Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Kristýna Sionová

Porovnání dvouvýběrového T-testu a jeho neparametrického ekvivalentu

Katedra pravděpodobnosti a matematické statistiky

Mgr. Alena Koubková
Studijní program: Matematika
Studijní obor: Obecná matematika

2006

Děkuji vedoucí bakalářské práce Mgr. Aleně Koubkové za cenné připomínky a rady poskytnuté při jejím vypracování.

Dále děkuji Ing. Vilému Sklenákovi, CSc. za konzultace týkající se používaného software.

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

V Praze dne 22. 5. 2006 Kristýna Sionová

Obsah

Název práce: Porovnání dvouvýběrového T-testu a jeho neparametrického ekvivalentu
Autor: Kristýna Sionová
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: Mgr. Alena Koubková
e-mail vedoucího: koubkova@karlin.mff.cuni.cz

Tato práce se zabývá porovnáním dvouvýběrového t testu a dvouvýběrového Wilcoxonova testu. Úvodní kapitola pojednává o problému testování hypotéz obecně. Jsou zmíněna důležitá rozdělení, jejich hustoty a základní vlastnosti. Dále jsou popsány způsoby konstrukce obou testů a jsou zdůrazněny předpoklady, které při jejich používání klademe na data. V hlavní části práce jsou testy aplikovány na simulovaná data, k čemuž je používán software R 2.2.0. V závěrečné kapitole jsou shrnuty výsledky, k nimž jsme v práci došli. Zjistili jsme, že pro výběry z normálního rozdělení dosahuje lepších výsledků t test, pro Cauchyovo a dvojitě exponenciální rozdělení je lepší Wilcoxonův test. Úspěšnost obou testů roste se zvyšujícím se rozsahem porovnávaných výběrů.

Klíčová slova: testování hypotéz, dvouvýběrový t test, dvouvýběrový Wilcoxonův test

Title: Comparison of two-sample T-test and its nonparametric equivalent
Author: Kristýna Sionová
Department: Department of Probability and Mathematical Statistics
Supervisor: Mgr. Alena Koubková
Supervisor's e-mail address: koubkova@karlin.mff.cuni.cz

The aim of this work is a comparison the two-sample t test with the two-sample Wilcoxon rank-sum test. We discuss the problem of hypothesis testing in general in the introductory part of the work. The important probability distributions, their densities and basic properties are mentioned. The methods of creating both tests are described, as well as the assumptions required. In the main part, both tests are applied to simulated data. The software R 2.2.0 is used for this purpose. The final part is focused on the summary of results obtained. We found out that t test is better than Wilcoxon test for samples from normal distribution and Wilcoxon test is better for samples from Cauchy and Laplace distribution. Both tests have better results for large sample sizes than for small sample sizes.

Keywords: hypothesis testing, two-sample t-test, two-sample Wilcoxon rank-sum test

Úvod

Ve své práci jsem porovnávala dvouvýběrový t test a dvouvýběrový Wilcoxonův test. Cílem bylo zjistit, kdy je použití kterého testu vhodnější a to v závislosti na rozdělení porovnávaných náhodných výběrů, na jejich rozsahu a na rozdílu mezi středními hodnotami výběrů.

V následujících odstavcích je uveden stručný přehled obsahu jednotlivých částí bakalářské práce.

V první kapitole se zabýváme problémem testování hypotéz obecně.

Druhá kapitola přináší přehled nejdůležitějších rozdělení pravděpodobností, která jsou zapotřebí ke konstrukci testů.

Ve třetí a čtvrté kapitole popisujeme konstrukci dvouvýběrového t testu a dvouvýběrového Wilcoxonova testu.

V páté kapitole práce jsou pomocí slovních popisů, tabulek a grafů zachyceny výsledky testů na simulovaných náhodných výběrech provedených v programu R.

Domnívám se, že právě pátá kapitola mé práce je nejdůležitější, neboť obsahuje mé vlastní výpočty. Proto jsem se snažila v ostatních kapitolách o co nejstručnější vyjádření a uvádím pouze důkazy nejdůležitějších tvrzení a na důkazy tvrzení pomocných pouze odkazuji.

Kapitola 1

Testování hypotéz

Buď $X_1, \dots, X_n$ náhodný výběr z rozdělení, jež závisí na neznámém parametru θ. O θ víme, že patří do nějaké množiny Θ, kterou nazýváme *parametrický prostor*. (Prvky množiny Θ mohou být například čísla, vektory, funkce aj.)

Dále budeme potřebovat následující označení. Symbolem $\mathbf{X}$ označíme vektor $(X_1, \dots, X_n)'$. Symbolem $\mathsf{P}_\theta(A)$ značíme pravděpodobnost, že nastane jev A za podmínky, že θ je skutečná hodnota zkoumaného parametru.

Na základě úvah o dané situaci se domníváme, že parametr náleží do nějaké neprázdné podmnožiny Θ_0 prostoru Θ. Tvrzení $\theta \in \Theta_0$ označíme H_0 a nazveme *nulová hypotéza*. Situaci, která nastane v případě, kdy H_0 neplatí, tj. $\theta \in \Theta \setminus \Theta_0$, nazveme *alternativní hypotéza* a označíme ji H_1.

Testem hypotézy H_0 proti hypotéze H_1 nazveme postup, na jehož základě se rozhodneme hypotézu H_0 buď zamítnout nebo nezamítnout. Test provádíme obvykle tak, že určíme vhodnou množinu $W \subset \mathbb{R}^n$ a pokud $\mathbf{X} \in W$, hypotézu H_0 zamítneme, v opačném případě ji nezamítneme. Množině W říkáme *kritický obor*.

Pokud H_0 platí a nezamítneme ji, nebo pokud platí H_1 a zamítneme H_0, je rozhodnutí správné. Zamítneme-li H_0, přestože platí, říkáme, že jsme se dopustili *chyby prvního druhu*. Nezamítneme-li H_0, přestože platí H_1, říkáme, že jsme se dopustili *chyby druhého druhu*. H_0 volíme tak, aby chyba prvního druhu byla závažnější než chyba druhého druhu. Obvykle požadujeme, aby pravděpodobnost chyby prvního druhu byla nejvýše rovna zvolenému malému číslu α, tj. aby

$$\alpha = \sup_{\theta \in \Theta_0} \mathsf{P}_\theta(\mathbf{X} \in W).$$

Při dodržení tohoto požadavku je nejlepší volba kritického oboru W taková, při níž je pravděpodobnost chyby druhého druhu minimální. Číslo α nazýváme *hladina testu*. Volíme obvykle $\alpha = 0,05$ nebo $\alpha = 0,01$. Číslo $\mathsf{P}_\theta(\mathbf{X} \in W)$ udává pravděpodobnost zamítnutí H_0 za podmínky, že hladina parametru je θ. Můžeme zavést tzv. *silofunkci testu* předpisem $\beta(\theta) = \mathsf{P}_\theta(\mathbf{X} \in W)$, $\theta \in \Theta$. Funkce $\beta(\theta)$ je zdola omezená nulou a shora jedničkou.

