Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra matematiky

BAKALÁŘSKÁ PRÁCE

Ověřování podpisů prezidentských kandidátů

Plzeň 2016 **Kateřina Filipová**

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím uvedených pramenů v seznamu literatury.

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

Poděkování

Ráda bych poděkovala vedoucímu své práce Mgr. Michalu Frieslovi, Ph.D. za jeho odborné vedení, vstřícný přístup a za jeho čas, který mi věnoval. Poděkování také patří rodině a přátelům, kteří mě podporovali nejen při psaní této práce ale i v dobách celého studia. Děkuji.

Abstrakt

Název práce: Ověřování podpisů prezidentských kandidátů

Katedra: Katedra matematiky

Vedoucí bakalářské práce: Mgr. Michal Friesl, Ph.D., KMA

Abstrakt: Bakalářská práce popisuje úlohu kontroly podpisů na petičních arších prezidentských kandidátů pro rok 2013 z pohledu matematické statistiky. Hlavním cílem je určit pravděpodobnosti chybného závěru (registraci či zamítnutí prezidentské kandidatury) při zákonem předepsaném postupu.

Klíčová slova: Prezidentské volby, pravděpodobnost, počet dodaných podpisů.

Abstract

Title: Verification of signatures presidential candidates

Department: Department of Mathematics

Supervisor: Mgr. Michal Friesl, Ph.D., KMA

Abstract: This bachelor's thesis describes the task of checking the signatures on the petition sheets of presidential candidates for 2013 in terms of mathematical statistics. The main objective is to determine the probability of an erroneous conclusion (registration or refusal presidential candidacy) according to the procedure prescribed by law.

Keywords: Presidental election, probability, the number of delivered signatures.

Obsah

Kapitola 1

Úvod

V lednu roku 2013 proběhly historicky první přímé volby prezidenta České republiky. Jedním ze způsobů, jak se mohl občan České republiky stát možným kandidátem, bylo odevzdání petičních archů obsahujících minimálně 50 000 podpisů občanů oprávněných volit. U těchto kandidátů, kteří odevzdali petiční archy s minimálně 50 000 podpisy, byla provedena Ministerstvem vnitra České republiky kontrola správnosti údajů pod peticí dle prováděcího zákona č. 275/2012 Sb., o volbě prezidenta republiky.

Cílem této bakalářské práce je určení pravděpodobnosti chybného závěru (registraci či zamítnutí prezidentské kandidatury) při zákonem předepsaném postupu pro různé hodnoty počtu neplatných, resp. platných podpisů v celkovém počtu dodaných podpisů pod peticí.

Kapitola 2

Řečí paragrafů

Jak již bylo řečeno, v lednu roku 2013 proběhly historicky první přímé volby prezidenta České republiky. Občané České republiky volili v prvním kole jednoho z devíti možných kandidátů. Prezidentským kandidátem se mohl stát občan České republiky, který má právo volit a dosáhl věku 40 let, jež byl nominován minimálně dvaceti poslanci parlamentu České republiky nebo minimálně deseti senátory České republiky nebo občanem České republiky starším 18 let, který dodal petiční archy s minimálně 50 000 podpisy občanů oprávněných volit. U prezidentských kandidátů, kteří odevzdali petici s minimálně 50 000 podpisy, byla provedena kontrola správnosti údajů pod peticí. Ministerstvo vnitra České republiky ověřovalo správnost údajů na dodaných petičních arších s minimálně 50 000 podpisy. Kontrola správnosti údajů na petici byla provedena dle prováděcího zákona č. 275/2012 Sb., o volbě prezidenta republiky, celé znění v [5], dle § 25, odstavce 5 a 6:

> „Ministerstvo vnitra ověří správnost údajů na peticích namátkově na náhodně vybraném vzorku údajů u 8 500 občanů podepsaných na každé petici. Zjistí-li nesprávné údaje u méně než 3 % podepsaných občanů, nezapočítá Ministerstvo vnitra tyto občany do celkového počtu občanů podepsaných na petici.“,

> „Zjistí-li Ministerstvo vnitra postupem podle odstavce 5 nesprávné údaje u 3 % nebo více než 3 % podepsaných občanů, provede kontrolu u dalšího vzorku stejného rozsahu (dále jen "druhý kontrolní vzorek"). Zjistí-li Ministerstvo vnitra, že druhý kontrolní vzorek vykazuje chybovost u méně než 3 % občanů podepsaných na petici, nezapočítá Ministerstvo vnitra občany z obou kontrolních vzorků do celkového počtu občanů podepsaných na petici. Zjistí-li Ministerstvo vnitra, že druhý kontrolní vzorek vykazuje chybovost u 3 % nebo více než 3 % občanů podepsaných na petici, odečte od celkového počtu občanů podepsaných na petici počet občanů, který procentuálně odpovídá chybovosti v obou kontrolních vzorcích.“

Zákon jasně nepopisuje, zda se první výběr vrací či nevrací zpět do celkového počtu podpisů pod peticí před realizací druhého výběru. Avšak více logické je, první výběr nevracet. Předpokládáme tedy, že když se v prvním vzorku zjistí nesprávnost údajů u 3 % nebo více občanů podepsaných pod peticí, první výběr se nevrací zpět a druhý kontrolní vzorek se vybírá už jen z celkového počtu dodaných podpisů bez prvního výběru. Díky tomu se druhý výběr stává závislý na výsledku prvního výběru.

Kapitola 3

Základní pojmy

3.1 Pravděpodobnost

Mějme množinu Ω, což je množina všech možných výsledků náhodného pokusu. *Náhodný pokus* je každý děj, jehož výsledek není předem jednoznačně určen podmínkami, za kterých probíhá. Výsledky tohoto náhodného pokusu se nazývají *elementární jevy* ω. Potom je tedy množina Ω množinou elementárních jevů ω. Většinou se nezabýváme konkrétními jednotlivými výsledky, ale nějakou jejich množinou, tj. podmnožinou množiny Ω. Každá podmnožina množiny Ω se nazývá *náhodný jev*.

Mějme množinu Ω. Pak σ - *algebrou* $\mathcal{A}$ nazveme neprázdný systém podmnožin množiny Ω splňující:

- $\Omega \in \mathcal{A}$,
- $A \in \mathcal{A} \Rightarrow \Omega \setminus A \in \mathcal{A}$,
- $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_i A_i \in \mathcal{A}$.

Náhodným jevům se připisují určité pravděpodobnosti pomocí pravděpodobnostních měr P. *Pravděpodobnostní míra* je reálná funkce, která je definovaná na σ - algebře náhodných jevů v množině elementárních jevů. Pravděpodobnostní míru lze také definovat jako zobrazení

$$P : \mathcal{A} \rightarrow \langle 0, 1 \rangle$$

takové, pro něž platí:

- $P(\Omega) = 1$,
- $P(\emptyset) = 0$,
- pro po dvou disjunktní (též neslučitelné, nemající společný prvek) množiny $A_1, A_2, \dots \in \mathcal{A}$, tj. $P(A_i \cap A_j) = \emptyset$ pro $i \neq j$, platí:

$$P(\bigcup_i A_i) = \sum_i P(A_i) \tag{3.1}$$

Trojice $(\Omega,\ \mathcal{A},\ P)$ se nazývá *pravděpodobnostní prostor*.

3.1.1 Podmíněná pravděpodobnost

Podmíněná pravděpodobnost jevu A za podmínky, že nastal jev B, s $P(B) > 0$, je definována jako

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{3.2}$$

kde P značí nepodmíněnou pravděpodobnost.

3.1.2 Věta o úplné pravděpodobnosti

Nechť pro jevy $H_1, H_2, \dots$ platí:

- náhodné jevy jsou po dvou disjunktní, tj. $P(H_i \cap H_j) = \emptyset \quad$ pro $i \neq j$
- $\bigcup_i H_i = \Omega$
- $P(H_i) > 0$ pro $i = 1, 2, \dots$

Pak pro každý jev $A \subset \Omega$ platí:

$$P(A) = \sum_i \underbrace{P(A \mid H_i)P(H_i)}_{P(A \cap H_i)} = \sum_i P(A \cap H_i) \tag{3.3}$$

3.2 Náhodná veličina

Náhodná veličina X je funkce, která každému výsledku náhodného pokusu přiřadí reálné číslo. Náhodná veličina je zobrazení

$$X : \Omega \to \mathbf{R},$$

které je definované na Ω se σ-algebrou $\mathcal{A}$ a je měřitelné. Že je zobrazení měřitelné znamená, že pro každé $B \in \mathcal{B}(\mathbf{R})$ platí

$$[X \in B] := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

Každé náhodné veličině $B \in \mathcal{B}(\mathbf{R})$ lze připsat pravděpodobnostní míru Q:

$$Q(B) = P\{X^{-1}(B)\},$$

kde

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$$

je vzor množiny $B \subset \mathbf{R}$ při zobrazení X. Potom se pravděpodobnostní míra Q nazývá *rozdělení náhodné veličiny X*. Častěji se používá zkráceně *rozdělení pravděpodobnosti* nebo dokonce jen *rozdělení*.

