

LATVIJAS UNIVERSITĀTE
FIZIKAS, MATEMĀTIKAS UN OPTOMETRIJAS FAKULTĀTE
MATEMĀTIKAS NODAĻA

**PARAMETRIKĀS UN EMPĪRISKĀS TICAMĪBAS FUNKCIJU
METOŽU KOMBINĀCIJAS**

MAĢISTRA DARBS

Autors: **Lauma Burkēviča**

Stud. apl. lb17110

Darba vadītājs: prof. Dr. Math. Jānis Valeinis

RĪGA 2019

Anotācija

Šajā darbā aplūkota parametriskās un empīriskās ticamības funkcijas metožu kombinācija, kas atrod kompromisu starp parametrisko un neparametrisko ticamības metodi. Abu metožu kombinācija rezultātā dod metodi, kas aizņemas parametriskā modeļa spējas kopā ar empīriskās ticamības metodes robustību. Darba gaitā tiek izpētītas metodes novērtējuma īpašības gan tās parametriskajai, gan neparametriskajai daļai. Šī darba mērķis ir izpētīt, kā darbojas šī metode, un salīdzināt ar parametriskās un empīriskās ticamības metodi. Tika izveidots apvienotais modelis un realizēts programmā R, kur tas tika izmantots reāliem un ģenerētiem datu piemēriem.

Atslēgas vārdi: parametriskā ticamība, empīriskā ticamība, kombinācija, robustība, fokusa parametrs.

Abstract

This thesis views combination of parametric and empirical likelihood function methods, which compromise between parametric and nonparametric likelihoods. Combination of both methods gives a model which borrows strength from parametric model with robustness from empirical likelihood method. Both properties of parametric and nonparametric part of the model are explored in this thesis. The aim of this thesis is to explore how this method works and compare it with parametric and empirical likelihood. The combined model was made and developed in program R, where it was used on real and generated datasets.

Keywords: parametric likelihood, empirical likelihood, combination, robustness, focus parameter.

Saturs

IEVADS	2
1 VISLIELĀKĀS TICAMĪBAS FUNKCIJAS METODE	4
1.1 Vislielākās ticamības funkcija un tās novērtējumi	4
1.2 Asimptotiskā normalitāte	5
1.3 Delta metode	7
2 EMPĪRISKĀS TICAMĪBAS FUNKCIJAS METODE	8
2.1 Neparametriskā vislielākās ticamības funkcija	8
2.2 Neparametriskās ticamības funkcijas attiecības tests	9
2.3 Empīriskās ticamības metodes vispārīgais gadījums	9
2.4 Ticamības intervālu konstruēšana kvantīlēm	12
3 PARAMETRISKĀS UN EMPĪRISKĀS TICAMĪBAS FUNKCIJU METOŽU KOMBINĀCIJAS	15
3.1 Ideja	15
3.2 Piemēri	16
4 HL NOVĒRTĒJUMA ASIMPTOTISKĀS ĪPAŠĪBAS	18
4.1 HL metodes pamatīpašības	18
4.2 Parametru ticamības intervāli	22
4.3 Efektivitātes zudums	23
5 HL METODE ĀRPUS PARAMETRISKĀ MODEĻA NOSACĪJUMIEM	24
6 BALANSA PARAMETRA REGULĒŠANA	27
6.1 Robustības un efektivitātes balansēšana	27
6.2 Vidējās kvadrātiskās kļūdas īpašības	27
6.3 Vidējās kvadrātiskās kļūdas novērtējums	29
7 HL REGRESIJAS MODEĻI	31
8 PRAKTISKIE PIEMĒRI	32
8.1 Ēģiptiešu dzīves ilgums Romas impērijas valdīšanas laikā	32
8.2 Normālā sadalījuma novērtēšana ar kvartīļu palīdzību	35
SECINĀJUMI	38
PIELIKUMS	41

IEVADS

Datu modelēšana ietver sevī daudzas dažādas metodes, sākot ar pilnīgi parametriskām un līdz pat pilnīgi neparametriskām. Piemēram, vislielākā ticamības metode ir viena no visbiežāk izmantotajām metodēm, ko izmanto parametru novērtēšanai parametriskos modeļos, kamēr empīriskās ticamības metode ir ļoti noderīga gadījumos, kad sadalījuma likumu ir grūti noteikt, attiecīgi tā pieder pie neparametriskajām metodēm. Eksistē arī neskaitāmi veidi, kā kombinēt parametriskās metodes ar neparametriskajām, piemēram, ņemot vidējo svērto starp parametriskiem un neparametriskiem novērtējumiem. 2017. gada publikācija [5] interesējas par parametriskās un empīriskās ticamības funkcijas metožu kombinācijām, piedāvājot izveidot "tiltu" starp parametrisku modeli un neparametrisko empīrisko ticamības attiecības metodi. Abu metožu kombinētā ticamības (*hybrid likelihood*, turpmāk tekstā HL) funkcija tiek konstruēta par pamatu izmantojot parasto ticamības funkciju un empīriskās ticamības funkciju. Publikācija piedāvā arī dažādus piemērus šīs metodes praktiskam izmantojumam, kas ietver kvantiļu novērtēšanu normāli sadalītai datu kopai, gan arī variantu, kas ļauj jebkādam zināmam sadalījumam atbilstošu datu kopu sadalīt vairākos datu intervālos, un veikt atsevišķu analīzi katram no tiem. Tiek arī parādīts, ka HL novērtējums bieži vien pārspēj vislielākās ticamības metodi, ja parametriskais modelis nespēj iegūt pareizus novērtējumus. HL metode piedāvā parametrisku robustību, tajā pašā laikā aizņemoties empīriskās ticamības funkcijas spējas caur izvēlētajiem kontroles parametriem.

Šī darba mērķis ir izpētīt, kā darbojas HL metode un iespēju robežās salīdzināt to ar parasto vislielākās ticamības metodi un neparametrisko empīriskās ticamības metodi. Lai sasniegtu darba mērķi, tika izvirzīti vairāki darba uzdevumi:

1. Iepazīties ar empīriskās attiecības funkcijas metodi,
2. Iepazīties ar HL funkcijas metodi un tās piemēriem,
3. Pielietot HL metodi ģenerētiem un reāliem datiem.

Darbs sastāv no ievada, 8 nodaļām, secinājumiem, literatūras avotu saraksta un pielikuma. 1. nodaļā aprakstīta vislielākās ticamības funkcijas metode, kas ir pamatā daļai no HL funkcijas metodes, 2. nodaļā aprakstīta empīriskās ticamības funkcijas metodes vispārīgais gadījums un ticamības intervālu konstruēšana kvantiļu gadījumā, kas veido otru pamatu daļai no HL funkcijas metodes, 3. nodaļā aprakstīta HL metodes ideja un tiek doti vairāki tās izmantošanas piemēri, 4. nodaļa aplūko HL novērtējuma asimptotiskās īpašības, t.i., tā apraksta HL parametriskās daļas

uzvedību, 5. nodaļa attiecīgi aplūko HL metodi ārpus parametriskā modeļa, t.i., tās empīriskās ticamības daļas īpašības, 6. nodaļa apraksta balansa parametra regulēšanas metodes, pētot tās robustību, efektivitāti un vidējo kvadrātisko kļūdu, 7. nodaļa sniedz īsu askatu HL regresijas modeļu idejai un 8. nodaļā aprakstīti praktiskās daļas rezultāti. Pielikumā atrodamas visas programmā R veidotās programmas, kuru iegūtie rezultāti redzami 8. nodaļā.

1. VISLIELĀKĀS TICAMĪBAS FUNKCIJAS METODE

Vislielākās ticamības metodes ieviešana bija viens no vissvarīgākajiem rezultātiem 20. gadsimta matemātiskajā statistikā, ko 1922. gadā ieviesa Britu statistiķis un ģenētiķis R.A. Fišers. [2] Tā ir visbiežāk lietotā metode parametru novērtēšanai parametriskā modelī.[15]

1.1. Vislielākās ticamības funkcija un tās novērtējumi

Pieņemsim, ka X_1, \dots, X_n ir vienādi sadalīti un neatkarīgi gadījuma lielumi ar varbūtības blīvuma vai varbūtības masas funkciju $f(x|\theta)$.

Definīcija 1. Ticamības funkcija ir definēta šādi:[15]

$$L_n(\theta) = \prod_{i=1}^n f(X_i|\theta). \quad (1.1)$$

Log-ticamības funkcija ir definēta šādi:

$$\ell_n(\theta) = \log L(\theta).$$

Ticamības funkcija ir datu kopīgais blīvums, vienīgi mēs to uztveram kā funkciju parametram θ . Tādējādi $L_n : \Theta \rightarrow [0, \infty)$. Ticamības funkcija nav tas pats, kas blīvuma funkcija, t.i., integrējot $L_n(\theta)$ neiegūst vērtību 1.

Definīcija 2. Maksimālās ticamības novērtējums $\hat{\theta}_n$ ir tāda θ vērtība, kas maksimizē funkciju $L_n(\theta)$.

$\ell_n(\theta)$ maksimums atrodas tajā pašā vietā, kur $L_n(\theta)$ maksimums, līdz ar to log-ticamības maksimizēšana dod to pašu rezultātu, ko parastās ticamības maksimizēšana. Bieži vien strādāt ar log-ticamības funkciju ir vienkāršāk.

Piemērs 1. Pieņemsim, ka $X_1, \dots, X_n \sim \text{Bernoulli}(p)$. Varbūtības funkcija ir $f(x|p) = p^x(1-p)^{1-x}$, kur $x = 0, 1$. Nezināmais parametrs ir p . Tad

$$L_n(p) = \prod_{i=1}^n f(X_i, p) = \prod_{i=1}^n p^{X_i}(1-p)^{1-X_i} = p^S(1-p)^{n-S},$$

kur $S = \sum_i X_i$. Līdz ar to

$$\ell_n(p) = S \log p + (n - S) \log(1 - p).$$

Nemot $\ell_n(p)$ atvasinājumu un pielīdzinot to 0, lai atrastu vislielākās ticamības novērtējumu $\hat{p} = S/n$.

Regularitātes nosacījumi

1. Gadījuma lielumi X_1, \dots, X_n ir neatkarīgi un vienādi sadalīti ar $f(x|\theta)$;
2. Parametru telpa Θ ir kompakta;
3. Īstā bet nezināmā parametra θ_0 vērtība ir $\theta_0 = \arg \max_{\theta \in \Theta} E_{\theta_0} \log f(X_i|\theta)$;
4. Logaritmiskā ticamības funkcija $\ell_n(\theta)$ ir nepārtraukta pēc θ ;
5. Eksistē $E_{\theta_0} \log f(X_i|\theta)$;
6. $\sup_{\theta \in \Theta} |n^{-1}L_n(\theta) - E_{\theta_0} \log f(X_i|\theta)| < \delta$ gandrīz droši kādai $\delta > 0$. [10]

Pie regularitātes nosacījumiem, vislielākās ticamības novērtējumam $\hat{\theta}_n$ piemīt vairākas īpašības, kas to padara par labu novērtējuma izvēli. Galvenās tā īpašības ir:

1. Tas ir būtisks: $\hat{\theta}_n \xrightarrow{P} \theta_*$, kur ar θ_* apzīmē parametra θ īsto vērtību;
2. Ja $\hat{\theta}_n$ ir vislielākās ticamības novērtējums parametram θ , tad $g(\hat{\theta}_n)$ ir vislielākās ticamības novērtējums $g(\theta)$;
3. Vislielākās ticamības novērtējums ir asimptotiski normāli: $\sqrt{n}(\hat{\theta} - \theta_*)/\hat{s.e.} \rightarrow N(0, 1)$;
4. Tas ir asimptotiski optimāls jeb efektīvs: tas nozīmē to, ka no visiem labajiem parametra θ novērtējumiem, vislielākās ticamības novērtējumam ir vismazākā dispersija pie liela datu apjoma. [15]

Tā kā tālākajās nodaļās tiks pievērsta uzmanība asmoptotiskajai normalitātei, tad šo tēmu aplūkosim nedaudz tuvāk.

1.2. Asimptotiskā normalitāte

Definīcija 3. Krosvalidācijas (*score*) funkcija tiek definēta šādi:

$$s(X|\theta) = \frac{\partial \log f(x|\theta)}{\partial \theta}. \quad (1.2)$$

Definīcija 4. Par Fišera informāciju sauc

$$I_n(\theta) = \mathbb{V}_\theta \left(\sum_{i=1}^n s(X_i|\theta) \right) = \sum_{i=1}^n \mathbb{V}_\theta(s(X_i|\theta)). \quad (1.3)$$

Var pierādīt, ka $\mathbb{E}_\theta(s(X|\theta)) = 0$. No tā seko $\mathbb{V}_\theta(s(X|\theta)) = \mathbb{E}_\theta(s^2(X|\theta))$.

Teorēma 1. $I_n(\theta) = nI(\theta)$ un

$$I(\theta) = -\mathbb{E}_\theta \left(\frac{\partial^2 \log f(X|\theta)}{\partial^2 \theta^2} \right) = - \int \left(\frac{\partial^2 \log f(x|\theta)}{\partial^2 \theta^2} \right) f(x|\theta) dx. \quad (1.4)$$

Teorēma 2. (Vislielākās ticamības funkcijas novērtējumu asimptotiskā normalitāte.) Pie regularitātes nosacījumiem ir spēkā sekojošais:

1. Ja $se = \sqrt{1/I_n(\theta)}$, tad

$$\frac{\hat{\theta}_n - \theta}{se} \rightarrow N(0, 1). \quad (1.5)$$

2. Ja $\hat{se} = \sqrt{1/I_n(\hat{\theta}_n)}$, tad

$$\frac{\hat{\theta}_n - \theta}{\hat{se}} \rightarrow N(0, 1). \quad (1.6)$$

Teorēmas pirmā daļa apgalvo, ka $\hat{\theta}_n \approx N(\theta, se)$, kur $\hat{\theta}_n$ standartnovirze ir $se = \sqrt{1/I_n(\theta)}$. Otrā teorēmas daļa apgalvo, ka tas joprojām ir spēkā, ja standartnovirzi aizstāj ar tās novērtējumu $\hat{se} = \sqrt{1/I_n(\hat{\theta}_n)}$. Būtībā maksimālās ticamības novērtējumu aptuveni atbilst $N(\theta, \hat{se}^2)$. To var izmantot, lai konstruētu asimptotisku ticamības intervālu.

Teorēma 3. Pieņemsim, ka

$$C_n = (\hat{\theta}_n - z_{\alpha/2} \hat{se}, \hat{\theta}_n + z_{\alpha/2} \hat{se}).$$

Tad $\mathbb{P}_\theta(\theta \in C_n) \rightarrow 1 - \alpha$ ja $n \rightarrow \infty$.

Pierādījums. Pieņemsim, ka Z apzīmē parasto normālo gadījuma lielumu. Tad

$$\begin{aligned} \mathbb{P}_\theta(\theta \in C_n) &= \mathbb{P}_\theta(\hat{\theta}_n - z_{\alpha/2} \hat{se} \leq \theta \leq \hat{\theta}_n + z_{\alpha/2} \hat{se}) \\ &= \mathbb{P}_\theta \left(-z_{\alpha/2} \leq \frac{\hat{\theta}_n - \theta}{\hat{se}} \leq z_{\alpha/2} \right) \\ &\rightarrow \mathbb{P}(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha. \end{aligned}$$

□

Ja $\alpha = 0.05$ un $z_{\alpha/2} = 1.96 \approx 2$, tad $\hat{\theta}_n \pm 2\hat{se}$ ir aptuvenais 95% ticamības intervāls.

Piemērs 2. Pieņemsim, ka $X_1, \dots, X_n \sim \text{Bernoulli}(p)$. Vislielākās ticamības novērtējums ir $\hat{p}_n = \sum_i X_i/n$ un $f(x|p) = p^x(1-p)^{1-x}$, $\log f(x|p) = x \log p + (1-x) \log(1-p)$, $s(X|p) = (x/p) - (1-x)/(1-p)$, un $-s'(X|p) = (x/p^2) + (1-x)/(1-p)^2$. Tādējādi

$$I(p) = \mathbb{E}_p(-s'(X|p)) = \frac{p}{p^2} + \frac{(1-p)}{(1-p)^2} = \frac{1}{p(1-p)}.$$

Līdz ar to

$$\hat{s}e = \frac{1}{\sqrt{I_n(\hat{p}_n)}} = \frac{1}{\sqrt{nI(\hat{p}_n)}} = \left\{ \frac{\hat{p}(1-\hat{p})}{n} \right\}^{1/2}.$$

Aptuvenais 95% ticamības intervāls ir

$$\hat{p}_n \pm 2 \left\{ \frac{\hat{p}(1-\hat{p})}{n} \right\}^{1/2}.$$

1.3. Delta metode

Pieņemsim, ka $\tau = g(\theta)$, kur g ir gluda funkcija. Vislielākās ticamības novērtējums τ ir $\hat{\tau} = g(\hat{\theta})$.

