma(4)s^3\int\limits_0^{s}\rho(t)dt = 2\bigg(\frac{1}{n^m}\sum_{i_1,\dots,i_m}\Big(\psi_1(Y_{i_1},...,Y_{i_m}) - \psi_2(Y_{i_1},...,Y_{i_m})\Big)\bigg)^2,$$

следи тврђење теореме. ■

У конкретном случају, када се користе Пури-Рубинова и Десуова карактеризација, применом претходне теореме добијају се следеће статистике

$$\lim_{a\to\infty} a^3 M_{n,a}^{\mathcal{P}}(\hat{\lambda}_n) = 2\bigg(\frac{1}{n^2}\sum_{i,j=1}^{n}|Y_i - Y_j| - \overline{Y}_n\bigg)^2,$$

$$\lim_{a\to\infty} a^3 M_{n,a}^{\mathcal{D}}(\hat{\lambda}_n) = 2\bigg(\frac{2}{n^2}\sum_{i,j=1}^{n}\min\{Y_i,Y_j\} - \overline{Y}_n\bigg)^2.$$

Емпиријске моћи ових тест статистика су такође приказане у табелама 3.1 и 3.2.

3.6 Поређење емпиријских моћи тестова

Да би одређени тестови били конкурентни међу великим бројем различитих тестова, потребно је да имају неке добре карактеристике. Критеријум који се најчешће предлаже за избор адекватног теста је да треба изабрати најмоћнији тест. Моћ теста је дефинисана као вероватноћа да ће тест одбацити нулту хипотезу када заправо није тачна и треба да буде одбачена. Према томе, статистички тест се разматра као добар ако има малу вероватноћу одбацивања нулте хипотезе када је она тачна, али велику вероватноћу одбацивања нулте хипотезе када је нетачна. Та мера квалитета тестова у случају малих узорака и узорака умерене величине ће бити коришћена у овом поглављу.

Емпиријске моћи су рачунате користећи Монте Карло симулације за узорке обима $n = 20$ и $n = 50$, са $N = 10000$ понављања и нивоом значајности $\alpha = 0.05$. Као алтернативне расподеле коришћене су

- Вејбулова $W(\theta)$ расподела са густином

$$g(x;\theta) = e^{-x^{\theta}}\theta x^{\theta-1}, x \geqslant 0, \theta > 0; \tag{3.30}$$

- гама $\Gamma(\theta)$ расподела са густином

$$g(x;\theta) = \frac{x^{\theta-1}e^{-x}}{\Gamma(\theta)}, x \geqslant 0, \theta > 0; \tag{3.31}$$

- полунормална HN расподела са густином

$$g(x) = \sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}, x \geqslant 0; \tag{3.32}$$

- униформна U расподела са густином

$$g(x) = 1, 0 \leqslant x \leqslant 1; \tag{3.33}$$

- Ченова $CH(\theta)$ расподела (*Chen (2000)*) са густином

$$g(x;\theta) = 2\theta x^{\theta-1}e^{x^{\theta}+2(1-e^{x^{\theta}})}, x \geqslant 0, \theta > 0; \tag{3.34}$$

- Расподела линеарне стопе отказа $LF(\theta)$ са густином

$$g(x;\theta) = e^{-x-\theta\frac{x^2}{2}}(1+\theta x), x \geqslant 0, \theta > 0; \tag{3.35}$$

- модификована расподела екстремних вредности $EV(\theta)$ са густином

$$g(x;\theta) = \frac{1}{\theta} e^{\frac{1-e^x}{\theta}+x}, x \geqslant 0, \theta > 0; \tag{3.36}$$

- логнормална $LN(\theta)$ расподела са густином

$$g(x;\theta) = \frac{1}{x\sqrt{2\pi\theta^2}} e^{-\frac{(\log x)^2}{2\theta^2}}, x \geqslant 0, \theta > 0; \tag{3.37}$$

- Дилонова $DL(\theta)$ расподела (*Dhillon (1981)*) са густином

$$g(x;\theta) = \frac{\theta+1}{x+1} (\log(x+1))^{\theta} e^{-(\log(x+1))^{\theta+1}}, x \geqslant 0, \theta > 0. \tag{3.38}$$

У табелама 3.1 и 3.2 приказане су емпиријске моћи тестова уведених у поглављима 3.2-3.4, као и моћи одговарајућих тестова са оптималном вредношћу параметра подешавања a добијеном користећи метод из поглавља 3.5. Такође, приказане су и емпиријске моћи граничних тест статистика, када $a \to \infty$, статистика L^2-типа из поглавља 3.3.

Прво што се може приметити из табела емпиријских моћи је да мера свих тестова је једнака нивоу значајности. Такође, може се приметити да су моћи за већину алтернатива умерене до високе. Ако се посматрају емпиријске моћи за мале узорке ($n = 20$) тестови засновани на Лапласовим трансформацијама имају међусобно блиске вредности емпиријских моћи и за одређену вредност параметра a код приближно истог броја алтернатива сваки од тих тестова има највећу вредност моћи у односу на остале. Ако се посматрају тестови засновани на емпиријској функцији расподеле, $W_n^{\mathcal{P}}$ има блиске вредности као и тестови засновани на Лапласовим трансформацијама, изузев код алтернатива $CH(0.5)$, $CH(1.5)$ и $\Gamma(0.4)$, међутим ни за једну алтернативу немају највећу вредност моћи. Са друге стране, $W_n^{\mathcal{D}}$ је углавном лошији од свих других тестова. Алтернатива код које сви тестови имају нешто слабије резултате је Вејбулова $W(0.8)$, што је алтернатива са тешким репом. Ако поредимо тестове према карактеризацијама, једини тестови где се уочава значајна разлика јесу тестови L^2-типа засновани на емпиријској функцији расподеле, где је тест $W_n^{\mathcal{P}}$ бољи од теста $W_n^{\mathcal{D}}$ изузев у случају алтернатива $CH(1.5)$, $LN(1.5)$ и $\Gamma(0.4)$. Код тестова заснованих на Лапласовој трансформацији, за алтернативе $CH(0.5)$, $W(0.7)$ и $\Gamma(0.4)$ може се приметити да су тестови засновани на Десуовој карактеризацији бољи, док код осталих алтернатива тестови засновани на различитим карактеризацијама имају блиске вредности. Наравно, у овој групи тестова може се уочити мањи или већи утицај параметра подешавања a на емпиријске моћи у зависности од алтернативе. У већини случајева са порастом параметра a расте и емпиријска моћ теста. Код тестова са Десуовом карактеризацијом то није случај једино код алтернатива $CH(0.5)$, $LN(0.8)$, $DL(1)$ и $\Gamma(0.4)$, док код тестова са Пури-Рубиновом карактеризацијом то није случај код алтернатива $\Gamma(2)$, $LN(0.8)$, $DL(1)$ и $DL(1.5)$. С обзиром на ту чињеницу, није зачуђујуће што је код тих алтернатива управо највећа

Табела 3.1: Проценат одбацивања нулте хипотезе за $n = 20$, $\alpha = 0.05$

Алт.	$Exp(1)$	$W(1.4)$	$\Gamma(2)$	HN	U	$CH(0.5)$	$CH(1)$	$CH(1.5)$	$LF(2)$	$LF(4)$	$EV(1.5)$	$LN(0.8)$	$LN(1.5)$	$DL(1)$	$DL(1.5)$	$W(0.8)$	$\Gamma(0.4)$
$W_n^{\mathcal{P}}$	5	47	60	27	75	2	20	61	35	50	55	48	0	34	76	1	0
$W_n^{\mathcal{D}}$	5	27	42	15	45	0	11	86	20	30	30	44	3	27	59	1	5
$M_{n,0.5}^{\mathcal{P}}$	5	46	66	25	64	15	18	84	35	49	46	57	1	39	81	1	30
$M_{n,1}^{\mathcal{P}}$	5	49	66	28	72	19	21	88	38	52	53	51	2	37	81	1	34
$M_{n,2}^{\mathcal{P}}$	5	50	67	31	75	21	23	89	40	55	56	45	6	37	81	2	36
$M_{n,5}^{\mathcal{P}}$	5	48	62	32	80	29	23	90	40	56	58	42	21	33	80	3	45
$M_{n,\widehat{a}}^{\mathcal{P}}$	5	47	63	27	74	33	19	88	37	52	53	52	2	39	80	2	46
$M_{n,\infty}^{\mathcal{P}}$	5	47	59	31	81	43	23	91	40	56	59	33	51	26	73	12	59
$M_{n,0.5}^{\mathcal{D}}$	5	34	55	15	33	67	11	66	20	29	26	61	14	36	76	15	81
$M_{n,1}^{\mathcal{D}}$	5	43	63	21	46	64	15	77	27	39	35	60	18	38	80	12	77
$M_{n,2}^{\mathcal{D}}$	5	44	64	23	56	60	18	83	31	44	42	56	24	38	83	11	74
$M_{n,5}^{\mathcal{D}}$	5	49	64	29	67	52	21	87	37	53	52	46	33	35	80	10	70
$M_{n,\widehat{a}}^{\mathcal{D}}$	5	41	57	22	64	55	18	85	31	46	44	56	16	34	78	13	74
$M_{n,\infty}^{\mathcal{D}}$	5	47	59	31	80	43	23	90	40	55	58	33	52	27	74	11	58
$L_n^{\mathcal{P}}$	5	47	64	25	67	12	17	84	34	46	47	55	1	39	82	1	27
$L_{n,0.5}^{\mathcal{P}}$	5	50	63	29	77	20	22	89	40	53	58	47	6	34	80	2	34
$L_{n,1}^{\mathcal{P}}$	5	50	62	31	79	26	22	90	41	55	55	40	12	32	81	3	42
$L_{n,2}^{\mathcal{P}}$	5	49	62	31	80	33	25	91	41	57	58	40	27	33	78	4	47
$L_{n,5}^{\mathcal{P}}$	5	51	62	31	80	37	23	91	42	57	60	34	39	29	76	7	53
$L_{n,\widehat{a}}^{\mathcal{P}}$	5	47	62	20	76	37	21	89	38	52	55	43	11	32	77	5	51
$L_n^{\mathcal{D}}$	5	36	57	16	45	61	12	73	24	33	32	61	14	37	77	12	75
$L_{n,0.5}^{\mathcal{D}}$	5	44	64	24	61	56	17	84	33	47	44	54	23	36	88	10	71
$L_{n,1}^{\mathcal{D}}$	5	47	64	27	66	54	20	86	35	50	50	49	30	35	81	10	69
$L_{n,2}^{\mathcal{D}}$	5	48	64	29	71	52	21	89	37	53	52	42	37	33	80	10	68
$L_{n,5}^{\mathcal{D}}$	5	50	64	32	78	49	23	91	41	57	58	40	48	30	77	12	65
$L_{n,\widehat{a}}^{\mathcal{D}}$	5	45	62	26	71	49	19	87	34	50	50	50	27	36	79	11	68

моћ тестова $M_{n,a}^{\mathcal{P}}$ и $M_{n,a}^{\mathcal{D}}$ добијена за тестове код којих $a \to \infty$, односно за граничне тест статистике добијене у поглављу 3.5.

Код узорака умерене величине ($n = 50$), резултати су врло слични са очекиваним повећањем вредности у односу на мале узорке.

Ако емпиријске моћи упоредимо са моћима неких ранијих тестова (видети на пример *Torabi et al. (2018)*) може се приметити да нови тестови имају или веће или сличне вредности моћи као ти тестови. Због тога се нови тестови препоручују за тестирање експонцијалности на основу резултата добијених у случају малих узорака и узорака умерене величине.

Табела 3.2: Проценат одбацивања нулте хипотезе за $n = 50$, $\alpha = 0.05$

Алт.	$Exp(1)$	$W(1.4)$	$\Gamma(2)$	HN	U	$CH(0.5)$	$CH(1)$	$CH(1.5)$	$LF(2)$	$LF(4)$	$EV(1.5)$	$LN(0.8)$	$LN(1.5)$	$DL(1)$	$DL(1.5)$	$W(0.8)$	$\Gamma(0.4)$
$W_n^{\mathcal{P}}$	5	85	95	54	99	16	38	100	69	86	90	89	1	70	99	0	44
$W_n^{\mathcal{D}}$	5	59	82	29	86	69	19	96	41	60	61	90	15	60	96	7	88
$M_{n,0.5}^{\mathcal{P}}$	5	84	97	48	95	90	33	100	65	83	81	94	36	77	100	17	97
$M_{n,1}^{\mathcal{P}}$	5	85	97	54	97	90	38	100	69	87	86	89	50	72	100	20	97
$M_{n,2}^{\mathcal{P}}$	5	86	96	57	98	91	41	100	73	89	90	83	65	67	100	24	98
$M_{n,5}^{\mathcal{P}}$	5	87	96	63	99	90	45	100	76	91	93	71	80	59	99	27	97
$M_{n,\widehat{a}}^{\mathcal{P}}$	5	85	97	57	99	92	43	100	73	89	91	95	51	76	100	23	97
$M_{n,\infty}^{\mathcal{P}}$	5	84	94	63	99	90	46	100	78	92	94	53	92	48	98	38	97
$M_{n,0.5}^{\mathcal{D}}$	5	70	93	25	65	97	16	97	38	59	52	98	42	77	99	35	99
$M_{n,1}^{\mathcal{D}}$	5	76	95	35	79	97	21	99	49	70	64	97	55	76	100	34	99
$M_{n,2}^{\mathcal{D}}$	5	81	96	43	90	96	29	100	59	79	76	92	66	73	100	35	99
$M_{n,5}^{\mathcal{D}}$	5	83	96	51	96	95	35	100	67	86	85	82	81	66	100	36	99
$M_{n,\widehat{a}}^{\mathcal{D}}$	5	81	95	47	97	96	36	100	66	85	82	97	56	75	100	36	99
$M_{n,\infty}^{\mathcal{D}}$	5	85	93	62	100	90	45	100	77	91	94	54	92	47	98	38	97
$L_n^{\mathcal{P}}$	5	84	96	49	96	88	33	100	64	85	81	93	36	77	100	18	97
$L_{n,0.5}^{\mathcal{P}}$	5	87	96	57	98	90	42	100	73	89	91	83	64	68	100	23	97
$L_{n,1}^{\mathcal{P}}$	5	86	96	60	99	90	45	100	75	91	92	76	74	63	99	27	97
$L_{n,2}^{\mathcal{P}}$	5	86	95	62	99	91	45	100	76	91	93	69	83	60	99	30	98
$L_{n,5}^{\mathcal{P}}$	5	86	94	62	99	91	46	100	78	92	93	60	89	52	99	33	97
$L_{n,\widehat{a}}^{\mathcal{P}}$	5	85	96	58	99	91	45	100	75	90	92	84	78	52	100	30	97
$L_n^{\mathcal{D}}$	5	73	94	28	81	96	18	98	46	67	63	97	47	77	99	32	99
$L_{n,0.5}^{\mathcal{D}}$	5	83	96	45	96	96	31	100	63	82	79	93	69	72	100	32	99
$L_{n,1}^{\mathcal{D}}$	5	85	96	51	96	95	34	100	66	86	85	87	78	69	100	36	99
$L_{n,2}^{\mathcal{D}}$	5	86	96	56	98	95	39	100	72	88	89	79	85	63	99	37	99
$L_{n,5}^{\mathcal{D}}$	5	86	95	59	99	94	43	100	74	90	92	65	89	56	99	39	98
$L_{n,\widehat{a}}^{\mathcal{D}}$	5	83	95	55	99	94	41	100	72	88	90	92	79	71	99	37	99

Глава 4

АСИМПТОТСКА ЕФИКАСНОСТ ТЕСТОВА САГЛАСНОСТИ

С обзиром да је током година развијен велики број тестова, једно од главних питања истраживача је који тест користити када треба утврдити да ли узорак потиче из неке конкретне расподеле. Другим речима, постоји потреба за утврђивањем који тест је најбољи од низа тестова. Већ је поменуто да се тестови најчешће пореде према њиховим способностима да детектују алтернативну хипотезу, односно моћима. Код параметарских тестова тај метод упоређивања често даје један тест који је оптималан, односно чија је моћ увек већа или једнака од моћи других тестова ако су им нивои значајности једнаки. Међутим, у непараметарској теорији су врло ретке ситуације у којима се неки тест издваја као униформно најмоћнији у односу на друге тестове.

Избору адекватног теста може допринети и познавање природе популације из које је узорак узет, као и испуњеност услова за примену неког модела. Међутим, тест који је на овај начин адекватан за популацију која се проучава не мора бити најмоћнији тест. Овај проблем се може превазићи избором теста који више одговара подацима, а онда повећањем обима узорка, ако је могуће, може се повећати и његова моћ.

Наравно, није увек могуће узети довољно велики узорак, али није ни погодно захтевати велики обим узорка ако се иста моћ може постићи са другим тестом и значајно мањим узорком. Појам који је повезан са тим колико треба повећати обим узорка за примену одређеног теста како би се постигла иста моћ као код најмоћнијег теста је релативна ефикасност. Постоји доста примера у којима је рачунањем релативне ефикасности, или њене асимптотске варијанте, довело до преиспитивања коришћења неких непараметарских тестова. На пример, за тест низа Валд-Волфовица (*Wald and Wolfowitz (1940)*), који је био врло популаран за тестирање хомогености, испоставило се да има ефикасност једнаку 0 у односу на параметарске тестове (за више детаља видети *Mood (1954)* и *Bahadur (1960)*). Због тога је тест нагло престао да се користи и замењен је ефикаснијим тестовима. Дакле, други начин за избор адекватног теста је коришћење релативне ефикасности. Редослед тестова који се добија поређењем њихових моћи не мора се поклапати са редоследом који се добија поређењем ефикасности (*Nikitin (1995)*).

4.1 Асимптотска релативна ефикасност

Нека је $X_1, ..., X_n$ узорак дефинисан на простору $(\Omega, \mathcal{A}, P_\theta)$, при чему је расподела случајних величина одређена параметром θ из параметарског простора Θ. Претпоставимо да се тестира $H_0: \theta \in \Theta_0 \subset \Theta$ против алтернативе $H_1: \theta \in \Theta_1 = \Theta \backslash \Theta_0$ користећи низ тест статистика $\{T_n\}$.

Без губитка општости претпоставимо да су значајне велике вредности тест статистике, односно да је критична област облика $\{T_n \geqslant q\}$, где је q неки реалан број. Ако су мале вредности тест статистике значајне, онда се може разматрати $-T_n$ или нека друга строго опадајућа функција од T_n, а ако су и мале и велике вредности од T_n значајне, онда се уместо T_n може користи на пример $|T_n - K_n|$, где је K_n константа, или се теорија приказана у наставку уопште не може применити уколико не постоји могућност трансформисања на наведени начин.

Означимо са $M(\theta) = P_\theta\{T_n \geqslant q\}$ функцију моћи овако дефинисаног теста. Нека је ниво значајности једнак мери теста $\sup\limits_{\theta \in \Theta_0} M(\theta)$. Нека је $\beta \in (0,1)$, $\theta \in \Theta_1$, и нека је $q_n := q_n(\beta, \theta)$ низ вредности тако да важи

$$P_\theta\{T_n > q_n\} \leqslant \beta \leqslant P_\theta\{T_n \geqslant q_n\}.$$

Дефинишемо са $N_T(\alpha, \beta, \theta)$ најмању величину узорка неопходну са нивоом α, $0 < \alpha < \beta$, на основу низа статистика $\{T_n\}$, да се достигне моћ не мања од β у тачки θ.

Претпоставимо да за тестирање H_0 против H_1 се користе два низа тест статистика $\{T_n^{(1)}\}$ и $\{T_n^{(2)}\}$. Релативна ефикасност низа $\{T_n^{(2)}\}$ у односу на $\{T_n^{(1)}\}$ је дефинисана на следећи начин

$$e_{T^{(1)},T^{(2)}} = \frac{N_{T^{(2)}}(\alpha, \beta, \theta)}{N_{T^{(1)}}(\alpha, \beta, \theta)}. \tag{4.1}$$

Вредност $e_{T^{(1)},T^{(2)}}(\alpha, \beta, \theta)$ већа од 1 показује да за дате α, β и θ треба предност дати низу $\{T_n^{(1)}\}$ у односу на низ $\{T_n^{(2)}\}$ јер први низ захтева мањи обим узорка за постизање исте моћи β са нивоом α и алтернативом одређеном са θ. Ако је $e_{T^{(1)},T^{(2)}}(\alpha, \beta, \theta)$ мања од 1, онда важи обрнуто. Међутим, релативну ефикасност је у општем случају врло тешко или чак немогуће израчунати. Тај проблем је могуће превазићи рачунањем граничне вредности од $e_{T^{(1)},T^{(2)}}(\alpha, \beta, \theta)$, ако она постоји, када један од параметара тежи својој граничној вредности док су преостала два фиксирана. Случајеви када је мали ниво значајности, висока моћ или када су блиске алтернативе су у суштини најзначајнији случајеви, због чега асимптотске релативне ефикасности дају довољно информација које омогућавају да се тестови рангирају. На тај начин се добијају три типа асимптотске релативне ефикасности:

- Ако за $\beta \in (0,1)$ и $\theta \in \Theta_1$ постоји гранична вредност

$$e^B_{T^{(1)},T^{(2)}}(\beta, \theta) = \lim_{\alpha \to 0} e_{T^{(1)},T^{(2)}}(\alpha, \beta, \theta),$$

онда се та гранична вредност зове Бахадурова асимптотска релативна ефикасност низа $\{T_n^{(2)}\}$ у односу на $\{T_n^{(1)}\}$.

- Ако за $\alpha \in (0,1)$ и $\theta \in \Theta_1$ постоји гранична вредност

$$e^{HL}_{T^{(1)},T^{(2)}}(\alpha,\theta) = \lim_{\beta\to 1} e_{T^{(1)},T^{(2)}}(\alpha,\beta,\theta),$$

онда се та гранична вредност зове Хаџис-Леманова асимптотска релативна ефикасност низа $\{T_n^{(2)}\}$ у односу на $\{T_n^{(1)}\}$.

- Ако за $0 < \alpha < \beta < 1$ и $\theta \to \theta_0 \in \partial\Theta_0$ постоји гранична вредност

$$e^{P}_{T^{(1)},T^{(2)}}(\alpha,\beta,\theta_0) = \lim_{\theta\to\theta_0} e_{T^{(1)},T^{(2)}}(\alpha,\beta,\theta),$$

онда се та гранична вредност зове Питманова асимптотска релативна ефикасност низа $\{T_n^{(2)}\}$ у односу на $\{T_n^{(1)}\}$.

Поред ове три асимптотске ефикасности које се најчешће користе, могуће је дефинисати и асимптотску ефикасност када два од три параметра конвергирају ка својим граничним вредностима док је трећи параметар фиксиран. Такве су, на пример, Черновљева (*Chernoff (1952)*) и Келембергова (*Kallenberg (1983)*) асимптотска релативна ефикасност.

Међутим, иако значајно једноставније од релативне ефикасности, и асимптотске релативне ефикасности су у општем случају тешке за израчунавање. Значајно је да: Бахадурова асимптотска релативна ефикасност обично не зависи од β; Хаџис-Леманова асимптотска релативна ефикасност обично не зависи од α; док Питманова асимптотска релативна ефикасност када је гранична расподела статистика нормална не зависи ни од α ни од β. У случају када гранична расподела статистика није нормална расподела, Питманова асимптотска релативна ефикасност зависи и од α и од β и сувише је компликована за рад. Пошто је то случај са тест статистикама приказаним у поглављу 3, овај тип асимптотске релативне ефикасности није погодно користити. Друга два типа асимптотске релативне ефикасности се базирају на грубој процени вероватноћа великих одступања. Рачунање Хаџис-Леманове асимптотске релативне ефикасности захтева проучавање великих одступања статистика при алтернативној хипотези при чему статистике више нису слободне од параметра. То отежава рад са овим типом ефикасности. Због тога ћемо се у наставку бавити Бахадуровом асимптотском релативном ефикасношћу. Штавише, испоставља се, на пример код велике класе тестова чије се тест статистике могу приказати као линеарне комбинације статистика ранга, да се локално све три асимптотске релативне ефикасности подударају (*Nikitin (1995)*).

Пошто се тест статистике приказане у поглављу 3 могу написати или као $V-$статистике или као супремум $V-$емпиријског процеса, одређивање Бахадурове релативне ефикасности ће бити приказано за класу $V-$статистика. Више детаља о Бахадуровој асимптотској релативној ефикасности у општем случају се може наћи у *Bahadur (1971)* или *Nikitin (1995)*.

4.2 Бахадуров нагиб $V-$статистика

Као што је већ раније поменуто, Бахадурова асимптотска релативна ефикасност подразумева да се фиксира моћ теста и пореди стопа опадања нивоа значајности како број опсервација расте.

Нека је простор $(\Omega, \mathcal{A})$ узорачки простор бесконачно много независних и једнако расподељених случајних величина $X_1, X_2, ...$, при чему је њихова расподела одређена параметром θ који узима вредности у простору Θ. Нека је Θ_0 подскуп од Θ, такав да се при нултој хипотези тестира да је параметар θ из Θ_0.

За свако $n \in \mathbb{N}$ нека је T_n произвољна тест статистика која зависи од случајних величина $X_1, X_2, ...$ само кроз узорак $\mathbf{X} = (X_1, ..., X_n)$. Такође, претпоставимо да су велике вредности од T_n значајне. Претпоставимо да постоји функција $G_n(t)$ таква да је

$$G_n(t) = \inf_{\theta \in \Theta_0} P_\theta\{T_n(\mathbf{X}) < t\},$$

за све $t \in \mathbb{R}$. Тада је $p-$вредност теста дефинисана са $L_n(\mathbf{X}) = 1 - G_n(T_n(\mathbf{X}))$. Наравно, L_n је случајна величина, односно статистика. Асимптотско понашање од L_n при алтернативној расподели је од великог значаја за поређење низова тест статистика. Кажемо да низ $\{T_n\}$ има тачан Бахадуров нагиб $c(\theta)$ за дато θ ако скоро сигурно важи

$$\lim_{n \to \infty} \frac{1}{n} \ln L_n(\mathbf{X}) = -\frac{1}{2} c_T(\theta). \tag{4.2}$$

Термин ,,тачан" Бахадуров нагиб има улогу да уведе разлику у односу на ,,приближни" Бахадуров нагиб, о којем ће касније бити речи.

Следећа теорема говори о повезаности величине $N_T(\alpha, \beta, \theta)$ са тачним Бахадуров нагибом. Доказ се може наћи у *Bahadur (1971)*.

Теорема 4.1 *Ако* (4.2) *важи и* $0 < c_T(\theta) < \infty$*, онда скоро сигурно*

$$\lim_{\alpha \to 0} \frac{N(\alpha, \beta, \theta)}{2 \ln \frac{1}{\alpha}} = \frac{1}{c_T(\theta)}.$$

Односно, за мало α, N_T је пропорционално инверзу тачног Бахадуровог нагиба $c_T(\theta)$.

Претпоставимо да су $\{T_n^{(1)}\}$ и $\{T_n^{(2)}\}$ два низа тест статистика, таква да тест статистика $T_n^{(k)}$, $k = 1, 2$, има тачан Бахадуров нагиб $c_k(\theta)$ и претпоставимо да за $\theta \in \Theta_1$ важи $0 < c_k(\theta) < \infty$. Тада, на основу (4.1) и теореме 4.1 следи да $e_{T^{(1)}, T^{(2)}} \xrightarrow{P_\theta} \frac{c_1(\theta)}{c_2(\theta)}$, кад $\alpha \to 0$. Односно количник тачних Бахадурових нагиба $\frac{c_1(\theta)}{c_2(\theta)}$ представља меру Бахадурове асимптотске ефикасности од $T_n^{(2)}$ у односу на $T_n^{(1)}$.

Следећа теорема даје користан метод за одређивање тачног Бахадуровог нагиба датог низа $\{T_n\}$ за који важи (4.2). Доказ теореме се такође може пронаћи у *Bahadur (1971)*.

Теорема 4.2 *Нека за низ* $\{T_n\}$ *важи*

$$T_n \overset{P_\theta}{\to} b(\theta),$$

за свако $\theta \in \Theta_1$, *где је* $-\infty < b(\theta) < \infty$, *и*

$$\lim_{n\to\infty} \frac{1}{n} \ln(1 - G_n(t)) = -h(t)$$

за свако t *из отвореног интервала* I *на којем је* h *непрекидна и* $\{b(\theta), \theta \in \Theta_1\} \subset I$. *Тада* (4.2) *важи и додатно, за свако* $\theta \in \Theta_1$,

$$c_T(\theta) = 2h(b(\theta)). \tag{4.3}$$

У случајевима када није могуће одредити функцију h, тачна расподела од $\{T_n\}$ се може заменити асимптотском расподелом и важиће аналогон претходне теореме, с тим што се у овом случају говори о приближном Бахадуровом нагибу а не тачном. Доказ следећег тврђења се може наћи у *Bahadur (1960)*.

Теорема 4.3 *Претпоставимо да су за низ статистика* $\{T_n\}$ *испуњени следећи услови*

1. *Постоји непрекидна функција расподеле* F *таква да, за свако* $\theta \in \Theta_0$,

$$\lim_{n\to\infty} P_\theta\{T_n \leqslant t\} = F(t), \text{ за свако } t.$$

2. *Постоји константа* a, $0 < a < \infty$, *таква да*

$$\ln(1 - F(t)) = -\frac{at^2}{2}(1 + o_p(1)) \text{ кад } t \to \infty.$$

3. *Постоји функција* b *на* Θ_1, *са* $0 < b < \infty$, *таква да, за свако* $\theta \in \Theta_1$, $T_n \overset{P}{\to} b(\theta)$, *кад* $n \to \infty$.

Тада важи

$$\lim_{n\to\infty} \frac{1}{n} \ln(1 - F(t)) = \frac{1}{2} c^{(a)}(\theta)$$

и, при том, $c^{(a)}(\theta) = ab^2(\theta)$.

Функција $c^{(a)}(\theta)$ се назива приближни Бахадуров нагиб. У пракси, разлози за одређивање приближног Бахадуровог нагиба, а не тачног, су вишеструки. На пример, ако тачна расподела тест статистике при нултој хипотези није позната (или је тешко одредити) или је n довољно велико па није неопходно радити са тачном расподелом или зато што тачан нагиб не постоји, односно за дато n расподела тест статистике се мења у зависности од θ из Θ_0. У случају када се тачни нагиб не може одредити, пожељно је да приближни нагиб да назнаку о понашању тачног нагиба а врло често се локално, односно у случају блиских алтернатива, они и поклапају. Тада приближна Бахадурова асимптотска релативна ефикасност даје представу о тачној Бахадуровој асимптотској релативној ефикасности.

У наставку ће бити приказана Бахадурова теорија примењена на специјалну класу статистика, односно оне статистике које се могу приказати као $V-$статистике. Прве теореме дају функције великих одступања за наведене статистике у недегенерисаном и слабо дегенерисаном случају, а на основу њих су одређене опште формуле за рачунање тачних Бахадурових нагиба.

Нека је $X_1, ..., X_n$ узорак са познатом непрекидном расподелом F. Нека је $\Phi(x_1, ..., x_m)$ симетрично језгро реда m и нека је V_n $V-$статистика са језгром Φ дефинисана као у (2.2). Уместо посматрања језгра $\Phi(X_1, .., X_m)$ увек се може посматрати језгро

$$\widetilde{\Phi}(U_1, ..., U_m) = \Phi(F^{-1}(U_1), ..., F^{-1}(U_m)),$$

где је $U_1, U_2, ...$ узорак из униформне расподеле на интервалу $[0, 1]$. То омогућава да се увек може претпоставити да узорак потиче из униформне расподеле на интервалу $[0, 1]$, па се услови везани за расподелу случајних величина редукују на услове везане за језгро.

Као што се може видети из (4.2), потребно је одредити функцију великих одступања. Дакле, ако је V_n $V-$статистика са очекивањем нула треба одредити граничну вредност

$$h_v(a) = \lim_{n\to\infty} \frac{1}{n} \ln P\{V_n \geqslant a\},$$

барем за мале вредности a.

Следеће две теореме дају понашање функције h_v у случају недегенерисане и слабо дегенерисане $V-$статистике. Њихови докази се могу наћи у *Nikitin and Ponikarov (2001).*

Теорема 4.4 *Претпоставимо да је језгро $\widetilde{\Phi}$ $V-$статистике ограничена функција на $[0, 1]^m$, $|\widetilde{\Phi}(x_1, ..., x_m)| \leqslant M$, $E(\widetilde{\Phi}(X_1, ..., X_m)) = 0$ и има ранг 1, односно $\sigma_1^2 = E(\widetilde{\varphi}_1^2(X_1)) > 0$ где је $\widetilde{\varphi}_1(x) = E(\widetilde{\Phi}(X_1, ..., X_m)|X_1 = x)$. Тада*

$$\lim_{n\to\infty} \frac{1}{n} \ln P\{V_n \geqslant a\} = \sum_{j=2}^{\infty} b_j a^j,$$

при чему је на десној страни конвергентан ред за довољно мало a. Поред тога, $b_2 = -\frac{1}{2m^2\sigma_1^2}$.

У суштини могуће је наћи било који коефицијент b_j реда са десне стране, међутим чак и одређивање b_3 захтева компликован рачун.

Теорема 4.5 *Претпоставимо да је језгро $\widetilde{\Phi}$ $V-$статистике ограничена функција на $[0, 1]^m$, $|\widetilde{\Phi}(x_1, ..., x_m)| \leqslant M$, $E(\widetilde{\Phi}(X_1, ..., X_m)) = 0$ и има ранг 2. Нека је v_1 највеће v које задовољава интегралну јеначину*

$$q(x) = \frac{1}{v} \int_0^1 \widetilde{\varphi}_2(x, y) q(y) dy,$$

где је $\widetilde{\varphi}_2(x,y) = E(\widetilde{\Phi}(X_1,...,X_m)|X_1 = x, X_2 = y)$, и претпоставимо да је υ_1 сопствена вредност линеарног интегралног оператора са језгром φ_2 који пресликава $L^2[0,1]$ у $L^2[0,1]$. Тада

$$\lim_{n\to\infty} \frac{1}{n} \ln P\{V_n \geqslant a\} = \sum_{j=2}^{\infty} b_j a^{\frac{j}{2}},$$

при чему је на десној страни конвергентан ред за довољно мало a. Поред тога, $b_2 = -\frac{\upsilon_1}{2\binom{m}{2}}$.

Користећи теорему 4.2 и теорему 4.4 у недегенерисаном случају, односно теорему 4.5 у слабо дегенерисаном случају, теоријски се добија израз за тачан Бахадуров нагиб. Међутим, у пракси је готово немогуће израчунати тај израз до краја. Због тога дефинишемо специјалну класу блиских алтернатива која омогућава да се тачни Бахадуров нагиб одреди локално.

Нека је $\mathcal{G} = \{G(x;\theta),\ \theta > 0\}$ фамилија алтернативних расподела, са густинама $\{g(x;\theta),\ \theta > 0\}$ и коначним очекивањем $\mu(\theta)$, за коју важи да је $\theta = 0$ онда и само онда када је $G(x;0)$ функција расподеле из нулте хипотезе. Нека је $X_1,...,X_n$ узорак из расподеле са густином $g(x;\theta)$. Тада желимо да тестирамо $H_0:\ \theta = 0$ против алтернативе $H_1:\ \theta > 0$. Претпоставимо да се тест статистика V_n која се користи за тестирање хипотезе може написати као $V-$статистика са језгром Φ и нека су значајне њене велике вредности. Такође, претпостављамо да је такав тест постојан, односно да је

$$b_\Phi(\theta) = E_\theta(\Phi(X_1,...,X_m)) > 0$$

за било које $\theta \in \Theta$, и нека $b_\Phi(\theta) \to 0$, кад $\theta \to 0$. Из закона великих бројева за $V-$статистике (теорема 2.5), следи да при H_1, за $\theta > 0$,

$$V_n \xrightarrow{P} b_\Phi(\theta). \tag{4.4}$$

Из ових претпоставки и теорема 4.4 и 4.5, следи да су услови теореме 4.2 задовољени, па следи да

$$c_V(\theta) \sim \frac{b_\Phi^2(\theta)}{m^2\sigma_1^2} \tag{4.5}$$

кад $\theta \to 0$ у недегенерисаном случају, односно

$$c_V(\theta) \sim \frac{2b_\Phi(\theta)}{m(m-1)\upsilon_1} \tag{4.6}$$

кад $\theta \to 0$ у слабо дегенерисаном случају. Ови нагиби се називају тачним локалним Бахадуровим нагибима.

