

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



PRAVDEPODOBNOSTNÉ MODELY V LUKOSTREĽBE

BAKALÁRSKA PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

**PRAVDEPODOBNOSTNÉ MODELY V LUKOSTREĽBE**

**BAKALÁRSKA PRÁCA**

Študijný program: Ekonomická a finančná matematika  
Študijný odbor: 1114 Aplikovaná matematika  
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky  
Veduci práce: Mgr. Martin Hurban



Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

---

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Lukáš Kováč  
**Študijný program:** ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)  
**Študijný odbor:** 9.1.9. aplikovaná matematika  
**Typ záverečnej práce:** bakalárska  
**Jazyk záverečnej práce:** slovenský  
**Sekundárny jazyk:** anglický

**Názov:** Pravdepodobnostné modely v lukostrelbe  
*Probabilistic models in archery*

**Cieľ:** Cieľom bakalárskej práce je analýza rozdelení zásahov v terči, odhady presnosti lukostrelcov, porovnanie súťažných systémov.

**Vedúci:** Mgr. Martin Hurban  
**Katedra:** FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky  
**Vedúci katedry:** prof. RNDr. Daniel Ševčovič, CSc.  
**Dátum zadania:** 19.10.2015

**Dátum schválenia:** 19.11.2015 doc. RNDr. Margaréta Halická, CSc.  
garant študijného programu

---

študent

---

vedúci práce

**Pod'akovanie** Rád by som sa poďakoval

- Ondrejovi Kováčovi, za pomoc pri spracovaní dát,
- Márii Mészárosovej (Hurbanovej), za korektúru mojich gramatických a štylistických zlyhaní,
- Martinovi Hurbanovi, za trpezlivosť, čas, ochotu a za rady a návrhy pri písaní tejto práce,
- Doc. RNDr. Margaréte Halickej, CSc. za poznámky ku korektnému písaniu bakalárskej práce a za uvedenie do matematickej byrokracie,
- Jánovi Somorčíkovi, za zrozumiteľné a zaujímavé prednášky z Ekonometrie, Štatistických metód a Počítačovej štatistiky.

## Abstrakt

KOVÁČ, Lukáš: Pravdepodobnostné modely v lukostreľbe [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: Mgr. Martin Hurban, Bratislava, 2016, 45 s.

V práci analyzujeme rozloženie šípov na terči vystrelených profesionálnymi lukostrelcami. Formulujeme spôsoby odhadov presnosti strelca pri predpoklade, že rozloženie výstrelov sa riadi dvojrozmerným normálnym rozdelením. Pozorovali sme, ako sa zmenil model po uvoľnení predpokladu, že výstrely strelca sú sústredené okolo stredu terča. Odhadovali sme parameter, ktorý by charakterizoval strelcovu presnosť a vychýlenosť od stredu terča. Odhady boli konštruované pomocou momentovej metódy a metódy maximálnej vierohodnosti. Požitím Monte Carlo simulácií sme porovnali ich teoretické kvality. Testami dobrej zhody sme overili mieru zhody teoretických a empirických početností zasiahnutia zón terča. Ukázalo sa, že testy na skúmanej vzorke strelcov pri všetkých troch rôznych odhadoch zaznamenali vysokú mieru zhody. Najlepšie umiestnený odhad momentovou metódou vynikal aj svojou výpočtovou nenáročnosťou.

**Kľúčové slová:** Odhad presnosti strelca, Metóda maximálnej vierohodnosti, Momentová metóda, Fischerova informácia, Monte Carlo simulácie, Test dobrej zhody

## Abstract

KOVÁČ, Lukáš: Probability models in archery [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: Mgr. Martin Hurban, Bratislava, 2016, 45 p.

In this study we analyze allocation of arrows on target shot by professional archers. We formalize methods to estimate archers accuracy under condition of normally distributed shots. The study observes how the models change after releasing the condition of arrows concentrating around the center of the target. The parameters we try to estimate should describe archers accuracy and deviation from the center. The estimations were based on maximum likelihood and moments methods. By using Monte Carlo simulations we compared theoretical qualities of different estimations. The theoretical and empirical accordance of frequencies, in which the individual target zones were hit, were tested by goodness of fit tests. The tests on sample of archers showed high rate of fit in three different estimations. The highest rated estimation by method of moments also scored best results in operation requirements comparison.

**Keywords:** Estimation of archer's accuracy, Maximum likelihood method, Method of moments, Fischer information, Monte Carlo simulation, Goodness of fit test

# Obsah

Úvod	8
<b>1 Pravidlá športovej lukostreľby</b>	<b>9</b>
1.1 Rozdelenie lukostreleckých disciplín . . . . .	9
1.2 Súťažný systém v lukostreľbe . . . . .	11
<b>2 Miera výkonnosti lukostrelcov</b>	<b>12</b>
2.1 Formulácia problému . . . . .	12
2.2 Odvodenie pravdepodobnosti zásahu medzikružia pri predpoklade nevy- chýlenosti . . . . .	13
2.2.1 Odhad $\sigma^2$ pomocou MMV . . . . .	13
2.2.2 Odhad $\sigma^2$ pomocou momentovej metódy . . . . .	15
2.3 Odvodenie pravdepodobnosti zásahu medzikružia bez predpokladu nevy- chýlenosti . . . . .	16
2.4 Výpočet Fisherovej informácie . . . . .	18
<b>3 Porovnávanie teoretickej kvality odhadov</b>	<b>21</b>
3.1 Monte Carlo simulácie . . . . .	21
3.2 Algoritmus odhadu $\sigma^2$ . . . . .	21
3.3 Výsledky MCS o kvalite odhadov . . . . .	23
3.4 Fischerova informácia . . . . .	29
<b>4 Porovnanie modelov na súťažných dátach</b>	<b>31</b>
4.1 Dátové vstupy . . . . .	31
4.2 Skutočné hodnoty pre jednotlivé odhady . . . . .	32
4.3 Testy dobrej zhody . . . . .	34
4.3.1 Chí-kvadrát test . . . . .	34
4.3.2 G-test . . . . .	36
4.4 Porovnanie odhadov . . . . .	36
<b>Záver</b>	<b>39</b>
<b>Príloha A</b>	<b>43</b>

**Príloha B**

**45**



## Úvod

Súťaživosť bola vždy hlboko zakorenená v povahe ľudí. Svoje zručnosti si medzi sebou porovnávali a predvádzali už od staroveku. Postupom času vznikali stále novšie a rozmanitejšie športy a aj počet účastníkov sa stále zvyšoval. V súčasnosti je šport nesmierne populárny a rozvinutý, počet rôznych súťažných disciplín pravdepodobne nikto nepozná. Do súťaží sa zapájajú ľudia z celého sveta.

A tak sa dostávame k nášmu problému. Tým, že je športovcov veľa, a že sú na vysokej úrovni sa sťažilo rozhodovanie, kto z nich je lepší. V pretekoch vyhraných o pár stotín sekúnd sa víťazstvo môže zdať náhodné. Znamená takáto výhra, že daný bežec dobehne prvý aj nabudúce? Práve z tohto dôvodu sa začali v športovom svete používať celkové hodnotenia športovcov za ich kariéru, nielen za jednu súťaž.

V našej práci sa budeme venovať lukostreľbe a pokúsime sa vytvoriť hodnotenie presnosti lukostrelca na základe jeho výsledkov zo súťažnej kvalifikácie [4]. Správanie výstrelov budeme popisovať rôznymi pravdepodobnostnými rozdeleniami, ako napríklad dvojrozmerným normálnym rozdelením. Naším cieľom bude odhadnúť parameter, ktorý by charakterizoval strelcovu presnosť a vychýlenosť jeho streľby od stredu terča. Analyzovať budeme postupy navrhnuté v prácach [3] a [7]. Následne budeme porovnávať teoretické kvality odhadov skonštruovaných na základe rôznych odhadovacích metód [13]. Na pozorovanie správania sa odhadov nám poslúžia Monte Carlo simulácie popísané v [2]. Tie nám poskytnú možnosť sledovať, akú výpovednú hodnotu o odhadoch v sebe dáta ukrývajú. Mieru výpovednej hodnoty bude popisovať Fischerova informácia, ktorá slúži aj na odhadnutie dolných hraníc disperzie odhadov [10].

Závislosti medzi veličinami z Monte Carlo simulácií budeme popisovať rôznymi regresnými modelmi uvedenými v [12]. Zaoberať sa budeme aj výsledkami reálnych lukostreleckých majstrovstiev a súťaží. Porovnáme strelcov v kvalifikačnej časti súťaží a odhadneme parametre charakterizujúce ich presnosť. Relevantnosť odhadov a ich schopnosť popísať strelcov výkon overíme testami dobrej zhody predstavenými v [6], [8] a [11]. Na záver zhrnieme a ohodnotíme rôzne druhy odhadov. Pozrieme sa, ako si počínali pri testoch dobrej zhody a porovnáme ich výpočtové nároky.

# 1 Pravidlá športovej lukostrelby

## 1.1 Rozdelenie lukostreleckých disciplín

V športovej lukostrelbe existuje mnoho rôznych disciplín ako: halová, terčová, terénna, 3D, . . . Záleži od pohlavia, typu luku, terča, hrúbky šípov a prostredia. Ďalej možno strelcov deliť podľa typu lukov (obr. 1): olympijský luk (*recurve bow*), kladkový luk (*compound bow*) a holý luk (*barebow*). Môže sa strieľať vnútri v hale na 18 alebo 25 metrov alebo vonku s nastaveniami uvedenými v nasledujúcej tabuľke:

vzdialenosť	muži	90 m	70 m	50 m	30 m
	ženy	70 m	60 m	50 m	30 m
šírka terča		122 cm		80 cm	

Lukostrelci sú ďalej delení do kategórií podľa veku a pohlavia. Najčastejšie zostavy sú:

- tzv. malá fita: streľba sady 2×36 šípov na vzdialenosť 70 m pre mužov a 60 m pre ženy.
- tzv. veľká fita: streľba sady 36 šípov na každú zo vzdialeností 30, 50, 70 a 90 m.

Strieľa sa 3 alebo 6 šípov za jedno kolo, po ktorom sa šípy pozbierajú a ohodnotia výsledky. Na každé z kôl má strelec obmedzený časový limit. Terče sú zložené z desiatich rovnako vzdialených sústredných kružníc s bodovaným ohodnotením od 1 do 10 (obr. 2). Vnútri najmenšieho kruhu sa nachádza kruh s polovičným polomerom s bodovým označením  $\times$ , ktorý slúži na určenie výfaza v prípade rovnakého bodového zisku. Postup pri určení výfazu je uvedený v práci [4]. Minutie terča sa označuje písmenom  $M$  a predstavuje zisk 0 bodov. Polomery terčov sa líšia v závislosti od vzdialenosti strelca od terča a prostredia (indoor, outdoor) a pohybujú sa na úrovni 40 cm až 122 cm. Podrobné delenia do vekových kategórií a dodatočné pravidlá sú uvedené v [4].

