MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Valeriya Plotnikova

Kombinování odhadů

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	RNDr. Jiří Dvořák, Ph.D.
Studijní program:	Matematika
Studijní obor:	Finanční matematika

Praha 2019

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

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

V .... dne .... Podpis autora

Na tomto místě bych ráda poděkovala vedoucímu RNDr. Jiřímu Dvořákovi, Ph.D. za odborné vedení bakalářské práce, cenné rady, trpělivost a čas věnovaný konzultačním hodinám. Mé poděkování patří také rodičům a bratrovi, kteří mi umožnili studium a vždy mě plně podporovali. V neposlední řadě děkuji svému příteli a spolubydlící za neustálou podporu během celého studia a za pomoc při gramatické kontrole práce.

Název práce: Kombinování odhadů

Autor: Valeriya Plotnikova

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Jiří Dvořák, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Zkoumáme obecnou metodu kombinování několika odhadů stejného reálného parametru v parametrickém modelu. Uvažujeme vážený průměr daných počátečních odhadů, kde součet váh je roven jedné. Metoda je založena na minimalizaci střední kvadratické chyby. Váhový vektor se vypočte pomocí matice čtvercové chyby. Danou matici můžeme teoreticky spočítát nebo numericky odhadnout na základě počátečních odhadů. Výsledný kombinovaný odhad je pak konstruován jako součin příslušného váhového vektoru a vektoru počatečních odhadů parametru. Metoda je aplikovatelná ve většině situacích, kde máme k dispozici více odhadů jednoho parametru. Často je používána v predikcích například časových řad.

Klíčová slova: odhad parametru, kombinování odhadů, lineární kombinace, střední čtvercová chyba

Title: Estimator averaging

Author: Valeriya Plotnikova

Department: Department of probability and mathematical statistics

Supervisor: RNDr. Jiří Dvořák, Ph.D., Department of probability and mathematical statistics

Abstract: We investigate general procedure to combine several estimators of the same real parameter in the parametric model. Considering weighted average of initial estimator, where weights are constrained to the sum of one, we obtain the final combined estimator. In that framework, the optimal vector of weights minimises the mean square error, providing better estimator of parameter. Combined estimator will be computed as the product of vector of weights and the vector of initial estimators of the parameter. This method is frequently used in numerous problems of modern statistics like forecast averaging for time series.

Keywords: parameter estimator, estimator averaging, linear combination, mean-square error

Obsah

Úvod

Cílem této práce je zkoumání metody kombinování odhadů nějakého reálného parametru. Uvažujme tedy parametrický pravděpodobnostní model s reálným parametrem θ a posloupnost $T_1, \dots, T_k$ několika počátečních odhadů stejného parametru. Úkolem je získat lepší odhad parametru θ na základě lineární kombinace

$$\hat{\theta}_\lambda = \boldsymbol{\lambda}^T \boldsymbol{T},$$

kde $\boldsymbol{\lambda} \in \mathbb{R}^k$ je vektor koeficientů pro příslušné jednotlivé počáteční odhady T_i, $i \in \{1, \dots, k\}$. Nejlepším odhadem parametru θ mezi lineárními kombinacemi počátečních odhadů rozumíme takový odhad $\hat{\theta}_\lambda$, který minimalizuje střední kvadratickou chybu (MSE). Konkrétně se snažíme minimalizovat

$$\mathrm{MSE}(\hat{\theta}_\lambda) = \mathsf{E}\,(\hat{\theta}_\lambda - \theta)^2.$$

Problémem kombinování odhadů stejného parametru se zabývá několik oblastí současné statistiky. Často se tato metoda vyskytuje v predikcích (forecasting) například časových řad (Timmermann (2006)).

Přesnost výsledného kombinovaného odhadu $\hat{\theta}_\lambda$ přímo závisí na volbě omezujících podmínek vektoru váh $\boldsymbol{\lambda}$. V literatuře (Bates a Granger (1969), Lavancier a Rochet (2016)) se často uvažuje o následujícím váženém průměru

$$\sum_{i=1}^{k} \lambda_i = 1.$$

Uvažujeme-li nestranné počáteční odhady, pak je výše uvedena podmínka přirozená, protože zajišťuje nestrannost výsledného kombinovaného odhadu. Váhový vektor $\boldsymbol{\lambda}$, který splňuje dánou podmínku, získáme na základě matice čtvercové chyby Σ, která je založená na počátečních odhadech T_i. Většinou není matice Σ explicitně známá a proto ji budeme odhadovat maticí $\hat{\Sigma}$, kterou dostaneme aplikací jednoho konzistentního odhadu z posloupnosti počátečních odhadů. Výsledný kombinovaný odhad potom je váženým průměrem počátečních odhadů

$$\hat{\theta}_\lambda = \boldsymbol{\lambda}^T \boldsymbol{T}.$$

V první kapitole zavedeme značení a vysvětlíme základní pojmy týkající se metody kombinování odhadů. Poté provedeme podrobné odvození tvaru váhového vektoru. Stručně ukážeme, jak by se změnila příslušná metoda kdyby se jednalo o vícerozměrný odhadovaný parametr $\boldsymbol{\theta}$. Nakonec zavedeme krátkou diskusi o principech odhadování matice čtvercové chyby Σ.

V druhé kapitole aplikujeme výše uvedenou metodu na dva pravděpodobnostní parametrické modely. Jako první budeme uvažovat normální rozdělení s nulovou střední hodnotou a pak Poissonovo rozdělení. Pro každý případ si zvolíme dva počáteční odhady a následně uvedeme postup k získání výsledného kombinovaného odhadu metodou kombinování odhadů pro jednotlivá rozdělení.

V poslední kapitole se budeme zaobírat simulací náhodných výběrů z rozdělení, která byla uvedená v druhé kapitole. Odhadneme parametry jednotlivých rozdělení pomocí předem zvolené posloupnosti počátečních odhadů a poté pomocí výsledného kombinovaného odhadu. Výsledky simulace uvedeme v tabulkách a vhodným způsobem je interpretujeme na základě již probrané teorie z druhé kapitoly.

1. Metoda kombinování odhadů

Postup kombinování odhadů se liší podle toho, máme-li za úkol odhadnout jeden nebo více parametrů.

V této práci se budeme zabývat metodou kombinování odhadů pro jednorozměrný parametr z prostoru $\mathbb{R}$. Poté velmi stručně ukážeme, jak se dá stejnou metodu aplikovat pro odhadování vícerozměrného parametru z prostoru $\mathbb{R}^d$ pro nějaké d.

1.1 Základní pojmy a značení

Předpokládejme, že máme k dispozici pozorování $\boldsymbol{X} \sim P_\theta$, například náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ pro $n \geq 2$ z parametrického modelu P_θ. Mějme pak posloupnost $\boldsymbol{T} = (T_1, \ldots, T_k)^T$ počátečních odhadů parametru θ, kde $\theta \in \mathbb{R}$. Dále nechť existují konečné druhé momenty odhadů, tzn. $\mathsf{E}\, T_i^2 < +\infty$ pro $i \in \{1, \ldots, k\}$. V celé práci uvažujeme sloupcové vektory.

Cílem je navrhnout pravidlo, podle kterého by se dalo vhodným způsobem zkombinovat jednotlivé odhady T_i, kde $i \in \{1, \ldots, k\}$, tak, abychom na výstupu dostali nový odhad parametru θ, který označíme jako odhad $\hat{\theta}_\lambda$. Tento odhad bude zároveň optimální v situaci, kdy se kvalita odhadu posuzuje pomocí hodnoty střední čtvercové chyby.

V literatuře se nejčastěji můžeme setkat s následujícím vyjádřením odhadu $\hat{\theta}_\lambda$ parametru θ pomocí lineární kombinace

$$\hat{\theta}_\lambda = \boldsymbol{\lambda}^T \boldsymbol{T} = \sum_{i=1}^{k} \lambda_i T_i, \tag{1.1}$$

kde $\boldsymbol{\lambda} \in \Lambda$ je k-složkový reálný vektor a Λ je nějaká neprázdná podmnožina prostoru $\mathbb{R}^k$. Vektor $\boldsymbol{\lambda}$ v tomto případě získáme minimalizací střední čtvercové (resp. kvadratické) chyby (MSE, mean squared error)

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda} \in \Lambda} \mathsf{E}\,[(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)^T]. \tag{1.2}$$

Hledáme tedy vektor váh $\boldsymbol{\lambda}$, minimalizující výše uvedené kritérium. K tomu potřebujeme znát analytické vyjádření matice $\mathsf{E}\,[(\boldsymbol{T} - \theta\mathbf{1})(\boldsymbol{T} - \theta\mathbf{1})^T]$, která může záviset na odhadovaném parametru θ. Proto se v praxi zavádí optimální „oracle“ odhad (Lavancier a Rochet (2016)), který je lineární kombinací

$$\hat{\theta}^* = \boldsymbol{\lambda}^{*T} \boldsymbol{T}, \tag{1.3}$$

kde vektor $\boldsymbol{\lambda}^*$ minimalizuje střední čtvercovou chybu (viz. (1.2)). V této situaci střední kvadratická chyba průměruje čtverce rozdílů uvažovaného odhadu a odhadovaného parametru (zde θ). Na základě střední čtvercové chyby můžeme odhady porovnávat a určovat, který je lepší.

V praxi se často může stát, že vektor $\boldsymbol{\lambda}^*$ není známý, tudíž potřebujeme daný vektor odhadnout pomocí nějakého nového vektoru, který označíme například jako $\hat{\boldsymbol{\lambda}}$. Tvar výsledného kombinovaného odhadu pak je

$$\hat{\theta}_\lambda = \hat{\boldsymbol{\lambda}}^T \boldsymbol{T}.$$

1.2 Odvození metody kombinování odhadů pomocí váženého průměru

Efektivita metody kombinování odhadů přímo závisí na volbě množiny Λ. Z toho důvodu je důležité navrhnout množinu Λ takovou, aby optimální „oracle“ odhad (1.3) měl dobré vlastnosti. Rozumíme tím přesnost výsledného odhadu a zároveň nenáročnost při výpočtu příslušného váhového vektoru. V literatuře, která se zabývá problémem kombinování odhadů (například Bates a Granger, 1969 a Lavancier a Rochet, 2016), se často můžeme setkat s následující podmínkou omezení pro vektor $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda}^T \mathbf{1} = 1, \tag{1.4}$$

kde $\boldsymbol{\lambda}$ je reálný k-složkový vektor váh a $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^k$. Je-li zapotřebí přidat další omezení pro váhový vektor, například nezápornost váh, platí uvedená metoda pro všechny uzavřené neprázdné podmnožiny množiny

$$\Lambda_{max} := \{\boldsymbol{\lambda} \in \mathbb{R}^k : \boldsymbol{\lambda}^T \mathbf{1} = 1\}. \tag{1.5}$$

Problematika volby množiny omezujících podmínek Λ_{max} byla rozsáhle studována v literatuře Lavancier a Rochet (2016, str. 179-180).

