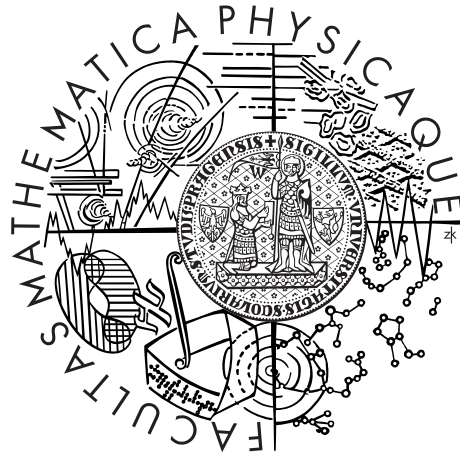


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## DIPLOMOVÁ PRÁCE



Tomáš Marcinko

# Konfidenční množiny v nelineární regresi

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: doc. RNDr. Karel Zvára CSc.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická  
statistika a ekonometrie

Praha 2012

Týmto by som sa chcel poďakovať doc. RNDr. Karlovi Zvárovi, CSc. za trpezlivosť, ochotu a cenné pripomienky pri tvorbe tejto diplomovej práce.

Prohlašuji, že jsem tuto diplomovou práci vypracoval 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 Praze dne 7.12.2012

Tomáš Marcinko

Název práce: Konfidenční množiny v nelineární regresii

Autor: Ing. Tomáš Marcinko

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: doc. RNDr. Karel Zvára, CSc.

Abstrakt: Cieľom tejto diplomovej práce je komplexné zhrnutie vlastností bodového odhadu parametrov normálneho nelineárneho modelu získaného metódou nelineárnych najmenších štvorcov a následný popis možností konštrukcie konfidenčných množín, resp. intervalov spoľahlivosti pre parametre tohto modelu. Vzhľadom k tomu, že na rozdiel od lineárneho modelu nie je možné tieto množiny či intervaly jednoducho a jednoznačne zostrojiť, musíme sa z praktického hľadiska uspokojiť hlavne s použitím približných metód. Súčasťou práce je i simulačná štúdia, ktorá porovnáva odhady pravdepodobností pokrytia skutočnej hodnoty parametrov jednotlivými konfidenčnými množinami a intervalmi spoľahlivosti pre rôzne stupne nelinearity modelu a rôzne rozsahy výberu. Ukázalo sa, že pri zanedbateľnej vnútornej krivosti modelu sa javí ako najvhodnejšia vierohodnostná konfidenčná množina.

Klíčová slova: nelineárny regresný model, vlastností bodového odhadu parametrov, konfidenčné množiny, parametrická a vnútorná krivosť

Title: Confidence regions in nonlinear regression

Author: Ing. Tomáš Marcinko

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Karel Zvára, CSc.

Abstract: The aim of this thesis is a comprehensive description of the properties of a nonlinear least squares estimator for a nonlinear regression model with normally distributed errors and thorough development of various methods for constructing confidence regions and confidence intervals for the parameters of the nonlinear model. Due to the fact that, unlike the case of linear models, there is no easy way to construct an exact confidence region for the parameters, most of these methods are only approximate. A short simulation study comparing observed coverage of various confidence regions and confidence intervals for models with different curvatures and sample sizes is also included. In case of negligible intrinsic curvature the use of likelihood-ratio confidence regions seems the most appropriate.

Keywords: nonlinear regression, properties of parameter estimates, confidence regions, parameter-effects and intrinsic curvature

# Obsah

<b>Úvod</b>	<b>1</b>
<b>1 Nelineárny model</b>	<b>3</b>
1.1 Motivácia	3
1.2 Bodový odhad parametrov	4
1.2.1 Metóda nelineárnych najmenších štvorcov	4
1.2.2 Metóda maximálnej vierohodnosti	6
1.2.3 Predpoklady konzistencie	8
1.2.4 Predpoklady asymptotickej normality	11
1.2.5 Boxov odhad vychýlenia	15
1.3 Miery krivosti	16
1.4 Nelineárna regresia v $\mathbb{R}$	19
<b>2 Konfidenčné množiny v nelineárnom modeli</b>	<b>23</b>
2.1 Waldovská konfidenčná množina	23
2.2 Vierohodnostná konfidenčná množina	25
2.3 Skórová konfidenčná množina	28
2.4 Asymptotické konfidenčné množiny	30
2.5 Presné konfidenčné množiny	33
2.6 Konfidenčné množiny v $\mathbb{R}$	35
<b>3 Simulačná štúdia</b>	<b>40</b>
3.1 Návrh simulačnej štúdie	40
3.2 Výsledky simulačnej štúdie	41
3.2.1 Logistický rastový model bez replikácií	41
3.2.2 Logistický rastový model s replikáciami	45
3.2.3 Model s vysokou vnútornou krivosťou	47
<b>Záver</b>	<b>49</b>
<b>Zoznam literatúry</b>	<b>51</b>
<b>Obsah priloženého CD</b>	<b>53</b>

„Although this may seem a paradox, all exact science is dominated by the idea of approximation.“

---

— Bertrand Russell

# Úvod

Pri riešení reálnych problémov v mnohých oblastiach vedy sa často stretávame s potrebou vysvetliť závislosť náhodnej veličiny nášho záujmu (vysvetľovaná premenná) na jednej alebo viacerých vopred daných veličinách (vysvetľujúce premenné). V matematickej štatistike či ekonometrii veľmi často popisujeme túto závislosť pomocou lineárneho modelu, čím implicitne prijímame predpoklad, že vplyv vysvetľujúcich premenných alebo ich transformácií na vysvetľovanú premennú je lineárny.

V reálnych aplikáciách vo fyzike, chémii, ekonómii či iných odvetviach sa však niekedy stáva, že vzťah medzi skúmanými premennými nemožno dostatočne dobre popísať pomocou lineárneho modelu, prípadne z príslušnej teórie vieme, že vzťah medzi sledovanými premennými je vskutku nelineárny. I v prípadoch, kedy aproximácia lineárnym modelom poskytuje dostatočne dobré výsledky, môže byť dôvodom pre použitie nelineárneho modelu zachovanie interpretácie parametrov teoretického modelu, ktoré chceme odhadnúť.

V tejto práci sa budeme zaoberať vlastnosťami bodového odhadu parametrov normálneho nelineárneho modelu, ktorý získame najčastejšie používanou metódou nelineárnych najmenších štvorcov, ako aj následnou konštrukciou konfidenčných množín, resp. intervalov spoľahlivosti, ktorá je na vlastnostiach bodového odhadu založená. Vzhľadom k tomu, že možnosti výstavby konfidenčných množín pre parametre nelineárneho modelu je hneď niekoľko, budú uvedené i výsledky simulačnej štúdie vyhodnocujúce kvalitu jednotlivých metód.

Prvá kapitola sa venuje hlavne bodovému odhadu parametrov nelineárneho modelu získanému metódou nelineárnych najmenších štvorcov, ktorý je v prípade normálneho modelu zhodný s maximálne vierohodným odhadom, a jeho základným vlastnostiam. Ukážeme, že pri splnení istých predpokladov regularity je tento odhad konzistentný a má asymptoticky normálne rozdelenie. Takisto sa budeme v krátkosti venovať problematike krivosti modelu, ktorá môže mať v prípade konečného výberu značný vplyv na validitu uvedených asymptotických či približných vlastností.

V druhej kapitole sú odvodené základné i menej používané metódy konštrukcie konfidenčných množín, resp. intervalov spoľahlivosti, pre parametre normálneho nelineárneho modelu. Okrem azda najčastejšie používanej približnej metódy založenej na Waldovom asymptotickom teste, si pomerne detailne popíšeme i vierohodnostné či skórové konfidenčné množiny, ktoré sú rovnako založené na známych asymptotických testových štatistikách. Takisto sa na základe asymptotickej teórie maximálnej vierohodnosti pokúsime odvodiť tzv. asymptotické konfidenčné množiny a v krátkosti sa budeme venovať i presnej

metóde konštrukcie konfidenčných množín, ktorá je založená na tzv. Hartleyho metóde. V závere prvých dvoch kapitol si na jednoduchom príklade ukážeme aplikáciu uvedených poznatkov v programovom prostredí R.

Záverečná tretia kapitola bude prezentovať výsledky simulačnej štúdie, ktorej cieľom je popísať vplyv hlavne parametrickej krivosti nelineárneho modelu, ako aj rozsahu výberu, na validitu príslušných konfidenčných množín s ohľadom na pravdepodobnosť pokrytia skutočného parametra modelu príslušnou konfidenčnou množinou. Pomocou výsledkov tejto simulačnej štúdie sa pokúsime rozhodnúť, ktorá z odvodených metód konštrukcie konfidenčných množín pre parametre nelineárneho modelu je z praktického hľadiska najvhodnejšia.

# 1. Nelineárny model

## 1.1 Motivácia

Predpokladajme, že stredné hodnoty nekorelovaných náhodných veličín  $Y_1, Y_2, \dots, Y_n$  môžeme vhodne popísať na základe modelu

$$E Y_i = f(x_{i1}, \dots, x_{ik}; \boldsymbol{\theta}),$$

kde  $x_{i1}, \dots, x_{ik}$  sú známe konštanty,  $k$  je počet regresorov a  $\boldsymbol{\theta}$  je  $p$ -rozmerný vektor parametrov.

Na rozdiel od lineárneho modelu však budeme uvažovať model, ktorý je v parametroch nelineárny, t.j. pre ktorý platí, že aspoň jedna z parciálnych derivácií regresnej funkcie  $f$  podľa jednotlivých parametrov nie je konštantná, resp. je funkciou aspoň jedného z parametrov. Jednoduchým príkladom modelu nelineárneho v parametroch je napríklad Cobb-Douglasov model produkčnej funkcie v obecnom tvare

$$f(x_{i1}, \dots, x_{ik}; \boldsymbol{\theta}) = \theta_0 x_{i1}^{\theta_1} \cdots x_{ik}^{\theta_k}.$$

Ak by sme na túto funkciu použili logaritmickú transformáciu, jednoducho by sme mohli získať model, ktorý by už bol v parametroch lineárny, a teda by bolo možné pre odhad modelu využiť techniky používané v prípade lineárneho modelu.

Možnosť správneho použitia linearizovaného modelu však úzko súvisí s charakterom náhodnej zložky v modeli. Ak by sme v uvedenom prípade uvažovali model s multiplikatívnou náhodnou zložkou v tvare

$$Y_i = \theta_0 x_{i1}^{\theta_1} \cdots x_{ik}^{\theta_k} e^{\varepsilon_i},$$

kde pre nekorelovanú náhodnú zložku platí  $E \varepsilon_i = 0$  a  $\text{var } \varepsilon_i = \sigma^2$ , tak po logaritmickej transformácii dostaneme lineárny model s náhodnou zložkou, ktorá spĺňa klasické požiadavky na ňu kladené. Ak by sme však uvažovali aditívnu náhodnú zložku (u ktorej je vzhľadom k charakteru vysvetľovanej premennej  $Y_i$  prirodzené predpokladať, že nerovnosť  $E Y_i + \varepsilon_i > 0$ , prípadne  $|\varepsilon_i| < E Y_i$ , platí s pravdepodobnosťou 1) a teda model v tvare

$$Y_i = \theta_0 x_{i1}^{\theta_1} \cdots x_{ik}^{\theta_k} + \varepsilon_i,$$

kde platí  $E \varepsilon_i = 0$  a  $\text{var } \varepsilon_i = \sigma^2$ , tak po linearizácii dostávame model

$$\log Y_i = \log \theta_0 + \theta_1 \log x_{i1} + \dots + \theta_k \log x_{ik} + \varepsilon_i^*,$$

ktorého náhodnú zložku možno vyjadriť ako

$$\varepsilon_i^* = \log(1 + \varepsilon_i (E Y_i)^{-1}).$$

Aproximáciou Taylorovým polynómom prvého rádu v bode  $E\varepsilon_i = 0$  dostaneme, že pre náhodnú zložku v transformovanom modeli  $\varepsilon_i^*$  približne platí  $E\varepsilon_i^* \approx E[\varepsilon_i (EY_i)^{-1}] = 0$  a  $\text{var} \varepsilon_i^* \approx \text{var}[\varepsilon_i (EY_i)^{-1}] = \sigma^2 (EY_i)^{-2}$ , a teda rozptyl náhodnej zložky  $\varepsilon_i^*$  už závisí na veľkosti  $EY_i$ . V prípade, že by sme u náhodnej zložky  $\varepsilon_i$  navyše uvažovali normálne rozdelenie, náhodná zložka linearizovaného modelu  $\varepsilon_i^*$  by normálne rozdelená byť nemohla. Napriek tomu, že z uvedených dôvodov nie je vhodné použiť odhady linearizovaného modelu bezprostredne, môžu byť tieto odhady využité vo forme prvotnej aproximácie pre iteratívny výpočet odhadu vektoru parametrov.

Je zrejmé, že idea linearizačnej transformácie nemôže byť v niektorých prípadoch vôbec použitá. Jednoduchým príkladom nelineárneho modelu, ktorý nemôže byť jednoducho prevedený na model lineárny, je

$$f(x_i; \boldsymbol{\theta}) = \theta_0 + \theta_1 x_i^{\theta_2}.$$

Nelineárny model, ktorý nie je možné linearizovať pomocou vhodnej transformácie, nazývame *vnútorne nelineárnym* a v takomto prípade je už pre odhad modelu nutné využiť techník nelineárnej regresie.

## 1.2 Bodový odhad parametrov

V ďalšom texte teda budeme pre vysvetľované náhodné veličiny  $Y_1, \dots, Y_n$  uvažovať model nelineárny v parametroch v tvare

$$Y_i = f(x_{i1}, \dots, x_{ik}; \boldsymbol{\theta}^*) + \varepsilon_i, \quad (1.1)$$

kde  $f$  je až na parametre známa funkcia,  $x_{i1}, \dots, x_{ik}$  sú známe konštanty,  $\boldsymbol{\theta}^*$  je skutočná hodnota neznámeho  $p$ -rozmerného vektorového parametra  $\boldsymbol{\theta}$  a náhodné zložky  $\varepsilon_i$  sú nezávislé a rovnako rozdelené s rozdelením  $N(0, \sigma^2)$ .

### 1.2.1 Metóda nelineárnych najmenších štvorcov

Jedným z najčastejšie používaných odhadov vektoru parametrov  $\boldsymbol{\theta}$  je odhad získaný *metódou nelineárnych najmenších štvorcov*, ktorý obdobne ako v lineárnom modeli minimalizuje reziduálny súčet štvorcov v tvare

$$\mathcal{S}(\boldsymbol{\theta}) = \sum_{i=1}^n [Y_i - f(x_{i1}, \dots, x_{ik}; \boldsymbol{\theta})]^2 = \sum_{i=1}^n [Y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2. \quad (1.2)$$

Na rozdiel od lineárnych najmenších štvorcov však nie je možné pre obecnú regresnú funkciu  $f$  nájsť príslušný odhad  $\hat{\boldsymbol{\theta}}$  minimalizujúci (1.2) analyticky, ale odhad je spravidla založený na opakovanej lineárnej aproximácii a je počítaný iteračne (napr. Gaussovým-Newtonovým algoritmom). Funkcia  $\mathcal{S}(\boldsymbol{\theta})$  navyše môže mať okrem globálneho minima i minima lokálne, čo môže spôsobovať ďalšie komplikácie v prípade nevhodnej počiatocnej aproximácie pre účely iteračného algoritmu.

Predpokladajme, že  $f(\mathbf{x}_i; \boldsymbol{\theta})$  sú dvakrát spojitely diferencovateľné a  $\hat{\boldsymbol{\theta}}$  je vnútorným bodom parametrického priestoru  $\Omega$ . Potom je pre bod minimalizujúci reziduálny súčet štvorcov vektoru parciálnych derivácií  $\mathcal{S}(\boldsymbol{\theta})$  nulový, t.j. platí

$$\left. \frac{\partial \mathcal{S}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\hat{\boldsymbol{\theta}}} \stackrel{\text{ozn}}{=} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\hat{\boldsymbol{\theta}}) = \mathbf{0}, \quad (1.3)$$

kde sme si za účelom zjednodušenia zaviedli alternatívne označenie hodnoty parciálnej derivácie v danom bode.

Pre nelineárny model budeme používať i nasledovné značenie: vektor hodnôt regresnej funkcie pri vektore parametrov  $\boldsymbol{\theta}$  budeme značiť ako

$$\mathbf{f}(\boldsymbol{\theta}) = (f(\mathbf{x}_1; \boldsymbol{\theta}), \dots, f(\mathbf{x}_n; \boldsymbol{\theta}))'$$

a maticu prvých derivácií regresnej funkcie podľa  $\boldsymbol{\theta}$  typu  $n \times p$  ako

$$\mathbf{F}_\bullet(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{f}(\boldsymbol{\theta}) = \left( \frac{\partial}{\partial \theta_j} f(\mathbf{x}_i; \boldsymbol{\theta}) \right).$$

O matici  $\mathbf{F}_\bullet(\boldsymbol{\theta})$  budeme predpokladať, že má prinajmenšom v okolí bodu  $\boldsymbol{\theta}^*$  hodnotu  $p$ . Takisto budeme používať skrátenejší zápis  $\mathbf{f}^* = \mathbf{f}(\boldsymbol{\theta}^*)$  a  $\mathbf{F}_\bullet^* = \mathbf{F}_\bullet(\boldsymbol{\theta}^*)$ .

Pri danom značení môžeme prepísať zápis modelu (1.1) do maticového tvaru

$$\mathbf{Y} = \mathbf{f}^* + \boldsymbol{\varepsilon},$$

kde  $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ , a funkciu (1.2) vyjadrujúcu reziduálny súčet štvorcov do tvaru

$$\mathcal{S}(\boldsymbol{\theta}) = [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})] = \|\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})\|^2. \quad (1.4)$$

Z požiadavky nulovosti vektoru parciálnych derivácií (1.3) dostávame *normálne rovnice* pre nelineárny model v tvare

$$-\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\hat{\boldsymbol{\theta}}) = \mathbf{F}_\bullet(\hat{\boldsymbol{\theta}})' [\mathbf{Y} - \mathbf{f}(\hat{\boldsymbol{\theta}})] = \mathbf{0}, \quad (1.5)$$

Pri skúmaní vlastností odhadu metódou nelineárnych najmenších štvorcov budeme využívať myšlienku lineárnej aproximácie. Tá vychádza z predpokladu, že pre  $\boldsymbol{\theta}$  dostatočne blízke skutočnej hodnote  $\boldsymbol{\theta}^*$  môžeme využiť aproximácie Taylorovým polynómom prvého radu

$$\mathbf{f}(\boldsymbol{\theta}) \approx \mathbf{f}^* + \mathbf{F}_\bullet^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*). \quad (1.6)$$

Navyše vzhľadom k predpokladanej spojitely diferencovateľnosti  $f(\mathbf{x}_i; \boldsymbol{\theta})$  približne platí aj  $\mathbf{F}_\bullet(\hat{\boldsymbol{\theta}}) \approx \mathbf{F}_\bullet^*$ , pričom dosadením týchto aproximácií do normálnych rovníc (1.5) dostávame približný vzťah

$$\mathbf{F}_\bullet^{*'} [\mathbf{Y} - \mathbf{f}^* - \mathbf{F}_\bullet^* (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)] = \mathbf{F}_\bullet^{*'} [\boldsymbol{\varepsilon} - \mathbf{F}_\bullet^* (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)] \approx \mathbf{0},$$

odkiaľ možno jednoducho získať približné vyjadrenie odhadu nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}$  v tvare

$$\hat{\boldsymbol{\theta}} \approx \boldsymbol{\theta}^* + (\mathbf{F}_\bullet^{*'} \mathbf{F}_\bullet^*)^{-1} \mathbf{F}_\bullet^{*'} \boldsymbol{\varepsilon}. \quad (1.7)$$

Analogicky pre reziduálny súčet štvorcov (1.4) platí

$$\mathcal{S}(\boldsymbol{\theta}) \approx \|\mathbf{Y} - \mathbf{f}^* - \mathbf{F}_*^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*)\|^2. \quad (1.8)$$

Dosadením  $\mathbf{F}_*^* (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) \approx \mathbf{F}_*^* (\mathbf{F}_*^{*'} \mathbf{F}_*^*)^{-1} \mathbf{F}_*^{*'} \boldsymbol{\varepsilon} = \mathbf{H}_F^* \boldsymbol{\varepsilon}$ , kde  $\mathbf{H}_F^* = \mathbf{F}_*^* (\mathbf{F}_*^{*'} \mathbf{F}_*^*)^{-1} \mathbf{F}_*^{*'}$  je symetrická idempotentná matica, dostaneme priblíženie reziduálneho súčtu štvorcov (1.8) v bode  $\hat{\boldsymbol{\theta}}$  v tvare

$$\mathcal{S}(\hat{\boldsymbol{\theta}}) \approx \|(\mathbf{I}_n - \mathbf{H}_F^*) \boldsymbol{\varepsilon}\|^2 = \boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{H}_F^*) \boldsymbol{\varepsilon}. \quad (1.9)$$

Dôležité vlastnosti odhadu metódou nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}$  sú jeho konzistencia a asymptoticky normálne rozdelenie, ktoré však platia iba pri splnení istých predpokladov regularity uvedených v nasledujúcich podkapitolách. Pri splnení predpokladov konzistencie odhadu  $\hat{\boldsymbol{\theta}}$  platí i konzistencia odhadu rozptylu  $\sigma^2$  v tvare

$$S^2 = \frac{\mathcal{S}(\hat{\boldsymbol{\theta}})}{n - p}.$$

## 1.2.2 Metóda maximálnej vierohodnosti

Ďalším často využívaným odhadom vektoru parametrov  $\boldsymbol{\theta}$  je odhad získaný metódou maximálnej vierohodnosti. Ak poznáme hustotu nezávislých náhodných veličín  $Y_1, \dots, Y_n$ , tak maximálne vierohodný odhad získame maximalizáciou logaritmickej vierohodnostnej funkcie modelu. V prípade, že hustotu  $Y_i$  môžeme vyjadriť ako  $p(y_i; \boldsymbol{\theta}, \sigma^2)$ , tak logaritmická vierohodnostná funkcia má tvar

$$L(\boldsymbol{\theta}, \sigma^2) = \log \prod_{i=1}^n p(y_i; \boldsymbol{\theta}, \sigma^2) = \sum_{i=1}^n \log p(y_i; \boldsymbol{\theta}, \sigma^2).$$

Avšak vzhľadom k tomu, že uvažujeme normálne rozdelenú náhodnú zložku, platí, že odhad vektoru parametrov  $\hat{\boldsymbol{\theta}}$  získaný metódou nelineárnych najmenších štvorcov je zhodný s maximálne vierohodným odhadom.