Teorēma 4. Ja $\tau = g(\theta)$, kur g ir atvasināma funkcija un $g'(\theta) \neq 0$, tad

$$\frac{\sqrt{n}(\hat{\tau}_n - \tau)}{\hat{s}e(\hat{\tau})} \rightarrow N(0, 1),$$

kur $\hat{\tau}_n = g(\hat{\theta}_n)$ un

$$\hat{s}e(\hat{\tau}_n) = |g'(\hat{\theta}_n)| \hat{s}e(\hat{\theta}_n).$$

Līdz ar to, ja

$$C_n = (\hat{\tau}_n - z_{\alpha/2} \hat{s}e(\hat{\tau}_n), \hat{\tau}_n + z_{\alpha/2} \hat{s}e(\hat{\tau}_n)),$$

tad $\mathbb{P}(\tau \in C_n) \rightarrow 1 - \alpha$, kad $n \rightarrow \infty$.

Piemērs 3. Pieņemsim, ka $X_1, \dots, X_n \sim \text{Bernoulli}(p)$ un $\psi = g(p) = \log(p/(1-p))$. Fišera informācijas funkcija ir $I(p) = 1/(p(1-p))$, tādējādi vislielākās ticamības metodes novērtējuma \hat{p} standartnovirze ir $se = \{\hat{p}(1-\hat{p})/n\}^{1/2}$. ψ vislielākās ticamības novērtējums ir $\hat{\psi} = \log \hat{p}/(1-\hat{p})$. Tā kā $g'(p) = 1/(p/(1-p))$, atbilstoši Deltas metodei

$$\hat{s}e(\hat{\psi}_n) = |g'(\hat{p}_n)| \hat{s}e(\hat{p}_n) = \frac{1}{\sqrt{n\hat{p}_n(1-\hat{p}_n)}}.$$

Aptuvenais 95% ticamības intervāls ir

$$\hat{\psi}_n \pm \frac{2}{\sqrt{n\hat{p}_n(1-\hat{p}_n)}}.$$

2. EMPĪRISKĀS TICAMĪBAS FUNKCIJAS METODE

Empīriskās ticamības funkcijas metode ir neparametriska statistikas metode, kas ļauj izmantot ticamības metodes nepieņemot, ka dati atbilst kādam zināmam sadalījumam. [11]

2.1. Neparametriskā vislielākās ticamības funkcija

Gadījuma lielumam $X \in \mathbb{R}$ kumulatīvā sadalījuma funkcija ir funkcija $F(x) = P(X \leq x)$, kur $-\infty < x < \infty$. Apzīmēsim $P(X = x) = F(x) - F(x-)$ ar $F(-x)$. [11]

Definīcija 5. X_1, \dots, X_n . Tad empīriskā kumulatīvā sadalījuma funkcija jeb empīriskā sadalījuma funkcija ir

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}, \quad (2.1)$$

kur $-\infty < x < \infty$ un I ir indikatorfunkcija.

Definīcija 6. Doti neatkarīgi un vienādi gadījuma lielumi X_1, \dots, X_n ar sadalījuma funkciju F_0 . Funkcijas F neparametriskā ticamības funkcija tiek definēta kā

$$L(F) = \prod_{i=1}^n (F(X_i) - F(X_i-)). \quad (2.2)$$

Piezīme 1. Ja F ir nepārtraukts sadalījums, $L(F) = 0$. Lai neparametriskās ticamības funkcija būtu pozitīva, sadalījuma funkcijai F jāuzliek pozitīva varbūtība visiem pētāmās datu kopas punktiem.

Teorēma 5. X_1, \dots, X_n ir neatkarīgi gadījuma lielumi ar sadalījuma funkciju F_0 . F_n ir to empīriskā sadalījuma funkcija un F ir kāda sadalījuma funkcija. Ja $F \neq F_n$, tad $L(F) < L(F_n)$.

Pierādījums. Pieņemsim, ka $z_1 < z_2 < \dots < z_m$ ir atšķirīgas X_1, \dots, X_n vērtības, $n_j \geq 1$ ir X_i skaits, kas sakrīt ar z_j , $p_j = F(z_j) - F(z_j-)$ un $\hat{p}_j = n_j/n$. Ja $p_j = 0 \forall j = 1, \dots, m$, tad $L(F) = 0 < L(F_n)$, līdz ar to, pieņemot, ka $\forall p_j > 0$ un ka vismaz vienam j $p_j \neq \hat{p}_j$, $\log(x) \leq x - 1 \forall x > 0$, kad $x = 1$. Tāpēc

$$\log\left(\frac{L(F)}{L(F_n)}\right) = \sum_{j=1}^m n_j \log\left(\frac{p_j}{\hat{p}_j}\right) = n \sum_{j=1}^m \hat{p}_j \log\left(\frac{p_j}{\hat{p}_j}\right) < n \sum_{j=1}^m \hat{p}_j \left(\frac{p_j}{\hat{p}_j} - 1\right) \leq 0,$$

kas dod nepieciešamo rezultātu: $L(F) < L(F_n)$. \square

Šī teorēma pierāda to, ka empīriskā sadalījuma funkcija maksimizē neparametrisko ticamības funkciju. [14]

2.2. Neparametriskās ticamības funkcijas attiecības tests

Neparametriskās ticamības funkcijas gadījumā var izmantot tās attiecību hipotēžu pārbaudes testos un ticamības intervālu iegūšanā. Sadalījumam F neparametriskās ticamības funkcijas attiecība izskatās šādi:

$$R(F) = \frac{L(F)}{L(F_n)},$$

kur $L(F)$ definēts (2.2). Pieņemsim, ka mūs interesē parametrs $\theta = T(F)$ kādai sadalījuma funkcijai T . F pieder kopai \mathcal{F} , dažos gadījumos var būt visu \mathbb{R} sadalījumu kopa.

Definīcija 7. Par profila ticamības attiecības funkciju sauc

$$R(\theta) = \sup\{R(F) | T(F) = \theta, F \in \mathcal{F}\}.$$

Empīriskās ticamības hipotēžu pārbaude noraida $H_0 : T(F_0) = \theta_0$, ja $\mathcal{R}(\theta_0) < r_0$, kur r_0 ir konstante. Empīriskās ticamības reģioni ir formā

$$\{\theta | \mathcal{R}(\theta) \geq r_0\}.$$

Bieži vien r_0 var noteikt ar Vilksa teorēmas 7 neparametrisko versiju.

2.3. Empīriskās ticamības metodes vispārīgais gadījums

Pieņemsim, ka X_1, \dots, X_n ir neatkarīgi un vienādi sadalīti gadījuma lielumi ar nezināmu sadalījuma funkciju F , kurai ir p -dimensionāls parametrs θ , un informācija par θ un F ir dota $r \geq p$ funkcionāli neatkarīgu nenovirzītu funkciju formā. [9] [14] Mēs maksimizēsim funkciju $L(F) = \prod_{i=1}^n p_i$ pie sekojošiem ierobežojumiem

$$p_i \geq 0, \sum_i p_i = 1, \sum_i p_i g(X_i, \theta) = 0, \quad (2.3)$$

kur $p_i = P(X = x_i)$.

Dotam θ eksistē unikāls maksimums, ja 0 ir iezliektajā čaulā, kura sastāv no $g(X_1, \theta), \dots, g(X_n, \theta)$. Maksimumu var atrast ar Lagranža reizinātāju palīdzību.

$$H = \sum_i \log p_i + \lambda_0 \left(1 - \sum_i p_i\right) - n\lambda^\tau \sum_i p_i g(X_i, \theta), \quad (2.4)$$

kur λ_0 un $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^\tau$ ir Lagranža reizinātāji. Atvasinot izteiksmi (2.4) pēc p_i , iegūst

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

un

$$p_i = \left(\frac{1}{n}\right) \frac{1}{1 + \lambda^\tau g(X_i, \theta)}. \quad (2.5)$$

Ņemot vērā 3. ierobežojumu no (2.3), iegūst

$$0 = \sum_i p_i g(X_i, \theta) = \frac{1}{n} \sum_i \frac{1}{1 + \lambda^\tau g(X_i, \theta)} g(X_i, \theta),$$

no kā var secināt, ka λ var noteikt kā funkciju no θ .

Piezīme 2. Nepieciešams, lai $0 \leq p_i \leq 1$, kas nozīmē to, ka λ un θ jāizpildās $1 + \lambda^\tau g(x_i, \theta) \geq 1/n \forall i$.

Fiksētam θ pieņem, ka $D_0 = \{\lambda : 1 + \lambda^\tau g(X_i, \theta) \geq 1/n\}$. D_0 ir slēgts un izliekts, kā arī ierobežots, ja 0 ir izliektajā čaulā, ko veido $g(X_i, \theta)$. Turklāt

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

ir negatīvi definita argumentam λ , nodrošinot to, ka $\sum_{i=1}^n g(X_i, \theta) g^\tau(X_i, \theta)$ ir pozitīvi definita. $\lambda = \lambda(\theta)$ ir nepārtraukti diferencējam funkcija pēc θ . Profila empīriskās ticamības funkcija parametram θ tagad tiek definēta šādi:

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

Attiecīgi logaritmiskā empīriskās ticamības funkcija tiek definēta kā

$$\ell_E(\theta) = \sum_{i=1}^n \log[1 + \lambda^\tau(\theta) g(X_i, \theta)]. \quad (2.6)$$

Lai iegūtu parametra θ novērtējumu $\tilde{\theta}$, varam minimizēt $\ell_E(\theta)$, un ar $\tilde{\theta}$ palīdzību iespējams novērtēt p_i no (2.5), kā arī iegūt sadalījuma funkcijas F novērtējumu

$$\tilde{F}_n(x) = \sum_{i=1}^n \tilde{p}_i I(X_i < x). \quad (2.7)$$

Ja $r = p$, tad $\tilde{\theta} = \hat{\theta}$ minimizē $\ell_E(\theta)$, kur $\hat{\theta}$ ir novērtējošā vienādojuma $\sum_{i=1}^n g(X_i, \theta) = 0$ atrisinājums. Turklāt $\tilde{p}_i = n^{-1}$ un (2.7) ir empīriskā kumulatīvā sadalījuma funkcija.

Lemma 1. ([9] [14]) Pieņemsim, ka $E[g(X, \theta_0) g^\tau(X, \theta_0)]$ ir pozitīvi definita, atvasinājums $\partial g(X, \theta) / \partial \theta$ ir nepārtraukts īstās vērtības θ_0 apkārtnē, $\|\partial g(X, \theta) / \partial \theta\|$ un $\|g(X, \theta)\|^3$ ierobežo kāda integrējama funkcija $G(x)$ šajā apkārtnē, un $E[\partial g(X, \theta_0) / \partial \theta]$ rangs ir p . Tad, ja $n \rightarrow \infty$, ar varbūtību

1 $\ell_E(\theta)$ sasniedz savu minimālo vērtību kādā punktā $\tilde{\theta}$ lodes $\|\theta - \theta_0\| \leq n^{-1/3}$ iekšienē, un $\tilde{\theta}$ un $\tilde{\lambda} = \lambda(\tilde{\theta})$ apmierina

$$Q_{1n}(\tilde{\theta}, \tilde{\lambda}) = 0, \quad Q_{2n}(\tilde{\theta}, \tilde{\lambda}) = 0,$$

kur

$$Q_{1n}(\theta, \lambda) = \frac{1}{n} \sum_i \frac{1}{1 + \lambda^\tau g(X_i, \theta)} g(X_i, \theta),$$

$$Q_{2n}(\theta, \lambda) = \frac{1}{n} \sum_i \frac{1}{1 + \lambda^\tau g(X_i, \theta)} \left(\frac{\partial g(X_i, \theta)}{\partial \theta} \right)^\tau \lambda.$$

Teorēma 6. Pie lemmas 1 nosacījumiem, pieņemsim, ka $\partial^2 g(X, \theta) / \partial \theta \partial \theta^\tau$ ir nepārtraukts pēc θ īstās vērtības θ_0 apkārtnē. Tad, ja $\|\partial^2 g(X, \theta) / \partial \theta \partial \theta^\tau\|$ šajā apkārtnē var ierobežot ar integrējamu funkciju $G(X)$, tad

$$\sqrt{n}(\tilde{\theta} - \theta_0) \rightarrow N(0, V), \quad \sqrt{n}(\tilde{\lambda} - 0) \rightarrow N(0, U),$$

$$\sqrt{n}(\tilde{F}_n(x) - F(x)) \rightarrow N(0, W(x)),$$

kur

$$\tilde{F}_n(x) = \sum_{i=1}^n \tilde{p}_i I(X_i < x), \quad \tilde{p}_i = \left(\frac{1}{n} \right) \frac{1}{1 + \tilde{\lambda}^\tau g(X_i, \tilde{\theta})},$$

$$V = \left[E \left(\frac{\partial g}{\partial \theta} \right)^\tau (E g g^\tau)^{-1} E \left(\frac{\partial g}{\partial \theta} \right) \right]^{-1},$$

$$W(x) = F(x)(1 - F(x)) - B(x) U B^\tau(x),$$

$$B(x) = E\{g(X_i, \theta_0) I_{X_i \leq x}\},$$

$$U = [E(g g^\tau)]^{-1} \left\{ I - E \left(\frac{\partial g}{\partial \theta} \right) V E \left(\frac{\partial g}{\partial \theta} \right)^\tau [E(g g^\tau)]^{-1} \right\}$$

un $\tilde{\theta}$ un $\tilde{\lambda}$ ir asimptotiski nekorolēti.

Teorēma 7. [14] Empīriskās ticamības attiecības statistika hipotēzei $H_0 : \theta = \theta_0$ ir

$$W_E(\theta_0) = 2\ell_E(\theta_0) - 2\ell_E(\tilde{\theta}),$$

kur $\ell_E(\theta)$ ir dota (2.6). Izpildoties lemmas 1 un teorēmas 6 nosacījumiem, $W_E(\theta_0) \rightarrow \chi_p^2$, kad $n \rightarrow \infty$ un H_0 ir spēkā.

Lemmas 1 un teorēmu 6 un 7 rezultāti ļauj novērtēt parametrus un konstruēt ticamības intervālus.

2.4. Ticamības intervālu konstruēšana kvantilēm

X_1, \dots, X_n ir gadījuma izlase, kur X_i ir vienādi sadalīti un neatkarīgi, ar nezināmu sadalījumu $F(x)$ un blīvuma funkciju $f(x)$. [7] [12] Dotam $0 < q < 1$ definēsim q -to kvantili kā $F^{-1}(q) = \inf\{x : F(x) \geq q\}$. Ticamības intervālus konstruēsim izlases kvantilei $\theta_q = F^{-1}(q)$.

Kvantile ir ļoti svarīgs populācijas raksturlielums. Dažos gadījumos pieeja ar kvantiļu palīdzību ir iespējama un noderīga, kad citas metodes nav derīgas. Piemēram, lai novērtētu Koši sadalījuma parametru ar blīvuma funkciju $f(x) = 1/\pi[1 + (x - \mu)^2]$, $-\infty < x < \infty$, izlases vidējā vērtība \bar{X} nav labs novērtējums lokācijas parametram μ . Šajā gadījumā izlases mediāna $\theta_{1/2}$ būtu daudz labāks novērtējums.

Pieņemsim, ka $X_{(1)}, \dots, X_{(n)}$ ir sakārtota izlase un kvantiles θ_0 novērtējums ir $\hat{\theta}_0 = X_{([nq])}$, kur $[nq]$ ir veselā daļa no skaitļa nq .

θ_q sakrīt ar M-vienādojumu [7] [8] [12]

$$\int_{-\infty}^{\infty} \phi(x - \theta) dF(x) = 0,$$

kur

$$\phi(z) = \begin{cases} -1, & \text{ja } z \leq 0, \\ q/(1 - q), & \text{ja } z \geq 0. \end{cases} \quad (2.8)$$

Empīriskās ticamības attiecība parametram θ ir

$$R(\theta_0) = \sup \left\{ \prod_{i=1}^n n p_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \phi(X_i - \theta_0) = 0 \right\}. \quad (2.9)$$

Logaritmiskā attiecība attiecīgi izskatās šādi:

$$\log R(\theta_0) = \sup \sum_{i=1}^n \log p_i + n \log n, \quad (2.10)$$

kur p_i , $i = 1, \dots, n$ apmierina tie paši nosacījumi, kas (2.9). Arī šoreiz var izmantot Lagranža reizinātāju metodi, lai maksimizētu $\sum_{i=1}^n \log p_i$. Maksimums tiek sasniegts, ja

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda(\theta_0) \phi(X_i - \theta_0)}, \quad i = 1, \dots, n, \quad (2.11)$$

kur $\lambda(\theta_0)$ var iegūt no vienādojuma

$$\frac{1}{n} \sum_{i=1}^n \frac{\phi(X_i - \theta_0)}{1 + \lambda(\theta_0) \phi(X_i - \theta_0)} = 0. \quad (2.12)$$

Ja $\theta_0 \in [X_{(1)}, \dots, X_{(n)}]$, tad (2.12) atrisinājums $\lambda(\theta_0)$ ir

$$\lambda(\theta_0) = (q - F_n(\theta_0))/q, \quad (2.13)$$

kur F_n ir empīriskā sadalījuma funkcija (2.1).