Předpokládejme, že jednotlivé odhady $T_1, \dots, T_k$ mají konečné momenty druhého řádu a zároveň jsou lineárně nezávislé ve smyslu, že Gramova matice

$$\Sigma = \mathsf{E}\left[(\boldsymbol{T} - \theta\mathbf{1})(\boldsymbol{T} - \theta\mathbf{1})^T\right] \tag{1.6}$$

je dobře definovaná a regulární. Matici Σ nazveme *maticí střední čtvercové chyby* (resp. maticí MSE). Optimální vektor vah, definující „oracle“ odhad

$$\hat{\theta}^* = \boldsymbol{\lambda}^{*T} \boldsymbol{T},$$

má potom tvar

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda} \in \Lambda} \mathsf{E}\left[(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)^T\right] = \arg\min_{\boldsymbol{\lambda} \in \Lambda} \boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda}, \tag{1.7}$$

kde $\Lambda \subseteq \Lambda_{max}$ je nějaká neprázdná podmnožina.

Poslední rovnost odvodíme úpravováním střední čtvercové chyby s ohledem na podmínku pro váhy (1.4), tzn.

$$\begin{aligned} &\mathsf{E}\left[(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T \boldsymbol{T} - \theta)^T\right] = \mathsf{E}\left[(\boldsymbol{\lambda}^T \boldsymbol{T} - \boldsymbol{\lambda}^T \mathbf{1}\theta)(\boldsymbol{\lambda}^T \boldsymbol{T} - \boldsymbol{\lambda}^T \mathbf{1}\theta)^T\right] = \\ &= \mathsf{E}\left[(\boldsymbol{\lambda}^T \boldsymbol{T} - \boldsymbol{\lambda}^T \mathbf{1}\theta)(\boldsymbol{T}^T \boldsymbol{\lambda} - \theta\mathbf{1}^T \boldsymbol{\lambda})\right] = \mathsf{E}\left[\boldsymbol{\lambda}^T (\boldsymbol{T} - \theta\mathbf{1})(\boldsymbol{T} - \theta\mathbf{1})^T \boldsymbol{\lambda}\right] = \boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda}. \end{aligned}$$

Chceme-li nalézt extrém (konkrétně minimum) diferencovatelné funkce za předpokladu platnosti diferencovatelné omezující podmínky, aplikujeme metodu Lagrangeových multiplikátorů. Označíme funkci

$$f(\boldsymbol{\lambda}) \equiv \boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda},$$

a vazbu (omezující podmínku $\boldsymbol{\lambda}^T \mathbf{1} = 1$) ve tvaru

$$g(\boldsymbol{\lambda}) \equiv \boldsymbol{\lambda}^T \mathbf{1} - 1.$$

Metoda Lagrangeových multiplikátorů začíná vytvořením Lagrangeovy funkce $\mathcal{L}$, závislé na dvou proměných $\boldsymbol{\lambda}$ a μ. Konkrétně

$$\mathcal{L}(\boldsymbol{\lambda},\, \mu) = f(\boldsymbol{\lambda}) - \mu g(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda} - \mu(\boldsymbol{\lambda}^T \mathbf{1} - 1),$$

kde neznámou proměnnou μ rozumíme takzvaný Lagrangeoův multiplikátor. Derivací Lagrangeovy funkce podle jednotlivých proměnných dostáváme gradient

$$\nabla_{\boldsymbol{\lambda},\mu} \mathcal{L}(\boldsymbol{\lambda},\mu) = \left(\frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \boldsymbol{\lambda}}, \frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \mu} \right)$$

a následným položením rovným nule získáme soustavu dvou rovnic o dvou neznámých $\boldsymbol{\lambda}$ a μ

$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \boldsymbol{\lambda}} = \frac{\partial (\boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} - \mu \frac{\partial (\boldsymbol{\lambda}^T \mathbf{1})}{\partial \boldsymbol{\lambda}} = 0, \tag{1.8}$$

$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \mu} = -(\boldsymbol{\lambda}^T \mathbf{1} - 1) = 0. \tag{1.9}$$

Pro počítání parciální derivace nějaké funkce (resp. formy) podle k-složkového vektoru $\boldsymbol{\lambda}$ potřebujeme zavést následující dva lemmata:

Lemma 1 (Gentle, 1943, str. 152-155)**.** *Nechť α je skalár definovaný následujícím způsobem*

$$\alpha = \boldsymbol{y}^T \boldsymbol{A} \boldsymbol{x},$$

kde $\boldsymbol{y} \in \mathbb{R}^{m \times 1}$, $\boldsymbol{x} \in \mathbb{R}^{n \times 1}$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ a navíc matice $\boldsymbol{A}$ nezávisí na $\boldsymbol{y}$ a $\boldsymbol{x}$. Pak platí

$$\frac{\partial \alpha}{\partial \boldsymbol{x}} = \boldsymbol{y}^T \boldsymbol{A}, \qquad \frac{\partial \alpha}{\partial \boldsymbol{y}} = \boldsymbol{x}^T \boldsymbol{A}^T.$$

Lemma 2 (Gentle, 1943, str. 152-155)**.** *Mějme skalár α definovaný pomocí kvadratické formy*

$$\alpha = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x},$$

kde $\boldsymbol{x} \in \mathbb{R}^{n \times 1}$, $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ a platí, že $\boldsymbol{A}$ nezávisí na $\boldsymbol{x}$. Potom platí

$$\frac{\partial \alpha}{\partial \boldsymbol{x}} = \boldsymbol{x}^T (\boldsymbol{A} + \boldsymbol{A}^T).$$

Postupně aplikujeme Lemmata 2 a 1 na členy vyrazu (1.8) a první člen dostáváme ve tvaru

$$\frac{\partial (\boldsymbol{\lambda}^T \Sigma \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \boldsymbol{\lambda}^T (\Sigma + \Sigma^T) = 2 \boldsymbol{\lambda}^T \Sigma.$$

Pro druhý člen následně dostáváme

$$\mu \frac{\partial (\boldsymbol{\lambda}^T \mathbf{1})}{\partial \boldsymbol{\lambda}} = \mu \frac{\partial (\boldsymbol{\lambda}^T \mathbb{I} \mathbf{1})}{\partial \boldsymbol{\lambda}} = \mu \mathbf{1}^T \mathbb{I}^T = \mu \mathbf{1}^T \mathbb{I},$$

kde $\mathbb{I}$ je jednotková matice řádu k. Jejím násobením se výsledek nezmění, ale umožní nám použit Lemma 1. Poslední rovnost je zřejmá ze symetrie jednotkové matice.

Řešíme potom soustavu dvou rovnic (1.8) a (1.9) o dvou neznámých $\boldsymbol{\lambda}$ a μ:

$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \boldsymbol{\lambda}} = 2\boldsymbol{\lambda}^T\Sigma - \mu\mathbf{1}^T\mathbb{I} = 0,$$
$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda},\mu)}{\partial \mu} = -\boldsymbol{\lambda}^T\mathbf{1} + 1 = 0.$$

Z první rovnice vypočítáme $\boldsymbol{\lambda}^T$ jako

$$\boldsymbol{\lambda}^T = \frac{1}{2}\mu\mathbf{1}^T\mathbb{I}(\Sigma^{-1}) = \frac{1}{2}\mu\mathbf{1}^T(\Sigma^{-1}). \tag{1.10}$$

Matice Σ je regulární, tudíž existuje její inverze. Dosazením spočítané hodnoty $\boldsymbol{\lambda}^T$ do rovnice (1.9) dostáváme

$$-(\boldsymbol{\lambda}^T\mathbf{1} - 1) = -\frac{1}{2}\mu\mathbf{1}^T\mathbb{I}(\Sigma^{-1})\mathbf{1} + 1 = 0.$$

Potřebujeme vyjádřit μ, tudíž

$$\mu = \frac{2}{\mathbf{1}^T\mathbb{I}(\Sigma^{-1})\mathbf{1}} = \frac{2}{\mathbf{1}^T(\Sigma^{-1})\mathbf{1}}.$$

Násobení jednotkovou maticí v jmenovateli můžeme vynechat, protože neovlivní konečný výsledek. Nakonec dosazením spočítané hodnoty μ do (1.10) získáváme

$$\boldsymbol{\lambda}^T = \frac{2(\mathbf{1}^T(\Sigma^{-1}))}{2(\mathbf{1}^T(\Sigma^{-1})\mathbf{1})} = \frac{\mathbf{1}^T\Sigma^{-1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}.$$

Z transpozice součinu matic a symetrie matice Σ obdržíme minimum hledané funkce

$$\boldsymbol{\lambda}^*_{max} = \left(\frac{\mathbf{1}^T\Sigma^{-1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}\right)^T = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}, \tag{1.11}$$

kde $\boldsymbol{\lambda}^*_{max} \in \Lambda_{max} = \Lambda$ a Λ_{max} je definovaná v (1.5).

V praxi se často může stát, že matice čtvercové chyby Σ je závislá na odhadovaném parametru θ. V tomto případě ji potřebujeme odhadnout pomocí jiného odhadu, řekněme $\hat{\Sigma}$, který umožní vypočítat výsledný zprůměrovaný odhad

$$\hat{\theta} = \hat{\boldsymbol{\lambda}}^T\boldsymbol{T}, \tag{1.12}$$

kde platí

$$\hat{\boldsymbol{\lambda}} = \arg\min_{\boldsymbol{\lambda}\in\Lambda} \boldsymbol{\lambda}^T\hat{\Sigma}\boldsymbol{\lambda}.$$

Potřebujeme tedy vhodným způsobem získat odhad $\hat{\Sigma}$. Problematikou odhadování matice Σ se budeme krátce zaobírat v podkapitole 1.3.

Ve výsledku máme zaručeno, že konečný kombinovaný odhad (1.12) vylepší původní posloupnost odhadů $\boldsymbol{T} = (T_1, \dots, T_k)^T$ parametru θ a jakoukoliv jejich lineární kombinaci.

1.3 Odhad matice čtvercové chyby

Přesnost odhadu $\hat{\Sigma}$ matice Σ považujeme za nejdůležitější faktor, který přímo ovlivňuje efektivitu procedury kombinování odhadů. Problematika volby odhadu $\hat{\Sigma}$ se rozlišuje podle toho, zda se jedná o parametrický nebo neparametrický (resp. semiparametrický) model. V každém případě odhadování matice Σ může být provedeno ze stejných dat, ze kterých byla vytvořena počáteční posloupnost odhadů $\boldsymbol{T} = (T_1, \ldots, T_k)^T$.

Máme-li k dispozici plně specifikovaný parametrický model, potom matici Σ můžeme odhadnout zapojením jednoho z počátečních odhadů parametru θ. Předpokládejme tedy, že matice čtvercové chyby Σ může být vyjádřena jako nějaké zobrazení parametru θ, tj.

$$\Sigma(.) : \mathbb{R}^d \mapsto \mathbb{R}^{k\times k}.$$

Tehdy si můžeme zvolit

$$\hat{\Sigma} = \Sigma(\hat{\theta}_0),$$

kde $\hat{\theta}_0$ je konzistentní odhad. Za nejvhodnější volbu odhadu $\hat{\theta}_0$ je považován libovolný z počátečních odhadů parametru θ, který je konzistentní.

