LATVIJAS UNIVERSITĀTE
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE
MATEMĀTIKAS NODAĻA

NEIMAŅA TESTU ANALĪZE UN PIELIETOJUMI

MAĢISTRA DARBS

Autors: **Sintija Cīrule**

Stud. apl. MaSt010038

Darba vadītājs: doc. Dr.math. Jānis Valeinis

RĪGA 2009

Anotācija

Maģistra darbs veltīts Neimaņa nogludināšanas testiem vienkāršām un saliktām hipotēzēm. Vienkāršas hipotēzes gadījumā testu izmanto vienmērīguma pārbaudei, bet saliktas hipotēzes gadījumā - ekponencialitātes un normalitātes pārbaudei. Darbā tiek apskatīta Neimaņu testu teorija apraksts, to asimptotiskā uzvedība pie nulles hipotēzes un alternatīvām. Veiktas Monte Karlo simulācijas, lai iegūtu kritiskās vērtības un jaudas pie plaša apjoma alternatīvām. Darbā empīriski salīdzināti Neimaņa testi ar matemātiskajā statistiskā plaši izmantoto Pīrsona χ^2 testu, Kolmogorova - Smirnova testu, Andersona - Darlinga un citiem testiem. Darbam pievienotas izveidotās programmas, lai iegūtu Neimaņa testa analīzei un pielietojumam nepieciešamos rezultātus.

Atslēgas vārdi: Nogludināšanas tests, Neimaņa tests, Švarca kritērijs, datu nosakošā metode

Abstract

This work is devoted to the Neyman smooth tests for simple and composite hypotheses. In case of the simple hypothesis test can be used for testing uniformity. For the composite hypothesis exponentiality and normality is cosidered. The asymptotic behavior of the Neyman test under null hypothesis and alternatives are discussed theoretically. Empirically Neyman test and classical Pearson χ^2 test, Kolmogorov - Smirnov test, Anderson Darling and some others tests have been compared.

Keywords: Smooth test, Neymans test, Schwarzs rule, data driven procedure

Saturs

Ievads **2**

1. NEIMAŅA TESTS VIENKĀRŠĀM HIPOTĒZĒM **4**

1.1. Neimaņa tests 4

1.2. Neimaņa testa asimptotiskā uzvedība pie nulles hipotēzes 6

1.3. Neimaņa testa asimptotiskā uzvedība pie alternatīvām 7

1.4. Simulācijas 8

1.4.1. Testu statistikas salīdzinājumiem 8

1.4.2. Alternatīvie sadalījumi 9

1.4.3. Kritiskās vērtības 9

1.4.4. Simulētās jaudas 10

2. NEIMAŅA TESTS SALIKTĀM HIPOTĒZĒM **13**

2.1. Neimaņa tests 13

2.2. Selekcijas kritēriji 16

2.2.1. Selekcijas kritēriju definēšana 16

2.2.2. Selekcijas kritēriju sākotnējais sadalījums 18

2.2.3. Selekcijas kritēriji pie alternatīvām 18

2.3. Datus nosakošie nogludināšanas testi saliktām hipotēzēm 19

2.4. Simulācijas 19

2.4.1. Alternatīvie sadalījumi 19

2.4.2. Kritiskās vērtības 20

2.4.3. Simulētās jaudas 21

Secinājumi **25**

Izmantotā literatūra un avoti **26**

Pielikums. Izveidotās programmas **27**

Ievads

Klasiska problēma matemātiskajā statistikā ir pārbaudīt izlases sadalījumu atbilstību kādai parametriskai sadalījumu saimei.

Angļu matemātiķis Karls Pīrsons 1900.gadā ieviesa Hī kvadrāta (χ^2) testu, kurš varbūt arī vēl joprojām ir populārākais zināmais atbilstības tests matemātiskajā statistikā.

Citi bieži izmantotie testi ir Kolmogorova - Smirnova tests, kas definēts 1933.gadā, Kramera - von Mises tests, kuru 1928.gadā ieviesa Kramers un Smirnovs 1936.gadā veica korekcijas. Šie testi var tikt pielietoti gan vienkāršām, gan saliktām hipotēzēm. Nogludināšanas testus vienmērīgajam sadalījumam intervālā [0,1] ieviesa Neimanis 1937.gadā [1].

Apskatīsim sekojošu vienkāršu hipotēžu pārbaudi par vienmērīgo sadalījumu

$$H_0 : X \sim v.s.[0,1] \text{ pret } H_1 : X \nsim v.s.[0,1].$$

Jebkuru citu vienkāršu hipotēžu pārbaudi par sadalījuma veidu var reducēt uz šo hipotēzi, sekojošā veidā. Ja $X_1, \ldots, X_n \sim F$, kur F ir sadalījuma funkcija, tad lieto transformāciju $F(X_1), \ldots, F(X_n) \sim v.s.[0,1]$.

Neimaņa testa statistika tiek bāzēta uz ortonormāliem polinomiem. Līdz pat 1994.gadam nebija skaidrs, cik daudz komponentes ortonormālā bāzē izvēlēties, lai testa jauda būtu maksimāla. Atkarībā no izejas datiem, bija dažādas rekomendācijas, tomēr izvēloties ne vairās kā divas, trīs komponentes. Neimaņa testu sauc par nogludināšanas testu, jo tas konstruēts kā maksimālās jaudas tests eksponenciālās saimes alternatīvām, kas modelē gludas nobīdes no nulles hipotēzes.

Literatūrā nozīmīgajam jautājumam par to, cik lielam jābūt komponentu skaitam, tiek piedāvātas trīs atšķirīgas pieejas. Pirmā pieeja piedāvā izmantot ortonormālas sistēmas tikai pirmās komponentes, un nulles hipotēzi pārbaudīt ar cik vien iespējams mazu komponenti. Šo pieeju plaši aprakstījis un pamatojis Inglots 1994.gadā [2].

Otrā pieeja ir vairāk analītiska pieeja. Gadījumā, kad nulles hipotēze tiek noraidīta, tiek izmantotas komponentes, lai secinātu par nobīdes iemeslu no nulles hipotēzes. Šo pieeju ir aprakstījis Rainers un Bests 1989.gadā [3].

Trešā pieeja saistīta ar automātisku komponenšu izvēli, kas balstās uz datiem. Šādu pieeju nodrošina Švarca selekcijas kritērija izmantošana. 1994.gada Ledvinas publikācijā [4] tiek piedāvāta jauna datu nosakošā metode komponentu skaita izvēlei Neimaņa testā.

Atšķirībā no iepriekšminētajiem priekšlikumiem, tiek ieteikta uz datiem balstīta komponentu izvēle. Šajā metodē tiek izmantots Švarca Baijesa informācijas kritērija (BIC) process, lai datiem izvēlētos eksponenciālā modeļa dimensiju, bet pēc tam izmantotu izvēlēto dimensiju kā komponentu skaitu. Šādā veidā jaunā procedūra paļaujas uz modeļa atbilstību datiem un pārbauda, vai ir būtiska atšķirība starp modeli un vienmērīgo sadalījumu.

Šo metodi izmanto gan vienkāršām, gan saliktām hipotēzēm. Pie saliktām hipotēzēm izmantojot Neimaņa testu pārbaudām eksponencialitāti un normalitāti. Kā arī tiek ieviestas Švarca likuma vienkāršākas modifikācijas [5].

Galvenais darba mērķis ir veikt empīrisku Neimaņa testa analīzi pie nulles un alternatīvās hipotēzes maza vai vidēja apjoma izlasēm. Darbā salīdzināsim matemātiskajā statistiskā plaši izmantotos Pīrsona χ^2 testu, Kolmogorova - Smirnova testu, Andersona Darlinga un citus ar mazāk zināmo Neimaņa nogludināšanas testu vienkāršu un saliktu hipotēžu gadījumos.

Darbs ir strukturizēts sekojoši. 1.nodaļa veltīta vienkāršām hipotēzēm, detalizēti apskatīta datu nosakošā metode, lai izvēlētos komponentu skaitu Neimaņa testā vienmērības pārbaudei. Aprakstīts Švarca selekcijas likums un asiptotika. Veiktas Monte Karlo simulācijas,lai iegūtu Neimaņa testa kritiskās vērtības un jaudas pie dažādām alternatīvām. Veikta Neimaņa testa salīdzināšana ar matemātiskajā statistiskā biežāk izmantotajiem testiem. 2.nodaļa veltīta saliktām hipotēzēm. Definēts Neimaņa tests normalitātes un eksponcialitātes pārbaudīšanai, apskatītas Švarca kritērija divas vienkāršākas modifikācijas. Definēti selekcijas kritēriji, pieņemot, apskatīta sākotnējā sadalījuma asiptotika un selekcijas kritēriju uzvedība pie alternatīvām. Veiktas Monte Karlo simulācijas eksponenciālā un normālā sadalījuma gadījumā, lai iegūtu kritiskās vērtības un jaudas pie plaša apjoma alternatīvām. Veikta Neimaņa testa salīdzināšana ar matemātiskajā statistiskā biežāk izmantotajiem testiem. Noslēdzot darbu, pielikumā programmas R kodi.

