

**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Adam Plavčan

**Maximální věrohodnost za přítomnosti
vedlejších parametrů**

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Michal Pešta, Ph.D.

Studijní program: Finanční matematika

Praha 2022

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Rád by som vyslovil úprimné podakovanie môjmu školiteľovi doc. RNDr. Michalovi Peš-
tovi, Ph.D. za cenné rady a odbornú pomoc, ktorú mi poskytol pri písaní bakalárskej
práce. Zároveň sa chcem poďakovať aj kamarátovi Bc. Martinovi Melicherovi za pomoc
a podporu. Ďakujem!

Název práce: Maximální věrohodnost za přítomnosti vedlejších parametrů

Autor: Adam Plavčan

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Michal Pešta, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Táto bakalárska práca popisuje rôzne metódy vierohodnosti, ktoré sa zaoberajú problematikou vedľajších parametrov a parametrov záujmu. Sú zadefinované pojmy profilová, marginálna, podmienená, integrálna a pseudo vierohodnosť, ktoré sú doprevádzané ilustračnými príkladmi. Praktická časť je venovaná aplikovaniu teoretických poznatkov na gamma rozdelení s parametrami a a p . Sú odvodené explicitné vzťahy pre štyri vierohodnostné metódy, pomocou ktorých získavame v simulačnej štúdii odhady parametrov záujmu, a následne ich porovnávame vrámci jednotlivých metód. Praktická časť bakalárskej práce je spracovaná v programovacom jazyku R.

Kľúčová slova: maximální věrohodnost, vedlejší parametry, rušivé parametry

Title: Maximum likelihood in the presence of incidental parameters

Author: Adam Plavčan

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Michal Pešta, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This bachelor's thesis describes the various methods of likelihood, which deal with the issue of incidental parameters and parameters of interest. Concepts such as profile, marginal, conditional, integrated and pseudo likelihood are defined and they are supplemented by illustrative examples. The practical part is devoted to the application of theoretical knowledge on the gamma distribution with parameters a and p . Explicit formulas for the four methods are derived, using them we obtain estimations of parameters of interest in a simulation study, and then we compare them within individual methods. The practical part of the bachelor's thesis was created in the programming language R.

Keywords: maximum likelihood, incidental parameters, nuisance parameters

Obsah

1	Úvod	2
	Zoznam použitých skratiek	3
2	Oboznámenie sa s problematikou	4
2.1	Motivácia	4
2.1.1	Motivačný príklad na analýzu párových dát	4
3	Teoretická časť	6
3.1	Definícia vierohodnosti	6
3.2	Definícia profilovej vierohodnosti	6
3.3	Ortogonálne parametre	6
3.3.1	Príklad na ortogonálne parametre	7
3.4	Marginálna a podmienená vierohodnosť	7
3.4.1	Ekvivalencia multinomického a Poissonovho modelu	8
3.4.2	Príklad na nekonzistentnosť vierohodných odhadov	9
3.5	Modifikovaná profilová vierohodnosť	10
3.5.1	Príklad na modifikovanú logaritmickú vierohodnosť	12
3.5.2	Príklad na postupnosť výstupov s normálnym rozdelením so strednou hodnotou a rozptylom	14
3.6	Integrálna vierohodnosť	14
3.7	Pseudo vierohodnosť	15
4	Praktická časť	17
4.1	Vierohodnostné metódy odhadovania parametrov pre gamma rozdelenie	17
4.1.1	Odhad maximálnou vierohodnosťou	17
4.1.2	Odhad marginálnou vierohodnosťou	19
4.1.3	Odhad integrálnou vierohodnosťou	21
4.1.4	Odhad pseudo vierohodnosťou	22
4.2	Porovnanie metód na základe simulovaných dát	23
4.2.1	Dáta zo simulácií	24
4.2.2	Výsledky simulácie	29
5	Záver	30
	Zoznam použitej literatúry	31
	Zoznam obrázkov	32
	Zoznam tabuliek	33
A.	Príloha	34

1. Úvod

V tejto bakalárskej práci sa budeme venovať metodológii, ktorá sa zaoberá odhadovaním parametrov pomocou špecifických foriem vierohodnosti. Tento problém sa prvýkrát objavil v článku s názvom „Konzistentné odhady založené na parciálne konzistentných pozorovaniach” Neyman a Scott (1948). Štatistik Jerzy Neyman a jeho študentka Elizabeth Scott nazvali danú problematiku ako problém vedľajších parametrov. Nanešťastie, vplyv článku nemal široký záber a nevenovala sa mu dostatočná pozornosť, čo neskôr priznávajú ďalší štatistici.

Hlavným cieľom teoretickej časti bakalárskej práce je vytvoriť ucelený súbor s jednotným značením, v ktorom sa čitateľ zoznámí s inými formami vierohodnosti, ako sú napríklad: marginálna vierohodnosť, podmienená vierohodnosť, pseudo vierohodnosť a pochopí ich základný princíp na ilustračných príkladoch. Teoretická časť vychádza najmä z prác Pawitan (2001) a Millar (2011) a má prevažne kompilačný charakter.

V praktickej časti sa pozrieme na 4 metódy (profilová, marginálna, integrálna a pseudo vierohodnosť) popísané v teoretickej časti aplikované na gamma rozdelenie a výpočtovo odvodíme vzťahy, ktoré neskôr aplikujeme v simulačnej štúdiu, na ktorú použijeme programovací jazyk R. Výsledky simulačnej štúdie spracujeme do prehľadných tabuliek a grafov s medzikvartilovým rozpätím. Tie majú za úlohu zistiť, ktorá zo spomínaných metód dáva konzistentnejšie výsledky k skutočnej hodnote parametra záujmu.

Zoznam použitých skratiek

iid	nezávislé a rovnako rozdelené
MLE	maximálne vierohodný odhad
RSS	reziduálny súčet štvorcov
\sim	mať pravdepodobnostné rozdelenie
\bar{X}	výberový priemer náhodného výberu \mathbf{X}
S_X^2	výberový rozptyl náhodného výberu \mathbf{X}
$\mathcal{N}(\mu, \sigma^2)$	normálne rozdelenie so strednou hodnotou μ a rozptylom σ^2
$Pois(\lambda)$	Poissonovo rozdelenie s parametrom λ
$Bin(n, p)$	Binomické rozdelenie s parametrami n a p
\xrightarrow{p}	konvergencia v pravdepodobnosti
χ_n^2	chí-kvadrát rozdelenie o n stupňoch voľnosti
\mathbf{M}^\top	transponované \mathbf{M}
\log	prirodzený logaritmus
$\ $	indikátor
$'$	derivácia

2. Oboznámenie sa s problematikou

2.1 Motivácia

Ako vhodný motivačný príklad nám slúži normálny model pre nezávislé a rovnako rozdelené (*iid*) náhodné veličiny

$$X_i \sim \mathcal{N}(\mu, \sigma^2), \quad i = 1, \dots, n.$$

Maximálne vierohodným odhadom rozptylu získame

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n},$$

čo nám oproti výberovému rozptylu, ktorý je nielen konzistentným, ale aj nestranným odhadom, dá v menovateli n namiesto $n - 1$. Dá sa povedať, že toto je nežiaduci dôsledok $\hat{\sigma}^2$, pretože daný odhad neberie do úvahy stratu jedného stupňa voľnosti z odhadu μ , a to konkrétne $\hat{\mu} = \bar{X}$.

Problémy, ktoré budeme riešiť v tejto práci, budú irelevantné v prípade, že dáta, ktoré skúmame sú nezávislé a veľkosť vzorky je dostatočne veľká, vtedy nastupuje na rad asymptotická teória. Pozornosť zameriame na prípady, kde sú limitované štatistické informácie o niektorých alebo všetkých parametroch. To môže nastať v prípade, že veľkosť vzorky je relatívne malá vzhľadom k počtu parametrov, dokonca to môže nastať aj v prípade, keď sa nám veľkosť vzorky javí dostatočná.

Komplexné modely s mnohými parametrami sú často potrebné, keď pozorovania nie sú nezávislé a model musí zachytiť závislostnú štruktúru. Závislostná štruktúra môže, ale nemusí byť predmetom záujmu, záleží na skúmanej otázke. Parametre, ktoré sú priamo relevantné vzhľadom k skúmanej otázke nazývame *parametre záujmu*. Ostatné parametre nazývame *vedľajšie parametre*. Vektorovo zapísané: $\boldsymbol{\theta} = (\boldsymbol{\psi}, \boldsymbol{\lambda})$, kde vektor $\boldsymbol{\theta}$ obsahuje všetky parametre v modeli, vektor $\boldsymbol{\psi}$ obsahuje parametre záujmu a vektor $\boldsymbol{\lambda}$ obsahuje vedľajšie parametre.

Ako už bolo spomenuté vyššie, tento problém sa prvýkrát objavil v článku s názvom „Konzistentné odhady založené na parciálne konzistentných pozorovaniach” od Neyman a Scott (1948). Problematiku ďalej rozvíja článok „Odhadovanie v prítomnosti vedľajších parametrov” publikovaný autorom Mak (1982). Samotný Lancaster (2000) spomína vo svojom diele „Problémy vedľajších parametrov od roku 1948”, že by sa mala vedľajším parametrom venovať väčšia pozornosť. Ďalšími relevantnými publikáciami sú napríklad „In All Likelihood: Statistical Modelling And Inference Using Likelihood” Pawitan (2001), „A maximum likelihood method for the incidental parameter problem” Moreira (2009), „Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB” Millar (2011) a magisterská práca študenta z Osla „Likelihood corrections for the incidental parameter problem in Poisson distributed panel data” Øren (2021).

2.1.1 Motivačný príklad na analýzu párových dát

Uvažujme model, Millar (2011), str. 190, kde pre $i = 1, \dots, n$,

$$Y_{i1} \sim \mathcal{N}\left(\mu_i - \frac{\delta}{2}, \tau^2\right), \quad (2.1)$$

$$Y_{i2} \sim \mathcal{N}\left(\mu_i + \frac{\delta}{2}, \tau^2\right), \quad (2.2)$$

sú vzájomne nezávislé. Predpokladajme, že parameter záujmu je v našom príklade δ , teda môžeme vysloviť nulovú hypotézu $H_0 : \delta = \delta_0$. Ostatné parametre $\mu_i, i = 1, \dots, n$ a τ^2 považujeme za vedľajšie parametre.

Analýza pomocov párových rozdielov

Definujme si rozdiely $D_i = Y_{i2} - Y_{i1}, i = 1, \dots, n$. Ďalej si označme $\sigma^2 = 2\tau^2$, potom sú rozdiely *iid* a $\mathcal{N}(\delta, \sigma^2)$. Na základe klasickej teórie nás to vedie k použitiu párového t-testu na rozdiely D_i s nulovou hypotézou $H_0 : \delta = \delta_0$. Rovnaké výsledky by sme dostali aplikovaním metódy *MLE* na D_i najmä kvôli n v menovateli pri *MLE* $\hat{\sigma}^2$.