Není-li zobrazení $\Sigma(.)$ explicitně známé, pak $\Sigma(\hat{\theta}_0)$ je možné aproximovat pomocí simulace Monte-Carlo a použitím počátečního odhadu $\hat{\theta}_0$. Daná metoda je často známá jako „*parametric bootstrap*“ a je důkladně popsána v práci Lavancier a Rochet (2016, str. 184-189).

1.4 Metoda kombinování odhadů pro vícerozměrný parametr

Nyní budeme předpokládát mnohorozměrný parametr

$$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^T \in \mathbb{R}^d,$$

kde pro každé θ_j uvažujeme posloupnost počátečních odhadů $\boldsymbol{T}_j$. Předpokládejme dále, že každá posloupnost $\boldsymbol{T}_j$ má minimálně jeden počáteční odhad příslušného parametru θ_j. Obecně posloupnosti $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_d$ můžou mít různé délky $k_1, \ldots, k_d$, pro $k_j \geq 1$. Mějme tedy posloupnosti $\boldsymbol{T}_1 \in \mathbb{R}^{k_1}, \ldots, \boldsymbol{T}_d \in \mathbb{R}^{k_d}$ počátečních odhadů a množinu $\boldsymbol{T} = (\boldsymbol{T}_1^T, \ldots, \boldsymbol{T}_d^T)^T \in \mathbb{R}^k$, kde $k = \sum_{j=1}^{d} k_j$. Výsledný vícerozměrný kombinovaný odhad potom má tvar

$$\hat{\boldsymbol{\theta}}_\lambda = \boldsymbol{\lambda}^T \boldsymbol{T} \in \mathbb{R}^d,$$

zde $\boldsymbol{\lambda}$ je reálná matice o velikosti $k \times d$. Ukážeme na jednoduchém příkladě postup kombinování odhadů pro vícerozměrný parametr.

Příklad: Uvažujeme parametrický model s parametrem

$$\boldsymbol{\theta} = (\theta_1, \theta_2)^T \in \mathbb{R}^2$$

Pro první složku θ_1 odhadovaného vícerozměrného parametru mějme posloupnost počátečních odhadů $\boldsymbol{T}_1$. Pro druhou složku θ_2 uvažujme obdobně posloupnost $\boldsymbol{T}_2$.

Pro $\boldsymbol{T}_1$ a $\boldsymbol{T}_2$ je dáno

$$\begin{aligned}\boldsymbol{T}_1 &= (T_{11}, T_{12}, T_{13})^T \in \mathbb{R}^3\\ \boldsymbol{T}_2 &= (T_{21}, T_{22})^T \in \mathbb{R}^2.\end{aligned}$$

Pak množina počátečních odhadů dvourozměrného parametru vypadá následovně

$$\boldsymbol{T} = (\boldsymbol{T}_1^T, \boldsymbol{T}_2^T)^T = (T_{11}, T_{12}, T_{13}, T_{21}, T_{22})^T \in \mathbb{R}^5.$$

Matice $\boldsymbol{\lambda} \in \mathbb{R}^{5\times 2}$ je tvaru

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_{11} & 0\\ \lambda_{12} & 0\\ \lambda_{13} & 0\\ 0 & \lambda_{21}\\ 0 & \lambda_{22}\end{pmatrix}.$$

Dvourozměrný kombinovaný odhad parametru $\hat{\boldsymbol{\theta}}_\lambda$ pak vychází ve tvaru

$$\hat{\boldsymbol{\theta}}_\lambda = \boldsymbol{\lambda}^T\boldsymbol{T} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & 0 & 0\\ 0 & 0 & 0 & \lambda_{21} & \lambda_{22}\end{pmatrix}\begin{pmatrix} T_{11}\\ T_{12}\\ T_{13}\\ T_{21}\\ T_{22}\end{pmatrix} = \begin{pmatrix} \lambda_{11}T_{11} + \lambda_{12}T_{12} + \lambda_{13}T_{13}\\ \\ \lambda_{21}T_{21} + \lambda_{22}T_{12}\end{pmatrix}.$$

Definujeme nyní matici J

$$J = \begin{pmatrix} \mathbf{1}_{k_1} & 0 & \dots & 0\\ 0 & \mathbf{1}_{k_2} & \ddots & 0\\ \vdots & \ddots & \ddots & 0\\ 0 & \dots & 0 & \mathbf{1}_{k_d}\end{pmatrix} \in \mathbb{R}^{k\times d},$$

kde $\mathbf{1}_{k_j}$ je sloupcový jednotkový vektor o délce k_j. Předpokládejme dále množinu omezujících podmínek. Konkrétně

$$\Lambda_{max} = \{\boldsymbol{\lambda} \in \mathbb{R}^{k\times d} : \boldsymbol{\lambda}^T J = \mathbb{I}_d\},$$

$\mathbb{I}_d$ zde označuje jednotkovou matici řádu d. Zvolíme-li $d = 1$, tak dostáváme dříve zmíněnou množinu Λ_{max} rovnou množině (1.5). Pro výše ilustrovaný příklad uvedeme nyní příslušnou množinu omezujících podmínek.

Příklad - pokračování: Matice J výchází následující

$$J = \begin{pmatrix} \mathbf{1}_3 & 0\\ 0 & \mathbf{1}_2\end{pmatrix} = \begin{pmatrix} 1 & 0\\ 1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 1\end{pmatrix}.$$

Podmínka množiny Λ_{max} tedy je

$$\boldsymbol{\lambda}^T J = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{21} & \lambda_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_{11} + \lambda_{12} + \lambda_{13} & 0 \\ 0 & \lambda_{21} + \lambda_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_d.$$

Zde se jedná o přirozené podmínky

$$\lambda_{11} + \lambda_{12} + \lambda_{13} = 1,$$
$$\lambda_{21} + \lambda_{22} = 1.$$

Zavedeme pak matici

$$\Sigma = \mathsf{E}\left[(\boldsymbol{T} - J\boldsymbol{\theta})(\boldsymbol{T} - J\boldsymbol{\theta})^T\right] \in \mathbb{R}^{k \times k},$$

která musí být dobře definována a regulární. Optimální „oracle“ odhad byl definován v Lavancier a Rochet (2016, str. 178) jako lineární transformaci $\hat{\boldsymbol{\theta}}^* = \boldsymbol{\lambda}^{*T}\boldsymbol{T}$, pro kterou

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda}\in\Lambda} \mathsf{E}\ ||\boldsymbol{\lambda}^T\boldsymbol{T} - \boldsymbol{\theta}||^2 = \arg\min_{\boldsymbol{\lambda}\in\Lambda}\ \mathrm{tr}(\boldsymbol{\lambda}^T\Sigma\boldsymbol{\lambda}),$$

kde $||.||$ je Euklidovská norma v prostoru $\mathbb{R}^d$, $\mathrm{tr}(\boldsymbol{\lambda}^T\Sigma\boldsymbol{\lambda})$ označuje stopu příslušné matice a $\Lambda \subseteq \Lambda_{max}$.

Na konci podkapitoly 1.2 jsme uvedli poznámku o odhadu $\hat{\Sigma}$ matice Σ, která je použitelná i na případ vícerozměrného parametru. Pro odhad vektoru $\boldsymbol{\lambda}$ platí

$$\hat{\boldsymbol{\lambda}} = \arg\min_{\boldsymbol{\lambda}\in\Lambda}\ \mathrm{tr}(\boldsymbol{\lambda}^T\hat{\Sigma}\boldsymbol{\lambda}).$$

Výsledný odhad vícerozměrného parametru získaný metodou kombinování odhadů je potom tvaru

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\lambda}}^T\boldsymbol{T}$$

a je nejlepším odhadem ve smyslu střední kvadratické chyby v rámci posloupnosti počátečních odhadů a jejich lineárních kombinací.

Ve skutečnosti platí, že metoda kombinování odhadů pro jednorozměrný odhadovaný parametr je speciálním případem metody kombinování odhadů pro parametr vícerozměrný, kde platí $d = 1$.

2. Příklady dvou odhadů stejného parametru

2.1 Diskuse volby modelů

V této kapitole se budeme zabývat aplikací metody kombinování odhadů pro jeden parametr a analyticky odvodíme výsledné kombinované odhady pro dva různé modely pravděpodobnostních rozdělení.

Nejprve se budeme zabývat odhadem rozptylu v normálním rozdělení s nulovou střední hodnotou. Pro odhad rozptylu si zvolíme dva velmi podobné odhady, které jsou založené na výběrovém rozptylu (liší se jen konstantou). Jako výsledek dostáváme váhy pro počáteční odhady parametru rozptylu v normálním rozdělení.

Zajímavým příkladem je model Poissonova rozdělení. Zde chceme odhadnout parametr λ, který je zároveň střední hodnotou a rozptylem náhodných veličin z daného rozdělení. V tomto případě za odhady parametru λ si zvolíme výběrový průměr a výběrový rozptyl. Ukáže se, že jeden z odhadů převládne a celá váha padne na něj. Následně vysvětlíme daný fakt pomocí tzv. *Lehmann-Scheffého* věty.

2.2 Příklad pro model normálního rozdělení

Pracujme v modelu $\mathcal{F} = \{\mathcal{N}(0,\sigma^2), \sigma^2 > 0\}$ (všechna normální rozdělení s nulovou střední hodnotou). Nechť $X_1, \ldots, X_n$, $n \geq 2$ je náhodný výběr z rozdělení $\mathcal{N}(0,\sigma^2)$. Cílem je odhadnout parametr σ^2, odpovídající rozptylu $\operatorname{var} X_i$ pro každé $i \in \{1, \ldots, n\}$.

Zvolíme si následující dva odhady parametru σ^2:

$$T_1 \equiv S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2, \tag{2.1}$$

$$T_2 \equiv \hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2, \tag{2.2}$$

kde pomocí S_n^2 značíme výběrový rozptyl a za $\bar{X}_n$ výběrový průměr. Odhad $\hat{\sigma}_n^2$ je získaný metodou maximální věrohodnosti. Posloupnost odhadů parametru σ^2 je potom tvaru $\mathbf{T} = (T_1,T_2)^T$. Z vlastnosti výběrového rozptylu víme, že odhad S_n^2 je nestranný a konzistentní odhad parametru σ^2. Odhad (2.2) je také konzistentní, ale již není nestranný.

Výsledný kombinovaný odhad, který označíme pomocí $\hat{\sigma}_\lambda^2$, byl popsán ve vzorci (1.1). Vyjádříme ho následujícím způsobem:

$$\hat{\sigma}_\lambda^2 = \boldsymbol{\lambda}^T\boldsymbol{T} = (\lambda_1,\lambda_2)\begin{pmatrix} S_n^2 \\ \hat{\sigma}_n^2 \end{pmatrix} = \lambda_1 S_n^2 + \lambda_2\hat{\sigma}_n^2,$$

kde vektor $\boldsymbol{\lambda} = (\lambda_1,\lambda_2)^T$ splňuje podmínku $\boldsymbol{\lambda}^T\boldsymbol{T} = 1$. Úkolem je tedy najít hodnotu váhového vektoru $\boldsymbol{\lambda}$ pomocí vyjádření (1.7). Čili

$$\boldsymbol{\lambda}^* = \arg\min_{\boldsymbol{\lambda}\in\Lambda_{max}} \mathsf{E}\left[(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)(\boldsymbol{\lambda}^T\boldsymbol{T} - \theta)^T\right] = \arg\min_{\boldsymbol{\lambda}\in\Lambda_{max}} \boldsymbol{\lambda}^T\Sigma\boldsymbol{\lambda},$$

kde Λ_{max} je definována v (1.5) a $\boldsymbol{\lambda}^*$ je optimální vektor váh, definující „oracle" odhad $(\hat{\sigma}^2)^* = \boldsymbol{\lambda}^{*T}\boldsymbol{T}$.