V našej práci budeme analyzovať dáta zodpovedajúce prípadu najčastejšie používanému na majstrovstvách a olympiádach:

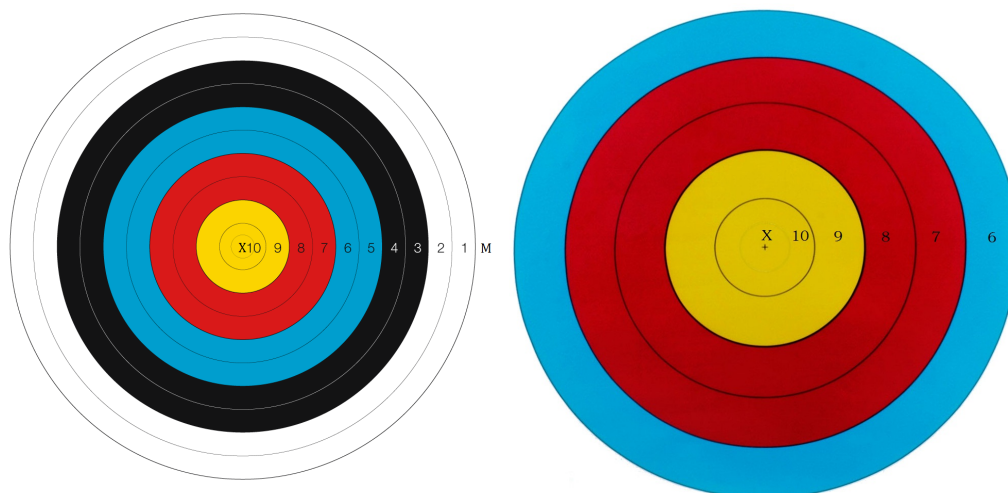
- typ luku: olympijský luk

- kategória: muži
- prostredie: vonku (outdoor)
- disciplína: terčová lukostreľba
- vzdialenosť od terča: 70 m
- veľkosť terča: 122 cm
- predpokladaná hrúbka šípu: 5,5 mm
- počet výstrelov: 72

Na nasledujúcich obrázkoch sú zobrazené rôzne typy lukov, ktoré sa používajú v súťažnej lukostreľbe a zobrazenie zón terča s príslušnými bodovými ohodnoteniami:



**Obrázok 1:** olympijský luk, kladkový luk a holý luk



Obrázok 2: Bodové ohodnotenia na 122 cm a 80 cm terčoch

## 1.2 Súťažný systém v lukostreľbe

V súťažnej lukostreľbe existujú rôzne pravidlá pre postup z kvalifikácie, cez vyraďovacie súboje až na stupeň víťazov. V súčasnosti sa najčastejšie používa systém kvalifikácie a vyraďovacích súbojov. Počas kvalifikácie sa pri disciplíne popísanej v časti (1.1) strieľajú 2 sady po 36-tich šípoch. Na základe výsledkov kvalifikácie sa určí poradie, pričom prví ôsmi majú zaručený priamy postup do neskoršej fázy vyraďovacích súbojov.

Zvyšní strelci, ktorí sa umiestnili nižšie ako na ôsmom mieste pokračujú do fázy vyraďovacích súbojov, kde prebiehajú súčasné rozstrely oboch strelcov, v sadách po troch šípoch. Po vystrelení troch výstrelcov sa šípy pozbierajú a nastrieľané hodnoty sa sčítajú a porovnávajú. Vyhrávajúci hráč získava 2 body, v prípade remízy získa každý strelec po jednom bode. Súboj pokračuje, až kým niektorý zo strelcov nezíska 6 bodov. Ak však nastane situácia, že obaja hráči nadobudli 5 bodov, nepokračuje sa v strieľaní 3 šípoch. Dochádza k rozstrelu, kde obaja strelci vystrelia iba jeden šíp, na základe ktorého sa určí víťaz.

Vyraďovacie súboje pokračujú až do momentu, v ktorom zostane už iba osem strelcov, ku ktorým sa pripoja prví ôsmi z kvalifikačnej časti. Vyraďovacie súboje pokračujú rovnakým spôsobom až do konečného určenia víťaza.

## 2 Miera výkonnosti lukostrelcov

### 2.1 Formulácia problému

Predpokladáme, že jednotlivé výstrely a teda aj celkové nastrielané bodové hodnoty strelcov sú navzájom nezávislé, rovnako rozdelené, z dvojrozmerného normálneho rozdelenia. Majme terč zobrazený na obrázku 2. Označme si  $S$  ako množinu všetkých možných realizácií výstrelcov:

$$S = \{1, 2, 3, \dots, 10, \times, M\}.$$

Pre výpočtové účely si označme bodovú hodnotu  $M$  (*missed*) predstavujúcu minútie terča ako číslo 0. Znak  $\times$  získaný po strelení najvnútornejšieho kruhu terča bude zastupovať číslo 10. Na transformáciu si zdefinujeme funkciu  $h(x) : S \rightarrow \mathbb{N}$ :

$$h(x) = \begin{cases} 0; & x = M \\ x; & x \in \mathbb{N} \\ 10; & x = \times \end{cases}$$

Množinu bodov z rovnakým bodovým ohodnotením označíme ako:

$$A_j = \{(x_1, x_2); r_j^2 \leq x_1^2 + x_2^2 \leq r_{j+1}^2\}; \quad j \in S \quad (2.1)$$

Množina  $A_j$  predstavuje medzikružie ohraničené danými polomerami. Do polomerov medzikruží je zahrnutá aj polovica hrúbky šípu. Polomery bodových zón v terči sú zobrazené v nasledovnej tabuľke:

Medzikružie	Bodový zisk $h_j$	Polomery zón
$A_\times$	10	0
$A_{10}$	10	3.325
$A_9$	9	6.375
$A_8$	8	12.475
$A_7$	7	18.575
$A_6$	6	24.675
$A_5$	5	30.775
$A_4$	4	36.875
$A_3$	3	42.975
$A_2$	2	49.075
$A_1$	1	55.175
$A_M$	0	61.275
		99999.275

Potom

$$P(\text{zasiahnutia } j\text{-teho medzikružia}) = \iint_{A_j} f(x_1, x_2) dx_1 dx_2. \quad (2.2)$$

## 2.2 Odvodenie pravdepodobnosti zásahu medzikružia pri predpoklade nevychýlenosti

Predpokladajme, že výstrely sú sústredené okolo stredu terča. Za hustotu rozdelenia  $f(x_1, x_2)$  v našom prípade dosadíme hustotu dvojrozmerného normálneho rozdelenia. Táto pravdepodobnosť bude závisieť od smerodajnej odchýlky  $\sigma$  a od polomerov medzikruží  $r_j$  a  $r_{j+1}$ , keďže integrujeme cez oblasti sústredných kruhov terča:

$$P(x \in A_j) = P(r_j^2 \leq x_1^2 + x_2^2 \leq r_{j+1}^2) = \iint_{A_j} \frac{1}{2\pi|\Sigma|^{-\frac{1}{2}}} e^{-\frac{1}{2}(\mathbb{X}-\mu)^T \Sigma^{-1}(\mathbb{X}-\mu)} d\mathbb{X}. \quad (2.3)$$

Po zahrnutí predpokladov  $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  a  $\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$ , čo predstavuje situáciu, že streľci mieria do stredu terča a rozptyl výstrelov je rovnaký v ľubovoľnom smere od stredu (homoskedasticita), dosadením do (2.3) dostávame

$$\iint_{A_j} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\frac{\mathbb{X}^T \mathbb{X}}{\sigma^2}} d\mathbb{X} = \frac{1}{2\pi\sigma^2} \iint_{A_j} e^{-\frac{1}{2}\frac{x_1^2+x_2^2}{\sigma^2}} d\mathbb{X}. \quad (2.4)$$

Využitím polárnej substitúcie  $x_1 = r \cos \varphi$ ,  $x_2 = r \sin \varphi$ , kde  $\varphi \in (0, 2\pi)$ ,  $r \in (r_j, r_{j+1})$ ,  $|J| = r$  získame

$$\frac{1}{2\pi\sigma^2} \int_0^{2\pi} d\varphi \int_{r_j}^{r_{j+1}} r e^{-\frac{r^2}{2\sigma^2}} dr = \frac{1}{\sigma^2} \int_{r_j}^{r_{j+1}} r e^{-\frac{r^2}{2\sigma^2}} dr. \quad (2.5)$$

Nakoniec použijeme substitúciu  $t = -\frac{r^2}{2\sigma^2}$ ,  $dt = -\frac{r}{\sigma^2} dr$  z čoho dostaneme explicitnú formulu pre pravdepodobnosť zásahu medzikružia:

$$P(x \in A_j) = \int_{-\frac{r_j^2}{2\sigma^2}}^{-\frac{r_{j+1}^2}{2\sigma^2}} e^t dt = e^{-\frac{r_j^2}{2\sigma^2}} - e^{-\frac{r_{j+1}^2}{2\sigma^2}}. \quad (2.6)$$

### 2.2.1 Odhad $\sigma^2$ pomocou MMV

Na odhad parametra  $\sigma^2$  použijeme metódu maximálnej vierohodnosti, ktorej odhad budeme označovať  $\hat{\sigma}^2$ .

**Vlastnosti odhadov MMV**

- asymptotická nevychýlenosť:

$$E(\hat{\sigma}^2) \longrightarrow \sigma^2 \text{ pre } n \longrightarrow \infty \quad (2.7)$$

- asymptotická efektívnosť:

$$\hat{\sigma}^2 \longrightarrow N(\sigma^2, \Sigma) \text{ pre } n \longrightarrow \infty, \quad (2.8)$$

kde  $\Sigma$  je kladne definitná matica zhodná s obrátenou hodnotou Fischerovej informácie  $I^{-1}(\sigma^2)$ , ako ukážeme v kapitole (2.4).

- konzistentnosť:

$$\forall \varepsilon \geq 0 : \lim_{n \rightarrow \infty} (|\hat{\sigma}^2 - \sigma^2| \geq \varepsilon) = 0 \quad (2.9)$$

Odhad parametra získaného MMV je definovaný ako argument maxima funkcie viero-hodnosti, ktorá je definovaná ako súčin pravdepodobností nastatia nami nameraných udalostí:

$$L(x_1, \dots, x_n, \sigma^2) = \prod_i^n P(X_i = x_i). \quad (2.10)$$

Náhodná premenná  $X_i$  v našom prípade predstavuje medzikružie zasiahnuté  $i$ -tým šípom.  $P(X_i = x_i)$  je pravdepodobnosť realizácie  $x_i$ . Označme si funkciu  $P_j(\sigma^2)$  ako pravdepodobnosť zasiahnutia  $j$ -teho medzikružia definovanú v kapitole (2.2), kde  $j \in S$ . Keďže to isté bodové hodnotenie môžeme dostať pri viacerých výstreloch, tieto pravdepodobnosti môžeme združiť:

$$L(x_1, \dots, x_n, \sigma^2) = \prod_{j \in S} (P_j(\sigma^2))^{n_j}, \quad (2.11)$$

kde  $n_j$  je počet zasiahnutí  $j$ -teho medzikružia. Pre zjednodušenie môžeme odteraz používať zlogaritmovanú  $l(x, \sigma^2)$  (*loglikelihood function*), pretože rastúca transformácia zachováva argument maxima:

$$l(x, \sigma^2) = \ln \left( \prod_{j \in S} (P_j(\sigma^2))^{n_j} \right) = \sum_{j \in S} n_j \ln (P_j(\sigma^2)). \quad (2.12)$$

Z podmienky prvého rádu pre bod extrémum platí:

$$\frac{\partial l(x, \hat{\sigma}^2)}{\partial \sigma^2} = \sum_{j \in S} n_j \frac{\frac{\partial P_j(\hat{\sigma}^2)}{\partial \sigma^2}}{P_j(\hat{\sigma}^2)} = 0. \quad (2.13)$$

Z predchádzajúceho odvodu máme

$$P_j(\hat{\sigma}^2) = e^{-\frac{r_j^2}{2\hat{\sigma}^2}} - e^{-\frac{r_{j+1}^2}{2\hat{\sigma}^2}}. \quad (2.14)$$

Potom

$$\frac{\partial P_j(\hat{\sigma}^2)}{\partial \sigma^2} = \frac{r_j^2}{2\hat{\sigma}^4} e^{-\frac{r_j^2}{2\hat{\sigma}^2}} - \frac{r_{j+1}^2}{2\hat{\sigma}^4} e^{-\frac{r_{j+1}^2}{2\hat{\sigma}^2}} \quad (2.15)$$

a od (2.13) dostaneme vzťah,

$$\sum_{j \in S} n_j \frac{\frac{r_j^2}{2\sigma_{ML}^4} e^{-\frac{r_j^2}{2\sigma_{ML}^2}} - \frac{r_{j+1}^2}{2\sigma_{ML}^4} e^{-\frac{r_{j+1}^2}{2\sigma_{ML}^2}}}{e^{-\frac{r_j^2}{2\sigma_{ML}^2}} - e^{-\frac{r_{j+1}^2}{2\sigma_{ML}^2}}} = 0, \quad (2.16)$$

z ktorého vieme numericky dopočítať odhad variancie strelca metódou maximálnej vierohodnosti.