Rozdělení pravděpodobnosti je pravidlo, které každé hodnotě (nebo intervalu hodnot) přiřazuje pravděpodobnost, že náhodná veličina nabude právě této hodnoty (nebo intervalu hodnot). Rozdělení náhodné veličiny popisuje její pravděpodobnostní chování při dané pravděpodobnostní míře P.

Funkce $F\colon \mathbf{R} \to \langle 0,1\rangle$, která je neklesající, zprava spojitá a platí pro ní

$$\lim_{x\to-\infty} F(x) = 0 \qquad \lim_{x\to\infty} F(x) = 1$$

se nazývá *distribuční funkcí*. Distribuční funkcí náhodné veličiny X rozumíme funkci

$$F(x) = P(X \le x), \quad x \in \mathbf{R} \tag{3.4}$$

Pro distribuční funkci platí následující:

- $P(X \in (a,b\rangle) = F(b) - F(a)$
- $P(X = a) = F(a) - F(a-)$, kde $F(a-)$ značí limitu F v bodě a zleva
- $P(X > a) = 1 - F(a)$

Náhodná veličina může mít rozdělení různých typů, nás bude zajímat rozdělení diskrétního typu.

Řekneme, že náhodná veličina X má pravděpodobnostní rozdělení diskrétního typu právě tehdy, když existuje konečná nebo spočetná množina reálných čísel $M = \{x_1, x_2, \ldots\} \subset \mathbf{R}$ taková, že

$$P(X = x_i) = p_i > 0 \quad \text{pro} \quad i = 1,2,\ldots \qquad \text{a} \qquad \sum_{x_i} P(X = x_i) = \sum_i p_i = 1$$

Pravděpodobnostní funkcí nazveme takovou funkci, která u diskrétní náhodné veličiny X přiřadí každému $x \in \mathbf{R}$ pravděpodobnost $P(X = x)$, zkráceně $P(x)$:

$$P(x) = \begin{cases} p_i & x = x_i, i = 1,2,\ldots \\ 0 & x \ne x_i \end{cases} \tag{3.5}$$

Distribuční funkce F diskrétní náhodné veličiny X je schodovitá, po úsecích konstantní, se skoky v bodech x_i, $i = 1,2,\ldots$. Pro distribuční funkci platí vztah

$$F(x) = \sum_{i, x_i \le x} p_i \tag{3.6}$$

Hypergeometrické rozdělení

Mějme konečnou množinu N prvků, z nichž K prvků má určitou sledovanou vlastnost. Z této množiny náhodně vybereme (bez vracení zpět) n prvků. Nechť veličina X značí počet prvků ve výběru mající sledovanou vlastnost.

Pak náhodná veličina X má *hypergeometrické* rozdělení pravděpodobnosti $HG(N,K,n)$ s parametry N, K, $n \in \mathbf{N}$, $n \le N$, $K \le N$, když pravděpodobnostní funkce má tvar:

$$P(x) = \begin{cases} \dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} & x \in \mathbf{N} \cup \{0\}, \text{ taková, že } \quad x \le K,\ 0 \le n - x \le N - K \\ 0 & \text{jinde} \end{cases} \tag{3.7}$$

Píšeme X $\sim$ HG(N, K, n).

Kapitola 4

Odvození postupu kontroly

V této kapitole se budeme věnovat výpočtu pravděpodobnosti zamítnutí při různém počtu neplatných podpisů v celkovém počtu dodaných podpisů pod peticí podle zákonem předepsaného postupu. V první části se zaměříme na konkrétní odvození výpočtu pravděpodobnosti zamítnutí, ve druhé části pak získáme pravděpodobnosti zamítnutí simulačně.

V následujících částech práce budeme pracovat s proměnnými, které si označíme:

N	...	celkový počet dodaných podpisů pod peticí,
n	...	velikost každého kontrolního vzorku = 8 500,
K	...	počet neplatných podpisů v N,
X_1	...	náhodná veličina značící počet neplatných podpisů v 1. kontrolním vzorku,
X_2	...	náhodná veličina značící počet neplatných podpisů v 2. kontrolním vzorku,
q	...	chybovost = 3 %,
h	...	mez chybovosti = $q \cdot n = 255$.

Dle zákonem předepsaného postupu může nastat zamítnutí prezidentské kandidatury třemi možnými způsoby.

První možností je, že se provede pouze první kontrolní vzorek, a na jeho základě se rozhoduje o zamítnutí či registraci prezidentské kandidatury. Pouze jeden kontrolní vzorek se provádí v případě, že nebudou splněny podmínky pro realizaci druhého, tj. počet neplatných podpisů v prvním vzorku x_1 je menší než mez chybovosti, což jsou 3 % z rozsahu náhodného výběru n. Potom se počet nalezených neplatných podpisů v prvním vzorku x_1 pouze odečte od celkového počtu dodaných podpisů N a kandidatura bude zamítnuta právě tehdy, když $N - x_1 < 50\,000$. Možnost zamítnutí kandidatury po prvním kontrolním vzorku záleží na velikosti N. Jelikož maximální možná hodnota x_1 je $h-1$ (má být menší než mez chybovosti), pak pro hodnotu N platí: $N < 50\,000 + h - 1$. Reálně může být kandidatura zamítnuta po uskutečnění pouze prvního kontrolního vzorku jen tehdy, když $N \leq 50\,000 + h - 2$, to jest konkrétně $N \leq 50\,253$.

Druhou možností zamítnutí prezidentské kandidatury je případ, kdy v prvním kontrolním vzorku bude počet neplatných podpisů x_1 větší nebo rovno h a ve druhém kontrolním vzorku bude počet neplatných podpisů $x_2 < h$. Potom se počet všech nalezených neplatných podpisů z obou kontrolních vzorků odečte od celkového počtu dodaných podpisů N a kandidatura bude zamítnuta právě tehdy, když $N - x_1 - x_2 < 50\,000$. Zamítnutí prezidentské kandidatury pouze odečtením nalezených neplatných podpisů v prvním a druhém kontrolním vzorku opět záleží na velikosti N. Maximálních možných hodnot, kterých můžou x_1 a x_2 pro tento případ nabývat, jsou $x_1 = n$ a $x_2 = h - 1$. Pak pro N platí: $N < 50000 + n + h - 1$. Reálné zamítnutí kandidatury v tomto případě může nastat jen tehdy, když $N \leq 50\,000 + n + h - 2$, to jest konkrétně $N \leq 58\,753$.

Třetí možností zamítnutí prezidentské kandidatury je opět po uskutečnění obou kontrolních vzorků za předpokladů, že x_1 i x_2 nabudou hodnot větších nebo rovno h. Pak se výsledný počet platných podpisů získá odečtením průměrné chybovosti z obou kontrolních vzorků z N od N a kandidatura bude zamítnuta právě tehdy, když $N - N \cdot (\frac{x_1+x_2}{2 \cdot n}) < 50\,000$. Zde nedokážeme určit velikost N, kdy dojde k zamítnutí kandidatury, neboť tento postup závisí na výši chybovosti jednotlivých kontrolních vzorků. Dokážeme ale naopak určit minimální hodnotu N, aby tato možnost - uskutečnění obou vzorků s chybovostmi minimálně 3 % v každém vzorku - nastala. Hodnoty x_1 a x_2 musí být rovny minimálně h, tedy pro N bude platit $N \geq 50\,000 + 2 \cdot h$, to jest konkrétně $N \geq 50\,510$.

Za zmínění stojí chybný postup, který byl pro kontrolu správnosti údajů na petičních arších v roce 2013 kontrolní komisí použit. V prvních dvou výše zmíněných možností zamítnutí byl postup shodný. Rozdílnost zamítnutí nastala ve třetí možnosti, kdy namísto zprůměrování chybovostí z obou kontrolních vzorků pouze tyto chybovosti sečetli. Díky tomu předpokládali, že v N je jednou tolik procent neplatných podpisů. Například, bylo-li by dodáno $N = 55\,000$ podpisů a byla-li by chybovost prvního kontrolního vzorku 4 % a druhého 6 %, tak správně mělo být odečteno 5 % z celkového počtu dodaných podpisů N (tj. 2750), kdežto chybným postupem odečetli 10 %, což je 5500. Z tohoto příkladu je vidět, že chybným postupem by byla kandidatura zamítnuta, ačkoliv by neměla. Je tedy zřejmé, že tento postup mohl prezidentské kandidáty poškodit.