**Tvrdenie 1.1.** *Nech platí nelineárny model (1.1), kde  $\varepsilon_i \stackrel{iid}{\sim} \mathbf{N}(0, \sigma^2)$ . Potom maximálne vierohodným odhadom vektoru parametrov  $\boldsymbol{\theta}^*$  je  $\hat{\boldsymbol{\theta}}$  a maximálne vierohodným odhadom rozptylu  $\sigma^2$  je  $\mathcal{S}(\hat{\boldsymbol{\theta}})/n$ .*

*Dôkaz.* Ak platí normálny nelineárny model v tvare (1.1), tak môžeme logaritmickú vierohodnostnú funkciu vyjadriť ako

$$\begin{aligned} L(\boldsymbol{\theta}, \sigma^2) &= \sum_{i=1}^n \log \left( \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[Y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2}{2\sigma^2}} \right) \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n [Y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2 \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \mathcal{S}(\boldsymbol{\theta}). \end{aligned} \quad (1.10)$$

Pre maximálne virohodné odhady parametrov normálneho nelineárneho modelu  $\hat{\boldsymbol{\theta}}$  a  $\hat{\sigma}^2$  zrejme platí

$$\frac{\partial}{\partial \boldsymbol{\theta}} L(\hat{\boldsymbol{\theta}}, \hat{\sigma}^2) = 0 \quad \text{a} \quad \frac{\partial}{\partial \sigma^2} L(\hat{\boldsymbol{\theta}}, \hat{\sigma}^2) = 0.$$

Je zjavné, že pri fixovaní  $\sigma^2$  je maximum logaritmickej virohodnostnej funkcie (1.10) dosiahnuté v bode minima reziduálneho súčtu štvorcov, teda v bode  $\hat{\boldsymbol{\theta}}$ . Naopak pri fixovaní vektoru  $\boldsymbol{\theta}$  v bode  $\hat{\boldsymbol{\theta}}$  platí

$$\frac{\partial}{\partial \sigma^2} L(\hat{\boldsymbol{\theta}}, \hat{\sigma}^2) = -\frac{n}{2\hat{\sigma}^2} + \frac{\mathcal{S}(\hat{\boldsymbol{\theta}})}{2(\hat{\sigma}^2)^2} = 0,$$

z čoho vyplýva, že

$$\hat{\sigma}^2 = \frac{\mathcal{S}(\hat{\boldsymbol{\theta}})}{n}.$$

Overíme, že odhady  $\hat{\boldsymbol{\theta}}$  a  $\hat{\sigma}^2 = \mathcal{S}(\hat{\boldsymbol{\theta}})/n$  skutočne maximalizujú logaritmickejšiu virohodnostnú funkciu  $L(\boldsymbol{\theta}, \sigma^2)$ . Vzhľadom k tomu, že  $\mathcal{S}(\boldsymbol{\theta}) \geq \mathcal{S}(\hat{\boldsymbol{\theta}}) = n\hat{\sigma}^2$ , platí

$$\begin{aligned} L(\hat{\boldsymbol{\theta}}, \hat{\sigma}^2) - L(\boldsymbol{\theta}, \sigma^2) &= -\frac{n}{2} \log \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2} \mathcal{S}(\hat{\boldsymbol{\theta}}) + \frac{n}{2} \log \sigma^2 + \frac{1}{2\sigma^2} \mathcal{S}(\boldsymbol{\theta}) \\ &= -\frac{n}{2} \log \frac{\hat{\sigma}^2}{\sigma^2} - \frac{n\hat{\sigma}^2}{2\hat{\sigma}^2} + \frac{1}{2\sigma^2} \mathcal{S}(\boldsymbol{\theta}) \\ &\geq -\frac{n}{2} \log \frac{\hat{\sigma}^2}{\sigma^2} - \frac{n}{2} + \frac{n\hat{\sigma}^2}{2\sigma^2} \\ &= \frac{n}{2} \left( \frac{\hat{\sigma}^2}{\sigma^2} - 1 - \log \frac{\hat{\sigma}^2}{\sigma^2} \right) \\ &\geq 0, \end{aligned}$$

pretože pre všetky  $a > 0$  platí  $a - 1 \geq \log a$  a zároveň  $\hat{\sigma}^2 > 0$  s. j. □

S využitím vyjadrenia logaritmickej virohodnostnej funkcie normálneho nelineárneho modelu (1.10) môžeme zároveň vyjadriť Fisherovu informačnú maticu pre neznámy vektor parametrov  $(\boldsymbol{\theta}^*, \sigma^2)'$  ako

$$\mathbf{J}_n(\boldsymbol{\theta}^*, \sigma^2) = \mathbf{E} \begin{pmatrix} -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} & -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \boldsymbol{\theta} \partial \sigma^2} \\ -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \sigma^2 \partial \boldsymbol{\theta}'} & -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix},$$

kde postupným derivovaním (1.10) a (1.4) dostaneme

$$\begin{aligned} \mathbf{E} \left[ -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] &= \mathbf{E} \left[ \frac{1}{2\sigma^2} \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathcal{S}(\boldsymbol{\theta}^*) \right] = \mathbf{E} \left[ -\frac{1}{\sigma^2} \frac{\partial \mathbf{F} \cdot (\boldsymbol{\theta})' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}^*} \right] \\ &= \frac{1}{\sigma^2} \mathbf{F} \cdot (\boldsymbol{\theta}^*)' \mathbf{F} \cdot (\boldsymbol{\theta}^*), \\ \mathbf{E} \left[ -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \boldsymbol{\theta} \partial \sigma^2} \right] &= \mathbf{E} \left[ -\frac{1}{2\sigma^4} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\boldsymbol{\theta}^*) \right] = \mathbf{E} \left[ \frac{1}{\sigma^4} \mathbf{F} \cdot (\boldsymbol{\theta}^*)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}^*)] \right] = \mathbf{0}, \\ \mathbf{E} \left[ -\frac{\partial^2 L(\boldsymbol{\theta}^*, \sigma^2)}{\partial \sigma^2 \partial \sigma^2} \right] &= \mathbf{E} \left[ -\frac{n}{2\sigma^4} + \frac{1}{\sigma^8} \mathcal{S}(\boldsymbol{\theta}^*) \right] = -\frac{n}{2\sigma^4} + \frac{n\sigma^2}{\sigma^8} = \frac{n}{2\sigma^4}. \end{aligned}$$

Príslušnú Fisherovu informačnú maticu teda možno vyjadriť ako

$$\mathbf{J}_n(\boldsymbol{\theta}^*, \sigma^2) = \begin{pmatrix} \frac{1}{\sigma^2} \mathbf{F}_\bullet(\boldsymbol{\theta}^*)' \mathbf{F}_\bullet(\boldsymbol{\theta}^*) & \mathbf{0} \\ \mathbf{0}' & \frac{n}{2\sigma^4} \end{pmatrix}. \quad (1.11)$$

Alternatívne možnosti odhadu vektoru parametrov  $\boldsymbol{\theta}^*$  nájde záujemca napríklad v monografii [18].

### 1.2.3 Predpoklady konzistencie

V nasledovných dvoch podkapitolách si zhrnieme postačujúce predpoklady pre konzistenciu a asymptotickú normalitu odhadu  $\hat{\boldsymbol{\theta}}$ . Pri výčte týchto predpokladov budeme vychádzať ako z prehľadu, ktorý uvádzajú Seber a Wild [18] či Wu [20], tak i z pôvodných článkov Jennrich [14] a Malinvaud [15].

Uvažujme opäť model v tvare

$$Y_i = f(\mathbf{x}_i; \boldsymbol{\theta}^*) + \varepsilon_i, \quad (1.12)$$

avšak zatiaľ bez bližších predpokladov o jednotlivých zložkách modelu.

Prvé rigorózne dokázané predpoklady konzistencie publikoval Jennrich [14]. Jeho závery aplikované na model (1.12) môžeme zhrnúť do nasledovnej vety.

**Veta 1.2** (Predpoklady konzistencie I). *Nech pre nelineárny model (1.12) platia nasledovné predpoklady:*

- J1. náhodné zložky  $\varepsilon_i$  sú nezávislé a rovnako rozdelené s nulovou strednou hodnotou a neznámym konečným rozptylom  $\sigma^2 > 0$ ;
- J2. skutočná hodnota vektoru neznámych parametrov  $\boldsymbol{\theta}^*$  patrí do kompaktného parametrického priestoru  $\Omega \subset \mathbb{R}^p$ ;
- J3.  $f(\mathbf{x}_i; \boldsymbol{\theta})$  sú známe spojité funkcie vektoru parametrov  $\boldsymbol{\theta} \in \Omega$ ;
- J4.  $n^{-1}D_n(\boldsymbol{\theta}, \boldsymbol{\theta}_1)$ , kde  $D_n(\boldsymbol{\theta}, \boldsymbol{\theta}_1) = \sum_{i=1}^n [f(\mathbf{x}_i; \boldsymbol{\theta}) - f(\mathbf{x}_i; \boldsymbol{\theta}_1)]^2$  a  $\boldsymbol{\theta}, \boldsymbol{\theta}_1 \in \Omega$ , konverguje rovnomerne ku spojitým funkciám  $D(\boldsymbol{\theta}, \boldsymbol{\theta}_1)$ , a zároveň platí  $D(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = 0$  vtedy a len vtedy, ak  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ .

Potom  $\hat{\boldsymbol{\theta}}_n$  a  $\hat{\sigma}_n^2 = \mathcal{S}(\hat{\boldsymbol{\theta}}_n)/n$  sú silne konzistentnými odhadmi neznámych parametrov  $\boldsymbol{\theta}^*$  a  $\sigma^2$ .

*Dôkaz.* Viď Seber a Wild [18], ods. 12.2.2. □

**Poznámka 1.3.** Povieme, že odhad  $\hat{\boldsymbol{\theta}}_n$  je silne konzistentným odhadom parametra  $\boldsymbol{\theta}^*$ , ak  $\hat{\boldsymbol{\theta}}_n \xrightarrow[n \rightarrow \infty]{s.j.} \boldsymbol{\theta}^*$ .

Intuitívne vysvetlenie postačiteľnosti predpokladu (J4) uvádza Amemiya [1]. Idea je založená na skutočnosti, že  $\hat{\boldsymbol{\theta}}_n$  je (slabo) konzistentným odhadom

neznámeho parametra  $\boldsymbol{\theta}^*$ , ak vektor  $\boldsymbol{\theta}^*$  ako jediný minimalizuje limitu v pravdepodobnosti výrazu  $n^{-1}\mathcal{S}(\boldsymbol{\theta})$ . Výraz  $n^{-1}\mathcal{S}(\boldsymbol{\theta})$  totiž môžeme vyjadriť ako

$$\begin{aligned}\frac{1}{n}\mathcal{S}(\boldsymbol{\theta}) &= \frac{1}{n} \sum_{i=1}^n [Y_i - f(\mathbf{x}_i; \boldsymbol{\theta})]^2 \\ &= \frac{1}{n} \sum_{i=1}^n [Y_i - f(\mathbf{x}_i; \boldsymbol{\theta}^*) + f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \boldsymbol{\theta})]^2 \\ &= \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 + \frac{2}{n} \sum_{i=1}^n \varepsilon_i [f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \boldsymbol{\theta})] + \frac{1}{n} D_n(\boldsymbol{\theta}^*, \boldsymbol{\theta}).\end{aligned}$$

Ďalej je zřejmé, že zo silného zákona veľkých čísel platí  $n^{-1} \sum_{i=1}^n \varepsilon_i^2 \xrightarrow[n \rightarrow \infty]{s.j.} \sigma^2$  a z Čebyševovej nerovnosti dostávame

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \varepsilon_i [f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \boldsymbol{\theta})] \right| > \epsilon \right\} \leq \frac{\sigma^2}{\epsilon^2 n^2} \sum_{i=1}^n [f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \boldsymbol{\theta})]^2.$$

Vzhľadom k tomu, že pravá strana nerovnosti konverguje pri platnosti predpokladu (J4) k nule, tak  $n^{-1} \sum_{i=1}^n \varepsilon_i [f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \boldsymbol{\theta})] \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0$ . Za platnosti predpokladu (J4) navyše platí, že  $n^{-1} D_n(\boldsymbol{\theta}^*, \boldsymbol{\theta}) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} D(\boldsymbol{\theta}^*, \boldsymbol{\theta})$ , a  $D(\boldsymbol{\theta}^*, \boldsymbol{\theta}) = 0$  vtedy a len vtedy, ak  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ . Z toho vyplýva, že vektor  $\boldsymbol{\theta}^*$  ako jediný minimalizuje limitu v pravdepodobnosti výrazu  $n^{-1}\mathcal{S}(\boldsymbol{\theta})$ , a teda  $\hat{\boldsymbol{\theta}}_n \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \boldsymbol{\theta}^*$ .

Nutnú a postačujúcu podmienku konzistencie odhadu  $\hat{\boldsymbol{\theta}}$  uvádza Wu [20].

**Veta 1.4.** *Nech pre nelineárny model (1.12) platia predpoklady (J1)–(J3), pričom parametrický priestor  $\Omega \subset \mathbb{R}^p$  je navyše konečný. Potom podmienka*

$$D_n(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = \sum_{i=1}^n [f(\mathbf{x}_i; \boldsymbol{\theta}) - f(\mathbf{x}_i; \boldsymbol{\theta}^*)]^2 \xrightarrow[n \rightarrow \infty]{} \infty \quad \forall \boldsymbol{\theta} \neq \boldsymbol{\theta}^* \quad (1.13)$$

je nutnou podmienkou pre silnú konzistenciu odhadu  $\hat{\boldsymbol{\theta}}$ .

Ak navyše platí, že náhodné veličiny  $\varepsilon_i$  nie sú ani zhora ani zdola obmedzené, potom  $D_n(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \rightarrow \infty$  pre všetky  $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$  je zároveň i postačujúcou podmienkou pre silnú konzistenciu odhadu  $\hat{\boldsymbol{\theta}}$ .

*Dôkaz.* Vid' Wu [20]. □

**Poznámka 1.5.** Wu [20] takisto uvádza, že predpoklad konečnosti parametrického priestoru  $\Omega$  nie je z praktického hľadiska obmedzujúci, pretože pri samotnom výpočte možno minimum reziduálneho súčtu štvorcov hľadať iba na konečnej množine hodnôt zvolením maximálneho počtu desiatinných miest.

Veta 1.4 nám okrem iného hovorí, že ak nosič príslušných náhodných veličín nie je obmedzený (platí napríklad pre  $\varepsilon \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ ) a zároveň existuje  $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$  tak, že  $\lim_{n \rightarrow \infty} D_n(\boldsymbol{\theta}, \boldsymbol{\theta}^*) < \infty$ , potom  $\hat{\boldsymbol{\theta}}_n$  nie je konzistentným odhadom. Nasledovný príklad, kedy odhad metódou nelineárnych najmenších štvorcov nie je konzistentným odhadom uvádza Malinvaud [15].

**Príklad 1.1.** Uvažujme model v tvare

$$Y_i = e^{-\alpha^* i} + \varepsilon_i, \quad i = 1, \dots, n,$$

kde  $\alpha^* > 0$  je skutočná hodnota neznámeho parametra a  $\varepsilon_i$  sú nezávislé a rovnako rozdelené veličiny,  $\varepsilon_i \in (-\infty, \infty)$ .

Vzhľadom k tomu, že platí

$$\begin{aligned} \lim_{n \rightarrow \infty} D_n(\alpha, \alpha^*) &= \sum_{i=1}^{+\infty} (e^{-\alpha i} - e^{-\alpha^* i})^2 \\ &= \frac{e^{-2\alpha}}{1 - e^{-2\alpha}} - \frac{2e^{-(\alpha+\alpha^*)}}{1 - e^{-(\alpha+\alpha^*)}} + \frac{e^{-2\alpha^*}}{1 - e^{-2\alpha^*}} < \infty \end{aligned}$$

pre ľubovoľné  $\alpha > 0$ , tak odhad metódou nelineárnych najmenších štvorcov  $\hat{\alpha}_n$  nie je konzistentný. Alternatívny dôkaz nekonzistencie  $\hat{\alpha}_n$ , ktorý nie je založený na nutnej a postačujúcej podmienke (1.13), uvádza Malinvaud [15].

△

Vzhľadom k tomu, že predpoklad (J4) často nie je možné overiť priamo, pretože nemáme k dispozícii nekonečnú postupnosť  $\{\mathbf{x}_i\}$ , bolo navrhnutých niekoľko alternatívnych predpokladov konzistencie, z ktorých si uvedieme verziu predpokladov, ktoré pôvodne navrhol Malinvaud [15] a ktoré uvádzajú Seber a Wild [18].

**Veta 1.6 (Predpoklady konzistencie II).** *Nech pre nelineárny model (1.12) platia nasledovné predpoklady:*

- M1. náhodné zložky  $\varepsilon_i$  sú nezávislé a rovnako rozdelené s nulovou strednou hodnotou a neznámym konečným rozptylom  $\sigma^2 > 0$ ;*
- M2. skutočná hodnota vektoru neznámych parametrov  $\boldsymbol{\theta}^*$  patrí do kompaktného parametrického priestoru  $\Omega \subset \mathbb{R}^p$ ;*
- M3.  $f(\mathbf{x}_i; \boldsymbol{\theta})$  sú známe spojité funkcie vektoru parametrov  $\boldsymbol{\theta} \in \Omega$ ;*
- M4. ak  $f(\mathbf{x}; \boldsymbol{\theta}) = f(\mathbf{x}; \boldsymbol{\theta}^*)$ , potom  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ ;*
- M5. funkcia  $H_n(\mathbf{x}) = n^{-1} \sum_{i=1}^n \mathbb{I}_{\{\mathbf{x}_i < \mathbf{x}\}} \xrightarrow[n \rightarrow \infty]{} H(\mathbf{x})$ , kde  $H(\mathbf{x})$  má vlastnosti distribučnej funkcie.*

*Potom  $\hat{\boldsymbol{\theta}}_n$  a  $\hat{\sigma}_n^2 = \mathcal{S}(\hat{\boldsymbol{\theta}}_n)/n$  sú silne konzistentnými odhadmi neznámych parametrov  $\boldsymbol{\theta}^*$  a  $\sigma^2$ .*

*Dôkaz.* Viď Seber a Wild [18], ods. 12.2.2.

□

Dôkaz vety 1.6 je založený na Hellyovej-Brayovej vete, podľa ktorej platí

$$\begin{aligned} \frac{1}{n} D_n(\boldsymbol{\theta}, \boldsymbol{\theta}') &= \int_{\mathbb{R}^k} [f(\mathbf{x}; \boldsymbol{\theta}) - f(\mathbf{x}; \boldsymbol{\theta}')]^2 dH_n(\mathbf{x}) \\ &\xrightarrow[n \rightarrow \infty]{} \int_{\mathbb{R}^k} [f(\mathbf{x}; \boldsymbol{\theta}) - f(\mathbf{x}; \boldsymbol{\theta}')]^2 dH(\mathbf{x}) = D(\boldsymbol{\theta}, \boldsymbol{\theta}'). \end{aligned}$$

**Poznámka 1.7.** Analogická veta s modifikovanými predpokladmi platí i v prípade náhodných vektorov  $\mathbf{x}_i$ , kedy  $H_n(\mathbf{x})$  je empirická distribučná funkcia.

### 1.2.4 Predpoklady asymptotickej normality

Predpokladajme, že odhad  $\hat{\boldsymbol{\theta}}$  získaný metódou nelineárnych najmenších štvorcov je aspoň slabo konzistentný. Už sme si ukázali, že  $\hat{\boldsymbol{\theta}}$  môžeme aproximovať výrazom (1.7). Intuitívne teda môžeme očakávať, že ako približné rozdelenie odhadu  $\hat{\boldsymbol{\theta}}$  môžeme použiť

$$\hat{\boldsymbol{\theta}} \sim \mathbf{N}\left(\boldsymbol{\theta}^*, \sigma^2(\mathbf{F}_*'\mathbf{F}_*)^{-1}\right).$$

Ak navyše nahradíme maticu  $(\mathbf{F}_*'\mathbf{F}_*)^{-1}$  príslušným konzistentným odhadom  $(\mathbf{F}(\hat{\boldsymbol{\theta}})'\mathbf{F}(\hat{\boldsymbol{\theta}}))^{-1}$  dostaneme približné rozdelenie

$$\hat{\boldsymbol{\theta}} \sim \mathbf{N}\left(\boldsymbol{\theta}^*, \sigma^2(\mathbf{F}(\hat{\boldsymbol{\theta}})'\mathbf{F}(\hat{\boldsymbol{\theta}}))^{-1}\right). \quad (1.14)$$

V tejto časti práce si stručne popíšeme predpoklady, za ktorých má odhad  $\hat{\boldsymbol{\theta}}$  limitné rozdelenie (1.14). Bol to práve Jennrich [14], ktorý odvodil a dokázal základné predpoklady asymptotickej normality odhadu  $\hat{\boldsymbol{\theta}}$  a ktorého závery aplikované na model (1.12) si zhrnieme do nasledovnej vety.

**Veta 1.8** (Predpoklady asymptotickej normality I). *Nech pre nelineárny model (1.12) platia predpoklady (J1)–(J4) vety 1.2 a zároveň nech platia i tieto dodatočné predpoklady:*

- J5. vektor  $\boldsymbol{\theta}^*$  je vnútorným bodom parametrického priestoru  $\Omega \subset \mathbb{R}^p$ ;
- J6. prvé a druhé parciálne derivácie funkcií  $f(\mathbf{x}_i; \boldsymbol{\theta})$  podľa parametrov  $\theta_1, \dots, \theta_p$  existujú a sú spojité pre všetky vektory  $\boldsymbol{\theta} \in \Theta^*$ , kde  $\Theta^*$  predstavuje okolie bodu  $\boldsymbol{\theta}^*$ ;
- J7a.  $n^{-1} \sum_{i=1}^n \left( \frac{\partial}{\partial \boldsymbol{\theta}} f(\mathbf{x}_i; \boldsymbol{\theta}) \right) \left( \frac{\partial}{\partial \boldsymbol{\theta}'} f(\mathbf{x}_i; \boldsymbol{\theta}) \right) = n^{-1} \mathbf{F}_*(\boldsymbol{\theta})' \mathbf{F}_*(\boldsymbol{\theta})$  konverguje rovnomerne k nejakej matici  $\mathbf{C}(\boldsymbol{\theta})$  na množine  $\Theta^*$ ;
- J7b.  $n^{-1} \sum_{i=1}^n \left( \frac{\partial^2}{\partial \theta_r \partial \theta_s} f(\mathbf{x}_i; \boldsymbol{\theta}) \right)^2$  konvergujú pre všetky  $r, s = 1, \dots, p$  rovnomerne na množine  $\Theta^*$ ;
- J8. matica  $\mathbf{C} = \mathbf{C}(\boldsymbol{\theta}^*)$  je regulárna.

Potom má náhodná veličina  $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*)$  asymptoticky normálne rozdelenie  $\mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{C}^{-1})$ , t.j. platí

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z, \quad Z \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{C}^{-1}). \quad (1.15)$$

Navyše, matica  $\mathbf{C}(\hat{\boldsymbol{\theta}}_n)$  je silne konzistentným odhadom matice  $\mathbf{C}$ .

*Dôkaz.* Viď Seber a Wild [18], ods. 12.2.3. □

**Poznámka 1.9.** Za platnosti predpokladov (J1)–(J8) nie je asymptoticky normálne rozdelený odhad  $\hat{\boldsymbol{\theta}}_n$  obecné asymptoticky efektívny. Jennrich [14] však poznamenal, že v prípade normálneho rozdelenia náhodných zložiek  $\varepsilon_i$  modelu (1.12) odhad  $\hat{\boldsymbol{\theta}}_n$  asymptoticky efektívny je.

Pre väčšiu názornosť si uvedieme náčrt hlavných bodov dôkazu asymptotickej normality odhadu metódou nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}_n$  na základe predpokladov (J1)–(J8), ktorý v inej podobe uvádza i Amemiya [1].

Vzhľadom k predpokladu (J6) dostávame na základe vety o strednej hodnote a na základe požiadavky nulovosti vektoru parciálnych derivácií (1.3) vzťah

$$\frac{\partial \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} = \frac{\partial \mathcal{S}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} + \frac{\partial^2 \mathcal{S}(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*) = \mathbf{0},$$

kde  $\tilde{\boldsymbol{\theta}}_n$  leží medzi  $\hat{\boldsymbol{\theta}}_n$  a  $\boldsymbol{\theta}^*$ . Z toho vyplýva, že pre náhodnú veličinu  $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*)$  platí

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*) = - \left[ \frac{1}{n} \frac{\partial^2 \mathcal{S}(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right]^{-1} \frac{1}{\sqrt{n}} \frac{\partial \mathcal{S}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}}. \quad (1.16)$$

Ukážeme, že výraz  $n^{-1} \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathcal{S}(\tilde{\boldsymbol{\theta}}_n)$  konverguje v pravdepodobnosti k regulárnej matici a  $-n^{-1/2} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\boldsymbol{\theta}^*)$  má asymptoticky normálne rozdelenie.