Výše popsaný postup se zabývá případem, kdy máme jeden výběr z nějakého rozdělení a rozhodujeme o hodnotě parametru, na němž toto rozdělení závisí. Můžeme se ovšem setkat i s obecnější situací. Mějme $X_{i1}, \dots, X_{in_i}$ náhodné výběry z rozdělení s distribuční funkcí F_i, $i = 1, \dots, k$. Za nulovou, respektive alternativní hypotézu volíme tvrzení, které popisuje nějaký vztah mezi těmito k rozděleními. Například zda jsou si jejich střední hodnoty (či rozptyly) rovny, zda mají všechna rozdělení stejnou distribuční funkci, zda jsou výběry navzájem nezávislé apod. Test provádíme podobně, jako kdybychom pracovali pouze s jediným výběrem: stanovíme vhodný kritický obor $W \subset \mathbb{R}^n$, kde $n = n_1 + \cdots + n_k$ a hypotézu H_0 zamítneme, pokud $\mathbf{X} \in W$, kde symbol $\mathbf{X}$ značí vektor $(X_{11}, \dots, X_{1n_1}, \dots, X_{k1}, \dots, X_{kn_k})'$.

V této práci se budeme zabývat situací, kdy $k = 2$ a zajímá nás porovnání středních hodnot výběrů.

Kapitola 2

Důležitá rozdělení

V hlavní části práce pojednáváme o dvouvýběrovém t testu a dvouvýběrovém Wilcoxonově testu. K jejich konstrukci je třeba znalost některých rozdělení a vztahů mezi nimi, o čemž pojednává tato kapitola.

Dále jsou zde uvedeny definice těch rozdělení, která budeme používat v páté kapitole zabývající se simulacemi.

2.1 Normální rozdělení

Řekneme, že náhodná veličina X má normální rozdělení (též Gaussovo) s parametry μ a σ^2, lze-li její hustotu vyjádřit takto:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}, \quad x \in \mathbb{R},\ \mu \in \mathbb{R},\ \sigma > 0.$$

Normální rozdělení označujeme symbolem $N(\mu, \sigma^2)$. Pro náhodnou veličinu $X \sim N(\mu, \sigma^2)$ platí $\mathsf{E}\, X = \mu$ a $\mathsf{var}\, X = \sigma^2$. Hustotu normálního rozdělení $N(0,1)$ (tzv. *normované normální rozdělení*) značíme φ a jeho distribuční funkci Φ. Hodnoty funkce Φ jsou tabelovány. Je-li $F(x)$ distribuční funkce nějakého normálního rozdělení $N(\mu, \sigma^2)$, platí: $F(x) = \Phi(\frac{x-\mu}{\sigma})$. Tento vztah je velmi užitečný, neboť funkci $F(x)$ nelze vyjádřit pomocí elementárních funkcí. Další důležitý vztah je $\Phi(-x) = 1 - \Phi(x)$. *Kritická hodnota* $u(\alpha)$ normovaného normálního rozdělení $N(0,1)$ je číslo definované vztahem $\mathsf{P}[X \geq u(\alpha)] = \alpha$, kde $X \sim N(0,1)$. Hodnoty $u(\alpha)$ jsou tabelovány (viz Anděl [2] str. 257).

Normované normální rozdělení je důležité i proto, že je limitním rozdělením normovaných součtů nezávislých náhodných veličin. Tvrzení, které konvergenci popisuje, nazýváme *centrální limitní věta*. Existuje více tvarů této věty, zde uvádíme verzi Ljapunovovu.

Tvrzení 2.1. *Nechť $X_{n1}, \ldots, X_{nk_n}$ jsou nezávislé náhodné veličiny. Nechť platí*

$$\lim_{n\to\infty} \frac{\left(\sum_{i=1}^{k_n} \mathsf{E}\, |X_{ni} - \mathsf{E}\, X_{ni}|^3 \right)^{\frac{1}{3}}}{\left(\sum_{i=1}^{k_n} \mathsf{var}\, X_{ni} \right)^{\frac{1}{2}}} = 0.$$

Pak pro každé $x \in \mathbb{R}$ platí:

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{\sum_{i=1}^{k_n}(X_{ni} - \mathsf{E}\,X_{ni})}{\sqrt{\sum_{i=1}^{k_n} \mathsf{var}\,X_{ni}}} \le x\right) = \Phi(x).$$

Důkaz. Důkaz lze najít například v knize Dupač, Hušková [3], str. 85. □

2.2 Dvojitě exponenciální rozdělení

Nechť $a \in \mathbb{R}$ a $\lambda > 0$. Řekneme, že náhodná veličina X má dvojitě exponenciální (též Laplaceovo) rozdělení (značíme $\mathsf{DEx}(a, \lambda)$), jestliže její hustotu lze vyjádřit ve tvaru

$$f(x) = \frac{\lambda}{2}\,\mathrm{e}^{\lambda|x-a|}.$$

Platí-li $X \sim \mathsf{DEx}(a, \lambda)$, pak $\mathsf{E}\,X = a$ a $\mathsf{var}\,X = \frac{2}{\lambda^2}$.

2.3 Cauchyovo rozdělení

Nechť $a \in \mathbb{R}$ a $b > 0$. Náhodná veličina X má Cauchyovo rozdělení, jestliže její hustota má tvar

$$f(x) = \frac{1}{\pi} \cdot \frac{b}{b^2 + (x-a)^2}.$$

Cauchyovo rozdělení značíme symbolem $\mathsf{C}(a, b)$. Cauchyovo rozdělení nemá střední hodnotu, ale jeho hustota je symetrická kolem bodu a.

2.4 Funkce Γ

Gama funkce $\Gamma(a)$ je definována předpisem

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \quad a > 0.$$

2.5 χ^2 rozdělení

Nechť $n \ge 1$. Řekneme, že náhodná veličina X má rozdělení χ^2 o n stupních volnosti, můžeme-li její hustotu vyjádřit ve tvaru

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0.$$

Rozdělení značíme symbolem χ^2_n.

2.6 t rozdělení

Rozdělení t o k stupních volnosti (též Studentovo rozdělení) je rozdělení, jehož hustotu lze vyjádřit takto:

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi k}} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \quad x \in \mathbb{R},\ k \geq 1.$$

Studentovo rozdělení označujeme symbolem t_k. Je-li $k > 1$, pak pro náhodnou veličinu $X \sim \mathsf{t}_k$ platí $\mathsf{E}\,X = 0$. Pro $k = 1$ získáme Cauchyovo rozdělení, které střední hodnotu nemá (viz odstavec 2.3).

V další textu se můžeme setkat se symbolem $Y \sim \mathsf{t}_k + a$. Rozumíme tím, že náhodná veličina $Y = X + a$, kde $X \sim \mathsf{t}_k$ a a je konstanta.

Následující tvrzení se zabývá limitním rozdělením t_k pro $k \to \infty$.

Tvrzení 2.2. *Označme $f_k(x)$ hustotu rozdělení t_k. Pak pro každé $x \in \mathbb{R}$ platí*

$$\lim_{k \to \infty} f_k(x) = \varphi(x).$$

Důkaz. Platnost tvrzení se snadno ověří přímým výpočtem. □

Kritická hodnota $t_k(\alpha)$ rozdělení t_k je číslo definované vztahem $\mathsf{P}\{|T| \geq t_k(\alpha)\} = \alpha$, kde náhodná veličina T má t rozdělení o k stupních volnosti. Povšimněme si, že kritické hodnoty Studentova rozdělení jsou definovány jinak než kritické hodnoty ostatních rozdělení. Hodnoty $t_k(\alpha)$ jsou tabelovány (viz Anděl [2], str. 259). Pro $k \to \infty$ hustota rozdělení t_k konverguje k φ, takže $t_k(\alpha) \to u\left(\frac{\alpha}{2}\right)$ pro $k \to \infty$. Hustota t rozdělení je sudá funkce, proto platí: $\mathsf{P}(T \geq t_k(\alpha)) = \frac{\alpha}{2}$ a $\mathsf{P}(T \leq t_k(\alpha)) = 1 - \frac{\alpha}{2}$.