С обзиром да су тестови уведени у поглављу 3 за тестирање експоненцијалности расподеле слободни од параметра расподеле при нултој хипотези, параметар θ је заправо дводимензионални параметар (ϑ, λ). У том случају је класа $\mathcal{G}_\mathcal{E}$ алтернативних расподела одређена са $\mathcal{G}_\mathcal{E} = \{G(x;\boldsymbol{\theta}), \boldsymbol{\theta} = (\vartheta,\lambda), \vartheta > 0, \lambda > 0\}$, са густинама $\{g(x;\boldsymbol{\theta}), \boldsymbol{\theta} = (\vartheta,\lambda), \vartheta > 0, \lambda > 0\}$, при чему је $g(x;\boldsymbol{\theta}) = \lambda e^{-\lambda x}, x \geqslant 0$, ако и само ако је $\boldsymbol{\theta} = (0,\lambda)$. У том случају се тестира

$H_0 : \boldsymbol{\theta} \in \{(0, \lambda), \lambda > 0\}$ против алтернативе $H_1 : \boldsymbol{\theta} \in \{(\vartheta, \lambda), \vartheta > 0, \lambda > 0\}$.

У следећим поглављима ће претходна теорија бити примењена на тестове сагласности. У општем случају ће бити кориђена класа $\mathcal{G}$, док ће за тестове (3.1), (3.10) и (3.21) бити коришћена класа $\mathcal{G}_{\mathcal{E}}$.

4.3 Бахадурова асимптотска ефикасност тестова сагласности L^2-типа

Тест статистике тестова сагласности L^2-типа одређене формулама (3.1) и (3.10) могу се представити као слабо дегенерисане $V-$статистике, као што је показано у поглављу 3, па се на њих могу применити резултати претходних теорема. Међутим, приликом рачунања асимптотске ефикасности оваквих тестова, разликоваћемо два случаја: слабо дегенерисане $V-$статистике без оцењеног параметра и слабо дегенерисане $V-$статистике са оцењеним параметром. Као што је показано раније, оцена параметра не мора имати утицаја на граничну расподелу статистике. Самим тим, уколико оцена нема утицаја на расподелу, Бахадуров нагиб ће имати значајно једноставнији облик него када параметар има утицаја на граничну расподелу.

Претпоставимо да је фамилија алтернативних расподела $\mathcal{G}$ дефинисана као у претходном поглављу. Такође, претпостављамо да су испуњени услови који су потребни да би важило (4.6).

Следећа теорема даје тачан Бахадуров нагиб, када је параметар θ близак нули (локално), $V-$статистика са језгром $\Phi(x_1, ..., x_m)$ које је слабо дегенерисано. Доказ ове теореме следи директно из (4.6) и Маклореновог развоја (за више детаља видети *Nikitin and Peaucelle (2004)*).

Теорема 4.6 *Нека је $\{T_n\}$ низ слабо дегенерисаних $V-$статистика са ограниченим језгрима и нека је $b_\Phi(\theta) > 0$. Нека је $g(x;\theta)$ алтернатива из $\mathcal{G}$ и нека постоји $\delta > 0$ тако да за $0 \leqslant \theta \leqslant \delta$ важи*

$$\int_{-\infty}^{+\infty} |g'_\theta(x;\theta)|dx < \infty, \int_{-\infty}^{+\infty} |g''_{\theta\theta}(x;\theta)|dx < \infty, \int_{-\infty}^{+\infty} |g'''_{\theta\theta\theta}(x;\theta)|dx < \infty, \tag{4.7}$$

где су g'_θ, $g''_{\theta\theta}$ и $g'''_{\theta\theta\theta}$ одговарајући изводи функције g по θ. Тада кад $\theta \to 0$

$$c_T(\theta) \sim \frac{1}{\upsilon_1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi_2(x,y) g'_\theta(x;0) g'_\theta(y;0) dxdy \cdot \theta^2, \tag{4.8}$$

где је $\varphi_2(x,y) = E(\Phi(X_1, ..., X_m)|X_1 = x, X_2 = y)$ друга пројекција језгра статистике Φ, а υ_1 највећа сопствена вредност оператора дефинисаног у (2.13).

Ако је тест статистика слабо дегенерисана $V-$статистика са оцењеним параметром, онда је ситуација нешто компликованија. У том случају се не могу одредити функције великих одступања, па се одређује приближни Бахадуров нагиб користећи граничну расподелу ових статистика. Следећа теорема је доказана у радовима аутора (*Cuparić (2019)* и *Cuparić et al. (2019a)*).

Због компактности записа, означимо са $\boldsymbol{x} = (x_1, ..., x_{2m})$ и $\boldsymbol{G}(\boldsymbol{x};\theta) = \prod_{i=1}^{2m} G(x_i;\theta)$.

Теорема 4.7 (Цупарић 2019) *Нека је $\{T_n(\widehat{\mu}_n)\}$ низ слабо дегенерисаних $V-$статистика са оцењеним параметром са ограниченим језгрима $\Phi(\boldsymbol{x};\widehat{\mu}_n)$ реда $2m$, таквим да постоји $\frac{\partial^2}{\partial\theta^2}\Phi(\boldsymbol{x};\mu(\theta))$, и нека је $b_\Phi(\theta) > 0$. Нека је $g(x;\theta)$ алтернатива из $\mathcal{G}$ и нека постоји $\delta > 0$ тако да за $0 \leqslant \theta \leqslant \delta$ важи (4.7). Тада кад $\theta \to 0$*

$$c_T^{(a)}(\theta) \sim \frac{1}{2m(2m-1)\upsilon_1}\Bigg(2m(2m-1)\int\limits_{\mathbb{R}^2}\varphi_2(x_1,x_2;1)\frac{\partial dG(x_1;0)}{\partial\theta}\frac{\partial dG(x_2;0)}{\partial\theta} + \left(\frac{\partial\mu(0)}{\partial\theta}\right)^2\int\limits_{\mathbb{R}^m}\frac{\partial^2\Phi(\boldsymbol{x},1)}{\partial\theta^2}d\boldsymbol{G}(\boldsymbol{x};0) + 4m\frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial\Phi(\boldsymbol{x},1)}{\partial\theta}\frac{\partial dG(x_1;0)}{\partial\theta}\prod_{i=2}^{m}dG(x_i;0)\Bigg)\theta^2, \tag{4.9}$$

где је $\varphi_2(x,y) = E(\Phi(X_1,...,X_{2m})|X_1 = x, X_2 = y)$ друга пројекција језгра статистике Φ, а υ_1 највећа сопствена вредност оператора дефинисаног у (2.25).

Пре доказа ове теореме увешћемо лему која је неопходна за доказивање теореме 4.7 (за више детаља видети *Zolotarev (1961)*).

Лема 4.1 *Нека су $Z_1, Z_2, ...$ независне случајне величине са стандардном нормалном расподелом и нека је*

$$\lambda_1^2 = \lambda_2^2 = \cdots = \lambda_{n_1}^2,$$
$$\lambda_{n_1+1}^2 = \lambda_{n_1+2}^2 = \cdots = \lambda_{n_1+n2}^2,$$
$$\cdots$$

низ позитивних бројева таквих да вредности из блокова чине строго опадајући низ и $\sum\limits_{k=1}^{\infty}\lambda_k^2 < \infty$. Нека је $\eta = \sum\limits_{k=1}^{\infty}\lambda_k^2 Z_k^2$. Тада следи да

$$1 - F_\eta(x) = \frac{K}{\Gamma(\frac{n_1}{2})}\left(\frac{x}{2\lambda_1^2}\right)^{\frac{n_1}{2}-1}e^{-\frac{x}{2\lambda_1^2}}(1+\varepsilon(x)), \tag{4.10}$$

где се K може одредити из израза

$$\frac{(2\lambda_1^2)^{n_1}\Gamma(n_1+1)}{K^2} = \frac{d^{n_1}}{ds^{n_1}}\mathcal{L}^{-2}(s)|_{s=-\frac{1}{2\lambda_1^2}} \tag{4.11}$$

при чему је $\mathcal{L} = E(e^{-s\eta})$ Лапласова трансформација од η, и $\varepsilon(x) \to 0$ кад $x \to \infty$.

Доказ теореме 4.7: Из теореме 2.9 следи да статистика $nT_n(\widehat{\mu}_n)$ конвергира у расподели ка случајној величини $\eta = \frac{2m(2m-1)}{2}\sum\limits_{i=1}^{\infty}\upsilon_i Z_i^2$, где је $\{Z_i\}$ низ независних случајних величина са стандардном нормалном расподелом и $\{\upsilon_i\}$ низ сопствених вредности оператора (2.25). Како је језгро оператора позитивно дефинитно, све сопствене вредности су позитивне. Нека је $\lambda_k = \frac{2m(2m-1)}{2}\upsilon_k,\ k=1,2,...$ Применом леме 4.1, добија се

$$1 - F_\eta(x) = \frac{K}{\Gamma(\frac{n_1}{2})}\Big(\frac{x}{2\lambda_1}\Big)^{\frac{n_1}{2}-1} e^{-\frac{x}{2\lambda_1}}(1+\varepsilon(x)),$$

где је K константа која се добија из (4.11). Односно

$$\ln(1-F_\eta(x)) = \ln\Big(\frac{K}{\Gamma(\frac{n_1}{2})}\Big) - \Big(\frac{n_1}{2}-1\Big)\ln\Big(\frac{x}{2\lambda_1}\Big) - \frac{x}{2\lambda_1} + \ln(1+\varepsilon(x)).$$

Следи да, када $x \to \infty$, важи

$$\ln(1-F_\eta(x)) = -\frac{x}{2\lambda_1} + o_p(x).$$

Посматрајмо сада статистику $\widetilde{T}_n = \sqrt{T_n}$. Из претходно доказаног, уз примену теореме о непрекидном пресликавању, важи

$$\ln(1-F_{\widetilde{T}}(s)) = -\frac{s^2}{2m(2m-1)\upsilon_1} + o_p(s^2),\ s \to \infty.$$

Такође, важи да $\widetilde{T}_n \xrightarrow{P} b_{\widetilde{T}}(\theta)$, где је $b_{\widetilde{T}}(\theta) = \sqrt{b_T(\theta)}$. Дакле, испуњена су сва три услова теореме 4.3, одакле следи да је $c^{(a)}(\theta) = \frac{2b^2_{\widetilde{T}}(\theta)}{2m(2m-1)\upsilon_1}$.

С обзиром да $\overline{X}_n$ конвергира скоро сигурно ка очекиваној вредности $\mu(\theta)$, користећи закон великих бројева за $V-$статистике са оцењеним параметром (теорема 2.7), $T_n(\widehat{\mu}_n)$ конвергира ка

$$b_T(\theta) = E_\theta(\Phi(\mathbf{X};\mu(\theta))).$$

Пошто се оцена параметра расподеле јавља у статистици како би статистика била слободна од параметра при нултој хипотези, можемо претпоставити да је $\mu(0)=1$. Ако се узме први извод функције $b_T(\theta)$ по θ и затим рестрихује на $\theta=0$ добија се

$$b'_T(0) = \int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},\mu(\theta))|_{\theta=0} d\boldsymbol{G}(\boldsymbol{x};0) + \int\limits_{\mathbb{R}^{2m}} \Phi(\boldsymbol{x},\mu(\theta))|_{\theta=0}\frac{\partial}{\partial\theta} d\boldsymbol{G}(\boldsymbol{x};\theta)|_{\theta=0} = 0,$$

при чему је први члан једнак нули због начина на који се формирају тест статистике, а други због дегенерисаности језгра. Пошто је овај члан једнак нули, потребно је одредити други извод функције. Тада

$$\begin{aligned} b''_T(\theta) = &\int\limits_{\mathbb{R}^{2m}} \frac{\partial^2}{\partial\theta^2}\Phi(\boldsymbol{x},\mu(\theta))\left(\frac{\partial\mu(\theta)}{\partial\theta}\right)^2 d\boldsymbol{G}(\boldsymbol{x},\theta) + 2\int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},\mu(\theta))\frac{\partial\mu(\theta)}{\partial\theta}\frac{\partial}{\partial\theta} d\boldsymbol{G}(\boldsymbol{x},\theta) \\ &+ \int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},\mu(\theta))\frac{\partial^2\mu(\theta)}{\partial\theta^2} d\boldsymbol{G}(\boldsymbol{x},\theta) + \int\limits_{\mathbb{R}^{2m}} \Phi(\boldsymbol{x},\mu(\theta))\frac{\partial^2}{\partial\theta^2} d\boldsymbol{G}(\boldsymbol{x},\theta). \end{aligned}$$

Када се рестрихује претходни израз на $\theta = 0$, трећи члан је једнак 0 из истог разлога као и код одређивања првог извода, а користећи слабу дегенерисаност језгра добија се

$$
\begin{aligned}
b''_T(0) &= \int\limits_{\mathbb{R}^{2m}} \frac{\partial^2}{\partial\theta^2}\Phi(\boldsymbol{x},1)\left(\frac{\partial\mu(0)}{\partial\theta}\right)^2 d\boldsymbol{G}(\boldsymbol{x},0) + 2\int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},1)\frac{\partial\mu(0)}{\partial\theta}\frac{\partial}{\partial\theta}d\boldsymbol{G}(\boldsymbol{x},0)\\
&+ \int\limits_{\mathbb{R}^{2m}} \Phi(\boldsymbol{x},1)\frac{\partial^2}{\partial\theta^2}d\boldsymbol{G}(\boldsymbol{x},0)\\
&= \left(\frac{\partial\mu(0)}{\partial\theta}\right)^2 \int\limits_{\mathbb{R}^{2m}} \frac{\partial^2}{\partial\theta^2}\Phi(\boldsymbol{x},1)d\boldsymbol{G}(\boldsymbol{x},0) + 4m\frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},1)\frac{\partial dG(x_1;0)}{\partial\theta}\prod_{i=2}^{2m} dG(x_i;0)\\
&+2m\int\limits_{\mathbb{R}^{2m}} \Phi(\boldsymbol{x},1)\left(\frac{\partial^2 dG(x_1;0)}{\partial\theta^2}\prod_{i=2}^{2m} dG(x_i;0)+(2m-1)\frac{\partial dG(x_1;0)}{\partial\theta}\frac{\partial dG(x_2;0)}{\partial\theta}\prod_{i=3}^{2m} dG(x_i;0)\right)\\
&= \left(\frac{\partial\mu(0)}{\partial\theta}\right)^2 \int\limits_{\mathbb{R}^{2m}} \frac{\partial^2}{\partial\theta^2}\Phi(\boldsymbol{x},1)d\boldsymbol{G}(\boldsymbol{x},0) + 4m\frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^{2m}} \frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},1)\frac{\partial dG(x_1;0)}{\partial\theta}\prod_{i=2}^{2m} dG(x_i;0)\\
&+ 2m(2m-1)\int\limits_{\mathbb{R}^2} \varphi_2(x_1,x_2;,1)\frac{\partial dG(x_1;0)}{\partial\theta}\frac{\partial dG(x_2;0)}{\partial\theta}.
\end{aligned}
$$

Развојем $b_T(\theta)$ у Маклоренов ред $b_T(\theta) = b_T(0) + b'_T(0)\theta + \frac{1}{2}b''_T(0)\theta^2 + \cdots$, и заменом претходно добијеног следи тврђење. ■

У ситуацијама у којима параметар нема утицаја на граничну расподелу статистике, задржава се само први члан у изразу за приближни нагиб док су преостала два члана једнака нули. Заправо, с обзиром да су у том случају оператори (2.13) и (2.25) једнаки, приближни Бахадуров нагиб тих тестова има исти облик као тачан Бахадуров нагиб у теореми 4.6. Овај случај је показан у раду аутора *Cuparić et al. (2019b)*.

Без обзира да ли статистика зависи или не зависи од оцењеног параметра, за добијање Бахадуровог нагиба неопходно је одредити сопствене вредности оператора, тачније највећу од њих. У општем случају то је врло захтевно, или чак немогуће одредити. Да би се превазишао тај проблем може се користити апроксимација предложена у *Božin et al. (2018)*. Посматрајмо оператор

$$
Aq(x) = \int\limits_{-\infty}^{+\infty} \varphi_2(x,y)q(y)dF(y) \tag{4.12}
$$

чију највећу сопствену вредност треба одредити за неку произвољну расподелу F. Размотримо ,,симетризовани" оператор

$$
Sq(x) = \sqrt{F'(x)}A\left(\frac{q(x)}{F'(x)}\right) = \int\limits_{-\infty}^{+\infty} \varphi_2(x,y)q(y)\sqrt{F'(x)}\sqrt{F'(y)}dy
$$

који има исти спектар као оператор A. Претпоставимо да је $\inf\limits_x\{F(x) = 1\} = B < \infty$.

Уколико ово није испуњено, односно $\inf_x\{F(x)=1\}=\infty$, може се посматрати засечени оператор $\overline{S}$ за функције дефинисане на $[-B,B]$, тако да је $F(B)>1-\varepsilon$ за неко $\varepsilon>0$, дефинисан са

$$\overline{S}q(x)=\int\limits_{-\infty}^{+\infty}\varphi_2(x,y)q(y)\sqrt{F'(x)}\sqrt{F'(y)}\mathrm{I}(y\leqslant B)F(B)^{-1}dy.$$

Дакле, за довољно велико B, $\overline{S}$ и S се разликују на скупу мере нула. Низ симетричних линеарних оператора дефинисаних са $(2m+1)\times(2m+1)$ матрицом $\mathcal{M}^{(m)}=||m_{i,j}^{(m)}||$, $0\leqslant i\leqslant m, 0\leqslant j\leqslant m$, где је

$$m_{i,j}^{(m)}=\varphi_2\Big(\frac{Bi}{m},\frac{Bj}{m}\Big)\sqrt{F\Big(\frac{B(i+1)}{m}\Big)-F\Big(\frac{Bi}{m}\Big)}\sqrt{F\Big(\frac{B(j+1)}{m}\Big)-F\Big(\frac{Bj}{m}\Big)}\frac{1}{F(B)},\qquad(4.13)$$

конвергира у норми ка $\overline{S}$, кад $m\to\infty$.

Оператори $\overline{S}$ и $\mathcal{M}^{(m)}$ су симетрични и самоадјунговани, па норма њихове разлике тежи нули кад $m\to\infty$. Тада спектри ова два оператора су на растојању које тежи нули (за више детаља видети *Božin et al. (2018)*). Према томе, низ највећих сопствених вредности оператора $\mathcal{M}^{(m)}$, $\upsilon_1^{(m)}$, мора конвергирати ка највећој сопственој вредности од $\overline{S}$. Када $B\to\infty$, сопствена вредност $\upsilon_1(B)$ тежи ка υ_1. Дакле, за довољно велико B, $\upsilon_1(B)$ и υ_1 ће се подударати са довољном тачношћу. Користећи матричне операторе и програм *Wolfram Mathematica 11.0*, могу се добити одговарајуће сопствене вредности. Аналогно важи и ако уместо оператора A (4.12), треба одредити највећу сопствену вредност оператора $A^\star$ (2.25).

Размотримо Бахадурове нагибе тестова L^2-типа приказаних у поглављима 3.2 и 3.3. У оба случаја, закључено је да се тестови могу написати као слабо дегенерисане $V-$статистике, што значи да су прве пројекције одговарајућих језгара једнаке нули док су друге пројекције језгара различите од нуле. Претпоставимо да густина $g(x;\boldsymbol{\theta})$ из класе $\mathcal{G}_{\mathcal{E}}$ задовољава услове (4.7). У овим случајевима можемо применити претходне теореме.

Тестови W_n, уведени у поглављу 3.2, не садрже оцену параметра па се у том случају примењује теорема 4.6. Тачан Бахадуров нагиб је

$$c_{W^{\mathcal{I}}}(\boldsymbol{\theta})\sim\frac{1}{\upsilon_1^{\mathcal{I}}}\int\limits_0^\infty\int\limits_0^\infty\varphi_2^{\mathcal{I}}(x_1,x_2;\lambda)g'_\vartheta(x_1;(0,\lambda))g'_\vartheta(x_2;(0,\lambda))dx_1dx_2\cdot\vartheta^2+o(\vartheta^2),\vartheta\to0,$$

где су $\varphi_2^{\mathcal{I}}(x_1,x_2)$, $\mathcal{I}\in\{\mathcal{P},\mathcal{D}\}$, друге пројекције дате у (3.8) и (3.9), а $\upsilon_1^{\mathcal{I}}$, $\mathcal{I}\in\{\mathcal{P},\mathcal{D}\}$, највеће сопствене вредности одговарајућих оператора. Да би се одредиле ове сопствене вредности користи се претходно описана процедура за апроксимацију. Вредности добијене за $B=10$ и различите вредности m приказане су у табели 4.1. Уз коришћење свега наведеног и израза (4.8), за одговарајуће алтернативе може се одредити тачан Бахадуров нагиб ових тестова. Добијени резултати, односно локалне апсолутне Бахадурове ефикасности (о чему ће касније бити речи), за конкретно изабране алтернативе су приказани на сликама 4.1-4.4 и у табели 4.6.

Табела 4.1: Највеће сопствене вредности матричног оператора $\mathcal{M}^{(m)}$ за $W_n^{\mathcal{P}}$ и $W_n^{\mathcal{D}}$

m	1000	2000	3000	4000	5000	6000
$\upsilon_1^{\mathcal{P}}$	$6.11 \cdot 10^{-3}$	$6.05 \cdot 10^{-3}$	$6.03 \cdot 10^{-3}$	$6.02 \cdot 10^{-3}$	$6.01 \cdot 10^{-3}$	$6.01 \cdot 10^{-3}$
$\upsilon_1^{\mathcal{D}}$	$4.37 \cdot 10^{-3}$	$4.42 \cdot 10^{-3}$	$4.44 \cdot 10^{-3}$	$4.45 \cdot 10^{-3}$	$4.46 \cdot 10^{-3}$	$4.46 \cdot 10^{-3}$

Са друге стране, тестови приказани у поглављу 3.3 у свом запису садрже оцењени параметар, међутим, као што је показано у доказу теореме 3.1, тај параметар нема утицаја на граничну расподелу тест статистике. Дакле, приближни Бахадурови нагиби у том случају имају облик

$$c_{M^{\mathcal{I}}}^{(a)}(\boldsymbol{\theta}) \sim \frac{1}{\upsilon_1^{\mathcal{I}}} \int_0^{\infty}\int_0^{\infty} \varphi_2^{\mathcal{I}}(x_1, x_2; \lambda) g'_{\vartheta}(x_1; (0,\lambda)) g'_{\vartheta}(x_2; (0,\lambda)) dx_1 dx_2 \cdot \vartheta^2, \ \vartheta \to 0,$$

где су $\varphi_2^{\mathcal{I}}(x_1, x_2)$, $\mathcal{I} \in \{\mathcal{P}, \mathcal{D}\}$, друге пројекције језгара дате у (3.19) и (3.20). Дакле за одређивање приближног Бахадуровог нагиба потребно је још одредити највећу сопствену вредност. Као и код претходних тестова, користи се наведена апроксимација. Вредности добијене за $B = 10$, различите вредности m и различите вредности параметра a су приказане у табели 4.2. Уз коришћење свега наведеног, за одговарајуће алтернативе може се одредити приближни Бахадуров нагиб ових тестова. Добијени резултати, односно локалне апсолутне Бахадурове ефикасности, за конкретно изабране алтернативе су приказани на сликама 4.1-4.4 и у табели 4.6.

Табела 4.2: Највеће сопствене вредности матричног оператора $\mathcal{M}^{(m)}$ за $M_n^{\mathcal{P}}$ и $M_n^{\mathcal{D}}$

	m	1000	2000	3000	4000	5000	6000
$\upsilon_1^{\mathcal{P}}$	$a = 0.5$	$1.36 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$	$1.32 \cdot 10^{-2}$	$1.32 \cdot 10^{-2}$	$1.32 \cdot 10^{-2}$
	$a = 1$	$5.44 \cdot 10^{-3}$	$5.36 \cdot 10^{-3}$	$5.33 \cdot 10^{-3}$	$5.31 \cdot 10^{-3}$	$5.30 \cdot 10^{-3}$	$5.30 \cdot 10^{-3}$
	$a = 2$	$1.77 \cdot 10^{-3}$	$1.75 \cdot 10^{-3}$	$1.74 \cdot 10^{-3}$	$1.73 \cdot 10^{-3}$	$1.73 \cdot 10^{-3}$	$1.73 \cdot 10^{-3}$
	$a = 5$	$2.87 \cdot 10^{-4}$	$2.84 \cdot 10^{-4}$	$2.83 \cdot 10^{-4}$	$2.82 \cdot 10^{-4}$	$2.82 \cdot 10^{-4}$	$2.82 \cdot 10^{-4}$
	$a = 10$	$5.70 \cdot 10^{-5}$	$5.65 \cdot 10^{-5}$	$5.63 \cdot 10^{-5}$	$5.62 \cdot 10^{-5}$	$5.62 \cdot 10^{-5}$	$5.62 \cdot 10^{-5}$
$\upsilon_1^{\mathcal{D}}$	$a = 0.5$	$6.48 \cdot 10^{-3}$	$6.20 \cdot 10^{-3}$	$6.12 \cdot 10^{-3}$	$6.07 \cdot 10^{-3}$	$6.05 \cdot 10^{-3}$	$6.05 \cdot 10^{-3}$
	$a = 1$	$2.88 \cdot 10^{-3}$	$2.80 \cdot 10^{-3}$	$2.77 \cdot 10^{-3}$	$2.76 \cdot 10^{-3}$	$2.75 \cdot 10^{-3}$	$2.75 \cdot 10^{-3}$
	$a = 2$	$1.06 \cdot 10^{-3}$	$1.04 \cdot 10^{-3}$	$1.03 \cdot 10^{-3}$	$1.03 \cdot 10^{-3}$	$1.02 \cdot 10^{-3}$	$1.02 \cdot 10^{-3}$
	$a = 5$	$2.00 \cdot 10^{-4}$	$2.00 \cdot 10^{-4}$	$1.99 \cdot 10^{-4}$	$1.99 \cdot 10^{-4}$	$1.98 \cdot 10^{-4}$	$1.98 \cdot 10^{-4}$
	$a = 10$	$4.52 \cdot 10^{-5}$	$4.47 \cdot 10^{-5}$	$4.45 \cdot 10^{-5}$	$4.44 \cdot 10^{-5}$	$4.44 \cdot 10^{-5}$	$4.44 \cdot 10^{-5}$

У поглављу 4.6 биће приказани неки тестови који се могу представити као слабо дегенерисане $V-$статистике и код којих оцењени параметар има утицаја на граничну расподелу. Њихови приближни Бахадурови нагиби се добијају коришћењем теореме 4.7, чије ће вредности, тачније локалне апсолутне Бахадурове ефикасности, бити упоређене са одговарајућим вредностима добијеним за претходно приказане тестове.

4.4 Бахадурова асимптотска ефикасност тестова сагласности L^∞−типа

Тест статистике тестова сагласности L^∞−типа могу се представити као супремум апсолутне вредности V−емпиријског процеса, односно

$$T_n = \sup_{t>0} |V_n(t)|, \tag{4.14}$$

што је показано у поглављу 3.4. За фиксирано t, V_n је недегенерисана V−статистика, па се на њу могу применити резултати поглавља 4.2 који се односе на такве статистике. Због тога прво размотримо Бахадуров нагиб недегенерисаних V−статистика.

Претпоставимо да је фамилија алтернативних расподела $\mathcal{G}$ дефинисана као у поглављу 4.2. Такође, претпостављамо да су испуњени услови који су потребни да би важило (4.5). Следећа теорема даје тачан Бахадуров нагиб недегенерисане V−статистике. Њен доказ следи из Маклореновог развоја и (4.5) и може се наћи у *Nikitin and Peaucelle (2004)*.

Теорема 4.8 *Нека је $\{V_n\}$ низ недегенерисаних V−статистика са ограниченим језгрима и нека је $b_\Phi(\theta) > 0$. Нека је $g(x;\theta)$ алтернатива из $\mathcal{G}$ и нека постоји $\delta > 0$ тако да за $0 \leqslant \theta \leqslant \delta$ важи*

$$\int_{-\infty}^{+\infty} |g'_\theta(x,\theta)|dx < \infty, \quad \int_{-\infty}^{+\infty} |g''_{\theta\theta}(x,\theta)|dx < \infty, \tag{4.15}$$

где су g'_θ и $g''_{\theta\theta}$ први и други извод функције g по θ. Тада, кад $\theta \to 0$,

$$c_V(\theta) \sim \frac{1}{D(\varphi_1(X))}\left(\int_{-\infty}^{+\infty} \varphi_1(x) g'_\theta(x;0)dx\right)^2 \theta^2, \tag{4.16}$$

где је $\varphi_1(x) = E(\Phi(X_1, ..., X_m)|X_1 = x)$ прва пројекција језгра Φ.

Следећи став говори о тачном Бахадуровом нагибу статистика T_n дефинисаних у (4.14) и следи из претходне теореме уз коришћење Гливенко-Кантелијеве теореме за V−статистике (за статистике за које ће у наставку бити примењен наведени став, одговарајућа Гливенко-Кантелијева теорема се може наћи у *Helmers et al. (1988)*).

Став 4.1 *Нека је $T_n = \sup_{t>0} |V_n(t)|$, где је V_n недегенерисана V−статистика. Нека су испуњени услови теореме 4.8. Тада, кад $\theta \to 0$,*

$$c_T(\theta) \sim \frac{1}{\sup_{t>0} K(t,t)}\left(\sup_{t>0}\left| m \int_{-\infty}^{+\infty} \varphi_1(x;t) g'_\theta(x;0)dx\right|\right)^2 \theta^2, \tag{4.17}$$

где је $\varphi_1(x,t) = E(\Phi(X_1, ..., X_m;t)|X_1 = x)$ прва пројекција језгра Φ и $K(s,t)$ коваријациона функција граничног Гаусовог процеса.

Међутим, тест статистике овог типа могу садржати и оцењен параметар фамилије расподела. Због тога се претходни став не може директно применити на њих. Следеће теореме дају приближан Бахадуров нагиб недегенерисаних V−статистика са оцењеним параметром, као и статистика облика

$$T_n(\widehat{\mu}_n) = \sup_{t>0} |V_n(\widehat{\mu}_n; t)|. \tag{4.18}$$

Због компактности записа, означимо са $\boldsymbol{x} = (x_1, ..., x_m)$ и $\boldsymbol{G}(\boldsymbol{x}; \theta) = \prod_{i=1}^m G(x_i; \theta)$.

Теорема 4.9 *Нека је $\{V_n(\widehat{\mu}_n)\}$ низ недегенерисаних V−статистика са оцењеним параметром са ограниченим језгрима $\Phi(\boldsymbol{x}; \widehat{\mu}_n)$ реда m, таквим да постоји $\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x}; \mu(\theta))$ и нека је $b_\Phi(\theta) > 0$. Нека је $g(x;\theta)$ алтернатива из $\mathcal{G}$ и нека постоји $\delta > 0$ тако да за $0 \leqslant \theta \leqslant \delta$ важи (4.15). Тада, кад $\theta \to 0$, следи*

$$c_V^{(a)}(\theta) \sim \frac{1}{m^2 D(\varphi_1(X))}\left(\frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x}; 1) d\boldsymbol{G}(\boldsymbol{x}; 1) + m\int\limits_{\mathbb{R}}\varphi_1(x;t)\frac{\partial g(x;0)}{\partial\theta}dx\right)^2\theta^2, \tag{4.19}$$

где је $\varphi_1(x) = E(\Phi(X_1, ..., X_m)|X_1 = x)$ прва пројекција језгра Φ.

Доказ: Према теореми 2.8 статистика $\sqrt{n}V_n(\widehat{\mu}_n)$ при нултој хипотези конвергира у расподели ка случајној величини са нормалном расподелом са очекивањем 0 и дисперзијом $m^2 D(\varphi_1(X))$. Нека је F_0 функција расподеле наведене граничне случајне величине, а f_0 одговарајућа густина. Тада важи (за више детаља видети *Feller (1968)*, страна 175.)

$$1 - F_0(t) \sim \frac{f_0(t)}{t},$$

када $t \to \infty$. За логаритам репа наведене расподеле важи

$$\ln(1 - F_0(t)) = -\frac{t^2}{2m^2 D(\varphi_1(X))} + o_p(t^2),\ t \to \infty.$$

Дакле, испуњен је услов 2 теореме 4.3 за $a = \frac{1}{m^2 D(\varphi_1(X))}$. Такође, из закона великих бројева за V−статистике са оцењеним параметром (теорема 2.7) следи да $V_n(\widehat{\mu}_n) \xrightarrow{P} b_\Phi(\theta)$, где је

$$b_\Phi(\theta) = E_\theta(\Phi(\mathbf{X}; \mu(\theta))).$$

Дакле, из теореме 4.3 закључујемо да је $c^{(a)}(\theta) = \frac{b_\Phi^2(\theta)}{m^2 D(\varphi_1(X))}$.

Остаје још да се одреди вредност $b_T(\theta)$. Без губитка општости можемо претпоставити да је $\mu(0) = 1$, с обзиром да су тестови са којима се ради слободни од параметра расподеле при нултој хипотези. Ако се узме први извод функције $b_V(\theta)$ по θ добија се

$$b'_\Phi(\theta) = \int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x}; \mu(\theta))\frac{\partial\mu(\theta)}{\partial\theta}d\boldsymbol{G}(\boldsymbol{x}; 0) + \int\limits_{\mathbb{R}^m}\Phi(\boldsymbol{x}; \mu(\theta))\frac{\partial}{\partial\theta}d\boldsymbol{G}(\boldsymbol{x}; \theta)$$

$$
\begin{aligned}
&= \frac{\partial\mu(\theta)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x};\mu(\theta))d\boldsymbol{G}(\boldsymbol{x};\mu(\theta)) + \int\limits_{\mathbb{R}^m}\Phi(\boldsymbol{x};\mu(\theta))\frac{\partial}{\partial\theta}\prod_{i=1}^m g(x_i;\theta)dx_i\\
&= \frac{\partial\mu(\theta)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x};\mu(\theta))d\boldsymbol{G}(\boldsymbol{x};\mu(\theta)) + m\int\limits_{\mathbb{R}^m}\Phi(\boldsymbol{x};\mu(\theta))\frac{\partial g(x;\theta)}{\partial\theta}dx_1\prod_{i=2}^m g(x_i;\theta)dx_i\\
&= \frac{\partial\mu(\theta)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x};\mu(\theta))d\boldsymbol{G}(\boldsymbol{x};\mu(\theta)) + m\int\limits_{\mathbb{R}}\varphi_1(x;\mu(\theta))\frac{\partial g(x;\theta)}{\partial\theta}dx.
\end{aligned}
$$

Рестриховањем на $\theta = 0$ добија се

$$
b'_\Phi(0) = \frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x};1)d\boldsymbol{G}(\boldsymbol{x};1) + m\int\limits_{\mathbb{R}}\varphi_1(x;t)\frac{\partial g(x;0)}{\partial\theta}dx.
$$

Развојем $b_\Phi(\theta)$ у Маклоренов ред $b_\Phi(\theta) = b_\Phi(0) + b'_\Phi(0)\theta + \cdots$, и заменом претходно добијеног следи тврђење. ■

Теорема 4.10 (Цупарић, Милошевић, Обрадовић 2019) *Нека је $T_n(\widehat{\mu}_n) = \sup\limits_{t>0}|V_n(\widehat{\mu}_n;t)|$, где је, за фиксирано t, $\{V_n(\widehat{\mu}_n;t)\}$ низ недегенерисаних $V-$статистика са оцењеним параметром са ограниченим језгрима реда m, таквим да постоји $\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},t;\mu(\theta))$ и нека је $b_\Phi(\theta) > 0$. Нека је $g(x;\theta)$ алтернатива из $\mathcal{G}$ и нека постоји $\delta > 0$ тако да за $0 \leqslant \theta \leqslant \delta$ важи* (4.15)*. Тада, кад $\theta \to 0$,*

$$
c_T(\theta) \sim \frac{1}{\sup\limits_{t>0}K(t,t)}\sup_{t>0}\left(\frac{\partial\mu(0)}{\partial\theta}\int\limits_{\mathbb{R}^m}\frac{\partial}{\partial\theta}\Phi(\boldsymbol{x},t;1)d\boldsymbol{G}(\boldsymbol{x};1)+m\int\limits_{\mathbb{R}}\varphi_1(x;t)\frac{\partial g(x;0)}{\partial\theta}dx\right)^2\theta^2, \quad (4.20)
$$

где је $\varphi_1(\boldsymbol{x},t) = E(\Phi(X_1,...,X_m;t)|X_1 = x)$ прва пројекција језгра Φ и $K(s,t)$ коваријационa функција граничног Гаусовог процеса.