V prípade prvého výrazu postupným derivovaním (1.4) dostaneme

$$\begin{aligned} \frac{1}{n} \frac{\partial^2 \mathcal{S}(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} &= \frac{1}{n} \frac{\partial}{\partial \boldsymbol{\theta}'} \left\{ -2 \frac{\partial \mathbf{f}'(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n)] \right\} \\ &= \frac{1}{n} \frac{\partial}{\partial \boldsymbol{\theta}'} \sum_{i=1}^n \left\{ -2 \frac{\partial f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} [Y_i - f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)] \right\} \\ &= \frac{2}{n} \sum_{i=1}^n \left\{ \frac{\partial f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}'} - \frac{\partial^2 f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} [Y_i - f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)] \right\}. \end{aligned}$$

Vzhľadom k tomu, že  $\hat{\boldsymbol{\theta}}_n$  je konzistentný odhad  $\boldsymbol{\theta}^*$ , t.j.  $\hat{\boldsymbol{\theta}}_n \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \boldsymbol{\theta}^*$ , zjavne platí aj  $\tilde{\boldsymbol{\theta}}_n \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \boldsymbol{\theta}^*$ , a navyše z predpokladu spojitosti funkcií  $f(\mathbf{x}_i; \boldsymbol{\theta})$  vyplýva  $f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} f(\mathbf{x}_i; \boldsymbol{\theta}^*)$ . Z tohto dôvodu za predpokladu (J7a) platí

$$\frac{2}{n} \sum_{i=1}^n \frac{\partial f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}'} = 2 \frac{1}{n} \mathbf{F}'(\tilde{\boldsymbol{\theta}}_n) \mathbf{F}(\tilde{\boldsymbol{\theta}}_n) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 2\mathbf{C}(\boldsymbol{\theta}^*) = 2\mathbf{C},$$

kde  $\mathbf{C}$  je v súlade s predpokladom (J8) regulárna matica typu  $p \times p$ . Obdobne za podmienky (J7b) platí

$$\begin{aligned} & - \frac{2}{n} \sum_{i=1}^n \frac{\partial^2 f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} [Y_i - f(\mathbf{x}_i; \boldsymbol{\theta}^*) + f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)] \\ &= - \frac{2}{n} \sum_{i=1}^n \frac{\partial^2 f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \varepsilon_i - \frac{2}{n} \sum_{i=1}^n \frac{\partial^2 f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} [f(\mathbf{x}_i; \boldsymbol{\theta}^*) - f(\mathbf{x}_i; \tilde{\boldsymbol{\theta}}_n)] \xrightarrow[n \rightarrow \infty]{\mathcal{P}} \mathbf{0}. \end{aligned}$$

Pre prvý výraz teda dostávame konvergenciu v pravdepodobnosti

$$\frac{1}{n} \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathcal{S}(\tilde{\boldsymbol{\theta}}_n) \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 2\mathbf{C}. \quad (1.17)$$

Pre druhý výraz platí vzťah

$$-\frac{1}{\sqrt{n}} \frac{\partial \mathcal{S}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} = 2 \frac{1}{\sqrt{n}} \mathbf{F} \cdot (\boldsymbol{\theta}^*)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}^*)] = \frac{2}{\sqrt{n}} \mathbf{F} \cdot \boldsymbol{\varepsilon} = \sum_{i=1}^n \frac{2}{\sqrt{n}} \frac{\partial f(\mathbf{x}_i; \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \varepsilon_i$$

a teda sa jedná o  $p$  lineárnych kombinácií nezávislých a rovnako rozdelených náhodných veličín  $\varepsilon_1, \dots, \varepsilon_n$ . S využitím predpokladu (J7a) máme

$$\sum_{i=1}^n \text{var} \left( \frac{2}{\sqrt{n}} \frac{\partial f(\mathbf{x}_i; \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \varepsilon_i \right) = \frac{4\sigma^2}{n} \sum_{i=1}^n \frac{\partial f(\mathbf{x}_i; \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{x}_i; \boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}'} \xrightarrow{n \rightarrow \infty} 4\sigma^2 \mathbf{C}.$$

Následne je možné s využitím viacrozmernej Fellerovej-Lindebergovej centrálnej limitnej vety ukázať, že  $-n^{-1/2} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\boldsymbol{\theta}^*)$  má asymptoticky normálne rozdelenie, konkrétne že platí

$$-\frac{1}{\sqrt{n}} \frac{\partial \mathcal{S}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z, \quad Z \sim \mathbf{N}(\mathbf{0}, 4\sigma^2 \mathbf{C}). \quad (1.18)$$

Na základe Cramérovej-Slutského vety teda z (1.16), (1.17) a (1.18) dostaneme

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z, \quad Z \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{C}^{-1}).$$

Vzhľadom k tomu, že  $n^{-1} \mathbf{F}'(\hat{\boldsymbol{\theta}}_n) \mathbf{F}(\hat{\boldsymbol{\theta}}_n)$  je pri splnení predpokladov (J1)-(J8) konzistentným odhadom matice  $\mathbf{C}$ , sme overili i asymptotickú platnosť vzťahu (1.14).

Iný systém predpokladov asymptotickej normality, ktorý nadväzuje na vetu 1.6, uvádza Malinvaud [16].

**Veta 1.10** (Predpoklady asymptotickej normality II). *Nech pre nelineárny model (1.12) platia predpoklady (M1)-(M5) vety 1.6 a zároveň nech platia tieto dodatočné predpoklady:*

M6. vektor  $\boldsymbol{\theta}^*$  je vnútorným bodom parametrického priestoru  $\Omega \subset \mathbb{R}^p$ ;

M7. funkcie  $f(\mathbf{x}_i; \boldsymbol{\theta})$  sú dvakrát spojitely diferencovateľné;

M8. matica typu  $p \times p$

$$\mathbf{C} = \left( \int_{\mathbb{R}^k} \frac{\partial f(\mathbf{x}; \boldsymbol{\theta}^*)}{\partial \theta_i} \frac{\partial f(\mathbf{x}; \boldsymbol{\theta}^*)}{\partial \theta_j} dH(\mathbf{x}) \right)$$

je regulárna.

Potom má náhodná veličina  $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*)$  asymptoticky normálne rozdelenie  $\mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{C}^{-1})$ , t.j. platí

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z, \quad Z \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{C}^{-1}).$$

Navyš, matica  $\mathbf{C}(\hat{\boldsymbol{\theta}}_n)$  je silne konzistentným odhadom matice  $\mathbf{C}$ .

*Dôkaz.* Viď Malinvaud [16]. □

Na záver tejto podkapitoly si zhrnieme asymptotické vlastnosti odhadu  $\hat{\boldsymbol{\theta}}$ , ktoré uvádzajú Seber a Wild [18]. Uvedené vlastnosti si záujemca jednoducho overí na základe lineárnych aproximácií (1.7) a (1.9).

**Veta 1.11** (Vlastnosti odhadu metódou nelineárnych najmenších štvorcov). *Nech pre nelineárny model (1.12) platí  $\varepsilon_i \stackrel{iid}{\sim} \mathbf{N}(0, \sigma^2)$  a nech platí systém predpokladov asymptotickej normality. Potom pre veľké  $n$  približne platia nasledovné vlastnosti:*

$$i) \hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^* \sim \mathbf{N}\left(\mathbf{0}, \sigma^2 (\mathbf{F}_*'^* \mathbf{F}_*^*)^{-1}\right);$$

$$ii) \mathcal{S}(\hat{\boldsymbol{\theta}}_n) / \sigma^2 = (n - p) S^2 / \sigma^2 \sim \chi_{n-p}^2;$$

iii) náhodné vektory  $\hat{\boldsymbol{\theta}}_n$  a  $S^2$  sú nezávislé;

iv) pre reziduálne súčty štvorcov platí

$$\frac{\mathcal{S}(\boldsymbol{\theta}^*) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\mathcal{S}(\hat{\boldsymbol{\theta}}_n)} \frac{n - p}{p} \sim F_{p, n-p}. \quad (1.19)$$

*Dôkaz.* Viď Seber a Wild [18], ods. 2.1.2, veta 2.1. □

**Poznámka 1.12.** Gallant [9] ukázal, že pre lineárne aproximácie (1.7) a (1.9), pomocou ktorých sú uvedené vlastnosti odvodené, platia vzťahy:

$$\hat{\boldsymbol{\theta}}_n = \boldsymbol{\theta}^* + (\mathbf{F}_*'^* \mathbf{F}_*^*)^{-1} \mathbf{F}_*'^* \boldsymbol{\varepsilon} + o_p\left(\frac{1}{\sqrt{n}}\right)$$

a

$$S^2 = \frac{\mathcal{S}(\hat{\boldsymbol{\theta}})}{n - p} = \frac{\boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{H}_F^*) \boldsymbol{\varepsilon}}{n - p} + o_p\left(\frac{1}{n}\right),$$

kde symbolom  $o_p(a_n)$ , kde  $\{a_n\}$  je postupnosť reálnych čísel, značíme náhodný vektor (prípadne náhodnú veličinu)  $X_n$ , pre ktorého zložky  $X_{in}$  platí

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\left|\frac{X_{in}}{a_n}\right| > \epsilon\right) = 0$$

pre každé  $\epsilon > 0$ , t.j. náhodné veličiny  $X_{in}/a_n$  konvergujú v pravdepodobnosti k nule.

### 1.2.5 Boxov odhad vychýlenia

Je veľmi dôležité si uvedomiť, že vlastnosti uvedené vo vete 1.11 približne platia až pre dostatočne veľký rozsah výberu  $n$ . V tomto ohľade, ako uvádza napríklad Ratkowsky [17], obecné neexistuje pravidlo, ktoré by určovalo, aký rozsah výberu postačuje k približnej platnosti asymptotických vlastností, resp. potrebný rozsah výberu môže byť pre rôzne modely a ich parametrizácie odlišný, pričom rozhodujúcim faktorom je stupeň nelinearity modelu.

Pre konečné výbery teda obecné nemá odhad metódou nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}$  normálne rozdelenie, nemusí mať najmenší rozptyl a dokonca nie je ani nestranný. V tejto časti si aspoň v krátkosti uvedme, ako je možné určiť veľkosť vychýlenia na základe aproximácie, ktorú navrhol Box [6].

Najprv si však rozšírime značenie zavedené v podkapitole 1.2.1. Pre každú funkciu  $f(\mathbf{x}_i; \boldsymbol{\theta})$ ,  $i = 1, \dots, n$  si zavedme si maticu príslušných druhých parciálnych derivácií typu  $p \times p$  ako

$$\mathbf{F}_{i..}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} f(\mathbf{x}_i; \boldsymbol{\theta})$$

a symbolom  $\mathbf{F}_{..}(\boldsymbol{\theta})$  budeme značiť trojrozmerné pole typu  $n \times p \times p$ , ktoré bude po vrstvách obsahovať matice  $\mathbf{F}_{i..}(\boldsymbol{\theta})$ ,  $i = 1, \dots, n$ , t.j.

$$\mathbf{F}_{..}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathbf{f}(\boldsymbol{\theta}) = \left( \mathbf{F}_{i..}(\boldsymbol{\theta}) \right)_{i=1, \dots, n}.$$

Takisto budeme používať skrátene značenie  $\mathbf{F}_{i..}(\boldsymbol{\theta}^*) = \mathbf{F}_{i..}^*$  a  $\mathbf{F}_{..}(\boldsymbol{\theta}^*) = \mathbf{F}_{..}^*$ .

S využitím tohto značenia môžeme vyjadriť kvadratickú aproximáciu regresnej funkcie Taylorovým polynómom druhého radu ako

$$\mathbf{f}(\boldsymbol{\theta}) \approx \mathbf{f}^* + \mathbf{F}_{..}^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*) + \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \mathbf{F}_{..}^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*),$$

pričom v prípade kvadratického členu budeme uvedeným značením chápať súčin cez druhý a tretí rozmer trojrozmerného poľa  $\mathbf{F}_{..}^*$ .

Box [6] odvodil na základe porovnania lineárnej a kvadratickej aproximácie regresnej funkcie odhad vychýlenia v tvare

$$\text{bias } \hat{\boldsymbol{\theta}} \doteq -\frac{\sigma^2}{2} (\mathbf{F}_{..}^{*'} \mathbf{F}_{..}^*)^{-1} \mathbf{F}_{..}^{*'} \mathbf{m}, \quad (1.20)$$

kde

$$\mathbf{m} = \left( \text{tr } \mathbf{F}_{1..}^* (\mathbf{F}_{..}^{*'} \mathbf{F}_{..}^*)^{-1}, \dots, \text{tr } \mathbf{F}_{n..}^* (\mathbf{F}_{..}^{*'} \mathbf{F}_{..}^*)^{-1} \right)',$$

pričom prostredníctvom  $\text{tr}$  značíme stopu matice. Odvodenie príslušného tvaru aproximácie vychýlenia odhadu  $\hat{\boldsymbol{\theta}}$  uvádza napríklad Zvára [21], obecnjšie odvodenie nájde záujemca v pôvodnom článku Box [6].

Ďalším špecifikom nelineárneho modelu je potreba jeho identifikovateľnosti. Tá vychádza zo skutočnosti, že rovnaký nelineárny regresný model s identickým priebehom regresnej funkcie možno vyjadriť v rôznych parametrizáciách.

Napríklad rastový model vyjadrený logistickou krivkou môže mať hneď niekoľko parametrizácií, z ktorých si uvedieme aspoň najčastejšie:

$$f(x; \alpha_1, \alpha_2, \alpha_3) = \frac{\alpha_1}{1 + e^{\alpha_2 - \alpha_3 x}}, \quad \alpha_1 > 0, \alpha_3 > 0, \quad (1.21)$$

$$f(x; \beta_1, \beta_2, \beta_3) = \frac{\beta_1}{1 + \beta_2 e^{-\beta_3 x}}, \quad \beta_1 > 0, \beta_2 > 0, \beta_3 > 0, \quad (1.22)$$

$$f(x; \gamma_1, \gamma_2, \gamma_3) = \frac{\gamma_1}{1 + \gamma_2 \gamma_3^x}, \quad \gamma_1 > 0, \gamma_2 > 0, 0 < \gamma_3 < 1. \quad (1.23)$$

Navyše je zjavné, že prostou transformáciou parametrov je možné previesť jednu parametrizáciu na druhú a naopak. V uvedených prípadoch zjavne platí

$$\begin{aligned} \beta_1 &= \alpha_1, & \beta_2 &= e^{\alpha_2}, & \beta_3 &= \alpha_3, \\ \gamma_1 &= \alpha_1, & \gamma_2 &= e^{\alpha_2}, & \gamma_3 &= e^{-\alpha_3}. \end{aligned}$$

Uvažujme teda regulárne a prosté zobrazenie  $\boldsymbol{\beta} = \mathbf{g}(\boldsymbol{\theta})$  a regulárnu maticu parciálnych derivácií  $\mathbf{G}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}'} \mathbf{g}(\boldsymbol{\theta})$  typu  $p \times p$ . Potom vychýlenie odhadu  $\hat{\boldsymbol{\beta}}$  v takto reparametrizovanom modeli možno aproximovať ako

$$\text{bias } \hat{\boldsymbol{\beta}} \doteq \mathbf{G}(\boldsymbol{\theta}^*) \text{bias } \hat{\boldsymbol{\theta}} + \frac{1}{2} \mathbf{m}_g, \quad (1.24)$$

kde

$$\mathbf{m}_g = \left( \text{tr} \frac{\partial^2 g_1(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \text{var } \hat{\boldsymbol{\theta}}, \dots, \text{tr} \frac{\partial^2 g_p(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \text{var } \hat{\boldsymbol{\theta}} \right)'.$$

Odhad vychýlenia dostaneme nahradením neznámeho parametra  $\boldsymbol{\theta}^*$  príslušným odhadom  $\hat{\boldsymbol{\theta}}$ . Odhad vychýlenia parametrov v percentuálnom vyjadrení vzhľadom k hodnote  $\hat{\boldsymbol{\theta}}$  je jeden z indikátorov stupňa krivosti modelu. To znamená, že čím väčšie je percentuálne vychýlenie odhadu, tým je stupeň nelinearity modelu väčší a tým väčší rozsah výberu potrebujeme k tomu, aby približne platili asymptotické vlastnosti vety 1.11. Ratkowsky [17] uvádza hrubé pravidlo, na základe ktorého je možné považovať percentuálny odhad vychýlenia, ktorý je v absolútnej hodnote väčší než 1 %, za indikátor významnej nelinearity modelu.

### 1.3 Miery krivosti

V predošlej kapitole sme uviedli, že platnosť asymptotických vlastností odhadu metódou nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}$  v prípade konečného rozsahu výberu závisí na stupni nelinearity modelu. Ten je možné kvantifikovať pomocou mier krivosti, ktoré navrhli Bates a Watts [3]. Vzhľadom k tomu, že konštrukcia a detailná geometrická interpretácia týchto mier krivosti zďaleka presahuje možnosti tejto práce, uvedieme si iba niektoré hlavné závery.

Je zrejmé, že platnosť asymptotických vlastností pre konečné výbery závisí na tom, ako dobrá je lineárna aproximácia

$$\mathbf{f}(\hat{\boldsymbol{\theta}}) \approx \mathbf{f}^* + \mathbf{F}^*(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)$$

v blízkosti bodu  $\theta^*$ . Vzhľadom k tomu, že skutočnú hodnotu vektoru parametrov  $\theta^*$  nepoznáme, nahradíme maticu  $\mathbf{F}_*$  príslušným odhadom  $\mathbf{F}_*(\hat{\theta})$ . V praxi teda využívame lineárnu aproximáciu v tvare

$$\mathbf{f}^* \approx \mathbf{f}(\hat{\theta}) + \mathbf{F}_*(\hat{\theta})(\theta^* - \hat{\theta}).$$

Táto aproximácia skutočnej hodnoty regresnej funkcie Taylorovým polynómom v bode  $\hat{\theta}$  je vhodná v prípade, že oblasť určená množinou všetkých možných stredných hodnôt  $\{\mathbf{f}(\theta) : \theta \in \Omega\}$  je v blízkosti bodu  $\hat{\theta}$  značne plochá, t.j. príslušná nadrovina ju dostatočne dobre aproximuje.

Ponímanie krivosti modelu je založené na porovnaní lineárnej a kvadratickej aproximácie. Konštrukciu mier krivosti si popíšeme aspoň intuitívne spôsobom, ktorý uvádzajú Seber a Wild [18]. Ak vyjadríme kvadratickú aproximáciu Taylorovým polynómom v bode  $\hat{\theta}$

$$\begin{aligned} \mathbf{f}^* &\approx \mathbf{f}(\hat{\theta}) + \mathbf{F}_*(\hat{\theta})(\theta^* - \hat{\theta}) + \frac{1}{2}(\theta^* - \hat{\theta})' \mathbf{F}_{**}(\hat{\theta})(\theta^* - \hat{\theta}) \\ &= \mathbf{f}(\hat{\theta}) + \mathbf{F}_*(\hat{\theta})\delta + \frac{1}{2}\delta' \mathbf{F}_{**}(\hat{\theta})\delta, \end{aligned}$$

kde

$$\delta = \theta^* - \hat{\theta},$$

je zrejmé, že lineárna aproximácia je vhodná, ak veľkosť kvadratického členu  $\delta' \mathbf{F}_{**}(\hat{\theta})\delta$  je v porovnaní s veľkosťou  $n$ -rozmerného vektoru  $\mathbf{F}_*(\hat{\theta})\delta$  malá.

Bates a Watts [3] navrhli dva typy krivosti, ktoré sú založené na rozklade kvadratického členu na dva ortogonálne vektory, z ktorých jeden je kolmý na príslušnú dotykovú nadrovinu (budeme značiť indexom  $N$ ) a druhý je s touto nadrovinou rovnobežný (budeme značiť indexom  $T$ ). Takúto dekompozíciu kvadratického členu dosiahneme použitím projekčnej matice v tvare  $\hat{\mathbf{H}}_F = \mathbf{F}_*(\hat{\theta})[\mathbf{F}_*(\hat{\theta})' \mathbf{F}_*(\hat{\theta})]^{-1} \mathbf{F}_*(\hat{\theta})'$ . Na základe tejto dekompozície boli definované krivosti v smere  $\delta$  vzťahmi

$$\begin{aligned} K_\delta^N &= \frac{\|(\mathbf{I}_n - \hat{\mathbf{H}}_F)[\delta' \mathbf{F}_{**}(\hat{\theta})\delta]\|}{\|\mathbf{F}_*(\hat{\theta})\delta\|^2} = \frac{\|\delta' \mathbf{F}_{**}^N(\hat{\theta})\delta\|}{\|\mathbf{F}_*(\hat{\theta})\delta\|^2}, \\ K_\delta^T &= \frac{\|\hat{\mathbf{H}}_F[\delta' \mathbf{F}_{**}(\hat{\theta})\delta]\|}{\|\mathbf{F}_*(\hat{\theta})\delta\|^2} = \frac{\|\delta' \mathbf{F}_{**}^T(\hat{\theta})\delta\|}{\|\mathbf{F}_*(\hat{\theta})\delta\|^2} \end{aligned}$$

Vzhľadom k tomu, že takto navrhnuté krivosti nie sú invariantné pri zmene merítka, používajú sa relatívne krivosti v tvare

$$\begin{aligned} \gamma_\delta^N &= \sigma \sqrt{p} K_\delta^N, \\ \gamma_\delta^T &= \sigma \sqrt{p} K_\delta^T, \end{aligned}$$

kde  $\gamma_\delta^N$  sa nazýva *vnútorná krivosť* v smere  $\delta$  a  $\gamma_\delta^T$  sa nazýva *parametrická krivosť* v smere  $\delta$ . Tieto názvy vyplývajú zo skutočnosti, že vnútorná krivosť v danom smere nezávisí na použitej parametrizácii modelu, tzn. nie je možné jej vplyv eliminovať, na druhej strane je však možné nadmernú parametrickú krivosť vhodnou reparametrizáciou znížiť.

Krivosť modelu však možno merať obecné v ľubovoľnom smere. Ak  $\mathbf{h}$  je nenulový  $p$ -rozmerný vektor, tak vnútornú a parametrickú krivosť v smere  $\mathbf{h}$  možno vyjadriť ako

$$\gamma_{\mathbf{h}}^N = \sigma\sqrt{p}K_{\mathbf{h}}^N = \sigma\sqrt{p} \frac{\|\mathbf{h}'\mathbf{F}_{\cdot\cdot}^N(\hat{\boldsymbol{\theta}})\mathbf{h}\|}{\|\mathbf{F}_{\cdot}(\hat{\boldsymbol{\theta}})\mathbf{h}\|^2},$$

$$\gamma_{\mathbf{h}}^T = \sigma\sqrt{p}K_{\mathbf{h}}^T = \sigma\sqrt{p} \frac{\|\mathbf{h}'\mathbf{F}_{\cdot\cdot}^T(\hat{\boldsymbol{\theta}})\mathbf{h}\|}{\|\mathbf{F}_{\cdot}(\hat{\boldsymbol{\theta}})\mathbf{h}\|^2}.$$

Je zjavné, že takto definované krivosti sú závislé na voľbe vektoru  $\mathbf{h}$ . Z tohto dôvodu sa v súlade s článkom Bates a Watts [3] ako miery krivosti spravidla používajú maximálne krivosti v tvare

$$\gamma^N = \max \gamma_{\mathbf{h}}^N,$$

$$\gamma^T = \max \gamma_{\mathbf{h}}^T,$$

prípadne priemerné krivosti v tvare

$$\gamma_{RMS}^N = \left[ \frac{1}{S_p} \int_{\|\mathbf{h}\|=1} (\gamma_{\mathbf{h}}^N)^2 dS \right]^{1/2},$$

$$\gamma_{RMS}^T = \left[ \frac{1}{S_p} \int_{\|\mathbf{h}\|=1} (\gamma_{\mathbf{h}}^T)^2 dS \right]^{1/2},$$

kde sa integruje cez povrch jednotkovej gule a  $S_p = \int_{\|\mathbf{h}\|=1} dS$  značí povrch  $p$ -rozmernej jednotkovej gule.

Bates a Watts [4] navrhujú porovnávať maximálne krivosti  $\gamma^N$ , resp.  $\gamma^T$  s hodnotou  $(\mathbf{F}_{p,n-p}^{1-\alpha})^{-1/2}$ , kde  $\mathbf{F}_{p,n-p}^{1-\alpha} = \mathbf{F}_{p,n-p}(1-\alpha)$  je príslušný kvantil  $\mathbf{F}$  rozdelenia s  $p$  a  $n-p$  stupňami voľnosti. Konkrétne, ak pre maximálnu vnútornú a parametrickú krivosť platí, že

$$\gamma^N < \frac{1}{2\sqrt{\mathbf{F}_{p,n-p}^{1-\alpha}}} \quad \text{a} \quad \gamma^T < \frac{1}{2\sqrt{\mathbf{F}_{p,n-p}^{1-\alpha}}},$$

potom je možné podľa autorov považovať krivosť modelu za nevýznamnú, a teda lineárnu aproximáciu za prijateľnú. V prípade priemerných krivostí navrhujú autori obdobné pravidlo, na základe ktorého možno krivosť modelu považovať za nevýznamnú, ak platí

$$\gamma_{RMS}^N \sqrt{\mathbf{F}_{p,n-p}^{1-\alpha}} < 0.3 \quad \text{a} \quad \gamma_{RMS}^T \sqrt{\mathbf{F}_{p,n-p}^{1-\alpha}} < 0.3.$$

Iné miery krivosti modelu navrhuje i Beale [5]. Ako však ukázali Guttman a Meeter [10], tieto miery môžu stupeň nelinearity modelu podhodnocovať. Detailný popis jednotlivých mier krivosti nelineárnych modelov uvádzajú Bates a Watts [4], prípadne Seber a Wild [18].

## 1.4 Nelineárna regresia v R

Na záver tejto kapitoly si budeme ilustrovať uvedené teoretické poznatky na odhade sigmoidálneho rastového modelu v programovom prostredí R.

