Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Petra Santnerová

Některé jednovýběrové testy

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Marie Hušková, DrSc.

Studijní program: Matematika, Obecná matematika

2009

Děkuji prof. RNDr. Marii Huškové, DrSc. za vedení mé bakalářské práce a za podnětné návrhy, které ji obohatily.

Prohlašuji, že jsem svou bakalářskou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 4. srpna 2009 Petra Santnerová

Obsah

Název práce: Některé jednovýběrové testy
Autor: Petra Santnerová
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: prof. RNDr. Marie Hušková, DrSc.
e-mail vedoucího: huskova@karlin.mff.cuni.cz

Abstrakt: Cílem této práce je seznámit čtenáře s jednovýběrovým problémem, nejpoužívanějšími jednovýběrovými testy a možnostmi aproximace kritických hodnot. Na začátku této práce jsou sepsány definice a věty, které jsou nezbytné pro řešení daných úloh. V každém testu je popsána testová statistika, kritická hodnota a nulová i alternativní hypotéza testu. Rozhodnutí o platnosti nulové hypotézy je provedeno na základě kritické hodnoty. Důležitou roli v této práci hraje aproximace kritických hodnot. U každého testu je uvedena aproximace pomocí limitní věty. V dalších kapitolách je ukázána aproximace pomocí metody zvané bootstrap, nebo pomocí simulace výběrů z různých rozdělení.

Klíčová slova: testová statistika, kritická hodnota, aproximace.

Title: Some one-sample tests
Author: Petra Santnerová
Department: Department of Probability and Mathematical Statistics
Supervisor: prof. RNDr. Marie Hušková, DrSc.
Supervisor's e-mail address: huskova@karlin.mff.cuni.cz

Abstract: The aim of my bachelor work is to acquaint a reader with the one-sample problem, with the most widely used one-sample tests and with possibilities of critical value approximation. There are quoted some definitions and thesis which are inevitable for solving given tasks in the beginning of this work. A test statistic, critical value and null and alternative hypothesis are described in every test. A decision of null hypothesis' validity is made on the base of the critical value. The critical value approximation is an important part of this work. In each test there is acquainted the approximation by a limit theorem. The approximation by the method called bootstrap, or by the simulation of samples from different distributions are shown in next chapters.

Keywords: test statistic, critical value, approximation.

Úvod

Jak již název napovídá, tato práce je věnována některým jednovýběrovým testům. Ty mají veliký význam v oblasti statistiky a finančnictví.

Jedním z typických motivačních příkladů pro jednovýběrové testy je testování hypotézy, zda daný lék je účinnější než jiný, zda má pohlaví vliv na dovednosti jedince atd.

Mezi příklady jednovýběrových testů patří testy o střední hodnotě, testy o mediánu či pořadové testy.

Cílem této práce je uvést přehled vybraných jednovýběrových testů, aproximovat kritické hodnoty pro dané testy a na simulační studii ukázat vhodnost různých metod aproximace.

Práce je rozdělena na 5 kapitol. V první kapitole se seznámíme se základními definicemi a větami, kterých budeme využívat. V následující kapitole je zformulován jednovýběrový problém. Třetí kapitola je věnována některým jednovýběrovým testům počínaje nejznámějším testem o střední hodnotě pro jednu proměnnou, t-testem, přes jeho asymptotickou verzi, z-test, po nejpoužívanější pořadové testy, jako jsou znaménkový test a jednovýběrový Wilcoxonův test. Tuto kapitolu ukončuje test dobré shody, Kolmogorovův-Smirnovův test. Ve čtvrté kapitole je popsána aproximace metodou bootstrap. Poslední kapitola obsahuje simulační studii pro jednotlivé testy a její výsledky. Simulace jsou učiněny ve statistickém programu R. Součástí této práce je i příloha, ve které najdeme tabulky kritických hodnot a kvantilů některých rozdělení. Elektronická verze práce má přílohu obohacenou o zdrojové kódy programu R.

Kapitola 1

Základní pojmy a definice

Teoretické základy, kterých budeme v této práci využívat, najdeme v prvních kapitolách knih [1] či [4]. Připomeňme některé z nich.

1.1 Náhodné veličiny

Definice 1.1.1 Měřitelnou reálnou funkci $X : (\Omega, \mathcal{A}) \to (\mathcal{R}^1, \mathcal{B}^1)$ nazýváme *náhodnou veličinou.*

Definice 1.1.2 Nechť X je náhodná veličina. Její distribuční funkce F_X je definována vztahem $F_X(x) = P(X \leq x),\ x \in \mathbf{R}$.

Věta 1.1.1 *Distribuční funkce F_X náhodné veličiny X má následující vlastnosti:*

(i) *F_X je neklesající.*

(ii) $\lim_{x\to\infty} F_X(x) = 1, \qquad \lim_{x\to-\infty} F_X(x) = 0.$

(iii) *F_X je zprava spojitá.*

Definice 1.1.3 Nechť F je nějaká distribuční funkce. Pak definujeme *kvantilovou funkci* F^{-1} příslušné distribuční funkce F předpisem

$$F^{-1}(u) = \inf\{x : F(x) \geq u\} \qquad 0 < u < 1.$$

Hodnotám $F^{-1}(u)$ se říká kvantily. Je-li F rostoucí, pak F^{-1} je obyčejná inverzní funkce k F.

1.2 Binomické rozdělení

Nechť n je přirozené číslo a $p \in (0,1)$. Jestliže rozdělení náhodné veličiny X je definováno předpisem

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

pak říkáme, že X má binomické rozdělení s parametry n, p.
Značení: $\mathrm{Bi}(n,p)$
Střední hodnota: $\mathrm{E}X = np$
Rozptyl: $\mathrm{var}X = np(1-p)$
Pro $n = 1$ hovoříme o tzv. alternativním rozdělení.

1.3 Rovnoměrné rozdělení

Nechť (a,b) je konečný nedegenerovaný interval. Hustota rovnoměrného rozdělení má tvar

$$f(x) = \frac{1}{b-a} \qquad \text{pro} \qquad a < x < b.$$

Značení: $\mathrm{R}(a,b)$
Střední hodnota: $\mathrm{E}X = \frac{a+b}{2}$
Rozptyl: $\mathrm{var}X = \frac{(b-a)^2}{12}$

1.4 Studentovo t-rozdělení

Nechť $n \geq 1$. Hustota Studentova t-rozdělení o n stupních volnosti je dána předpisem

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \qquad x \in \mathbf{R}.$$

Značení: t_n
Střední hodnota: $\mathrm{E}X = 0$ pro $n > 1$
Rozptyl: $\mathrm{var}X = \frac{n}{n-2}$ pro $n > 2$
Pro $n = 1$ dostáváme Cauchyovo rozdělení C(0,1).

1.5 Normální rozdělení

Nechť $\mu \in \mathbf{R}$, $\sigma > 0$ jsou nějaké konstanty. Pak hustota normálního rozdělení je dána předpisem

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbf{R}.$$

Distribuční funkce je

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \Phi\Big(\frac{x-\mu}{\sigma}\Big).$$

Značení: $\mathrm{N}(\mu, \sigma^2)$
Střední hodnota: $\mathrm{E}X = \mu$
Rozptyl: $\mathrm{var}X = \sigma^2$
Normálnímu rozdělení se někdy říká Gaussovo rozdělení. Speciálním případem je tzv. normované normální rozdělení $N(0,1)$, kde $\mu = 0$ a $\sigma = 1$. Hustotu toho rozdělení značíme φ a příslušnou distribuční funkci označujeme Φ.

1.6 Dvojitě exponenciální (Laplaceovo) rozdělení

Nechť $a \in \mathbf{R},\ b > 0$ jsou nějaké konstanty. Hustota dvojitě exponenciálního rozdělení má tvar

$$f(x) = \frac{1}{2b}\, e^{-\frac{|x-a|}{b}}, \qquad x \in \mathbf{R}.$$

Značení: DEx(a,b)
Střední hodnota: $\mathrm{E}X = a$
Rozptyl: $\mathrm{var}X = 2b^2$
Dvojitě exponenciálnímu rozdělení se také říká Laplaceovo rozdělení.

1.7 Konvergence a limitní věty

Definice 1.7.1 Posloupnost náhodných veličin $X_1, X_2, \ldots$ *konverguje v pravděpodobnosti* k náhodné veličině X pro $n \to \infty$ (značíme $X_n \xrightarrow{P} X$), pokud pro každé $\epsilon > 0$ platí

$$\lim_{n\to\infty} P[|X_n - X| \geq \epsilon] = 0.$$

Definice 1.7.2 Nechť $F_1, F_2, \ldots$ je posloupnost distribučních funkcí a nechť F je distribuční funkce. Jestliže $F_n(x) \to F(x)$ v každém bodě x, který je bodem spojitosti funkce F, pak říkáme, že posloupnost F_n *slabě konverguje* k F.

Definice 1.7.3 Nechť náhodné veličiny $X_1, X_2, \ldots$ mají distribuční funkci $F_1, F_2, \ldots$ a nechť X má distribuční funkci F. Říkáme, že X_n *konvergují* k X *v distribuci*, jestliže F_n konvergují slabě k F. Značíme $X_n \xrightarrow{d} X$.

Definice 1.7.4 Posloupnost náhodných veličin $X_1, X_2, \ldots$ *konverguje skoro jistě* k náhodné veličině X (značíme $X_n \xrightarrow{s.j.} X$), pokud platí

$$\lim_{n\to\infty} P[X_n \to X] = 1.$$

Věta 1.7.1 ***(Kolmogorovova)*** *Nechť $X_1, X_2, \ldots$ jsou nezávislé stejně rozdělené náhodné veličiny s konečnou střední hodnotou μ. Pak $\overline{X_n} \xrightarrow{s.j.} \mu$ pro $n \to \infty$, kde $\overline{X_n} = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$.*
Důkaz: [4] věta 4.8 □

Věta 1.7.2 ***(Lindebergova CLV)*** *Nechť $X_1, X_2, \ldots$ jsou nezávislé stejně rozdělené náhodné veličiny se střední hodnotou μ a s konečným rozptylem $\sigma^2 > 0$. Pak pro $n \to \infty$ platí*

$$\frac{X_1 + \ldots + X_n - n\mu}{\sqrt{n}} \xrightarrow{d} N(0, \sigma^2).$$

Důkaz: [4] věta 4.12 □

Věta 1.7.3 ***(o spojité transformaci)*** *Nechť g je reálná spojitá funkce.*

- *Jestliže $X_n \xrightarrow{P} X$ pro $n \to \infty$, pak $g(X_n) \xrightarrow{P} g(X)$ pro $n \to \infty$;*
- *jestliže $X_n \xrightarrow{d} X$ pro $n \to \infty$, pak $g(X_n) \xrightarrow{d} g(X)$ pro $n \to \infty$,*

kde X a X_n jsou náhodné veličiny.
Důkaz: [1] str. 332. □

Věta 1.7.4 ***(Sluckého)*** *Nechť $X_1, X_2, \ldots$ je posloupnost náhodných veličin s distribučními funkcemi $F_1, F_2, \ldots$. Nechť F je distribuční funkce a $c > 0$ je konstanta. Nechť F_n konvergují slabě k F. Nechť $Y_1, Y_2, \ldots$ je posloupnost náhodných veličin taková, že $Y_n \xrightarrow{P} c$ pro $n \to \infty$. Definujme $S_n = X_n Y_n$. Nechť F_n^S je distribuční funkce S_n. Pak $F_n^S(x)$ konvergují slabě k $F(\frac{x}{c})$.*
Důkaz: [1] str. 333. □

1.8 Testování hypotéz

Definice 1.8.1 Nechť $X_1, \ldots, X_n$ je posloupnost nezávislých stejně rozdělených náhodných veličin s distribuční funkcí F. Pak říkáme, že $X_1, \ldots, X_n$ je *náhodný výběr* z rozdělení s distribuční funkcí F.

Definice 1.8.2 Nechť Θ je množina funkcí, která má aspoň dva různé prvky. Pak Θ nazveme *parametrický prostor*.

Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor s rozdělením, které závisí na parametru $\boldsymbol{\theta} \in \Theta$ takovém, že $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$. Nechť ω je neprázdná vlastní podmnožina prostoru Θ. Předpokládejme, že jsme po rozboru dané situace došli k domněnce, že $\boldsymbol{\theta} \in \omega$. Jelikož tuto domněnku zatím nemáme potvrzenou, nazýváme tvrzení $\boldsymbol{\theta} \in \omega$ *nulovou hypotézou*, značíme $H_0 : \boldsymbol{\theta} \in \omega$. Zbývající možnosti zahrnuje *alternativní hypotéza* $H_1 : \boldsymbol{\theta} \in \omega_1 = \Theta - \omega$. Řekneme, že hypotéza H_0 je *jednoduchá*, jestliže je množina ω jednobodová.