Velmi užitečné je tvrzení, které popisuje vztah mezi rozděleními t_n, $N(0,1)$ a χ^2_n:

Tvrzení 2.3. *Nechť $X \sim N(0,1)$, $Y \sim \chi^2_n$ jsou nezávislé náhodné veličiny. Pak má náhodná veličina*

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

rozdělení t o n stupních volnosti.

Důkaz. Lze nalézt v knize Anděl [1], str. 74. □

2.7 F rozdělení

Rozdělení F o m a n stupních volnosti (též Fisherovo-Snedecorovo rozdělení) je rozdělení, jehož hustotu lze vyjádřit ve tvaru

$$f_{m,n}(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1} \left(1 + \frac{m}{n}x\right)^{-\frac{m+n}{2}}, \quad x > 0,\ m \geq 1,\ n \geq 1.$$

F rozdělení značíme symbolem $\mathsf{F}_{m,n}$. *Kritická hodnota* $F_{m,n}(\alpha)$ rozdělení F je číslo definované vztahem $\mathsf{P}\{Z \geq F_{m,n}(\alpha)\} = \alpha$, kde náhodná veličina Z má F rozdělení o m a n stupních volnosti. Platí: $F_{m,n}(\alpha) = \frac{1}{F_{n,m}(1-\alpha)}$.

Kapitola 3

Dvouvýběrový t test

Nyní přistupme k samotným testům, jimiž porovnáváme střední hodnotu ve dvou nezávislých výběrech. Buď $X_1, \dots, X_n$ výběr z $N(\mu_1, \sigma^2)$, buď $Y_1, \dots, Y_m$ výběr z $N(\mu_2, \sigma^2)$. Nechť $m \geq 2$, $n \geq 2$, $\sigma^2 > 0$. Chceme testovat hypotézu H_0 : $\mu_1 - \mu_2 = \Delta$ proti alternativě $H_1 : \mu_1 - \mu_2 \neq \Delta$ (případně $H_1 : \mu_1 - \mu_2 < \Delta$ nebo $\mu_1 - \mu_2 > \Delta$). Při testování budeme používat toto označení:

$$\begin{aligned} \overline{X} &= \tfrac{1}{n} \textstyle\sum_{i=1}^n X_i, & \overline{Y} &= \tfrac{1}{m} \textstyle\sum_{i=1}^m Y_i, \\ S_X^2 &= \tfrac{1}{n-1} \textstyle\sum_{i=1}^n (X_i - \overline{X})^2, & S_Y^2 &= \tfrac{1}{m-1} \textstyle\sum_{i=1}^m (Y_i - \overline{Y})^2. \end{aligned}$$

Veličiny $\overline{X}$ a $\overline{Y}$ nazýváme *výběrový průměr* a veličiny S_X^2 a S_Y^2 nazveme *výběrový rozptyl*. Platí následující:

Tvrzení 3.1. *Nechť $X_1, \dots, X_n$ je výběr z $N(\mu_1, \sigma^2)$ a $Y_1, \dots, Y_m$ je výběr z $N(\mu_2, \sigma^2)$. Nechť $m \geq 2$, $n \geq 2$, $\sigma^2 > 0$. Pak*

$$\begin{aligned} \overline{X} &\sim N\left(\mu_1, \tfrac{\sigma^2}{n}\right), & \overline{Y} &\sim N\left(\mu_2, \tfrac{\sigma^2}{m}\right), \\ \tfrac{(n-1)S_X^2}{\sigma^2} &\sim \chi^2_{n-1}, & \tfrac{(m-1)S_Y^2}{\sigma^2} &\sim \chi^2_{m-1}. \end{aligned}$$

Důkaz. Viz Anděl [1], str. 70. □

Pro další postup je důležité, zda jsou výběry na sobě nezávislé. Pokud ano, je nejvhodnější k rozhodování o platnosti H_0 užít dvouvýběrový t test. Pokud ne, používá se párový t test. Ve své práci se zabývám dvouvýběrovým testem, párový test je pro úplnost stručně popsán v odstavci 3.4.

Abychom mohli použít dvouvýběrový t test, musí být splněny tři výše zmíněné předpoklady, výběry musí pocházet z normálního rozdělení, musí být nezávislé a musí mít stejné rozptyly. Postupy, kterými můžeme testovat první dvě vlastnosti, jsou uvedeny například v knize Anděl [1]. Test shody rozptylů je stručně popsán v následujícím odstavci.

3.1 Test shody rozptylů

Chceme testovat hypotézu $H_0 : \sigma_1^2 = \sigma_2^2$ oproti alternativě $H_1 : \sigma_1^2 \neq \sigma_2^2$. Pokud H_0 platí, je $\frac{\sigma_1^2}{\sigma_2^2} = 1$ a zřejmě i podíl výběrových rozptylů bude blízko jedné.

Hypotézu H_0 zamítneme, bude-li $\frac{S_X^2}{S_Y^2} \geq k_1 > 1$ nebo $\frac{S_X^2}{S_Y^2} \leq k_2 < 1$. Lze spočítat, že bude $k_1 = F_{n-1,m-1}\left(\frac{\alpha}{2}\right)$ a $k_2 = F_{n-1,m-1}\left(1-\frac{\alpha}{2}\right)$.

Zdůvodnění těchto rovností jakož i potřebná tvrzení a jejich důkazy lze nalézt v knize Anděl [1], str. 77-79.

3.2 Dvouvýběrový t test

Dvouvýběrový t test je založen na následujícím tvrzení.

Tvrzení 3.2. *Buď $X_1, \ldots, X_n$ výběr z $N(\mu_1, \sigma^2)$, buď $Y_1, \ldots, Y_m$ výběr z $N(\mu_2, \sigma^2)$. Nechť $m \geq 2$, $n \geq 2$, $\sigma^2 > 0$. Nechť oba výběry jsou na sobě nezávislé. Pak náhodná veličina T definovaná vztahem*

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

má $\mathtt{t}$ *rozdělení o $m + n - 2$ stupních volnosti.*

Důkaz. Již víme, že $\overline{X} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$ a $\overline{Y} \sim N\left(\mu_2, \frac{\sigma^2}{m}\right)$. Protože výběry $X_1, \ldots, X_n$ a $Y_1, \ldots, Y_m$ jsou nezávislé, jsou i veličiny $\overline{X}$ a $\overline{Y}$ nezávislé. Lze ukázat, že rozdíl nezávislých normálně rozdělených náhodných veličin má také normální rozdělení, jehož střední hodnota je rovna rozdílu středních hodnot původních dvou veličin a rozptyl je roven součtu jejich rozptylů (viz Anděl [1], str. 65, lemma 4.7). Takže platí $\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m}\right)$, z čehož plyne:

