LATVIJAS UNIVERSITĀTE
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE
MATEMĀTIKAS NODAĻA

BĀRTLETA KOREKCIJA EMPĪRISKĀS TICAMĪBAS METODEI

MAĢISTRA DARBS

Autors: **Sandra Vucāne**

Stud. apl. sv08090

Darba vadītājs: doc. Dr.math. Jānis Valeinis

RĪGA 2010

Anotācija

Empīriskā ticamības metode ir vienīgā neparametriskā metode, kas pieļauj Bārtleta korekciju. Vienas izlases gadījumā Bārtleta korekcijas iespējamība empīriskās ticamības metodei ir parādīta gan vidējai vērtībai, gan gludām funkcijām no vidējās vērtības, gan kvantilēm, gan arī citām problēmām. To pašu gan pagaidām nevar teikt par divu izlašu gadījumu, jo tikai 2008.gadā tika parādīta Bārtleta korekcija divu izlašu vidējo vērtību starpībai. Izvirzītais darba mērķis bija iegūt Bārtleta korekciju vispārējām divu izlašu problēmām, balstoties uz nesen iegūto Bārtleta korekciju divu izlašu vidējo vērtību starpībai.

Atslēgas vārdi: Edgeworth izvirzījumi, empīriskā ticamības metode, Bārtleta korekcija

Abstract

The empirical likelihood method is the only nonparametric method that admits Bartlett adjustment. In one sample case Bartlett correction has been shown for mean, for smooth functions of mean, for quantiles and also for some other problems. But we can't say the same about two sample case, because only in 2008 Bartlett correction was shown for the two sample mean difference. The main goal of this thesis was to obtain Bartlett correction for empirical likelihood for general two sample problem.

Keywords: Edgeworth expansions, empirical likelihood, Bartlett correction

Saturs

Ievads

Empīriskās ticamības metode (turpmāk rakstīsim arī EL), kuru ieviesa Owen [1, 2, 3, 4] ir viena no svarīgākajām neparametriskajām metodēm.

Konstruējot ticamības intervālus, galvenā šīs metodes priekšrocība ir, ka EL ticamības intervālu forma nebalstās uz pieņēmumu par pētāmā sadalījuma simetriju, bet gan tikai uz pētāmo datu kopu, t.i., EL ticamības intervāli nav simetriski. Vēl viena no šīs metodes priekšrocībām ir tāda, ka, salīdzinot ar citām neparametriskajām metodēm, EL metode ir vienīgā neparametriskā metode, kas pieļauj Bārtleta korekciju.

Bārtleta korekcija būtībā izriet no Edgeworth izvirzījumiem, ar kuru palīdzību var iegūt statistikas sadalījuma aproksimāciju, izmantojot tā kumulantus. Bārtleta korekcijas būtība ir, koriģējot empīriskās logaritmiskās ticamības attiecību ar tās vidējo vērtību, kuras izvirzījums ir formā $1 - an^{-1}$, uzlabot tās χ^2 sadalījuma aproksimācijas kļūdu no $O(n^{-1})$ uz $O(n^{-2})$, kur n ir izlases apjoms. Bārtleta korekcijai ir pārsteidzošs efekts, jo ar šo vienkāršo skalāro transformāciju iespējams uzlabot ticamības intervālu pārklājuma kļūdu pie visiem ticamības līmeņiem.

Sākotnēji Bārtleta korekcija tika parādīta ticamības attiecības testiem parametriskajai ticamības funkcijai, t.i., maksimālās ticamības metodei. DiCiccio, Hall un Romano [5] parādīja, ka gadījumā, kad nezināmo parametru θ var izteikt kā gludu funkciju no vidējās vērtības, pastāv cieša saistība starp parametrisko un neparametrisko ticamības funkciju, kas kalpoja par pamatu tam, ka arī empīriskajai ticamības funkcijai var veikt Bārtleta korekciju.

Tas, ka empīriskās ticamības metodei ir iespējama Bārtleta korekcija, ir parādīts samērā plašam problēmu lokam. Vienas izlases gadījumā Hall un La Scala [6] parādīja Bārtleta korekciju vidējās vērtības gadījumā, DiCiccio, Hall un Romano [7] – gadījumā, kad nezināmo parametru var izteikt kā gludu funkciju no vidējās vērtības. Pāris gadus vēlāk Chen un Hall [8] parādīja, ka arī gadījumā, kad novērtējamo parametru nevar izteikt kā gludu funkciju no vidējās vērtības, proti, kvantiļu gadījumā, gludinātajai empīriskās ticamības metodei ir iespējama Bārtleta korekcija. Chen [9, 10] parādīja Bārtleta korekciju lineārās regresijas koeficientiem. Lazar un Mykland [11] parādīja, ka gadījumā ar traucējošiem parametriem, EL metodei varētu arī nebūt iespējama Bārtleta korekcija, taču vēlāk Chen un Cui [12] pierādīja, ka arī šajā gadījumā EL metodei ir iespējama Bārtleta korekcija.

Kopš tika parādīta Bārtleta korekcijas iespējamība EL metodei dažādām problēmām vienas izlases gadījumā, ir bijuši vairāki mēģinājumi to vispārināt divu izlašu gadījumā. Pirmais no mēģinājumiem bija apskatīt divu izlašu vidējo vērtību problemātiku, ko paveica Jing [13], kurš apskatīja divu izlašu vidējo vērtību starpību un parādīja, ka šajā gadījumā ir iespējams pielietot Bārtleta korekciju, tādējādi uzlabojot ticamības intervālu pārklājuma precizitāti no $O(n^{-1})$ uz $O(n^{-2})$. 2008.gadā Liu, Zou un Zhang [14] parādīja, ka Jing [13] piedāvātā Bārtleta korekcija divu izlašu vidējo vērtību starpībai ir kļūdaina un ieguva pareizo variantu, taču darba izstrādes gaitā tika konstatētas neprecizitātes arī šajā publikācijā. 2010.gadā Liu un Yu [15] piedāvāja Bārtleta korekciju koriģētajai EL divu izlašu vidējo vērtību starpībai, šajā publikācijā galvenokārt tiek izmantota Liu, Zou un Zhang [14] pieeja.

Qin un Zhao [16] aprakstīja empīriskās ticamības metodi divu izlašu vidējo vērtību un sadalījuma funkciju starpībām un Valeinis [17] parādīja, ka to var vispārināt un uzrakstīt formā, kas der ne tikai divu izlašu vidējo vērtību un sadalījuma funkciju starpībām, bet arī kvantiļu starpībām, P-P, Q-Q grafikiem, ROC līknēm un strukturālajiem attiecību modeļiem. Balstoties uz šo rezultātu, radās jautājums, vai arī Bārtleta korekciju nevar parādīt šai vispārinātajai problēmai divu izlašu gadījumā.

Un no tā arī izriet šī darba mērķis: iegūt Bārtleta korekciju vispārinātajai divu izlašu problēmai, balstoties uz nesen iegūto Bārtleta korekciju divu izlašu vidējo vērtību starpībai.

Lai sasniegtu izvirzīto mērķi tika uzstādīti vairāki darba uzdevumi:

- Veikt Bārtleta korekcijas padziļinātu teorētisko analīzi gan vienas, gan divu izlašu gadījumā. Teorētiski detalizēti izanalizēt pierādījumus vienas izlases gadījumā par vidējo vērtību un kvantilēm.

- Implementēt Bārtleta korekcijas programmā R, pārbaudot empīrisko pārklājuma precizitāti gan vienas, gan divu izlašu gadījumā.

- Pārstrādāt Liu, Zou un Zhang 2008.gada publikācijā veikto pierādījumu un pārbaudīt, vai pierādījuma tehniku var lietot arī citiem divu izlašu gadījumiem.

- Pierādīt Bārtleta korekciju P-P, Q-Q grafikiem, ROC līknēm, kvantiļu un sadalījuma funkciju starpībām.

Darbs sastāv no ievada, 4 nodaļām, secinājumiem un izmantotās literatūras un avotu saraksta. 1.nodaļā aprakstīti Edgeworth izvirzījumi, 2.nodaļā aprakstīta empīriskās ticamības metode vienai izlasei, uzmanību pievēršot vidējās vērtības un kvantiļu gadījumiem, kuriem tiek parādīta Bārtleta korekcija un arī simulāciju rezultāti šajos abos gadījumos, lai varētu praktiski novērtēt Bārtleta korekcijas sniegto uzlabojumu. Nākamajā nodaļā tiek iepazīstināts ar empīriskās ticamības metodi divām izlasēm, tiek apskatīta Bārtleta korekcija divu izlašu vidējo vērtību starpībai, kurā izlabotas [14] pieļautās neprecizitātes, ko izrādījās ne mazums, kā arī šajā nodaļā pieejami simulāciju rezultāti, lai praktiski novērtētu Bārtleta korekcijas efektivitāti šajā gadījumā. 4.nodaļā apskatīta gludinātā empīriskā ticamības metode divu izlašu gadījumā, kā arī mēģinājums iegūt Bārtleta korekciju vispārinātajai divu izlašu problēmai. Visas darbā atspoguļotās simulācijas tika veiktas, izmantojot programmu R.

1. Edgeworth izvirzījumi

Kā jau ievadā tika minēts, tad Bārtleta korekcijas iespējamība ir cieši saistīta ar Edgeworth izvirzījumiem. Ar Edgeworth izvirzījumu palīdzību var iegūt statistikas sadalījuma aproksimāciju caur tā kumulantiem. Bhattacharya un Gosh [18] pie zināmiem nosacījumiem pierādīja, ka parametram, ko var izteikt kā funkciju no vidējās vērtības ir spēkā Edgeworth izvirzījumi.

Edgeworth izvirzījumu aprakstīšanai, galvenokārt, tiks izmantota Hall grāmata [19], kur papildus iespējams apskatīt Edgeworth izvirzījumus kontekstā ar butstrapa metodi.

Lai labāk izprastu Edgeworth izvirzījumu būtību, vispirms tiks apskatīts Edgeworth izvirzījums statistikai $S_n = n^{1/2}(\hat{\theta} - \theta_0)/\sigma$, kur nezināmais parametrs ir vidējā vērtība un tā novērtējums ir izlases vidējā vērtība.

1.1. Edgeworth izvirzījums vidējās vērtības gadījumā

Pieņemsim, ka $X_1, \ldots, X_n$ ir neatkarīgi un vienādi sadalīti (turpmāk to rakstīsim, izmantojot arī anglisko saīsinājumu *i.i.d.*) gadījuma lielumi ar vidējo vērtību $\theta_0 = \mu = \mathbb{E}X$ un galīgu dispersiju $\sigma^2 = \mathbb{D}X$. Par parametra θ_0 novērtējumu izvēlamies izlases vidējo vērtību, t.i., $\hat{\theta} = \bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, kuras dispersija ir $n^{-1}\sigma^2$. No Centrālās robežteorēmas seko, ka statistika S_n asimptotiski ir normāli sadalīta ar vidējo vērtību 0 un dispersiju 1. Šo aproksimāciju varam izmantot, piemēram, ticamības intervālu konstruēšanai, proti, ja x_α ir vienādojuma

$$\mathbb{P}(|N(0,1)| \leq x_\alpha) = \alpha$$

atrisinājums, tad $I = (\hat{\theta} - n^{-1/2}\sigma x_\alpha, \hat{\theta} + n^{-1/2}\sigma x_\alpha)$ ir ticamības intervāls parametram θ_0 pie ticamības līmeņa α. Kas savukārt nozīmē, ka

$$\mathbb{P}(\theta_0 \in I) \to \alpha,$$

kad $n \to \infty$, taču ļoti liela nozīme ir tam, cik precīza ir šī normālā sadalījuma aproksimācija. To iespējams noteikt ar ticamības intervālu konverģences precizitātes palīdzību. Vieglākais veids, kā aprakstīt kļūdas aproksimācijā ar normālo sadalījumu, ir izmantot raksturīgās funkcijas.

Definīcija 1. [20] Par gadījuma lieluma X raksturīgo funkciju sauc kompleksu funkciju

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \int e^{itX} dF(x), \tag{1.1}$$

kur $e^{it} = \cos(t) + i\sin(t)$, $t \in \mathbb{R}$. Ja sadalījuma funkcijai $F(x)$ eksistē blīvuma funkcija $f(x)$, tad (1.1) var pārrakstīt

$$\varphi_X(t) = \int e^{itX} f(x)dx.$$

Tālāk apskatīsim dažas no raksturīgo funkciju īpašībām [20], kas būs nepieciešamas tālākajos aprēķinos:

1. $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$;

2. Ja $X_1, \ldots, X_n$ ir neatkarīgi gadījuma lielumi, tad

$$\varphi_{\sum_{i=1}^n X_i}(t) = \varphi_{X_1}(t) \cdot \ldots \cdot \varphi_{X_n}(t);$$

Plašāka informācija par raksturīgajām funkcijām, to īpašībām, īpašību pierādījumiem atrodama, piemēram, [20].

Definīcija 2. [20] Par gadījuma lieluma X kumulatīvo ģenerējošo funkciju sauc

$$K_X(t) = \ln \varphi_X(t) \tag{1.2}$$

un izvirzot (1.2) Teilora rindā punkta $t_0 = 0$ apkārtnē, var iegūt

$$K_X(t) = \ln \varphi_X(t) = \sum_{r=1}^{n} \kappa_r (it)^r / r! + o(t^n), \tag{1.3}$$

kur koeficientus $\kappa_1, \ldots, \kappa_n$ sauc par X kumulantiem.

Dažas kumulantu īpašības:

1. ja κ_r ir X $r-$tais kumulants, tad $a + bX$ $r-$tais kumulants ir $b^r\kappa_r$;

2. ja $X_1, \ldots, X_n$ ir neatkarīgi gadījuma lielumi, tad

$$K_{\sum_{i=1}^n X_i}(t) = \sum_{i=1}^{n} K_{X_i}(t) \tag{1.4}$$

   un šīs summas $r-$tais kumulants ir vienāds ar atbilstošo kumulantu summu;

3. ja $X_1, \ldots, X_n$ ir vienādi sadalīti gadījuma lielumi, tad (1.4) $r-$tais kumulants ir $n\kappa_r$;

4. ja $X_1, \ldots, X_n$ ir neatkarīgi un vienādi sadalīti, tad $n^{-1/2}(X_1 + \ldots + X_n)$ $r-$tais kumulants ir $n^{1-r/2}\kappa_r$.

Ar šo definīciju palīdzību esam iepazinušies ar dažiem no turpmāk nepieciešamajiem apzīmējumiem un varam atgriezties pie statistikas $S_n = n^{1/2}(\hat{\theta} - \theta_0)/\sigma$.

Tā kā S_n asimptotiski ir standartnormāli sadalīta, tad pie $n \to \infty$ statistikas S_n raksturīgā funkcija φ_{S_n} konverģē uz standartnormālā sadalījuma raksturīgo funkciju, t.i.,

$$\varphi_{S_n}(t) = \mathbb{E}\{\exp(itS_n)\} \to \mathbb{E}\{\exp(itN(0,1))\} = e^{-t^2/2}, \ -\infty < t < \infty.$$

Izmantojot raksturīgo funkciju 1. un 2.īpašību, varam iegūt, ka

$$\varphi_{S_n}(t) = \mathbb{E}e^{it\sqrt{n}\frac{(\bar{X}-\mu)}{\sigma}} = e^{-it\sqrt{n}\frac{\mu}{\sigma}}\mathbb{E}e^{\frac{it}{\sqrt{n}\sigma}\sum_{i=1}^{n} X_i} = e^{-it\sqrt{n}\frac{\mu}{\sigma}}\left[\varphi_X\left(t/(\sqrt{n}\sigma)\right)\right]^n$$

un ja definējam $Y = (X-\mu)/\sigma$, tad gadījuma lieluma Y raksturīgā funkcija ir

$$\varphi_Y(t) = \mathbb{E}e^{it(X-\mu)\sigma^{-1}} = e^{-it\mu\sigma^{-1}}\mathbb{E}e^{it\sigma^{-1}X}$$

un apskatot tās vērtību punktā $t/\sqrt{n}$,

$$\varphi_Y(t/\sqrt{n}) = e^{-itn^{-1/2}\mu\sigma^{-1}}\varphi_X\left(t/(\sqrt{n}\sigma)\right),$$

varam iegūt, ka

$$\varphi_{S_n}(t) = (\varphi_Y(t/\sqrt{n}))^n. \tag{1.5}$$

Izsakām Y raksturīgo funkciju ar kumulatīvās ģenerējošas funkcijas palīdzību,

$$\varphi_Y(t) = \exp\{\kappa_1 it + \frac{1}{2}\kappa_2(it)^2 + \ldots + \frac{1}{j!}\kappa_j(it)^j + \ldots\} \tag{1.6}$$

un, izvirzot Y raksturīgo funkciju $\varphi_Y(t)$ Teilora rindā, iegūstam

$$\varphi_Y(t) = 1 + \mathbb{E}(Y)it + \frac{1}{2}\mathbb{E}(Y^2)(it)^2 + \ldots + \frac{1}{j!}\mathbb{E}(Y^j)(it)^j + \ldots. \tag{1.7}$$

Apvienojot (1.6) un (1.7), iegūstam

$$\begin{aligned}\sum_{j\geq 1}\frac{1}{j!}\kappa_j(it)^j &= \ln\left[1 + \sum_{j\geq 1}\frac{1}{j!}\mathbb{E}(Y^j)(it)^j\right] \\ &= \sum_{k\geq 1}\left[(-1)^{k+1}\frac{1}{k}\left(\sum_{j\geq 1}\frac{1}{j!}\mathbb{E}(Y^j)(it)^j\right)^k\right].\end{aligned} \tag{1.8}$$

Pielīdzinot izteiksmē (1.8) koeficientus pie $(it)^j$, iegūstam

$$\begin{aligned}
\kappa_1 &= \mathbb{E}(Y),\\
\kappa_2 &= \mathbb{E}(Y^2) - (\mathbb{E}Y)^2 = \mathbb{D}(Y),\\
\kappa_3 &= \mathbb{E}(Y^3) - 3\mathbb{E}(Y^2)\mathbb{E}(Y) + 2(\mathbb{E}Y)^3 = \mathbb{E}(Y-\mathbb{E}Y)^3,\\
\kappa_4 &= \mathbb{E}(Y^4) - 4\mathbb{E}(Y^3)\mathbb{E}(Y) - 3(\mathbb{E}Y^2)^2 + 12\mathbb{E}(Y^2)(\mathbb{E}Y)^2 - 6(\mathbb{E}Y)^4\\
&= \mathbb{E}(Y-\mathbb{E}Y)^4 - 3(\mathbb{D}Y)^2,\\
&\dots . \qquad (1.9)
\end{aligned}$$

Izmantojot (1.6), (1.9) un kumulatīvo ģenerējošo funkciju īpašības un to, ka $\kappa_1 = \mathbb{E}(Y) = 0$ un $\kappa_2 = \mathbb{D}(Y) = 1$, vienādojumu (1.5) var pārrakstīt

$$\varphi_{S_n}(t) = \exp\{-\frac{1}{2}t^2 + n^{-1/2}\frac{1}{3!}\kappa_3(it)^3 + \dots + n^{-(j-2)/2}\frac{1}{j!}\kappa_j(it)^j + \dots\} \qquad (1.10)$$

$$= \exp\{-t^2/2\}\exp\{n^{-1/2}(3!)^{-1}\kappa_3(it)^3\}\exp\{n^{-1}(4!)^{-1}\kappa_4(it)^4\}\cdot\ldots$$
$$\cdot\ldots\cdot\exp\{n^{-(j-2)/2}(j!)^{-1}\kappa_j(it)^j\}\cdot\ldots \qquad (1.11)$$

Izmantojot eksponentes izvirzījumu Teilora rindā, (1.11) var pārrakstīt

$$\begin{aligned}
\varphi_{S_n}(t) =& e^{-t^2/2}\left[1 + n^{-1/2}(3!)^{-1}\kappa_3(it)^3 + [n^{-1/2}(3!)^{-1}\kappa_3(it)^3]^2/2 + \ldots\right]\\
&\cdot[1 + n^{-1}(4!)^{-1}\kappa_4(it)^4 + [n^{-1}(4!)^{-1}\kappa_4(it)^4]^2/2 + \ldots]\\
&\cdot[1 + n^{-3/2}(5!)^{-1}\kappa_5(it)^5 + [n^{-3/2}(5!)^{-1}\kappa_5(it)^5]^2/2 + \ldots]\\
&\cdot\ldots\cdot[1 + n^{-(j-2)/2}(j!)^{-1}\kappa_j(it)^j + [n^{-(j-2)/2}(j!)^{-1}\kappa_j(it)^j]^2/2 + \ldots]\cdot\ldots
\end{aligned}$$

Sareizinot un savelkot kopā visus saskaitāmos pie $n^{-(j-2)/2}$ pakāpēm, iegūstam

$$\begin{aligned}
\varphi_{S_n}(t) =& e^{-t^2/2}[1 + \{(3!)^{-1}\kappa_3(it)^3\}n^{-1/2} + \{(4!)^{-1}\kappa_4(it)^4\\
&+ [(3!)^{-1}\kappa_3(it)^3]^2/2\}n^{-1} + \{(3!)^{-1}\kappa_3(it)^3(4!)^{-1}\kappa_4(it)^4\\
&+ (5!)^{-1}\kappa_5(it)^5\}n^{-3/2} + \ldots]. \qquad (1.12)
\end{aligned}$$