**Príklad 1.2.** Za účelom ilustrácie bodového odhadu parametrov nelineárneho modelu použijeme dáta popisujúce obsah vody v bunkách v koreni rastliny v závislosti na vzdialenosti od špičky koreňa (dáta uvádza Ratkowsky [17], str. 88, data set 4). Vzhľadom k charakteru dát využijeme na popis danej závislosti logistický rastový model (1.21). Za účelom výpočtu prvých a druhých parciálnych derivácií využijeme funkciu `deriv3`.

```
# načítanie dát
> Ratkowsky4 <-
+ data.frame(x = seq(0.5, 14.5, by = 1.0),
+           y = c( 1.3,  1.3,  1.9,  3.4,  5.3,  7.1, 10.6, 16.0,
+               16.4, 18.3, 20.9, 20.5, 21.3, 21.2, 20.9))
# parametrizácia logistického rastového modelu
> nls.function_A <-
+ deriv3(~ a1 / (1 + exp(a2 - a3 * x)), namevec = c("a1","a2","a3"),
+       function.arg = function(x, a1, a2, a3){})
```

Pred tým, než pristúpime k samotnému odhadu tohto modelu, budeme potrebovať vhodnú počiatočnú aproximáciu neznámych parametrov pre iteratívny algoritmus. Z priebehu logistickej funkcie (1.21) je zrejmé, že parameter  $\alpha_1$  vyjadruje asymptotu tejto funkcie, ktorú pre účely počiatočnej aproximácie môžeme odhadnúť ako maximálnu hodnotu, ktorú vysvetľovaná premenná dosahuje pri veľkých hodnotách vysvetľujúcej premennej a ktorú označíme  $a_1^*$ . Počiatočné hodnoty neznámych parametrov  $\alpha_2$  a  $\alpha_3$  je možné určiť napríklad na základe odhadu linearizovaného modelu v tvare

$$\log\left(\frac{a_1^*}{f(x; a_1^*, \alpha_2, \alpha_3)} - 1\right) = \alpha_2 - \alpha_3 x.$$

Získané hodnoty použijeme ako počiatočné riešenie pre Gaussov-Newtonov algoritmus v rámci štandardnej R-kovej funkcie `nls`.

```
# model A
> a1_init = ceiling(max(Ratkowsky4$y))
> a2_init = lm(log((a1_init/y)-1) ~ x, data = Ratkowsky4)$coef[[1]]
> a3_init = -lm(log((a1_init/y)-1) ~ x, data = Ratkowsky4)$coef[[2]]
> model_A <-
+ nls(y ~ nls.function_A(x, a1, a2, a3), data = Ratkowsky4,
+     start = c(a1 = a1_init, a2 = a2_init, a3 = a3_init))
> summary(model_A, correlation = TRUE)
Formula: y ~ nls.function_A(x, a1, a2, a3)
Parameters:
      Estimate Std. Error t value Pr(>|t|)
a1  21.5089     0.4154   51.78 1.77e-15 ***
a2   3.9573     0.2619   15.11 3.57e-09 ***
a3   0.6222     0.0446   13.95 8.89e-09 ***
Residual standard error: 0.7194 on 12 degrees of freedom
```

Correlation of Parameter Estimates:

```
      a1      a2
a2 -0.43
a3 -0.60  0.95
```

Hodnoty odhadu asymptotickej variančnej matice odhadu parametrov vyplývajúcej z asymptotickej normality odhadu (1.15) môžeme určiť na základe matice prvých parciálnych derivácií v bode odhadu parametrov.

```
> F_hat = model_A$m$gradient()
> S = sqrt(sum(residuals(model_A)^2) / summary(model_A)$df[2])
> S^2 * solve(t(F_hat) %*% F_hat)
      a1      a2      a3
a1  0.17252487 -0.04679431 -0.011097446
a2 -0.04679431  0.06857917  0.011131115
a3 -0.01109745  0.01113111  0.001989346
```

Záujemca sa môže jednoducho presvedčiť, že odhadnuté štandardné chyby či korelácie odhadnutých parametrov z vyššie uvedeného výstupu funkcie `nls` odpovedajú danému odhadu asymptotickej variančnej matice.

Za účelom určenia Boxovho odhadu vychýlenia daného odhadu parametrov, ako aj maximálnych či priemerných vnútorných a parametrických krivostí modelu, využijeme funkciu `Mrms.curv`, ktorú uvádza Zvára [21].

```
> source("MODIFrms.R")
> Mrms.curv(model_A)
Parameter effects:          0.1782 (max 0.3717)
      Intrinsic:          0.0608 (max 0.1071)
Parameter effects (x sqrt(F)): 0.3329 (max 0.6944)
      Intrinsic (x sqrt(F)): 0.1136 (max 0.2)
      Estimate      Bias  Rel. Bias
a1      21.509  0.015359  0.071 %
a2       3.957  0.017338  0.438 %
a3       0.622  0.002762  0.444 %
```

Vidíme, že relatívne vychýlenie v absolútnej hodnote neprekračuje 1%, a tak podľa tohto hrubého kritéria nie je nelinearita modelu príliš veľká. Podľa oboch kritérií, ktoré pre jednotlivé miery krivosti navrhujú Bates a Watts [4], možno síce považovať vnútornú krivosť modelu za zanedbateľnú, hodnoty mier parametrickej krivosti už príslušné hranice mierne prekračujú, a teda pri interpretácii záverov vyplývajúcich z lineárnej aproximácie je potreba byť opatrnejší.

Uvedme si ešte odhady ďalších dvoch parametrizácií logistického rastového modelu, tzn. (1.22) a (1.23).

```
# model B
> nls.function_B <-
+ deriv3(~ b1 / (1 + b2 * exp(-b3*x)), namevec = c("b1", "b2", "b3"),
+       function.arg = function(x, b1, b2, b3){})
> b1_init = a1_init
> b2_init = exp(a2_init)
```

```

> b3_init = a3_init
> model_B <-
+ nls(y ~ nls.function_B(x, b1, b2, b3), data = Ratkowsky4,
+     start = c(b1 = b1_init, b2 = b2_init, b3 = b3_init))
> summary(model_B)
Formula: y ~ nls.function_B(x, b1, b2, b3)
Parameters:
      Estimate Std. Error t value Pr(>|t|)
b1  21.5089     0.4154  51.784 1.77e-15 ***
b2  52.3168    13.7005   3.819 0.002446 **
b3   0.6222     0.0446  13.950 8.89e-09 ***

Residual standard error: 0.7194 on 12 degrees of freedom

# model C
> nls.function_C <-
+ deriv3(~ c1 / (1 + c2 * (c3 ^ x)), namevec = c("c1","c2","c3"),
+       function.arg = function(x, c1, c2, c3){})
> c1_init <- a1_init
> c2_init <- exp(a2_init)
> c3_init <- exp(-a3_init)
> model_C <-
+ nls(y ~ nls.function_C(x, c1, c2, c3), data = Ratkowsky4,
+     start = c(c1 = c1_init, c2 = c2_init, c3 = c3_init))
> summary(model_C)
Formula: y ~ nls.function_C(x, c1, c2, c3)
Parameters:
      Estimate Std. Error t value Pr(>|t|)
c1  21.5089     0.41536  51.784 1.77e-15 ***
c2  52.3168    13.70053   3.819 0.002446 **
c3   0.5368     0.02394  22.420 3.66e-11 ***

Residual standard error: 0.7194 on 12 degrees of freedom

```

Môžeme sa presvedčiť, že pre bodové odhady rôznych parametrizácií logistického rastového modelu (až na zanedbateľné odchýlky spôsobené iteračným algoritmom) platí

$$\begin{aligned}
 b_1 &= a_1, & b_2 &= e^{a_2}, & b_3 &= a_3, \\
 c_1 &= a_1, & c_2 &= e^{a_2}, & c_3 &= e^{-a_3}.
 \end{aligned}$$

Na záver si overíme, že reparametrizácia nelineárneho modelu nemá vplyv na jeho vnútornú krivosť. Zároveň si môžeme všimnúť, že parametrická krivosť týchto parametrizácií logistického rastového modelu je značná, rovnako ako aj odhad vychýlenia pre neznámy parameter  $\beta_2$ , resp.  $\gamma_2$ . Nelinearitu týchto modelov teda môžeme považovať za významnú, a teda pre testovanie významnosti príslušných parametrov by mali byť preferované testy, ktoré na parametrickej krivosti nezávisia.

```

> Mrms.curv(model_B)
Parameter effects:          0.8156 (max 1.8372)
      Intrinsic:          0.0608 (max 0.1071)
Parameter effects (x sqrt(F)): 1.5237 (max 3.4323)
      Intrinsic (x sqrt(F)): 0.1136 (max 0.2)
      Estimate      Bias      Rel. Bias
b1      21.509  0.015359  0.071 %
b2      52.317  2.700987  5.163 %
b3       0.622  0.002762  0.444 %

> Mrms.curv(model_C)
Parameter effects:          0.9477 (max 2.1309)
      Intrinsic:          0.0608 (max 0.1071)
Parameter effects (x sqrt(F)): 1.7706 (max 3.981)
      Intrinsic (x sqrt(F)): 0.1136 (max 0.2)
      Estimate      Bias      Rel. Bias
c1      21.509  0.015359  0.071 %
c2      52.317  2.700984  5.163 %
c3       0.537 -0.000949 -0.177 %

```

△

## 2. Konfidenčné množiny v nelineárnom modeli

Okrem bodového odhadu vektoru neznámych parametrov  $\theta^*$  v nelineárnom modeli (1.1) nás často zaujímajú i príslušné intervalové odhady vo forme intervalov spoľahlivosti pre jednotlivé parametre, prípadne konfidenčných množín pre zvolenú podmnožinu  $\theta^*$ .

Vzhľadom k tomu, že pre konečné výbery obecné nepoznáme presné rozdelenie odhadu  $\hat{\theta}$ , nie je na rozdiel od normálneho lineárneho modelu možné priamočiaro skonštruovať žiaduce presné intervaly spoľahlivosti, resp. konfidenčné oblasti. Z tohto dôvodu bolo navrhnutých niekoľko možností, ako intervalový odhad parametrov so zvolenou spoľahlivosťou aproximovať. Táto kapitola sa teda bude venovať základným možnostiam konštrukcie konfidenčných množín v nelineárnom modeli, ktoré sú zväčša odvodené z asymptotickej teórie metódy maximálnej vierohodnosti.

V tejto kapitole budeme opäť uvažovať normálny nelineárny model (1.1) definovaný v podkapitole 1.2. Rovnako predpokladáme, že tento model spĺňa príslušné predpoklady regularity.

### 2.1 Waldovská konfidenčná množina

Azda najčastejšie využívaná forma konfidenčných množín je založená na lineárnej aproximácii (1.6), a teda na asymptotickej normalite (1.14) odhadu získaného metódou nelineárnych najmenších štvorcov.

**Tvrdenie 2.1.** *Za platnosti nulovej hypotézy  $H_0 : \theta^* = \theta_0$  proti alternatíve  $H_1 : \theta^* \neq \theta_0$  má štatistika*

$$F_1 = \frac{(\hat{\theta}_n - \theta_0)' \mathbf{F}(\hat{\theta}_n)' \mathbf{F}(\hat{\theta}_n) (\hat{\theta}_n - \theta_0)}{pS^2} \quad (2.1)$$

*približne F rozdelenie s  $p$  a  $n - p$  stupňami voľnosti.*

*Dôkaz.* Vzhľadom k tomu v normálnom nelineárnom modeli je odhad metódou nelineárnych najmenších štvorcov  $\hat{\theta}$  zároveň maximálne vierohodným odhadom, môžeme k testovaniu hypotézy  $H_0 : \theta^* = \theta_0$  využiť asymptotický Waldov test, ktorého testová štatistika má obecný tvar

$$W(\theta_0) = n(\hat{\theta}_n - \theta_0)' \mathbf{J}(\hat{\theta}_n) (\hat{\theta}_n - \theta_0)' = (\hat{\theta}_n - \theta_0)' \mathbf{J}_n(\hat{\theta}_n) (\hat{\theta}_n - \theta_0)'$$

a po dosadení  $\mathbf{J}_n(\hat{\theta}_n)$  na základe (1.11) dostávame

$$W(\theta_0) = \frac{(\hat{\theta}_n - \theta_0)' \mathbf{F}(\hat{\theta}_n)' \mathbf{F}(\hat{\theta}_n) (\hat{\theta}_n - \theta_0)}{\sigma^2}.$$

Z asymptotickej teórie maximálne vierohodného odhadu vieme, že Waldova štatistika  $W(\boldsymbol{\theta}_0)$  má za platnosti  $H_0$ , ako aj príslušných predpokladov regularity, asymptoticky rozdelenie  $\chi_p^2$ . Približné rozdelenie štatistiky  $F_1$  vyplýva teda z vety 1.11, konkrétne zo skutočností, že pre veľké  $n$  má  $(n-p)S^2/\sigma^2$  približne rozdelenie  $\chi_{n-p}^2$  a náhodné vektory  $\hat{\boldsymbol{\theta}}_n$  a  $S^2$  sú nezávislé. Potom totiž približne platí

$$F_1 = \frac{(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \mathbf{F}(\hat{\boldsymbol{\theta}}_n)' \mathbf{F}(\hat{\boldsymbol{\theta}}_n) (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) / \sigma^2}{\frac{p}{(n-p)S^2/\sigma^2}} \underset{\sim}{\sim} F_{p, n-p}.$$

□

Na základe približného rozdelenia testovej štatistiky  $F_1$  tak dostávame *waldovskú* konfidenčnú množinu s koeficientom spoľahlivosti  $1 - \alpha$  v tvare

$$K_W(\boldsymbol{\theta}^*) = \left\{ \boldsymbol{\theta} \in \Omega : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})' \mathbf{F}(\hat{\boldsymbol{\theta}})' \mathbf{F}(\hat{\boldsymbol{\theta}}) (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) < pS^2 F_{p, n-p}^{1-\alpha} \right\}.$$

Je teda zřejmé, že waldovská množina je ohraničená  $p$ -rozmerným elipsoidom so stredom v bode  $\hat{\boldsymbol{\theta}}$ , obdobne ako je tomu v prípade lineárneho modelu.

Tento prístup je možné zobecniť i pre prípad konštrukcie waldovskej konfidenčnej množiny pre zvolenú podmnožinu parametrov. Uvažujme rozklad vektorového parametra  $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*)'$ , pričom nás bude zaujímať konfidenčná množina iba pre neznámy vektor  $\boldsymbol{\theta}_1^*$  dĺžky  $p_1$ . Ďalej označme rozklad regulárnej matice  $\hat{\mathbf{V}} = \mathbf{F}(\hat{\boldsymbol{\theta}})' \mathbf{F}(\hat{\boldsymbol{\theta}})$  typu  $p \times p$  ako

$$\hat{\mathbf{V}} = \begin{pmatrix} \hat{\mathbf{V}}_{11} & \hat{\mathbf{V}}_{12} \\ \hat{\mathbf{V}}_{21} & \hat{\mathbf{V}}_{22} \end{pmatrix}.$$

Pre inverziu  $\hat{\mathbf{V}}^{-1}$ , ktorej rozklad budeme značiť ako

$$\hat{\mathbf{V}}^{-1} = \begin{pmatrix} \hat{\mathbf{V}}^{11} & \hat{\mathbf{V}}^{12} \\ \hat{\mathbf{V}}^{21} & \hat{\mathbf{V}}^{22} \end{pmatrix},$$

platí napríklad  $\hat{\mathbf{V}}^{11} = (\hat{\mathbf{V}}_{11} - \hat{\mathbf{V}}_{12} \hat{\mathbf{V}}_{22}^{-1} \hat{\mathbf{V}}_{21})^{-1}$ . Pri danom značení má príslušná testová štatistika pre test hypotézy  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$  proti alternatíve  $H_1: \boldsymbol{\theta}_1^* \neq \boldsymbol{\theta}_{10}$  tvar

$$\tilde{F}_1 = \frac{(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{10})' (\hat{\mathbf{V}}^{11})^{-1} (\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{10})}{p_1 S^2}$$

a za platnosti testovanej hypotézy má  $\tilde{F}_1$  približne rozdelenie  $F$  s  $p_1$  a  $n - p$  stupňami voľnosti. Príslušná waldovská konfidenčná množina teda nadobúda tvar

$$\tilde{K}_W(\boldsymbol{\theta}_1^*) = \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : (\boldsymbol{\theta}_1 - \hat{\boldsymbol{\theta}}_1)' (\hat{\mathbf{V}}^{11})^{-1} (\boldsymbol{\theta}_1 - \hat{\boldsymbol{\theta}}_1) < p_1 S^2 F_{p_1, n-p}^{1-\alpha} \right\},$$

kde  $\Omega_1$  značí parametrický priestor vektoru  $\boldsymbol{\theta}_1^*$ . V prípade  $p_1 = 1$  dostaneme pre ľubovoľný parameter  $\theta_j$ , kde  $j = 1, \dots, p$ , waldovský interval spoľahlivosti v tvare  $\left\{ \theta_j \in \Omega_j : (\theta_j - \hat{\theta}_j)^2 (\hat{v}^{jj})^{-1} < S^2 F_{1, n-p}^{1-\alpha} \right\}$ , ktorý je taktiež možno vyjadriť ako

$$CI_W(\theta_j^*) = \left( \hat{\theta}_j - S\sqrt{\hat{v}^{jj}} t_{n-p}^{1-\alpha/2}; \hat{\theta}_j + S\sqrt{\hat{v}^{jj}} t_{n-p}^{1-\alpha/2} \right),$$

kde  $t_{n-p}^{1-\alpha/2} = t_{n-p}(1 - \alpha/2)$  je príslušný kvantil  $t$  rozdelenia s  $n - p$  stupňami voľnosti.

Tento interval spoľahlivosti je samozrejme špeciálnym prípadom konfidénčného intervalu pre lineárnu kombináciu  $\mathbf{a}'\boldsymbol{\theta}^*$ , kde  $\mathbf{a}$  je  $p$ -rozmerný nenulový vektor, ktorý plynie z asymptotickej normality (1.14) a teda je založený na štatistike s približným rozdelením  $t_{n-p}$  v tvare

$$T = \frac{\mathbf{a}'\hat{\boldsymbol{\theta}} - \mathbf{a}'\boldsymbol{\theta}^*}{S(\mathbf{a}'\hat{\mathbf{V}}^{-1}\mathbf{a})^{1/2}}.$$

Príslušný interval spoľahlivosti pre  $\mathbf{a}'\boldsymbol{\theta}^*$  má tvar

$$CI(\mathbf{a}'\boldsymbol{\theta}^*) = \left( \mathbf{a}'\hat{\boldsymbol{\theta}} - S\sqrt{\mathbf{a}'\hat{\mathbf{V}}^{-1}\mathbf{a}} t_{n-p}^{1-\alpha/2}; \mathbf{a}'\hat{\boldsymbol{\theta}} + S\sqrt{\mathbf{a}'\hat{\mathbf{V}}^{-1}\mathbf{a}} t_{n-p}^{1-\alpha/2} \right).$$

Poznamenajme, že okrem vyjadrenia waldovskej konfidénčnej množiny  $K_W(\boldsymbol{\theta}^*)$  sa naskytá možnosť vyjadriť simultánne intervaly spoľahlivosti aj použitím Bonferroniho metódy, a to nahradením kvantilu  $t_{n-p}^{1-\alpha/2}$  vo vyjadrení individuálnych intervalov spoľahlivosti kvantilom  $t_{n-p}^{1-\alpha/(2p)}$ .

## 2.2 Vierohodnostná konfidénčná množina

Druhá často používaná metóda konštrukcie konfidénčnej množiny pre vektorový parameter  $\boldsymbol{\theta}^*$  je založená na pomere vierohodností.

**Tvrdenie 2.2.** *Za platnosti nulovej hypotézy  $H_0 : \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1 : \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  má štatistika*

$$F_2 = \frac{\mathcal{S}(\boldsymbol{\theta}_0) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\mathcal{S}(\hat{\boldsymbol{\theta}}_n)} \frac{n-p}{p} = \frac{\mathcal{S}(\boldsymbol{\theta}_0) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{pS^2} \quad (2.2)$$

*približne F rozdelenie s  $p$  a  $n - p$  stupňami voľnosti.*

*Dôkaz.* Predpokladajme, že rozptyl náhodnej zložky modelu  $\sigma^2$  je známy a uvažujme štatistiku  $LR(\boldsymbol{\theta}_0)$  v tvare

$$LR(\boldsymbol{\theta}_0) = 2[L(\hat{\boldsymbol{\theta}}_n) - L(\boldsymbol{\theta}_0)],$$

kde odhad metódou nelineárnych najmenších štvorcov  $\hat{\boldsymbol{\theta}}_n$  maximalizuje logaritmickej vierohodnostnú funkciu  $L(\boldsymbol{\theta}, \sigma^2)$  v tvare (1.10). Po dosadení a jednoduchej úprave dostávame

$$LR(\boldsymbol{\theta}_0) = \frac{\mathcal{S}(\boldsymbol{\theta}_0) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\sigma^2}.$$

Z teórie maximálnej vierohodnosti vieme, že štatistika  $LR(\boldsymbol{\theta}_0)$  má asymptoticky  $\chi_p^2$  rozdelenie. Približné rozdelenie štatistiky  $F_2$  opäť vyplýva zo skutočností, že pre veľké  $n$  má  $(n-p)S^2/\sigma^2$  približne rozdelenie  $\chi_{n-p}^2$  a náhodné vektory  $\hat{\boldsymbol{\theta}}_n$  a  $S^2$  sú nezávislé, t.j. platí

$$F_2 = \frac{\frac{(\mathcal{S}(\boldsymbol{\theta}_0) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n))/\sigma^2}{p}}{\frac{(n-p)S^2/\sigma^2}{n-p}} \sim \mathbf{F}_{p, n-p}$$

□

Približné rozdelenie štatistiky  $F_2$  je možné v súlade s (1.19) odvodiť pomocou lineárnej aproximácie (1.9). Výraz (2.2) je totiž možné za platnosti  $\mathbf{H}_0$  aproximovať ako

$$F_2 = \frac{[\mathcal{S}(\boldsymbol{\theta}_0) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)]/p}{\mathcal{S}(\hat{\boldsymbol{\theta}}_n)/(n-p)} \approx \frac{[\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}'(\mathbf{I}_n - \mathbf{H}_{\mathbf{F}}^*)\boldsymbol{\varepsilon}]/p}{\boldsymbol{\varepsilon}'(\mathbf{I}_n - \mathbf{H}_{\mathbf{F}}^*)\boldsymbol{\varepsilon}/(n-p)} = \frac{\boldsymbol{\varepsilon}'\mathbf{H}_{\mathbf{F}}^*\boldsymbol{\varepsilon}/p}{\boldsymbol{\varepsilon}'(\mathbf{I}_n - \mathbf{H}_{\mathbf{F}}^*)\boldsymbol{\varepsilon}/(n-p)},$$

kde vzhľadom k vlastnostiam idempotentnej matice  $\mathbf{H}_{\mathbf{F}}^*$  a rozdeleniu náhodnej zložky modelu  $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$  platí, že  $\boldsymbol{\varepsilon}'\mathbf{H}_{\mathbf{F}}^*\boldsymbol{\varepsilon}/\sigma^2$  a  $\boldsymbol{\varepsilon}'(\mathbf{I}_n - \mathbf{H}_{\mathbf{F}}^*)\boldsymbol{\varepsilon}/\sigma^2$  sú nezávislé s rozdeleniami  $\chi_p^2$ , resp.  $\chi_{n-p}^2$ .