Pierādījums. Lai pierādītu, ka $\lambda(\theta_0)$ var aprēķināt pēc (2.13), sākumā (2.12) sadala divās daļās: pirmajā $\phi(z) = 1$ un otrajā $\phi(z) = q(1 - q)$. [1]

$$\begin{aligned} \frac{1}{n} \sum_{i=n}^m \frac{\phi(X_i - \theta)}{1 + \lambda\phi(X_i - \theta)} + \frac{1}{n} \sum_{i=m+1}^n \frac{\phi(X_i - \theta)}{1 + \lambda\phi(X_i - \theta)} &= 0 \\ \frac{1}{n} \sum_{i=n}^m \frac{-1}{1 - \lambda} + \frac{1}{n} \sum_{i=m+1}^n \frac{q(1 - q)^{-1}}{1 + \lambda q(1 - q)^{-1}} &= 0 \\ \frac{m}{n} \frac{-1}{1 - \lambda} + \frac{n - m}{n} \frac{q(1 - q)^{-1}}{1 + \lambda q(1 - q)^{-1}} &= 0 \\ \frac{m}{n} \left(\frac{-1}{1 - \lambda} - \frac{n - m}{n} \frac{q(1 - q)^{-1}}{1 + \lambda q(1 - q)^{-1}} \right) &= - \left(\frac{n - m}{n} \frac{q(1 - q)^{-1}}{1 + \lambda q(1 - q)^{-1}} \right) \\ F_n(\theta) \left(\frac{-1}{1 - \lambda} - \frac{q}{(1 - q) + \lambda q} \right) &= - \frac{q}{(1 - q) + \lambda q} \\ F_n(\theta) \left(\frac{-1}{1 - \lambda} \right) = -q \Rightarrow \lambda &= \frac{q - F_n(\theta)}{q} \end{aligned}$$

□

No tā izriet, ka logaritmisko empīrisko ticamības attiecību var uzrakstīt šādi:

$$\ell(\theta) = 2 \log R(\theta) = 2n \left(F_n(\theta) \log \frac{F_n(\theta)}{q} + (1 - F_n(\theta)) \log \frac{1 - F_n(\theta)}{1 - q} \right). \quad (2.14)$$

Pierādījums. Tāpat kā iepriekšējā pierādījumā, arī šajā sadalīsim vienādojumu divās daļās pēc tā paša principa. [1]

$$\begin{aligned} 2 \sum_{i=1}^n \log \left(1 + \frac{q - F_n(\theta)}{q} \phi(X_i - \theta) \right) \\ = 2 \sum_{i=1}^m \log \left(1 + \frac{q - F_n(\theta)}{q} \phi(-1) \right) + 2 \sum_{i=1+m}^n \log \left(1 + \frac{q - F_n(\theta)}{q} \frac{q}{1 - q} \right) \\ = 2m \log \left(\frac{F_n(\theta)}{q} \right) + 2(n - m) \log \left(1 + \frac{q - F_n(\theta)}{1 - q} \right) \\ = 2n F_n(\theta) \left(\log \left(\frac{F_n(\theta)}{q} \right) - \log \left(1 + \frac{q - F_n(\theta)}{1 - q} \right) \right) + \log \left(1 + \frac{q - F_n(\theta)}{1 - q} \right) \\ = 2n \left(F_n(\theta) \log \left(\frac{F_n(\theta)}{q} \right) + \log \left(1 + \frac{q - F_n(\theta)}{1 - q} \right) (1 - F_n(\theta)) \right) \\ = 2n \left(F_n(\theta) \log \frac{F_n(\theta)}{q} + (1 - F_n(\theta)) \log \frac{1 - F_n(\theta)}{1 - q} \right). \end{aligned}$$

□

$\log R(\theta)$ no (2.14) ir pakāpienveida funkcija ar lēcieniem datu punkts. $\log R(\theta)$ var pieņemt tikai galīga skaita vērtības, tāpēc χ_1^2 aproksimācija nav pārāk precīza. Lai novērstu šo problēmu, var aizstāt $F_n(\cdot)$ ar kādu tās nogludinātu versiju.

Lai nogludinātu $F_n(\cdot)$ ar kodolu metodes palīdzību, pieņemsim, ka joslas platums $h = h_n > 0$, kur $h \rightarrow 0$, ja $n \rightarrow \infty$ [7]. Izvēloties kādu Boreiļa mērojamu funkciju $K(x)$ kā kodolu un definējot $G(t) = \int_{-\infty}^t K(x)dx$, empīrisko sadalījuma funkciju F_n var aizstāt ar nogludināto empīriskā sadalījuma funkciju

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n G\left(\frac{c - X_i}{h}\right). \quad (2.15)$$

Attiecīgi, aizstājot $F_n(\theta)$ ar $\hat{F}_n(\theta)$ (2.14), iegūsim pielāgoto empīriskās ticamības attiecību

$$\hat{\ell}(\theta) = 2n \left(\hat{F}_n(\theta) \log \frac{\hat{F}_n(\theta)}{q} + (1 - \hat{F}_n(\theta)) \log \frac{1 - \hat{F}_n(\theta)}{1 - q} \right). \quad (2.16)$$

Pirms nākamās teorēmas, nepieciešams atzīmēt dažus regularitātes nosacījumus:

1. $f(x) = F'(x)$. Kādam veselam skaitlim $r \geq 2$ θ_0 apkārtnē eksistē f un $f^{(r-1)}$, turklāt tie ir nepārtraukti punktā θ_0 . $f(\theta_0) > 0$,
2. Kodola funkcija $K(\cdot)$ ir ierobežota un tai ir kompakts definīcijas apgabals $[a, b]$. Kātai dekompozīcijai $a = u_0 < u_1 < \dots < u_m = b$, $K(\cdot)$ ir vai nu stingri pozitīva, vai stingri negatīva katrā intervālā (u_{j-1}, u_j) , kur $j = 1, \dots, m$. Pieņemsim, ka

$$\int uK(u)G(u)du = 0,$$

$$\int u^j K(u)du = \begin{cases} 1, & j = 0, \\ 0, & 1 \leq j \leq r_1 - 1, \\ C, & j = r_1, \end{cases}$$

kur C ir kāda galīga konstante un $r_1 \geq r$,

3. $nh / \log n \rightarrow \infty$, nh^2 ir ierobežots, ja $n \rightarrow \infty$.

Teorēma 8. Ja $\hat{\ell}(\theta_0)$ ir definēts pēc (2.16) un spēkā ir visi iepriekš minētie regularitātes nosacījumi. Tad, ja $n \rightarrow \infty$,

$$P(\hat{\ell}(\theta_0) \leq x) - P(\chi_1^2 \leq x) = O(n^{-1}) \quad (2.17)$$

katram fiksētam x . [14]

No šīs teorēmas seko

$$\lim_{n \rightarrow \infty} P\{\theta_0 \in I_{hc}\} = P(\chi_1^2 \leq c),$$

kur $I_{hc} = \{\theta : \hat{\ell}(\theta) \leq c\}$. Ja c ir izvēlēts tā, lai $P(\chi_1^2 \leq c) = \alpha$, tad I_{hc} pārklājuma varbūtība aproksimēs α ar kļūdu $O(n^{-1})$, ja $n \rightarrow \infty$.

3. PARAMETRISKĀS UN EMPĪRISKĀS TICAMĪBAS FUNKCIJU METOŽU KOMBINĀCIJAS

3.1. Ideja

Datu modelēšanā ir dažādas metodes no pilnībā parametriskām, semiparametriskām līdz pat pilnībā neparametriskām. Eksistē arī neskaitāmi veidi, kā kombinēt parametriskās metodes ar neparametriskajām, piemēram, ņemot vidējo svērtu starp parametriskiem un neparametriskiem kvantiļu novērtējumiem. Publikācija [5] interesējas par tiltu starp parametrisku modeli un neparametrisko ticamības attiecības metodi. Abu metožu kombinētā ticamības funkcija tiek konstruēta balstoties uz 1) parasto ticamības funkciju parametriskam modelim, t.i., $L_n(\theta)$, kur n ir datu kopas izmērs; 2) empīriskās ticamības funkciju dotai kontroles parametru kopai, t.i., $R_n(\mu)$, kur interesējošie μ parametri tiek "grūsti cauri" parametriskajam modelim, veidojot $R_n(\mu(\theta))$. HL funkcija $H_n(\theta)$ definēta (3.2) izmantota strādājoša modeļa parametru vektora novērtēšanai; ar $\hat{\theta}_{hl}$ tiek apzīmēts interesējošais maksimālās HL novērtējums. Tas savukārt liek domāt par citu interesējošo lielumu novērtējumiem. Ja ψ ir tāds fokusa parametrs, kas izteikts caur strādājošu modeli kā $\psi = \psi(\theta)$, tad tas tiks novērtēts izmantojot $\hat{\psi}_{hl} = \psi(\hat{\theta}_{hl})$.

Ja strādājošs parametriskais modelis ir pareizs, šie HL novērtējumi zaudē noteiktu daudzumu savas efektivitātes, salīdzinot ar parasto vislielākās ticamības novērtējumu. Tomēr var pamatot, ka ideālos modeļa apstākļos efektivitātes zudums ir mazs, un HL novērtējumi bieži vien pārspēj vislielākās ticamības metodi gadījumos, kad strādājošs modelis nav pareizs. Līdz ar to HL metode drīz varēs piedāvāt parametrisku robustumu, aizsardzību pret modeļa neprecizitāti, aizņemoties empīriskās ticamības funkcijas spēku caur izvēlētajiem kontroles parametriem.

Lai gan šo konstrukciju un metodi var pacelt līdz regresijas modeļiem, daudz praktiskāk ir sākt ar vienkāršāku neatkarīgu, vienādi sadalītu struktūru, lai parādītu galveno ideju un attīstītu teoriju. Tāpēc pieņemsim, ka Y_1, \dots, Y_n ir neatkarīgi un vienādi sadalīti novērojumi ar nezināmu blīvuma funkciju. Mēs vēlamies piemērot datus kādai parametriskai saimei, kur $f_\theta(y) = f(y, \theta)$ ar $\theta = (\theta_1, \dots, \theta_p)^t \in \Theta$, kur Θ ir vaļēja \mathbb{R}^p apakškopa. Datu pielāgošanas (*fitting*) mērķis parasti ir novērtēt konkrētus lielumus $\psi = \psi(f)$, kas saukti par fokusa parametriem, bet ne obligāti pilnībā uzticoties modelim. Publikācijas [5] metode robustu novērtējumu konstruēšanai šiem

fokusa parametriem ietver sevī noteiktus papildus parametrus, kurus sauksim par kontroles parametriem: $\mu_j = \mu_j(f)$, $j = 1, \dots, q$. Šie parametri izvēlēti, lai nodrošinātos pret neprecizitātēm, un var gan ietvert, gan neietvert fokusa parametrus. Pieņem, ka $\mu = (\mu_1, \dots, \mu_q)$ tiek identificēti caur novērtēšanas vienādojumiem (*estimating equations*) $E_f m_j(Y, \mu) = 0$, kur $j = 1, \dots, q$. Tagad aplūkosim

$$R_n(\mu) = \max \left\{ \prod_{i=1}^n (nw_i) : \sum_{i=1}^n w_i = 1, \sum_{i=1}^n w_i m(Y_i, \mu) = 0, \forall w_i > 0 \right\}. \quad (3.1)$$

Šī ir empīriskā ticamības funkcija parametram μ [11]. Atbilstošam $g = (g_1, \dots, g_q)$ var izvēlēties $m(Y, \mu) = g(Y) - \mu$, un empīriskās ticamības metode iegūst ticamības intervālus parametriem $\mu = E_f g_j(Y)$. HL funkcija attiecīgi izskatās šādi:

$$H_n(\theta) = L_n(\theta)^{1-a} R_n(\mu(\theta))^a, \quad (3.2)$$

kur $L_n(\theta) = \prod_{i=1}^n f(Y_i, \theta)$ ir parastā ticamības funkcija, $R_n(\mu)$ ir empīriskā ticamības funkcija parametram μ , kas šeit aprēķināta pie vērtības $\mu(\theta)$, kas ir μ novērtēts pie f_θ , un $a \in [0, 1)$ ir balansa parametrs. Atbilstošais vislielākais HL novērtējums ir $\hat{\theta}_{hl}$, $H_n(\theta)$ maksimizētājs. Ja $\psi = \psi(f)$ ir interesējošais parametrs, tas tiek novērtēts kā $\hat{\psi}_{hl} = \psi(f(\cdot, \hat{\theta}_{hl}))$. Tas nozīmē to, ka vispirms jāizsaka ψ attiecībā uz modeļa parametriem, t.i., $\psi = \psi(f(\cdot, \theta_{hl})) = \psi(\theta)$, un tad jāievieto vislielākajā HL novērtējumā.

Piezīme 3. Vispārīgais paņēmieni 3.2 strādā daudzdimensiju gadījumā ar vektoriem Y_i , līdz ar to funkcija g_j var veidot tā, lai tā atspoguļo kovariācijas. Viendimensionālā gadījumā iespējas ietver variantu $m_j(Y, \mu_j) = I\{Y \leq \mu_j\} - j/q$, kur $j = 1, \dots, q$, kad modelis tiek veidots ar kvantiļu palīdzību.

3.2. Piemēri

Piemērs 4. Pieņem, ka f_θ ir normālā blīvuma funkcija ar parametriem (ξ, σ^2) un $m_j(y, \mu_j) = I\{y \leq \mu_j\} - j/4$, $j = 1, 2, 3$. Tad 3.1 ar sekojošu $\mu_j(\xi, \sigma^2) = \xi + \sigma F^{-1}(j/4)$, $j = 1, 2, 3$, var uztvert kā normālā sadalījuma parametru novērtēšanu, modificējot parasto vislielākās ticamības metodi ņemot vērā vēlmi labi pielāgot modeli trim kvartilēm. Alternatīvi uz to var skatīties kā darbošanos ar kvartilēm, aizņemoties spēku no normālās sadalījumu saimes, lai iegūtu labāku rezultātu, nekā izmantojot tikai trīs empīriskās kvartiles [5]. Normālā sadalījuma (ξ, σ^2) gadījumā, mēs varētu gribēt, lai izvēlētajam modelim ne tikai vidējā vērtība un standartnovirze būtu labi novērtēta, bet arī to, lai $\hat{\xi} - 0.675\hat{\sigma}$, $\hat{\xi} + 0.675\hat{\sigma}$ saprātīgās robežās atbilstu kvartilēm $F_n^{-1}(\frac{1}{4})$, $F_n^{-1}(\frac{3}{4})$. Kā arī mēs varētu gribēt, lai modelis dod labu novērtējumu tieši fokusa parametram

$\psi = \psi(f)$ caur $\hat{\psi}_{HL} = \psi(f(\cdot, \hat{\theta}_{HL}))$. Tad, ņemot $m(y, \mu) = I\{y \leq \mu\} - 3/4$ un izmantojot $\hat{\xi}_{HL} + 0.675\sigma_{HL}$, var novērtēt $F^{-1}(\frac{3}{4})$. [6]

Piemērs 5. Pieņem, ka f_θ pieder pie beta sadalījumu saimes ar parametriem (b, c) , kur vislielākās ticamības metode novērtē $\log Y$ un $\log(1-Y)$ momentus. Pievieno tam funkcijas $m_j(y, \mu_j) = y^j m$, $j = 1, 2$. Šoreiz tā ir betas pielāgošana ar modifikācijām, lai iegūtu pareizu vedējo vērtību un dispersiju. Momentu novērtēšana aizņemoties spēku no betas sadalījumu saimes. [5]

Piemērs 6. Izvēlas jebkādu parametrisko sadalījumu saimi $f(y, \theta)$ un atbilstoši datu kopai precizē nozīmīgu intervālu vai reģionu A . Tad izmanto $m(y, p) = I\{y \in A\} - p$ kā kontroles vienādojumu, kur $p = P\{Y \in A\} = \int_A f(y, \theta) dy$. Efektes būs tāds, ka parametriskās maksimālās ticamības novērtējumi tiks "pagrūsti", viegli vai ne tik viegli, atkarībā no parametra a lieluma, lai parūpētos par to, ka empīriskais binomiālais novērtējums $\hat{p} = n^{-1} \sum_{i=1}^n I\{Y_i \in A\}$ nebūtu tālu no novērtētā $p(A, \hat{\theta}) = \int_A f(y, \hat{\theta}) dy$. Šo ideju var paplašināt izmantojot sadalītu A_1, \dots, A_k ar kontroles vienādojumu $m_j(y, p) = I\{y \in A_j\} - p_j$, $j = 1, \dots, k - 1$. [5]

4. HL NOVĒRTĒJUMA ASIMPTOTISKĀS ĪPAŠĪBAS

4.1. HL metodes pamatīpašības

Šīs nodaļas mērķis ir iepazīties ar HL novērtējuma asimptotiskajām īpašībām pie parametriska modeļa $f(\cdot) = f(\cdot, \theta_0)$ atbilstošam un patiesam θ_0 . HL novērtējums $\hat{\theta}_{\text{hl}}$ maksimizē

$$h_n(\theta) = \log H_n(\theta) = (1 - a)\ell_n(\theta) + a \log R_n(\mu(\theta)) \quad (4.1)$$

pie θ (pieņemot, ka tas ir unikāls), kur $\ell_n(\theta) = \log L_n(\theta)$. Aplūkosim $h_n(\cdot)$ abu daļu lokālo uzvedību.

