



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Marcel Vagner

Prípustnosť a neprípustnosť odhadu

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Matúš Maciak, Ph.D.

Studijní program: Finanční matematika
(B0542A170001)

Studijní obor: MFMP (0542RA170001)

Praha 2023

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

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

Rád by som sa poďakoval profesorovi RNDr. Matúšovi Maciakovi, Ph.D. za konštruktívne pripomienky a cenné rady.

Název práce: Prípustnosť a neprípustnosť odhadu

Autor: Marcel Vagner

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Matúš Maciak, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Kvalita odhadu parametra sa často posudzuje pomocou strednej kvadratickej odchýlky (MSE). Pre jednorozmerný parameter je odhad konštruovaný pomocou metódy najmenších štvorcov najlepší. Pri odhadovaní viac ako dvojrozmerného parametra sa však odhad stane neprípustný, teda vždy existuje vzhľadom k MSE iný odhad, ktorý bude lepší ako MSE odhad bez ohľadu na hodnotu parametra. Tento záver je známy ako James Stein paradox. V rámci tejto bakalárskej práce najprv zadefinujeme prípustnosť a neprípustnosť odhadu, definujeme James-Steinov odhad a porovnáme chovania rôznych odhadov.

Klíčová slova: Prípustnosť, stredná kvadratická odchýlka, Steinov odhad, odhad parametra

Title: Admissibility and Inadmissibility of an Estimate

Author: Marcel Vagner

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Matúš Maciak, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The quality of a parameter estimate is usually assessed using the mean squared error (MSE). For one dimensional parameter, the estimate constructed using the least squares method is the best. However, for a vector parameter with more than two dimensions this estimator becomes inadmissible. There is always some different estimator which dominates the least squares estimate regardless of the parameter value. This phenomenon is well known as the Stein Paradox. The aim of this bachelor thesis is to describe admissibility and inadmissibility of an estimator, define the James-Stein estimator and perform a simulation study to compare different estimators.

Keywords: Admissibility, Stein's estimate, mean squared error, parameter estimate

Obsah

Úvod	2
1 Odhad	3
1.1 Základné pojmy	3
1.2 James-Steinov odhad	5
2 Dvojmerný prípad	8
2.1 Dvojmerný parameter	8
3 Neprípustnosť VP odhadu	11
3.1 Vlastnosti James-Steinovho odhadu	11
3.2 Dôkaz neprípustnosti	11
4 Simulácia	14
4.1 Generovanie dát	14
4.2 Chovanie rizikovej funkcie	14
Záver	18
Zoznam použitej literatúry	19
A Prílohy	20
A.1 Simulačný kód	20
A.2 Numerické výstupy simulácií	23

Úvod

James-Steinov paradox patrí k zaujímavým zisteniam minulého storočia v oblasti štatistiky. Pri odhade strednej hodnoty normálneho rozdelenia je v prípade jediného dostupného pozorovania intuitívne a matematicky opodstatnené použiť ako odhad práve dané pozorovanie. Jedná sa totiž o odhad strednej hodnoty výberovým priemerom, čo je neustranný a konzistentný odhad strednej hodnoty. V prípade normálneho rozdelenia je to zároveň odhad získaný metódou maximálnej vierohodnosti a spomedzi všetkých možných odhadov je najvhodnejší z hľadiska strednej kvadratickej odchýlky (MSE). Ak pozorujeme realizáciu z dvojrozmerného normálneho rozdelenia, je takto zvolený odhad taktiež najlepší. Pokiaľ však máme k dispozícii jedno pozorovanie pochádzajúce z viacrozmerného normálneho rozdelenia s dimenziou vektora stredných hodnôt väčšou ako dva, prestáva byť z hľadiska MSE tento odhad ideálny, stáva sa tzv. neprípustným odhadom. Existuje totiž iný odhad, ktorý pre ľubovoľnú skutočnú hodnotu vektora stredných hodnôt dosahuje nižšie hodnoty strednej kvadratickej odchýlky. Tento prekvapujúci záver sa nazýva James-Steinov paradox.

Cieľom tejto bakalárskej práce je popísať prípustnosť a neprípustnosť odhadu, definovať James-Steinov odhad a vyšetriť chovanie rôznych odhadov pomocou numerických simulácií. Práca je členená na štyri kapitoly. V prvej kapitole zavedieme potrebné pojmy. Definujeme mimo iné odhad, prípustnosť, respektíve neprípustnosť odhadu a strednú kvadratickú odchýlku. Na záver prvej kapitoly predstavíme James-Steinov odhad a uvedieme jeho vybrané vlastnosti. V druhej kapitole analyticky dokážeme prípustnosť výberového priemeru z jedného pozorovania pre dimenziu vektora stredných hodnôt menšiu ako tri. V tretej kapitole diskutujeme detaily James-Steinovho odhadu a dokážeme, že dosahuje nižšie hodnoty strednej kvadratickej odchýlky než výberový priemer. Štvrtá a zároveň posledná kapitola tejto práce obsahuje numerické simulácie vykonané v rámci vypracovania bakalárskej práce. Na náhodne generovaných dátových súboroch budeme skúmať chovanie stredných kvadratických odchýlok uvažovaných odhadov a porovnáme ich správanie pri zmene skutočných hodnôt vektorov stredných hodnôt a ich dimenzií. Numericky získané hodnoty spracujeme vo forme grafov a budeme diskutovať zhodu s teoretickými vlastnosťami.

1. Odhad

1.1 Základné pojmy

V praxi často vyvstáva potreba odhadnúť teoretické vlastnosti pozorovaných hodnôt. V prípade, že chápeme pozorované hodnoty ako realizácie nezávislých rovnako rozdelených náhodných veličín, alebo náhodných vektorov, pracujeme s náhodným výberom. Definície v tejto sekcii spolu so sprievodným komentárom sme čerpali z (Kulich, 2014).

Definícia 1. Postupnosť $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$ nezávislých rovnako rozdelených náhodných vektorov s distribučnou funkciou \mathbf{F} , nazývame *náhodný výber z rozdelenia \mathbf{F}* . Konštantu $p \in \mathbb{N}$ nazývame *rozsah výberu*.

Prvky náhodného výberu môžeme označovať aj ako pozorovania alebo dáta. Distribučnú funkciu \mathbf{F} nepoznáme. Často nás však zaujíma vybraná vlastnosť daného rozdelenia, ako napríklad stredná hodnota alebo rozptyl. Na to, aby sme sa o danej vlastnosti niečo bližšie dozvedeli, využívame pozorované hodnoty a predpoklad, že daná distribučná funkcia \mathbf{F} spadá do predom zvolenej množiny rozdelení \mathcal{F} , ktorú voláme *model*.

Definícia 2. *Model* pre náhodný výber $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$ je predom stanovená množina rozdelení \mathcal{F} , o ktorej predpokladáme, že do nej spadá aj neznáme rozdelenie \mathbf{F} .

Vlastnosti rozdelenia \mathbf{F} , ktoré nás zaujímajú, nazývame *parametre*. Jedná sa o nejakú konštantu, prípadne vektor konštant $\boldsymbol{\xi} \in \mathbb{R}^k$, ktorú by sme vedeli zistiť, ak by sme poznali \mathbf{F} . Parameter teda vo všeobecnosti môžeme zapísať ako $\boldsymbol{\xi} \equiv t(\mathbf{F})$, kde t je nejaký funkcionál. Množinu všetkých prípustných hodnôt parametra $\boldsymbol{\xi} \equiv t(\mathbf{F}) : \mathbf{F} \in \mathcal{F}$ značíme $\Theta \subseteq \mathbb{R}^k$. Konkrétnu hodnotu hľadaného parametra potrebujeme odhadnúť z pozorovaných dát.

Definícia 3. *Odhadom parametra $\boldsymbol{\xi}$* rozumieme merateľnú funkciu náhodného výberu $\hat{\boldsymbol{\xi}} \equiv \mathbf{T}_p(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)$.

Nakoľko je odhad parametra merateľnou funkciou pozorovaných dát, je taktiež náhodnou veličinou.

Definícia 4. Odhad $\hat{\boldsymbol{\xi}} \equiv \mathbf{T}_p(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)$ nazývame *nestranný odhad* parametra $\boldsymbol{\xi}$ ak platí $\mathbb{E}\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}$ pre všetky $\boldsymbol{\xi} \in \Theta$.

Nestranné odhady nazývame aj nevychýlené. Odhady s konečnou strednou hodnotou, ktoré nie sú nestranné nazývame vychýlené.

Definícia 5. Nech odhad $\hat{\boldsymbol{\xi}} \equiv \mathbf{T}_p(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)$ parametra $\boldsymbol{\xi}$ má konečný vektor stredných hodnôt. Rozdiel $\mathbb{E}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})$ nazývame *vychýlenie* odhadu $\hat{\boldsymbol{\xi}}$.

Odhady ktoré sú nestranné majú teda nulové vychýlenie. V rámci tejto práce sa zameriame na odhady strednej hodnoty. Jedným z najbežnejších je *výberový priemer*.

Definícia 6. Veličina $\bar{\mathbf{X}}_p = \frac{1}{p} \sum_{i=1}^p \mathbf{X}_i$ sa nazýva výberový priemer náhodného výberu $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$.

Výberový priemer je nestranný odhad vektora stredných hodnôt. Nakoľko existuje veľké množstvo možných odhadov, potrebujeme spôsob ako rôzne odhady hľadaného parametra porovnať z hľadiska kvality. Pre tento účel zadefinujeme pojem stratovej a rizikovej funkcie (Nagy, 2023).