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} \sim N(0,1).$$

Z tvrzení 3.1 víme, že $\frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2_{n-1}$ a $\frac{(m-1)S_Y^2}{\sigma^2} \sim \chi^2_{m-1}$. Protože výběry $X_1, \ldots, X_n$ a $Y_1, \ldots, Y_m$ jsou nezávislé, jsou i veličiny $\frac{(n-1)S_X^2}{\sigma^2}$ a $\frac{(m-1)S_Y^2}{\sigma^2}$ nezávislé a lze ukázat, že pro jejich součet platí:

$$\frac{(n-1)S_X^2 + (m-1)S_Y^2}{\sigma^2} \sim \chi^2_{n+m-2}.$$

Z tvrzení 2.3 víme, že náhodná veličina, jež vznikne podílem náhodné veličiny s normovaným normálním rozdělením a odmocniny z veličiny s rozdělením χ^2, kterou jsme podělili počtem jejích stupňů volnosti, má $\mathtt{t}$ rozdělení o stejném počtu stupňů volnosti, jako mělo rozdělení χ^2. Sestavme takový podíl a zkusme jej upravit:

$$\begin{aligned} T &\stackrel{ozn.}{=} \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} \left(\sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{\sigma^2(n+m-2)}}\right)^{-1} \\ &= \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}. \end{aligned}$$

Přesně takto byla veličina T v tvrzení definována, takže má opravdu rozdělení t_{n+m-2}. $\square$

3.2.1 Test proti oboustranné alternativě

Chceme testovat hypotézu $H_0 : \mu_1 - \mu_2 = \Delta$ oproti alternativě $H_1 : \mu_1 - \mu_2 \neq \Delta$ na hladině α. Pokud H_0 platí, bude $\mu_1 - \mu_2 - \Delta = 0$ a zřejmě i $\overline{X} - \overline{Y} - \Delta$ bude blízko 0. Hypotézu H_0 proto zamítneme, bude-li $\left|\overline{X} - \overline{Y} - \Delta\right| \geq k$ pro nějaké k. Hodnotu k určíme z podmínky, že požadujeme, aby hladina testu (tj. pravděpodobnost chyby prvního druhu) byla α:

$$
\begin{aligned}
\alpha &= \mathsf{P}\left(\text{zamítneme } H_0 | H_0 \text{ platí}\right) \\
&= \mathsf{P}(\left|\overline{X} - \overline{Y} - \Delta\right| \geq k | \mu_1 - \mu_2 = \Delta) \\
&= \mathsf{P}\left(\frac{\left|\overline{X} - \overline{Y} - \Delta\right|}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}}\sqrt{\frac{mn(n+m-2)}{n+m}} \geq l \middle| \mu_1 - \mu_2 = \Delta\right),
\end{aligned}
$$

kde

$$
l = \frac{k}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}}\sqrt{\frac{mn(n+m-2)}{n+m}}. \tag{3.1}
$$

Označme

$$
T = \frac{\overline{X} - \overline{Y} - \Delta}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}}\sqrt{\frac{mn(n+m-2)}{n+m}}. \tag{3.2}
$$

Pak $\alpha = \mathsf{P}\left(|T| \geq l | \mu_1 - \mu_2 = \Delta\right)$. Z tvrzení 3.2 víme, že za platnosti nulové hypotézy má náhodná veličina T rozdělení t_{n+m-2}, takže $l = t_{n+m-2}(\alpha)$. Rozhodovací pravidlo tedy je, že hypotézu H_0 zamítneme na hladině α, bude-li $|T| \geq t_{n+m-2}(\alpha)$.

3.2.2 Test proti jednostranné alternativě

Pokud bychom hypotézu $H_0 : \mu_1 - \mu_2 = \Delta$ chtěli testovat proti alternativě $H_1 : \mu_1 - \mu_2 > \Delta$, bude postup podobný. Platí-li H_0, bude $\mu_1 - \mu_2 - \Delta = 0$ a i $\overline{X} - \overline{Y} - \Delta$ bude blízko 0. Hypotézu H_0 zamítneme, bude-li $\overline{X} - \overline{Y} - \Delta \geq k$ pro nějaké k. Hodnotu k určíme jako v předešlém případě:

$$
\begin{aligned}
\alpha &= \mathsf{P}\left(\text{zamítneme } H_0 | H_0 \text{ platí}\right) \\
&= \mathsf{P}(\overline{X} - \overline{Y} - \Delta \geq k | \mu_1 - \mu_2 = \Delta) \\
&= \mathsf{P}(T \geq l | \mu_1 - \mu_2 = \Delta),
\end{aligned}
$$

kde T a l jsou definovány vztahy 3.2 a 3.1. Již víme, že veličina T má rozdělení t_{n+m-2}, takže $l = t_{n+m-2}(2\alpha)$. (Tento vztah platí pro $\alpha \leq \frac{1}{2}$. Protože požadujeme, aby pravděpodobnost chyby prvního druhu byla malá, nemusíme se výpočtem pro $\alpha > \frac{1}{2}$ zabývat.) Hypotézu H_0 zamítneme na hladině α, bude-li $T \geq t_{n+m-2}(2\alpha)$.

Pro opačnou jednostrannou alternativu $H_1 : \mu_1 - \mu_2 < \Delta$ se zcela analogickým postupem dostaneme k závěru, že H_0 zamítneme na hladině α, bude-li $T \leq -t_{n+m-2}(2\alpha)$.(Stejně jako v předešlém případě, i tento vztah platí pouze pro $\alpha \leq \frac{1}{2}$.)

3.3 Síla testů

V úvodu textu byla zmíněna silofunkce testu $\beta(\theta)$, která je pro $\theta \in \Theta$ rovna pravděpodobnosti, že zamítneme H_0, je-li skutečná hodnota parametru θ. Zabýváme se testováním hypotézy $H_0 : \mu_1 - \mu_2 = \Delta$, takže možné hodnoty parametru nebudeme značit θ, ale pro přehlednost Δ_1. Množina Θ je buď celé $\mathbb{R}$ (testujeme-li proti oboustranné alternativě), nebo jeden z intervalů $[\Delta, \infty)$, $(-\infty, \Delta]$ (testujeme-li proti jednostranné alternativě).

Výpočet hodnot funkce $\beta(\Delta_1)$ pro jednotlivé případy je uveden v následujících třech odstavcích. Pro zjednodušení počítání označme v těchto třech odstavcích

$$c = \frac{1}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}.$$

Připomeňme, že symbol $t_k(\alpha)$ značí číslo definované vztahem $\mathsf{P}\{|T| \geq t_k(\alpha)\} = \alpha$, kde $T \sim \mathsf{t}_k$.