Tātad (1.12) varam pārrakstīt formā

$$\varphi_{S_n}(t) = e^{-t^2/2}\{1 + n^{-1/2}r_1(it) + n^{-1}r_2(it) + \ldots + n^{-j/2}r_j(it) + \ldots\}, \qquad (1.13)$$

kur

$$r_1(u) = \frac{1}{6}\kappa_3u^3,\ r_2(u) = \frac{1}{24}\kappa_4u^4 + \frac{1}{72}\kappa_3^2u^6,$$
$$r_3(u) = \frac{1}{1296}\kappa_3u^9 + \frac{1}{144}\kappa_3\kappa_4u^7 + \frac{1}{120}\kappa_5u^5,\ \text{utt.} \qquad (1.14)$$

Pārrakstot (1.13) formā

$$\varphi_{S_n}(t) = e^{-t^2/2} + n^{-1/2} r_1(it) e^{-t^2/2} + n^{-1} r_2(it) e^{-t^2/2} + \ldots + n^{-j/2} r_j(it) e^{-t^2/2} + \ldots \quad (1.15)$$

un izmantojot, ka

$$\varphi_{S_n}(t) = \int_{-\infty}^{\infty} e^{itx} dP(S_n \leq x)$$

un

$$e^{-t^2/2} = \int_{-\infty}^{\infty} e^{itx} d\Phi(x), \quad (1.16)$$

kur Φ ir standartnormālā sadalījuma funkcija, apvienojot (1.15) un (1.16), varam iegūt inverso izvirzījumu

$$\mathbb{P}(S_n \leq x) = \Phi(x) + n^{-1/2} R_1(x) + n^{-1} R_2(x) + \ldots + n^{-j/2} R_j(x) + \ldots, \quad (1.17)$$

kur $R_j(x)$ ir funkcija, kuras Furjē – Stiltjesa transformācija ir $r_j(it)e^{-t^2/2}$, t.i.,

$$\int_{-\infty}^{\infty} e^{itx} dR_j(x) = r_j(it) e^{-t^2/2}.$$

Lai pilnībā noteiktu statistikas S_n sadalījuma aproksimāciju, vēl jāatrod $R_j(x)$. Atkārtoti parciāli integrējot (1.16), var iegūt

$$\begin{aligned} e^{-t^2/2} &= (-it)^{-1} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(1)}(x) \\ &= (-it)^{-2} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(2)}(x) \\ &\vdots \\ &= (-it)^{-j} \int_{-\infty}^{\infty} e^{itx} d\Phi^{(j)}(x), \end{aligned}$$

kur $\Phi^{(j)} = (d/dx)^j \Phi(x)$. Tādējādi

$$\int_{-\infty}^{\infty} e^{itx} d\{(-D)^j \Phi(x)\} = (it)^j e^{-t^2/2}, \quad (1.18)$$

kur D ir diferenciāloperators d/dx, $r_j(-D)$ tiek interpretēts kā polinoms, kas atkarīgs no D tā, lai arī $r_j(-D)$ ir diferenciāloperators un izmantojot (1.18), iegūstam

$$\int_{-\infty}^{\infty} e^{itx} d\{r_j(-D)\Phi(x)\} = r_j(it)^j e^{-t^2/2}. \quad (1.19)$$

Tādējādi no (1.19) izriet, ka meklējamā funkcija R_j ir formā

$$R_j(x) = r_j(-D)\Phi(x). \quad (1.20)$$

Un $j \geq 1$,

$$(-D)^j \Phi(x) = -H_{e_{j-1}}(x)\phi(x),$$

kur funkcijas H_{e_j} ir Hermīta polinomi.

Definīcija 3. [19] Par $k-$tās kārtas Hermīta polinomu sauc ortogonālu polinomu formā

$$H_{e_k} = (-1)^n e^{x^2/2} \frac{\partial^k}{\partial x^k} e^{-x^2/2}.$$

Piemēram, pirmie 5 Hermīta polinomi ir

$$H_{e_0}(x) = 1, \; H_{e_1}(x) = x, \; H_{e_2}(x) = x^2 - 1,$$
$$H_{e_3}(x) = x(x^2 - 3), \; H_{e_4}(x) = x^4 - 6x^2 + 3.$$

Jāatzīmē, ka H_{e_j} ir $j-$tās pakāpes polinoms un tas ir nepāra pakāpes, ja j ir nepāra skaitlis un pāra pakāpes, ja j ir pāra skaitlis. Izmantojot (1.14) un (1.20), varam iegūt, ka

$$R_1(x) = -\frac{1}{6}\kappa_3(x^2 - 1)\phi(x),$$
$$R_2(x) = -x\{\frac{1}{24}\kappa_4(x^2 - 3) + \frac{1}{72}\kappa_3^2(x^4 - 10x^2 + 15)\}\phi(x), \ldots.$$

Tātad $j \geq 1$,

$$R_j(x) = p_j(x)\phi(x),$$

kur p_j ir $(3j-1)-$ās pakāpes polinomi un polinoma pakāpe ir pāra vai nepāra atkarībā no tā, vai j ir pāra vai nepāra skaitlis. Tagad (1.17) var pārrakstīt formā

$$\mathbb{P}(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \ldots + n^{-j/2}p_j(x)\phi(x) + \ldots, \quad (1.21)$$

kas ir Edgeworth izvirzījums sadalījumam $\mathbb{P}(S_n \leq x)$. Tātad esam ieguvuši alternatīvu $\mathbb{P}(S_n \leq x)$ aproksimācijai, proti, Edgeworth izvirzījumu, kas uzlabo normālā sadalījuma aproksimāciju, jo (1.21) koeficients pie $n^{-1/2}$ uzlabo asimetrijas (skewness) efektu un pie n^{-1} – ekscesa (kurtosis) efektu.

Parasti izvirzījums (1.21) ir pieejams tikai kā asimptotisks izvirzījums, kas nozīmē, ka izvirzījumā tiek iekļauti saskaitāmie līdz noteiktai kārtai un atlikušo (neiekļauto) saskaitāmo kārta ir mazāka par pēdējā iekļautā saskaitāmā kārtu,

$$\mathbb{P}(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \ldots + n^{-j/2}p_j(x)\phi(x) + o(n^{j-2}). \quad (1.22)$$

Izvirzījums (1.22) ir spēkā fiksētam j, kad $n \to \infty$.

1.2. Edgeworth izvirzījums vispārējā gadījumā

Tagad apskatīsim Edgeworth izvirzījumus vispārīgākā gadījumā. Ar S_n apzīmēsim statistiku, kas asimptotiski ir standartnormāli sadalīta, un kas ir formā $S_n = n^{1/2}(\hat{\theta}-\theta_0)/\sigma$ vai $S_n = n^{1/2}(\hat{\theta}-\theta_0)/\hat{\sigma}$, kur $\hat{\sigma}^2$ ir būtisks dispersijas σ^2 novērtējums. Ar $\varphi_n(t)$ apzīmēsim S_n raksturīgo funkciju un ar $\kappa_{j,n}$ apzīmēsim S_n $j-$to kumulantu.

$$\begin{aligned}\varphi_n(t) &= \mathbb{E}\{\exp(itS_n)\} \\ &= \exp\{\kappa_{1,n}it + \frac{1}{2}\kappa_{2,n}(it)^2 + \ldots + \frac{1}{j!}\kappa_{j,n}(it)^j + \ldots\}.\end{aligned} \tag{1.23}$$

Vairumā gadījumu $\kappa_{j,n}$ ir ar kārtu $n^{-(j-2)/2}$ un var tikt izvirzīts sekojošā veidā

$$\kappa_{j,n} = n^{-(j-2)/2}\left(k_{j,1} + n^{-1}k_{j,2} + n^{-1}k_{j,2}\right),\ j \geq 1, \tag{1.24}$$

skatīt [19], kur $k_{1,1} = 0$ un $k_{2,1} = 1$, statistika S_n ir centrēta un normēta tā, lai $\kappa_{1,n} = \mathbb{E}(S_n) \to 0$ un $\kappa_{2,n} = \mathbb{D}(S_n) \to 1$. Kad S_n ir normalizēta neatkarīgu un vienādi sadalītu gadījuma lielumu summa, tad (1.24) netieši ir izteiksmē (1.10) un

$$\kappa_{j,n} = n^{-(j-2)/2}\kappa_j,\ j \geq 2,$$

kur κ_j ir $j-$tais Y kumulants. Apvienojot (1.23) un (1.24), varam iegūt, ka

$$\begin{aligned}\varphi_n(t) =& \exp\{-\frac{1}{2}t^2 + n^{-1/2}\left[k_{1,2}it + \frac{1}{3!}k_{3,1}(it)^3\right] + n^{-1}\left[\frac{1}{2}k_{2,2}(it)^2 + \frac{1}{4!}k_{4,1}(it)^4\right] + \ldots\} \\ =& e^{-t^2/2}\{1 + n^{-1/2}\left[k_{1,2}it + \frac{1}{6}k_{3,1}(it)^3\right] \\ &+ n^{-1}\left[\frac{1}{2}k_{2,2}(it)^2 + \frac{1}{24}k_{4,1}(it)^4 + \frac{1}{2}\left[k_{1,2}it + \frac{1}{6}k_{3,1}(it)^3\right]^2\right] + O(n^{-3/2})\}\end{aligned} \tag{1.25}$$

Vispārinot (1.25), var iegūt, ka

$$\varphi_n(t) = e^{-t^2/2} + n^{-1/2}r_1(it)e^{-t^2/2} + n^{-1}r_2(it)e^{-t^2/2} + \ldots + n^{-j/2}r_j(it)e^{-t^2/2} + \ldots,$$

kur r_j ir polinomi ar kārtu ne lielāku par $3j$. Šis izvirzījums formāli sakrīt ar (1.15). Darbojoties līdzīgi kā vidējās vērtības gadījumā, varam iegūt (1.21) analogu

$$\mathbb{P}(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \ldots + n^{-j/2}p_j(x)\phi(x) + \ldots, \tag{1.26}$$

kur p_j ir polinomi ar kārtu ne lielāku par $3j-1$. No iepriekšējās nodaļas atsauksim atmiņā, ka

$$\int_{-\infty}^{\infty} e^{itx}d\{-H_{e_{j-1}}(x)\phi(x)\} = (it)^j e^{-t^2/2},\ j \geq 1,$$

kur H_{e_j} ir j-tais Hermīta polinoms, tādējādi izmantojot (1.25), var iegūt, ka

$$r_1(it) = k_{1,2}it + \frac{1}{6}k_{3,1}(it)^3$$

un

$$r_2(it) = \frac{1}{2}(k_{2,2} + k_{1,2}^2)(it)^2 + \frac{1}{24}(k_{4,1} + 4k_{1,2}k_{3,1})(it)^4 + \frac{1}{72}k_{3,1}^2(it)^6,$$

savukārt

$$p_1(x) = -\{k_{1,2} + \frac{1}{6}k_{3,1}H_{e_2}(x)\} = -\{k_{1,2} + \frac{1}{6}k_{3,1}(x^2 - 1)\}$$

un

$$\begin{aligned} p_2(x) =& -\{\frac{1}{2}(k_{2,2} + k_{1,2}^2)H_{e_1}(x) + \frac{1}{24}(k_{4,1} + 4k_{1,2}k_{3,1})H_{e_3}(x) + \frac{1}{72}k_{3,1}^2 H_{e_5}(x)\} \\ =& -x\{\frac{1}{2}(k_{2,2} + k_{1,2}^2) + \frac{1}{24}(k_{4,1} + 4k_{1,2}k_{3,1})(x^2 - 3) \\ &+ \frac{1}{72}k_{3,1}^2(x^4 - 10x^2 + 15).\} \end{aligned}$$

Tāpat kā iepriekš varam iegūt, ka $j \geq 1$,

$$\mathbb{P}(S_n \leq x) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + n^{-1}p_2(x)\phi(x) + \ldots + n^{-j/2}p_j(x)\phi(x) + o(n^{-j/2}). \tag{1.27}$$

2. Empīriskās ticamības metode vienas izlases gadījumā

Empīriskās ticamības metode ir samērā jauna metode, kuras pamatlicējs ir Owen [1, 2, 3, 4]. EL metodes vispārējais gadījums vienas izlases gadījumā ļoti labi izklāstīts [21], kur tiek apskatīti *i.i.d.* $d-$dimensionāli gadījuma lielumi $X_1, \ldots, X_n$ ar nezināmu sadalījuma funkciju F un p–dimensionālu parametru θ.

Lai vienkāršotu pierakstu un lai metodes būtību padarītu uzskatāmāku, apskatīsim gadījumu, kad $d = 1$ un $p = 1$. Vēl sīkāku un precīzāku metodes aprakstu skatīt diplomdarbā [22].

Pieņemsim, ka informācija par F un θ ir dota funkcijas $g(X, \theta)$ formā, kur $E\{g(X, \theta)\} = 0$, piemēram, vidējās vērtības gadījumā $g(X_i, \theta) = X_i - \mu$.

Empīriskā ticamības funkcija ir formā

$$L(G) = \prod_{i=1}^{n} \mathbb{P}(X = X_i) = \prod_{i=1}^{n} p_i,$$

tas nozīmē, ka vislielākās ticamības funkcija tiek aprakstīta diskrētiem sadalījumiem ar punktveida diskrētām varbūtībām datu punktos. Funkciju $L(G)$ maksimizē empīriskā sadalījuma funkcija F_n. Līdzīgi kā parametriskajā gadījumā, arī empīriskās ticamības attiecību var izmantot hipotēžu pārbaudei un ticamības intervālu konstruēšanai. Sadalījuma funkcijai F definēsim empīriskās ticamības attiecību:

$$R(G) = \frac{L(G)}{L(F_n)} = \prod_{i=1}^{n} np_i.$$

Šajā gadījumā funkciju $L(F) = \prod_{i=1}^{n} p_i$ maksimizēsim pie ierobežojumiem

$$p_i \geq 0, \ \sum_i p_i = 1, \ \sum_i p_i g(X_i, \theta) = 0. \tag{2.1}$$

Ja 0 pieder izliektajai čaulai, ko veido punkti $g(X_1, \theta), \ldots, g(X_n, \theta)$, θ eksistē viens vienīgs maksimums. Maksimumu var atrast ar Lagranža reizinātāju palīdzību.

$$H = \sum_i \ln(p_i) + \lambda_0 \left(1 - \sum_i p_i\right) - n\lambda \sum_i p_i g(X_i, \theta), \tag{2.2}$$

kur λ_0 un λ ir Lagranža reizinātāji. Atvasinot izteiksmi (2.2) pēc p_i, iegūstam

$$\frac{\partial H}{\partial p_i} = \frac{1}{p_i} - \lambda_0 - n\lambda g(X_i, \theta) = 0, \ \sum_i p_i \frac{\partial H}{\partial p_i} = n - \lambda_0 = 0 \Rightarrow \lambda_0 = n$$

un

$$p_i = \frac{1}{n}\frac{1}{1+\lambda g(X_i,\theta)}. \tag{2.3}$$

Ņemot vērā trešo (2.1) ierobežojumu, iegūstam, ka

$$\sum_i p_i g(X_i,\theta) = \frac{1}{n}\sum_i \frac{g(X_i,\theta)}{1+\lambda g(X_i,\theta)} = 0. \tag{2.4}$$

Lagranža reizinātājs λ var tikt noteikts kā funkcija no θ. Tā kā $0 \le p_i \le 1$, tad λ un θ jāizpildās $1+\lambda g(X_i,\theta) \ge 1/n$ katram i.

Fiksētam θ, $D_\theta = \{\lambda : 1+\lambda g(X_i,\theta) \ge 1/n\}$. Kopa D_θ ir izliekta, slēgta un ierobežota, ja 0 ir izliektajā čaulā, ko veido $g(X_i,\theta)$. Kā arī

$$\frac{\partial}{\partial\lambda}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i,\theta)}{1+\lambda g(X_i,\theta)}\right\} = -\frac{1}{n}\sum_{i=1}^{n}\frac{(g(X_i,\theta))^2}{(1+\lambda g(X_i,\theta))^2}$$

ir negatīvs pēc argumenta λ, nodrošinot to, ka $\sum_{i=1}^{n}(g(X_i,\theta))^2$ ir pozitīva. $\lambda=\lambda(\theta)$ ir nepārtraukti diferencējama funkcija pēc θ.

Empīrisko ticamību parametram θ var pārrakstīt

$$L(\theta) = \prod_{i=1}^{n}\left\{\frac{1}{n}\frac{1}{1+\lambda(\theta)g(X_i,\theta)}\right\}$$

un empīriskās ticamības attiecība ir

$$R(\theta) = \frac{L(\theta)}{L(\hat{\theta})} = n^n\prod_{i=1}^{n}\left\{\frac{1}{n}\frac{1}{1+\lambda(\theta)g(X_i,\theta)}\right\} = \prod_{i=1}^{n}\{1+\lambda(\theta)g(X_i,\theta)\}^{-1},$$

taču praksē biežāk empīriskās ticamības vietā tiek izmantota logaritmiskā empīriskā ticamība un logaritmiskā empīriskās ticamības attiecība

$$W(\theta) = -2\log R(\theta) = 2[l(\theta)-l(\hat{\theta})] = 2\sum_{i=1}^{n}\log\left[1+\lambda(\theta)g(X_i,\theta)\right]. \tag{2.5}$$

Minimizējot $l(\theta)$, var iegūt parametra θ novērtējumu $\tilde{\theta}$, t.i., $\tilde{\theta} = \arg\min_\theta l(\theta)$, kur $\tilde{\theta}$ ir empīriskās ticamības novērtējums. Turklāt ar $\tilde{\theta}$ palīdzību var noteikt p_i novērtējumus $\tilde{p}_i$ un sadalījuma funkcijas F novērtējumu $\tilde{F}_n$. Šajā gadījumā $\tilde{\theta} = \hat{\theta}$ minimizē $l(\theta)$, kur $\hat{\theta}$ ir vienādojumu $\sum_{i=1}^{n} g(X_i,\theta) = 0$ atrisinājums. Turklāt $\tilde{p}_i = n^{-1}$ un $\tilde{F}_n$ ir empīriskā sadalījuma funkcija. Qin un Lawless [21] pie zināmiem nosacījumiem pierādīja, ka logaritmiskā empīriskās ticamības attiecības statistika

$$W(\theta_0) \to_d \chi_1^2,$$

kad $n \to \infty$ un $H_0 : \theta = \theta_0$ ir spēkā, kur $W(\theta)$ ir uzdots ar (2.5). Qin un Lawless [21] iegūtie rezultāti dod iespēju novērtēt parametrus un konstruēt ticamības intervālus. Tāpat kā parametriskajā gadījumā, arī neparametriskajā gadījumā logaritmiskajai attiecības testa statistikai ir χ^2 sadalījums un empīriskās ticamības novērtējums $\tilde{\theta}$ ir asimptotiski normāli sadalīts.

2.1. EL metode vidējai vērtībai

Gadījumā, kad aplūkojam vidējās vērtības problemātiku, proti, nezināmais parametrs, ko vēlamies novērtēt ir vidējā vērtība μ, $g(X_i, \theta) = X_i - \mu$ un tādējādi (2.4) var pārrakstīt formā

$$\frac{1}{n} \sum_i \frac{X_i - \mu}{1 + \lambda(X_i - \mu)} = 0 \tag{2.6}$$

un (2.5) var pārrakstīt formā

$$W(\mu) = 2 \sum_{i=1}^{n} \log\{1 + \lambda(X_i - \mu)\}. \tag{2.7}$$

2.2. Bārtleta korekcija vidējai vērtībai

Viena no galvenajām EL priekšrocībām, salīdzinot ar butstrapu un citām neparametriskajām metodēm, ir tieši Bārtleta korekcijas iespējamība. Līdzīgi kā parametriskajā gadījumā, arī empīriskajai logaritmiskajai ticamības attiecībai ir pieejama Bārtleta korekcija, ar kuras palīdzību var uzlabot χ^2 sadalījuma aproksimācijas kārtu.

Izmantojot [23], apskatīsim Bārtleta korekciju vidējās vērtības gadījumā.

Ja $Y_1, Y_2, \ldots, Y_n$ ir gadījuma lielumi, kuriem $Y_i = \sigma^{-1}(X_i - \mu_0)$, tad (2.7) var pārrakstīt

$$W(\mu) = 2 \sum_{i=1}^{n} \log\{1 + u(Y_i - \nu)\}, \tag{2.8}$$

kur $u = \sigma\lambda$, $\nu = \sigma^{-1}(\mu - \mu_0)$ un u ir spēkā

$$\frac{1}{n} \sum_i \frac{Y_i - \nu}{1 + u(Y_i - \nu)} = 0. \tag{2.9}$$

Salīdzinot (2.6) un (2.7) ar (2.8) un (2.9), redzams, ka nezaudējot problēmas vispārīgumu, varam pieņemt, ka $\mu_0 = 0$ un $\sigma^2 = 1$.