Definícia 7. Stratová funkcia je ľubovoľné merateľné zobrazenie

$$L : \Theta \times \Theta \rightarrow [0, \infty).$$

kde Θ je množina všetkých možných hodnôt parametra.

Hodnota stratovej funkcie $L(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}})$ má svojím spôsobom odzrkadľovať, ako blízko je náš odhad skutočnej hodnote parametra. Preto v praxi volíme L tak, aby platilo

$$\begin{aligned} L(\boldsymbol{\xi}_1, \boldsymbol{\xi}_1) &= 0 \\ \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\| < \|\boldsymbol{\xi}_3 - \boldsymbol{\xi}_4\| &\implies L(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) < L(\boldsymbol{\xi}_3, \boldsymbol{\xi}_4) \end{aligned}$$

pre ľubovoľné $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3, \boldsymbol{\xi}_4 \in \Theta$. Keďže hodnota stratovej funkcie $L(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}})$ závisí od odhadu $\hat{\boldsymbol{\xi}}$, je taktiež náhodnou veličinou.

Definícia 8. Riziková funkcia odhadu $\hat{\boldsymbol{\xi}}$ parametra $\boldsymbol{\xi} \in \Theta$ je definovaná ako

$$R(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}) = \mathbb{E}_{\boldsymbol{\xi}} L(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}), \quad \boldsymbol{\xi} \in \Theta.$$

Jedná sa o strednú hodnotu stratovej funkcie $L(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}})$ konkrétneho odhadu $\hat{\boldsymbol{\xi}}$ v závislosti od skutočnej hodnoty parametra $\boldsymbol{\xi}$. Chovanie rizikovej funkcie nám teda dáva istú predstavu o tom, ako spoľahlivý je náš odhad pri zmene skutočnej hodnoty hľadaného parametra. Podobne ako pri voľbe odhadov, existuje veľké množstvo stratových, respektíve rizikových funkcií, na základe ktorých porovnávame konkrétne odhady. V tejto práci však budeme využívať jednu z najčastejšie volených rizikových funkcií, a to *strednú kvadratickú odchýlku*.

Definícia 9. Nech odhad $\hat{\boldsymbol{\xi}}$ (v zmysle definície 3) parametra $\boldsymbol{\xi} \in \Theta \subseteq \mathbb{R}^k$ má konečný rozptyl. Funkciu

$$\text{MSE}_{\hat{\boldsymbol{\xi}}} = \mathbb{E} \|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|^2$$

voláme *stredná kvadratická odchýlka* odhadu $\hat{\boldsymbol{\xi}}$.

Na základe zvolenej rizikovej funkcie, vieme posúdiť kvalitu rôznych odhadov hľadaného parametra. Cieľom je nájsť taký odhad, ktorý dosahuje spomedzi všetkých možných odhadov najmenšiu hodnotu rizikovej funkcie.

Definícia 10. Hovoríme, že odhad $\hat{\boldsymbol{\xi}}_1$ dominuje odhadu $\hat{\boldsymbol{\xi}}_2$ ak sú splnené nasledujúce podmienky

$$\begin{aligned} \forall \boldsymbol{\xi} \in \Theta : R(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}_1) &\leq R(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}_2) \\ \exists \boldsymbol{\xi} \in \Theta : R(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}_1) &< R(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}_2). \end{aligned}$$

Inak povedané, v prípade, že odhad $\hat{\xi}_1$ dominuje odhadu $\hat{\xi}_2$, odhad $\hat{\xi}_2$ stráca svojím spôsobom opodstatnenie. Vždy totiž môžeme voliť odhad $\hat{\xi}_1$, ktorý bude dosahovať rovnakých alebo lepších výsledkov. Tým sa dostávame k otázke prípustnosti odhadu.

Definícia 11. *Pokiaľ existuje ľubovoľný odhad $\hat{\xi}_1$, ktorý dominuje odhadu $\hat{\xi}_2$, teda, že platí*

$$\begin{aligned} \forall \xi \in \Theta : R(\xi, \hat{\xi}_1) &\leq R(\xi, \hat{\xi}_2) \\ \exists \xi \in \Theta : R(\xi, \hat{\xi}_1) &< R(\xi, \hat{\xi}_2). \end{aligned}$$

voláme odhad $\hat{\xi}_2$ neprípustný. V prípade, že žiadny taký odhad neexistuje, voláme odhad $\hat{\xi}_2$ prípustný.

1.2 James-Steinov odhad

V tejto práci sa zameriame na špecifický prípad, kde budeme pozorovať náhodný výber z n -rozmerného normálneho rozdelenia s vektorom stredných hodnôt $\xi = (\xi_1, \dots, \xi_n)^T$ a kovariančnou maticou $\Sigma = I_n$, to jest n -rozmernou jednotkovou maticou. Ako bolo spomenuté vyššie, jedným z najčastejšie používaných odhadov strednej hodnoty je výberový priemer. V prípade normálneho rozdelenia je to taktiež odhad získaný pomocou metódy maximálnej vierohodnosti.

Predpokladajme, že rozsah nášho náhodného výberu je $p = 1$, t.j., máme len jedno pozorovanie $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$. Ako odhad vektoru stredných hodnôt použijeme práve výberový priemer. Máme:

$$\begin{aligned} \text{Jednoprvkový náhodný výber: } \mathbf{X} &\sim \mathbf{F}, \\ \text{Model: } \mathbf{F} \in \mathcal{F} &= \{\mathcal{N}_n(\xi, \Sigma), \xi \in \mathbb{R}^n, \Sigma = I_n\}, \\ \text{Odhadovaný parameter: } \xi &= \mathbb{E}\mathbf{X} \\ \text{Odhad parametra: } \hat{\xi}_0 &= \bar{\mathbf{X}}_p \end{aligned}$$

Nakoľko máme len jedno pozorovanie, výberový priemer bude rovný práve danému pozorovaniu, teda $\hat{\xi}_0(\mathbf{X}) = \bar{\mathbf{X}}_p = \mathbf{X}$. V ďalšom budeme odhad $\hat{\xi}_0$ označovať VP odhad. Ukážeme, že v prípade, že pri porovnaní s inými odhadmi využijeme ako rizikovú funkciu strednú kvadratickú odchýlku, to jest

$$R(\xi, \hat{\xi}) = \mathbb{E}\|\xi - \hat{\xi}\|^2 = \mathbb{E}\sum_{i=1}^n (\xi_i - \hat{\xi}_i)^2,$$

zistíme, že odhad $\hat{\xi}_0(\mathbf{X}) = \mathbf{X}$ je prípustný pre $n \leq 2$, no pre $n \geq 3$ sa stáva neprípustným. Existuje totiž iný odhad, ktorý pre ľubovoľnú skutočnú hodnotu hľadaného parametra ξ dosahuje nižšiu hodnotu strednej kvadratickej odchýlky. Tento odhad sa nazýva James-Steinov odhad, viz. (James a Stein, 1992, str. 363), a má predpis:

$$\hat{\xi}_1(\mathbf{X}) = \left(1 - \frac{(n-2)}{\|\mathbf{X}\|^2}\right) \mathbf{X}$$

Fakt, že odhad $\hat{\boldsymbol{\xi}}_0(\mathbf{X}) = \mathbf{X}$ prestáva byť prípustný pre dimenziu n väčšiu ako dva sa nazýva James-Steinov paradox. Hodnota $\text{MSE}_{\hat{\boldsymbol{\xi}}_0}$ sa pri zmene skutočnej hodnoty vektoru stredných hodnôt nemení. Platí:

$$\text{MSE}_{\hat{\boldsymbol{\xi}}_0} = \mathbb{E}\|\hat{\boldsymbol{\xi}}_0 - \boldsymbol{\xi}\|^2 = n$$

pre ľubovoľnú dimenziu $n \in \mathbb{N}$ a vektor stredných hodnôt $\boldsymbol{\xi} \in \mathbb{R}^n$.

Tvrdenie 1. *Nech $\mathbf{X} = (X_1, \dots, X_n)^T \sim \mathcal{N}_n(\boldsymbol{\xi}, I_n)$ kde $X_i \sim \mathcal{N}(\xi_i, 1)$ pre všetky $i \in \{1, \dots, n\}$. Potom pre všetky $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ platí $\mathbb{E}\|\mathbf{X} - \boldsymbol{\xi}\|^2 = n$*

Dôkaz. Uvažujme náhodný vektor \mathbf{Y} definovaný ako

$$\begin{aligned} \mathbf{Y} &= (Y_1, \dots, Y_n)^T = \mathbf{X} - \boldsymbol{\xi} \\ Y_i &= X_i - \xi_i \sim \mathcal{N}(0, 1), \forall i \in \{1, \dots, n\}. \end{aligned}$$

Kedže platí:

$$\mathbb{E}Y_i^2 = (\mathbb{E}Y_i)^2 + \text{Var}Y_i = 0 + 1 = 1, \forall i \in \{1, \dots, n\},$$

tak z linearity strednej hodnoty dostávame

$$\mathbb{E}\|\mathbf{X} - \boldsymbol{\xi}\|^2 = \mathbb{E}\|\mathbf{Y}\|^2 = \mathbb{E}\sum_{i=1}^n Y_i^2 = \sum_{i=1}^n \mathbb{E}Y_i^2 = \sum_{i=1}^n 1 = n.$$

□

Najväčšie zlepšenie oproti VP odhadu, respektíve najväčší pokles v hodnote strednej kvadratickej odchýlky, dosahuje James-Steinov odhad pre vektor stredných hodnôt $\boldsymbol{\xi} = \mathbf{0}$. V prípade, že $\|\boldsymbol{\xi}\|^2 = 0$, bude pre ľubovoľnú dimenziu $n > 2$ platíť (James a Stein, 1992)

$$\text{MSE}_{\hat{\boldsymbol{\xi}}_1} = \mathbb{E}\left\|\left(1 - \frac{(n-2)}{\|\mathbf{X}\|^2}\right)\mathbf{X}\right\|^2 = 2, \quad (1.1)$$

pričom $\text{MSE}_{\hat{\boldsymbol{\xi}}_1}$ bude pre $\|\boldsymbol{\xi}\|^2 \rightarrow \infty$ konvergovať k $\text{MSE}_{\hat{\boldsymbol{\xi}}_0} = n$.