Test hypotéz je popsán *kritickým oborem* $W \subset \mathbf{R^n}$. Pokud $\boldsymbol{X} \in W$, pak hypotézu H_0 zamítneme. V opačném případě ji nezamítneme.

Při rozhodování o platnosti některé hypotézy se můžeme dopustit chyby. Pokud H_0 platí a my ji zamítneme, dopustíme se *chyby prvního druhu*. Pokud H_0 neplatí a my ji nezamítneme, dopustíme se *chyby druhého druhu*. Zpravidla volíme kritický obor tak, aby pravděpodobnost chyby prvního druhu byla menší nebo rovna předem danému kladnému číslu α, tj. aby $P_{\boldsymbol{\theta}}(\boldsymbol{X} \in W) \leq \alpha$ pro každé $\boldsymbol{\theta} \in \omega$. Hodnotu $\sup\limits_{\boldsymbol{\theta} \in \omega} P_{\boldsymbol{\theta}}(\boldsymbol{X} \in W)$ nazýváme *hladinou testu*. Při dané hladině testu se také snažíme o co nejmenší pravděpodobnost chyby druhého druhu. Funkce $\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\boldsymbol{X} \in W)$, kde $\boldsymbol{\theta} \in \Theta$, se nazývá *síla testu*. Číslo $\beta(\boldsymbol{\theta})$ udává pravděpodobnost, že zamítneme H_0, když hodnota parametru je $\boldsymbol{\theta}$.

Kapitola 2

Jednovýběrový problém

Jedním z nejběžnějších testových problémů se zdá být testování parametru polohy náhodného výběru $X_1, \ldots, X_n$ z nějakého daného rozdělení. Tento problém je jednoduchý, jestliže můžeme předpokládat, že rozdělení je symetrické kolem neznámého parametru θ, tj.

$$P(X \leq \theta - x) = P(X \geq \theta + x) \qquad \forall x.$$

Příkladem takového rozdělení jsou např. normální, Laplaceovo, Cauchyovo či rovnoměrné rozdělení. Podobným způsobem, jaký je uveden v knize [6], ukážeme, že rozdíl dvou nezávislých pozorování ze stejného výběru má symetrické rozdělení.

Máme dány pokusné jednotky $Y_1, \ldots, Y_n$, na které budeme chtít aplikovat nějaké „ošetření", např. nás bude zajímat, zda nový lék nějak působí na daného jedince Y_i. Utvoříme dvojice pozorování

$$\begin{pmatrix} Y_1 \\ Z_1 \end{pmatrix}, \ldots, \begin{pmatrix} Y_n \\ Z_n \end{pmatrix},$$

kde Y_i je stav jedince před podáním léku a Z_i je stav jedince po podání léku. Dvojice pozorování můžeme chápat jako náhodný výběr z dvourozměrného rozdělení se sdruženou distribuční funkcí $F_{YZ}(y, z)$. Předpokládejme, že tato distribuční funkce je spojitá. Hypotéza, že daný lék nemá na jedince vliv, odpovídá předpokladu, že distribuční funkce $F_{YZ}(y, z)$ je symetrická kolem přímky $y = z$, tj.

$$F_{YZ}(y, z) = F_{YZ}(z, y) \qquad \forall y, z. \tag{2.1}$$

Po transformaci $X_i = Z_i - Y_i$ se pro výše popsanou situaci vžil pojem *jednovýběrový problém*. Nyní máme dán náhodný výběr $X_1, \ldots, X_n$ se spojitou distribuční funkcí F. Je-li X_i symetrické kolem nějakého $\lambda_X \in \mathbf{R}$ pro $i = 1, \ldots, n$, pak $X_i - \lambda_X$ má stejné rozdělení jako $-(X_i - \lambda_X)$,

$$P(X_i - \lambda_X \leq x) = P(-(X_i - \lambda_X) \leq x)$$

$$\begin{aligned} P(X_i \leq x + \lambda_X) &= P(-X_i \leq x - \lambda_X) \\ P(X_i \leq x + \lambda_X) &= 1 - P(X_i \leq -x + \lambda_X) \\ F(x + \lambda_X) &= 1 - F(-x + \lambda_X) \qquad \forall x. \end{aligned} \tag{2.2}$$

Pro hypotézu, že daný lék nemá na jedince vliv, jsme předpokládali symetrii sdružené distribuční funkce $F_{YZ}(y, z)$ kolem přímky $y = z$. Z tohoto předpokladu plyne, že $\lambda_X = 0$ a tedy

$$F(x) + F(-x) = 1 \qquad \forall x. \tag{2.3}$$

Úloha vede na testování hypotézy symetrie

$$H_0 : F(x) + F(-x) = 1 \qquad \forall x \in \mathbf{R} \tag{2.4}$$

proti alternativě H_1 : existuje x, pro které rovnost (2.3) neplatí.

Kapitola 3

Některé jednovýběrové testy

V této kapitole se seznámíme se základními jednovýběrovými testy, jejich popisem, testovou statistikou a aproximací kritických hodnot. Teoretické kritické hodnoty, které budeme při výpočtech potřebovat, najdeme v tabulkách v příloze A. Nejprve uvedeme testy o střední hodnotě pro jednu proměnnou. Nejznámějším z nich je test využívající Studentovo t-rozdělení. Odtud název t-test. Dále uvedeme jeho asymptotickou modifikaci, asymptotický z-test. Poté se seznámíme se dvěma nejpoužívanějšími pořadovými testy, znaménkovým a jednovýběrovým Wilcoxonovým testem. Na konci této kapitoly nalezneme test dobré shody, Kolmogorovův-Smirnovův test.

3.1 Testy o rovnosti středních hodnot

3.1.1 Jednovýběrový t-test

Mějme náhodný výběr $X_1, ..., X_n$ z normálního rozdělení $N(\mu, \sigma^2)$, kde $\mu \in \mathbf{R}$, $\sigma^2 > 0$. Budeme testovat hypotézu $H_0 : \mu = \mu_0$ proti alternativě $H_1 : \mu \neq \mu_0$, kde $\mu_0 \in \mathbf{R}$ dané. Jde vlastně o test hypotézy symetrie kolem μ_0. Metodou podílem věrohodností dostaneme následující testovou statistiku:

$$T_n = \sqrt{n}\,\frac{\overline{X_n} - \mu_0}{S_n}, \tag{3.1}$$

kde

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2.$$

Hypotézu H_0 zamítáme pro velké $|T_n|$, tj. $|T_n| \geq c_n(\alpha)$, kde $c_n(\alpha)$ je vhodně zvolená konstanta. Poznamenejme, že za platnosti hypotézy H_0, můžeme $c_n(\alpha)$ určit z věty níže.

Věta 3.1.1 *Nechť $X_1, ..., X_n$ je náhodný výběr z $N(\mu_0, \sigma^2)$, kde $n \geq 2$ a $\sigma^2 > 0$. Pak testová statistika T_n má Studentovo rozdělení o $n-1$ stupních volnosti.*
Důkaz: Viz. [1] str. 74. □

Tedy H_0 zamítáme na hladině α právě tehdy, když $|T_n| \geq t_{n-1}(1-\frac{\alpha}{2})$, kde $t_{n-1}(1-\frac{\alpha}{2})$ je $(1-\frac{\alpha}{2})-$tý kvantil Studentova rozdělení o $(n-1)$ stupních volnosti.

Poznámka Jak uvidíme v nadcházející větě, pro velká n (30 a více) se Studentovo rozdělení blíží normovanému normálnímu rozdělení.

Věta 3.1.2 *Platí:*

$$\lim_{n\to\infty} H_n(x) = \Phi(x), \qquad x \in \mathbf{R},$$

kde H_n je distribuční funkce Studentova rozdělení o n stupních volnosti.
Důkaz: [4] str. 108. □

Poznámka Obdobně můžeme uvažovat i jednostranné testy. Kvantily Studentova rozdělení jsou uvedeny v tab. A.4.

Pro levostranný test budeme testovat hypotézu $H_0 : \mu_X = \mu_0$ proti alternativě $H_1 : \mu_X > \mu_0$. Pak hypotézu H_0 zamítneme právě tehdy, když $T_n \geq t_{n-1}(1-\alpha)$.

Pro pravostranný test budeme testovat hypotézu $H_0 : \mu_X = \mu_0$ proti alternativě $H_1 : \mu_X < \mu_0$. Pak hypotézu H_0 zamítneme právě tehdy, když $T_n \leq -t_{n-1}(1-\alpha)$.

Příklad 3.1.1 V prodejně domácích potřeb obdrželi zásilku 20 košťat. Košťata měla mít délku 100 cm, ale přeměřením se zjistilo, že skutečné délky byly tyto:

$$102,\ 104,\ 95,\ 104,\ 101,\ 99,\ 101,\ 103,\ 96,\ 101$$

$$97,\ 96,\ 103,\ 102,\ 101,\ 99,\ 98,\ 105,\ 102,\ 97.$$

Odchylky od požadované délky tedy jsou (v centimetrech):

$$2,\ 4,\ -5,\ 4,\ 1,\ -1,\ 1,\ 3,\ -4,\ 1,-3,\ -4,\ 3,\ 2,\ 1,\ -1,\ -2,\ 5,\ 2,\ -3.$$

Je třeba zjistit, zda se výrobce nedopustil nějaké systematické odchylky od požadované hodnoty.
Jednotlivé odchylky považujeme za realizaci náhodného výběru z $N(\mu, \sigma^2)$, kde $\sigma^2 > 0$ je neznámý parametr. Budeme testovat hypotézu $H_0 : \mu = 0$ proti alternativě $H_1 : \mu \neq 0$ na hladině $\alpha = 0,05$. Pro $n = 20$ nám po dosazení do vzorců vyjde: $\overline{X_{20}} = 0,3$ a $S_{20}^2 = 7,9216$. Tedy testová statistika $T_{20} = 0,4767 < 2,0930 = t_{19}(0,975)$. H_0 nezamítneme na hladině $\alpha = 0,05$. Nemůžeme říct, zda se výrobce dopustil nějaké chyby, neboť námi získaná data to nepotvrzují.

$\triangle$

3.1.2 Asymptotický z-test

Mějme dán náhodný výběr $X_1, ..., X_n$ z rozdělení se střední hodnotou μ_X a konečným kladným rozptylem σ_X^2. Budeme testovat hypotézu $H_0 : \mu_X = \mu_0$ proti alternativě $H_1 : \mu_X \neq \mu_0$. Uvažujme testovou statistiku

$$Z_n = \sqrt{n}\,\frac{\overline{X_n} - \mu_0}{S_n}, \tag{3.2}$$

kde

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad {S_n}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2.$$

Z Lindebergovy věty 1.7.2 víme, že pro $n \to \infty$

$$\sqrt{n}(\overline{X_n} - \mu_X) \xrightarrow{d} N(0, \sigma^2),$$

čili

$$\sqrt{n}\frac{(\overline{X_n} - \mu_X)}{\sigma} \xrightarrow{d} N(0,1).$$

S_n^2 je nestranný a konzistentní odhad σ^2 (viz [4] věta 5.1.), tj. $S_n^2 \xrightarrow{P} \sigma^2$ pro $n \to \infty$. Dle věty 1.7.3 o spojité transformaci

$$\frac{\sigma}{S_n} \xrightarrow{P} 1 \qquad \text{pro } n \to \infty.$$

Ze Sluckého věty 1.7.4 pro $n \to \infty$ plyne

$$\underbrace{\frac{\sigma}{S_n}}_{\xrightarrow{P} 1} \underbrace{\sqrt{n}\frac{(\overline{X_n} - \mu_X)}{\sigma}}_{\xrightarrow{d} N(0,1)} \xrightarrow{d} N(0,1)$$

Předpokládejme rovnost $\mu_X = \mu_0$, pak H_0 zamítneme na hladině α právě tehdy, když $|Z_n| \geq u_{1-\frac{\alpha}{2}}$, kde u_p je p-kvantil normálního rozdělení, který nalezneme v tab. A.3.