Analýza neupravených dát

Teraz predpokladajme, že by nebolo zrejmé, že H_0 vieme testovať pomocou rozdielov D_i a namiesto toho postavíme inferenciu na základe pôvodného modelu $Y_{ij}, i = 1, \dots, n, j = 1, 2$, tak ako je zadané (2.1) a (2.2). Tento model vieme vyhladiť použitím analýzy rozptylu s vysvetľujúcou premennou, ktorá je faktorová premenná majúca n odlišných levelov a váhy pri $j = 1$ s hodnotou -0.5 a $j = 2$ s hodnotou 0.5 . Tento postup využíva presnú klasicкую teóriu, teda správne determinuje, že máme $n - 1$ stupňov voľnosti pri $2n$ pozorovaniach a $n + 1$ regresných koeficientov.

Pokiaľ by sme na dané dáta aplikovali metódu maximálnej vierohodnosti, odhady by boli nepresné, pretože strata $n + 1$ stupňov voľnosti nebola zohľadnená. *MLE* parametra τ^2 je

$$\hat{\tau}^2 = \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^2 (Y_{ij} - \hat{\mu}_{ij})^2,$$

zatiaľ čo nestranný odhad $\hat{\tau}^2$ má menovateľa $n - 1$. Teda $\mathbb{E}[\hat{\tau}^2] < \frac{\tau^2}{2}$ a $\hat{\tau}^2$ je nekonzistentný odhad τ^2 .

3. Teoretická časť

3.1 Definícia vierohodnosti

Nech máme náhodný výber $\mathbf{X} = (X_1, \dots, X_N)$ s rozsahom N z rozdelenia s hustotou f závislou na parametri $\boldsymbol{\theta}$, tj. $f(x; \boldsymbol{\theta})$, potom vierohodnosťou nazývame funkciu

$$L(\boldsymbol{\theta}) \equiv L(\boldsymbol{\theta}; \mathbf{X}) = f_{\mathbf{X}}(\mathbf{X}; \boldsymbol{\theta}) = \prod_{i=1}^N f(X_i; \boldsymbol{\theta}),$$

kde $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ je hustota náhodného vektoru \mathbf{X} pri realizácii $\mathbf{x} = (x_1, \dots, x_N)$.

3.2 Definícia profilovej vierohodnosti

Nech máme vierohodnosť $L(\boldsymbol{\psi}, \boldsymbol{\lambda})$, potom definujeme profilovú vierohodnosť $\boldsymbol{\psi}$ ako

$$L_p(\boldsymbol{\psi}) = \max_{\boldsymbol{\lambda}} L(\boldsymbol{\psi}, \boldsymbol{\lambda}), \quad (3.1)$$

kde maximalizujeme so zafixovanou hodnotou parametra $\boldsymbol{\psi}$. Poznamenajme, že profilová vierohodnosť nadobúda maximum pre rovnaké $\boldsymbol{\psi}$, ktoré nájdeme klasickou vierohodnosťou.

3.3 Ortogonálne parametre

Ideálny scenár napríklad nastáva, ak máme náhodný výber (X, Y) s rozsahom 2, kde náhodná veličina X pochádza z modelu s hustotou $f(x; \boldsymbol{\psi})$ a náhodná veličina Y pochádza z modelu s hustotou $f(y; \boldsymbol{\lambda})$ ¹, kde X a Y sú nezávislé náhodné veličiny a nemáme žiadnu logickú spojitost medzi parametrami $\boldsymbol{\psi}$ a $\boldsymbol{\lambda}$, vtedy definujeme vierohodnosť parametrov $(\boldsymbol{\psi}, \boldsymbol{\lambda})$

$$L(\boldsymbol{\psi}, \boldsymbol{\lambda}) = f(X; \boldsymbol{\psi})f(Y; \boldsymbol{\lambda}) = L(\boldsymbol{\psi}; X)L(\boldsymbol{\lambda}; Y). \quad (3.2)$$

Keďže Y neobsahuje žiadnu informáciu o parametri $\boldsymbol{\psi}$, niet pochyb o tom, že skutočná vierohodnosť parametra $\boldsymbol{\psi}$ je

$$L(\boldsymbol{\psi}) = L(\boldsymbol{\psi}; X). \quad (3.3)$$

Taktiež v ideálnom prípade vieme faktorizovať vierohodnosť, a to formou

$$L(\boldsymbol{\psi}, \boldsymbol{\lambda}) = L_1(\boldsymbol{\psi})L_2(\boldsymbol{\lambda}), \quad (3.4)$$

kde nám nezáleží, ako dáta vstúpia do $L_1(\cdot)$ a $L_2(\cdot)$, ale je zrejmé, že informácie o $\boldsymbol{\psi}$ sú zachytené v $L_1(\boldsymbol{\psi})$ a informácie o parametri $\boldsymbol{\lambda}$ sú v $L_2(\boldsymbol{\lambda})$. Ak takáto faktorizácia existuje, potom hovoríme o ortogonálnych parametroch $\boldsymbol{\psi}$ a $\boldsymbol{\lambda}$.

¹V celej práci značíme hustotu náhodnej veličiny všeobecne $f(\cdot)$. Symboly $f(x; \boldsymbol{\psi})$ a $f(y; \boldsymbol{\lambda})$ tu teda značia dve rôzne hustoty. Rozlíšenie medzi jednotlivými hustotami bude vždy jasné z použitých parametrov.

3.3.1 Príklad na ortogonálne parametre

Uvažujme úmrtia spôsobené dopravnými nehodami, Pawitan (2001), str. 277, kde počet obetí X a Y sú nezávislé náhodné veličiny, ktoré sú poissonovsky rozdelené s parametrami λ_X a λ_Y . Potom vierohodnostná funkcia je

$$L(\lambda_X, \lambda_Y) = e^{-(\lambda_X + \lambda_Y)} \lambda_X^X \lambda_Y^Y.$$

Označme si parameter záujmu $\psi = \lambda_Y / \lambda_X$ a vedľajší parameter $\lambda = \lambda_X + \lambda_Y$. Teda

$$L(\psi, \lambda) = \left(\frac{\psi}{1 + \psi} \right)^Y \left(\frac{1}{1 + \psi} \right)^X \lambda^{X+Y} e^{-\lambda} \equiv L_1(\psi) L_2(\lambda),$$

kde

$$L_1(\psi) \equiv \left(\frac{\psi}{1 + \psi} \right)^Y \left(\frac{1}{1 + \psi} \right)^X,$$

čo je profilová vierohodnosť parametra ψ . Všeobecne však platí, že ak existuje ortogonálny parameter k parametru ψ , potom bez bližšej špecifikácie parametra λ , profilová vierohodnosť automaticky vypočíta $L_1(\psi)$.

Častokrát ideálny prípad nenastane a my dostaneme:

$$L(\boldsymbol{\psi}, \boldsymbol{\lambda}) = L_1(\boldsymbol{\psi}) L_2(\boldsymbol{\psi}, \boldsymbol{\lambda}).$$

3.4 Marginálna a podmienená vierohodnosť

Notácie, ktoré používame, predpokladajú, že $(\mathbf{U}, \mathbf{V}) = h(\mathbf{Y})$ je vhodná invertibilná transformácia náhodného výberu \mathbf{Y} . Napríklad náhodný výber $\mathbf{Y} = (Y_1, \dots, Y_n)$ je transformovaný do náhodného vektora (\mathbf{U}, \mathbf{V}) , kde $\mathbf{U} = \bar{Y}$ a $\mathbf{V} = (Y_1 - \bar{Y}, \dots, Y_{n-1} - \bar{Y})$.

Hustota $f(\mathbf{U}, \mathbf{V}; \boldsymbol{\theta})$ môže byť vyjadrená ako súčin marginálnej hustoty \mathbf{U} a podmienenej hustoty \mathbf{V} pri podmienke \mathbf{U}

$$\begin{aligned} L(\boldsymbol{\theta}) &= L(\boldsymbol{\psi}, \boldsymbol{\lambda}) = f(\mathbf{U}, \mathbf{V}; \boldsymbol{\theta}) \\ &= f(\mathbf{V}|\mathbf{U}; \boldsymbol{\theta}) f(\mathbf{U}; \boldsymbol{\theta}) \\ &\equiv f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi}, \boldsymbol{\lambda}) f(\mathbf{U}; \boldsymbol{\psi}, \boldsymbol{\lambda}) = L_1(\boldsymbol{\psi}, \boldsymbol{\lambda}) L_2(\boldsymbol{\psi}, \boldsymbol{\lambda}). \end{aligned} \tag{3.5}$$

V (3.5) sme si definovali všeobecný vzťah, ktorý platí vždy. Teraz sa pozrieme na dva špeciálne prípady, ktoré môžu nastať:

1. $\mathbf{V}|\mathbf{U}$ závisí iba na $\boldsymbol{\psi}$. V tom prípade sa dá $f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi}, \boldsymbol{\lambda})$ zapísať jednoducho ako $f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})$. V takom prípade $f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})$ nazývame podmienenou vierohodnosťou.
2. \mathbf{U} závisí iba na $\boldsymbol{\psi}$. V tom prípade sa dá $f(\mathbf{U}; \boldsymbol{\psi}, \boldsymbol{\lambda})$ zapísať ako $f(\mathbf{U}; \boldsymbol{\psi})$, kde vzťah $f(\mathbf{U}; \boldsymbol{\psi})$ nazývame marginálnou vierohodnosťou.

Najideálnejší prípad, ktorý môže nastať je, že marginálna hustota \mathbf{U} závisí iba na $\boldsymbol{\lambda}$ a podmienená hustota \mathbf{V} pri podmienke \mathbf{U} závisí iba na $\boldsymbol{\psi}$. Teda modifikujeme (3.5) na

$$\begin{aligned} f(\mathbf{U}, \mathbf{V}; \boldsymbol{\theta}) &= f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})f(\mathbf{U}; \boldsymbol{\lambda}) \\ &\equiv L_1(\boldsymbol{\psi})L_2(\boldsymbol{\lambda}). \end{aligned} \quad (3.6)$$

Môžeme povedať, že \mathbf{U} je čiastočne postačujúce na parameter $\boldsymbol{\lambda}$, pretože podmienená distribúcia \mathbf{V} pri podmienke \mathbf{U} nezávisí na $\boldsymbol{\lambda}$. Z toho vyplýva, že všetky informácie o parametroch záujmu $\boldsymbol{\psi}$, sú obsiahnuté vo vzťahu $f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})$.

Najčastejší prípad, s ktorým sa stretávame je, že podmienená vierohodnosť nezávisí na $\boldsymbol{\lambda}$, ale marginálna hustota \mathbf{U} závisí na obidvoch parametroch $\boldsymbol{\psi}$ a $\boldsymbol{\lambda}$,

$$\begin{aligned} L(\boldsymbol{\theta}) &= L(\boldsymbol{\psi}, \boldsymbol{\lambda}) = f(\mathbf{U}, \mathbf{V}; \boldsymbol{\theta}) \\ &= f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})f(\mathbf{U}; \boldsymbol{\psi}, \boldsymbol{\lambda}) \\ &\equiv L_1(\boldsymbol{\psi})L_2(\boldsymbol{\psi}, \boldsymbol{\lambda}). \end{aligned} \quad (3.7)$$

Marginálne a podmienené vierohodnosti sú užitočné v týchto prípadoch:

- $f(\mathbf{V}|\mathbf{U}; \boldsymbol{\psi})$ alebo $f(\mathbf{U}; \boldsymbol{\psi})$ sú jednoduchšie oproti pôvodnému modelu $f(\mathbf{Y}; \boldsymbol{\psi}, \boldsymbol{\lambda})$.
- Ignorovaním $L_1(\boldsymbol{\psi}, \boldsymbol{\lambda})$ alebo $L_2(\boldsymbol{\psi}, \boldsymbol{\lambda})$ nestratíme veľa informácií.
- Použitie plnej vierohodnosti je nekonzistentné.