Tvrdenie 2. *Nech $\mathbf{X} = (X_1, \dots, X_n)^T \sim \mathcal{N}_n(\boldsymbol{\xi}, I_n)$ kde $X_i \sim \mathcal{N}(\xi_i, 1)$ pre všetky $i \in \{1, \dots, n\}$. Potom pre všetky $\boldsymbol{\xi} \in \mathbb{R}^n$ platí*

$$\mathbb{E}\left\|\left(1 - \frac{(n-2)}{\|\mathbf{X}\|^2}\right)\mathbf{X}\right\|^2 = 2. \quad (1.2)$$

Dôkaz.

$$\begin{aligned} \mathbb{E}\left\|\left(1 - \frac{(n-2)}{\|\mathbf{X}\|^2}\right)\mathbf{X}\right\|^2 &= \mathbb{E}\|\mathbf{X}\|^2 - \mathbb{E}[2(n-2)] + \mathbb{E}\frac{(n-2)^2}{\|\mathbf{X}\|^2} \\ &= n - 2(n-2) + (n-2)^2\mathbb{E}\frac{1}{\|\mathbf{X}\|^2} \\ &= n - 2(n-2) + (n-2)^2(n-2) \\ &= 2 \end{aligned} \quad (1.3)$$

Tretia rovnosť vyplýva z toho, že $\|\mathbf{X}\|^2$ má centrálné chí-kvadrát rozdelenie s $n \geq 3$ stupňami voľnosti, takže $1/\|\mathbf{X}\|^2$ má inverzné chí-kvadrát rozdelenie so strednou hodnotou $n - 2$. □

V prípade, že $\|\mathbf{X}\|^2 > (n - 2)$, tak James-Steinov odhad zmenší každý prvok vektora \mathbf{X} smerom k počiatku. Toto zmenšenie, z angličtiny *shrinkage effect*, je práve dôvodom, prečo zlepšenie James-Steinovho odhadu oproti VP odhadu klesá s rastúcou normou vektora stredných hodnôt $\|\boldsymbol{\xi}\|^2$.

Poznámka. James-Steinov odhad je možné zapísať v tvare

$$\hat{\boldsymbol{\xi}}_2 = \left(1 - \frac{(n-2)}{\|\mathbf{X} - \boldsymbol{\nu}\|^2}\right) (\mathbf{X} - \boldsymbol{\nu}) + \boldsymbol{\nu}$$

kde $\boldsymbol{\nu} \in \mathbb{R}^n$ je ľubovoľný vektor. Odhad $\hat{\boldsymbol{\xi}}_1$ je teda špecifický prípad odhadu $\hat{\boldsymbol{\xi}}_2$ pri voľbe $\boldsymbol{\nu} = \mathbf{0}$. Odhad $\hat{\boldsymbol{\xi}}_2$ bude dominovať odhadu $\hat{\boldsymbol{\xi}}_0$ bez ohľadu na zvolený vektor $\boldsymbol{\nu}$, no zlepšenie oproti výberovému priemeru bude najvýraznejšie práve pre vektor stredných hodnôt $\boldsymbol{\xi} = \boldsymbol{\nu}$. Podobne ako v prípade $\hat{\boldsymbol{\xi}}_1$, bude zlepšenie oproti výberovému priemeru, nepriamo úmerné vzdialenosti vektora $\boldsymbol{\xi}$ od $\boldsymbol{\nu}$. V tomto všeobecnom prípade bude odhad $\hat{\boldsymbol{\xi}}_2$ znižovať pozorovaný vektor \mathbf{X} smerom k $\boldsymbol{\nu}$.

2. Dvojrozmerný prípad

V tejto kapitole sa zameriame na prípustnosť VP odhadu $\hat{\xi}_0(\mathbf{X}) = \mathbf{X}$ v prípade, že dimenzia n vektora stredných hodnôt ξ je menšia ako tri. Dokážeme, že je prípustný pre $n = 2$, pričom prípustnosť pre $n = 1$ bude z daného dôkazu taktiež vyplývať.

2.1 Dvojrozmerný parameter

Ako bolo naznačené na začiatku kapitoly, pre rozmery parametra $n \leq 2$ je VP odhad prípustný. Zvyšok kapitoly, vrátane referencií, bude sledovať prácu Charlesa Steina, viz. (Stein, 1956). Pre účely dôkazu nižšie uvedeného tvrdenia zadefinujeme *sféricky symetrický odhad*.

Definícia 12. *Hovoríme, že odhad $\hat{\xi}$ je sféricky symetrický okolo počiatku, ak ho vieme napísať ako*

$$\hat{\xi}(\mathbf{X}) = [1 - h(\|\mathbf{X}\|^2)]\mathbf{X},$$

kde h je reálna funkcia. Ekvivalentne môžeme povedať, že odhad spĺňa

$$\hat{\xi} = g \circ \hat{\xi} \circ g^{-1}$$

pre ľubovoľnú ortogonálnu transformáciu g . Teda pre všetky x platí

$$g[\hat{\xi}(g^{-1}\mathbf{X})] = \hat{\xi}(\mathbf{X}).$$

Tvrdenie 3. *Nech \mathbf{X} n -rozmerný náhodný vektor, s vektorom stredných hodnôt $\xi \in \mathbb{R}^n$ a kovariančnou maticou $I_n \in \mathbb{R}^{n \times n}$, ktorá je n -rozmerná jednotková matica. Potom je odhad $\hat{\xi}_0(\mathbf{X}) = \mathbf{X}$ vektora stredných hodnôt prípustný pre $n \leq 2$.*

Dôkaz. Pri dokazovaní prípustnosti pre nižšie rozmery budeme postupovať obdobne ako v (Hodges a Lehmann, 1951) a (Girshick a kol., 1951). Nech $\hat{\xi}$ je ľubovoľný odhad ξ , s konečnou hodnotou rizikovej funkcie R pre ľubovoľnú hodnotu parametra. Ďalej zadefinujme funkciu b ako

$$b(\xi) = \mathbf{E}\hat{\xi}(\mathbf{X}) - \xi.$$

Označme $b_{ij}(\xi) = \frac{\partial}{\partial \xi_j} b_i(\xi)$, kde $b_i(\xi)$ je i -ta súradnica $b(\xi)$ a ξ_j je j -ta súradnica vektora ξ . Potom platí

$$R(\xi) \geq \|b(\xi)\|^2 + \sum_i \left(\sum_j \eta_{ij} [\delta_{ij} + b_{ij}(\xi)] \right)^2 \quad (2.1)$$

pre ľubovoľnú voľbu koeficientov η_{ij} , ktorá spĺňa $\sum_i \eta_{ij}^2 = 1$ pre všetky i a kde $\delta_{ij} = 1$ ak $i = j$, inak 0. Voľbou

$$\eta_{ij} = \frac{\delta_{ij} + b_{ij}(\xi)}{\sqrt{\sum_i [\delta_{ij} + b_{ij}(\xi)]^2}}$$

maximalizujeme pravú stranu nerovnosti (2.1) a dostávame

$$\begin{aligned} R(\boldsymbol{\xi}) &\geq \|b(\boldsymbol{\xi})\|^2 + \sum_i \sum_j [\delta_{ij} + b_{ij}(\boldsymbol{\xi})]^2 \\ &= \|b(\boldsymbol{\xi})\|^2 + n + 2 \sum_i b_{ii}(\boldsymbol{\xi}) + \sum_i \sum_j b_{ij}^2(\boldsymbol{\xi}). \end{aligned} \quad (2.2)$$

V prípade, že odhad $\hat{\boldsymbol{\xi}}$ je sféricky symetrický, má b tvar

$$b(\boldsymbol{\xi}) = -\phi(\|\boldsymbol{\xi}\|^2)\boldsymbol{\xi},$$

kde ϕ je reálna diferencovateľná funkcia. V tomto prípade vynechaním posledného člena dostávame z (2.2)

$$R(\boldsymbol{\xi}) \geq n + \|\boldsymbol{\xi}\|^2 \phi^2(\|\boldsymbol{\xi}\|^2) - 2n\phi(\|\boldsymbol{\xi}\|^2) - 4\|\boldsymbol{\xi}\|^2 \phi'(\|\boldsymbol{\xi}\|^2). \quad (2.3)$$