1. Lai noteiktu parametru λ, izvirzīsim (2.6) Teilora rindā gadījumā, kad $\mu = 0$ un $\sigma^2 = 1$, tātad šajā gadījumā (2.6) pārrakstās formā

$$T = n^{-1} \sum_i X_i \{1 + \lambda X_i\}^{-1} = 0. \tag{2.10}$$

Izvirzot T Teilora rindā, iegūstam

$$\begin{aligned} T &= n^{-1} \sum_i X_i \{1 + \lambda X_i\}^{-1} \\ &= n^{-1} \sum_i \left[1 - \lambda X_i + (\lambda X_i)^2 - (\lambda X_i)^3 + \ldots\right] X_i. \end{aligned} \tag{2.11}$$

Ievedam apzīmējumus, $\alpha_k = \mathbb{E}(X^k)$ un $A_k = n^{-1} \sum_i X_i^k - \alpha_k$, tātad $n^{-1} \sum_i X_i^k = A_k + \alpha_k$.

Piezīmēsim, ka būtisku lomu spēlē fakts, ka $\lambda = O_p(n^{-1/2})$, kas pierādīts, piemēram, [2].

Izmantojot ievestos apzīmējumus, (2.11) var pārrakstīt

$$T = (A_1 + \alpha_1) - \lambda(A_2 + \alpha_2) + \lambda^2(A_3 + \alpha_3) - \lambda^3(A_4 + \alpha_4) + O_p(n^{-2}).$$

Šajā gadījumā $\alpha_1 = 0$ un $\alpha_2 = 1$, tāpēc T var pārrakstīt

$$T = A_1 - \lambda A_2 - \lambda + \lambda^2 A_3 + \lambda^2 \alpha_3 - \lambda^3 A_4 - \lambda^3 \alpha_4 + O_p(n^{-2})$$

un tā kā $\lambda = O_p(n^{-1/2})$, tad $\lambda^3 A_4 = O_p(n^{-2})$ un

$$T = A_1 - \lambda A_2 - \lambda + \lambda^2 A_3 + \lambda^2 \alpha_3 - \lambda^3 \alpha_4 + O_p(n^{-2}).$$

Pielīdzinot $T = 0$, jāizsaka λ. Izmantojot [24], λ izvirzījumu var uzrakstīt formā $\lambda = \lambda_1 + \lambda_2 + \lambda_3 + O_p(n^{-2})$, kur λ_i ir $O_p(n^{-i/2}), i = \overline{1,3}$ un

$$\begin{aligned} &\lambda_1 = A_1,\ \lambda_2 = \alpha_3 A_1^2 - A_1 A_2, \\ &\lambda_3 = A_1 A_2^2 + A_1^2 A_3 + 2\alpha_3^2 A_1^3 - 3\alpha_3 A_1^2 A_2 - \alpha_4 A_1^3, \end{aligned}$$

ko var iegūt izmantojot izteiksmi $T = 0$ pārrakstot formā

$$\lambda = A_1 - \lambda A_2 + \lambda^2 A_3 + \lambda^2 \alpha_3 - \lambda^3 \alpha_4 + O_p(n^{-2})$$

Apskatīsim, piemēram, kā var iegūt λ_1 un λ_2.

$$\begin{aligned} &A_1 = O_p(n^{-1/2}) \Rightarrow \lambda_1 = A_1, \\ &\lambda A_2 = O_p(n^{-1}), \lambda^2 \alpha_3 = O_p(n^{-1}) \Rightarrow \lambda_2 = -\lambda A_2 + \lambda^2 \alpha_3 \Rightarrow \lambda_2 = -A_1 A_2 + A_1^2 \alpha_3, \end{aligned}$$

Tātad

$$\lambda = A_1 + \alpha_3 A_1^2 - A_1 A_2 + A_1 A_2^2 + A_1^2 A_3 + 2\alpha_3^2 A_1^3 - 3\alpha_3 A_1^2 A_2 - \alpha_4 A_1^3 + O_p(n^{-2}). \tag{2.12}$$

Izvirzot $\sum_i \lambda X_i(1+\lambda X_i)^{-1}=0$ Teilora rindā, izmantosim izvirzījumu (2.10), tādējādi iegūstot, ka

$$\sum_i \lambda X_i = \sum_i ((\lambda X_i)^2 - (\lambda X_i)^3 + (\lambda X_i)^4 - \ldots). \tag{2.13}$$

Turpinām ar empīriskās logaritmiskās attiecības izvirzīšanu Teilora rindā

$$\begin{aligned} W(0) = W_0 &= 2\sum_i \log(1+\lambda X_i) \\ &= 2\sum_i \{\lambda X_i - \frac{1}{2}(\lambda X_i)^2 + \frac{1}{3}(\lambda X_i)^3 - \frac{1}{4}(\lambda X_i)^4\} + O_p(n^{-3/2}). \end{aligned} \tag{2.14}$$

Pareizinot W_0 ar n^{-1} un aizstājot $\sum_i \lambda X_i$ ar (2.13), kā arī izmantojot iepriekš ieviestos apzīmējumus α_k un A_k, varam iegūt

$$\begin{aligned} n^{-1}W_0 &= n^{-1}\sum_i \left[(\lambda X_i)^2 - \frac{4}{3}(\lambda X_i)^3 + \frac{3}{2}(\lambda X_i)^4\right] + O_p(n^{-5/2}) \\ &= \lambda^2 A_2 + \lambda^2 - \frac{4}{3}\lambda^3 A_3 - \frac{4}{3}\lambda^3\alpha_3 + \frac{3}{2}\lambda^4\alpha_4 + O_p(n^{-5/2}) \end{aligned} \tag{2.15}$$

un ievietojot λ izvirzījumu (2.12) izteiksmē (2.15), iegūstam

$$\begin{aligned} n^{-1}W_0 =& A_1^2 - A_2A_1^2 + \frac{2}{3}\alpha_3 A_1^3 + A_2^2A_1^2 + \frac{2}{3}A_3A_1^3 - 2\alpha_3A_2A_1^3 \\ &+ \alpha_3^2A_1^4 - \frac{1}{2}\alpha_4A_1^4 + O_p(n^{-5/2}). \end{aligned} \tag{2.16}$$

2. $n^{-1}W_0$ var uzrakstīt formā $n^{-1}W_0 = R^2 + O_p(n^{-5/2})$, kur $R = R_1+R_2+R_3+O_p(n^{-2})$ un R_i ir $O_p(n^{-i/2})$, skatīt [23]. Var noteikt, ka

$$\begin{aligned} &R_1 = A_1,\ R_2 = -\frac{1}{2}A_2A_1 + \frac{1}{3}\alpha_3A_1^2, \\ &R_3 = \frac{3}{8}A_2^2A_1 + \frac{1}{3}A_3A_1^2 - \frac{5}{6}\alpha_3A_2A_1^2 + \frac{4}{9}\alpha_3^2A_1^3 - \frac{1}{4}\alpha_4A_1^3. \end{aligned}$$

Kā rezultātā $n^{-1}W_0$ var pārrakstīt formā

$$n^{-1}W_0 = R_1^2 + 2R_1R_2 + 2R_1R_3 + R_2^2 + O_p(n^{-5/2}).$$

Johnson un Kotz [25] pierādīja, ka $s-$tais nR^2 kumulants ir

$$\kappa_s = 2^{s-1}(s-1)!\{\mathbb{E}(nR^2)\}^s + O(n^{-3/2}).$$

Neņemot vērā $O(n^{-3/2})$, varam iegūt, ka $s-$tais $(nR^2)\{\mathbb{E}(nR^2)\}^{-1}$ kumulants ir $2^{s-1}(s-1)!$, kas savukārt ir arī $s-$tais χ_1^2 kumulants.

No tā, ka $n^{1/2}R$ sadalījumam var uzrakstīt Edgeworth izvirzījumu, kas ir līdzīgā formā kā (1.27) un tā kā $W_0 = nR^2 + O_p(n^{-3/2})$, tad

$$\begin{aligned} \mathbb{P}\left[W_0\{\mathbb{E}(nR^2)\}^{-1} \le z\right] &= \mathbb{P}\left[(nR^2)\{\mathbb{E}(nR^2)\}^{-1} \le z\right] \\ &= \mathbb{P}(\chi_1^2 \le z) + O(n^{-3/2}). \end{aligned} \tag{2.17}$$

Balstoties pamatā uz pāra un nepāra pakāpju polinomiem Edgeworth izvirzījumā, empīriskajai ticamībai tāpat kā parametriskajai ticamībai var parādīt, ka saskaitāmais $O(n^{-3/2})$ izteiksmē (2.17) patiesībā ir $O(n^{-2})$, skatīt [26].

Tātad

$$\mathbb{P}\left[W_0\{\mathbb{E}(nR^2)\}^{-1} \le z\right] = \mathbb{P}(\chi_1^2 \le z) + O(n^{-2}).$$

3. Tagad jāatrod $\mathbb{E}(nR^2)$. Definējot, ka $t_1 = \alpha_3^2$ un $t_2 = \alpha_4$, var pierādīt, ka

$$\begin{aligned} &\mathbb{E}(R_1^2) = n^{-1}, \\ &\mathbb{E}(R_1R_2) = n^{-2}\left(\frac{1}{3}t_1 - \frac{1}{2}t_2\right) + O(n^{-3}), \\ &\mathbb{E}(R_1R_3) = n^{-2}\left(-\frac{5}{12}t_1 + \frac{5}{8}t_2\right) + O(n^{-3}), \\ &\mathbb{E}(R_2^2) = n^{-2}\left(-\frac{1}{6}t_1 + \frac{1}{4}t_2\right) + O(n^{-3}). \end{aligned}$$

Tādējādi

$$\begin{aligned} \mathbb{E}(nR^2) &= n\{\mathbb{E}(R_1^2) + 2\mathbb{E}(R_1R_2) + 2\mathbb{E}(R_1R_3) + \mathbb{E}(R_2^2)\} + O(n^{-2}) \\ &= 1 + n^{-1}\left(-\frac{1}{3}t_1 + \frac{1}{2}t_2\right) + O(n^{-2}) \\ &= 1 + n^{-1}\left(-\frac{1}{3}\alpha_3^2 + \frac{1}{2}\alpha_4\right) + O(n^{-2}) \end{aligned} \tag{2.18}$$

un

$$\{\mathbb{E}(nR^2)\}^{-1} = 1 - n^{-1}\left(-\frac{1}{3}\alpha_3^2 + \frac{1}{2}\alpha_4\right) + O(n^{-2}). \tag{2.19}$$

Ja nekoriģēto EL ticamības intervālu parametram μ definējam

$$I_\alpha = \{\mu : W(\mu) \le c_\alpha\},$$

kur c_α ir tāds, ka $\mathbb{P}(\chi_1^2 \le c_\alpha) = 1 - \alpha$, tad EL ticamības intervālu ar Bārtleta korekciju varam definēt

$$I'_\alpha = \{\mu : W(\mu) \le c_\alpha(1 + n^{-1}a)\},$$

kur a vidējās vērtības gadījumā var iegūt no (2.18). Praksē īsto a vērtību nevarēs noteikt, tāpēc tā vietā jāizmanto $\hat{a}$,

$$\hat{a} = -\frac{1}{3}\hat{\alpha}_3^2 + \frac{1}{2}\hat{\alpha}_4, \tag{2.20}$$

kas novērtēts no datiem.

2.3. Simulācijas Bārtleta korekcijai vidējai vērtībai

Lai noskaidrotu Bārtleta korekcijas ietekmi uz pārklājuma precizitāti ticamības intervāliem, kas konstruēti vidējai vērtībai, izmantojot korekcijas parametra īsto vērtību a un tā novērtējumu $\hat{a}$, kas noteikts no datiem, izmantosim diplomdarbā [22] iegūtās simulāciju tabulas. Dažādiem sadalījumiem, ticamības līmeņiem un izlašu apjomiem tika uzģenerētas 10000 izlases, lai konstruētu divpusējus ticamības intervālus, izmantojot 3 atšķirīgas metodes. Pirmā metode EL ir nekoriģēta empīriskās ticamības metode, kas izmanto χ^2 kritisko vērtību c_α. Otrā metode $EL_{B_{teo}}$, konstruējot ticamības intervālus ar EL metodes palīdzību, χ^2 kritiskās vērtības c_α vietā izmanto $c_\alpha/(1-an^{-1})$. Šajā gadījumā jābūt zināmām populācijas īstajām momentu vērtībām, kas praksē parasti nav zināmas. Pēdējā metode $EL_{B_{nov}}$ a vietā izmanto $\hat{a}$, kas definēts ar (2.20).

1. tabula Bārtleta korekcija $N(0,1)$ sadalījuma datiem

		Ticamības līmenis		
		90%	95%	99%
$n=10$	EL	0.8429	0.8975	0.9552
	$EL_{B_{teo}}$	0.8674	0.9162	0.9650
	$EL_{B_{nov}}$	0.8616	0.9118	0.9633
$n=20$	EL	0.8786	0.9334	0.9787
	$EL_{B_{teo}}$	0.8924	0.9420	0.9830
	$EL_{B_{nov}}$	0.8900	0.9410	0.9825
$n=50$	EL	0.8910	0.9456	0.9867
	$EL_{B_{teo}}$	0.8952	0.9486	0.9879
	$EL_{B_{nov}}$	0.8948	0.9482	0.9877

1.tabulā redzamie rezultāti ir iegūti, izmantojot standartnormālā sadalījuma $N(0,1)$ datus, 2.tabulā redzamie rezultāti ir iegūti, izmantojot χ_1^2 sadalījuma datus un 3.tabulā redzamie rezultāti ir iegūti, izmantojot t_5 sadalījuma datus. Īstās parametra a vērtības apskatītajiem sadalījumiem ir attiecīgi: 3/2, 29/6 un 9/2.

2. tabula Bārtleta korekcija χ_1^2 sadalījuma datiem

		Ticamības līmenis		
		90%	95%	99%
$n=10$	EL	0.7704	0.8329	0.8975
	$EL_{B_{teo}}$	0.8389	0.8826	0.9297
	$EL_{B_{nov}}$	0.7906	0.8480	0.9074
$n=20$	EL	0.8394	0.8925	0.9523
	$EL_{B_{teo}}$	0.8741	0.9198	0.9669
	$EL_{B_{nov}}$	0.8540	0.9042	0.9587
$n=50$	EL	0.8710	0.9265	0.9776
	$EL_{B_{teo}}$	0.8876	0.9387	0.9818
	$EL_{B_{nov}}$	0.8797	0.9328	0.9793

Aplūkojot iegūtos rezultātus, jāsecina, ka pie maziem izlašu apjomiem ar nekoriģētās empīriskās ticamības metodes palīdzību konstruēto ticamības intervālu pārklājuma precizitāte diezgan ievērojami atšķiras no uzdotā ticamības līmeņa. $EL_{B_{teo}}$ un $EL_{B_{nov}}$ ievērojami uzlabo pārklājuma precizitāti, pie tam palielinoties izlašu apjomiem, $EL_{B_{teo}}$ rezultāti ir tikai nedaudz labāki par $EL_{B_{nov}}$. Vissarežģītākā situācija ir ar χ^2 sadalījumu (2.tabula), ar $EL_{B_{teo}}$ iegūtie rezultāti ir diezgan apmierinoši, taču $EL_{B_{nov}}$ minimāli uzlabo empīrisko ticamību, ko var izskaidrot ar χ^2 sadalījuma asimetriskumu.

Salīdzinot 1. un 3.tabulu, redzams, ka, ja pie ļoti maziem izlašu apjomiem ($n=10$) $N(0,1)$ ir minimāla atšķirība starp $EL_{B_{teo}}$ un $EL_{B_{nov}}$, tad to pašu nevar teikt par vēl vienu simetrisku sadalījumu t_5, kuram šī atšķirība starp $EL_{B_{teo}}$ un $EL_{B_{nov}}$ ir ievērojama.

3. tabula Bārtleta korekcija t_5 sadalījuma datiem

		Ticamības līmenis		
		90%	95%	99%
$n = 10$	EL	0.8216	0.8862	0.9492
	$EL_{B_{teo}}$	0.8890	0.9334	0.9735
	$EL_{B_{nov}}$	0.8442	0.9025	0.9575
$n = 20$	EL	0.8611	0.9159	0.9753
	$EL_{B_{teo}}$	0.8957	0.9431	0.9864
	$EL_{B_{nov}}$	0.8741	0.9267	0.9794
$n = 50$	EL	0.8881	0.9408	0.9834
	$EL_{B_{teo}}$	0.9040	0.9505	0.9885
	$EL_{B_{nov}}$	0.8951	0.9451	0.9860

2.4. EL metode kvantilēm

Parastajiem EL ticamības intervāliem, kas konstruēti kvantilēm ir relatīvi liela pārklājuma kļūda ar kārtu $n^{-1/2}$ pat divpusējo intervālu gadījumā. Savukārt ar kodolu metožu palīdzību nogludinātās EL ticamības intervālu pārklājuma kļūda ir ar kārtu n^{-1}, pie samērā plašas nogludināšanas parametra izvēles. Pirmie EL gludināšanas metodi piedāvāja Chen un Hall [8], vēlāk Zhou un Jing [27] piedāvāja vēl vienu metodi. Šajā gadījumā akcentēsimies uz Chen un Hall [8] metodi, kas pieļauj arī Bārtleta korekciju.

Chen un Hall [8] izmanto kodolu gludināšanas metodi, lai panāktu, ka ticamības intervālu pārklājuma kļūda ir ar kārtu n^{-1} un parāda, ka pie zināmiem nosacījumiem, izmantojot Bārtleta korekciju, var panākt, ka ticamības intervālu pārklājuma kļūda ir ar kārtu n^{-2}.

Pieņemsim, ka $X_1, X_2, \ldots, X_n$ ir gadījuma izlase no sadalījuma ar sadalījuma funkciju F un mēs vēlamies novērtēt nezināmo parametru $\theta = \theta_q$, kur $\theta_q = F^{-1}(q)$ ir $q-$tā kvantile. Ar G_h apzīmēsim deģenerētās sadalījuma funkcijas G_0 gludināto versiju,

$$G_0(x) = \begin{cases} 1, \, x \geq 0 \\ 0, \text{ pretējā gadījumā} \end{cases} . \tag{2.21}$$

Īsi apskatīsim kodolu gludināšanas metodi.

Pieņemsim, ka $X_1, \ldots, X_n$ ir $i.i.d$ izlase ar blīvuma funkciju f. Kodols tiek definēts kā gluda funkcija K, kurai ir spēkā, ka $K(x) \geq 0$, $\int K(x)dx = 1$, $\int xK(x)dx = 0$ un $\int x^2K(x)dx > 0$.

Minēsim dažus no populārākajiem kodoliem: Epanečnikova kodols

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2/5)/\sqrt{5}, \ |x| < \sqrt{5} \\ 0, \text{ pretējā gadījumā} \end{cases},$$

normālais jeb Gausa kodols

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Definīcija 4. [20] Ja K ir kodols un joslas platums h ir pozitīvs skaitlis, tad

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{x - X_i}{h}\right)$$

ir kodolu blīvuma funkcijas novērtējums. Ar K tiek apzīmēts $r-$tās kārtas kodols, kuram pie $r \geq 2$ un $C \neq 0$ ir spēkā

$$\int u^j K(u)du = \begin{cases} 1, \ j = 0, \\ 0, \ 1 \leq j \leq r, \\ C, \ j = r, \end{cases}, \tag{2.22}$$

un

$$G(t) = \int_{-\infty}^{t} K(x)dx. \tag{2.23}$$

Gadījums, kad $r = 2$, ir visbiežāk sastopamais, tad K ir simetriska blīvuma funkcija un $G(t)$ ir sadalījuma funkcija.

Izmantojot vispārīgo problēmas nostādni vienas izlases gadījumā, piezīmēsim, ka kvantiļu gadījumā $g(X_i, \theta) = G_h(\theta - X_i) - q$, kur $G_h(x) = G(x/h)$ un h ir joslas platums.

Izpildoties zināmiem nosacījumiem Chen un Hall [8] pierādīja, ka $W(\theta)$, kas definēta ar (2.5) ir asimptotisks χ^2_1 sadalījums un ka $\mathbb{P}(\theta \in I_{c_\alpha}) = \alpha + O(n^{-1})$, kur
$I_{c_\alpha} = \{\theta : W(\theta) \leq c\}$ un c pozitīva konstante. Iepriekš jau tika minēts, ka šajā gadījumā ticamības intervālu pārklājuma precizitāti vēl var uzlabot ar Bārtleta korekcijas palīdzību.

2.5. Bārtleta korekcija kvantilēm

Īsi aprakstīsim Chen un Hall publikācijā [8] atrodamo pieeju Bārtleta korekcijai kvantiļu gadījumā. Ar g_i apzīmēsim $g(X_i, \theta) = G_h(\theta - X_i) - q$ un atcerēsimies (2.4), ko var pārrakstīt

$$\frac{1}{n}\sum_i \frac{g_i}{1+\lambda g_i} = 0. \tag{2.24}$$

1. Var parādīt, ka $\lambda = O_p(n^{-1/2} + h^r)$, skatīt [8].

2. Lai no (2.24) noteiktu parametru λ, izvirzīsim (2.24) Teilora rindā, $j \geq 1$,

$$\begin{aligned} 0 &= n^{-1}\sum_{i=1}^{n} g_i\{1 - \lambda g_i + (\lambda g_i)^2 - (\lambda g_i)^3 + \ldots\} \\ &= \sum_{k=1}^{j}(-\lambda)^{k-1}\bar{g}_k + O_p\{(n^{-1/2} + h^r)^j\}, \end{aligned} \tag{2.25}$$

kur $\bar{g}_k = n^{-1}\sum_i g_i^j, j \geq 1$. Atrisinot (2.25) priekš λ, var iegūt, ka

$$\begin{aligned} \lambda = &\bar{g}_2^{-1}\bar{g}_1 + \bar{g}_2^{-3}\bar{g}_3\bar{g}_1^2 + \{2\bar{g}_2^{-5}\bar{g}_3^2 - \bar{g}_2^{-4}\bar{g}_4\}\bar{g}_1^3 \\ &+ \sum_{k=4}^{j} R_{1k}\bar{g}_1^k + O_p\{(n^{-1/2} + h^r)^{j+1}\}, \end{aligned} \tag{2.26}$$

kur ar R_{1k} apzīmē $\bar{g}_2^{-(2k-1)}$, kas pareizināts ar polinomu no $\bar{g}_2, \ldots, \bar{g}_{k+1}$ ar konstantiem koeficientiem.