### 2.2.2 Odhad $\sigma^2$ pomocou momentovej metódy

Odhad parametra  $\sigma^2$  možno skonštruovať aj postupom uvedeným v [7]. Idea momentovej metódy sformulovanej v [13] spočíva v porovnaní očakávaných (teoretických) a nameraných momentov  $M_k(X)$ :

$$M_k(X) = E(X^k), \quad (2.17)$$

kde  $E(X^k)$  je  $k$ -ty očakávaný moment. Použitím výberového priemeru, ktorý je podľa [12] najlepším lineárnym nevychýleným odhadom, ako odhadu prvého momentu získavame rovnosť:

$$M_1(X) = \bar{X} = \frac{Z}{N} = E(X), \quad (2.18)$$

kde  $Z$  nameraný súčet bodov, ktoré strelec dosiahol za všetkých  $N = 72$  výstrelcov. Pravdepodobnosť nadobudnutia bodového ohodnotenia  $h(j)$ , čo je ekvivalentné s pravdepodobnosťou zasiahnutia medzikružia  $A_j$  odvodenou v (2.6) je:

$$P(x \in A_j) = P_j(\sigma^2) = e^{-\frac{r_j^2}{2\sigma^2}} - e^{-\frac{r_{j+1}^2}{2\sigma^2}}, \quad (2.19)$$

kde množinu zásahov sme zadefinovali v kapitole (2.1) ako

$$S = \{1, 2, 3, \dots, 10, \times, M\}.$$



Zónu  $j$  zasiahneme s pravdepodobnosťou  $P_j(\sigma^2)$  a za každý výstrel získame  $h(j)$  bodov, potom z definície strednej hodnoty:

$$E(X) = \sum_{j \in S} h(j)P_j(\sigma^2). \quad (2.20)$$

Po využití (2.18) a dosadení za  $P_j(\sigma^2)$  z (2.19) vzniká

$$Z = N \sum_{j \in S} h(j) \left( e^{-\frac{r_j^2}{2\sigma^2}} - e^{-\frac{r_{j+1}^2}{2\sigma^2}} \right), \quad (2.21)$$

čím získavame implicitný vzťah pre odhad  $\tilde{\sigma}^2$  parametra  $\sigma^2$  pomocou momentovej metódy.

### 2.3 Odvodenie pravdepodobnosti zásahu medzikružia bez predpokladu nevychýlenosti

V kapitole (2.2) sme predpokladali, že strelci sa v priemere triafajú presne do stredu terča. V tejto kapitole budeme analyzovať situáciu, ktorá nastane po uvoľnení tohto predpokladu. Predpokladajme, že výstrely strelca sú mierne vychýlené, či už vplyvom vonkajších podmienok, alebo čisto vlastnosťami strelca. Pokúsime sa skonštruovať model, ktorý by vychýlenosť od stredu dokázal popísať a zahrnúť do výpočtov. Predpokladajme, že výstrely sú opäť z normálneho rozdelenia, situované rovnomerne okolo bodu  $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ , kde  $\mu_1$  a  $\mu_2$  predstavujú výchylky od osí  $x_1$  a  $x_2$ . Pre súradnice zasiahnutého bodu platí

$$\mathbb{X} \sim \mathbb{N} \left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right). \quad (2.22)$$

Pre zložky vektora  $\mathbb{X}$  platí

$$x_i \sim N(\mu_i, \sigma^2); i \in \{1, 2\}, \quad (2.23)$$

čo môžeme upraviť na

$$\frac{x_i}{\sigma} \sim N(\mu_i, 1); i \in \{1, 2\}. \quad (2.24)$$

Náhodné premenné  $x_i$  a následne aj  $\frac{x_i}{\sigma}$  sú nezávislé z normálneho rozdelenia. Potom pre súčet ich kvadrátov platí

$$\frac{x_1^2}{\sigma^2} + \frac{x_2^2}{\sigma^2} \sim \chi_{2,\lambda}^2, \quad (2.25)$$

kde  $\chi_{2,\lambda}^2$  je necentrálne chí-kvadrát rozdelenie s dvomi stupňami voľnosti a parametrom necentrálnosti  $\lambda = \mu_1^2 + \mu_2^2$  definované v [9]. Všimnime si, že premenná  $x_1^2 + x_2^2$  v skutočnosti predstavuje polomer  $r^2$ . Parameter  $\sqrt{\lambda} = \sqrt{\mu_1^2 + \mu_2^2}$  sa tiež dá interpretovať ako odchýlka od stredu súradnicovej sústavy. Funkciou  $F_\lambda(x)$  označíme distribučnú funkciu necentrálneho chí-kvadrát rozdelenia s dvomi stupňami voľnosti a parametrom necentrálnosti  $\lambda$ :

$$F_\lambda(x) = e^{-\frac{\lambda}{2}} \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{2})^i}{i!} G_{2+2i}(x), \quad (2.26)$$

kde  $G_k(x)$  je distribučná funkcia centrálného chí-kvadrát rozdelenia s  $k$  stupňami voľnosti:

$$G_k(x) = \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}. \quad (2.27)$$

Funkcia  $\Gamma(t)$  je Gamma funkcia definovaná v [1] ako:

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx, \quad (2.28)$$

funkcia  $\gamma(a, t)$  predstavuje dolnú nekompletnú (*lower incomplete*) Gamma funkciu:

$$\gamma(t, a) = \int_0^a x^{t-1} e^{-x} dx, \quad (2.29)$$

Podobne pre hustotu necentrálneho chí-kvadrátu platí

$$f_\lambda(x) = e^{-\frac{\lambda}{2}} \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{2})^i}{i!} g_{2+2i}(x) = e^{-\frac{\lambda}{2}} \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{2})^i}{i!} \frac{x^i e^{-\frac{x}{2}}}{2^{i+1} \Gamma(i+1)}, \quad (2.30)$$

kde  $g_k(x)$  je hustota centrálného chí-kvadrát rozdelenia s  $k$  stupňami voľnosti:

$$g_k(x) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}. \quad (2.31)$$

Pre pravdepodobnosť zasiahnutia  $j$ -teho medzikružia  $A_j$  platí

$$P(x \in A_j) = P(r_j^2 \leq x_1^2 + x_2^2 \leq r_{j+1}^2) = P\left(\frac{r_j^2}{\sigma^2} \leq \frac{x_1^2 + x_2^2}{\sigma^2} \leq \frac{r_{j+1}^2}{\sigma^2}\right). \quad (2.32)$$

Pravdepodobnosť z (2.32) môžeme upraviť na

$$P\left(\frac{x_1^2 + x_2^2}{\sigma^2} \leq \frac{r_{j+1}^2}{\sigma^2}\right) - P\left(\frac{x_1^2 + x_2^2}{\sigma^2} < \frac{r_j^2}{\sigma^2}\right) \quad (2.33)$$

Využitím (2.25) dostávame

$$P(x \in A_j) = F_\lambda\left(\frac{r_{j+1}^2}{\sigma^2}\right) - F_\lambda\left(\frac{r_j^2}{\sigma^2}\right). \quad (2.34)$$

Podobne ako v kapitole (2.2.1) budeme za účelom odhadu parametrov  $\sigma^2$  a  $\lambda$  maximalizovať zlogaritmovanú funkciu vierohodnosti

$$l(x, \sigma^2) = \sum_{j \in S} n_j \ln(P_j(\sigma^2, \lambda)), \quad (2.35)$$

kde  $P_j(\sigma^2, \lambda)$  bude pravdepodobnosť zasiahnutia  $A_j$  z (2.34). Dosadením dostávame funkciu

$$l(x, \sigma^2) = \sum_{j \in S} n_j \ln\left(F_\lambda\left(\frac{r_{j+1}^2}{\sigma^2}\right) - F_\lambda\left(\frac{r_j^2}{\sigma^2}\right)\right). \quad (2.36)$$

Maximalizáciou získame odhady pre  $\sigma^2$  a  $\lambda$ , ktoré si označíme ako  $\bar{\sigma}^2$  a  $\bar{\lambda}$ . Presnosť streľca bude teda charakterizovať dvojica parametrov  $[\lambda, \sigma^2]$ . Analyzovaný model korešponduje s označeniami v [3], pričom parameter  $\lambda$  možno stotožniť s výrazom *accuracy* popisujúcim vzdialenosť stredu terča od bodu, okolo ktorého sú výstrely sústredené. Výraz *precision* hovorí o miere rozptýlenia zásahov  $\sigma^2$  okolo tohoto bodu. Odhady  $[\bar{\lambda}, \bar{\sigma}^2]$  sa už nebudú dať vyjadriť implicitne, ako pri odhade MMV a momentovou metódou pri predpoklade nevychýlenosti v kapitole (2.2). Pre numerické riešenie využijeme metódy nelineárnej optimalizácie zabudované v programe R, konkrétne funkciu  $nlm(f(x), x_0)$ , ktorá hľadá minimá funkcie  $f(x)$  s počiatočným bodom optimalizácie  $x_0$ .

## 2.4 Výpočet Fisherovej informácie

**Veta 2.1** (Cramér–Raova nerovnosť[10]). *Nech  $f_\theta(x)$  je hustota pravdepodobnosti spojitá v  $\theta$ ,  $\mathbb{X} = (x_1, x_2, \dots, x_n)$  je náhodný vektor so združenou hustotou  $g_\theta(\mathbb{X}) = \prod_{i=1}^n f_\theta(x_i)$  a  $\hat{\theta}(\mathbb{X}) = \hat{\theta}$  je nevychýlený odhad parametra  $\theta$ . Ak  $g_\theta(\mathbb{X})$  spĺňa nasledovné predpoklady:*

$$\frac{\partial}{\partial \theta} \int \hat{\theta} g_\theta(\mathbb{X}) d\mathbb{X} = \int \hat{\theta} \frac{\partial g_\theta(\mathbb{X})}{\partial \theta} d\mathbb{X},$$

*pričom sú splnené podmienky na deriváciu parametrického integrálu definované v [10] a pre každé  $\theta$  je disperzia  $D(\hat{\theta})$  konečná. Potom*

$$D(\hat{\theta}) \geq \frac{1}{nE\left(\left(\frac{\partial \log f_\theta(x)}{\partial \theta}\right)^2\right)}, \quad (2.37)$$

*kde výraz  $nE\left(\left(\frac{\partial \log f_\theta(x)}{\partial \theta}\right)^2\right)$  sa často označuje ako Fisherova informácia  $I(\theta)$ :*

$$D(\hat{\theta}) \geq \frac{1}{I(\theta)}. \quad (2.38)$$

V našom prípade budeme skúmať informáciu o presnosti strelca  $\sigma^2$  obsiahnutú v dátach. Tvar upravenej Cramér–Raovej nerovnosti bude

$$D(\hat{\sigma}^2) \geq \frac{1}{I(\sigma^2)} \quad (2.39)$$

Nerovnosť nám hovorí o kvalite nášho odhadu. Čím nižšia je hodnota  $D(\hat{\sigma}^2)$ , tým presnejšie možno z dát odhadnúť  $\sigma^2$ . Taktiež platí, že v nerovnosti (2.38) nastáva rovnosť práve vtedy, keď je náš odhad ten najlepší možný, t. j. efektívny. Z vlastností odhadu metódou maximálnej vierohodnosti (kap 2.2.1) však vieme, že odhad je asymptoticky efektívny (2.8) a pre dostatočne veľké  $n$  platí

$$\lim_{n \rightarrow \infty} D(\hat{\sigma}^2) = \frac{1}{I(\sigma^2)}. \quad (2.40)$$

Fischerova informácia nám hovorí o výpovednej hodnote dát, ktoré v našom prípade predstavujú bodový zisk strelca z každého výstrelu. Je daná vzťahom

$$I(\sigma^2) = E(\text{score}^2), \quad (2.41)$$

kde *score* je funkcia závislá od  $\sigma^2$  a bodovom ohodnotení  $x$ .

$$\text{score}(x, \sigma^2) = \frac{\partial \ln(L(x, \sigma^2))}{\partial \sigma^2} \frac{\partial^2 \ln(L(x, \hat{\sigma}^2))}{(\partial \hat{\sigma}^2)^2} = \frac{\partial^2 l(x, \hat{\sigma}^2)}{(\partial \hat{\sigma}^2)^2}, \quad (2.42)$$

Podľa [12] je za predpokladu, že je funkcia  $\ln(L(x, \sigma^2))$  dvakrát diferencovateľná a za istých podmienok regularity definovaných v [10], môžeme vzťah (2.41) nahradiť

$$I(\sigma^2) = -E \left( \frac{\partial^2 \ln(L(x, \sigma^2))}{(\partial \sigma^2)^2} \right). \quad (2.43)$$