Najprv využijeme (2.3), aby sme ukázali, že pre $n = 2$ je VP odhad prípustný. Zo (Stein, 1956, str. 201) vyplýva, že ak odhad $\hat{\boldsymbol{\xi}}_0$ nie je prípustný, tak existuje sféricky symetrický odhad $\hat{\boldsymbol{\xi}}$, ktorý je vždy lepší a teda existuje funkcia ψ , ktorá nenadobúda nulovú hodnotu, pre ktorú platí

$$2 \geq R(\boldsymbol{\xi}) \geq 2 + \|\boldsymbol{\xi}\|^2 \phi^2(\|\boldsymbol{\xi}\|^2) - 4\phi(\|\boldsymbol{\xi}\|^2) - 4\|\boldsymbol{\xi}\|^2 \phi'(\|\boldsymbol{\xi}\|^2)$$

pre všetky $\|\boldsymbol{\xi}\|^2 > 0$. Ak označíme $t = \|\boldsymbol{\xi}\|^2$ a $\psi(t) = t\phi(t)$ vidíme, že platí

$$0 \geq \frac{1}{t}\psi^2(t) - 4\psi'(t) \quad (2.4)$$

pre $t > 0$. Teda ψ je neklesajúca funkcia. Ukážeme, že (2.4) implikuje, že ψ je rovné nule. Najprv uvažujme, že $\psi(t_0) < 0$ pre nejaké $t_0 > 0$. Integrovaním nerovnosti

$$\frac{\psi'(t)}{\psi^2(t)} \geq \frac{1}{4t}$$

od $t < t_0$ do t_0 dostávame

$$-\frac{1}{\psi(t_0)} + \frac{1}{\psi(t)} \geq \frac{1}{4} \log \frac{t_0}{t}. \quad (2.5)$$

Pre $t \rightarrow 0$ bude ľavá strana nerovnosti (2.5) obmedzená, no pravá strana bude divergovať do nekonečna. Teda nastane spor. V prípade, že $\psi(t_0) > 0$ pre nejaké t_0 , tak platí

$$\frac{1}{\psi(t_0)} - \frac{1}{\psi(t)} \geq \frac{1}{4} \log \frac{t}{t_0}$$

pre všetky $t > t_0$. Pre $t \rightarrow \infty$, bude ľavá strana obmedzená, no pravá strana bude divergovať do nekonečna, teda opäť nastane spor.

Využitím (2.2) ukážeme, že pre $n \geq 3$ neexistuje žiadne $c > (n-2)^2$ a $\|\boldsymbol{\xi}_0\|^2$ také, aby pre všetky $\|\boldsymbol{\xi}\|^2 \geq \|\boldsymbol{\xi}_0\|^2$ platilo

$$R(\boldsymbol{\xi}) \leq n - \frac{c}{\|\boldsymbol{\xi}\|^2}. \quad (2.6)$$

Na základe (Stein, 1956, str. 199-200) vieme, že existuje odhad, ktorého zlepšenie oproti VP odhadu je asymptotické voči $(n-2)^2/\|\xi\|^2$ pre $\|\xi\|^2 \rightarrow \infty$. Stačí teda ukázať, že nerovnosť

$$n - \frac{c}{\|\xi\|^2} \geq n + \|\xi\|^2 \phi^2(\|\xi\|^2) - 2n\phi(\|\xi\|^2) - 4\|\xi\|^2 \phi'(\|\xi\|^2), \quad (2.7)$$

ktorú dostaneme kombináciou (2.3) a (2.6) nemá riešenie, ktoré by platilo pre všetky $\|\xi\|^2 \geq \|\xi_0\|^2$. Na dôkaz tejto vlastnosti, nech

$$\phi(\|\xi\|^2) = \frac{n-2}{\|\xi\|^2} + f(\|\xi\|^2).$$

Z (2.7) tak dostávame

$$-\frac{c - (n-2)^2}{\|\xi\|^2} \geq \|\xi\|^2 f^2(\|\xi\|^2) - 4f(\|\xi\|^2) - 4\|\xi\|^2 f'(\|\xi\|^2).$$

Nech $t = \|\xi\|^2, \psi(t) = t(f(t))$. Potom platí

$$\begin{aligned} -\frac{c - (n-2)^2}{t} &\geq \frac{1}{t} \psi^2(t) - 4\psi'(t) \\ \frac{\psi'(t)}{a^2 + \psi^2(t)} &\geq \frac{1}{4t}, \end{aligned} \quad (2.8)$$

kde $a^2 = c - (n-2)^2$. Z nerovnosti (2.4), ktorá je slabšia ako nerovnosť (2.8) vidíme, že $\psi(t) \leq 0$. Tým pádom pre $t_0 < t$ dostávame

$$\tan^{-1} \frac{\psi(t)}{a} - \tan^{-1} \frac{\psi(t_0)}{a} \geq \frac{1}{4a} \log \frac{t}{t_0}. \quad (2.9)$$

Ľavá strana nerovnosti (2.9) je obmedzená, no pravá strana diverguje do nekonečna pre $t \rightarrow \infty$, čo je opäť spor. Tým je dôkaz dokončený. □

3. Neprípustnosť VP odhadu

V tejto kapitole formálne dokážeme neprípustnosť VP odhadu $\hat{\boldsymbol{\xi}}_0(\mathbf{X}) = \mathbf{X}$ pre dimenziu $n \geq 3$ a odôvodníme člen $(n-2)$ v James-Steinovom odhade. Podobne ako predošlá kapitola bude táto kapitola sledovať prácu Jamesa a Steina, viď. (James a Stein, 1992). Z dôvodu zhody so štruktúrou tejto práce došlo k miernej zmene značenia a zápisu, no obsahová stránka zostala nezmenená.

3.1 Vlastnosti James-Steinovho odhadu

Za predpokladu, že dimenzia n vektora stredných hodnôt je väčšia ako tri, platí pre strednú kvadratickú odchýlku James-Steinovho odhadu

$$\mathbb{E} \left\| \left(1 - \frac{(n-2)}{\|\mathbf{X}\|^2} \right) \mathbf{X} - \boldsymbol{\xi} \right\|^2 = n - \mathbb{E} \frac{(n-2)^2}{n-2+2Y} < n, \quad (3.1)$$

kde Y je náhodná veličina s Poissonovým rozdelením so strednou hodnotou $\|\boldsymbol{\xi}\|^2/2$. Vzťah 3.1 budeme dokazovať v ďalšej sekcii.

Zo vzťahu 3.1 vyplýva, že hodnota strednej kvadratickej odchýlky James-Steinovho odhadu je ostro menšia než v prípade VP odhadu pre ľubovoľnú hodnotu $\boldsymbol{\xi}$. Voľba člena $(n-2)$ v menovateli James-Steinovho odhadu nie je náhodná. Uvažujme odhad

$$\phi_2(\mathbf{X}) = \left(1 - \frac{b}{\|\mathbf{X}\|^2} \right) \mathbf{X}, \quad (3.2)$$

kde $b > 0$ je kladná reálna konštanta. Hodnota strednej kvadratickej odchýlky odhadu $\phi_2(\mathbf{X})$ potom dosahuje minima cez všetky $\boldsymbol{\xi}$ práve pre hodnotu $b = n-2$.

3.2 Dôkaz neprípustnosti

V tejto sekcii dokážeme vlastnosti uvedené v sekcii 3.1, to jest neprípustnosť VP odhadu a zároveň voľbu konštanty $b = n-2$. Uvažujme odhad

$$\phi_2(\mathbf{X}) = \left(1 - \frac{b}{\|\mathbf{X}\|^2} \right) \mathbf{X}, \quad (3.3)$$

kde $b > 0$ je kladná reálna konštanta. Hodnota strednej kvadratickej odchýlky tohto odhadu je

$$\begin{aligned} \text{MSE}_{\phi_2} &= \mathbb{E} \left\| \left(1 - \frac{b}{\|\mathbf{X}\|^2} \right) \mathbf{X} - \boldsymbol{\xi} \right\|^2 \\ &= \mathbb{E} \|\mathbf{X} - \boldsymbol{\xi}\|^2 - 2b \mathbb{E} \frac{(\mathbf{X} - \boldsymbol{\xi})^T \mathbf{X}}{\|\mathbf{X}\|^2} + b^2 \mathbb{E} \frac{1}{\|\mathbf{X}\|^2} \\ &= n - 2b \mathbb{E} \frac{(\mathbf{X} - \boldsymbol{\xi})^T \mathbf{X}}{\|\mathbf{X}\|^2} + b^2 \mathbb{E} \frac{1}{\|\mathbf{X}\|^2}. \end{aligned} \quad (3.4)$$

$\|\mathbf{X}\|^2$ je náhodná veličina s necentrálnym chi-kvadrát rozdelením s n stupňami voľnosti a parametrom $\|\boldsymbol{\xi}\|^2$. Nech W je náhodná veličina s podmieneným centrálnym chi-kvadrát rozdelením s $n + 2Y$ stupňami voľnosti, kde Y je náhodná veličina s Poissonovým rozdelením so strednou hodnotou $1/2\|\boldsymbol{\xi}\|^2$. $\|\mathbf{X}\|^2$ a W sú rovnako rozdelené a platí

$$\mathbb{E} \frac{1}{\|\mathbf{X}\|^2} = \mathbb{E} \frac{1}{\chi_{n+2Y}^2} = \mathbb{E} \left(\mathbb{E} \frac{1}{\chi_{n+2Y}^2} \middle| Y \right) = \mathbb{E} \frac{1}{n-2+2Y}. \quad (3.5)$$

Na výpočet strednej hodnoty prostredného výrazu na pravej strane (3.4) označíme

$$\begin{aligned} U &= \frac{\boldsymbol{\xi}^T \mathbf{X}}{\|\boldsymbol{\xi}\|} \\ V &= \left\| \mathbf{X} - \frac{\boldsymbol{\xi}^T \mathbf{X}}{\|\boldsymbol{\xi}\|^2} \boldsymbol{\xi} \right\|^2. \end{aligned} \quad (3.6)$$