1. NEIMAŅA TESTS VIENKĀRŠĀM HIPOTĒZĒM

1.1. Neimaņa tests

Nogludinošais Neimaņa tests vienmērības pārbaudei intervālā [0,1] noraida nulles hipotēzi H_0 pie lielām sekojošas statistikas vērtībām

$$N_k = \sum_{j=1}^{k} \left(n^{-1/2} \sum_{i=1}^{n} \phi_j(X_i) \right)^2, \tag{1.1.1}$$

kur $X_1, X_2, \ldots, X_n$ ir neatkarīgi un vienādi sadalīti gadījuma lielumi un $\phi_1, \phi_2, ...$ ir normalizēti Ležandra polinomi slēgtā intervālā $L_2([0,1])$ ar $\phi_0 \equiv 1$. Pirmie četri Ležandra polinomi ir šādi

$$\phi_1 = (x - 0.5)\sqrt{12};$$

$$\phi_2 = \sqrt{5}(6(x - 0.5)^2 - 0.5);$$

$$\phi_3 = \sqrt{7}(20(x - 0.5)^3 - 3(x - 0.5));$$

$$\phi_4 = 210(x - 0.5)^4 - 45(x - 0.5)^2 + 9/8.$$

Kad konstruē Neimaņa nogludināšanas testu, tiek aizstāta vienmērīguma pārbaude intervālā [0,1] (hipotēze H_0) ar parametrisku problēmu $H : \theta = 0$ pret $A : \theta \neq 0$ eksponenciālā saimē ar kārtu k, kas ir

$$p_\theta(x) = \exp\{\theta \circ \phi(x) - \psi_k(\theta)\}, \tag{1.1.2}$$

kur

$$\theta = (\theta_1, \ldots, \theta_k), \quad \phi = (\phi_1, \ldots, \phi_k),$$

$$\psi_k(\theta) = \log \int_0^1 \exp\{\theta \circ \phi(x)\} dx \tag{1.1.3}$$

un $\circ$ ir R^k iekšējais reizinājums.

Vienkārša eksponenciālā modeļa (1.1.2) vietā, tiek piedāvāts apskatīt eksponenciālā modeļa saimi $p_{\theta^{(k)}}(x), 1 \leq k \leq K$, K-fiksēts un izvēlas šajā saimē atbilstošu biežumu esošajiem datiem ar Švarca BIC. Tā kā tiek izmantoti jau eksistējoši rezultāti no BIC, tad daudz piemērotāks ir K-dimensionāla eksponenciālā modeļa apakšmodelis. Šī iemesla dēļ jāaplūko modeļi $\mu_s, 1 \leq s \leq K$, kas definēti ar sekojošu blīvuma funkciju intervālā [0,1],

$$\mu_s : p_\theta(x) = \exp\{\theta \circ \phi(x) - \psi_k(\theta)\}, \tag{1.1.4}$$

kur $x \in [0,1]$, $\theta = (\theta_1, \ldots, \theta_K) \in \Omega_s \subset \mathbb{R}^k$, $\Omega_s = \{\theta \in \mathbb{R}^k : \theta_i = 0, i = s+1, ..., K\}$, un ar BIC no $\cup_{s=1}^{K} \mu_s$ izvēlas atbilstošu biežumu esošajiem datiem.

Tālāk definē

$$Y_n = (\bar{\phi}_1, \ldots, \bar{\phi}_k), \qquad \bar{\phi}_j = n^{-1} \sum_{i=1}^{n} \phi_j(X_i), \tag{1.1.5}$$

$$L_s = n \sup_{\theta \in \ \Omega_s} \{Y_n \circ \theta - \psi_s(\theta)\}, \quad \mathcal{L}_s = L_s - \frac{1}{2} s \log n. \tag{1.1.6}$$

Švarca likums izvēlās modeli ar indeksu S, definētu sekojoši

$$S = \min\{j, 1 \leq j \leq K : \mathcal{L}_j = \max_{1 \leq s \leq K} \mathcal{L}_s.\} \tag{1.1.7}$$

Bet vēlāk 1996.gadā Inglot un Ledvina [5] ievēroja, ka ticamības funkcijas novirzes statistiskai $n \sup_{\theta \in \ \mathbb{R}^k} \{Y_n \circ \theta - \psi_k(\theta)\}$ var iegūt, analizējot daudz vienkāršāku statistiku $\frac{1}{2} n \|\bar{\phi}\|^2_{(k)}$, kur ar $\| \cdot \|$ apzīmē Eiklīda normu. Švarca likums S, kas uzdots ar (1.1.7), salīdzina maksimizētās ticamības funkcijas vērtības. Maksimizētā ticamības funkcija, ir ekvivalenta $\frac{1}{2} n \|Y_n(\beta)\|^2$. Līdz ar to parasti tiek lietota sekojoša Švarca likuma modifikācija

$$\begin{aligned} S2 = \min\{k : 1 \leq k \leq K, n \left\|Y_n\right\|^2_{(k)} - k \log n) \\ \geq n \left\|Y_n\right\|^2_{(j)} - j \log n, j = 1, ..., K\}, \end{aligned} \tag{1.1.8}$$

kur normas indekss nozīmē dimensiju. Neimaņa datu nosakošā nogludinošā testa statistika tiek definēta kā

$$N_S = \sum_{j=1}^{S} \left(n^{-1/2} \sum_{i=1}^{n} \phi_j(X_i) \right)^2 \tag{1.1.9}$$

ar S, kas uzdots formulā (1.1.7).

Izvēloties S dimensijas modeli, ir iegūta jauna Neimaņa testa versija N_S. Kā pamatojums šai izvēlei, tiek dotas likuma (1.1.9) un N_S galvenās īpašības teorēmu veidā.

Iepriekš tiek pieņemts, ka $p(\cdot,\theta) \in \mu_s$ dažiem $s = 1,...,K$. Ar P_θ apzīmē eksponenciālās saimes novēroto neatkarīgi un vienādi sadalīto mainīgo ticamības likumu, un kopa P_θ^n apzīmē $X_1,\ldots,X_n$ kopīgo likumu.

Teorēma 1. *Pieņemot, ka novēroto gadījuma lielumu likums pieder pie modeļa μ_s pie $s = 1,...,K$, tad ar $P_\theta^n \to 1$, kad $n \to \infty$, likums (1.1.7) noved pie zemākās dimensijas modeļa izvēles, saturot īsto θ vērtību.*

Teorēma 2. *Ja $\theta \in \Omega_s \setminus \Omega_{s-1}$ pie $s = 1,...,K$, tad*

$$P_\theta^n(S < s) = O(e^{-n^\beta}) \text{ pie } \beta > 0,$$

un pie $1 \le i \le K - s$ tas nozīmē, ka

$$P_\theta^n(S \ge s + i) = O(n^{-i/2}(\log n)^{(s+i-2)/2}).$$

Teorēma 3. *Tiek pieņemts, ka H_0 ir patiesa. Tad pie $n \to \infty$ N_S sadalījums aproksimē χ^2 sadalījumu ar vienu brīvības pakāpi.*

Teorēmu pierādījumus skatīties [6].

1.2. Neimaņa testa asimptotiskā uzvedība pie nulles hipotēzes

Lai pierādītu Neimaņa testa N_S atbilstību, nepieciešama informācija par N_S uzvedību gan pie nulles hipotēzes, gan pie alternatīvām. Aplūkojam sākotnējo gadījumu un S asimptotiku, S var būt neatkarīgs. P_0 nozīmē, ka X_i ir vienmērīgi sadalīts intervālā [0,1].

$$P_0(S = 1) = 1 - \sum_{k=2}^{d(n)} P_0(S = k) \tag{1.2.1}$$

un izmantojot, ka $\psi_1(0) = 0$,

$P_0(S = k) \le P_0(L_k \ge L_1)$

$$\le P_0\left(n \sup_{v \in R^k}\{Y_n \circ v - \psi_k(v)\} \ge \frac{1}{2}(k-1)\log n\right). \tag{1.2.2}$$

Nākamo definē

$$\mathbb{V}_k = \max_{1 \le j \le k} \sup_{x \in [0,1]} |\phi_j(x)|. \tag{1.2.3}$$

Ortonormāliem Ležandra polinomiem intervālā [0,1]

$$\mathbb{V}_k = (2k+1)^{1/2}. \tag{1.2.4}$$

Sekojošā teorēmā ir dota S asimptotika.

Teorēma 4. *Pieņem, ka*

$$\lim_{n\to\infty} d(n)V_{d(n)}(n^{-1}\log n)^{1/2} = 0 \tag{1.2.5}$$

Tad

$$\lim_{n\to\infty} P_0(S=1) = 1. \tag{1.2.6}$$

Teorēma 5. *Ja* (1.2.5) *spēkā, tad pie* H_0

$$N_S \rightarrow_d \chi_1^2. \tag{1.2.7}$$

Teorēmu pierādījumi [7], 1598., 1599.lpp.

1.3. Neimaņa testa asimptotiskā uzvedība pie alternatīvām

Dots $X_1, X_2, \ldots, X_n$ - neatkarīgi un vienādi sadalīti gadījuma lielumi, katrs ar sadalījumu P intervālā [0,1]. Pieņem, ka

$$E_P\phi_1(X) = ... = E_P\phi_{K-1}(X) = 0, \quad E_P\phi_K(X) \neq 0, \tag{1.3.1}$$

$K = K(P)$. N_S atbilstība pierādīta formas (1.3.1) jebkurai alternatīvai. Pieņem, ka

$$\lim_{n\to\infty} \inf d(n) \geq K, \tag{1.3.2}$$

ja $\lim_{n\to\infty} d(n) = \infty$, pie fiksēta K. Nākošā teorēmā ir aprakstīta S asimptotiskā uzvedība.

Teorēma 6. *Ja* (1.3.1) *spēkā, tad*

$$\lim_{n\to\infty} P(S \geq K) = 1. \tag{1.3.3}$$

Teorēma 7. *Ja* (1.3.1) *spēkā, tad*

$$N_S \rightarrow_P \infty, kad\ n \rightarrow \infty. \tag{1.3.4}$$

Teorēma 8. *Ja N_S ir ierobežots ar varbūtību pie H_0, kad $n \to \infty$, tad katrai fiksētai alternatīvai no* (1.3.1)*, testa jauda, kuras pamatā ir N_S, tiecās uz 1, kad $n \to \infty$.*

Teorēma 9. *Ja* (1.2.5) *spēkā, tad tests, kura pamatā ir N_S ir atbilstošs pret jebkuru alternatīvu no* (1.3.1).

Teorēmu pierādījumi [7], 1600., 1601.lpp.