3.3.1 Test proti oboustranné alternativě $\mu_1 - \mu_2 \neq \Delta$

$$\begin{aligned}
\beta(\Delta_1) &= \mathsf{P}\left(\text{zamítneme } H_0 | \mu_1 - \mu_2 = \Delta_1\right) \\
&= \mathsf{P}\left(\frac{\left|\overline{X} - \overline{Y} - \Delta\right|}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}} \geq t_{n+m-2}(\alpha) \Big| \mu_1 - \mu_2 = \Delta_1\right) \\
&= 1 - \mathsf{P}\left(c \left|\overline{X} - \overline{Y} - \Delta_1 + \Delta_1 - \Delta\right| < t_{n+m-2}(\alpha) \left| \mu_1 - \mu_2 = \Delta_1\right.\right) \\
&= 1 - (\zeta - \eta), \qquad \Delta_1 \in \mathbb{R},
\end{aligned}$$

kde pro ζ a η platí:

$$c(\Delta - \Delta_1) - t_{n+m-2}(\alpha) = \begin{cases} t_{n+m-2}(2 - 2\eta) & \text{pro} \quad c(\Delta - \Delta_1) - t_{n+m-2}(\alpha) \geq 0, \\ -t_{n+m-2}(2\eta) & \text{pro} \quad c(\Delta - \Delta_1) - t_{n+m-2}(\alpha) < 0, \end{cases}$$

$$c(\Delta - \Delta_1) + t_{n+m-2}(\alpha) = \begin{cases} t_{n+m-2}(2 - 2\zeta) & \text{pro} \quad c(\Delta - \Delta_1) + t_{n+m-2}(\alpha) \geq 0, \\ -t_{n+m-2}(2\zeta) & \text{pro} \quad c(\Delta - \Delta_1) + t_{n+m-2}(\alpha) < 0. \end{cases}$$

S rostoucím $|\Delta - \Delta_1|$ klesá hodnota $\zeta - \eta$, takže $1 - (\zeta - \eta)$ roste. To znamená, že čím větší je absolutní hodnota rozdílu mezi skutečnou a hypotetickou hodnotou $\mu_1 - \mu_2$, tím větší je pravděpodobnost, že neplatnou hypotézu H_0 zamítneme.

3.3.2 Test proti jednostranné alternativě $\mu_1 - \mu_2 > \Delta$

Zcela analogickým postupem jako v předcházejícím odstavci lze dospět k výsledku

$$\beta(\Delta_1) = \zeta, \qquad \Delta_1 \in [\Delta, \infty) \, ,$$

kde pro ζ platí:

$$t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) = \begin{cases} t_{n+m-2}(2\zeta) & \text{pro} \quad t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) \geq 0, \\ -t_{n+m-2}(2 - 2\zeta) & \text{pro} \quad t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) < 0. \end{cases}$$

S rostoucím Δ_1 klesá hodnota výrazu na levé straně rovnosti, takže ζ roste. To znamená, že čím větší je rozdíl mezi skutečnou a hypotetickou hodnotou $\mu_1 - \mu_2$, tím větší je pravděpodobnost, že neplatnou hypotézu H_0 zamítneme.

3.3.3 Test proti jednostranné alternativě $\mu_1 - \mu_2 < \Delta$

Podobně jako u opačné jednostranné alternativy získáme výpočtem výsledek

$$\beta(\Delta_1) = \zeta, \qquad \Delta_1 \in (-\infty, \Delta] \, ,$$

kde pro ζ platí:

$$-t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) = \begin{cases} t_{n+m-2}(2 - 2\zeta) & \text{pro} \quad -t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) \geq 0, \\ -t_{n+m-2}(2\zeta) & \text{pro} \quad -t_{n+m-2}(2\alpha) - c\,(\Delta_1 - \Delta) < 0. \end{cases}$$

S klesajícím Δ_1 roste hodnota výrazu na pravé straně rovnosti, takže ζ roste. To znamená, že čím větší je rozdíl mezi hypotetickou a skutečnou hodnotou $\mu_1 - \mu_2$, tím větší je pravděpodobnost, že neplatnou hypotézu H_0 zamítneme.

3.4 Párový t test

Mějme dva náhodné výběry stejného rozsahu $X_1, \ldots, X_n$ z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, \ldots, Y_n$ z rozdělení $N(\mu_2, \sigma_2^2)$. Chceme testovat hypotézu $H_0 : \mu_1 - \mu_2 = \Delta$ proti alternativě $H_1 : \mu_1 - \mu_2 \neq \Delta$. Předpokládejme, že mezi náhodnými veličinami X_i a Y_i, $i = 1 \ldots n$ existuje jistá souvislost, například mohou představovat rychlost běžce na začátku a na konci závodu. Je zřejmé, že pokud se obě náhodné veličiny vztahují k témuž objektu, nemůžou být nezávislé, k testování H_0 proto nelze užít dvouvýběrový t test. Můžeme-li ovšem předpokládat, že dvojice (X_i, Y_i) jsou navzájem nezávislé, lze definovat nové veličiny $Z_i = X_i - Y_i$, $i = 1, \ldots, n$. Pak platí $Z_i \sim N(\mu, \sigma^2)$ $\forall i$ kde $\mu = \mu_1 - \mu_2$. (Podobný vztah byl užit už v důkazu tvrzení 3.2, kde se ale odečítaly nezávislé náhodné veličiny s normálním rozdělením. Pro veličiny, které nejsou nezávislé, platí totéž, s výjimkou vztahu mezi rozptyly.) Po takové úpravě přejde nulová hypotéza do tvaru $H_0 : \mu = \Delta$ a alternativní hypotéza bude znít $H_1 : \mu \neq \Delta$. Lze ukázat, že rozhodovací pravidlo v takovém případě je: hypotézu H_0 zamítneme na hladině α, pokud bude $\frac{|\overline{Z} - \Delta|\sqrt{n}}{S_Z} \geq t_{n-1}(\alpha)$.

Kapitola 4

Dvouvýběrový Wilcoxonův test

Buď $X_1, \dots, X_m$ náhodný výběr ze spojitého rozdělení s distribuční funkcí F, buď $Y_1, \dots, Y_n$ náhodný výběr ze spojitého rozdělení s distribuční funkcí G. Nechť oba výběry jsou na sobě nezávislé. Zvolme nulovou hypotézu $H_0 : F = G$ a alternativní hypotézu $H_1 : F \neq G$. Veličiny $X_1, \dots, X_m, Y_1, \dots, Y_n$ (tzv. *sdružený náhodný výběr*) uspořádáme vzestupně podle velikosti. Označme T_1 součet pořadí hodnot $X_1, \dots, X_m$ a T_2 součet pořadí hodnot $Y_1, \dots, Y_n$. Zřejmě platí

$$T_1 + T_2 = 1 + 2 + \cdots + (m+n) = \frac{1}{2}(m+n)(m+n+1). \tag{4.1}$$

Pro další výpočty budeme potřebovat první dva momenty statistik T_1 a T_2.