Пре доказа теореме наводимо теорему која говори о понашању логаритма репа супремума Гаусовог процеса (за више детаља видети *Nikitin (1995)* и *Ledoux and Talagrand (2013)*).

Теорема 4.11 *Нека је $X(t)$ Гаусов процес дефинисан на $[0,\infty)$, са $E(X(t)) = 0$ и $K(t,s) = E(X(t)X(s))$, $t,s > 0$, такав да*

$$
P\{\sup_{t>0}|X(t)| < \infty\} = 1.
$$

Означимо са $\sigma^2 = \sup\limits_{t>0}K(t,t)$. Тада

$$
\lim_{z\to\infty}\ln P\{\sup_{t>0}|X(t)| \geqslant z\} = -\frac{1}{2\sigma^2}.
$$

Поред тога, за доказ теореме 4.10, неопходно је и тврђење аналогно Гливенко-Кантелијевој теореми, чији исказ и доказ је наведен у наставку.

Теорема 4.12 *Нека је $\mathcal{L}_n(t)$ $V-$емпиријска Лапласова трансформација дефинисана у* (2.38) *и $\mathcal{L}(t)$ теоријска Лапласова трансформација дефинисана у* (2.37)*. Тада важи*

$$\sup_{t>0}|\mathcal{L}_n(t)-\mathcal{L}(t)| \xrightarrow{P} 0,\ n\to\infty.$$

Доказ: Јасно је да је $\mathcal{L}(t)$ непрекидна монотоно опадајућа функција таква да је $\mathcal{L}(0)=1$ и $\mathcal{L}(+\infty)=0$. Нека је $\varepsilon>0$ фиксирано. За неко фиксирано $N=N(\varepsilon)<+\infty$, постоји подела тако да $0=t_0<t_1<\cdots<t_N=+\infty$ и $\mathcal{L}(t_{j-1})-\mathcal{L}(t_j)\leqslant\frac{\varepsilon}{2},\ j\in\{1,...,N\}$.

Нека је $t>0$ произвољна тачка таква да $t\in[t_{j-1},t_j]$. Тада

$$\mathcal{L}_n(t)-\mathcal{L}(t)\leqslant\mathcal{L}_n(t_{j-1})-\mathcal{L}(t_j)\leqslant|\mathcal{L}_n(t_{j-1})-\mathcal{L}(t_{j-1})|+\frac{\varepsilon}{2}\leqslant\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|+\frac{\varepsilon}{2}.$$

Са друге стране, на сличан начин важи

$$\mathcal{L}_n(t)-\mathcal{L}(t)\geqslant\mathcal{L}_n(t_j)-\mathcal{L}(t_{j-1})\geqslant\mathcal{L}_n(t_j)-\mathcal{L}(t_j)-\frac{\varepsilon}{2}\geqslant-\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|-\frac{\varepsilon}{2}.$$

Комбинујући претходна два израза, следи

$$|\mathcal{L}_n(t)-\mathcal{L}(t)|\leqslant\frac{\varepsilon}{2}+\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|.$$

Односно,

$$\sup_{t>0}|\mathcal{L}_n(t)-\mathcal{L}(t)|\leqslant\frac{\varepsilon}{2}+\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|.$$

Из закона великих бројева за $V-$статистике 2.5, следи да $\mathcal{L}_n(t_j)\xrightarrow{P}\mathcal{L}(t_j)$. Према томе, за фиксирано $j\in\{0,...,m\}$ за неко $\delta>0$ постоји n_j тако да за све $n>n_j$

$$P\left\{|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|\geqslant\frac{\varepsilon}{2}\right\}\leqslant\delta.$$

Одатле следи да

$$P\left\{\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|\geqslant\frac{\varepsilon}{2}\right\}=P\left\{\exists j\in\{0,...,N\}:\ |\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|\geqslant\frac{\varepsilon}{2}\right\}\leqslant(N+1)\delta,$$

за све $n\geqslant\max\limits_{j\in\{0,...,N\}}n_j$. Према томе, за неко $\delta'>0$ постоји $n'\in\mathbb{N}$ тако да

$$P\left\{\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t_j)-\mathcal{L}(t_j)|\geqslant\frac{\varepsilon}{2}\right\}\leqslant\delta',$$

за све $n\geqslant n'$. Комбинујући добијене резултате, следи да

$$\begin{aligned}P\left\{\sup_{t>0}|\mathcal{L}_n(t)-\mathcal{L}(t)|>\frac{\varepsilon}{2}\right\}&\leqslant P\left\{\frac{\varepsilon}{2}+\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t)-\mathcal{L}(t)|>\varepsilon\right\}\\&\leqslant P\left\{\max_{j\in\{0,...,N\}}|\mathcal{L}_n(t)-\mathcal{L}(t)|>\frac{\varepsilon}{2}\right\}\to 0,\ n\to\infty,\end{aligned}$$

чиме је доказано тврђење. ■

Претходна теорема и њен доказ ће важити и у случају $V-$емпиријских Лапласових трансформација са оцењеним параметром уз једину измену да се уместо закона великих бројева за $V-$статистике 2.5 користи закон великих бројева за $V-$статистике са оцењеним параметром (Хефдингова репрезентација (2.5) и теорема 2.7).

Доказ теореме 4.10: Из последице 3.1 следи да статистика $\sqrt{n}T_n(\widehat{\mu}_n)$ конвергира у расподели ка $\sup\limits_{t>0}|\eta(t)|$, где је $\eta(t)$ центрирани Гаусов процес, а из теореме 4.11 следи да је

$$\ln P\{\sup_{t>0}|V_n(\widehat{\mu}_n;t)|>z\}=-\frac{z^2}{2\sup\limits_{t>0}K(t,t)}+o_p(z^2),\ z\to\infty. \tag{4.21}$$

Односно, испуњен је услов 2. теореме 4.3 за $a=\frac{1}{\sup\limits_{t>0}K(t,t)}$. Из теореме 4.12 следи

$$\begin{aligned}\sup_{t>0}|V_n(\widehat{\mu}_n;t)-b_\Phi(\theta)|&=\sup_{t>0}|(\mathcal{L}_n^{(1)}(t)-\mathcal{L}_n^{(2)}(t))e^{-at}-E_\theta(\Phi(\mathbf{X};\mu(\theta))|\\&=\sup_{t>0}|(\mathcal{L}_n^{(1)}(t)-\mathcal{L}^{(1)}(t))e^{-at}-(\mathcal{L}_n^{(2)}(t)-\mathcal{L}^{(2)}(t))e^{-at}|\\&\leqslant\sup_{t>0}|(\mathcal{L}_n^{(1)}(t)-\mathcal{L}^{(1)}(t))e^{-at}|+\sup_{t>0}|(\mathcal{L}_n^{(2)}(t)-\mathcal{L}^{(2)}(t))e^{-at}|\\&\overset{P}{\to}0,\ n\to\infty.\end{aligned}$$

Према томе $T_n(\widehat{\mu}_n)\overset{P}{\to}\sup\limits_{t>0}b_\Phi(\theta)$ при алтернативној расподели. Дакле, испуњена су сва три услова теореме 4.3, одакле следи да је $c^{(a)}(\theta)=\frac{b_\Phi^2(\theta)}{\sup\limits_{t>0}K(t,t)}$.

Одређивање $b_\Phi(\theta)$ следи на исти начин као у доказу теореме 4.9, па одатле следи тврђење. ■

Тестови приказани у поглављу 3.4 у свом запису садрже оцењени параметар, међутим, као што је напоменуто у доказу теореме 3.3, тај параметар нема утицаја на граничну расподелу тест статистике. Према томе, приближни Бахадуров нагиб ових тестова је специјалан случај претходне теореме. Претпоставимо да густина $g(x;\boldsymbol{\theta})$ из класе $\mathcal{G}_\mathcal{E}$ задовољава услове (4.15). Тада је приближни Бахадуров нагиб

$$c_{L^\mathcal{I}}(\boldsymbol{\theta})\sim\frac{1}{\sup\limits_{t>0}K^\mathcal{I}(t,t)}\left(\sup_{t>0}\left|m\int\limits_{-\infty}^{+\infty}\varphi_1^\mathcal{I}(x;t)g'_\vartheta(x;(0,\lambda))dx\right|\right)^2\vartheta^2,\ \vartheta\to0,$$

где су

$$\begin{aligned}\varphi_1^\mathcal{P}(x;t)&=\frac{1}{2}e^{-at}\left(\frac{1-2e^{-x}}{t+1}+e^{-tx}-\frac{2\left(e^{-x}-e^{-tx}\right)}{t-1}\right),\\\varphi_1^\mathcal{D}(x;t)&=\frac{1}{2}e^{-at}\left(\frac{1}{t+1}+e^{-tx}-\frac{2(2te^{-(2t+1)x}+1)}{2t+1}\right),\end{aligned}$$

а $K^\mathcal{I}(t,t)$, $\mathcal{I}\in\{\mathcal{P},\mathcal{D}\}$, коваријације приказане у (3.26) и (3.27). Локалне апсолутне Бахадурове ефикасности су приказани на сликама 4.1-4.4 и у табели 4.6.

4.5 Локална апсолутна Бахадурова ефикасност

Још један значајан резултат Бахадурове теорије јесте постојање горње границе за тачан Бахадуров нагиб. Претпоставимо да је за свако θ, расподела случајних величина $X_1, .., X_n$ одређена са $g(x;\theta)$. За неке θ и θ_0 из Θ Кулбак-Лајблеров информациони број K је дефинисан са

$$K(\theta, \theta_0) = E_\theta\Big(\ln \Big(\frac{g(X;\theta)}{g(X;\theta_0)} \Big) \Big).$$

При чему је $0 \leqslant K(\theta, \theta_0) \leqslant \infty$, и $K(\theta, \theta_0) = 0$ ако и само ако су у питању исте расподеле. За свако $\theta \in \Theta$, нека је

$$K(\theta, \Theta_0) = \inf_{\theta_0 \in \Theta_0} K(\theta, \theta_0).$$

Тада је $K(\theta, \Theta_0)$ добро дефинисано на Θ, $0 \leqslant K(\theta, \Theta_0) \leqslant \infty$ и $K(\theta, \Theta_0) = 0$ за $\theta \in \Theta_0$. Следеће тврђење даје неједнакост познату као Бахадур-Рагавачаријева (за више детаља видети *Bahadur (1967)*).

Став 4.2 *Ако је c_T тачан нагиб низа $\{T_n\}$ онда $c_T(\theta) \leqslant 2K(\theta, \Theta_0)$ за свако $\theta \in \Theta_1$.*

Ако у претходном ставу важи да је $c_T(\theta) = 2K(\theta, \Theta_0)$ за свако $\theta \in \Theta_1$, онда је низ $\{T_n\}$ асимптотски оптималан у Бахадуровом смислу. Класа статистика за које то важи има мало. При одређеним условима, то ће бити испуњено за статистике теста количника веродостојности.

Нека је Λ_n тест статистика теста количника веродостојности, односно

$$\Lambda_n = \frac{\sup\limits_{\theta \in \Theta_0} \prod\limits_{i=1}^{n} g(x_i;\theta)}{\sup\limits_{\theta \in \Theta} \prod\limits_{i=1}^{n} g(x_i;\theta)} \tag{4.22}$$

и нека је $\widehat{T_n}$ нека строго опадајућа функција од Λ_n, на пример $\widehat{T_n} = -2\ln \Lambda_n$. Доказ следећег става се може наћи у *Bahadur (1967)*.

Став 4.3 *Тачан нагиб од $\{\widehat{T}_n\}$ постоји и једнак је $2K(\theta, \Theta_0)$ за свако $\theta \in \Theta_1$.*

Неке статистике различите од $\widehat{T}_n$ које су такође оптималне, могу се наћи у *Bahadur and Bickel (2009)*. Занимљиво је да се заправо испоставља да неки тестови који су асимптотски еквивалентни са тестом количника веродостојности (на пример, χ^2 тест за мултиномну расподелу) имају ефикасност мању од $2K(\theta, \Theta_0)$ за већину вредности $\theta \in \Theta_1$ (за више детаља видети *Abrahamson (1965)*).

Пошто је, према ставу 4.2, $2K(\theta, \Theta_0)$ горње ограничење за тачан Бахадуров нагиб, природно је да се у општем случају дефинише количник $\frac{c_T(\theta)}{2K(\theta,\Theta_0)}$ који се назива апсолутна Бахадурова ефикасност. Међутим, за θ различито од нуле, тешко је одредити наведени количник. Због тога се чешће одређује локална апсолутна Бахадурова ефикасност, која представља граничну вредност наведеног количника када $\theta \to 0$.

С обзиром да је за неке тестове могуће одредити само приближни Бахадуров нагиб, пожељно је да постоји горње ограничење и приближног Бахадуровог нагиба. Међутим, у општем случају не постоји аналогон Бахадур-Рагавачаријеве неједнакости у случају приближног Бахадуровог нагиба. Пошто су локално приближни и тачни нагиб често еквивалентни, у том случају би приближни Бахадуров нагиб теста количника веродостојности могао да се узме као горње ограничење приближног нагиба. Следећа теорема даје услове при којима су приближни и тачни нагиби тестова количника веродостојности једнаки. За више детаља видети *Rublík (1989)*.

Теорема 4.13 *Нека је тест статистика* $\widehat{T}_n = -2\ln\Lambda_n$, *где је* Λ_n *дефинисано у* (4.22), *и нека су испуњени следећи услови*

1. *важи* $\lim\limits_{n\to\infty}\frac{\widehat{T}_n(s)}{n} = 2K(\theta,\Theta_0)$ *скоро сигурно;*
2. *ако* γ *и* γ_n, $n = 1, 2, ...,$ *припадају* Θ, *онда* $\lim\limits_{n\to\infty}K(\gamma,\gamma_n) = 0$ *повлачи да* $\lim\limits_{n\to\infty}\gamma_n = \gamma$*;*
3. *за* $\theta\in\Theta_1$ *важи* $\lim\limits_{n\to\infty}\frac{2}{n}\ln L_n(s) = -c(\theta)$, *скоро сигурно, и* $c(\theta) = 2K(\theta,\Theta_0)$*;*
4. $\theta\in\Theta\backslash\overline{\Theta}_0$.

Тада

(I) *Ако за* $G(t) = \inf\limits_{\theta\in\Theta_0}\lim\limits_{n\to\infty}F_n(t;\theta)$ *важи*

$$\ln(1-G(t)) = -\frac{1}{2}t(1+o_p(1)),\ t\to\infty,$$

онда постоји приближни нагиб и важи $c^{(a)}(\theta) = c(\theta)$.

(II) *Ако је* $G(t)$ *коначна мешавина случајних величина са* χ^2 *расподелом, онда приближни нагиб постоји и важи* $c^{(a)}(\theta) = c(\theta)$.

Претпоставимо да је нулта расподела експоненцијална, а алтернативна расподела нека расподела са густином $g(x,\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\vartheta,\lambda)$, из фамилије $\mathcal{G}_{\mathcal{E}}$ дефинисане у поглављу 4.2 и нека важи (4.7). Показано је да у том случају двоструко Кулбак-Лајблерово растојање, за мало ϑ ($\vartheta\to 0$), се може изразити као

$$2K((\vartheta,\lambda),(0,\lambda)) = \left(\int\limits_0^\infty (g'_\vartheta(x;(0,\lambda)))^2\frac{e^{\lambda x}}{\lambda}dx - \lambda^2\left(\int\limits_0^\infty x g'_\vartheta(x,0)dx\right)^2\right)\vartheta^2 + o_p(\vartheta^2).$$

(за више детаља видети *Nikitin (1996)*).

Размотримо следеће четири класе алтернатива које припадају фамилији $\mathcal{G}$

- Вејбулова $W(\vartheta,\lambda)$ расподела са густином

$$g(x;\vartheta,\lambda) = e^{-(\lambda x)^{1+\vartheta}}(1+\vartheta)\lambda^{\vartheta+1}x^{\vartheta},\ x\geqslant 0, \vartheta > 0,\ \lambda > 0; \tag{4.23}$$

- гама $\Gamma(\vartheta, \lambda)$ расподела са густином

$$g(x;\vartheta,\lambda) = \frac{x^{\vartheta}\lambda^{\vartheta+1}e^{-\lambda x}}{\Gamma(\vartheta+1)},\ x \geqslant 0, \vartheta > 0,\ \lambda > 0; \tag{4.24}$$

- расподела линеарне стопе отказа $LFR(\theta)$ са густином

$$g(x;\vartheta,\lambda) = e^{-\lambda x - \vartheta\lambda^2\frac{x^2}{2}}(\lambda + \vartheta\lambda^2 x),\ x \geqslant 0, \vartheta > 0,\ \lambda > 0; \tag{4.25}$$

- мешавина експоненцијалних расподала $EMNW(\theta, \beta)$ (*Jevremovic (1991)*) са густином

$$g(x;\vartheta,\lambda) = (1+\vartheta)\lambda e^{-\lambda x} - \vartheta\beta\lambda e^{-\beta\lambda x},\ x \geqslant 0, \vartheta \in \left(0, \frac{1}{\beta-1}\right],\ \lambda > 0. \tag{4.26}$$

Вредности двоструког Кулбак-Лајблеровог растојања за ове класе алтернатива дате су у табели 4.3.

Табела 4.3: Двоструко Кулбак-Лајблерово растојање (коефицијент уз ϑ^2) за мало ϑ

Алт.	$W(\vartheta,\lambda)$	$\Gamma(\vartheta,\lambda)$	$LFR(\vartheta,\lambda)$	$EWNM(\vartheta,\lambda,3)$
$2KL((\vartheta,\lambda),(0,\lambda))$	0.64493	1.64493	1	0.35556

4.6 Поређење тестова сагласности

Поређење тестова сагласности у односу на њихову моћ може се наћи у великом броју радова (на пример *Torabi et al. (2018)*). Теоријски, редослед тестова који се добија користећи њихове емпиријске моћи је близак редоследу тестова који се добија користећи њихове Хаџис-Леманове асимптотске ефикасности (*Nikitin (1995)*) и не мора се увек поклапати са редоследом тестова добијеним коришћењем Бахадурове асимптотске ефикасности. У овом поглављу ће бити приказани Бахадурови нагиби класичних тестова прилагођених за тестирање сагласности са експоненцијалном расподелом, као и неких новијих тестова формираних баш за тестирање сагласности са експоненцијалном расподелом. Претпоставимо да је фамилија алтернативних расподела $\mathcal{G}_{\mathcal{E}}$ дефинисана као у поглављу 4.2 која задовољава услове (4.7). Такође, претпостављамо да су испуњени услови који су потребни да би важило (4.5) или (4.6). Све вредности су израчунате локално, односно кад $\vartheta \to 0$. У даљем тексту претпоставићемо да је $\lambda = 1$.

Прву групу тестова чине тестови чије тест статистике се могу написати као недегенерисане $V-$статистике. Ове тест статистике као граничну расподелу имају нормалну расподелу. У зависности од тога да ли статистика у свом запису нема оцену параметра или је има, тачни Бахадурови нагиби се могу добити применом теореме 4.8, односно 4.9. За тестове за које није раније одређен Бахадуров нагиб, одређен је у раду аутора (*Cuparić et al. (2019a)*). У ову групу спадају следећи тестови:

- Тестови интегралног типа засновани на карактеризацијама једнаке расподељености, односно тестови облика

$$I_n = \int_0^\infty \Big(H_n^{(1)}(t) - H_n^{(2)}(t)\Big) dF_n(t),$$

где су $H_n^{(1)}(t)$ и $H_n^{(2)}(t)$ V-емпиријске функције расподела случајних величина из карактеризације дефинисане као у (3.2) и (3.3), а F_n је емпиријска функција расподеле. Размотрићемо овај тип тестова за конкретне карактеризације. Тачни Бахадурови нагиби наведених тестова су одређени у радовима у којима су тестови и предложени, а у општем случају нагиб се може одредити коришћењем теореме 4.8.

  - Тест предложен у *Jovanović et al. (2015)*, заснован на карактеризацији Арнолд-Виласењор (*Arnold and Villasenor (2013)*), са тест статистиком

$$I_{n,k}^{(1)} = \int_0^\infty \frac{1}{n^k} \sum_{i_1,\dots,i_k} \Bigg(\mathrm{I}\{\max(X_{i_1},\dots,X_{i_k}) < t\} - \frac{1}{k!} \sum_{\pi(j_1,\dots,j_k)} \mathrm{I}\left\{ \frac{X_{i_1}}{j_1},\dots,\frac{X_{i_k}}{j_k} < t \right\} \Bigg) dF_n(t)$$

  и тачним Бахадуровим нагибом, за $k = 2$,

$$c(\vartheta) \sim 2721.60 \Big(\int_0^\infty \Big(\frac{4}{9}e^{-x} - \frac{1}{3}e^{-2x} - \frac{1}{6}e^{-\frac{x}{2}} \Big) g'_\vartheta(x;0,1)dx \Big)^2 \vartheta^2,$$

  односно, за $k = 3$,

$$c(\vartheta) \sim 2107.51 \Big(\int_0^\infty \Big(\frac{5}{8}e^{-x} - \frac{3}{4}e^{-2x} + \frac{1}{4}e^{-3x} - \frac{3}{32}e^{-\frac{x}{2}} - \frac{1}{12}e^{-\frac{x}{3}} \Big) g'_\vartheta(x;0,1)dx \Big)^2 \vartheta^2.$$

  - Тест предложен у *Milošević and Obradović (2016b)*, заснован на карактеризацији Милошевић-Обрадовић (*Milošević and Obradović (2016)*) са тест статистиком

$$I_n^{(2)} = \int_0^\infty \left(\frac{1}{n^2} \sum_{j,k=1}^n \mathrm{I}\{\max(X_j, X_k) < t\} - \frac{1}{n^3} \sum_{i,j,k=1}^n \mathrm{I}\{X_i + \min(X_j, X_k) < t\} \right) dF_n(t)$$

  и тачним Бахадуровим нагибом

$$c(\vartheta) \sim 1080 \Big(\int_0^\infty \Big(\frac{1}{12} - \frac{1}{8}e^{-2x} - \frac{1}{6}e^{-x} \Big) g'_\vartheta(x;0,1)dx \Big)^2 \vartheta^2.$$

  - Тест предложен у *Milošević (2016)*, заснован на Обрадовићевој карактеризацији (*Obradović (2015b)*), са тест статистиком

$$I_n^{(3)}=\int_0^\infty\left(\frac{1}{n^3}\sum_{j,k,l=1}^{n}\mathrm{I}\{\max(X_j,X_k,X_l)<t\}-\frac{1}{n^4}\sum_{i,j,k,l=1}^{n}\mathrm{I}\{X_i+med(X_j,X_k,X_l)<t\}\right)dF_n(t)$$

и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 1448.28\Big(\int_0^\infty\Big(-\frac{1}{20}+\frac{2}{5}e^{-3x}-\frac{9}{10}e^{-2x}+\frac{1}{2}e^{-x}\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

– Тест предложен у *Volkova (2015)*, заснован на Јанев-Чакрабортијевој карактеризацији (*Yanev and Chakraborty (2013)*), са тест статистиком

$$I_n^{(4)}=\int_0^\infty\frac{1}{n^3}\sum_{i,j,k=1}^{n}\left(\mathrm{I}\{\max(X_i,X_j,X_k)<t\}-\frac{1}{6}\sum_{\pi(i,j,k)}\mathrm{I}\left\{\frac{X_i}{3}+\max(X_j,X_k)<t\right\}\right)dF_n(t)$$

и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 7596.52\Big(\int_0^\infty\Big(-\frac{1}{4}e^{-3x}+\frac{9}{16}e^{-2x}-\frac{3}{8}e^{-x}+\frac{1}{12}e^{-\frac{x}{3}}+\frac{1}{2}\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

– Тест заснован на Пури-Рубиновој карактеризацији (карактеризација 3.1), са тест статистиком

$$I_n^{\mathcal{P}}=\int_0^\infty\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}\{X_i<t\}-\frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{I}\{|X_i-X_j|<t\}\right)dF_n(t)$$

и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 194.40\Big(\int_0^\infty\Big(\frac{2x}{3}e^{-x}-\frac{1}{6}\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

– Тест предложен у *Nikitin (2017)*, заснован на Десуовој карактеризацији (карактеризација 3.2), са тест статистиком

$$I_n^{\mathcal{D}}=\int_0^\infty\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}\{X_i<t\}-\frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{I}\{\min(X_i,X_j)<t\}\right)dF_n(t)$$

и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 343.64\Big(\int_0^\infty\Big(\frac{4}{9}e^{-3x}-\frac{1}{3}e^{-x}+\frac{1}{18}\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

- Тест заснован на емпиријској карактеристичној функцији предложен у *Epps and Pulley (1986)* са тест статистиком

$$EP_n = \sqrt{48}\Big(\frac{1}{n}\sum_{j=1}^{n} e^{-\frac{X_j}{\overline{X}_n}} - \frac{1}{2}\Big)$$

  и приближаним Бахадуров нагибом

$$c^{(a)}(\vartheta) \sim 3\Bigg(\int_0^\infty \Big(4e^{-x} + x\Big) g'_\vartheta(x;0,1)dx\Bigg)^2 \cdot \vartheta^2.$$

- Скор тест за параметар облика код Вејбулове расподеле предложен у *Cox and Oakes (1984)* са тест статистиком

$$CO_n = 1 + \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{X_i}{\overline{X}_n}\right)\log\frac{X_i}{\overline{X}_n}$$

  и приближаним Бахадуров нагибом

$$c^{(a)}(\vartheta) \sim \frac{6}{\pi^2}\Bigg(\int_0^\infty \Big((1-x)\log x + (1-\gamma)x\Big) g'_\vartheta(x;0,1)dx\Bigg)^2 \cdot \vartheta^2.$$

- Тест заснован на Ђинијевом коефицијенту предложен у *Gail and Gastwirth (1978)* са тест статистиком

$$G_n^* = \Big|\frac{1}{2n(n-1)\overline{X}_n}\sum_{i,j=1}^{n}|X_i - X_j| - \frac{1}{2}\Big|$$

  и приближаним Бахадуров нагибом (видети *Nikitin and Tchirina (1996)*)

$$c^{(a)}(\vartheta) \sim 12\Bigg(\int_0^\infty \Big(2e^{-x} + \frac{x}{2}\Big) g'_\vartheta(x;0,1)dx\Bigg)^2 \vartheta^2.$$

- Скор тест за параметар облика код гама расподеле предложен у *Moran (1951)* и *Tchirina (2005)* са тест статистиком

$$MO_n = \Big|\gamma + \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{\overline{X}_n}\Big|$$

  и приближаним Бахадуров нагибом (видети *Tchirina (2005)*)

$$c^{(a)}(\vartheta) \sim \Big(\frac{\pi^2}{6} - 1\Big)^{-1}\Bigg(\int_0^\infty (\log x - x) g'_\vartheta(x;0,1)dx\Bigg)^2 \vartheta^2.$$

- Тестови интегралног типа засновани на Лапласовој трансформацији и карактеризацијама једнаке расподељености са тест статистикама

$$J_{n,a} = \int_0^\infty \Big(\mathcal{L}_n^{(1)}(t) - \mathcal{L}_n^{(2)}(t)\Big)e^{-at}dt, \tag{4.27}$$

где су $\mathcal{L}_n^{(1)}(t)$ и $\mathcal{L}_n^{(2)}(t)$ V-емпиријске Лапласове трансформације случајних величина из карактеризације ψ_1 и ψ_2 дефинисане у (3.11) и (3.12), примењене на скалиран узорак и $a > 0$. Ови тестови су предложени у *Milošević and Obradović (2016a)*. Код ових тестова оцена параметра нема утицаја на граничну расподелу тест статитике. Приближни Бахадуров нагиб ових тестова за одговарајуће језгро Φ и Десуову карактеризацију је

$$\begin{aligned} c^{(a)}(\vartheta) \sim \frac{1}{\sigma^2_{\mathcal{D}}(a)}\Big(\int_0^\infty \Big(&\frac{1}{2(a+x)} - \frac{1}{2(a+2x)}e^{-x}\Big(2 + ae^{\frac{a}{2}+x}\Gamma\Big(0,\frac{a}{2}\Big) + 2e^{\frac{a}{2}+x}x\Gamma\Big(0,\frac{a}{2}\Big)\Big) \\ &- \frac{1}{2}e^a Ei(-a) - ae^{\frac{a}{2}+x}\Gamma\Big(0,\frac{a}{2}+x\Big) - 2e^{\frac{a}{2}+x}x\Gamma\Big(0,\frac{a}{2}+x\Big)\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2, \end{aligned}$$

односно Пури-Рубинову карактеризацију

$$\begin{aligned} c^{(a)}(\vartheta) \sim \frac{1}{\sigma^2_{\mathcal{P}}(a)}\Big(\int_0^\infty \Big(&\frac{1}{2(a+x)} - e^{-x-a}\Big(e^{2a}Ei(-a) + Ei(a) - Ei(a+x)\Big) \\ &- \frac{1}{2}e^a Ei(-a))\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2, \end{aligned}$$

где је $Ei(z) = -\int_{-z}^\infty \frac{e^{-t}}{t}dt$ и $\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}dt$, а $\sigma^2_{\mathcal{D}}(a)$ и $\sigma^2_{\mathcal{P}}$ се могу израчунати за конкретно задате a (за више детаља видети *Milošević and Obradović (2016a)*).

У другу групу спадају тестови L^2-типа, односно тестови чије се тест статистике могу приказати као слабо дегенерисане $V-$статистике са језгром Φ облика (2.21). Граничне расподеле следећих тест статистика су одређене у радовима у којима су и предложени, али се могу одредити и коришћењем теореме 2.9. Бахадурови нагиби тих тестова су одређени у радовима аутора (*Cuparić (2019)* и *Cuparić et al. (2019a)*), а с обзиром да сви у свом запису имају оцењен параметар могу се израчунати користећи теорему 4.7. Због компликованости израза за Бахадурове нагибе, биће приказана језгра одговарајућих $V-$статистика, на основу којих се заменом у изразе за Бахадурове нагибе добијају одговарајући изрази. С обзиром да су све тест статистике реда 2, друга пројекција језгра је једнака језгру статистике. Сопствене вредности одговарајућих оператора су рачунате уз помоћ апроксимације приказане у поглављу 4.3 и дата је највећа од њих за сваки од тестова. Размотрићемо следеће тестове:

- Лилифорсова модификација Крамер-фон Мизесовог теста са тест статистиком

$$\omega_n^2 = \int\limits_0^\infty (F_n(t) - (1 - e^{-\frac{t}{\overline{X}_n}}))^2 \frac{1}{\overline{X}_n} e^{-\frac{t}{\overline{X}_n}} dt.$$

Језгро одговарајуће $V-$статистике је

$$\Phi_{\omega^2}(x, y; \mu(\vartheta, 1)) = e^{-\max\left(\frac{x}{\mu(\vartheta,1)}, \frac{y}{\mu(\vartheta,1)}\right)} - e^{-\frac{x}{\mu\vartheta,1}} - e^{-\frac{y}{\mu(\vartheta,1)}} + \frac{1}{2}\left(e^{-\frac{2x}{\mu(\vartheta,1)}} + e^{-\frac{2y}{\mu(\vartheta,1)}}\right) + \frac{1}{3},$$

а највећа сопствена вредност је $\upsilon_1 = \frac{21}{250}$.

- Лилифорсова модификација теста Андерсон-Дарлинга са тест статистиком

$$\mathrm{AD}_n = \int\limits_0^\infty \frac{(F_n(t) - (1 - e^{-\frac{t}{\overline{X}_n}}))^2}{\overline{X}_n(1 - e^{-\frac{t}{\overline{X}_n}})} dt.$$

Језгро одговарајуће $V-$статистике је

$$\Phi_{\mathrm{AD}}(x, y; \mu(\vartheta, 1)) = \frac{x}{\mu(\vartheta, 1)} + \frac{y}{\mu(\vartheta, 1)} - 1 - \log(e^{\max(\frac{x}{\mu(\vartheta,1)}, \frac{y}{\mu(\vartheta,1)})} - 1),$$

а највећа сопствена вредност је $\upsilon_1 = \frac{23}{50}$.

- Тест предложен у *Baringhaus and Henze (1991)* са тест статистиком

$$\mathrm{BH}_n = \int\limits_0^\infty \Big((1 + t)\mathcal{L}_n'(t) + \mathcal{L}_n(t)\Big)^2 e^{-at} dt,$$

где је $\mathcal{L}_n(t) = \frac{1}{n}\sum\limits_{i=1}^{n} e^{-t\frac{X_i}{\overline{X}_n}}$ емпиријска Лапласова трансформација. Језгро одговарајуће $V-$статистике је

$$\Phi_{\mathrm{BH}}(x, y; \mu(\vartheta, 1)) = \frac{(1 - x)(1 - y)}{x + y + a\mu(\vartheta, 1)} - \frac{x\mu(\vartheta, 1) + y\mu(\vartheta, 1) - 2xy}{(a\mu(\vartheta, 1) + x + y)^2} + \frac{2xy\mu(\vartheta, 1)}{(a\mu(\vartheta, 1) + x + y)^3}.$$

Код овог теста сопствене вредности зависе од вредности параметра a и њихове вредности су приказане у табели 4.4 за различите вредности овог параметра.

- Тест предложен у *Henze (1992)* са тест статистиком

$$\mathrm{HE}_n = \int\limits_0^\infty \Big(\mathcal{L}_n(t) - \frac{1}{1 + t}\Big)^2 e^{-at} dt,$$

где је $\mathcal{L}_n(t) = \frac{1}{n}\sum\limits_{i=1}^{n} e^{-t\frac{X_i}{\overline{X}_n}}$ емпиријска Лапласова трансформација. Језгро одговарајуће $V-$статистике је

$$\Phi_{\mathrm{HE}}(x,y;\mu(\vartheta,1)) = 1 + \frac{\mu(\vartheta,1)}{a\mu(\vartheta,1)+x+y} + ae^{a}Ei(-a) + e^{a+\frac{x}{\mu(\vartheta,1)}}Ei\left(-\left(a+\frac{x}{\mu(\vartheta,1)}\right)\right)$$
$$+ e^{a+\frac{y}{\mu(\vartheta,1)}}Ei\left(-\left(a+\frac{y}{\mu(\vartheta,1)}\right)\right),$$

где је $Ei(z) = -\int\limits_{-z}^{\infty}\frac{e^{-t}}{t}dt$. Као и код претходног теста, и код овог теста сопствене вредности зависе од вредности параметра a и њихове вредности су приказане у табели 4.4 за различите вредности овог параметра.

- Тест предложен у *Henze and Meintanis (2002a)* са тест статистиком

$$HM_n = \int\limits_{0}^{\infty}\Big(\mathcal{L}_n(t) - \frac{1}{1+t}\Big)^2(1+t)^2e^{-at}dt,$$

где је $\mathcal{L}_n(t) = \frac{1}{n}\sum\limits_{i=1}^{n}e^{-t\frac{X_i}{\overline{X}_n}}$ емпиријска Лапласова трансформација. Језгро одговарајуће $V-$статистике је

$$\Phi_{HM}(x,y;\mu(\vartheta,1)) = \frac{1}{a} - \frac{\mu(\vartheta,1)(\mu(\vartheta,1)(1+a)+x)}{(a\mu(\vartheta,1)+x)^2} - \frac{\mu(\vartheta,1)(\mu(\vartheta,1)(1+a)+y)}{(a\mu(\vartheta,1)+y)^2}$$
$$+ \frac{2\mu^3(\theta)}{(a\mu(\vartheta,1)+x+y)^3} + \frac{2\mu^2(\theta)}{(a\mu(\vartheta,1)+x+y)^2} + \frac{\mu(\vartheta,1)}{a\mu(\vartheta,1)+x+y},$$

а сопствене вредности такође зависе од параметра a и приказане су у табели 4.4.