1.4. Simulācijas

1.4.1. Testu statistikas salīdzinājumiem

Pīrsona hī kvadrāta (χ^2) statistika

$X_1, \ldots, X_n$ ir gadījuma izlase ar sadalījuma funkciju $F(x)$. Virkne X_j tiek sadalīta M intrvālos, piemēram $E_1, ..., E_M$. Ja $N_1, \ldots, N_M$ ir X_i novērtējumi šajā intervālā, tad N_i ir bionimāli sadalīts ar parametru n un $p_i = P(X_j \mid E_i) = \int_{E_i} dF(x)$, kad spēkā ir nulles hipotēze. Pīrsona χ^2 statistika ir sekojoša

$$T = \sum_{i=1}^{M} \frac{(N_i - np_i)^2}{np_i}. \tag{1.4.1}$$

Kolmogorova - Smirnova statistika

Empīriskā un teorētiskā sadalījuma salīdzināšanai pēc Kolmogorova - Smirnova metodes aprēķina šādu satitistiku

$$K(x) = \sqrt{n} \sup_x |F_n(x) - F(x)|, \tag{1.4.2}$$

kur $F_n(x)$ - empīriskā sadalījuma funkcija.

Modificētā Watsona statistika

$$U^2 = n \int_0^1 (F_n(t) - t)^2 dt - n \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{2} \right)^2, \tag{1.4.3}$$

kur F_n ir empīriska sadalījuma funkcija, tas ir $F_n(x) = n^{-1} \sum_{i=1}^{n} 1_{\{X_i \leq X\}}$. Turklāt, izmantojam modificētu statistikas U^2 formu, kuru apzīmē ar W un uzdod ar

$$W = \left(U^2 - \frac{0.1}{n} + \frac{0.1}{n^2} \right) \left(1 + \frac{0.8}{n} \right). \tag{1.4.4}$$

Statistikas W priekšrocība ir, ka tās kritiskās vērtības praktiski ir neatkarīgas no n.

Andersona-Darlinga statistika

Anderson-Darling tests ir Kolmogorova-Smirnova testa modifikācija.

$$AD = n \int_0^1 \frac{(F_n(t) - t)^2}{t(1-t)} dt, \tag{1.4.5}$$

arī šeit F_n ir empīriska sadalījuma funkcija.

Neuhausa statistika

Neuhaus šo statistiku ieviesa 1988.gadā.

$$N = \sum_{j=1}^{\infty} \lambda_j \left(n^{-1/2} \sum_{i=1}^{n} \sqrt{2} \cos(\pi j X_i) \right)^2, \tag{1.4.6}$$

kur

$$\lambda_j = \left\{ \left(\frac{4}{\pi j a} \right) \sin \left(\frac{\pi j a}{4} \right) \right\}^4, \quad a = 0.4$$

1.4.2. Alternatīvie sadalījumi

Simulācijās izmantotas šādas daudzveidīgas alternatīvas ar dažāda veida blīvuma funkcijām

$$g_1(x) = \frac{1}{2} + \frac{1}{20}(x^{-4/5} + (1-x)^{-4/5})$$
$$g_2(x) = \frac{1}{4}(x^{-1/2} + (1-x)^{-1/2})$$
$$g_3(x) = 1 + d\cos(\pi f x),$$
$$g_4(x) = \exp\left\{ \sum_{j=1}^{k} \theta_j \phi_j(x) - \psi_k(\theta) \right\}$$

ar ortonormāliem Ležendra polinomiem intervālā [0,1]

1.4.3. Kritiskās vērtības

Sākumā pārbaudīsim ar Monte Karlo simulāciju palīdzību Teorēmas 7. apgalvojumu, ka selekcijas likums $S2$ pie nulles hipotēzes konverģē ar varbūtību 1. Aplūkosim selekcijas likuma $S2$ sadalījumu pie alternatīvas $g_4(x)$ ar dažāda veida parametriem. Visas simulācijas tika veiktas 10 000 reižu. Simulēšanas rezultāti attēloti 1.1. tabulā. Redzams, ka patiešām, ja izlases apjoms $n = 100$, tad notikumam $\{S = k\}$ ir augstāka varbūtība par gadījumu, ja izlases apjoms $n = 50$, tādā veidā redzama konverģence. Izņēmums ir gadījums, kad $k = 4$, $\theta^{(4)} = (0, -0.5, 0, -0.2)$, pat pie $n = 100$, kad dominē pie otrās pakāpes polinoma.

1.1. tabula Novērtētais $P(S2 = s)$, $d(50) = 10$, $d(100) = 10$ alternatīva g_4

		s									
Parametri	n	1	2	3	4	5	6	7	8	9	10
$k = 1$, $\theta^{(1)} = 0.4$	50	94	4	1	1						
	100	96	2	1	1						
$k = 2$, $\theta^{(2)} = (0, -0.5)$	50	2	94	3	1						
	100	6	91	2	1						
$k = 2$, $\theta^{(2)} = (0.25, -0.35)$	50	46	50	3	1						
	100	21	75	4							
$k = 3$, $\theta^{(3)} = (0, 0, 0.6)$	50	32	1	53	7	6			1		
	100	4		68	11	14	1	1	1		
$k = 4$, $\theta^{(4)} = (0, -0.5, 0, -0.2)$	50	25	72	2	1						
	100	5	93	1	1						
$k = 4$, $\theta^{(4)} = (0.1, 0.2, -0.3, -0.4)$	50	15		6	74	4		1			
	100	18	1	3	74	2					
$k = 5$, $\theta^{(5)} = (0, 0, 0, 0, 0.5)$	50	49	2			40	5	2	1		
	100	16				77	4	1	1		1
$k = 8$, $\theta^{(8)} = (0, 0, 0, 0, 0, 0, 0, -0.7)$	50	61	4	1					31	2	1
	100	16							81	3	

1.2. tabula satur testa N_{S2} simulētās kritiskās vērtības dažādiem izlases apjomiem pie būtiskuma līmeņa $\alpha = 0.05$. Pēc simulāciju rezultātiem redzams, ka pie nulles hipotēzes, dominējošā ir zemākā dimensija, un palielinoties selekcijas likuma dimensijai, kritiskā vērtība kļūst lielāka. Kallenberga un Ledvina [7] secināja, ka maziem vai vidējiem izlašu apjomiem neskatoties uz to, ka gadījumi, kad $S2 > 1$ ir relatīvi reti, tomēr tiem ir nozīmīga ietekme, jo selekcijas kritērijs un testa statistiska ir savstarpēji saistīti. Kā rezultātā simulētās kritiskās vērtības ir būtiski lielākas nekā asimptotiski bāzētās uz χ_1^2 attiecīgo kvantīli, tas ir 3.84.

1.4.4. Simulētās jaudas

Lai redzētu N_{S2} uzvedību pie mazām izlasēm, simulējām jaudas pie dažām iepriekš definētajām alternatīvām pie izlases apjoma $n = 20$ un $1 < K < 6$ un pie $n = 50$ un $1 < K < 10$. Iegūtās jaudas attēlotas 1.3. tabulā. Abos izlases apjoma izvēles gadījumos N_{S2} jauda izturas pieņemami. Jauda ir maza, kamēr simulētajiem datiem eksponenciālās

1.2. tabula Testa N_{S2} kritiskās vērtības pie $\alpha = 0.05$, $d(50) = 10$

	$d(50)$									
n	1	2	3	4	5	6	7	8	9	10
20	3.762	5.464	6.635	7.245	7.526	7.631				
30	3.824	5.444	6.226	6.673	7.046	7.086	7.146			
50	3.890	5.315	5.929	6.047	6.120	6.121	6.161	6.182	6.194	6.203
80	3.799	5.461	5.498	5.515	5.526	5.621	5.662	5.692	5.742	5.838
100	3.782	5.085	5.176	5.285	5.364	5.410	5.418	5.538	5.612	5.659
120	3.818	5.122	5.223	5.275	5.436	5.445	5.484	5.498	5.553	5.599

saimes dimensija nav pareizi izvēlēta. Kad K sakrīt ar attiecīgo dimensiju, jauda būtiski palielinās un lieliem K ir gandrīz nepieciešamajā līmenī. Izņēmumi ir alternatīvas, kurām $\exp\{\theta \circ \phi(x) - \psi_k(\theta)\}$ datus apraksta labi, piemēram, alternatīvas g_3 ar $f = 1$ vai g_4 ar $s = 1$. Pēc simulācijām redzams, ka maksimālā jauda ir sasniegta, kad $K = 1$.

1.3. tabula Testa N_{S2} jaudas pie $\alpha = 0.05$, $d(50) = 10$

		Jaudas (%)										
	Alternatīva un parametri	n	1	2	3	4	5	6	7	8	9	10
g_2		20	10	58	63	62	64		-	-	-	-
		50	9.4	85	86	86	86	86	88	88	88	88
g_3	$d = 0.5$, $f = 1$	20	35	40	41	42	40	42	-	-	-	-
		50	73	73	74	76	74	73	73	71	73	72
g_4	$s = 5$, $\theta^{(5)} = (0, 0, 0, 0, 0.5)$	20	6	15	16	17	34	33	-	-	-	-
		50	6	15	17	55	50	55	54	51	53	54

1.4.tabulā salīdzināts Neimaņa tests ar pārējiem četriem testiem konkrētā gadījumā, kad izlases apjoms $n = 50$ un $K = 10$. Vairākos gadījumu Neimaņa tests dominē pār salīdzināmajiem testiem. N un W testiem ir diezgan līdzīgas kvalitātes. AD tests pie dažām alternatīvām ir jaudīgāks, bet pie citām jauda ir mazāka nekā pārējiem testiem. Kā arī ir redzams, ka salīdzināmo testu jaudas ir zemākas pie lielāka biežuma alternatīvām, piemēram, $g_3(x)$ ar $f = 5$ un $g_4(x)$ ar $s \geq 5$ nekā Neimaņa testam.

1.5.tabulā salīdzināts Neimaņa tests ar χ^2 testa jaudām, kuram apskatīti dažādi intervālu skaiti. Neimaņa tests pārsvarā dominē pār χ^2 testu neatkarīgi no intervālu skaita.