Uvažujme model z kapitoly (2.2) s pravdepodobnosťou 2.6. Potom podľa (2.12) platí:

$$l(x, \sigma^2) = \sum_{j \in S} n_j \ln(P_j(\sigma^2)). \quad (2.44)$$

Derivovaním získame:

$$\frac{\partial l(x, \sigma^2)}{\partial \sigma^2} = \sum_{j \in S} n_j \frac{\partial P_j(\sigma^2)}{P_j(\sigma^2)}, \quad (2.45)$$

$$\frac{\partial^2 l(x, \sigma^2)}{(\partial \sigma^2)^2} = \sum_{j \in S} n_j \frac{\frac{\partial^2 P_j(\sigma^2)}{(\partial \sigma^2)^2} P_j(\sigma^2) - \left( \frac{\partial P_j(\sigma^2)}{\partial \sigma^2} \right)^2}{P_j^2(\sigma^2)}, \quad (2.46)$$

kde podľa (2.14) a (2.15)

$$P_j(\sigma^2) = e^{-\frac{r_j^2}{2\sigma^2}} - e^{-\frac{r_{j+1}^2}{2\sigma^2}} \quad (2.47)$$

a

$$\frac{\partial P_j(\sigma^2)}{\partial \sigma^2} = \frac{r_j^2}{2\sigma^4} e^{-\frac{r_j^2}{2\sigma^2}} - \frac{r_{j+1}^2}{2\sigma^4} e^{-\frac{r_{j+1}^2}{2\sigma^2}}. \quad (2.48)$$

Po druhej derivácii:

$$\frac{\partial^2 P_j(\sigma^2)}{(\partial \sigma^2)^2} = -\frac{r_j^2}{\sigma^6} e^{-\frac{r_j^2}{2\sigma^2}} + \left(\frac{r_j^2}{2\sigma^4}\right)^2 e^{-\frac{r_j^2}{2\sigma^2}} + \frac{r_{j+1}^2}{\sigma^6} e^{-\frac{r_{j+1}^2}{2\sigma^2}} - \left(\frac{r_{j+1}^2}{2\sigma^4}\right)^2 e^{-\frac{r_{j+1}^2}{2\sigma^2}}. \quad (2.49)$$

Po zjednodušení dostaneme

$$\frac{\partial^2 P_j(\sigma^2)}{(\partial \sigma^2)^2} = \frac{r_j^2}{\sigma^6} \left(\frac{r_j^2}{4\sigma^2} - 1\right) e^{-\frac{r_j^2}{2\sigma^2}} - \frac{r_{j+1}^2}{\sigma^6} \left(\frac{r_{j+1}^2}{4\sigma^2} - 1\right) e^{-\frac{r_{j+1}^2}{2\sigma^2}}. \quad (2.50)$$

Dosadením do

$$I(\sigma^2) = -E \left( \sum_{j \in S} n_j \frac{\frac{\partial^2 P_j(\sigma^2)}{(\partial \sigma^2)^2} P_j(\sigma^2) - \left(\frac{\partial P_j(\sigma^2)}{\partial \sigma^2}\right)^2}{P_j^2(\sigma^2)} \right) \quad (2.51)$$

získavame vzťah pre výpočet Fischerovej informácie, ktorý budeme riešiť simulačnými metódami.

### 3 Porovnávanie teoretickej kvality odhadov

#### 3.1 Monte Carlo simulácie

Na určenie kvality navrhnutých odhadov  $\sigma^2$  použijeme simulácie Monte Carlo (MCS), ktoré dokážu našu situáciu popísať napriek nedostatku dát. Simulácie sa využívajú najmä vtedy, keď model obsahuje priveľa neznámych parametrov. Výsledky simulácií dokážu bližšie opísať správanie sa modelu a jeho pravdepodobnostného rozdelenia. Princíp MCS spočíva v generovaní sád dát náhodne alebo na základe rozdelenia, podľa ktorého predpokladáme, že sa správajú. MCS využívajú závery centrálnej limitnej vety uvedej v [5] a slabý zákon veľkých čísel.

**Veta 3.1** (Slabý zákon veľkých čísel). *Uvažujme nekonečnú postupnosť  $x_1, x_2, \dots$  nezávislých, rovnako rozdelených, Lebesgue-ovskych integrovateľných náhodných premenlivých so strednými hodnotami  $E(x_1) = E(x_2) = \dots = \mu$ . Potom výberový priemer konverguje v pravdepodobnosti k strednej hodnote*

$$\bar{X}_n \xrightarrow[n \rightarrow \infty]{P} \mu. \quad (3.1)$$

*Ekvivalentne*

$$\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P(|\bar{X} - \mu| > \varepsilon) = 0 \quad (3.2)$$

Podrobné popísanie princípu MCS a iné príklady využitia sú uvedené v [2].

V našich modeloch predpokladáme, že výstrely strelca sa riadia dvojrozmerným normálnym rozdelením:

$$x \sim N(\mu, \Sigma), \quad (3.3)$$

kde  $X$  sú súradnice výstrelu,  $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  je stredná hodnota a  $\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$  kovariančná matica rozdelenia. V kapitole (2.2) sme odvodili pravdepodobnosť s akou strelec zasiahne bodovú zónu  $A$ :

$$P_j(x \in A) = e^{-\frac{r_j^2}{2\sigma^2}} - e^{-\frac{r_{j+1}^2}{2\sigma^2}}. \quad (3.4)$$

#### 3.2 Algoritmus odhadu $\sigma^2$

V ďalšom kroku zvolíme nasledujúci postup:

1. Diskretizácia  $\sigma^2$ :

Rovnomerným diskretizovaním  $\sigma^2$  na intervale  $I = (25, 400)$  vytvoríme  $d$ -počet diskretizácií  $\sigma_i^2$ . Rozsah intervalu  $I$  sme určili na základe presností reálnych strelcov počítaných v kapitole (4).

## 2. Generovanie dát:

Simulujeme 72 výstrelov imaginárneho  $i$ -teho strelca s presnosťou  $\sigma_i^2$ ;  $i \in \{1, \dots, d\}$ . Výstrely sú generované pomocou funkcie *sample()* v programe R (v. 3.2.2).

3. Výpočet odhadov  $\hat{\sigma}_i^2$  a  $\tilde{\sigma}_i^2$ :

Pomocou vzťahov

$$\sum_{j \in S} n_j \frac{\partial P_j(\hat{\sigma}^2)}{\partial \hat{\sigma}^2} = 0 \quad (3.5)$$

a

$$Z = N \sum_{j \in S} h(j) P_j(\tilde{\sigma}^2), \quad (3.6)$$

kde  $n_j$ ,  $N$  a  $Z$  sú známe hodnoty a  $\sigma_i^2$  je diskretizovaná premenná.

## 4. Opakovanie experimentov:

Parameter  $\sigma_i^2$  odhadneme pre všetkých  $d$  diskretizácií. Následne kroky (2.) a (3.)  $k$ -krát zopakujeme. Za  $k$  sme si v našich výpočtoch zvolili číslo 100.

## 5. Výpočet priemeru a štandardnej odchýlky odhadov:

pre každé z diskretizovaných  $\sigma_i^2$  máme  $k$  počet odhadov  $\hat{\sigma}_i^2$  a  $\tilde{\sigma}_i^2$ . Z nich spočítame priemery  $M(\sigma^2, N)$  závislé od počtu výstrelov  $N$  a presnosti  $\sigma^2$ . Použijeme skrátene značenie  $\widehat{M}$  pre odhad MMV a  $\widetilde{M}$  pre odhad momentovou metódou. Analogicky vypočítame štandardné odchýlky  $S^2(\sigma^2, N)$  so značím  $\widehat{S}^2$  a  $\widetilde{S}^2$ .

```

#1.1:Počítanie priemeru a odchylky sigmyB, sigmyML
for (k in 1:kexp) {      #počet experimentov
  for (i in 1:d) {      #počet delení v diskretizácii
    for (j in 1:12) {  #počet medzikruží
      prob[i,j]<-Psigmad(sigmad,i,j) #pravdepodobnosť zasiahnutia medzikružia
    }
    #generovanie dát
    X[i,]<-sample(x=c(11:0),size=72,replace=TRUE,prob=prob[i,])
    for (j in 1:12) {
      expn[i,j]<-length(X[i,][X[i,]==12-j])
    }
    #experimentálna sigmaML pre sigmad[i] v k-tom experimente
    expsigmaML[k,i]<-uniroot.all(PMLi, c(0, 500))
    #experimentálna sigmaB pre sigmad[i] v k-tom experimente
    expsigmaB[k,i]<-uniroot.all(PBi,c(0,500))
  }
}
for (i in 1:d) {
  MB[i]<-mean(expsigmaB[,i])
  S2B[i]<-var(expsigmaB[,i])
  MML[i]<-mean(expsigmaML[,i])
  S2ML[i]<-var(expsigmaML[,i])
}

```

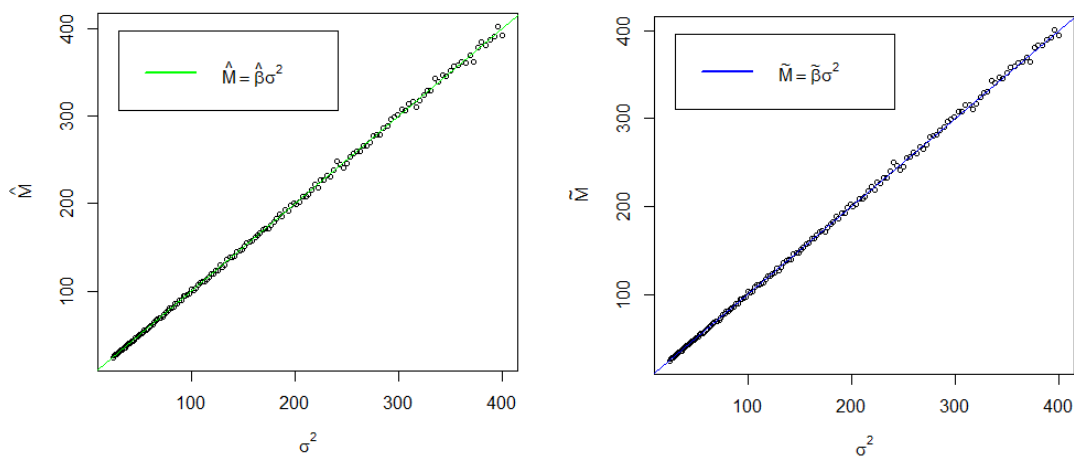
Obrázok 3: Výňatok zo zdrojového kódu programu používajúceho MCS

Obrázok 3 predstavuje implementáciu krokov 1-5 v prostredí R. Funkcia  $prob(x, size, replace, prob)$  zabezpečuje náhodný výber z  $x \in S$  s návratom. Vyberá 72krát s pravdepodobnosťou zasiahnutia jednotlivých bodových zón  $P_j(\sigma^2)$  odvodenou v časti (2.2). Funkcia  $uniroot.all()$  hľadá korene rovníc (3.5) a (3.6) na zvolenom intervale, funkcie  $mean()$  a  $var()$  počítajú priemery a štandardné odchýlky odhadov  $\sigma_i^2$  z  $k$  experimentov.

### 3.3 Výsledky MCS o kvalite odhadov

Odhady  $\hat{\sigma}_i^2$  a  $\tilde{\sigma}_i^2$  a následne aj ich priemery  $\widehat{M}$  a  $\widetilde{M}$  by sa mali nachádzať blízko pôvodnej  $\sigma_i^2$ .