Odhady T_1 a T_2 mají konečné momenty druhého řádu a zároveň jsou lineárně nezávislé tak, že Gramova matice Σ je dobře definována a je regulární. Tudíž můžeme spočítat matici čtvercové chyby uvedenou v (1.6)

$$\Sigma = \mathsf{E}\,[(\boldsymbol{T} - \sigma^2\mathbf{1})(\boldsymbol{T} - \sigma^2\mathbf{1})^T] = \mathsf{E}\left[\begin{pmatrix} S_n^2 - \sigma^2 \\ \hat{\sigma}_n^2 - \sigma^2 \end{pmatrix}(S_n^2 - \sigma^2, \hat{\sigma}_n^2 - \sigma^2)\right] =$$

$$= \mathsf{E}\begin{pmatrix} (S_n^2 - \sigma^2)^2 & (S_n^2 - \sigma^2)(\hat{\sigma}_n^2 - \sigma^2) \\ (\hat{\sigma}_n^2 - \sigma^2)(S_n^2 - \sigma^2) & (\hat{\sigma}_n^2 - \sigma^2)^2 \end{pmatrix} =$$

$$= \begin{pmatrix} \mathsf{E}\,(S_n^2)^2 - 2\sigma^2\,\mathsf{E}\,S_n^2 + \sigma^4 & \mathsf{E}\,(S_n^2\hat{\sigma}_n^2) - \sigma^2(\mathsf{E}\,S_n^2 + \mathsf{E}\,\hat{\sigma}_n^2) + \sigma^4 \\ \mathsf{E}\,(S_n^2\hat{\sigma}_n^2) - \sigma^2(\mathsf{E}\,S_n^2 + \mathsf{E}\,\hat{\sigma}_n^2) + \sigma^4 & \mathsf{E}\,(\hat{\sigma}_n^2)^2 - 2\sigma^2\,\mathsf{E}\,\hat{\sigma}_n^2 + \sigma^4 \end{pmatrix}.$$

Nyní vypočítáme jednotlivé střední hodnoty, které se vyskytují v prvcích matice Σ. Pro odvození některých momentů budeme potřebovat následující větu:

Věta 3 (Anděl, 2007, str. 70)**.** *Nechť* $\mathbf{X} = (X_1, \ldots, X_n)^T$ *je náhodný výběr. Pak nechť pro každé* $i = 1, \ldots, n$ *platí* $X_i \sim \mathcal{N}(\mu, \sigma^2)$*. Je-li* $n > 1$ *a* $\sigma^2 > 0$*, pak*

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Důkaz. Jednotlivé kroky důkazu jsou popsány v práci Anděl (2007, str. 70-71). ❑

Položíme-li

$$Y \equiv \sum_{i=1}^{n}(X_i - \bar{X}_n)^2,$$

potom z Věty 3 plyne

$$\frac{(n-1)S_n^2}{\sigma^2} = \frac{Y}{\sigma^2} \sim \chi^2_{n-1}. \tag{2.3}$$

Příslušné střední hodnoty v prvcích matice Σ jsou:

- $\mathsf{E}\,S_n^2 = \sigma^2$ a to platí z nestrannosti výběrového rozptylu.
- $\mathsf{E}\,(S_n^2)^2 = \operatorname{var} S_n^2 + (\mathsf{E}\,S_n^2)^2 = \frac{2\sigma^4}{n-1} + \sigma^4 = \frac{(n+1)\sigma^4}{n-1}$ (viz. Věta 3).
- $\mathsf{E}\,\hat{\sigma}_n^2 = \frac{1}{n}\,\mathsf{E}\,Y = \frac{(n-1)\sigma^2}{n}$ (dle důsledku (2.3)).
- $\mathsf{E}\,(\hat{\sigma}_n^2)^2 = \operatorname{var}\hat{\sigma}_n^2 + (\mathsf{E}\,\hat{\sigma}_n^2)^2 = \frac{1}{n^2}\operatorname{var} Y + (\mathsf{E}\,\hat{\sigma}_n^2)^2 = \frac{(n^2-1)\sigma^4}{n^2}$, opět využíváme důsledku (2.3).
- $\mathsf{E}\,(S_n^2\hat{\sigma}_n^2) = \mathsf{E}\left(\frac{1}{n-1}Y\frac{1}{n}Y\right) = \frac{n-1}{n}\,\mathsf{E}\,(S_n^2)^2 = \frac{(n+1)\sigma^4}{n}$.

Spočítané střední hodnoty dosadíme zpět do matice Σ a tím dostáváme

$$\Sigma = \begin{pmatrix} \frac{2\sigma^4}{n-1} & \frac{2\sigma^4}{n} \\ \frac{2\sigma^4}{n} & \frac{(2n-1)\sigma^4}{n^2} \end{pmatrix}.$$

V podkapitole 1.2 jsme odvodili vyjádření vektoru váh pomocí Gramovy matice Σ (viz. (1.11)), konkrétně

$$\boldsymbol{\lambda}^*_{max} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} = \begin{pmatrix} \frac{1-n}{1+n} \\ \frac{2n}{1+n} \end{pmatrix},$$

kde inverzní matici Σ^{-1} jsme vypočítali pomocí adjungové matice. Pro takový váhový vektor $\boldsymbol{\lambda}^*_{max}$ platí podmínka (1.4).

Zde si můžeme všimnout, že vypočítaný vektor váh $\boldsymbol{\lambda}^*_{max}$ je přímo optimální váhový vektor, definující „oracle“ odhad. Proto nemusíme odhadovat daný vektor, resp. odhadovat matici Σ. V našem případě platí, že se z každého prvku matice Σ dá vytknout σ^2 tak, že se v jednotlivých prvcích matice již nevyskytuje parametr σ^2. Následně, díky tvaru (1.11), by se odhadovaný parametr vykrátil. Proto nemusíme odhadovat příslušnou matici Σ pomocí nějakého konzistentního počátečního odhadu. Často se může stát, že není možné vyjádřit matici Σ pomocí nějakého násobku odhadovaného parametru. V této situaci je důležité odhadnout Σ pomocí $\hat{\Sigma}$, obsahující konzistentní počáteční odhad.

Ve výsledku jsme vypočítali příslušné hodnoty váh pro odhady S_n^2 a $\hat{\sigma}_n^2$ parametru σ^2, definované v (2.1) a (2.2). Výsledný kombinovaný odhad $\hat{\sigma}^2_\lambda$ má tvar

$$\hat{\sigma}^2_\lambda = (\boldsymbol{\lambda}^*_{max})^T\boldsymbol{T} = \left(\frac{1-n}{1+n}\right)S_n^2 + \left(\frac{2n}{1+n}\right)\hat{\sigma}_n^2 = \left(\frac{n-1}{n+1}\right)S_n^2.$$

Střední čtvercovou chybu kombinovaného odhadu můžeme spočítat na základě střední čtvercové chyby výběrového rozptylu, která je prvkem matice Σ na místě [1,1]. Tedy

$$\mathrm{MSE}(\hat{\sigma}^2_\lambda) = \mathrm{MSE}\left(\left(\frac{n-1}{n+1}\right)S_n^2\right) = \frac{2\sigma^4}{n+1}.$$

Kombinovaný odhad $\hat{\sigma}^2_\lambda$ můžeme vyjádřit pomocí c-násobku výběrového rozptylu S_n^2, tj.

$$\hat{\sigma}^2_\lambda = cS_n^2, \qquad c = \left(\frac{n-1}{n+1}\right). \tag{2.4}$$

Chceme nyní zjistit, zda hodnota konstanty c je opravdovým minimem mezi násobky S_n^2.

Zvolíme-li si nový odhad $\hat{A}_n$, který je kladným c-násobkem výběrového rozptylu S_n^2, čili

$$\hat{A}_n = cS_n^2,$$

potom příslušná střední čtvercová chyba má tvar

$$\mathrm{MSE}(\hat{A}_n) = \mathrm{var}\,\hat{A}_n + (\mathrm{Bias}\,\hat{A}_n)^2,$$

kde Bias $= \mathsf{E}\,(\hat{A}_n - \sigma^2)$ je vychýlení. Odtud plyne

$$\begin{aligned}\mathrm{MSE}(\hat{A}_n) &= c^2 \operatorname{var} S_n^2 + (\mathsf{E}\,(cS_n^2 - \sigma^2))^2 = \\ &= c^2\left(\frac{2(\sigma^2)^2}{n-1}\right) + ((c-1)\sigma^2)^2 = \left(\frac{2c^2}{n-1} + c^2 - 2c + 1\right)(\sigma^2)^2.\end{aligned}$$

Chceme zjistit pro jakou hodnotu c bude střední čtvercová chyba odhadu $\hat{A}_n$ minimální, tj. pro jaké c bude odhad $\hat{A}_n$ nejbližší k odhadovanému parametru σ^2. Položíme-li

$$f(c,n) = \frac{2c^2}{n-1} + c^2 - 2c + 1,$$

pro nějaké pevně zvolené $n \geq 2$, potom chceme minimalizovat funkci $f(c,n)$. Hledáme c_0 takové, že

$$c_0 = \arg \min_{c \in (0,+\infty)} f(c,n).$$

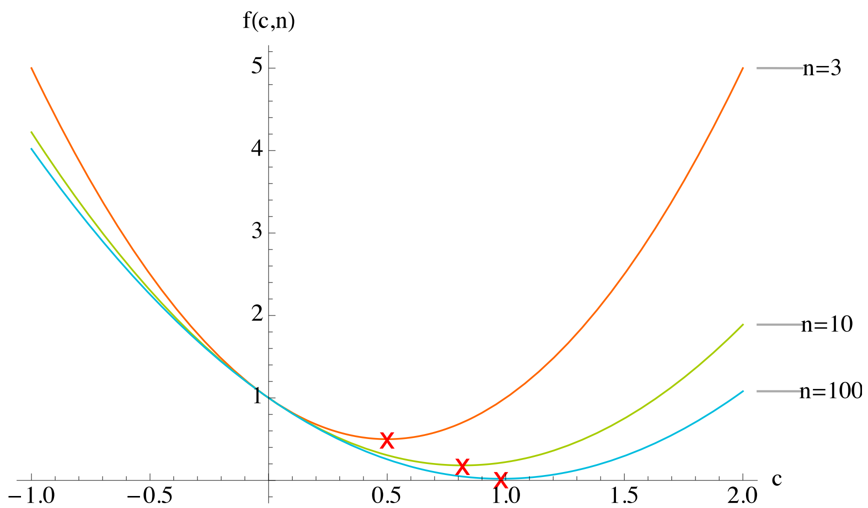
Zderivujeme funkci $f(c,n)$ podle c a položíme derivaci rovnou nule

$$\frac{\partial f(c,n)}{\partial c} = \frac{4c}{n-1} + 2c - 2 = 0,$$

pak dostáváme hodnotu příslušného c_0

$$c_0 = \left(\frac{n-1}{n+1}\right), \qquad n \geq 2,$$

což odpovídá výsledku kombinovaného odhadu uvedeného v (2.4). Výsledný kombinovaný odhad $\hat{\sigma}^2_\lambda$ je tedy skutečně nejlepší odhad ve smyslu střední čtvercové chyby v uvažovaném modelu. Pro lepší představu jsme graficky znázornili funkci $f(c,n)$ pro případy, kdy n může nabývat hodnot 3, 10 a 100, a také příslušná minima jednotlivých funkcí (viz. Obrázek 2.1).