Ievēro lokālo empīrisko ticamības attiecību $R_n(\mu(\theta_0 + s/\sqrt{n}))$, kur s pieder kādai patvaļīgai kompakatai kopai $S \subset \mathbb{R}^p$. Pieraksta vienkāršošanai ņemsim $m_{i,n}(s) = m(Y_i, \mu(\theta_0 + s/\sqrt{n}))$. Jāņem vērā arī funkcijas

$$G_n(\lambda, s) = \sum_{i=1}^n 2 \log\{1 + \lambda^t m_{i,n}(s)/\sqrt{n}\},$$

$$G_n^*(\lambda, s) = 2\lambda^t V_n(s) - \lambda^t W_n(s)\lambda$$

q -dimensionālam λ , kur

$$V_n(s) = n^{-1/2} \sum_{i=1}^n m_{i,n}(s) \quad \text{un} \quad W_n(s) = n^{-1} \sum_{i=1}^n m_{i,n}(s)m_{i,n}(s)^t.$$

Piezīme 4. G_n^* ir G_n izvirzījums Teilora rindā.

Tagad izteiksim empīriskās ticamības statistiku ar Lagranža reizinātāju palīdzību $\hat{\lambda}_n$:

$$-2 \log R_n(\mu(\theta_0 + s/\sqrt{n})) = \max_{\lambda} G_n(\lambda, s) = G_n(\hat{\lambda}_n(s), s)$$

ar $\hat{\lambda}_n(s)$ atrisinājumu vienādojumam

$$\sum_{i=1}^n \frac{m_{i,n}(s)}{1 + \lambda^t m_{i,n}(s)\sqrt{n}} = 0 \quad \forall s$$

Sekojošā lemma ir ļoti svarīga, lai saprastu HL metodes pamatīpašības. Ar $A = (a_{j,k})$, $\|A\| = (\sum_{j,k} a_{j,k}^2)^{1/2}$ apzīmēsim Eiklīda normu.

Lemma 2. $S \subset \mathbb{R}^p$ ir kompakta. Pieņem, ka

(i) $\sup_{s \in S} \|V_n(s)\| = O_{pr}(1)$;

(ii) $\sup_{s \in S} \|W_n(s) - W\| \rightarrow_{pr} 0$, kur $W = \text{Var } m(Y, \mu(\theta_0))$ ir pilna ranga;

(iii) $n^{-1/2} \sup_{s \in S} \max_{i \leq n} \|m_{i,n}(s)\| \rightarrow_{pr} 0$.

Tad maksimizētāji $\hat{\lambda}_n(s) = \text{argmax}_{\lambda} G_n(\lambda, s)$ un $\lambda_n^*(s) = \text{argmax}_{\lambda} G_n^*(\lambda, s) = W_n^{-1}(s)V_n(s)$ abi ir $O_{pr}(1) \forall s \in S$, un $\sup_{s \in S} |\max_{\lambda} G_n(\lambda, s) - \max_{\lambda} G_n^*(\lambda, s)| = \sup_{s \in S} |G_n(\hat{\lambda}_n(s), s) - G_n^*(\lambda_n^*(s), s)| \rightarrow_{pr} 0$.

Pierādījums. $G_n(\cdot, s)$ maksimizētāju pieraksta kā $\hat{\lambda}_n(s) = \| \hat{\lambda}_n(s) \| u(s)$, kur $u(s)$ ir vienības garuma vektors.

$$\| \hat{\lambda}_n(s) \| \{ u(s)^t W_n(s) u(s) - E_n(s) u(s)^t V_n(s) \} \leq u(s)^t V_n(s),$$

kur $E_n(s) = n^{-1/2} \max_{i \leq n} \|m_{i,n}(s)\|$, kas pēc varbūtības vienādi visiem s tiecas uz nulli pēc pieņēmuma (iii). No pieņēmuma (i) seko, ka $\sup_{s \in S} \|u(s)^t V_n(s)\| = O_{pr}(1)$. Turklāt

$$u(s)^t W_n(s) u(s) \geq e_{n,\min}(s)$$

ir mazākā $W_n(s)$, īpašvērtība, kas pēc varbūtības konverģē uz mazāko W īpašvērtību, tas ir norobežots no nulles caur (ii) pieņēmumu.. No tā seko $\| \hat{\lambda}_n(s) \| = O_{pr}(1) \forall s$. Arī $\lambda_n^*(s) = W_n(s)^{-1} V_n(s)$ ir ierobežots pēc varbūtības $\forall s$. Caur $\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 h(x)$, kur $|h(x)| \leq 2, |x| \leq \frac{1}{2}$ var uzrakstīt

$$G_n(\lambda, s) = 2\lambda^t V_n(s) - \frac{1}{2} \lambda^t W_n(s) \lambda + r_n(\lambda, s) = G_n^*(\lambda, s) + r_n(\lambda, s).$$

Tad patvaļīgam $c > 0$ aplūko jebkuru λ , kam $\|\lambda\| \leq c$ un varam noteikt

$$\begin{aligned} |r_n(\lambda, s)| &\leq \frac{2}{3} \sum_{i=1}^n |\lambda^t m_{i,n}(s) / \sqrt{n}|^3 |h(\lambda^t m_{i,n}(s) / \sqrt{n})| \\ &\leq \frac{4}{3} E_n(s) \|\lambda\| \lambda^t W_n(s) \lambda \leq \frac{4}{3} E_n(s) c^3 e_{n,\max}(s) \end{aligned}$$

no lielākās $W_n(s)$ īpašvērtības, kamēr $c E_n(s) \leq \frac{1}{2}$. Izvēlas pietiekami lielu c , lai gan $\hat{\lambda}(s)$, gan $\lambda_n^*(s)$ atrastos šajā lodē, ko veido visi s ar varbūtību $1 - \varepsilon'$ iepriekš izvēlētam mazam ε' . Tad

$$\begin{aligned} &P\left(\sup_{s \in S} |\max_{\lambda} G_n(\lambda, s) - \max_{\lambda} G_n^*(\lambda, s)| \geq \varepsilon\right) \\ &\leq P\left(\sup_{s \in S} \sup_{\|\lambda\| \leq c} |G_n(\lambda, s) - G_n^*(\lambda, s)| \geq \varepsilon\right) \\ &\leq \left((4/3)c^3 \sup_{s \in S} (E_n(s) e_{n,\max}(s)) \geq \varepsilon\right) + P\left(\sup_{s \in S} \| \hat{\lambda}_n(s) \| > c\right) \\ &+ P\left(\sup_{s \in S} \| \lambda_n^*(s) \| > c\right) + P\left(c \sup_{s \in S} E_n(s) > 1/2\right). \end{aligned}$$

Augšējā robeža secīgajām varbūtībām no kreisās puses ir ierobežota ar $4\varepsilon'$. Līdz ar to ir pierādīts, ka $\sup_{s \in S} |\max_{\lambda} G_n(\lambda, s) - \max_{\lambda} G_n^*(\lambda, s)| \rightarrow_{\text{pr}} 0$. \square

Piezīme 5. Mums ir precīza izteiksme G_N^* maksimizētājiem, līdz ar to tā maksimums ir izsakāms šādi:

$$\max_{\lambda} G_n^*(\lambda, s) = V_n(s)^t W_n^{-1}(s) V_n(s).$$

No iepriekš iegūtā seko, ka situācijās, kad spēkā lemma (2),

$$-2 \log R_n(\mu(\theta_0 + s/\sqrt{n})) = V_n(s)^t W_n^{-1}(s) V_n(s) + o_{\text{pr}}(1),$$

vienādi visiem $s \in S$.

Piezīme 6. Pēc lielo skaitļu likuma lemmas (2) nosacījums (ii) ir spēkā $\forall s$, ja $\sup_s \|W_n(s) - W_n(0)\| \rightarrow_{\text{pr}} 0$.

Nākamajai teorēmai nepieciešams pieņēmums par funkciju $m(y, \mu)$ no izteiksmes 3.1, un par to, kā $\mu = \mu(f_{\theta}) = \mu(\theta)$ uzvedas θ_0 tuvumā. Lai $Em(Y, \mu(\theta_0)) = 0$, pieņemam, ka

$$\sup_{s \in S} \|V_n(s) - V_n(0) - \xi_n s\| = o_{\text{pr}}(1), \quad (4.2)$$

kur ξ_n ar dimensiju $q \times p$ pēc varbūtības tiecas uz ξ_0 .

Pārrakstīsim logaritmisko HL funkciju kā lokālu $1/\sqrt{(n)}$ mērojamu perturbāciju ap θ_0 :

$$\begin{aligned} A_n(s) &= h_n(\theta_0 + s/\sqrt{(n)}) - h_n(\theta_0) \\ &= (1 - a)\{\ell(\theta_0 + s/\sqrt{n}) - \ell_n(\theta_0)\} + a\{\log R_n(\mu(\theta_0 + s/\sqrt{n})) - \log R_n(\mu(\theta_0))\} \end{aligned} \quad (4.3)$$

Pēc daudzddimensiju centrālās robežteorēmas

$$\begin{pmatrix} U_{n,0} \\ V_{n,0} \end{pmatrix} = \begin{pmatrix} n^{-1/2} \sum_{i=1}^n u(Y_i, \theta_0) \\ n^{-1/2} \sum_{i=1}^n m(Y_i, \mu(\theta_0)) \end{pmatrix} \rightarrow_d \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} \sim N_{p+q}(0, \Sigma), \quad \text{kur } \Sigma = \begin{pmatrix} J & C \\ C^t & W \end{pmatrix}. \quad (4.4)$$

Šeit $u(y, \theta) = \partial \log f(y, \theta) / \partial \theta$ ir krosvalidācijas (*score*) funkcija, $J = \text{Var } u(Y, \theta_0)$ ir Fišera informācijas matrica ar dimensiju $p \times p$, $C = Eu(Y, \theta)m(Y, \mu(\theta_0))^t$ ar dimensiju $p \times q$, un $W = \text{Var } m(Y, \mu(\theta_0))$ tāpat kā iepriekš. $(p + q) \times (p + q)$ dispersiju matrica Σ ir pozitīvi definēta pēc pieņēmuma. Tas nodrošina to, ka parametriksā un empīriskā ticamības funkcija nepārklājas savā starpā.

Teorēma 9. Pieņem, ka $f(y, \theta)$ gluduma nosacījumi ir spēkā (precīzāk aprakstīts pierādījumā); arī lemmas 2 nosacījumi ir spēkā kopā ar (4.2) nosacījumu pie atbilstoša ξ_0 jebkurai kompaktai

$S \subset \mathbb{R}^p$; Σ ir pilna ranga matrica. Tad katrai kompaktai S , $A_n(s)$ vāji konverģē uz $A(s) = s^t U^* - \frac{1}{2} s^t J^* s$ funkciju telpā $\ell^\infty(s)$, kur

$$U^* = (1 - a)U_0 - a\xi_0^t W^{-1} V_0 \quad \text{un} \quad J^* = (1 - a)J + a\xi_0^t W^{-1} \xi_0.$$

Šeit $U^* \sim N_p(0, K^*)$ ar dispersiju matricu

$$K^* = (1 - a)^2 J + a^2 \xi_0^t W^{-1} \xi_0 - a(1 - a)(C W^{-1} \xi_0 + \xi_0^t W^{-1} C^t).$$

Pierādījums. Mēs strādāsim ar divām komponentēm no (4.3) atsevišķi. Vispirms ar $U_n = n^{-1/2} \sum_{i=1}^n u(Y_i, \theta_0)$, kas tiecas uz $U_0 \sim N_p(0, J)$ (4.4):

$$\ell_n(\theta_0 + s/\sqrt{n}) - \ell_n(\theta_0) = s^t U_n - \frac{1}{2} s^t J s + \varepsilon_n(s), \quad \text{kur} \quad \sup_{s \in S} |\varepsilon_n(s)| \rightarrow_{\text{pr}} 0, \quad (4.5)$$

pie dažādiem mērenās regularitātes nosacījumiem. Ja $\log f(y, \theta)$ ir ieliekts, citi nosacījumi nav nepieciešami, ja neskaita galīgumu un Fišera informācija matricu J . Bez ieliektības, pieņemot, ka eksistē trešās kārtas atvasinājumi $D_{i,j,k}(y, \theta) = \partial^3 \log f(y, \theta) / \partial \theta_i \partial \theta_j \partial \theta_k$, ar Teilora rindas palīdzību iespējams pārbaudīt (4.5) pie nosacījuma, ka $\sup_{\theta \in \mathcal{N}} \max_{i,j,k} |D_{i,j,k}(Y, \theta)|$ ir galīga vidējā vērtība ar \mathcal{N} apkārtni ap θ_0 . Šis nosacījums parasti ir sastopams pie visām parametriskajām saimēm. Jāatzīmē, ka (4.5) var tikt izveidots bez trešās pakāpes atvasinājumiem, ar mērenu nepārtrauktības nosacījumu pie otrās pakāpes atvasinājumiem.

Tālāk redzams, ka var izmantot lemmu 2, ietverot

$$\log R_n(\mu(\theta_0 + s/\sqrt{n})) = -\frac{1}{2} V_n(s)^t W_n(s)^{-1} V_n(s) + o_{\text{pr}}(1), \quad (4.6)$$

$\forall s \in S$. Lai tas būtu spēkā, jāpārbauda lemmas 2 nosacījums (i) (mēs pieņemām (ii) un (iii) par spēkā esošiem). Šeit (i) seko izmantojot (4.2), jo

$$\sup_s \|V_n(s)\| = \sup_s \|V_n(0) + \xi_n s\| + o_{\text{pr}}(1) \rightarrow_d \sup_s \|V_0 + \xi_0 s\|.$$

Līdz ar to $\sup_s \|V_n(s)\| = O_{\text{pr}}(1)$

No iepriekš iegūtā

$$\begin{aligned} \log R_n(\mu(\theta_0 + s/\sqrt{n})) - \log R_n(\mu(\theta_0)) &\rightarrow_d -\frac{1}{2} (V_0 + \xi_0 s)^t W^{-1} (V_0 + \xi_0 s) + \frac{1}{2} V_0^t W^{-1} V_0 \\ &= -V_0^t W^{-1} \xi_0 s - \frac{1}{2} s^t \xi_0^t W^{-1} \xi_0 s. \end{aligned}$$

Šī konverģence notiek arī ar (4.5), ņemot vērā (4.4), ar ko mēs arī nonākam pie nepieciešamā rezultāta. \square

Pieņemot, ka μ ir atvasināms punktā $\theta = \theta_0$, C var pārrakstīt izmantojot parciālo atvasinājumu vektoru $\frac{\partial \mu}{\partial \theta}$ ar atvasinājumiem punktā θ_0 .

$$C = -\frac{\partial \mu}{\partial \theta} E_{\theta} m^*(Y, \mu(\theta_0)),$$

kur $m^*(y, \mu) = \partial m(y, \mu) / \partial \mu$, ja m ir atvasināma pēc μ . Ja, piemēram, $m(y, \mu) = g(y) - \mu$, atsaucoties uz $\mu = Eg(Y)$, iegūst $C = \frac{\partial \mu}{\partial \theta}$. Izmantojot (4.2), $\xi_0 = -\left(\frac{\partial \mu}{\partial \theta}\right)^t$ ar dimensiju $q \times p$. Teorēma 9 dod $U^* = (1 - a)U_0 + a\frac{\partial \mu}{\partial \theta} W^{-1} V_0$ kopā ar

$$J^* = (1 - a)J + a\frac{\partial \mu}{\partial \theta} W^{-1} \left(\frac{\partial \mu}{\partial \theta}\right)^t \quad \text{un} \quad K^* = (1 - a)^2 J + \{1 - (1 - a)^2\} \frac{\partial \mu}{\partial \theta} W^{-1} \left(\frac{\partial \mu}{\partial \theta}\right)^t \quad (4.7)$$

4.2. Parametru ticamības intervāli

Nākamā teorēma ir teorēmas 9 sekas, kam nepieciešama funkcija $\Gamma_n(\theta) = n^{-1}\{h_n(\theta) - h_n(\theta_0)\}$ kopā ar tās versiju populācijai

$$\Gamma(\theta) = -(1 - a)\text{KL}(f_{\theta_0}, f_{\theta}) - aE \log(1 + \xi(\theta)^t m(Y, \mu(\theta))); \quad (4.8)$$

Piezīme 7. $\hat{\theta}_{\text{hl}}$ ir $\text{argmax}_{\Gamma_n(\cdot)}$. Šeit $\text{KL}(f, f_{\theta}) = \int f \log(f/f_{\theta})$ ir Kulbaka-Leiblera starpība, šajā gadījumā no f_{θ_0} uz f_{θ} , ar $Em(Y, \mu(\theta)) / \{1 + \xi(\theta)^t m(Y, \mu(\theta))\} = 0$ atrisinājumu $\xi(\theta)$.