Na základe približného rozdelenia testovej štatistiky  $F_2$  dostaneme tzv. *vierohodnostnú* konfidenčnú množinu s koeficientom spoľahlivosti  $1 - \alpha$  v tvare

$$K_{LR}(\boldsymbol{\theta}^*) = \left\{ \boldsymbol{\theta} \in \Omega : \mathcal{S}(\boldsymbol{\theta}) < \mathcal{S}(\hat{\boldsymbol{\theta}}) \left( 1 + \frac{p}{n-p} \mathbf{F}_{p, n-p}^{1-\alpha} \right) \right\}.$$

Obdobne, ako je tomu v prípade waldovskej konfidenčnej množiny, je možné skonštruovať vierohodnostné konfidenčné množiny iba pre zvolenú podmnožinu parametrov. Opäť budeme uvažovať rozklad vektoru  $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*)'$ , pričom nás bude zaujímať konfidenčná množina pre neznámy vektor  $\boldsymbol{\theta}_1^*$  dĺžky  $p_1$ , tzn. vektor parametrov  $\boldsymbol{\theta}_2^*$  budeme považovať za rušivý. Za účelom konštrukcie príslušnej konfidenčnej množiny budeme teda potrebovať maximálne vierohodný odhad vektorového parametra  $\boldsymbol{\theta}_2^*$  za podmienky  $\mathbf{H}_0 : \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$ , pričom tento odhad označíme ako  $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10})$ . Následným dosadením vektoru  $\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}) = (\boldsymbol{\theta}_{10}, \tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10}))$  do logaritmickej vierohodnostnej funkcie dostaneme tzv. *profilovú* logaritmickú vierohodnostnú funkciu  $L(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))$ . Príslušný pomer vierohodností bude mať pri znalosti rozptylu  $\sigma^2$  tvar

$$\widetilde{LR}(\boldsymbol{\theta}_{10}) = 2[L(\hat{\boldsymbol{\theta}}_n) - L(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))] = \frac{\mathcal{S}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\sigma^2}$$

a po analogickom dosadení konzistentného odhadu parametra  $\sigma^2$  dostaneme testovú štatistiku

$$\tilde{F}_2 = \frac{\mathcal{S}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) - \mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{p_1 S^2},$$

ktorá má za platnosti nulovej hypotézy približne rozdelenie  $F$  s  $p_1$  a  $n - p$  stupňami voľnosti. Príslušná vierohodnostná konfidenčná množina pre  $\theta_1^*$  má tak tvar

$$\tilde{K}_{LR}(\theta_1^*) = \left\{ \theta_1 \in \Omega_1 : \mathcal{S}(\tilde{\theta}(\theta_1)) < \mathcal{S}(\hat{\theta}) \left( 1 + \frac{p_1}{n-p} F_{p_1, n-p}^{1-\alpha} \right) \right\}.$$

V špeciálnom prípade  $p_1 = 1$  dostaneme pre ľubovoľný parameter  $\theta_j$ , kde  $j = 1, \dots, p$ , vierohodnostný interval spoľahlivosti v tvare

$$CI_{LR}(\theta_j^*) = \left\{ \theta_j \in \Omega_j : \mathcal{S}(\tilde{\theta}(\theta_j)) < \mathcal{S}(\hat{\theta}) \left( 1 + \frac{1}{n-p} F_{1, n-p}^{1-\alpha} \right) \right\}.$$

Pre vyjadrenie simultánnych intervalov spoľahlivosti možno opäť využiť Bonferroniho metódu.

Vzhľadom k tomu, že obe popísané metódy konštrukcie približných konfidenčných množín sú založené na asymptotickej teórii maximálne vierohodného odhadu, sa naskytá prirodzená otázka, aké sú medzi týmito metódami hlavné rozdiely.

V prvom rade, waldovské konfidenčné množiny sú analógiou konfidenčných množín pre parametre normálneho lineárneho modelu, kde matica modelu  $\mathbf{X}$  je nahradená maticou  $\mathbf{F}(\hat{\theta})$ , a teda waldovské množiny silne závisia na platnosti lineárnej aproximácie regresnej funkcie (1.6) v blízkosti bodu  $\theta^*$ , ktorá je platná iba asymptoticky. Ďalšou nevýhodou waldovského prístupu je, že príslušné množiny závisia na parametrizácii modelu a ich presnosť je tak ovplyvnená aj parametrickou krivosťou modelu. Veľkou výhodou tohto prístupu je však pomerná jednoduchosť výpočtu, ako i znázornenia tejto konfidenčnej množiny.

Naopak, vierohodnostné konfidenčné množiny vyplývajú z reziduálneho súčtu štvorcov, a tak viac odpovedajú intuitívnej požiadavke blízkosti pozorovaní vysvetľovanej premennej  $\mathbf{Y}$  a odhadu  $\mathbf{f}(\hat{\theta})$ . Nespornou výhodou tohto prístupu je i nezávislosť na parametrizácii modelu, tzn. tieto množiny sú ovplyvnené iba vnútornou krivosťou modelu, ktorá je však často zanedbateľná. Na druhej strane, hlavnou nevýhodou vierohodnostného prístupu je výrazne vyššia výpočtová náročnosť spojená s určením celej množiny  $K_{LR}$ , ako i komplikácie pri grafickom znázornení tejto množiny.

Ďalším zaujímavým faktom je, že waldovské i vierohodnostné konfidenčné množiny sú asymptoticky ekvivalentné. Na túto skutočnosť môžeme poukázať napríklad využitím aproximácie  $\mathbf{F}(\hat{\theta})(\hat{\theta} - \theta^*) \approx \mathbf{F}^*(\hat{\theta} - \theta^*) \approx \mathbf{H}_{\mathbf{F}}^* \varepsilon$  vo vzťahu

$$(\hat{\theta} - \theta^*)' \mathbf{F}(\hat{\theta})' \mathbf{F}(\hat{\theta}) (\hat{\theta} - \theta^*) \approx \varepsilon' \mathbf{H}_{\mathbf{F}}^* \mathbf{H}_{\mathbf{F}}^* \varepsilon = \varepsilon' \mathbf{H}_{\mathbf{F}}^* \varepsilon \approx \mathcal{S}(\theta^*) - \mathcal{S}(\hat{\theta}).$$

Asymptotická ekvivalencia týchto vzťahov je zabezpečená poznámkou 1.12. Napriek tomu sa môže waldovská a vierohodnostná konfidenčná množina v prípade konečného výberu výrazne líšiť, pričom v takomto prípade je z uvedených dôvodov preferovaná práve vierohodnostná množina.

**Poznámka 2.3.** Vzhľadom k skutočnosti, že vierohodnostné konfidenčné množiny môžu byť negatívne ovplyvnené významnou vnútornou krivosťou modelu, Hamilton a Wiens [12] navrhli korekčný člen, ktorý na vnútornej krivosti závisí, a pomocou ktorého je možné určiť korigovanú vierohodnostnú konfidenčnú množinu pre danú podmnožinu parametrov. Konkrétny tvar korekčného člena, ako i korigovanej konfidenčnej množiny nájde záujemca napríklad v monografii Seber a Wild [18], ods. 5.3.

## 2.3 Skórová konfidenčná množina

Ukázali sme, že dve najčastejšie používané konfidenčné množiny pre parametre nelineárneho modelu sú vo svojej podstate založené na vhodných úpravách známych asymptotických testových štatistík - v prvom prípade Waldovho testu a v prípade vierohodnostných množín testu pomerom vierohodností. Prirodzene sa teda núka možnosť konštrukcie konfidenčnej množiny na základe testovej štatistiky skórového testu.

**Tvrdenie 2.4.** *Za platnosti nulovej hypotézy  $H_0 : \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1 : \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  má štatistika*

$$F_3 = \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}(\boldsymbol{\theta}_0) [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]}{pS^2} \quad (2.3)$$

približne  $F$  rozdelenie s  $p$  a  $n - p$  stupňami voľnosti.

*Dôkaz.* Opäť predpokladajme, že rozptyl náhodnej zložky modelu  $\sigma^2$  je známy a uvažujme skórovú štatistiku  $LM(\boldsymbol{\theta}_0)$  v tvare

$$LM(\boldsymbol{\theta}_0) = \frac{1}{n} \mathbf{U}(\boldsymbol{\theta}_0)' [\mathbf{J}(\boldsymbol{\theta}_0)]^{-1} \mathbf{U}(\boldsymbol{\theta}_0) = \mathbf{U}(\boldsymbol{\theta}_0)' [\mathbf{J}_n(\boldsymbol{\theta}_0)]^{-1} \mathbf{U}(\boldsymbol{\theta}_0),$$

kde  $\mathbf{U}(\boldsymbol{\theta}_0)$  je  $p$ -rozmerný vektor prvých parciálnych derivácií logaritmickej vierohodnostnej funkcie  $L(\boldsymbol{\theta})$  v bode  $\boldsymbol{\theta}_0$ , t.j.

$$\mathbf{U}(\boldsymbol{\theta}_0) = \frac{\partial}{\partial \boldsymbol{\theta}} L(\boldsymbol{\theta}_0) = -\frac{1}{2\sigma^2} \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{S}(\boldsymbol{\theta}_0) = \frac{1}{\sigma^2} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)].$$

Po dosadení  $\mathbf{U}(\boldsymbol{\theta}_0)$  a Fisherovej informačnej matice (1.11) v bode  $\boldsymbol{\theta}_0$  dostávame

$$LM(\boldsymbol{\theta}_0) = \frac{1}{\sigma^2} [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}(\boldsymbol{\theta}_0) [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)].$$

Vieme, že štatistika  $LM(\boldsymbol{\theta}_0)$  má asymptoticky  $\chi_p^2$  rozdelenie, a po analogickom nahradení neznámeho rozptylu náhodnej zložky  $\sigma^2$  príslušným konzistentným odhadom  $S^2$  dostaneme s využitím vlastností z vety 1.11 približné rozdelenie štatistiky  $F_3$ , t.j.

$$F_3 = \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}(\boldsymbol{\theta}_0) [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] / \sigma^2}{\frac{p}{(n-p) S^2 / \sigma^2}} \sim F_{p, n-p}.$$

□

Analogicky, ako tomu bolo v prípade waldovskej a vierohodnostnej množiny, dostávame na základe približného rozdelenia testovej štatistiky  $F_3$  skórovú konfidenčnú množinu s koeficientom spoľahlivosti  $1 - \alpha$  v tvare

$$K_{LM}(\boldsymbol{\theta}^*) = \left\{ \boldsymbol{\theta} \in \Omega : [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' \mathbf{H}_F(\boldsymbol{\theta}) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})] < pS^2 F_{p,n-p}^{1-\alpha} \right\},$$

kde  $\mathbf{H}_F(\boldsymbol{\theta}) = \mathbf{F}_\bullet(\boldsymbol{\theta}) [\mathbf{F}_\bullet(\boldsymbol{\theta})' \mathbf{F}_\bullet(\boldsymbol{\theta})]^{-1} \mathbf{F}_\bullet(\boldsymbol{\theta})'$ .

V prípade konštrukcie skórovej konfidenčnej množiny pre zvolenú podmnožinu parametrov, t.j. pre vektor  $\boldsymbol{\theta}_1^*$  dĺžky  $p_1$ , budeme rovnako ako v prípade vierohodnostnej množiny potrebovať maximálne vierohodný odhad rušivého vektorového parametra  $\boldsymbol{\theta}_2^*$  za podmienky  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$ , pre ktorý sme zaviedli značenie  $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10})$ . Ak budeme opäť uvažovať vektor  $\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}) = (\boldsymbol{\theta}_{10}, \tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10}))$ , tak skórový test hypotézy  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$  proti  $H_1: \boldsymbol{\theta}_1^* \neq \boldsymbol{\theta}_{10}$  v prípade rušivého parametra  $\boldsymbol{\theta}_2^*$  a pri známom rozptyle  $\sigma^2$  je založený na štatistike

$$\widetilde{LM}(\boldsymbol{\theta}_{10}) = \mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))' (\sigma^2 \tilde{\mathbf{V}}^{11}) \mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})),$$

kde  $\tilde{\mathbf{V}}^{11}$  je submatice matice  $\tilde{\mathbf{V}}^{-1} = [\mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))' \mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]^{-1}$  príslušná vektoru  $\boldsymbol{\theta}_1$  a  $\mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))$  je  $p_1$ -rozmerný vektor prvých parciálnych derivácií logaritmickej vierohodnostnej funkcie  $L(\boldsymbol{\theta})$  podľa  $\boldsymbol{\theta}_1$  v bode  $\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})$ , t.j.

$$\mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) = \frac{\partial}{\partial \boldsymbol{\theta}_1} L(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) = -\frac{1}{2\sigma^2} \frac{\partial}{\partial \boldsymbol{\theta}_1} \mathcal{S}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})).$$

Štatistika  $\widetilde{LM}(\boldsymbol{\theta}_{10})$  má za platnosti  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$  opäť asymptoticky  $\chi_{p_1}^2$  rozdelenie.

Ak si uvedomíme, že odhad  $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10})$  vyhovuje normálnym rovniciam (1.5), t.j.  $\mathbf{U}_2(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))$  je nulový vektor, možno štatistiku  $\widetilde{LM}(\boldsymbol{\theta}_{10})$  vyjadriť v známejšom tvare (viď Seber a Wild [18], ods. 5.3)

$$\begin{aligned} \widetilde{LM}(\boldsymbol{\theta}_{10}) &= \begin{pmatrix} \mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) \\ \mathbf{U}_2(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) \end{pmatrix}' \left[ \mathbf{J}_n(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) \right]^{-1} \begin{pmatrix} \mathbf{U}_1(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) \\ \mathbf{U}_2(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) \end{pmatrix} \\ &= \frac{1}{\sigma^2} [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]' \tilde{\mathbf{H}}_F(\boldsymbol{\theta}_{10}) [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))], \end{aligned}$$

kde  $\tilde{\mathbf{H}}_F(\boldsymbol{\theta}_{10}) = \mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) [\mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))' \mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]^{-1} \mathbf{F}_\bullet(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))'$ . Po nahradení neznámeho rozptylu  $\sigma^2$  jeho konzistentným odhadom  $S^2$  teda dostaneme štatistiku

$$\tilde{F}_3 = \frac{[\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]' \tilde{\mathbf{H}}_F(\boldsymbol{\theta}_{10}) [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]}{p_1 S^2},$$

ktorá má pri platnosti testovanej hypotézy približne rozdelenie  $F$  s  $p_1$  a  $n - p$  stupňami voľnosti. Príslušná konfidenčná množina má teda tvar

$$\tilde{K}_{LM}(\boldsymbol{\theta}_1^*) = \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))] \tilde{\mathbf{H}}_F(\boldsymbol{\theta}_1) [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))] < p_1 S^2 F_{p_1, n-p}^{1-\alpha} \right\}.$$

V prípade konštrukcie individuálneho skórového intervalu spoľahlivosti pre ľubovoľný parameter  $\theta_j$ , kde  $j = 1, \dots, p$ , dostaneme

$$CI_{LM}(\theta_j^*) = \left\{ \theta_j \in \Omega_j : [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\theta_j))]'\tilde{\mathbf{H}}_{\mathbf{F}}(\theta_j)[\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\theta_j))] < S^2 \mathbf{F}_{1, n-p}^{1-\alpha} \right\},$$

prípadne v tvare založenom na  $U_j(\tilde{\boldsymbol{\theta}}_n(\theta_j)) = \frac{\partial}{\partial \theta_j} L(\tilde{\boldsymbol{\theta}}_n(\theta_j)) = \frac{-1}{2\sigma^2} \frac{\partial}{\partial \theta_j} \mathcal{S}(\tilde{\boldsymbol{\theta}}_n(\theta_j))$

$$\begin{aligned} CI_{LM}(\theta_j^*) &= \left\{ \theta_j \in \Omega_j : \left[ \frac{\partial}{\partial \theta_j} \mathcal{S}(\tilde{\boldsymbol{\theta}}(\theta_j)) \right]^2 \tilde{v}^{jj} < 4S^2 \mathbf{F}_{1, n-p}^{1-\alpha} \right\} \\ &= \left\{ \theta_j \in \Omega_j : \left| \mathbf{F}_{\cdot j}(\tilde{\boldsymbol{\theta}}(\theta_j))' [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\theta_j))] \right| < \frac{S}{\sqrt{\tilde{v}^{jj}}} t_{n-p}^{1-\alpha/2} \right\}, \end{aligned}$$

kde  $\mathbf{F}_{\cdot j}(\tilde{\boldsymbol{\theta}}_n(\theta_j))$  je  $j$ -tý stĺpec matice  $\mathbf{F}_{\cdot}$  v bode  $\tilde{\boldsymbol{\theta}}_n(\theta_j)$ . Podobne, ako tomu je v prípade vierohodnostného prístupu, je miernym problémom skórových konfidenčných množín presné určenie, ako i znázornenie tejto množiny. Navyše sa môže stať, že množina  $CI_{LM}(\theta_j^*)$ , rovnako ako  $CI_{LR}(\theta_j^*)$ , bude tvorená niekoľkými disjunktnými intervalmi.

**Poznámka 2.5.** Je vhodné si uvedomiť, že v reálnych aplikáciách nebude z dôvodu iteračného výpočtu odhadu  $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_{10})$  vektor  $\mathbf{U}_2(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))$  rovný nulovému vektoru. Jeho zložky však budú blízke nule, a teda použitie rozšíreného tvaru štatistiky  $\tilde{LM}(\boldsymbol{\theta}_{10})$  by týmto faktom nemalo byť ovplyvnené.

## 2.4 Asymptotické konfidenčné množiny

Obecnou metódou, ako zostrojiť konfidenčné množiny pre vektor neznámych parametrov  $\boldsymbol{\theta}^*$ , je konštrukcia vhodnej testovej štatistiky za účelom testu hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$ , ktorá bude mať za platnosti hypotézy  $H_0$  známe rozdelenie, a následné vytvorenie konfidenčnej množiny z bodov  $\boldsymbol{\theta}_0$ , pre ktoré tento test príslušnú nulovú hypotézu na danej hladine a pri daných dátach nezamietne. V tejto kapitole si uvedieme *asymptotické* konfidenčné množiny, ktoré predpokladajú dostatočný rozsah výberu a sú založené na asymptotických testoch hypotézy  $H_0$ , u ktorých je obecné neznámy rozptyl  $\sigma^2$  považovaný za rušivý parameter. Opäť je možné využiť analógie Waldovho testu, skórového testu či testu pomerom vierohodností.

**Tvrdenie 2.6.** *Za platnosti nulovej hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  majú štatistiky*

$$i) W^{as.}(\boldsymbol{\theta}_0) = \frac{(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \mathbf{F}_{\cdot}(\hat{\boldsymbol{\theta}}_n) \mathbf{F}_{\cdot}(\hat{\boldsymbol{\theta}}_n) (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)}{\mathcal{S}(\hat{\boldsymbol{\theta}}_n)/n}$$

$$ii) LR^{as.}(\boldsymbol{\theta}_0) = n[\log \mathcal{S}(\boldsymbol{\theta}_0) - \log \mathcal{S}(\hat{\boldsymbol{\theta}}_n)]$$

$$iii) LM^{as.}(\boldsymbol{\theta}_0) = \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}_{\cdot}(\boldsymbol{\theta}_0) [\mathbf{F}_{\cdot}(\boldsymbol{\theta}_0) \mathbf{F}_{\cdot}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}_{\cdot}(\boldsymbol{\theta}_0) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]}{\mathcal{S}(\boldsymbol{\theta}_0)/n}$$

asymptoticky  $\chi^2$  rozdelenie s  $p$  stupňami voľnosti.

*Dôkaz.* V podkapitole 1.2.2 sme si ukázali, že  $\hat{\boldsymbol{\theta}}$  je maximálne vierohodným odhadom neznámeho vektoru parametrov  $\boldsymbol{\theta}^*$  a  $\hat{\sigma}^2 = \mathcal{S}(\hat{\boldsymbol{\theta}})/n$  je maximálne vierohodným odhadom rozptylu  $\sigma^2$ . Na základe vierohodnosti modelu (1.10) je možné jednoducho ukázať, že pri obmedzení  $\boldsymbol{\theta} = \boldsymbol{\theta}_0$  sa dá vyjadriť maximálne vierohodný odhad rozptylu  $\sigma^2$  ako  $\hat{\sigma}_0^2 = \mathcal{S}(\boldsymbol{\theta}_0)/n$ .

Ukážeme, že príslušné testové štatistiky sú vskutku testovými štatistikami známych asymptotických testov s rušivým parametrom  $\sigma^2$ . K tomu však budeme potrebovať rozklad inverzie Fisherovej informačnej matice (1.11)

$$[\mathbf{J}_n(\boldsymbol{\theta}^*, \sigma^2)]^{-1} = \begin{pmatrix} \mathbf{J}_n^{11}(\boldsymbol{\theta}^*, \sigma^2) & \mathbf{J}_n^{12}(\boldsymbol{\theta}^*, \sigma^2) \\ \mathbf{J}_n^{21}(\boldsymbol{\theta}^*, \sigma^2) & \mathbf{J}_n^{22}(\boldsymbol{\theta}^*, \sigma^2) \end{pmatrix},$$

pričom z vlastností pozitívne semidefinitnej blokovej matice platí

$$\begin{aligned} \mathbf{J}_n^{11}(\boldsymbol{\theta}^*, \sigma^2) &= \sigma^2 [\mathbf{F} \cdot (\boldsymbol{\theta}^*)' \mathbf{F} \cdot (\boldsymbol{\theta}^*)]^{-1}; & \mathbf{J}_n^{12}(\boldsymbol{\theta}^*, \sigma^2) &= \mathbf{0}; \\ \mathbf{J}_n^{21}(\boldsymbol{\theta}^*, \sigma^2) &= \mathbf{0}'; & \mathbf{J}_n^{22}(\boldsymbol{\theta}^*, \sigma^2) &= 2n^{-1}\sigma^4. \end{aligned}$$

i) Waldov test

Waldova testová štatistika testu hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  s rušivým parametrom  $\sigma^2$  má tvar (viď Anděl [2])

$$\begin{aligned} W(\boldsymbol{\theta}_0) &= (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' [\mathbf{J}_n^{11}(\hat{\boldsymbol{\theta}}_n, \hat{\sigma}^2)]^{-1} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \\ &= (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' [\hat{\sigma}^{-2} \mathbf{F} \cdot (\hat{\boldsymbol{\theta}}_n)' \mathbf{F} \cdot (\hat{\boldsymbol{\theta}}_n)] (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \\ &= \frac{(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \mathbf{F} \cdot (\hat{\boldsymbol{\theta}}_n)' \mathbf{F} \cdot (\hat{\boldsymbol{\theta}}_n) (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)'}{\mathcal{S}(\hat{\boldsymbol{\theta}})/n} \end{aligned}$$

a pri platnosti testovanej hypotézy konverguje v distribúcii k náhodnej veličine s rozdelením  $\chi_p^2$ .

ii) test pomerom vierohodností

Pomer vierohodností v prípade testu hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  s rušivým parametrom  $\sigma^2$  má tvar

$$\begin{aligned} LR(\boldsymbol{\theta}_0) &= 2[L(\hat{\boldsymbol{\theta}}_n, \hat{\sigma}^2) - L(\boldsymbol{\theta}_0, \hat{\sigma}_0^2)] \\ &= 2 \left[ \frac{n}{2} \log \hat{\sigma}_0^2 + \frac{1}{2\hat{\sigma}_0^2} \mathcal{S}(\boldsymbol{\theta}_0) - \frac{n}{2} \log \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2} \mathcal{S}(\hat{\boldsymbol{\theta}}_n) \right] \\ &= n \log \frac{\mathcal{S}(\boldsymbol{\theta}_0)}{n} + \frac{n\mathcal{S}(\boldsymbol{\theta}_0)}{\mathcal{S}(\boldsymbol{\theta}_0)} - n \log \frac{\mathcal{S}(\hat{\boldsymbol{\theta}})}{n} - \frac{n\mathcal{S}(\hat{\boldsymbol{\theta}}_n)}{\mathcal{S}(\hat{\boldsymbol{\theta}}_n)} \\ &= n[\log \mathcal{S}(\boldsymbol{\theta}_0) - \log \mathcal{S}(\hat{\boldsymbol{\theta}}_n)] \end{aligned}$$

a pri platnosti testovanej hypotézy konverguje v distribúcii k náhodnej veličine s rozdelením  $\chi_p^2$ .

iii) skórový test

Testová štatistika skórového testu hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  s rušivým parametrom  $\sigma^2$  má tvar

$$\begin{aligned} LM(\boldsymbol{\theta}_0) &= \mathbf{U}_1(\boldsymbol{\theta}_0, \hat{\sigma}_0^2)' \mathbf{J}_n^{11}(\boldsymbol{\theta}_0, \hat{\sigma}_0^2) \mathbf{U}_1(\boldsymbol{\theta}_0, \hat{\sigma}_0^2) \\ &= \left[ \frac{\partial}{\partial \boldsymbol{\theta}} L(\boldsymbol{\theta}_0, \hat{\sigma}_0^2) \right]' \hat{\sigma}_0^2 [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} L(\boldsymbol{\theta}_0, \hat{\sigma}_0^2) \right] \\ &= \frac{1}{\hat{\sigma}_0^2} [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}(\boldsymbol{\theta}_0) \hat{\sigma}_0^2 [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \frac{1}{\hat{\sigma}_0^2} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] \\ &= \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{F}(\boldsymbol{\theta}_0) [\mathbf{F}(\boldsymbol{\theta}_0)' \mathbf{F}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}(\boldsymbol{\theta}_0)' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]}{\mathcal{S}(\boldsymbol{\theta}_0)/n} \end{aligned}$$

a pri platnosti testovanej hypotézy takisto konverguje v distribúcii k náhodnej veličine s rozdelením  $\chi_p^2$ .  $\square$

Na základe rozdelenia príslušných asymptotických testových štatistík je možné určiť asymptotické konfidenčné množiny pre vektorový parameter  $\boldsymbol{\theta}^*$  v tvare

$$\begin{aligned} K_W^{as.}(\boldsymbol{\theta}^*) &= \left\{ \boldsymbol{\theta} \in \Omega : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})' \mathbf{F}(\hat{\boldsymbol{\theta}})' \mathbf{F}(\hat{\boldsymbol{\theta}}) (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) < \hat{\sigma}^2 \chi_p^2(1 - \alpha) \right\}', \\ K_{LR}^{as.}(\boldsymbol{\theta}^*) &= \left\{ \boldsymbol{\theta} \in \Omega : n [\log \mathcal{S}(\boldsymbol{\theta}) - \log \mathcal{S}(\hat{\boldsymbol{\theta}})] < \chi_p^2(1 - \alpha) \right\}, \\ K_{LM}^{as.}(\boldsymbol{\theta}^*) &= \left\{ \boldsymbol{\theta} \in \Omega : [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' \mathbf{H}_F(\boldsymbol{\theta}) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})] < \hat{\sigma}_{\boldsymbol{\theta}}^2 \chi_p^2(1 - \alpha) \right\}, \end{aligned}$$

kde  $\hat{\sigma}^2 = \mathcal{S}(\hat{\boldsymbol{\theta}})/n$ ,  $\hat{\sigma}_{\boldsymbol{\theta}}^2 = \mathcal{S}(\boldsymbol{\theta})/n$ ,  $\mathbf{H}_F(\boldsymbol{\theta}) = \mathbf{F}(\boldsymbol{\theta}) [\mathbf{F}(\boldsymbol{\theta})' \mathbf{F}(\boldsymbol{\theta})]^{-1} \mathbf{F}(\boldsymbol{\theta})'$  a  $\chi_p^2(1 - \alpha)$  je príslušný kvantil  $\chi^2$  rozdelenia s  $p$  stupňami voľnosti.