Obrázek 2.1: Grafické znázornění funkce $f(c,n)$ pro tři různé volby rozsahu výběru n: červenými kříži jsou zde vyznačená minima c_0 příslušných funkcí

.

2.3 Příklad pro model Poissonova rozdělení

Nyní uvažujeme náhodný výběr $\mathbf{X} = (X_1, \dots, X_n)^T$, $n \geq 2$ v modelu Poissonova rozdělení $\mathcal{F} = \{\text{Pois}(\lambda), \lambda > 0\}$. Odhadujeme parametr $\lambda = \mathsf{E}\ X_i = \text{var}\, X_i$, kde $i \in \{1, \dots, n\}$.

Zvolíme si následující dva odhady parametru λ:

$$T_1 \equiv \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

$$T_2 \equiv S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2,$$

kde $\bar{X}$ je výběrový průměr a S_n^2 je výběrový rozptyl. Oba odhady jsou nestranné a zároveň konzistentní. Posloupnost odhadů parametru λ je potom $\mathbf{T} = (T_1, T_2)^T$. Pro lepší náhled na problém označíme tentokrát vektor váh jako $\boldsymbol{w} = (w_1, w_2)^T$ (weights). Přeformulujeme-li problém (1.1), pak kombinovaný odhad λ bude vyjádřen následovně

$$\hat{\lambda}_{\boldsymbol{w}} = \boldsymbol{w}^T \boldsymbol{T} = w_1 \bar{X}_n + w_2 S_n^2$$

a pro vektor váh musí platit podmínka $\boldsymbol{w}^T \mathbf{1} = 1$.

Obdobně jako u předchozího případu, spočítáme hodnotu váhového vektoru $\boldsymbol{w}$ pomocí optimálního vektoru váh $\boldsymbol{w}^*$, který definuje „oracle“ odhad $\hat{\lambda}^* = \boldsymbol{w}^{*T} \boldsymbol{T}$. To jest

$$\boldsymbol{w}^* = \arg \min_{\boldsymbol{w} \in \Lambda_{max}} \mathsf{E}\left[(\boldsymbol{w}^T \boldsymbol{T} - \lambda)(\boldsymbol{w}^T \boldsymbol{T} - \lambda)^T\right] = \arg \min_{\boldsymbol{w} \in \Lambda_{max}} \boldsymbol{w}^T \Sigma \boldsymbol{w},$$

kde

$$\Lambda_{max} = \{\boldsymbol{w} \in \mathbb{R}^k : \boldsymbol{w}^T \mathbf{1} = 1\}.$$

Pro Gramovu matici jsou splněny předpoklady: pro každé $i \in \{1,2\}$ platí $\mathsf{E}\ T_i^2 < +\infty$ a zároveň T_1 a T_2 jsou lineárně nezávislé. Tudíž můžeme spočítat matici čtvercové chyby Σ dle (1.6)

$$\Sigma = \mathsf{E}\left[(\boldsymbol{T} - \sigma^2 \mathbf{1})(\boldsymbol{T} - \sigma^2 \mathbf{1})^T\right] = \mathsf{E}\left[\begin{pmatrix} \bar{X}_n - \lambda \\ S_n^2 - \lambda \end{pmatrix} (\bar{X}_n - \lambda, S_n^2 - \lambda)\right] =$$

$$= \mathsf{E}\begin{pmatrix} (\bar{X}_n - \lambda)^2 & (\bar{X}_n - \lambda)(S_n^2 - \lambda) \\ (S_n^2 - \lambda)(\bar{X}_n - \lambda) & (S_n^2 - \lambda)^2 \end{pmatrix} =$$

$$= \begin{pmatrix} \mathsf{E}\,(\bar{X}_n)^2 - 2\lambda\,\mathsf{E}\ \bar{X}_n + \lambda^2 & \mathsf{E}\,(\bar{X}_n S_n^2) - \lambda(\mathsf{E}\ \bar{X}_n + \mathsf{E}\ S_n^2) + \lambda^2 \\ \mathsf{E}\,(\bar{X}_n S_n^2) - \lambda(\mathsf{E}\ \bar{X}_n + \mathsf{E}\ S_n^2) + \lambda^2 & \mathsf{E}\,(S_n^2)^2 - 2\lambda\,\mathsf{E}\ S_n^2 + \lambda^2 \end{pmatrix}.$$

Jednotlivé střední hodnoty v prvcích matice Σ jsou:

- $\mathsf{E}\ \bar{X}_n = \mathsf{E}\ S_n^2 = \lambda$, díky nestrannosti obou odhadů.
- $\mathsf{E}\,(\bar{X}_n)^2 = \text{var}\,\bar{X}_n + (\mathsf{E}\ \bar{X}_n)^2 = \frac{1}{n}\lambda + \lambda^2$.

- $\mathsf{E}\,(S_n^2)^2 = \operatorname{var} S_n^2 + (\mathsf{E}\ S_n^2)^2 = \frac{1}{n}\lambda + \left(\frac{n+1}{n-1}\right)\lambda^2$. Rozptyl S_n^2 jsme získali pomocí faktoriálního momentu s předpokladem, že existují konečné momenty druhého a čtvrtého řádu pro každé X_i, kde $i \in \{1, \ldots, n\}$, čili

$$\operatorname{var} S_n^2 = \mathsf{E}\,(S_n^2)^2 - (\mathsf{E}\ S_n^2)^2 = \frac{1}{n}\lambda + \Big(\frac{2}{n-1}\Big)\lambda^2.$$

Pro střední hodnotu smíšeného členu $\bar{X}_n S_n^2$ využíváme následující věty:

Věta 4 (Dahiya a Gurland, 1969, str. 172). *Nechť $\mathbf{X} = (X_1, \ldots, X_n)^T$ je náhodný výběr z Poissonova rozdělení se střední hodnotou λ. Pro každé $\delta \geq 0$ platí*

$$\mathsf{E}((\bar{X}_n)^\delta S_n^2) = E((\bar{X}_n)^{\delta+1}),$$

kde $\bar{X}_n$ je výběrový průměr a S_n^2 je výběrový rozptyl.

Důkaz. Důkaz věty je detailně vypracován v článku Dahiya a Gurland (1969, str. 172).

❑

Odtud získáváme

- $\mathsf{E}\,(\bar{X}_n S_n^2) = \mathsf{E}\,(\bar{X}_n)^2 = \frac{1}{n}\lambda + \lambda^2.$

Máme k dispozici všechny střední hodnoty jednotlivých členů matice Σ, a nakonec dostáváme

$$\Sigma = \begin{pmatrix} \frac{\lambda}{n} & \frac{\lambda}{n} \\ \frac{\lambda}{n} & \frac{\lambda}{n} + \frac{2\lambda^2}{n-1} \end{pmatrix}. \tag{2.5}$$

Pro odvození vektoru váh použijeme vztah (1.11). Matice Σ je regulární, a proto existuje inverzní matice Σ^{-1}. Váhový vektor je

$$\boldsymbol{w}^*_{max} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

protože

$$\Sigma^{-1}\mathbf{1} = \left(\frac{n}{\bar{X}_n}, 0\right)^T \qquad \text{a} \qquad \mathbf{1}^T\Sigma^{-1}\mathbf{1} = \frac{n}{\bar{X}_n},$$

a zřejmě platí $\boldsymbol{w}^T\mathbf{1} = 1$. Váhový vektor $\boldsymbol{w}^*_{max}$ je opět optimální váhový vektor, který definuje „oracle“ odhad (viz. (1.2)). Ve výsledku platí, že vektor váh nezávisí na parametrech modelu a tedy máme přímo dostupný „oracle“ odhad.

Nakonec pro odhad parametru λ v modelu $\mathcal{F} = \{\text{Pois}(\lambda), \lambda > 0\}$ obdržíme následující kombinovaný odhad

$$\hat{\lambda}_{\boldsymbol{w}} = (\boldsymbol{w}^*_{max})^T\boldsymbol{T} = 1\bar{X}_n + 0S_n^2 = \bar{X}_n.$$

Střední čtvercová chyba kombinovaného odhadu se rovná střední čtvercové chybě výběrového průměru, která je prvkem matice Σ na místě [1,1]. To jest

$$\text{MSE}(\hat{\lambda}_w) = \text{MSE}(\bar{X}_n) = \frac{\lambda}{n}.$$

Výsledek metody kombinování odhadů v modelu Poissona rozdělení s výše uvedenými odhady by neměl být překvapivý. Poznámenáme-li rozptyly výběrového průměru a výběrového rozptylu, pak zjevně platí následující nerovnost

$$\operatorname{var} S_n^2 = \frac{\lambda}{n} + \frac{2\lambda^2}{n-1} \quad > \quad \frac{\lambda}{n} = \operatorname{var} \bar{X}_n.$$

Můžeme tuto nerovnost interpretovat jako fakt, že variabilita, tedy i odchylka pro S_n^2 je v tomto případě větší než variabilita $\bar{X}_n$. To je důvod, proč celá váha odhadovaného parametru λ padne na odhad $\bar{X}_n$.

Chování váhového vektoru $\hat{\boldsymbol{w}}^*_{max}$ se teoreticky dá vysvětlit pomocí *Lehmann-Scheffeovy věty*, která upřesní pojem *nejlepšího nestranného odhadu*. Na začátek zavedeme dvě pomocné definice.

Definice 1 (Anděl, 2007, str. 136)**.** *Statistika* $\mathbf{S}$ *se nazývá* suficientní *pro parametr* $\boldsymbol{\theta}$*, jestliže podmíněné rozdělení vektoru* $\mathbf{X} = (X_1, \dots, X_n)^T$ *při daném* $\mathbf{S}$ *nezávisí na* $\boldsymbol{\theta}$*.*

Jinými slovy by se dalo říct, že statistika je *suficientní*, pokud vyčerpává každou informaci o parametru θ, která byla v příslušném náhodném výběru $X_1, \dots, X_n$ původně obsažená.