Teorēma 10. Pie teorēmas 9 nosacījumiem

- (i) $\hat{\theta}_{\text{hl}}$ ir konsistents tiecoties uz θ_0 ;
- (ii) $\sqrt{n}(\hat{\theta}_{\text{hl}} - \theta_0) \rightarrow_d (J^*)^{-1} U^* \sim N_p(0, (J^*)^{-1} K^* (J^*)^{-1})$
- (iii) $2\{h_n(\hat{\theta}_{\text{hl}}) - h_n(\theta_0)\} \rightarrow_d (U^*)^t (J^*)^{-1} U^*$.

Šie rezultāti ļauj konstruēt θ ticamības reģionus un tā komponentu ticamības intervālus. Tā kā mēs interesējamies par fokusa parametriem, pieņemsim, ka $\psi = \psi(\theta) = \psi(\theta_1, \dots, \theta_p)$ ir tāds parametrs, kuram ψ ir diferencējams punktā θ_0 , un apzīmēsim $c = \partial \psi(\theta_0) / \partial \theta$. HL novērtējums parametram ψ ir *plug-in* novērtējums $\hat{\psi}_{\text{hl}} = \psi(\hat{\theta}_{\text{hl}})$. Ar $\psi_0 = \psi(\theta_0)$ kā īsto parametra vērtību, ar delta metodi iegūst

$$\sqrt{n}(\hat{\psi}_{\text{hl}} - \psi_0) \rightarrow_d c^t (J^*)^{-1} U^* \sim N(0, \kappa^2), \quad \text{kur } \kappa^2 = c^t (J^*)^{-1} K^* (J^*)^{-1} c. \quad (4.9)$$

Fokusa parametrs ψ varētu būt viena no $\mu = \mu(\theta)$ komponentēm izmantota HL modeļa empīriskās ticamības daļā, piemēram, μ_j , kurai $\sqrt{n}(\hat{\mu}_{j,\text{hl}} - \mu_{0,j})$ ir normāla robeža ar dispersiju $\left(\frac{\partial \mu_j}{\partial \theta}\right)^t (J^*)^{-1} K^* (J^*)^{-1} \frac{\partial \mu_j}{\partial \theta}$, kur $\frac{\partial \mu_j}{\partial \theta} = \partial \mu_j(\theta_0) / \partial \theta$. Izmantojot teorēmā 10 iegūtos rezultātus, iespējams iegūt Valda (*Wald*) un ticamības attiecības tipa ticamības intervālus un testus parametram θ un ticamības intervālus parametram ψ . Tad būtu nepieciešami konsekventi novērtējumi \hat{J}^* un \hat{K}^* , kurus iespējams iegūt ar *plug-in* palīdzību.

4.3. Efektivitātes zudums

Lai noskaidrotu, cik daudz tiek zaudēts efektivitātes ziņā, izmantojot HL novērtējumu, jāaplūko gadījums, kad a ir mazs. Mums ir $J^* = J + aA_1$ un $K^* = J + aA_2 + O(a^2)$ ar $A_1 = \xi_0^t W^{-1} \xi_0 - J$ un $A_2 = -2J - CW^{-1} \xi_0 - \xi_0^t W^{-1} C^t$. Funkciju klasei formā $m(y, \mu) = g(y) - T(\mu)$ atbilstoši $\mu = T^{-1}(Eg(Y))$, mums ir $A_2 = 2A_1$. Pieņemsim, ka $T(\cdot)$ eksistē nepārtraukta inversitāte pie $\mu(\theta)$ parametram θ θ_0 apkārtnē. Uzrakstot $(J^*)^{-1} K^* (J^*)^{-1}$ kā

$$(J^{-1} - aJ^{-1}A_1J^{-1})(J + aA_2)(J^{-1} - aJ^{-1}A_1J^{-1}) + O(a^2),$$

var secināt, ka tas ir $J^{-1} + O(a^2)$, kas nozīmē to, ka efektivitātes zudums ir ļoti mazs, ja a ir mazs.

5. HL METODEDE ĀRPUS PARAMETRISKĀ MODEĻA NOSACĪJUMIEM

Iepriekšējā nodaļā mēs aplūkojām HL novērtējuma stratēģiju pie parametriskā modeļa nosacījumiem, kad pie atbilstošiem nosacījumiem, HL metode ir konsekventa un asimptotiski normāla ar noteiktu nelielu efektivitātes zudumu, salīdzinot ar parasto ticamības metodi, t.i., gadījumu, kad $a = 0$. Šajā nodaļā noskaidrosim HL metodes uzvedību ārpus parametriskā modeļa nosacījumiem, kuru tagad sauksim par strādājošu modeli. Izrādās, ka HL metode bieži vien pārspēj vislielākās ticamības metodi, samazinot modeļi novirzi, kas vidējās kvadrātiskās kļūdas izteiksmē kompensē vairāk nekā nelielā mainīguma (*variability*) pieaugums.

Lai izpētītu iepriekš minētās īpašības, nepieciešams paplašināt $f(y, \theta)$ modeli uz $f(y, \theta, \gamma)$, kur $\gamma = (\gamma_1, \dots, \gamma_r)$ ir papildu parametru vektors. Eksistē tāda nulles vērtība $\gamma = \gamma_0$, kas paplašināto modeli pārvērš par iepriekš aplūkoto strādājošo modeli. Aplūkosim vislielākās ticamības un HL shēmas, kad γ atrodas γ_0 apkārtnē. Pieņemsim, ka

$$f_{true}(y) = f(y, \theta_0, \gamma_0 + \delta/\sqrt{n}), \quad (5.1)$$

kur $\delta = (\delta_1, \dots, \delta_r)$ nosaka relatīvo attālumu no nulles modeļa. Pieņemsim arī to, ka $\hat{\theta}$ piemīt šāda īpašība

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow_d N_p(B\delta, \Omega), \quad (5.2)$$

ar atbilstošu $p \times r$ matricu B , kas parāda, cik daudz modeļa novirze itekmē novērtējumu θ , un dispersijas robežas matricu Ω . Tad interesējošais parametrs $\psi = \psi(f)$ var tikt izteikts kā $\psi = \psi(\theta, \gamma)$ ar tā īsto vērtību $\psi_{true} = \psi(\theta_0, \gamma_0 + \delta/\sqrt{n})$. Sekojošais ψ parametra novērtējums līdz ar to ir $\hat{\psi} = \psi(\hat{\theta}, \gamma_0)$. Tad delta metode noved pie

$$\sqrt{n}(\hat{\psi} - \psi_{true}) \rightarrow_d N(b^t\delta, \tau^2), \quad (5.3)$$

kur

$$b = B^t \frac{\partial \psi}{\partial \theta} - \frac{\partial \psi}{\partial \gamma} \quad \text{un} \quad \tau^2 = \left(\frac{\partial \psi}{\partial \theta} \right)^t \Omega \frac{\partial \psi}{\partial \theta},$$

un kur parciālaie atvasinājumi iegūti pie strādājoša modeļa, t.i., punktos θ_0, γ_0 . Ierobežojošā (*limiting*) vidējā kvadrātiskā kļūda novērtējumam μ būtu $mse(\delta) = (b^t\delta)^2 + \tau^2$. Starp ierobežoti

strādājoša modeļa izmantošanas sekām, kad tas ir mēreni nepareizs pie $\gamma = \gamma_0 + \delta/\sqrt{n}$, ir novirze $b^t \delta$.

Piezīme 8. Šīs novirzes apmēri atkarīgi no fokusa parametra, un tā varētu būt 0 pat tad, ja modelis ir nepareizs.

Tagad aplūkosim gan vislielākās ticamības, gan HL metodes iepriekš izveidotajā ietvarā, eksponējot ar to saistītās matricas B un Ω , tādējādi arī vidējo kvadrātisko kļūdu caur (5.3). Vispirms aplūkosim parametriskās vislielākās ticamības metodes novērtējumu $\hat{\theta}_{ml}$. Lai parādītu nepieciešamos rezultātus, aplūko $(p+r) \times (p+r)$ Fišera informācijas matricu

$$J_{wide} = \begin{pmatrix} J_{00} & J_{01} \\ J_{10} & J_{11} \end{pmatrix} \quad (5.4)$$

modelim $f(y, \theta, \gamma)$, kas aprēķināta pie sākuma vērtībām (θ_0, γ_0) . It īpaši $p \times p$ daļu J_{00} , kas atbilst tieši modelim tikai ar θ , bez γ , un ir ekvivalenta iepriekš dotajai matricai J no (4.4). Tieši tāpat kā (5.2) un (5.3), $\hat{\psi}_{ml} = \psi(\hat{\theta}_{ml})$ gadījumā ir spēkā

$$\sqrt{n}(\hat{\psi}_{ml} - \psi_{true}) \rightarrow_d N(\omega^t \delta, \tau_0^2), \quad (5.5)$$

kur

$$\omega = J_{10} J_{00}^{-1} \frac{\partial \psi}{\partial \theta} - \frac{\partial \psi}{\partial \gamma} \quad \text{un} \quad \tau_0^2 = \left(\frac{\partial \psi}{\partial \theta} \right)^t J_{00}^{-1} \frac{\partial \psi}{\partial \theta}.$$

Šajā gadījumā ar interesējošo parametru ψ , izmantosim HL metodi, lai iegūtu $\psi_{hl} = \psi(\hat{\theta}_{hl}, \gamma_0)$. Noskaidrosim, kas notiek ar $\hat{\theta}_{hl}$, vispārinot to, kas tika iegūts iepriekšējā nodaļā. Tālāk esošie vienādojumi nepeieciešami nākamajai teorēmai.

$$S(y) = \partial \log f(y, \theta_0, \gamma_0) / \partial \gamma,$$

ir krosvalidācijas funkcija,

$$K_{01} = \int f(y, \theta_0) m(y, \mu(\theta_0)) S(y) dy$$

ar dimensiju $q \times r$,

$$L_{01} = (1-a) J_{01} - a \left(\frac{\partial \psi}{\partial \theta} \right) W^{-1} K_{01}$$

ar dimensiju $p \times r$ un transponēto matricu $L_{10} = L_{01}^t$.

Teorēma 11. Pieņemsim, ka dati oriģināli nāk no paplašināta $p+r$ dimensiju modeļa (5.1) un teorēmas 10 nosacījumi ir spēkā. HL metodei ar fokusa parametru $\psi = \psi(f)$ ar vienādojuma (3.2) konstrukciju, rezultāti (5.2) un (5.3) ir spēkā ar $B = (J^*)^{-1} L_{01}$ un $\Omega = (J^*)^{-1} K^* (J^*)^{-1}$.

Pierādījums. Lai pierādītu šo teorēmu, atkārtoti jāaplūko $A_n(\cdot) \rightarrow_d A(\cdot)$ no teorēmas 9, tikai tagad nepieciešams tos paplašināt gadījumam ar parametru δ . Pirmkārt,

$$\ell_n(\theta_0 + s\sqrt{n}) - \ell_n(\theta_0) = U_n s - \frac{1}{2} s^t J_n s + o_{pr}(1) \rightarrow_d (U + J_{01}\delta)^t s - \frac{1}{2} s^t J_{00} s.$$

Tas būtībā ir spēkā kopš noskaidrots, ka $U_n = n^{-1/2} \sum_{i=1}^n u(Y_i, \theta_0)$ vidējā vērtība ir $J_{01}\delta$, bet dispersija saglabājusies tā pati. Mums nepieciešams paralēls rezultāts $V_{n,0} = n^{-1/2} \sum_{i=1}^n m(Y_i, \mu(\theta_0))$ gadījumā pie f_{true} . Šeit

$$\begin{aligned} E_{true} m(Y, \mu(\theta_0)) &= \int m(y, \mu(\theta_0)) f(y, \theta_0) \{1 + S(y)^t \delta / \sqrt{n} + o(1/\sqrt{n})\} dy \\ &= 0 + K_{01}\delta / \sqrt{n} + o(1/\sqrt{n}), \end{aligned}$$

dodot $V_{n,0} \rightarrow_d V_0 + K_{01}\delta$. Tas noved pie nepieciešamās $A_n \rightarrow_d A$ daļas no teorēmas 9 un tās pierādījuma, un pēc tam tālāk pie tagadējā lokālās apkārtnes modeļa stāvokļa:

$$A_n(s) = h_n(\theta_0 + s/\sqrt{n}) - h_n(\theta_0) \rightarrow_d A(s) = s^t U_{plus}^* - \frac{1}{2} s^t J^* s,$$

kur

$$U_{plus}^* = (1 - a)(U + J_{01}\delta) - a\xi_0^t W^{-1}(V_0 + K_{01}\delta) = U^* + L_{01}\delta.$$

Ar teorēmas 10 palīdzību, nonākam pie

$$\sqrt{n}(\hat{\theta}_{hl} - \theta_0) \rightarrow_d (J^*)^{-1}(U^* + L_{01}\delta) \sim N_p((J^*)^{-1}K^*(J^*)^{-1}),$$

kas arī bija prasīts. □

Tagad ierobežojošais sadalījums $\hat{\psi}_{hl} = \psi(\hat{\theta}_{hl})$ var tikt secināts gluži kā (5.3), kas seko no (5.2):

$$\sqrt{n}(\hat{\psi}_{hl} - \psi_{true}) \rightarrow_d N(\omega_{hl}^t \delta, \tau_{0,hl}^2), \tag{5.6}$$

kur

$$\omega_{hl} = L_{10}(J^*)^{-1} \frac{\partial \psi}{\partial \theta} - \frac{\partial \psi}{\partial \gamma} \quad \text{un} \quad \tau_{0,hl}^2 = \left(\frac{\partial \psi}{\partial \theta} \right)^t (J^*)^{-1} K^* (J^*)^{-1} \frac{\partial \psi}{\partial \theta}.$$

Piezīme 9. Aplūkotie lielumi, kas apraksta HL novērtējuma īpašības, ir atkarīgi no parametra a no HL metodes pamata konstrukcijas (3.2). Gadījumā, kad $a = 0$, runa ir par parasto vislielākās ticamības metodi ar (5.6) sašaurinoties uz (5.5). Ja a aizvīrās tālāk no nulles, lielāks uzsvars tiek dots empīriskās ticamības funkcijas daļai, lai "pagrūstu" θ tā, lai iegūtu $n^{-1} \sum_{i=1}^n m(Y_i, \mu(\theta))$ tuvāk nullei. Rezultāts parasti ir ar mazāku novirzi $|\omega_{hl}(a)^t \delta|$ nekā $|\omega^t \delta|$, un nedaudz lielāku standartnovirzi $\tau_{0,hl}$ salīdzinot ar τ_0 . Līdz ar to labas a vērtības atrašana kļūst par novirzes un dispersijas balansēšanas spēli, kas tiks aprakstīta nākamajā nodaļā.

6. BALANSA PARAMETRA REGULĒŠANA

Pamata HL metodes konstrukcija (3.2) pirmkārt liek izvēlēties situācijai atbilstošus kontroles parametrus μ , un pēc tam fokusa parametru ψ . Var būt arī speciālgadījums, kurā fokuss ψ ir vienīgais kontroles parametrs. Abos gadījumos vēl ir arī balansa parametrs a , kas jāizvēlas. Tālāk piedāvāti vairāki veidi, kā to izdarīt.