Poznámka: Jestliže H_0 neplatí, pak

$$\sqrt{n}\ \frac{\overline{X_n} - \mu_0}{S_n} = \underbrace{\sqrt{n}\ \frac{\overline{X_n} - \mu_X}{S_n}}_{\xrightarrow{d} N(0,1)} \quad + \quad \sqrt{n}\ \frac{\mu_X - \mu_0}{S_n}$$

a síla testu má hodnotu

$$\beta(\mu_X) = P_{\mu_X}(|Z_n| > u_{1-\frac{\alpha}{2}}) = P_{\mu_X}(Z_n < -u_{1-\frac{\alpha}{2}}) + P_{\mu_X}(Z_n > u_{1-\frac{\alpha}{2}}) =$$

$$= P_{\mu_X}\Big(\sqrt{n}\ \frac{\overline{X_n} - \mu_X}{S_n} < -u_{1-\frac{\alpha}{2}} - \sqrt{n}\ \frac{\mu_X - \mu_0}{S_n}\Big) \quad +$$

$$+ \quad P_{\mu_X}\Big(\sqrt{n}\ \frac{\overline{X_n} - \mu_X}{S_n} > u_{1-\frac{\alpha}{2}} - \sqrt{n}\ \frac{\mu_X - \mu_0}{S_n}\Big)$$

$$\xrightarrow{n\to\infty} \Phi\Big(-u_{1-\frac{\alpha}{2}} - \sqrt{n}\ \frac{\mu_X - \mu_0}{\sigma}\Big) + 1 - \Phi\Big(u_{1-\frac{\alpha}{2}} - \sqrt{n}\ \frac{\mu_X - \mu_0}{\sigma}\Big),$$

kde Φ je distribuční funkce normovaného normálního rozdělení.

Příklad 3.1.2 Na výše popsaný příklad s košťaty aplikujeme asymptotický z-test. Hodnotu testové statistiky pro z-test již známe, neboť testová statistika z-testu a t-testu je stejná. Rozdíl je v zamítacím pravidle.

$n = 20,\ |Z_{20}| \geq u_{1-\frac{\alpha}{2}}$. Víme, že $Z_{20} = 0,4767 < 1,960 = u_{0,975}$, proto H_0 nezamítneme na asymptotické hladině $\alpha = 0,05$.

$\triangle$

3.2 Nejpoužívanější pořadové testy

Pořadové testy užívají místo hodnot náhodných veličin jejich indexy z řady, v níž jsou seřazeny dle velikosti od nejmenší po největší. Výhodou těchto testů je, že nepotřebujeme znát rozdělení, ze kterého máme daný náhodný výběr. Obvykle se předpokládá spojitost distribuční funkce. Níže si uvedeme pouze dva pořadové testy, znaménkový a jednovýběrový Wilcoxonův test.

Poznámka: Pokud platí hypotéza symetrie (2.4), pak medián je roven nule. Pokud existuje střední hodnota, pak je také rovna nule.

3.2.1 Znaménkový test

Mějme dán náhodný výběr $X_1, \ldots, X_n$ z rozdělení s distribuční funkcí, která je spojitá na okolí m_X, kde m_X je medián toho rozdělení, tj.

$$P[X_i \leq m_X] = \frac{1}{2} = P[X_i \geq m_X].$$

Budeme testovat hypotézu $H_0 : m_X = m_0$ proti hypotéze $H_1 : m_X \neq m_0$. Testovou statistiku založíme na znaménkách veličin $X_i - m_0$. Označme proto Y_n počet rozdílů s kladným znaménkem, tj.

$$Y_n = \sum_{i=1}^{n} I_{(0,\infty)}(X_i - m_0). \tag{3.3}$$

Platí-li nulová hypotéza, pak $P_{H_0}[X_i \leq m_0] = \frac{1}{2} = P_{H_0}[I_{(0,\infty)}(X_i - m_0) = 1]$, tedy Y_n je součtem alternativních navzájem nezávislých náhodných veličin a má binomické rozdělení s parametry $n, \frac{1}{2}$, tj.

$$Y_n \sim Bi\Big(n, \frac{1}{2}\Big), \qquad P[Y_n = k] = \binom{n}{k}\frac{1}{2^n} \qquad k = 0, \ldots, n$$

H_0 zamítneme na hladině α právě tehdy, když $Y_n \leq a_n(\alpha)$ nebo $Y_n \geq b_n(\alpha)$, kde $a_n(\alpha)$ je největší celé číslo takové, že $\sum\limits_{k=0}^{a_n(\alpha)} \binom{n}{k} \frac{1}{2^n} \leq \frac{\alpha}{2}$ a $b_n(\alpha)$ je nejmenší celé číslo takové, že $\sum\limits_{k=b_n(\alpha)}^{n} \binom{n}{k} \frac{1}{2^n} \leq \frac{\alpha}{2}$. Pro malá n najdeme hodnoty v tab. A.2.

Pro velká n použijeme CLV pro binomické rozdělení (viz. [4] str. 87) a pro $n \to \infty$ za platnosti H_0 dostaneme:

$$\frac{2}{\sqrt{n}} Y_n - \sqrt{n} \xrightarrow{d} N(0,1).$$

Pak hypotézu H_0 zamítneme na asymptotické hladině α právě tehdy, když $|\frac{2}{\sqrt{n}} Y_n - \sqrt{n}| \geq u_{1-\frac{\alpha}{2}}$, kde u_p je p-kvantil normálního rozdělení.

Poznámka: Větu lze použít i v obecnější podobně, kde se testuje rovnost mediánů u nezávislých, ne nutně stejně rozdělených náhodných veličin.

Příklad 3.2.1 Použijeme znaménkový test na příklad s košťaty. Délky košťat (v centimetrech) jsou následující:

$$102,\ 104,\ 95,\ 104,\ 101,\ 99,\ 101,\ 103,\ 96,\ 101$$

$$97,\ 96,\ 103,\ 102,\ 101,\ 99,\ 98,\ 105,\ 102,\ 97.$$

Testujeme hypotézu, že medián rozdělení je $m_0 = 100\,\mathrm{cm}$, tj. $H_0 : m_X = m_0$. Spočteme veličiny $X_i - m_0$:

$$2,\ 4,\ -5,\ 4,\ 1,\ -1,\ 1,\ 3,\ -4,\ 1,$$

$$-3,\ -4,\ 3,\ 2,\ 1,\ -1,\ -2,\ 5,\ 2,\ -3.$$

$n = 20$. Počet rozdílů s kladným znaménkem $Y_{20} = 12$. Dle tabulky kritických hodnot znaménkového testu pro hladinu $\alpha = 0,05$ je $a_{20}(0,05) = 5 < Y_{20} < 15 = b_{20}(0,05)$. Podle těchto výsledků nulovou hypotézu nezamítáme na hladině $\alpha = 0,05$.

$\triangle$

3.2.2 Jednovýběrový Wilcoxonův test

Mějme náhodný výběr $X_1, \ldots, X_n$ z rozdělení se spojitou distribuční funkcí, která je symetrická, tj. $F(x) + F(-x) = 1 \ \forall x \in \mathbf{R}$. Ze spojitosti distribuční funkce plyne, že $P(X_i = 0) = 0$ pro $i = 1, \ldots, n$.

Budeme testovat hypotézu $H_0 : m_X = 0$ proti alternativě $H_1 : m_x \neq 0$. Označme R_i^+ pořadí veličiny $|X_i|$ mezi $|X_1|, \ldots, |X_n|$ a definujme

$$S_n^+ = \sum_{X_i \geq 0} R_i^+, \qquad S_n^- = \sum_{X_i < 0} R_i^+.$$

Za testovou statistiku zvolíme

$$\min(S_n^+, S_n^-). \tag{3.4}$$

Označíme-li $S_n = \sum_{i=1}^{n} R_i^+ \operatorname{sgn} X_i$, pak platí

$$S_n^+ - S_n^- = S_n, \ S_n^+ + S_n^- = \frac{n(n+1)}{2}. \tag{3.5}$$

H_0 zamítneme na hladině α, pokud $\min(S_n^+, S_n^-) \leq c_n(\alpha)$, kde $c_n(\alpha)$ je celé číslo, které lze pro malé n najít v tabulce A.1. Pro velká n za platnosti H_0 platí (viz. Šidák [5] věta 1.7 str. 166)

$$W_n^+ = \frac{S_n^+ - E_{H_0} S_n^+}{\sqrt{\operatorname{var}_{H_0} S_n^+}} \xrightarrow{d} N(0,1) \text{ pro } n \to \infty,$$

kde $E_{H_0} S_n^+ = \frac{n(n+1)}{4}$ a $\operatorname{var}_{H_0} S_n^+ = \frac{n(n+1)(2n+1)}{24}$.

H_0 zamítneme na hladině, která se s rostoucím n blíží k α, právě tehdy, když $|W_n^+| \geq u_{1-\frac{\alpha}{2}}$, kde u_p je p-kvantil normálního rozdělení.

Vyjádříme-li S_n^+ ze vzorce (3.5), pak po dosazení do W_n^+ dostaneme

$$W_n^+ = \frac{(\frac{n(n+1)}{2} - S_n^-) - E_{H_0}(\frac{n(n+1)}{2} - S_n^-)}{\sqrt{\operatorname{var}_{H_0}(\frac{n(n+1)}{2} - S_n^-)}} = -\frac{S_n^- - E_{H_0} S_n^-}{\sqrt{\operatorname{var}_{H_0} S_n^-}} = -W_n^-$$

Opět si ukážeme aplikaci testu na příkladu s košťaty.

Příklad 3.2.2 Máme dány odchylky od požadované délky košťat (v centimetrech):

$$2,\ 4, -5,\ 4,\ 1, -1,\ 1,\ 3, -4,\ 1, -3, -4,\ 3,\ 2,\ 1, -1, -2,\ 5,\ 2, -3.$$

Budeme testovat hypotézu, že medián rozdělení odchylek je $m_0 = 0$, tj. $H_0 : m_X = 0$ proti alternativě $H_1 : m_x \neq 0$. Odchylky seřadíme vzestupně podle jejich absolutní hodnoty. Máme

$$-1, -1,\ 1,\ 1,\ 1,\ 1, -2, 2, 2, 2, -3, -3, 3, 3, -4, -4, 4, 4, -5, 5.$$

Sečteme pořadí kladných čísel. $S_{20}^{+} = 127$. Pak $S_{20}^{-} = \frac{20(20+1)}{2} - S_{20}^{+} = 210 - 127 = 83$. Dostaneme, že $\min(S_{20}^{+}, S_{20}^{-}) = 83 > 52 = c_{20}(0,05)$. Nulovou hypotézu nezamítneme na hladině $\alpha = 0,05$.

$\triangle$

3.3 Test dobré shody

Testy dobré shody se používají ke zjištění, zda daný výběr pochází z určitého rozdělení. Jako příklad testu dobré shody uvedeme Kolmogorovův-Smirnovův test. Existuje celá řada jeho modifikací. My zde předvedeme pouze základní model.

3.3.1 Kolmogorovův-Smirnovův test

Mějme náhodný výběr $X_1, \ldots, X_n$ z rozdělení se spojitou distribuční funkcí F_X. Budeme testovat hypotézu $H_0 : F_X(x) = F_0(x) \quad \forall x$, proti alternativě $H_1 : \exists x : F_X(x) \neq F_0(x)$, kde F_0 je zcela specifikováno. Testová statistika je založená na empirické distribuční funkci

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i < x),$$

kde $I(X_i < x)$ je indikátor daného jevu, který nabývá hodnot

$$I(X_i < x) = \begin{cases} 1 & X_i < x \\ 0 & \text{jinak.} \end{cases}$$

Předpis statistiky uvedený Kolmogorovem a Smirnovem má tvar

$$D_n = \sup_{x \in \mathbf{R}} |F_n(x) - F_0(x)|. \tag{3.6}$$

Pokud platí H_0, pak zřejmě D_n je malé a naopak, pokud H_0 neplatí, pak D_n očekáváme velké. Hypotézu H_0 zamítneme na hladině α právě tehdy, když $D_n \geq c_n(\alpha)$, kde $P(D_n \leq c_n(\alpha)) = 1 - \alpha$. Pro malá $n \leq 20$ lze $c_n(\alpha)$ najít v tabulce A.5.

Lze ukázat (viz. [1] str. 241), že pro $x > 0$ platí:

$$\lim_{n \to \infty} P[\sqrt{n} D_n < x] = K(x), \tag{3.7}$$

kde

$$K(x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2 x^2}. \tag{3.8}$$

Do vzorce (3.8) dosadíme $x = \sqrt{n} D_n$ a vypočteme hodnotu $K(\sqrt{n} D_n)$. Pokud $K(\sqrt{n} D_n) \geq 1 - \alpha$, pak hypotézu H_0 zamítneme na hladině, která se pro $n \to \infty$ blíží k číslu α. Pro velká n se kritická hodnota pro veličinu D_n aproximuje číslem $D_n^*(\alpha) = \sqrt{\frac{1}{2n} \ln \frac{2}{\alpha}}$, tj. hypotézu H_0 zamítneme, když $D_n \geq D_n^*(\alpha)$.