Definice 2 (Anděl, 2007, str. 124)**.** *Řekneme, že statistika* $\mathbf{S}$ *je* úplná*, platí-li pro každou její měřitelnou funkci* $w(\mathbf{S})$*, která má konečnou střední hodnotu, implikace*

$$\{\mathsf{E}\, w(\mathbf{S}) = 0 \text{ pro každé } \theta \in \Omega\} \implies \{w(\mathbf{S}) = 0 \text{ skoro jistě pro každé } \theta \in \Omega\}.$$

Věta 5 (Anděl, 2007, str. 131)**.** *Nechť* $\mathbf{S}$ *je úplná suficientní statistika. Nechť statistika* $W = g(\mathbf{S})$ *je nestranný odhad parametrické funkce* $a(\boldsymbol{\theta})$*, kde* $\mathsf{E}\, W^2 < \infty$ *pro každé* $\boldsymbol{\theta} \in \Omega$*. Pak* W *je nejlepší nestranný odhad* $a(\boldsymbol{\theta})$*, a to jediný.*

Důkaz. Jednotlivé kroky důkazu jsou popsány v práci Anděl (2007, str. 136). ❑

Nyní pomocí Věty 5 chceme ukázat, že odhad výběrovým rozptylem $\bar{X}_n$ je nejlepším nestranným odhadem parametru λ v modelu Poissonova rozdělení. Daný příklad byl detailně popsán v práci Anděl (2007, str.138). Pro úplnost a lepší náhled na problematiku nejlepšího nestranného odhadu přebíráme vypracovaný příklad z výše uvedeného zdroje.

Uvažujeme náhodný výběr $\mathbf{X} = (X_1, \dots, X_n)^T$ z $\text{Pois}(\lambda)$, kde $\lambda > 0$ je neznámý parametr. Označme

$$\mathbf{S} = \sum_{i=1}^{n} X_i$$

a platí $\mathbf{S} \sim \text{Pois}(n\lambda)$. Nejprve chceme ukázat, že statistika $\mathbf{S}$ je *suficientní.*

Nechť $x_1, \dots, x_n$ jsou nezáporná celá čísla. Je-li $\sum_{i=1}^{n} x_i \neq s$, pak

$$P[X_1 = x_1, \dots, X_n = x_n | S = s] = 0.$$

Nechť potom $\sum_{i=1}^{n} x_i = s$. Odtud

$$P[X_1 = x_1, \ldots, X_n = x_n | S = s] =$$

$$= \frac{P[X_1 = x_1, \ldots, X_n = x_n]}{P[S = s]} = \frac{\prod_{i=1}^{n} P[X_i = x_i]}{P[S = s]} =$$

$$= \frac{\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}}{\frac{(n\lambda)^s e^{-n\lambda}}{s!}} = \frac{s!}{x_1! \ldots x_n!} \left(\frac{1}{n}\right)^s.$$

Výsledek nezávisí na λ, takže statistika $\mathbf{S}$ je *suficientní.*

Nyní budeme řešit druhou část úlohy, tj. dokážeme, že statistika $\mathbf{S}$ je *úplná.* Nechť $\mathsf{E}\ w(\mathbf{S}) = 0$ pro všechna $\lambda > 0$. Tudíž

$$\sum_{s=0}^{\infty} w(s) \frac{(n\lambda)^s}{s!} e^{-n\lambda} = 0,$$

takže

$$\sum_{s=0}^{\infty} \frac{w(s) n^s}{s!} \lambda^s = 0.$$

Mocninná řada uvedená v posledním vzorci je rovná nule pro všechna $\lambda > 0$, a tak všechny její koeficienty musí být rovny nule. Proto $w(s) = 0$ pro $s = 0,1, \ldots,$ tedy $w(\mathbf{S}) = 0$ skoro jistě, a tudíž statistika $\mathbf{S}$ je *úplná.* Tím je převzatý příklad dokončen.

Pro $\bar{X}_n = \mathbf{S}/n$ platí $\mathsf{E}\ \bar{X}_n = \lambda$, takže podle Lehmann-Scheffého věty (viz. Věta 5) je $\bar{X}_n$ *nejlepší nestranný* odhad parametru λ.

Tím jsme podtvrdili správnost výsledku pro váhový vektor $\hat{\boldsymbol{w}}^*_{ma\underline{x}}$ a zdůvodnili jsme, proč celá váha odhadovaného parametru λ padne na odhad $\bar{X}_n$.

3. Simulace

V této kapitole se budeme zaobírat simulací náhodných výběrů pro normální a Poissonovo rozdělení. Poté odhadneme parametry těchto rozdělení pomocí předem dané posloupnosti odhadů a následně pomocí metody kombinování odhadů. Budeme uvažovat různé hodnoty odhadovaných parametrů a různé rozsahy výběru. Na závěr uvedeme tabulky s jednotlivými výsledky, které pak vhodným způsobem interpretujeme.

3.1 Simulace pro model normálního rozdělení

Uvažujme model normálního rozdělení, tj. $\mathcal{N}(0, \sigma^2)$. Budeme generovat K náhodných výběrů z daného rozdělení o pevně daném rozsahu výběru n, tedy

$$\begin{gathered} X_1^1, \ldots, X_n^1, \\ X_1^2, \ldots, X_n^2, \\ \ldots \\ X_1^K, \ldots, X_n^K. \end{gathered}$$

Pro každý z jednotlivých náhodných výběrů odhadneme parametr rozptylu σ^2 postupně pomocí dvou počátečních odhadů uvedených v podkapitole 2.2, konkrétně (2.1) a (2.2). Položíme-li $\theta = \sigma^2$, pak můžeme definovat

$$\hat{\theta}_{n,1}^i = S_{n,i}^2 \equiv \frac{1}{n-1} \sum_{j=1}^{n} (X_j^i - \bar{X}_n^i)^2, \tag{3.1}$$

$$\hat{\theta}_{n,2}^i = \hat{\sigma}_{n,i}^2 \equiv \frac{1}{n} \sum_{j=1}^{n} (X_j^i - \bar{X}_n^i)^2, \tag{3.2}$$

pro každé $i \in \{1, \ldots, K\}$. V tomto případě $\bar{X}_n^i$ označuje výběrový průměr náhodného výběru $X_1^i, \ldots, X_n^i$, tzn.

$$\bar{X}_n^i = \frac{1}{n} \sum_{j=1}^{n} X_j^i.$$

Dostáváme tak systém dvojic počátečních odhadů parametru θ

$$\begin{gathered} \hat{\theta}_{n,1}^1, \quad \hat{\theta}_{n,2}^1, \\ \hat{\theta}_{n,1}^2, \quad \hat{\theta}_{n,2}^2, \\ \ldots \\ \hat{\theta}_{n,1}^K, \quad \hat{\theta}_{n,2}^K, \end{gathered}$$

pro nějaké pevně dané n. Pro každý náhodný výběr platí, že výsledný váhový vektor má následující tvar

$$\boldsymbol{\lambda}_{max}^* = \begin{pmatrix} \frac{1-n}{1+n} \\ \frac{2n}{1+n} \end{pmatrix}. \tag{3.3}$$

Počáteční odhady parametru rozptylu (viz. 3.1, 3.2) a vypočítaný váhový vektor (viz. 3.3) nám umožňují konstruovat výsledný odhad pomocí metody kombinování odhadů. Pro každé $i \in \{1, \ldots, K\}$ tedy položíme

$$\hat{\theta}^i_{n,comb} \equiv (\boldsymbol{\lambda}^*_{max})^T \begin{pmatrix} \hat{\theta}^i_{n,1} \\ \hat{\theta}^i_{n,2} \end{pmatrix} = \left(\frac{n-1}{n+1}\right)\hat{\theta}^i_{n,1}.$$

Takovým způsobem získáváme posloupnost výsledných kombinovaných odhadů

$$\begin{gathered} \hat{\theta}^1_{n,comb}, \\ \hat{\theta}^2_{n,comb}, \\ \ldots \\ \hat{\theta}^K_{n,comb}. \end{gathered}$$

V dané situaci budeme hodnotit odhady pomocí relativní střední čtvercové chyby a relativního vychýlení, které jsou definované následovně

$$\widehat{\mathrm{rMSE}}(\hat{\theta}_{n,l}) = \frac{1}{K}\sum_{i=1}^{K} \frac{\left(\hat{\theta}^i_{n,l} - \theta\right)^2}{\theta^2}, \tag{3.4}$$

$$\widehat{\mathrm{rBias}}(\hat{\theta}_{n,l}) = \frac{1}{K}\sum_{i=1}^{K} \frac{\left(\hat{\theta}^i_{n,l} - \theta\right)}{\theta}, \tag{3.5}$$

kde K je počet simulací. Takovým způsobem jsme schopní spočítat příslušné charakteristiky pro jednotlivé počáteční odhady a také pro výsledný kombinovaný odhad. V tomto případě $l \in \{1, 2, comb\}$ a platí $\theta = \sigma^2$. Například pro odhad $\hat{\theta}_{n,1}$ (viz. (3.1)) dostáváme:

$$\widehat{\mathrm{rMSE}}(\hat{\theta}_{n,1}) = \frac{1}{K}\sum_{i=1}^{K} \frac{\left(\hat{\theta}^i_{n,l} - \sigma^2\right)^2}{(\sigma^2)^2},$$

$$\widehat{\mathrm{rBias}}(\hat{\theta}_{n,1}) = \frac{1}{K}\sum_{i=1}^{K} \frac{\left(\hat{\theta}^i_{n,l} - \sigma^2\right)}{\sigma^2},$$

Zde platí, že čím je menší hodnota relativní střední čtvercové chyby, tím je příslušný odhad bližší k odhadovanému parametru, tedy je přesnější.

Nadále předpokládejme, že počet simulací je vždy konstantní, konkrétně $K = 1000$. Budeme chtít zkoumat chování jednotlivých odhadů

$$\begin{aligned} \hat{\theta}_{n,1} &= S^2_n, \\ \hat{\theta}_{n,2} &= \hat{\sigma}^2_n, \\ \hat{\theta}_{n,comb} &= \left(\frac{n-1}{n+1}\right)S^2_n \end{aligned}$$

ve smyslu odhadů relativní střední čtvecové chyby a relativního vychýlení (tj. $\widehat{\mathrm{rMSE}}$ a $\widehat{\mathrm{rBias}}$). Pro úplnost výsledků uvedeme také hodnoty odhadů příslušných rozptylů, definovaných následujícím způsobem

$$\widehat{\mathrm{rvar}}(\hat{\theta}_{n,l}) = \widehat{\mathrm{rMSE}}(\hat{\theta}_{n,l}) - (\widehat{\mathrm{Bias}}(\hat{\theta}_{n,l}))^2. \tag{3.6}$$

$\theta = \sigma^2 = 1$									
	$\hat{\theta}_{\mathbf{n,1}}$			$\hat{\theta}_{\mathbf{n,2}}$			$\hat{\theta}_{\mathbf{n,comb}}$		
	Bias[1]	**Var**[2]	**MSE**[3]	**Bias**[1]	**Var**[2]	**MSE**[3]	**Bias**[1]	**Var**[2]	**MSE**[3]
n=10	−0,028	0,220	0,221	−0,125	0,178	0,194	−0,204	0,147	0,189
n=25	−0,008	0,085	0,085	−0,048	0,078	0,080	−0,084	0,072	0,079
n=50	0,002	0,045	0,045	−0,018	0,043	0,043	−0,038	0,041	0,043
n=100	−0,001	0,021	0,021	−0,011	0,021	0,021	−0,021	0,020	0,021
n=250	−0,002	0,008	0,008	−0,006	0,008	0,008	−0,010	0,008	0,008
n=500	0,001	0,004	0,004	−0,006	0,004	0,004	−0,003	0,004	0,004
n=1000	0,001	0,002	0,002	0,000	0,002	0,002	−0,001	0,002	0,002

Pozn: [1] Odhad rBias (viz. (3.5)), [2] Odhad rvar (viz. (3.6)), [3] Odhad rMSE (viz. (3.4)).