- Два теста предложена у *Henze and Meintanis (2002b)* са тест статистикама

$$\mathrm{HM}_n^{(i)} = \int\limits_{0}^{\infty}(s_n(t) - tc_n(t))^2\omega_i(t)dt, \quad i = 1,2,$$

где су $c_n(\cdot)$ и $s_n(\cdot)$ реални и имагирани део емпријске карактеристичне функције $\Psi_n(t) = \frac{1}{n}\sum\limits_{j=1}^{n}e^{it\frac{X_j}{\overline{X}_n}}$, а $\omega_1(t) = e^{-at}$ и $\omega_2(t) = e^{-at^2}$. Њихова језгра су

$$\Phi^{(1)}_{\mathrm{HM}}(x,y;\mu(\vartheta,1)) = \frac{a\mu^2(\vartheta,1)}{2(a^2\mu^2(\vartheta,1)+(x-y)^2)} - \frac{a\mu^2(\vartheta,1)}{2(a^2\mu^2(\vartheta,1)+(x+y)^2)}$$
$$+a\frac{a^2\mu^2(\vartheta,1)-3(x-y)^2}{(a^2\mu^2(\vartheta,1)+(x-y)^2)^3} + a\frac{a^2\mu^2(\vartheta,1)-3(x+y)^2}{(a^2\mu^2(\vartheta,1)+(x+y)^2)^3} - \frac{2a\mu^3(\theta)(x+y)}{(a^2\mu^2(\vartheta,1)+(x+y)^2)^2}$$

и

$$\Phi^{(2)}_{\mathrm{HM}}(x,y;\mu(\vartheta,1)) = \frac{\sqrt{\pi}}{4\sqrt{a}}\Bigg(\Big(\frac{1}{2a} - \frac{x+y}{a\mu(\vartheta,1)} - \frac{(x-y)^2}{4a^2\mu^2(\vartheta,1)} - 1\Big)e^{-\frac{(x+y)^2}{4a\mu^2(\vartheta,1)}}$$
$$+\Big(1 + \frac{1}{2a} - \frac{(x-y)^2}{4a^2\mu^2(\vartheta,1)}\Big)e^{-\frac{(x-y)^2}{4a\mu^2(\vartheta,1)}}\Bigg).$$

Сопствене вредности такође зависе од параметра a и приказане су у табели 4.4.

Табела 4.4: Сопствене вредности одговарајућих оператора за различите вредности параметра a

a	0.5	1	2	5	10
BH_n	$1.42\cdot10^{-1}$	$7.89\cdot10^{-2}$	$3.61\cdot10^{-2}$	$8.98\cdot10^{-3}$	$2.32\cdot10^{-3}$
HE_n	$4.86\cdot10^{-2}$	$1.58\cdot10^{-2}$	$3.77\cdot10^{-3}$	$3.24\cdot10^{-4}$	$3.32\cdot10^{-5}$
HM_n	$9.17\cdot10^{-1}$	$1.47\cdot10^{-1}$	$1.92\cdot10^{-2}$	$8.80\cdot10^{-3}$	$6.22\cdot10^{-5}$
$HM_n^{(1)}$	2.30	$4.33\cdot10^{-1}$	$6.11\cdot10^{-2}$	$2.18\cdot10^{-3}$	$8.28\cdot10^{-5}$
$HM_n^{(2)}$	$3.68\cdot10^{-1}$	$1.20\cdot10^{-1}$	$3.33\cdot10^{-2}$	$4.53\cdot10^{-3}$	$7.89\cdot10^{-4}$

У последњу групу спадају тестови $L^\infty-$типа, односно тестови који се могу представити као супремум апсолутне вредности недегенерисане $V-$статистике. Размотрићемо прво тестове $L^\infty-$типа засноване на $V-$емпиријској функцији и карактеризацијама једнаке расподељености облика

$$D_n = \sup_{t>0}\Big|H_n^{(\omega_1)}(t) - H_n^{(\omega_2)}(t)\Big|.$$

Бахадурови нагиби ових тестова се могу добити користећи став 4.1. У ту групу спадају следећи тестови, чији су Бахадурови нагиби одређени у радовима у којима су тестови и предложени:

- Тест предложен у *Jovanović et al. (2015)*, заснован на карактеризацији Арнолд-Виласињор (*Arnold and Villasenor (2013)*) са тест статистиком

  $$D_{n,k}^{(1)} = \sup_{t\geqslant 0}\left|\frac{1}{n^k}\sum_{i_1,\dots,i_k=1}^{n}\left(I\{\max(X_{i_1},\dots,X_{i_k})<t\} - \frac{1}{k!}\sum_{\pi(j_1,\dots,j_k)} I\left\{\frac{X_{i_1}}{j_1},\dots,\frac{X_{i_k}}{j_k}<t\right\}\right)\right|$$

  и тачним Бахадуровим нагибом, за $k=2$,

  $$\begin{aligned}c(\vartheta)\sim 44.76\sup_{t>0}\Big(\int_0^\infty\Big(&I\{x<t\}F(t)-\frac{1}{2}I\{x<t\}F(2(t-x))\\&-\frac{1}{2}I\{x<2t\}F\Big(t-\frac{x}{2}\Big)\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2,\end{aligned}$$

  односно, за $k=3$,

  $$\begin{aligned}c(\vartheta)\sim 44.62\sup_{t>0}\Big(\int_0^\infty\Big(&\mathrm{I}\{x<\ t\}\Big(F^2(t)-F(2(t-x))+\frac{2}{3}F(3(t-x))\Big)\\&-I\{x<2t\}\Big(\frac{1}{2}F(t-\frac{x}{2})-\frac{1}{6}F(3(t-\frac{x}{2}))\Big)\\&-I\{x<3t\}\Big(\frac{2}{3}F(t-\frac{x}{3})-\frac{1}{2}F(2(t-\frac{x}{3}))\Big)\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.\end{aligned}$$

- Тест предложен у *Milošević and Obradović (2016b)*, заснован на карактеризацији Милошевић-Обрадовић (*Milošević and Obradović (2016)*) са тест статистиком

$$D_n^{(2)} = \sup_{t\geqslant 0}\left|\left(\frac{1}{n^2}\sum_{j,k=1}^{n} I\{\max(X_j,X_k)<t\} - \frac{1}{n^3}\sum_{i,j,k=1}^{n} I\{X_i+\min(X_j,X_k)<t\}\right)\right|$$

  и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 64.10\sup_{t>0}\Big(\int\limits_0^\infty\Big(\frac{1}{3}e^{-2t}-\frac{1}{3}+\frac{2}{3}te^{-t}+\frac{1}{3}I\{x<t\}(e^{-2t}+2x+1-2e^{-t}(1-x+t))\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

- Тест предложен у *Milošević (2016)*, заснован на Обрадовићевој карактеризацији (*Obradović (2015b)*) са тест статистиком

$$D_n^{(3)} = \sup_{t\geqslant 0}\left|\left(\frac{1}{n^3}\sum_{j,k,l=1}^{n} I\{\max(X_j,X_k,X_l)<t\} - \frac{1}{n^4}\sum_{i,j,k,l=1}^{n} I\{X_i+med(X_j,X_k,X_l)<t\}\right)\right|$$

  и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 58.82\sup_{t>0}\Big(\int\limits_0^\infty\Big(\frac{1}{4}I\{x<t\}e^{-x-3t}\Big(-e^x-2e^{4x}+6e^{2t}+3e^x+t+3e^{3x}+t-e^x$$
$$+2t(9-6x)\Big)+\frac{1}{4}I\{x\geqslant t\}e^{-3t}\Big(-1+6e^t-2e^{3t}-e^{2t}(3-6t)\Big)\Big)gg'_\vartheta(x;0,1)dx\Big)^2\vartheta^2.$$

- Тест предложен у *Volkova (2015)*, заснован на Јанев-Чакрабортијевој карактеризацији (*Yanev and Chakraborty (2013)*) са тест статистиком

$$D_n^{(4)} = \sup_{t\geqslant 0}\left|\frac{1}{n^3}\sum_{i,j,k=1}^{n}\left(I\{\max(X_i,X_j,X_k)<t\} - \frac{1}{6}\sum_{\pi(i,j,k)} I\left\{\frac{X_i}{3}+\max(X_j,X_k)<t\right\}\right)\right|$$

  и тачним Бахадуровим нагибом

$$c(\vartheta)\sim 89.38\sup_{t>0}\Big(\int\limits_0^\infty\Big(\varphi_1(x;t)\Big)g'_\vartheta(x;0,1)dx\Big)^2\vartheta^2,$$

  где је $\varphi_1(x;t)=E(\Phi(X_1,X_2,X_3,t)|X_1=x)$ и

$$\Phi(x_1,x_2,x_3,t)=\frac{1}{3}\Big(-I\{3\max(x_1,x_2)+x_3<3t\}-I\{3\max(x_1,x_3)+x_2<3t\}$$
$$-I\{3\max(x_2,x_3)+x_1<3t\}+3I\{\max(x_1,x_2,x_3)<t\}\Big).$$

  Због сложености израза није дат експлицитан облик прве пројекције.

- Тест заснован на Пури-Рубиновој карактеризацији (карактеризација 3.1) са тест статистиком

$$D_n^{\mathcal{P}} = \sup_{t \geqslant 0} \left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant t\} - \frac{1}{n^2} \sum_{i,j=1}^{n} \mathrm{I}\{|X_i - X_j| \leqslant t\} \right|$$

и тачним Бахадуровим нагибом

$$c(\vartheta) \sim 22.72 \sup_{t>0} \Big(\int_0^{\infty} \frac{1}{2} \big((2e^{t-x} - 2) I\{x > t\} - I\{x < t\} - 2e^{-x-t} + e^{-t} + 1 \big) \, g'_{\vartheta}(x; 0, 1) dx \Big)^2 \vartheta^2.$$

- Тест предложен у *Nikitin (2017)* заснован на Десуовој карактеризацији (карактеризација 3.2) са тест статистиком

$$D_n^{\mathcal{D}} = \sup_{t \geqslant 0} \left| \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant t\} - \frac{1}{n^2} \sum_{i,j=1}^{n} \mathrm{I}\{2 \min(X_i, X_j) \leqslant t\} \right|$$

и тачним Бахадуровим нагибом

$$c(\vartheta) \sim 16 \sup_{t>0} \Big(\int_0^{\infty} \frac{1}{2} \Big(-I\{x < t\} - 2e^{-\frac{t}{2}} I\{2x > t\} + e^{-t} + 1 \Big) \, g'_{\vartheta}(x; 0, 1) dx \Big)^2 \vartheta^2.$$

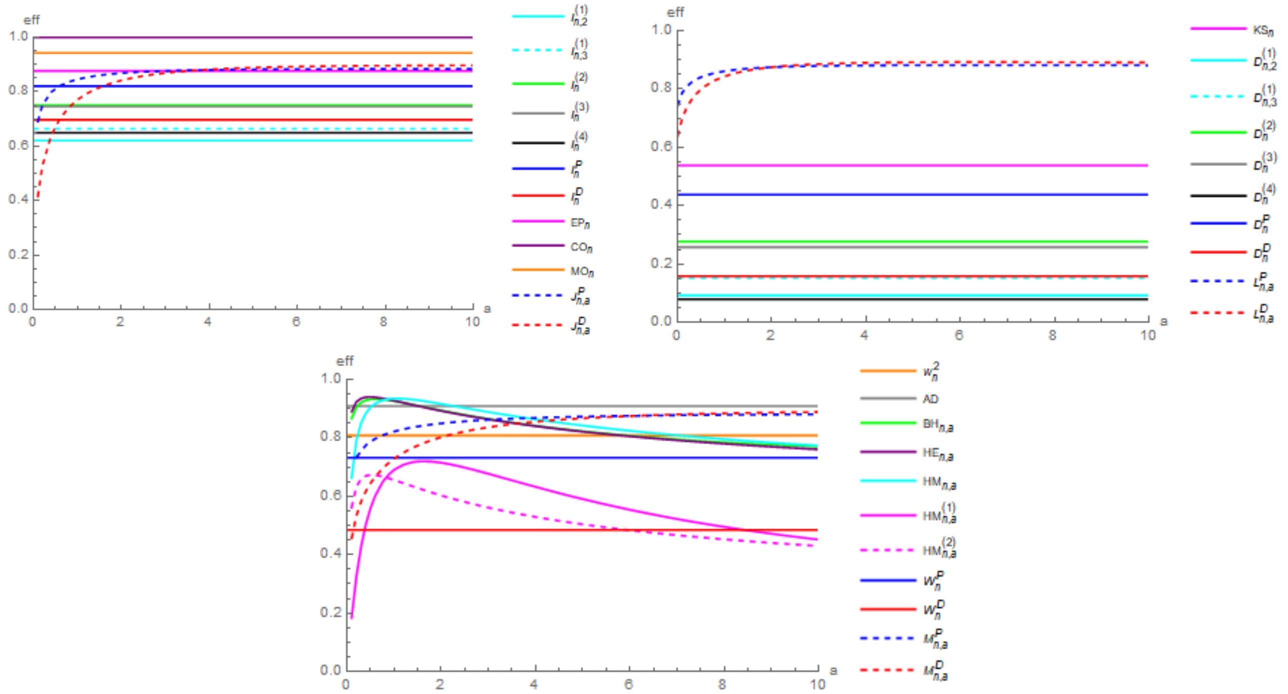
Поред наведених тестова у ову групу спада и Лилифорсова модификација теста Колмогоров-Смирнова (*Lilliefors (1969)*)

$$KS_n = \sup |F_n(t) - (1 - e^{-\frac{t}{\bar{X}_n}})|.$$

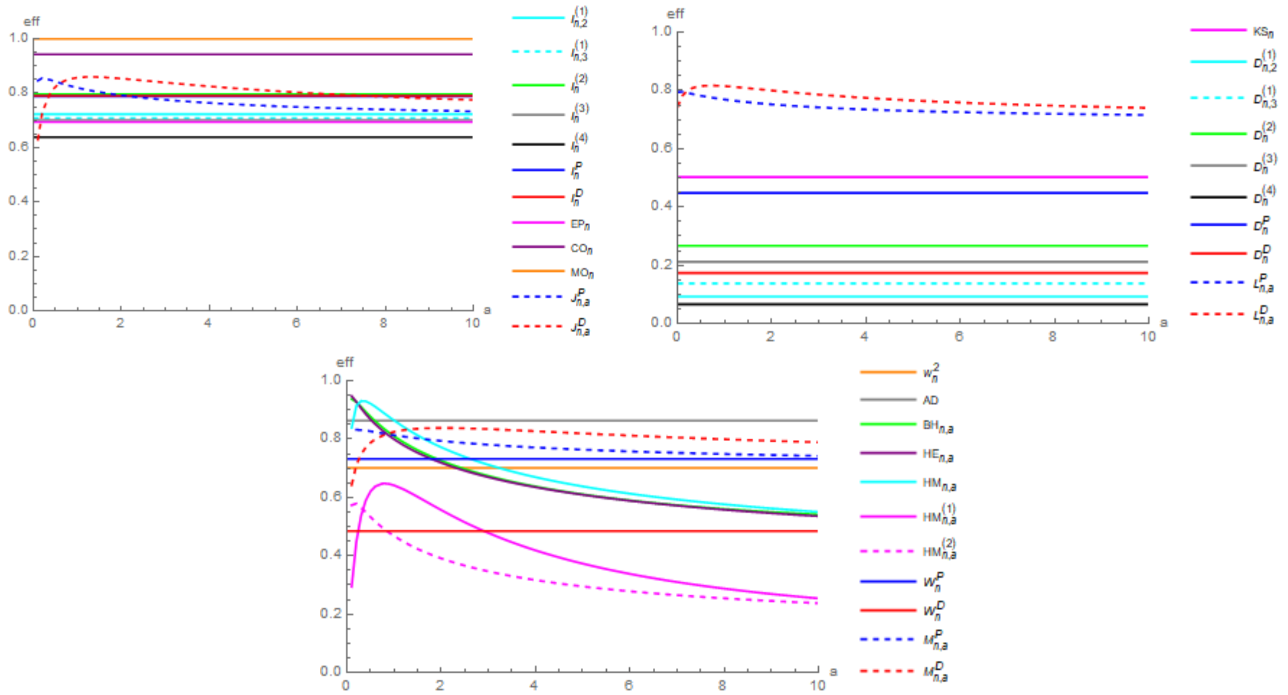
Овај тест садржи оцену параметра у свом запису па се његов приближни Бахадуров нагиб рачуна користећи теорему 4.10, израчунат је у раду *Nikitin and Tchirina (2007)* и дат је изразом

$$c^{(a)}(\vartheta) \sim 6.70 \sup_{x \geqslant 0} \Big(x e^{x} \int_0^{\infty} G'_{\vartheta}(u; 0, 1) du - G(x; 0, 1) dx \Big)^2 \cdot \vartheta^2.$$

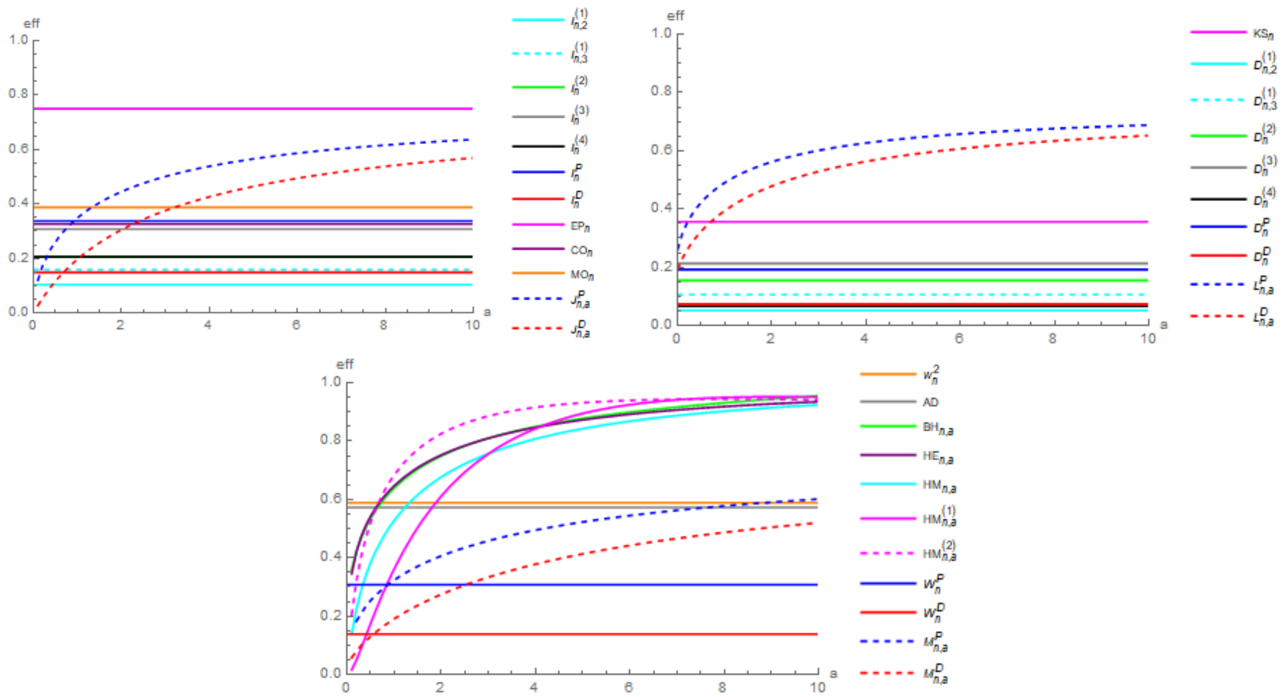
На сликама 4.1-4.4 приказане су локалне Бахадурове релативне ефикасности претходно наведених тестова, као и тестова поменутих у претходна два поглавља, у односу на тест количника веродостојности у функцији од параметра подешавања a. Код тестова који не зависе од параметра подешавања a вредности су представљене правом линијом. На сликама су тестови подељени у три групе, као што су и раније били подељени, за сваку од четири алтернативе (4.23)-(4.26).



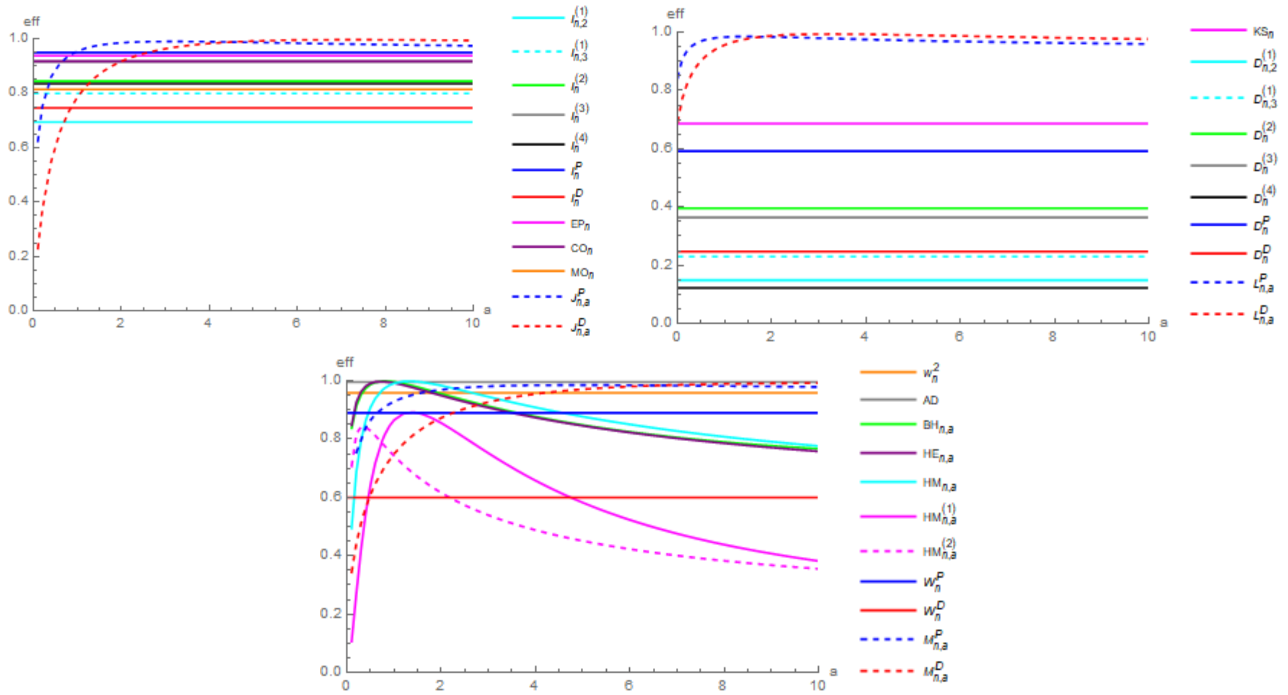
Слика 4.1: Локална апсолутна приближна Бахадурова ефикасност за Вејбулову расподелу као алтернативну расподелу



Слика 4.2: Локална апсолутна приближна Бахадурова ефикасност за гама расподелу као алтернативну расподелу



Слика 4.3: Локална апсолутна приближна Бахадурова ефикасност за расподелу линеарне стопе отказа као алтернативну расподелу



Слика 4.4: Локална апсолутна приближна Бахадурова ефикасност за мешавину експоненцијалних расподала $EMNW(3)$ као алтернативну расподелу

Оно што се свакако може приметити код свих алтернатива и тестова заснованих на L^∞ растојањима је да су тестови $L^{\mathcal{P}}_{n,a}$ и $L^{\mathcal{D}}_{n,a}$, приказани у поглављу 3.4, значајно бољи од осталих тестова наведеног типа. Ако се пореде та два теста међусобно, разлика међу њима је мала, а који је бољи зависи од алтернативе и вредности параметра a.

Код осталих група тестова, не можемо рећи да је неки тест најефикаснији за све алтернативе. Такође, утицај параметра a је приметан и код различитих тестова другачије утиче на монотоност функције. Познато је да су тестови CO_n и MO_n локално оптимални тестови за Вејбулову и гама алтернативу, редом, па су и најефикаснији у тим случајевима. У случају расподеле линеарне стопе отказа као алтернативне расподеле, најефикаснији су тестови EP_n, $HM^{(1)}_{n,a}$ и $HM^{(1)}_{n,a}$. Међутим, за све остале алтернативе тестови $HM^{(1)}_{n,a}$ и $HM^{(1)}_{n,a}$ су међу најлошијима. Код $EMNW(3)$ алтернативе, погодним избором параметра a може се постићи релативна ефикасност блиска 1 код свих тестова заснованих на Лапласовим трансформацијама случајних величина из карактерзација ($J_{n,a}$, $M_{n,a}$ и $L_{n,a}$).

У табелама 4.5 и 4.6 приказане су апсолутне Бахадурове релативне ефикасности за све претходно наведене тестове, алтернативе (4.23)-(4.26) и за $a \in \{0.5, 1, 2, 5, 10\}$, за оне тестове који зависе од параметра a.

Табела 4.5: Локалне апсолутне приближне Бахадурове ефикасности

	W	Γ	LFR	$EMNW(3)$
$I_{n,2}^{(1)}$	0.621	0.723	0.104	0.694
$I_{n,3}^{(1)}$	0.664	0.708	0.159	0.799
$I_n^{(2)}$	0.750	0.796	0.208	0.844
$I_n^{(3)}$	0.746	0.701	0.308	0.916
$I_n^{(4)}$	0.649	0.638	0.206	0.835
$I_n^{\mathcal{P}}$	0.821	0.788	0.337	0.949
$I_n^{\mathcal{D}}$	0.697	0.790	0.149	0.746
EP_n	0.876	0.694	0.750	0.937
CO_n	1	0.943	0.326	0.917
G_n^*	0.876	0.694	0.750	0.937
MO_n	0.943	1	0.388	0.814
$J_{n,0.5}^{\mathcal{P}}$	0.812	0.843	0.262	0.888
$J_{n,1}^{\mathcal{P}}$	0.846	0.820	0.349	0.955
$J_{n,2}^{\mathcal{P}}$	0.868	0.792	0.445	0.985
$J_{n,5}^{\mathcal{P}}$	0.882	0.756	0.566	0.987
$J_{n,10}^{\mathcal{P}}$	0.884	0.733	0.637	0.974
$J_{n,0.5}^{\mathcal{D}}$	0.674	0.826	0.117	0.608
$J_{n,1}^{\mathcal{D}}$	0.771	0.857	0.198	0.786
$J_{n,2}^{\mathcal{D}}$	0.842	0.854	0.305	0.917
$J_{n,5}^{\mathcal{D}}$	0.889	0.813	0.465	0.991
$J_{n,10}^{\mathcal{D}}$	0.896	0.775	0.569	0.994
ω_n^2	0.808	0.701	0.588	0.958
AD_n	0.909	0.863	0.573	0.996
$BH_{n,0.5}$	0.932	0.877	0.534	0.987
$BH_{n,1}$	0.926	0.810	0.638	0.996
$BH_{n,2}$	0.894	0.726	0.749	0.956
$BH_{n,5}$	0.823	0.611	0.878	0.848
$BH_{n,10}$	0.771	0.542	0.956	0.767
$HE_{n,0.5}$	0.940	0.868	0.542	0.991
$HE_{n,1}$	0.928	0.799	0.647	0.992
$HE_{n,2}$	0.893	0.719	0.752	0.949
$HE_{n,5}$	0.822	0.609	0.873	0.846
$HE_{n,10}$	0.761	0.536	0.935	0.758
$HM_{n,0.5}$	0.905	0.922	0.382	0.909
$HM_{n,1}$	0.935	0.864	0.528	0.991
$HM_{n,2}$	0.917	0.772	0.677	0.983
$HM_{n,5}$	0.842	0.638	0.842	0.877
$HM_{n,10}$	0.774	0.550	0.924	0.776
$HM_{n,0.5}^{(1)}$	0.560	0.621	0.174	0.643
$HM_{n,1}^{(1)}$	0.691	0.642	0.361	0.865
$HM_{n,2}^{(1)}$	0.715	0.557	0.612	0.855
$HM_{n,5}^{(1)}$	0.591	0.373	0.895	0.582
$HM_{n,10}^{(1)}$	0.452	0.254	0.951	0.382

Табела 4.6: Локалне апсолутне приближне Бахадурове ефикасности

	W	Γ	LFR	$EMNW(3)$
$HM^{(2)}_{n,0.5}$	0.673	0.533	0.520	0.828
$HM^{(2)}_{n,1}$	0.656	0.468	0.683	0.742
$HM^{(2)}_{n,2}$	0.603	0.391	0.825	0.616
$HM^{(2)}_{n,5}$	0.504	0.295	0.931	0.451
$HM^{(2)}_{n,10}$	0.430	0.238	0.942	0.355
$D^{(1)}_{n,2}$	0.092	0.093	0.052	0.149
$D^{(1)}_{n,3}$	0.152	0.138	0.106	0.230
$D^{(2)}_{n}$	0.277	0.267	0.155	0.396
$D^{(3)}_{n}$	0.258	0.212	0.213	0.364
$D^{(4)}_{n}$	0.079	0.066	0.067	0.122
$D^{\mathcal{P}}_{n}$	0.437	0.448	0.192	0.592
$D^{\mathcal{D}}_{n}$	0.158	0.174	0.073	0.247
KS	0.538	0.503	0.356	0.686
$W^{\mathcal{P}}_{n}$	0.733	0.719	0.308	0.891
$W^{\mathcal{D}}_{n}$	0.485	0.544	0.138	0.600
$M^{\mathcal{P}}_{n,0.5}$	0.787	0.827	0.253	0.865
$M^{\mathcal{P}}_{n,1}$	0.822	0.814	0.324	0.929
$M^{\mathcal{P}}_{n,2}$	0.850	0.794	0.407	0.969
$M^{\mathcal{P}}_{n,5}$	0.873	0.764	0.523	0.985
$M^{\mathcal{P}}_{n,10}$	0.881	0.742	0.601	0.979
$M^{\mathcal{P}}_{n,\infty}$	0.876	0.694	0.750	0.937
$M^{\mathcal{D}}_{n,0.5}$	0.645	0.788	0.130	0.610
$M^{\mathcal{D}}_{n,1}$	0.729	0.825	0.191	0.750
$M^{\mathcal{D}}_{n,2}$	0.803	0.838	0.275	0.873
$M^{\mathcal{D}}_{n,5}$	0.867	0.820	0.413	0.971
$M^{\mathcal{D}}_{n,10}$	0.889	0.789	0.520	0.992
$M^{\mathcal{D}}_{n,\infty}$	0.876	0.694	0.750	0.937
$L^{\mathcal{P}}_{n}$	0.743	0.795	0.261	0.843
$L^{\mathcal{P}}_{n,0.5}$	0.841	0.782	0.423	0.970
$L^{\mathcal{P}}_{n,1}$	0.861	0.769	0.490	0.983
$L^{\mathcal{P}}_{n,2}$	0.875	0.752	0.562	0.984
$L^{\mathcal{P}}_{n,5}$	0.882	0.730	0.644	0.972
$L^{\mathcal{P}}_{n,10}$	0.881	0.716	0.688	0.960
$L^{\mathcal{D}}_{n}$	0.639	0.744	0.192	0.691
$L^{\mathcal{D}}_{n,0.5}$	0.799	0.815	0.323	0.902
$L^{\mathcal{D}}_{n,1}$	0.844	0.815	0.394	0.957
$L^{\mathcal{D}}_{n,2}$	0.875	0.800	0.479	0.988
$L^{\mathcal{D}}_{n,5}$	0.892	0.766	0.588	0.990
$L^{\mathcal{D}}_{n,10}$	0.891	0.740	0.652	0.976

Глава 5

ТЕСТОВИ САГЛАСНОСТИ ЗА ЦЕНЗУРИСАНЕ ПОДАТКЕ

Највећи део статистичке литературе се бави потпуним узорцима, односно подацима у којима су доступне све вредности. Међутим, у већини реалних ситуација истраживачи су суочени са тим да су њихови узорци засечени или цензурисани. У неким ситуацијама, када је проценат таквих елемената у узорку мали, они се могу занемарити. Са друге стране, ако тај проценат није занемарљив, односно ако су рестрикције на узорак значајне, то се мора размотрити на адекватан начин како би узорак могао да се користи за даљу анализу.

Дакле, разматрамо узорак који је рестрихован на неки део узорачког простора. У зависности од природе рестриховања разликују се засечени и цензурисани узорци. Засечени узорци су они код којих су одређене вредности из популације у потпуности изостављене. Коен је у свом делу (*Cohen (1991)*) навео да је вероватно тачније рећи да је сама популација засечена, а да се узорци називају засеченим јер се узимају из засечене популације. Са друге стране, цензурисани узорци су они у којима се вредности бележе само за оне елементе који се налазе у одговарајућем делу узорачког простора, а за остале се не врши мерење. У првим радовима, цензурисани узорци су описивани као засечени са познатим бројем недостајућих (немерених) података. Може се десити да подаци имају и обе одлике, односно да буду засечени са једне стране а цензурисане са друге стране (најчешћи случај је засечености са леве и цензурисаности са десне стране).

Пример 5.1 *У раду Lagakos et al. (1988) су приказани подаци везани за време инфекције и инкубациони периода 258 одраслих (група 0) и 37 деце (група 1) који су заражени вирусом $HIV-$а кроз трансфузију контаминираном крвљу, што је довело до развоја AIDS-а, закључно са 30. јуном 1986. Подаци садрже датуме инфекције и инкубациони период за развој AIDS-а у виду тромесечних периода, са почетком 1. априла 1978. Деца су се заразила или пре порођаја или на самом порођају, тако да је код њих инкубациони период од 1. априла 1978. до рођења. Само оне особе код којих се развио AIDS пре завршетка проучавања су укључене у истраживање. Заражене особе код којих се није развила болест пре завршетка проучавања нису укључене у проучавање. Овакви подаци* *одговарају засеченим подацима са десне стране.*

Пример 5.2 *Резерват Амбосели у Кенији се састоји углавном од савана, са одређеним пределима са дрвећем и водом. Бабуни у овом подручју спавају на дрвећу, али сваког јутра силазе са дрвећа и одлазе у потрагу за храном често далеко од дрвећа где спавају. Посебно је посматрана једна група бабуна у Амбосељу. У тој групи време (од доласка истраживача) до силаска подразумева када средишњи члан групе (члан који представља медијану) сиђе на земљу. Ова времена се бележе јер расподела ових времена има значај за одређивање како бабуни распоређују време између различитих активности или стања неопходних за њихово преживљавање. Међутим, дешава се да истраживач када дође на локацију на којој се бабуни налазе, више од половине бабуна (или готово сви) се већ налази на земљи. У том случају се бележи наведено време, али је познато да се жељени догађај десио раније. Група је посматрана 152 дана, од чега је 58 пута истраживач био присутан када је циљни бабун сишао на земљу, док је 94 пута време силаска било пре доласка истраживача (Wagner and Altmann (1973)). Овакви подаци одговарају подацима цензурисаним са леве стране.*

Пример 5.3 *Механичка својства материјала који се користе су битна у многим наукама и индустријама. На пример, механичка својства стена и минерала одређују енергију која ће бити утрошена у рударским пословима. Често су механичка својства великог броја различитих материјала дефинисана њиховом структуром на микроскопском нивоу која се одражава на макроскопска својства као што је компресиона чврстоћа. Компресиона чврстоћа се одређује помоћу појединачног узорка дефинисаног облика, који се подвргава компресији до тачке пуцања. Један начин да се изврши испитивање компресионе чврстоће је квази-статички једноосни тест. Приликом извођења овог теста, узорак се поставља између две паралелне плоче које се међусобно постепено приближавају све док узорак не пукне, а континуално се бележи оптерећење и померање плоча током теста. Проучавана је компресиона чврстоћа кварца уз коришћење еласто-пластичних плоча са три прилагођене чврстоће. Тестирањем је добијено 50 вредности чврстоће. Уколико је елемент узорка успео да издржи компресију плочом максималне чврстоће, он је десно цензурисан јер код њега је позната само доња граница издржљивости и таквих елемената је 27. Међутим, елементи узорка су такође потенцијално засечени са леве стране пре почетка теста површинском снагом, што доводи до тога да је компресиона чврстоћа кварца мања од минималне компресионе чврстоће плоче. За више детаља видети Pejchal et al. (2017).*

Први резултати везани за обраду засечених података појавили су се релативно рано у раду Франсиса Галтона (*Galton (1898)*). Затим су такви подаци обрађени у неколико радова Пирсона (*Pearson (1902)*), Пирсона и Лија (*Pearson and Lee (1908)*), Лија (*Lee (1914)*) и Фишера (*Fisher (1931)*). У додатку рада *Bliss and Stevens (1937)*, Стивенс је дао једначине веродостојности за оцењивање параметара нормалне расподеле код узорка у којем је појављивање елемента испод одређене вредности забележено али појединачне вредности тих елемената нису одређене. Он их у том раду назива засеченим подацима, док Халд (*Hald (1949)*) истиче да те податке треба разликовати од засечених и назвати ,,цензурисаним". Сам појам је предложио Џон Е. Кирик. Дакле, главна разлика која се истиче између тих група података у првим радовима је да је у Стивенсовим подацима број неизмерених елемената познат, док тај податак није био познат у ранијим радовима. Након тога појмови засечених и цензурисаних података јављају се у многим радовима, а велики број радова објављених до средине прошлог века са том тематиком наведен је у *Mendenhall (1958)*. Данас, постоје многе монографије које се баве различитим облицима цензурисаних података, као и њиховом статистичком анализом (на пример *Andersen et al. (2012), Kalbfleisch and Prentice (2011), Fleming and Harrington (2011), Lawless (2011)*).