Poznámka: Kolmogorovův-Smirnovův test nemusí být zcela specifikovaný, tj. neznáme parametry rozdělení s distribuční funkcí F_0. V takovém případě použijeme odhady těchto parametrů, např. střední hodnotu μ odhadneme výběrovým průměrem $\overline{X}_n$ a rozptyl σ^2 odhadneme výběrovým rozptylem S_n^2.

Použití Kolmogorovova-Smirnovova testu ukážeme na již známém příkladu s košťaty.

Příklad 3.3.1 V prodejně domácích potřeb obdrželi zásilku 20 košťat. Košťata měla mít délku 100 cm, ale přeměřením se zjistilo, že skutečné délky jsou tyto:

$$102,\ 104,\ 95,\ 104,\ 101,\ 99,\ 101,\ 103,\ 96,\ 101$$

$$97,\ 96,\ 103,\ 102,\ 101,\ 99,\ 98,\ 105,\ 102,\ 97.$$

Pro jednoduchost budeme předpokládat, že Kolmogorovův-Smirnovův test je zcela specifikovaný ($\mu = 100$ a $\sigma^2 = 3$). V obecném případě bychom test aplikovali na náhodné veličiny $Y_i = \frac{X_i - \overline{X_n}}{S_n}$, kde X_i jsou jednotlivé délky košťat, $\overline{X_n}$ je průměrná délka košťat a S_n^2 je výběrový rozptyl délek košťat.

Na hladině $\alpha = 0,05$ budeme testovat hypotézu, zda délky košťat pocházejí z $N(100, 3)$ oproti alternativě, že z tohoto rozdělení nepocházejí. Označme X_i délku jednotlivých košťat. Testujeme tedy $H_0 : F(x) = F_0(x)$, kde $F_0(x)$ je distribuční funkce normálního rozdělení o parametrech $\mu = 100$ a $\sigma^2 = 3$ proti $H_1 : \exists x : F_X(x) \neq F_0(x)$. Pro snadnější počítání upravíme testovou statistiku (3.6) následovně: $D_n = \max(D_1^*, \ldots, D_n^*)$, kde $D_i^* = \max\Big\{ \underbrace{\Big|F_0(x_i) - \frac{i-1}{n}\Big|}_{D_i'}, \underbrace{\Big|\frac{i}{n} - F_0(x_i)\Big|}_{D_i''} \Big\}$. Jak je vidět z tabulky 3.1, $D_{20} = 0,231 < 0,29408 = c_{20}(0,05)$, proto hypotézu H_0 nezamítneme na hladině $\alpha = 0,05$.

$\triangle$

Poznámka: Obdobně jako Kolmogorovův-Smirnovův test se používá Smirnovův test pro symetrii: Máme dány nezávislé stejně rozdělené náhodné veličiny $X_1, \ldots, X_n$ se spojitou distribuční funkcí F, pro kterou existuje x takové, že $F(x) + F(-x) \neq 1$. Smirnovem navržená statistika je dána předpisem

$$U_n = \sqrt{\frac{n}{2}} \sup_x |F_n(x) + F_n(-x) - 1|,$$

kde $F_n(x)$ je empirická distribuční funkce výběru $X_1, \ldots, X_n$. Další informace o Smirnovově testu lze nalézt v knize [7] str. 68 a v článku [2].

Tabulka 3.1: Hodnoty pro výpočet statistiky D_n

Pořadí i	seřazené hodnoty $x_{(i)}$	$\frac{i-1}{n}$	$\frac{i}{n}$	$F_0(x_{(i)})$	$D_i^{'}$	$D_i^{''}$	D_i^{*}
1	95	0	0,05	0,048	0,048	0,002	0,048
2	96	0,05	0,1	0,091	0,041	0,009	0,041
3	96	0,1	0,15	0,091	0,009	0,059	0,059
4	97	0,15	0,2	0,159	0,009	0,041	0,041
5	97	0,2	0,25	0,159	0,041	0,091	0,041
6	98	0,25	0,3	0,252	0,002	0,048	0,048
7	99	0,3	0,35	0,369	0,069	0,019	0,069
8	99	0,35	0,4	0,369	0,019	0,031	0,031
9	101	0,4	0,45	0,631	0,231	0,181	**0,231**
10	101	0,45	0,5	0,631	0,181	0,131	0,181
11	101	0,5	0,55	0,631	0,131	0,081	0,131
12	101	0,55	0,6	0,631	0,081	0,031	0,081
13	102	0,6	0,65	0,748	0,148	0,098	0,148
14	102	0,65	0,7	0,748	0,098	0,048	0,098
15	102	0,7	0,75	0,748	0,048	0,002	0,048
16	103	0,75	0,8	0,841	0,091	0,041	0,091
17	103	0,8	0,85	0,841	0,041	0,009	0,041
18	104	0,85	0,9	0,909	0,059	0,009	0,059
19	104	0,9	0,95	0,909	0,009	0,041	0,041
20	105	0,95	1	0,952	0,002	0,048	0,048

Kapitola 4

Aproximace kritických hodnot

V předchozí kapitole jsme se věnovali testovým statistikám a v některých případech jsme uvedli i kritické hodnoty. Zde se budeme hlouběji zabývat aproximacemi kritických hodnot. Vyjdeme opět ze situace, kde máme dán výběr z rozdělení s distribuční funkcí splňující (2.3). Jelikož nemáme k dispozici žádné další informace o daném rozdělení, pro učinění závěru o platnosti hypotézy budeme muset získat kritické hodnoty některou z metod aproximace, přesněji metodou simulací kritických hodnot a metodou zvanou bootstrap. Simulace jsou vhodné např. u pořadových testů. Bootstrapovou metodu předvedeme na z-testu.

4.1 Aproximace bootstrapem

Metoda bootstrap je příkladem tzv. intenzivní počítačové metody pro statistickou analýzu dat. První článek o této metodě napsal kalifornský univerzitní profesor Bradley Efron v roce 1979. I když existují bootstrapové verze pro závislé nestejně rozdělené náhodné veličiny, my se omezíme pouze na výběr nezávislých stejně rozdělených náhodných veličin s blíže neurčenou distribuční funkcí F. Nechť $G_n(X_1, \ldots, X_n)$ je statistika založená na výběru $X_1, \ldots, X_n$, jejíž rozdělení chceme aproximovat. Algoritmus bootstrapové metody je následovný:

- Pro $n \in \mathbf{N}$ zkonstruujeme empirickou distribuční funkci daného výběru

$\mathbf{X} = (X_1, \ldots, X_n)$, označíme ji F_n, čili

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i < x),$$

kde $I(X_i < x)$ je indikátor daného jevu, který nabývá hodnot

$$I(X_i < x) = \begin{cases} 1 & X_i < x \\ 0 & \text{jinak} \end{cases}$$

- Provedeme náhodný výběr o rozsahu n z rozdělení s distribuční funkcí F_n, který označíme $X^* = (X_1^*, \ldots, X_n^*)$ a budeme ho nazývat *bootstrapovým výběrem*. Tedy n-krát vybíráme s vracením z množiny $X_1 = x_1, \ldots, X_n = x_n$.

Označme $G_n^* = G_n(X_1^*, \ldots, X_n^*)$. Pravděpodobnost konkrétního bootstrapového výběru je pro všechny bootstrapové výběry stejná a je rovna

$$P^*(X_1^* = X_1, \ldots, X_n^* = X_n | X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{n} \cdots \frac{1}{n} = \frac{1}{n^n}.$$

Pro marginální pravděpodobnosti platí

$$P^*(X_i^* = X_j | X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{n} \qquad \forall\, i, j = 1, \ldots, n.$$

Často se podmíněná pravděpodobnost $P^*(X_i^* \leq x | X_1 = x_1, \ldots, X_n = x_n)$ značí pouze $P^*(X_i^* \leq x)$.

Abychom mohli přesně stanovit charakteristiky bootstrapového rozdělení, museli bychom provést všech n^n možných výběrů s vracením z náhodného výběru $X_1, \ldots, X_n$, což je pro malá n velmi náročné a pro velká n takřka nemožné. Při výběru o rozsahu 8, bychom museli provést $8^8 = 16777216$ výběrů a z nich určit příslušné charakteristiky.

Proto se obvykle na bootstrapový výběr $X_1^*, \ldots, X_n^*$ a empirickou distribuční funkci F_n používá metoda *Monte Carlo*, kde B-krát generujeme nezávislý náhodný výběr z rozdělení F_n a pro každý vygenerovaný výběr spočteme hodnotu statistiky G_n^*.

Nyní uvedeme větu z článku [9], která udává důkaz konzistence pro nestandardizované statistiky, rychlost této konvergence a dále rychlost konvergence bootstrapové aproximace rozdělení standardizovaného výběrového průměru.

Věta 4.1.1 *Nechť $X_1, \dots, X_n$ jsou nezávislé stejně rozdělené náhodné veličiny s distribuční funkcí F. Označme*

$$\mathrm{E}X_1 = \mu, \qquad 0 < \mathrm{var}X_1 = \sigma^2 < \infty, \qquad \mu_n^* = \overline{X}_n,$$

$$\rho_\infty(H, G) = \sup_{x \in \boldsymbol{R}} |H(x) - G(x)|, \qquad \sigma^{2*} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2,$$

$$H_n(x) = P(\sqrt{n}(\overline{X}_n - \mu) \le x), \qquad H_n^* = P^*(\sqrt{n}(\overline{X}_n^* - \mu^* \le x),$$

$$\widetilde{H}_n(x) = P(\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \le x), \qquad \widetilde{H}_n^* = P^*(\sqrt{n}\frac{\overline{X}_n^* - \mu^*}{\sigma^*} \le x).$$

(i) *Jestliže $EX_1^2 < \infty$, pak*

$$\rho_\infty(H_n^*, H_n) \stackrel{s.j.}{\to} 0 \qquad \text{pro } n \to \infty.$$

(ii) *Jestliže $EX_1^4 < \infty$, pak*

$$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}\rho_\infty(H_n^*, H_n)}{\sqrt{\log\log n}} = \frac{\sqrt{\mathrm{var}(X_1 - \mu)^2}}{2\sigma^2\sqrt{2\pi e}} \qquad \text{s.j.}$$

(iii) *Jestliže $E|X_1|^3 < \infty$ a F je řešetovitá, tj. existují konstanty c, h takové, že $\sum_{k=-\infty}^{\infty} P(X_1 = c + kh) = 1$, pak*

$$\overline{\lim_{n\to\infty}} \sqrt{n}\rho_\infty(\widetilde{H}_n^*, \widetilde{H}_n) = \frac{h}{\sqrt{2\pi}\sigma} \qquad \text{s.j.}$$

(iv) *Jestliže $E|X_1|^3 < \infty$ a F není řešetovitá, pak*

$$\sqrt{n}\rho_\infty(\widetilde{H}_n^*, \widetilde{H}_n) \stackrel{s.j.}{\to} 0 \qquad \text{pro } n \to \infty.$$

Aplikujeme-li tuto větu na testovou statistiku z-testu (3.2), získáme aproximaci rozdělení $P(Z_n \le x)$ pomocí $P^*(Z_n^* \le x)$. Pak kritickou hodnotu odpovídající Z_n aproximujeme $1 - \alpha$ kvantilem příslušného rozdělení Z_n^*.

Další informace o bootstrapové metodě lze získat v knize [3].

Kapitola 5

Simulační studie

Simulaci provedeme ve statistickém programu R. Vygenerujeme náhodné výběry o rozsahu $n_1 = 20$ a $n_2 = 60$. Budeme chtít odhadnout kritické hodnoty statistik jednotlivých testů. Pro z-test budeme kritické hodnoty aproximovat bootstrapovou metodou. V ostatních testech použijeme simulaci. Testy provedeme pro hypotézu $H_0 : \mu = 0$ proti alternativě $H_1 : \mu \neq 0$ na hladině $\alpha = 0,05$.

V případě Kolmogorovova-Smirnovova testu budeme, na hladině $\alpha = 0,05$, testovat hypotézu $H_0 : F_X(x) = \Phi(x) \quad \forall x$ proti alternativě $H_1 : \exists x$ takové, že $F_X(x) \neq \Phi(x)$, kde Φ je distribuční funkce rozdělení $\mathrm{N}(0,1)$.