1.4. tabula Testu N_{S2}, N, W, AD un $K-S$ jaudas pie $\alpha = 0.05$, $d(50) = 10$

	Alternatīva un parametri	Jaudas (%)				
		N_{S2}	N	W	AD	$K-S$
g_1		98	86	87	93	54
g_2		88	62	62	64	73
g_3	$d = 0.5,\ f = 1$	72	62	49	70	35
	$d = 0.5,\ f = 2$	70	55	62	33	73
	$d = 0.7,\ f = 4$	56	51	28	17	86
	$d = 0.9,\ f = 5$	62	34	29	13	86
g_4	$k = 1,\ \theta^{(1)} = 0.4$	71	59	52	68	24
	$k = 2,\ \theta^{(2)} = (0, -0.5)$	98	89	92	56	89
	$k = 2,\ \theta^{(2)} = (0.25, -0.35)$	67	53	56	36	47
	$k = 3,\ \theta^{(3)} = (0, 0, 0.6)$	70	71	73	31	58
	$k = 4,\ \theta^{(4)} = (0, -0.5, 0, -0.2)$	76	46	56	16	90
	$k = 4,\ \theta^{(4)} = (0.1, 0.2, -0.3, -0.4)$	85	75	77	37	55
	$k = 5,\ \theta^{(5)} = (0, 0, 0, 0, 0.5)$	54	21	19	14	88
	$k = 8,\ \theta^{(8)} = (0, 0, 0, 0, 0, 0, 0, -0.7)$	42	12	15	11	89

1.5. tabula Testu N_{S2}, χ^2_5, χ^2_6, χ^2_7, χ^2_8, χ^2_9 un χ^2_{10} jaudas pie $\alpha = 0.05$, $d(50) = 10$

	Alternatīva un parametri	Jaudas (%)						
		N_{S2}	χ^2_5	χ^2_6	χ^2_7	χ^2_8	χ^2_9	χ^2_{10}
g_1		98	54	63	63	69	72	75
g_2		88	39	45	42	47	46	50
g_3	$d = 0.5,\ f = 1$	72	47	44	83	38	37	35
	$d = 0.5,\ f = 2$	70	43	41	43	36	33	34
	$d = 0.7,\ f = 4$	56	51	69	38	62	67	60
	$d = 0.9,\ f = 5$	62	54	75	57	86	85	86
g_4	$k = 1,\ \theta^{(1)} = 0.4$	71	46	42	42	34	36	34
	$k = 2,\ \theta^{(2)} = (0, -0.5)$	98	76	75	76	75	76	69
	$k = 2,\ \theta^{(2)} = (0.25, -0.35)$	67	41	39	39	34	36	30
	$k = 3,\ \theta^{(3)} = (0, 0, 0.6)$	70	61	61	65	60	64	62
	$k = 4,\ \theta^{(4)} = (0, -0.5, 0, -0.2)$	76	38	37	40	34	37	34
	$k = 4,\ \theta^{(4)} = (0.1, 0.2, -0.3, -0.4)$	85	51	59	73	77	74	78
	$k = 5,\ \theta^{(5)} = (0, 0, 0, 0, 0.5)$	54	21	23	23	25	33	35
	$k = 8,\ \theta^{(8)} = (0, 0, 0, 0, 0, 0, 0, -0.7)$	42	28	6	33	65	56	59

2. NEIMAŅA TESTS SALIKTĀM HIPOTĒZĒM

2.1. Neimaņa tests

Pieņemam, ka $X_1, X_2, \ldots, X_n$ ir neatkarīgi un vienādi sadalīti gadījuma lielumi ar blīvuma funkciju $f(x)$. Pārbaudīsim nulles hipotēzi

$$H_0 : f(x) \in \{f(x;\beta), \beta \in B\},$$

kur

$$\{f(x;\beta), \beta \in B\}$$

ir dota blīvuma funkciju kopa, piemēram, normālā vai eksponenciālā blīvuma funkciju kopa. Normālā blīvuma funkciju kopa atbilst

$$f(x;\beta) = (\sqrt{2\pi\sigma})^{-1} \exp\left\{-\frac{1}{2}(x-\mu)^2/\sigma^2\right\}, \tag{2.1.1}$$

kur $\beta = (\mu, \sigma)$ un $\hat{\beta} = \left(\bar{X}, \left\{n^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right\}^{\frac{1}{2}}\right)$ ar $\bar{X} = n^{-1}\sum_{i=1}^{n}$.
Eksponenciālā blīvuma funkciju kopa atbilst

$$f(x;\beta) = \begin{cases} 0, & x < 0 \\ \beta^{-1}\exp(-\beta^{-1}x), & x \geq 0 \end{cases} \tag{2.1.2}$$

un $\hat{\beta} = n^{-1}\sum_{i=1}^{n} X_i$.

Pārbaudīsim hipotēzi H_0 pret plašu alternatīvu skaitu. Pirmkārt, apskatīsim vienkāršāko situāciju, kad β ir zināms. Pirmais solis ir transformēt uz vienmērīgumu pie nulles hipotēzes, aizstājot X_i ar $F(X_i;\beta)$, kur $F(x;\beta)$ ir X_i sadalījuma funkcija. Otrajā solī, nogludināšanas testu definēšana, tiek ieviesta eksponenciālā saime

$$g_k(x;\theta;\beta) = \exp\left(\sum_{i=1}^{k}\theta_j\phi_j(F(x;\beta)) - \psi_k(\theta)\right) f(x;\beta), \tag{2.1.3}$$

kur $\phi_0, \phi_1, \dots, \phi_k$ ir ortonormāli Ležandra polinomi telpā $L_2([0,1])$ ar $\phi_0 \equiv 1$, $\theta = (\theta_1, \dots, \theta_k)$ un $\psi_k(\theta)$ ir normalizēta konstante, kas definēta ar

$$\psi_k(\theta) = \log \int_0^1 \exp\left(\sum_{j=1}^k \theta_j \phi_j(y)\right) dy.$$

Hipotēze H_0 saimē (2.1.3) reducējas uz hipotēzi $H : \theta = 0$. Kad β ir zināma, nogludināšanas testa statistikas tiek uzdotas ar

$$T_k(\beta) = \sum_{j=1}^k \left(n^{-1/2} \sum_{i=1}^n (\phi_j(F(X_i;\beta))\right)^2, k = 1, 2, ... \tag{2.1.4}$$

Uzdots, ka $\{f(x;\beta) : \beta \in B\}$ ir blīvumu kopa ar atbilstošām sadalījumu funkcijām $\{F(x;\beta) : \beta \in B\}$, kur $B \subset R^q$. Definētas eksponenciālās saimes $k = 1, 2,$ ar to blīvumiem

$$g_k(x;\theta;\beta) = exp\{\theta \circ \phi(F(x;\beta)) - \psi_k(\theta)\} f(x;\beta), \tag{2.1.5}$$

kur

$$\theta = (\theta_1, \dots, \theta_k)', \phi = (\phi_0, \dots, \phi_k)', \psi_k(\theta) = \log \int_0^1 exp\{\theta \circ \phi(y)\} dy,$$

un $\circ$ ir R^k iekšējais reizinājums. Ar " $'$ " apzīmē transponento matricu vai vektoru.

$$Y_n(\beta) = (\bar{\phi}_1(\beta), \dots, \bar{\phi}_j(\beta))' = n^{-1} \sum_{i=1}^n (\phi_1(F(X_i;\beta)), \dots, \phi_j(F(X_i;\beta)))'$$

ar atkarīgu j. Ticamības funkcija neatkarīgiem $X_1, \dots, X_n$ ar blīvuma funkcijām (2.1.5) ir uzdota ar

$$\exp\{n(\theta \circ Y_n(\beta) - \psi_k(\theta))\} \prod_{i=1}^n f(X_i;\beta).$$

Katrai $\beta \in B$ Švarca likums, lai izvēlētos atbilstošus apakšmodeļus sekojošām dimensijām, ir

$$S(\beta) = \min\{k; 1 \le k \le d(n), L_k(\beta) \ge L_j(\beta), j = 1, ..., d(n)\}, \tag{2.1.6}$$

kur

$$L_k(\beta) = n \sup_{\theta \in \Omega_s} \{\theta \circ Y_n(\beta) - \psi_s(\theta)\} - \frac{1}{2} k \log n.$$

Neskatoties uz to, ka nav pieminēts apzīmējumā, $S(\beta)$ ir atkarīgs no eksponenciālās saimes augšējās robežas $d(n)$. Ar $\hat{\beta}$ apzīmē β maksimālās ticamības funkcijas novērtējumu pie H_0. Definē

$$S = S(\hat{\beta}). \tag{2.1.7}$$

Kad β ir zināma un k tiek fiksēts, testa statistiska (2.1.4) ir modeļa (2.1.3) pamata statistika, skatīties Teorēmu 4.2.1 [3]. Tādēļ, kad β ir nezināma un k tiek fiksēts, tāpat izmantojam pamata statistiku, pārbaudot saliktu hipotēzi $H : \theta^{(k)} = 0$ pret $A : \theta^{(k)} \neq 0$ [3].

Apzīmēsim $k \times k$ vienības matricu ar I un matemātisko cerību ar $E_{(0,\beta)}$, kad X blīvuma funkcija ir $f(x;\beta)$. Tālāk definē

$$I_\beta = \left(-E_{(0,\beta)} \frac{\partial}{\partial \beta_t} \phi_j(F(X;\beta))\right)_{t=1,\ldots,q,\ j=1,\ldots,k},$$

$$I_{\beta\beta} = \left(-E_{(0,\beta)} \frac{\partial^2}{\partial \beta_t \partial \beta_u} \log f(X;\beta)\right)_{t=1,\ldots,q,\ u=1,\ldots,q}.$$

$$R(\beta) = I'_\beta (I_{\beta\beta} - I_\beta I'_\beta)^{-1} I_\beta.$$

$$W_k = n Y'_n(\hat{\beta}) \{I + R(\hat{\beta})\} Y_n(\hat{\beta}). \tag{2.1.8}$$

Tagad datu nosakošā testa statistika definēta sekojošā veidā

$$W_S = W_{S(\hat{\beta})} \tag{2.1.9}$$

ar W_k, kas uzdots vienādojumā (2.1.8), $S(\beta)$ uzdots (2.1.6) un $\hat{\beta}$ ir β maksimālās ticamības funkcijas novērtējums. Pie lielām W_S vērtībām nulles hipotēze tiek noraidīta.