Засечени и цензурисани подаци се могу јавити у разним областима: индустрији, економији, инжењерству, контроли квалитета, медицини, као и свим природним наукама. Заправо, примери засечених података се најчешће јављају у индустрији када се узима узорак у производњи где су већ на почетку уклоњени елементи изнад и испод одређене вредности. Са друге стране, цензурисани подаци се најчешће јављају у проучавањима дужине живота и експериментима тренутка реаговања, где се често посматрање завршава пре него што дође до отказа или реакције свих опсервација. Већина метода који се примењују на овакве податке је заправо резултат природне модификације опште познатих техника за примену на непотпуни узорак.

У наставку ће бити размотрени цензурисани подаци и прилагођавање тестова уведених у претходним поглављима тако да се могу применити на такве податке.

5.1 Типови цензурисања

Као што је већ поменуто, честа је ситуација да је вредност приликом реализације неког догађаја на одређеним елементима позната само ако се налази у одређеном интервалу. Субјекти који су учествовали у истраживању, али чија мерења нису позната, се називају цензурисаним. Велики број различитих начина цензурисања се јавља у пракси, међутим они се могу груписати у категорије на основу неколико критеријума.

Прва подела, која је и најчешћа у литератури, је на податке

- цензурисане са десне стране,
- цензурисане са леве стране,
- двострано цензурисане.

У реалним ситуацијама најчешће се јављају подаци цензурисани са десне стране. Типични примери цензурисања са десне стране су управо они у којима се посматра време до појаве неког догађаја. У том случају код неких јединки појава циљног догађаја може бити спречена услед неког другог догађаја. На пример, ако се посматра ефикасност неког лека у смислу успоравања болести (одлагања смрти) може се десити да особа напусти истраживање или да премине од неке друге болести која није обухваћена истраживањем. Са друге стране, цензурисање са леве стране се јавља када се циљни догађај већ десио пре почетка посматрања одређене јединке. На пример, ако се посматра време када деца уче одређене способности (ход, говор и слично), може се десити да у истраживање уђе дете које је пре почетка истраживања већ имало развијену одређену способност. Овом типу цензурисања одговара и пример 5.2. Чест пример у пракси где се јавља цензурисање са леве стране, а који није повезан са временом до појаве неког догађаја, је када због немогућности мерног инструмента да измери сувише мале вредности се бележи доња мерна граница инстумента. Међутим, постоје и ситуације када се јавља цензурисање и са десне и са леве стране. У том случају се подаци називају двострано цензурисаним. Рецимо у примеру са развојем способности код деце, може се десити да неко дете пре краја проучавања неће развити неке способности. Уопштење сва три претходно наведена типа цензурисања је интервално цензурисање (за више детаља видети *Sun (2007)*).

Друга подела је урађена према томе на који начин се врши цензурисање на

- Тип I код којег се стварна вредност измерена на одговарајућој јединки бележи само ако се десила пре (односно после) унапред одређене вредности. Разлози за појављивање овог типа цензурисања су вишеструки, а неки од њих су немогућност да се истраживање почне баш у тренутку природног почетка (рецимо ако се посматра развој способности деце и истраживач није у могућности да посматра дете од самог рођења) или недовољно времена или средстава да се истраживање изведе до краја па се прекине у неком тренутку. Овом типу цензурисања одговара пример 1.1 из увода.

- Тип II код којег је фиксиран број највећих (односно најмањих) вредности цензурисан. Овај тип цензурисања се најчешће јавља код тестирања опреме. На пример, n истих уређаја је пуштено у рад и њихов рад се обуставља када њих r откаже. Овај начин вршења експеримента је посебно погодан због уштеде у времену. Треба истаћи да код овог типа цензурисања, број цензурисаних и број нецензурисаних елемената су фиксирани, док је време цензурисања случајно.

- Случајно цензурисање код којег су времена цензурисања случајне величине. Ако је време реализације циљног догађаја за неку јединку мање од вредности променљиве којом се цензурише, онда је тачно време реализације догађаја познато, док у супротном није. У том случају најчешће се претпоставља да је вероватноћа да се циљни догађај деси код оних елемената који су цензурисани иста као код оних који су остали нецензурисани. Ако размотримо раније наведени пример у којем се посматра ефикасност неког лека у смислу успоравања болести (одлагања времена смрти), у том случају наведено цензурисање, поред тога што је цензурисање са десне стране, је и случајно цензурисање.

Поред ова три типа цензурисања, постоје и неке хибридне варијанте цензурисања од којих се прва јавила у раду Епштајна (*Epstein (1954)*).

Размотримо цензурисање са десне стране. Аналогно се могу размотрити и друга два типа цензурисања. Основни проблем који се јавља је да се услед цензурисања са десне стране могу изменити карактеристике циљног догађаја. На пример, ако се у проучавању неке болести одређени пацијенти који се осећају посебно добро уклоне из истраживања, онда пацијенти који остају под ризиком нису више репрезентативни за узорак пацијената који би био да није било цензурисања. У том смислу, ако се изузму из истраживања такви пацијенти, то ће довести до тога да изгледа као да преостали пацијенти имају већу стопу смртности него да цензурисања није било. Механизам цензурисања који чува репрезентативност узорка у односу на оно што би био узорак да нема цензурисања се назива независним. Цензурисања типа I и типа II су такође и независна цензурисања. Случајно цензурисање код кога се претпоставља да су времена реализације циљног догађаја и цензурисања независне случајне величине је такође независно цензурисање и заправо представља јачи услов од независности цензурисања. Поставља се питање како у претходно наведеном примеру довести до тога да цензурисање буде независно. То је могуће урадити увођењем додатних информација путем променљивих, предиктора, као што су пол или старост особе приликом уласка у испитивање. Ако се испостави да је искључивање из истраживања (цензурисање) у свакој подгрупи случајно, онда је у питању независно цензурисање. Са друге стране, ако вероватноћа цензурисања у свакој подгрупи зависи од предиктора, цензурисање се назива зависним. Зависно цензурисање се често јавља у медицинским проучавањима услед различитих третмана пацијената (на пример, ако се једној групи пацијената даје плацебо а другима лек за одређену болест). У следећем поглављу ће бити дат формални начин испитивања независности цензурисања.

Разликоваћемо још два типа цензурисања, а то су неинформативно и информативно. Неинформативно цензурисање је оно код којег расподела реализације циљног догађаја не даје никакве информације о расподели цензурисања и обрнуто. Са друге стране, код информативног цензурисања расподеле реализације циљног догађаја и цензурисања су повезане на неки начин. Пример информативног цензурисања који се најчешће јавља у литератури је Козиол-Гринов модел (*Koziol and Green (1976)*) код којег су интензитети реализације циљног догађаја и цензурисања пропорционални.

У наставку, изузев ако није другачије наглашено, претпостављаћемо да је цензурисање са десне стране, случајно и неинформативно.

5.2 Основни појмови и модели за случајно цензурисане податке с десне стране

Два основна појма која се користе у анализи цензурисаних података су вероватноћа преживљавања и функција хазарда. Нека је T ненегативна случајна величина. С обзиром да у највећем броју случајева у пракси случајна величина T представља време од или до неког догађаја, она је непрекидна. Међутим, случајна величина T се може јавити и као дискретна случајна величина најчешће услед заокруживања мерења, груписања времена у интервале или ако мерења представљају број циклуса.

Нека је $F(t)$ функција расподеле случајне величине T. Вероватноћа да се догађај деси након тренутка t је одређена функцијом преживљавања дефинисаном са

$$S(t) = P\{T > t\} = \int_t^\infty dF(u).$$

Ако је случајна величина T непрекидна, функција преживљавања је непрекидна строго опадајућа функција и важи $S(t) = 1 - F(t)$, $S(0) = 1$ и $S(\infty) = \lim_{t\to\infty} S(t) = 0$. У случају дискретне случајне величине, функција преживљавања је нерастућа део по део константна функција, непрекидна с лева за коју важи да је $S(0) = 1$ и $S(\infty) = 0$.

Појам стопе хазарда се користи у многим областима где се користе и различити називи: интензитет смртности у демографији, функција интензитета у случајним процесима, условни степен отказа у теорији поузданости и слично. Стопа хазарда је дефинисана са

$$\lambda(t) = \lim_{dt\to 0} \frac{P\{t \leqslant T < t + dt | T \geqslant t\}}{dt}. \tag{5.1}$$

Она представља ризик за опстанак или отказ у тренутку t, оних јединки које су преживеле до тренутка t. Према томе $\lambda(t)dt$ представља приближну вероватноћу да се код јединке која је доживела време t, у следећем тренутку деси догађај који се прати.

Ако је T непрекидна случајна величина, тада је

$$\lambda(t) = -\frac{d \ln S(t)}{dt} = \frac{f(t)}{S(t)},$$

где је $f(t)$ одговарајућа густина расподеле случајне величине T. У дискретном случају

$$\lambda(t_j) = \frac{P\{T = t_j\}}{S(t_j)},\ j = 1, 2, ...$$

У оба случаја, одговарајућа функција хазарда Λ је дефинисана са

$$\Lambda(t) = -\ln S(t). \tag{5.2}$$

Једини услов за стопу хазарда је да мора бити ненегативна.

Пример 5.4 *Нека случајна величина T има експоненцијалну расподелу са параметром $\frac{1}{\mu}$. Тада, функција преживљавања ове случајне величине је $S(t) = e^{-\frac{1}{\mu}t}$, $t > 0$, $\mu > 0$, а стопа хазарда $\lambda(t) = \frac{1}{\mu}$.*

Нека је $X'_1, ..., X'_n$ узорак времена реализације циљног догађаја. Претпоставићемо да су у питању непрекидне случајне величине са функцијом расподеле F. Обично у анализи оваквих догађаја није познат потпун узорак, већ постоји неки други узорак $C_1, ..., C_n$ са функцијом расподеле G, који онемогућава праћење одређених елемената низа $X_1', .., X_n'$ до реализације циљног догађаја. Дакле, често је познат само узорак облика

$$(X_1, \delta_1), ..., (X_n, \delta_n), \tag{5.3}$$

где је, за свако $i = 1, ..., n$, $X_i = \min\{X_i', C_i\}$ и δ_i такозвани индикатор цензурисања који говори да ли је X_i једнако са X_i' или је познато само да је вредност од X_i мања од X_i', односно $\delta_i = I\{X_i' \leqslant C_i\}$. Ако још претпоставимо да су случајне величине X_i' и C_i независне, расподела случајне величине X_i чије су вредности познате је одређена са

$$\begin{aligned} H(x) = P\{X_1 \leqslant x\} &= P\{\min\{X_1', C_1\} \leqslant x\} \\ &= 1 - P\{\min\{X_1', C_1\} > x\} \\ &= 1 - P\{X_1' > x, C_1 > x\} \\ &= 1 - (1 - F(x))(1 - G(x)). \end{aligned} \tag{5.4}$$

Претходно описани модел одговара моделу случајног цензурисања са десне стране.

Често се узорак цензурисан са десне стране записује користећи случајне процесе. Основни непараметарски проблеми цензурисаних података се могу проучавати у терминима интензитета бројачког процеса који бележи нецензурисане догађаје као временске приносе. Први се овим приступом бавио Ален у својој дисертацији *Statistical inference for a family of counting processes (Aalen (1975))*, који је комбинујући стохастичку интеграцију, теорију мартингала и теорију бројачких процеса развио методологију којом се једноставно може доћи до закључака у вези са преживљавањем код цензурисаних и засечених података. Највећи део тих резултата као и других резултата из ове области могу се наћи у *Andersen et al. (2012)* или *Fleming and Harrington (2011)*.

Пре увођена самих процеса, дефинишемо филтрацију. Нека је $(\Omega, \mathcal{F}, P)$ простор вероватноћа. Филтрација $\{\mathcal{F}_t,\ t \geqslant 0\}$ је растућа непрекидна здесна фамилија $\sigma-$алгебри таквих да је $\mathcal{F}_t \subset \mathcal{F}$. Најприроднија филтрација је историја случајног процеса, јер је сваки процес мерљив у односу на своју историју. У складу са тим, у случају цензурисаних података, дефинишемо филтрацију $\{\mathcal{F}_t,\ t \geqslant 0\}$ тако да садржи вредности од X_i и δ_i за све i такве да је $X_i \leqslant t$, а иначе само даје информацију да је $X_i > t$. Може се дефинисати и такозвана ,,пре-t" $\sigma-$алгебра $\mathcal{F}_{t-}$ која је најмања $\sigma-$алгебра која садржи све $\mathcal{F}_s$, такве да је $s < t$.

За развој теорије цензурисаних података уз помоћ случајних процеса, претпоставља се само да је цензурисање независно, односно као што је наведено у претходном поглављу да у неком тренутку t преживљавање у будућности није статистички измењено цензурисањем (у односу на оно што би било без цензурисања) и преживљавањем у будућности. У општем случају овај услов је слабији од независности случајних величина X_i и C_i, за $i = 1, ..., n$, што се може видети на примеру цензурисања типа II код којег су случајне величине X_i и $C_i = X_{(r)}$ зависне али је цензурисање независно (за више детаља видети

Bagdonavicius et al. (2013)). Условом независног цензурисања захтева се да приближна вероватноћа да се у следећем тренутку деси циљни догађај код оних јединки које су доживеле време t иста као да цензурисања није било, односно

$$P\{X_i \in [t, t+dt), \delta_i = 1 | \mathcal{F}_t\} = \begin{cases} \lambda(t)dt + o_p(dt) & , \ X_i \geqslant t, \\ 0 & , \ X_i < t. \end{cases}$$

Ако се упореди добијени резултат са изразом у (5.1), види се да је функција хазарда остала непромењена услед цензурисања.

Дефинишемо бројачки процес $\{N(t),\ t \geqslant 0\}$, са

$$N(t) = \sum_{i=1}^{n} N_i(t) = \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant t, \delta_i = 1\}.$$

Процес $\{N(t),\ t \geqslant 0\}$ броји реализације циљног догађаја у узорку пре или у тренутку t. Ако се са $dN(t)$ означи прираштај $N((t+dt)-)-N(t-)$ процеса N током малог интервала времена $[t, t+dt)$, онда је његово очекивање

$$\begin{aligned} E(dN(t)|\mathcal{F}_{t-}) &= E\Big(\sum_{i=1}^{n} \mathrm{I}\{X_i \in [t, t+dt), \delta_i = 1 | \mathcal{F}_t\}\Big) \\ &= \sum_{i=1}^{n} \mathrm{I}\{X_i \geqslant t\}\lambda(t)dt \\ &= Y(t)\lambda(t)dt, \end{aligned} \tag{5.5}$$

где је $\{Y(t),\ t \geqslant 0\}$ процес дефинисан са $Y(t) = \sum\limits_{i=1}^{n} Y_i(t) = \sum\limits_{i=1}^{n} \mathrm{I}\{X_i \geqslant t\}$, број елемената који су под ризиком баш пред тренутак t. Процес $\{Y(t)\lambda(t),\ t \geqslant 0\}$ се назива интензитет бројачког процеса. Стопа хазарда $\lambda(t)$ даје условну просечну стопу промене у процесу N током кратког интевала времена, при услову да су и време реализације циљног догађаја и време цензурисања већи од времена t, па индиректно одређује условну вероватноћу са којом N прави скок у малом интервалу времена.

Означимо разлику између бројачког процеса и његовог кумулативног интензитета са M, односно

$$M(t) = N(t) - Y(t)\Lambda(t).$$

За фиксирано t, $Y(t)\Lambda(t)$ је случајна величина која даје приближни број скокова процеса N у интервалу $(0, t]$. Из претходног важи

$$dM(t) = dN(t) - Y(t)\lambda(t)dt. \tag{5.6}$$

Ако се размотри условно очекивање при филтрацији $\{\mathcal{F}_t,\ t \geqslant 0\}$, дефинисаној са

$$\mathcal{F}_t = \sigma(N_i(s), Y_i(s), 0 \leqslant s \leqslant t, i = 1, ..., n),$$

претходног израза, следи да је

$$E(dM(t)|\mathcal{F}_t) = E(dN(t) - Y(t)\lambda(t)dt|\mathcal{F}_t) = E(dN(t)|\mathcal{F}_t) - Y(t)\lambda(t)dt = 0,$$

при чему је коришћено да је $Y(t)\lambda(t)dt$ неслучајно у односу на филтрацију $\mathcal{F}_t$.

Случајни процес $M(t)$ је мартингал ако за свако $t \geqslant 0$ важи

1) $E(|M(t)|) < \infty$,

2) $E(M(t)|\mathcal{F}_s) = M(s)$, за све $s < t$.

Да је својство $E(dM(t)|\mathcal{F}_t) = 0$, за свако t, еквивалентно са условом 2) за мартингале следи из

$$\begin{aligned} E(M(t)|\mathcal{F}_s) - M(s) &= E(M(t) - M(s)|\mathcal{F}_s) \\ &= E\Big(\int_s^t dM(u)|\mathcal{F}_s\Big) \\ &= \int_s^t E(E(dM(u)|\mathcal{F}_{u-})|\mathcal{F}_s) = 0. \end{aligned}$$

Дакле, уместо записивања десно цензурисаног узорка као $(X_1, \delta_1), ..., (X_n, \delta_n)$, може се записати као $(N_1(t), Y_1(t), t \geqslant 0), ..., (N_n(t), Y_n(t), t \geqslant 0)$. Ако су познати (X_i, δ_i) онда се из дефиниција одговарајућих процеса могу одредити $(N_i(t), Y_i(t), t \geqslant 0)$. Са друге стране, ако су познати $(N_i(t), Y_i(t), t \geqslant 0)$, онда се X_i може одредити као тачка скока од Y_i, а ако N_i има скок у X_i онда је $X_i = X_i'$ и $\delta_i = 1$. Ако је $N_i(t) = 0$, за све $t \geqslant 0$, односно N_i нема скокове, онда је $X_i = C_i$ и $\delta_i = 0$. Предност записа узорка преко случајних процеса је у томе што процеси показују историју времена реализације догађаја и цензурисања током трајања експеримента.

Када су познати сви подаци, функција расподеле се може оценити емпиријском функцијом расподеле која свакој вредности додељује исту вероватноћу. У поглављу 2.1 показано је и зашто је та оцена погодна. На основу ње, могу се затим оценити и функција преживљавања и функција хазарда. Међутим, када постоје цензурисани подаци, те оцене нису адекватне, односно не треба све податке третирати као једнако вероватне, већ у разматрања треба увести и информацију о томе да ли је податак цензурисан.

Стандардна оцена која се користи за функцију преживљавања, па самим тим и функцију расподеле, за податке међу којима су присутни и они који су цензурисани са десне стране јесте Каплан-Мајерова оцена предложена у раду *Kaplan and Meier (1958)*. Ова оцена се добија дељењем временске осе у мање интервале, оцењивањем условне вероватноће преживљавања у сваком интервалу пропорцијом преживљавања, а затим множењем ових оцена, чиме се добија неусловна вероватноћа преживљавања. Односно, ако је дат узорак облика (5.3), онда је Каплан-Мајерова оцена функције преживљавања одређена са

$$\widehat{S}_{KM}(t) = \prod_{X_i \leqslant t} \left(1 - \frac{\delta_i}{\sum\limits_{k=1}^{n} I\{X_k \geqslant X_i\}}\right). \tag{5.7}$$

Ако нема цензурисаних вредности, односно $\delta_i = 1$, за све $i = 1, ..., n$, ова оцена је заправо емпиријска функција расподеле.

Проблем код ове оцене се јавља када је највећа вредност цензурисана, односно када је највеће посматрано време цензурисано. Сами аутори рада *Kaplan and Meier (1958)* наглашавају да у том случају не би требало користити наведену формулу и да је вредност од $\widehat{S}_{KM}(t)$ негде између 0 и $\widehat{S}_{KM}(X_{(n)})$ али да се не зна њена тачна вредност. Ефрон (*Efron (1967)*) је модификовао ову оцену тако да њена вредност буде једнака нули након времена $X_{(n)}$ чак и када је тај елемент цензурисан.

Са друге стране, оцена функције преживљавања се може записати и коришћењем бројачког процеса. У том случају се користе веза између функције хазарда и функције преживљавања (5.2). На овај начин се могу добити две оцене које су асимптотски еквивалентне. У сваком случају, полази се од Нелсон-Аленове оцене функције хазарда

$$\widehat{\Lambda}(t) = \int_0^t \frac{dN(s)}{Y(s)}. \tag{5.8}$$

Код података цензурисаних са десне стране, бројачки процес $\{N(t), t \geqslant 0\}$ даје тежину 1 сваком времену у којем се десио скок, односно није било цензурисања, а 0 иначе. У том случају оцена (5.8) се може написати као

$$\widehat{\Lambda}(t) = \sum_{X_i \leqslant t} \frac{\Delta N(X_i)}{Y(X_i)}.$$

Ова оцена је предложена од стране Алена (*Aalen (1978)*) и представља уопштење независно предложених оцена Нелсона (*Nelson (1969)*) и Алтшулера (*Altshuler (1970)*). Најчешће се ова оцена назива Нелсон-Аленовом оценом. Нелсон је оцену увео као графичку технику за функцију хазарда цензурисаних података како би се добила информација о облику расподеле времена отказа. Према томе једна од могућих примена Нелсон-Аленове оцене је графичка провера да ли случајне величине X_i припадају одређеној параметарској фамилији расподела.

Увођењем Нелсон-Аленове оцене у израз

$$S(t) = 1 - \int_0^t S(s-) d\Lambda(s),$$

добија се

$$\widehat{S}_G(t) = \prod_{X_i \leqslant t} \left(1 - \frac{\Delta N(X_i)}{Y(X_i)}\right) \tag{5.9}$$

(за више детаља видети *Fleming and Harrington (2011)*). Ову оцену је први пут увео Гил (*Gill (1980)*). Оцена (5.9) је једнака са (5.7) за $t < X_{(n)}$ и обе су једнаке нули након $X_{(n)}$ када тај елемент узорка није цензурисан. Када је највећа вредност узорка цензурисана, онда је $\widehat{S}_G(t) = \widehat{S}_G(X_{(n)})$ за $t > X_{(n)}$. Дакле, наведене оцене се разликују само када је највећа вредност у узорку цензурисана.

Када се упореде пристрасности и средње квадратне грешке испоставља се да је Гилова модификација Каплан-Мајерове оцене боља за велике и умерене вредности функције преживљавања у крајњој тачки, док је Ефронова модификација боља за мале вредности функције преживљавања (за више детаља видети *Geurts (1987)* или *Klein (1991)*). Међутим, у практичним применама, разлика између ове две модификације није од великог значаја.

Са друге стране, Нелсон-Аленова оцена се може користити и у изразу $S(t) = e^{-\Lambda(t)}$. У том случају оцена функције преживљавања је одређена са

$$\widehat{S}_{NA}(t) = e^{-\widehat{\Lambda}(t)} = \prod_{X_i \leqslant t} e^{-\frac{\Delta N(X_i)}{Y(X_i)}}. \tag{5.10}$$

Ова оцена се најчешће назива Нелсон-Аленовом функцијом преживљавања. Када се ради са малим узорцима, Каплан-Мајерова оцена функције преживљавања (5.7) има мање вредности него Нелсон-Аленова оцена преживљавања (5.10). Међутим, за велике обиме узорка, обе оцене су конвергентне и њихова разлика је занемарљива. За више детаља о поређењу оцена видети *Murray and Sandercock (2020)* или *Colosimo et al. (2002)*.

Особине Каплан-Мајерове оцене у случају великих узорака први пут су проучаване у раду *Efron (1967)* и касније у *Breslow and Crowley (1974)*. Гил (*Gill (1983)*) је први користио бројачке процесе како би доказао униформну постојаност и асимптотску расподелу Каплан-Мајерове оцене. Следеће две теореме дају унапређене Гилове резултате, приказане у радовима *Wang (1987)* и *Ying (1989)* редом. У оба случаја претпоставља се да су случајне величине C_i и X_i' независне, при чему је расподела података цензурисаних са десне стране одређена са (5.4). Такође, $\widehat{F}(t) = 1 - \widehat{S}(t)$.

Теорема 5.1 *Нека је* $\tau_H = \sup_t \{H(t) < 1\}$. *Тада важи*

$$\sup_{t < \tau_H} |\widehat{F}(t) - F(t)| \xrightarrow{P} 0,$$

$$\sup_{t \leqslant X_{(n)}} |\widehat{F}(t) - F(t)| \xrightarrow{P} 0,$$

кад $n \to \infty$.

Теорема 5.2 *Нека је* $t \in \tau_H$, *где је* $\tau_H = \sup_t \{H(t) < 1\}$. *Нека је*

$$Q(t) = \int_0^t \frac{d\Lambda(s)}{1 - H(s-)}$$

и

$$J(t)=\frac{Q(t)}{1+Q(t)}.$$

Тада, кад $n\to\infty$,

$$\sqrt{n}\frac{1-J}{1-F}(\widehat{F}_n-F)\xrightarrow{D}Z_1 \text{ у } D[0,\tau_H], \tag{5.11}$$

īде је Z_1 *Гаусов īроцес са коваријационом функцијом*

$$cov(Z_1(t_1),Z_1(t_2))=J(t_1\wedge t_2)(1-J(t_1\vee t_2)).$$

Додаѿно, ако важи $\int_0^{\tau_H}\frac{dF(t)}{1-G(t-)}<\infty$, *онда*

$$\sqrt{n}(\widehat{F}_n-F)\xrightarrow{D}Z_2 \text{ у } D[0,\tau_H], \text{ кад } n\to\infty,$$

īде је Z_2 *Гаусов īроцес са коваријационом функцијом*

$$cov(Z_2(t_1),Z_2(t_2))=(1-F(t_1))(1-F(t_2))Q(t_1\wedge t_2),$$

и

$$\sqrt{n}\frac{1-\widehat{J}_n}{1-\widehat{F}_n}(\widehat{F}_n-F)\xrightarrow{D}Z_1 \text{ у } D[0,\tau_H], \text{ кад } n\to\infty,$$

īде је $\widehat{J}_n(t)=\frac{\widehat{Q}_n(t)}{1+\widehat{Q}_n(t)}$ *и* $\widehat{Q}_n(t)=\int_0^t\frac{ndN(s)}{Y(s)(Y(s)-1)}$.

Процес Z_1 из претходне теореме је заправо Браунов мост са наведеном коваријационом функцијом. У раду *Hall and Wellner (1980)* показано је да се (5.11) може еквивалентно записати као

$$\sqrt{n}\frac{\widehat{F}_n-F}{1-F}\xrightarrow{D}B(Q) \text{ у } D[0,\tau_H],$$

кад $n\to\infty$, где је $B(Q)$ Брауново кретање са коваријационом функцијом

$$cov(B(Q(t_1)),B(Q(t_2)))=Q(t_1\wedge t_2).$$

У наставку, у теоријским деловима ће се користити Каплан-Мајерова оцена одређена са (5.9), без експлицитног навођења индекса G, док у практичним применама вредности обе варијанте Каплан-Мајерове оцене ће бити једнаке јер ће се вредности рачунати за t такво да је $t\leqslant X_{(n)}$.

5.3 $V-$статистике скалиране инверзом вероватноће цензурисања

Основне статистичке анализе углавном третирају све елементе узорка једнако важним. Међутим, у одређеним ситуацијама постоји потреба да се одређеним елементима додели већа вероватноћа избора него осталим. То се обично постиже увођењем одговарајућих тежина.

Постоје случајеви у којима само истраживање захтева узорковање у којем није иста вероватноћа избора сваког елемента. На пример, ако се проучава одређена болест која погађа одређени део популације (хипертензија се чешће јавља код гојазних људи), истраживач жели да већи део његовог узорка припада том делу популације. У том случају би стандардне (неотежане) вредности оцена на основу узорка биле пристрасне оцене теоријских величина. Ова пристрасност избора се може елиминисати тако што се користи тежинска оцена, која сваком елементу узорка даје тежину једнаку инверзу вероватноће избора. Тиме се ублажава тежина оних јединки које су чешће узорковане и формално се ствара проучавање у којем нема разлике у избору елемената. Још неки примери у којима се користи скалирање одређеним тежинама се могу наћи на пример у *Mansournia and Altman (2016)*.

Са друге стране, сличан принцип се може користити и када постоје цензурисани подаци у узорку. У том случају се елементи узорка разликују у вероватноћи да не дају потпуну информацију, односно да буду цензурисани. У овим ситуацијама ће стандардне оцене добијене на основу постојећег узорка такође бити пристрасне. Због тога се може применити скалирање инверзом вероватноће да елемент узорка не буде цензурисан. Модел скалирања инверзом вероватноће цензурисања смањује утицај цензурисаних података у узорку који је цензурисан са десне стране тако што даје додатну тежину подацима који нису цензурисани. На овај начин се може редуковати пристрасност оцене.

Овај метод је први пут за цензурисане податке употребљен за зависно цензурисане податке у Коксовом моделу (*Robins and Rotnitzky (1992)*). У наставку ће модел бити примењен како би се тестови представљени у поглављу 3 прилагодили за примену на податке цензурисане са десне стране. С обзиром да се и тестови L^2-типа и тестови $L^\infty-$типа могу представити као одговарајуће трансформације $V-$емпиријских процеса, прво ћемо размотрити $V-$емпиријске процесе трансформисане уз помоћ модела скалирања инверзом.

Претпоставимо да је познат узорак облика $(X_1, \delta_1), ..., (X_n, \delta_n)$, описан у претходном поглављу, који одговара случајном цензурисању са десне стране. Такође, претпостављамо да су одговарајуће расподеле одређене са F и G непрекидне. На основу овог узорка дефинишемо одговарајуће $V-$емпиријске процесе

$$V_{n,c}(t) = \frac{1}{n^m} \sum_{i_1,...,i_m=1}^{n} \frac{\Phi(X_{i_1}, ..., X_{i_m}; t) \prod\limits_{k=1}^{m} \delta_{i_k}}{\prod\limits_{k=1}^{m} K_c(X_{i_k}-)}, \tag{5.12}$$

где је $\Phi(x_1, ..., x_m; t)$ симетрична функција по својим аргументима, а $K_c(x) = 1 - G(x)$ функција преживљавања цензора таква да је $K_c(x) > 0$, за свако x, са вероватноћом 1. С обзиром на претпоставку да су цензорске величине непрекидне, јасно је да

$K_c(x) = K_c(x-)$, међутим због каснијег увођења одговарајућих оцена наглашава се да је у питању вредност слева.

Приметимо да у случају нецензурисаног узорка, односно када је $\delta_i = 1$ и $K_c(X_i-) = 1$, за све i, $\{V_{n,c}(t), t \geqslant 0\}$ је V-емпиријски процес са језгром $\Phi(X_1', ..., X_m'; t)$ чија је функција средње вредности $\theta(t) = E(\Phi(X_1', ..., X_m'; t))$.

Означимо са

$$\widetilde{\Phi}((X_1, \delta_1), ..., (X_m, \delta_m); t) = \frac{\Phi(X_1, ..., X_m; t) \prod\limits_{k=1}^{m} \delta_k}{\prod\limits_{k=1}^{m} K_c(X_k-)}, \tag{5.13}$$

језгро наведеног процеса. Може се приметити да овакав модел чува средњу вредност језгра, односно важи

$$\begin{aligned} E(\widetilde{\Phi}((X_1, \delta_1), ..., (X_m, \delta_m); t)) &= E\bigg(\frac{\Phi(X_1, ..., X_m; t) \prod_{k=1}^{m} \delta_k}{\prod_{k=1}^{m} K_c(X_k-)}\bigg) \\ &= E\bigg(E\bigg(\frac{\Phi(X_1', ..., X_m'; t) \prod_{k=1}^{m} I\{X_k' \leqslant C_k\}}{\prod_{k=1}^{m} K_c(X_k'-)}\bigg| X_1', ..., X_m'\bigg)\bigg) \\ &= E\bigg(\frac{\Phi(X_1', ..., X_m'; t)}{\prod_{k=1}^{m} K_c(X_k'-)} \prod_{k=1}^{m} K_c(X_k'-)\bigg) \\ &= E\bigg(\Phi(X_1', ..., X_m'; t)\bigg) = \theta(t). \end{aligned}$$

У наставку ћемо претпостављати да је $\theta(t) = 0$, за свако t, пошто је то случај који се јавља у статистикама размотреним у поглављу 3 када важи нулта хипотеза.

Проблем код наведених емпиријских процеса се јавља јер код неинформативног цензурисања функција преживљавања цензора није позната. Међутим, она се може заменити одговарајућом Каплан-Мајеровом оценом функције преживљавања цензора, у којој цензурисани и нецензурисани подаци мењају места, чиме се добија

$$\widehat{V}_{n,c}(t) = \frac{1}{n^m} \sum_{i_1,...,i_m=1}^{n} \frac{\Phi(X_{i_1}, ..., X_{i_m}; t) \prod\limits_{k=1}^{m} \delta_{i_k}}{\prod\limits_{k=1}^{m} \widehat{K}_c(X_{i_k}-)}, \tag{5.14}$$

где је $\widehat{K}_c(x-) = \prod\limits_{X_i<x} \Bigg(1 - \frac{1-\delta_i}{\sum\limits_{k=1}^{n} I\{X_k \geqslant x\}}\Bigg)$. Треба нагласити да овако добијени процеси више нису $V-$емпиријски процеси, јер је уведена Каплан-Мајерова оцена заснована на свим елементима узорака.

Следећа теорема даје асимптотска својства процеса дефинисаног у (5.14) у случају када је функција Φ непрекидна на $[0, \infty)$. Пре увођења теореме треба нагласити да, слично дефиницији бројачког процеса $N(t)$ из претходног поглавља, може се дефинисати бројачки процес који одговара цензурисаним елементима, односно

$$N_i^c(t) = \mathrm{I}\{X_i \leqslant t, \delta_i = 0\}.$$

Њему одговарајући мартингал је $M_i^c(t) = N_i^c(t) - \int_0^t Y_i d\Lambda_c(u)$, у односу на филтрацију $\mathcal{F}_t = \sigma(N_i^c(s), Y_i(s), 0 \leqslant s \leqslant t, i = 1, ..., n)$, где је $Y_i = \mathrm{I}\{X_i \geqslant t\}$ дефинисано као у претходном поглављу, односно не зависи од тога да ли је елемент цензурисан или не, а Λ_c функција хазарда цензора.

Теорема 5.3 *Нека $X_1', ..., X_n'$ узорак из ненегативне расподеле F и $(X_1, \delta_1), ..., (X_n, \delta_n)$ одговарајући узорак случајно цензурисан са десне стране. Претпоставимо да је $\Phi(\cdot; t)$ функција из $C[0, \infty)$ таква да је, за фиксирано t, $E\Phi(X_1', ..., X_m'; t) = 0$ и да за неке позитивне константе A_1, A_2 и A_3 важи*

(1) *$E\widetilde{\Phi}^2((X_{i_1}, \delta_{i_1}), ..., (X_{i_m}, \delta_{i_m}); t) < A_1 < \infty, 1 \leqslant i_1, ..., i_m \leqslant n$, где је $\widetilde{\Phi}$ дефинисано у (5.13);*

(2) *$\int\limits_0^\infty \frac{\varphi_1^2(x;t)}{K_c(x-)} dF(x) < A_2 < \infty$, где је $\varphi_1(x; t) = E(\Phi(X_1', ..., X_m'; t) | X_1' = x)$ прва пројекција језгра Φ;*

(3) *$\int\limits_0^\infty \omega^2(u; t) y(u) \lambda_c(u) du < A_3 < \infty$, где је $\omega(u; t) = \frac{1}{y(u)} \int\limits_u^\infty \varphi_1(x; t) dF(x)$ и $y(u) = P\{X_1 \geqslant u\}$, $u \geqslant 0$;*

(4) *$\max\limits_{0 \leqslant t \leqslant T} E\left(\left(\frac{\partial \varphi_1(X_1'; t)}{\partial t}\right)^2 \frac{1}{K_c(X_1'-)}\right) < \infty$, за свако $T \geqslant 1$;*

(5) *$\max\limits_{0 \leqslant t \leqslant T} \int\limits_0^\infty \left(\int\limits_u^\infty \left|\frac{\partial \varphi_1(x; t)}{\partial t}\right| dF(x)\right)^2 \frac{g(u)}{P\{X \geqslant u\} K_c(u)} du < \infty$, за свако $T \geqslant 1$.*

Тада $\{\sqrt{n}\widehat{V}_{n,c}(t), t \geqslant 0\}$ конвергира у $C[0, \infty)$ ка центрираном Гаусовом процесу $\{\eta(t), t \geqslant 0\}$ са коваријационом функцијом

$$cov(\eta(t_1), \eta(t_2)) = m^2 E\left(\left(\frac{\varphi_1(X_1; t_1)\delta_1}{K_c(X_1-)} + \int\limits_0^\infty \omega(u; t_1) dM_1^c(u)\right)\left(\frac{\varphi_1(X_1; t_2)\delta_1}{K_c(X_1-)} + \int\limits_0^\infty \omega(u; t_2) dM_1^c(u)\right)\right), \tag{5.15}$$

за $t_1, t_2 \in [0, \infty)$.