4.1 Odvození výpočtu pravděpodobnosti

Označme si jev zamítnutí prezidentské kandidatury dle zákonem předepsaného postupu Z. Dále si označíme výše zmíněné možnosti zamítnutí prezidentské kandidatury dle zákonem předepsaného postupu. Nechť jev zamítnutí kandidatury na základě uskutečnění pouze prvního kontrolního vzorku je značen A, jev zamítnutí kandidatury na základě uskutečnění obou kontrolních vzorků, kdy $h \leq x_1 \leq n$ a $x_2 < h$, je značen B a jev zamítnutí kandidatury na základě uskutečnění obou kontrolních vzorků, kdy $h \leq x_1 \leq n$ a $h \leq x_2 \leq n$ je značen C. Pravděpodobnost zamítnutí kandidatury $P(Z)$ je rovna pravděpodobnosti sjednocení jevů A, B a C.

$$P(Z) = P(A \cup B \cup C) \tag{4.1}$$

Jelikož jsou tyto tři varianty zamítnutí kandidatury disjunktními jevy, potom pro pravděpodobnost zamítnutí platí, že se rovná pravděpodobnosti sjednocení disjunktních jevů:

$$P(Z) = P(A) + P(B) + P(C) \tag{4.2}$$

V následující části rozebereme jednotlivé sčítance pro získání pravděpodobnosti zamítnutí kandidatury $P(Z)$.

Pravděpodobnost zamítnutí kandidatury ve všech třech případech záleží na počtu neplatných podpisů x_1 v prvním kontrolním vzorku. Pravděpodobnosti $P(A), P(B)$ a $P(C)$ lze tedy vyjádřit podmíněnými pravděpodobnostmi za podmínky, že nastane jev $X_1 = x_1$. Aplikujeme větu o úplné pravděpodobnosti, neboť nabývání hodnot $X_1 = x_1$ jsou disjunktními jevy a získáme

$$P(A) = \sum_{x_1} P(A \mid X_1 = x_1) \cdot P(X_1 = x_1) \tag{4.3}$$

$$P(B) = \sum_{x_1} P(B \mid X_1 = x_1) \cdot P(X_1 = x_1) \tag{4.4}$$

$$P(C) = \sum_{x_1} P(C \mid X_1 = x_1) \cdot P(X_1 = x_1) \tag{4.5}$$

Náhodná veličina X_1 se řídí hypergeometrickým rozdělením s parametry $N, K, n \in \mathbf{N}, n \leq N, K \leq N$, tj. $X_1 \sim HG(N, K, n)$. Náhodná veličina X_2 je závislá na výsledcích X_1. Také se řídí hypergeometrickým rozdělením, ale s parametry $N - n, K - x_1, n \in \mathbf{N}, n \leq N - n, K - x_1 \leq N - n$, to znamená, že $(X_2 \mid X_1 = x_1) \sim HG(N - n, K - x_1, n)$.

Začneme odvozením výpočtu pro $P(A)$. Podmíněná pravděpodobnost jevu A z rovnice (4.3) za podmínky, že nastane $X_1 = x_1$, bude rovna 0 nebo 1, neboť se po uskutečnění pouze prvního vzorku rozhodujeme o zamítnutí či nikoliv. Z $0 \leq x_1 \leq K$ a $0 \leq n - x_1 \leq N - K$ plyne obor hodnot hypergeometrického rozdělení pro x_1: $\max(0,\ n - N + K) \leq x_1 \leq \min(K, n)$. K získanému oboru hodnot přidáme podmínku $0 \leq x_1 \leq h - 1$, neboť nemá být proveden druhý vzorek. Kandidatura bude zamítnuta právě tehdy, když $N - x_1 < 50\,000$. Z toho vyplývá, že $x_1 \geq N - 50\,000 + 1$. Díky přidaným podmínkám k oboru hodnot hypergeometrického rozdělení získáme hodnoty x_1, pro které bude kandidatura zamítnuta, tzn. $P(A \mid X_1 = x_1) = 1$. Výsledný obor hodnot pro x_1 je:
$\max(0,\ n - N + K, N-\ 50\,000\ +\ 1) \leq x_1 \leq \min(K, h - 1, n)$ a pro X_1 platí $X_1\ \sim\ HG(N, K, n)$. Potom pro $P(A)$ platí:

$$\begin{aligned} P(A) &= \sum_{\max(0,n-N+K,N-50000+1)}^{\min(K,h-1,n)} P(A \mid X_1 = x_1) \cdot P(X_1 = x_1) = \\ &= \sum_{\max(0,n-N+K,N-50000+1)}^{\min(K,h-1,n)} 1 \cdot P(HG(N, K, n) = x_1) \end{aligned} \tag{4.6}$$

Pro jevy B a C nastává rozhodnutí o zamítnutí kandidatury až po realizaci druhého vzorku. Obor hodnot hypergeometrického rozdělení pro x_1 je $\max(0, n-N+K) \leq x_1 \leq \min(K, n)$. Pro oba jevy platí $h \leq x_1 \leq n$. Tuto informaci zahrneme do oboru hodnot proměnné x_1. Potom pro x_1 platí:
$\max(0, n-N+K, h) \leq x_1 \leq \min(K, n)$ a pravděpodobnost zamítnutí kandidatury po realizaci druhého kontrolního vzorku má tvar:

$$P(B) = \sum_{\max(0,n-N+K,h)}^{\min(K,n)} P(B \mid X_1 = x_1) \cdot P(X_1 = x_1) \tag{4.7}$$

$$P(C) = \sum_{\max(0,n-N+K,h)}^{\min(K,n)} P(C \mid X_1 = x_1) \cdot P(X_1 = x_1) \tag{4.8}$$

Nyní odvodíme výpočet pro $P(B)$. Podmíněná pravděpodobnost $P_{X_1=x_1}(B)$ znamená, že za platnosti $X_1 = x_1$ bude existovat více hodnot $X_2 = x_2$ splňujících podmínky jevu B, tj. $0 \leq x_2 \leq h-1$. Proto podmíněné pravděpodobnosti $P(B \mid X_1 = x_1)$ sečteme a poté vynásobíme $P(X_1 = x_1)$. Nyní potřebujeme odvodit x_2, které pro dané $X_1 = x_1$ nastanou. Obor hodnot hypergeometrického rozdělení pro x_2 doplníme o podmínku $0 \leq x_2 \leq h-1$, která platí pro jev B. Kandidatura bude zamítnuta právě tehdy, když $N - x_1 - x_2 < 50\,000$, tj. $x_2 \geq N - 50\,000 - x_1 + 1$. Díky přidaným podmínkám je výsledný obor hodnot:
$\max(0, n+K-N, N-50\,000-x_1+1) \leq x_2 \leq \min(K-x_1, n-x_1, h-1)$. Pro X_2 platí $X_2 \sim HG(N-n,\ K-x_1,\ n)$. Potom výsledný tvar pro výpočet $P(B)$ je

$$P(B) = \sum_{x_1 \in M_1} \sum_{x_2 \in M_2} P(HG(N-n, K-x_1, n) = x_2) \cdot P(HG(N, K, n) = x_1), \tag{4.9}$$

kde $M_1 = \{\max(0,\ n-N+K, h), \ldots, \min(K, n)\}$ a
$M_2 = \{\max(0,\ n-N+K, N-50000-x_1+1), \ldots, \min(K-x_1, n-x_1, h-1)\}$.

Nakonec odvodíme výpočet pro $P(C)$. Pro jev C doplníme obor hodnot hypergeometrického rozdělení pro x_2 podmínkou $h \leq x_2 \leq n$. Dále platí, že kandidatura bude zamítnuta právě tehdy, když $N - N \cdot (\frac{x_1+x_2}{2 \cdot n}) < 50\,000$, tj. $x_2 \geq \frac{2 \cdot n \cdot (N-50\,000)}{N} - x_1 + 1$. Po přidání zmíněných podmínek je výsledný obor hodnot pro x_2
$\max(0, n-N+K, h, \frac{2n(N-50\,000)}{N} - x_1 + 1) \leq x_2 \leq \min(K-x_1, n-x_1, n)$. Tvar pro výpočet $P(C)$ je pak

$$P(C) = \sum_{x_1 \in M_1} \sum_{x_2 \in M_3} P(HG(N-n, K-x_1, n) = x_2) \cdot P(HG(N, K, n) = x_1), \tag{4.10}$$

kde $M_1 = \{\max(0,\ n-N+K, h, \ldots, \min(K, n)\}$ a
$M_3 = \{\max(0,\ n-N+K, h, \frac{2n(N-50000)}{N} - x_1 + 1)), \ldots, \min(K-x_1, n-x_1, n)\}$.