Potom

$$W = \|\mathbf{X}\|^2 = U^2 + V \quad (3.7)$$

kde U má normálne rozdelenie so strednou hodnotou $\|\boldsymbol{\xi}\|$ a rozptylom 1 a V je nezávislé od U a má χ^2 rozdelenie s $n - 1$ stupňami voľnosti. Združená hustota U a V je

$$\frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2}(u - \|\boldsymbol{\xi}\|)^2 \right] \frac{1}{2^{(n-1)/2} \Gamma[(n-1)/2]} v^{(n-3)/2} e^{-v/2} \quad (3.8)$$

pre $v \geq 0$ a 0 pre $v < 0$. Združená hustota U a W je potom

$$\frac{1}{\sqrt{2\pi} 2^{(n-1)/2} \Gamma[(n-1)/2]} (w - u^2)^{(n-3)/2} \exp \left[-\frac{1}{2}\|\boldsymbol{\xi}\|^2 + \|\boldsymbol{\xi}\|u - \frac{1}{2}w \right], \quad (3.9)$$

ak $u^2 \leq w$ a 0 inak. Platí

$$\begin{aligned} \mathbb{E} \frac{U}{W} &= \frac{\exp(-\frac{1}{2}\|\boldsymbol{\xi}\|^2)}{\sqrt{2\pi} 2^{(n-1)/2} \Gamma[(n-1)/2]} \int_0^\infty dw \\ &\int_{-\sqrt{w}}^{\sqrt{w}} \frac{u}{w} (w - u^2)^{(n-3)/2} \exp \left(\|\boldsymbol{\xi}\|u - \frac{1}{2}w \right) du. \end{aligned} \quad (3.10)$$

Substitúciou $t = \frac{u}{\sqrt{w}}$ dostávame

$$\begin{aligned}
& \int_0^\infty dw \int_{-\sqrt{w}}^{\sqrt{w}} \frac{u}{w} (w - u^2)^{(n-3)/2} \exp(\|\boldsymbol{\xi}\|u - \frac{1}{2}w) du \\
&= \int_0^\infty w^{(n-3)/2} \exp(-\frac{1}{2}w) dw \int_{-1}^1 t(1-t^2)^{(n-3)/2} \exp(\|\boldsymbol{\xi}\|t\sqrt{w}) dt \\
&= \int_0^\infty w^{(n-3)/2} \exp(-\frac{1}{2}w) dw \sum_{i=0}^\infty \frac{(\|\boldsymbol{\xi}\|\sqrt{w})^i}{i!} \int_{-1}^1 t^{i+1} (1-t^2)^{(n-3)/2} dt \\
&= \sum_{j=0}^\infty \frac{\|\boldsymbol{\xi}\|^{2j+1}}{(2j+1)!} \int_0^\infty w^{n/2+j-1} \exp(-\frac{1}{2}w) dw \int_{-1}^1 t^{2(j+1)} (1-t^2)^{(n-3)/2} dt \\
&= \sum_{j=0}^\infty \frac{\|\boldsymbol{\xi}\|^{2j+1}}{(2j+1)!} 2^{n/2+j} \Gamma\left[\frac{n}{2} + j\right] B\left[j + \frac{3}{2}, \frac{n-1}{2}\right] \\
&= \sum_{j=0}^\infty \frac{\|\boldsymbol{\xi}\|^{2j+1} 2^{n/2+j} \Gamma\left[j + \frac{3}{2}\right] \Gamma\left[\frac{n-1}{2}\right]}{(2j+1)! \left(\frac{n}{2} + j\right)} \\
&= \Gamma\left[\frac{n-1}{2}\right] 2^{n/2} \sum_{j=0}^\infty \frac{2^j \Gamma\left[\frac{1}{2}\right] \|\boldsymbol{\xi}\|^{2j+1}}{2^{2j+1} \Gamma[j+1] \left(\frac{n}{2} + j\right)}.
\end{aligned} \tag{3.11}$$

Z rovníc (3.6), (3.7), (3.10) a (3.11) vidíme, že platí

$$\begin{aligned}
\mathbb{E} \frac{(\mathbf{X} - \boldsymbol{\xi})^T \mathbf{X}}{\|\mathbf{X}\|^2} &= 1 - \|\boldsymbol{\xi}\| \mathbb{E} \frac{U}{W} \\
&= 1 - \|\boldsymbol{\xi}\| \frac{\exp(-\frac{1}{2}\|\boldsymbol{\xi}\|^2)}{\sqrt{\pi}} \sum_{j=0}^\infty \frac{\Gamma\left[\frac{1}{2}\right] \|\boldsymbol{\xi}\|^{2j+1}}{2^{j+1} \Gamma[j+1] \left(\frac{n}{2} + j\right)} \\
&= \exp(-\frac{1}{2}\|\boldsymbol{\xi}\|^2) \left[\sum_{i=0}^\infty \frac{(\frac{1}{2}\|\boldsymbol{\xi}\|^2)^i}{i!} - \|\boldsymbol{\xi}\|^2 \sum_{j=0}^\infty \frac{(\frac{1}{2}\|\boldsymbol{\xi}\|^2)^j}{j!(n+2j)} \right] \\
&= \exp(-\frac{1}{2}\|\boldsymbol{\xi}\|^2) \sum_{i=0}^\infty \frac{(\frac{1}{2}\|\boldsymbol{\xi}\|^2)^i}{i!} \frac{n-2}{n-2+2i} \\
&= (n-2) \mathbb{E} \frac{1}{n-2+2Y},
\end{aligned} \tag{3.12}$$

kde Y má Poissonovo rozdelenie so strednou hodnotou $\mathbb{E}Y = \|\boldsymbol{\xi}\|^2/2$. Kombináciou (3.4), (3.5) a (3.12) dostávame

$$\begin{aligned}
\text{MSE}_{\phi_2} &= \mathbb{E} \left\| \left(1 - \frac{b}{\|\mathbf{X}\|^2}\right) \mathbf{X} - \boldsymbol{\xi} \right\|^2 \\
&= n - 2(n-2)b \mathbb{E} \frac{1}{n-2+2Y} + b^2 \mathbb{E} \frac{1}{n-2+2Y}.
\end{aligned} \tag{3.13}$$

Hodnotu tejto funkcie minimalizujeme pre všetky $\boldsymbol{\xi}$, ak za b dosadíme $n-2$, čím dostaneme práve pravú stranu výrazu 3.1.

4. Simulácia

V tejto kapitole bude naším cieľom demonštrovať chovanie strednej kvadratickej odchýlky pre VP odhad $\hat{\xi}_0$ a James-Steinov odhad $\hat{\xi}_1$ definované v 1.2 na konkrétnych dátových súboroch. Zameriame sa na chovanie James-Steinovho odhadu oproti VP odhadu pri obmene dĺžky a dimenzie vektorov stredných hodnôt ξ . Budeme pracovať s konkrétnymi realizáciami náhodného výberu z príslušných normálnych rozdelení. Transformáciou prvkov tohoto náhodného výberu a následným spriemerovaním, dostaneme numerické aproximácie strednej kvadratickej odchýlky.

Numerické simulácie boli vykonané v programe Wolfram Mathematica. Príklad používaného kódu vid. príloha A.1

4.1 Generovanie dát

Simulácie prebiehali podľa nasledovného algoritmu.

Zvolíme dimenziu $n \in \mathbb{N}$ vektora stredných hodnôt.

Zvolíme vektor stredných hodnôt $\xi \in \mathbb{R}^n$.

Vygenerujeme postupnosť $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p \sim \mathcal{N}_n(\xi, I_n)$ s rozsahom $p = 10000$. Z tejto postupnosti následne vypočítame hodnotu

$$\frac{1}{p} \sum_{i=1}^p \|\hat{\xi}_1(\mathbf{X}_i) - \xi\|^2 = \frac{1}{p} \sum_{i=1}^p \left\| \left(1 - \frac{n-2}{\|\mathbf{X}_i\|^2}\right) \mathbf{X}_i - \xi \right\|^2. \quad (4.1)$$

Tým dostaneme numerickú aproximáciu strednej kvadratickej odchýlky James-Steinovho odhadu pre uvažovaný vektor stredných hodnôt ξ .

Z postupnosti $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$ vygenerovanej vyššie vypočítame taktiež hodnotu

$$\frac{1}{p} \sum_{i=1}^p \|\hat{\xi}_0(\mathbf{X}_i) - \xi\|^2 = \frac{1}{p} \sum_{i=1}^p \|\mathbf{X}_i - \xi\|^2, \quad (4.2)$$

čím získame aproximáciu strednej kvadratickej odchýlky VP odhadu.

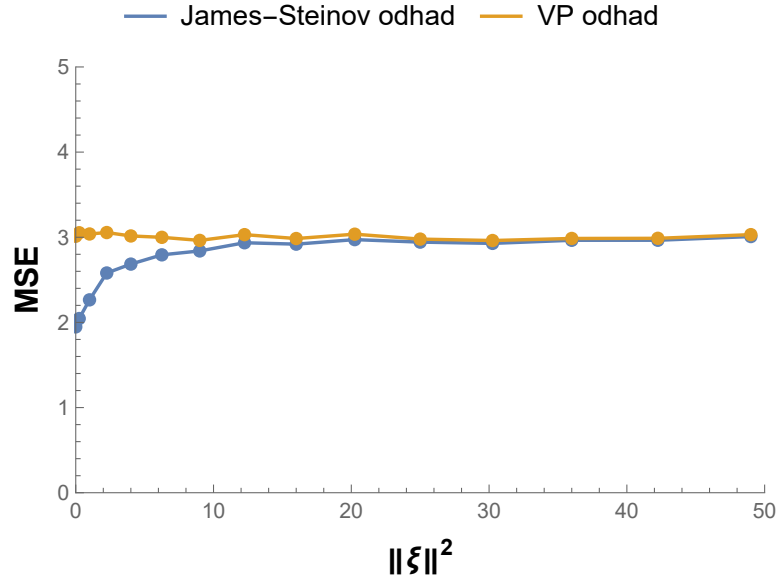
Simuláciu opakujeme pre rôzne hodnoty parametra ξ .