Доказ: Циљ је показати да се процес $\sqrt{n}\widehat{V}_{c,n}(t)$ може приказати као

$$\sqrt{n}\widehat{V}_{n,c}(t) = \frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n(t) + \sqrt{n} R_n(t), \tag{5.16}$$

где $\{\mathcal{Z}_n(t), t \geqslant 0\}$ конвергира ка центрираном Гаусовом процесу и $\sqrt{n}R_n(t)$ конвергира униформно ка 0. Доказ се састоји из три дела

I $\sqrt{n}\widehat{V}_c(t)$ представити у облику (5.16) где је $\mathcal{Z}_n(t)$ сума независних и једнако расподељених случајних величина чији сабирци не зависе од $\widehat{K}_c$;

II показати да процес $\{\mathcal{Z}_n(t), t \geqslant 0\}$ конвергира ка центрираном Гаусовом процесу користећи теорему 2.12, односно показати да

(а) коначнодимензионалне расподеле од $\mathcal{Z}_n(t)$ конвергирају ка вишедимензионалној нормалној расподели,

(б) $\{\mathcal{Z}_n(t), t \geqslant 0\}$ је густ;

III применити теорему Слуцког (теорема 2.14).

Корак I:

С обзиром да $\widehat{V}_{n,c}(t) = V_{n,c}(t) + \widehat{V}_{n,c}(t) - V_{n,c}(t)$, где је $V_{n,c}(t)$ дефинисано у (5.12), желимо да пронађемо одговарајуће репрезентације за $\sqrt{n}V_{n,c}(t)$ и $\sqrt{n}(\widehat{V}_{n,c}(t) - V_{n,c}(t))$.

Користећи везу (2.5) између $U-$ и $V-$статистика, за фиксирано t, познато је да

$$V_{n,c}(t) = \frac{\binom{n}{m}m!}{n^m}U_{n,c}(t) + \frac{1}{n^m}\sum_{(i_1,\dots,i_m)} \widetilde{\Phi}((X_{i_1}, \delta_{i_1}), \dots, (X_{i_m}, \delta_{i_m}); t), \tag{5.17}$$

где је сума узета по свим $m-$торкама $(i_1, \dots, i_m)$ таквим да се бар два индекса поклапају и $U_{n,c}(t)$ одговарајући $U-$емпиријски процес. Из услова (1) можемо закључити да други сабирак у изразу (5.17) униформно конвергира ка 0. Први члан израза (5.17), на основу (2.10), се може записати као

$$\sqrt{n}U_{n,c}(t) = \frac{m}{\sqrt{n}}\sum_{i=1}^{n} \widetilde{\varphi}_1(X_i, \delta_i; t) + \sqrt{n}R_n'(t), \tag{5.18}$$

где је $\widetilde{\varphi}_1$ прва пројекција језгра $\widetilde{\Phi}$, односно

$$\begin{aligned}
\widetilde{\varphi}_1(x_1, \delta_1; t) &= E(\widetilde{\Phi}((X_1, \delta_1), \dots, (X_m, \delta_m); t)|X_1 = x_1, \delta_1) \\
&= E\left(\frac{\Phi(X_1, X_2, \dots, X_m; t)\prod_{k=1}^m \delta_k}{\prod_{k=1}^m K_c(X_k-)}\middle| X_1 = x_1, \delta_1\right) \\
&= \frac{\delta_1}{K_c(x_1-)}E\Big(\frac{\Phi(x_1, X_2, \dots, X_m; t)\prod_{k=2}^m \delta_k}{\prod_{k=2}^m K_c(X_k-)}\Big) \\
&= \frac{\delta_1}{K_c(x_1-)}E\left(\frac{\Phi(x_1, X_2', \dots, X_m'; t)}{\prod_{k=2}^m K_c(X_k'-)}\prod_{k=2}^m E(I\{X_k' \leqslant C_k\}|X_k')\right) \\
&= \frac{\delta_1}{K_c(x_1-)}E\left(\frac{\Phi(x_1, X_2', \dots, X_m'; t)}{\prod_{k=2}^m K_c(X_k'-)}\prod_{k=2}^m K_c(X_k'-)\right) \\
&= \frac{\delta_1}{K_c(x_1-)}E(\Phi(x_1, X_2', \dots, X_m'; t)) = \frac{\varphi_1(x_1; t)\delta_1}{K_c(x_1-)}.
\end{aligned}$$

Тада

$$R_n'(t) = U_{n,c}(t) - \frac{m}{n}\sum_{i=1}^{n}\widetilde{\varphi}_1(X_i, \delta_i; t) = \sum_{r=2}^{n}\frac{\binom{m}{r}}{\binom{n}{r}}\sum_{1\leqslant i_1<\dots<i_r\leqslant n} g_r(X_{i_1}, \dots, X_{i_r}; t),$$

где је

$$g_r((x_1,\delta_1),...,(x_r,\delta_r);t)=\widetilde{\varphi}_r((x_1,\delta_1),...,(x_r,\delta_r);t)-\sum_{i=1}^{r}g_1(x_i,\delta_i;t)$$
$$-\sum_{1\leqslant i_1<i_2\leqslant r}g_2((x_{i_1},\delta_1),(x_{i_2},\delta_2);t)-\cdots-\sum_{1\leqslant i_1<\cdots<i_{r-1}\leqslant r}g_{r-1}((x_{i_1},\delta_1),...,(x_{i_{r-1}},\delta_{r-1});t)$$

и $\widetilde{\varphi}_r((x_1,\delta_1),...,(x_r,\delta_r);t)=E(\widetilde{\Phi}((X_1,\delta_1),...,(X_m,\delta_m);t)|(X_1=x_1,\delta_1),...,(X_r=x_r,\delta_r))$ r-та пројекција језгра $\widetilde{\Phi}$. Према томе, користећи (2.9),

$$E(R_n'(t))^2=\sum_{c=2}^{n}\frac{\binom{m}{r}^2}{\binom{n}{r}}Eg_r^2(X_1,...,X_r;t)=O_p\left(\frac{1}{n^2}\right),\ n\to\infty,$$

ако је $E(\widetilde{\Phi}((X_1,\delta_1),...,(X_m,\delta_m);t))^2<\infty$, што следи из услова (1) теореме. Дакле, $\sqrt{n}R_n'(t)$ конвергира униформно ка 0. Комбиновањем овог резултата са (5.17) и (5.18) добија се

$$\sqrt{n}V_{n,c}(t)=\frac{\binom{n}{m}m!}{n^m}\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{\varphi}_1(X_i,\delta_i;t)+\sqrt{n}\widetilde{R}_n(t),\tag{5.19}$$

где $\sqrt{n}\widetilde{R}_n(t)$ униформно конвергира ка 0.

Са друге стране

$$\sqrt{n}(\widehat{V}_{n,c}(t)-V_{n,c}(t))=\frac{\sqrt{n}}{n^m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\left(\prod_{k=1}^{m}\frac{\delta_{i_k}}{\widehat{K}_c(X_{i_k}-)}-\prod_{k=1}^{m}\frac{\delta_{i_k}}{K_c(X_{i_k}-)}\right)$$
$$=\frac{\sqrt{n}}{n^m}\sum_{k=1}^{m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\prod_{j=1}^{m}\delta_{i_j}\frac{\widehat{K}_c(X_{i_k}-)-K_c(X_{i_k}-)}{\prod\limits_{l\leqslant k}\widehat{K}_c(X_{i_l}-)\prod\limits_{l'\geqslant k}K_c(X_{i_{l'}}-)}.\tag{5.20}$$

У имениоцу претходног израза, $\widehat{K}_c$ се може заменити њеном граничном вредношћу K_c. Грешка која се том приликом прави је

$$\sqrt{n}R_n''(t)=\frac{\sqrt{n}}{n^m}\sum_{k=1}^{m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}\frac{\widehat{K}_c(X_{i_k}-)-K_c(X_{i_k}-)}{\prod\limits_{l\leqslant k}\widehat{K}_c(X_{i_l}-)\prod\limits_{l'\geqslant k}K_c(X_{i_{l'}}-)}$$
$$-\frac{\sqrt{n}}{n^m}\sum_{k=1}^{m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}\frac{\widehat{K}_c(X_{i_k}-)-K_c(X_{i_k}-)}{K_c^2(X_{i_k}-)\prod\limits_{l\neq k}K_c(X_{i_l}-)}$$
$$=\frac{\sqrt{n}}{n^m}\sum_{k=1}^{m}\sum_{i_1,...,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}(\widehat{K}_c(X_{i_k}-)-K_c(X_{i_k}-))$$
$$\times\frac{\prod\limits_{l\leqslant k}K_c(X_{i_l}-)-\prod\limits_{l\leqslant k}\widehat{K}_c(X_{i_l}-)}{\prod\limits_{l\geqslant k}K_c(X_{i_l}-)\prod\limits_{l'\leqslant k}K_c(X_{i_{l'}}-)\widehat{K}_c(X_{i_{l'}}-)}$$

$$= \frac{\sqrt{n}}{n^m}\sum_{k=1}^{m}\sum_{i_1,\dots,i_m=1}^{n}\Phi(X_{i_1},\dots,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}\frac{\widehat{K}_c(X_{i_k}-)-K_c(X_{i_k}-)}{K_c(X_{i_k}-)}\sum_{j=1}^{k}\frac{K_c(X_{i_j}-)-\widehat{K}_c(X_{i_j}-)}{\prod\limits_{l\geqslant j}K_c(X_{i_l}-)\prod\limits_{l'\leqslant j}\widehat{K}_c(X_{i_{l'}}-)}.$$

У циљу да се покаже да $\sqrt{n}R''_n(s)$ униформно конвергира ка 0, довољно је показати да важи

$$\int\limits_{\mathbb{R}} B_n(x_1,...,x_m)\frac{\Phi(x_1,...,x_m;t)}{\prod\limits_{k=1}^{n}\widehat{K}_c(x_i-)}\prod_{i=1}^{n}dW_n(x_i),$$

где је

$$B_n(x_1,...,x_m)=\sqrt{n}\sum_{k=1}^{m}\frac{\widehat{K}_c(x_k-)-K_c(x_k-)}{K_c(x_k-)}\sum_{j=1}^{k}\frac{K_c(x_j-)-\widehat{K}_c(x_j-)}{\prod\limits_{l\geqslant j}K_c(x_l-)}\prod_{l'\geqslant j}\widehat{K}_c(x_{l'}-)$$

и W_n је емпиријска расподела X_i-ева који су нецензурисани, односно за које је $\delta_i=1$. Тада, применом Коши-Шварцове неједнакости, се добија

$$\begin{aligned}|\sqrt{n}R''_n(t)| &\leqslant \Big(\int\limits_{\mathbb{R}} B_n^2(x_1,...,x_m)\prod_{i=1}^{m}dW_n(x_i)\Big)^{\frac{1}{2}}\Big(\int\limits_{\mathbb{R}}\frac{\Phi^2(x_1,...,x_m;t)}{\prod\limits_{i=1}^{m}\widehat{K}_c^2(x_i-)}\prod_{i=1}^{m}dW_n(x_i)\Big)^{\frac{1}{2}}\\ &\leqslant \sup_{x_1,\dots,x_m\geqslant 0}|B_n^2(x_1,...,x_m)|\Big(\frac{1}{n^m}\sum_{i_1,\dots,i_m=1}^{n}\frac{\Phi^2(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}}{\prod\limits_{k=1}^{m}\widehat{K}_c^2(X_{i_k}-)}\Big)^{\frac{1}{2}}\end{aligned}$$

Процес B_n слабо конвергира нула процесу на основу теореме 5.2 и

$$\frac{1}{n^m}\sum_{i_1,\dots,i_m=1}^{n}\frac{\Phi^2(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}}{\prod\limits_{k=1}^{m}\widehat{K}_c^2(X_{i_k}-)}\xrightarrow{P}E\left(\frac{\Phi^2(X_1,...,X_m;t)\delta_1\cdots\delta_m}{\prod\limits_{k=1}^{m}K_c^2(X_k-)}\right)<A_1.$$

Дакле, $\sqrt{n}R''_n(t)$ конвергира униформно ка 0.

Користећи теорему о средњој вредности, следи да за свако u

$$\begin{aligned}\sqrt{n}(\widehat{K}_c(u-)-K_c(u-)) &= \sqrt{n}(e^{-\widehat{\Lambda}_c(u-)}-e^{-\Lambda_c(u-)})+o_p(1)\\ &= \sqrt{n}K_c(u-)(\Lambda_c(u-)-\widehat{\Lambda}_c(u-))+o_p(1),\ n\to\infty,\end{aligned} \tag{5.21}$$

где је $\widehat{\Lambda}_c$ Нелсон-Аленова оцена функције хазарда Λ_c дефинисана у (5.8). Заменом израза (5.21) у (5.20), следи да

$$\sqrt{n}(\widehat{V}_{n,c}(t)-V_{n,c}(t))=\frac{\sqrt{n}}{n^m}\sum_{i_1,\dots,i_m=1}^{n}\Phi(X_{i_1},...,X_{i_m};t)\frac{\widehat{\Lambda}_c(X_{i_k}-)-\Lambda_c(X_{i_k}-)}{\prod_{k=1}^{m}K_c(X_{i_k}-)}\prod_{k=1}^{m}\delta_{i_k}+\sqrt{n}R'''_n(t), \tag{5.22}$$

где $\sqrt{n}R'''_n(t)$ конвергира униформно ка 0.

Поново користећи везу између $U-$ и $V-$статистика, као и $U-$статистике и прве пројекције језгра (2.10), претходни израз се може апроксимирати са

$$\frac{\binom{n}{m}m!}{n^m}\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}(\widehat{\Lambda}_c(X_i-)-\Lambda(X_i-)). \tag{5.23}$$

Грешка која се прави приликом апроксимације $U-$статистике је

$$\begin{aligned}
&\Bigg|\frac{\sqrt{n}}{\binom{n}{m}}\sum_{1\leqslant i_1<\ldots<i_m\leqslant n}\frac{\Phi(X_{i_1},...,X_{i_m};t)\prod\limits_{k=1}^{m}\delta_{i_k}}{\prod\limits_{k=1}^{m}K_c(X_{i_k}-)}(\widehat{\Lambda}_c(X_{i_k}-)-\Lambda_c(X_{i_k}-))\\
&\qquad-\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}(\widehat{\Lambda}_c(X_i-)-\Lambda_c(X_i-))\Bigg|\\
&\leqslant m\Bigg|\frac{1}{n\binom{n-1}{m-1}}\sum_{i_1<\ldots<i_{m-1}\in\{1,\ldots,n\}\setminus j}\frac{\Phi(X_j,X_{i_1},...,X_{i_{m-1}};t)\delta_j\prod\limits_{k=1}^{m-1}\delta_{i_k}}{K_c(X_j-)\prod\limits_{k=1}^{m-1}K_c(X_{i_k}-)}-\frac{1}{n}\sum_{i=1}^{n}\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}\Bigg|\\
&\qquad\times\sup_{u\geqslant 0}\sqrt{n}|\widehat{\Lambda}_c(u)-\Lambda_c(u)|\\
&\leqslant m\Bigg|\frac{1}{n}\sum_{j=1}^{n}\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{i_1<\ldots<i_{m-1}\in\{1,\ldots,n\}\setminus j}\frac{\Phi(X_j,X_{i_1},...,X_{i_{m-1}};t)\delta_j\prod\limits_{k=1}^{m-1}\delta_{i_k}}{K_c(X_j-)\prod\limits_{l=1}^{m-1}K_c(X_{i_l}-)}-\frac{\varphi_1(X_j;t)\delta_j}{K_c(X_j-)}\Bigg)\Bigg|\\
&\qquad\times\sup_{u\geqslant 0}\sqrt{n}|\widehat{\Lambda}_c(u)-\Lambda_c(u)|\\
&\leqslant m\Bigg(\frac{1}{n}\sum_{j=1}^{n}\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{i_1<\ldots<i_{m-1}\in\{1,\ldots,n\}\setminus j}\frac{\Phi(X_j,X_{i_1},...,X_{i_{m-1}};t)\delta_j\prod\limits_{k=1}^{m-1}\delta_{i_k}}{K_c(X_j-)\prod\limits_{l=1}^{m-1}K_c(X_{i_l}-)}-\frac{\varphi_1(X_j;t)\delta_j}{K_c(X_j-)}\Bigg)^2\Bigg)^{\frac{1}{2}}\\
&\qquad\times\sup_{u\geqslant 0}\sqrt{n}|\widehat{\Lambda}_c(u)-\Lambda_c(u)|.
\end{aligned}\tag{5.24}$$

За први члан претходног израза важи

$$\begin{aligned}
&E\Bigg(\frac{1}{n}\sum_{j=1}^{n}\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{i_1<\ldots<i_{m-1}\in\{1,\ldots,n\}\setminus j}\frac{\Phi(X_j,X_{i_1},...,X_{i_{m-1}};t)\delta_j\prod_{k=1}^{m-1}\delta_{i_k}}{K_c(X_j-)\prod_{l=1}^{m-1}K_c(X_{i_l}-)}-\frac{\varphi_1(X_j;t)\delta_j}{K_c(X_j-)}\Bigg)^2\Bigg)\\
&=E\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{2\leqslant i_1<\ldots<i_{m-1}\leqslant n}\frac{\Phi(X_1,X_{i_1},...,X_{i_{m-1}};t)\delta_1\prod_{k=1}^{m-1}\delta_{i_k}}{K_c(X_1-)\prod_{l=1}^{m-1}K_c(X_{i_l}-)}-\frac{\varphi_1(X_1;t)\delta_1}{K_c(X_1-)}\Bigg)^2\\
&=E\Bigg[E\Bigg(\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{2\leqslant i_1<\ldots<i_{m-1}\leqslant n}\frac{\Phi(X_1,X_{i_1},...,X_{i_{m-1}};t)\delta_1\prod_{k=1}^{m-1}\delta_{i_k}}{K_c(X_1-)\prod_{l=1}^{m-1}K_c(X_{i_l}-)}-\frac{\varphi_1(X_1;t)\delta_1}{K_c(X_1-)}\Bigg)^2\Bigg|X_1,\delta_1\Bigg)\Bigg].
\end{aligned}$$

За фиксирано t, $\frac{1}{\binom{n-1}{m-1}}\sum\limits_{2\leqslant i_1<\ldots<i_{m-1}\leqslant n}\frac{\Phi(x_1,X_{i_1},...,X_{i_{m-1}};s)\delta_1\prod\limits_{k=1}^{m-1}\delta_{i_k}}{K_c(x_1-)\prod_{k=1}^{m-1}K_c(X_{i_k}-)}$ је $U-$статистика са средњом

вредношћу $\frac{\varphi_1(x_1;s)\delta_1}{K_c(x_1-)}$. Према томе

$$\begin{aligned}
&E\Bigg[D\Bigg(\Bigg(\frac{1}{\binom{n-1}{m-1}}\sum_{2\leqslant i_1<\ldots<i_{m-1}\leqslant n}\frac{\Phi(X_1,X_{i_1},...,X_{i_{m-1}};t)\delta_1\prod_{k=1}^{m-1}\delta_{i_k}}{K_c(X_1-)\prod_{k=1}^{m-1}K_c(X_{i_k}-)}\Bigg)\Bigg|X_1,\delta_1\Bigg)\Bigg]\\
&\leqslant E\Bigg[\frac{m-1}{n}D\Bigg(\frac{\Phi(X_1,X_2,...,X_m;t)\prod_{k=1}^{m}\delta_{i_k}}{\prod_{k=1}^{m}K_c(X_{i_k}-)}\Bigg|X_1,\delta_1\Bigg)\Bigg]\\
&=\frac{m-1}{n}E[D(\widetilde{\Phi}((X_1,\delta_1),...,(X_m,\delta_m);t)|X_1,\delta_1)].
\end{aligned}$$

Из услова (2) теореме следи да је дисперзија у горњем изразу ограничена. На основу тога, и с обзиром да $\sqrt{n}(\Lambda_c(u-)-\widehat{\Lambda}_c(u-))=O_p(1),\ n\to\infty$, следи да грешка апроксимације приказана на почетку израза (5.24) униформно конвергира ка 0. Додатно, остатак приликом апроксимације $V-$статистике $U-$статистиком конвергира униформно ка 0, слично као што је показано раније.

Разлика између Нелсон-Аленове оцене функције хазарда и стварне вредности функције хазарда може се изразити користећи мартингале, при чему важи

$$\widehat{\Lambda}_c(s)-\Lambda_c(s)=\int_0^s\frac{dN^c(u)}{Y(u)}-\int_0^s d\Lambda_c(u)=\int_0^s\frac{Y(u)d\Lambda_c(u)+dM^c(u)}{Y(u)}-\int_0^s d\Lambda_c(u)=\int_0^s\frac{dM^c(u)}{Y(u)}.$$

Тада израз (5.23) постаје

$$\begin{aligned}
\frac{\binom{n}{m}m!}{n^m}\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}\int_0^{X_i-}\frac{dM^c(u)}{Y^c(u)}&=\frac{\binom{n}{m}m!}{n^m}\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}\int_0^{X_i-}\frac{\frac{1}{n}d\sum_{j=1}^{n}M_j^c(u)}{\frac{1}{n}\sum_{j=1}^{n}Y_j(u)}\\
&=\frac{\binom{n}{m}m!}{n^m}m\int_0^{\infty}\varphi_1(x;t)\int_0^{x-}\frac{d\sum_{j=1}^{n}\frac{1}{\sqrt{n}}M_j^c(u)}{y(u)}dF(x)+o_p(1)\\
&=\frac{\binom{n}{m}m!}{n^m}\frac{m}{\sqrt{n}}\sum_{j=1}^{n}\int_0^{\infty}\omega(u;t)dM_j^c(u)+o_p(1),\ n\to\infty,
\end{aligned}\tag{5.25}$$

где је $y(u)=EY_1(u)$ и $\omega(u;t)=\frac{1}{y(u)}\int\limits_u^{\infty}\varphi_1(x;t)dF(x)$.

Комбинујући (5.19), (5.22) и (5.25), добија се

$$\sqrt{n}\widehat{V}_{n,c}(t)=\frac{\binom{n}{m}m!}{n^m}\mathcal{Z}_n(t)+\sqrt{n}R_n(t),$$

где је

$$\mathcal{Z}_n(t)=\frac{m}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)}+\int_0^{\infty}\omega(u;t)dM_i^c(u)\right)\tag{5.26}$$

и $\sqrt{n}R_n(t)$ конвергира униформно ка 0. Овим је комплетиран доказ првог корака.

Корак II**:**

Из услова (1) и (2) теореме следи да, за фиксирано t, $\frac{1}{\sqrt{n}}\mathcal{Z}_n(t)$ је сума независних и једнако расподељених случајних величина са коначним дисперзијама. Према томе, гранична расподела те суме је нормална. Аналогно, применом вишедимензионалне централне граничне теореме, следи да су све коначнодимензионалне расподеле нормалне. Дакле, одатле следи да важи корак II(а).

У овом кораку треба још доказати да је процес $\{\mathcal{Z}_n(t), t \geqslant 0\}$ густ (корак II(б)). За то се користи последица 2.1. Услов 1. наведене последице следи из услова (2) и (3) теореме и неједнакости

$$\begin{aligned} E|\mathcal{Z}_n(0)|^2 &\leqslant \frac{m}{n}E\bigg(\frac{\varphi_1(X_i;0)\delta_i}{K_c(X_i-)} + \int_0^\infty \omega(t;0)dM_i^c(t)\bigg)^2 \\ &\leqslant 2\frac{m}{n}E\bigg(\frac{\varphi_1(X_i;0)\delta_i}{K_c(X_i-)}\bigg)^2 + 2\frac{m}{n}E\bigg(\int_0^\infty \omega(t;0)dM_i^c(t)\bigg)^2 \\ &\leqslant 2A_2 + 2A_3 < \infty. \end{aligned}$$

Тиме је показано да $\sup\limits_{n\geqslant 1} E|\mathcal{Z}_n(0)|^2 < \infty$. Да би се показао услов 2. наведене последице, показаћемо да постоје α, β и функција Υ такви да

$$\sup_{n\geqslant 1} E|\mathcal{Z}_n(s) - \mathcal{Z}_n(t)|^\alpha \leqslant \Upsilon(T)|t-s|^{1+\beta},\ \forall T>0, 0\leqslant s,t\leqslant T.$$

Претпоставимо да је $s = t+\zeta,\ \zeta \in [0,T]$. Тада следи

$$\mathcal{Z}_n(t+\zeta) - \mathcal{Z}_n(t) = \frac{2}{\sqrt{n}}\sum_{i=1}^n\bigg(\frac{\varphi_1(X_i;t+\zeta)\delta_i}{K_c(X_i-)} - \frac{\varphi_1(X_i;t)\delta_i}{K_c(X_i-)} + \int_0^\infty(\omega(u;t+\zeta)-\omega(u;t))dM_i(u)\bigg),$$

односно

$$\begin{aligned} E(\mathcal{Z}_n(t+\zeta) - \mathcal{Z}_n(t))^2 = E\bigg(\frac{4}{n}\sum_{i,j=1}^n&\bigg(\frac{(\varphi_1(X_i;t+\zeta)-\varphi_1(X_i;t))\delta_i}{K_c(X_i-)} + \int_0^\infty(\omega(u;t+\zeta)-\omega(u;t))dM_i(u)\bigg) \\ &\times\bigg(\frac{(\varphi_1(X_j;t+\zeta)-\varphi_1(X_j;t))\delta_i}{K_c(X_j-)} + \int_0^\infty(\omega(u;t+\zeta)-\omega(u;t))dM_j(u)\bigg)\bigg). \end{aligned}$$

У случају када $i \neq j$, елементи суме су једнаки 0 због независности и из чињенице да је, за свако $t>0$, $E(\mathcal{Z}_n(t)) = 0$. Према томе

$$E(\mathcal{Z}_n(t+\zeta) - \mathcal{Z}_n(t))^2 = \frac{m^2}{n}\sum_{i=1}^n E\bigg(\frac{(\varphi_1(X_i;t+\zeta)-h_1(X_i;t))\delta_i}{K_c(X_i-)} + \int_0^\infty(\omega(u;t+\zeta)-\omega(u;t))dM_i^c(u)\bigg)^2$$

$$= m^2 E\bigg(\frac{(\varphi_1(X_1;t+\zeta)-h_1(X_1;t))\delta_1}{K_c(X_1-)} + \int\limits_0^\infty (\omega(u;t+\zeta)-\omega(u;t))dM_1^c(u)\bigg)^2$$

$$\leqslant 2m^2 E\bigg(\frac{(\varphi_1(X_1;t+\zeta)-\varphi_1(X_1;t))\delta_1}{K_c(X_1^*-)}\bigg)^2 + 2m^2 E\bigg(\int\limits_0^\infty (\omega(u;t+\zeta)-\omega(u;t))dM_1^c(u)\bigg)^2,$$

где је коришћено да је $(a+b)^2 \leqslant 2a^2+2b^2$.

За први члан горњег израза, користећи теорему о средњој вредности, важи

$$\begin{aligned} E\bigg(\frac{(\varphi_1(X_1;t+\zeta)-\varphi_1(X_1;t))\delta_1}{K_c(X_1-)}\bigg)^2 &= E\bigg(\frac{(\varphi_1(X_1';t+\zeta)-\varphi_1(X_1';t))^2}{K_c(X_1'-)}\bigg) \\ &= E\bigg(\Big(\zeta\frac{\partial\varphi_1(X_1';t)}{\partial t}|_{t=\xi}\Big)^2\frac{1}{K_c(X_1'-)}\bigg) \\ &= \zeta^2 E\bigg(\Big(\frac{\partial\varphi_1(X_1';t)}{\partial t}|_{t=\xi}\Big)^2\frac{1}{K_c(X_1'-)}\bigg) \\ &\leqslant \zeta^2\Upsilon_1(T), \end{aligned}$$

где је $\xi\in[t,t+\zeta]$ и $\Upsilon_1(T)=\max\limits_{0\leqslant t\leqslant T} E\bigg(\Big(\frac{\partial\varphi_1(X_1';t)}{\partial t}|_{t=\xi}\Big)^2\frac{1}{K_c(X_1'-)}\bigg)$. Функција $\Upsilon_1(T)$ је коначна из услова (4) теореме.

Слично, користећи теорему о средњој вредности, за други члан горњег израза важи

$$\begin{aligned} |\omega(u;t+\zeta)-\omega(u;t)| &= \bigg|\frac{1}{P\{X\geqslant u\}}\int\limits_u^\infty \varphi_1(x;t+\zeta)dF(x) - \frac{1}{P\{X\geqslant t\}}\int\limits_u^\infty \varphi_1(x;t)dF(x)\bigg| \\ &\leqslant \frac{1}{P\{X\geqslant t\}}\int\limits_u^\infty |\varphi_1(x;t+\zeta)-\varphi_1(x;t)|dF(x) \\ &= \frac{|\zeta|}{P\{X\geqslant t\}}\int\limits_u^\infty \bigg|\frac{\partial\varphi_1(x;t)}{\partial t}|_{t=\xi}\bigg| dF(x), \end{aligned}$$

где је $\xi\in[t,t+\zeta]$. Тада

$$\begin{aligned} E\bigg(\int\limits_0^\infty (\omega(u;t+\zeta)-\omega(u;t))dM_1^c(u)\bigg)^2 &= \int\limits_0^\infty (\omega(u;t+\zeta)-\omega(u;t))^2 y(u)\lambda_c(u)du \\ &\leqslant \int\limits_0^\infty\bigg(\frac{|\zeta|}{P\{X\geqslant u\}}\int\limits_u^\infty\bigg|\frac{\partial\varphi_1(x;t)}{\partial t}|_{t=\xi}\bigg|dF(x)\bigg)^2 P\{X\geqslant u\}\frac{g(u)}{K_c(u)}du \\ &= \zeta^2\int\limits_0^\infty\bigg(\int\limits_u^\infty\bigg|\frac{\partial\varphi_1(x;t)}{\partial t}|_{t=\xi}\bigg|dF(x)\bigg)^2\frac{g(u)}{P\{X\geqslant u\}K_c(u)}du \leqslant \zeta^2\Upsilon_2(T), \end{aligned}$$

где је $\Upsilon_2(T)=\max\limits_{0\leqslant t\leqslant T}\int\limits_0^\infty\bigg(\int\limits_u^\infty\bigg|\frac{\partial\varphi_1(x;t)}{\partial t}|_{t=\xi}\bigg|dF(x)\bigg)^2\frac{g(u)}{P\{X\geqslant u\}K_c(u)}du<\infty$ што следи из услова (5) теореме.

Дакле, важи да је

$$E(\mathcal{Z}_n(t+\zeta)-\mathcal{Z}_n(t))^2 \leqslant 2m^2\zeta^2(\Upsilon_1(T)+\Upsilon_2(T)).$$

Тиме је показано да је процес $\{\mathcal{Z}_n(t), t\geqslant 0\}$ густ, чиме је комплетиран доказ корака II.

Корак III:

Применом теореме Слуцког за случајне процесе (теорема 2.14) на (5.16) и с обзиром да $\frac{\binom{n}{m}m!}{n^m}$ конвергира ка 1 следи тврђење теореме. ■

Претходна теорема у случају у потпуности одређеног процеса $\{\widehat{U}_{n,c}(t), t \geqslant 0\}$, насталог прилагођавањем $U-$емпиријског процеса за случај цензурисаних података, доказана је у раду аутора *Cuparić and Milošević (2020)*.

Претпоставимо да је расподела F позната, што одговара случају када се одређује асимптотска расподела тест статистике при нултој хипотези. Са друге стране, с обзиром да је реч о неинформативном цензурисању, расподела цензурисања није позната. Самим тим коваријацију (5.15) треба оценити. Природна оцена је одређена са

$$\widehat{cov}(\eta(t_1),\eta(t_2))=\frac{m^2}{n}\sum_{i=1}^{n}\Bigg(\frac{\varphi_1(X_i;t_1)\delta_i}{\widehat{K}_c(X_i-)}+\int_0^\infty\widehat{\omega}(u;t_1)d\widehat{M}_i^c(u)\Bigg)\Bigg(\frac{\varphi_1(X_i;t_2)\delta_i}{\widehat{K}_c(X_i-)}+\int_0^\infty\widehat{\omega}(u;t_2)d\widehat{M}_i^c(u)\Bigg).$$

Можемо приметити да је

$$\begin{aligned}\int_0^\infty \widehat{\omega}(u;t)d\widehat{M}_i^c(u) &= \int_0^\infty \widehat{\omega}(u;t)dN_i^c - \int_0^\infty \widehat{\omega}(u;t)Y_i(u)d\widehat{\Lambda}_c(u)\\ &= \widehat{\omega}(X_i;t)(1-\delta_i) - \int_0^\infty \widehat{\omega}(u;t)I\{X_i \geqslant u\}\frac{dN^c(u)}{Y(u)}\\ &= \widehat{\omega}(X_i;t)(1-\delta_i) - \sum_{j=1}^{n}\widehat{\omega}(X_j;t)I\{X_i \geqslant X_j\}\frac{1-\delta_j}{Y(X_j)},\end{aligned} \tag{5.27}$$

где је

$$\widehat{\omega}(u;t) = \frac{1}{(1-F(u))\widehat{K}_c(u-)}\int_0^\infty \varphi_1(x;t)\mathrm{I}\{x>u\}dF(x),$$

постојана оцена од

$$\omega(u;t) = \frac{1}{P\{X_1 \geqslant u\}}\int_0^\infty \varphi_1(x;t)\mathrm{I}\{x>u\}dF(x).$$

У случају када није позната ни расподела F, потребно је оценити и ту расподелу и прву пројекцију језгра.

5.4 Тестови сагласности

У овом поглављу ће бити примењени резултати претходног поглавља на тестове уведене у поглављима 3.3 и 3.4 за тестирање нулте хипотезе експоненцијалности. Приметимо да се раније дефинисане тест статистике могу трансформисати уз помоћ метода скалирања инверзом вероватноће цензурисања тако да важи

$$\widehat{M}_{c,a} = \int_0^\infty (\widehat{V}_{n,c}(t))^2 e^{-at} dt = \int_0^\infty \left(\frac{1}{n^m} \sum_{i_1,..,i_m=1}^{n} \frac{\Phi(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}}{\prod_{k=1}^m \widehat{K}_c(X_{i_k}-)} \right)^2 e^{-at} dt, \tag{5.28}$$

$$\widehat{L}_{c,a} = \sup_{t>0} |\widehat{V}_{n,c}(t)e^{-at}| = \sup_{t>0} \left| \frac{1}{n^m} \sum_{i_1,..,i_m=1}^{n} \frac{\Phi(X_{i_1},...,X_{i_m};t)\delta_{i_1}\cdots\delta_{i_m}}{\prod_{k=1}^m \widehat{K}_c(X_{i_k}-)} e^{-at} \right|, \tag{5.29}$$

где је $\Phi(x_1,...,x_m;t) = e^{-t\psi_1(x_1,...,x_m)} - e^{-t\psi_2(x_1,...,x_m)}$ и $\psi_i,\ i \in \{1,2\}$, су случајне величине из карактеризација. Без губитка општости претпоставимо да је функција Φ симетрична по својим аргументима (ако није, Φ ће бити одговарајућа симетризација наведеног језгра). Такође, због карактеризације, можемо приметити да је $\theta(t) = E(\Phi(X_1',...,X_m';t)) = 0$, за свако $t > 0$. Према томе, уз важење услова теореме 5.3, знамо да ће гранични процес $V-$емпријског процеса чијом L^2, односно L^∞, трансформацијом се добијају претходне статистике, бити Гаусов са коваријационом функцијом (5.15).

У наставку ће бити приказане теореме које дају граничне резултате тест статистика (5.28) и (5.29), а доказане су у раду аутора (*Cuparić and Milošević (2020)*).