Pro chybný postup kontroly je odvození výpočtu pravděpodobnosti zamítnutí stejný až do vyjádření si oboru hodnot x_2 pro výpočet $P(C)$. Samozřejmě platí, že obor hodnot bude doplněn o podmínku $h \leq x_2 \leq n$. Ale kandidatura má být zamítnuta právě tehdy, když $N - N \cdot (\frac{x_1+x_2}{n}) < 50\,000$, tj.
$x_2 \geq \frac{n \cdot (N-50\,000)}{N} - x_1 + 1$. Potom tvar výpočtu $P(C)$ pro chybný postup kontroly je

$$P(C) = \sum_{x_1 \in M_1} \sum_{x_2 \in M_4} P(HG(N-n, K-x_1, n) = x_2) \cdot P(HG(N, K, n) = x_1), \tag{4.11}$$

kde $M_1 = \{\max(0, n-N+K, h, \ldots, \min(K, n)\}$ a
$M_4 = \{\max(0, n-N+K, h, \frac{n(N-50000)}{N} - x_1 + 1)), \ldots, \min(K-x_1, n-x_1, n)\}$.

Číselné výsledky $P(Z)$ pro zvolená N budou v kapitole 4.

4.2 Simulace a intervaly spolehlivosti

V této části popíšeme, jak budeme získávat pravděpodobnosti zamítnutí kandidatury simulačně a odvodíme si, jak budeme vytvářet intervaly spolehlivosti pro některé hodnoty. Konkrétní číselné výsledky pak budou uvedeny v kapitole 4.

Při získávání pravděpodobností simulací potřebujeme, aby hodnoty x_1 a x_2 byly hodnotami pocházejícími z hypergeometrického rozdělení. Na základě těchto vygenerovaných hodnot x_1 a x_2 rozhodujeme o zamítnutí či nezamítnutí prezidentské kandidatury. Jelikož vygenerování náhodných čísel s hypergeometrickým rozdělením pomocí „matlabovské“ funkce 'hyge' je časově náročné, zvolíme pro vygenerování hodnot jiný přístup.

Hlavní myšlenkou je vygenerování náhodných permutací. Náhodnou permutaci získáme seřazením posloupnosti náhodně vygenerovaných neopakujících se čísel z intervalu (0, 1). Číslice od 1 do K v náhodné permutaci značí neplatné podpisy. Pro 1. kontrolní vzorek vezmeme prvních n prvků této permutace, pro druhý kontrolní vzorek pak $n+1$ až $2n$ prvků, a v každém vzorku porovnáváme, kolik čísel je menších či rovno K. Potom počet takových hodnot jsou hodnoty x_1 a x_2.

Počet simulací, tj. kolikrát tento postup opakujeme pro každé zvolené K, je 20 000. Výsledná pravděpodobnost zamítnutí kandidatury pomocí simulace je odhadnuta jako podíl počtu zamítnutí ku počtu simulací.

Nyní si popíšeme, jak budeme získávat intervaly spolehlivosti. Mějme počet simulací $s = 20\,000$. Simulací získáme pouze dva výsledky, buď bude kandidatura zamítnuta, nebo nikoliv, a my sledujeme, kolikrát zamítnuta bude. Stane-li se to r - krát, pak relativní četnost zamítnutí v simulaci je $\hat{p} = \frac{r}{s}$. Taktéž lze říct, že $\hat{p}$ je odhadem pravděpodobnosti zamítnutí kandidatury $P(Z)$ získaným simulací. Budeme počítat interval spolehlivosti pro parametr p alternativního rozdělení, kde p znamená skutečnou pravděpodobnost zamítnutí kandidatury, pomocí parametru $\hat{p}$. Pro dostatečně velké s, to jest takové, pro které platí $s\hat{p}(1 - \hat{p}) \geq 9$ bude mít přibližně 100(1-α)%-ní interval spolehlivosti tvar

$$\left(\hat{p} - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{s}}, \hat{p} + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{s}}\right), \tag{4.12}$$

kde $\alpha \in (0,1)$ a $u_{1-\frac{\alpha}{2}}$ je kvantil normovaného normálního rozdělení.

Šířka (taktéž absolutní chyba) intervalu spolehlivosti e charakterizuje přesnost lokalizace skutečné hodnoty. Máme-li interval spolehlivosti (a, b), potom šířka intervalu $e = b - a$. Pro šířku e námi získaného přibližně 100(1-α)% - ní intervalu spolehlivosti platí

$$e = \hat{p} + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{s}} - \hat{p} + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{s}} = 2 \cdot u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{s}} \tag{4.13}$$

Kapitola 5

Výsledky

V této kapitole jsou v grafickém znázornění uvedeny výsledky pravděpodobností zamítnutí $P(Z)$ získané použitím odvozeného vzorce pro výpočet $P(Z)$ z předchozí kapitoly. Dále je zde graficky zpracované srovnání hodnot pravděpodobností zamítnutí získané simulací se skutečnými hodnotami $P(Z)$. Výpočet $P(Z)$ i simulace jsou provedeny pouze pro vybrané hodnoty K.

5.1 Volba hodnot K

Jak již bylo řečeno, výpočty a simulace jsou provedeny jen na některých vybraných hodnotách K, proto si zde uvedeme, na jakém principu jsou hodnoty K voleny.

Znamená-li K počet neplatných podpisů v celkovém počtu dodaných podpisů N, potom $N - K$ znamená počet platných podpisů v N. Budeme počítat pravděpodobnosti zamítnutí podle toho, kolik platných podpisů je v N, tj. zvolíme si pro všechna N stejné hodnoty $N - K$. Jelikož minimální počet platných podpisů pro registraci kandidatury je 50 000, budeme volit $N - K$ kolem této hodnoty. Zvolíme interval od 49 000 do 51 000 platných podpisů. V intervalech $\langle 49\,000,\ 49\,900\rangle$ a $\langle 50\,100,\ 51\,000\rangle$ použijeme rozložení hodnot s krokem 30 a v intervalu $\langle 49\,900,\ 50\,100\rangle$ použijeme krok 8, tj. volíme $N - K = 49\,000, 49\,030, \ldots, 49\,900, 49\,908, \ldots, 50\,092, 50\,100, 50\,130, \ldots, 50\,970, 51\,000$ pro výpočet $P(Z)$. Pro lepší přehlednost a porovnání výsledků při různých hodnotách N jsou výsledné hodnoty $P(Z)$ graficky znázorněné v závislosti na $N - K$, nikoliv na K.

Pro porovnávání hodnot získaných pomocí simulace a skutečných hodnot $P(Z)$ volíme menší počet hodnot s jiným rozložením. Interval platných podpisů $N - K$ zůstává stejný, tj. 49 000 až 51 000. Avšak volíme rozdílné rozložení hodnot platných podpisů, a to takto: $N - K = 49\,000, 49\,070, \ldots, 49\,700, 49\,720, \ldots, 49\,900, 49\,910, \ldots, 50\,100, 50\,120, \ldots, 50\,300, 50\,370, \ldots, 51\,000$.

5.2 Vstupní data

V roce 2013, kdy byly konány historicky první přímé volby prezidenta České republiky, bylo podáno pouze 8 peticí splňujících podmínku na minimální počet 50 000 podpisů, jak je uvedeno v [6], které měly podpořit osobnosti jakožto prezidentské kandidáty. U těchto peticí byla provedena Ministerstvem vnitra České republiky kontrola správnosti údajů, na jejímž základě se poté rozhodlo o zamítnutí či registraci prezidentské kandidatury. Počty dodaných podpisů N na petičních arších každého z potenciálních prezidentských kandidátů jsou uvedeny v tabulce 5.1. Dále jsou v této tabulce uvedeny jisté hodnoty K. Pro tyto hodnoty platí, že pokud by v N bylo minimálně tolik neplatných podpisů, tak by v ideálním případě měly být kandidatury zamítnuty, tj. pravděpodobnost zamítnutí by měla být rovna jedné.