6.1. Robustības un efektivitātes balansēšana

Atļaujot empīriskās ticamības un parametriskās ticamības metožu kombināciju, robustību var iegādāties caur kontroles parametriem μ no HL metodes konstrukcijas par noteikta mērena efektivitātes zuduma cenu. Viens veids kā regulēt balansu pēc kontroles parametru izvēles, ir izraudzīties a tā lai efektivitātes zudums pie strādājoša parametriskā modeļa būtu ierobežots ar fiksētu, nelielu lielumu, piemēram, 5%. To var sasniegt izmantojot 4. nodaļas sekas, salīdzinot inverso Fišera informācijas matricu J^{-1} (asociētu ar vislielākās ticamības novērtējumu) ar matricu $(J_a^*)^{-1} K_a^* (J_a^*)^{-1}$ (HL novērtējumam). Šeit mēs atsaucamies uz 4. nodaļas sekām, piemēram, (4.7), pievienojot a apakšindeksu uzsvaram. Ja ir kāda specifiska interese atiecībā uz kādu fokusa parametru ψ , a var izraudzīties šādi:

$$\kappa_a = \{c^t (J_a^*)^{-1} K_a^* (J_a^*)^{-1} c\}^{-1/2} \leq (1 + \varepsilon) \kappa_0 = (1 + \varepsilon) (c^t J^{-1} c)^{1/2}, \quad (6.1)$$

kur ε ir pieprasītais sliekšnis. Ja $\varepsilon = 0.05$, piemēram, tas nodrošina to, ka ticamības intervāli ir tikai 5% platāki nekā vislielākās ticamības funkcijas gadījumā, bet tas dod arī papildus drošību, ka μ parametri tiek labi kontrolēti procesa laikā, piemēram, robustības iemeslu dēļ. Izteiksmē (4.9) patiesībā arī $\theta_{0,a}$ $c_a = \partial\psi(\theta_{0,a})/\partial\theta$ ir atkarīgs no a , bet runājot par parametriskā modeļa efektivitāti, $\theta_{0,a}$ ir tāds pats kā θ_0 , līdz ar to c_a ir tāds pats kā $c = \partial\psi(\theta_0)/\partial\theta$.

6.2. Vidējās kvadrātiskās kļūdas īpašības

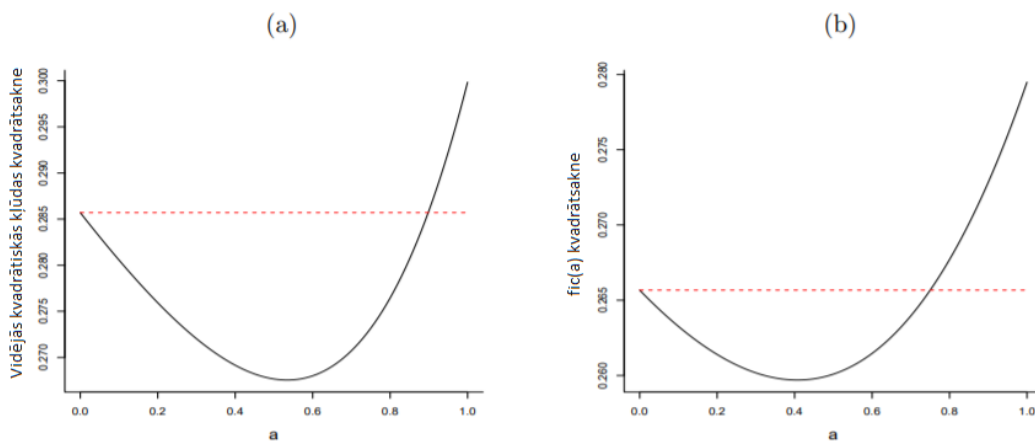
Iepriekšējā metode ar (6.1) ir balstīta uz teoriju no 4. nodaļas pie strādājoša parametriskā modeļa nosacījumiem. Sekojošajam nepieciešama teorija no 5. nodaļas, kas izpētīja HL novērtējuma uzvedību strādājoša modeļa apkārtņē. Rezultātus vispirms var izmantot, lai aplūkotu

ierobežojošās vidējās kvadrātiskās kļūdas īpašības vislielākās ticamības un HL novērtējumiem, kur būs redzams, ka HL metode parasti uzvedas labāk - nedaudz lielāka dispersija tiek kompensēta ar mazāku modelēšanas novirzi. Arīdzan vidējās kvadrātiskās kļūdas līkne kā funkcija no balansa parametra a var tikt novērtēta no datiem. Tas noved pie idejas izvēlēties tādu a , lai tas minimizētu šo riska līkni.

Dotam fokusa parametram ψ ierobežojošā vidējā kvadrātiskā kļūda izmantojot HL metodi ar parametru a ir iegūstama no (5.6):

$$\text{mse}(a) = \{\omega_{\text{hl}}(a)^t \delta\}^2 + \tau_{0,\text{hl}}(a)^2. \quad (6.2)$$

Vispirms nepieciešams novērtēt šo līkni kā funkciju no balansa parametra a situācijās, kad dots modeļa paplašinājuma parametrs δ . Ja $a = 0$, $\text{mse}(a)$ atbilst vislielākās ticamības novērtējuma vidējai kvadrātiskajai kļūdai. Ja $\text{mse}(a)$ ir mazāka nekā $\text{mse}(0)$ kādam a , tad HL metode strādā labāk nekā maksimālās ticamības metode.



Att. 6.1: (a) Punktotā horizontālā līnija ir vidējās kvadrātiskās kļūdas sakne vislielākās ticamības novērtējumam, pilnā līnija - HL novērtējumam kā funkcija no balansa parametra a , (b) $\text{fic}(a)$ kvadrātsakne kā funkcija no a , kad $n=100$ un $\gamma = 1 + \delta\sqrt{n}$ [5]

Attēlā 6.1(a) redzama $\text{mse}(a)$ kvadrātsaknes līkne ļoti vienkāršā izpildījumā, kur parametrisks sākuma modelis ir $\text{Beta}(\theta, 1)$ ar blīvumu $\theta y^{\theta-1}$ un fokusa parametru $\psi = EY^2$ HL konstrukcijā, kas ir $\theta/(\theta + 2)$ pie modeļa nosacījumiem. Paplašinātais modelis, pie kura mēs aplūkojam vidējās kvadrātiskās kļūdas īpašības vislielākās ticamības un HL metodēm, ir $\text{Beta}(\theta, \gamma)$ ar $\gamma = 1 + \delta/\sqrt{n}$ no (5.1). Izvēlētā δ šajā gadījumā ir $Q^{1/2} = (J^{11})^{1/2}$ no (6.3), ko var interpretēt kā vienu standartnovirzi tālāk no nulles modeļa. $\text{mse}(a)$ kvadrātsaknes līkne parāda, ka HL novērtējums $\hat{\theta}_{\text{hl}}/(\hat{\theta}_{\text{hl}} + 2)$ ir labāks nekā $\hat{\theta}_{\text{ml}}/(\hat{\theta}_{\text{ml}} + 2)$, izņemot, ja a ir tuvu 1. Ja neatbilstības parametrs δ ir ļoti mazs, t.i., kad strādājošs modelis ir gandrīz pareizs, šie rezultāti reizēm var norādīt, ka vislielākās ticamības metode ir labāka nekā HL metode.

Jāatzīmē, ka $\omega_{\text{hl}}(a)$ no (5.6) sākas pie $a = 0$ ar $J_{01}J_{00}^{-1}\frac{\partial\psi}{\partial\theta} - \frac{\partial\psi}{\partial\gamma}$ no (5.5), kas ir saistīts ar vislielākās ticamības metodi, bet tas samazinās un tuvojas nullei, pieaugot a vērtībai. Līdz ar to, ja HL metode izmanto tikai mazu daļu no parastās logaritmiskās ticamības tās konstrukcijas ziņā, no tā sekojošajam $\hat{\psi}_{\text{hl},a}$ ir maza novirze, bet potenciāli lielāka dispersija nekā vislielākās ticamības gadījumā. Līdz ar to HL metodi var uztvert kā novirzes samazināšanas procesu kontroles un fokusa parametriem gadījumos, kad parametriskais modelis $f(\cdot, \theta)$ nav pietiekami uzticams.

6.3. Vidējās kvadrātiskās kļūdas novērtējums

Ideja, lai atrastu labu a vērtību, ir novērtēt $\text{mse}(a)$ un tad izvēlēties tā minimizētāju. $\text{mse}(a)$ ietver sevī ω_{hl} un $\tau_{0,\text{hl}}$ no (5.6), un tos var novērtēt ar *plug-in* pieeju. Grūtības sagādā parametra δ daļa, precīzāk, $\delta\delta^t$ no $\omega_{\text{hl}}(a)\delta\delta^t\omega_{\text{hl}}(a)$. Šim parametram, kas ir $O(1/\sqrt{n})$ mērojams caur $\gamma = \gamma_0 + \delta/\sqrt{n}$, svarīgākā informācija atrodama $D_n = \sqrt{n}(\hat{\gamma}_{\text{ml}} - \gamma_0)$ caur parametriskās vislielākās ticamības novērtējumu paplašinātajā $f(y, \theta, \gamma)$ modelī. Pēc [4] saistības ar fokusēto informācijas kritēriju (FIC), mums ir

$$D_n \rightarrow_d D \sim N_r(\delta, Q), \quad \text{kur} \quad Q = J^{11} = (J_{11} - J_{10}J_{00}^{-1}J_{01})^{-1}. \quad (6.3)$$

Faktors δ/\sqrt{n} no $O(1/\sqrt{n})$ nevar tikt novērtēts vienmēr. Tā kā DD^t vidējā vērtība ir $\delta\delta^t + Q$ robežās, mēs novērtēsim kvadrātisko novirzes parametru $(b^t\delta)^2 = b\delta\delta^tb$ veidā izmantojot $\{b^t(D_nD_n^t - \hat{Q})b\}_+$, kur ar \hat{Q} novērtē $Q = J_{11}$, un $x_+ = \max(x, 0)$. Mēs konstruējam $r \times r$ matricu \hat{Q} no novērtētā un invertējam pilno $(p+r) \times (p+r)$ Fišera informācijas matricu J_{wide} no (5.4). Tas noved pie $\text{mse}(a)$ novērtēšanas izmantojot

$$\begin{aligned} \text{fic}(a) &= \{\hat{\omega}_{\text{hl}}(a)^t(D_nD_n^t - \hat{Q})\hat{\omega}_{\text{hl}}(a)\}_+ + \hat{\tau}_{0,\text{hl}}(a)^2 = \\ &[n\hat{\omega}_{\text{hl}}(a)^t\{(\hat{\gamma} - \gamma_0)(\hat{\gamma} - \gamma_0)^t - \hat{Q}\}\hat{\omega}_{\text{hl}}(a)]_+ + \hat{\tau}_{0,\text{hl}}(a)^2. \end{aligned} \quad (6.4)$$

Piezīme 10. Fokusētās informācijas kritērijs FIC tiek izmantots, lai atrastu konkrētu labāko modeli dotajam interesējošajam fokusa parametram, kamēr, piemēram, AIC izvēlas vienu labāko modeli neatkarīgi no tā, kam modelis paredzēts. Tas ir atkarīgs no fokusa parametra, līdz ar to dod dažādas vērtības pie dažādiem fokusa parametriem. *Claeskens* un *Hjort* savā publikācijā [3] parāda, ka bieži vien FIC dod labākus rezultātus nekā AIC.

Attēls 6.1(b) parāda šādu fic -kvadrātsaknes līkni, novērtēto $\text{mse}(a)$ kvadrātsakni. Kamēr $\text{mse}(a)$ kvadrātsaknes līkne attēlā 6.1(a) nāk tikai un vienīgi no no pieņēmumiem un skaitliskiem pētījumiem no paplašinātā $f(y, \theta, \gamma)$ modeļa, tikmēr $\text{fic}(a)$ kvadrātsaknes līkne ir iegūta no dotiem datiem. Sākuma modelis un tā paplašinājums no attēla 6.1(a) ir $\text{Beta}(\theta, 1)$ no $\text{Beta}(\theta, \gamma)$ ar $n = 100$ simulētiem datu punktiem, izmantojot $\gamma = 1 + \delta/\sqrt{n}$, kur δ izvēlēts atbilstoši attēlā

6.1(a) redzamajam. HL metode tika izmantota ar otro momentu $\psi = EY^2$ gan kontroles, gan fokusa gadījumā. Novērtētais risks ir vismazākais gadījumā, kad $a = 0.41$.

7. HL REGRESIJAS MODEĻI

HL ideju var pacelt no neatkarīgu un vienādi sadalīgu izlašu ietvara uz regresijas modeli. Aplūkosim normālo lineāro regresijas modeli datiem (x_i, y_i) ar d -dimensionālu neatkarīgo mainīgo vektoru x_i un y_i , kam vidējā vērtība ir $x_i^t \beta$. Vislielākās ticamības metodes risinājums būtu saistīts ar vienādojumu $Em(Y, X, \beta) = 0$, kur $m(y, x, \beta) = (y - x^t \beta)x$. Pamatā esošo regresijas parametru var izteikt kā $\beta = (EXX^t)^{-1}EXY$, iesaistot arī neatkarīgo mainīgo sadalījumu. Tagad aplūkosim apakšvektoru x_0 ar dimensiju $d_0 < d$ un ar to asociēto novērtēšanas vienādojumu $m_0(y, x, \gamma) = (y - x_0^t \gamma)x + 0$. Tā tiek ieviesta HL konstrukcija $(1 - a)\ell_n(\beta) + a \log R_n(\gamma(\beta))$. Šeit $\ell_n(\beta)$ ir parastā logaritmiskā ticamība, $R_n(\gamma)$ ir empīriskā ticamība saistīta ar m_0 un $\gamma(\beta)$ ir $(EX_0X_0^t)^{-1}EX_0Y$ redzams caur mazākās regresijas lēcu, kur $EX_0Y = X_0X^t\beta$. Tas noved pie secinājuma par β , kur tiek ņemts vērā, ka regresija, ņemotvērā x_0 komponentes, var būt gana svarīga.

8. PRAKTISKIE PIEMĒRI

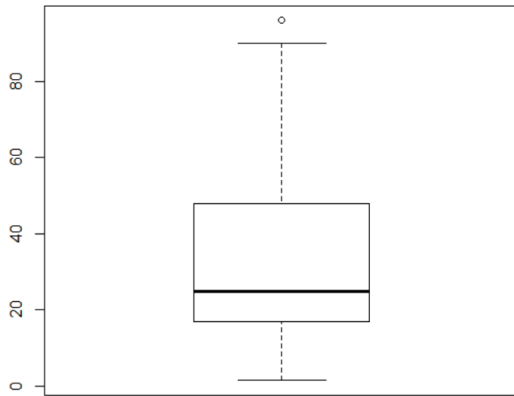
8.1. Ēģiptiešu dzīves ilgums Romas impērijas valdīšanas laikā

Pīrsons (1902) aplūkojis interesantu datu kopu $n = 141$ un tās sadalījumus, kurā redzams Ēģiptiešu dzīves ilgums sievietēm un vīriešiem Romas impērijas valdīšana laikā (pirms apm. 2000 gadu). [13] Šo pašu piemēru HL kontekstā aplūko arī pamatā izmantotajā publikācija [5], līdz ar to šajā nodaļā būs redzams tā paša piemēra iegūtie praktiskie rezultāti, kā arī nedaudz smalkāk aplūkots pats process, kā HL modelis nonāk pie šiem rezultātiem. Salīdzināsim arī dažādus HL modeļus grafiski. Jāatzīmē, ka praktiskie iegūtie rezultāti pilnībā sakrīt ar [5] iegūto, tāpēc var pieņemt, ka uzprogrammētais modelis ir pareizs.

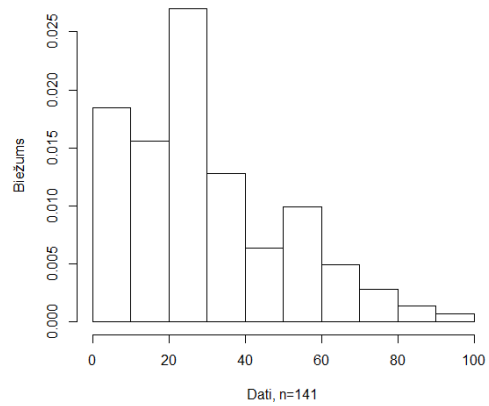
Tabula 8.1: Seno ēģiptiešu dzīves ilgums

Dzīves ilgums gados	V	S	Dzīves ilgums gados	V	S	Dzīves ilgums gados	V	S
1.5	1	1	21	1	6	50	5	1
1.83	1	-	22	2	2	52	3	1
2	2	-	23	1	2	54	-	1
3	3	2	24	2	-	55	2	1
4	4	3	25	5	4	59	1	-
5	3	-	26	4	1	60	4	1
6	-	2	27	1	1	62	1	-
7	1	-	28	-	1	63	1	-
8	-	-	29	1	1	65	2	-
9	-	1	30	1	2	68	2	-
10	1	1	32	1	-	70	-	1
11	1	1	33	2	1	70.5	1	-
14	2	1	35	-	3	72	3	-
16	1	1	36	2	2	84	1	-
17	2	3	37	1	-	90	1	-
18	-	2	40	3	3	96	-	1
19	1	2	46	1	-			
20	2	3	48	2	-			

Tabulā 8.1 redzami šajā piemērā izmantotie dati, kur V apzīmē vīriešu skaitu (kopā 82) un S - sieviešu skaitu (kopā 59). Aplūkosim dzīves ilguma visus datus kopā bez to grupēšanas, kur $n = 141$.