Zcela analogickým postupom je možné zostrojiť i asymptotické konfidenčné množiny pre zvolenú podmnožinu parametrov  $\boldsymbol{\theta}_1$ , ktorých špeciálnym prípadom ( $p_1 = 1$ ) sú aj individuálne intervaly spoľahlivosti a ktoré možno zapísať v tvare

$$\begin{aligned} \tilde{K}_W^{as.}(\boldsymbol{\theta}_1^*) &= \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : (\boldsymbol{\theta}_1 - \hat{\boldsymbol{\theta}}_1)' (\hat{\mathbf{V}}^{11})^{-1} (\boldsymbol{\theta}_1 - \hat{\boldsymbol{\theta}}_1) < \hat{\sigma}^2 \chi_{p_1}^2(1 - \alpha) \right\}, \\ \tilde{K}_{LR}^{as.}(\boldsymbol{\theta}_1^*) &= \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : n [\log \mathcal{S}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1)) - \log \mathcal{S}(\hat{\boldsymbol{\theta}})] < \chi_{p_1}^2(1 - \alpha) \right\}, \\ \tilde{K}_{LM}^{as.}(\boldsymbol{\theta}_1^*) &= \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))] \tilde{\mathbf{H}}_F(\boldsymbol{\theta}_1) [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))] < \tilde{\sigma}_{\boldsymbol{\theta}_1}^2 \chi_{p_1}^2(1 - \alpha) \right\}, \end{aligned}$$

kde  $\tilde{\mathbf{H}}_F(\boldsymbol{\theta}_1) = \mathbf{F}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1)) [\mathbf{F}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))' \mathbf{F}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))]^{-1} \mathbf{F}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))'$  a  $\tilde{\sigma}_{\boldsymbol{\theta}_1}^2 = \mathcal{S}(\tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}_1))/n$ .

**Poznámka 2.7.** Je zjavné, že napríklad waldovská a asymptotická waldovská konfidenčná množina sa líšia iba výrazom na pravej strane príslušných nerovností, pričom platí  $\hat{\sigma}^2 \chi_p^2(1 - \alpha) < p S^2 \mathbf{F}_{p, n-p}^{1-\alpha}$ , tzn. asymptotická verzia waldovskej množiny je podmnožinou približnej waldovskej množiny. Samozrejme sú tieto množiny asymptoticky ekvivalentné – ak má totiž premenná  $X$  rozdelenie  $F$  s  $\nu_1$  a  $\nu_2$  stupňami voľnosti, tak  $\nu_1 X \xrightarrow[\nu_2 \rightarrow \infty]{\mathcal{D}} Y$ , kde náhodná veličina  $Y$  má  $\chi^2$  rozdelenie s  $\nu_1$  stupňami voľnosti, a teda platí  $\lim_{n \rightarrow \infty} p \mathbf{F}_{p, n-p}^{1-\alpha} = \chi_p^2(1 - \alpha)$ .

## 2.5 Presné konfidenčné množiny

Doposiaľ sme sa zaoberali konfidenčnými množinami, ktoré boli založené na štatistikách s približným  $F$  rozdelením, resp. asymptotickým  $\chi^2$  rozdelením. V tejto podkapitole si uvedieme obecnú metódu konštrukcie štatistiky pre test hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$ , ktorá má presne  $F$  rozdelenie. Najprv však uvedieme špeciálny prípad takejto štatistiky, ktorú je možné získať úpravou skórovej testovej štatistiky (2.3).

**Tvrdenie 2.8.** *Za platnosti nulovej hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti alternatíve  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  má štatistika*

$$F_4 = \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{H}_F(\boldsymbol{\theta}_0) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] / p}{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' [\mathbf{I}_n - \mathbf{H}_F(\boldsymbol{\theta}_0)] [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] / (n - p)} \quad (2.4)$$

presne  $F$  rozdelenie s  $p$  a  $n - p$  stupňami voľnosti.

*Dôkaz.* Za platnosti nulovej môžeme štatistiku  $F_4$  vyjadriť ako

$$F_4 = \frac{\boldsymbol{\varepsilon}' \mathbf{H}_F^* \boldsymbol{\varepsilon} / p}{\boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{H}_F^*) \boldsymbol{\varepsilon} / (n - p)},$$

kde  $\mathbf{H}_F^* = \mathbf{F}_*^* (\mathbf{F}_*^{*'} \mathbf{F}_*^*)^{-1} \mathbf{F}_*^{*'}$  je symetrická idempotentná matica s hodnotou  $p$ . Je zrejmé, že matica  $(\mathbf{I}_n - \mathbf{H}_F^*)$  je takisto idempotentná s hodnotou  $n - p$  a navyše, vzhľadom k tomu, že  $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$  a zároveň  $\mathbf{H}_F^* (\mathbf{I}_n - \mathbf{H}_F^*) = \mathbf{0}$ , sú náhodné veličiny  $\boldsymbol{\varepsilon}' \mathbf{H}_F^* \boldsymbol{\varepsilon}$  a  $\boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{H}_F^*) \boldsymbol{\varepsilon}$  nezávislé s  $\chi^2$  rozdelením s  $p$ , resp.  $n - p$  stupňami voľnosti. Štatistika  $F_4$  teda má  $F$  rozdelenie s  $p$  a  $n - p$  stupňami voľnosti.  $\square$

Na základe štatistiky  $F_4$  tak dostávame presnú konfidenčnú množinu pre parameter  $\boldsymbol{\theta}$  v tvare

$$\begin{aligned} K_F(\boldsymbol{\theta}^*) &= \left\{ \boldsymbol{\theta} \in \Omega : \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' \mathbf{H}_F(\boldsymbol{\theta}) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]}{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' [\mathbf{I}_n - \mathbf{H}_F(\boldsymbol{\theta})] [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]} < \frac{p}{n - p} \mathbf{F}_{p, n-p}^{1-\alpha} \right\} \\ &= \left\{ \boldsymbol{\theta} \in \Omega : [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})]' \mathbf{H}_F(\boldsymbol{\theta}) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta})] < \hat{\sigma}_\theta^2 \frac{np \mathbf{F}_{p, n-p}^{1-\alpha}}{n - p + p \mathbf{F}_{p, n-p}^{1-\alpha}} \right\}. \end{aligned}$$

Je teda zrejmé, že táto presná konfidenčná množina sa líši od skórovej konfidenčnej množiny iba pravou stranou príslušných nerovností. Navyše, porovnaním s asymptotickou verziou skórovej konfidenčnej množiny zistíme, že  $K_F(\boldsymbol{\theta}^*) \subset K_{LM}^{as}(\boldsymbol{\theta}^*)$  a že tieto množiny sú asymptoticky ekvivalentné.

Obecnú metódu konštrukcie štatistiky pre test hypotézy  $H_0: \boldsymbol{\theta}^* = \boldsymbol{\theta}_0$  proti  $H_1: \boldsymbol{\theta}^* \neq \boldsymbol{\theta}_0$  s presným  $F$  rozdelením navrhol na základe rozkladu reziduálneho súčtu štvorcov Hartley [13]. Nech  $\mathbf{P}$  je ľubovoľná symetrická idempotentná matica typu  $n \times n$  a s hodnotou  $p$ . Je zrejmé, že matica  $(\mathbf{I}_n - \mathbf{P})$  je taktiež idempotentná s hodnotou  $n - p$  a ďalej vzhľadom k tomu, že  $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$  a zároveň  $\mathbf{P} (\mathbf{I}_n - \mathbf{P}) = \mathbf{0}$ , sú náhodné veličiny  $\boldsymbol{\varepsilon}' \mathbf{P} \boldsymbol{\varepsilon}$  a  $\boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{P}) \boldsymbol{\varepsilon}$  nezávislé

a majú  $\chi^2$  rozdelenie s  $p$ , resp.  $n - p$  stupňami voľnosti. Za platnosti nulovej hypotézy má teda štatistika

$$F_{\mathbf{P}} = \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' \mathbf{P} [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] / p}{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' (\mathbf{I}_n - \mathbf{P}) [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] / (n - p)}$$

F rozdelenie s  $p$  a  $n - p$  stupňami voľnosti. Hypotézu  $H_0$  môžeme zamietnuť, ak je hodnota štatistiky  $F_{\mathbf{P}}$  väčšia než príslušný kvantil F rozdelenia. Je zjavné, že štatistika  $F_4$  je špeciálnym prípadom štatistiky  $F_{\mathbf{P}}$  pre

$$\mathbf{P} = \mathbf{H}_{\mathbf{F}}(\boldsymbol{\theta}_0) = \mathbf{F}_{\bullet}(\boldsymbol{\theta}_0) [\mathbf{F}_{\bullet}(\boldsymbol{\theta}_0)' \mathbf{F}_{\bullet}(\boldsymbol{\theta}_0)]^{-1} \mathbf{F}_{\bullet}(\boldsymbol{\theta}_0)'$$

Jednou z ďalších možností je voľba takej matice  $\mathbf{P}$ , ktorá nezávisí na vektore parametrov  $\boldsymbol{\theta}^*$  a zároveň, pre ktorú približne platí  $(\mathbf{I}_n - \mathbf{P}) \mathbf{f}(\boldsymbol{\theta}^*) \approx \mathbf{0}$  tak, aby menovateľ štatistiky pokiaľ možno najmenej závisel na vektore parametrov. V takomto prípade totiž jednoduchou úpravou dostaneme

$$F_{\mathbf{P}} \approx \frac{[\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)]' [\mathbf{Y} - \mathbf{f}(\boldsymbol{\theta}_0)] - \mathbf{Y}' (\mathbf{I}_n - \mathbf{P}) \mathbf{Y}}{\mathbf{Y}' (\mathbf{I}_n - \mathbf{P}) \mathbf{Y}} \frac{n - p}{p},$$

tzn. veľkosť štatistiky bude závisieť hlavne na reziduálnom súčte štvorcov  $\mathcal{S}(\boldsymbol{\theta}_0)$ . Jednu z metód konštrukcie takejto matice na základe Lagrangeovej interpolácie uvádza Hartley [13].

Nespornou výhodou Hartleyho metódy je, že poskytuje presný test hypotézy  $H_0$ , a teda umožňuje konštrukciu presnej konfidenčnej množiny pre neznámy vektor parametrov  $\boldsymbol{\theta}^*$ . Rozdelenie štatistiky  $F_{\mathbf{P}}$  navyše nezávisí na splnení predpokladov regularity. Problémom však je, že sila testu založenom na  $F_{\mathbf{P}}$  výrazne závisí na voľbe matice  $\mathbf{P}$ . Presný test s malou silou je totiž z praktického hľadiska menej užitočný, než približný test s výrazne vyššou silou. Ďalšou mierou nevýhodou je i skutočnosť, že tento postup nie je možné zobecniť pre test hypotézy  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$ , kde vektor  $\boldsymbol{\theta}_1^*$  opäť vyjadruje podmnožinu vektorového parametra  $\boldsymbol{\theta}^*$ .

**Poznámka 2.9.** Za účelom testovania hypotézy  $H_0: \boldsymbol{\theta}_1^* = \boldsymbol{\theta}_{10}$ , kde  $\boldsymbol{\theta}_1^*$  je vektor zvolenej podmnožiny parametrov dĺžky  $p_1 < p$ , navrhol Hamilton [11] analógiu štatistiky  $F_4$  v tvare

$$\tilde{F}_4(\boldsymbol{\theta}_{10}) = \frac{[\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]' [\tilde{\mathbf{H}}_{\mathbf{F}}(\boldsymbol{\theta}_{10}) - \tilde{\mathbf{H}}_{\mathbf{F}_2}(\boldsymbol{\theta}_{10})] [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))] / p_1}{[\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]' [\mathbf{I}_n - \tilde{\mathbf{H}}_{\mathbf{F}}(\boldsymbol{\theta}_{10})] [\mathbf{Y} - \mathbf{f}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))] / (n - p)},$$

kde  $\tilde{\mathbf{H}}_{\mathbf{F}_2}(\boldsymbol{\theta}_{10}) = \mathbf{F}_{2\bullet}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10})) [\mathbf{F}_{2\bullet}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))' \mathbf{F}_{2\bullet}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))]^{-1} \mathbf{F}_{2\bullet}(\tilde{\boldsymbol{\theta}}_n(\boldsymbol{\theta}_{10}))'$ , pričom  $\mathbf{F}_{1\bullet}(\boldsymbol{\theta})$  značí maticu prvých parciálnych derivácií regresnej funkcie podľa  $\boldsymbol{\theta}_2$  typu  $n \times p_2$ . Štatistika  $\tilde{F}_4$  má však iba približné F rozdelenie s  $p_1$  a  $n - p$  stupňami voľnosti. Približná konfidenčná množina založená na štatistike  $\tilde{F}_4$  má teda tvar

$$K_H(\boldsymbol{\theta}_1^*) = \left\{ \boldsymbol{\theta}_1 \in \Omega_1 : \tilde{F}_4(\boldsymbol{\theta}_1) < F_{p_1, n-p}^{1-\alpha} \right\}.$$

## 2.6 Konfidenčné množiny v R

Na záver nadviažeme na príklad 1.2 a jednotlivé možnosti konštrukcie konfidenčných množín porovnáme s využitím programovacieho jazyka R. Vzhľadom k tomu, že štandardné R-kové knižnice obsahujú iba základné metódy konštrukcie konfidenčných množín, využijeme i vlastné R-kové funkcie, ktoré záujemca nájde na priloženom CD.

**Príklad 2.1.** Uvažujme opäť logistický rastový model (1.21) odhadnutý v príklade 1.2. Ako prvé nás budú zaujímať individuálne intervaly spoľahlivosti pre parametre  $\alpha_1$ ,  $\alpha_2$  a  $\alpha_3$ . Waldovské intervaly spoľahlivosti  $CI_W(\alpha_j)$  pre  $j = 1, 2, 3$  môžeme jednoducho určiť s využitím výstupu funkcie `nls`. Príslušné výpočty dosiahneme napríklad pomocou funkcie `confint.Wald` (inou možnosťou je i funkcia `confint2` knižnice `nlrwr`).

```
> source("NLM-funkcie.R")
> confint.Wald(model_A)
           2.5 %      97.5 %
a1  20.6039095  22.4138979
a2   3.3867374   4.5278961
a3   0.5250053   0.7193645
```

Ako sme si už uviedli, platnosť waldovských intervalov spoľahlivosti závisí hlavne na platnosti lineárnej aproximácie (1.6) a môže teda byť negatívne ovplyvnená vnútornou aj parametrickou krivosťou modelu. Okrem už uvedených mier krivosti či odhadu vychýlenia odhadov parametrov získame užitočnú informáciu o vplyve nelinearity modelu aj na základe tzv. profilových diagramov (viď obr. 2.1), ktoré znázorňujú nezáporné hodnoty štatistiky

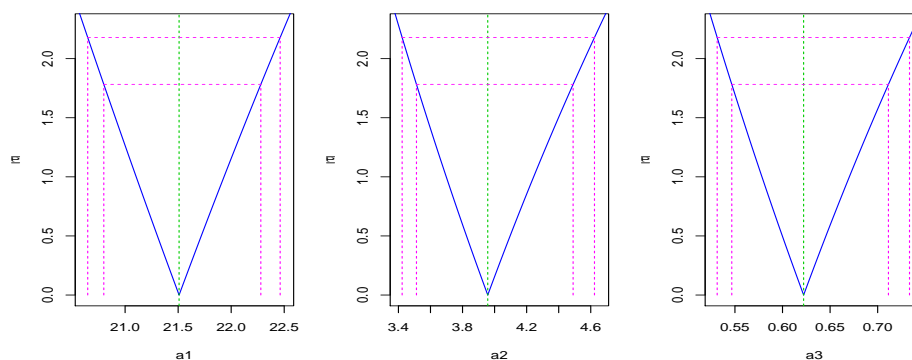
$$|\tau_j(\theta_j)| = \sqrt{\frac{\mathcal{S}(\tilde{\theta}_n(\theta_j)) - \mathcal{S}(\hat{\theta}_n)}{S^2}}$$

pre rôzne hodnoty parametra  $\theta_j$ . Profilové diagramy znázorníme pomocou R-kovej funkcie `plot.profile.nls`.

```
> par(mfrow = c(1, 3), bty = "o")
> plot(profile(model_A), conf = c(0.90, 0.95))
```

S ohľadom na tvar štatistiky  $\tilde{F}_2$  je zrejmé, že body  $\theta_j$ , pre ktoré nadobúda štatistika  $|\tau_j(\theta_j)|$  hodnoty menšie než kvantil  $t_{n-p}^{1-\alpha/2}$ , budú tvoriť vierohodnostný interval spoľahlivosti pre príslušné parametre modelu. Z profilového diagramu môžeme vidieť, že vierohodnostné intervaly spoľahlivosti budú približne symetrické okolo bodového odhadu, čo takisto poukazuje na malý stupeň krivosti tejto parametrizácie modelu.

Príslušné vierohodnostné intervaly spoľahlivosti môžeme jednoducho vypočítať napríklad pomocou štandardnej R-kovej funkcie `confint`. Vytvorená funkcia `confint.LR` je taktiež založená na profilovej logaritmickej vierohodnostnej funkcii, na rozdiel od funkcie `confint` však priamočiaro umožňuje detailnejšie profilovanie vhodným nastavením parametra `delta.t` funkcie `profile`, ako aj využíva na pohľad presnejšiu metódu interpolácie bodov získaných funkciou `profile`, ktorá je prebraná práve z funkcie `plot.profile.nls`.



Obr. 2.1: Profilové diagramy pre model (1.21)

```
> confint.LR(model_A, delta.t = 0.2)
      2.5 %    97.5 %
a1  20.647585 22.461760
a2   3.423286  4.623209
a3   0.531361  0.733467
```

Miernym nedostatkom funkcií `confint` a `confint.LR` sa môže zdať, že sa v oboch prípadoch jedná iba o aproximáciu intervalov spoľahlivosti  $CI_{LR}(\alpha_j)$  pre  $j = 1, 2, 3$ , ktorá je založená na interpolácii tzv. profilových stôp, presnosť tejto aproximácie je však veľmi dobrá. Ukážeme, že aproximácia vierohodnostného intervalového odhadu pre parameter  $\alpha_1$  (platí to ale aj pre ostatné parametre modelu) na základe funkcie `confint.LR` platí v podstate presne.

```
> S = summary(model_A)$sigma
> df = summary(model_A)$df
> a1 = confint.LR(model_A) ["a1", 1]
> model_A1 <- nls(y ~ a1 / (1 + exp(a2 - a3 * x)), data = Ratkowsky4,
+               start = c(a2 = a2_init, a3 = a3_init))
> pf((deviance(model_A1) - deviance(model_A)) / S^2, 1, df[2])
[1] 0.95
> a1 = confint.LR(model_A) ["a1", 2]
> model_A1 <- nls(y ~ a1 / (1 + exp(a2 - a3 * x)), data = Ratkowsky4,
+               start = c(a2 = a2_init, a3 = a3_init))
> pf((deviance(model_A1) - deviance(model_A)) / S^2, 1, df[2])
[1] 0.95
```

V rámci štandardných R-kových knižníc žiaľ nie sú všeobecne známe žiadne funkcie, ktoré by určili ostatné typy intervalov spoľahlivosti či konfidenčných množín. Na priloženom CD je však možné presné výpočty jednotlivých konfidenčných množín pre uvažovaný model nájsť. Vzhľadom k dĺžke príslušných R-kových skriptov uvedieme iba príslušné hodnoty intervalových odhadov s koeficientom spoľahlivosti 0,95. Porovnaním jednotlivých intervalov si je možné všimnúť, že pre tento model sú približne vierohodnostné, skórové i Hamiltonove intervaly spoľahlivosti navzájom podobné a takisto nie sú na rozdiel od waldovských intervalov symetrické okolo príslušného bodového odhadu. Na druhej strane všetky asymptotické intervaly sú

užšie než ich približné verzie, a teda je možné očakávať, že v prípade malého počtu pozorovaní budú mať asymptotické intervaly pravdepodobnosť pokrytia skutočného parametra nižšiu než je nominálny koeficient spoľahlivosti.

```
# približné intervaly spoľahlivosti
> print(CI_LR)
      2.5 %      97.5 %
a1 20.647585 22.461760
a2  3.423286  4.623209
a3  0.531361  0.733467
> print(CI_LM)
      2.5 %      97.5 %
a1 20.651488 22.457556
a2  3.406613  4.660794
a3  0.529531  0.738348
> print(CI_Hamilton)
      2.5 %      97.5 %
a1 20.653008 22.455900
a2  3.399573  4.677781
a3  0.528776  0.740512

# asymptotické intervaly spoľahlivosti
> confint.Wald.asym(model_A)
      2.5 %      97.5 %
a1 20.780757 22.237050
a2  3.498236  4.416398
a3  0.543995  0.700374
> print(CI_LR_asym)
      2.5 %      97.5 %
a1 20.764816 22.320154
a2  3.492145  4.519431
a3  0.543189  0.716219
> print(CI_LM_asym)
      2.5 %      97.5 %
a1 20.708168 22.388687
a2  3.434223  4.620180
a3  0.534593  0.731193
```

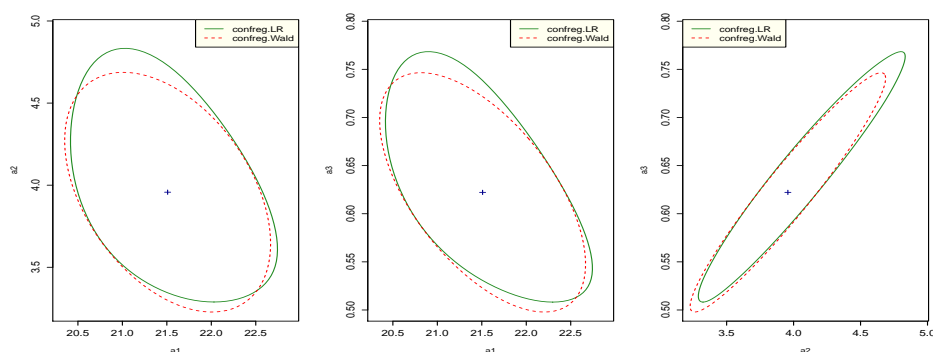
Pri interpretácii individuálnych intervalov spoľahlivosti pre parametre nelineárneho modelu je potreba byť obozretný v tom zmysle, že jednotlivé odhady parametrov môžu byť často výrazne korelované, a teda nemožno vnímať tieto intervaly spoľahlivosti samostatne. Napríklad v prípade, že by sme nahradili hodnotu jediného parametra nejakou hodnotou z príslušného intervalu spoľahlivosti (hodnoty zvyšných parametrov by sme teda ponechali na hodnotách získaných metódou nelineárnych najmenších štvorcov), môžeme veľmi ľahko dostať odhad regresnej funkcie, ktorý bude iba málo odpovedať pôvodným dátam. Preto je často vhodné nejakým spôsobom zobrazíť aj konfidenčné množiny, prinajmenšom pre jednotlivé dvojice parametrov.