Tabulka 5.1: Přehled počtu dodaných podpisů N prezidentskými kandidáty

Jméno a příjmení kandidáta	N	K
Ing. Jana Bobošíková	56 191	6 192
Ing. Vladimír Dlouhý, CSc.	59 165	9 166
Tomio Okamura	61 966	11 967
Taťana Fischerová	72 434	22 435
MUDr. Zuzana Roithová, MBA	81 199	31 200
Prof. JUDr. Vladimír Franz	87 782	37 783
Ing. Jan Fischer, CSc.	101 261	51 262
Ing. Miloš Zeman	106 081	56 082

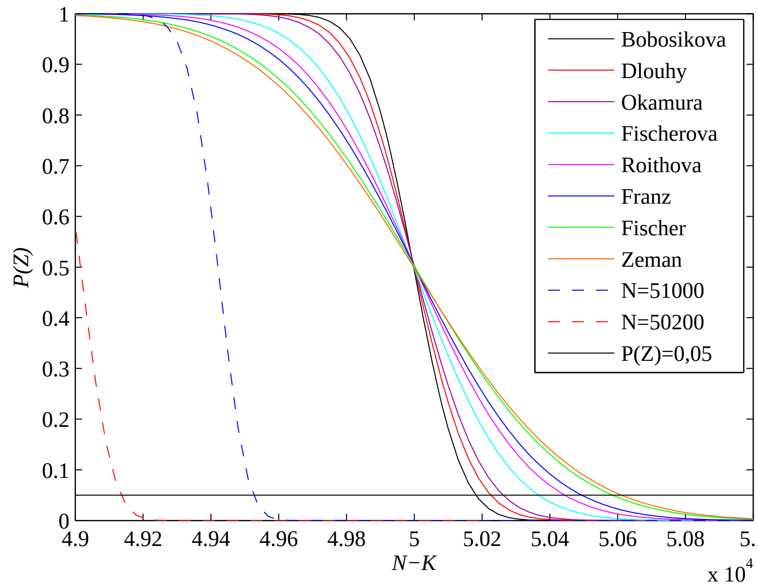
Dále budeme pracovat s hodnotami N blízkými 50 000. Jako první volíme $N = 51\,000$. Dále zvolíme takové N, aby mohlo dojít k zamítnutí už po prvním vzorku, tj. $N \leq 50\,253$, volíme $N = 50\,200$.

5.3 Skutečné pravděpodobnosti zamítnutí

V této části nejprve graficky porovnáme hodnoty pravděpodobností zamítnutí $P(Z)$ pro zvolená N, a potom graficky porovnáme pravděpodobnosti zamítnutí $P(Z)$ získané podle zákonem předepsaného postupu a získané podle chybného postupu, který byl v roce 2013 použit. Všechny hodnoty pravděpodobností zamítnutí prezidentských kandidatur $P(Z)$ jsou získány z předpisu pro výpočet $P(Z)$ odvozený v 4.1.

Na obrázku 5.1 jsou znázorněné křivky pravděpodobností zamítnutí $P(Z)$ v závislosti na počtu platných podpisů $N - K$ pro zvolená N z předchozí podkapitoly.

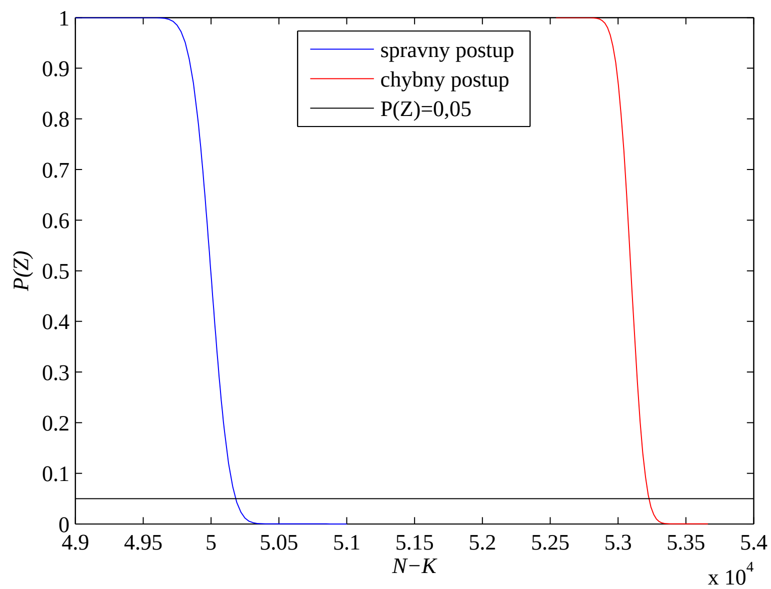
Obrázek 5.1: Srovnání $P(Z)$ pro všechna zvolená N



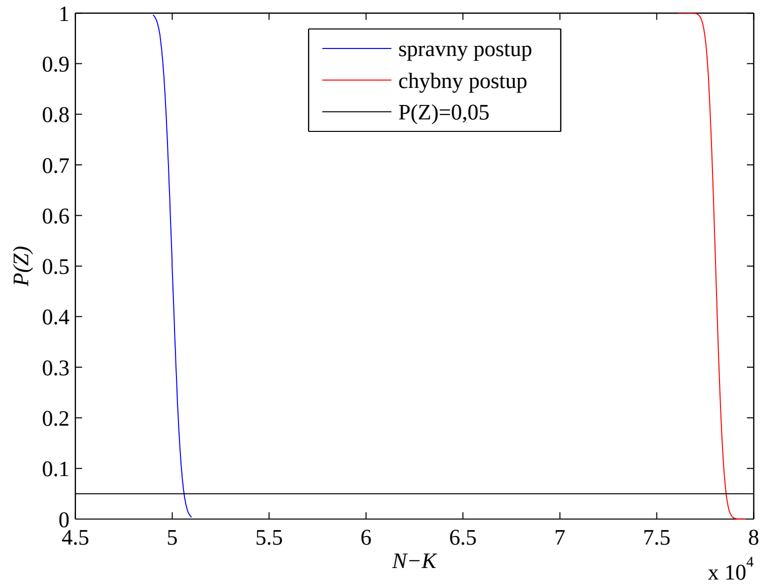
Vodorovná černá přímka v obrázku 5.1 značí 5%-ní pravděpodobnost, že prezidentská kandidatura bude zamítnuta při daném počtu platných podpisů v N. Obráceně lze říci, že pokud dodáme N podpisů, a chceme mít alespoň 95%-ní pravděpodobnost, že naše kandidatura bude přijata, mělo by být v celkovém počtu dodaných podpisů alespoň $N - K$ platných podpisů. Z obrázku vyplývá, že čím více podpisů dodáme, tím více platných podpisů musí petice obsahovat, aby naše kandidatura byla s 95%-ní pravděpodobností přijata. Například v $N = 56\,191$ by mělo být přibližně 50 180 platných podpisů a v $N = 106\,081$ by mělo být přibližně 56 610 podpisů, aby došlo s 95%-ní pravděpodobností k registraci kandidatury. Z toho jasně vyplývá, že větší počet dodaných podpisů by měl obsahovat více platných podpisů. Avšak když se podíváme na rozdíl počtu platných podpisů (tj. konkrétně na příkladě 6 430), který je podstatně menší než rozdíl počtu dodaných podpisů (49 890), bylo by asi lepší se zabývat procentuálním vyjádřením počtu platných podpisů v N. Z toho pak vyplývá, že z celkového počtu dodaných podpisů N rovno 56 191 by mělo být přibližně 89,3 % platných podpisů. Kdežto již zmíněných 56 610 platných podpisů v 106 081 dodaných podpisů činí „jen“ zhruba 53,4 %. Je tedy zřejmé, že pokud se pro přijetí prezidentské kandidatury požaduje (minimálně) 50 000 platných podpisů, tak větší pravděpodobnost budeme mít s větším počtem dodaných podpisů.

Na obrázcích 5.2 a 5.3 jsou znázorněné dvě křivky, které představují závislost pravděpodobnosti zamítnutí prezidentské kandidatury $P(Z)$ na počtu platných podpisů $N - K$ pro paní Ing. Bobošíkovou a pana Ing. Zemana. Modrá křivka představuje závislost pravděpodobnosti zamítnutí kandidatury $P(Z)$ získané podle zákonem předepsaného postupu na $N - K$. Červená křivka pak značí závislost $P(Z)$ získané podle chybného postupu na $N - K$. Vodorovné přímky v obou obrázcích opět představují 5%-ní pravděpodobnost, že bude kandidatura zamítnuta při daném počtu platných podpisů, resp. 95%-ní pravděpodobnost, že bude prezidentská kandidatura přijata.