Izvedumu tam, ka λ ir formā (2.26) skatīt pie Bārtleta korekcijas gludinātajai EL divu izlašu gadījumā.

Piezīmēsim, ka [8] iespējams pieļauta neprecizitāte, tāpēc (2.26) atšķiras no [8] atrodamās λ izteiksmes.

3. Izvirzot $W(\theta)$ Teilora rindā un ievietojot (2.26), varam iegūt

$$\begin{aligned} W(\theta) =&2\sum_{i=1}^{n}\log(1 + \lambda g_i) \\ =&2n\sum_{k=1}^{j+1}(-1)^{k+1}k^{-1}\lambda^k\bar{g}_k + O_p\{(n^{-1/2} + h^r)^{j+2}\}, \\ =&n\{\bar{g}_2^{-1}\bar{g}_1^2 + \frac{2}{3}\bar{g}_2^{-3}\bar{g}_3\bar{g}_1^3 + \left(\bar{g}_2^{-5}\bar{g}_3^2 - \frac{1}{2}\bar{g}_2^{-4}\bar{g}_4\right)\bar{g}_1^4 \\ &- \left(2\bar{g}_2^{-6}\bar{g}_3\bar{g}_4 - 2\bar{g}_2^{-7}\bar{g}_3^3 - \frac{2}{5}\bar{g}_2^{-5}\bar{g}_5\right)\bar{g}_1^5\} \\ &+ n\sum_{k=5}^{j} R_{2k}\bar{g}_1^{k+1} + O_p\{n(n^{-1/2} + h^r)^{j+2}.\} \end{aligned} \tag{2.27}$$

Tā kā jau λ izvirzījumā Chen un Hall [8] pieļāva neprecizitāti, tad arī (2.27) atšķiras no [8] atrodamā $W(\theta)$ izvirzījuma.

4. $W(\theta)$ no (2.27) var pārrakstīt formā

$$W(\theta) = (n^{1/2} S_j')^2,$$

kur

$$\begin{aligned} S_j' = &\bar{g}_2^{-1/2} \{ \bar{g}_1 + \frac{1}{3}\bar{g}_2^{-2}\bar{g}_3\bar{g}_1^2 + \left(\frac{4}{9}\bar{g}_2^{-4}\bar{g}_3^2 - \frac{1}{4}\bar{g}_2^{-3}\bar{g}_4\right)\bar{g}_1^3 \\ &+ \left(\frac{23}{27}\bar{g}_2^{-6}\bar{g}_3^3 - \frac{11}{12}\bar{g}_2^{-5}\bar{g}_3\bar{g}_4 + \frac{1}{5}\bar{g}_2^{-4}\bar{g}_5\right)\bar{g}_1^4 \\ &+ \sum_{k=5}^{j} T_k \bar{g}_1^k \} + U_{1j} = S_j + U_{1j}, \end{aligned} \tag{2.28}$$

kur ar T_k apzīmē $\bar{g}_2^{-2(k-1)}$, kas pareizināts ar polinomu no $\bar{g}_2, \ldots, \bar{g}_k$ ar konstantiem koeficientiem un $U_{1j} = O_p\{(n^{-1/2} + h^r)^{j+1}\}$.

Dēļ jau iepriekš minētajām neprecizitātēm [8], jāpiezīmē, ka (2.28) atšķiras no [8] atrodamās S_j' izteiksmes.

5. Nākamais solis ir iegūt Edgeworth izvirzījumu $n^{1/2}S_j$ sadalījumam.

   Apzīmēsim $\mu_k = \mathbb{E}(\bar{g}_k)$ un $V_k = \bar{g}_k - \mu_k$, tad izmantojot [8] var iegūt, ka

$$S_j = p(V) + U_{2j},$$

   kur $U_{2j} = O_p(n^{-3})$.

   Chen un Hall [8] parādīja, ka nosakot $n^{1/2}p(V)$ kumulantus un pieņemot, ka $nh^r \to 0$, var iegūt formālu Edgeworth izvirzījumu $p(V)$ sadalījumam

$$\begin{aligned} \mathbb{P}\{n^{1/2}p(V) \le x\} = &\Phi(x) - \frac{1}{12n}\{6\mu_2^{-1}(n\mu_1)^2 + 3\mu_2^{-2}\mu_4 - 2\mu_2^{-3}\mu_3^2\}x\phi(x) \\ &+ (\text{pāra pakāpes polinoms no } x)\phi(x) + o(nh^{2r}) \\ &+ O(n^{-2}), \end{aligned} \tag{2.29}$$

   kur Φ un ϕ ir $N(0,1)$ sadalījuma un blīvuma funkcijas.

6. Gala rezultātā Chen un Hall [8] parāda, ka ja $nh^r \to 0$, tad

$$\mathbb{E}\{W(\theta)\} = 1 + n^{-1}\beta + o(n^{-1}),$$

   kur $\beta = \frac{1}{6}(3\mu_2^{-2}\mu_4 - 2\mu_2^{-3}\mu_3^2)$ un $\mu_j = \mathbb{E}\left[G\{(\theta - X_i)/h\} - q\right]^j$.

Tātad $W(\theta)$ un χ^2 sadalījumu vidējās vērtības atšķiras par $n^{-1}\beta$ un $W(\theta)/(1+n^{-1}\beta)$ sadalījums no χ^2_1 sadalījuma atšķirsies ar kārtu n^{-2}, t.i.,

$$P(\theta \in I_{d(c,\gamma)}) = \alpha + O(n^{-2}),$$

kur $\gamma = \beta$ vai $\gamma = \hat{\beta}$ un $d(c,\gamma) = c(1 + n^{-1}\gamma)$.

Ja β nav zināma, tad to novērtē no datiem,

$$\hat{\beta} = \frac{1}{6}\left(3\hat{\mu}_2^{-2}\hat{\mu}_4 - 2\hat{\mu}_2^{-3}\hat{\mu}_3^2\right),$$

kur

$$\hat{\mu}_j = n^{-1}\sum_i \left[G\left\{(\hat{\theta} - X_i)/h\right\} - q\right]^j$$

un $\hat{\theta}$ ir θ novērtējums.

2.6. Simulācijas Bārtleta korekcijai kvantilēm

Novērtēsim Bārtleta korekcijas praktisko ietekmi kvantiļu gadījumā. Lai salīdzinātu pārklājuma precizitāti ticamības intervāliem, kas konstruēti mediānai, izmantosim simulāciju tabulas no diplomdarba [22]. Simulācijās tika apskatīti 2 sadalījumi $N(0,1)$ un χ^2_3. Abiem sadalījumiem 10000 reižu tika ģenerēti dati pie izlašu apjomiem $n = 20$ un $n = 50$, un izmantojot EL (negludinātā EL metode), EL_{CH} (Chen un Hall gludinātā EL metode) un EL_{CH_B} (Chen un Hall gludinātā EL metode ar Bārtleta korekciju) tika konstruēti ticamības intervāli pie dažādiem nozīmības līmeņiem un noteikta ticamības intervālu pārklājuma precizitāte.

Kodolu gludināšanās metodēs joslas platuma h izvēlei ir ļoti liela nozīme, tāpēc, veicot simulācijas, tika izvēlētas vairākas h vērtības (n^{-1}, $n^{-3/4}$, $n^{-1/2}$ un $n^{-1/4}$). 4.tabulā ir redzami simulāciju rezultāti χ^2_3 sadalījuma mediānai pie nozīmības līmeņiem $\alpha = 0.05, 0.1$ un 0.01. Aplūkojot rezultātus, jāsecina, ka jau pie tik nelieliem izlašu apjomiem, apskatītās metodes strādā ļoti labi un ka ar gludinātās metodes palīdzību var iegūt labākus rezultātus kā ar negludināto. Bārtleta korekcijas uzlabojums ir samērā minimāls, taču rezultāti rāda, ka, palielinoties n, Bārtleta korekcijas uzlabojums samazinās un palielinoties h Bārtleta korekcijas uzlabojums palielinās. Izvērtējot kā mainās ticamības intervālu pārklājuma precizitāte, mainoties h, jāsecina, ka EL_{CH} un EL_{CH_B} minimāli ietekmē izvēlētā h amplitūda.

4. tabula Pārklājuma precizitāte χ^2_3 mediānai

			Ticamības līmenis		
			90%	95%	99%
$n = 20$	h	EL	0.8879	0.9609	0.9888
	n^{-1}	EL_{CH}	0.8901	0.9563	0.9887
		EL_{CH_B}	0.8922	0.9574	0.9890
	$n^{-3/4}$	EL_{CH}	0.8945	0.9524	0.9889
		EL_{CH_B}	0.8981	0.9541	0.9897
	$n^{-1/2}$	EL_{CH}	0.8956	0.9490	0.9895
		EL_{CH_B}	0.9000	0.9509	0.9902
	$n^{-1/4}$	EL_{CH}	0.8933	0.9447	0.9887
		EL_{CH_B}	0.8986	0.9476	0.9899
$n = 50$	h	EL	0.8798	0.9347	0.9930
	n^{-1}	EL_{CH}	0.8859	0.9425	0.9914
		EL_{CH_B}	0.8866	0.9437	0.9917
	$n^{-3/4}$	EL_{CH}	0.8927	0.9464	0.9906
		EL_{CH_B}	0.8942	0.9482	0.9912
	$n^{-1/2}$	EL_{CH}	0.8953	0.9490	0.9906
		EL_{CH_B}	0.8971	0.9493	0.9908
	$n^{-1/4}$	EL_{CH}	0.8926	0.9463	0.9893
		EL_{CH_B}	0.8943	0.9483	0.9903

Diplomdarbā [22] ir pieejami simulāciju rezultāti arī Zhou un Jing gludinātajai EL metodei, kas ir daudz jūtīgāka pret h izmaiņām.

5.tabulā ir redzami simulāciju rezultāti $N(0,1)$ sadalījuma mediānai pie nozīmības līmeņiem $\alpha = 0.05, 0.1$, un 0.01. Līdzīgi kā gadījumā ar asimetrisko χ^2_3 sadalījumu, jāsecina, ka arī standartnormālā sadalījuma gadījumā Chen un Hall gludinātā metode uzrāda labus rezultātus un Bārtleta korekcija tos vēl tikai uzlabo.

5. tabula Pārklājuma precizitāte $N(0, 1)$ mediānai

			Ticamības līmenis		
			90%	95%	99%
$n = 20$	h	EL	0.8837	0.9589	0.9884
	n^{-1}	EL_{CH}	0.8898	0.9498	0.9885
		EL_{CH_B}	0.8942	0.9509	0.9889
	$n^{-3/4}$	EL_{CH}	0.8951	0.9468	0.9883
		EL_{CH_B}	0.8996	0.9507	0.9897
	$n^{-1/2}$	EL_{CH}	0.8977	0.9465	0.9881
		EL_{CH_B}	0.9026	0.9505	0.9895
	$n^{-1/4}$	EL_{CH}	0.8940	0.9474	0.9883
		EL_{CH_B}	0.9007	0.9504	0.9891
$n = 50$	h	EL	0.8848	0.9361	0.9941
	n^{-1}	EL_{CH}	0.8959	0.9461	0.9905
		EL_{CH_B}	0.8973	0.9476	0.9910
	$n^{-3/4}$	EL_{CH}	0.8982	0.9492	0.9902
		EL_{CH_B}	0.9002	0.9500	0.9907
	$n^{-1/2}$	EL_{CH}	0.8988	0.9480	0.9903
		EL_{CH_B}	0.9005	0.9496	0.9905
	$n^{-1/4}$	EL_{CH}	0.8958	0.9484	0.9897
		EL_{CH_B}	0.8984	0.9503	0.9901

3. Empīriskās ticamības metode divu izlašu gadījumā

Definēsim empīrisko ticamības metodi divu izlašu gadījumā, izmantojot [28, 16] lietotos apzīmējumus.

Pieņemsim, ka $X_1, \ldots, X_{n_1}$ ir *i.i.d.* gadījuma lielumi ar nezināmu sadalījuma funkciju F_1 un $Y_1, \ldots, Y_{n_2}$ ir *i.i.d.* gadījuma lielumi ar nezināmu sadalījuma funkciju F_2. Interesējoties par funkciju $\Delta(t)$ (turpmāk$-\Delta$), kas definēta intervālā T, pieņemsim, ka θ_0 ir parametrs, kas saistīts ar vienu no sadalījuma funkcijām F_1 vai F_2. Tālāk mēs uzskatām, ka visa informācija par nezināmajiem īstajiem parametriem θ_0 un Δ_0 ir pieejama ar funkcijām $w_1(X_i, \theta_0, \Delta_0, t)$ un $w_2(Y_j, \theta_0, \Delta_0, t)$, kurām ir spēkā

$$\mathbb{E}_{F_1} w_1(X, \theta_0, \Delta_0, t) = 0, \tag{3.1}$$

$$\mathbb{E}_{F_2} w_2(Y, \theta_0, \Delta_0, t) = 0. \tag{3.2}$$

$\Delta_0 = \theta_1 - \theta_0$, kur θ_0 un θ_1 ir parametri, kas attiecīgi ir saistīti ar F_1 un F_2. Lai konstruētu ticamības intervālus funkcijai Δ fiksētam $t \in T$, definēsim empīriskās ticamības attiecību

$$R(\Delta, \theta) = \sup_{\theta, p, q} \prod_{i=1}^{n_1} (n_1 p_i) \prod_{j=1}^{n_2} (n_2 q_j), \tag{3.3}$$

kur $p = (p_1, \ldots, p_{n_1})$ un $q = (q_1, \ldots, q_{n_2})$ nosaka ierobežojumi

$$p_i \geq 0, \quad \sum_{i=1}^{n_1} p_i = 1, \quad \sum_{i=1}^{n_1} p_i w_1(X_i, \theta, \Delta, t) = 0, \tag{3.4}$$

$$q_j \geq 0, \quad \sum_{j=1}^{n_2} q_j = 1, \quad \sum_{j=1}^{n_2} q_j w_2(Y_j, \theta, \Delta, t) = 0. \tag{3.5}$$

Ja 0 pieder izliektajai čaulai, ko veido $w_1(X_i, \theta, \Delta, t)$ un izliektajai čaulai, ko veido $w_2(Y_j, \theta, \Delta, t)$, tad eksistē viens vienīgs (3.3) atrisinājums.

Izmantojot Lagranža reizinātāju metodi, var iegūt, ka

$$p_i = \frac{1}{n_1(1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t))}, \quad i = 1, \ldots, n_1. \tag{3.6}$$

un

$$q_j = \frac{1}{n_2(1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t))}, \quad j = 1, \ldots, n_2, \tag{3.7}$$

kur Lagranža reizinātājus $\lambda_1(\theta)$ un $\lambda_2(\theta)$ var noteikt no vienādojumiem

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \frac{w_1(X_i, \theta, \Delta, t)}{1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t)} = 0, \tag{3.8}$$

un

$$\frac{1}{n_2} \sum_{j=1}^{n_2} \frac{w_2(Y_j, \theta, \Delta, t)}{1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t)} = 0. \tag{3.9}$$

Definējam empīrisko logaritmisko ticamības attiecību

$$\begin{aligned} W(\Delta, \theta) = &- 2 \log R(\Delta, \theta) = 2 \sum_{i=1}^{n_1} \log(1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t)) \\ &+ 2 \sum_{j=1}^{n_2} \log(1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t)). \end{aligned} \tag{3.10}$$

Parametra θ novērtējumu $\hat{\theta} = \hat{\theta}(\Delta)$, kas maksimizē (3.3) fiksētam parametram Δ, var iegūt no vienādojuma

$$\frac{\partial W(\Delta,\theta)}{\partial\theta} = \sum_{i=1}^{n_1} \frac{\lambda_1(\theta)\alpha_1(X_i,\theta,\Delta,t)}{1+\lambda_1(\theta)w_1(X_i,\theta,\Delta,t)} + \sum_{j=1}^{n_2} \frac{\lambda_2(\theta)\alpha_2(Y_j,\theta,\Delta,t)}{1+\lambda_2(\theta)w_2(Y_j,\theta,\Delta,t)} = 0, \quad (3.11)$$

kur

$$\alpha_1(X_i,\theta,\Delta,t) = \frac{\partial w_1(X_i,\theta,\Delta,t)}{\partial\theta} \;\text{ un }\; \alpha_2(Y_i,\theta,\Delta,t) = \frac{\partial w_2(Y_i,\theta,\Delta,t)}{\partial\theta}.$$

Teorēma 1. *[16] Izpildoties standarta gluduma nosacījumiem funkcijām w_1, w_2, α_1, α_2, vienādojumam* (3.11) *eksistē atrisinājums $\hat{\theta}$ tāds, ka $\hat{\theta}$ ir parametra θ_0 būtisks novērtējums un $R(\Delta,\theta)$ sasniedz lokālo maksimumu punktā $\hat{\theta}$ un*

$$\sqrt{n}(\hat{\theta}-\theta_0) \to_d N\left(0, \frac{\beta_1\beta_2}{\beta_2\beta_{10}^2 + k\beta_1\beta_{20}^2}\right), \quad (3.12)$$

kur $k < \infty$ ir pozitīva konstante tāda, ka $n_2/n_1 \to k$, kad $n_1, n_2 \to \infty$,

$$-2\log R(\Delta_0,\hat{\theta}) \to_d \chi_1^2, \quad (3.13)$$

kad $n_1, n_2 \to \infty$, katram fiksētam $t \in T$,

$$\beta_1 = \mathbb{E}_{F_1} w_1^2(X,\theta_0,\Delta_0,t), \quad \beta_2 = \mathbb{E}_{F_2} w_2^2(Y,\theta_0,\Delta_0,t),$$

$$\beta_{10} = \mathbb{E}_{F_1} \alpha_1(X,\theta_0,\Delta_0,t), \quad \beta_{20} = \mathbb{E}_{F_2} \alpha_2(Y,\theta_0,\Delta_0,t).$$

Ticamības intervāli, kas konstruēti balstoties uz EL metodi, parametram Δ katram fiksētam $t \in T$ ir formā $\Delta : \{R(\Delta,\hat{\theta}) > c\}$, kur $\hat{\theta}$ ir vienādojuma (3.11) atrisinājums. Konstanti c var noteikt izmantojot Teorēmu 1.

3.1. Bārtleta korekcija vidējo vērtību starpībai

Arī divu izlašu gadījumā Bārtleta korekcija empīriskās ticamības metodei vispirms tika parādīta vidējās vērtības problemātikai, proti, vidējo vērtību starpībai. Pirmais to parādīja Jing [13], taču vēlāk Liu, Zou un Zhang [14] parādīja, ka [13] piedāvātajā Bārtleta korekcijā ir pieļautas kļūdas, un piedāvāja savu pierādījumu.

Vēloties vispārināt Bārtleta korekciju 2 izlašu problemātikai, sākotnēji tika plānots balstīties tieši uz nesen parādīto Liu, Zou un Zhang [14] rezultātu.

Lai sasniegtu vēlamo mērķi, sākumā bija rūpīgi jāizpēta Liu, Zou un Zhang [14] rezultāts, taču, tā kā šajā publikācijā bija izlaists ļoti daudz detaļu, tad, lai labāk saprastu pierādījuma tehniku, nācās rekonstruēt pierādījumu, kā rezultātā [14] tika konstatētas neprecizitātes, kuras tagad centīsimies novērst.

Uzskatāmības dēļ, pamatā izmantosim apzīmējumus no [14]. Jāpiezīmē arī, ka [14] tika apskatīti $d-$dimensionāli gadījuma lielumi, taču, lai vienkāršotu pierakstu, apskatīsim gadījumu, kad $d = 1$.

Pieņemsim, ka $X_1, X_2, \ldots, X_{n_1}$ ir $i.i.d.$ gadījuma lielumi ar nezināmu sadalījuma funkciju F un $X_{n_1+1}, X_{n_1+2}, \ldots, X_{n_1+n_2}$ ir $i.i.d.$ gadījuma lielumi ar nezināmu sadalījuma funkciju G. Vēlamies pārbaudīt hipotēzi

$$H_0 : \Delta = \int XdG(x) - \int XdF(x) = \Delta_0 \text{ pret alternatīvu } H_1 : \Delta \neq \Delta_0. \tag{3.14}$$

Empīriskā ticamība parametram Δ ir formā

$$L(\Delta_0) = \sup\{\prod_i p_i \prod_j q_j | \sum_i p_i(X_i - \mu_0) = \sum_j q_j(Y_j - \mu_0) = 0\},$$

kur $Y_j = X_j - \Delta_0$, $\Delta_0 = \mu_1 - \mu_0$, μ_0 ir pirmās izlases vidējā vērtība un μ_1 ir otrās izlases vidējā vērtība.