5.1 Teoretické hodnoty

Tabulka 5.1: Teoretické kritické hodnoty

Test	Hodnota pro n_1	Hodnota pro n_2
t-test	2,093	2,001
z-test	1,960	1,960
znaménkový test	5; 15	21; 39
Wilcoxonův test	52	648

V předchozí tabulce 5.1 jsou uvedeny potřebné teoretické kritické hodnoty pro hladinu $\alpha = 0,05$, kterých budeme využívat k zamítání hypotéz v dalších sekcích této práce. Tyto hodnoty jsme získali z knihy [8], jejíž část je uvedena v příloze A.

V tabulce 5.2 jsou uvedeny teoretické kritické hodnoty Kolmogorovova-Smirnovova testu pro hladinu $\alpha = 0,05$. Kritické hodnoty jsme získali z knihy [8], jejíž část je také uvedena v příloze A.

Tabulka 5.2: Teoretické kritické hodnoty Kolmogorovova-Smirnovova testu

Test	Hodnota pro n_1	Hodnota pro n_2
K-S test	0,29408	0,17231

5.2 Bootstrap

Uvažujme následující výběry o rozsahu $n_1 = 20$ a $n_2 = 60$:

A výběr z normálního rozdělení $\mathrm{N}(0, 1)$;

B výběr z normálního rozdělení $\mathrm{N}(\frac{1}{2}, 1)$;

C výběr z normálního rozdělení $\mathrm{N}(1, 1)$;

D výběr z Laplaceova rozdělení $\mathrm{DEx}(0, 1)$;

E výběr z Laplaceova rozdělení $\mathrm{DEx}(\frac{1}{2}, 1)$;

F výběr z Laplaceova rozdělení $\mathrm{DEx}(1, 1)$;

Pro každý z výše popsaných výběrů provedeme $N = 500$ bootstrapových výběrů, ke kterým spočteme hodnotu statistiky příslušného testu. Výsledné hodnoty seřadíme dle absolutní hodnoty od nejmenší po největší. Hledaná kritická hodnota bude $N \cdot (1 - \alpha)$-tý prvek námi pozorovaných hodnot.

Nyní uvedeme předpis pro statistiku spočtenou z bootstrapového výběru. Nechť $X_1, \ldots, X_n$ je výběr z daného rozdělení. K němu vytvoříme bootstrapový výběr $X_1^*, \ldots, X_n^*$. Pak testová statistika pro z-test je daná předpisem:

$$Z_n^* = \sqrt{n}\,\frac{\overline{X_n^*} - \overline{X_n}}{S_n^*},$$

kde

$$\overline{X_n^*} = \frac{1}{n}\sum_{i=1}^{n} X_i^*, \qquad S_n^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^* - \overline{X_n^*})^2.$$

Tabulka 5.3: Bootstrapové kritické hodnoty pro z-test

Rozsah	A	B	C	D	E	F
n_1	2,033	1,996	1,948	2,170	1,955	1,978
n_2	2,016	1,910	1,954	2,049	1,879	1,907

Tabulka 5.4: Rozdíl teoretické a bootstrapové kritické hodnoty pro z-test

Rozsah	A	B	C	D	E	F
n_1	-0,073	-0,036	0,012	-0,21	0,005	-0,018
n_2	-0,056	0,05	0,006	-0,089	0,081	0,053

5.3 Simulace

Simulaci provedeme pro t-test, pořadové testy a Kolmogorovův-Smirnovův test.

Začneme **t-testem**. Vytvoříme si náhodný výběr z Laplaceova rozdělení s parametry (0; 1), (0,5; 1) a (1; 1) o rozsahu $n_1 = 20$ a $n_2 = 60$. Chceme získat testovou statistiku pro nulovou hypotézu. Spočteme testovou statistiku t-testu

$$T_n = \sqrt{n}\,\frac{\overline{X_n} - \mu_i}{S_n} \qquad \mu_i = 0; 0,5; 1,$$

jejíž absolutní hodnotu uložíme do vektoru V. Tento krok opakujeme 500-krát. Jako kritickou hodnotu zvolíme takové číslo x_α, pro které platí

$$\frac{\text{počet } V[i] \leq x_\alpha}{500} = 1 - \alpha, \qquad i = 1, \dots, 500,$$

tj. kritická hodnota je $500 \cdot (1-\alpha)$-tá nejmenší složka vektoru V. Výsledky nalezneme v tabulce 5.5.

Tabulka 5.5: Simulované kritické hodnoty pro t-test

Výběr	n_1	n_2
DEx(0,1)	2,103	1,943
DEx($\frac{1}{2}$,1)	2,010	2,016
DEx(1,1)	2,058	2,098

Tabulka 5.6: Rozdíl teoretické a simulované kritické hodnoty pro t-test

Výběr	n_1	n_2
DEx(0,1)	-0,01	0,058
DEx($\frac{1}{2}$,1)	0,083	-0,015
DEx(1,1)	0,0350	-0,097

Pokračujeme **znaménkovým testem**. Víme, že statistika znaménkového testu má za platnosti nulové hypotézy binomické rozdělení. Vygenerujeme tedy náhodný výběr z binomického rozdělení s parametry $(n, \frac{1}{2})$, kde $n = n_1 = 20$. Daný výběr uložíme do vektoru V. Celý krok opět opakujeme 500-krát. Jako kritické hodnoty zvolíme taková čísla $a_n(\alpha)$ a $b_n(\alpha)$, pro která platí

$$\frac{\text{počet } V[i] \leq a_n(\alpha)}{500} = \frac{\alpha}{2}, \qquad i = 1, \ldots, 500,$$

a

$$\frac{\text{počet } V[i] \leq b_n(\alpha)}{500} = 1 - \frac{\alpha}{2}, \qquad i = 1, \ldots, 500,$$

tj. kritické hodnoty jsou $500 \cdot (\frac{\alpha}{2})$-tá a $500 \cdot (1 - \frac{\alpha}{2})$-tá nejmenší složka vektoru V. Stejný postup zopakujeme pro $n = n_2 = 60$. Výsledky jsou uvedeny v tabulce 5.7.

Tabulka 5.7: Simulované kritické hodnoty pro znaménkový test

Výběr	n_1	n_2
$\mathrm{Bi}(n_i, \frac{1}{2})$	5; 15	22; 38

Tabulka 5.8: Rozdíl teoretické a simulované kritické hodnoty pro znaménkový test

Výběr	n_1	n_2
$\mathrm{Bi}(n_i, \frac{1}{2})$	0; 0	-1; 1

Nyní budeme simulovat kritickou hodnotu pro **jednovýběrový Wilcoxonův test**. Vygenerujeme náhodný výběr z rovnoměrného rozdělení na $(-1, 1)$ o rozsahu $n_1 = 20$ a $n_2 = 60$, spočteme statistiku jednovýběrového Wilcoxonova testu (3.4), kterou uložíme do vektoru V. Tento postup opakujeme 500-krát. Jako kritickou hodnotu zvolíme takové číslo x_α s vlastností

$$\frac{\text{počet } V[i] \leq x_\alpha}{500} = \alpha, \qquad i = 1, \ldots, 500,$$

tj. kritická hodnota je $500 \cdot \alpha$-tá nejmenší složka vektoru V. Výsledné kritické hodnoty jsou uvedeny v tabulce 5.9.

Tabulka 5.9: Simulované kritické hodnoty pro jednovýběrový Wilcoxonův test

Výběr	n_1	n_2
R(-1,1)	51	646

Tabulka 5.10: Rozdíl teoretické a simulované kritické hodnoty pro jednovýběrový Wilcoxonův test

Výběr	n_1	n_2
R(-1,1)	1	2

Na závěr provedeme simulaci kritické hodnoty pro **Kolmogorovův-Smirnovův test**. Vygenerujeme náhodný výběr z rovnoměrného rozdělení na $(-1, 1)$ o rozsahu $n_1 = 20$ a $n_2 = 60$, spočteme statistiku Kolmogorova-Smirnovova testu (3.6), kterou uložíme do vektoru V. Opět celý postup opakujeme 500-krát. Jako kritickou hodnotu zvolíme takové číslo x_α, pro jež platí

$$\frac{\text{počet } V[i] \leq x_\alpha}{500} = 1 - \alpha, \qquad i = 1, \ldots, 500,$$

tj. kritická hodnota je $500 \cdot (1 - \alpha)$-tá nejmenší složka vektoru V. Výsledné kritické hodnoty jsou uvedeny v tabulce 5.11.

Tabulka 5.11: Simulované kritické hodnoty pro Kolmogorovův-Smirnovův test

Výběr	n_1	n_2
R(-1,1)	0,2975	0,2020

Tabulka 5.12: Rozdíl teoretické a simulované kritické hodnoty pro Kolmogorovův-Smirnovův test

Výběr	n_1	n_2
R(-1,1)	-0,00342	-0,02969

5.4 Síla testů

Nyní budeme simulovat sílu výše popsaných testů. Budeme uvažovat tyto náhodné výběry o rozsahu $n_1 = 20$ a $n_2 = 60$:

1. Náhodný výběr z Laplaceova rozdělení s parametry $(a, 1)$.
2. Náhodný výběr ze Studentova rozdělení o 4 stupních volnosti s posunem o a.

Postup pro spočtení síly testu:

1. Spočteme testovou statistiku daného testu.
2. Rozhodneme, zda zamítneme či nezamítneme nulovou hypotézu.
3. Krok 1 a 2 opakujeme N-krát.

Sílu testu určíme jako podíl

$$\frac{\text{počet případů, kdy zamítneme } H_0}{\text{N}}.$$

Graf 5.1 reprezentuje výsledky simulace síly testů pro Laplaceovo rozdělení o rozsahu n_1, jejíž hodnoty jsou uvedeny v tabulce 5.13. Obdobně je tomu pro rozsah n_2, kde hodnoty simulace pro sílu testů, uložené v tabulce 5.14, jsou vykresleny v grafu 5.2.

Graf 5.3 reprezentuje výsledky simulace síly testů pro Studentovo rozdělení s posunem o a o rozsahu n_1. Výsledky této simulace jsou uvedeny v tabulce 5.15. Hodnoty simulace pro sílu testů pro rozsah n_2, uložené v tabulce 5.16, jsou vykresleny v grafu 5.4.

Námi simulované hodnoty ukazují, že Kolmogorovův-Smirnovův test se zdá být nejsilnější z vybraných testů. Mezi silné testy dále patří Wilcoxonův test. Podíváme-li se na výběr z Laplaceova rozdělení, jako nejslabší test uvidíme t-test, ale u Studentova rozdělení je to znaménkový test.

Tabulka 5.13: Síla testů pro Laplaceovo rozdělení s parametry (a,1) o rozsahu n_1

Test	a=0,2	a=0,4	a=0,6	a=0,8	a=1	a=1,2
t-test	0,124	0,234	0,484	0,670	0,828	0,932
z-test	0,142	0,274	0,530	0,712	0,860	0,944
znaménkový test	0,130	0,272	0,508	0,686	0,866	0,924
Wilcoxonův test	0,138	0,276	0,546	0,734	0,878	0,946
K-S test	0,126	0,326	0,656	0,834	0,958	0,982

Tabulka 5.14: Síla testů pro Laplaceovo rozdělení s parametry (a,1) o rozsahu n_2

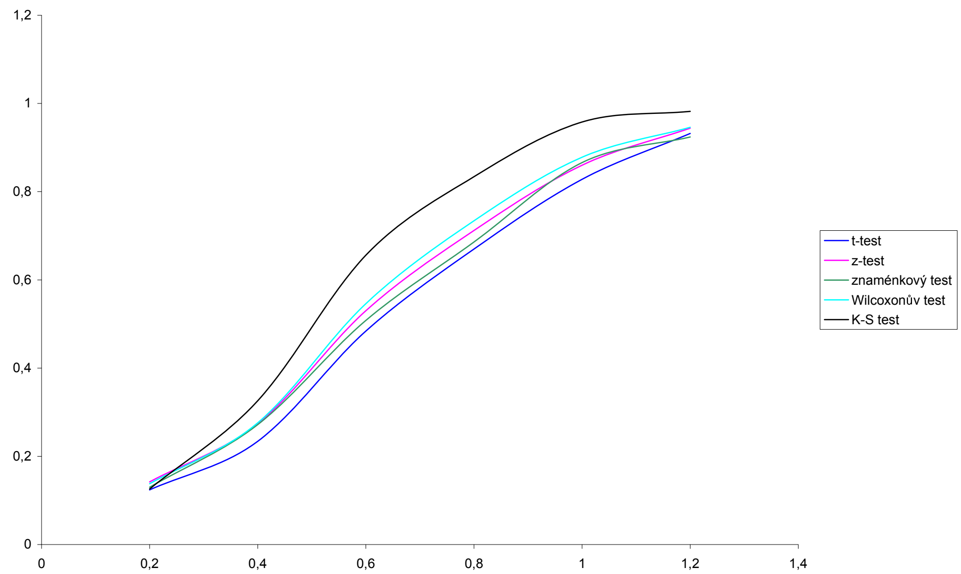
Test	a=0,2	a=0,4	a=0,6	a=0,8	a=1
t-test	0,188	0,588	0,902	0,994	1
z-test	0,200	0,606	0,906	0,994	1
znaménkový test	0,206	0,666	0,928	0,992	0,998
Wilcoxonův test	0,252	0,726	0,958	0,998	1
K-S test	0,286	0,766	0,976	1	1