Vypočítané hodnoty vynesieme do grafu závislosti strednej kvadratickej odchýlky uvažovaných odhadov od kvadrátu normy vektora $\|\xi\|^2$. Hodnoty strednej kvadratickej odchýlky James-Steinovho odhadu budú závisieť od konkrétnych vektorov ξ a dimenzie n . Hodnoty rizikovej funkcie VP odhadu však budú závisieť len od dimenzie n .

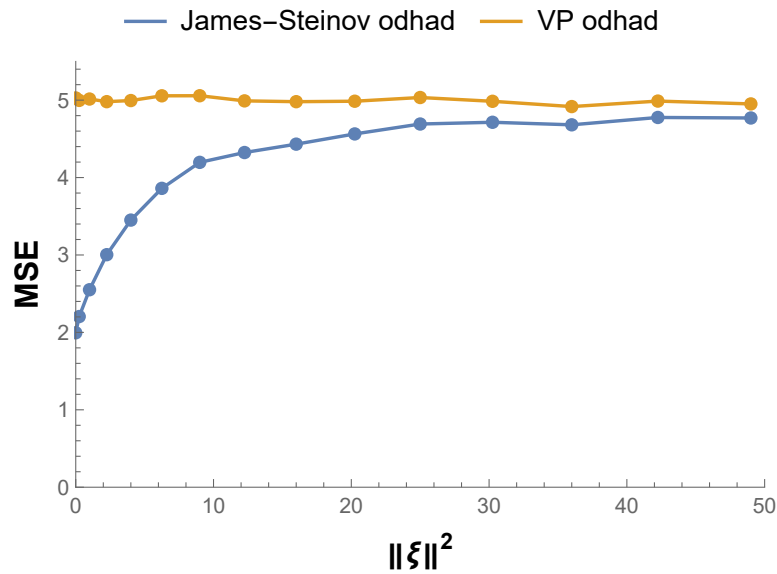
4.2 Chovanie rizikovej funkcie

Naším cieľom bolo demonštrovať chovanie stredných kvadratických odchýlok uvažovaných odhadov. Dimenzie vektora stredných hodnôt sme volili postupne $n = 3, 5, 10, 20$. Pre zvolenú dimenziu sme následne menili normu $\|\xi\|^2$ v rozsahu $[0; 49]$.

V ďalšom prezentované numerické výsledky vykazujú drobné odchýlky od analytických očakávaní. Jedná sa o štatistické chyby spôsobené obmedzeným počtom



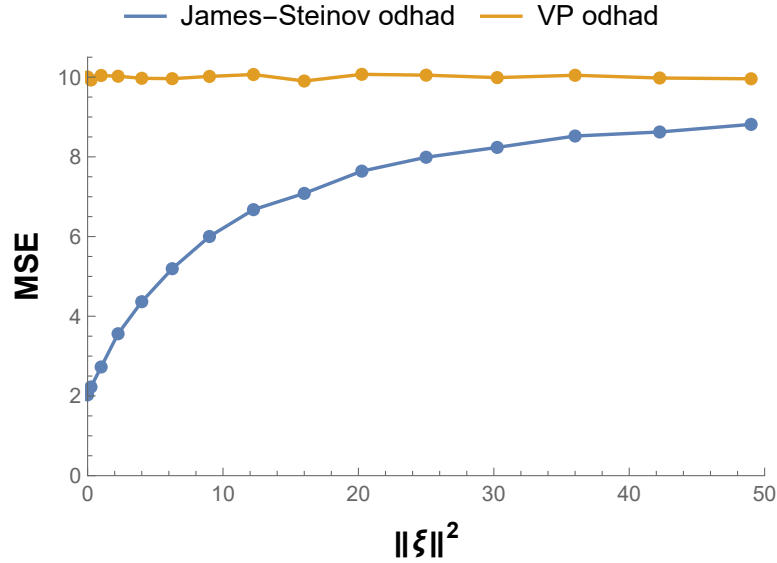
Obrázek 4.1: Stredná kvadratická odchýlka odhadov pri zmene kvadrátu normy vektora stredných hodnôt $\|\xi\|^2$ a dimenzii $n = 3$. Konkrétne vypočítané hodnoty, vrátane štandardných chýb, vid. príloha A.2, tabuľka A.1.



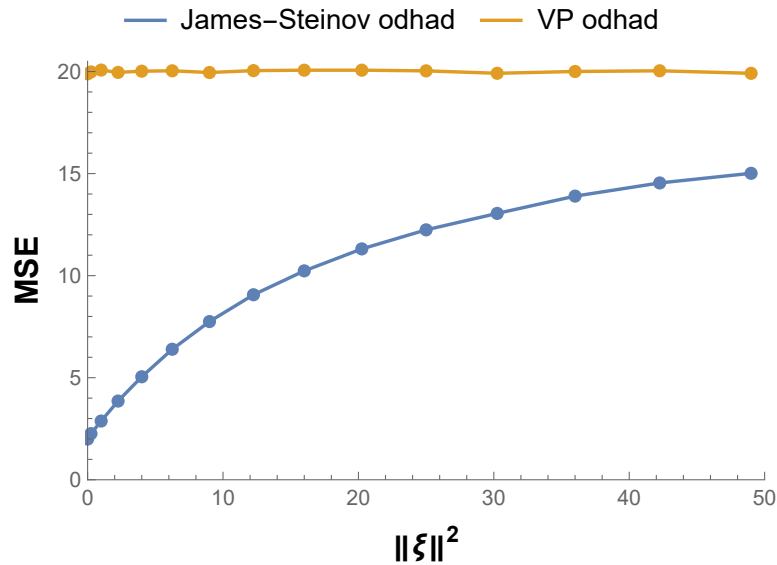
Obrázek 4.2: Stredná kvadratická odchýlka pri zmene kvadrátu normy vektora stredných hodnôt $\|\xi\|^2$ a dimenzii $n = 5$. Konkrétne vypočítané hodnoty, vrátane štandardných chýb, vid. príloha A.2, tabuľka A.2.

prvkov (p) generovaných postupností. Hodnotu $p = 10000$ sme zvolili ako vhodný kompromis medzi štatistickou presnosťou simulácií a časovou náročnosťou výpočtov. Pozorované štatistické chyby nemajú vplyv na demonštrované kvalitatívne vlastnosti skúmaných odhadov. Kvantitatívne uvádzame hodnoty štatistických chýb pre jednotlivé simulácie v tabuľkách v prílohe A.2.

Výstup simulácie s dimenziou vektora stredných hodnôt $n = 3$ je zobrazený na grafe 4.1. Numerický výpočet strednej kvadratickej odchýlky VP odhadu dosahoval s minimálnou fluktuáciou hodnotu rovnú práve zvolenej dimenzii, teda tri.



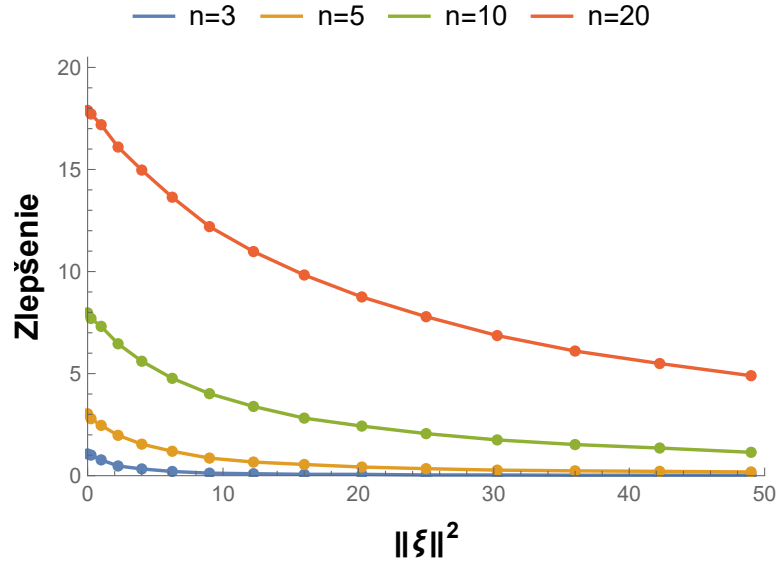
Obrázek 4.3: Stredná kvadratická odchýlka odhadov pri zmene kvadrátu normy vektora stredných hodnôt $\|\xi\|^2$ a dimenzii $n = 10$. Konkrétne vypočítané hodnoty, vrátane štandardných chýb, vid. príloha A.2, tabuľka A.3.



Obrázek 4.4: Stredná kvadratická odchýlka odhadov pri zmene kvadrátu normy vektora stredných hodnôt $\|\xi\|^2$ a dimenzii $n = 20$. Konkrétne vypočítané hodnoty, vrátane štandardných chýb, vid. príloha A.2, tabuľka A.4.

V prípade James-Steinovho odhadu vidíme očakávaný pokles hodnoty strednej kvadratickej odchýlky. Najväčšie zlepšenie je pozorovateľné práve pre hodnotu $\|\xi\|^2 = 0$. V súlade s teoretickým predpokladom uvedeným v 1.1 bola nami spočítaná hodnota strednej kvadratickej odchýlky James-Steinovho odhadu pre $\|\xi\|^2 = 0$ rovná dvom v rámci štatistickej chyby simulácie. Pri náraste hodnôt $\|\xi\|^2$ potom stredná kvadratická odchýlka James-Steinovho odhadu konvergovala k hodnotám vypočítaným pre VP odhad.