Tvrzení 4.1. *Nechť náhodné veličiny T_1 a T_2 jsou definovány jako výše. Nechť hypotéza H_0 platí. Pak*

$$\begin{aligned} \mathsf{E}\, T_1 &= \frac{1}{2}m(m+n+1), & \mathsf{var}\, T_1 &= \frac{1}{12}mn(m+n+1), \\ \mathsf{E}\, T_2 &= \frac{1}{2}n(m+n+1), & \mathsf{var}\, T_2 &= \frac{1}{12}mn(m+n+1). \end{aligned}$$

Důkaz. Označme R_i pořadí i-té náhodné veličiny ze sdruženého výběru $X_1, \dots, X_m, Y_1, \dots, Y_n$. Platí-li $H_0 : F = G$, pak sdružený výběr je náhodným výběrem z rozdělení s distribuční funkcí F a náhodná veličina R_i nabývá každé z hodnot $1, \dots, (m+n)$ se stejnou pravděpodobností a sice $\frac{1}{m+n}$. Platí tedy:

$$
\begin{aligned}
\mathsf{E}\, R_i &= \sum_{j=1}^{m+n} \frac{1}{m+n} \cdot j = \frac{1}{m+n} \cdot \frac{1}{2}(m+n)(m+n+1) \\
&= \frac{1}{2}(m+n+1), \qquad i = 1, \dots, (m+n), \\
\mathsf{E}\, R_i^2 &= \sum_{j=1}^{m+n} \frac{1}{m+n} \cdot j^2 = \frac{1}{m+n} \cdot \frac{(m+n)(m+n+1)(2m+2n+1)}{6} \\
&= \frac{(m+n+1)(2m+2n+1)}{6}, \qquad i = 1, \dots, (m+n) \\
\text{a}\, \mathsf{E}\, R_i R_j &= \sum\sum_{k \neq l} \frac{1}{(m+n)(m+n-1)} \cdot kl \\
&= \frac{1}{(m+n)(m+n-1)} \cdot \left(\sum_k \sum_l kl - \sum_k k^2 \right) \\
&= \frac{(m+n+1)(3m+3n+2)}{12}, \quad 1 \le i \le m+n, 1 \le j \le m+n, i \neq j.
\end{aligned}
$$

Náhodná veličina T_1 je součtem pořadí veličin $X_1, \dots, X_m$, lze ji tedy vyjádřit jako $T_1 = \sum_{i=1}^m R_i$. Pak

$$
\begin{aligned}
\mathsf{E}\, T_1 &= \mathsf{E} \left(\sum_{i=1}^m R_i \right) = \sum_{i=1}^m \mathsf{E}\, R_i = \sum_{i=1}^m \frac{1}{2}(m+n+1) = \frac{1}{2} m(m+n+1), \\
\mathsf{E}\, T_1^2 &= \mathsf{E} \left(\sum_{i=1}^m R_i \right)^2 = \mathsf{E} \left(\sum_{i=1}^m R_i^2 + \sum\sum_{i \neq j} R_i R_j \right) \\
&= \sum_{i=1}^m \mathsf{E}\, R_i^2 + \sum\sum_{i \neq j} \mathsf{E}\, R_i R_j = \frac{1}{12} m(m+n+1)(3m^2+3m+n+3mn) \\
\text{a}\, \mathsf{var}\, T_1 &= \mathsf{E}\, T_1^2 - (\mathsf{E}\, T_1)^2 = \frac{1}{12} mn(m+n+1).
\end{aligned}
$$

Z rovnosti 4.1 vidíme, že pro T_2 platí $T_2 = \frac{1}{2}(m+n)(m+n+1) - T_1$, proto

$$
\mathsf{E}\, T_2 = \mathsf{E} \left(\frac{1}{2}(m+n)(m+n+1) - T_1 \right) = \frac{1}{2}(m+n)(m+n+1) - \mathsf{E}\, T_1 = \frac{1}{2} n(m+n+1)
$$

a

$$
\mathsf{var}\, T_2 = \mathsf{var} \left(\frac{1}{2}(m+n)(m+n+1) - T_1 \right) = \mathsf{var}\, T_1 = \frac{1}{12} mn(m+n+1).
$$

$\square$

Namísto testových statistik T_1, T_2 se pro test H_0 častěji používají statistiky U_1 a U_2 definované těmito vztahy:

$$U_1 = mn + \frac{1}{2}m(m+1) - T_1 \tag{4.2}$$

$$U_2 = mn + \frac{1}{2}n(n+1) - T_2 \tag{4.3}$$

Tvrzení 4.2. *Buďte U_1 a U_2 náhodné veličiny definované jako výše. Platí-li H_0, pak*

$$\mathsf{E}\,U_1 = \frac{1}{2}mn, \qquad \mathsf{var}\,U_1 = \frac{1}{12}mn(m+n+1),$$

$$\mathsf{E}\,U_2 = \frac{1}{2}mn, \qquad \mathsf{var}\,U_2 = \frac{1}{12}mn(m+n+1).$$

Důkaz. Z rovnosti 4.2 plyne

$$\mathsf{E}\,U_1 = \mathsf{E}\left(mn + \frac{1}{2}m(m+1) - T_1\right) = mn + \frac{1}{2}m(m+1) - \mathsf{E}\,T_1 = \frac{1}{2}mn,$$

$$\mathsf{var}\,U_1 = \mathsf{var}\left(mn + \frac{1}{2}m(m+1) - T_1\right) = \mathsf{var}\,T_1 = \frac{1}{12}mn(m+n+1).$$

Podobně pro U_2 z rovnosti 4.3 plyne

$$\mathsf{E}\,U_2 = \mathsf{E}\left(mn + \frac{1}{2}n(n+1) - T_2\right) = mn + \frac{1}{2}n(n+1) - \mathsf{E}\,T_2 = \frac{1}{2}mn,$$

$$\mathsf{var}\,U_2 = \mathsf{var}\left(mn + \frac{1}{2}n(n+1) - T_2\right) = \mathsf{var}\,T_2 = \frac{1}{12}mn(m+n+1).$$

□

Z čísel $1, \dots, m+n$ lze utvořit celkem $\binom{m+n}{m}$ kombinací bez opakování o m prvcích. Platí-li $H_0 : F = G$, pak sdružený výběr je náhodným výběrem z rozdělení s distribuční funkcí F a pořadí jeho prvků je zcela náhodné. Jinými slovy, prvky množiny $\{R_1, \dots, R_m\}$ může být každá z oněch $\binom{m+n}{m}$ kombinací se stejnou pravděpodobností. S touto znalostí můžeme snadno určit rozdělení veličin U_1 a U_2. Ve prospěch $H_1 : F \neq G$ bude svědčit to, pokud se po uspořádání sdruženého náhodného výběru ocitne „mnoho“ prvků výběru $X_1, \dots, X_m$ na příliš „nízkých“ nebo naopak „vysokých“ příčkách tohoto uspořádání. Nepřesné pojmy označené uvozovkami lze snadno nahradit přesným rozhodovacím pravidlem. Příliš nízkým či vysokým pozicím totiž odpovídají velké, resp. malé hodnoty veličin U_1 a U_2, jejichž rozdělení je známé.

Definujeme *kritickou hodnotu* $W(\alpha)$ pro dvouvýběrový Wilcoxonův test vztahem $\mathsf{P}[min(U_1, U_2) \leq W(\alpha)] \leq \alpha$. Hypotézu H_0 zamítneme, jestliže $min(U_1, U_2)$ je menší nebo rovno příslušné tabelované kritické hodnotě (viz Statistické metody [2] str. 266-267).

Povšimněme si, že v definici kritické hodnoty se objevuje $\mathsf{P}[\dots] \leq \alpha$ a nikoli $\mathsf{P}[\dots] = \alpha$ jako tomu bylo u definice kritické hodnoty $\mathtt{t}$ rozdělení. Je tomu tak proto, že veličiny U_1 a U_2 mají diskrétní rozdělení a α proto nemůže nabývat všech hodnot z intervalu $[0, 1]$.