Obrázek 5.2: Srovnání $P(Z)$ získaných dle správného a chybného postupu pro Ing. Bobošíkovou



Obrázek 5.3: Srovnání $P(Z)$ získaných dle správného a chybného postupu pro Ing. Zemana



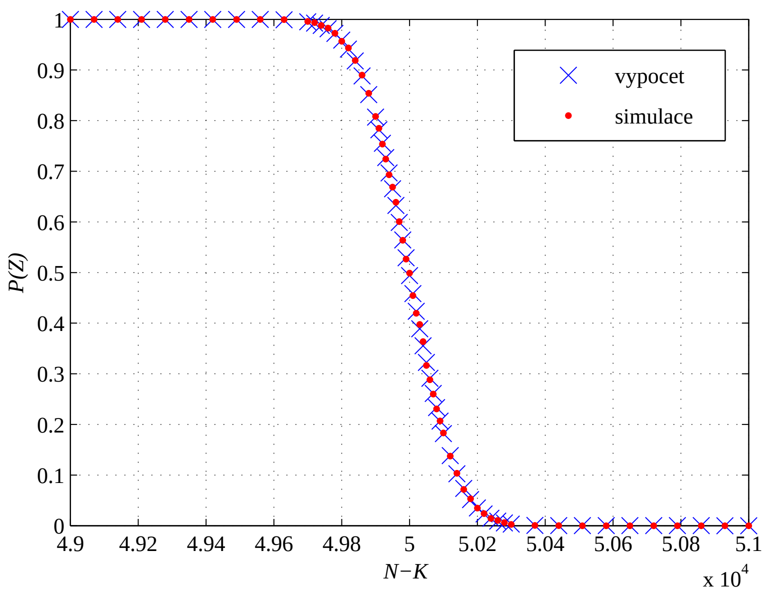
Pro chybný postup vyplývá, stejně jako pro ten správný, že pro větší N je požadovaný větší počet $N - K$ v N, aby byla kandidatura s 95%-ní pravděpodobností přijata. I nyní si vyjádříme počet platných podpisů v N procentuálně pro 95%-ní pravděpodobnost zamítnutí kandidatury, tentokrát ale pro hodnoty získané chybným postupem kontroly. Získáme, že v $N = 56\,191$ by mělo být přibližně 94,7 % platných podpisů a v $N = 106\,081$ přibližně 74,1 %. Obě hodnoty $\frac{K}{N}$ se zvýšily, každá řádově jinak, což bude mít nejspíš za následek velikost N. Je zřejmé, že chybný postup kontroly mohl poškodit (a skutečně tomu tak i bylo) prezidentské kandidáty. Například pro Ing. Zemana mohla být získaná chybovost z kontrolních vzorků podle správného postupu kontroly dokonce lehce přes 50 %, neboť odečtením $0,5N$ od N získáme stále více než 50 000 podpisů. Kdežto u chybného postupu kontroly je maximální chybovost pro nezamítnutí kandidatury velmi přibližně 24,6 %, což jest zhruba poloviční. Což by mohlo plynout z faktu, že chybovosti získané z kontrolních vzorků nebyly zprůměrovány, ale pouze sečteny.

5.4 Výsledky simulace a intervaly spolehlivosti

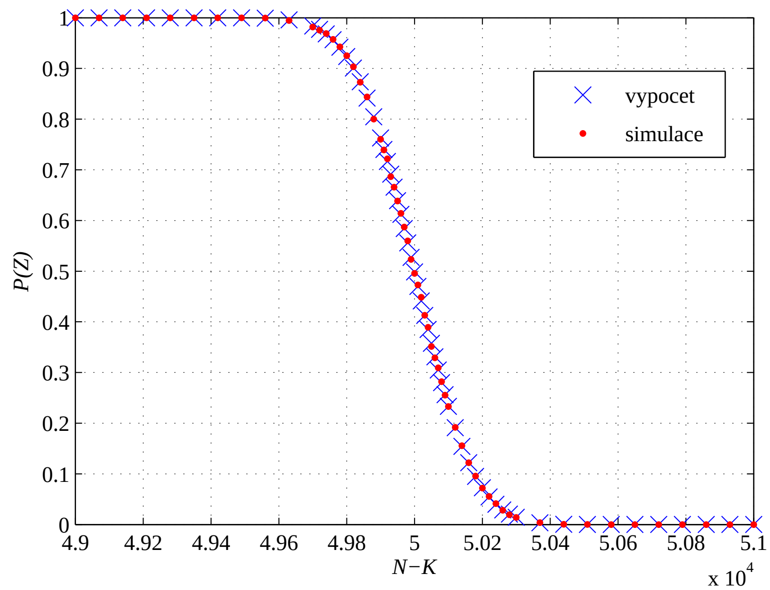
V této části budeme opět graficky porovnávat a poté vytvářet intervaly spolehlivosti. Tentokrát srovnáme skutečné pravděpodobnosti zamítnutí $P(Z)$ s pravděpodobnostmi získanými simulací. Pak pro několik hodnot $P(Z)$ vyjádříme intervaly spolehlivosti pomocí hodnot získaných simulací a porovnáme, jaký vliv na ně má velikost N. Konkrétně budeme zjišťovat 95%-ní interval spolehlivosti pro $N - K = 49\,700, 50\,000$ a $50\,300$, když počet simulací s je roven $20\,000$.

Obrázky 5.4 až 5.11 znázorňují srovnání skutečných hodnot pravděpodobností zamítnutí $P(Z)$ (modré křížky) a pravděpodobností zamítnutí získané simulací (červené body) pro potenciální prezidentské kandidáty. Na obrázcích 5.12 a 5.13 je pak znázorněné srovnání pravděpodobností pro obecně zvolené hodnoty N.

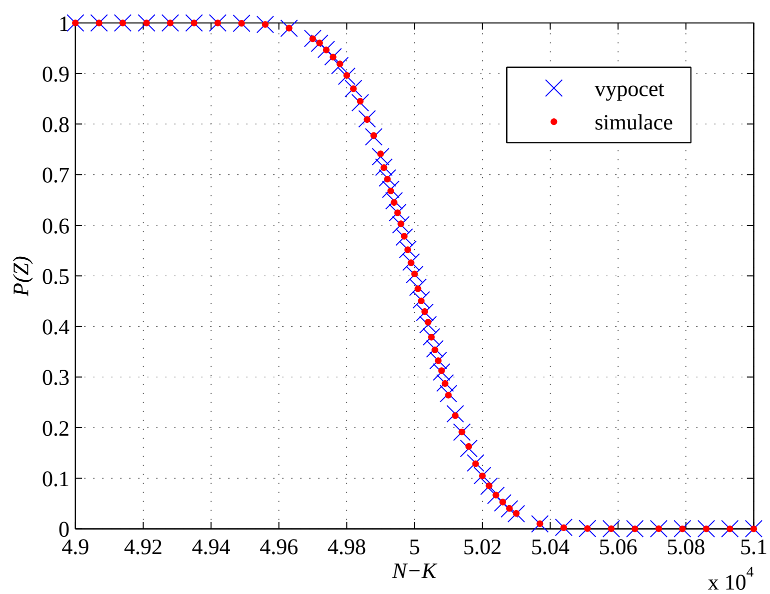
Obrázek 5.4: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Ing. Janu Bobošíkovou



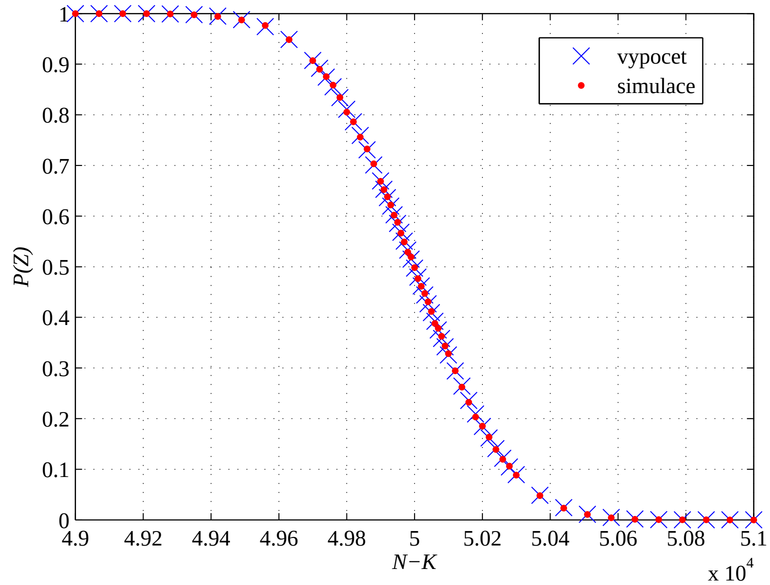
Obrázek 5.5: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Ing. Vladimíra Dlouhého, CSc.