Ďalšie grafy ukazujú vývoj strednej kvadratickej odchýlky oboch odhadov pre



Obrázok 4.5: Rozdiel strednej kvadratickej odchýlky VP odhadu a James-Steinovho odhadu.

vyššie dimenzie. Obdobne ako na obrázku 4.1 ostáva hodnota strednej kvadratickej odchýlky VP odhadu prakticky konštantná, vždy rovná uvažovanej dimenzii (opäť v rámci štandardnej chyby simulácie). Chovanie James-Steinovho odhadu je taktiež podobné v každej simulácii. Stredná kvadratická odchýlka pre $\|\xi\|^2 = 0$ nadobúda hodnotu dva a následne rastie smerom k hodnote n .

Obrázok 4.5 ukazuje rozdiel v strednej kvadratickej odchýlke porovnávaných odhadov pre rôzne dimenzie n . Rozdiel v hodnotách strednej kvadratickej odchýlky konverguje k nule pri rastúcej hodnote $\|\xi\|^2$. Obrázok 4.5 znázorňuje aj nárast rozdielu strednej kvadratickej odchýlky uvažovaných odhadov pri pevnej norme vektora stredných hodnôt a rôznych dimenziách n . Tento rozdiel, respektíve pokles v hodnote strednej kvadratickej odchýlky James-Steinovho odhadu oproti VP odhadu, v grafe označujeme ako zlepšenie.

Záver

Cieľom tejto bakalárskej práce bolo popísať prípustnosť a neprípustnosť odhadu, definovať James-Steinov odhad a vyšetriť chovanie rôznych odhadov pomocou numerických simulácií.

V teoretickom úvode sme definovali potrebné pojmy, ako odhad, jeho prípustnosť a strednú kvadratickú odchýlku. Predstavili sme VP odhad (výberový priemer, viď. úvod kapitoly 1.2) a James-Steinov odhad, uviedli ich základné vlastnosti a popísali James-Steinov paradox. Následne sme dokázali prípustnosť VP odhadu pre dimenziu (n) vektora stredných hodnôt menšiu ako tri. Diskutovali sme člen $(n - 2)$ v predpise James-Steinovho odhadu. Formálne sme dokázali neprípustnosť VP odhadu pre dimenziu vektora stredných hodnôt väčšiu ako dva.

Vlastným prínosom tejto práce bola príprava a vykonanie numerických simulácií, ktorými sme demonštrovali správanie sa stredných kvadratických odchýlok uvažovaných odhadov. Na náhodne generovaných dátových súboroch sme numericky počítali strednú kvadratickú odchýlku VP odhadu a James-Steinovho odhadu. Pre zvolenú dimenziu vektora stredných hodnôt sme do grafu vyniesli hodnoty stredných kvadratických odchýlok v závislosti od kvadrátu normy vektora stredných hodnôt. Demonštrovali sme neprípustnosť VP odhadu pre $n \geq 3$ a diskutovali charakteristické trendy oboch uvažovaných odhadov na základe numericky získaných hodnôt. Diskutovali sme štatistické chyby v generovaných hodnotách. Konkrétne numerické chyby sme uviedli v prílohe. Numerické simulácie boli vykonané v programe Wolfram Mathematica.

Zoznam použitej literatúry

- GIRSHICK, M., SAVAGE, L. A. KOL. (1951). Bayes and minimax estimates for quadratic loss functions. In *Proc. Second Berkeley Symp. Math. Statist. Probab.*, volume 1, pages 53–74. Univ. of California Press, Berkeley.
- HODGES, J. a LEHMANN, E. (1951). Some applications of the cramer-rao inequality. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 13–23. University of California Press.
- JAMES, W. a STEIN, C. (1992). Estimation with quadratic loss. In *Breakthroughs in statistics: Foundations and basic theory*, pages 443–460. Springer.
- KULICH, M. (2014). Přehledový větník. https://www.karlin.mff.cuni.cz/~pesta/NMFM301/statistika_fm.pdf. Sprístupnená: 20-7-2023.
- NAGY, S. (2023). Mathematical Statistics II. <https://www2.karlin.mff.cuni.cz/~nagy/NMSA332/NMSA332.pdf>. Sprístupnená: 12-4-2023.
- STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Technical report, Stanford University Stanford United States.

A. Prílohy

A.1 Simulačný kód

Príklad kódu z aplikácie wolfram mathematica využitý pri simuláciách

```
(* inicializácia nahodneho generatora na konstantny seed *)
SeedRandom[1118]

(* Pomocne funkcie na realizáciu odhadov *)
stein[x_, b_] := 1 - (b - 2)/x . x
mse[x_, y_, b_] :=
Total[MapThread[(stein[x, b]*#1 - #2)^2 &, {x, y}]]
bezny[x_, y_] := Total[MapThread[(#1 - #2)^2 &, {x, y}]]

(* Numerická simulácia pre dimenziu n=1 *)
(* Generovanie realizácii normalneho rozdelenia pre rozne stredne
hodnoty*)
vybery1 =
Table[RandomVariate[NormalDistribution[i, 1], 1000]
, {i, 0, 7, 0.5}];
stredneh1 = Table[{i}, {i, 0, 7, 0.5}];

(* Numericky vypocet strednej kvadratickej odchylky nad
vygenerovanim datasetom pre bezny odhad a James Steinov odhad *)
baseline1 =
Table[Mean[Map[bezny[#, stredneh1[[i]]] &, vybery1[[i]]]], {i,
Length[stredneh1]}];
xi1 = Table[
Mean[Map[mse[#, stredneh1[[i]], 2.5] &, vybery1[[i]]]], {i,
Length[stredneh1]}];

(* Simulacia pre n=3, rovnaké kroky ako pre dimenziu 1 *)
vybery3 =
Table[RandomVariate[
MultinormalDistribution[{i, 0, 0}, IdentityMatrix[3]], 10000]
, {i, 0, 7, 0.5}];
stredneh3 = Table[{i, 0, 0}, {i, 0, 7, 0.5}];
baseline3 =
Table[Mean[Map[bezny[#, stredneh3[[i]]] &, vybery3[[i]]]], {i,
Length[stredneh3]}];
xi3 = Table[
Mean[Map[mse[#, stredneh3[[i]], 3] &, vybery3[[i]]]], {i,
Length[stredneh3]}];

(* Simulácia pre n=5 az n=20, rovnaké kroky ako pre dimenziu 1 *)
vybery5 =
Table[RandomVariate[
MultinormalDistribution[{i, 0, 0, 0, 0}, IdentityMatrix[5]],
10000], {i, 0, 7, 0.5}];
```



```

stredneh5 = Table[{i, 0, 0, 0, 0}, {i, 0, 7, 0.5}];
baseline5 =
  Table[Mean[Map[bezny[#, stredneh5[[i]]] &, vybery5[[i]]]], {i,
    Length[stredneh5]};
xi5 = Table[
  Mean[Map[mse[#, stredneh5[[i]], 5] &, vybery5[[i]]]], {i,
    Length[stredneh5]};
vybery10 =
  Table[RandomVariate[
    MultinormalDistribution[{i, 0, 0, 0, 0, 0, 0, 0, 0, 0},
      IdentityMatrix[10]], 10000], {i, 0, 7, 0.5}];
stredneh10 = Table[{i, 0, 0, 0, 0, 0, 0, 0, 0, 0},
  {i, 0, 7, 0.5}];
baseline10 =
  Table[Mean[Map[bezny[#, stredneh10[[i]]] &, vybery10[[i]]]], {i,
    Length[stredneh10]};
xi10 = Table[
  Mean[Map[mse[#, stredneh10[[i]], 10] &, vybery10[[i]]]], {i,
    Length[stredneh10]};
vybery20 =
  Table[RandomVariate[
    MultinormalDistribution[{i, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
      IdentityMatrix[20]], 10000], {i, 0, 7, 0.5}];
stredneh20 =
  Table[{i, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
    {i, 0, 7, 0.5}];
baseline20 =
  Table[Mean[Map[bezny[#, stredneh20[[i]]] &, vybery20[[i]]]], {i,
    Length[stredneh20]};
xi20 = Table[
  Mean[Map[mse[#, stredneh20[[i]], 20] &, vybery20[[i]]]], {i,
    Length[stredneh20]};

```

```

(* Pomocna funkcia pre vykreslenie grafov, parovanie hodnot
numericky vypocitanych rizikovych funkcii odhadov s normami
vektorov strednych hodnot *)
combo[x_] := MapThread[{#1, #2} &, {Table[i^2, {i, 0, 7, 0.5}], x}];

```

```

(* Generovanie vystupnych grafov *)
ListLinePlot[{combo[xi3], combo[baseline3]}, AxesOrigin -> {0, 0},
  PlotLegends -> Placed[{"Steinov odhad", "MSE odhad"}, Above]
, Frame -> {{True, False}, {True, False}},
  FrameLabel -> {{Style["Riziková funkcia", Bold, Black, 15],
    None}, {Style[DoubleBracketingBar[\[Xi]]^2, 15, Bold, Black],
    None}}, PlotRangePadding -> None, Mesh -> All,
  PlotRange -> {{0, 50}, {0, 5}}]
ListLinePlot[{combo[xi5], combo[baseline5]}, AxesOrigin -> {0, 0},
  PlotLegends -> Placed[{"Steinov odhad", "MSE odhad"}, Above]
, Frame -> {{True, False}, {True, False}},