3.4.1 Ekvivalencia multinomického a Poissonovho modelu

Nech máme nezávislé Y_1 a Y_2 poissonovsky distribuované so strednými hodnotami μ_1 a μ_2 , Millar (2011), str. 194, a predpokladajme, že parameter záujmu je $p = \frac{\mu_1}{\mu_1 + \mu_2}$. Hustota (Y_1, Y_2) je

$$\begin{aligned} f(\mathbf{y}; \mu_1, \mu_2) &= \frac{\mu_1^{y_1} e^{-\mu_1}}{y_1!} \frac{\mu_2^{y_2} e^{-\mu_2}}{y_2!} \\ &= \frac{(\mu_1 + \mu_2)^{(y_1 + y_2)} e^{-(\mu_1 + \mu_2)}}{(y_1 + y_2)!} \frac{(y_1 + y_2)!}{y_1! y_2!} \left(\frac{\mu_1}{\mu_1 + \mu_2} \right)^{y_1} \left(\frac{\mu_2}{\mu_1 + \mu_2} \right)^{y_2}. \end{aligned} \quad (3.8)$$

Použitím reparametrizácie $p = \frac{\mu_1}{\mu_1 + \mu_2}$, $\lambda = \mu_1 + \mu_2$ a $n = y_1 + y_2$ do vzťahu (3.8) dostaneme kompaktnjší tvar

$$f(\mathbf{y}; p, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!} \times \frac{n!}{y_1! y_2!} p^{y_1} (1-p)^{y_2} = Pois(\lambda) \times Bin(n, p), \quad (3.9)$$

kde $Pois(\lambda)$ je Poissonovo rozdelenie s parametrom λ a $Bin(n, p)$ je binomické rozdelenie s parametrami n a p . Keďže y_1 a y_2 nadobúdajú diskkrétne hodnoty, hustota $\mathbf{y} = (y_1, y_2)$ je pravdepodobnosť pozorovania (y_1, y_2) , čo je identické ako pravdepodobnosť (hustota) pre (n, y_2) . To znamená, že to môžeme prepísať ako

$$f(\mathbf{y}; p, \lambda) = f(n, y_1; p, \lambda) = f(n; \lambda) f(y_1|n; p). \quad (3.10)$$

Toto je práve ideálny scenár, ako sme popisovali vo vzťahu (3.6), kde vieme explicitne vyjadriť informácie ohľadom parametra p práve podmienenou pravdepodobnosťou y_1 s podmienkou n .

3.4.2 Príklad na nekonzistentnosť vierohodných odhadov

Uvažujme dáta, Pawitan (2001), str. 274–275, kde Y_{i1} a Y_{i2} sú *iid* z $\mathcal{N}(\mu_i, \sigma^2)$, pre $i = 1, \dots, N$, všetky sú nezávislé cez index i . Parameter záujmu je σ^2 . Počet neznámych parametrov je $N + 1$ a počet pozorovaní je $2N$. Majme nasimulované dáta (tabuľka 10.1, Pawitan (2001), str. 275), čo nám poskytuje výhodu, že vieme presnú hodnotu parametra $\sigma^2 = 1$ a stredné hodnoty μ_i , takže môžeme ukázať skutočnú vierohodnosť parametra σ^2 . Aby sme mohli pozorovať strannosť/nestrannosť odhadu, potrebujeme aby veľkosť vzorky N bolo dostačujúca, teraz nám bude postačovať $N = 20$.

Nech $\boldsymbol{\theta} = (\mu_1, \dots, \mu_N, \sigma^2)$, potom skutočná vierohodnosť je

$$L(\boldsymbol{\theta}) = \prod_{i=1}^N \prod_{j=1}^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{Y_{ij} - \mu_i}{\sigma}\right)^2\right)$$

$$\log L(\boldsymbol{\theta}) = -N \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{j=1}^2 (Y_{ij} - \mu_i)^2.$$

Ako sme už vyššie poznamenali, keďže poznáme hodnoty μ_i (pretože boli nasimulované), potom plná vierohodnosť je skutočná vierohodnosť parametra σ^2 . Aby sme mohli vypočítať profilovú vierohodnosť parametra σ^2 , pre každé σ^2 ukážeme, že $\hat{\mu}_i = \bar{Y}_i$. Označme si reziduálny súčet štvorcov $RSS = \sum_{i=1}^N \sum_{j=1}^2 (Y_{ij} - \bar{Y}_i)^2$, potom profilová vierohodnosť parametra σ^2 je

$$\log L(\sigma^2) = \max_{\mu_1, \dots, \mu_N} \log L(\mu_1, \dots, \mu_N, \sigma^2) = -N \log(\sigma^2) - \frac{RSS}{2\sigma^2}$$

a *MLE* je

$$\hat{\sigma}^2 = \frac{RSS}{2N}.$$

Poznamenajme, že to, čo sme dostali, je analýza rozptylu, kde reziduálne stupne voľnosti majú hodnotu N a je zjavné, že RSS je $\sigma^2 \chi_N^2$, teda $E[\hat{\sigma}^2] = \frac{\sigma^2}{2}$ pre všetky N . Ďalej platí, že $\text{var}[\hat{\sigma}^2] = \frac{\sigma^4}{2N}$, z čoho vyplýva

$$\hat{\sigma}^2 \xrightarrow{p} \frac{\sigma^2}{2},$$

teda odhad nie je konzistentný.

Aby sme dostali nestrannú inferenciu pre parameter σ^2 , zvolíme nasledujúcu transformáciu, Pawitan (2001), str. 279,

$$U_i = \frac{Y_{i1} - Y_{i2}}{\sqrt{2}}$$

$$V_i = \frac{Y_{i1} + Y_{i2}}{\sqrt{2}}.$$

Je zrejmé, že všetky U_i sú *iid* a pochádzajú z rozdelenia $\mathcal{N}(0, \sigma^2)$ a všetky V_i sú *iid* a pochádzajú z rozdelenia $\mathcal{N}(\mu_i, \sigma^2)$. Vierohodnosť parametra σ^2 založená na U_i je marginálna vierohodnosť, ktorá má formu

$$L(\sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^N U_i^2\right\},$$

U_i a V_i sú nezávislé, je to tiež podmienená vierohodnosť. MLE z marginálnej vierohodnosti je

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N U_i^2 = \frac{RSS}{N},$$

čo je rovnaký nestranný odhad z analýzy rozptylu.

3.5 Modifikovaná profilová vierohodnosť

Sú situácie, v ktorých nie je možné použiť marginálnu a podmienenú vierohodnosť, aj keď by to teória umožňovala, ale je veľmi náročné odvodiť ich presnú formu. Približné formy marginálnej a podmienenej vierohodnosti vieme získať modifikovaním profilovej vierohodnosti.

Dá sa ukázať, Pawitan (2001), str. 286–288, že približná hustota $\hat{\psi}$ (v prípade, že nemáme vedľajšie parametre a dimenzia parametra $\hat{\psi}$ je 1) je

$$f_{\hat{\psi}}(\hat{\psi}; \psi) \approx \frac{1}{\sqrt{2\pi}} |I(\hat{\psi})|^{1/2} \frac{L(\psi)}{L(\hat{\psi})}. \quad (3.11)$$

Ešte lepšiu aproximáciu, Pawitan (2001), str. 286, je možné dosiahnuť pomocou Barndorff-Nielsenovej p^* -formuly

$$c(\psi) \frac{1}{\sqrt{2\pi}} |I(\hat{\psi})|^{1/2} \frac{L(\psi)}{L(\hat{\psi})},$$

ktorá má nastavenú normalizačnú konštantu $c(\psi)$ tak, aby sa hustota vyintegrovala na jednotku. My budeme používať jednoduchšiu p -formulu, ktorej normalizačná konštanta neobsahuje parameter ψ .

Nech máme $(\hat{\psi}, \hat{\lambda})$ MLE odhady (ψ, λ) , potom približná hustota je

$$f_{\hat{\psi}, \hat{\lambda}}(\hat{\psi}, \hat{\lambda}; \psi, \lambda) \approx c |I(\hat{\psi}, \hat{\lambda})|^{1/2} \frac{L(\psi, \lambda)}{L(\hat{\psi}, \hat{\lambda})}, \quad (3.12)$$

kde $c = (2\pi)^{-p/2}$ a p je dimenzia (ψ, λ) .

Profilová vierohodnosť sa prirodzene vyskytuje v približnej marginálnej distribúcii $\hat{\lambda}$. V skutočnosti je to teoretický základ pre skonštruovanie modifikovanej profilovej vierohodnosti. Majme $\hat{\lambda}_\psi$, čo je MLE parametra λ so zafixovanou hodnotou parametra ψ a $I(\hat{\lambda}_\psi)$ príslušnú napozorovanú Fisherovú informáciu. Profilová vierohodnosť parametra ψ je

$$L_p(\psi) = L(\psi, \hat{\lambda}_\psi).$$

Marginálna hustota parametra $\hat{\lambda}$ je

$$\begin{aligned} f(\hat{\lambda}) &= f(\hat{\lambda}_\psi) \left| \frac{\partial \hat{\lambda}_\psi}{\partial \hat{\lambda}} \right| \\ &\approx c |I(\hat{\lambda}_\psi)|^{1/2} \frac{L(\psi, \lambda)}{L(\psi, \hat{\lambda}_\psi)} \left| \frac{\partial \hat{\lambda}_\psi}{\partial \hat{\lambda}} \right|. \end{aligned} \quad (3.13)$$

Podmienená hustota $\hat{\psi}$ s podmienkou $\hat{\lambda}$ je

$$\begin{aligned} f(\hat{\psi}|\hat{\lambda}) &= \frac{f(\hat{\psi}, \hat{\lambda})}{f(\hat{\lambda})} \\ &\approx c \left| I(\hat{\lambda}_\psi) \right|^{-1/2} \frac{L(\psi, \hat{\lambda}_\psi)}{L(\hat{\psi}, \hat{\lambda}_\psi)} \left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right|. \end{aligned} \quad (3.14)$$

Použili sme p -formulu na menovateľ a čitateľ, čo nás vedie k približnej podmienenej logaritmickej vierohodnosti parametra ψ , ktorá má tvar

$$\begin{aligned} \log L_m(\psi) &= \log L(\psi, \hat{\lambda}_\psi) - \frac{1}{2} \log \left| I(\hat{\lambda}_\psi) \right| + \log \left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right| \\ &= \log L_p(\psi) - \frac{1}{2} \log \left| I(\hat{\lambda}_\psi) \right| + \log \left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right|, \end{aligned} \quad (3.15)$$

kde je $L_m(\psi)$ požadovaná modifikovaná profilová vierohodnosť. Výraz $\frac{1}{2} \log \left| I(\hat{\lambda}_\psi) \right|$ môžeme interpretovať ako chybový člen, ktorý odčíta z profilovej logaritmickej vierohodnosti neistú informáciu z vedľajšieho parametra λ . Jakobián $\left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right|$ slúži ako zachovávateľ invariance, čo drží modifikovanú profilovú vierohodnosť invariantnú voči transformáciám vedľajších parametrov. Rovnako by sme vzťah (3.15) vedeli odvodiť použitím marginálnej distribúcie $\hat{\psi}$.