Теорема 5.4 (Цупарић, Милошевић 2020) *При условима теореме 5.3, следи да*

$$n\widehat{M}_{c,a} \xrightarrow{D} \sum_{k=1}^{\infty} \upsilon_k Z_k^2,$$

где је $\upsilon_k,\ k = 1,2,...,$ низ сопствених вредности оператора A, дефинисаног за функције $q \in C[0,\infty)$ за које је $\int_0^\infty q(t)^2 w(t)dt < \infty$, са

$$Aq(t_1) = \int_0^\infty cov(\eta(t_1),\eta(t_2))q(t_2)e^{-at_2}dt_2,$$

где је $cov(\eta(t_1),\eta(t_2))$ дата у (5.15), и $Z_k,\ k = 1,2,...,$ независне случајне величине са стандардном нормалном расподелом.

Пре доказа теореме, увешћемо Мерцерову теорему (више детаља видети у *Van Trees and Bell (2013)*).

Теорема 5.5 *Претпоставимо да је $K(s,t)$ симетрична, непрекидна и ненегативно дефинитна функција дефинисана на $[a,b]\times[a,b]$. Тада постоји ортонормирани скуп* *сопствених вектора* $q_j(s)$ *и сопствених вредности* υ_j *таквих да важи*

$$K(s,t) = \sum_{j=1}^{\infty} \upsilon_j q_j(s) q_j(t)$$

при чему је конвергенција апсолутна и униформна по s *и* t*, где сопствене вредности и сопствени вектори задовољавају интегралну једначину*

$$\upsilon_j q_j(s) = \int_a^b K(s,t) q_j(t) dt.$$

Следеће теореме дају Карунен-Ловов развој случајног процеса и његов специјалан случај за Гаусов процес. Докази ових теорема се могу наћи у *Ash and Gardner (1975).*

Теорема 5.6 *Нека је* $\{X(t),\ a \leqslant t \leqslant b\}$*,* a *и* b *коначни,* L^2 *процес са очекивањем 0 и непрекидном коваријационом функцијом* K*. Нека је* $\{q_n, n = 1, 2, ...\}$ *ортонормирана база простора одређена сопственим векторима којима одговарају ненула сопствене вредности* υ_n *интегралног оператора повезаног са* K*, где је* q_n *сопствени вектор који одговара сопственој вредности* υ_n*. Тада*

$$X(t) = \sum_{n=1}^{\infty} Z_n q_n(t),\ a \leqslant t \leqslant b, \tag{5.30}$$

где су Z_n *ортогоналне случајне величине такве да је* $E(Z_n) = 0$ *и* $E|Z_n|^2 = \upsilon_n$*. Ред конвергира у* L^2 *униформно по* t*, односно*

$$E\left|X(t) - \sum_{k=1}^{n} Z_k q_k(t)\right|^2 \to 0,$$

кад $n \to \infty$ *униформно за* $t \in [a,b]$*.*

Теорема 5.7 *У Карунен-Ловововом развоју* (5.30) *Гаусовог процеса, случајне величине* Z_k *су нормалне, односно за свако* r*,* $Z_1, ..., Z_r$ *имају вишедимензионалну нормалну расподелу. Ако су случајне величине процеса реалне, онда су* Z_k *независне случајне величине.*

Доказ теореме 5.4: Доказ се састоји из три дела:

I показујемо да важи

$$\int_0^{\infty} \mathcal{Z}_n^2(t) e^{-at} dt \xrightarrow{D} \int_0^{\infty} \eta^2(t) e^{-at} dt, \tag{5.31}$$

где је $\{\eta(t), t \geqslant 0\}$ центрирани Гаусов процес са коваријационом функцијом (5.15), II показујемо да важи

$$\int_0^\infty (\sqrt{n}\widehat{V}_c(t))^2 e^{-at}dt \xrightarrow{D} \int_0^\infty \eta^2(t)e^{-at}dt, \tag{5.32}$$

III показујемо да важи

$$\int_0^\infty \eta^2(t)e^{-at}dt = \sum_{i=1}^\infty \upsilon_i Z_i^2.$$

Корак I:

Пошто је

$$\begin{aligned} cov(\eta(t),\eta(t)) &= m^2 E\biggl(\frac{\varphi_1(X_1;t)\delta_1}{K_c(X_1-)} + \int_0^\infty \omega(u;t)dM_1^c(u)\biggr)^2 \\ &\leqslant 2m^2 E\biggl(\frac{\varphi_1(X_1;t)\delta_1}{K_c(X_1-)}\biggr)^2 + 2m^2 E\biggl(\int_0^\infty \omega(u;t)dM_1^c(u)\biggr)^2 \\ &\leqslant 2m^2 E\biggl(\frac{\varphi_1^2(X_1';t)}{K_c(X_1'-)}\biggr) + 2m^2 \int_0^\infty \omega^2(u;t)y(t)\lambda_c(u)du < 2m^2(A_2+A_3) < \infty \end{aligned}$$

што следи из услова (2) и (3) теореме (5.3), следи да је

$$\int_0^\infty cov(\eta(t),\eta(t))e^{-at}dt < \infty.$$

Користећи Фубинијеву теорему следи да

$$E\int_0^\infty \eta^2(t)e^{-at}dt = \int_0^\infty cov(\eta(t),\eta(t))e^{-at}dt < \infty.$$

Дакле, можемо закључити да је

$$\int_0^\infty \eta^2(t)e^{-at}dt < \infty \text{ с.с.}$$

С обзиром да је

$$\mathcal{Z}_n(t) = \frac{2}{\sqrt{n}}\sum_{i=1}^n \biggl(\frac{h_1(X_i;t)\delta_i}{K_c(X_i-)} + \int_0^\infty \omega(u;t)dM_i^c(u)\biggr),$$

користећи независност и једнаку расподељеност случајних величина, следи да је

$$
\begin{aligned}
cov(\mathcal{Z}_n(t_1),\mathcal{Z}_n(t_2)) &= E\Bigg[\Bigg(\frac{2}{\sqrt{n}}\sum_{i=1}^{n}\Bigg(\frac{h_1(X_i;t_1)\delta_i}{K_c(X_i-)}+\int_0^\infty \omega(u;t_1)dM_i^c(u)\Bigg)\Bigg)\\
&\quad\times\Bigg(\frac{2}{\sqrt{n}}\sum_{i=1}^{n}\Bigg(\frac{h_1(X_i;t_2)\delta_i}{K_c(X_i-)}+\int_0^\infty \omega(u;t_2)dM_i^c(u)\Bigg)\Bigg)\Bigg]\\
&= \frac{4}{n}\sum_{i,j=1}^{n}E\Bigg[\Bigg(\frac{h_1(X_i;t_1)\delta_i}{K_c(X_i-)}+\int_0^\infty \omega(u;t_1)dM_i^c(u)\Bigg)\Bigg(\frac{h_1(X_j;t_2)\delta_i}{K_c(X_j-)}+\int_0^\infty \omega(u;t_2)dM_j^c(u)\Bigg)\Bigg]\\
&= \frac{4}{n}\sum_{i=1}^{n}E\Bigg[\Bigg(\frac{h_1(X_i;t_1)\delta_i}{K_c(X_i-)}+\int_0^\infty \omega(u;t_1)dM_i^c(u)\Bigg)\Bigg(\frac{h_1(X_i;t_2)\delta_i}{K_c(X_i-)}+\int_0^\infty \omega(u;t_2)dM_i^c(u)\Bigg)\Bigg]\\
&= cov(\eta(t_1),\eta(t_2)).
\end{aligned}
$$

Дакле, коваријације од $\mathcal{Z}_n(t)$ и $\eta(t)$ су једнаке. Нека су

$$T_{n,1}=\int_{R^+\setminus S}\mathcal{Z}_n^2(t)e^{-at}dt, n\geqslant 1$$

и

$$T_1=\int_{R^+\setminus S}\eta^2(t)e^{-at}dt,$$

где је S компактан скуп такав да за фиксирано $\varepsilon>0$ важи $E(T_{n,1})=E(T_1)\leqslant\varepsilon^2$. Из Марковљеве неједнакости, пошто су у питању ненегативне случајне величине, следи

$$P\{T_{n,1}\geqslant\varepsilon\}\leqslant\frac{ET_{n,1}}{\varepsilon}\leqslant\varepsilon. \tag{5.33}$$

Нека су

$$T_{n,2}=\int_S\mathcal{Z}_n^2(t)e^{-at}dt, n\geqslant 1$$

и

$$T_2=\int_S\eta^2(t)e^{-at}dt.$$

Тада је

$$\int_0^\infty\mathcal{Z}_n^2(t)e^{-at}dt=T_{n,1}+T_{n,2},\ \int_0^\infty\eta^2(t)e^{-at}dt=T_1+T_2.$$

У теореми 2.1 је показано да $\mathcal{Z}_n(t)\stackrel{D}{\to}\eta(t)$. На компакту S важи да је $\rho(\mathcal{Z}_n(t),\eta(t))$ ограничено. Према томе, пресликавање $f\mapsto\int_S f^2(t)e^{-at}dt$ постоји и непрекидно је за свако f из $C(S)$. Из теореме о непрекидном пресликавању следи

$$\int_S (\mathcal{Z}_n(t) - \eta(t))^2 e^{-at} dt \xrightarrow{D} 0.$$

С обзиром да, из неједнакости троугла, важи

$$\left| \int_S \mathcal{Z}_n^2(t) e^{-at} dt - \int_S \eta^2(t) e^{-at} dt \right| \leqslant \int_S (\mathcal{Z}_n(t) - \eta(t))^2 e^{-at} dt,$$

следи да $T_{n,2} \xrightarrow{D} T_2$. Дакле, из претходног, (5.33) и непрекидности расподеле од T_2 следи

$$\begin{aligned}
P\{T_1 + T_2 \leqslant s - \varepsilon\} - \varepsilon \leqslant P\{T_2 \leqslant s - \varepsilon\} - \varepsilon &= \lim_{n\to\infty} P\{T_{n,2} \leqslant s - \varepsilon\} - \varepsilon \\
&\leqslant \lim_{n\to\infty} (P\{T_{n,1} + T_{n,2} \leqslant s\} + P\{T_{n,1} \geqslant \varepsilon\}) - \varepsilon \\
&\leqslant \liminf_{n\to\infty} P\{T_{n,1} + T_{n,2} \leqslant s\} \\
&\leqslant \limsup_{n\to\infty} P\{T_{n,1} + T_{n,2} \leqslant s\} \\
&\leqslant \lim_{n\to\infty} P\{T_{n,1} + T_{n,2} \leqslant s\} \\
&\leqslant \lim_{n\to\infty} P\{T_{n,2} \leqslant s\} \\
&= P\{T_2 \leqslant s\} \\
&\leqslant P\{T_1 + T_2 \leqslant s + \varepsilon\} + P\{T_1 \geqslant \varepsilon\} \\
&\leqslant P\{T_1 + T_2 \leqslant s + \varepsilon\} + \varepsilon.
\end{aligned}$$

Када $\varepsilon \to 0$, следи (5.31).

Корак II:

С обзиром да, $\left| \sqrt{n}\widehat{V}_{n,c}(t) - \frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n(t) \right| = \sqrt{n} R_n(t)$ конвергира ка 0, следи да

$$\int_0^\infty \left| \sqrt{n}\widehat{V}_{n,c}(t) - \frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n(t) \right|^2 e^{-at} dt \xrightarrow{D} 0.$$

Користећи неједнакост троугла добија се

$$\left| \int_0^\infty (\sqrt{n}\widehat{V}_{n,c}(t))^2 e^{-at} dt - \int_0^\infty \left(\frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n(t) \right)^2 e^{-at} dt \right| \leqslant \int_0^\infty \left| \sqrt{n}\widehat{V}_{n,c}(t) - \frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n(t) \right|^2 dt,$$

и према томе важи (5.32).

Корак III:

С обзиром да је коваријациона функција ограничена, симетрична и позитивно дефинитна, из Мерцерове теореме (теорема 5.5) следи декомпозиција

$$cov(\eta(t_1), \eta(t_2)) = \sum_{i=1}^{\infty} \upsilon_i q_i(t_1) q_i(t_2), \tag{5.34}$$

где су υ_i сопствене вредности и $q_i(t)$ одговарајући сопствени вектори интегралног оператора A дефинисаног са

$$Aq(t_1) = \int_0^{\infty} cov(\eta(t_1), \eta(t_2)) q(t_2) e^{-at_2} dt_2. \tag{5.35}$$

Из теореме 5.7 следи

$$\eta(t) = \sum_{i=1}^{\infty} \sqrt{\upsilon_i} Z_i q_i(t),$$

где су Z_i независне случајне величине са $\mathcal{N}(0,1)$ расподелом. Према томе следи

$$\int_0^{\infty} \eta^2(t) e^{-at} dt = \int_0^{\infty} \Big(\sum_{i=1}^{\infty} \sqrt{\upsilon_i} Z_i q_i(t) \Big)^2 e^{-at} dt = \sum_{i=1}^{\infty} \upsilon_i Z_i^2.$$

Последња једнакост следи јер су $\{q_i\}$ ортонормирани вектори. Тиме је доказ теореме комплетиран. ■

Теорема 5.8 *При условима теореме 5.3, следи да*

$$\sqrt{n} \widehat{L}_{c,a} \xrightarrow{D} \sup_{t>0} |\eta(t)|,$$

где је $\{\eta(t), t \geqslant 0\}$ *центрирани Гаусов процес са коваријационом функцијом* (5.15).

Доказ: Увођењем смене $s = e^{-t} \in [0,1]$, језгро процеса $\{\sqrt{n}\widehat{V}_{n,c}(s), s \in [0,1]\}$ има облик

$$\Phi(x_1, ..., x_m; s) = s^{\psi_1(x_1,...,x_m)} - s^{\psi_2(x_1,...,x_m)}.$$

Дакле, наведени процес представља процес дефинисан на простору $C[0,1]$ са супремум нормом. Користећи сличне аргументе као у доказу теореме 5.3, може се показати да процес $\{\sqrt{n}\widehat{V}_{n,c}(s), s \in [0,1]\}$ конвергира слабо у $C[0,1]$ ка центрираном Гаусовом процесу $\{\eta(s), s \in [0,1]\}$ са коваријационом функцијом која се добија из (5.15) увођењем исте смене. С обзиром да је супремум непрекидна функција на простору $C[0,1]$, из теореме о непрекидном пресликавању и коришћењем трансформације $t = -\ln s$, следи тврђење теореме. ■

У наставку ћемо проучавати перформансе тест статистика за тестирање просте хипотезе $H_0 : F(x) = 1 - e^{-\frac{x}{\mu}}, x \geqslant 0$, са Пури-Рубиновом и Десуовом карактеризацијом. У том случају, процес $\{\widehat{V}_{n,c}(t), t \geqslant 0\}$ чијом се L^2, односно L^{∞}, трансформацијом добијају статистике (5.28) и (5.29), је одређен са

$$\widehat{V}^{\mathcal{I}}_{n,c}(t) = \frac{1}{n^2}\sum_{i,j}\frac{\Phi^{\mathcal{I}}(X_i,X_j;t)\delta_i\delta_j}{\widehat{K}_c(X_i-)\widehat{K}_c(X_j-)},$$

где је $\mathcal{I} \in \{\mathcal{P},\mathcal{D}\}$ и

$$\Phi^{\mathcal{I}}(x_1,x_2;t) = \frac{1}{2}\Big(e^{-tx_1} + e^{-tx_2} - 2e^{-t\psi^{\mathcal{I}}(x_1,x_2)}\Big) \tag{5.36}$$

за

$$\psi^{\mathcal{P}}(x,y) = |x-y|,\ \psi^{\mathcal{D}}(x,y) = 2\min(x,y).$$

Прве пројекције језгара $\Phi^{\mathcal{P}}$ и $\Phi^{\mathcal{D}}$, за $\mu = 1$, су

$$\varphi^{\mathcal{P}}_1(x;t) = \frac{1}{2(t+1)} - \frac{2t}{t^2-1}e^{-x} + \frac{1+t}{2(t-1)}e^{-tx}, \tag{5.37}$$

$$\varphi^{\mathcal{D}}_1(x;t) = -\frac{2te^{-(2t+1)x}}{2t+1} + \frac{e^{-tx}}{2} - \frac{1}{2(t+1)(2t+1)}. \tag{5.38}$$

Прво треба проверити да ли те статистике задовољавају услове претходних теорема. Приметимо да

$$\begin{aligned} E(\widetilde{\Phi}^{\mathcal{I}}((X_1,\delta_1),(X_2,\delta_2);t))^2 &= E\left(\frac{\Phi^{\mathcal{I}}(X_1,X_2;t)\delta_1\delta_2}{K_c(X_1-)K_c(X_2-)}\right)^2 = E\Bigg(\frac{(\Phi^{\mathcal{I}}(X_1',X_2';s))^2}{K_c(X_1'-)K_c(X_2'-)}\Bigg) \\ &\leqslant 4E\Bigg(\frac{1}{K_c(X_1'-)K_c(X_2'-)}\Bigg) = 4\left(E\Bigg(\frac{1}{K_c(X_1'-)}\Bigg)\right)^2, \end{aligned}$$

што следи из ограничености језгра (без обзира на карактеризацију) и независности. Даље,

$$\int\limits_0^\infty \frac{(\varphi^{\mathcal{I}}_1(x;t))^2}{K_c(x-)}dF(x) \leqslant 4E\Bigg(\frac{1}{K_c(X_1'-)}\Bigg),$$

што следи из ограничености првих пројекција (5.37) и (5.38). Дакле, да би били испуњени услови (1) и (2) теореме 5.3 потребно је да важи $E\Big(\frac{1}{K_c(X_1'-)}\Big) < \infty$. Што се тиче услова (3), важи да је

$$\begin{aligned} Var\left(\int\limits_0^\infty \omega(u;t)dM^c_1(u)\right) &= \int\limits_0^\infty \omega^2(u;t)y(u)\lambda_c(u)du \\ &= \int\limits_0^\infty \left(\frac{1}{P\{X \geqslant u\}}\int\limits_u^\infty \varphi^{\mathcal{I}}_1(x;t)dF(x)\right)^2 P\{X \geqslant u\}\frac{g(u)}{K_c(u)}du \end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{(1-F(u))K_c(u)} \bigg(\int_u^\infty \varphi_1^{\mathcal{I}}(x;t) f(x) dx \bigg)^2 \frac{g(u)}{K_c(u)} du \\
&\leqslant \int_0^\infty \frac{4(1-F(u))g(u)}{K_c^2(u)} du.
\end{aligned}
$$

Према томе, услов (3) теореме 5.3 ће бити испуњен ако важи $\int_0^\infty \frac{1-F(u)}{K_c^2(u)} dG(u) < \infty$.

Пошто је

$$
\frac{\partial \varphi_1^{\mathcal{P}}(x;t)}{\partial t} = -\frac{1}{2(t+1)^2} + \frac{2e^{-x}(t^2+1)}{(t^2-1)^2} - \frac{e^{-tx}}{(t-1)^2} - \frac{e^{-tx}x(t+1)}{2(t-1)},
$$

може се показати да је, за $t \in [0, T]$,

$$
\frac{\partial \varphi_1^{\mathcal{P}}(x;t)}{\partial t} \leqslant 36x + 17.
$$

Према томе следи да је

$$
\begin{aligned}
\max_{0 \leqslant t \leqslant T} E\bigg(\Big(\frac{\partial \varphi_1^{\mathcal{P}}(X_1';t)}{\partial s} \Big)^2 \frac{1}{K_c(X_1'-)} \bigg) &\leqslant E\bigg(\max_{0 \leqslant t \leqslant T} \bigg| \Big(\frac{\partial \varphi_1^{\mathcal{P}}(X_1';t)}{\partial s} \Big)^2 \frac{1}{K_c(X_1'-)} \bigg| \bigg) \\
&\leqslant E\bigg(\frac{(36X_1'+17)^2}{K_c(X_1'-)} \bigg) \\
&\leqslant 2 \cdot 36^2 E\bigg(\frac{(X_1')^2}{K_c(X_1'-)} \bigg) + 2 \cdot 17^2 E\bigg(\frac{1}{K_c(X_1'-)} \bigg).
\end{aligned}
$$

Дакле, услов (4) ће бити испуњен за свако $T \geqslant 1$, ако је $\int_0^\infty \frac{u^2}{K_c(u-)} dF(u) < \infty$ и $\int_0^\infty \frac{1}{K_c(u-)} dF(u) < \infty$.

Слично

$$
\begin{aligned}
&\max_{0 \leqslant t \leqslant T} \int_0^\infty \bigg(\int_u^\infty \Big| \frac{\partial \varphi_1^{\mathcal{P}}(x;t)}{\partial t} |_{t=\xi} \Big| dF(x) \bigg)^2 \frac{g(u)}{P\{X \geqslant u\} K_c(u)} du \\
&\leqslant \int_0^\infty \bigg(\int_u^\infty \max_{0 \leqslant t \leqslant T} \Big| \frac{\partial \varphi_1^{\mathcal{P}}(x;t)}{\partial t} |_{t=\xi} \Big| dF(x) \bigg)^2 \frac{g(u)}{P\{X \geqslant u\} K_c(u)} du \\
&\leqslant \int_0^\infty \bigg(\int_u^\infty (36x+17) dF(x) \bigg)^2 \frac{g(u)}{e^{-u} K_c^2(u)} du \\
&= \int_0^\infty \big(e^{-u}(36u+53) \big)^2 \frac{g(u)}{e^{-u} K_c^2(u)} du \\
&\leqslant 2 \cdot 36^2 \int_0^\infty e^{-u} u^2 \frac{g(u)}{K_c^2(u)} du + 2 \cdot 53^2 \int_0^\infty e^{-u} \frac{g(u)}{K_c^2(u)} du.
\end{aligned}
$$

Према томе, услов (5) је испуњен ако је $\int\limits_0^\infty \frac{u^2(1-F(u))}{K_c^2(u)}dG(u) < \infty$ и $\int\limits_0^\infty \frac{1-F(u)}{K_c^2(u)}dG(u) < \infty$.

Што се тиче Десуове карактеризације, важи да је

$$\frac{\partial \varphi_1^{\mathcal{D}}(x;t)}{\partial t} = \frac{4txe^{-(2t+1)x}}{2t+1} - \frac{1}{2}xe^{-tx} - \frac{2e^{-(2t+1)x}}{(2t+1)^2} + \frac{4t+3}{2(t+1)^2(2t+1)^2}$$

и може се показати да је, за $t \in [0, T]$,

$$\frac{\partial \varphi_1^{\mathcal{D}}(x;t)}{\partial t} \leqslant \frac{7+5x}{2}.$$

Према томе следи да је

$$\begin{aligned}\max_{0\leqslant t\leqslant T} E\Bigg(\left(\frac{\partial \varphi_1^{\mathcal{D}}(X_1';t)}{\partial s}\right)^2 \frac{1}{K_c(X_1'-)}\Bigg) &\leqslant E\Bigg(\max_{0\leqslant t\leqslant T}\Bigg|\left(\frac{\partial \varphi_1^{\mathcal{D}}(X_1';t)}{\partial s}\right)^2 \frac{1}{K_c(X_1'-)}\Bigg|\Bigg)\\ &\leqslant E\Bigg(\left(\frac{7+5X_1'}{2}\right)^2 \frac{1}{K_c(X_1'-)}\Bigg)\\ &\leqslant \frac{25}{2}E\left(\frac{(X_1')^2}{K_c(X_1'-)}\right) + \frac{49}{2}E\left(\frac{1}{K_c(X_1'-)}\right).\end{aligned}$$

Слично

$$\begin{aligned}&\max_{0\leqslant t\leqslant T}\int\limits_0^\infty\left(\int\limits_u^\infty\left|\frac{\partial \varphi_1^{\mathcal{D}}(x;t)}{\partial t}|_{t=\xi}\right|dF(x)\right)^2\frac{g(u)}{P\{X\geqslant u\}K_c(u)}du\\ &\leqslant \int\limits_0^\infty\left(\int\limits_u^\infty\max_{0\leqslant t\leqslant T}\left|\frac{\partial \varphi_1^{\mathcal{D}}(x;t)}{\partial t}|_{t=\xi}\right|dF(x)\right)^2\frac{g(u)}{P\{X\geqslant u\}K_c(u)}du\\ &\leqslant \int\limits_0^\infty\left(\int\limits_u^\infty\frac{7+5x}{2}dF(x)\right)^2\frac{g(u)}{e^{-u}K_c^2(u)}du\\ &= \int\limits_0^\infty\left(\frac{1}{2}(12+5u)\right)^2\frac{e^{-u}g(u)}{K_c^2(u)}du\\ &\leqslant 72\int\limits_0^\infty\frac{e^{-u}g(u)}{K_c^2(u)}du + \frac{25}{2}\int\limits_0^\infty\frac{u^2e^{-u}g(u)}{K_c^2(u)}du.\end{aligned}$$

Дакле, да би услови (4) и (5) били испуњени треба да важи $\int\limits_0^\infty \frac{u^2}{K_c(u-)}dF(u) < \infty$, $\int\limits_0^\infty \frac{1}{K_c(u-)}dF(u) < \infty$, $\int\limits_0^\infty \frac{u^2(1-F(u))}{K_c^2(u)}dG(u) < \infty$ и $\int\limits_0^\infty \frac{1-F(u)}{K_c^2(u)}dG(u) < \infty$.

Према томе, без обзира на карактеризацију, услови претходних теорема се своде на услове

$$\int\limits_0^\infty \frac{dF(u)}{K_c(u-)} < \infty,\ \int\limits_0^\infty \frac{1-F(u)}{K_c^2(u)}dG(u) < \infty,\ \int\limits_0^\infty \frac{u^2dF(u)}{K_c(u-)} < \infty,\ \int\limits_0^\infty \frac{u^2(1-F(u))}{K_c^2(u)}dG(u) < \infty.$$

Наравно, ове услове формално можемо проверити само код информативног цензурисања, односно када је расподела цензурисања позната. Стога, ако за модел цензурисања узмемо Козиол-Гринов модел, који је поменут раније, онда је

$$K_c(x) = (1 - F(x))^{\beta},$$

где је $\beta > 0$ непознати параметар. У том случају вероватноћа цензурисања је

$$p = P\{X' > C\} = \int_0^{\infty}\int_y^{\infty} dF(x)dG(y) = \int_0^{\infty}\int_y^{\infty} f(x)\beta(1 - F(y))^{\beta-1}dxdy = \frac{\beta}{1+\beta}. \tag{5.39}$$

У случају нулте хипотезе експоненцијалности, следи да је $K_c(x) = e^{-\beta x},\ x \geqslant 0$. Тада су сви услови испуњени ако је $\beta < 1$. С обзиром да из (5.39) $\beta = \frac{p}{1-p}$, услови теореме 5.3 ће бити испуњени при овом моделу ако је проценат цензурисања мањи од 50%, што је разуман захтев.

У случају великих узорака могу се користити резултати претходних теорема уз коришћење постојаних оцена одговарајућих коваријација за апроксимацију $p-$вредности тестова. Међутим, у случају узорака мале и умерене величине неопходно је коришћење неког од метода реузорковања за апроксимацију тих вредности. Такав један метод је приказан у следећем поглављу.

5.5 Метод реузорковања за апроксимацију $p-$вредности теста

У наставку ћемо размотрити метод реузорковања у случају узорака умерене величине ($n = 50$) и за одређени проценат цензурисања, који се може користити за апроксимацију $p-$вредности тестова L^2 и L^{∞} типа размотрених у претходном поглављу. Разликоваћемо случај када се тестира проста хипотеза и случај када се тестира сложена хипотеза. Претпостављамо да расподела цензурисања није позната, а желимо да испитамо како се емпиријске моћи тестова мењају у зависности од процента цензурисања.

С обзиром да расподела тест статистика при нултој хипотези зависи од расподеле цензурисања, у случају узорка умерене величине, неопходно је коришћење неког метода поновног узорковања. Биће приказан прилагођен Ефронов бутстреп (*Efron (1981)*) при чему се користи информација да се поновно узорковање врши при нултој хипотези, односно да је расподела полазног (нецензурисаног дела) узорка позната.

Размотримо прво тестирање просте хипотезе H_0 да узорак има експоненцијалну $\mathcal{E}(1)$ расподелу. Емпиријске моћи се одређују коришћењем следеће бутстреп процедуре:

1. На основу узорка $(X_1, \delta_1), ..., (X_n, \delta_n)$ одредити тест статистику T_n;

2. Емпиријска критична вредност $q^*_{n,\alpha}$ се добија на следећи начин:

(а) генерисати узорак $C_1^*, ..., C_n^*$ из Каплан-Мајерове функције расподеле G_n;

(б) генерисати нови узорак $X_1', ..., X_n'$ из $\mathcal{E}(1)$ расподеле;

(в) користећи (а) и (б) формирати бутстрепован узорак $(X_1^*, \delta_1^*), ..., (X_n^*, \delta_n^*)$, где је $X_i^* = \min(X_i', C_i^*)$ и $\delta_i^* = I\{X_i' \leqslant C_i^*\}$;

(г) на основу узорка из (в) одредити вредност тест статистике

$$T_n^* = T_n((X_1^*, \delta_1^*), ..., (X_n^*, \delta_n^*));$$

(д) поновити кораке (а)-(г) B пута;

(ђ) на основу добијеног низа бутстрепованих тест статистика оценити критичну вредност $q_{n,\alpha}^*$;

3. Одбацити H_0 ако је $T_n \geqslant q_{n,\alpha}^*$;

4. Поновити кораке 1-3. N пута. Оцењена емпиријска моћ је проценат одбацивања H_0.

Треба нагласити да бутстреп који прати кораке претходног алгоритма, приликом реузорковања не чува проценат цензурисања, већ расподелу цензора.

Нека су $\{\sqrt{n}\widehat{V}_c(t), t \geqslant 0\}$ и $\{\eta(t), t \geqslant 0\}$ процеси размотрени у теореми 5.3. Следећа теорема, која је доказана у раду аутора (*Cuparić and Milošević (2020)*), даје асимптотско понашање одговарајућег бутстрепованог процеса $\{\sqrt{n}\widehat{V}_c^*(t), t \geqslant 0\}$ што оправдава коришћење наведене процедуре.

Теорема 5.9 (Цупарић, Милошевић 2020) *Претпоставимо да су задовољени услови теореме 5.3. Тада, условно у односу на узорак* $(X_1, \delta_1),...,(X_n, \delta_n)$, *следи да* $\{\sqrt{n}\widehat{V}_c^*(t), t\geqslant 0\}$ *слабо конвергира ка процесу* $\{\eta(t), t \geqslant 0\}$ *у* $C[0, \infty)$.

Доказ: Нека је $\{\mathcal{Z}_n^*(t), t \geqslant 0\}$, бутстрепована верзија процеса $\{\mathcal{Z}_n(t), t \geqslant 0\}$ дефинисана са

$$\mathcal{Z}_n^*(t) = \frac{m}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{\varphi_1(X_i^*; t)\delta_i^*}{\widehat{K}_c(X_i^*-)} + \int_0^\infty \widehat{\omega}(u; t) dM_i^{c*}(u) \right),$$

где је $dM_i^{c*}(u) = dN_i^{c*}(u) + Y^*(u)d\widehat{\Lambda}_c(u)$ и $M_i^{c*}(u)$ мартингал у односу на филтрацију $\mathcal{F}_t = \sigma(N_i^c(s), Y_i(s), N_i^{c*}(s), Y_i^*(s), 0 \leqslant s \leqslant t, i = 1, ..., n)$. Желимо да покажемо да

1. $\{\mathcal{Z}_n^*(t), t \geqslant 0\}$, условно у односу на узорак $(X_1, \delta_1), ..., (X_n, \delta_n)$, конвергира слабо ка $\{\eta(t), t \geqslant 0\}$;

2. $\sqrt{n}\widehat{V}_{n,c}^*(t) = \frac{\binom{n}{m} m!}{n^m} \mathcal{Z}_n^*(t) + \sqrt{n}R_n^*(t)$, где је $||\sqrt{n}R_n^*(t)|| \xrightarrow{P^*} 0$.

Доказ првог дела се састоји у показивању да све коначнодимензионалне расподеле у односу на полазни узорак су вишедимензионалне нормалне, да је $\{\mathcal{Z}_n^*(t), t \geqslant 0\}$ густ и да се његова гранична коваријациона функција поклапа са (5.15).

Први део следи из Линдберг-Фелерове централне граничне теореме. Специјално, с обзиром да је $\frac{\delta_i^*}{\widehat{K}_c(X_i^*)} = 1 - \int_0^\infty \frac{dM_i^{c*}(u)}{\widehat{K}_c(u)}$ (што је једно од основних идентитета показаних у *Robins and Rotnitzky (1992)*, 313. страна), следи да се $\mathcal{Z}_n^*(t)$ може написати као

$$\mathcal{Z}_n^*(t) = \frac{m}{\sqrt{n}} \sum_{i=1}^n \left(\varphi_1(X_i'; t) - \int_0^\infty \left(\varphi_1(X_i'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM_i^{c*}(u)}{\widehat{K}_c(u-)} \right).$$

Узимајући у обзир мартингална својства од M^{c*}, за свако $t > 0$, $E^*\left(\mathcal{Z}_n^*(t)\right) = 0$.

Даље, следи

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n\sigma_*^2(t)} \sum_{i=1}^n E^* \Bigg[\left(\varphi_1(X_i'; t) - \int_0^\infty \left(\varphi_1(X_i'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM_i^{c*}(u)}{\widehat{K}_c(u-)} \right)^2 \\
&\times I\left\{ \left| \varphi_1(X_i'; t) - \int_0^\infty \left(\varphi_1(X_i'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM_i^{c*}(u)}{\widehat{K}_c(u-)} \right| > \sqrt{n}\varepsilon\sigma_*(t) \right\} \Bigg] \\
&= \lim_{n\to\infty} \frac{1}{\sigma_*^2(t)} E^* \Bigg[\left(\varphi_1(X_1'; t) - \int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM_1^{c*}(u)}{\widehat{K}_c(u-)} \right)^2 \\
&\times I\left\{ \left| \varphi_1(X_1'; t) - \int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM_1^{c*}(u)}{\widehat{K}_c(u-)} \right| > \sqrt{n}\varepsilon\sigma_*(t) \right\} \Bigg], \qquad (5.40)
\end{aligned}$$

где је

$$\begin{aligned}
&\sigma_*^2(t) = Var^*(\mathcal{Z}_n^*(t)) \\
&= E^* \left(\varphi_1(X_1'; t) - \int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; s) dF(v) \right) \frac{dM^{c*}(u)}{\widehat{K}_c(u-)} \right)^2 \\
&= E^* \varphi_1^2(X_1'; t) + E^* \left(\int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM^{c*}(u)}{\widehat{K}_c(u-)} \right)^2 \\
&- 2E^* \left(\varphi_1(X_1'; t) \int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right) \frac{dM^{c*}(u)}{\widehat{K}_c(u-)} \right) \\
&= E^* \varphi_1^2(X_1'; t) + E^* \left(\int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; t) dF(v) \right)^2 I\{X_1' \leqslant u\} \frac{d\widehat{\lambda}_c(u)}{\widehat{K}_c(u-)} \right) \\
&- 2E^* \left(\varphi_1(X_1'; t) \int_0^\infty \left(\varphi_1(X_1'; t) - \frac{1}{1-F(u)} \int_u^\infty \varphi_1(v; s) dF(v) \right) \frac{dM^{c*}(u)}{\widehat{K}_c(u-)} \right).
\end{aligned}$$

С обзиром да $\widehat{K}_c(x) \xrightarrow{P} K_c(x)$ и $\max\limits_{i=1,\dots,n} \frac{K_c(X_i-)}{\widehat{K}_c(X_i-)} = O_p(1),\ n \to \infty$, (теорема 5.1) следи да

$$\frac{1}{\widehat{K}_c(X_i-)} = \frac{1}{K_c(X_i-)} + o_p(1),\ n \to \infty.$$

Додатно, $\widehat{\Lambda}_c(t) \xrightarrow{P} \Lambda_c(t)$. Према томе $\sigma_*^2(t) \xrightarrow{P} \sigma^2(t) = cov(\eta(t), \eta(t)) < \infty$, и треба доказати да

$$\begin{aligned}\lim_{n\to\infty} E^*\Bigg[&\left(\varphi_1(X_1';t) - \int\limits_0^\infty \left(\varphi_1(X_1';t) - \frac{1}{1-F(u)} \int\limits_u^\infty \varphi_1(v;t)dF(v)\right) \frac{dM_1^{c*}(u)}{\widehat{K}_c(u-)}\right)^2 \\ &\times I\left\{\left|\varphi_1(X_1';t) - \int\limits_0^\infty \left(\varphi_1(X_1';t) - \frac{1}{1-F(u)} \int\limits_u^\infty \varphi_1(v;t)dF(v)\right) \frac{dM_1^{c*}(u)}{\widehat{K}_c(u-)}\right| > \sqrt{n}\varepsilon\right\}\Bigg] = 0 \qquad (5.41)\end{aligned}$$

што директно следи из L^2-ограничености (у односу на меру dF).