Tabulka 3.1: Výsledky simulace jednotlivých odhadů v modelu $\mathcal{N}(0,1)$, uvedené pro různé rozsahy výběru n.

Výsledky simulace pro volbu parametru $\sigma^2 = 1$ může čtenář najít ve výše uvedené tabulce (viz. Tabulka 3.1).

Ukázalo se, že hodnoty relativní střední čtvercové chyby a relativního vychýlení nezávisí na hodnotě parametru σ^2. Použití obecného rozpylu σ^2 místo 1 odpovídá přenásobení pozorovaných hodnot konstantou σ, která se pak při výpočtech relativní střední čtvercové chyby a relativního vychýlení vykrátí. Při použití stejného *seed* v simulacích by tedy výsledky vyšly shodné, proto je neuvádíme.

V každém případě platí, že hodnota relativní střední čtvercové chyby pro kombinovaný odhad $\hat{\theta}_{n,comb}$ je vždy menší než hodnoty relativních středních čtvercových chyb pro dva počáteční odhady (tj. $\hat{\theta}_{n,1}$ a $\hat{\theta}_{n,2}$). Platí to pro různé rozsahy výběrů $n \in \{10, 25, 50, 100, 250, 500, 1000\}$.

Graficky znázorníme teoretickou hodnotu relativní střední čtvercové chyby jednotlivých odhadů jako funkce rozsahu výběru n (viz. Obrázek A.1 v Příloze A.1). Také označíme na grafech hodnoty rMSE pro jednotlivé rozsahy výběru n, které byly uvedené v Tabulce 3.1.

Tím jsme potvrdili fakt, že výsledný kombinovaný odhad $\hat{\theta}_{n,comb}$, vypočítaný pomocí metody kombinování odhadů, je v uvažovaném modelu nejlepším odhadem parametru σ^2 mezi lineárními kombinacemi počátečních odhadů, s podmínkou že součet váh je roven jedna.

3.2 Simulace pro model Poissonova rozdělení

Podobně jako v předchozí kapitole budeme generovat K náhodných výběrů tentokrát z Poissonova rozdělení Pois(λ), kde $\lambda > 0$. Tedy uvažujme

$$X_1^1, \ldots, X_n^1,$$
$$X_1^2, \ldots, X_n^2,$$
$$\ldots$$
$$X_1^K, \ldots, X_n^K,$$

kde $n \geq 2$ je předem daný rozsah náhodného výběru. Pro každý z jednotlivých náhodných výběrů obdobně odhadneme parametr λ pomocí dvou počátečních odhadů, které jsme uvedli v podkapitole 2.3. Tentokrát jsou počátečními odhady výběrový průměr a výběrový rozptyl. Položíme-li tedy $\theta = \lambda$, potom

$$\hat{\theta}^i_{n,1} = \bar{X}^i_n \equiv \frac{1}{n}\sum_{j=1}^{n} X^i_j, \tag{3.7}$$

$$\hat{\theta}^i_{n,2} = S^2_{n,i} \equiv \frac{1}{n-1}\sum_{j=1}^{n}(X^i_j - \bar{X}^i_n)^2, \tag{3.8}$$

pro každé $i \in \{1,\ldots,K\}$. Posloupnost všech počátečních odhadů paramertu λ pro jednotlivé náhodné výběry je pak tvaru

$$\begin{gathered} \hat{\theta}^1_{n,1},\quad \hat{\theta}^1_{n,2}, \\ \hat{\theta}^2_{n,1},\quad \hat{\theta}^2_{n,2}, \\ \cdots \\ \hat{\theta}^K_{n,1},\quad \hat{\theta}^K_{n,2}. \end{gathered}$$

V kapitole 2.3 jsme odvodili tvar váhového vektoru pro výše uvedené počáteční odhady, konkrétně

$$\boldsymbol{w}^*_{max} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Pro každé $i \in \{1,\ldots,K\}$ poté položíme

$$\hat{\theta}^i_{n,comb} \equiv (\boldsymbol{w}^*_{max})^T \begin{pmatrix} \hat{\theta}^i_{n,1} \\ \hat{\theta}^i_{n,2} \end{pmatrix} = \hat{\theta}^i_{n,1}.$$

Výsledné kombinované odhady náhodných výběrů jsou

$$\begin{gathered} \hat{\theta}^1_{n,comb} = \hat{\theta}^1_{n,1}, \\ \hat{\theta}^2_{n,comb} = \hat{\theta}^2_{n,1}, \\ \cdots \\ \hat{\theta}^K_{n,comb} = \hat{\theta}^K_{n,1}. \end{gathered}$$

Takovým způsobem jsme schopni spočítat relativní střední kvadratické chyby a relativní vychýlení pro jednotlivé případy. Například pro výběrový průměr, a tedy i pro výsledný kombinovaný odhad dostáváme

$$\widehat{\text{rMSE}}(\hat{\theta}_{n,comb}) = \widehat{\text{rMSE}}(\hat{\theta}_{n,1}) = \frac{1}{K}\sum_{i=1}^{K}\frac{\left(\hat{\theta}^i_{n,1} - \lambda\right)^2}{\lambda^2},$$

$$\widehat{\text{rBias}}(\hat{\theta}_{n,comb}) = \widehat{\text{rBias}}(\hat{\theta}_{n,1}) = \frac{1}{K}\sum_{i=1}^{K}\frac{\left(\hat{\theta}^i_{n,1} - \lambda\right)}{\lambda}.$$

Nyní budeme chtít simulovat K náhodných výběrů pro různé volby parametru λ a rozsahu výběrů n. Následně výsledky simulace interpretujeme vhodným způsobem.

Uvažujme nadále 1000 různých simulací náhodného výběru z Poissonova rozdělení, kde parametr λ může nabývat hodnot z množiny $\{1, 5, 20\}$. Vypočítané relativní střední kvadratické chyby a relativní vychýlení pro jednotlivé odhady najdeme v tabulce (viz. Tabulka 3.2). Již nebudeme uvádět hodnoty rozptylů pro jednotlivé odhady, protože příslušné hodnoty relativního vychýlení vycházejí velmi malé. Ve výsledku bychom pro rozptyly dostali skoro stejné hodnoty jako pro relativní střední kvadratické chyby (plyne to ze vzorce (3.6)).

$\theta = \lambda = 1$				
	$\hat{\theta}_{\mathbf{n,1}} = \hat{\theta}_{\mathbf{n,comb}}$		$\hat{\theta}_{\mathbf{n,2}}$	
Rozsah výběru	**Bias**[1]	**MSE**[2]	**Bias**[1]	**MSE**[2]
n=10	0,0010	0,1000	−0,0200	0,2983
n=25	0,0122	0,0424	−0,0067	0,1247
n=50	0,0019	0,0210	−0,0036	0,0591
n=100	0,0015	0,0106	0,0005	0,0302
n=250	0,0008	0,0042	0,0006	0,0115
n=500	0,0012	0,0020	0,0038	0,0059
n=1000	0,0001	0,0009	0,0038	0,0031
$\theta = \lambda = 5$				
	$\hat{\theta}_{\mathbf{n,1}} = \hat{\theta}_{\mathbf{n,comb}}$		$\hat{\theta}_{\mathbf{n,2}}$	
Rozsah výběru	**Bias**[1]	**MSE**[2]	**Bias**[1]	**MSE**[2]
n=10	−0,0006	0,0194	−0,0174	0,2284
n=25	0,0047	0,0083	−0,0092	0,0920
n=50	0,0006	0,0042	−0,0030	0,0428
n=100	0,0007	0,0021	0,0028	0,0222
n=250	0,0005	0,0008	0,0002	0,0086
n=500	0,0006	0,0004	0,0024	0,0043
n=1000	−0,0001	0,0002	0,0031	0,0022
$\theta = \lambda = 20$				
	$\hat{\theta}_{\mathbf{n,1}} = \hat{\theta}_{\mathbf{n,comb}}$		$\hat{\theta}_{\mathbf{n,2}}$	
Rozsah výběru	**Bias**[1]	**MSE**[2]	**Bias**[1]	**MSE**[2]
n=10	−0,0008	0,0050	−0,0136	0,2286
n=25	−0,0000	0,0021	0,0044	0,0863
n=50	0,0008	0,0010	0,0061	0,0417
n=100	0,0007	0,0005	0,0007	0,0110
n=250	0,0006	0,0002	0,0021	0,0083
n=500	0,0003	0,0001	0,0007	0,0041
n=1000	0,0002	0,0001	0,0014	0,0022

Pozn: [1] Odhad relativního vychýlení, [2] Odhad relativní střední čtvercové chyby.

Tabulka 3.2: Výsledky simulace jednotlivých odhadů v modelu Pois(λ) pro různé rozsahy výběru n.

V každém případě platí, že hodnota relativní střední čtvercové chyby pro kombinovaný odhad $\hat{\theta}_{n,comb}$, tedy i pro odhad výběrovým průměrem $\hat{\theta}_{n,1}$, je vždy menší než hodnota relativní střední čtvercové chyby pro odhad výběrovým rozptylem $\hat{\theta}_{n,2}$.

Můžeme si všimnout, že s rostoucí hodnotou odhadovaného parametru λ hodnota $\widehat{\text{rMSE}}$ pro kombinovaný (tedy i odhad výběrovým průměrem) výrazně klesá. Odpovídá to tvaru $\text{rMSE}(\hat{\theta}_1)$:

$$\text{rMSE}(\hat{\theta}_{n,1}) = \frac{\text{MSE}(\hat{\theta}_{n,1})}{\lambda^2},$$

kde

$$\text{MSE}(\hat{\theta}_{n,1}) = \frac{\lambda}{n}$$

je prvkem matice Σ na místě $[1, 1]$ (viz. (2.5)). Z toho plyne

$$\text{rMSE}(\hat{\theta}_{n,1}) = \frac{1}{\lambda n}.$$

Pro lepší přehled graficky znázorníme teoretickou hodnotu relativní střední čtvercové chyby kombinovaného odhadu $\text{rMSE}(\hat{\theta}_{n,comb})$ jako funkce rozsahu výběru n pro různé volby parametru λ (viz. Obrázek A.2 v Příloze A.2). Navíc uvedeme na grafech jednotlivé hodnoty $\text{rMSE}(\hat{\theta}_{n,comb})$ pro rozsahy výběru n, které byly zmíněné v Tabulce 3.2.