Atbilstošā empīriskā logaritmiskā ticamības attiecība ir $l(\Delta_0) = -2\ln L(\Delta_0)/L(\hat{\Delta})$, kur $L(\hat{\Delta}) = n_1^{-n_1} n_2^{-n_2}$. Izmantojot Lagranža metodi, var iegūt, ka

$$p_i = \frac{1}{n_1(1 + \frac{n}{n_1}\lambda_1(X_i - \mu))},\ q_j = \frac{1}{n_2(1 + \frac{n}{n_2}\lambda_2(Y_j - \mu))}, \tag{3.15}$$

tādējādi empīrisko logaritmisko ticamības attiecību var pārrakstīt

$$l(\lambda_1, \lambda_2, \mu) = 2\left[\sum_i \ln\{1 + \frac{n}{n_1}\lambda_1(X_i - \mu)\} + \sum_j \ln\{1 + \frac{n}{n_2}\lambda_2(Y_j - \mu)\}\right]. \tag{3.16}$$

Atrisinot sistēmu

$$\begin{cases} \dfrac{1}{n_1}\displaystyle\sum_i \dfrac{X_i-\mu}{1+\frac{n}{n_1}\lambda_1(X_i-\mu)} = 0 \\ \dfrac{1}{n_2}\displaystyle\sum_j \dfrac{Y_j-\mu}{1+\frac{n}{n_2}\lambda_2(Y_j-\mu)} = 0 \\ \lambda_1+\lambda_2=0, \end{cases} \tag{3.17}$$

varam iegūt atrisinājumu $(\lambda_1^*, \lambda_2^*, \mu^*)$.

Piezīme 2. *Lai arī izskatās, ka Liu, Zou un Zhang [14] aprakstītā problēmas nostādne atšķiras no vispārīgās, kas tika aprakstīta iepriekšējā nodaļā, taču īstenībā tās ir ekvivalentas gadījumā, kad* $w_1(X_i, \theta, \Delta, t) = X_i - \mu$ *un* $w_2(Y_j, \theta, \Delta, t) = Y_j - \mu$.

Tātad, ja $w_1(X_i, \theta, \Delta, t) = X_i-\mu$ *un* $w_2(Y_j, \theta, \Delta, t) = Y_j-\mu$, *tad tā kā* $\alpha_1(X_i, \theta, \Delta, t) = \alpha_2(y_j, \theta, \Delta, t) = -1$, *tad* (3.11) *var pārrakstīt*

$$-\sum_{i=1}^{n_1}\frac{\lambda_1(\theta)}{1+\lambda_1 w_1(X_i,\theta,\Delta,t)} - \sum_{j=1}^{n_2}\frac{\lambda_2(\theta)}{1+\lambda_2 w_2(Y_j,\theta,\Delta,t)} = 0.$$

Izmantojot (3.4), (3.5), (3.6) *un* (3.7), *iegūstam*

$$-\lambda_1(\theta)n_1 - \lambda_2(\theta)n_2 = 0 \Rightarrow \frac{n_1}{n}\lambda_1(\theta) + \frac{n_2}{n}\lambda_2(\theta) = 0.$$

Apzīmējot $\tilde{\lambda}_1(\theta) = (n_1/n)\lambda_1$ *un* $\tilde{\lambda}_2(\theta) = (n_2/n)\lambda_2$, *iegūstam sistēmas* (3.17) *3. vienādojumu. Aizstājot izteiksmēs* (3.4), (3.5), (3.6) *un* (3.7), $\lambda_1(\theta)$ *ar* $n\tilde{\lambda}_1(\theta)/n_1$ *un* $\lambda_2(\theta)$ *ar* $n\tilde{\lambda}_2(\theta)/n_2$, *varam iegūt arī sistēmas* (3.17) *pirmos divus vienādojumus, tādējādi parādot, ka Liu, Zou un Zhang [14] problēmas nostādne neatšķiras no vispārīgās, kas definēta 3.nodaļā.*

Ievedīsim apzīmējumus $v_1 = n\mathbb{D}(x)/n_1$, $v_2 = n\mathbb{D}(y)/n_2$, $v = v_1 + v_2$ un $w = v_1 v^{-1} v_2$. Ar μ_k un ξ_k apzīmēsim attiecīgi μ^* un λ^* aproksimācijas līdz kārtai $O(n^{(k+1)/2}), k = \overline{1,3}$. Piezīmēsim, ka gala rezultātā μ_k, ξ_k un to izteiksmēm ir būtiska nozīme, lai iegūtu empīriskās logaritmiskās ticamības attiecības izvirzījumu ar noteiktu kārtu.

Izteiksmes $n^{-1}\sum_{i=1}^n w_i(1+\bar{\lambda}w_i)^{-1} = 0$ Teilora rindas izvirzījums ir

$$0 = n^{-1}\sum_{i=1}^{n} w_i\{1 - \bar{\lambda}w_i + (\bar{\lambda}w_i)^2 - (\bar{\lambda}w_i)^3 + \ldots\}.$$

Šajā gadījumā $w_i = X_i - \mu$ $(w_j = Y_j - \mu)$ un $\bar{\lambda} = n\lambda_1/n_1$ $(\bar{\lambda} = n\lambda_2/n_2)$

$$\frac{1}{n_1}\sum_i (X_i-\mu^*) - \frac{n}{n_1^2}\sum_i \lambda_1^*(X_i-\mu^*)^2 + \frac{n^2}{n_1^3}\sum_i \lambda_1^{*2}(X_i-\mu^*)^3$$
$$-\frac{n^3}{n_1^4}\sum_i \lambda_1^{*3}(X_i-\mu^*)^4 + \ldots = 0$$

1. Apskatām Teilora rindas izvirzījuma pirmos divus saskaitāmos

$$\frac{1}{n_1}\sum_i (X_i - \mu^*) - \frac{n}{n_1^2}\sum_i \lambda_1^*(X_i-\mu^*)^2 + O_p(n^{-1}) = 0.$$

Ievedam apzīmējumus

$$C_{11} = \frac{1}{n_1}\sum_i (X_i - \mu_0),\ C_{12} = \frac{1}{n_2}\sum_j (Y_j - \mu_0),$$

un veicam algebriskus pārveidojumus

$$\frac{1}{n_1}\sum_i (X_i - \mu^*) - \left(\frac{n}{n_1^2}\sum_i (X_i-\mu^*)^2 - v_1\right)\lambda_1^* - v_1\lambda_1^* + O_p(n^{-1}) = 0,$$
$$\frac{1}{n_1}\sum_i (X_i - \mu^*) - v_1\lambda_1^* + O_p(n^{-1/2}) = 0,$$
$$\frac{1}{n_1}\sum_i (X_i - \mu_0 + \mu_0 - \mu^*) - v_1\lambda_1^* + O_p(n^{-1/2}) = 0,$$
$$\frac{1}{n_1}\sum_i (X_i - \mu_0) + \mu_0 - \mu^* - v_1\lambda_1^* + O_p(n^{-1/2}) = 0,$$
$$C_{11} + \mu_0 - \mu^* - v_1\lambda_1^* + O_p(n^{-1/2}) = 0 \mid \cdot\ v_1^{-1},$$
$$\lambda_1^* = v_1^{-1}(C_{11} + \mu_0 - \mu^*) + O_p(n^{-1/2}). \tag{3.18}$$

Līdzīgi iegūst, ka

$$\lambda_2^* = v_2^{-1}(C_{12} + \mu_0 - \mu^*) + O_p(n^{-1/2}). \tag{3.19}$$

Iegūtās izteiksmes (3.18) un (3.19) atšķiras no tām, kas atrodamas [14]. Liu, Zou un Zhang [14] noteikuši, ka $\lambda_1^* = v_1^{-1}(C_{11} - \mu^*) + O_p(n^{-1/2})$ un $\lambda_2^* = v_2^{-1}(C_{12} - \mu^*) + O_p(n^{-1/2})$, taču, ņemot vērā ievestos apzīmējumus un parādītos algebriskos pārveidojumus, jāsecina, ka [14] pieļautas neprecizitātes.

Jāpiezīmē, ka arī tālākajos soļos iegūtās λ_1^*, λ_2^* un μ^* aproksimācijas atšķiras no [14] atrodamajām.

Ievietojam iegūtās λ_1^*, λ_2^* izteiksmes (3.18) un (3.19) sistēmas (3.17) 3.vienādojumā

$$\lambda_1^* + \lambda_2^* = 0,$$
$$v_1^{-1}(C_{11} + \mu_0 - \mu^*) + v_2^{-1}(C_{12} + \mu_0 - \mu^*) + O_p(n^{-1/2}) = 0,$$
$$v_1^{-1}C_{11} + v_1^{-1}(\mu_0 - \mu^*) + v_2^{-1}C_{12} + v_2^{-1}(\mu_0 - \mu^*) + O_p(n^{-1/2}) = 0,$$
$$v_1^{-1}C_{11} + v_2^{-1}C_{12} + (\mu_0 - \mu^*)\underbrace{(v_1^{-1} + v_2^{-1})}_{w^{-1}} + O_p(n^{-1/2}) = 0,$$
$$v_1^{-1}C_{11} + v_2^{-1}C_{12} + (\mu_0 - \mu^*)w^{-1} + O_p(n^{-1/2}) = 0 \mid \cdot\ w,$$
$$\mu^* - \mu_0 = w(v_1^{-1}C_{11} + v_2^{-1}C_{12}) + O_p(n^{-1/2}). \tag{3.20}$$

(3.20) var pārrakstīt arī formā

$$\mu_1 - \mu_0 = w(v_1^{-1}C_{11} + v_2^{-1}C_{12}).$$

Izmantojot to, ka $\lambda_1^* = -\lambda_2^*$ no sistēmas (3.17) 3.vienādojuma, varam iegūt

$$\lambda_1^* = -v_2^{-1}(C_{12} + \mu_0 - \mu^*) + O_p(n^{-1/2}),$$
$$\lambda_1^* = -v_2^{-1}C_{12} - v_2^{-1}(\mu_0 - \mu^*) + O_p(n^{-1/2}).$$

Ievietojam iegūtajā izteiksmē (3.20)

$$\lambda_1^* = -v_2^{-1}C_{12} - v_2^{-1}(-w(v_1^{-1}C_{11} + v_2^{-1}C_{12}) + O_p(n^{-1/2})) + O_p(n^{-1/2}),$$
$$\lambda_1^* = -v_2^{-1}C_{12} + \underbrace{v_2^{-1}wv_1^{-1}}_{v^{-1}}C_{11} + v_2^{-1}wv_2^{-1}C_{12} + O_p(n^{-1/2}),$$
$$\lambda_1^* = \underbrace{(v_2^{-1}wv_2^{-1} - v_2^{-1})}_{v^{-1}}C_{12} + v^{-1}C_{11} + O_p(n^{-1/2}),$$
$$\lambda_1^* = v^{-1}(C_{11} - C_{12}) + O_p(n^{-1/2}). \tag{3.21}$$

(3.21) var pārrakstīt arī formā

$$\xi_1 = v^{-1}(C_{11} - C_{12}) + O_p(n^{-1/2}) \equiv v^{-1}D_1 + O_p(n^{-1/2}).$$

2. Apskatām Teilora rindas izvirzījuma pirmos trīs saskaitāmos

$$\frac{1}{n_1}\sum_i (X_i - \mu^*) - \frac{n}{n_1^2}\sum_i \lambda_1^*(X_i - \mu^*)^2 + \frac{n^2}{n_1^3}\sum_i \lambda_1^{*2}(X_i - \mu^*)^3 + O_p(n^{-3/2}) = 0,$$
$$\frac{1}{n_1}\sum_i (X_i - \mu^*) - \left(\frac{n}{n_1^2}\sum_i (X_i - \mu^*)^2 - v_1\right)\lambda_1^* - v_1\lambda_1^* + \frac{n^2}{n_1^3}\sum_i \lambda_1^{*2}(X_i - \mu^*)^3$$
$$+ O_p(n^{-3/2}) = 0. \tag{3.22}$$

Ievedot apzīmējumus

$$C_{21} = -\left[\frac{n}{n_1^2}\sum_i (X_i - \mu_0)^2 - v_1\right]\xi_1,\ C_{31} = \frac{n^2}{n_1^3}\sum_i \xi_1^2(X_i - \mu_0)^3$$

un

$$C_{22} = \left[\frac{n}{n_2^2}\sum_j (Y_j - \mu_0)^2 - v_2\right]\xi_1,\ C_{32} = \frac{n^2}{n_2^3}\sum_j \xi_1^2(Y_j - \mu_0)^3,$$

izteiksmi (3.22) var pārrakstīt formā

$$C_{11} + \mu_0 - \mu^* + C_{21} - v_1\lambda_1^* + C_{31} + O_p(n^{-3/2}) = 0.$$

Pareizinot izteiksmi (3.22) ar v_1^{-1}, var iegūt

$$\lambda_1^* = v_1^{-1}(C_{11} + C_{21} + C_{31} + \mu_0 - \mu^*) + O_p(n^{-3/2})$$

un līdzīgi var iegūt

$$\lambda_2^* = v_2^{-1}(C_{12} + C_{22} + C_{32} + \mu_0 - \mu^*) + O_p(n^{-3/2}).$$

Ievietojot abas šīs izteiksmes sistēmas (3.17) 3.vienādojumā, iegūstam

$$\mu^* - \mu_0 = w[v_1^{-1}(C_{11} + C_{21} + C_{31}) + v_2^{-1}(C_{12} + C_{22} + C_{32})] + O_p(n^{-3/2}),$$

ko var pārrakstīt formā

$$\mu_2 - \mu_0 = w[v_1^{-1}(C_{11} + C_{21} + C_{31}) + v_2^{-1}(C_{12} + C_{22} + C_{32})],$$

kā arī

$$\begin{aligned}\lambda_1^* &= v^{-1}[(C_{11} - C_{12}) + (C_{21} - C_{22}) + (C_{31} - C_{32})] + O_p(n^{-3/2}) \\ &\equiv v^{-1}(D_1 + D_2 + D_3) + O_p(n^{-3/2}),\end{aligned} \tag{3.23}$$

ko var pārrakstīt formā

$$\xi_2 = v^{-1}(D_1 + D_2 + D_3) + O_p(n^{-3/2}).$$

3. Apskatot Teilora rindas izvirzījuma pirmos četrus saskaitāmos

$$\begin{aligned}&\frac{1}{n_1}\sum_i (X_i - \mu^*) - \frac{n}{n_1^2}\sum_i \lambda_1^*(X_i - \mu^*)^2 + \frac{n^2}{n_1^3}\sum_i \lambda_1^{*2}(X_i - \mu^*)^3 \\ &- \frac{n^3}{n_1^4}\sum_i \lambda_1^{*3}(X_i - \mu^*)^4 + O_p(n^{-2}) = 0,\end{aligned}$$

līdzīgā veidā kā iepriekšējā solī var iegūt, ka

$$\lambda_1^* = v_1^{-1}(C_{11} + C_{21}^* + C_{31}^* + C_{41} + \mu_0 - \mu^*) + O_p(n^{-2})$$

un

$$\lambda_2^* = v_2^{-1}(C_{12} + C_{22}^* + C_{32}^* + C_{42} + \mu_0 - \mu^*) + O_p(n^{-2}),$$

kur

$$\begin{aligned}C_{41} &= -\frac{n^3}{n_1^4}\sum_i \xi_1^3(X_i - \mu_0)^4,\ C_{42} = \frac{n^3}{n_2^4}\sum_j \xi_1^3(Y_j - \mu_0)^4, \\ C_{21}^* &= -\left[\frac{n}{n_1^2}\sum_i (X_i - \mu_1)^2 - v_1\right]\xi_2,\ C_{31}^* = \frac{n^2}{n_1^3}\sum_i \xi_2^2(X_i - \mu_1)^3, \\ C_{22}^* &= \left[\frac{n}{n_2^2}\sum_j (Y_j - \mu_1)^2 - v_2\right]\xi_2 \text{ un } C_{32}^* = \frac{n^2}{n_2^3}\sum_j \xi_2^2(Y_j - \mu_1)^3.\end{aligned}$$

Darbojoties tāpat kā iepriekšējos soļos un izmantojot sistēmas (3.17) 3.vienādojumu, iegūstam

$$\mu^* - \mu_0 = w[v_1^{-1}(C_{11} + C_{21}^* + C_{31}^* + C_{41}) + v_2^{-1}(C_{12} + C_{22}^* + C_{32}^* + C_{42})] + O_p(n^{-2}),$$

ko var pārrakstīt formā

$$\mu_3 - \mu_0 = w[v_1^{-1}(C_{11} + C_{21}^* + C_{31}^* + C_{41}) + v_2^{-1}(C_{12} + C_{22}^* + C_{32}^* + C_{42})],$$

un

$$\begin{aligned}\lambda_1^* &= v^{-1}[(C_{11} - C_{12}) + (C_{21}^* - C_{22}^*) + (C_{31}^* - C_{32}^*) + (C_{41} - C_{42})] + O_p(n^{-2}) \\ &\equiv v^{-1}(D_1 + D_2^* + D_3^* + D_4) + O_p(n^{-2}),\end{aligned}$$

ko savukārt var pārrakstīt

$$\xi_3 = v^{-1}(D_1 + D_2^* + D_3^* + D_4) + O_p(n^{-2}).$$

4. Izvirzot $l(\lambda_1^*, \lambda_2^*, \mu^*)$ Teilora rindā, iegūstam

$$\begin{aligned}l(\lambda_1^*, \lambda_2^*, \mu^*) = &\frac{2n}{n_1}\sum_i \lambda_1^*(X_i - \mu^*) + \frac{2n}{n_2}\sum_j \lambda_2^*(Y_j - \mu^*) \\ &- \frac{n^2}{n_1^2}\sum_i \lambda_1^{*2}(X_i - \mu^*)^2 - \frac{n^2}{n_2^2}\sum_j \lambda_2^{*2}(Y_j - \mu^*)^2 \\ &+ \frac{4n^3}{3n_1^3}\sum_i \lambda_1^{*3}(X_i - \mu^*)^3 + \frac{4n^3}{3n_2^3}\sum_j \lambda_2^{*3}(Y_j - \mu^*)^3 \\ &- \frac{n^4}{2n_1^4}\sum_i \lambda_1^{*4}(X_i - \mu^*)^4 - \frac{n^4}{2n_2^4}\sum_j \lambda_2^{*4}(Y_j - \mu^*)^4 + O_p(n^{-3/2}).\end{aligned}$$

Un pareizinot $l(\lambda_1^*, \lambda_2^*, \mu^*)$ ar n^{-1} un aizstājot μ^* un λ_1^* ar to aproksimācijām, iegūstam

$$\begin{aligned}l(\lambda_1^*, \lambda_2^*, \mu^*)n^{-1} = &\frac{2}{n_1}\sum_i \xi_3(X_i - \mu_3) - \frac{2}{n_2}\sum_j \xi_3(Y_j - \mu_3) \\ &- \frac{n}{n_1^2}\sum_i \xi_3^2(X_i - \mu_1)^2 - \frac{n}{n_2^2}\sum_j \xi_3^2(Y_j - \mu_1)^2 \\ &+ \frac{2n^2}{3n_1^3}\sum_i \xi_2^3(X_i - \mu_1)^3 - \frac{2n^2}{3n_2^3}\sum_j \xi_2^3(Y_j - \mu_1)^3 \\ &- \frac{n^3}{2n_1^4}\sum_i \xi_1^4(X_i - \mu_0)^4 - \frac{n^3}{2n_2^4}\sum_j \xi_1^4(Y_j - \mu_0)^4 \\ &+ O_p(n^{-5/2}) \\ = &\, 2I_1 - I_2 + \frac{2}{3}I_3 - \frac{1}{2}I_4 + O_p(n^{-5/2}).\end{aligned}$$

No sistēmas (3.17) pirmajiem diviem vienādojumiem seko, ka

$$C_{11} - \mu_3 + \mu_0 - \left(\frac{n}{n_1^2} \sum_i (X_i - \mu_1)^2 - v_1 \right) \xi_3 - v_1 \xi_3 + C_{31}^* + C_{41}$$
$$+ O_p(n^{-2}) = 0 \text{ un}$$
$$C_{12} + \mu_3 + \mu_0 - \left(\frac{n}{n_2^2} \sum_j (Y_j - \mu_1)^2 + v_2 \right) \xi_3 - v_2 \xi_3 + C_{32}^* + C_{42}$$
$$+ O_p(n^{-2}) = 0$$

jeb

$$\frac{n}{n_1^2} \sum_i \xi_3 (X_i - \mu_1)^2 = C_{11} - \mu_3 + \mu_0 + C_{31}^* + C_{41} + O_p(n^{-2}), \tag{3.24}$$
$$-\frac{n}{n_2^2} \sum_j \xi_3 (Y_j - \mu_1)^2 = C_{12} - \mu_3 + \mu_0 + C_{32}^* + C_{42} + O_p(n^{-2}). \tag{3.25}$$

Ievietojot (3.24) un (3.25) izteiksmēs I_1, I_2, I_3 un I_4, iegūstam

$$\begin{aligned} I_1 &= \xi_3(C_{11} - C_{12}) = v^{-1}(D_1 + D_2^* + D_3^* + D_4)D_1, \\ I_2 &= \xi_3(C_{11} - C_{12} + C_{31}^* - C_{32}^* + C_{41} - C_{42}) \\ &= (D_1 + D_3^* + D_4)v^{-1}(D_1 + D_2^* + D_3^* + D_4), \\ I_3 &= \xi_2(C_{31}^* - C_{32}^*) = v^{-1}(D_1 + D_2 + D_3)D_3^*, \\ I_4 &= -\xi_1(C_{41} - C_{42}) = -v^{-1}D_1 D_4. \end{aligned}$$

Ievietojot $l(\lambda_1^*, \lambda_2^*, \mu^*)n^{-1}$ izteiksmes I_1, I_2, I_3 un I_4 un izmantojot, ka

$$D_1 = O_p(n^{-1/2}),\ D_2 = O_p(n^{-1}),\ D_2^* = O_p(n^{-1}),\ D_2^* - D_2 = O_p(n^{-3/2}),$$
$$D_3 = O_p(n^{-1}),\ D_3^* = O_p(n^{-1}),\ D_3^* - D_3 = O_p(n^{-3/2}),\ D_4 = O_p(n^{-3/2}),$$

kā arī atstājot tikai locekļus līdz kārtai $n^{-5/2}$, iegūstam

$$\begin{aligned} l(\lambda_1^*, \lambda_2^*, \mu^*)n^{-1} &= v^{-1}D_1^2 + v^{-1}D_2^* D_1 - \frac{1}{3}D_3 v^{-1} D_2 - \frac{1}{3} v^{-1} D_3^2 \\ &\quad + \frac{2}{3} D_1 v^{-1} D_3^* + \frac{1}{2} D_1 v^{-1} D_4 + O_p(n^{-5/2}). \end{aligned} \tag{3.26}$$

Tas, ka, piemēram, $D_1 = O_p(n^{-1/2})$, seko no centrālās robežteorēmas, jo

$$\begin{aligned} D_1 = C_{11} - C_{12} &= \frac{1}{n_1} \sum_i (X_i - \mu_0) - \frac{1}{n_2} \sum_j (Y_j - \mu_0) \\ &= \left(\frac{1}{n_1} \sum_i X_i - \mu_0 \right) - \left(\frac{1}{n_2} \sum_j Y_j - \mu_0 \right) \\ &= O_p(n^{-1/2}) - O_p(n^{-1/2}) = O_p(n^{-1/2}). \end{aligned}$$

Līdzīgā veidā var iegūt arī pārējās izteiksmes.