Švarca likums $S(\beta)$, kas uzdots ar (2.1.6), salīdzina maksimizētās ticamības funkcijas vērtības. Pierādīts [8], ka maksimizētā ticamības funkcija (kura faktiski ir ticamības funkcijas attiecības statistika, lai pārbaudītu hipotēzi $H : \theta = 0$ pret $A : \theta \neq 0$, kad β ir zināma), ir ekvivalenta $\frac{1}{2} n \|Y_n(\beta)\|^2$, kur ar $\|\cdot\|$ apzīmē Eiklīda normu.

Ievietojot novērtējumu $\hat{\beta}$, rodas papildus $R(\beta)$ nosacījums - asimptotiska inversā kovariācijas matrica. Ņemot to vērā, iegūstam Švarca likuma modifikāciju

$$\begin{aligned} S1 = S1(\hat{\beta}) = \min\{k : 1 \leq k \leq d(n), \\ W_k - k \log n \geq W_j - j \log n, j = 1, ..., d(n)\}, \end{aligned} \tag{2.1.10}$$

kuru ir vieglāk aprēķināt. Atbilstošā testa statistiska ir

$$W_{S1} = W_{S1(\hat{\beta})}. \tag{2.1.11}$$

Bet vēl vienkāršāka modifikācija, kuru ir vieglāk aprēķināt, ir

$$S2 = S2(\hat{\beta}) = \min\{k : 1 \le k \le d(n), n\|Y_n\hat{\beta})\|^2_{(k)} - k\log n$$
$$\ge n\|Y_n\hat{\beta})\|^2_{(j)} - j\log n, \quad j = 1, ..., d(n)\}, \qquad (2.1.12)$$

kur normas indekss nozīmē dimensiju.

Atbilstošā testa statistika ir

$$W_{S2} = W_{S2(\hat{\beta})}. \qquad (2.1.13)$$

2.2. Selekcijas kritēriji

2.2.1. Selekcijas kritēriju definēšana

Ar P_β apzīmēsim, ka X_i ir blīvuma funkcija $f(x;\beta)$ un ar E_β un var_β attiecīgi atbilstošā matemātiskā cerība un dispersija. Saimei $\{f(x;\beta) : \beta \in B\}$ nepieciešami sekojoši nosacījumi.

(R1) Ja $t, u = 1, ..., q,\ (\partial/\partial\beta_t)f(x;\beta)$ un $(\partial^2/\partial\beta_t\partial\beta_u)f(x;\beta)$ eksistē gandrīz visur un ir tādi, ka katrai $\beta_0 \in B_0$ vienmērīgums β_0 apkārtnē

$$|(\partial/\partial\beta_t)f(x;\beta)| \le H_t(x)$$

un

$$\left|(\partial^2/\partial\beta_t\partial\beta_u)f(x;\beta)\right| \le G_{tu}(x),$$

kur $\int_R H_t(x)dx < \infty$ un $\int_R G_{tu}(x)\ dx. < \infty$.

(R2) Ja $t, u = 1, ..., q,\ (\partial/\partial\beta_t)\log f(x;\beta)$ un $(\partial^2/\partial\beta_t\partial\beta_u)\log f(x;\beta)$ eksistē gandrīz visur un ir tādi, ka Fišera informatīvā matrica

$$I_{\beta\beta} = E_\beta\left\{\left(\frac{\partial}{\partial\beta}\log f(X_1;\beta)\right)\left(\frac{\partial}{\partial\beta}\log f(X_1;\beta)\right)'\right\}$$

ir ierobežota, pozitīvi definēta un nepārtraukta, un kad $\delta \to 0$

$$E_\beta\left\{\sup_{h:\|h\|\le\delta}\left\|\frac{\partial^2}{\partial\beta\partial\beta'}\log f(X_1;\beta+h) - \frac{\partial^2}{\partial\beta\partial\beta'}\log f(X_1;\beta)\right\|\right\} \to 0.$$

(R3) Katrai $\beta_0 \in B_0$ eksistē $\eta = \eta(\beta_0) > 0$ ar

$$\sup_{\|\beta-\beta_0\|<\eta} \sup_{x\in R} \left| \frac{\partial^2}{\partial\beta\partial\beta_u} F(x;\beta) \right| < \infty, t, u = 1, ..., q.$$

un

$$\sup_{x\in R} \left| \frac{\partial}{\partial\beta_t} F(x;\beta) \right|_{\beta=\beta_0} < \infty, t = 1, ..., q.$$

(R4)

$$P_\beta(\sqrt{n}\|\hat{\beta} - \beta\| \geq r) \leq c_1 \exp(-c_2 r^2)$$

visiem $r = \sqrt{\log n}$ ar $0 < \rho \leq \rho_1$ un $n \geq n_1$.

Nākošie nosacījumi attiecās uz ortonormālu sistēmu $\{\phi_j\}_{j=0}^{\infty}$.

(S1)

$$\sup_{x\in[0,1]} \left|\phi_j'(x)\right| \leq c_3 j^{m_1}$$

katram $j = 1, 2, ..., d(n)$ un $c_3 > 0$, $m_1 > 0$.

(S2)

$$\sup_{x\in[0,1]} \left|\phi_j''(x)\right| \leq c_4 j^{m_2}$$

katram $j = 1, 2, ..., d(n)$ un $c_4 > 0$, $m_2 > 0$.

Šie nosacījumi attiecās uz eksponenciālās saimes dimensiju $d(n)$.

(D1) $\{d(n)V_{d(n)}\}^2 n^{-1} \log n \to 0$ kad $n \to \infty$, kur $V_k = \max_{1\leq j\leq k} \sup_{x\in[0,1]} \|\phi_j(x)\|$.

(D2) $d(n) = o(\{n/\log n\}^{(2m)^{-1}})$ kad $n \to \infty$, kur $\max(m_1, m_2)$.

(D3) $d(n) = o(n^c)$ kad $n \to \infty$

$$c < c_2 b^{-2},$$

ja $\rho_1 b \geq 1$, un ar $c = c_2\rho_1^2$, pretējā gadījumā, kur c_2 un ρ_1 ir uzdots ar (R4) un

$$b = \left(\sum_{t=1}^{q} var_\beta \frac{\partial}{\partial\beta_t} \log f(X:\beta) \right)^{1/2}.$$

2.2.2. Selekcijas kritēriju sākotnējais sadalījums

Sekojošās teorēmās parādīta selekcijas kritēriju sākotnējā sadalījuma asimptotiska. Varbūtība P_β nozīmē, ka X_i ir blīvuma funkcija $f(x;\beta)$. Pirmā teorēma parāda cik tuvu $Y_n(\hat{\beta})$ ir $Y_n(\beta)$.

Teorēma 10. *Pieņem (R1) - (R4), (S1), (S2), (D2), (D3). Tad eksistē $\varepsilon < 0$ tāds, ka*

$$\lim_{n\to\infty} \sum_{k=2}^{d(n)} P_\beta(\|Y_n(\hat{\beta} - Y_n(\beta)\|) \geq (1-\varepsilon)\{(k-1)n^{-1}logn\}^{1/2}) = 0. \tag{2.2.1}$$

Teorēmas pierādījums [9], 1240.lpp.

Nākošā teorēma parāda, ka pie H_0 Švarca atlases kritērijs un tā modifikācija asiptotiski koncentrēta uz dimensiju 1.

Teorēma 11. *Pieņem (R1) - (R4), (S1), (S2), (D1), (D2), (D3), tad*

$$\lim_{n\to\infty} P_\beta(S1(\hat{\beta} \geq 2) = 0. \tag{2.2.2}$$

Teorēmas pierādījums [9], 1232.lpp.

2.2.3. Selekcijas kritēriji pie alternatīvām

Šajā apakšnodaļā aplūkosim selekcijas kritēriju uzvedību pie alternatīvām. Pieņemsim, ka $X_1, X_2, \ldots$ - neatkarīgi un vienādi sadalīti gadījuma lielumi, katrs ar sadalījumu P. Pie alternatīvām ir tikai β novērtējums $\hat{\beta}$. Novērtējums $\hat{\beta}$ pie atlternatīvā sadalījuma parasti konverģēs uz kādu B_0 elementu, šis elements arī būs β.

Aplūkosim P kā saimes $\{f(x;\beta), \beta \in B\}$ alternatīvi, ja eksistē $K(\beta)$ tāds, ka

$$E_P\phi_1(F(X;\beta)) = ... = E_P\phi_{K(\beta)-1}(F(X;\beta)) = 0, \quad E_P\phi_{K(\beta)}(F(X;\beta)) \neq 0. \tag{2.2.3}$$

Pieņem, ka

$$\lim_{n\to\infty} \inf d(n) \geq K. \tag{2.2.4}$$

Ja (2.2.3) nav spēkā, tad $E_P\phi_{K(\beta)}(F(X;\beta)) = 0$ visiem j un jebkurai alternatīvai.

Pie nulles hipotēzes selekcijas kritēriji koncentrējas uz dimensiju 1, bet pie alternatīvām nozīme ir arī augstākām dimensijām.

Teorēma 12. *Pieņem, ka (2.9) ir spēkā un β ir sasaistīta ar $\hat{\beta}$ un P ir tāds, ka*

$$\|\hat{\beta} - \beta\| \xrightarrow{P} 0. \tag{2.2.5}$$

Teorēmas pierādījums [9], 1234.lpp.

2.3. Datus nosakošie nogludināšanas testi saliktām hipotēzēm

Pārbaudot saliktu hipotēzi H_0 datus nosakošo nogludināšanas testu statistika ir uzdotas ar

$$W_{S1(\hat{\beta})}(\hat{\beta}) \tag{2.3.1}$$

ar $W_k(\hat{\beta})$ uzdotu ar (2.1.8), $S_1(\hat{\beta})$ ar (2.1.11) un (2.1.10). Nulles hipotēzi noraida pie lielām testa statistikas vētībām. Sekojošā teorēmā parādīta testu statistiku sākotnējā sadalījuma asimptotika.