V prípade, že je počet parametrov v nelineárnom modeli väčší než dva, môžeme pre každú dvojicu parametrov (teda pre podmnožinu parametrov dĺžky  $p_1 = 2$ ) zostrojiť tzv. *párové* konfidenčné množiny, pri ktorých konštrukcii budeme považovať zvyšné parametre modelu za rušivé. V programovom prostredí R možno znázorniť waldovskú párovú konfidenčnú množinu  $\tilde{K}_W(\alpha_i, \alpha_j)$  pre ľubovoľnú dvojicu parametrov pomocou funkcie `ellipse.nls` knižnice `ellipse`. Na druhej strane, vierohodnostnú párovú konfidenčnú množinu  $\tilde{K}_{LR}(\alpha_i, \alpha_j)$  znázorníme pomocou funkcie `ellipse.profile.nls`, ktorá je založená na tzv. štvorbodovej aproximácii, ktorú uvádzajú Bates a Watts [4]. Je však potreba upozorniť na skutočnosť, že táto štvorbodová aproximácia môže zlyhávať v prípade vysokej krivosti modelu.

```
# párové konfidenčné množiny
> library(ellipse)
> par(mfrow = c(1, 3), bty = "o")
> plot(ellipse(profile(model_A, alphamax = 0.05, delta.t = 0.11),
+           which = c(1, 2)), type = 'l', lty = 1, col = "forestgreen")
> lines(ellipse(model_A, which = c(1, 2)), col = "red", lty = 2)
> points(coef(model_A)[1], coef(model_A)[2], col="darkblue", pch = 3)

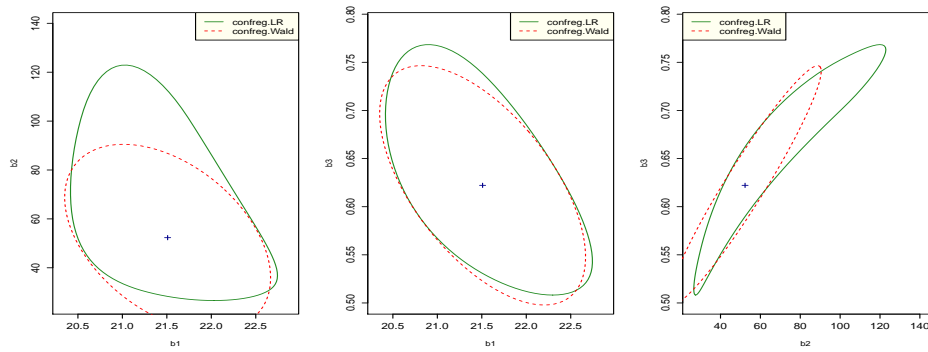
> plot(ellipse(profile(model_A, alphamax = 0.05, delta.t = 0.11),
+           which = c(1, 3)), type = 'l', lty = 1, col = "forestgreen")
> lines(ellipse(model_A, which = c(1, 3)), col = "red", lty = 2)
> points(coef(model_A)[1], coef(model_A)[3], col="darkblue", pch = 3)

> plot(ellipse(profile(model_A, alphamax = 0.05, delta.t = 0.11),
+           which = c(2, 3)), type = 'l', lty = 1, col = "forestgreen")
> lines(ellipse(model_A, which = c(2, 3)), col = "red", lty = 2)
> points(coef(model_A)[2], coef(model_A)[3], col="darkblue", pch = 3)
```



Obr. 2.2: Párové konfidenčné množiny pre model (1.21)

S cieľom porovnať vplyv parametrickej krivosti modelu si uvedieme párovú waldovskú konfidenčnú množinu a aproximáciu párovej vierohodnostnej konfidenčnej množiny aj pre parametrizáciu (1.22). Očakávane, pri väčšej parametrickej krivosti je tvar waldovskej a vierohodnostnej konfidenčnej množiny odlišnejší.



Obr. 2.3: Párové konfidenčné množiny pre model (1.22)

Možnosti zobrazenia celej konfidenčnej množiny v prípade viac než dvojparametrického modelu sú z praktického hľadiska obmedzené. V prípade modelu s tromi parametrami sa síce naskytá možnosť vyobrazenia trojrozmernej konfidenčnej množiny napríklad pomocou funkcie `contour3d` R-kovej knižnice `misc3d`, pri väčšom počte parametrov však zobrazenie celej konfidenčnej množiny nie je možné. Funkcia `nlsContourRSS` knižnice `nlstools` nám ale umožňuje pre každú dvojicu parametrov znázorniť rez viacrozmerného telesa, ktoré je určené vierohodnostnou konfidenčnou množinou pre celý vektor parametrov modelu, a to rovinou, ktorá je daná súradnicami príslušnej dvojice parametrov a ktorá zároveň prechádza bodovým odhadom získaným metódou nelineárnych najmenších štvorcov.

Naopak, funkcia `nlsConfRegions` takisto knižnice `nlstools` nám umožňuje zobraziť pre každú dvojicu parametrov stanovený počet bodov, ktoré tvoria náhodný výber z 95% vierohodnostnej konfidenčnej množiny, podľa kritéria, ktoré navrhol Beale [5]. V tomto prípade sa teda jedná o formu znázornenia kolmého priemetu telesa určeného vierohodnostnou konfidenčnou množinou na rovinu, ktorá odpovedá súradniciam príslušnej dvojice parametrov.

△

## 3. Simulačná štúdia

Na záver tejto práce si uvedieme výsledky simulačnej štúdie, ktorej cieľom je preskúmať vplyv krivosti modelu, ako aj počtu pozorovaní na pravdepodobnosť pokrytia skutočnej hodnoty parametrov jednotlivými konfidenčnými množinami popísanými v predošlej kapitole.

### 3.1 Návrh simulačnej štúdie

Simulačná štúdia založená na metóde Monte Carlo bola navrhnutá tak, aby bolo možné empiricky overiť, ako často je skutočný parameter nelineárneho modelu obsiahnutý v jednotlivých konfidenčných množinách, prípadne intervaloch spoľahlivosti, pričom budeme uvažovať koeficient spoľahlivosti 0,95. Budeme teda predpokladať platnosť nelineárneho modelu v tvare (1.1), ako aj znalosť skutočnej hodnoty  $p$ -rozmerného vektorového parametra  $\theta^*$ , a simulované hodnoty náhodných veličín  $Y_1, \dots, Y_n$  získame generovaním pseudonáhodných hodnôt nezávislých a rovnako rozdelených náhodných zložiek modelu  $\varepsilon_1, \dots, \varepsilon_n$ , ktoré pochádzajú z rozdelenia  $N(0, \sigma^2)$ . Hodnoty z normálneho rozdelenia budeme generovať pomocou R-kovej funkcie `rnorm`.

Pre účely tejto simulačnej štúdie budeme opäť vychádzať z logistického rastového modelu, ktorý sme uvažovali v príklade 1.2, pričom za skutočné hodnoty vektorového parametra budeme považovať uvedené hodnoty získané metódou nelineárnych najmenších štvorcov (samozrejme je možné zvoliť ľubovoľné hodnoty z parametrického priestoru). S cieľom porovnať vplyv parametrickej krivosti modelu na pravdepodobnosť pokrytia skutočnej hodnoty parametra jednotlivými konfidenčnými množinami budeme uvažovať štyri rôzne parametrizácie logistickej krivky – okrem parametrizácií (1.21), (1.22) a (1.23) budeme uvažovať i parametrizáciu

$$f(x; \delta_1, \delta_2, \delta_3) = \frac{\delta_1}{1 + e^{(\delta_2 - x)/\delta_3}}, \quad \delta_1 > 0, \delta_3 > 0, \quad (3.1)$$

ktorú používa R-ková funkcia `SSlogis` (túto parametrizáciu budeme označovať ako model D). Ukážeme, že práve táto parametrizácia máva najmenšiu parametrickú krivosť spomedzi všetkých štyroch uvažovaných možností. Porovnaním príslušných parametrov v modeloch (1.21) a (3.1) sa môžeme jednoducho presvedčiť, že platí

$$\delta_1 = \alpha_1, \quad \delta_2 = \frac{\alpha_2}{\alpha_3}, \quad \delta_3 = \frac{1}{\alpha_3}.$$

Vzhľadom k tomu, že by sme chceli na základe simulačnej štúdie vyhodnotiť aj vplyv počtu pozorovaní, budeme okrem pôvodného počtu pozorovaní  $n = 15$  uvažovať i simulované dáta rozsahu 30, 60 a 120. Zvýšenie počtu pozorovaní

dosiahneme buď replikáciami hodnôt vysvetľujúcej premennej v pôvodnom dátovom súbore alebo rovnomerným zahustením siete hodnôt vysvetľovanej premennej. Pre každú z parametrizácií (1.21), (1.22), (1.23) a (3.1) a pre každý rozsah výberu (vrátane odlišenia spôsobu zvýšenia rozsahu výberu) spracujeme 10 000 simulácií, pričom pre jednotlivé parametrizácie budú simulované dáta daného rozsahu vždy rovnaké.

Následne pre každý simulovaný výber odhadneme parametre príslušného nelineárneho modelu Gaussovým-Newtonovým algoritmom s využitím R-kovej funkcie `nls` (ako počiatočnú aproximáciu pre účely iteračného algoritmu použijeme skutočné hodnoty parametrov). S využitím odhadnutého modelu napokon rozhodneme, či skutočná hodnota vektorového parametra naozaj patrí do príslušnej konfidenčnej množiny pre celý vektor parametrov (budeme pritom uvažovať všetky uvedené spôsoby konštrukcie približnej konfidenčnej množiny, t.j.  $K_W(\boldsymbol{\theta}^*)$ ,  $K_{LR}(\boldsymbol{\theta}^*)$ ,  $K_{LM}(\boldsymbol{\theta}^*)$ , ďalej všetky asymptotické konfidenčné množiny  $K_W^{as.}(\boldsymbol{\theta}^*)$ ,  $K_{LR}^{as.}(\boldsymbol{\theta}^*)$ ,  $K_{LM}^{as.}(\boldsymbol{\theta}^*)$ , ako aj presnú konfidenčnú množinu  $K_F(\boldsymbol{\theta}^*)$ ). Navyše, pre každý z parametrov daného modelu rozhodneme, či skutočná hodnota parametra patrí do príslušného intervalu spoľahlivosti (opäť budeme uvažovať všetky približné a asymptotické intervaly spoľahlivosti vrátane Hamiltonovho intervalu určeného na základe štatistiky  $\tilde{F}_4(\theta_j)$ ). Napokon odhadneme pravdepodobnosť pokrytia skutočného parametra príslušnou konfidenčnou množinou, resp. intervalom spoľahlivosti, relatívnu početnosť výskytu skutočného parametra v danej konfidenčnej množine, resp. v danom intervale spoľahlivosti. Na základe presného prístupu k jednovýberovému testu o parametre alternatívneho rozdelenia založenému na binomickom rozdelení (viď Clopper a Pearson [7]) sa môžeme presvedčiť, že pri danom rozsahu simulačnej štúdie ( $N = 10\,000$ ) nezamietneme na hladine 0,05 hypotézu, že pravdepodobnosť pokrytia skutočného parametra príslušnou konfidenčnou množinou sa rovná nominálnemu koeficientu spoľahlivosti, ak napozorovaná relatívna početnosť výskytu skutočného parametra v danej konfidenčnej množine bude medzi hodnotami 0,9457 a 0,9542 vrátane.

## 3.2 Výsledky simulačnej štúdie

### 3.2.1 Logistický rastový model bez replikácií

Ako prvé si uvedieme výsledky simulačnej štúdie pre jednotlivé parametrizácie logistického rastového modelu s rozsahmi výberu 15, 30, 60 a 120 bez replikácií. Jednotlivé odhady pravdepodobnosti pokrytia príslušnými intervalmi spoľahlivosti, resp. konfidenčnými množinami sú uvedené v tabuľke 3.2, resp. 3.4. Hodnoty, ktoré sú menšie ako 0,9457 alebo väčšie ako 0,9542 označujeme hviezdičkou. Informácia o odhade vychýlenia odhadov jednotlivých parametrov a o mierach krivosti sú uvedené v tabuľkách 3.1 a 3.3. Hviezdičkou opäť označujeme hodnoty, ktoré poukazujú na významnú nelinearitu modelu na základe kritérií uvedených v kapitolách 1.2.5 a 1.3. Naopak skutočné hodnoty parametrov budeme za účelom zjednodušeného zápisu značiť  $\theta_j$ .

Tab. 3.1: Priemerné hodnoty odhadu vychýlenia odhadu parametrov

$\theta_j$ model	$\theta_1$ A až D	A	$\theta_2$ B a C	D	A a B	$\theta_3$ C	D
$n = 15$	0,072 %	0,440 %	5,266 %*	0,057 %	0,445 %	-0,178 %	0,071 %
$n = 30$	0,033 %	0,216 %	2,587 %*	0,026 %	0,219 %	-0,087 %	0,035 %
$n = 60$	0,016 %	0,107 %	1,285 %*	0,013 %	0,109 %	-0,043 %	0,017 %
$n = 120$	0,008 %	0,053 %	0,639 %	0,006 %	0,054 %	-0,022 %	0,009 %

Tab. 3.2: Odhad pravdepodobnosti pokrytia intervalmi spoľahlivosti

$\theta_j$ model	$\theta_1$ A až D	A	$\theta_2$ B a C	D	A a B	$\theta_3$ C	D	
$n = 15$	$CI_W(\theta_j)$	0,9505	0,9474	0,9354*	0,9494	0,9489	0,9479	0,9471
	$CI_{LR}(\theta_j)$	0,9494	0,9476	0,9476	0,9492	0,9482	0,9482	0,9482
	$CI_{LM}(\theta_j)$	0,9495	0,9480	0,9480	0,9496	0,9480	0,9480	0,9480
	$CI_H(\theta_j)$	0,9498	0,9482	0,9482	0,9498	0,9479	0,9479	0,9479
	$CI_W^{as.}(\theta_j)$	0,8965*	0,8901*	0,8923*	0,8980*	0,8929*	0,8908*	0,8906*
	$CI_{LR}^{as.}(\theta_j)$	0,9145*	0,9100*	0,9100*	0,9148*	0,9113*	0,9113*	0,9113*
	$CI_{LM}^{as.}(\theta_j)$	0,9352*	0,9337*	0,9337*	0,9353*	0,9335*	0,9335*	0,9335*
$n = 30$	$CI_W(\theta_j)$	0,9527	0,9519	0,9439*	0,9565*	0,9520	0,9513	0,9518
	$CI_{LR}(\theta_j)$	0,9518	0,9532	0,9532	0,9566*	0,9526	0,9526	0,9526
	$CI_{LM}(\theta_j)$	0,9519	0,9538	0,9538	0,9565*	0,9526	0,9526	0,9526
	$CI_H(\theta_j)$	0,9519	0,9537	0,9537	0,9569*	0,9524	0,9524	0,9524
	$CI_W^{as.}(\theta_j)$	0,9293*	0,9273*	0,9244*	0,9315*	0,9302*	0,9292*	0,9282*
	$CI_{LR}^{as.}(\theta_j)$	0,9373*	0,9381*	0,9381*	0,9398*	0,9372*	0,9372*	0,9372*
	$CI_{LM}^{as.}(\theta_j)$	0,9456*	0,9464	0,9464	0,9494	0,9469	0,9469	0,9469
$n = 60$	$CI_W(\theta_j)$	0,9457	0,9509	0,9506	0,9484	0,9498	0,9493	0,9494
	$CI_{LR}(\theta_j)$	0,9460	0,9505	0,9505	0,9484	0,9492	0,9492	0,9492
	$CI_{LM}(\theta_j)$	0,9468	0,9507	0,9507	0,9485	0,9496	0,9496	0,9496
	$CI_H(\theta_j)$	0,9470	0,9505	0,9505	0,9485	0,9496	0,9496	0,9496
	$CI_W^{as.}(\theta_j)$	0,9359*	0,9414*	0,9382*	0,9371*	0,9383*	0,9380*	0,9386*
	$CI_{LR}^{as.}(\theta_j)$	0,9396*	0,9431*	0,9431*	0,9428*	0,9427*	0,9427*	0,9427*
	$CI_{LM}^{as.}(\theta_j)$	0,9434*	0,9463	0,9463	0,9451*	0,9466	0,9466	0,9466
$n = 120$	$CI_W(\theta_j)$	0,9555*	0,9510	0,9473	0,9554*	0,9512	0,9503	0,9511
	$CI_{LR}(\theta_j)$	0,9540	0,9517	0,9517	0,9555*	0,9502	0,9502	0,9502
	$CI_{LM}(\theta_j)$	0,9542	0,9519	0,9519	0,9552*	0,9503	0,9503	0,9503
	$CI_H(\theta_j)$	0,9542	0,9519	0,9519	0,9552*	0,9502	0,9502	0,9502
	$CI_W^{as.}(\theta_j)$	0,9494	0,9446*	0,9429*	0,9500	0,9452*	0,9448*	0,9452*
	$CI_{LR}^{as.}(\theta_j)$	0,9514	0,9485	0,9485	0,9511	0,9465	0,9465	0,9465
	$CI_{LM}^{as.}(\theta_j)$	0,9528	0,9503	0,9503	0,9541	0,9486	0,9486	0,9487

Pre začiatok si je vhodné všimnúť, že ak v rôznych parametrizáciách toho istého modelu sa nejaký parameter nemení (napríklad platí  $\alpha_1 = \beta_1 = \gamma_1 = \delta_1$ ) pravdepodobnosť pokrytia tohto parametra v rámci príslušných intervalov spoľahlivosti sa takisto pri zmene parametrizácie nemení. To okrem iného znamená, že v prípade odhadu vplyvu parametrickej krivosti na validitu waldovského intervalu spoľahlivosti je vhodnejšie sa orientovať podľa veľkosti odhadu vychýlenia odhadu príslušného parametra než podľa celkovej miery parametrickej krivosti modelu. Môžeme si všimnúť, že významná hodnota odhadu

vychýlenia u daného parametra predznamenáva, že pravdepodobnosť pokrytia príslušného parametra waldovským intervalom bude pri menších rozsahoch výberu nižšia než nominálny koeficient spoľahlivosti.

Pokiaľ ide o vplyv rozsahu výberu, tak môžeme pozorovať, že s rastúcim počtom pozorovaní klesá priemerná hodnota odhadu vychýlenia odhadov parametrov a takisto klesá i parametrická (ale aj vnútorná) krivosť modelu. To znamená, že s rastúcim rozsahom výberu vplyv krivosti modelu na waldovské množiny skutočne klesá, a teda v prípade veľkého rozsahu je možné využiť i túto jednoduchú metódu konštrukcie intervalu spoľahlivosti.

Čo sa týka iných približných intervalov spoľahlivosti, tak je vhodné poznamenať, že ak nejaký parameter nahradíme v inej parametrizácii iba nejakou jeho funkciou, tak sa pravdepodobnosť jeho pokrytia vierohodnostným, skórovým alebo Hamiltonovým intervalom spoľahlivosti nemení (to samozrejme neplatí napríklad pre  $\delta_2 = \alpha_2/\alpha_3$ ). Tento záver je samozrejme v súlade s tým, že parametrická krivosť nemá na validitu týchto intervalov spoľahlivosti žiaden vplyv, čo je v porovnaní s waldovským prístupom značná výhoda. Vidíme, že práve tieto tri metódy konštrukcie intervalu spoľahlivosti dávajú najlepšie výsledky, pričom odhad pravdepodobnosti pokrytia týmito intervalmi odpovedá nominálnemu koeficientu spoľahlivosti.

Naopak pravdepodobnosť pokrytia asymptotickými intervalmi spoľahlivosti je i v prípade stredného rozsahu výberu menšia než nominálny koeficient spoľahlivosti, a tak v praxi nemožno použitie týchto intervalov spoľahlivosti odporučiť. Najlepšie výsledky však dáva asymptotický interval založený na skórovom teste.

Pokiaľ ide o pravdepodobnosť pokrytia skutočnej hodnoty celého vektorového parametra waldovskou konfidenčnou množinou, tak je negatívny vplyv parametrickej krivosti modelu na túto pravdepodobnosť ešte výraznejší, a dokonca aj pri rozsahu výberu  $n = 120$  je odhad tejto pravdepodobnosti v modeloch s vyššou parametrickou krivosťou stále mierne nižší než nominálny koeficient spoľahlivosti. V prípade konštrukcie konfidenčných množín či testovania hypotéz o celom vektore parametrov je teda vhodnejšie využiť približných metód, ktoré na parametrickej krivosti nezávisia. Vzhľadom k tomu, že odhady pravdepodobností pokrytia vierohodnostnou, skórovou i presnou konfidenčnou množinou sú veľmi podobné, je možné v praxi odporučiť ľubovoľnú z nich. Vzhľadom k mierne jednoduchšej konštrukcii, ako aj intuitívnejšej interpretácii založenej na reziduálnom súčte štvorcov je teda možné pre modely s nízkou vnútornou krivosťou preferovať častejšie používané vierohodnostné približné konfidenčné množiny, resp. intervaly spoľahlivosti.

Tab. 3.3: Priemerné hodnoty mier krivosti modelu

model		A	B	C	D
$n = 15$	$\gamma_{RMS}^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,3278*	1,5174*	1,7629*	0,1593
	$\gamma^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,6840*	3,4161*	3,9626*	0,3115
	$\gamma_{RMS}^N(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,1114	0,1114	0,1114	0,1114
	$\gamma^N(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,1965	0,1965	0,1965	0,1965
$n = 30$	$\gamma_{RMS}^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,2083	0,9881*	1,1469*	0,1016
	$\gamma^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,4334	2,2253*	2,5783*	0,1973
$n = 60$	$\gamma_{RMS}^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,1407	0,6764*	0,7848*	0,0688
	$\gamma^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,2925	1,5236*	1,7643*	0,1332
$n = 120$	$\gamma_{RMS}^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,0974	0,4702*	0,5454*	0,0477
	$\gamma^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,2024	1,0591*	1,2261*	0,0921

Tab. 3.4: Odhad pravdepodobnosti pokrytia konfidencnými množinami

model		A	B	C	D
$n = 15$	$K_W(\boldsymbol{\theta})$	0,9468	0,9027*	0,8903*	0,9478
	$K_{LR}(\boldsymbol{\theta})$	0,9501	0,9501	0,9501	0,9501
	$K_{LM}(\boldsymbol{\theta})$	0,9495	0,9495	0,9495	0,9495
	$K_F(\boldsymbol{\theta})$	0,9497	0,9497	0,9497	0,9497
	$K_W^{as.}(\boldsymbol{\theta})$	0,8422*	0,7868*	0,7724*	0,8397*
	$K_{LR}^{as.}(\boldsymbol{\theta})$	0,9099*	0,9099*	0,9099*	0,9099*
	$K_{LM}^{as.}(\boldsymbol{\theta})$	0,9714*	0,9714*	0,9714*	0,9714*
	$K_F^{as.}(\boldsymbol{\theta})$	0,9714*	0,9714*	0,9714*	0,9714*
$n = 30$	$K_W(\boldsymbol{\theta})$	0,9523	0,9197*	0,9106*	0,9526
	$K_{LR}(\boldsymbol{\theta})$	0,9535	0,9535	0,9535	0,9535
	$K_{LM}(\boldsymbol{\theta})$	0,9536	0,9536	0,9536	0,9536
	$K_F(\boldsymbol{\theta})$	0,9539	0,9539	0,9539	0,9539
	$K_W^{as.}(\boldsymbol{\theta})$	0,9093*	0,8738*	0,8635*	0,9076*
	$K_{LR}^{as.}(\boldsymbol{\theta})$	0,9361*	0,9361*	0,9361*	0,9361*
	$K_{LM}^{as.}(\boldsymbol{\theta})$	0,9643*	0,9643*	0,9643*	0,9643*
	$K_F^{as.}(\boldsymbol{\theta})$	0,9643*	0,9643*	0,9643*	0,9643*
$n = 60$	$K_W(\boldsymbol{\theta})$	0,9468	0,9305*	0,9259*	0,9460
	$K_{LR}(\boldsymbol{\theta})$	0,9468	0,9468	0,9468	0,9468
	$K_{LM}(\boldsymbol{\theta})$	0,9473	0,9473	0,9473	0,9473
	$K_F(\boldsymbol{\theta})$	0,9473	0,9473	0,9473	0,9473
	$K_W^{as.}(\boldsymbol{\theta})$	0,9294*	0,9107*	0,9047*	0,9267*
	$K_{LR}^{as.}(\boldsymbol{\theta})$	0,9395*	0,9395*	0,9395*	0,9395*
	$K_{LM}^{as.}(\boldsymbol{\theta})$	0,9515	0,9515	0,9515	0,9515
	$K_F^{as.}(\boldsymbol{\theta})$	0,9515	0,9515	0,9515	0,9515
$n = 120$	$K_W(\boldsymbol{\theta})$	0,9477	0,9413*	0,9378*	0,9493
	$K_{LR}(\boldsymbol{\theta})$	0,9489	0,9489	0,9489	0,9489
	$K_{LM}(\boldsymbol{\theta})$	0,9492	0,9492	0,9492	0,9492
	$K_F(\boldsymbol{\theta})$	0,9491	0,9491	0,9491	0,9491
	$K_W^{as.}(\boldsymbol{\theta})$	0,9387*	0,9297*	0,9274*	0,9391*
	$K_{LR}^{as.}(\boldsymbol{\theta})$	0,9451*	0,9451*	0,9451*	0,9451*
	$K_{LM}^{as.}(\boldsymbol{\theta})$	0,9503	0,9503	0,9503	0,9503
	$K_F^{as.}(\boldsymbol{\theta})$	0,9503	0,9503	0,9503	0,9503

### 3.2.2 Logistický rastový model s replikáciami

V prípade modelu s replikáciami sú odhady pravdepodobností pokrytia intervalmi spoľahlivosti či konfidenčnými množinami veľmi podobné odhadom získaným v prípade zahustenia siete hodnôt vysvetľujúcej premennej. Replikované dáta teda nemajú na použitie príslušných metód konštrukcie konfidenčných množín takmer žiadny vplyv. Závbery uvedené v predošlom odstavci sa teda nemenia.