Vzťah modifikovanej profilovej vierohodnosti (3.15) platí vo všeobecnosti aj tam, kde $(\hat{\psi}, \hat{\lambda})$ nie je postačujúce. Nech máme prostú funkciu dát $x \mapsto (\hat{\psi}, \hat{\lambda}, a(x))$, kde $a(x)$ je doplnkový člen. Aby sme explicitne ukázali závislosť na dátach x , vyjadríme

$$L(\psi, \lambda) \equiv L(\psi, \lambda; x) = L(\psi, \lambda; \hat{\psi}, \hat{\lambda}, a).$$

Pre zafixovaný parameter ψ , MLE $\hat{\lambda}_\psi$ platí

$$\frac{\partial}{\partial \hat{\lambda}_\psi} L(\psi, \hat{\lambda}_\psi; \hat{\psi}, \hat{\lambda}, a) = 0.$$

Po zderivovaní podľa $\hat{\lambda}_\psi$ dostávame

$$\frac{\partial^2}{\partial \hat{\lambda}_\psi^2} L(\psi, \hat{\lambda}_\psi; \hat{\psi}, \hat{\lambda}, a) + \frac{\partial^2}{\partial \hat{\lambda}_\psi \partial \hat{\lambda}} L(\psi, \hat{\lambda}_\psi; \hat{\psi}, \hat{\lambda}, a) \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} = 0,$$

teda

$$\left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right| = \frac{\left| I(\hat{\lambda}_\psi) \right|}{\left| \frac{\partial^2}{\partial \hat{\lambda}_\psi \partial \hat{\lambda}} L(\psi, \hat{\lambda}_\psi; \hat{\psi}, \hat{\lambda}, a) \right|}. \quad (3.16)$$

Menovateľ je niekedy veľmi náročné dostať, pretože nemusíme mať explicitnú závislosť vierohodnosti na $(\hat{\psi}, \hat{\lambda})$.

Sú situácie, kedy nastane $\hat{\lambda}_\psi = \hat{\lambda}$, čo implikuje $\left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right| = 1$, vtedy môžeme posledný člen vzťahu (3.15) vynechať. V prípade, že máme ψ skalár, je možné nastaviť vedľajší parameter λ tak, aby sa $\left| \frac{\partial \hat{\lambda}}{\partial \hat{\lambda}_\psi} \right| \approx 1$, vtedy volíme λ , aby platila rovnosť

$$\mathbb{E} \left[\frac{\partial^2}{\partial \psi \partial \lambda} \log L(\psi, \lambda) \right] = 0.$$

Takéto parametre nazývame informačne ortogonálne.

Je zaujímavé porovnať, Pawitan (2001), str. 288, tieto výsledky s Bayesovskou formulou, teda ako vypočítať modifikovanú vierohodnosť použitím Bayesovských výpočtových metód. Pre skalárne parametre je kvadratická aproximácia

$$\log \frac{L(\psi)}{L(\hat{\psi})} \approx -\frac{1}{2} I(\hat{\psi})(\psi - \hat{\psi})^2,$$

čo implikuje

$$\begin{aligned} \int L(\psi) d\psi &\approx L(\hat{\psi}) \int e^{-\frac{1}{2} I(\hat{\psi})(\psi - \hat{\psi})^2} d\psi \\ &= L(\hat{\psi}) (2\pi)^{1/2} |I(\hat{\psi})|^{-1/2}. \end{aligned}$$

Je to známe ako Laplaceova integrálna aproximácia.

Pre model s dvoma parametrami máme ihneď pre fixné ψ integrálnu vierohodnosť

$$L_{int}(\psi) \equiv \int L(\psi, \lambda) d(\lambda) \approx c L(\psi, \hat{\lambda}_\psi) |I(\hat{\lambda}_\psi)|^{-1/2},$$

kde c neobsahuje parameter ψ , teda

$$\log L_{int}(\psi) \approx \log L_p(\psi) - \frac{1}{2} \log |I(\hat{\lambda}_\psi)|,$$

čo je presne modifikovaná profilová vierohodnosť v prípade ortogonálnych parametrov.

3.5.1 Príklad na modifikovanú logaritmickú vierohodnosť

Nech máme náhodné veličiny (X_1, \dots, X_{11}) , ktoré sú *iid* z rozdelenia $\mathcal{N}(\mu, \sigma^2)$ s rozsahom $n = 11$, Pawitan (2001), str. 289, kde obidva parametre sú neznáme. Logaritmická vierohodnosť je

$$\log L(\mu, \sigma^2) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2.$$

V prvom rade nájdeme modifikovanú profilovú vierohodnosť parametra μ . Nech máme μ *MLE* parametra σ^2 je

$$\begin{aligned}
\hat{\sigma}_\mu^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \\
&= \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i\mu + \mu^2) \\
&= \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 + (\bar{X} - \mu)^2 \\
&= \frac{1}{n} \left(\sum_{i=1}^n X_i^2 - 2 \sum_{i=1}^n X_i \bar{X} + n\bar{X}^2 \right) + (\bar{X} - \mu)^2 \\
&= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 + (\bar{X} - \mu)^2 \\
&= \hat{\sigma}^2 + (\bar{X} - \mu)^2.
\end{aligned}$$

Ľhneď dostávame

$$\log L(\mu, \hat{\sigma}_\mu^2) = -\frac{n}{2} \log \hat{\sigma}_\mu^2 - \frac{n}{2},$$

zároveň

$$I(\hat{\sigma}_\mu^2) = \frac{n}{2\hat{\sigma}_\mu^2}$$

a

$$\frac{\partial \hat{\sigma}_\mu^2}{\partial \hat{\sigma}_\mu^2} = 1.$$

Až na konštantu je modifikovaná profilová vierohodnosť

$$\begin{aligned}
\log L_m(\mu) &= -\frac{n}{2} \log \hat{\sigma}_\mu^2 - \frac{1}{2} \log I(\hat{\sigma}_\mu^2) \\
&= -\frac{n-2}{2} \log \hat{\sigma}_\mu^2,
\end{aligned}$$

rovnaká ako profilová vierohodnosť založená na $(n-1)$ pozorovaniach.

Teraz uvažujme parameter záujmu σ^2 . Pre každé σ^2 získame

$$\hat{\mu}_{\sigma^2} = \bar{X} = \hat{\mu},$$

takže $\frac{\partial \hat{\mu}}{\partial \hat{\mu}_{\sigma^2}} = 1$. Fisherova informácia je

$$I(\hat{\mu}_{\sigma^2}) = \frac{n}{\sigma^2},$$

teda modifikovaná profilová vierohodnosť je

$$\log L_m(\sigma^2) = -\frac{n-1}{2} \log \sigma^2 - \frac{(n-1)S_X^2}{2\sigma^2}$$

úplne rovnaká ako marginálna vierohodnosť založená na S_X^2 , kde S_X^2 je výberový rozptyl.

3.5.2 Príklad na postupnosť výstupov s normálnym rozdelením so strednou hodnotou a rozptylom

Nech máme \mathbf{Y} n -dimenzionálny náhodný vektor pochádzajúci z viacrozmerneho normálneho rozdelenia $\mathbf{Y} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde

$$\begin{aligned} \mathbb{E}[\mathbf{Y}] &= \mathbf{X}\boldsymbol{\beta} \\ &= \boldsymbol{\mu}, \end{aligned}$$

\mathbf{X} má rozmery $n \times p$ hodnosti p a rozptylová matica je $\boldsymbol{\Sigma} \equiv \boldsymbol{\Sigma}(\boldsymbol{\psi})$ s rozmermi $n \times n$. Vierohodnosť je

$$\log L(\boldsymbol{\beta}, \boldsymbol{\psi}) = -\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

Nech máme $\boldsymbol{\psi}$, potom MLE $\boldsymbol{\beta}$ je klasický vážený odhad metódou najmenších štvorcov

$$\hat{\boldsymbol{\beta}}_\psi = (\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}.$$

Profilová vierohodnosť parametra $\boldsymbol{\psi}$ je

$$\log L_p(\boldsymbol{\psi}) = -\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_\psi)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_\psi).$$

Fisherova očakávaná informácia je

$$I(\hat{\boldsymbol{\beta}}_\psi) = \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}.$$

Musí platiť

$$\mathbb{E} \left\{ \frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\psi}_i} \log L(\boldsymbol{\beta}, \boldsymbol{\psi}) \right\} = \mathbb{E} \left\{ \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \boldsymbol{\psi}_i} \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right\} = 0,$$

pre ľubovoľné $\boldsymbol{\psi}_i$, teda $\boldsymbol{\beta}$ a $\boldsymbol{\psi}$ sú informačne ortogonálne a platí, že Jakobián $\left| \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \hat{\boldsymbol{\beta}}_\psi} \right| \approx 1$. Potom je modifikovaná profilová vierohodnosť

$$\log L_m(\boldsymbol{\psi}) = \log L_p(\boldsymbol{\psi}) - \frac{1}{2} \log |\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}|,$$

čo presne odpovedá reziduálnej maximálnej vierohodnosti.

3.6 Integrálna vierohodnosť

Dá sa ukázať, že približnú podmienenú vierohodnosť môžeme získať použitím integrálnej vierohodnosti, ktorú definujeme ako

$$L_{\partial \boldsymbol{\lambda}}(\boldsymbol{\psi}) = \int L(\boldsymbol{\psi}, \boldsymbol{\lambda}) d\boldsymbol{\lambda}. \quad (3.17)$$

Odvozenie daného vzťahu: Millar (2011), str. 199–200, pre jednoduchosť predpokladajme, že MLE $\hat{\boldsymbol{\lambda}}$ je čiastočne postačujúce na parametri $\boldsymbol{\lambda}$, teda $\hat{\boldsymbol{\lambda}}$ vo vzťahu (3.6) zohráva úlohu \mathbf{U} . Pawitan (2001) ukázal, že podmienená logaritmickej vierohodnosť môže byť aproximovaná ako

$$l_c(\boldsymbol{\psi}) \approx l(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}_\psi) - \frac{1}{2} \log |-\mathbf{H}(\hat{\boldsymbol{\lambda}}_\psi)| + c(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\lambda}}_\psi), \quad (3.18)$$

kde $l(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}_\psi) = l^*(\boldsymbol{\psi})$ je profilová logaritmická vierohodnosť a

$$\mathbf{H}(\hat{\boldsymbol{\lambda}}_\psi) = \left. \frac{\partial^2 l(\boldsymbol{\psi}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}^2} \right|_{\boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}_\psi} \quad (3.19)$$

je záporná Fisherová informačná matica pre $\boldsymbol{\lambda}$, vyčíslená v $\hat{\boldsymbol{\lambda}}_\psi$, kde $\boldsymbol{\psi}$ je známe.