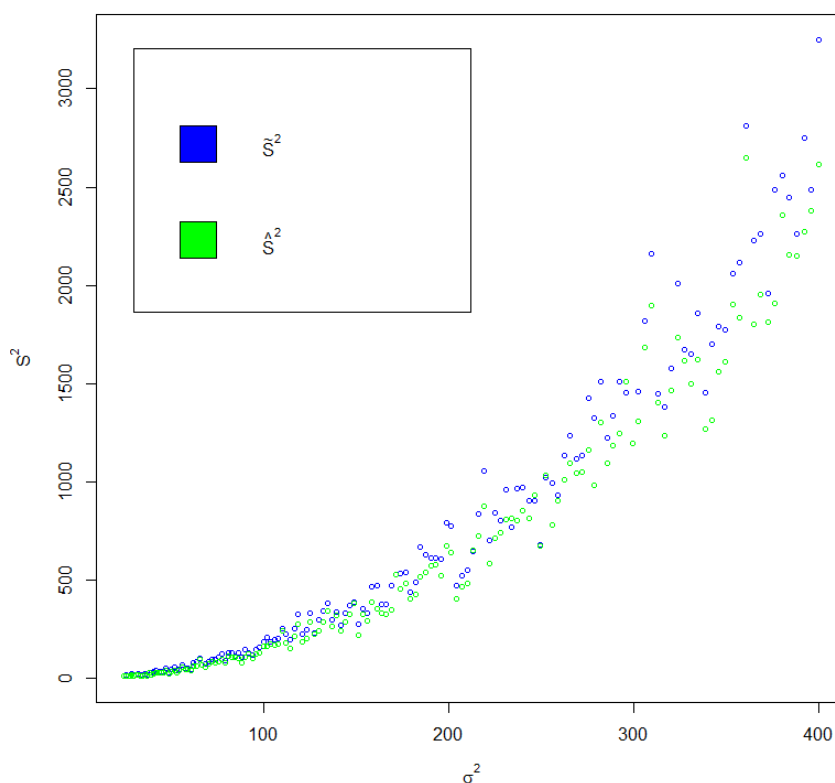




Obrázok 4:  $\hat{M}$  a  $\tilde{M}$  pri konštantnom počte výstrelov  $N = 72$

Na obrázku 4 vidíme, že už pre počet simulácií  $k = 100$  sa stredná hodnota odhadov pohybuje veľmi blízko k pôvodnej hodnote  $\sigma^2$ , z ktorej boli dáta generované. Parametre  $\beta$  sú odhadnuté koeficienty v lineárnej regresii a ich hodnoty su dopočítané v programe R:  $\hat{\beta} = 0.9979$ ,  $\tilde{\beta} = 0.9997$ .

Použitie MCS nám umožňuje skúmať aj disperzie odhadov, ktoré tiež charakterizujú kvalitu. Čím lepší odhad, tým si je istejší a jeho disperzia je nižšia. Pri MCS bude disperziu nahrádzať jej aproximácia, štandardná odchýlka  $S^2$ .



**Obrázok 5:** Závislosť štandardných odchýlok odhadov  $\hat{S}^2$  a  $\tilde{S}^2$  od  $\sigma^2$

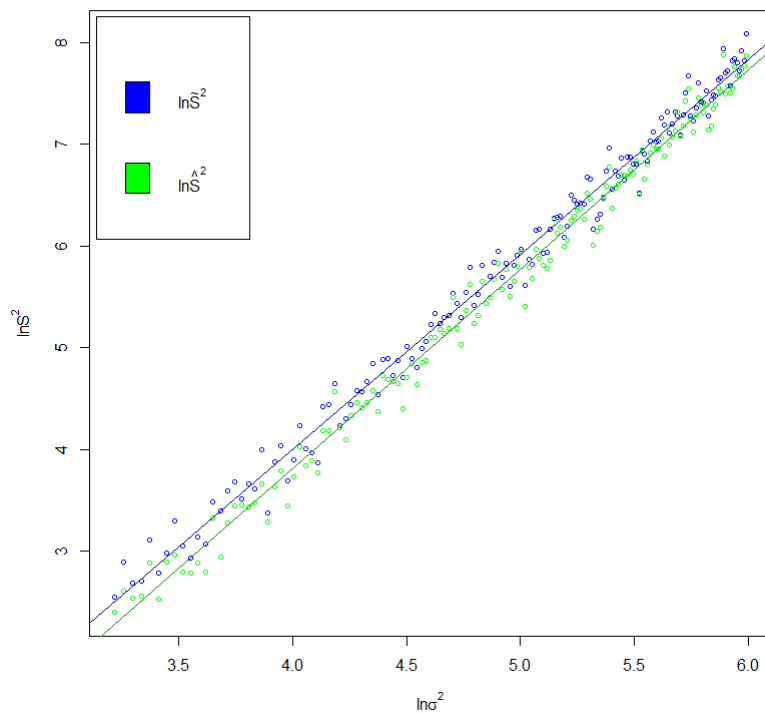
Z obrázku 5 je zrejmé, že odhad  $\hat{\sigma}^2$  je kvalitnejší, pretože je nižšia ako v prípade odhadu  $\tilde{\sigma}^2$ . Toto pozorovanie nie je prekvapivé, pretože  $\hat{\sigma}^2$  bol odhad metódou maximálnej vierohodnosti. Vzťah medzi  $\sigma^2$  a disperziou jej odhadov naznačuje polynomiálny rast. Očakávame, že štandardná odchýlka odhadov by sa mohla riadiť funkciou

$$S^2(\sigma^2) \approx (\sigma^2)^a, \quad (3.7)$$

kde koeficient  $a$  chceme odhadnúť. Zlogaritmuje obe strany (3.7):

$$\ln(S^2(\sigma^2)) \approx \ln(\sigma^2)^a = a \ln \sigma^2 \quad (3.8)$$

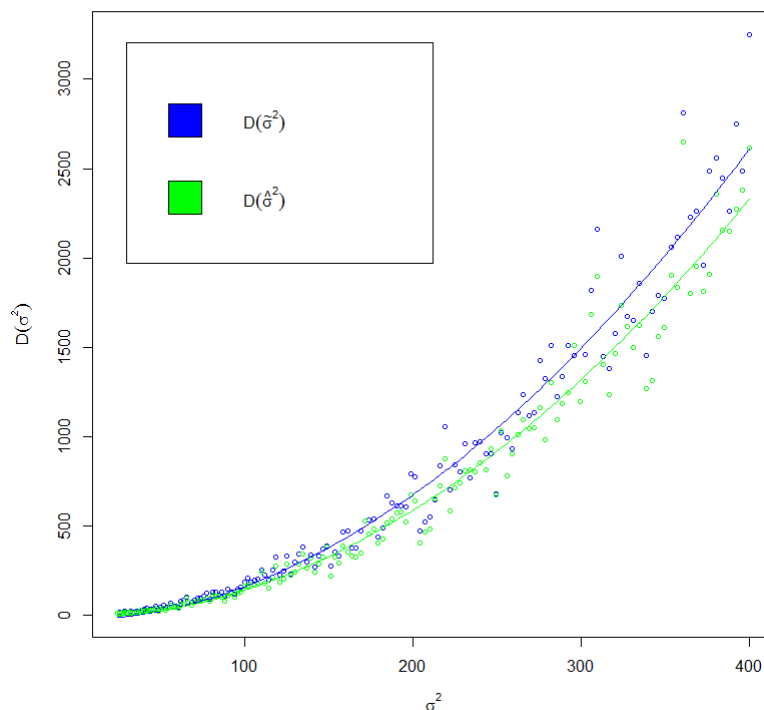
Na obrázku 6 je zobrazený vzťah medzi  $\ln \sigma^2$  a  $\ln(S^2(\sigma^2))$ .



**Obrázok 6:** Závislosť  $S^2(\sigma^2)$  od  $\sigma^2$

Podľa obrázku 6 je vzťah lineárny, čo nám potvrdzuje polynomiálny rast  $S^2(\sigma^2)$  od  $\sigma^2$ . Mocninu polynómu určíme použitím lineárnej regresie  $\ln(\sigma^2)$  na  $\ln(S^2(\sigma^2))$ , kde sklon regresnej priamky je

$$\ln(S^2(\sigma^2)) = a \ln \sigma^2. \quad (3.9)$$

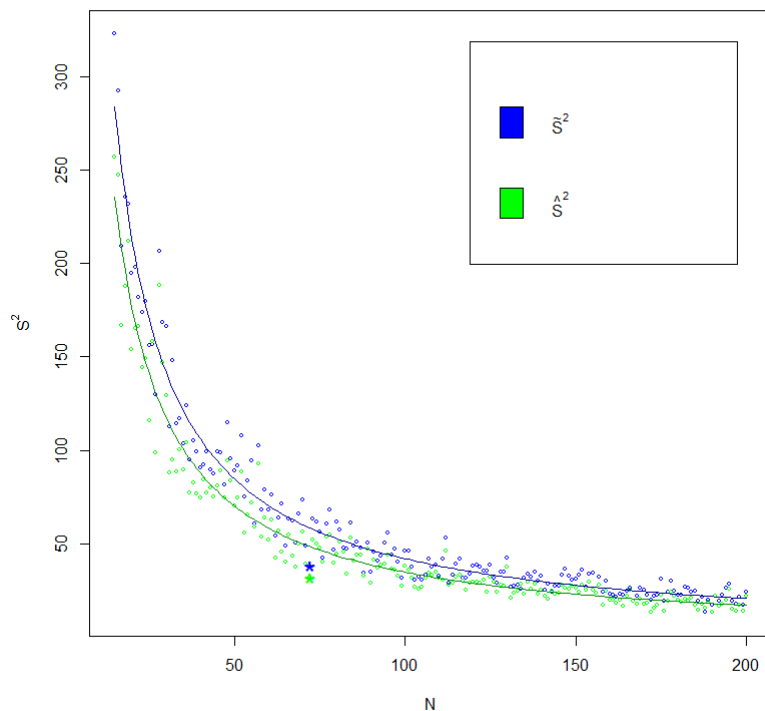


**Obrázok 7:** Závislosť disperzie odhadu od  $\sigma^2$  s regresnou krivkou

Regresné krivky na obrázku 7 sú určené rovnicami typu

$$S^2(\sigma^2) = b(\sigma^2)^a + c,$$

kde použijeme odhadnutý parameter  $a$ , koeficienty  $b$  a  $c$  sú konštanty odhanuté regresiou (3.7). Vypočítané koeficienty predpokladaných vzťahov sú uvedené v prílohe A. Vyššie uvedený postup použijeme pri viacerých reláciach, v ktorých bude domnievaný polynomiálny charakter. Následne vzniká nová motivácia sledovať vývoj disperzie v závislosti od počtu dát. Intuitívne by závislosť mala byť klesajúca funkcia, keďže zvýšením množstva dát majú odhady k dispozícii viacej informácií o strelcovi a jeho presnosti.



**Obrázok 8:** Závislosť  $S^2(\sigma^2)$  od  $N$

Generovanie dát na obrázku 8 prebiehalo rovnako ako v postupe na začiatku kapitoly (3.2). Diskretizovali sme však premennú  $N$  predstavujúcu počet výstrelů, zatiaľ čo vektor pravdepodobností bol fixovaný pre  $\sigma^2 = 56.25$ , za ktorú sme dosadili priemernú presnosť reálnych strelcov vypočítanú v kapitole (4). Regresné krivky predstavujú závislosť

$$S^2(\sigma^2) \approx \frac{1}{\sigma^2}.$$

Na obrázku 8 sú zvýraznené hodnoty odhadov  $S^2(\sigma^2)$  pre  $N = 72$ , čo predstavuje počet šířov vystrelených pri „malej fite“.