Jsou-li n, m velká ($n > 10, m > 10$), použijeme pro testování hypotézy H_0 následující postup. Lze ukázat, že pro $m \to \infty$ a $n \to \infty$ má veličina U_1 asymptoticky normální rozdělení, takže veličina U definovaná vztahem

$$U = \frac{U_1 - \mathsf{E}\, U_1}{\sqrt{\mathsf{var}\, U_1}}$$

má rozdělení $N(0, 1)$ a H_0 zamítneme na hladině blížící se α, pokud $|U| \geq u(\frac{\alpha}{2})$, kde $u(\frac{\alpha}{2})$ je kritická hodnota pro normální rozdělení $N(0, 1)$.

Kapitola 5

Výsledky testů

V této kapitole se zabýváme následující situací. Mějme $X_1, \ldots, X_n$ náhodný výběr ze spojitého rozdělení se střední hodnotou μ_1 a $Y_1, \ldots, Y_m$ náhodný výběr ze spojitého rozdělení se střední hodnotou μ_2. Chceme testovat hypotézu $H_0 : \mu_1 - \mu_2 = 0$ proti oboustranné alternativě $H_1 : \mu_1 - \mu_2 \neq 0$ na hladině $\alpha = 0,05$. K testování užijeme jednak dvouvýběrový t test, jednak dvouvýběrový Wilcoxonův test. K tomuto účelu používáme program R 2.2.0, ve kterém jsou k dispozici příkazy `t.test` a `wilcox.test` pro provádění obou testů. Pomocí tohoto programu rovněž simulujeme náhodné výběry, na nichž testy provádíme. Pro daná rozdělení náhodných veličin X a Y jsem provedla vždy 1 000 opakování obou testů a zaznamenala jsem si, s jakou úspěšností probíhaly. Pojmem úspěšnost rozumím podíl testů se správným výsledkem (tj. těch, které zamítly neplatnou, nebo nezamítly platnou hypotézu H_0) na celkovém počtu uskutečněných testů vyjádřený v procentech. Zkoumala jsem, zda a jak je průběh ovlivněn splněním (resp. porušením) podmínek, které na data klademe při užívání t testu, tj. shody rozptylů a normality, zda jsou výsledky ovlivněny rozsahem výběrů a zda se úspěšnost testů liší při testování platné či neplatné nulové hypotézy.

Výsledky všech testů jsou ilustrovány grafy vytvořenými v programu R 2.2.0. Červenou barvou jsou vždy vyznačeny výsledky t testu, modrou výsledky Wilcoxonova testu. Výsledky některých testů jsou navíc zaznamenány v tabulkách, uvádím v nich pouze ty nejdůležitější a sice: testy, v nichž byla úspěšnost t testu nejnižší a nejvyšší, test, v němž úspěšnost t testu poprvé překročila hodnotu $50,0\%$ a $95,0\%$ (resp. pod tyto hodnoty klesla) a test, v němž úspěšnost t testu poprvé dosáhla hranice $100,0\%$ (resp. pod ni klesla). Dále testy, ve kterých výše uvedených výsledků dosáhl Wilcoxonův test. V těch případech, kdy není zřejmá závislost mezi hodnotou proměnlivého parametru a výsledky testů (viz obr. 5.2, 5.4, 5.6, 5.8, 5.9, 5.10, 5.13 a 5.15), uvádím testy s nejvyšší a nejnižší úspěšností a též prvý, padesátý, stý, stý padesátý a poslední test.

Abychom nemuseli pořád opakovat „rozdělení, z něhož pochází náhodný výběr $X_1, \ldots, X_n$ (resp. $Y_1, \ldots, Y_m$)“, budeme v následujícím textu místo toho někdy používat úspornější vyjádření „rozdělení náhodné veličiny X (resp. Y)“.

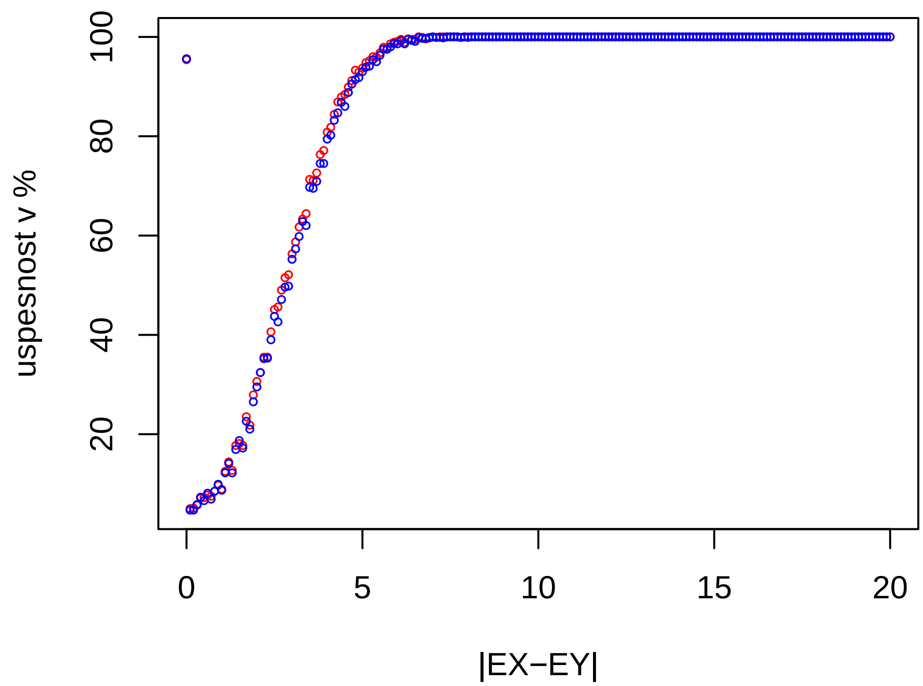
5.1 Splněny předpoklady dvouvýběrového t testu

V prvním souboru testů měla jak náhodná veličina X, tak náhodná veličina Y rozdělení $N(0, 10)$, t test skončil úspěšně v 95,4% případů, Wilcoxonův test skončil úspěšně v 95,6%. To odpovídá tomu, že testy probíhaly na hladině 5%, tj. že pravděpodobnost zamítnutí platné hypotézy H_0 měla být právě pětiprocentní.

Ve druhém až dvoustém prvém souboru testů platilo $X \sim N(0, 10)$ a $Y \sim N(\frac{i}{10}, 10)$, $i = 1, \ldots, 200$. Úspěšnost testů (tj. podíl těch, které neplatnou hypotézu H_0 zamítly na celkovém počtu proběhlých testů) se s rostoucí absolutní hodnotou rozdílu $\mu_1 - \mu_2$ zvyšovala, což odpovídá výpočtu, který byl proveden v odstavci 3.3.1. Pro $|\mu_1 - \mu_2| \geq 8,1$ zamítly oba testy neplatnou hypotézu H_0 ve $100,0\%$ případů.

Výsledky dvouvýběrového t testu a Wilcoxonova testu se v žádném z proběhlých souborů testů od sebe nelišily o více než 3%, ale t test byl většinou lepší, pouze sedmkrát dosáhl Wilcoxonův test vyšší úspěšnosti než t test.

Obrázek 5.1 znázornňuje závislost úspěšnosti testů na hodnotě výrazu $|\mu_1 - \mu_2|$.