Слично се доказује да су све коначнодимензионалне расподеле од $\{\mathcal{Z}_n^*(t), t \geqslant 0\}$ вишедимензионалне нормалне. Доказ да је процес густ се изводи на исти начин као за процес $\{\mathcal{Z}_n(t), t \geqslant 0\}$ чиме се комплетира доказ дела 1.

Остатак доказа се изводи на исти начин као доказ теореме 5.3 (корак I). ■

За илустрацију резултата претходно описане процедуре као алтернативну расподелу користићемо Вејбулову $W(\theta)$ расподелу са густином

$$g(x;\theta) = e^{-x^\theta}\theta x^{\theta-1}, x \geqslant 0, \theta > 0.$$

Узорак је умерене величине ($n = 50$), ниво значајности $\alpha = 0.05$, а проценат цензурисања полазног узорка је $10\%, 20\%$ или 30%. Пошто расподела цензора није позната, да би се постигао одређени проценат цензурисања, користи се Козиол-Гринов модел. Задавањем процента цензурисања p, из (5.39) се може одредити β у поменутом моделу. Узимањем узорака из алтернативне расподеле и узорка из расподеле цензора добијене из Козиол-Гриновог модела (5.39), добија се полазни узорак са одређеним процентом цензурисања на који је затим примењена бутстреп процедура, са $B = 1000$ и $N = 1000$. Такође су размотрене различите вредности параметра подешавања a, $a \in \{1, 2, 5\}$. Резултати су приказани у табели 5.1.

Табела 5.1: Проценат одбацивања (просте) нулте хипотезе

p	Алт.	$\widehat{M}_{c,1}^{\mathcal{P}}$	$\widehat{M}_{c,2}^{\mathcal{P}}$	$\widehat{M}_{c,5}^{\mathcal{P}}$	$\widehat{M}_{c,1}^{\mathcal{D}}$	$\widehat{M}_{c,2}^{\mathcal{D}}$	$\widehat{M}_{c,5}^{\mathcal{D}}$	$\widehat{L}_{c,1}^{\mathcal{P}}$	$\widehat{L}_{c,2}^{\mathcal{P}}$	$\widehat{L}_{c,5}^{\mathcal{P}}$	$\widehat{L}_{c,1}^{\mathcal{D}}$	$\widehat{L}_{c,2}^{\mathcal{D}}$	$\widehat{L}_{c,5}^{\mathcal{D}}$
0.1	$Exp(1)$	5	5	4	5	5	4	5	4	4	5	4	4
	$W(1.4)$	80	76	69	78	77	76	72	67	66	72	74	73
0.2	$Exp(1)$	5	5	5	5	5	5	4	5	5	5	6	5
	$W(1.4)$	78	70	70	76	78	75	67	66	66	68	72	74
0.3	$Exp(1)$	5	8	7	5	5	5	4	5	6	6	7	7
	$W(1.4)$	72	67	67	74	74	74	63	63	63	68	72	72

У случају тестирања сложене хипотезе, односно хипотезе облика $H_0: \ X \in \mathcal{E}(\frac{1}{\mu})$ за произвољно $\mu > 0$, може се користи бутстреп алгоритам сличан претходном, са јединим разликом у тачки 2. (б), у којој се сада генерише нови узорак из $\mathcal{E}\left(\frac{1}{\widehat{\mu}}\right)$, где је $\widehat{\mu}$ оцена максималне веродостојности параметра μ.

Функција веродостојности у случају узорка цензурисаног са десне стране је одређена са (за детаље видети на пример *Lawless (2011)*)

$$L(\mu) = \prod_{i=1}^{n} f^{\delta_i}(X_i;\mu)(1 - F(X_i;\mu))^{1-\delta_i} K_c^{\delta_i}(X_i) g^{1-\delta_i}(X_i), \ \mu > 0.$$

Пошто је у питању неинформативно цензурисање, расподела цензора не зависи од параметра μ. Заменом расподеле из нулте хипотезе у претходном изразу добија се

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\mu^{\delta_i}} e^{-\frac{X_i}{\mu}} K_c^{\delta_i}(x_i) g^{1-\delta_i}(x_i), \ \mu > 0.$$

Из претходног израза се на стандардан начин добија оцена максималне веродостојности и она је једнака $\widehat{\mu} = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} \delta_i}$. Постојаност ове оцене следи применом закона великих бројева на количник. Општи услови за постојаност оцене максималне веродостојности у случају података цензурисаних са десне стране могу се наћи у *Stute (1992)*.

Постојаност оцене омогућава да се покаже постојаност бутстрепа у случају сложене хипотезе на исти начин као и у теореми 5.9.

У табели 5.2 приказани су резултати процедуре у случају тестирања сложене хипотезе експоненцијалности и при истим условима као у случају просте хипотезе.

Табела 5.2: Проценат одбацивања (сложене) нулте хипотезе

p	Алт.	$\widehat{M}^{\mathcal{P}}_{c,1}$	$\widehat{M}^{\mathcal{P}}_{c,2}$	$\widehat{M}^{\mathcal{P}}_{c,5}$	$\widehat{M}^{\mathcal{D}}_{c,1}$	$\widehat{M}^{\mathcal{D}}_{c,2}$	$\widehat{M}^{\mathcal{D}}_{c,5}$	$\widehat{L}^{\mathcal{P}}_{c,1}$	$\widehat{L}^{\mathcal{P}}_{c,2}$	$\widehat{L}^{\mathcal{P}}_{c,5}$	$\widehat{L}^{\mathcal{D}}_{c,1}$	$\widehat{L}^{\mathcal{D}}_{c,2}$	$\widehat{L}^{\mathcal{D}}_{c,5}$
0.1	$Exp(1)$	5	6	4	5	5	5	5	4	5	5	5	5
	$W(1.4)$	86	78	75	77	79	80	72	73	74	75	80	81
0.2	$Exp(1)$	6	6	6	5	6	6	6	6	6	5	6	6
	$W(1.4)$	82	81	82	76	78	84	78	76	76	78	78	80
0.3	$Exp(1)$	5	7	8	5	8	6	8	9	9	7	8	8
	$W(1.4)$	85	83	87	74	82	82	77	78	78	81	81	80

Наведену процедуру можемо применити за тестирање експоненцијалности на реалним подацима. Подаци представљају времена преживљавања (у данима) 51 пацијента који су примили одређени третман за туморе главе и врата. Ови подаци су први пут размотрени у раду Ефрона (*Efron (1988)*) и приказани су у табели 5.3 где су са звездицом означени подаци који су цензурисани (њих 9 или приближно 18%).

Ефрон је ове податке поделио у 47 интервала, сваки дужине један месец, и користећи стандардну логистичку регресију оценио је стопу хазарда. Показано је да је оцењена стопа хазарда за ове податке унимодална. Према томе експоненцијална расподела у овом случају не би требало да буде адекватан избор.

Табела 5.3: Времена преживљавања (у данима) пацијената

7	34	42	63	64	74⋆	83	84	91	108	112	129	133
133	139	140	140	146	149	154	157	160	160	165	173	176
185⋆	218	225	241	248	273	277	279⋆	297	319⋆	405	417	420
440	523⋆	523	583	594	1101	1116⋆	1146	1226⋆	1349⋆	1412⋆	1417	

У табели 5.4 приказане су $p-$вредности тестова предложених у претходном поглављу за наведене податке одређене на три начина

I коришћењем бутстреп метода и тест статистика добијених на основу метода скалирања инверзом расподеле цензурисања,

II занемаривањем информације о цензурисању и третирањем свих елемената као да су нецензурисани и коришћењем Монте Карло метода,

III брисањем цензурисаних елемената и рачунањем вредности тест статистике на основу узорка мањег обима и коришћењем Монте Карло метода.

Табела 5.4: $p-$вредности тестова

метод	$\widehat{M}_{c,1}^{\mathcal{P}}$	$\widehat{M}_{c,2}^{\mathcal{P}}$	$\widehat{M}_{c,5}^{\mathcal{P}}$	$\widehat{M}_{c,1}^{\mathcal{D}}$	$\widehat{M}_{c,2}^{\mathcal{D}}$	$\widehat{M}_{c,5}^{\mathcal{D}}$	$\widehat{L}_{c,1}^{\mathcal{P}}$	$\widehat{L}_{c,2}^{\mathcal{P}}$	$\widehat{L}_{c,5}^{\mathcal{P}}$	$\widehat{L}_{c,1}^{\mathcal{D}}$	$\widehat{L}_{c,2}^{\mathcal{D}}$	$\widehat{L}_{c,5}^{\mathcal{D}}$
I	0.000	0.000	0.000	0.164	0.170	0.122	0.012	0.013	0.015	0.013	0.009	0.027
II	0.007	0.005	0.008	0.194	0.172	0.123	0.013	0.013	0.016	0.012	0.012	0.012
III	0.003	0.004	0.005	0.211	0.182	0.131	0.006	0.007	0.008	0.016	0.014	0.014

Можемо приметити да већина тестова одбацује нулту хипотезу и да су у овом случају закључци сва три начина у већини случајева исти. У раду аутора (*Cuparić and Milošević (2020)*) одређене су $p-$вредности још неких тестова користећи бутстреп процедуру. Наведени пример је још један показатељ значаја постојања великог броја тестова заснованих на различитим својствима.

Треба нагласити да у општем случају на то да ли ће се једноставним уклањањем елемената узорака, или евентуално третирањем свих као да су нецензурисани, добити исти закључак као када се примени претходно описана процедура скалирања инверзом цензурисања утиче и проценат цензурисања као и врста цензурисања. У табели 5.5 приказан је симулиран узорак обима 50 који потиче из Вејбулове $W(1.4)$ расподеле са 40% цензурисаних елемената (цензурисани елементи су означени звездицом).

Табела 5.5: Симулирани узорак

0.011⋆	0.040⋆	0.043⋆	0.060⋆	0.094	0.095	0.096⋆	0.098⋆	0.133	0.138
0.170	0.183⋆	0.195⋆	0.211	0.213	0.275⋆	0.291⋆	0.293	0.341	0.363⋆
0.367	0.367	0.379	0.384⋆	0.410⋆	0.424	0.434⋆	0.440	0.486	0.493
0.496⋆	0.502	0.541	0.614⋆	0.627⋆	0.654	0.660	0.678	0.690	0.708
0.759⋆	0.794	0.824⋆	0.909	1.060	1.288	1.371	1.420	1.457	1.740⋆

У табели 5.6 приказане су $p-$вредности тестова за симулиран узорак одређене на раније наведена три начина (методи I – III). Ови методи су примењени у случају тестирања просте хипотезе $H_0 : X \in \mathcal{E}(1)$ и у случају тестирања сложене хипотезе $H_0 : X \in \mathcal{E}(\frac{1}{\mu}),\ \mu > 0$.

Табела 5.6: $p-$вредности тестова за симулирани узорак

H_0	метод	$\widehat{M}^{\mathcal{P}}_{c,1}$	$\widehat{M}^{\mathcal{P}}_{c,2}$	$\widehat{M}^{\mathcal{P}}_{c,5}$	$\widehat{M}^{\mathcal{D}}_{c,1}$	$\widehat{M}^{\mathcal{D}}_{c,2}$	$\widehat{M}^{\mathcal{D}}_{c,5}$	$\widehat{L}^{\mathcal{P}}_{c,1}$	$\widehat{L}^{\mathcal{P}}_{c,2}$	$\widehat{L}^{\mathcal{P}}_{c,5}$	$\widehat{L}^{\mathcal{D}}_{c,1}$	$\widehat{L}^{\mathcal{D}}_{c,2}$	$\widehat{L}^{\mathcal{D}}_{c,5}$
	I	0.019	0.013	0.016	0.005	0.004	0.003	0.014	0.018	0.013	0.029	0.033	0.028
проста	II	0.1182	0.161	0.230	0.137	0.151	0.209	0.199	0.241	0.298	0.199	0.229	0.285
	III	0.0372	0.052	0.084	0.022	0.034	0.063	0.074	0.092	0.124	0.061	0.080	0.113
	I	0.005	0.014	0.004	0.001	0.004	0.002	0.003	0.006	0.006	0.007	0.003	0.001
сложена	II	0.058	0.053	0.057	0.084	0.063	0.058	0.058	0.060	0.063	0.066	0.062	0.062
	III	0.015	0.017	0.017	0.012	0.012	0.015	0.017	0.018	0.020	0.014	0.015	0.019

Може се приметити да, ако је ниво значајности $\alpha = 0.05$ и тестира се проста нулта хипотеза, метод I даје тачан закључак о одбацивању нулте хипотезе експоненцијалности. Третирањем свих елемената као да су нецензурисани (метод II) добија се да не треба одбацити нулту хипотезу, док у случају када се избришу цензурисани елементи (метод III) различити тестови дају различите закључке. У случају када се тестира сложена хипотеза, методи I и II дају исте закључке које су давали и у случају тестирања просте хипотезе, док метод III у случају свих тест статистика даје закључак да не треба прихватити нулту хипотезу. С обзиром да када се избришу цензурисани елементи узорка добија се узорак од 30 елемената, а да су се наведени тестови показали као добри за тестирање сложене хипотезе у случају малих узорака (табела 3.1), одатле је и очекиван адекватан закључак коришћењем овог начина тестирања.

Глава 6

ЗАКЉУЧАК

Фокус дисертације је на формирању нових класа тестова сагласности са расподелом. У том циљу су предложене три нове класе тестова, две засноване на L^2 и једна заснована на L^∞ растојању између случајних величина из карактеризације. Да би се одредила асимптотска својства тестова L^2-типа коришћен је добијени теоријски резултат у вези са граничном расподелом дегенерисаних $V-$статистика са оцењеним параметром. Такође одређена је и асимптотска расподела тестова $L^\infty-$типа користећи својства $V-$емпријских процеса. За илустрацију резултата везаних за тестове коришћена је експоненцијална расподела и две карактеризације, Пури-Рубинова и Десуова. Као прва мера квалитета тестова коришћене су емпиријске моћи.

Користећи резултате везане за асимптотске расподеле, добијени су теоријски резултати везани за Бахадурову асимптотску ефикасност $V-$статистика са оцењеним параметром. На основу њих, одређена је друга мера квалитета нових тестова, приближна Бахадурова асимптотска релативна ефикасност. Поред тога, израчунате су и асимптотске ефикасности неких класичних и неких новијих тестова експоненцијалности.

Последњи део дисертације се односи на примену наведених тестова у случају цензурисаних података. Трансформисане су наведене класе тестова тако да могу да се користе у те сврхе и добијени теоријски резултати везани за граничну расподелу тих класа. Такође, развијена је модификација бутстрепа како би могао да се користи у новодобијеним условима за добијање емпиријских моћи. Коришћење тог новог метода је теоријски потврђено одређивањем асимптотских својстава.

Могући правци даљег истраживања су:

- формирање тестова L^2- и $L^\infty-$типа за друге карактеризације, као и за друге расподеле,
- одређивање асимптотских својстава у случају када узорак није прост случајан,
- адаптација тестова и на неке друге типове цензурисања, као што је рецимо интервално или зависно цензурисање,
- одређивање асимптотске ефикасности тестова примењених на цензурисане податке.

Литература

Aalen, O. (1975). *Statistical inference for a family of counting processes.* Ph. D. thesis, Department of Statistics, University of California, Berkeley.

Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics 6*(4), 701–726.

Abrahamson, I. G. (1965). *On the stochastic comparison of tests of hypotheses.* Ph. D. thesis, University of Chicago, Department of Statistics.

Ahmad, I. A. and I. A. Alwasel (1999). A goodness-of-fit test for exponentiality based on the memoryless property. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61*(3), 681–689.

Allison, J. and L. Santana (2015). On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests. *Journal of Statistical Computation and Simulation 85*(16), 3276–3288.

Altshuler, B. (1970). Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences 6*, 1–11.

Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (2012). *Statistical models based on counting processes.* Springer Science & Business Media.

Anderson, T. W. and D. A. Darling (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics 23*(2), 193–212.

Angus, J. E. (1982). Goodness-of-fit tests for exponentiality based on a loss-of-memory type functional equation. *Journal of Statistical Planning and Inference 6*(3), 241–251.

Arcones, M. A. (2007). Two tests for multivariate normality based on the characteristic function. *Mathematical Methods of Statistics 16*(3), 177–201.

Arnold, B. C. and J. A. Villasenor (2013). Exponential characterizations motivated by the structure of order statistics in samples of size two. *Statistics & Probability Letters 83*(2), 596–601.

Ash, R. B. and M. F. Gardner (1975). *Topics in Stochastic Processes.* Probability & Mathematical Statistics Monograph. New York: Academic Press.

Bagdonavicius, V., J. Kruopis, and M. S. Nikulin (2013). *Nonparametric tests for complete data.* John Wiley & Sons.

Bahadur, R. R. (1960). Stochastic comparison of tests. *Annals of Mathematical Statistics 31*(2), 276–295.

Bahadur, R. R. (1967). An optimal property of the likelihood ratio statistic. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, pp. 13–26.

Bahadur, R. R. (1971). *Some limit theorems in statistics.* SIAM, Philadelphia.

Bahadur, R. R. and P. J. Bickel (2009). An optimality property of Bayes' test statistics. *Lecture Notes-Monograph Series 57*, 18–30.

Baringhaus, L. and N. Henze (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics 43*(3), 551–564.

Billingsley, P. (1968). *Convergence of probability measures.* John Wiley & Sons.

Bliss, C. I. and W. Stevens (1937). The calculation of the time-mortality curve. *Annals of Applied Biology 24*(4), 815–852.

Božin, V., B. Milošević, Ya. Yu. Nikitin, and M. Obradović (2018). New characterization-based symmetry tests. *Bulletin of the Malaysian Mathematical Sciences Society 43*(1), 1–24.

Breslow, N. and J. Crowley (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of statistics 2*(3), 437–453.

Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics & Probability Letters 49*(2), 155–161.

Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics 23*(4), 493–507.

Choi, B., K. Kim, and S. Heun Song (2004). Goodness-of-fit test for exponentiality based on Kullback–Leibler information. *Communications in Statistics-Simulation and Computation 33*(2), 525–536.

Cohen, A. C. (1991). *Truncated and censored samples: theory and applications.* CRC press.

Colosimo, E., F. v. Ferreira, M. Oliveira, and C. Sousa (2002). Empirical comparisons between Kaplan-Meier and Nelson-Aalen survival function estimators. *Journal of Statistical Computation and Simulation 72*(4), 299–308.

Cox, D. and D. Oakes (1984). *Analysis of survival data.* Chapman and Hall, New York.

Cuparić, M. (2019). Approximate Bahadur efficiency of Henze-Meintanis exponentiality tests with comparison. *Matematički vesnik 71*, 169–179.

Cuparić, M. and B. Milošević (2020). New characterization based exponentiality tests for randomly censored data. *Preprint, arXiv:2011.07998.*

Cuparić, M., B. Milošević, Ya. Yu. Nikitin, and M. Obradović (2020). Some consistent exponentiality tests based on Puri-Rubin and Desu characterizations. *Applications of Mathematics 65*(3), 245–255.

Cuparic, M., B. Miloševic, and M. Obradovic (2019). New class of suprem-type exponentiality tests based on V-empirical Laplace transforms and Puri-Rubin characterization. In *Proceedings of the 21st European Young Statisticians Meeting*, pp. 15–19.

Cuparić, M., B. Milošević, and M. Obradović (2019a). New consistent exponentiality tests based on V-empirical Laplace transforms with comparison of efficiencies. *Preprint, arXiv:1904.00840.*

Cuparić, M., B. Milošević, and M. Obradović (2019b). New L^2-type exponentiality tests. *SORT 43*(1), 25–50.

Cuparić, M., B. Milošević, and M. Obradović (2020). Asymptotic distribution of certain degenerate V- and U-statistics with estimated parameters. *Preprint, arXiv:2012.11020.*

D'Agostino, R. B. and M. A. Stephens (1986). *Goodness-of-fit-techniques*, Volume 68. Marcel Dekker, INC.

de Wet, T. and R. H. Randles (1987). On the effect of substituting parameter estimators in limiting χ^2 U and V- statistics. *The Annals of Statistics 15*(1), 398–412.

Desu, M. M. (1971). A characterization of the exponential distribution by order statistics. *The Annals of Mathematical Statistics. 42*(2), 837–838.

Dhillon, B. (1981). Lifetime distributions. *IEEE Transactions on Reliability 30*, 457–459.

Ebrahimi, N., M. Habibullah, and E. S. Soofi (1992). Testing exponentiality based on Kullback-Leibler information. *Journal of the Royal Statistical Society: Series B (Methodological) 54*(3), 739–748.

Efron, B. (1967). The two sample problem with censored data. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Volume 4, pp. 831–853. University of California Press, Berkeley, CA.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association 76*(374), 312–319.

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association 83*(402), 414–425.

Epps, T. W. and L. B. Pulley (1983). A test for normality based on the empirical characteristic function. *Biometrika 70*(3), 723–726.

Epps, T. W. and L. B. Pulley (1986). A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function. *Journal of the Royal Statistical Society. Series B (Methodological) 48*(2), 206–213.

Epstein, B. (1954). Truncated life tests in the exponential case. *The Annals of Mathematical Statistics 25*(3), 555–564.

Feller, W. (1968). *An introduction to probability theory and its applications*, Volume I. John Wiley & Sons.

Fernholz, L. T. (2012). *Von Mises calculus for statistical functionals*, Volume 19. Springer Science & Business Media.

Fisher, R. A. (1931). The truncated normal distribution. *British Association for the Advancement of Science I*, XXXIII–XXXIV.

Fisher, R. A. and L. H. C. Tippett (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 24, pp. 180–190. Cambridge University Press.

Fleming, T. R. and D. P. Harrington (2011). *Counting processes and survival analysis*, Volume 169. John Wiley & Sons.

Fortiana, J. and A. Grané (2002). A scale-free goodness-of-fit statistic for the exponential distribution based on maximum correlations. *Journal of statistical planning and inference 108*(1-2), 85–97.

Fromont, M. and B. Laurent (2006). Adaptive goodness-of-fit tests in a density model. *The Annals of Statistics 34*(2), 680–720.

Gail, M. and J. Gastwirth (1978). A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic. *Journal of the Royal Statistical Society: Series B (Methodological) 40*(3), 350–357.

Galambos, J. and S. Kotz (1978). *Characterizations of Probability Distributions.: A Unified Approach with an Emphasis on Exponential and Related Models.*, Volume 675. Springer.

Galton, F. (1898). An examination into the registered speeds of american trotting horses, with remarks on their value as hereditary data. *Proceedings of the Royal Society of London 62*(379-387), 310–315.

Geurts, J. H. (1987). On the small-sample performance of Efron's and of Gill's version of the product limit estimator under nonproportional hazards. *Biometrics 43*(3), 683–692.

Gill, R. D. (1980). *Censoring and Stochastw Integrals.* Mathematical Centre Tracts, Vol. 124. Mathematisch Centrum, Amsterdam.

Gill, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *The Annals of Statistics 11*(1), 49–58.

Hald, A. (1949). Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point. *Scandinavian Actuarial Journal 1949*(1), 119–134.

Hall, W. J. and J. A. Wellner (1980). Confidence bands for a survival curve from censored data. *Biometrika 67*(1), 133–143.

Helmers, R., P. Janssen, and R. Serfling (1988). Glivenko-Cantelli properties of some generalized empirical df's and strong convergence of generalized L-statistics. *Probability theory and related fields 79*(1), 75–93.

Henze, N. (1992). A new flexible class of omnibus tests for exponentiality. *Communications in Statistics-Theory and Methods 22*(1), 115–133.

Henze, N. and S. Koch (2020). On a test of normality based on the empirical moment generating function. *Statistical Papers 61*(1), 17–29.

Henze, N. and S. Meintanis (2002a). Tests of fit for exponentiality based on the empirical Laplace transform. *Statistics: A Journal of Theoretical and Applied Statistics 36*(2), 147–161.

Henze, N. and S. G. Meintanis (2002b). Goodness-of-fit tests based on a new characterization of the exponential distribution. *Communications in Statistics-Theory and Methods 31*(9), 1479–1497.

Henze, N. and S. G. Meintanis (2005). Recent and classical tests for exponentiality: a partial review with comparisons. *Metrika 61*(1), 29–45.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *The annals of mathematical statistics 19*(3), 293–325.

Iverson, H. and R. Randles (1989). The effects on convergence of substituting parameter estimates into U-statistics and other families of statistics. *Probability Theory and Related Fields 81*(3), 453–471.

Jevremovic, V. (1991). A note on mixed exponential distribution with negative weights. *Statistics & probability letters 11*(3), 259–265.

Jiménez-Gamero, M., B. Milošević, and M. Obradović (2020). Exponentiality tests based on basu characterization. *Statistics 54*(4), 714–736.

Jovanović, M., B. Milošević, Ya. Yu. Nikitin, M. Obradović, and K. Yu. Volkova (2015). Tests of exponentiality based on Arnold–Villasenor characterization and their efficiencies. *Computational Statistics & Data Analysis 90*, 100–113.

Kalbfleisch, J. D. and R. L. Prentice (2011). *The statistical analysis of failure time data*, Volume 360. John Wiley & Sons.

Kallenberg, W. C. (1983). Intermediate efficiency, theory and examples. *The Annals of Statistics 11*(1), 170–182.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association 53*(282), 457–481.

Karatzas, I. and S. Shreve (1991). *Brownian Motion and Stochastic Calculus*, Volume 113. Springer Science & Business Media.

Klar, B. (2001). Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function. *Annals of the Institute of Statistical Mathematics 53*(2), 338–353.

Klein, J. P. (1991). Small sample moments of some estimators of the variance of the Kaplan-Meier and Nelson-Aalen estimators. *Scandinavian Journal of Statistics 18*(4), 333–340.

Kondo, T. (1930). A theory of the sampling distribution of standard deviations. *Biometrika 22*(1/2), 36–64.

Korolyuk, V. S. and Y. V. Borovskikh (1994). *Theory of U-statistics.* Kluwer, Dordrecht.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference.* Springer.

Koziol, J. A. and S. B. Green (1976). A Cramer-von Mises statistic for randomly censored data. *Biometrika 63*(3), 465–474.

Lagakos, S. W., L. M. Barraj, and V. d. Gruttola (1988). Nonparametric analysis of truncated survival data, with application to aids. *Biometrika 75*(3), 515–523.

Lawless, J. F. (2011). *Statistical models and methods for lifetime data*, Volume 362. John Wiley & Sons.

Ledoux, M. and M. Talagrand (2013). *Probability in Banach Spaces: isoperimetry and processes.* Springer Science & Business Media.

Lee, A. (1914). Table of the Gaussian "tail" functions; when the "tail" is larger than the body. *Biometrika 10*(2/3), 208–214.

Lehmann, E. L. (2004). *Elements of large-sample theory.* Springer Science & Business Media.

Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association 64*(325), 387–389.

Mansournia, M. A. and D. G. Altman (2016). Inverse probability weighting. *Bmj 352*, i189.

Meeker, W. Q. and L. A. Escobar (2014). *Statistical methods for reliability data.* John Wiley & Sons.

Mendenhall, W. (1958). A bibliography on life testing and related topics. *Biometrika 45*(3/4), 521–543.

Milošević, B. (2016). Asymptotic efficiency of new exponentiality tests based on a characterization. *Metrika 79*(2), 221–236.

Miloševic, B. (2017). Some recent characterization based goodness of fit tests. In *the Proceedings of the 20th European Young Statisticians Meeting*, pp. 67–73.

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

Milošević, B. and M. Obradović (2016b). Some characterization based exponentiality tests and their Bahadur efficiencies. *Publications de L'Institut Mathematique 100*(114), 107–117.

Milošević, B. and M. Obradović (2016). Some characterizations of the exponential distribution based on order statistics. *Applicable Analysis and Discrete Mathematics 10*(2), 394–407.

Mood, A. M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. *The Annals of Mathematical Statistics 25*(3), 514–522.

Moran, P. (1951). The random division of an interval – Part II. *Journal of the Royal Statistical Society: Series B (Methodological) 13*(1), 147–150.

Murray, D. L. and B. K. Sandercock (2020). *Population Ecology in Practice: Underused, Misused and Abused Methods.* John Wiley & Sons, Incorporated.

Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology 1*(1), 27–52.

Nikitin, Ya. Yu. (1995). *Asymptotic efficiency of nonparametric tests.* Cambridge University Press, New York.

Nikitin, Ya. Yu. (1996). Bahadur efficiency of a test of exponentiality based on a loss-of-memory type functional equation. *Journal of Nonparametric Statistics 6*(1), 13–26.

Nikitin, Ya. Yu. (2017). Tests based on characterizations, and their efficiencies: a survey. *Preprint, arXiv:1707.01522.*

Nikitin, Ya. Yu. and I. Peaucelle (2004). Efficiency and local optimality of nonparametric tests based on U- and V-statistics. *Metron 62*(2), 185–200.

Nikitin, Ya. Yu. and E. Ponikarov (2001). Rough asymptotics of probabilities of Chernoff type large deviations for von Mises functionals and U-statistics. *Translations of the American Mathematical Society-Series 2 203*, 107–146.

Nikitin, Ya. Yu. and A. Tchirina (1996). Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic. *Journal of the Italian Statistical Society 5*(1), 163–175.

Nikitin, Ya. Yu. and A. Tchirina (2007). Lilliefors test for exponentiality: large deviations, asymptotic efficiency, and conditions of local optimality. *Mathematical Methods of Statistics 16*(1), 16–24.

Obradović, M. (2015a). On asymptotic efficiency of goodness of fit tests for pareto distribution based on characterizations. *Filomat 29*(10), 2311–2324.

Obradović, M. (2015b). Three characterizations of exponential distribution involving median of sample of size three. *Journal of Statistical Theory and Applications 14*(3), 257–264.

Obradović, M., M. Jovanović, and B. Milošević (2015). Goodness-of-fit tests for Pareto distribution based on a characterization and their asymptotics. *Statistics 49*(5), 1026–1041.

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50*(302), 157–175.

Pearson, K. (1902). On the systematic fitting of curves to observations and measurements. *Biometrika 1*(3), 265–303.

Pearson, K. and A. Lee (1908). On the generalised probable error in multiple normal correlation. *Biometrika 6*(1), 59–68.

Pejchal, V., G. Žagar, R. Charvet, C. Dénéréaz, and A. Mortensen (2017). Compression testing spherical particles for strength: Theory of the meridian crack test and implementation for microscopic fused quartz. *Journal of the Mechanics and Physics of Solids 99*, 70–92.

Puri, P. S. and H. Rubin (1970). A characterization based on the absolute difference of two iid random variables. *The Annals of Mathematical Statistics 41*(6), 2113–2122.

Randles, R. H. (1982). On the asymptotic normality of statistics with estimated parameters. *The Annals of Statistics 10*(2), 462–474.

Robins, J. M. and A. Rotnitzky (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS epidemiology*, pp. 297–331. Springer.

Rublík, F. (1989). On optimality of the LR tests in the sense of exact slopes. I. General case. *Kybernetika 25*(1), 13–14.

Serfling, R. J. (1984). Generalized L-, M-, and R-statistics. *The Annals of Statistics 12*(1), 76–86.

Serfling, R. J. (2009). *Approximation theorems of mathematical statistics*, Volume 162. John Wiley & Sons.

Silverman, B. W. (1983). Convergence of a class of empirical distribution functions of dependent random variables. *The Annals of Probability 11*(3), 745–751.

Stute, W. (1992). Strong consistency of the MLE under random censoring. *Metrika 39*(1), 257–267.

Sukhatme, P. V. (1937). Tests of significance for samples of the χ^2-population with two degrees of freedom. *Annals of Eugenics 8*(1), 52–56.

Sun, J. (2007). *The statistical analysis of interval-censored failure time data*. Springer.

Tchirina, A. (2005). Bahadur efficiency and local optimality of a test for exponentiality based on the Moran statistics. *Journal of Mathematical Sciences 127*(1), 1812–1819.

Tenreiro, C. (2011). An affine invariant multiple test procedure for assessing multivariate normality. *Computational statistics & data analysis 55*(5), 1980–1992.

Tenreiro, C. (2019). On the automatic selection of the tuning parameter appearing in certain families of goodness-of-fit tests. *Journal of Statistical Computation and Simulation 89*(10), 1780–1797.

Torabi, H., N. H. Montazeri, and A. Grané (2018). A wide review on exponentiality tests and two competitive proposals with application on reliability. *Journal of Statistical Computation and Simulation 88*(1), 108–139.

Van Trees, H. L. and K. L. Bell (2013). *Detection estimation and modulation theory*. Wiley.

Volkova, K. Yu. (2015). Goodness-of-fit tests for exponentiality based on Yanev-Chakraborty characterization and their efficiencies. In *the Proceedings of the 19th European Young Statisticians Meeting, Prague*, pp. 156–159.

von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *The Annals of Mathematical Statistics 18*(3), 309–348.

Wagner, S. S. and S. A. Altmann (1973). What time do the baboons come down from the trees?(an estimation problem). *Biometrics 29*(4), 623–635.

Wald, A. and J. Wolfowitz (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics 11*(2), 147–162.

Wang, J. G. (1987). A note on the uniform consistency of the Kaplan-Meier estimator. *The Annals of Statistics 15*(3), 1313–1316.

Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer.

Widder, D. V. (1946). *The Laplace transform*. Princeton university press.

Yanev, G. P. and S. Chakraborty (2013). Characterizations of exponential distribution based on sample of size three. *Pliska Studia Mathematica Bulgarica 22*(1), 237p–244p.

Ying, Z. (1989). A note on the asymptotic properties of the product-limit estimator on the whole line. *Statistics & probability letters 7*(4), 311–314.

Zolotarev, V. M. (1961). Concerning a certain probability problem. *Theory of Probability & Its Applications 6*(2), 201–204.

Биографија аутора

Марија Цупарић (рођена Радичевић) рођена је 27.3.1992. године у Чачку, где је завршила основну школу ,,Вук Караџић" и Гимназију. Основне студије на Математичком факултету, Универзитета у Београду, на смеру Статистика, актуарска и финансијска математика завршила је 2014. године са просечном оценом 9.02. Исте године је уписала мастер студије на Математичком факултету, студијски програм Математика, модул Статистика, актуарска и финансијска математика и положила све испите са просечном оценом 10. Мастер рад под називом ,,Анализа ризика у животном осигурању" одбранила је 2015. године, под руководством проф. др Павла Младеновића. Докторске студије на Математичком факултету у Београду, студијски програм Математика, уписала је 2015. године. Све испите предвиђене планом положила је са оценом 10.

Од 2014. године запослена је на Математичком факултету, прво као сарадник у настави (2014-2016) а затим и као асистент (2016-до данас) за ужу научну област Вероватноћа и статистика. Држала је вежбе на курсевима: Вероватноћа и статистика А, Вероватноћа и статистика Б, Статистички софтвер 3, Статистички софтвер 4, Математичка статистика, Одабрана поглавља математичке статистике, Биостатистика, Животно осигурање, Линеарни статистички модели и Елементи финансијске математике.

Објавила је три научна рада. Имала је саопштења на седам конференција, од чега су две по позиву. Учествовала је на две летње школе и била учесник пројекта 174012 Министарства просвете, науке и технолошког развоја Републике Србије.

Прилог 1.

Изјава о ауторству

Име и презиме аутора Марија Цупарић

Број индекса 2008/2015

Изјављујем

да је докторска дисертација под насловом

Тестови сагласности засновани на L^2 и L^∞ растојањима и њихова асимптотска ефикасност

- резултат сопственог истраживачког рада,
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 10.03.2021.

М Цупарић

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Марија Цупарић

Број индекса 2008/2015

Студијски програм Математика

Наслов рада Тестови сагласности засновани на L^2 и L^∞ растојањима и њихова асимптотска ефикасност

Ментор др Бојана Милошевић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, 10.03.2021

МЦупарић

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Тестови сагласности засновани на L^2 и L^∞ растојањима и њихова асимптотска ефикасност

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прераде (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прераде (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 10.03.2021.

Мијупарић

1. **Ауторство**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство - некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство - некомерцијално - без прераде**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство - некомерцијално - делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство - без прераде**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство - делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.