Teorēma 13. *Pie Teorēmas 11. nosacījumiem*

$$W_{S1(\hat{\beta})}(\hat{\beta}) \xrightarrow{P_\beta} \chi_1^2. \tag{2.3.2}$$

Teorēmas pierādījums [9], 1237.lpp.

Teorēma 14. *Pie Teorēmas 12 nosacījumiem*

$$W_{S1(\hat{\beta})} \xrightarrow{P} \infty. \tag{2.3.3}$$

Teorēmas pierādījums [9], 1238.lpp.

2.4. Simulācijas

Simulācijas tiek veiktas līdzīgā ceļā kā aprakstīts literatūrā [9]. Saliktu hipotēžu gadījumā tiek pārbaudīta eksponencialitāte un normalitāte. Lai iegūtu nepieciešamos testu statistikas rezultātus, jāveic virkne aprēķinu. Pirmajā solī jāaprēķina maksimālās ticamības funkcijas novērtējums pie nulles hipotēzes, tad jāizvēlas dimensija. Dimensijām $j = 1, ..., d(n)$ jāaprēķina $Y_n(\hat{\beta})$, un tad aprēķina statistiku pie katras dimensijas. Visbeidzot aprēķina selekcijas kritēriju pie maksimālās ticamības funkcijas novērtējuma un statistiku pie šī selekcijas kritērija. Visas simulācijas tika veiktas 10 000 reižu.

2.4.1. Alternatīvie sadalījumi

Saliktu hipotēžu gadījumā pārbaudot eksponencialitāti un normalitāti simulācijās izmantotas šādas daudzveidīgas alternatīvas ar dažāda veida blīvuma funkcijām.

2.1. tabula Alternatīvas

Alternatīva	Blīvuma funkcija
Weibull $(b;k)$	$bk(bx)^{k-1}\exp\{-(bx)^k\},\ x>0$
Gamma $(p;q)$	$q^{-p}\{\Gamma(p)\}^{-1}x^{p-1}\exp(-x/q),\ x>0$
χ^2_k=Gamma $(\frac{1}{2}k;2)$	$\{2^{\frac{1}{2}k}\Gamma(k/2)\}^{-1}x^{\frac{1}{2}k-1}\exp(-\frac{1}{2}x),\ x>0$
$LN(g;d)$	$d(x\sqrt{2\pi})^{-1}\exp\{-\frac{1}{2}(dlogx+g)^2\},\ x>0$
Beta $(p;q)$	$x^{p-1}(1-x)^{q-1}\{B(p,q)\}^{-1},\ 0\le x\le 1$
Vienmērīgais $(p;q)$	$(b-a)^{-1},\ a\le x\le b$

2.4.2. Kritiskās vērtības

2.2. tabulā parādītas testa W_{S1} simulētās kritiskās vērtības pārbaudot eksponencialitāti dažādiem izlases apjomiem pie būtiskuma līmeņa $\alpha = 0.05$. Ir redzams, ka kritiskās vērtības daudz neatšķiras starp $d(n)$ apgabalā no 2 līdz 10, bet starp pirmo un otro ir liels lēciens.

2.2. tabula Testa W_{S1} 5% kritiskās vērtības pārbaudot eksponencialitāti

	$d(n)$									
n	1	2	3	4	5	6	7	8	9	10
20	3.556	5.236	5.689	5.905	6.437	6.816	6.909	7.324	7.324	7.692
30	3.558	4.347	4.957	6.686	6.851	6.992	7.207	7.485	7.637	7.810
50	4.002	4.872	5.032	5.369	5.599	5.674	5.811	5.996	6.574	6.595

2.3. tabulā parādītas testa W_{S1} simulētās kritiskās vērtības pārbaudot normalitāti dažādiem izlases apjomiem pie būtiskuma līmeņa $\alpha = 0.05$. Kritiskās vērtības, daudz nemainās pie $d(n) = 3$ līdz 10.

2.3. tabula Testa W_{S1} 5% kritiskās vērtības pārbaudot normalitāti

	$d(n)$									
n	1	2	3	4	5	6	7	8	9	10
20	3.211	4.533	4.606	4.833	4.938	5.038	5.095	5.302	5.431	5.488
50	3.616	4.900	5.005	5.136	5.254	5.299	5.308	5.369	5.404	5.483

2.4.3. Simulētās jaudas

Pārbaudot eksponencialitāti un normalitāti, jaudas salīdzināšanai aplūkojam plaši izmantotos Pīrsona χ^2 testu un Kolmogorova - Smirnova testu. Literatūrā Neimaņa testi nav salīdzināti ar šiem testiem.

2.4. un 2.5. tabulās parādītas jaudas pārbaudot normalitāti. Lai redzētu izmaiņas, pieaugot izlases apjomam, apskatīti gadījumi, kad $n = 20$ un $n = 50$. Testi salīdzināti pie plaša apjoma alternatīvām. Lai varētu izdarīt secinājumus par χ^2 testa jaudu, apskatīti dažādi intervālu skaiti. No iegūtajiem rezultātiem redzams, ka Neimaņa tests dominē pār salīdzināmajiem testiem.

2.4. tabula: Testu W_{S1}, $K-S$, χ^2_5, χ^2_6, χ^2_7, χ^2_8, χ^2_9 un χ^2_{10} novērtētās jaudas pārbaudot normalitāti pie $\alpha = 0.05$, $d(20) = 6$

	Jaudas (%)							
Alternatīva	W_{S1}	$K-S$	χ^2_5	χ^2_6	χ^2_7	χ^2_8	χ^2_9	χ^2_{10}
χ^2_3	62	39	33	38	43	44	32	27
χ^2_4	54	34	23	25	27	29	25	19
Beta(0.5,0.5)	66	31	29	23	27	41	49	53
Beta(0.5,1)	63	40	29	36	48	63	61	50
Beta(2,1)	30	19	12	13	17	22	21	22
Beta(2,2)	9	5	9	6	5	7	5	5
Beta(3,2)	13	7	8	7	6	6	7	6
Gamma(0.4,1)	98	93	89	97	99	97	92	92
Gamma(0.6,1)	93	81	71	86	92	87	73	73
Gamma(0.7,1)	91	74	64	77	86	82	67	66
Gamma(1.5,1)	71	41	31	36	40	43	31	28
Gamma(2.4,1)	50	29	22	22	22	26	20	17
Gamma(3,0.333)	41	20	19	19	16	21	14	13
Gamma(4,1)	32	18	15	17	15	16	12	10
LN(0,1)	91	79	73	80	83	80	70	68
LN(0,0.8)	83	66	58	65	67	61	49	48
LN(0,1.2)	95	89	85	89	92	88	81	83
LN(-0.2,0.633)	72	48	41	40	41	46	34	34
LN(-0.3,0.775)	79	62	52	57	63	61	50	48
Logistic(1)	17	9	12	8	7	8	5	5
Vienmērīgais(0,2)	30	10	14	8	7	12	13	15
Weibull(0.5,1)	100	99	95	99	100	98	96	97
Weibull(0.6,1)	99	95	90	95	99	97	90	92
Weibull(0.8,1)	92	80	69	81	86	85	72	71
Weibull(1.4,1)	50	26	18	22	24	31	20	18
Weibull(1.6,1)	36	20	14	15	15	20	15	18
Weibull(2)	16	11	11	10	9	10	8	8
Weibull(2,1)	26	10	10	7	7	10	8	6
Weibull(2.2,1)	12	7	10	7	8	7	8	5
Weibull(3.6,1)	16	4	8	5	6	6	4	4
Weibull(4,1)	73	4	8	6	5	8	4	6

2.6. un 2.7. tabulās parādītas jaudas pārbaudot eksponencialitāti. Līdzīgi kā pār-

2.5. tabula: Testu W_{S1}, $K-S$, χ^2_5, χ^2_6, χ^2_7, χ^2_8, χ^2_9 un χ^2_{10} novērtētās jaudas pārbaudot normalitāti pie $\alpha = 0.05$, $d(50) = 10$

Alternatīva	Jaudas (%)							
	W_{S1}	$K-S$	χ^2_5	χ^2_6	χ^2_7	χ^2_8	χ^2_9	χ^2_{10}
χ^2_1	100	88	97	100	100	100	100	100
χ^2_3	97	39	53	65	72	79	82	86
χ^2_4	91	34	43	47	55	61	63	63
Beta(0.5,0.5)	99	31	81	75	64	62	64	72
Beta(0.5,1)	100	40	57	59	70	84	92	94
Beta(2,1)	81	19	21	22	27	33	40	49
Beta(2,2)	28	5	12	10	5	8	8	8
Beta(3,2)	25	7	10	10	8	10	9	9
Gamma(0.4,1)	100	93	99	100	100	100	100	100
Gamma(0.6,1)	100	81	95	98	100	100	100	100
Gamma(0.7,1)	100	74	89	97	100	100	100	100
Gamma(1.5,1)	97	41	55	66	72	78	84	85
Gamma(2.4,1)	53	29	36	39	45	49	51	53
Gamma(3,0.333)	83	20	28	31	35	38	39	40
Gamma(4,1)	68	18	24	23	26	26	30	20
LN(0,1)	100	79	96	98	99	99	100	100
LN(0,0.8)	93	66	86	91	95	96	95	97
LN(0,1.2)	99	89	99	100	100	100	100	100
LN(-0.2,0.633)	99	48	68	75	80	83	84	83
LN(-0.3,0.775)	99	62	81	91	93	95	95	97
Logistic(1)	26	9	14	12	9	9	8	9
Vienmērīgais(0,2)	80	10	28	22	17	17	18	19
Weibull(0.5,1)	100	99	100	100	100	100	100	100
Weibull(0.6,1)	100	95	100	100	100	100	100	100
Weibull(0.8,1)	100	80	93	98	100	100	100	100
Weibull(1.4,1)	88	26	33	39	44	53	60	63
Weibull(1.6,1)	83	20	21	25	29	30	34	36
Weibull(2)	47	11	13	13	13	14	15	14
Weibull(2,1)	40	10	13	10	13	13	13	13
Weibull(2.2,1)	35	7	10	10	9	11	10	10
Weibull(3.6,1)	7	4	7	7	6	6	7	6
Weibull(4,1)	9	4	7	6	6	6	6	5

baudot normalitāti apskatīti gadījumi, kad $n = 20$ un $n = 50$. Testi salīdzināti pie tām plašām alternatīvām. Šajā gadījījumā Pīrsona χ^2 tests un Kolmogorova - Smirnova tests vairākums gadījumu dominē pār Neimaņa testu, lai gan ir alternatīvas, kur Nemaņa tests ir vienlīdzīgs.