Tabulka 5.15: Síla testů pro Studentovo rozdělení o 4 stupních volnosti s posunem o a o rozsahu n_1

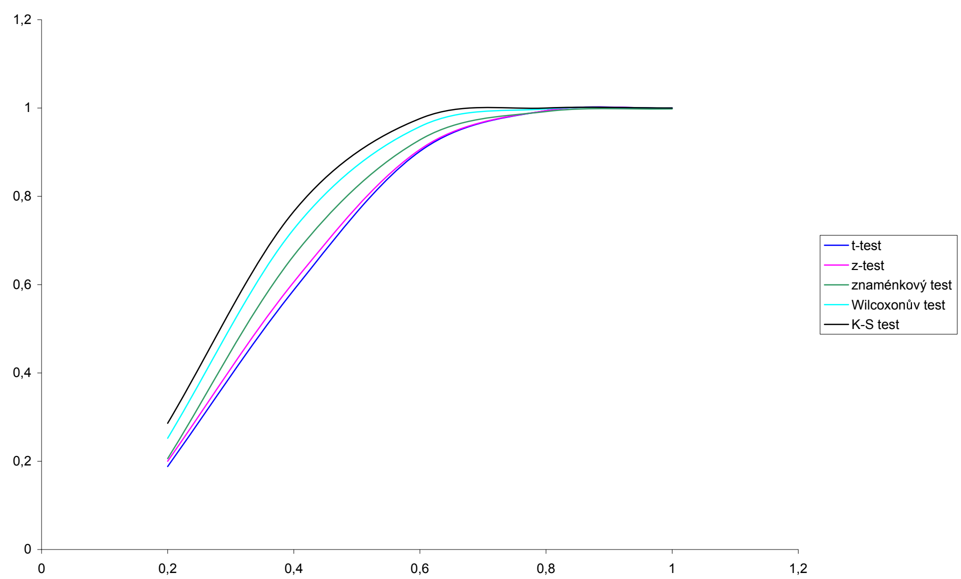
Test	a=0,2	a=0,4	a=0,6	a=0,8	a=1	a=1,2
t-test	0,080	0,248	0,490	0,694	0,874	0,944
z-test	0,102	0,278	0,534	0,746	0,886	0,958
znaménkový test	0,066	0,188	0,413	0,624	0,812	0,922
Wilcoxonův test	0,100	0,254	0,522	0,750	0,900	0,972
K-S test	0,104	0,274	0,574	0,812	0,956	0,992

Tabulka 5.16: Síla testů pro Studentovo rozdělení o 4 stupních volnosti s posunem o a o rozsahu n_2

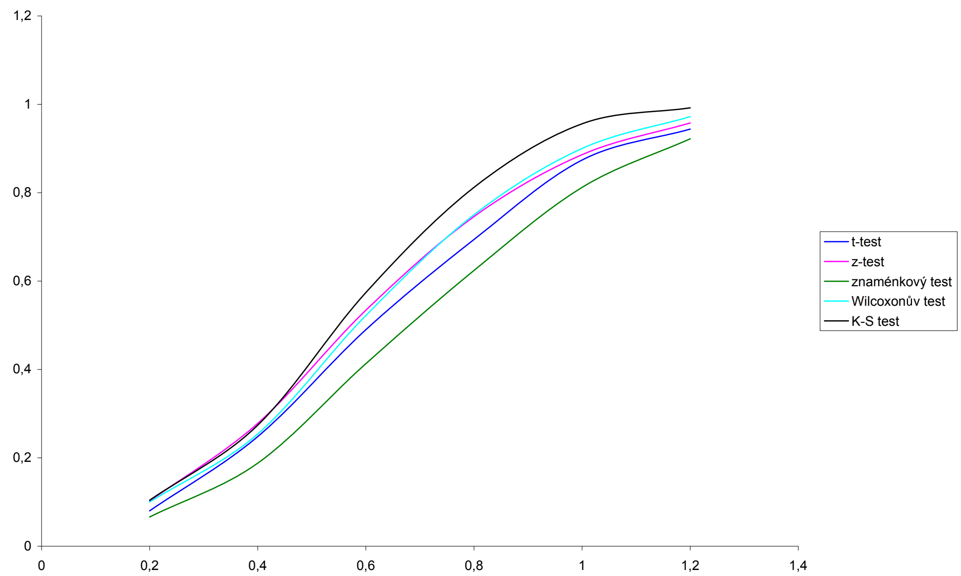
Test	a=0,2	a=0,4	a=0,6	a=0,8	a=1
t-test	0,190	0,598	0,904	0,998	1
z-test	0,210	0,606	0,910	0,984	1
znaménkový test	0,148	0,538	0,858	0,982	0,996
Wilcoxonův test	0,230	0,714	0,956	0,996	1
K-S test	0,270	0,762	0,956	1	1



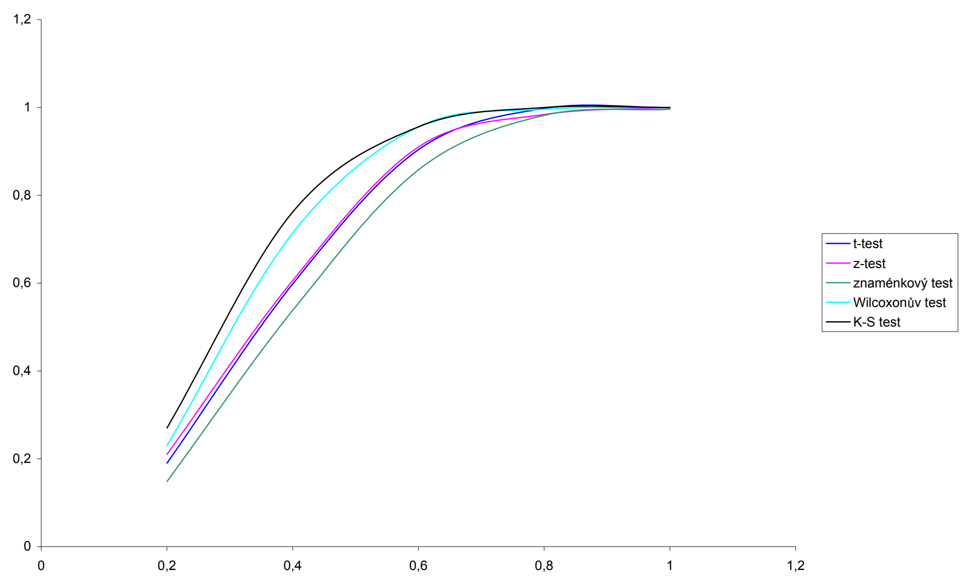
Obrázek 5.1: Síla testů pro Laplaceovo rozdělení s parametry (a,1) o rozsahu n_1



Obrázek 5.2: Síla testů pro Laplaceovo rozdělení s parametry (a,1) o rozsahu n_2



Obrázek 5.3: Síla testů pro Studentovo rozdělení o 4 stupních volnosti s posunem o a o rozsahu n_1



Obrázek 5.4: Síla testů pro Studentovo rozdělení o 4 stupních volnosti s posunem o a o rozsahu n_2

Závěr

V této práci jsme se zabývali jednovýběrovým problémem, některými jednovýběrovými testy a aproximací kritických hodnot. Na začátku jsme představili několik jednovýběrových testů, jejich testovou statistiku a někdy i kritickou hodnotu. Jednotlivé testy jsme aplikovali na tentýž příklad s košťaty. Později jsme ukázali, jak se dá kritická hodnota aproximovat a provedli jsme simulační studii, na které jsme aproximaci ilustrovali. Pro aproximaci kritických hodnot jsme použili bootstrapovou metodu a simulaci z různých rozdělení. Získané výsledky jsme uvedli v tabulkách. Závěr simulační studie byl věnován síle jednotlivých testů, jejíž výsledky jsme uvedli jak v tabulkách, tak v grafech.

Simulační studii jsme provedli ve statistickém programu R. Zdrojové kódy lze nalézt v příloze elektronické verze této práce.

Literatura

[1] Jiří Anděl (2005): *Základy matematické statistiky*, Matfyzpress.

[2] C. C. Butler (1969): *A test for symmetry using the sample distribution function*, Ann. Math. Statist. 40, 2209-2210.

[3] A. C. Davison, D. V. Hinklex (1997): *Bootstrap Methods and Their Application*, Cambridge University Press.

[4] Václav Dupač, Marie Hušková (2001): *Pravděpodobnost a matematická statistika*, Karolinum.

[5] Jaroslav Hájek, Zbyněk Šidák (1967): *Theory of Rank Tests*, Academia.

[6] Jana Jurečková (1981): *Pořadové testy*, Univerzita Karlova v Praze.

[7] P. R. Krishnaiah, P. K. Sen (1984): *Handbook of Statistics*, North Holland.

[8] Jiří Likeš, Josef Laga (1978): *Základní statistické tabulky*, SNTL - Nakladatelství technické literatury.

[9] Zuzana Prášková (2004): *METODA BOOTSTRAP*, ROBUST [online], [cit. 2009-07-09], s. 1-16. Dostupný z WWW: `<http://www.statspol.cz/robust/robust2004/praskova.pdf>`.

Příloha A

Tabulky kritických hodnot a kvantilů

V této části práce uvedeme jen některé kvantily a kritické hodnoty pro jednotlivé testy. Tabulky jsou převzaty z knihy [8].

Tabulka A.1: Kritické hodnoty pro jednovýběrový Wilcoxonův test

n	$c_n(0,05)$	$c_n(0,01)$	n	$c_n(0,05)$	$c_n(0,01)$
6	0	-	20	52	37
7	2	-	21	58	42
8	3	0	22	65	48
9	5	1	23	73	54
10	8	3	24	81	61
11	10	5	25	89	68
12	13	7	26	98	75
13	17	9	27	107	83
14	21	12	28	116	91
15	25	15	29	126	100
16	29	19	30	137	109
17	34	23	40	264	220
18	40	27	50	434	373
19	46	32	60	648	567

Tabulka A.2: Kritické hodnoty pro znaménkový test

n	$a_n(0,05)$	$b_n(0,05)$	$a_n(0,01)$	$b_n(0,01)$
3	0	3	0	3
4	0	4	0	4
5	0	5	0	5
6	0	6	0	6
7	0	7	0	7
8	0	8	0	8
9	1	8	0	9
10	1	9	0	10
11	1	10	0	11
12	2	10	1	11
13	2	11	1	12
14	2	12	1	13
15	3	12	2	13
16	3	13	2	14
17	4	13	2	15
18	4	14	3	15
19	4	15	3	16
20	5	15	3	17
21	5	16	4	17
22	5	17	4	18
23	6	17	4	19
24	6	18	5	19
25	7	18	5	20
26	7	19	6	20
27	7	20	6	21
28	8	20	6	22
29	8	21	7	22
30	9	21	7	23

Tabulka A.3: Kvantily rozdělení N(0,1)