Výsledky simulace pro Poissonovo rozdělení s parametrem λ a se dříve uvedenými počátečními odhady parametru λ potvrdily, že v dané situaci výsledný kombinovaný odhad $\hat{\theta}_{n,comb} = \hat{\theta}_{n,1}$ nejlépe odhaduje parametr λ ve smyslu střední kvadratické chyby.

Závěr

Daná práce je prvním materiálem v českém jazyce týkající se problematiky metody kombinování odhadů stejného parametru.

Za hlavní příspěvek této práce považujeme zahrnutí podrobného odvození váhového vektoru pomocí matice kvadratické chyby pro jednorozměrný odhadovaný parametr. Výše uvedené odvození, dokud je nám známo, nebylo detailně vypracováno v žádném ze zdrojů použité literatury.

Dalším přínosem do problematiky kombinování odhadů by mohla být praktická aplikace metody na příkladech dvou pravděpodobnostních rozdělení. V práci se uvádí příklad normálního rozdělení s nulovou střední hodnotou a příklad Poissonova rozdělení. Zde bylo ilustrováno použití metody kombinování předem zvolených odhadů pro získání lepšího výsledného odhadu parametru ve smyslu střední kvadratické chyby. Ukázalo se, že výsledek pro kombinovaný odhad Poissonova rozdělení má zajímavé vlastnosti, které byly v práci podpořeny teorií o nestranném odhadu s minimálním rozptylem.

Simulační studie dříve zmíněných pravděpodobnostních rozdělení ukázala, že výsledný kombinovaný odhad, dosahnutý metodou kombinování odhadů, dává skutečně nejlepší možný odhad parametru. V tomto rámci byl za nejlepší možný odhad považován ten, který minimalizuje střední kvadratickou chybu.

Seznam použité literatury

Anděl, J. (2007). *Základy matematické statistiky.* Druhé opravené vydání. Matfyzpress, Praha. ISBN 80-7378-001-1.

Bates, J. a Granger, C. (1969). The combination of forecasts. *Operational Research Society*, **20(4)**, 451–468.

Dahiya, R. a Gurland, J. (1969). Functions of the sample mean and sample variance of a poisson variate. *International Biometric Society*, **25**, 171–173.

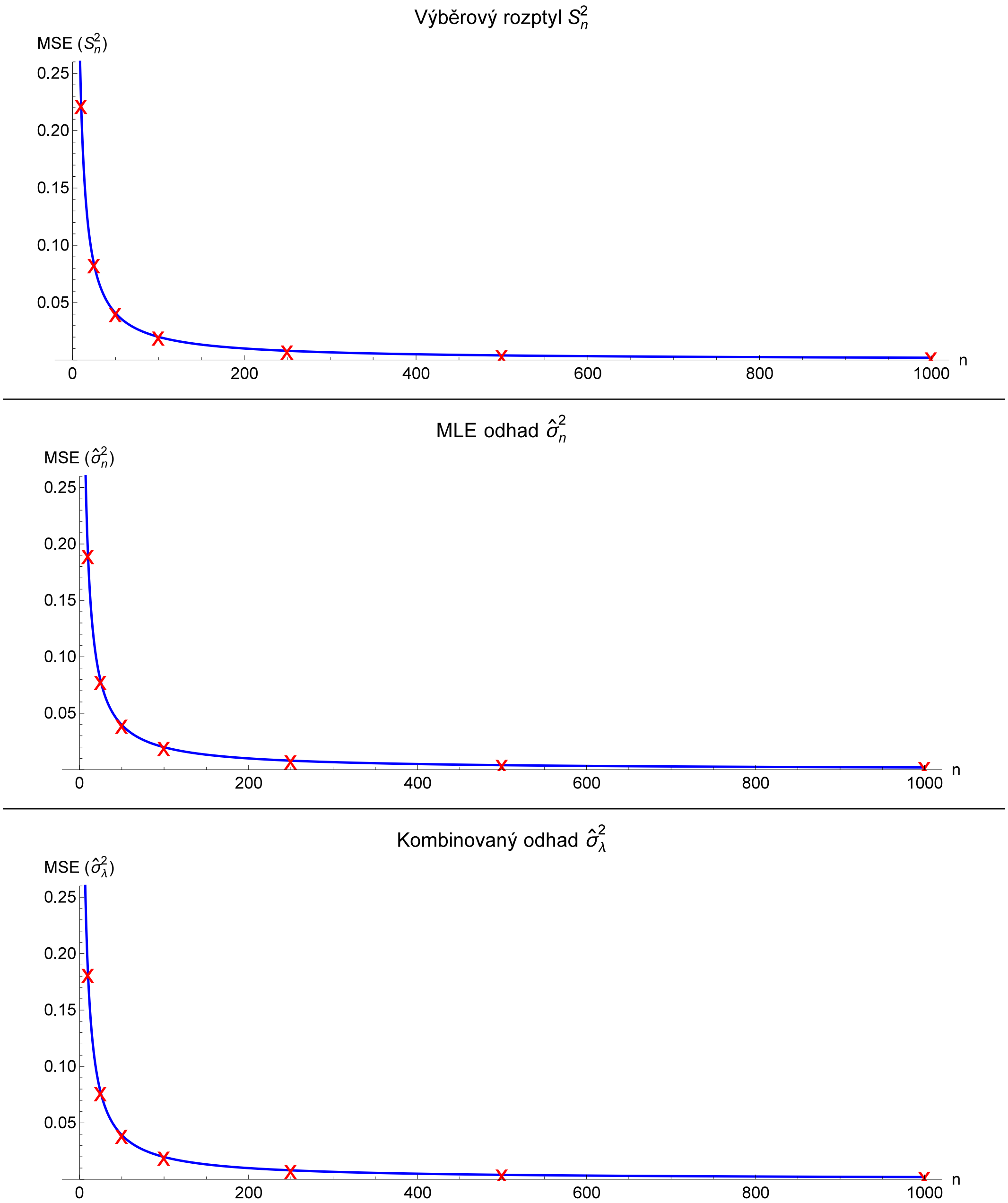
Gentle, J. E. (1943). *Matrix Algebra: Theory, Computations, and Applications in Statistics.* 2007th Edition. Springer Texts in Statistics. ISBN 978-0387708720.

Lavancier, F. a Rochet, P. (2016). A general procedure to combine estimators. *Computational Statistics and Data Analysis*, **94**, 175–192.

Timmermann, A. (2006). Forecast combinations. *Handbook of Economic Forecasting*, pages 135–196.

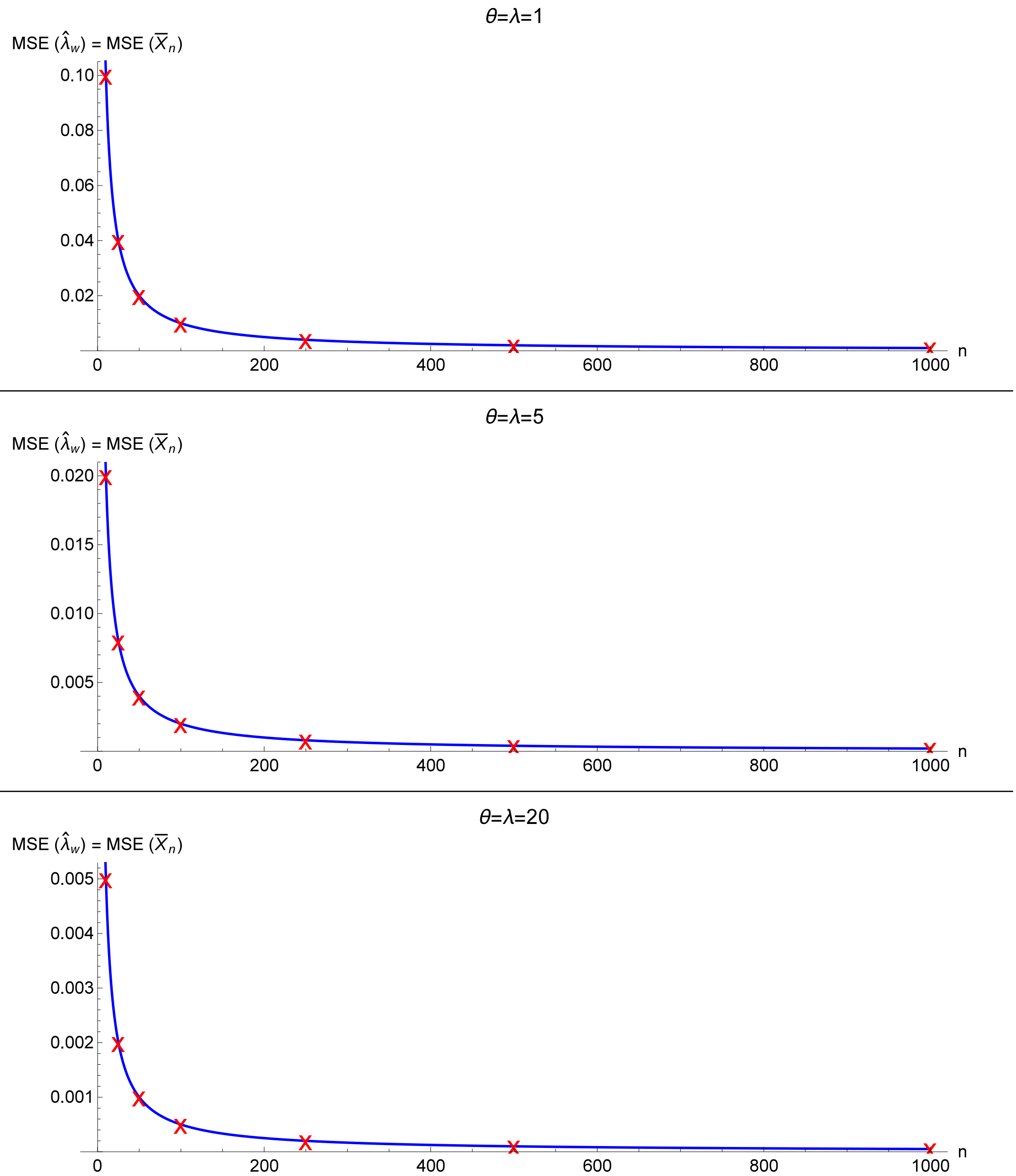
A. Přílohy

A.1 První příloha



Obrázek A.1: Grafické znázornění teoretických hodnot rMSE pro jednotlivé odhady S_n^2, $\hat{\sigma}_n^2$ a $\hat{\sigma}_\lambda^2$: červenými kříži jsou zde označené příslušné hodnoty rMSE pro uvažované rozsahy výběrů $n \in \{10, 25, 50, 100, 250, 500, 1000\}$.

A.2 Druhá příloha



Obrázek A.2: Grafické znázornění teoretických hodnot rMSE($\hat{\theta}_{n,comb}$) pro různé volby λ v závislosti na rozsahu výběru n: červené kříže označují jednotlivé hodnoty rMSE($\hat{\theta}_{n,comb}$) pro uvažované rozsahy výběru $n \in \{10, 25, 50, 100, 250, 500, 1000\}$.