### 3.4 Fischerova informácia

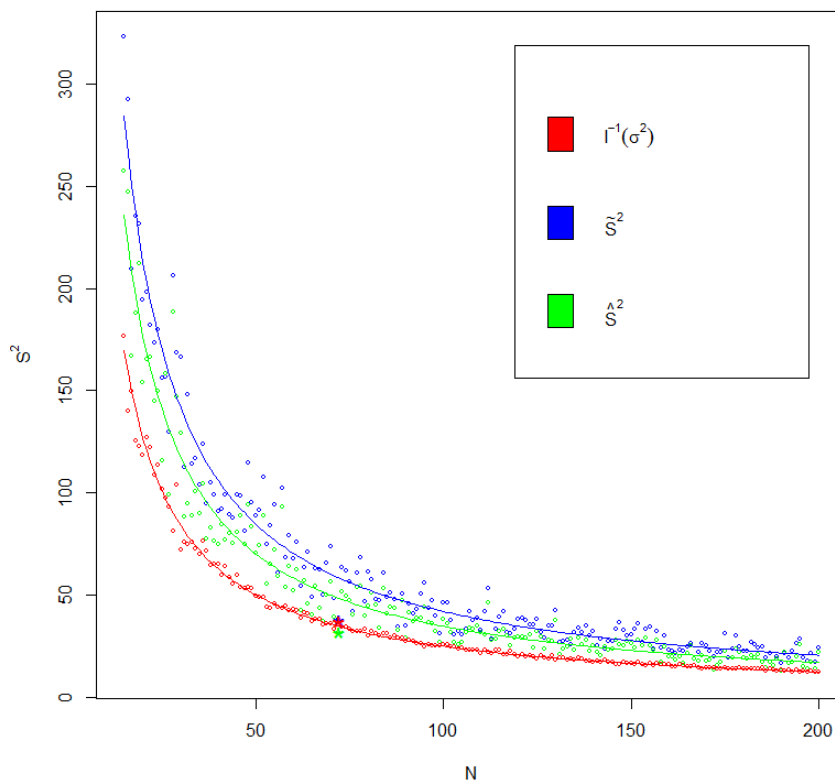
Monte Carlo simulácie možno použiť na výpočet Crámer-Raovej medze. V kapitole (2.4) sme mali definovaný vzťah

$$D(\hat{\sigma}^2) \geq \frac{1}{I(\sigma^2)}, \quad (3.10)$$

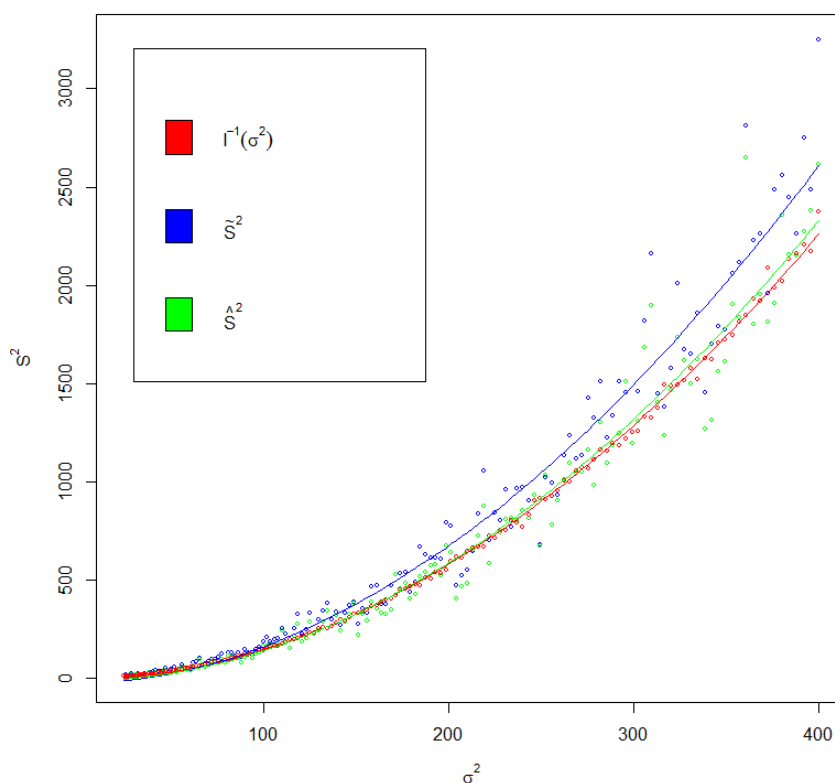
kde  $I(\sigma^2)$  predstavovala Fischerovú informáciu. V kapitole (2.4) sme uviedli aj vzorec pre jej výpočet:

$$I(\sigma^2) = -E \left( \frac{\partial^2 l(x, \sigma^2)}{(\partial \sigma^2)^2} \right). \quad (3.11)$$

Je zaujímavé pozorovať vzťah medzi teoretickou hranicou  $I^{-1}(\sigma^2)$  a štandardnými odchýlkami odhadov. V simuláciach v predchádzajúcej kapitole sme generovali dáta na odhad vývoja  $S^2(\sigma^2)$  pri zmene  $\sigma^2$  a  $N$ . Tie isté dáta môžeme použiť na načrtnutie Fischerovej informácie. V obrázkoch 9 a 10 je zobrazený vývoj  $S^2(\sigma^2)$  i  $I(\sigma^2)$ , respektíve  $I^{-1}(\sigma^2)$ .



**Obrázok 9:** Porovnanie  $I^{-1}(\sigma^2)$  a odhadov  $S^2(\sigma^2)$  v závislosti od  $N$



**Obrázok 10:** Porovnanie  $I^{-1}(\sigma^2)$  a odhadov  $S^2(\sigma^2)$  v závislosti od  $\sigma^2$

Na obrázkoch 9 a 10 si môžeme všimnúť, že  $I(\sigma^2)$  predstavuje dolné ohraničenie, na obrázku 9 možno pozorovať, že ohraničenie sa asymptoticky dosahuje. Tento fakt vyplýva aj z asymptotickej efektívnosti odhadu MMV. Je zjavné, že regresná krivka  $\tilde{S}^2$  je vyššie nad úrovňou  $I^{-1}(\sigma^2)$  ako krivka  $\hat{S}^2$ , čo znamená vyššiu odhadu. Treba poznamenať, že odhad  $\tilde{\sigma}^2$  je možné zostrojiť z nižšieho množstva dát. Zatiaľ čo  $\hat{\sigma}^2$  potreboje bodové ohodnotenie každého z  $N$  výstrelov, odhad  $\tilde{\sigma}^2$  používa celkové získané body za  $N$  výstrelov. Konkrétne, odhad  $\tilde{\sigma}^2$  využil 12-krát menej dát. Kvalitu odhadov budeme porovnávať na reálnych dátach v kapitole (4).

## 4 Porovnanie modelov na súťažných dátach

### 4.1 Dátové vstupy

V tejto kapitole budeme analyzovať dáta z lukostreleckých majstrovstiev a olympiád (pozri príloha B). Sústrediť sa budeme na skupinu strieľajúcich v kategórii a disciplíne definovanej v kapitole (1.1).

70m-1		1	2	3	Sum	Tot.		10+X	X
1	X	10	9	29			2	1	
	9	9	9	27	56	56			
2	X	X	10	30			3	2	
	9	9	9	27	57	113			
3	X	9	9	28			1	1	
	9	9	8	26	54	167			
4	X	10	9	29			2	1	
	9	9	8	26	55	222			
5	X	10	10	30			3	1	
	9	9	9	27	57	279			
6	10	9	9	28			1		
	9	8	8	25	53	332			
					<b>Total</b>	<b>332</b>	<b>12</b>	<b>6</b>	

70m-2		1	2	3	Sum	Tot.		10+X	X
1	10	9	9	28				1	
	8	8	8	24	52	52			
2	X	10	9	29			2	1	
	9	9	8	26	55	107			
3	X	10	10	30			3	1	
	9	9	9	27	57	164			
4	10	9	9	28			1		
	9	8	8	25	53	217			
5	X	10	9	29			2	1	
	9	9	9	27	56	273			
6	X	9	9	28			1	1	
	9	8	8	25	53	326			
					<b>Total</b>	<b>326</b>	<b>10</b>	<b>4</b>	
					<b>Signed Total</b>	<b>326</b>	<b>10</b>	<b>3</b>	

Obrázok 11: Ukážka súboru dát

Na obrázku 11 je ukážka súboru dát, ktorý sme použili na verifikáciu modelov. Ide o výsledky jedného strelca pri streľbe sady  $2 \times 36$  šípov na vzdialenosť 70 m. Z tabuliek výsledkov pre každého strelca využijeme stĺpce 1, 2 a 3, ktoré obsahujú nastrieľané bodové hodnoty. Analyzovať budeme celkovo 597 elitných strelcov zo siedmich rôznych súťaží konajúcich sa v roku 2015 (pozri príloha B). Budeme odhadovať ich presnosti pomocou rôznych modelov odvodených v kapitole (2). Následne otestujeme, ako dobre bola presnosť u jednotlivých strelcov odhadnutá.



## 4.2 Skutočné hodnoty pre jednotlivé odhady

V našej práci sme popísali celkovo tri odhady parametrov charakterizujúcich strelca:

- odhad  $\hat{\sigma}^2$  metódou maximálnej vierohodnosti charakterizovaný vzťahom:

$$\sum_{j \in S} n_j \frac{\frac{\partial P_j(\hat{\sigma}^2)}{\partial \sigma^2}}{P_j(\hat{\sigma}^2)} = 0. \quad (4.1)$$

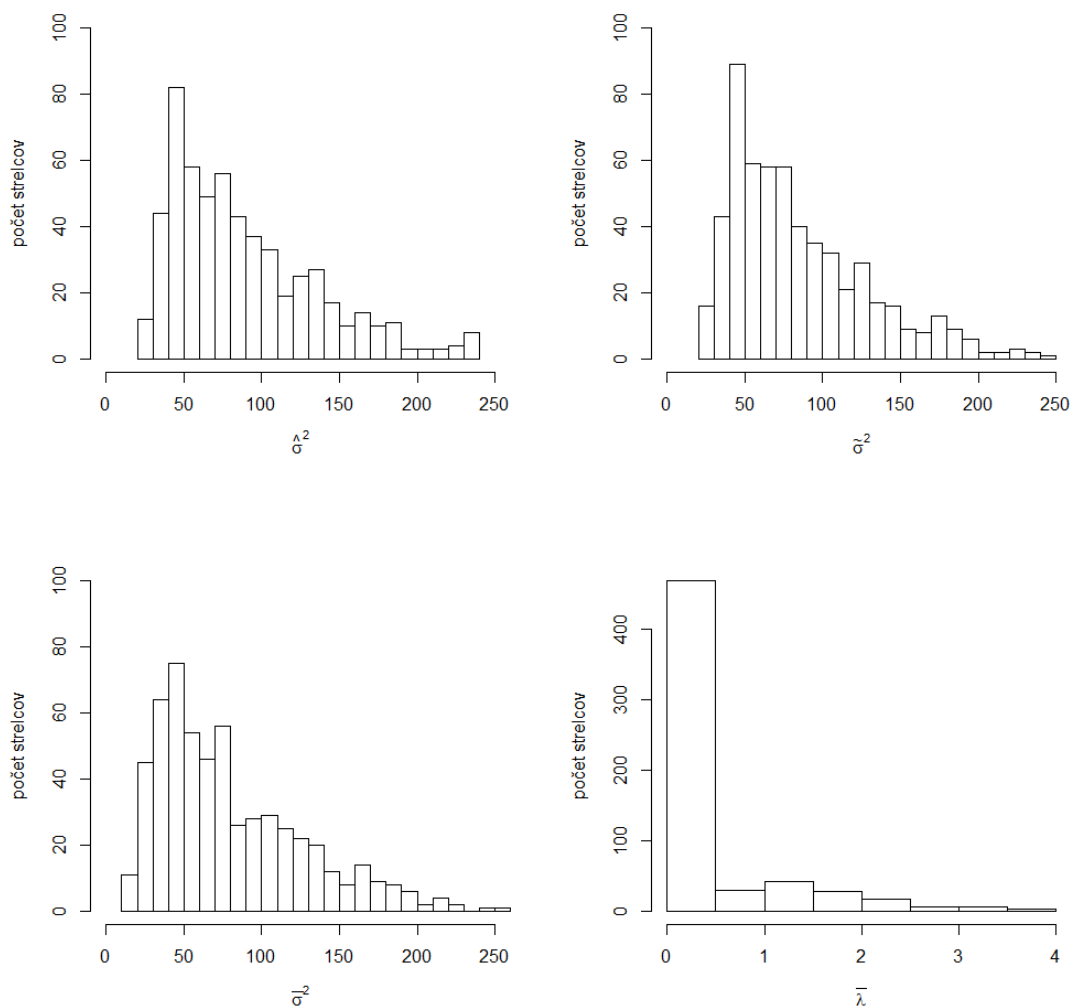
- odhad  $\tilde{\sigma}^2$  podľa [7] na základe očakávaného bodového zisku  $Z$ :

$$Z = N \sum_{j \in S} h(j) P_j(\tilde{\sigma}^2). \quad (4.2)$$

- dvojica odhadov  $[\bar{\lambda}, \bar{\sigma}^2]$  maximalizujúca funkciu vierohodnosti:

$$[\bar{\lambda}, \bar{\sigma}^2] = \operatorname{argmax} \left( \sum_{j \in S} n_j \ln \left( F_{\lambda} \left( \frac{r_{j+1}^2}{\sigma^2} \right) - F_{\lambda} \left( \frac{r_j^2}{\sigma^2} \right) \right) \right), \quad (4.3)$$

kde  $\sqrt{\lambda}$  predstavuje vychýlenie (v cm) od stredu terča a  $\sigma^2$  rovnomerný rozptyl výstrelů od bodu vychýlenia.