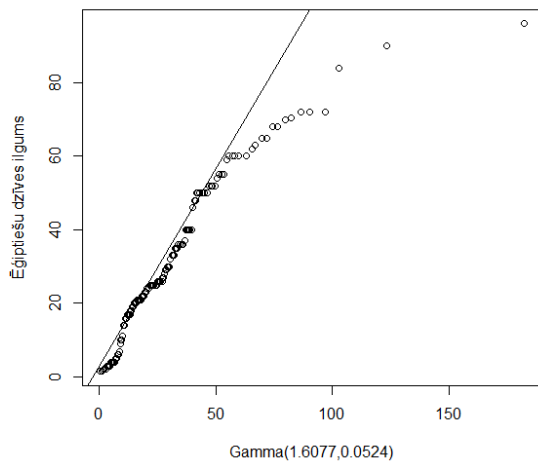
(a) Ēģiptiešu dzīves ilguma kastu grafiks



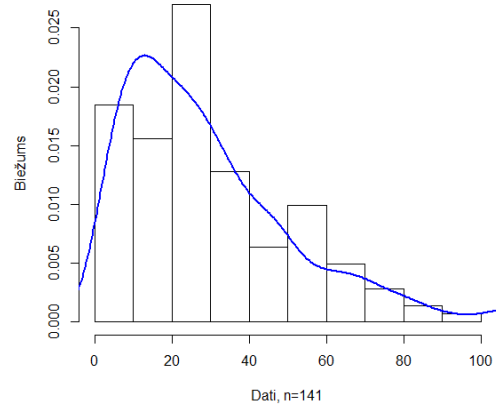
(b) Ēģiptiešu dzīves ilguma histogramma

Att. 8.1: Ēģiptiešu dzīves ilgums

Publikācija [5] piedāvā izmantot $\text{Gamma}(b, c)$ sadalījumu pēc 6. piemēra parauga 3. nodaļā, kur b ir formas parametrs (*shape*) un c ir attiecības parametrs (*rate*). Novērtējot ar parasto vislielākās ticamības metodi iegūst $b = 1.6077$ un $c = 0.0524$, kas grafiski izskatās šādi:



(a) Kvantīļu-kvantīļu grafiks, kurā salīdzināti ēģiptiešu dzīves ilgumi ar $\text{Gamma}(b, c)$, kur b, c iegūti ar vislielākās ticamības novērtējuma palīdzību.



(b) Ēģiptiešu dzīves ilguma histogramma ar vislielākās ticamības novērtējuma blīvuma funkciju.

Att. 8.2: Ēģiptiešu dzīves ilgums - vislielākās ticamības novērtējums

Esam uzzinājuši, ka $\text{Gamma}(b, c)$ sadalījums diezgan labi atbilst datiem, bet ne to lielākajām vērtībām. Tāpēc HL modelī šajā gadījumā iespējams izvēlēties būtisku intervālu, tādējādi

padarot to robustāku un dodot vairāk kontroles pār modeli. Šajā piemērā mūs interesē varbūtība

$$p = p(b, c) = P\{Y \in [L_1, L_2]\} = \int_{L_1}^{L_2} f(y, b, c) dy,$$

kur $[L_1, L_2]$ ir mūsu interesējošās vecuma grupas. Attiecīgi logaritmiskā HL izskatās šādi:

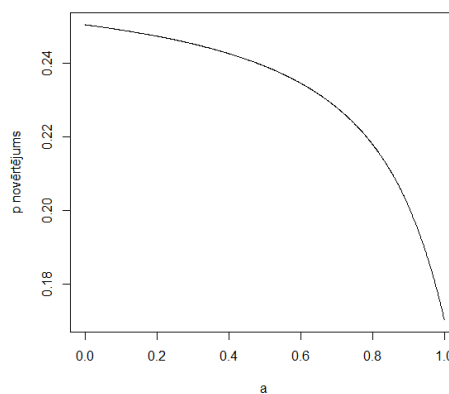
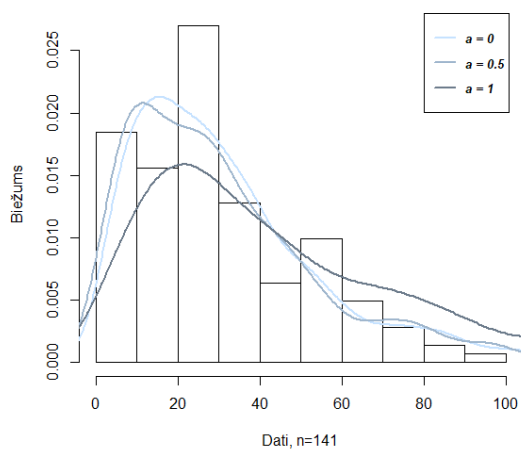
$$h_n(b, c) = (1 - a)\ell_n(b, c) + a \log R_n(p(b, c)), \quad (8.1)$$

kur $R_n(p)$ ir empīriskā ticamības attiecība, kas saistīta ar $m(y, p) = I\{y \in [L_1, L_2]\} - p$, kas principā atbilst empīriskās ticamības metodes kvantiļu gadījumam, aizstājot kvantili q ar varbūtību p , kā arī šajā gadījumā jāņem vērā, ka y pieder intervālam, kas ierobežots no abām pusēm, tāpēc sadalījuma funkciju jāaprēķina intervālam.

Tagad maksimizēsim funkciju (8.1) katram a un nolāsīsim abus HL novērtējumus (\hat{b}_a, \hat{c}_a) un attiecīgi arī $\hat{p}_a = p(\hat{b}_a, \hat{c}_a)$. Izvēlētais vecuma intervāls ir $[9.5, 20.5]$. Tabulā 8.2 redzami rezultāti dažām a vērtībām, kas aptuveni parāda atšķirību starp dažādām balansa parametra a izvēlēm.

Tabula 8.2: HL novērtējumi ēģiptiešu dzīves ilguma datiem

a	b	c	p
0	1.61342	0.05257	0.25034
0.25	1.59775	0.05114	0.24626
0.5	1.57275	0.04875	0.0.2391
0.75	1.53009	0.04399	0.2234
1	1.64179	0.03597	0.17021

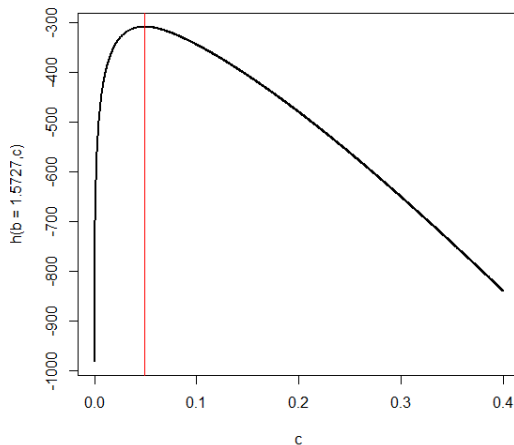


(a) Histogramma un blīvuma funkcijas dažādiem novērtējumiem, ja $a = 0$, $a = 0.5$ un $a = 1$

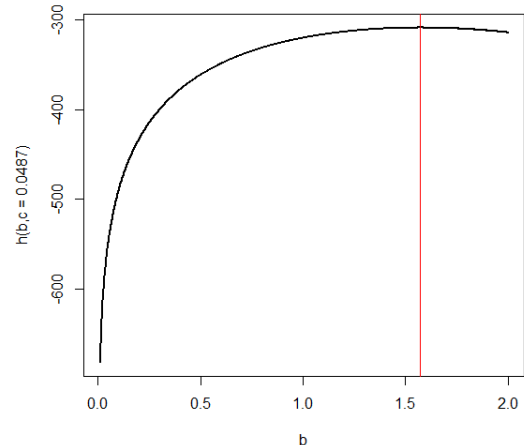
(b) p novērtējuma atkarība no balansa parametra a

Att. 8.3: Iegūtie HL modeļu rezultāti

Attēlā 8.3(a) redzams, ka HL gadījums, kad $a = 0.5$ īpaši neatšķiras no vislielākās ticamības modeļa iegūtā rezultāta $a = 0$, kamēr empīriskās ticamības modelis $a = 1$ dod manāmi atšķirīgu rezultātu. Tikmēr attēla 8.3(b) daļa parāda \hat{p}_a atkarību no balansa parametra a . Pārbaudīsim šos rezultātus vienam no a grafiski. Piemēram, ja $a = 0.5$, izvēlēsimies $b = 1.57275$ un virknei dažādu c atradīsim vislielāko vērtību. Tad aplūkosim $c = 0.04875$ un atkārtosim to pašu tikai ar dažādām b vērtībām. Rezultātā iegūstam šādus grafikus (vertikālā sarkanā līnija norāda uz vislielāko vērtību).



(a) $b = 1.57275$ ir fiksēts, vislielākā funkcijas vērtība tiek sasniegta, ja $c = 0.0487$.



(b) $c = 0.04875$ ir fiksēts, vislielākā funkcijas vērtība tiek sasniegta, ja $b = 1.572$.

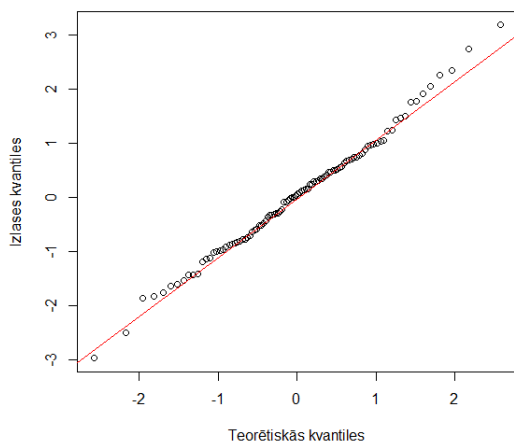
Att. 8.4: b un c maksimizēšana

Redzams, ka grafikiem nav nepieciešama gludināšana atšķirībā no parastās empīriskās ticamības metodes kvantiļu gadījuma, grafiki ir gludi, līdz ar to vislielākās vērtības noteikšana nesagādā nekādas grūtības.

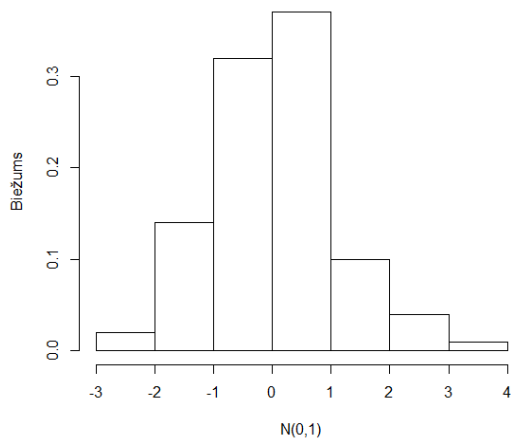
Publikācija [5] norāda, ka šajā gadījumā efektivitātes zudums ir pie mazākām a vērtībām. To pašu varētu atkārtot dažādām vecuma grupām tiem pašiem datiem, lai pēc tam apkopotu rezultātus, un iegūtu labākus secinājumus par fokusa parametru, nekā izmantojot rezultātus tikai par vienu vecuma intervālu.

8.2. Normālā sadalījuma novērtēšana ar kvartiļu palīdzību

Šoreiz aplūkosim praktiski pēc 4. piemēru no 3. nodaļas, mēģinot novērtēt normālo sadalījumu ar izlecējiem un tā kvartiles. Ņemsim $N(0, 1)$ ģenerētus datus, kur $n = 100$ un 5% no datiem ir izlecēji ar standartnovirzi 5. Šajā piemērā HL empīriskās ticamības daļa arī aprēķināta izmantojot kvantilēm paredzēto metodi, taču šoreiz bija nepieciešama gludināšana, tāpēc sadalījuma funkcija aizstāta ar sadalījuma funkcijas novērtējumu, kas iegūts atbilstoši (2.15).



(a) Kvantiļu-kvantiļu grafiks



(b) Histogramma ģenerētajiem $N(0, 1)$ datiem

Att. 8.5: Ģenerētie dati $N(0,1)$ ar izlēcējiem, kur $n=100$

Attēlā 8.5(a) redzams kvantiļu-kvantiļu grafiks ģenerētajiem datiem pret teorētisko sadalījumu. Vidū tas ļoti precīzi atbilst teorētiskajam sadalījumam, kamēr galos manāmas nelielas izlēcēju veidotās atšķirības. Novērtējumi atkarībā no balansa parametra a mainās šādi:

Tabula 8.3: HL novērtējumi ģenerētajiem $N(0, 1)$ datiem

a	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi} + 0.675\hat{\sigma}$
0	0.029	1.103	0.773
0.25	0.023	1.101	0.766
0.5	0.016	1.099	0.757
0.75	0.008	1.096	0.747
1	-0.773	2.237	0.736

Tabula no 8.3 kreisajā pusē apkopo rezultātus par $F_n^{-1}(\frac{3}{4})$ kvartiles novērtējumu. Redzams, ka, ja $a = 1$, tad novērtējums ir vistālākais no teorētiskajām vērtībām, līdz ar to var secināt, ka empīriskā ticamības metode šiem datiem nav pati piemērotākā. Salīdzinoši $a < 1$ dod ļoti tuvus rezultātus visām a vērtībām, jo šoreiz daudz lielāks svars sanāk vislielākās ticamības daļai. Palielinoties a vērtībai, kamēr tā ir mazāka nekā 1, samazinās vidējās vērtības ξ novērtējums, taču nedaudz palielinās standartnovirze. Izmaiņas gan ir minimālas un visi novērtējumi ir diezgan tuvi teorētiskajam sadalījumam, bet jāņem vērā arī tas, ka, palielinoties a vērtībai, samazinās kvartiles novērtējums, līdz ar to tas tuvojas teorētiskajai vērtībai 0.675 pie lielākām a vērtībām. Protams, svarīgi būtu arī uzzināt vidējo kvadrātisko kļūdu, lai noteiktu labāko a vērtību, bet attiecībā uz kvartiles novērtējumu var pieļaut iespēju, ka HL modelis varētu dot precīzāku novērtējumu šiem datiem, izvēloties $0 < a < 1$.

Palielinot izlecēju skaitu rezultātu būtība nemainās, taču izmainot arī izlecēju standartnovirzi, empīriskās ticamības modelis $a = 1$ dod nesalīdzināmi labāku rezultātu gan par vislielākās ticamības metodi, kad $a = 0$, gan arī par HL ar $0 < a < 1$. Piemēram, ja izlecēju standartnovirzi palielina no 5 uz 10, un to skaitu palielina līdz 25% no visa datu apjoma, tad $N(0, 1)$ ģenerētiem datiem ar $n = 100$, empīriskās ticamības modelis dod iegūst vismazāko standartnovirzi (apm. 3), kamēr citām a vērtībām tā ir daudz lielāka (no 4.3 līdz 4.5), kas nozīmē to, ka ne vislielākās ticamības metode, ne arī HL modelis, kad $a < 1$ nav tik robusts kā parastais empīriskās ticamības modelis, runājot par šiem datiem. Arī kvantiļu novērtējums ir vistuvākais teorētiskajam ar empīriskās ticamības metodi. Tas nozīmē to, ka nākotnē, turpinot HL modeļa izpēti, jāņem vērā, ka pie liela mazāk izteiktu izlecēju skaita empīriskās ticamības metode dod gana labus rezultātus, un vairāk jākoncentrējas uz tādiem piemēriem, kas līdzīgi attēlā 8.5 redzajiem datiem, kam HL modelis tiešām varētu sniegt kādus uzlabojumus.

SECINĀJUMI

Darbā tika apskatīta parametriksās un empīriskās ticamības funkciju metožu kombinācija - HL metode. Tā tika aplūkots arī dažos praktiskos piemēros, salīdzinot dažādos tās iegūtos novērtējumus un to atkarību no balansa parametra. Simulētajiem normālā sadalījuma datiem ar izlēcējiem tika novērtēta vidējā vērtība, standartnovirze un attiecīgi arī viena no kvartilēm. Iegūtie rezultāti liecina, ka pie neliela skaita izlēcēju vislielākās ticamības metode un HL metode strādā labāk par empīriskās ticamības funkcijas metodi, kuras iegūtie rezultāti bija ļoti tāli no teorētiskajiem lielumiem, tomēr grūti noteikt, vai HL metodes piedāvātie rezultāti tiešām dod labus uzlabojumus kvartiļu novērtējumiem. Nākotnē, turpinot pētīt šo metodi, nākamais uzdevums noteikti būtu vidējās kvadrātiskās kļūdas aprēķins, kas ļautu atrast vislabāko modeli, kas ir diezgan sarežģīts uzdevums ar šobrīd pieejamo literatūras daudzumu par šo metodi.