α	0,5	0,90	0,95	0,975	0,99	0,995	0,999
u_α	0	1,282	1,645	1,960	2,326	2,576	3,090

Tabulka A.4: Kvantily pro Studentovo rozdělení $t_n(1-\frac{\alpha}{2})$

n	$t_n(0,875)$	$t_n(0,95)$	$t_n(0,975)$	$t_n(0,9875)$	$t_n(0,995)$	$t_n(0,9975)$
1	2,4142	6,3138	12,706	25,452	63,657	127,32
2	1,6036	2,9200	4,3027	6,2053	9,9248	14,089
3	1,4226	2,3534	3,1825	4,1765	5,8409	7,4533
4	1,3444	2,1318	2,7764	3,4954	4,6041	5,5976
5	1,3009	2,0150	2,5706	3,1634	4,0321	4,7733
6	1,2733	1,9432	2,4469	2,9687	3,7074	4,3168
7	1,2543	1,8946	2,3646	2,8412	3,4995	4,0293
8	1,2403	1,8595	2,3060	2,7515	3,3554	3,8325
9	1,2297	1,8331	2,2622	2,6850	3,2498	3,6897
10	1,2213	1,8125	2,2281	2,6338	3,1693	3,5814
11	1,2145	1,7959	2,2010	2,5931	3,1058	3,4966
12	1,2089	1,7823	2,1788	2,5600	3,0545	3,4284
13	1,2041	1,7709	2,1604	2,5326	3,0123	3,3725
14	1,2001	1,7613	2,1448	2,5096	2,9768	3,3257
15	1,1967	1,7530	2,1315	2,4899	2,9467	3,2860
16	1,1937	1,7459	2,1199	2,4729	2,9208	3,2520
17	1,1910	1,7396	2,1098	2,4581	2,8982	3,2225
18	1,1887	1,7341	2,1009	2,4450	2,8784	3,1966
19	1,1866	1,7291	2,0930	2,4334	2,8609	3,1737
20	1,1848	1,7247	2,0860	2,4231	2,8453	3,1534
21	1,1831	1,7207	2,0796	2,4138	2,8314	3,1352
22	1,1816	1,7171	2,0739	2,4055	2,8188	3,1188
23	1,1802	1,7139	2,0687	2,3979	2,8073	3,1040
24	1,1789	1,7109	2,0639	2,3910	2,7969	3,0905
25	1,1777	1,7081	2,0595	2,3846	2,7874	3,0782
26	1,1766	1,7056	2,0555	2,3788	2,7787	3,0669
27	1,1757	1,7033	2,0518	2,3734	2,7707	3,0565
28	1,1748	1,7011	2,0484	2,3685	2,7633	3,0469
29	1,1739	1,6991	2,0452	2,3638	2,7564	3,0380
30	1,1731	1,6973	2,0423	2,3596	2,7500	3,0298
40	1,1673	1,6839	2,0211	2,3289	2,7045	2,9712
60	1,1616	1,6707	2,0003	2,2991	2,6603	2,9146
120	1,1559	1,6577	1,9799	2,2699	2,6174	2,8599
∞	1,1503	1,6449	1,9600	2,2414	2,5758	2,8070

Tabulka A.5: Kvantily statistiky Kolmogorovova-Smirnovova testu

n	$\alpha = 0{,}01$	$\alpha = 0{,}02$	$\alpha = 0{,}05$	$\alpha = 0{,}1$	$\alpha = 0{,}2$
1	0,99500	0,99000	0,97500	0,95000	0,90000
2	0,92929	0,90000	0,84189	0,77639	0,68377
3	0,82900	0,78456	0,70760	0,63604	0,56481
4	0,73424	0,68887	0,62394	0,56522	0,49265
5	0,66853	0,62718	0,56328	0,50945	0,44698
6	0,61661	0,57741	0,51926	0,46799	0,41037
7	0,57581	0,53844	0,48342	0,43607	0,38148
8	0,54179	0,50654	0,45427	0,40962	0,35831
9	0,51332	0,47960	0,43001	0,38746	0,33910
10	0,48893	0,45662	0,40925	0,36866	0,32260
11	0,46770	0,43670	0,39122	0,35242	0,30829
12	0,44905	0,41918	0,37543	0,33815	0,29577
13	0,43247	0,40362	0,36143	0,32549	0,28470
14	0,41762	0,38970	0,34890	0,31417	0,27481
15	0,40420	0,37713	0,33760	0,30397	0,26588
16	0,39201	0,36571	0,32733	0,29472	0,25778
17	0,38086	0,35528	0,31796	0,28627	0,25039
18	0,37062	0,34569	0,30936	0,27851	0,24360
19	0,36117	0,33685	0,30143	0,27136	0,23735
20	0,35241	0,32866	0,29408	0,26473	0,23156
30	0,28986	0,27023	0,24170	0,21756	0,19032
40	0,25205	0,23494	0,21012	0,18913	0,16547
50	0,22604	0,21068	0,18841	0,16959	0,14840
60	0,20673	0,19267	0,17231	0,15511	0,13573
70	0,19167	0,17863	0,15975	0,14381	0,12586

Příloha B

Zdrojové kódy

Bootstrap pro z-test s rozsahem$= 20; 60$**:**

```
N=500
V=rep(NA,N)
n=rozsah
x=rnorm(n,0,1)
for (i in 1:N){
xboot=sample(x,replace=T)
V[i]=t.test(xboot, mu=mean(x), alternative ="two.sided",
            conf.level =0.95)$statistic
}
K=sort(abs(V))
j=N*(1-0.05)
K[j]
```

Simulace pro t-test se střední hodnotou EX$=0; \frac{1}{2}; 1$
a rozsahem$= 20; 60$**:**

```
N=500
V=rep(NA,N)
n=rozsah
mi=EX
for (i in 1:N){
x=rlaplace(n,mi,1)
```

```
y=x-mi
V[i]=t.test(y, mu=0, alternative ="two.sided",
            conf.level =0.95)$statistic
}
K=sort(V)
j=N*(1-0.05)
K[j]
```

Simulace pro znaménkový test pro rozsah$= 20; 60$**:**

```
N=500
n=rozsah
V=rbinom(N,n,0.5)
K=sort(V)
j=N*(1-0.05)
K[j]
```

Simulace pro Wilcoxonův test pro rozsah$= 20; 60$**:**

```
N=500
V=rep(NA,N)
n=rozsah
for (i in 1:N){
x=runif(n,-1,1)
Wklad=wilcox.test(x, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor=n*(n+1)/2 - Wklad
V[i]=min(Wklad,Wzapor)
}
K=sort(V)
j=N*(1-0.05)
K[j]
```

Simulace pro Kolmogorovův-Smirnovův test s rozsahem$= 20; 60$**:**

```
N=500
V=rep(NA,N)
n=rozsah
```

```
for (i in 1:N){
x=runif(n,-1,1)
V[i]=ks.test(x,y="pnorm",mean=0,
             sd=1)$statistic
}
K=sort(V)
j=N*(1-0.05)
K[j]
```

Síla testů pro Studentovo rozdělení o 4 stupních volnosti o rozsahu $n_1 = 20$:

```
N=500
n=20
P1KS=0
P2KS=0
P3KS=0
P4KS=0
P5KS=0
P6KS=0
P1W=0
P2W=0
P3W=0
P4W=0
P5W=0
P6W=0
P1S=0
P2S=0
P3S=0
P4S=0
P5S=0
P6S=0
P1T=0
P2T=0
P3T=0
P4T=0
P5T=0
P6T=0
```

```
P1Z=0
P2Z=0
P3Z=0
P4Z=0
P5Z=0
P6Z=0
k1=5
k2=15
kriticT=2.093
kriticZ=1.96
kriticW=52
kriticKS=0.29408
for (i in 1:N){
x=rt(n,4,0)
x1=x+0.2
x2=x+0.4
x3=x+0.6
x4=x+0.8
x5=x+1
x6=x+1.2
Tn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn1) >= kriticT) P1T=P1T+1
Tn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn2) >= kriticT) P2T=P2T+1
Tn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn3) >= kriticT) P3T=P3T+1
Tn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn4) >= kriticT) P4T=P4T+1
Tn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn5) >= kriticT) P5T=P5T+1
Tn6=t.test(x6, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn6) >= kriticT) P6T=P6T+1
```

```
Zn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn1) >= kriticZ) P1Z=P1Z+1
Zn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn2) >= kriticZ) P2Z=P2Z+1
Zn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn3) >= kriticZ) P3Z=P3Z+1
Zn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn4) >= kriticZ) P4Z=P4Z+1
Zn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn5) >= kriticZ) P5Z=P5Z+1
Zn6=t.test(x6, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn6) >= kriticZ) P6Z=P6Z+1

Yn1=sign.test(x1,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn1 >= k2) P1S=P1S+1
if (Yn1 <= k1) P1S=P1S+1
Yn2=sign.test(x2,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn2 >= k2) P2S=P2S+1
if (Yn2 <= k1) P2S=P2S+1
Yn3=sign.test(x3,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn3 >= k2) P3S=P3S+1
if (Yn3 <= k1) P3S=P3S+1
Yn4=sign.test(x4,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn4 >= k2) P4S=P4S+1
if (Yn4 <= k1) P4S=P4S+1
Yn5=sign.test(x5,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
```

```
if (Yn5 >= k2) P5S=P5S+1
if (Yn5 <= k1) P5S=P5S+1
Yn6=sign.test(x6,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn6 >= k2) P6S=P6S+1
if (Yn6 <= k1) P6S=P6S+1

Wklad1=wilcox.test(x1, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor1=n*(n+1)/2 - Wklad1
if (min(Wklad1,Wzapor1) <= kriticW) P1W=P1W+1
Wklad2=wilcox.test(x2, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor2=n*(n+1)/2 - Wklad2
if (min(Wklad2,Wzapor2) <= kriticW) P2W=P2W+1
Wklad3=wilcox.test(x3, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor3=n*(n+1)/2 - Wklad3
if (min(Wklad3,Wzapor3) <= kriticW) P3W=P3W+1
Wklad4=wilcox.test(x4, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor4=n*(n+1)/2 - Wklad4
if (min(Wklad4,Wzapor4) <= kriticW) P4W=P4W+1
Wklad5=wilcox.test(x5, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor5=n*(n+1)/2 - Wklad5
if (min(Wklad5,Wzapor5) <= kriticW) P5W=P5W+1
Wklad6=wilcox.test(x6, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor6=n*(n+1)/2 - Wklad6
if (min(Wklad6,Wzapor6) <= kriticW) P6W=P6W+1

KS1=ks.test(x1,y="pnorm",mean=0,
           sd=1)$statistic
if (KS1 >= kriticKS) P1KS=P1KS+1
KS2=ks.test(x2,y="pnorm",mean=0,
           sd=1)$statistic
if (KS2 >= kriticKS) P2KS=P2KS+1
```

```
KS3=ks.test(x3,y="pnorm",mean=0,
           sd=1)$statistic
if (KS3 >= kriticKS) P3KS=P3KS+1
KS4=ks.test(x4,y="pnorm",mean=0,
           sd=1)$statistic
if (KS4 >= kriticKS) P4KS=P4KS+1
KS5=ks.test(x5,y="pnorm",mean=0,
           sd=1)$statistic
if (KS5 >= kriticKS) P5KS=P5KS+1
KS6=ks.test(x6,y="pnorm",mean=0,
           sd=1)$statistic
if (KS6 >= kriticKS) P6KS=P6KS+1
}
P1T/N
P2T/N
P3T/N
P4T/N
P5T/N
P6T/N

P1Z/N
P2Z/N
P3Z/N
P4Z/N
P5Z/N
P6Z/N

P1S/N
P2S/N
P3S/N
P4S/N
P5S/N
P6S/N

P1W/N
P2W/N
P3W/N
P4W/N
```

```
P5W/N
P6W/N

P1KS/N
P2KS/N
P3KS/N
P4KS/N
P5KS/N
P6KS/N
```