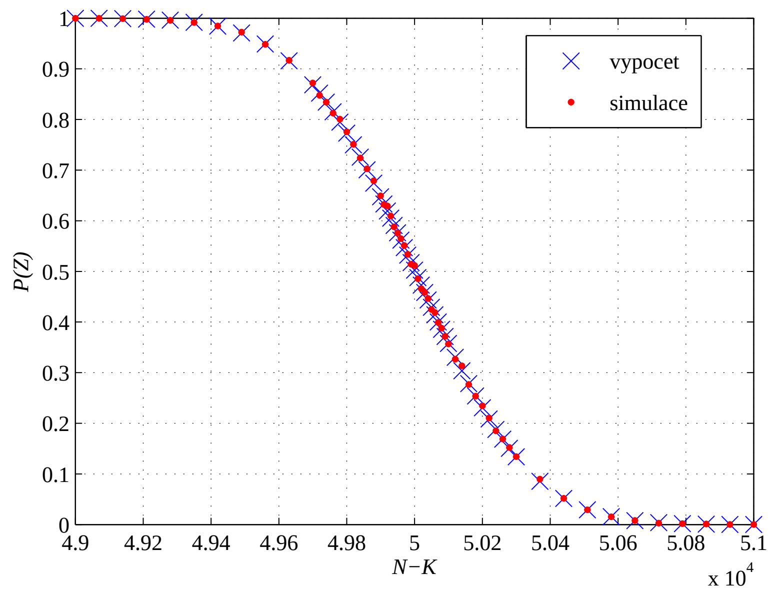
Obrázek 5.6: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Tomia Okamuru



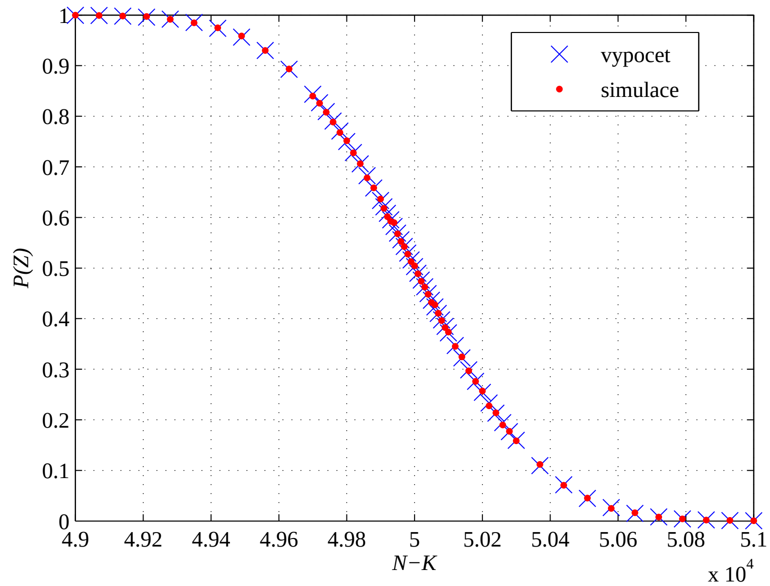
Obrázek 5.7: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Taťanu Fischerovou



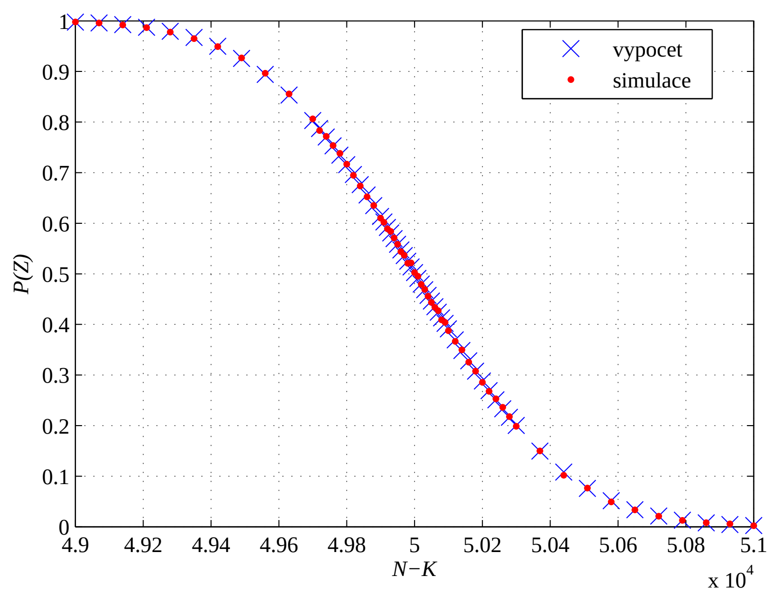
Obrázek 5.8: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro MUDr. Zuzanu Roithovou, MBA



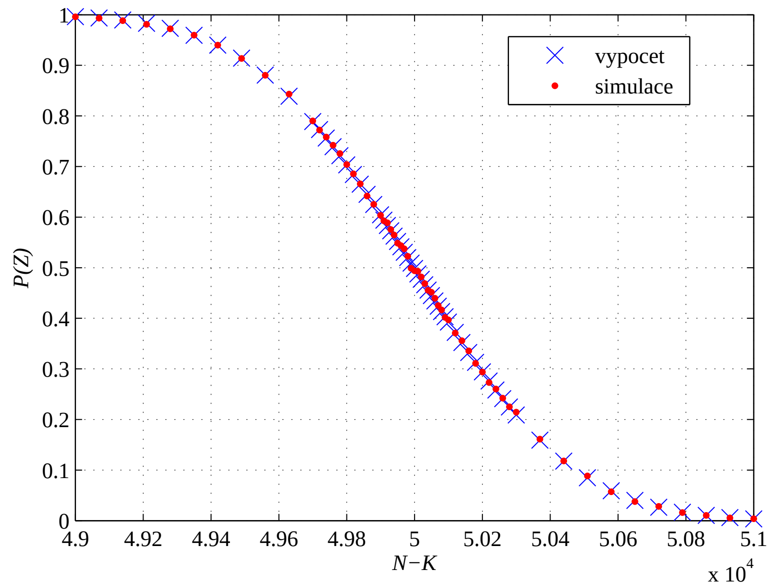
Obrázek 5.9: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Prof. JUDr. Vladimíra Franze



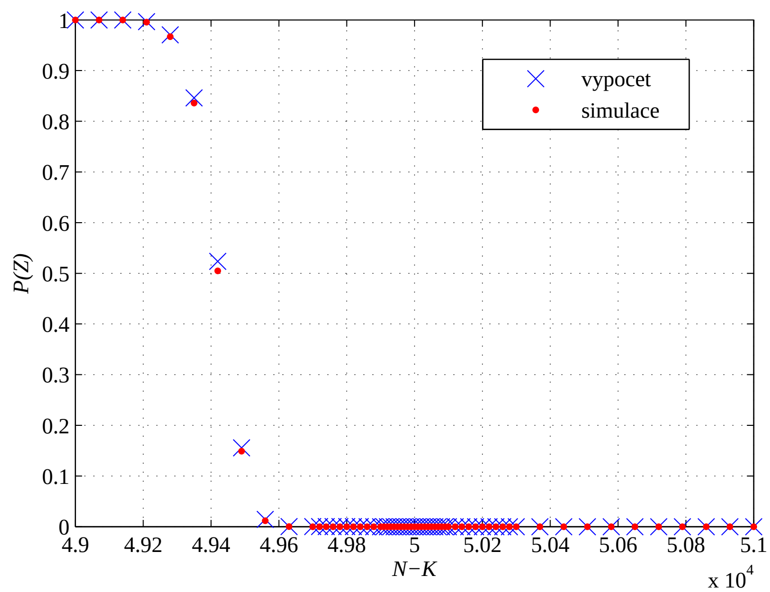
Obrázek 5.10: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Ing. Jana Fischera, CSc.