Otrs aplūkotais piemērs tikai aplūkots arī darba pamatā izmantotajā publikācijā [5], un parādījas, kā aprēķināt HL vērtību atsevišķam datu intervālam gadījumā, kad datu sadalījums atbilst $\text{Gamma}(b, c)$ sadalījumam, un aplūkoja dažādu parametru atkarību vienam no otra, ļaujot labāk izprast, kā tieši notiek vislielāko vērtību atrašana, kā arī parādot to, ka šiem rezultātiem tiešām var uzticēties. Šie rezultāti parāda to, ka diezgan viegli un ērti HL metodi var pielietot dažādiem datu intervāliem atsevišķi, lai pēc tam tos apkopotu un iegūtu daudz labākus secinājumus, nekā veicot analīzi visai datu izlasei vienlaicīgi.

Tādējādi, izpildot darba ievadā uzstādītos uzdevumus, tika sasniegts darba mērķis - izpētīt, kā darbojas HL metode un iespēju robežās salīdzināt to ar parasto vislielākās ticamības metodi un neparametrisko empīriskās ticamības metodi. Tā kā praktisku iemeslu dēļ, HL metode tika aplūkota ar vienkāršu neatkarīgu, vienādi sadalītu struktūru, tad arī tas paver iespējas tālākiem pētījumiem, piemēram, regresijas līmenī, kas šajā darbā tika aplūkots pavisam nedaudz no teorētiskās puses. Iespējams, ka HL metode kopā ar regresiju varētu sniegt noderīgus un interesantus rezultātus, bet tas, tāpat kā balansa parametra regulēšana un vidējā kvadrātiskā kļūda, būtu labs pamats daudz lielākam un nopietnākam pētījumam nākotnē.

IZMANTOTĀ LITERATŪRA UN AVOTI

- [1] Adimari G., *An empirical likelihood statistic for quantiles*. J. Statist. Comput. Simul., 1998, Vol. 60, p. 85-95.
- [2] Aldrich J., *R. A. Fisher and the Making of Maximum Likelihood 1912 - 1922*. Statistical Science, 1997, Vol. 12, No. 3, 162-176.
- [3] Hjort N. L., Claeskens G., *Frequentist Model Average Estimators*. Journal of the American Statistical Association, 98, 879-899.
- [4] Hjort N. L., Claeskens G., *Model Selection and Model Averaging*. Cambridge University Press, Cambridge.
- [5] Hjort N. L., Keilegom I.V., McKeague I. W., *Hybrid combinations of parametric and empirical likelihoods*. University of Oslo, Columbia University, and KU Leuven, 2017.
- [6] Hjort N. L., Keilegom I.V., McKeague I. W., *The Hybrid Likelihood: Combining Parametric and Empirical Likelihoods*. Tehnologihuset, 2017.
- [7] Jing B., Zhou W., *Adjusted empirical likelihood method for quantiles*. The Institute of Statistical Mathematics, 2003.
- [8] Kozek A., *On M-estimators and normal quantiles*. The Annals of Statistics, Vol. 31, No. 4. 1170-1185, 2003.
- [9] Lawless J., Qin J., *Empirical likelihood and general estimating equations*. The Annals of Statistics, 1994.
- [10] Maximum Likelihood Estimation (MLE). <https://courses.cit.cornell.edu/econ620/reviewm5.pdf>. skatīts 25.05.2019.
- [11] Owen A. B., *Empirical Likelihood*. Chapman & Hall, 2001.
- [12] Owen A. B., *Empirical likelihood ratio confidence intervals for a single functional*. Biometrika, Vol. 75, No. 2, p. 237-249, 1988.

- [13] Pearson K., *On the change in expectation of life in man during a period of circa 2000 years*. Biometrika, 1, 261-264, 1902.
- [14] Vucāne S., *Empīriskās ticamības funkcijas metodes analīze un tās pielietojumi statistikā*. diplomdarbs, 2008
- [15] Wasserman L., *All of statistics A Concise Course in Statistical inference*. Springer, 2004.

PIELIKUMS

Ēģiptiešu dzīves ilgums Romas impērijas valdīšanas laikā

```
1 ## Dati
2 dati <- read.table(file="egyptian.txt", header = T)
3 dati$count <- dati$M + dati$F
4 dati <- rep(dati$ageAtDeath, times = dati$count)
5 dati <- data.frame(dati)
6 dati <- dati$dati
7
8 b <- 1.6077
9 c <- 0.0524
10 l1 <- 9.5
11 l2 <- 20.5
12 n <- 141
13
14 ## Varbutiba
15 indicator <- ifelse(dati >= l1 & dati <= l2, 1, 0)
16 g_fun <- function(x) {(c^b * x^(b-1) * exp(-c*x)) / (gamma(b))}
17 g_fun <- Vectorize(g_fun)
18 p <- integrate(g_fun, l1, l2)$value
19
20 ## ML dalja
21 ll <- function(dati, b, c) {
22   R <- dgamma(dati, shape = b, rate = c)
23   sum(log(R))
24 }
25
26 ll(dati, b, c)
27
28 ll_max <- function(dati, x) -ll(dati, x[1], x[2])
29
30 op <- optim(par=c(1.5, 0.05), ll_max, dati=dati, method = "L-BFGS-B", lower
   = c(0.00001, 0.00001), upper = c(2, 0.5))
31 b <- op$par[1]
32 c <- op$par[2]
33 integrate(g_fun, 9.5, 20.5)
34
35 ## EL dalja
36 ## Bez gludinasanas
```

```

37 f <- ecdf(dati)
38 f_int <- f(12) - f(11)
39
40 ## Ar gludinasanu
41 f_nov <- function(x, dati) {
42   h <- n^(-1/4)
43   (1 / n) * sum(pnorm((x - dati) / h))
44 }
45
46 fn1 <- f_nov(11, dati)
47 fn2 <- f_nov(12, dati)
48 f_int <- fn2 - fn1
49
50 ## logR EL daljai
51 lnR_h <- function(dati, b, c) {
52   g_fun <- function(x) {(c^b * x^(b-1) * exp(-c*x)) / (gamma(b))}
53   g_fun <- Vectorize(g_fun)
54   p <- integrate(g_fun, lower = 9.5, upper = 20.5)$value
55   -n * (f_int * log(f_int / p) + (1 - f_int) * log((1 - f_int) / (1 - p)))
56 }
57
58 R_max <- function(dati, x) -lnR_h(dati, x[1], x[2])
59
60 op <- optim(par=c(1.5, 0.05), R_max, dati=dati, method = "L-BFGS-B", lower
  = c(0.00001,0.00001), upper = c(2, 0.5))
61 b <- op$par[1]
62 c <- op$par[2]
63 integrate(g_fun, 9.5, 20.5)
64
65
66
67 ##### lnH - galvena funkcija #####
68 fn1 <- f_nov(11, dati)
69 fn2 <- f_nov(12, dati)
70 f_int <- fn2 - fn1
71
72 lnH <- function(dati, b, c) {
73   R <- dgamma(dati, shape = b, rate = c)
74   ll <- sum(log(R))
75

```

```

76 g_fun <- function(x) {(c^b * x^(b-1) * exp(-c*x)) / (gamma(b))}
77 g_fun <- Vectorize(g_fun)
78 p <- integrate(g_fun, lower = 9.5, upper = 20.5)$value
79 e1 <- -n * (f_int * log(f_int / p) + (1 - f_int) * log((1 - f_int) / (1 -
      p)))
80
81 (1 - a) * l1 + a * e1
82 }
83
84 H_max <- function(dati, x) -lnH(dati, x[1], x[2])
85
86 ## Visi a
87 a_values <- seq(0, 1, by = 0.01)
88 gamma_p <- c()
89 j <- 0
90
91 for (i in a_values) {
92   a <- i
93   j <- j + 1
94   op <- optim(par=c(1.5, 0.05), H_max, dati=dati, method = "L-BFGS-B",
      lower = c(0.00001,0.00001), upper = c(2, 0.5))
95   b <- op$par[1]
96   c <- op$par[2]
97   gamma_p[j] <- integrate(g_fun, 9.5, 20.5)$value
98   print(paste0("a = ", a, " b = ", round(b, 5), " c = ", round(c, 5), " p =
      ", round(gamma_p[j], 5)))
99 }
100 plot(a_values, gamma_p, type = "l", xlab = "a", ylab = "p novertejums")
101
102
103 ## Grafiki – max parbaude
104 a <- 0.5
105 b_seq <- seq(0.01, 2, by = 0.001)
106 c_seq <- seq(0.0001, 0.4, by = 0.0002)
107 k <- 0
108 l <- 0
109 p_b_fixed <- c()
110 p_c_fixed <- c()
111 lRb_fixed <- c()
112 lRc_fixed <- c()

```

```

113
114 b <- 1.6077 # a = 0
115 b <- 1.5727 # a = 0.5
116 b <- 1.6418 # a = 1
117 for (i in c_seq) {
118   k <- k + 1
119   c <- i
120   p_b_fixed[k] <- integrate(g_fun, l1, l2)$value
121   lRb_fixed[k] <- lnH(dati, b, c)
122 }
123
124 c <- 0.0524 # a = 0
125 c <- 0.0487 # a = 0.5
126 c <- 0.0360 # a = 1
127 for (j in b_seq) {
128   l <- l + 1
129   b <- j
130   p_c_fixed[l] <- integrate(g_fun, l1, l2)$value
131   lRc_fixed[l] <- lnH(dati, b, c)
132 }
133
134 plot(c_seq, lRb_fixed, type = "l", lwd = 2, xlab = "c", ylab = "h(b =
135   1.5727, c)")
136 abline(v = c_seq[which(lRb_fixed == max(lRb_fixed))], col = "red")
137 plot(b_seq, lRc_fixed, type = "l", lwd = 2, xlab = "b", ylab = "h(b, c =
138   0.0487)")
139 abline(v = b_seq[which(lRc_fixed == max(lRc_fixed))], col = "red")
140
141 ## max
142 b_seq[which(lRc_fixed == max(lRc_fixed))]
143 c_seq[which(lRb_fixed == max(lRb_fixed))]
144
145 ## Histogramma
146 set.seed(27)
147 hist(dati, prob = T, main = "", ylab = "Biezums", xlab = "Dati, n=141")
148 lines(density(rgamma(1000, shape = 1.6077, rate = 0.0524)), lwd = 2, col =
149   "blue")
150 lines(density(rgamma(1000, shape = 1.5727, rate = 0.0487)), lwd = 2, col =
151   "slategray3")
152 lines(density(rgamma(1000, shape = 1.6418, rate = 0.0360)), lwd = 2, col =

```

```

    "slategray4")
149 legend(80, 0.028, legend=c("a = 0", "a = 0.5", "a = 1"),
150       col=c("slategray1", "slategray3", "slategray4"), lty = c(1,1,1), lwd
       = c(2,2,2), cex=0.8,
151       text.font=4)

```

Normālā sadalījuma novērtēšana ar kvartiļu palīdzību

```

1 ## gludinasanas funkcija
2 f_nov <- function(x, dati) {
3   h <- n^(-1/4)
4   (1 / n) * sum(pnorm((x - dati) / h))
5 }
6
7 ## lnH
8 lnH <- function(dati, mean, sd) {
9   R <- dnorm(dati, mean, sd)
10  ll <- sum(log(R))
11
12  mu <- mean + sd * qnorm(q)
13  f <- ecdf(dati)
14  fn <- f(mu)
15  n <-<- length(dati)
16  fn <- f_nov(mu, dati)
17  el <- -n * (fn * log(fn / q) + (1 - fn) * log((1 - fn) / (1 - q)))
18
19  (1 - a) * ll + a * el
20 }
21
22 H_max <- function(dati, x) -lnH(dati, x[1], x[2])
23
24 ## dati ar izlecejiem
25 ## N(0,1) 5%, sd=5 izlecejiem
26 library(npsm)
27 dati <- rcn(100, 0.05, 5)
28 hist(dati, prob = T, ylab = "Biezums", xlab = "N(0,1)", main = "")
29
30 qqnorm(dati, xlab = "Teoretiskas kvantiles", ylab = "Izlasses kvantiles",
       main = "")

```

```

31 qqline(dati, col = "red")
32
33 mean(dati)
34 sd(dati)
35 q <- 0.75
36 summary(dati)
37
38 a_values <- seq(0, 1, by = 0.05)
39 j <- 0
40
41 for (i in a_values) {
42   a <- i
43   j <- j + 1
44   op <- optim(par=c(quantile(dati)[2],1), H_max, dati=dati, method = "L-
      BFGS-B", lower = c(quantile(dati)[1],1), upper = c(quantile(dati)[5],
      3))
45   mean <- op$par[1]
46   sd <- op$par[2]
47   print(paste0("a = ", a, " mean = ", round(mean, 3), " sd = ", round(sd,
      3), " quart = ", round(mean + sd * qnorm(q), 3)))
48 }
49
50 ## N(0,1) 25%, sd=10 izlecejiem
51 dati <- rcn(100, 0.25, 10)
52 hist(dati, prob = T, ylab = "Biezums", xlab = "N(0,1)", main = "")
53
54 qqnorm(dati, xlab = "Teoretiskas kvantiles", ylab = "Izlases kvantiles",
      main = "")
55 qqline(dati, col = "red")
56
57 mean(dati)
58 sd(dati)
59 q <- 0.75
60 summary(dati)
61
62 a_values <- seq(0, 1, by = 0.05)
63 j <- 0
64
65 for (i in a_values) {
66   a <- i

```

```

67 j <- j + 1
68 op <- optim(par=c(quantile(dati)[2],3), H_max, dati=dati, method = "L-
    BFGS-B", lower = c(quantile(dati)[1],2), upper = c(quantile(dati)[5],
    5))
69 mean <- op$par[1]
70 sd <- op$par[2]
71 print(paste0("a = ", a, " mean = ", round(mean, 3), " sd = ", round(sd,
    3), " quart = ", round(mean + sd * qnorm(q), 3)))
72 }

```

HL funkcija pēc vispārējā gadījuma formulām

```

1 indicator <- ifelse(dati >= 11 & dati <= 12, 1, 0)
2 g_fun <- function(x) {(x^(b - 1) * exp(-x / c)) / (c^b * gamma(b))}
3 g_fun <- Vectorize(g_fun)
4 p <- integrate(g_fun, 9.5, 20.5)
5
6 ## max lambda
7 f <- function(dati, p, indicator) {
8   ff <- function(lambda) {
9     (sum((1 + lambda * (indicator - p))^( -1) * (indicator - p))) / length(
10      dati_izv)
11   }
12   ff <- Vectorize(ff)
13   l_min <- -5
14   l_max <- 5
15   xx <- seq(l_min, l_max, 0.5)
16   plot(xx, ff(xx))
17   uniroot(ff, c(l_min, l_max))$root
18 }
19 ## EL dalja
20 R <- function(dati, p, indicator, lambda) {
21   prod(1 + lambda * (indicator - p))^( -1)
22 }
23
24 R(dati, p, indicator, lambda)
25
26 ## ML dalja
27 ll <- function(dati, b, c) {

```

```

28 R <- dgamma(dati, shape = b, rate = c)
29 -sum(log(R))
30 }
31
32 ll(dati, b, c)
33 ll_max <- function(dati, x) -ll(dati, x[1], x[2])
34
35 op <- optim(par=c(1.5, 0.05), ll_max, dati = dati)
36 b <- op$par[1]
37 c <- op$par[2]
38 integrate(g_fun, 9.5, 20.5)
39
40 ## Abi kopaa - HL
41 H <- function(dati, a, b, c) {
42   p <- integrate(g_fun, 9.5, 20.5)
43   lambda <- f(dati, p, indicator)
44   (1 - a) * ll(dati_izv, b, c) + a * log(R(dati, p$value, indicator, lambda
45   ))
46 }
47 H_max <- function(dati, a, x) -H(dati, a, x[1], x[2])
48
49 ## max rezultats
50 op <- optim(par=c(1.5, 0.05), H_max, dati=dati, a = 0)
51 b <- op$par[1]
52 c <- op$par[2]
53 integrate(g_fun, 9.5, 20.5)

```

Maģistra darbs "Parametriksās un empīriskās ticamības funkciju metožu kombinācijas" izstrādāts LU Fizikas, Matemātikas un Optometrijas fakultātē.

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

Autors: _____ Lauma Burkēviča

Rekomendēju darbu aizstāvēšanai

Vadītājs: prof. Dr. Math. Jānis Valeinis _____ .06.2019.

Recenzents: Māra Delesa-Vēliņa

Darbs iesniegts Matemātikas nodaļā ____ .06.2019.

Dekāna pilnvarotā persona: vecākā metodiķe Dzintra Holsta

Darbs aizstāvēts maģistra gala pārbaudījuma komisijas sēdē

____ .06.2019. prot. Nr _____

Komisijas sekretāre: lektore Maruta Avotiņa