Ak je člen $c(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\lambda}}_\psi)$ vo vzťahu (3.19) zanedbateľný, potom platí

$$l_c(\boldsymbol{\psi}) \approx l(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}_\psi) - \frac{1}{2} \log |-\mathbf{H}(\hat{\boldsymbol{\lambda}}_\psi)|. \quad (3.20)$$

Vzťah (3.20) je často označovaný ako modifikovaná profilová logaritmická vierohodnosť, ktorá až na aditívnu konštantu korešponduje s log Laplaceovej aproximácie $L_{\partial\lambda}(\boldsymbol{\psi})$. Táto aproximácia je zadaná vzťahom

$$L(\boldsymbol{\theta}; \mathbf{Y}) \approx f(\mathbf{Y}, \hat{\mathbf{U}}_\theta; \boldsymbol{\theta}) (2\pi)^{\frac{q}{2}} |\mathbf{O}(\hat{\mathbf{u}}_\theta)|^{-\frac{1}{2}}, \quad (3.21)$$

kde $\mathbf{O}(\hat{\mathbf{u}}_\theta)$ je matica

$$\mathbf{O}(\hat{\mathbf{u}}_\theta) = - \left. \frac{\partial^2 \log f(\mathbf{y}, \mathbf{u}; \boldsymbol{\theta})}{\partial \mathbf{u}^2} \right|_{\mathbf{u}=\hat{\mathbf{u}}_\theta}.$$

Potom Laplaceova aproximácia je

$$L_{\partial\lambda}(\boldsymbol{\psi}) = \int L(\boldsymbol{\psi}, \boldsymbol{\lambda}) d\boldsymbol{\lambda} \approx L(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}_\psi) (2\pi)^{\frac{q}{2}} |-\mathbf{H}(\hat{\boldsymbol{\lambda}}_\psi)|^{-\frac{1}{2}}, \quad (3.22)$$

kde q je dimenzia parametra $\boldsymbol{\lambda}$.

3.7 Pseudo vierohodnosť

Tento pojem sa v literatúre zvykne označovať aj ako odhadovaná vierohodnosť. My sa budeme držať označenia pseudo vierohodnosť. Nech máme parametrický priestor $(\boldsymbol{\psi}, \boldsymbol{\lambda})$, kde parameter záujmu je $\boldsymbol{\psi}$ a máme ľubovoľný “rozumný” odhad $\hat{\boldsymbol{\lambda}}$ parametra $\boldsymbol{\lambda}$, nie nutne *MLE*. Pseudo vierohodnosť je

$$L_{ps}(\boldsymbol{\psi}) = L(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}). \quad (3.23)$$

Inferencia zo pseudo vierohodnosti sa typicky spolieha na asymptotickú distribúciu odhadu $\hat{\boldsymbol{\psi}}$, teda platí rovnosť

$$S(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}) = \frac{\partial}{\partial \boldsymbol{\psi}} L(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}) = 0. \quad (3.24)$$

Veta 1. *Nech $(\boldsymbol{\psi}_0, \boldsymbol{\lambda}_0)$ sú skutočné hodnoty parametrov, potom za platnosti podmienok regularity, Pawitan (2001), str. 241, platí:*

- Ak je $\hat{\boldsymbol{\lambda}}$ konzistentné, potom existuje konzistentná postupnosť riešení $\hat{\boldsymbol{\psi}}$.

- Ak sú $\sqrt{n}S(\boldsymbol{\psi}_0, \boldsymbol{\lambda}_0)$ a $\sqrt{n}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)$ asymptoticky normálne s nulovou strednou hodnotou a kovariačnou maticou

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

potom $\sqrt{n}(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0)$ je asymptoticky normálny s nulovou strednou hodnotou a rozptylom

$$\sigma^2 = I_{11}^{-1} + I_{11}^{-2}I_{12}(\sigma_{22}I_{12} - 2\sigma_{12}).$$

- Ak je odhad $\hat{\boldsymbol{\lambda}}$ asymptoticky ekvivalentný MLE parametra $\boldsymbol{\lambda}_0$, potom platí $\sigma_{12} = 0$, $\sigma_{22} = (I_{22} - I_{21}I_{11}^{-1}I_{12})^{-1}$ a $\sigma^2 = (I_{11} - I_{12}I_{22}^{-1}I_{21})^{-1}$, potom $\hat{\boldsymbol{\psi}}$ je asymptoticky ekvivalentné MLE.

Najjednoduchší prípad nastane, ak budeme brať do úvahy MLE parametra $\hat{\boldsymbol{\lambda}}$ a $I_{12} = 0$, vtedy sú parametre $\boldsymbol{\psi}$ a $\boldsymbol{\lambda}$ informačne ortogonálne. Asymptoticky je možné považovať $L(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}})$ za klasickú vierohodnosť.

4. Praktická časť

V praktickej časti bakalárskej práce sa budeme venovať aplikácii teoretických poznatkov z tretej kapitoly na simulované dáta a uplatníme rôzne vyššie spomínané metódy.

4.1 Vierohodnostné metódy odhadovania parametrov pre gamma rozdelenie

Nech máme náhodný výber $\mathbf{X} = (X_1, \dots, X_N)$ s rozsahom N , kde $X_i \stackrel{\text{iid}}{\sim} \Gamma(a, p)$. Parameter a považujeme za vedľajší parameter a parameter p za parameter záujmu. Hustota gamma rozdelenia má tvar

$$f(x; a, p) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} \mathbb{1}_{(0, \infty)}(x); \quad a > 0, \quad p > 0, \quad x \in \mathbb{R}, \quad (4.1)$$

kde $\mathbb{1}_{(0, \infty)}(x)$ je indikátor.

4.1.1 Odhad maximálnou vierohodnosťou

Pomocou metódy maximálnej vierohodnosti budeme odhadovať obidva parametre. Takto získaný odhad parametra p je rovný odhadu, ktorý by sme získali profilovou vierohodnosťou, ktorá odpovedá vzťahu (3.1). Nech X_1, \dots, X_N tvoria náhodný výber z rozdelenia $\Gamma(a, p)$.

Zostavíme si vierohodnostné rovnice

$$\begin{aligned} L(a, p) &= \prod_{i=1}^N f(X_i; a, p) \\ &= \prod_{i=1}^N \frac{a^p}{\Gamma(p)} X_i^{p-1} e^{-aX_i} \\ &= \frac{a^{Np}}{(\Gamma(p))^N} \left(\prod_{i=1}^N X_i \right)^{p-1} \exp \left\{ -a \sum_{i=1}^N X_i \right\}; \end{aligned} \quad (4.2)$$

$$\begin{aligned} \log L(a, p) &= \log \prod_{i=1}^N f(X_i; a, p) \\ &= \log \left(\frac{a^{Np}}{(\Gamma(p))^N} \left(\prod_{i=1}^N X_i \right)^{p-1} \exp \left\{ -a \sum_{i=1}^N X_i \right\} \right) \\ &= Np \log a + (p-1) \log \left(\prod_{i=1}^N X_i \right) - a \sum_{i=1}^N X_i - \log \left((\Gamma(p))^N \right) \\ &= Np \log a + (p-1) \sum_{i=1}^N \log X_i - a \sum_{i=1}^N X_i - N \log(\Gamma(p)). \end{aligned} \quad (4.3)$$

Rovnica pre parameter a . Zderivujeme podľa parametra a a následne položíme nule.

$$\begin{aligned}\frac{\partial \log L(a, p)}{\partial a} &= \frac{Np}{a} - \sum_{i=1}^N X_i := 0 \\ \frac{Np}{a} &= \sum_{i=1}^N X_i \Rightarrow \hat{a} = \frac{N\hat{p}}{\sum_{i=1}^N X_i} = \frac{\hat{p}}{\bar{X}}\end{aligned}\tag{4.4}$$

\hat{a} je maximálne vierohodný odhad parametra a .

Rovnica pre parameter p . Zderivujeme podľa parametra p

$$\begin{aligned}\frac{\partial \log L(a, p)}{\partial p} &= N \log a + \sum_{i=1}^N \log X_i - N \frac{\Gamma'(p)}{\Gamma(p)} \\ &= N \log a + \sum_{i=1}^N \log X_i - N \text{Digamma}(p),\end{aligned}\tag{4.5}$$

kde $\text{Digamma}(p)$ je funkcia definovaná vzťahom $\frac{\Gamma'(p)}{\Gamma(p)}$. Nahradíme a maximálne vierohodným odhadom \hat{a} vo vzťahu (4.5) a položíme rovno nule

$$\begin{aligned}\frac{\partial \log L(a, p)}{\partial p} &= N \log \frac{\hat{p}}{\bar{X}} + \sum_{i=1}^N \log X_i - N \text{Digamma}(p) \\ &= N \log \hat{p} - N \log \bar{X} + \sum_{i=1}^N \log X_i - N \text{Digamma}(p) \\ &= N(\log \hat{p} - \log \bar{X} + \overline{\log(X)} - \text{Digamma}(p)) := 0.\end{aligned}\tag{4.6}$$

Teraz sme sa dostali do bodu, kde potrebujeme nájsť riešenie numericky pomocou softvéru. Návod ako numericky vyrátať MLE parametra p (v prílohe k bakalárskej práci sme to aplikovali na 1000 pozorovaní z rozdelenia $\Gamma(0.2, 2)$) a odhady parametrov a a p vyšli $\hat{a} = 0.209$ a $\hat{p} = 2.08$), je nasledovný:

1. vygenerujeme si ľubovoľnú veľkosť vzorky (čím väčšia vzorka, tým bude odhad presnejší) z gamma rozdelenia s parametrami a a p ;
2. vypočítame empirickú verziu strednej hodnoty z dát, ktoré sme vygenerovali (očakávame hodnotu relatívne blízku hodnote $\frac{p}{a}$);
3. zlogaritmuje dáta a vypočítame empirickú verziu strednej hodnoty zlogaritmovaných dát;
4. definujeme si funkciu, ktorú sme dostali v (4.6);
5. vykreslíme do grafu a odhadneme, kde by sa mohol nachádzať koreň rovnice;
6. pomocou funkcie `uniroot` jazyka R nájdeme daný koreň (musí byť relatívne blízko hodnote p);
7. odhad parametra a už len jednoducho vypočítame dosadením do vzťahu (4.4).