5. Turpmākajiem aprēķiniem ieviesīsim sekojošus apzīmējumus

$$\begin{aligned} &z_{i0} = v^{-1/2}(X_i - \mu_0), z_{j0} = v^{-1/2}(Y_j - \mu_0), z_{i1} = v^{-1/2}(X_i - \mu_1), \\ &z_{j1} = v^{-1/2}(Y_j - \mu_1), \alpha_l = \left(\frac{n}{n_1}\right)^{l-1} \mathbb{E}z_{i0}^l + (-1)^l \left(\frac{n}{n_2}\right)^{l-1} \mathbb{E}z_{j0}^l, \\ &A_{0l} = \left(\frac{n^{l-1}}{n_1^l}\right) \sum_i z_{i0}^l + (-1)^l \left(\frac{n^{l-1}}{n_2^l}\right) \sum_j z_{j0}^l - \alpha_l, \\ &A_{1l} = \left(\frac{n^{l-1}}{n_1^l}\right) \sum_i z_{i1}^l + (-1)^l \left(\frac{n^{l-1}}{n_2^l}\right) \sum_j z_{j1}^l - \alpha_l. \end{aligned} \tag{3.27}$$

Izmantojot tikko ieviestos apzīmējumus, aprēķināsim visus (3.26) saskaitāmos un salīdzināsim ar rezultātiem, kas atrodami [14].

Sāksim ar $v^{-1}D_1^2$,

$$v^{-1}D_1^2 = v^{-1}(C_{11} - C_{12})^2 = v^{-1}\left(\frac{1}{n_1}\sum_i (X_i - \mu_0) - \frac{1}{n_2}\sum_j (Y_j - \mu_0)\right)^2,$$

ko izmantojot (3.27) var pārrakstīt formā

$$v^{-1}D_1^2 = v(A_{01} + \alpha_1)^2$$

un tā kā $\alpha_1 = 0$, jo $\mathbb{E}(X_i - \mu_0) = \mathbb{E}(Y_j - \mu_0) = 0$, tad

$$v^{-1}D_1^2 = v^{-1}vA_{01}^2 = A_{01}^2. \tag{3.28}$$

Tagad apskatīsim $v^{-1}D_2^*D_1$, ko izmantojot iegūto $v^{-1}D_1^2$ izteiksmi, var pārrakstīt

$$v^{-1}D_2^*D_1 = v^{-1/2}D_1 v^{-1/2}D_2^* = A_{01}v^{-1/2}D_2^*, \tag{3.29}$$

tāpēc jāatrod tikai $v^{-1/2}D_2^*$.

$$\begin{aligned} v^{-1/2}D_2^* =& v^{-1/2}(C_{21}^* - C_{22}^*) = -v^{-1/2}\left[\frac{n}{n_1^2}\sum_i (X_i - \mu_1)^2 - v_1\right]\xi_2 \\ & - v^{-1/2}\left[\frac{n}{n_2^2}\sum_j (Y_j - \mu_1)^2 - v_2\right]\xi_2 \\ =& - v^{-1/2}\xi_2\left[\frac{n}{n_1^2}\sum_i (X_i - \mu_1)^2 + \frac{n}{n_2^2}\sum_j (Y_j - \mu_1)^2 - (v_1 + v_2)\right], \end{aligned}$$

ko izmantojot (3.27) un to, ka $\xi_2 = v^{-1}(D_1 + D_2 + D_3)$, var pārrakstīt

$$\begin{aligned} v^{-1/2}D_2^* &= -v^{-1/2}\xi_2 v A_{12} = -v^{-1/2}v^{-1}(D_1 + D_2 + D_3)vA_{12} \\ &= -v^{-1/2}(D_1 + D_2 + D_3)A_{12}. \end{aligned} \tag{3.30}$$

Iepriekš jau ieguvām, ka $D_1 = v^{1/2}A_{01}$, tāpēc jāatrod vēl D_2 un D_3.

$$\begin{aligned} D_2 &= C_{21} - C_{22} \\ &= -\left[\frac{n}{n_1^2}\sum_i (X_i - \mu_0)^2 + \frac{n}{n_2^2}\sum_j (Y_j - \mu_0)^2 - (v_1 + v_2)\right]\xi_1 \\ &= -\xi_1 v A_{02}, \end{aligned}$$

Tā kā $\xi_1 = v^{-1}D_1 = v^{-1/2}A_{01}$, tad $D_2 = -v^{-1/2}A_{01}vA_{02} = -v^{1/2}A_{01}A_{02}$.

$$\begin{aligned} D_3 = C_{31} - C_{32} &= \frac{n^2}{n_1^3}\sum_i \xi_1^2 (X_i - \mu_0)^3 - \frac{n^2}{n_2^3}\sum_j \xi_1^2 (Y_j - \mu_0)^3 \\ &= \xi_1^2\left(\frac{n^2}{n_1^3}\sum_i (X_i - \mu_0)^3 - \frac{n^2}{n_2^3}\sum_j (Y_j - \mu_0)^3\right) \\ &= \xi_1^2 (A_{03} + \alpha_3) v^{3/2} = (v^{-1/2}A_{01})^2 (A_{03} + \alpha_3) v^{3/2} \\ &= v^{1/2}A_{01}^2 (A_{03} + \alpha_3). \end{aligned}$$

Tātad (3.30) var pārrakstīt

$$\begin{aligned} v^{-1/2}D_2^* &= -v^{-1/2}(v^{1/2}A_{01} - v^{1/2}A_{01}A_{02} + v^{1/2}A_{01}^2(A_{03} + \alpha_3))A_{12} \\ &= (-A_{01} + A_{01}A_{02} - A_{01}^2(A_{03} + \alpha_3))A_{12} \\ &= -A_{01}A_{12} + A_{01}A_{02}A_{12} - A_{01}^2(A_{03} + \alpha_3)A_{12} \\ &= -A_{01}A_{12} + A_{01}A_{02}A_{12} - A_{01}^2\alpha_3 A_{12} + O_p(n^{-2}). \end{aligned} \tag{3.31}$$

Ievietojot (3.31) izteiksmē (3.29), iegūstam

$$\begin{aligned} v^{-1}D_2^* D_1 &= A_{01}v^{-1/2}D_2^* \\ &= A_{01}(-A_{01}A_{12} + A_{01}A_{02}A_{12} - A_{01}^2\alpha_3 A_{12} + O_p(n^{-2})) \\ &= -A_{01}^2 A_{12} + A_{01}^2 A_{02}A_{12} - A_{01}^3\alpha_3 A_{12} + O_p(n^{-5/2}). \end{aligned} \tag{3.32}$$

Lai aprēķinātu $D_1 v^{-1} D_3^*$, jāatrod $v^{-1/2} D_3^*$.

$$\begin{aligned}
v^{-1/2} D_3^* &= v^{-1/2}(C_{31}^* - C_{32}^*) = \frac{n^2}{n_1^3}\sum_i \xi_2^2 (x_i - \mu_1)^3 - \frac{n^2}{n_2^3}\sum_j \xi_2^2 (y_j - \mu_1)^3 \\
&= v^{-1/2}\xi_2^2 \left(\frac{n^2}{n_1^3}\sum_i (x_i - \mu_1)^3 - \frac{n^2}{n_2^3}\sum_j (y_j - \mu_1)^3 \right) \\
&= v^{-1/2}\xi_2^2 (A_{13} + \alpha_3) v^{3/2} = \xi_2^2 (A_{13} + \alpha_3) v. \qquad (3.33)
\end{aligned}$$

Pirms (3.33) noteikšanas aprēķināsim ξ_2^2.

$$\begin{aligned}
\xi_2^2 &= (v^{-1}(D_1 + D_2 + D_3))^2 = v^{-2}(v^{1/2}A_{01} - v^{1/2}A_{01}A_{02} + D_3)^2 \\
&= v^{-2}(v^{1/2}A_{01} - v^{1/2}A_{01}A_{02} + v^{1/2}A_{01}^2(A_{03} + \alpha_3))^2 \\
&= v^{-1}(A_{01} - A_{01}A_{02} + A_{01}^2(A_{03} + \alpha_3))^2. \qquad (3.34)
\end{aligned}$$

Ievietosim (3.34) izteiksmē $v^{-1/2} D_3^*$,

$$\begin{aligned}
v^{-1/2} D_3^* =& \xi_2^2 (A_{13} + \alpha_3) v \\
=& v^{-1}(A_{01} - A_{01}A_{02} + A_{01}^2(A_{03} + \alpha_3))^2 (A_{13} + \alpha_3) v \\
=& (A_{01} - A_{01}A_{02} + A_{01}^2(A_{03} + \alpha_3))^2 (A_{13} + \alpha_3) \\
=& (A_{01} - A_{01}A_{02})^2 (A_{13} + \alpha_3) + A_{01}^4 (A_{03} + \alpha_3)^2 (A_{13} + \alpha_3) \\
& + 2A_{01}^2 (A_{03} + \alpha_3)(A_{01} - A_{01}A_{02})(A_{13} + \alpha_3) \\
=& A_{01}^2 A_{13} + A_{01}^2 \alpha_3 + A_{01}^2 A_{02}^2 A_{13} + A_{01}^2 A_{02}^2 \alpha_3 - 2A_{01}^2 A_{02} A_{13} \\
& - 2A_{01}^2 A_{02} \alpha_3 + 2A_{01}^3 A_{03} A_{13} + 2A_{01}^3 A_{03} \alpha_3 - 2A_{01}^3 A_{03} A_{02} A_{13} \\
& - 2A_{01}^3 A_{03} A_{02} \alpha_3 + 2A_{01}^3 \alpha_3 A_{13} + 2A_{01}^3 \alpha_3^2 - 2A_{01}^3 \alpha_3 A_{02} A_{13} \\
& - 2A_{01}^3 \alpha_3^2 A_{02} + A_{01}^4 A_{03}^2 A_{13} + A_{01}^4 A_{03}^2 \alpha_3 + 2A_{01}^4 A_{03} \alpha_3 A_{13} \\
& + 2A_{01}^4 A_{03} \alpha_3^2 + A_{01}^4 \alpha_3^2 A_{13} + A_{01}^4 \alpha_3^3 \\
=& A_{01}^2 A_{13} + A_{01}^2 \alpha_3 - 2A_{01}^2 A_{02} \alpha_3 + 2A_{01}^3 \alpha_3^2 + O_p(n^{-2}). \qquad (3.35)
\end{aligned}$$

Ievietojot (3.35) izteiksmē $D_1 v^{-1} D_3^*$, iegūsim

$$\begin{aligned}
D_1 v^{-1} D_3^* &= A_{01} v^{-1/2} D_3^* \\
&= A_{01}(A_{01}^2 A_{13} + A_{01}^2 \alpha_3 - 2A_{01}^2 A_{02} \alpha_3 + 2A_{01}^3 \alpha_3^2 + O_p(n^{-2})) \\
&= A_{01}^3 A_{13} + A_{01}^3 \alpha_3 - 2A_{01}^3 A_{02} \alpha_3 + 2A_{01}^4 \alpha_3^2 + O_p(n^{-5/2}). \qquad (3.36)
\end{aligned}$$

Tagad atradīsim $D_2 v^{-1} D_3$,

$$\begin{aligned} D_2 v^{-1} D_3 &= -v^{-1} v^{1/2} A_{01} A_{02} v^{1/2} A_{01}^2 (A_{03} + \alpha_3) \\ &= -A_{02} A_{01}^3 (A_{03} + \alpha_3) = -A_{02} A_{01}^3 A_{03} - A_{02} A_{01}^3 \alpha_3 \\ &= -A_{02} A_{01}^3 \alpha_3 + O_p(n^{5/2}). \end{aligned} \tag{3.37}$$

Vēl jānosaka $D_3^2 v^{-1}$ un $D_1 v^{-1} D_4$.

$$\begin{aligned} D_3^2 v^{-1} &= v^{-1} (v^{1/2} A_{01}^2 (A_{03} + \alpha_3))^2 = A_{01}^4 (A_{03} + \alpha_3)^2 \\ &= A_{01}^4 A_{03}^2 + A_{01}^4 \alpha_3^2 + 2 A_{01}^4 A_{03} \alpha_3 = A_{01}^4 \alpha_3^2 + O_p(n^{-5/2}) \end{aligned} \tag{3.38}$$

un lai aprēķinātu $D_1 v^{-1} D_4 = A_{01} v^{-1/2} D_4$, jāatrod D_4.

$$\begin{aligned} D_4 = C_{41} - C_{42} &= -\frac{n^3}{n_1^4} \sum_i \xi_1^3 (X_i - \mu_0)^4 - \frac{n^3}{n_2^4} \sum_j \xi_1^3 (Y_j - \mu_0)^4 \\ &= -\xi_1^3 \left(\frac{n^3}{n_1^4} \sum_i (X_i - \mu_0)^4 - \frac{n^3}{n_2^4} \sum_j (Y_j - \mu_0)^4 \right) \\ &= -\xi_1^3 (A_{04} + \alpha_4) v^2 = -v^{1/2} A_{01}^3 (A_{04} + \alpha_4). \end{aligned} \tag{3.39}$$

Ievietojam (3.39) izteiksmē $D_1 v^{-1} D_4$,

$$\begin{aligned} D_1 v^{-1} D_4 &= A_{01} v^{-1/2} D_4 = -A_{01}^4 (A_{04} + \alpha_4) \\ &= -A_{01}^4 A_{04} - A_{01}^4 \alpha_4 = -A_{01}^4 \alpha_4 + O_p(n^{-5/2}). \end{aligned} \tag{3.40}$$

Salīdzinot visas šeit iegūtās (3.26) saskaitāmo izteiksmes, jāsecina, ka ir atrasta vēl viena neprecizitāte [14], jo izteiksmei $D_1 v^{-1} D_2^*$ jābūt savādākā formā kā atrodama [14]. Izteiksme $v^{-1} D_2^*$ iegūta pareizi, neprecizitāte radusies pareizinot $v^{-1} D_2^*$ ar D_1. Tagad varam ievietot (3.28), (3.32), (3.36), (3.37), (3.38) un (3.40) $l(\lambda_1^*, \lambda_2^*, \mu^*) n^{-1}$ izvirzījumā (3.26).

$$\begin{aligned} l(\lambda_1^*, \lambda_2^*, \mu^*) n^{-1} &= A_{01}^2 - A_{01}^2 A_{12} + A_{01}^2 A_{02} A_{12} - A_{01}^3 \alpha_3 A_{12} + \frac{2}{3} A_{01}^3 A_{13} \\ &+ \frac{2}{3} A_{01}^3 \alpha_3 - A_{01}^3 A_{02} \alpha_3 + A_{01}^4 \alpha_3^2 - \frac{1}{2} A_{01}^4 \alpha_4 + O_p(n^{-5/2}). \end{aligned} \tag{3.41}$$

Pārrakstīsim (3.41) formā

$$l(\lambda_1^*, \lambda_2^*, \mu^*) n^{-1} = \Delta_1 + \Delta_2 + O_p(n^{-5/2}),$$

kur

$$\Delta_1 = A_{01}^2 + \frac{2}{3}A_{01}^3\alpha_3 + A_{01}^4\alpha_3^2 - \frac{1}{2}A_{01}^4\alpha_4 - 2A_{01}^3A_{02}\alpha_3 - A_{02}A_{01}^2$$
$$+ A_{02}^2A_{01}^2 + \frac{2}{3}A_{03}A_{01}^3, \tag{3.42}$$
$$\Delta_2 = \frac{2}{3}A_{01}^3(A_{13} - A_{03}) + (A_{02} - A_{12})(A_{01}^2 + \alpha_3A_{01}^3 - A_{01}^2A_{02}). \tag{3.43}$$

Piezīmēsim, ka iegūtais saskaitāmais Δ_1 (3.42) sakrīt ar logaritmiskās empīriskās ticamības attiecības izvirzījumu vienas izlases gadījumā (2.16), tāpēc izmantojot iegūtos rezultātus vienas izlases gadījumā (skatīt nodaļu 2.2.), varam rakstīt, ka

$$n\mathbb{E}\Delta_1 = 1 - n^{-1}\left(\frac{1}{3}\alpha_3^2 - \frac{1}{2}\alpha_4\right). \tag{3.44}$$

Gadījumā, ja

$$\Delta_2^* = \frac{2}{3}A_{01}^3(A_{13} - A_{03}) + (A_{02} - A_{12})A_{01}^2$$

Liu, Zou un Zhang iegūst, ka

$$n\mathbb{E}\Delta_2 = \frac{n}{n_1n_2}\frac{v_1v_2}{v^2}. \tag{3.45}$$

Jāpiezīmē, ka šeit iegūtā Δ_2 atšķiras no Δ_2^*, taču $\mathbb{E}\Delta_2$ noteikšana ir samērā laikietilpīgs process, tāpēc pagaidām vēl nav izdarīts. Taču tā kā Δ_2 var izteikt

$$\Delta_2 = \Delta_2^* + (A_{02} - A_{12})(\alpha_3A_{01}^3 - A_{01}^2A_{02}),$$

tad

$$\mathbb{E}\Delta_2 = \frac{n}{n_1n_2}\frac{v_1v_2}{v^2} + \delta,$$

kas nozīmē, ka atšķirībai starp Δ_2 un Δ_2^* nevajadzētu būt pārāk nozīmīgai.