Tab. 3.5: Priemerné hodnoty odhadu vychýlenia odhadu parametrov

$\theta_j$ model	$\theta_1$ A až D	$\theta_2$			$\theta_3$		
		A	B a C	D	A a B	C	D
$n = 30$	0,036 %	0,219 %	2,605 %*	0,028 %	0,222 %	-0,089 %	0,035 %
$n = 60$	0,018 %	0,110 %	1,298 %*	0,014 %	0,111 %	-0,044 %	0,018 %
$n = 120$	0,009 %	0,055 %	0,646 %	0,007 %	0,055 %	-0,022 %	0,009 %

Tab. 3.6: Odhad pravdepodobnosti pokrytia intervalmi spoľahlivosti

$\theta_j$ model	$\theta_1$ A až D	$\theta_2$			$\theta_3$			
		A	B a C	D	A a B	C	D	
$n = 30$ ( $r = 2$ )	$CI_W(\theta_j)$	0,9513	0,9511	0,9467	0,9572	0,9538	0,9540	0,9525
	$CI_{LR}(\theta_j)$	0,9503	0,9516	0,9516	0,9561	0,9541	0,9541	0,9541
	$CI_{LM}(\theta_j)$	0,9504	0,9518	0,9518	0,9559	0,9538	0,9538	0,9538
	$CI_H(\theta_j)$	0,9502	0,9522	0,9522	0,9562	0,9539	0,9539	0,9539
	$CI_W^{as}(\theta_j)$	0,9274*	0,9295*	0,9253*	0,9333*	0,9306*	0,9304*	0,9321*
	$CI_{LR}^{as}(\theta_j)$	0,9353*	0,9384*	0,9384*	0,9412*	0,9403*	0,9403*	0,9403*
	$CI_{LM}^{as}(\theta_j)$	0,9426*	0,9455*	0,9455*	0,9495	0,9488	0,9488	0,9488
$n = 60$ ( $r = 4$ )	$CI_W(\theta_j)$	0,9465	0,9508	0,9500	0,9483	0,9493	0,9495	0,9489
	$CI_{LR}(\theta_j)$	0,9467	0,9490	0,9490	0,9484	0,9490	0,9490	0,9490
	$CI_{LM}(\theta_j)$	0,9466	0,9485	0,9485	0,9480	0,9497	0,9497	0,9497
	$CI_H(\theta_j)$	0,9467	0,9485	0,9485	0,9481	0,9495	0,9495	0,9495
	$CI_W^{as}(\theta_j)$	0,9364*	0,9396*	0,9402*	0,9380*	0,9394*	0,9382*	0,9385*
	$CI_{LR}^{as}(\theta_j)$	0,9414*	0,9419*	0,9419*	0,9408*	0,9423*	0,9423*	0,9423*
	$CI_{LM}^{as}(\theta_j)$	0,9444*	0,9458	0,9458	0,9456*	0,9457	0,9457	0,9457
$n = 120$ ( $r = 8$ )	$CI_W(\theta_j)$	0,9537	0,9473	0,9483	0,9545*	0,9515	0,9502	0,9501
	$CI_{LR}(\theta_j)$	0,9532	0,9487	0,9487	0,9542	0,9498	0,9498	0,9498
	$CI_{LM}(\theta_j)$	0,9536	0,9489	0,9489	0,9546*	0,9497	0,9497	0,9498
	$CI_H(\theta_j)$	0,9536	0,9489	0,9489	0,9546*	0,9497	0,9497	0,9497
	$CI_W^{as}(\theta_j)$	0,9491	0,9433*	0,9443*	0,9490	0,9456*	0,9450*	0,9444*
	$CI_{LR}^{as}(\theta_j)$	0,9510	0,9461	0,9461	0,9503	0,9470	0,9470	0,9470
	$CI_{LM}^{as}(\theta_j)$	0,9523	0,9479	0,9479	0,9525	0,9485	0,9485	0,9485

Tab. 3.7: Priemerné hodnoty mier parametrickej krivosti modelu

model		A	B	C	D
$n = 30$	$\gamma_{RMS}^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,2153	0,9906*	1,1510*	0,1047
	$\gamma^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,4490	2,2309*	2,5878*	0,2047
$n = 60$	$\gamma_{RMS}^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,1477	0,6784*	0,7883*	0,0718
	$\gamma^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,3081	1,5281*	1,7725*	0,1405
$n = 120$	$\gamma_{RMS}^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,1030	0,4716*	0,5480*	0,0501
	$\gamma^T (\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,2148	1,0623*	1,2322*	0,0980

Tab. 3.8: Odhad pravdepodobnosti pokrytia konfidenčnými množinami

model		A	B	C	D
$n = 30$ ( $r = 2$ )	$K_W(\boldsymbol{\theta})$	0,9518	0,9207*	0,9108*	0,9512
	$K_{LR}(\boldsymbol{\theta})$	0,9513	0,9513	0,9513	0,9513
	$K_{LM}(\boldsymbol{\theta})$	0,9514	0,9514	0,9514	0,9514
	$K_F(\boldsymbol{\theta})$	0,9518	0,9518	0,9518	0,9518
$n = 60$ ( $r = 4$ )	$K_W^{as}(\boldsymbol{\theta})$	0,9069*	0,8758*	0,8659*	0,9081*
	$K_{LR}^{as}(\boldsymbol{\theta})$	0,9360*	0,9360*	0,9360*	0,9360*
	$K_{LM}^{as}(\boldsymbol{\theta})$	0,9627*	0,9627*	0,9627*	0,9627*
	$K_W(\boldsymbol{\theta})$	0,9478	0,9312*	0,9255*	0,9462
$n = 120$ ( $r = 8$ )	$K_{LR}(\boldsymbol{\theta})$	0,9483	0,9483	0,9483	0,9483
	$K_{LM}(\boldsymbol{\theta})$	0,9476	0,9476	0,9476	0,9476
	$K_F(\boldsymbol{\theta})$	0,9476	0,9476	0,9476	0,9476
	$K_W^{as}(\boldsymbol{\theta})$	0,9279*	0,9108*	0,9046*	0,9263*
$n = 30$ ( $r = 2$ )	$K_{LR}^{as}(\boldsymbol{\theta})$	0,9403*	0,9403*	0,9403*	0,9403*
	$K_{LM}^{as}(\boldsymbol{\theta})$	0,9513	0,9513	0,9513	0,9513
	$K_W(\boldsymbol{\theta})$	0,9479	0,9386*	0,9361*	0,9476
	$K_{LR}(\boldsymbol{\theta})$	0,9477	0,9477	0,9477	0,9477
$n = 60$ ( $r = 4$ )	$K_{LM}(\boldsymbol{\theta})$	0,9483	0,9483	0,9483	0,9483
	$K_F(\boldsymbol{\theta})$	0,9483	0,9483	0,9483	0,9483
	$K_W^{as}(\boldsymbol{\theta})$	0,9391*	0,9291*	0,9264*	0,9389*
	$K_{LR}^{as}(\boldsymbol{\theta})$	0,9439*	0,9439*	0,9439*	0,9439*
$n = 120$ ( $r = 8$ )	$K_{LM}^{as}(\boldsymbol{\theta})$	0,9501	0,9501	0,9501	0,9501

**Poznámka 3.1.** Seber a Wild [18] uvádzajú i špeciálnu možnosť konštrukcie konfidenčných množín v prípade replikovaných dát. Táto možnosť je založená na alternatívnom odhade rozptylu  $\sigma^2$  (označme  $S_r^2$ ), ktorý nezávisí na platnosti modelu (1.1) a ktorý je možné vyjadriť ako súčet štvorcov odchýlok jednotlivých pozorovaní vysvetľovanej premennej od príslušného priemeru získaného pri rovnakých hodnotách regresorov podelený celkovým počtom pozorovaní zníženým o počet unikátnych hodnôt, ktoré vektor regresorov v príslušnom výbere nadobúda (označme  $m$ ). Potom je možné napríklad vierohodnostnú konfidenčnú množinu vyjadriť ako

$$\left\{ \boldsymbol{\theta} \in \Omega : \mathcal{S}(\boldsymbol{\theta}) < \mathcal{S}(\hat{\boldsymbol{\theta}}) + pS_r^2 \mathbf{F}_{p,n-m}^{1-\alpha} \right\}.$$

### 3.2.3 Model s vysokou vnútornou krivosťou

Doposiaľ sme uvažovali model, ktorého vnútorná krivosť bola zanedbateľná. Napriek tomu, že táto situácia je v reálnych aplikáciách veľmi častá, uvedieme si na záver tejto práce výsledky simulačnej štúdie modelu, ktorý má vnútornú krivosť významnú. Vytvorený model bol inšpirovaný článkom St. Laurent a Cook [19].

Uvažujme teda model v tvare

$$Y_i = \theta_1 x_{i1} + e^{\theta_2 x_{i2}} + \varepsilon_i,$$

pre ktorý budeme uvažovať skutočné hodnoty parametrov  $\theta_1 = 2$  a  $\theta_2 = 0,15$  a rozptyl nezávislej a normálne rozdelenej náhodnej zložky  $\sigma^2 = 0,6$ . Hodnoty regresorov nadobúdajú hodnoty

$$x_1 = c(-4.0, -3, -2.5, -2.0, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2.0, 2.5, 3, 4.0)$$

$$x_2 = c(-2.5, -2, -1.7, -1.5, -1.2, -1, -0.5, 0, 0.5, 1, 1.2, 1.5, 1.7, 2, 2.5)$$

pričom za účelom zvýšenia počtu pozorovaní využijeme replikácie.

Opäť spracujeme simulačnú štúdiu za účelom odhadu pravdepodobnosti pokrytia príslušných parametrov jednotlivými intervalmi spoľahlivosti, resp. konfidenčnými množinami, pričom tentoraz nebudeme uvažovať asymptotické verzie týchto intervalov, resp. množín. Vzhľadom k veľkej variabilite jednotlivých odhadov vychýlenia či mier krivosti budeme uvádzať v tabuľkách 3.9 a 3.11 namiesto priemerných ich mediánové hodnoty. Poznamenajme, že z dôvodu konvergenčných problémov pri odhade parametrov sme museli približne 0,8 % simulovaných výberov zo simulačnej štúdie vylúčiť.

Môžeme si ihneď všimnúť extrémne vysokú hodnotu odhadu vychýlenia odhadu parametra  $\theta_2$ . Príslušné odhady pravdepodobnosti pokrytia jednotlivých skutočných hodnôt parametrov príslušnými intervalmi spoľahlivosti sú zaznamenané v tabuľke 3.10.

Tab. 3.9: Mediánové hodnoty odhadu vychýlenia odhadu parametrov

rozsah	$n = 15$	$n = 60$	$n = 120$
$\theta_1$	1,311%*	0,794%	0,595%
$\theta_2$	-83,09%*	-35,26%*	-20,16%*

Tab. 3.10: Odhad pravdepodobnosti pokrytia intervalmi spoľahlivosti

rozsah	$n = 15$ ( $r = 1$ )		$n = 60$ ( $r = 4$ )		$n = 120$ ( $r = 8$ )	
$\theta_j$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
$CI_W(\theta_j)$	0,9419*	0,8976*	0,9136*	0,8792*	0,9023*	0,8829*
$CI_{LR}(\theta_j)$	0,9781*	0,9793*	0,9721*	0,9732*	0,9626*	0,9685*
$CI_{LM}(\theta_j)$	0,9641*	0,9665*	0,9479	0,9501	0,9514	0,9516
$CI_H(\theta_j)$	0,9492	0,9537	0,9462	0,9472	0,9507	0,9512

Vidíme, že vysoká hodnota odhadu vychýlenia v prípade odhadu parametra  $\theta_2$  má opäť vplyv na nižšiu pravdepodobnosť pokrytia skutočnej hodnoty tohto

parametra waldovským intervalom spoľahlivosti. To odpovedá negatívne mu vplyvu parametrickej krivosti na waldovské intervaly spoľahlivosti, ktorý sme pozorovali už v prípade logistického rastového modelu.

Zaujímavé ale je, že aj keď parametrická krivosť nemá vplyv na ostatné približné intervaly spoľahlivosti, jedinou metódou konštrukcie, ktorej pravdepodobnosť pokrytia skutočnej hodnoty oboch parametrov odpovedá nominálnemu koeficientu spoľahlivosti je Hamiltonova metóda. Vysoká vnútorná krivosť modelu pri nižšom rozsahu výberu má za následok vyššiu pravdepodobnosť pokrytia skutočnej hodnoty parametrov u vierohodnostného i skórového intervalu spoľahlivosti. Negatívny vplyv sa pritom prejavuje viac u vierohodnostného prístupu. Tento záver sa teda líši od výsledku simulačnej štúdie, ktorú publikovali Donaldsonová a Schnabel [8] a podľa ktorej nemala vnútorná krivosť modelu na pravdepodobnosť pokrytia vierohodnostným intervalom vplyv – je však nutné podotknúť, že aj v uvedenom článku bol uvažovaný jediný model s vysokou mierou vnútornej krivosti, a teda na potvrdenie či vyvrátenie tejto hypotézy by bolo potrebné túto simulačnú štúdiu rozšíriť aj o iné modely s vysokou vnútornou krivosťou.

Analogické odhady pravdepodobnosti pokrytia vektorového parametra jednotlivými konfidenčnými množinami sú uvedené v tabuľke 3.12. Jedinou konfidenčnou množinou, ktorá i pri nižšom rozsahu výberu dodržiava nominálny koeficient spoľahlivosti je presná konfidenčná množina. Jednotlivé metódy konštrukcie konfidenčných množín sú síce asymptoticky ekvivalentné, ich postupné približovanie je však v takomto prípade pomerne pomalé. Z približných konfidenčných množín dosahuje v modeli s vysokou vnútornou krivosťou najlepšie výsledky skórová konfidenčná množina, najhoršie opäť konfidenčná množina waldovského typu.

Tab. 3.11: Mediánové hodnoty mier krivosti modelu

rozsah	$n = 15$	$n = 60$	$n = 120$
$\gamma_{RMS}^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	2,1121*	1,2476*	0,9443*
$\gamma^T(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	3,4491*	2,0373*	1,5421*
$\gamma_{RMS}^N(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	0,7668*	0,4976*	0,3813*
$\gamma^N(\mathbf{F}_{p,n-p}^{1-\alpha})^{1/2}$	1,2521*	0,8125*	0,6227*

Tab. 3.12: Odhad pravdepodobnosti pokrytia konfidenčnými množinami

rozsah	$n = 15$	$n = 60$	$n = 120$
$K_W(\boldsymbol{\theta})$	0,9086*	0,8954*	0,8929*
$K_{LR}(\boldsymbol{\theta})$	0,9686*	0,9627*	0,9604*
$K_{LM}(\boldsymbol{\theta})$	0,9634*	0,9514	0,9538
$K_F(\boldsymbol{\theta})$	0,9519	0,9486	0,9528

# Záver

Cieľom tejto práce bolo popísať a porovnať rôzne možnosti konštrukcie konfidenčných množín, prípadne intervalov spoľahlivosti, pre parametre normálneho nelineárneho modelu. Vzhľadom k skutočnosti, že na rozdiel od lineárneho modelu nie je možné v prípade konečného výberu jednoducho a jednoznačne skonštruovať konfidenčné množiny, ktoré budú presne s danou spoľahlivosťou pokrývať skutočné hodnoty parametrov modelu, boli odvodené viaceré približné konfidenčné množiny, ktoré sú síce asymptoticky ekvivalentné, ale pri malom rozsahu výberu sa môžu navzájom pomerne výrazne líšiť. Okrem približných konfidenčných množín boli na základe teórie maximálnej vierohodnosti odvodené i asymptotické konfidenčné množiny, ktoré však nemusia byť vhodné v prípade menšieho rozsahu výberu. Takisto sme si v krátkosti uviedli i najčastejšie používanú presnú konfidenčnú množinu založenú na Hartleyho metóde.

Na základe výsledkov simulačnej štúdie sme si overili, že v prípade azda najčastejšie používanej waldovskej konfidenčnej množiny je pravdepodobnosť pokrytia skutočného parametra touto konfidenčnou množinou, prípadne intervalmi spoľahlivosti, často štatisticky významne nižšia než je zvolený koeficient spoľahlivosti. Tento fakt je samozrejme v súlade s výsledkami rozsiahlej simulačnej štúdie, ktorú prezentovali Donaldsonová a Schnabel [8] a v ktorej porovnávali waldovskú, vierohodnostnú a presnú konfidenčnú množinu. Obmedzením waldovskej metódy je hlavne eliptický tvar konfidenčnej množiny, ktorý často nemusí odpovedať tvaru žiaducej konfidenčnej množiny pre parametre nelineárneho modelu, čo nastáva hlavne v prípade vyššej parametrickej krivosti modelu. So zvyšujúcim sa rozsahom výberu sa však pravdepodobnosť pokrytia waldovskou konfidenčnou množinou blíži nominálnej hodnote koeficientu spoľahlivosti.

Z uvedených dôvodov by mali byť v reálnych aplikáciách preferované približné alebo presné metódy, ktoré nie sú parametrickou krivosťou modelu ovplyvnené. Simulačná štúdia potvrdila, že ako vierohodnostné, tak i skórové a presné konfidenčné množiny majú pravdepodobnosť pokrytia skutočného parametra veľmi blízku nominálnemu koeficientu spoľahlivosti i v prípade malého rozsahu výberu. Naopak, asymptotické konfidenčné množiny či intervaly spoľahlivosti sa neodporúča používať v prípade malého, ani stredne veľkého rozsahu výberu, kedy môže byť príslušná pravdepodobnosť pokrytia i výrazne nižšia než je zvolený koeficient spoľahlivosti.

V prípade významnej vnútornej krivosti modelu však môžu byť vierohodnostné a čiastočne i skórové konfidenčné množiny či intervaly spoľahlivosti mierne konzervatívne, t.j. môžu mať pravdepodobnosť pokrytia skutočnej hodnoty parametra vyššiu než je zvolený koeficient spoľahlivosti. V prípade snahy o udržanie pravdepodobnosti pokrytia v blízkosti hodnoty príslušného koefici-

entu spoľahlivosti sa tak odporúča použitie presnej konfidenčnej množiny založenej na Hartleyho metóde, resp. približnej Hamiltonovej metódy konštrukcie intervalov spoľahlivosti.

Na záver je však potreba poznamenať, že vnútorná krivosť modelu je v reálnych aplikáciách často zanedbateľná. A tak vzhľadom k tomu, že vierohodnostná, skórová a presná konfidenčná množina, resp. vierohodnostné, skórové a Hamiltonove intervaly spoľahlivosti, dávajú často veľmi podobné výsledky, možno považovať každú z týchto metód za vhodnú pre praktické aplikácie. Vzhľadom k najviac intuitívnej interpretácii založenej na reziduálnom súčte štvorcov, by som osobne odporúčal použitie vierohodnostných konfidenčných množín či intervalov.

# Zoznam literatúry

- [1] AMEMIYA, T. Non-Linear Regression Models. In: GRILICHES, Z., INTRILIGATOR, M.D. *Handbook of Econometrics*. Amsterdam: North-Holland, 1983, s. 333-389
- [2] ANDĚL, J. *Základy matematické statistiky*. Praha: Matfyzpress, 2007
- [3] BATES, D.M., WATTS, D.G. Relative Curvature Measures of Nonlinearity. *Journal of the Royal Statistical Society*, 1980, roč. 42, č. 1, s. 1-25
- [4] BATES, D.M., WATTS, D.G. *Nonlinear regression analysis and its applications*. New York: Wiley, 1988
- [5] BEALE, E.M.L. Confidence Regions in Non-Linear Estimation. *Journal of the Royal Statistical Society*, 1960, roč. 22, č. 1, s. 41-88
- [6] BOX, M.J. Bias in Nonlinear Estimation. *Journal of the Royal Statistical Society*, 1971, roč. 33, č. 2, s. 171-201
- [7] CLOPPER, C.J., PEARSON, E.S. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, 1934, roč. 26, č. 4, s. 404-413
- [8] DONALDSON, J.R., SCHNABEL, R.B. Computational Experience with Confidence Regions and Confidence Intervals for Nonlinear Least Squares. *Technometrics*, 1987, roč. 29, č. 1, s. 67-82
- [9] GALLANT, A.R. *Nonlinear Statistical Models*. New York: Wiley, 1987
- [10] GUTTMAN, I., MEETER, D.A. On Beale's Measures of Non-Linearity. *Technometrics*, 1965, roč. 7, č. 4, s. 623-637
- [11] HAMILTON, D. Confidence regions for parameter subsets in nonlinear regression. *Biometrika*, 1986, roč. 73, č. 1, s. 57-64
- [12] HAMILTON, D., WIENS, D. Correction factors for F ratios in nonlinear regression. *Biometrika*, 1987, roč. 74, č. 2, s. 423-425
- [13] HARTLEY, H.O. Exact confidence regions for the parameters in non-linear regression laws. *Biometrika*, 1964, roč. 51, č. 3/4, s. 347-353
- [14] JENNRICH, R.I. Asymptotic Properties of Non-Linear Least Squares Estimators. *The Annals of Mathematical Statistics*, 1969, roč. 40, č. 2, s. 633-643
- [15] MALINVAUD, E. The Consistency of Nonlinear Regressions. *The Annals of Mathematical Statistics*, 1970, roč. 41, č. 3, s. 956-969

- [16] MALINVAUD, E. *Statistical methods of econometrics*. Amsterdam: North-Holland, 1970
- [17] RATKOWSKY, D.A. *Nonlinear regression modeling: a unified practical approach*. New York: Marcel Dekker, 1983
- [18] SEBER, G.A.F., WILD, C.J. *Nonlinear Regression*. Hoboken: Wiley, 2006
- [19] ST.LAURENT, R.T., COOK R.D. Leverage, Local Influence and Curvature in Nonlinear Regression. *Biometrika*, 1993, roč. 80, č. 1, s. 99-106
- [20] WU, C.-F. The Consistency of Nonlinear Regressions. *The Annals of Statistics*, 1981, roč. 9, č. 3, s. 501-513
- [21] ZVÁRA, K. *Regrese*. Praha: Matfyzpress, 2008

# Obsah priloženého CD

K tejto práci je priložené CD, na ktorom je možné nájsť:

- elektronickú verziu diplomovej práce vo formáte PDF,
- pomocné funkcie v programovacom jazyku R,
- zdrojové kódy k jednotlivým príkladom, ako i simulačnej štúdiu.

Všetky výpočty boli spracované v programovom prostredí R, verzia 2.12.2.