```

```

FrameLabel -> {{Style["Riziková funkcia", Bold, Black, 15],
  None}, {Style[DoubleBracketingBar[\[Xi]]^2, 15, Bold, Black],
  None}}, PlotRangePadding -> None, Mesh -> All,
PlotRange -> {{0, 50}, {0, 5.5}}]
ListLinePlot[{combo[xi10], combo[baseline10]}, AxesOrigin -> {0, 0},
PlotLegends -> Placed[{"Steinov odhad", "MSE odhad"}, Above]
, Frame -> {{True, False}, {True, False}},
FrameLabel -> {{Style["Riziková funkcia", Bold, Black, 15],
  None}, {Style[DoubleBracketingBar[\[Xi]]^2, 15, Bold, Black],
  None}}, PlotRangePadding -> None, Mesh -> All,
PlotRange -> {{0, 50}, {0, 10.5}}]
ListLinePlot[{combo[xi20], combo[baseline20]}, AxesOrigin -> {0, 0},
PlotLegends -> Placed[{"Steinov odhad", "MSE odhad"}, Above]
, Frame -> {{True, False}, {True, False}},
FrameLabel -> {{Style["Riziková funkcia", Bold, Black, 15],
  None}, {Style[DoubleBracketingBar[\[Xi]]^2, 15, Bold, Black],
  None}}, PlotRangePadding -> None, Mesh -> All,
PlotRange -> {{0, 50}, {0, 20.5}}]
zlepsenie = {baseline3 - xi3, baseline5 - xi5, baseline10 - xi10,
  baseline20 - xi20};
ListLinePlot[{combo[zlepsenie[[1]]], combo[zlepsenie[[2]]],
  combo[zlepsenie[[3]]], combo[zlepsenie[[4]]]}, AxesOrigin -> {0, 0},
PlotLegends -> Placed[{"n=3", "n=5", "n=10", "n=20"}, Above]
, Frame -> {{True, False}, {True, False}},
FrameLabel -> {{Style["Zlepšenie", Bold, Black, 15],
  None}, {Style[DoubleBracketingBar[\[Xi]]^2, 15, Bold, Black],
  None}}, PlotRangePadding -> None, Mesh -> All,
PlotRange -> {{0, 50}, {0, 20.5}}]
pomoc[x_, y_] := MapThread[#1 - #2 &, {x, y}]

(*vypocet standardnych odchylok*)
js3 = Map[StandardDeviation,
  Table[Map[mse[#, stredneh3[[i]], 3] &, vybery3[[i]]], {i,
  Length[stredneh3]}]]
js5 = Map[StandardDeviation,
  Table[Map[mse[#, stredneh5[[i]], 5] &, vybery5[[i]]], {i,
  Length[stredneh5]}]]
js10 = Map[StandardDeviation,
  Table[Map[mse[#, stredneh10[[i]], 10] &, vybery10[[i]]], {i,
  Length[stredneh10]}]]
js20 = Map[StandardDeviation,
  Table[Map[mse[#, stredneh20[[i]], 20] &, vybery20[[i]]], {i,
  Length[stredneh20]}]]
be3 = Map[StandardDeviation,
  Table[Map[bezny[#, stredneh3[[i]]] &, vybery3[[i]]], {i,
  Length[stredneh3]}]]
be5 = Map[StandardDeviation,
  Table[Map[bezny[#, stredneh5[[i]]] &, vybery5[[i]]], {i,
  Length[stredneh5]}]]
be10 = Map[StandardDeviation,
  Table[Map[bezny[#, stredneh10[[i]]] &, vybery10[[i]]], {i,

```

```

Length[stredneh10}}]]
be20 = Map[StandardDeviation,
Table[Map[bezny[#, stredneh20[[i]]] &, vybery20[[i]]], {i,
Length[stredneh20}}]]

```

A.2 Numerické výstupy simulácií

V tejto časti uvádzame numerické výstupy simulácií prezentovaných v grafoch v sekcii 4.2.

$\ \xi\ ^2$	J-S	SE J-S	VP	SE VP
0.00	1.948	0.0328	3.012	0.0248
0.25	2.045	0.0318	3.053	0.0247
1.00	2.265	0.0292	3.039	0.0249
2.25	2.580	0.0455	3.055	0.0248
4.00	2.685	0.0240	3.015	0.0243
6.25	2.794	0.0229	2.999	0.0245
9.00	2.839	0.0229	2.963	0.0236
12.25	2.936	0.0241	3.031	0.0248
16.00	2.919	0.0240	2.985	0.0246
20.25	2.972	0.0242	3.036	0.0247
25.00	2.943	0.0239	2.978	0.0242
30.25	2.928	0.0240	2.962	0.0243
36.00	2.965	0.0241	2.986	0.0243
42.25	2.965	0.0242	2.988	0.0244
49.00	3.008	0.0249	3.031	0.0250

Tabulka A.1: Numericky vypočítané hodnoty a štandardné chyby stredných kvadratických odchýlok James-Steinovho odhadu a VP odhadu pre rôzne hodnoty normy vektora stredných hodnôt $\|\xi\|^2$ pre dimenziu $n = 3$. Skratka J-S značí strednú kvadratickú odchýlku James-Steinovho odhadu. SE J-S potom značí štandardnú chybu pri výpočte J-S. Obdobne VP predstavuje strednú kvadratickú odchýlku VP odhadu a SE VP potom štandardnú chybu.

$\ \xi\ ^2$	J-S	SE J-S	VP	SE VP
0.00	1.995	0.0337	5.028	0.0319
0.25	2.205	0.0495	4.997	0.0317
1.00	2.551	0.0298	5.014	0.0318
2.25	3.004	0.0271	4.980	0.0316
4.00	3.451	0.0257	4.995	0.0316
6.25	3.861	0.0309	5.056	0.0318
9.00	4.196	0.0276	5.057	0.0318
12.25	4.323	0.0270	4.991	0.0311
16.00	4.432	0.0282	4.980	0.0317
20.25	4.564	0.0289	4.986	0.0320
25.00	4.692	0.0295	5.034	0.0317
30.25	4.714	0.0295	4.985	0.0311
36.00	4.681	0.0299	4.917	0.0313
42.25	4.776	0.0301	4.988	0.0316
49.00	4.770	0.0299	4.951	0.0311

Tabulka A.2: Numericky vypočítané hodnoty a štandardné chyby stredných kvadratických odchýlok James-Steinovho odhadu a VP odhadu pre rôzne hodnoty normy vektora stredných hodnôt $\|\xi\|^2$ pre dimenziu $n = 5$. Skratka J-S značí strednú kvadratickú odchýlku James-Steinovho odhadu. SE J-S potom značí štandardnú chybu pri výpočte J-S. Obdobne VP predstavuje strednú kvadratickú odchýlku VP odhadu a SE VP potom štandardnú chybu.

$\ \xi\ ^2$	J-S	SE J-S	VP	SE VP
0.00	2.027	0.0309	10.003	0.0453
0.25	2.226	0.0310	9.928	0.0447
1.00	2.727	0.0297	10.042	0.0444
2.25	3.561	0.0286	10.023	0.0455
4.00	4.365	0.0258	9.971	0.0448
6.25	5.193	0.0269	9.962	0.0447
9.00	6.001	0.0282	10.019	0.0452
12.25	6.675	0.0303	10.065	0.0451
16.00	7.083	0.0314	9.902	0.0447
20.25	7.642	0.0334	10.070	0.0449
25.00	7.990	0.0355	10.050	0.0448
30.25	8.237	0.0366	9.988	0.0441
36.00	8.523	0.0379	10.047	0.0449
42.25	8.624	0.0382	9.978	0.0446
49.00	8.814	0.0388	9.958	0.0440

Tabulka A.3: Numericky vypočítané hodnoty a štandardné chyby stredných kvadratických odchýlok James-Steinovho odhadu a VP odhadu pre rôzne hodnoty normy vektora stredných hodnôt $\|\xi\|^2$ pre dimenziu $n = 10$. Skratka J-S značí strednú kvadratickú odchýlku James-Steinovho odhadu. SE J-S potom značí štandardnú chybu pri výpočte J-S. Obdobne VP predstavuje strednú kvadratickú odchýlku VP odhadu a SE VP potom štandardnú chybu.

$\ \xi\ ^2$	J-S	SE J-S	VP	SE VP
0.00	1.994	0.0294	19.886	0.0629
0.25	2.261	0.0289	19.979	0.0638
1.00	2.877	0.0276	20.072	0.0637
2.25	3.857	0.0285	19.957	0.0636
4.00	5.047	0.0271	20.017	0.0630
6.25	6.393	0.0274	20.035	0.0641
9.00	7.750	0.0286	19.952	0.0639
12.25	9.065	0.0304	20.044	0.0640
16.00	10.234	0.0325	20.064	0.0643
20.25	11.310	0.0353	20.065	0.0639
25.00	12.244	0.0368	20.033	0.0625
30.25	13.049	0.0399	19.913	0.0624
36.00	13.896	0.0424	20.002	0.0629
42.25	14.542	0.0447	20.036	0.0635
49.00	15.014	0.0463	19.911	0.0631

Tabulka A.4: Numericky vypočítané hodnoty a štandardné chyby stredných kvadratických odchýlok James-Steinovho odhadu a VP odhadu pre rôzne hodnoty normy vektora stredných hodnôt $\|\xi\|^2$ pre dimenziu $n = 20$. Skratka J-S značí strednú kvadratickú odchýlku James-Steinovho odhadu. SE J-S potom značí štandardnú chybu pri výpočte J-S. Obdobne VP predstavuje strednú kvadratickú odchýlku VP odhadu a SE VP potom štandardnú chybu.