2.6. tabula: Testu W_{S1}, $K-S$, χ^2_5, χ^2_6, χ^2_7, χ^2_8, χ^2_9 un χ^2_{10} novērtētās jaudas pārbaudot eksponencialitāti pie $\alpha = 0.05$, $d(20) = 6$

Alternatīva	Jaudas (%)							
	W_{S1}	$K-S$	χ^2_5	χ^2_6	χ^2_7	χ^2_8	χ^2_9	χ^2_{10}
χ^2_1	95	88	81	92	94	92	82	81
χ^2_1	75	17	54	52	55	52	56	54
χ^2_3	21	100	100	100	100	100	100	100
χ^2_4	49	100	100	100	100	100	100	100
Beta(0.5,0.5)	56	100	100	100	100	100	100	100
Beta(0.5,1)	44	100	100	100	100	100	100	100
Beta(2,1)	97	100	100	100	100	100	100	100
Beta(2,2)	82	100	100	100	100	100	100	100
Beta(3,2)	100	100	100	100	100	100	100	100
Gamma(0.4,1)	87	96	99	67	99	99	100	100
Gamma(0.6,1)	39	59	87	83	85	83	84	79
Gamma(0.7,1)	26	37	68	65	60	62	62	61
Gamma(1.5,1)	6	55	84	82	82	80	78	75
Gamma(2.4,1)	25	100	100	100	100	100	100	100
Gamma(3,0.333)	47	100	100	100	100	100	100	100
Gamma(4,1)	80	100	100	100	100	100	100	100
LN(0,1)	22	45	75	74	74	71	71	71
LN(0,0.8)	40	43	77	75	70	70	68	65
LN(0,1.2)	47	53	81	81	79	82	80	79
LN(-0.2,0.633)	66	8	72	66	63	59	59	54
LN(-0.3,0.775)	33	5	53	43	41	43	39	35
Vienmērīgais(0,2)	32	9	48	48	55	62	58	51
Weibull(0.5,1)	94	24	73	75	78	78	79	80
Weibull(0.6,1)	69	16	60	56	58	60	64	63
Weibull(0.8,1)	36	7	36	33	35	35	33	30
Weibull(1.4,1)	10	4	40	32	29	28	29	27
Weibull(1.6,1)	22	3	53	44	45	41	39	38
Weibull(2)	95	4	85	79	74	74	73	68
Weibull(2,1)	63	3	83	79	76	75	73	67
Weibull(2.2,1)	76	3	92	88	83	85	82	78
Weibull(3.6,1)	100	14	100	100	100	100	100	100
Weibull(4,1)	100	19	100	100	100	100	100	100

2.7. tabula: Testu W_{S1}, $K-S$, χ^2_5, χ^2_6, χ^2_7, χ^2_8, χ^2_9 un χ^2_{10} novērtētās jaudas pārbaudot eksponencialitāti pie $\alpha = 0.05$, $d(50) = 10$

Alternatīva	Jaudas (%)							
	W_{S1}	$K-S$	χ^2_5	χ^2_6	χ^2_7	χ^2_8	χ^2_9	χ^2_{10}
χ^2_1	97	30	79	82	81	85	82	85
χ^2_3	46	100	90	100	100	100	100	100
χ^2_4	58	100	100	100	100	100	100	100
Beta(0.5,0.5)	100	100	100	100	100	100	100	100
Beta(0.5,1)	95	100	100	100	100	100	100	100
Beta(2,1)	41	100	100	100	100	100	100	100
Beta(2,2)	100	100	100	100	100	100	100	100
Beta(3,2)	100	100	100	100	100	100	100	100
Gamma(0.4,1)	70	100	100	100	100	100	100	100
Gamma(0.6,1)	36	94	99	99	99	99	100	98
Gamma(0.7,1)	87	71	92	91	92	91	88	89
Gamma(1.5,1)	99	92	99	99	99	98	100	98
Gamma(2.4,1)	94	100	100	100	100	100	100	100
Gamma(3,0.333)	71	100	100	100	100	100	100	100
Gamma(4,1)	100	100	100	100	100	100	100	100
LN(0,1)	100	84	99	98	97	97	97	96
LN(0,0.8)	88	85	99	99	99	98	98	96
LN(0,1.2)	61	91	98	99	98	98	98	98
LN(-0.2,0.633)	100	31	99	100	99	98	98	97
LN(-0.3,0.775)	94	9	85	84	81	78	77	74
Logistic(1)	100	53	100	100	100	100	100	100
Vienmērīgais(0,2)	100	23	79	95	99	97	94	92
Weibull(0.5,1)	100	60	97	98	98	99	99	99
Weibull(0.6,1)	59	30	85	89	88	91	89	92
Weibull(0.8,1)	100	10	43	47	44	48	48	45
Weibull(1.4,1)	98	5	63	64	59	56	56	55
Weibull(1.6,1)	98	8	90	90	89	87	85	83
Weibull(2)	93	18	100	100	100	100	100	100
Weibull(2,1)	100	16	100	100	100	100	100	100
Weibull(2.2,1)	100	27	100	100	100	100	100	100
Weibull(3.6,1)	100	98	100	100	100	100	100	100
Weibull(4,1)	93	100	100	100	100	100	100	100

Secinājumi

Darbā vienkāršas un saliktas hipotēzes gadījumā tika veikta Neimaņa testa teorētiskā analīze un empīriskās jaudas simulācijas, salīdzinot to ar citiem literatūrā labi pazīstamajiem testiem.

Šajā darbā tika veiktas Monte Karlo simulācijas maza un vidēja apjoma izlasēm. Izmantojot Neimaņa testus svarīgi izvēlēties pareizo komponentes dimensiju, pretējā gadījumā notiek ievērojams jaudas zudums.

Pēc simulēšanas rezultātiem redzams, ka Pīrsona χ^2 testa jaudas ir atkarīgas no izvēlēto intervālu skaita un tās nav īpaši lielas, tests ir piemērotāks vairāk diskrētiem gadījumiem.

Darbā plaši veiktie simulēšanas rezultāti un teorētiskie secinājumi rāda, ka Neimaņa testi labi piemēroti plaša loka alternatīvām. Vairākumā aplūkotajos gadījumos Neimaņa testi ir konkurētspējīgi ar matemātiskajā statistiskā plaši izmantoto Pīrsona χ^2 testu, Kolmogorova Smirnova, Andersona - Darlinga, kā arī ar Watsona un Neuhausa testu.

Simulāciju analīze tika veikta ar statistikas programmu R. Darbs tika rakstīts ar programmu Latex, kura ir daudz piemērotāka matemātiskajai un statistiskajai terminoloģijai.

Izmantotā literatūra un avoti

[1] J. Neyman. Smooth test for goodness of fit. *Skandinavske Aktuarietiddkrift*, 20:150 – 199, 1937.

[2] W. C. M Kallenberg T. Inglot and T.Ledwina. Power approximations to and power comparison of certain goodness-of-fit tests. *scand. J. Statist.*, 21:131–145, 1994a.

[3] J. C. W. Rayner and D. J. Best. *Smooth Tests of Goodness of fit.* Oxford University Press, New York, 1989.

[4] T. Ledwina. Data driven version of neyman's smooth testof fit. *J.Amer.Statist.Association*, 89:1000 – 1005, 1994.

[5] T. Inglot and T. Ledwina. Asymtotic optimality of data driven neyman's tests for uniformity. *Ann. Statist.*, 24(5):1982 – 2019, 1996.

[6] D. M. A. Haughton. On the choice of a model to fit data from an exponentialfamily. *Ann. Statist.*, 16(1):342–355, 1988.

[7] W. C. M and T. Ledwina. Consystenc'y and monte carlo simulation of a data driven version of smooth goodness-of-fit tests. *Ann. Statist.*, 23(5):1594 – 1608, 1995.

[8] D. R. Cox and D. V. Hinkley. *Theoretical Statistics.* Chapman and Hall, London, 1974.

[9] W. C. M Kallenberg T. Inglot and T. Ledwina. Data driven smoooth tests for composite hypotheses. *Ann.Statist.*, 25:1222–1250, 1997.