4.1.2 Odhad marginálnou vierohodnosťou

Majme náhodný výber X_1, \dots, X_N z rozdelenia $\Gamma(a, p)$. Uvažujme transformáciu

$$G(x_1, \dots, x_N) = \left(\frac{x_1}{S}, \dots, \frac{x_{N-1}}{S}, S \right), \quad (4.7)$$

kde $S = \sum_{i=1}^N x_i$. Transformácia G je bijekcia z $(0, \infty)^N$ do $(0, \infty)^N$. Inverz transformácie G je

$$G^{-1}(y_1, \dots, y_N) = \left(y_1 y_N, \dots, y_{N-1} y_N, y_N \left(1 - \sum_{i=1}^{N-1} y_i \right) \right), \quad (4.8)$$

to sa ľahko overí dosadením. Združenú hustotu $\mathbf{x} = (x_1, \dots, x_n)$ poznáme. Môžeme teda spočítať združenú hustotu $\mathbf{y} = (y_1, \dots, y_N)$. Využijeme nato vzťah

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(G^{-1}(\mathbf{y})) \left| \det(dG^{-1}) \right|, \quad (4.9)$$

kde $f_{\mathbf{X}}$ je združená hustota \mathbf{X} , $\mathbf{Y} = G(\mathbf{X})$ a dG^{-1} je matica parciálnych derivácií G^{-1} vyhodnotených v \mathbf{y} . Po použití vzťahu (4.9) dostávame

$$\begin{aligned} & \left(\prod_{i=1}^{N-1} \frac{a^p}{\Gamma(p)} (y_i y_N)^{p-1} e^{-a y_i y_N} \right) \left(\frac{a^p}{\Gamma(p)} \left(y_N \left(1 - \sum_{i=1}^{N-1} y_i \right) \right) \right)^{p-1} \\ & \quad \times \exp \left\{ -a y_N \left(1 - \sum_{i=1}^{N-1} y_i \right) \right\} \left| \det(dG^{-1}) \right|. \end{aligned}$$

Po úpravách

$$\frac{a^{Np}}{(\Gamma(p))^N} y_N^{N(p-1)} \prod_{i=1}^{N-1} (y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} y_i \right)^{p-1} e^{-a y_N} \left| \det(dG^{-1}) \right|.$$

Teraz vypočítame maticu parciálnych derivácií G^{-1}

$$\begin{pmatrix} \frac{\partial y_1 y_N}{\partial y_1} & \cdots & \frac{\partial y_1 y_N}{\partial y_{N-1}} & \frac{\partial y_1 y_N}{\partial y_N} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial y_{N-1} y_N}{\partial y_1} & \cdots & \frac{\partial y_{N-1} y_N}{\partial y_{N-1}} & \frac{\partial y_{N-1} y_N}{\partial y_N} \\ \frac{\partial y_N (1 - \sum_{i=1}^{N-1} y_i)}{\partial y_1} & \cdots & \cdots & \frac{\partial y_N (1 - \sum_{i=1}^{N-1} y_i)}{\partial y_N} \end{pmatrix},$$

po zderivovaní dostávame

$$\begin{pmatrix} y_N & 0 & \cdots & 0 & y_1 \\ 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & y_N & y_{N-1} \\ -y_N & \cdots & \cdots & -y_N & 1 - \sum_{i=1}^{N-1} y_i \end{pmatrix}.$$

Na výpočet determinantu bude najlepšie robiť stĺpcové úpravy. Keď od posledného stĺpca odčítame $\frac{y_i}{y_N}$ násobok i -teho stĺpca, tak sa determinant nezmení. Keď to urobíme pre každé $i < N$, tak posledný stĺpec sa takmer vynuluje, potom bude výpočet determinantu triviálny. Po spomínaných úpravách dostaneme

$$\begin{pmatrix} y_N & 0 & \dots & 0 & 0 \\ 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \dots & 0 & y_N & 0 \\ -y_N & \dots & \dots & -y_N & 1 \end{pmatrix}.$$

Keďže máme dolnú trojuholníkovú maticu, potom jej determinant je rovný súčinu prvkov na diagonále, teda dostávame

$$|\det(dG^{-1})| = \prod_{i=1}^{N-1} y_N = y_N^{N-1}.$$

Získavame teda združenú hustotu (y_1, \dots, y_N)

$$\frac{a^{Np}}{(\Gamma(p))^N} y_N^{Np-1} \prod_{i=1}^{N-1} (y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} y_i\right)^{p-1} e^{-ay_N}. \quad (4.10)$$

Označme $\mathbf{U} = (Y_1, \dots, Y_{N-1})$ a $\mathbf{V} = Y_N$. Zintegrovaním hustoty \mathbf{Y} podľa y_N dostávame hustotu \mathbf{U}

$$f(\mathbf{u}) = \frac{a^{Np}}{(\Gamma(p))^N} \prod_{i=1}^{N-1} (y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} y_i\right)^{p-1} \int_0^\infty y_N^{Np-1} e^{-ay_N} dy_N.$$

Po vyčíslení

$$f(\mathbf{u}) = \frac{a^{Np}}{(\Gamma(p))^N} \prod_{i=1}^{N-1} (y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} y_i\right)^{p-1} a^{-Np} \Gamma(Np).$$

Upravíme

$$f(\mathbf{u}) = \frac{\Gamma(Np)}{(\Gamma(p))^N} \prod_{i=1}^{N-1} (y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} y_i\right)^{p-1}, \quad (4.11)$$

je dôležité si všimnúť, že $f(\mathbf{u})$ je závislé iba na parametri p , parameter a tam chýba. Použili sme vzťah $f(\mathbf{u}, \mathbf{v}; a, p) = f(\mathbf{u}; p) f(\mathbf{v}|\mathbf{u}; a, p)$. Hustotu $f(\mathbf{v}|\mathbf{u})$ vieme už ľahko dopočítať

$$f(\mathbf{v}|\mathbf{u}) = \frac{f(\mathbf{u}, \mathbf{v})}{f(\mathbf{u})} = \frac{a^{Np}}{\Gamma(Np)} y_N^{Np-1} e^{-ay_N},$$

čo je de facto hustota rozdelenia $\Gamma(a, Np)$. Takže to, čo budeme vo výpočtoch zanedbávať je jedna náhodná veličina s rozdelením $\Gamma(a, Np)$. Keďže o parametri a z ostatných dát nevieme nič (\mathbf{u} nezávisí od a), tak toto nám naozaj o parametri p nič nehovorí. Vyškrtneme člen $f(\mathbf{v}|\mathbf{u})$, pretože neobsahuje veľa informácií o parametri p . Náš odhad parametra p bude taký, ktorý maximalizuje marginálnu vierohodnosť $f(\mathbf{U})$. Postupujeme klasicky

$$\log f(\mathbf{U}) = \log \left(\frac{\Gamma(Np)}{(\Gamma(p))^N} \prod_{i=1}^{N-1} (Y_i)^{p-1} \left(1 - \sum_{i=1}^{N-1} Y_i\right)^{p-1} \right).$$

Dostaneme

$$\log \Gamma(Np) - N \log \Gamma(p) + (p-1) \sum_{i=1}^{N-1} \log Y_i + (p-1) \log \left(1 - \sum_{i=1}^{N-1} Y_i \right).$$

Zderivujeme a postavíme rovno nule

$$NDigamma(Np) - NDigamma(p) + \sum_{i=1}^{N-1} \log Y_i + \log \left(1 - \sum_{i=1}^{N-1} Y_i \right) := 0.$$

Tu ale opäť potrebujeme nájsť riešenie numericky pomocou softvéru (viď. príloha). Ukázalo sa, že marginálna vierohodnosť použitá na rovnaké dáta, pri ktorých sme numericky počítali maximálnu vierohodnosť, je vhodná (odhad parametra p , vyšiel $\hat{p} = 2.079$ oproti skutočnej hodnote 2).

4.1.3 Odhad integrálnou vierohodnosťou

Teraz budeme odhadovať náš parameter záujmu p integrálnou vierohodnosťou, a to presne podľa vzťahu (3.17). Pre lepší prehľad si zavedieme značenie $b = \left(\prod_{i=1}^N X_i \right)^{p-1}$, $c = \sum_{i=1}^N X_i$, $d = (\Gamma(p))^N$, potom si vzťah (4.2) môžeme prepísať na

$$\frac{a^{Np}}{(\Gamma(p))^N} \left(\prod_{i=1}^N X_i \right)^{p-1} \exp \left\{ -a \sum_{i=1}^N X_i \right\} = \frac{a^{Np}}{d} b e^{-ac}.$$

Budeme integrovať cez parameter a na intervale $(0, \infty)$

$$\begin{aligned} L_{\partial a}(p) &= \int_0^\infty \frac{a^{Np}}{d} b e^{-ac} da = \frac{b}{d} \int_0^\infty a^{Np} e^{-ac} da \\ &= \frac{b}{d} c^{-Np-1} \Gamma(Np+1) \\ &= \frac{\left(\prod_{i=1}^N X_i \right)^{p-1}}{(\Gamma(p))^N} \left(\sum_{i=1}^N X_i \right)^{-Np-1} \Gamma(Np+1). \end{aligned} \tag{4.12}$$

Máme teda funkciu (4.12), ktorá závisí čisto na parametri p a ideme nájsť jej maximum. Označme $f = \prod_{i=1}^N X_i$, potom prvá derivácia

$$\frac{\partial}{\partial p} \left(\frac{f^{p-1} c^{-Np-1} \Gamma(Np+1)}{(\Gamma(p))^N} \right)$$

sa rovná

$$\begin{aligned} &f^{p-1} (-c^{-Np-1}) (\Gamma(p))^{-N} \Gamma(Np+1) \\ &\quad \times (N \log(c) - \log(f) + NDigamma(p) - NDigamma(Np+1)), \end{aligned}$$

keďže sú prvé štyri členy nenulové, potom nulu nám vie dať iba posledný člen $(N \log(c) - \log(f) + NDigamma(p) - NDigamma(Np+1))$. Avšak, tu opäť potrebujeme nájsť riešenie numericky pomocou softvéru (viď. príloha). Ukázalo sa, že integrálna vierohodnosť použitá na rovnaké dáta, pri ktorých sme numericky počítali maximálnu vierohodnosť, je vhodná (odhad parametra p , vyšiel $\hat{p} = 2.082$ oproti skutočnej hodnote 2).

4.1.4 Odhad pseudo vierohodnosťou

V prvom rade si vypočítame nejaký rozumný odhad vedľajšieho parametra a . Vhodnou metódou na výpočet môže byť napríklad momentová metóda. Keďže platí $E[X] = \frac{p}{a}$ a $\text{var}[X] = \frac{p}{a^2}$, potom $E[X^2] = \frac{p(p+1)}{a^2}$, čo vieme dobre odhadnúť $\frac{1}{n} \sum_{i=1}^N X_i^2$. Strednú hodnotu odhadneme výberovým priemerom.

1. rovnica

$$\frac{p}{a} = \frac{1}{N} \sum_{i=1}^N X_i, \quad (4.13)$$

2. rovnica

$$\frac{p(p+1)}{a^2} = \frac{1}{N} \sum_{i=1}^N X_i^2. \quad (4.14)$$

Z 1. rovnice (4.13) vyplýva

$$p = a\bar{X} \Rightarrow \hat{p} = \hat{a}\bar{X}. \quad (4.15)$$

Dosadením do 2. rovnice (4.14), dostávame

$$\frac{a\bar{X}(a\bar{X} + 1)}{a^2} = \frac{\bar{X}(a\bar{X} + 1)}{a} = \frac{1}{N} \sum_{i=1}^N X_i^2. \quad (4.16)$$

Vyjadrením parametra a zo vzťahu (4.16) dostávame odhad parametra a , \hat{a} , pomocou momentovej metódy

$$\hat{a} = \frac{\bar{X}}{\frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2} = \frac{\bar{X}}{\frac{N-1}{N} \left(\frac{N}{N-1} \left(\frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2 \right) \right)} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}.$$

Teraz už len použitím vzťahu 4.15 dostávame odhad parametra p , \hat{p} , pomocou momentovej metódy

$$\hat{p} = \frac{\bar{X}^2}{\frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2} = \frac{\bar{X}^2}{\frac{N-1}{N} \left(\frac{N}{N-1} \left(\frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2 \right) \right)} = \frac{\bar{X}^2}{\frac{N-1}{N} S_X^2}.$$

Vypočítali sme si vhodný odhad vedľajšieho parametra a a následne, na základe definície (3.23) zostavíme pseudo vierohodný odhad parametra záujmu p .