Síla testů pro Studentovo rozdělení o 4 stupních volnosti o rozsahu $n_2 = 60$:

```
N=500
n=60
P1KS=0
P2KS=0
P3KS=0
P4KS=0
P5KS=0
P1W=0
P2W=0
P3W=0
P4W=0
P5W=0
P1S=0
P2S=0
P3S=0
P4S=0
P5S=0
P1T=0
P2T=0
P3T=0
P4T=0
P5T=0
P1Z=0
P2Z=0
P3Z=0
```

```
P4Z=0
P5Z=0
k1=21
k2=39
kriticT=2.001
kriticZ=1.96
kriticW=648
kriticKS=0.17231
for (i in 1:N){
x=rt(n,4,0)
x1=x+0.2
x2=x+0.4
x3=x+0.6
x4=x+0.8
x5=x+1
Tn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn1) >= kriticT) P1T=P1T+1
Tn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn2) >= kriticT) P2T=P2T+1
Tn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn3) >= kriticT) P3T=P3T+1
Tn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn4) >= kriticT) P4T=P4T+1
Tn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn5) >= kriticT) P5T=P5T+1

Zn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn1) >= kriticZ) P1Z=P1Z+1
Zn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn2) >= kriticZ) P2Z=P2Z+1
Zn3=t.test(x3, mu=0, alternative ="two.sided",
```

```
          conf.level =0.95)$statistic
if (abs(Zn3) >= kriticZ) P3Z=P3Z+1
Zn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn4) >= kriticZ) P4Z=P4Z+1
Zn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn5) >= kriticZ) P5Z=P5Z+1

Yn1=sign.test(x1,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn1 >= k2) P1S=P1S+1
if (Yn1 <= k1) P1S=P1S+1
Yn2=sign.test(x2,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn2 >= k2) P2S=P2S+1
if (Yn2 <= k1) P2S=P2S+1
Yn3=sign.test(x3,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn3 >= k2) P3S=P3S+1
if (Yn3 <= k1) P3S=P3S+1
Yn4=sign.test(x4,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn4 >= k2) P4S=P4S+1
if (Yn4 <= k1) P4S=P4S+1
Yn5=sign.test(x5,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn5 >= k2) P5S=P5S+1
if (Yn5 <= k1) P5S=P5S+1

Wklad1=wilcox.test(x1, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor1=n*(n+1)/2 - Wklad1
if (min(Wklad1,Wzapor1) <= kriticW) P1W=P1W+1
Wklad2=wilcox.test(x2, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor2=n*(n+1)/2 - Wklad2
if (min(Wklad2,Wzapor2) <= kriticW) P2W=P2W+1
```

```
Wklad3=wilcox.test(x3, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor3=n*(n+1)/2 - Wklad3
if (min(Wklad3,Wzapor3) <= kriticW) P3W=P3W+1
Wklad4=wilcox.test(x4, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor4=n*(n+1)/2 - Wklad4
if (min(Wklad4,Wzapor4) <= kriticW) P4W=P4W+1
Wklad5=wilcox.test(x5, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor5=n*(n+1)/2 - Wklad5
if (min(Wklad4,Wzapor4) <= kriticW) P5W=P5W+1

KS1=ks.test(x1,y="pnorm",mean=0,
           sd=1)$statistic
if (KS1 >= kritic) P1KS=P1KS+1
KS2=ks.test(x2,y="pnorm",mean=0,
           sd=1)$statistic
if (KS2 >= kritic) P2KS=P2KS+1
KS3=ks.test(x3,y="pnorm",mean=0,
           sd=1)$statistic
if (KS3 >= kritic) P3KS=P3KS+1
KS4=ks.test(x4,y="pnorm",mean=0,
           sd=1)$statistic
if (KS4 >= kritic) P4KS=P4KS+1
KS5=ks.test(x5,y="pnorm",mean=0,
           sd=1)$statistic
if (KS5 >= kritic) P5KS=P5KS+1
}
P1T/N
P2T/N
P3T/N
P4T/N
P5T/N

P1Z/N
P2Z/N
P3Z/N
```

```
P4Z/N
P5Z/N

P1S/N
P2S/N
P3S/N
P4S/N
P5S/N

P1W/N
P2W/N
P3W/N
P4W/N
P5W/N

P1KS/N
P2KS/N
P3KS/N
P4KS/N
P5KS/N
```

Síla testů pro Laplaceovo rozdělení o rozsahu $n_1 = 20$:

```
N=500
n=20
P1KS=0
P2KS=0
P3KS=0
P4KS=0
P5KS=0
P6KS=0
P1W=0
P2W=0
P3W=0
P4W=0
P5W=0
P6W=0
P1S=0
```

```
P2S=0
P3S=0
P4S=0
P5S=0
P6S=0
P1T=0
P2T=0
P3T=0
P4T=0
P5T=0
P6T=0
P1Z=0
P2Z=0
P3Z=0
P4Z=0
P5Z=0
P6Z=0
k1=5
k2=15
kriticT=2.093
kriticZ=1.96
kriticW=52
kriticKS=0.29408
for (i in 1:N){
x1=rlaplace(n,0.2,1)
x2=rlaplace(n,0.4,1)
x3=rlaplace(n,0.6,1)
x4=rlaplace(n,0.8,1)
x5=rlaplace(n,1,1)
x6=rlaplace(n,1.2,1)
Tn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn1) >= kriticT) P1T=P1T+1
Tn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn2) >= kriticT) P2T=P2T+1
Tn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
```

```
if (abs(Tn3) >= kriticT) P3T=P3T+1
Tn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn4) >= kriticT) P4T=P4T+1
Tn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn5) >= kriticT) P5T=P5T+1
Tn6=t.test(x6, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn6) >= kriticT) P6T=P6T+1

Zn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn1) >= kriticZ) P1Z=P1Z+1
Zn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn2) >= kriticZ) P2Z=P2Z+1
Zn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn3) >= kriticZ) P3Z=P3Z+1
Zn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn4) >= kriticZ) P4Z=P4Z+1
Zn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn5) >= kriticZ) P5Z=P5Z+1
Zn6=t.test(x6, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn6) >= kriticZ) P6Z=P6Z+1

Yn1=sign.test(x1,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn1 >= k2) P1S=P1S+1
if (Yn1 <= k1) P1S=P1S+1
Yn2=sign.test(x2,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn2 >= k2) P2S=P2S+1
if (Yn2 <= k1) P2S=P2S+1
```

```
Yn3=sign.test(x3,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn3 >= k2) P3S=P3S+1
if (Yn3 <= k1) P3S=P3S+1
Yn4=sign.test(x4,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn4 >= k2) P4S=P4S+1
if (Yn4 <= k1) P4S=P4S+1
Yn5=sign.test(x5,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn5 >= k2) P5S=P5S+1
if (Yn5 <= k1) P5S=P5S+1
Yn6=sign.test(x6,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn6 >= k2) P6S=P6S+1
if (Yn6 <= k1) P6S=P6S+1

Wklad1=wilcox.test(x1, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor1=n*(n+1)/2 - Wklad1
if (min(Wklad1,Wzapor1) <= kriticW) P1W=P1W+1
Wklad2=wilcox.test(x2, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor2=n*(n+1)/2 - Wklad2
if (min(Wklad2,Wzapor2) <= kriticW) P2W=P2W+1
Wklad3=wilcox.test(x3, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor3=n*(n+1)/2 - Wklad3
if (min(Wklad3,Wzapor3) <= kriticW) P3W=P3W+1
Wklad4=wilcox.test(x4, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor4=n*(n+1)/2 - Wklad4
if (min(Wklad4,Wzapor4) <= kriticW) P4W=P4W+1
Wklad5=wilcox.test(x5, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor5=n*(n+1)/2 - Wklad5
if (min(Wklad5,Wzapor5) <= kriticW) P5W=P5W+1
Wklad6=wilcox.test(x6, mu=0, alternative ="two.sided",
```

```
                  conf.level =0.95)$statistic
Wzapor6=n*(n+1)/2 - Wklad6
if (min(Wklad6,Wzapor6) <= kriticW) P6W=P6W+1

KS1=ks.test(x1,y="pnorm",mean=0,
           sd=1)$statistic
if (KS1 >= kriticKS) P1KS=P1KS+1
KS2=ks.test(x2,y="pnorm",mean=0,
           sd=1)$statistic
if (KS2 >= kriticKS) P2KS=P2KS+1
KS3=ks.test(x3,y="pnorm",mean=0,
           sd=1)$statistic
if (KS3 >= kriticKS) P3KS=P3KS+1
KS4=ks.test(x4,y="pnorm",mean=0,
           sd=1)$statistic
if (KS4 >= kriticKS) P4KS=P4KS+1
KS5=ks.test(x5,y="pnorm",mean=0,
           sd=1)$statistic
if (KS5 >= kriticKS) P5KS=P5KS+1
KS6=ks.test(x6,y="pnorm",mean=0,
           sd=1)$statistic
if (KS6 >= kriticKS) P6KS=P6KS+1
}
P1T/N
P2T/N
P3T/N
P4T/N
P5T/N
P6T/N

P1Z/N
P2Z/N
P3Z/N
P4Z/N
P5Z/N
P6Z/N

P1S/N
```

```
P2S/N
P3S/N
P4S/N
P5S/N
P6S/N

P1W/N
P2W/N
P3W/N
P4W/N
P5W/N
P6W/N

P1KS/N
P2KS/N
P3KS/N
P4KS/N
P5KS/N
P6KS/N
```

Síla testů pro Laplaceovo rozdělení o rozsahu $n_2 = 60$:

```
N=500
n=60
P1KS=0
P2KS=0
P3KS=0
P4KS=0
P5KS=0
P1W=0
P2W=0
P3W=0
P4W=0
P5W=0
P1S=0
P2S=0
P3S=0
P4S=0
```

```
P5S=0
P1T=0
P2T=0
P3T=0
P4T=0
P5T=0
P1Z=0
P2Z=0
P3Z=0
P4Z=0
P5Z=0
k1=21
k2=39
kriticT=2.001
kriticZ=1.96
kriticW=648
kriticKS=0.17231
for (i in 1:N){
x1=rlaplace(n,0.2,1)
x2=rlaplace(n,0.4,1)
x3=rlaplace(n,0.6,1)
x4=rlaplace(n,0.8,1)
x5=rlaplace(n,1,1)
Tn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn1) >= kriticT) P1T=P1T+1
Tn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn2) >= kriticT) P2T=P2T+1
Tn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn3) >= kriticT) P3T=P3T+1
Tn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn4) >= kriticT) P4T=P4T+1
Tn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Tn5) >= kriticT) P5T=P5T+1
```

```
Zn1=t.test(x1, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn1) >= kriticZ) P1Z=P1Z+1
Zn2=t.test(x2, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn2) >= kriticZ) P2Z=P2Z+1
Zn3=t.test(x3, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn3) >= kriticZ) P3Z=P3Z+1
Zn4=t.test(x4, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn4) >= kriticZ) P4Z=P4Z+1
Zn5=t.test(x5, mu=0, alternative ="two.sided",
          conf.level =0.95)$statistic
if (abs(Zn5) >= kriticZ) P5Z=P5Z+1

Yn1=sign.test(x1,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn1 >= k2) P1S=P1S+1
if (Yn1 <= k1) P1S=P1S+1
Yn2=sign.test(x2,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn2 >= k2) P2S=P2S+1
if (Yn2 <= k1) P2S=P2S+1
Yn3=sign.test(x3,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn3 >= k2) P3S=P3S+1
if (Yn3 <= k1) P3S=P3S+1
Yn4=sign.test(x4,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn4 >= k2) P4S=P4S+1
if (Yn4 <= k1) P4S=P4S+1
Yn5=sign.test(x5,y = NULL,md=0,alternative="two.sided",
             conf.level = 0.95)$rval$statistic
if (Yn5 >= k2) P5S=P5S+1
if (Yn5 <= k1) P5S=P5S+1
```

```
Wklad1=wilcox.test(x1, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor1=n*(n+1)/2 - Wklad1
if (min(Wklad1,Wzapor1) <= kriticW) P1W=P1W+1
Wklad2=wilcox.test(x2, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor2=n*(n+1)/2 - Wklad2
if (min(Wklad2,Wzapor2) <= kriticW) P2W=P2W+1
Wklad3=wilcox.test(x3, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor3=n*(n+1)/2 - Wklad3
if (min(Wklad3,Wzapor3) <= kriticW) P3W=P3W+1
Wklad4=wilcox.test(x4, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor4=n*(n+1)/2 - Wklad4
if (min(Wklad4,Wzapor4) <= kriticW) P4W=P4W+1
Wklad5=wilcox.test(x5, mu=0, alternative ="two.sided",
                  conf.level =0.95)$statistic
Wzapor5=n*(n+1)/2 - Wklad5
if (min(Wklad5,Wzapor5) <= kriticW) P5W=P5W+1

KS1=ks.test(x1,y="pnorm",mean=0,
           sd=1)$statistic
if (KS1 >= kriticKS) P1KS=P1KS+1
KS2=ks.test(x2,y="pnorm",mean=0,
           sd=1)$statistic
if (KS2 >= kriticKS) P2KS=P2KS+1
KS3=ks.test(x3,y="pnorm",mean=0,
           sd=1)$statistic
if (KS3 >= kriticKS) P3KS=P3KS+1
KS4=ks.test(x4,y="pnorm",mean=0,
           sd=1)$statistic
if (KS4 >= kriticKS) P4KS=P4KS+1
KS5=ks.test(x5,y="pnorm",mean=0,
           sd=1)$statistic
if (KS5 >= kriticKS) P5KS=P5KS+1
}
P1T/N
```

P2T/N
P3T/N
P4T/N
P5T/N

P1Z/N
P2Z/N
P3Z/N
P4Z/N
P5Z/N

P1S/N
P2S/N
P3S/N
P4S/N
P5S/N

P1W/N
P2W/N
P3W/N
P4W/N
P5W/N

P1KS/N
P2KS/N
P3KS/N
P4KS/N
P5KS/N