**Obrázok 12:** Histogramy odhadnutých  $\sigma^2$  strelcov

Všimnime si rozloženie strelcov na obrázku 12. Strelci v histograme pre  $\tilde{\sigma}^2$  vykazujú pozitívnejší sklon, čo naznačuje nižšie hodnoty  $\sigma^2$  ako pri odhade  $\hat{\sigma}^2$ . Dôvodom je fakt, že odhad pomocou celkového nástrelu je robustnejší ako odhad MMV, ktorý využíval bodové ohodnotenia každého výstrelu a je preto citlivý na veľmi presné alebo nepresné výstrely. Predstavme si strelca, ktorý zasiahne vysoké aj nízke bodové zóny. Iný strelca bude mať vyrovnané priemerné výstrely, pričom celkovo získa rovnaký počet bodov ako prvý strelca. Odhad momentovou metódou využíva iba súčet bodových ohodnotení a preto medzi oboma strelcami nezachytí rozdiel a ich presnosť  $\sigma^2$  odhadne rovnako. Odhad MMV ale tento rozdiel zaznamená a strelcovi s nevyrovnanými výstrelmi priradí vyššiu mieru nepresnosti. Odhad  $\tilde{\sigma}^2$  je preto robustnejší a nie je ovplyvnený extrém-

nymi hodnotami, čo by mohlo odôvodniť nazdávané nižšie hodnoty  $\tilde{\sigma}^2$ . Skonstruujeme hypotézu:

$$\mathbf{H0: } \tilde{\sigma}^2 \geq \hat{\sigma}^2$$

$$\mathbf{H1: } \tilde{\sigma}^2 < \hat{\sigma}^2$$

Odhady momentovou metódou sú v priemere vyššie ako odhady MMV  
 Odhady  $\tilde{\sigma}$  sú v priemere nižšie

Odhady  $\tilde{\sigma}$  a  $\hat{\sigma}^2$  sú závislé, pretože pochádzajú z dát o tých istých strelcoch. Na overenie hypotézy **H0** použijeme párový t-test definovaný spolu s podmienkami na možnosť jeho použitia v [12]. P-hodnota (*P-value*) vypočítaná v programe R nadobúda hodnotu  $\rho = 2,2 \cdot 10^{-16}$ . Párový t-test na hladine významnosti 5% zamietá hypotézu **H0** a potvrdzuje domnienku, že odhad  $\tilde{\sigma}^2$  nadobúda v priemere nižšie hodnoty ako  $\hat{\sigma}^2$ .

V modeli bez predpokladu nevychýlenosti sme získali dvojicu odhadov  $[\bar{\lambda}, \bar{\sigma}^2]$ . Odhad  $\hat{\sigma}^2$  je v skutočnosti iba špeciálnym podprípacom tohto modelu a maximalizuje tú istú funkciu vierohodnosti ako  $\bar{\sigma}^2$  pri predpoklade  $\lambda = 0$ . Odhad  $\hat{\sigma}^2$  predstavuje odchýlku šípov od stredu terča, zatiaľ čo  $\bar{\sigma}^2$  hovorí o rozptyle od bodu, okolo ktorého sú centrované. Preto pri strelcoch s nízkou mierou vychýlenosti vyšiel odhad  $\hat{\sigma}^2$  zhodný s  $\bar{\sigma}^2$ . Opačne, pre strelcov s vysokou mierou vychýlenosti  $\bar{\lambda}$  by mal byť odhad  $\hat{\sigma}^2$  výrazne vyšší ako  $\bar{\sigma}^2$ , čo potvrdujú aj naše merania.

### 4.3 Testy dobrej zhody

Na overenie, či naše modely dobre popisujú skutočné pravdepodobnostné rozdelenie použijeme testy dobrej zhody. Testy fungujú na princípe, že porovnávajú teoretickú distribúciu zodpovedajúcu predpokladom a empirickú distribúciu pozorovanú na dátovom súbore.

#### 4.3.1 Chí-kvadrát test

Chí-kvadrát test dobrej zhody porovnáva rozdelenie diskretných náhodných premenných s konečným počtom tried  $k$ . Test analyzuje rozdiely medzi očakávanými a pozorovanými hodnotami v každej z  $k$  tried a podľa [11] platí nasledujúca veta:

**Veta 4.1.** *Uvažujme súbor  $n$  nezávislých náhodných premenných  $x_1, \dots, x_n$  predstavu-*

júcich náhodný výber z množín  $A_1, \dots, A_k$  s pravdepodobnosťami

$$P(x_i \in A_j) = p_j, \quad j \in \{1, \dots, k\}.$$

Premenná  $v_j$  označuje nameraný počet udalostí  $x_i \in A_j$  pre  $i \in \{1, \dots, n\}$ . Potom náhodná premenná

$$\sum_{j=1}^k \frac{(v_j - np_j)^2}{np_j} \rightarrow \chi_{k-r}^2 \quad (4.4)$$

konverguje k chí-kvadrát rozdeleniu s  $k-r$  stupňami voľnosti, kde  $r$  je počet odhadovaných parametrov v modeli.

Podľa [6] je aproximácia chí-kvadrát rozdelením adekvátna, iba ak sú splnené nasledovné podmienky:

- počet nameraných udalostí  $v_j \geq 10$
- počet tried  $k \geq 3$
- očakávané hodnoty  $np_j \geq 0.25$

V práci pozorujeme profesionálnych strelcov, ktorý zasahujú prevažne iba prvé štyri medzikružia. Aby bola splnená prvá podmienka a počet zásahov každého medzikružia bol aspoň 10, zjednotíme zóny s nízkym počtom zasiahnutí do väčšej zóny. Upravené množiny označíme ako

$$B_j = \begin{cases} A_j; & j \in \{\times, 10, 9\} \\ \bigcup_{l=1}^8 A_l \cup A_M; & j = 8 \end{cases} \quad (4.5)$$

Premenná  $v_j$  bude v našom prípade znamenať nameraný počet zasiahnutí medzikružia  $B_j$ ;  $j \in \{\times, 10, 9, 8\}$ . Súčin  $np_j$  bude očakávaný počet zasiahnutí pri pravdepodobnosti  $P_j(x \in B_j)$  a počte výstrelcov  $N = 72$ . Dostávame

$$\sum_{j=1}^4 \frac{(v_j - NP_j(\sigma^2))^2}{NP_j(\sigma^2)} \rightarrow \chi_3^2 \quad (4.6)$$

pre odhady  $\tilde{\sigma}$ ,  $\hat{\sigma}^2$  a

$$\sum_{j=1}^4 \frac{(v_j - NP_j(\bar{\lambda}, \bar{\sigma}^2))^2}{NP_j(\bar{\lambda}, \bar{\sigma}^2)} \rightarrow \chi_2^2 \quad (4.7)$$

pre model bez predpokladu nevychýlenosti, ktorý odhadoval až dva parametre  $[\lambda, \sigma^2]$ .

Test zamietne zhodu s modelom, ak je rozdiel medzi očakávanými a nameranými početnosťami príliš veľký. V nasledujúcej tabuľke sú percentuálne podiely z 597 strelcov, ktorých odhady test na hladine významnosti 5% nezamietol:

Odhady	$\hat{\sigma}^2$	$\tilde{\sigma}^2$	$[\bar{\lambda}, \bar{\sigma}^2]$
Percento nezamietnutých	0.9280	0.9665	0.9129

### 4.3.2 G-test

Podobne ako v (4.8) podľa [8] platí

$$2 \sum_{j=1}^k v_j \ln \left( \frac{v_j}{np_j} \right) \rightarrow \chi_{k-r}^2. \quad (4.8)$$

Štatistika (4.8) sa často označuje G-štatistika, z čoho vznikol názov G-test. Narozdiel od chí-kvadrát testu, ktorý porovnáva rozdiely medzi očakávanými a nameranými hodnotami, G-test pozoruje logaritmus ich podielov, ktorý by tiež mal byť blízky k nule, ak model dobre popisuje skutočnú situáciu.

Po skonštruovaní testu na hladine významnosti 5% dostávame:

Odhady	$\hat{\sigma}^2$	$\tilde{\sigma}^2$	$[\bar{\lambda}, \bar{\sigma}^2]$
Percento nezamietnutých	0.9414	0.9765	0.9196

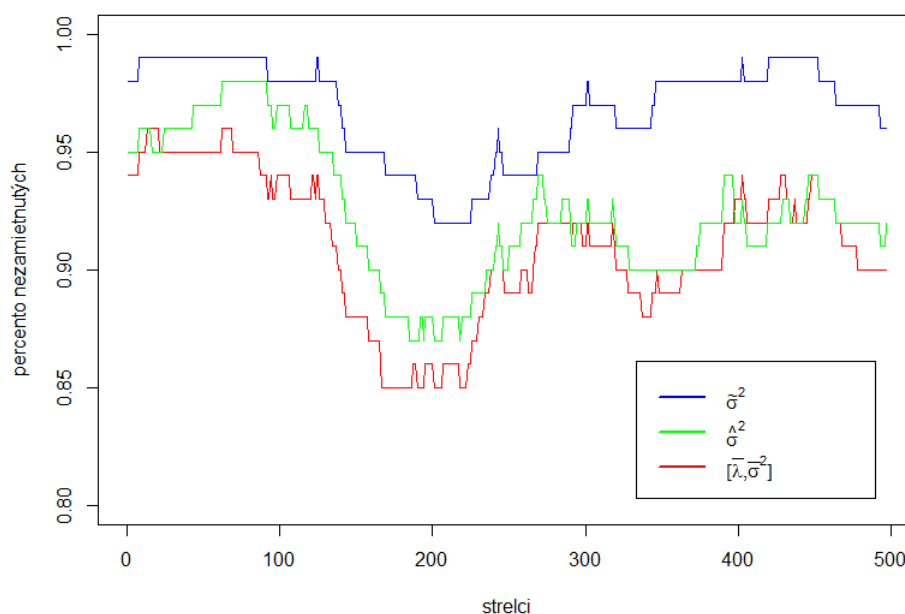
## 4.4 Porovnanie odhadov

Zaujímavé bolo, že napriek tomu, že model uvažujúci nevychýlenosť dát bol najvšeobecnejší, najlepší fit dát bol dosiahnutý najjednoduchším odhadom pomocou momentovej vety. Pre skúmaný počet strelcov 597 by tento výsledok nemal byť dielom náhody.

Možným vysvetlením je podobný postup odhadu  $\tilde{\sigma}^2$  a konštrukcie testov dobrej zhody. Kým odhad  $\tilde{\sigma}^2$  využíval zhodnosť očakávaných a nameraných celkových nastrieaných hodnôt, testy dobrej zhody pozorovali rozdiel medzi teoretickými a nameranými hodnotami v štyroch medzikružiach terča. Odhad momentovou metódou je tiež robustnejší menej ovplyvňovaný extrémnymi hodnotami. Nepriaznivé výsledky odhadov  $[\bar{\lambda}, \bar{\sigma}^2]$  môžu byť spôsobené zvýšenej ochotou zamietania testov dobrej zhody kvôli zvýšenému počtu parametrov  $r$  alebo z dôvodu numerických nepresností pri maximalizácii komplikovanej funkcie vierohodnosti (2.36). Iným zdrojom potenciál-

nej zmeny je pozorovanie miery zhody iba na štyroch medzikružiach  $B_j$ , zatiaľ čo odhady  $\hat{\sigma}^2$  a  $[\bar{\lambda}, \bar{\sigma}^2]$  brali do úvahy početnosti vo všetkých medzikružiach  $A_j$ .

Iným zaujímavým pozorovaním je, či sú niektoré z modelov vhodnejšie pre presnejších, respektíve menej presných strelcov. Zoraďme si strelcov podľa bodového zisku za všetkých 72 výstrelov od najvýkonnejšieho až po strelca s najnižším ziskom bodov. Postupne vyberajme skupiny 100 strelcov z tohto poradia a pozorujme percentuálnu početnosť strelcov nezamietnutých testami dobrej zhody.



**Obrázok 13:** Percentá nezamietnutých strelcov v skupinách zoradených podľa výšky bodového zisku

Na obrázku 13 sú zaznamenané výsledky chí-kvadrát testu pre skupiny 100 strelcov s postupne znižujúcim sa bodovým ziskom. Všimnime si, že odhad  $\hat{\sigma}^2$  dominuje hlavne u strelcov s nižšou výkonnosťou. Taktiež sa ukazuje istá korelácia v zamietaní medzi jednotlivými modelmi. V praxi to znamená, že modely zle popisujú, respektíve zlyhávajú na tých istých strelcoch.