Pseudo vierohodnosť je

$$\begin{aligned} L_{ps}(p) &= L(p, \hat{a}) = \prod_{i=1}^N f(X_i; p, \hat{a}) \\ &= \prod_{i=1}^N \frac{\left(\frac{\bar{X}}{\frac{N-1}{N} S_X^2} \right)^p}{\Gamma(p)} X_i^{p-1} \exp \left\{ -\frac{\bar{X}}{\frac{N-1}{N} S_X^2} X_i \right\} \\ &= \frac{\left(\frac{\bar{X}}{\frac{N-1}{N} S_X^2} \right)^{Np}}{(\Gamma(p))^N} \left(\prod_{i=1}^N X_i \right)^{p-1} \exp \left\{ -\frac{\bar{X}}{\frac{N-1}{N} S_X^2} \sum_{i=1}^N X_i \right\}. \end{aligned} \quad (4.17)$$

Zlogaritmuje

$$\begin{aligned}
\log L_{ps}(p) &= \log \prod_{i=1}^N f \left(X_i; p, \frac{\bar{X}}{\frac{N-1}{N} S_X^2} \right) \\
&= \log \left(\frac{\left(\frac{\bar{X}}{\frac{N-1}{N} S_X^2} \right)^{Np}}{(\Gamma(p))^N} \left(\prod_{i=1}^N X_i \right)^{p-1} \exp \left\{ -\frac{\bar{X}}{\frac{N-1}{N} S_X^2} \sum_{i=1}^N X_i \right\} \right) \\
&= Np \log \frac{\bar{X}}{\frac{N-1}{N} S_X^2} + (p-1) \log \left(\prod_{i=1}^N X_i \right) - \frac{\bar{X}}{\frac{N-1}{N} S_X^2} \sum_{i=1}^N X_i - \log \left((\Gamma(p))^N \right) \\
&= Np \log \frac{\bar{X}}{\frac{N-1}{N} S_X^2} + (p-1) \sum_{i=1}^N \log X_i - \frac{\bar{X}}{\frac{N-1}{N} S_X^2} \sum_{i=1}^N X_i - N \log(\Gamma(p)).
\end{aligned} \tag{4.18}$$

Zderivujeme a postavíme rovno nule

$$\begin{aligned}
\frac{\partial \log L_{ps}(p)}{\partial p} &= N \log \frac{\bar{X}}{\frac{N-1}{N} S_X^2} + \sum_{i=1}^N \log X_i - N \frac{\Gamma'(p)}{\Gamma(p)} \\
&= N \log \frac{\bar{X}}{\frac{N-1}{N} S_X^2} + \sum_{i=1}^N \log X_i - N \text{Digamma}(p) := 0.
\end{aligned} \tag{4.19}$$

Opäť potrebujeme nájsť riešenie numericky pomocou softvéru (viď. príloha). Ukázalo sa, že pseudo vierohodnosť použitá na rovnaké dáta, pri ktorých sme numericky počítali maximálnu, integrálnu a marginálnu vierohodnosť, je vhodná (odhad parametra p , vyšiel $\hat{p} = 2.088$ oproti skutočnej hodnote 2).

Poznámka: V prílohe boli takisto ilustrované grafy vierohodnostných funkcií a logaritmic-kých vierohodnostných funkcií jednotlivých metód pre gamma rozdelenie s parametrom $a = 0.2$ a parametrom $p = 2$ v rozsahu $N = 30$ (väčší rozsah už bol pre vierohodnostné funkcie, najmä integrálnu, náročný na výpočet).

4.2 Porovnanie metód na základe simulovaných dát

Na základe predošlých výpočtov sa pokúsime zistiť, ktorá metóda je na dáta z rozdelenia gamma najvhodnejšia. Vygenerujeme si 100 a 1000 náhodných výberov z rozdelenia gamma s parametrom $a = 0.2$ a parametrom $p = 2$ o rozsahu 100 a 1000, vygenerujeme všetky 4 kombinácie týchto čísel. Simulácia 2 pozostáva z rovnakých kombinácií, ale náhodné výbery pochádzajú z rozdelenia gamma s parametrom $a = 0.6$ a parametrom $p = 5$. Parameter a považujeme za vedľajší parameter a parameter p za parameter záujmu. Následne budeme odhadovať parameter záujmu (aj keď jeho skutočnú hodnotu poznáme). Vypočítame $\widehat{MSE}(p)$ (empirická MSE pre parameter p), ktorá je definovaná vzťahom

$$\widehat{MSE}(p) = \frac{1}{N} \sum_{i=1}^N (\hat{p}_i - p)^2,$$

čiže odhad teoretickej $MSE(p)$, definovanej vzťahom

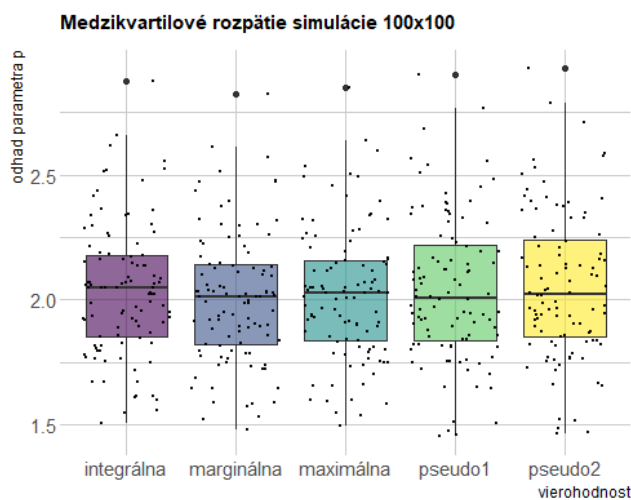
$$MSE(p) = \mathbf{E} \left[(\hat{p} - p)^2 \right],$$

zo simulácie. Vykreslíme si aj medzikvartilové rozpätie, kde bude vyznačený medián a jednotlivé odhady parametra p .

4.2.1 Dáta zo simulácií

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	2.021016	0.06962580
marginálna	2.003548	0.06785777
integrálna	2.038473	0.07201599
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	2.039632	0.08551491
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	2.055694	0.08870445

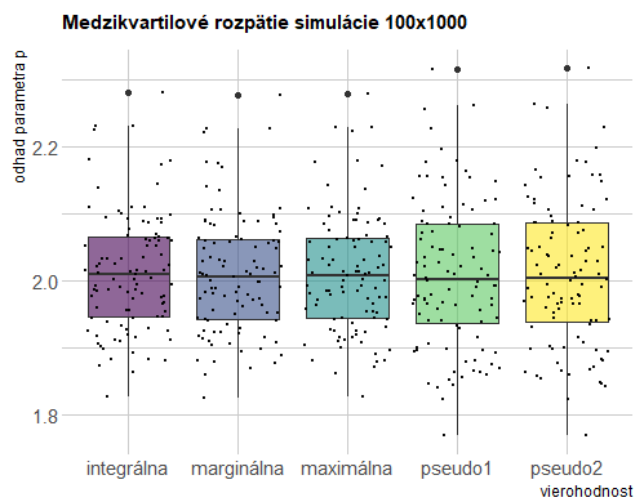
Tabuľka 4.1: 100×100 simulácií so seedom 333



Obr. 4.1: Medzikvartilové rozpätie simulácie 100×100

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	2.012468	0.007973338
marginálna	2.010731	0.007917864
integrálna	2.014205	0.008034864
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	2.009002	0.011845100
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	2.010563	0.011898530

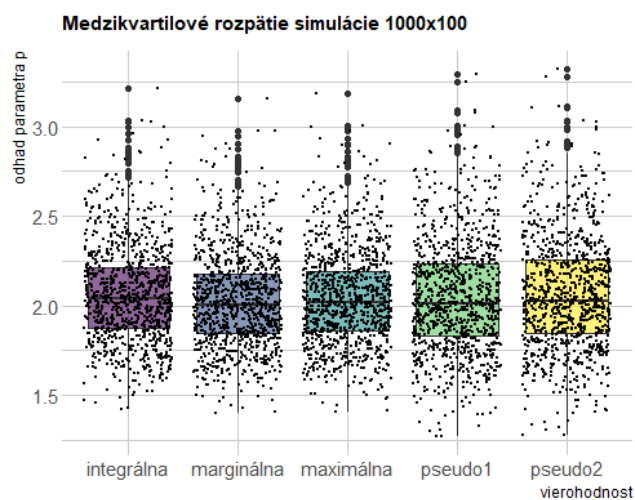
Tabuľka 4.2: 100×1000 simulácií so seedom 33333



Obr. 4.2: Medzikvartilové rozpätie simulácie 100×1000

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	2.045927	0.07346494
marginálna	2.028217	0.07077008
integrálna	2.063625	0.07680111
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	2.047671	0.10144180
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	2.063813	0.10520090

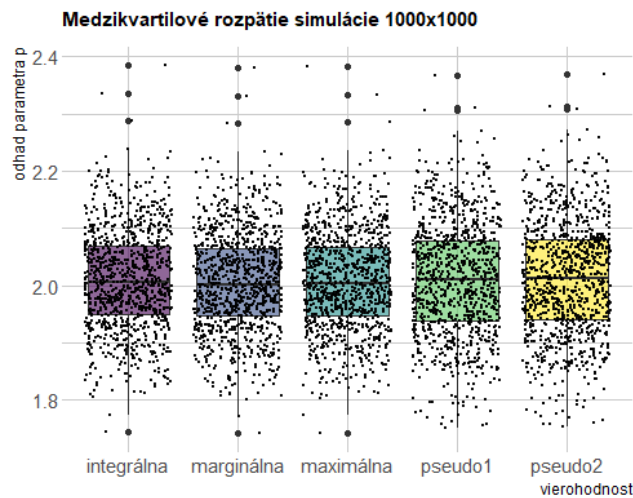
Tabuľka 4.3: 1000×100 simulácií so seedom 3333



Obr. 4.3: Medzikvartilové rozpätie simulácie 1000×100

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	2.007483	0.007575635
marginálna	2.005751	0.007538098
integrálna	2.009215	0.007619186
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	2.010012	0.009915463
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	2.011574	0.009968321