Noteiktās matemātiskās cerības (3.44) un (3.45) palīdz iegūt [14] aprakstīto Bārtleta korekciju

$$b = n[n\mathbb{E}\Delta_1 + n\mathbb{E}\Delta_2 - 1] = -\frac{1}{3}\alpha_3^2 + \frac{1}{2}\alpha_4 + \frac{n^2}{n_1n_2}\frac{v_1v_2}{v^2}. \tag{3.46}$$

Teorēma 3. *[14] Ja funkciju F un G ceturtie momenti ir galīgi un ja H_0 ir spēkā, tad*

$$\mathbb{P}(l(\Delta_0) < \chi_1^2(\alpha)(1 + b/n)) = \alpha + O(n^{-2}),$$

kur b definēts ar (3.46). Un ticamības intervālu divu izlašu vidējo vērtību starpībai ar Bārtleta korekciju var definēt

$$I_\alpha = \{\Delta : l(\Delta) < \chi_1^2(\alpha)(1 + b/n)\}.$$

3.2. Simulācijas Bārtleta korekcijai vidējo vērtību starpībai

Lai praktiski novērtētu Bārtleta korekcijas efektivitāti, konstruēsim ticamības intervālus 2 izlašu vidējo vērtību starpībai dažādiem sadalījumiem un noteiksim ticamības intervālu pārklājumu precizitāti EL metodei (EL) un EL metodei ar Bārtleta korekciju ($EL_{B_{teo}}$, $EL_{B_{nov}}$), izmantojot gan īsto, gan novērtēto [14] noteikto b vērtību. 20000 reižu simulējot katras izlases datus, tika konstruēti ticamības intervāli sekojošu sadalījumu pāru vidējo vērtību starpībai:

1. $N(0,1)$ un $N(2,1)$,
2. $\exp(1)$ un $\exp(2)$,
3. χ_3^2 un $\exp(1)$,
4. $N(0,1)$ un t_5

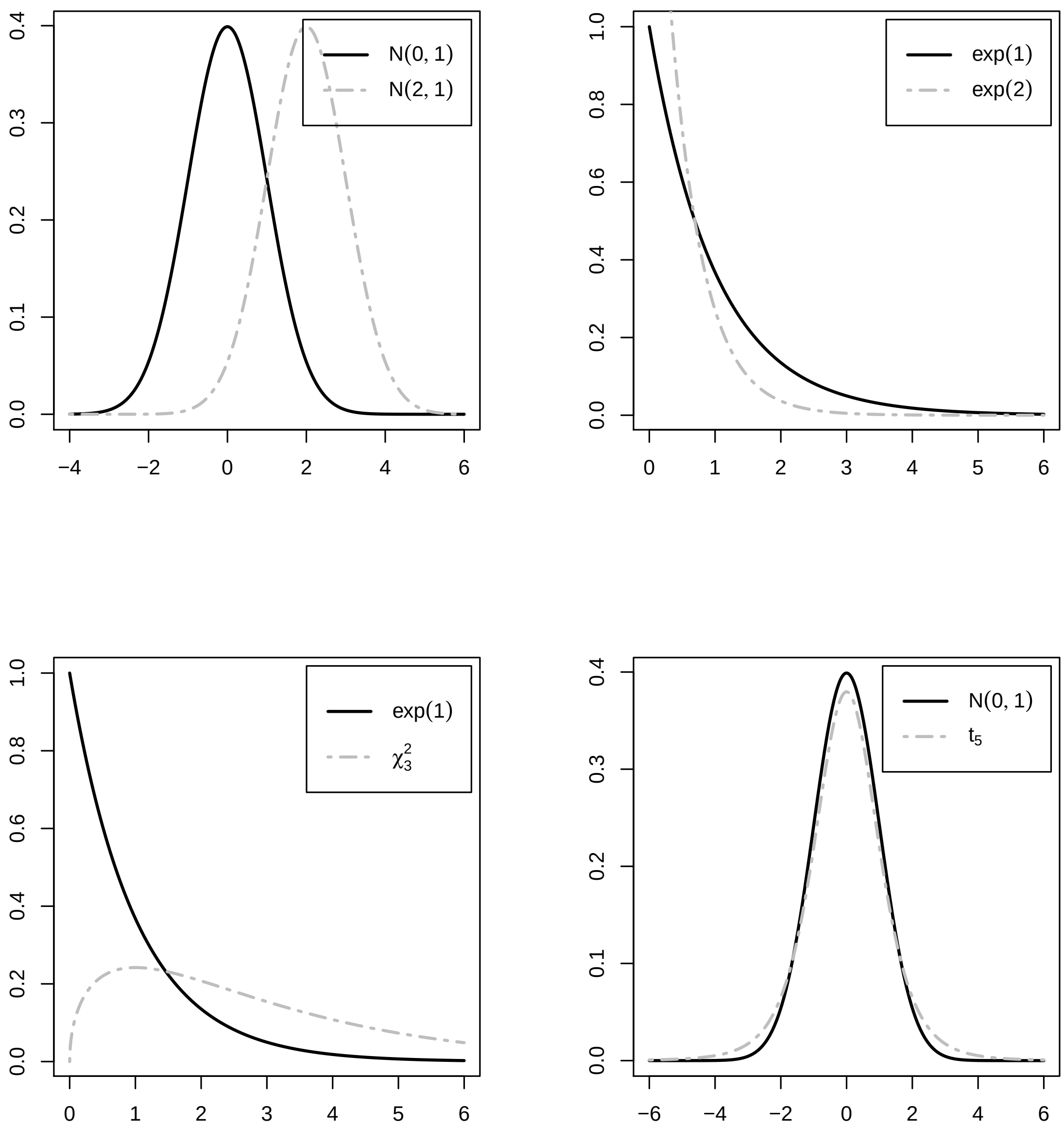
pie ticamības līmeņa $\alpha = 0.05$.

1.attēlā redzamas visos 4 gadījumos apskatīto sadalījumu pāru blīvuma funkcijas.

6. tabula Bārtleta korekcija χ_3^2 un $\exp(1)$ sadalījumu datiem

		$n_2 = 10$	$n_2 = 20$	$n_2 = 30$
$n_1 = 10$	EL	0.8877	0.9210	0.9284
	$EL_{B_{teo}}$	0.9183	0.9364	0.9412
	$EL_{B_{nov}}$	0.9057	0.9352	0.9374
$n_1 = 20$	EL	0.8915	0.9213	0.9313
	$EL_{B_{teo}}$	0.9162	0.9358	0.9425
	$EL_{B_{nov}}$	0.9056	0.9318	0.9403
$n_1 = 30$	EL	0.8838	0.9251	0.9342
	$EL_{B_{teo}}$	0.9119	0.9355	0.9427
	$EL_{B_{nov}}$	0.9007	0.9286	0.9379

6., 7., 8., 9.tabulā redzams, ka nekoriģētās EL ticamības intervālu pārklājuma precizitāte ievērojami atšķiras no ticamības līmeņa $\alpha = 0.05$ un ar Bārtleta korekcijas palīdzību var panākt būtisku uzlabojumu un palielinoties n_1 un n_2, atšķirība starp $EL_{B_{teo}}$ un $EL_{B_{nov}}$ samazinās.



1. att. Simulācijās apskatīto sadalījumu blīvuma funkcijas

7. tabula Bārtleta korekcija $N(0,1)$ un $N(2,1)$ sadalījumu datiem

		$n_2 = 10$	$n_2 = 20$	$n_2 = 30$
$n_1 = 10$	EL	0.9204	0.9288	0.9255
	$EL_{B_{teo}}$	0.9376	0.9376	0.9338
	$EL_{B_{nov}}$	0.9354	0.9338	0.9352
$n_1 = 20$	EL	0.9258	0.9390	0.9408
	$EL_{B_{teo}}$	0.9378	0.9490	0.9460
	$EL_{B_{nov}}$	0.9373	0.9458	0.9457
$n_1 = 30$	EL	0.9241	0.9420	0.9429
	$EL_{B_{teo}}$	0.9348	0.9449	0.9498
	$EL_{B_{nov}}$	0.9308	0.9483	0.9466

8. tabula Bārtleta korekcija $\exp(1)$ un $\exp(2)$ sadalījumu datiem

		$n_2 = 10$	$n_2 = 20$	$n_2 = 30$
$n_1 = 10$	EL	0.8862	0.8803	0.8757
	$EL_{B_{teo}}$	0.9186	0.9134	0.9167
	$EL_{B_{nov}}$	0.9015	0.8962	0.8946
$n_1 = 20$	EL	0.9170	0.9163	0.9201
	$EL_{B_{teo}}$	0.9379	0.9378	0.9343
	$EL_{B_{nov}}$	0.9305	0.9257	0.9280
$n_1 = 30$	EL	0.9200	0.9261	0.9289
	$EL_{B_{teo}}$	0.9389	0.9396	0.9430
	$EL_{B_{nov}}$	0.9339	0.9378	0.9374

Normālo sadalījumu gadījumā (7.tabula) pat pie nelieliem izlašu apjomiem ir neliela atšķirība starp $EL_{B_{teo}}$ un $EL_{B_{nov}}$, ko nevar teikt par gadījumu ar $N(0,1)$ un t_5 (9.tabula). 1.attēlā redzams, ka blīvuma funkciju atšķirība starp $N(0,1)$ un t_5 nav liela, stjūdenta sadalījumam ir "smagākas" astes, taču tieši šim sadalījumu pārim ir vislielākā atšķirība starp $EL_{B_{teo}}$ un $EL_{B_{nov}}$ pie nelieliem izlašu apjomiem, kā arī tieši šajā gadījumā ar $EL_{B_{teo}}$ var sasniegt labākās pārklājuma precizitātes.

9. tabula Bārtleta korekcija $N(0,1)$ un t_5 sadalījumu datiem

		$n_2 = 10$	$n_2 = 20$	$n_2 = 30$
$n_1 = 10$	EL	0.9141	0.9258	0.9255
	$EL_{B_{teo}}$	0.9435	0.9426	0.9366
	$EL_{B_{nov}}$	0.9270	0.9383	0.9368
$n_1 = 20$	EL	0.9161	0.9347	0.9355
	$EL_{B_{teo}}$	0.9468	0.9452	0.9476
	$EL_{B_{nov}}$	0.9294	0.9418	0.9456
$n_1 = 30$	EL	0.9105	0.9323	0.9390
	$EL_{B_{teo}}$	0.9469	0.9505	0.9498
	$EL_{B_{nov}}$	0.9196	0.9429	0.9445

4. Gludinātā empīriskās ticamības metode divu izlašu gadījumā

Vienas izlases gadījumā Chen un Hall [8] parādīja, ka kvantilēm ar gludināto EL ticamības intervālu konverģences precizitāti var uzlabot no $O(n^{-1/2})$ uz $O(n^{-1})$, kā arī viņi pierādīja, ka gludinātajai EL metodei var veikt Bārtleta korekciju, tādējādi uzlabojot konverģences precizitāti līdz $O(n^{-2})$.

Gludinātā EL metode ir svarīga ne tikai kvantiļu gadījumā, jo gan P-P, gan Q-Q grafikus, gan ROC līknes var izteikt kā kvantiļu funkcijas (skatīt Piemērus $1-5$), tāpēc arī šīm problēmām lietderīgi izmantot gludināto EL.

Tāpat kā vienas izlases kvantiļu gadījumā, arī divām izlasēm izmantosim kodolu gludināšanas metodi. Ar $G_j = \int_{u \le t} K_j(u)du$, $j = 1,2$ apzīmēsim gludināto deģenerēto sadalījuma funkciju G_0, kas definēta ar (2.21). Atgādināsim, ka K_j ir $r-$tās kārtas kodols, kas tika definēts vienai izlasei kvantiļu gadījumā un kuram jāizpildās (2.22). Definēsim vēl $G_{b_j}(t) = G_j(t/b_j)$, kur $b_1 = b_1(n_1)$ un $b_2 = b_2(n_2)$ ir joslu platumi, kas tiecas uz 0, n_1 un n_2 tiecoties uz bezgalību.

Pieņemsim, ka $p = (p_1, \ldots, p_{n_1})$ un $q = (q_1, \ldots, q_{n_2})$ ir divi varbūtību vektori, kas sastāv no nenegatīvām komponentēm, kas summā dod 1. Definēsim sekojošus sadalījuma funkciju novērtējumus

$$\hat{F}_{b_1,p}(x) = \sum_{i=1}^{n_1} p_i G_{b_1}(x - X_i) \quad \text{un} \quad \hat{F}_{b_2,q}(y) = \sum_{j=1}^{n_2} q_j G_{b_2}(y - Y_j). \tag{4.1}$$

Parametram Δ definēsim gludināto divu izlašu empīriskās ticamības attiecību

$$R(\Delta, \theta) = \sup_{p,q} \prod_{i=1}^{n_1} (n_1 p_i) \prod_{j=1}^{n_2} (n_2 q_j). \tag{4.2}$$

Kā jau tika minēts gludinātajai EL metodei ir plašs pielietojums, jo lielai daļai 2 izlašu problēmu funkcijas $w_1(X, \theta, \Delta, t)$ un $w_1(X, \theta, \Delta, t)$ izsakās kā funkcijas no idikatorfunkcijām. Apstiprināsim to ar dažu uzskatāmu piemēru palīdzību.

Piemērs 1. Divu sadalījuma funkciju starpība. $\theta_0 = F_1(t)$ un $\Delta_0 = F_2(t) - F_1(t)$. Šajā gadījumā (3.1) un (3.2) ir spēkā, ja

$$w_1(X, \theta_0, \Delta_0, t) = I_{\{X \leq t\}} - \theta_0, \quad w_2(Y, \theta_0, \Delta_0, t) = I_{\{Y \leq t\}} - \theta_0 - \Delta_0.$$

Piemērs 2. Divu kvantiļu funkciju starpība. Zhou un Jing [27] nodemonstrēja divu kvantiļu starpības problēmu vienas izlases gadījumā un Valeinis [28] vispārināja to 2 izlašu gadījumam. $\theta_0 = F_1^{-1}(t)$ and $\Delta_0 = F_2^{-1}(t) - F_1^{-1}(t)$. Šajā gadījumā, lai izpildītos (3.1) un (3.2),

$$w_1(X, \theta_0, \Delta_0, t) = I_{\{X \leq \theta_0\}} - t, \quad w_2(Y, \theta_0, \Delta_0, t) = I_{\{Y \leq \theta_0 + \Delta_0\}} - t.$$

Piemērs 3. P-P grafiks. Izmantojot gludināto EL metodi Claeskens [29] būtībā vēlreiz pierādīja Qin un Zhao [16] rezultātu P-P grafikiem. $\theta_0 = F_2^{-1}(t)$ un $\Delta_0 = F_1(F_2^{-1}(t))$, kas ir funkciju F_1 un F_2 P-P grafiks. Šajā gadījumā

$$w_1(X, \theta_0, \Delta_0, t) = I_{\{X \leq \theta_0\}} - \Delta_0, \quad w_2(Y, \theta_0, \Delta_0, t) = I_{\{Y \leq \theta_0\}} - t.$$

Piemērs 4. ROC līkne. ROC līkne tiek definēta kā $\Delta = 1 - F_1(F_2^{-1}(1-t))$ un tai ir ļoti būtiska nozīme medicīniskajā statistikā. ROC līknes ir cieši saistītas ar P-P grafikiem, tādējādi P-P grafiku rezultāti var tikt pielietoti ROC līkņu gadījumā. Claeskens [29] pētīja ROC līknes saistībā ar EL metodi. $\theta_0 = F_2^{-1}(1-t)$ un $\Delta_0 = 1 - F_1(F_2^{-1}(1-t))$. Šajā gadījumā

$$w_1(X, \theta_0, \Delta_0, t) = I_{\{X \leq \theta_0\}} + \Delta_0 - 1, \quad w_2(Y, \theta_0, \Delta_0, t) = I_{\{Y \leq \theta_0\}} + t - 1.$$

Piemērs 5. Q-Q grafiks. Einmahl un Mckeague [30], izmantojot EL metodi, analizēja Q-Q grafikus cenzorētiem datiem. $\theta_0 = F_2(t)$ un $\Delta_0 = F_1^{-1}(F_2(t))$, kas pazīstama kā Q-Q grafiks. Šajā gadījumā

$$w_1(X, \theta_0, \Delta_0, t) = I_{\{X \leq \Delta_0\}} - \theta_0, \quad w_2(Y, \theta_0, \Delta_0, t) = I_{\{Y \leq t\}} - \theta_0.$$

Valeinis [28] parādīja, ka visas šis 2 izlašu problēmas var apvienot.

10. tabula Funkcijas w_1 un w_2 Piemēriem 1–5.

Piemērs	$w_1(X_i, \theta, \Delta, t)$	$w_2(X_i, \theta, \Delta, t)$
1	$\hat{F}_{b_1,p}(t) - \theta_0$	$\hat{F}_{b_2,q}(t) - \theta_0 + \Delta_0$
2	$\hat{F}_{b_1,p}(\theta_0) - t$	$\hat{F}_{b_2,q}(\theta_0 + \Delta_0) - t$
3	$\hat{F}_{b_1,p}(\theta_0) - \Delta_0$	$\hat{F}_{b_2,q}(\theta_0) - t$
4	$\hat{F}_{b_1,p}(\theta_0) - 1 + \Delta_0$	$\hat{F}_{b_2,q}(\theta_0) - 1 + t$
5	$\hat{F}_{b_1,p}(\Delta_0) - \theta_0$	$\hat{F}_{b_2,q}(t) - \theta_0$

10. tabulā redzamās funkcijas w_1 un w_2, kas izteiktas caur $\hat{F}_{b_1}$ un $\hat{F}_{b_2}$. Izmantojot (4.1) 10. tabulā redzamās funkcijas w_1 un w_2 var pārrakstīt formā, kas redzama 11. tabulā.

11. tabula Funkcijas w_1 un w_2 Piemēriem 1–5.

Piemērs	$w_1(X_i, \theta, \Delta, t)$	$w_2(X_i, \theta, \Delta, t)$
1	$H_{b_1}(t - X_i) - \theta_0$	$H_{b_2}(t - Y_j) - \theta_0 - \Delta_0$
2	$H_{b_1}(\theta_0 - X_i) - t$	$H_{b_2}(\theta_0 + \Delta_0 - Y_j) - t$
3	$H_{b_1}(\theta_0 - X_i) - \Delta_0$	$H_{b_2}(\theta_0 - Y_j) - t$
4	$H_{b_1}(\theta_0 - X_i) - 1 + \Delta_0$	$H_{b_2}(\theta_0 - Y_j) - 1 + t$
5	$H_{b_1}(\Delta_0 - X_i) - \theta_0$	$H_{b_2}(t - Y_j) - \theta_0$

Ņemot vērā, to ka visiem apskatītajiem piemēriem funkcijas w_1 un w_2 izsakās līdzīgā veidā, Valeinis [28] parādīja, ka novērtējošos vienādojumus w_1 un w_2 gludinātās EL metodes gadījumā var pārrakstīt vispārīgā formā

$$w_1(X_i, \theta, \Delta, t) = G_{b_1}(\xi_1(\theta) - X_i) - \xi_2(\theta), \tag{4.3}$$

$$w_2(Y_j, \theta, \Delta, t) = G_{b_2}(\psi_1(\theta) - X_i) - \psi_2(\theta), \tag{4.4}$$

kuriem

$$\mathbb{E}_{F_1}(w_1(X_i, \theta_0, \Delta, t)) = 0, \quad \mathbb{E}_{F_2}(w_2(Y_j, \theta_0, \Delta, t)) = 0. \tag{4.5}$$

Arī gludinātās EL gadījumā ir spēkā Teorēma 1, taču, lai noteiktu joslas platumu b_1 un b_2 asimptotiskās kārtas, jāpārskata Teorēmas 1 pierādījums šajā konkrētajā gadījumā. Atzīmēsim, ka Claeskens [29] šo kārtu ir noteikusi ROC līkņu gadījumā un Valeinis [17] – strukturālo attiecību modeļu vispārinātajam P-P grafikam.

4.1. Bārtleta korekcija gludinātajai EL metodei

Mēģināsim iegūt Bārtleta korekciju vispārinātajai 2 izlašu problēmai gludinātajai EL. Izmantojot (3.8), (3.9) un (3.11), varam definēt mūs interesējošu sistēmu

$$\begin{cases} \dfrac{1}{n_1}\displaystyle\sum_{i=1}^{n_1} \dfrac{w_1(X_i, \theta, \Delta, t)}{1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t)} = 0, \\ \dfrac{1}{n_2}\displaystyle\sum_{j=1}^{n_2} \dfrac{w_2(Y_j, \theta, \Delta, t)}{1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t)} = 0, \\ \displaystyle\sum_{i=1}^{n_1} \dfrac{\lambda_1(\theta)\alpha_1(X_i, \theta, \Delta, t)}{1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t)} + \displaystyle\sum_{j=1}^{n_2} \dfrac{\lambda_2(\theta)\alpha_2(Y_j, \theta, \Delta, t)}{1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t)} = 0, \end{cases} \tag{4.6}$$

kur $w_1(X_i, \theta, \Delta, t)$ un $w_2(X_i, \theta, \Delta, t)$ definēts ar (4.3) un (4.4). Lai nedaudz vienkāršotu pierakstu, apzīmēsim

$$\lambda_1 = \lambda_1(\theta), n = n_1 + n_2, \lambda_2 = \lambda_2(\theta), w_1 = w_1(X_i, \theta, \Delta, t) \text{ un } w_2 = w_2(Y_j, \theta, \Delta, t).$$

Sākotnēji, lai iegūtu Bārtleta korekciju vispārīgai 2 izlašu problēmai, bija plānots balstīties uz Liu, Zou un Zhang [14] izdarīto divu izlašu vidējo vērtību starpībai, taču sīkāk iepazīstoties ar pierādījuma detaļām (skatīt nodaļu 3.1.), nācās secināt, ka šo pieeju nevar pielietot aprakstītajai vispārinātajai divu izlašu problēmai, jo sistēmas (4.6) 3.vienādojums neizsakās tik vienkāršā formā kā sistēmas (3.17) 3.vienādojums.

Tā vietā izvirzītā mērķa sasniegšanai nepieciešams vairāk balstīties uz Chen un Hall [8], kuri parādīja Bārtleta korekciju gludinātajai EL kvantilēm vienas izlases gadījumā.

Izvirzīsim sistēmas (4.6) 1. vienādojumu Teilora rindā

$$
\begin{aligned}
0 &= n_1^{-1}\sum_{i=1}^{n_1}\frac{w_1}{1+\lambda_1 w_1} = n_1^{-1}\sum_{i=1}^{n_1} w_1(1-\lambda_1 w_1+(\lambda_1 w_1)^2-(\lambda_1 w_1)^3+\ldots) \\
&= n_1^{-1}\sum_{k=1}^{\infty}\sum_{i=1}^{n_1}(-\lambda_1)^{k-1}w_1^k = \sum_{k=1}^{\infty}(-\lambda_1)^{k-1}n_1^{-1}\sum_{i=1}^{n_1}w_1^k \\
&= \sum_{k=1}^{\infty}(-\lambda_1)^{k-1}\bar{w}_{1k},
\end{aligned} \tag{4.7}
$$

kur $\bar{w}_{1k} = n_1^{-1}\sum_{i=1}^{n_1} w_1^k$.