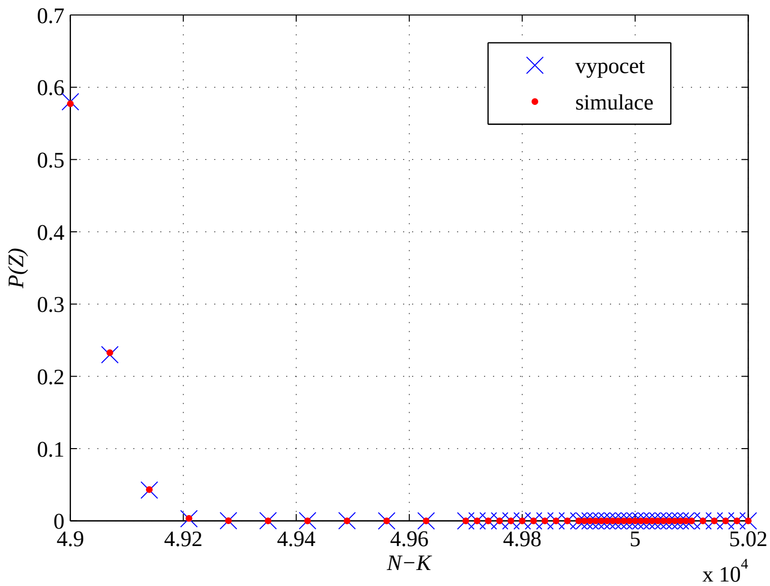
Obrázek 5.11: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro Ing. Miloše Zemana



Obrázek 5.12: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro $N = 51\,000$



Obrázek 5.13: Porovnání hodnot $P(Z)$ získaných výpočtem a simulací pro $N = 50\,200$



Jak již bylo řečeno, chceme si vyjádřit 95%-ní interval spolehlivosti pro 49 700, 50 000 a 50 300 platných podpisů. To znamená, že volíme $\alpha = 0,05$, což nám říká, že s 95%-ní pravděpodobností bude daný interval obsahovat skutečnou hodnotu $P(Z)$. Počet simulací je 20 000. Ve skutečnosti pomocí vzorce (4.12) získáme (pro $\alpha = 0,05$) přibližně 95%-ní interval spolehlivosti. Chceme-li zjistit, jak přesná byla simulace, vyjádříme si ze získaných intervalů spolehlivosti jejich šířky e. Čím je šířka e kratší (pro stejné α), tím je interval spolehlivosti pro skutečnou hodnotu přesnější. Šířku e získáme odečtením krajních hodnot intervalu spolehlivosti a dosazením za $\alpha = 0,05$ a $s = 20\,000$. Výsledný předpis pro e je závislý už jen na odhadu pravděpodobnosti zamítnutí $\hat{p}$

$$e = 2 \cdot u_{1-\frac{0,05}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{20000}} = \frac{2 \cdot 1,96}{\sqrt{20000}} \cdot \sqrt{\hat{p} \cdot (1-\hat{p})}$$

Pokud hodnoty $\hat{p}$ budou blízké 0, resp. 1, bude šířka e menší než pro hodnoty $\hat{p}$ kolem 0,5. Což potvrzují i tabulky 5.4 pro $N - K = 50\,300$, kdy hodnoty $\hat{p}$ jsou blízká nule, resp. 5.2, a 5.3 pro $N - K = 50\,000$, kdy hodnoty $\hat{p}$ se pohybují kolem 0,05.

Nyní se podívejme, jaký vliv má velikost N na šířku intervalu. Z tabulek 5.2 a 5.4 vyplývá, že větší přesnost simulace pro skutečnou hodnotu $P(Z)$ se dosahuje pro menší N. Kdežto pro počet platných podpisů $N - K = 50\,000$ nemá velikost N, při zaokrouhlení na 4 desetinná místa, žádný vliv na rozdílnost přesností, jak vyplývá z tabulky 5.3.

Tabulka 5.2: Přehled intervalů spolehlivosti pro $N - K = 49\,700$

N	Interval spolehlivosti	Délka intervalu
56 191	(0,9951, 0,9969)	0,0017
59 165	(0,9799, 0,9836)	0,0037
61 966	(0,9661, 0,9709)	0,0048
72 434	(0,9030, 0,9110)	0,0081
81 199	(0,8674, 0,8767)	0,0093
87 782	(0,8347, 0,8448)	0,0102
101 261	(0,8010, 0,8119)	0,0110
106 081	(0,7844, 0,7956)	0,0113

Tabulka 5.3: Přehled intervalů spolehlivosti pro $N - K = 50\,000$

N	Interval spolehlivosti	Šířka intervalu
56 191	(0,4921, 0,5059)	0,0139
59 165	(0,4887, 0,5026)	0,0139
61 966	(0,4970, 0,5108)	0,0139
72 434	(0,4914, 0,5053)	0,0139
81 199	(0,5047, 0,5186)	0,0139
87 782	(0,4974, 0,5113)	0,0139
101 261	(0,4955, 0,5093)	0,0139
106 081	(0,4870, 0,5008)	0,0139

Tabulka 5.4: Přehled intervalů spolehlivosti pro $N - K = 50\,300$

N	Interval spolehlivosti	Šířka intervalu
56 191	(0,0018, 0,0031)	0,0014
59 165	(0,0122, 0,0155)	0,0032
61 966	(0,0279, 0,0327)	0,0048
72 434	(0,0844, 0,0922)	0,0079
81 199	(0,1294, 0,1388)	0,0094
87 782	(0,1533, 0,1635)	0,0101
101 261	(0,1931, 0,2041)	0,0111
106 081	(0,2089, 0,2202)	0,0114

Kapitola 6

Závěr

Bakalářská práce se zabývala kontrolou správnosti údajů na dodaných petičních arších prováděnou Ministerstvem vnitra České republiky při historicky první přímé volbě prezidenta České republiky. Cílem této bakalářské práce bylo seznámit se se zákonem předepsaným postupem pro kontrolu správnosti údajů na dodaných petičních arších a vypočítat pravděpodobnosti zamítnutí prezidentské kandidatury, když celkový počet dodaných podpisů obsahoval určitý počet neplatných, resp. platných, podpisů. Pro různé hodnoty celkového počtu dodaných podpisů pod peticí potenciálními prezidentskými kandidáty byly vypočítány skutečné pravděpodobnosti zamítnutí prezidentské kandidatury. Také byla provedena simulace, na jejímž základě byly získány odhady pravděpodobnosti zamítnutí prezidentské kandidatury.

Nejprve byl odvozen předpis pro výpočet skutečných pravděpodobností zamítnutí prezidentské kandidatury pomocí definice sjednocení disjunktních jevů a následnou aplikací věty o úplné pravděpodobnosti. Poté byla uvedena myšlenka postupu simulace a byl vyjádřen tvar přibližně $100(1-\alpha)\%$-ního intervalu spolehlivosti pro odhady pravděpodobností zamítnutí prezidentské kandidatury získané právě touto simulací.

V závěrečné kapitole byly uvedeny číselné výsledky získané postupem uvedeným v předcházející kapitole. Pro získání výsledků byl použit matematický software MATLAB od společnosti The MathWorks.

Nejprve jsme porovnali závislost pravděpodobností zamítnutí prezidentských kandidatur na počtu platných podpisů v celkovém počtu dodaných podpisů pro všechny potenciální kandidáty. Porovnáním krajních hodnot celkového počtu dodaných podpisů jsme ověřili, že větší pravděpodobnosti registrace kandidatury docílíme tím, že dodáme větší počet podpisů, což jsme intuitivně předpokládali.

V závěru jsme pomocí simulace získaly odhady pravděpodobností zamítnutí prezidentských kandidatur. Na jejich základě jsme si vyjádřily 95%-ní intervaly spolehlivosti, když celkový počet dodaných podpisů pod peticí obsahoval 49 700, 50 000 a 50 300 platných podpisů. Zjistili jsme, že absolutní chyba intervalů spolehlivosti pro 50 000 platných podpisů v celkovém počtu dodaných je pro všechny zvolené počty dodaných podpisů stejná. Zatímco pro 50 300. resp. 49 700, platných podpisů ve všech dodaných se absolutní chyba s větším počtem dodaných popisů zvětšuje.

Literatura

[1] Anděl, J.: *Matematická statistika.* SNTL, Praha, 1985.

[2] Johnson, N. L., Kemp, A. W., Kotz, S.: *Univariate discrete distributions.* Third edition, Wiley, Hoboken, 2005.

[3] Reif, J.: *Metody matematické statistiky.* Západočeská univerzita, Plzeň, 2004.

[4] Vencálek, O.: *Volba prezidenta: problém kontroly podpisů.* Informační bulletin České statistické společnosti 25 (2014), no. 2, 1-11.

[5] ČESKO. Zákon č. 275 ze dne 18. července 2012 o volbě prezidenta republiky a o změně některých zákonů (zákon o volbě prezidenta republiky). In: *Sbírka zákonů České republiky.* 2012, částka 95, s. 3554-3582. ISSN 1211-1244.

[6] MINISTERSTVO VNITRA ČR, *Ministerstvo vnitra České republiky* [online]. [cit. 2016-05-20]. Dostupné z: http://www.mvcr.cz/clanek/rozhodnuti.aspx

Seznam obrázků

Seznam tabulek

Přílohy

Přílohy na CD

- Bakalářská práce
- Vstupní data
- Zdrojové kódy v MATLABu
- Získané výsledky v MS Excel