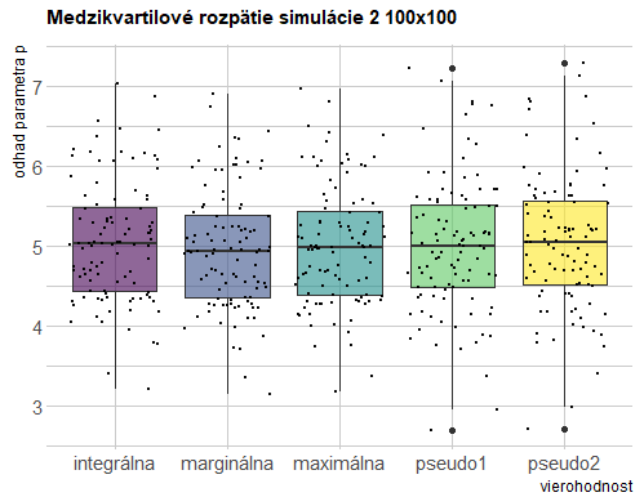
Tabuľka 4.4: 1000 × 1000 simulácií so seedom 33



Obr. 4.4: Medzikvartilové rozpätie simulácie 1000 × 1000

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	4.997621	0.5645202
marginálna	4.950736	0.5557625
integrálna	5.044496	0.5777879
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	4.997341	0.7041403
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	5.042961	0.7202118

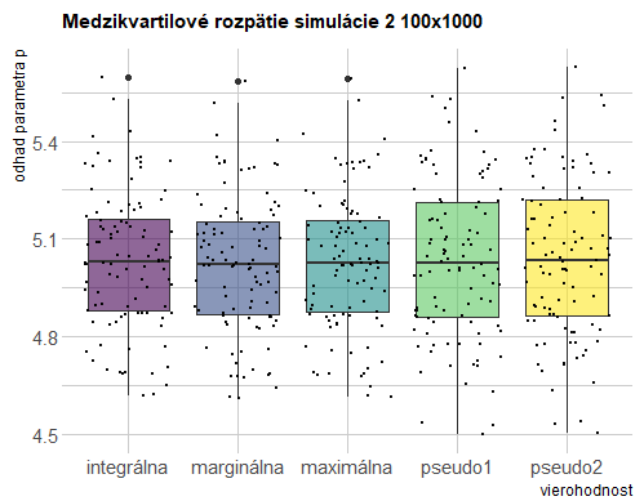
Tabuľka 4.5: 100 × 100 simulácií 2 so seedom 333



Obr. 4.5: Medzikvartilové rozpätie simulácie 2 100×100

vierochnost	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	5.029348	0.04402400
marginálna	5.024629	0.04368366
integrálna	5.034068	0.04440894
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	5.034463	0.05892012
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	5.039020	0.05937021

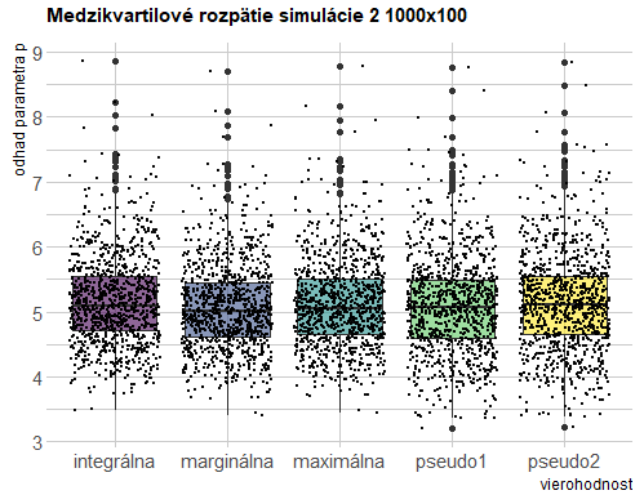
Tabuľka 4.6: 100×1000 simulácií 2 so seedom 33333



Obr. 4.6: Medzikvartilové rozpätie simulácie 2 100×1000

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	5.117277	0.4770170
marginálna	5.069202	0.4588751
integrálna	5.165341	0.4998742
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	5.090658	0.5795245
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	5.137213	0.6016888

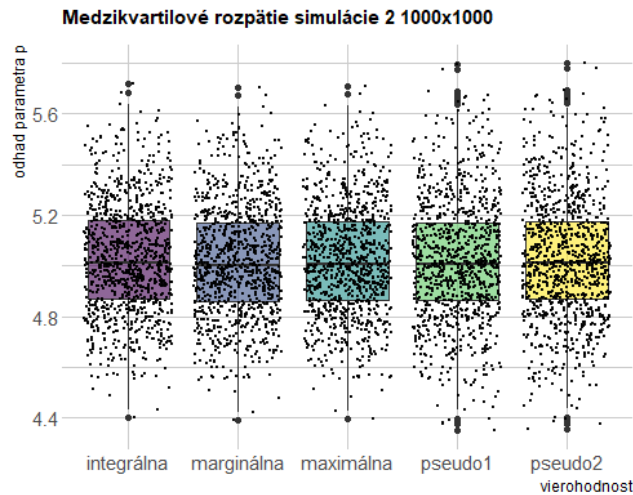
Tabuľka 4.7: 1000×100 simulácií 2 so seedom 3333



Obr. 4.7: Medzikvartilové rozpätie simulácie 2 1000×100

vierohodnosť	priemer odhadov (p)	$\widehat{MSE}(p)$
maximálna/profilová	5.022914	0.04808826
marginálna	5.018201	0.04779970
integrálna	5.027628	0.04842073
pseudo $\hat{a} = \frac{\bar{X}}{S_X^2}$	5.020409	0.05817557
pseudo $\hat{a} = \frac{\bar{X}}{\frac{N-1}{N} S_X^2}$	5.024952	0.05849684

Tabuľka 4.8: 1000×1000 simulácií 2 so seedom 33



Obr. 4.8: Medzikvartilové rozpätie simulácie 2 1000×1000

4.2.2 Výsledky simulácie

Zameriame sa na hodnoty \widehat{MSE} , pretože nám dávajú porovnanie odhadov zo simulovaných dát. V prvom rade zdôrazníme, že so zvyšujúcim rozsahom výberu sa \widehat{MSE} znižuje, čo sme samozrejme očakávali, pretože zo zvyšujúcim počtom dát narastá presnosť odhadu. Z výsledkov, ktoré sú uvedené v tabuľkách vyplýva, že v každej simulácii sa najlepšie darilo marginálnej vierohodnosti. S veľmi malými rozdielmi zaostávala maximálna/profilová vierohodnosť a integrálna vierohodnosť. Výrazne horšie obstála pseudo vierohodnosť, a to aj napriek tomu, že sme vyskúšali dva rôzne momentové odhady parametra a , čo dokazujú aj medzikvartilové rozpätia simulácií, v ktorých vidíme mediány a jednotlivé hodnoty odhadov parametra záujmu.

Prečo je dobré vymýšľať nové metódy vierohodnosti, keď u nás majú veľmi podobné výsledky maximálna, marginálna a integrálna vierohodnosť? Dôvod je ten, že maximálna, prípadne iná, vierohodnosť sa nie vždy (a to hlavne v komplikovanejších modeloch) dá numericky vypočítať, prípadne to trvá veľmi dlho, alebo je to numericky nestabilné.

5. Záver

Bakalárska práca sa venovala rôznym formám vierohodnosti a za pomoci dostupných zdrojov sme vytvorili ucelený súbor poznatkov s jednotným značením, ktorý slúži ako vhodná literatúra pre prvotné oboznámenie sa s problematikou vedľajších parametrov.

V úvode práce sme uviedli, ako budeme označovať parametre záujmu a vedľajšie parametre. Motivačným príkladom sme ilustrovali ako sú formulované problémy s obsahom vedľajších parametrov.

V teoretickej časti bakalárskej práce sme predstavili jednotlivé techniky a pre lepšie pochopenie ilustrovali na praktických príkladoch.

Ďalším prínosom práce je praktická časť, kde sú dopodrobna odvodené vzťahy na maximálnu/profilovú, marginálnu, integrálnu a pseudo vierohodnosť použité na gamma rozdelenie. Pre marginálnu vierohodnosť bola skonštruovaná transformácia, v ktorej sa podarilo jasne ukázať, že hustota $f(\mathbf{v}|\mathbf{u})$ neobsahuje veľa informácií o parametri p , teda je možné ju zanedbať. Následne sme pomocou programovacieho jazyka R, využili odvodené vzťahy a zisťovali na štyroch rôznych simuláciách, ktorá vierohodnostná metóda dosahuje najpresnejšie výsledky. Porovnávali sme \widehat{MSE} odhadov, a prehľadne ich spracovali do tabuliek. Vykreslili sme medzikvartilové rozpätie, kde bol vyznačený medián a jednotlivé odhady parametra p .

Téma maximálnej vierohodnosti za prítomnosti vedľajších parametrov sa dá ďalej rozšíriť o viaceré formy vierohodnosti ako napríklad reziduálna vierohodnosť. Bolo by zaujímavé porovnať výsledky simulačnej štúdie s inými vierohodnosťami, či už na gamma rozdelení, alebo na ľubovoľných dátach.

Zoznam použitej literatúry

- LANCASTER, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, **95**, 391–413.
- MAK, T. K. (1982). Estimation in the presence of incidental parameters. *Statistical Society of Canada*, **10**, 121–132.
- MILLAR, R. B. (2011). *Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB*. John Wiley and Sons, Chichester. ISBN 978-0-470-09483-9.
- MOREIRA, M. J. (2009). A maximum likelihood method for the incidental parameter problem. *The Annals of Statistics*, **37**(6A), 3660–3696.
- NEYMAN, J. a SCOTT, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, **16**(1), 1–32.
- PAWITAN, Y. (2001). *In All Likelihood: Statistical Modelling And Inference Using Likelihood*. Oxford University Press, New York. ISBN 978-0-19-850765-9.
- ØREN, J. (2021). Likelihood corrections for the incidental parameter problem in poisson distributed panel data. *University of Oslo*. URL https://www.duo.uio.no/bitstream/handle/10852/87899/JonasOeren_masteroppgave.pdf?sequence=1&isAllowed=y. Date of the latest version: Spring 2021; Accessed: July 21, 2022; Master Thesis.

Zoznam obrázkov

4.1	Medzikvartilové rozpätie simulácie 100×100	24
4.2	Medzikvartilové rozpätie simulácie 100×1000	25
4.3	Medzikvartilové rozpätie simulácie 1000×100	25
4.4	Medzikvartilové rozpätie simulácie 1000×1000	26
4.5	Medzikvartilové rozpätie simulácie $2 \ 100 \times 100$	27
4.6	Medzikvartilové rozpätie simulácie $2 \ 100 \times 1000$	27
4.7	Medzikvartilové rozpätie simulácie $2 \ 1000 \times 100$	28
4.8	Medzikvartilové rozpätie simulácie $2 \ 1000 \times 1000$	29

Zoznam tabuliek

4.1	100 × 100 simulácií so seedom 333	24
4.2	100 × 1000 simulácií so seedom 33333	24
4.3	1000 × 100 simulácií so seedom 3333	25
4.4	1000 × 1000 simulácií so seedom 33	26
4.5	100 × 100 simulácií 2 so seedom 333	26
4.6	100 × 1000 simulácií 2 so seedom 33333	27
4.7	1000 × 100 simulácií 2 so seedom 3333	28
4.8	1000 × 1000 simulácií 2 so seedom 33	28

A. Príloha

Príloha

Elektronická príloha obsahuje simulačnú štúdiu použitú v bakalárskej práci, bola vytvorená v programovacom jazyku R.