Pozrime sa, ako dobre modely popisujú strelcov, u ktorých sa prejavila vychýlenosť od stredu terča. Strelcov, pre ktorých platí  $\bar{\lambda} \geq 0.05$  je dokopy 156. Percentuálne počty nezamietnutí sú:

Odhady	$\hat{\sigma}^2$	$\tilde{\sigma}^2$	$[\bar{\lambda}, \bar{\sigma}^2]$
Chí-kvadrát test	0.9679	0.9872	0.9679
G-test	0.9744	0.9808	0.9744

Pri popisovaní vychýlených strelcov si model bez predpokladu nevychýlenosti počína lepšie, avšak stále zaostáva za odhadom pomocou momentovej metódy. Všetky prístupy však veľmi dobre popisujú realitu. Pozrime sa na odhady z pohľadu výpočtovej náročnosti. V nasledujúcej tabuľke sú uvedené časové intervaly, za ktoré sú vypočítané odhady pre našu vzorku 597 strelcov:

Odhady	$\hat{\sigma}^2$	$\tilde{\sigma}^2$	$[\bar{\lambda}, \bar{\sigma}^2]$
Výpočtové časy	1.40 s	0.50 s	47.38 s
Čas odhadu 1. strelca	0.0023506 s	0.0008375 s	0.0793635 s

Všimnime si takmer trojnásobne kratšiu dĺžku trvania výpočtu odhadov  $\tilde{\sigma}^2$  oproti  $\hat{\sigma}^2$ . Znamená to, že odhad  $\tilde{\sigma}^2$  nevyhniká len mierou zhody, ale aj výpočtovou nenáročnosťou. Všetky výpočty a odhady boli vykonané na počítači s nasledovnými parametrami:

program	R (v. 3.2.2)
OS	Windows 7
CPU	Intel Core 2 Quad Q6600, 2.4 GHz
RAM	4 GB DDR2, 800 MHz

## Záver

V práci sme popísali niekoľko spôsobov odhadu presnosti lukostrelcov za rôznych predpokladov, menovite: odhad momentovou metódou a metódou maximálnej vierohodnosti. Skúmali sme model s predpokladom, že strelec mieri na stred terča. Neskôr sme zostrojili všeobecnejší model, ktorý predpoklad nevychýlenosti vypustil. Pre oba modely sme boli schopní vypočítať pravdepodobnosti zásahov medzikruží.

V kapitole (2) sme odvodili postup dvoch odhadov parametra presnosti pre model s predpokladom nevychýlenosti: odhad momentovou metódou  $\tilde{\sigma}^2$  a odhad metódou maximálnej vierohodnosti  $\hat{\sigma}^2$ . Pre model bez predpokladu nevychýlenosti presnosť charakterizovala dvojica odhadov  $[\bar{\lambda}, \bar{\sigma}^2]$ . V kapitole (3) sme pomocou Monte Carlo simulácií ohraničili teoretické kvality modelov Fischerovou informáciou. Teoreticky najkvalitnejší sa ukázal byť odhad  $\hat{\sigma}^2$ .

V kapitole (4) sme odhady porovnávali na dátach pochádzajúcich z rôznych lukostreleckých podujatí. Na overenie kvality modelov sme použili testy dobrej zhody, ktoré porovnávajú teoretické a empirické početnosti zásahov jednotlivých medzikruží terča. Pozorovali sme aj časy, za ktoré boli odhady vypočítané. Najvšeobecnejší odhad  $[\bar{\lambda}, \bar{\sigma}^2]$  bol podľa očakávaní výpočtovo najnáročnejší. Testy dobrej zhody nezamietli primeranosť odhadu vo vyše 91% zo vzorky takmer 600 strelcov. Odhad  $\hat{\sigma}^2$  testy nezamietli v cca 93%-tách prípadoch a jeho výpočtový čas bol 1.4 s. Odhad momentovou metódou bol z teoretického hľadiska najjednoduchší, čo sa prejavilo aj na dĺžke výpočtu, 0.5 s. Napriek tomu sa ukázal byť najlepší z hľadiska miery zhody teoretických a skutočných početností a dosiahol primeranú mieru zhody až v cca 97% prípadoch.

V budúcnosti by mohlo byť zaujímavé skúmanie spoločných vlastností strelcov, pri ktorých odhady zlyhali. Takýmto postupom by sa mohol podariť nájsť zdroj chybivosti pre každý druh odhadu. Zaujímavé by bolo zozbierať dáta od menej skúsených strelcov a porovnať kvalitu uvádzaných modelov na danej skupine. Ďalším možným vylepšením je zjednodušenie funkcie vierohodnosti pri modeli bez predpokladu nevychýlenosti. Vhodné by boli úpravy, ktoré by zvýšili rýchlosť jej konvergenencie, prípadne patrične zvolené počiatočné body nelineárnej optimalizácie.

Prínos práce vidíme hlavne v možnej implementácii v súťažnej lukostrelbe. Uvádzané metódy odhadu parametra presnosti môžu byť použité pri prepočte lukostreleckých



výkonov pri zmene vzdialenosti. S využitím popísaných modelov možno analyzovať rozhodovacie vlastnosti herných systémov. Spravodlivosť pravidiel by spočívala v umožnení výhry toho najlepšieho strelca a obmedzení prvku náhody. Charakterizovala by ju pravdepodobnosť víťazstva najpresnejšieho strelca.

## Zoznam použitej literatúry

- [1] Abramowitz M., Stegun I.A.: *Handbook of Mathematical Functions*, Dover (1972)  
Section 26.4.25
- [2] James F.: Monte Carlo theory and practice, Reports on Progress in Physics, (1980),  
Volume 43, Number 9
- [3] Ertan Hyari: *Exploratory spatial analysis of hit distribution in archery*. International Journal of Academic Research Part A, (2013), Vol. 5(6), 112-118
- [4] Hurban V., Durný M.: *Stanovy a pravidlá FITA*, pravidlá a stanovy, Slovenský lukostrelecký zväz, Bratislava, 2010, dostupné na internete (8.5.2016):  
[http://www.slz.sk/images/dokumenty/2010/FITARules2010\\_SVK\\_1.pdf](http://www.slz.sk/images/dokumenty/2010/FITARules2010_SVK_1.pdf)
- [5] Janková Katarína: *Pravdepodobnosť a štatistika*, učebné texty, (2014)
- [6] Koehler Kenneth J., Larnzt Kinley: *An Empirical Investigation of Goodness-of-Fit Statistics for Sparse Multinomials*, Journal of the American Statistical Association Vol. 75, No. 370 (Jun., 1980), pp. 336-344
- [7] Kundracik František: *Prepočet skóre na rôznych terčoch*, interný materiál, (2015)
- [8] McDonald John H.: *G-test*, učebné texty, Department of Biological Sciences University of Delaware, dostupné na internete (21.11.2015):  
<http://www.dmi.units.it/inverniz/adir/statgtestgof.html>
- [9] Murison Bob: *Distribution Theory of the Non-Central Chi-Square*, učebné texty, dostupné na internete (30.4.2016): <https://turing.une.edu.au/stat354/notes/node47.html>
- [10] Suhasini Subba Rao: *Advanced Statistical Inference*, učebné texty, dostupné na internete (10.4.2016): <https://www.stat.tamu.edu/suhasini/teaching613/chapter1.pdf>
- [11] Panchenko Dmitry: *Pearson's Theorem*, učebné texty, Massachusetts Institute of Technology, dostupné na internete (21.11.2015): <http://ocw.mit.edu/>

courses/mathematics/18-443-statistics-for-applications-fall-2003/lecture-  
notes/lec23.pdf

- [12] Somorčík Ján: *Prednášky zo štatistických metód*, učebné texty, (2015)
- [13] Watkins Joseph C.: *Introduction to the Science of Statistics*, učebné texty, dostupné na internete (9.5.2016): <http://math.arizona.edu/~jwatkins/statbook.pdf>

## Príloha A

### Koeficienty regresných kriviek

V kapitole (3) sme pre vzťahy medzi disperziami, diskretizovanou presnosťou  $\sigma^2$  a počtom výstrelov  $N$  skonštruovali polynomiálne regresie a následne regresní krivky. V nasledujúcich tabuľkách uvedieme koeficienty, ktoré sme pri MCS vypočítali. Regresné modely odhadovali parametre vo vzťahoch tvaru:

$$D(\sigma^2) = b(\sigma^2)^a + c.$$

Tabuľka pre závislosť od  $\sigma^2$ :

	$b$	$a$	$c$	$R^2$	$\bar{R}^2$
$D(\hat{\sigma}^2)$	0.0186	1.959924	-12.9009	0.968	0.9678
$D(\tilde{\sigma}^2)$	0.02729	1.91561	-23.6973	0.9663	0.9661
$I^{-1}(\sigma^2)$	0.01955	1.946249	-6.95374	0.9988	0.9988
$I(\sigma^2)$	50.5	-1.946249	0.000212		

Za účelom porovnania je uvedený aj riadok pre  $I^{-1}(\sigma^2)$ . V stĺpci  $R^2$  sa nachádza koeficient determinácie, ktorý hovorí o kvalite modelu. Čím je bližší k číslu 1, tým lepšie popisuje a odhaduje reálne vzťahy medzi veličinami. Upravený (*adjusted*) koeficient determinácie v sebe zahŕňa aj počet parametrov, ktoré sme v modeli použili.

Podobne ako pre  $\sigma^2$  sformulujeme závislosť od  $N$ :

$$D(\sigma^2) = bN^a + c.$$

Tabuľka pre závislosť od  $N$ :

	$b$	$a$	$c$	$R^2$	$\bar{R}^2$
$D(\hat{\sigma}^2)$	3542.8624	-1	-0.5927	0.9459	0.9456
$D(\tilde{\sigma}^2)$	4271.1667	-1	-0.6794	0.9603	0.9601
$I^{-1}(\sigma^2)$	2486.1505	-1	0.2982	0.9959	0.9959
$I(\sigma^2)$	0.0004005	1	-0.0001168	0.9901	0.99

Všimnime si zaujímavú hodnotu parametra  $a$ . Regresný model nám tvrdí, že medzi Fisherovou informáciou  $I(\sigma^2)$  a počtom dát  $N$  existuje priama úmernosť, respektíve nepriama úmernosť medzi  $N$  a  $D(\sigma^2)$ . Pomocný regresný model nám v skutočnosti

odhadol koeficient  $a = -1.007$  pri oboch odhadoch, ale test o kontraste sformulovaný v [12] nezamietol hypotézu, že  $a = -1$ . Iný test o kontraste nezamietol hypotézu, že parameter  $c$  (*intercept*) je zanedbateľný. Regresné modely sa teda líšia iba v jedinom parametri  $b$ , ktorý predstavuje „výšku hladiny“ regresnej krivky  $N^{-1}$ . Koeficient  $b$  je najvyšší pre disperziu odhadu  $\tilde{\sigma}^2$ , čo naznačuje nižšiu kvalitu odhadu, avšak asymptoticky

$$\lim_{N \rightarrow \infty} D(\hat{\sigma}^2) = \lim_{N \rightarrow \infty} D(\tilde{\sigma}^2) \approx 0,$$

keďže parameter  $c$  je zanedbateľný. Dochádzame k záveru, že pre dostatočný počet výstrelov  $N$  sú oba odhady ekvivalentné.

## Príloha B

### Zoznam lukostreleckých súťaží

V kapitole (4) sme používali výsledky strelcov z nasledujúcich lukostreleckých podujatí:

1. European Games 2015, Baku, Azerbajdžan, 16.06-22.06.2015
2. Toronto 2015 Parapan Am Games, Toronto, Kanada, 08.08-10.08.2015
3. Archery World Cup 2015 - Stage 3, Wroclaw, Poľsko, 11.08-16.08.2015
4. World Archery Para Championships, Donaueschingen, Nemecko, 22.08-30.08.2015
5. Archery World Cup 2015 - Stage 4, Medellín, Kolumbia, 08.09-13.09.2015
6. Campionati Italiani Targa 2015, Torino, Taliansko, 25.09-27.09.2015
7. 8th European Club Team Cup, Riom, Francúzsko, 02.10-04.10.2015