Izmantojot Lemmu 3 no [28] (4.7) var pārrakstīt, $s \geq 1$,

$$
\sum_{k=1}^{\infty}(-\lambda_1)^{k-1}\bar{w}_{1k} = \sum_{k=1}^{s}(-\lambda_1)^{k-1}\bar{w}_{1k} + O_p((n_1^{-1/2}+b_1^r)^s) = 0,
$$

Neņemot vērā $O_p((n_1^{-1/2}+b_1^r)^s)$, iegūstam

$$
\sum_{k=1}^{s}(-\lambda_1)^{k-1}\bar{w}_{1k} = 0.
$$

Tagad iteratīvi risināsim šo vienādojumu, lai atrastu λ_1 katram $s \geq 1$. Gadījumā, kad $s = 1$, mēs nevaram iegūt vienu unikālu atrisinājumu.

$s = 2:$ $\bar{w}_{11} - \lambda_1\bar{w}_{12} = 0 \Rightarrow \lambda_1 = \bar{w}_{12}^{-1}\bar{w}_{11}$;

$s = 3:$ $\bar{w}_{11} - \lambda_1\bar{w}_{12} + \lambda_1^2\bar{w}_{13} = 0$ un izmantojot $\lambda_1 = \bar{w}_{12}^{-1}\bar{w}_{11}$, iegūstam

$w_1 - \lambda_1\bar{w}_{12} + (\bar{w}_{12}^{-1}\bar{w}_{11})^2\bar{w}_{13} = 0 \Rightarrow \lambda_1 = \bar{w}_{12}^{-1}\bar{w}_{11} + \bar{w}_{13}\bar{w}_{12}^{-3}\bar{w}_{11}^2$;

$s = 4:$ $\bar{w}_{11} - \lambda_1\bar{w}_{12} + \lambda_1^2\bar{w}_{13} - \lambda_1^3\bar{w}_{14} = 0$ un ievietojot $\lambda_1 = \bar{w}_{12}^{-1}\bar{w}_{11} + \bar{w}_{13}\bar{w}_{12}^{-3}\bar{w}_{11}^2$,

iegūstam

$$
\begin{aligned}
\lambda_1 =& \bar{w}_{12}^{-1}\bar{w}_{11} + \bar{w}_{13}\bar{w}_{12}^{-3}\bar{w}_{11}^2 + (2\bar{w}_{12}^{-5}\bar{w}_{13}^2 - \bar{w}_{12}^{-4}\bar{w}_{14})\bar{w}_{11}^3 \\
&+ (\bar{w}_{13}^3\bar{w}_{12}^{-7} - 3\bar{w}_{12}^{-6}\bar{w}_{13}\bar{w}_{14})\bar{w}_{11}^4 - 3\bar{w}_{13}^2\bar{w}_{12}^{-8}\bar{w}_{14}\bar{w}_{11}^5 - \bar{w}_{13}^3\bar{w}_{12}^{-10}\bar{w}_{14}\bar{w}_{11}^6,
\end{aligned}
$$

ko var arī pārrakstīt formā

$$
\begin{aligned}
\lambda_1 =& \bar{w}_{12}^{-1}\bar{w}_{11} + \bar{w}_{13}\bar{w}_{12}^{-3}\bar{w}_{11}^2 + (2\bar{w}_{12}^{-5}\bar{w}_{13}^2 - \bar{w}_{12}^{-4}\bar{w}_{14})\bar{w}_{11}^3 \\
&+ \sum_{k=4}^{j} R_{1k}\bar{w}_{11}^k + O_p((n_1^{-1/2}+b_1^r)^{j+1}).
\end{aligned} \tag{4.8}
$$

No Lemmas 4 [28] seko, ka

$$
\lambda_2 = -\frac{1}{\kappa}\frac{\beta_{10}}{\beta_{20}}\lambda_1 + o_p(n^{-1/2}), \tag{4.9}
$$

kur $\kappa \to n_2/n_1$, kad $n_1, n_2 \to \infty$, un $\beta_{10} = f_1(\xi(\theta_0))\xi_1(\theta_0)'$, $\beta_{20} = f_2(\phi(\theta_0))\phi_1(\theta_0)'$. Neņemot vērā $o_p(n^{-1/2})$ (4.9), to var pārrakstīt

$$\lambda_2 = c\lambda_1 \tag{4.10}$$

Tagad apskatīsim empīrisko logaritmisko ticamības attiecību W no (3.10). Izvirzot Teilora rindā $\log(1 + \lambda_1 w_1)$ un $\log(1 + \lambda_2 w_2)$ un izmantojot (4.10), iegūstam

$$\begin{aligned}
W &= 2\left(\sum_{i=1}^{n_1} \log(1 + \lambda_1 w_1) + \sum_{j=1}^{n_2} \log(1 + \lambda_2 w_2)\right) \\
&= 2\left(\sum_{i=1}^{n_1}\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k w_1^k + \sum_{j=1}^{n_2}\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}(c\lambda_1)^k w_2^k\right) \\
&= 2\left(n_1\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k \bar{w}_{1k} + n_2\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}(c\lambda_1)^k \bar{w}_{2k}\right) \\
&= 2\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k\left(n_1\bar{w}_{1k} + n_2 c^k \bar{w}_{2k}\right) \\
&= 2\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k \tilde{v}_k \qquad (4.11) \\
&= 2n\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k \bar{v}_k \qquad (4.12)
\end{aligned}$$

kur $\tilde{v}_k = n_1\bar{w}_{1k} + n_2 c^k \bar{w}_{2k}$ un $\bar{v}_k = (n_1/n)\bar{w}_{1k} + (n_2/n)c^k\bar{w}_{2k}$.

Iegūtā izteiksme (4.12) ir līdzīgā formā kā vienas izlases kvantiļu gadījumā (2.27). Atgriezīsimies pie W izvirzījuma

$$\begin{aligned}
W &= 2\sum_{k=1}^{\infty}(-1)^{k+1}k^{-1}\lambda_1^k \tilde{v}_k \\
&= 2\left(\lambda_1\tilde{v}_1 - \frac{1}{2}\lambda_1^2\tilde{v}_2 + \frac{1}{3}\lambda_1^3\tilde{v}_3 - \frac{1}{4}\lambda_1^4\tilde{v}_4 + \frac{1}{5}\lambda_1^5\tilde{v}_5 - \dots\right) \qquad (4.13)
\end{aligned}$$

Pieņemot, ka $\lambda_1 = \bar{w}_{12}^{-1}\bar{w}_{11} + \bar{w}_{13}\bar{w}_{12}^{-3}\bar{w}_{11}^2 + (2\bar{w}_{12}^{-5}\bar{w}_{13}^2 - \bar{w}_{12}^{-4}\bar{w}_{14})\bar{w}_{11}^3$, iegūsim λ_1^2, λ_1^3, λ_1^4 un λ_1^5.

$$\begin{aligned}
\lambda_1^2 =& \bar{w}_{12}^{-2}\bar{w}_{11}^2 + 5\bar{w}_{13}^2\bar{w}_{12}^{-6}\bar{w}_{11}^4 + 2\bar{w}_{12}^{-4}\bar{w}_{11}^3\bar{w}_{13} + 4\bar{w}_{12}^{-10}\bar{w}_{13}^4\bar{w}_{11}^6 + \bar{w}_{12}^{-8}\bar{w}_{14}^2\bar{w}_{11}^6 \\
&- 4\bar{w}_{12}^{-9}\bar{w}_{13}^2\bar{w}_{14}\bar{w}_{11}^6 - 2\bar{w}_{12}^{-5}\bar{w}_{14}\bar{w}_{11}^4 + 4\bar{w}_{12}^{-8}\bar{w}_{13}^3\bar{w}_{11}^5 - 2\bar{w}_{12}^{-7}\bar{w}_{13}\bar{w}_{14}\bar{w}_{11}^5,
\end{aligned}$$

izteiksmes priekš λ_1^3, λ_1^4 un λ_1^5 jau ir daudz garākas, tāpēc tās šeit neuzrādīsim.

Ievietojot iegūtās $\lambda_1^i,\ i=\overline{1,5}$ (4.13), iegūstam

$$\begin{aligned}
W &= 2\tilde{v}_1\bar{w}_{12}^{-1}\bar{w}_{11} + (2\tilde{v}_1\bar{w}_{13}\bar{w}_{12}^{-3} - \tilde{v}_2\bar{w}_{12}^{-2})\bar{w}_{11}^2 \\
&+ \left(\frac{2}{3}\tilde{v}_3\bar{w}_{12}^{-3} - 2\tilde{v}_2\bar{w}_{13}\bar{w}_{12}^{-4} - 2\tilde{v}_1\bar{w}_{14}\bar{w}_{12}^{-4} + 4\tilde{v}_1\bar{w}_{13}^2\bar{w}_{12}^{-5}\right)\bar{w}_{11}^3 \\
&+ \left(2\tilde{v}_3\bar{w}_{12}^{-5}\bar{w}_{13} - \frac{1}{2}\tilde{v}_4\bar{w}_{12}^{-4} + 2\tilde{v}_2\bar{w}_{14}\bar{w}_{12}^{-5} - 5\tilde{v}_2\bar{w}_{13}^2\bar{w}_{12}^{-6}\right)\bar{w}_{11}^4 \\
&+ \left(\frac{2}{5}\tilde{v}_5\bar{w}_{12}^{-5} + 2\tilde{v}_2\bar{w}_{13}\bar{w}_{14}\bar{w}_{12}^{-7} - 2\tilde{v}_3\bar{w}_{14}\bar{w}_{12}^{-6} + 6\tilde{v}_3\bar{w}_{13}^2\bar{w}_{12}^{-7} - 4\tilde{v}_2\bar{w}_{13}^3\bar{w}_{12}^{-8} - 2\tilde{v}_4\bar{w}_{13}\bar{w}_{12}^{-6}\right)\bar{w}_{11}^5 \\
&+ \sum_{k=5}^{j} R_{2k}\bar{w}_{11}^{k+1} + (o_p(b) + O_p(\delta + l^{-1/2}))^{j+2}, \qquad (4.14)
\end{aligned}$$

kur $b=\min(b_1,b_2)$, $l=\min(n_1,n_2)$ un $\delta=O(b^r)$.

W izvirzījumā ievietojam $\tilde{v}_k = n_1\bar{w}_{1k} + n_2c^k\bar{w}_{2k}$ iegūstot,

$$\begin{aligned}
W =& n_1\{\bar{w}_{11}^2\bar{w}_{12}^{-1} + \frac{2}{3}\bar{w}_{11}^3\bar{w}_{13}\bar{w}_{12}^{-3} + \left(\bar{w}_{13}^2\bar{w}_{12}^{-5} - \frac{1}{2}\bar{w}_{14}\bar{w}_{12}^{-4}\right)\bar{w}_{11}^4 \\
&+ \left(2\bar{w}_{13}^3\bar{w}_{12}^{-7} - 2\bar{w}_{13}\bar{w}_{14}\bar{w}_{12}^{-6} + \frac{2}{5}\bar{w}_{15}\bar{w}_{12}^{-5}\right)\bar{w}_{11}^5\} + n_1\sum_{k=5}^{j} R_{3k}\bar{w}_{11}^{k+1} \\
&+ n_2\{2c\bar{w}_{21}\bar{w}_{12}^{-1}\bar{w}_{11} + \left(2c\bar{w}_{21}\bar{w}_{12}^{-3}\bar{w}_{13} - c^2\bar{w}_{22}\bar{w}_{12}^{-2}\right)\bar{w}_{11}^2 \\
&+ \left(4c\bar{w}_{21}\bar{w}_{12}^{-5}\bar{w}_{13}^2 - 2c\bar{w}_{21}\bar{w}_{12}^{-4}\bar{w}_{14} - 2c^2\bar{w}_{22}\bar{w}_{12}^{-4}\bar{w}_{13} + \frac{2}{3}c^3\bar{w}_{23}\bar{w}_{12}^{-3}\right)\bar{w}_{11}^3 \\
&+ \left(2c^2\bar{w}_{14}\bar{w}_{12}^{-5}\bar{w}_{22} + 2c^3\bar{w}_{13}\bar{w}_{12}^{-5}\bar{w}_{23} - 5c^2\bar{w}_{13}^2\bar{w}_{12}^{-6}\bar{w}_{22} - \frac{1}{2}c^4\bar{w}_{24}\bar{w}_{12}^{-4}\right)\bar{w}_{11}^4 \\
&+ \{6c^3\bar{w}_{23}\bar{w}_{12}^{-7}\bar{w}_{13}^2 - 2c^4\bar{w}_{24}\bar{w}_{12}^{-6}\bar{w}_{13} - 2c^3\bar{w}_{23}\bar{w}_{12}^{-6}\bar{w}_{14} \\
&- 4c^2\bar{w}_{13}^3\bar{w}_{12}^{-8}\bar{w}_{22} + 2c^2\bar{w}_{13}\bar{w}_{12}^{-7}\bar{w}_{14} + \frac{2}{5}c^5\bar{w}_{25}\bar{w}_{12}^{-5}\}\bar{w}_{11}^5\} \\
&+ n_2\sum_{k=5}^{j} R_{4k}\bar{w}_{11}^{k+1} + (o_p(b) + O_p(\delta + l^{-1/2}))^{j+2}. \qquad (4.15)
\end{aligned}$$

Apskatot (4.15), redzams, ka empīriskās logaritmiskās ticamības attiecības izvirzījumu, esam ieguvuši formā

$$W = K_1 + K_2,$$

kur K_1 ir atkarīgs tikai no 1.izlases un sakrīt ar W izvirzījumu vienas izlases kvantiļu gadījumā (2.27).

Tālāk jādarbojas līdzīgi kā tika aprakstīts vienas izlases kvantiļu gadījumā, taču tā kā šajā – vispārīgās 2 izlašu problēmas – gadījumā, W izteiksme ir daudz komplicētākā formā, tad šo soļu izpilde ir daudz sarežģītāka un laikietilpīgāka.

Secinājumi

Darbā tika apskatīta empīriskās ticamības metode – vienīgā neparametriskā metode, kurai iespējama Bārtleta korekcija. Darba izstrādes laikā tika veikta padziļināta teorētiska Bārtleta korekcijas analīze vienas izlases gadījumā – vidējai vērtībai un kvantilēm un divu izlašu gadījumā – vidējo vērtību starpībai, kas prasīja arī padziļinātu iepazīšanos ar Edgeworth izvirzījumiem.

Veicot praktisku Bārtleta korekcijas efektivitātes pārbaudi ar empīrisko pārklājuma precizitāti vienas izlases gadījumā – vidējai vērtībai un mediānai un divu izlašu gadījumā – vidējo vērtību starpībai, nācās secināt, ka, jau pie nelieliem izlašu apjomiem, ar Bārtleta korekcijas palīdzību iespējams ievērojami uzlabo pārklājuma precizitāti un, palielinoties izlases apjomam, atšķirība starp uzlabojumu, ko sniedz teorētiskā un novērtētā Bārtleta korekcija, samazinās.

Sākotnēji bija plānots, lai iegūtu Bārtleta korekciju empīriskajai ticamības metodei [28] iegūtajai vispārinātajai 2 izlašu problēmai, balstīties uz Liu, Zou un Zhang [14] izdarīto divu izlašu vidējo vērtību starpībai, taču pārstrādājot publikācijā atrodamo pierādījumu un izlabojot tur konstatētās neprecizitātes, nācās secināt, ka šo pieeju 3.vienādojuma sarežģītās formas dēļ nevar pielietot aprakstītajai vispārinātajai divu izlašu problēmai.

Tādēļ izvirzītā mērķa sasniegšanai, tika nolemts vairāk balstīties uz Chen un Hall [8] pieeju, kuri parādīja Bārtleta korekciju gludinātajai EL kvantilēm vienas izlases gadījumā. Izmantojot šo pieeju tika iegūts, empīriskās logaritmiskās ticamības attiecības izvirzījums vispārinātajai 2 izlašu problēmai.

Jau vienai izlasei Bārtleta korekcijas iegūšana kvantiļu gadījumā ir ļoti sarežģīts uzdevums un divu izlašu vispārinātajai problēmai tas kļūst vēl tikai sarežģītāks. Idejiski nepieciešamie darba posmi ir skaidri, taču problēma pagaidām ir praktiskajā realizācijā, kam nepieciešama padziļināta papildus izpēte.

Izmantotā literatūra un avoti

[1] Art B. Owen. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75(2):237–249, 1988.

[2] Art B. Owen. Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18(1):90–120, 1990.

[3] Art Owen. Empirical likelihood for linear models. *The Annals of Statistics*, 19:1725–1747, 1991.

[4] *Empirical Likelihood.* CHAPMAN & HALL/CRC, 2001.

[5] Thomas J. DiCiccio, Peter Hall, and Joseph P. Romano. Comparison of parametric and empirical likelihood functions. *Biometrika*, 76(3):465–476, 1989.

[6] Peter Hall and Barbara La Scala. Methodology and algorithms of empirical likelihood. *International Statistical Review*, 58(2):109–127, 1990.

[7] Thomas J. DiCiccio, Peter Hall, and Joseph P. Romano. Empirical likelihood is Bartlett-correctable. *Annals of Statistics*, 19(2):1053–1061, 1991.

[8] Song Xi Chen and Peter Hall. Smoothed empirical likelihood confidence intervals for quantiles. *Annals of Statistics*, 21(3):1166–1181, 1993.

[9] Song Xi Chen. On the accuracy of empirical likelihood confidence regions for linear regression model. *Ann. Inst. Statist. Math.*, 45(4):621–637, 1993.

[10] Song Xi Chen. Empirical likelihood confidence intervals for linear regression coefficients. *Journal of Multivariate analysis*, 49:24–40, 1994.

[11] Nicole A. Lazar and Per Aslak Mykland. Empirical likelihood in the presence of nuisance parameters. *Biometrika*, 86(1):203–211, 1999.

[12] Song Xi Chen and Hengjiang Cui. On Bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika*, 93(1):215–220, 2006.

[13] Bing-Yi Jing. Two-sample empirical likelihood method. *Statistics & Probability Letters*, 24:315–319, 1995.

[14] Yukun Liu, Changliang Zou, and Runchu Zhang. Empirical likelihood for the two-sample mean problem. *Statistics & Probability Letters*, 78:548–556, 2008.

[15] Yukun Liu and Chi Wai Yu. Bartlett correctable two-sample adjusted empirical likelihood. *Journal of Multivariate Analysis*, 101:1701–1711, 2010.

[16] Y. S. Qin and L. C. Zhao. Empirical likelihood ratio confidence intervals for various differences of two populations. *Systems Sci. Math. Sci.*, 13, 2000.

[17] Jānis Valeinis. *Confidence bands for structural relationship models*. PhD thesis, Georg-August-Universitat zu Gottingen, 2007.

[18] R. N. Bhattacharya and J. K. Ghosh. On the validity of the formal Edgeworth expansion. *The Annals of Statistics*, 6(2):434–451, 1978.

[19] *The Bootstrap and Edgeworth expansion*. Springer-Verlag, 1997.

[20] *All of statistics. A Concise Course in Statistical inference*. Springer, 2004.

[21] Jing Qin and Jerry Lawless. Empirical likelihood and general estimating equations. *Annals of Statistics*, 22(1):300–325, 1994.

[22] Sandra Vucāne. Empīriskās ticamības funkcijas metodes analīze un tās pielietojumi statistikā.

[23] Thomas J. DiCiccio, Peter Hall, and Joseph P. Romano. Bartlett adjustement for empirical likelihood. Technical Report 298, Stanford university, 1988.

[24] Bing-Yi Jing un Andrew T. A. Wood. Exponential empirical likelihood is not Bartlett correctable. *The Annals of Statistics*, 24(1):365–369, 1996.

[25] *Continuous univariate distributions*. 2. John Wiley & Sons, New York, 1970.

[26] O. E. Barndorff-Nielsen and Peter Hall. On the level-error after Bartlett adjustement of the likelihood ratio statistic. *Biometrika*, 75(2):374–378, 1988.

[27] Wang Zhou and Bing-Yi Jing. Adjusted empirical likelihood method for quantiles. *Ann. Inst. Statist. Math.*, 55(4):689–703, 2003.

[28] Jānis Valeinis. General two sample empirical likelihood.

[29] Gerda Claeskens, Bing-Yi Jing, Liang Peng, and Wang Zhou. Empirical likelihood confidence regions for comparison distributions and ROC curves.

[30] John H. J. Einmahl and Ian W. McKeague. Confidence tubes for multiple quantile plots via empirical liekelihood. *The Annals of Statistics*, 27(4):1348–1367, 1999.

Maģistra darbs “Bārtleta korekcija empīriskās ticamības metodei” 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: Sandra Vucāne

____________________ ________________

(paraksts) (datums)

Rekomendēju darbu aizstāvēšanai.

Vadītājs: doc. Dr.math. Jānis Valeinis

____________________ ________________

(paraksts) (datums)

Recenzents: doc. Dr.math. Nadežda Siņenko

____________________ ________________

(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)