Pielikums. Izveidotās programmas

Neimaņa testa kritiskās vērtības

```
n=50
T<-numeric(10)
TT<-numeric()
for (dn in 1:10){
##print(dn)
SK=10000
while (SK>0) {
P<-runif(n)
## Defineejam Polinomu
P1<-function(x) (x-0.5)*sqrt(12);
P2<-function(x) sqrt(5)*(6*(x-0.5)^2-0.5);
P3<-function(x) sqrt(7)*(20*(x-0.5)^3-3*(x-0.5));
P4<-function(x) 210*(x-0.5)^4-45*(x-0.5)^2+9/8;
P5<-function(x) sqrt(11)*(63*(2*x-1)^5-70*(2*x-1)^3+15*(2*x-1))/8;
P6<-function(x) sqrt(13)*(231*(2*x-1)^6-315*(2*x-1)^4+105*(2*x-1)^2-5)/16;
P7<-function(x) sqrt(15)*(429*(2*x-1)^7-693*(2*x-1)^5+315*(2*x-1)^3-35*(2*x-1))/16;
P8<-function(x) sqrt(17)*(6435*(2*x-1)^8-12012*(2*x-1)^6+6930*(2*x-1)^4-1260*(2*x-1)^2+35)/128;
P9<-function(x) sqrt(19)*(12155*(2*x-1)^9-25740*(2*x-1)^7+18018*(2*x-1)^5-4620*(2*x-1)^3+315*(2*x-1))/128;
P10<-function(x) sqrt(21)*(46189*(2*x-1)^10-109395*(2*x-1)^8+90090*(2*x-1)^6-30030*(2*x-1)^4+3465*(2*x-1)^2-63)/256
S<-numeric(dn)
## Katram k=1,2,3,4,... aprekina Phi_, Yn un S2
for (k in 1:dn){
if (k==1)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2)
S[1]=(n*(Yn_2)^2-k*log(n))
}
if (k==2)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2 )
S[2]=(n*(Yn_2)^2-k*log(n))
}
if (k==3)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2)
S[3]=(n*(Yn_2)^2-k*log(n))
##print(S[3])
}
if (k==4)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2)
S[4]=(n*(Yn_2)^2-k*log(n))
##print(S[4])
}
if (k==5)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2)
S[5]=(n*(Yn_2)^2-k*log(n))
##print(S[5])
}
if (k==6)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2)
S[6]=(n*(Yn_2)^2-k*log(n))
##print(S[6])
}
if (k==7)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2)
S[7]=(n*(Yn_2)^2-k*log(n))
##print(S[7])
}
if (k==8)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2)
S[8]=(n*(Yn_2)^2-k*log(n))
##print(S[8])
}
if (k==9)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2+(n^(-1)*sum(P9(P)))^2)
S[9]=(n*(Yn_2)^2-k*log(n))
##print(S[9])
}
if (k==10)
{Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2+(n^(-1)*sum(P9(P)))^2+(n^(-1)*sum(P10(P)))^2)
S[10]=(n*(Yn_2)^2-k*log(n))
##print(S[10])
}}
K<-order(S)
d<-K[dn]
##print(d)
dati<-P
T[1]=(n^(-1/2)*sum(P1(dati)))^2
T[2]=T[1]+(n^(-1/2)*sum(P2(dati)))^2
T[3]=T[2]+(n^(-1/2)*sum(P3(dati)))^2
T[4]=T[3]+(n^(-1/2)*sum(P4(dati)))^2
T[5]=T[4]+(n^(-1/2)*sum(P5(dati)))^2
T[6]=T[5]+(n^(-1/2)*sum(P6(dati)))^2
T[7]=T[6]+(n^(-1/2)*sum(P7(dati)))^2
T[8]=T[7]+(n^(-1/2)*sum(P8(dati)))^2
T[9]=T[8]+(n^(-1/2)*sum(P9(dati)))^2
T[10]=T[9]+(n^(-1/2)*sum(P10(dati)))^2
TT[SK]<-T[d]
SK=SK-1
}}
Krit<-sort(TT)
Krit[9500]
```

Neimaņa testa jauda

```
n=50
T<-numeric(10)
s<-numeric(10)
for (i in 1:10){
s[i]<-0
}

for (dn in 1:10){
##print(dn)
SK=10000

while (SK>0) {

vert=runif(n)
g1 <- function (y) {(1/2)+(1/20)*((y^(-4/5))+((1-y)^(-4/5)))}
x<-seq(0.01,1,by=0.01)
X<-numeric(n)
f<-numeric(length(x))
for(i in 1:length(x)) {
f[i] <-integrate(g1,0,x[i])value}

##Matricas katra kolonna ir dati punkts
RES <- matrix (data=vert,nrow=length(x),ncol=n,byrow=TRUE)
RES <- f-RES
for (j in 1:n){
i<-1
if (RES[i,j]>0) {
##print(x[i])
X[j]<-x[i]
##print(X[j])
}
else{
if (RES[i,j]<0) {
while (RES[i,j]<0){
i<-i+1}
##print(x[i])
X[j]<-(x[i-1]+x[i])/2
##print(X[j])
}}}
P<-X
## Defineejam Polinomu

P1<-function(x) (x-0.5)*sqrt(12);
P2<-function(x) sqrt(5)*(6*(x-0.5)^2-0.5);
P3<-function(x) sqrt(7)*(20*(x-0.5)^3-3*(x-0.5));
P4<-function(x) 210*(x-0.5)^4 - 45*(x-0.5)^2 + 9/8;
P5<-function(x) sqrt(11)*(63*(2*x-1)^5-70*(2*x-1)^3+15*(2*x-1))/8;
P6<-function(x) sqrt(13)*(231*(2*x-1)^6-315*(2*x-1)^4+105*(2*x-1)^2-5)/16;
P7<-function(x) sqrt(15)*(429*(2*x-1)^7-693*(2*x-1)^5+315*(2*x-1)^3-35*(2*x-1))/16;
P8<-function(x) sqrt(17)*(6435*(2*x-1)^8-12012*(2*x-1)^6+6930*(2*x-1)^4-1260*(2*x-1)^2+35)/128;
P9<-function(x) sqrt(19)*(12155*(2*x-1)^9-25740*(2*x-1)^7+18018*(2*x-1)^5-4620*(2*x-1)^3+315*(2*x-1))/128;
P10<-function(x) sqrt(21)*(46189*(2*x-1)^{10}-109395*(2*x-1)^8+90090*(2*x-1)^6-30030*(2*x-1)^4+3465*(2*x-1)^2-63)/256


S<-numeric(dn)
## Katram k=1,2,3,4,... aprekina Phi_, Yn un S2
for (k in 1:dn){
##print(k)
if (k==1)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2)
S[1]=(n*(Yn_2)^2-k*log(n))
##print(S[1])
}

if (k==2)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2 )
S[2]=(n*(Yn_2)^2-k*log(n))
##print(S[2])
}

if (k==3)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2)
S[3]=(n*(Yn_2)^2-k*log(n))
##print(S[3])
}

if (k==4)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2)
S[4]=(n*(Yn_2)^2-k*log(n))
##print(S[4])
}

if (k==5)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2)
S[5]=(n*(Yn_2)^2-k*log(n))
##print(S[5])
}

if (k==6)
{
```

```
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2)
S[6]=(n*(Yn_2)^2-k*log(n))
##print(S[6])
}

if (k==7)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2)
S[7]=(n*(Yn_2)^2-k*log(n))
##print(S[7])
}

if (k==8)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2)
S[8]=(n*(Yn_2)^2-k*log(n))
##print(S[8])
}


if (k==9)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2+(n^(-1)*sum(P9(P)))^2)
S[9]=(n*(Yn_2)^2-k*log(n))
##print(S[9])
}

if (k==10)
{
Yn_2=sqrt((n^(-1)*sum(P1(P)))^2+(n^(-1)*sum(P2(P)))^2+(n^(-1)*sum(P3(P)))^2+(n^(-1)*sum(P4(P)))^2+(n^(-1)*sum(P5(P)))^2+(n^(-1)*sum(P6(P)))^2+
(n^(-1)*sum(P7(P)))^2+(n^(-1)*sum(P8(P)))^2+(n^(-1)*sum(P9(P)))^2+(n^(-1)*sum(P10(P)))^2)
S[10]=(n*(Yn_2)^2-k*log(n))
##print(S[10])
}
}
##print(Yn_2)

K<-order(S)
d<-K[dn]

dati<-X

T[1]=(n^(-1/2)*sum(P1(dati)))^2
T[2]=T[1]+(n^(-1/2)*sum(P2(dati)))^2
T[3]=T[2]+(n^(-1/2)*sum(P3(dati)))^2
T[4]=T[3]+(n^(-1/2)*sum(P4(dati)))^2
T[5]=T[4]+(n^(-1/2)*sum(P5(dati)))^2
T[6]=T[5]+(n^(-1/2)*sum(P6(dati)))^2
T[7]=T[6]+(n^(-1/2)*sum(P7(dati)))^2
T[8]=T[7]+(n^(-1/2)*sum(P8(dati)))^2
T[9]=T[8]+(n^(-1/2)*sum(P9(dati)))^2
T[10]=T[9]+(n^(-1/2)*sum(P10(dati)))^2

if (T[d]>qchisq(0.95,1)){s[dn]=s[dn]+1}

SK=SK-1
}}
Jauda=s
Jauda
```

Kolmogorova-Smirnova testa jauda

```
n<-50
SK<-10000
ss<-numeric()
ss<-0

while (SK>0) {

vert<-runif(50,0,1)
n<-length(vert)
g1 <- function (y) {(1/2)+(1/20)*((y^(-4/5))+((1-y)^(-4/5)))}
x<-seq(0.01,1,by=0.01)
X<-numeric(vert)
f<-numeric(length(x))
for(i in 1:length(x)) {
f[i] <-integrate(g1,0,x[i] )value}

##Matricas katra kolonna ir dati punkts
RES <- matrix (data=vert,nrow=length(x),ncol=n,byrow=TRUE)
RES <- f-RES

for (j in 1:length(vert)){
i<-1
if (RES[i,j]>0) {
##print(x[i])
X[j]<-x[i]
##print(X[j])
}

else{

if (RES[i,j]<0) {
while (RES[i,j]<0){
```

```
i<-i+1}
##print(x[i])
X[j]<-(x[i-1]+x[i])/2
##print(X[j])
}}}
P<-X
P
KS<-ks.test(P,"punif",0,1)statistic[[1]]

if (KS<0.05){ss<-ss+1}

SK=SK-1
}
Jauda=ss
Jauda
```

Maģistra darbs “Neimaņa testu analīze un pielietojumi” izstrādāts LU Fizikas un Matemātikas fakultātē.

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: Sintija Cīrule

____________________ ________________

(paraksts) (datums)

Rekomendēju darbu aizstāvēšanai.

Vadītājs: doc. Dr.math. Jānis Valeinis

____________________ ________________

(paraksts) (datums)

Recenzents: lektors Jānis Smotrovs

____________________ ________________

(paraksts) (datums)

Darbs iesniegts Matemātikas nodaļā ________________

(datums)

____

(darbu pieņēma)

Darbs aizstāvēts maģistra gala pārbaudījuma komisijas sēdē

________________ prot. Nr. ________, vērtējums________

(datums)

Komisijas sekretārs/-e: ________________ ________________

(Vārds, Uzvārds) (paraksts)