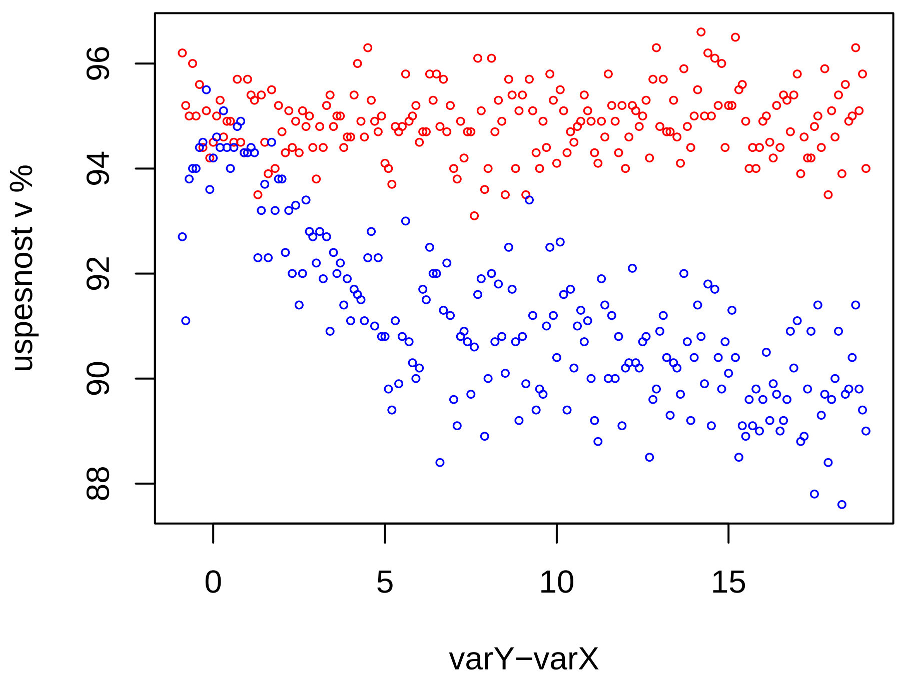
Obrázek 5.1: Splněny předpoklady t testu

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,10)$	$N(0,\ 10)$	100	100	0	95,4	95,6
2	$N(0,10)$	$N(0{,}1,\ 10)$	100	100	-0,1	5,0	4,7
29	$N(0,10)$	$N(2{,}8,\ 10)$	100	100	-2,8	51,5	49,6
31	$N(0,10)$	$N(3{,}0,\ 10)$	100	100	-3,0	56,3	55,2
53	$N(0,10)$	$N(5{,}2,\ 10)$	100	100	-5,2	95,2	94,1
54	$N(0,10)$	$N(5{,}3,\ 10)$	100	100	-5,3	96,0	95,4
67	$N(0,10)$	$N(6{,}6,\ 10)$	100	100	-6,6	100,0	99,9
71	$N(0,10)$	$N(7{,}0,\ 10)$	100	100	-7,0	100,0	100,0

5.2 Porušen předpoklad shody rozptylů

V následujících souborech testů bylo pro rozdělení náhodných veličin nejprve $X \sim N(0,1)$ a $Y \sim N\left(0, \frac{i}{10}\right)$, $i = 1, \dots, 200$, poté byla zkoumána rozdělení $X \sim N(0,1)$ a $Y \sim N\left(5, \frac{i}{10}\right)$, $i = 1, \dots, 200$.

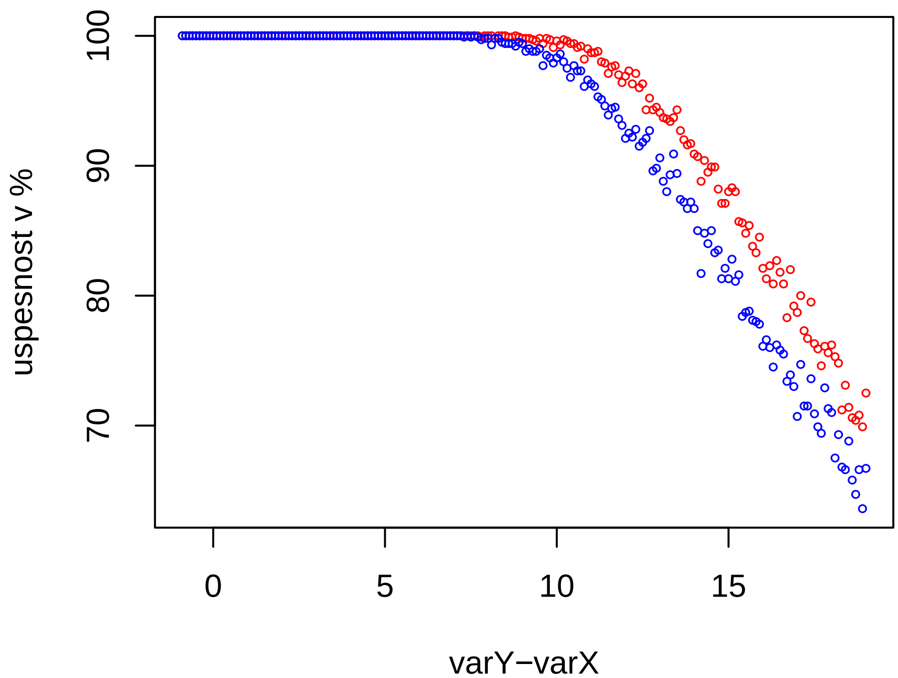
V prvním případě byla úspěšnost t testu i Wilcoxonova testu velmi vysoká pro všechny hodnoty $\mathsf{var}\, Y$, pohybovala se v rozmezí 87,6−96,6%, přičemž t test většinou dopadl lépe než Wilcoxonův test. Se zvyšující se hodnotou výrazu $|\mathsf{var}\, X - \mathsf{var}\, Y|$ úspěšnost Wilcoxonova testu nepatrně klesá, tento pokles je velmi dobře patrný na obrázku 5.2.



Obrázek 5.2: Porušen předpoklad shody rozptylů, H_0 platí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(0,\ 0{,}1)$	100	100	0	96,2	92,7
8	$N(0,1)$	$N(0,\ 0{,}8)$	100	100	0	95,1	95,5
50	$N(0,1)$	$N(0,\ 5{,}0)$	100	100	0	94,6	91,1
86	$N(0,1)$	$N(0,\ 8{,}6)$	100	100	0	93,1	90,6
100	$N(0,1)$	$N(0,\ 10{,}0)$	100	100	0	95,4	90,8
150	$N(0,1)$	$N(0,\ 15{,}0)$	100	100	0	95,0	90,4
152	$N(0,1)$	$N(0,\ 15{,}2)$	100	100	0	96,6	90,8
193	$N(0,1)$	$N(0,\ 19{,}3)$	100	100	0	93,9	87,6
200	$N(0,1)$	$N(0,\ 20{,}0)$	100	100	0	94,0	89,0

Ve druhém případě byla pro hodnoty výrazu $|\mathsf{var}\,X - \mathsf{var}\,Y|$ menší nebo jen nepatrně vyšší než hodnota $|\mu_1 - \mu_2|$ úspěšnost t testu i Wilcoxonova testu stoprocentní, to znamená, že oba testy vždy v každém z tisíci případů zamítly neplatnou hypotézu H_0. Se zvyšující se hodnotou výrazu $|\mathsf{var}\,X - \mathsf{var}\,Y|$ ovšem úspěšnost obou testů velmi rychle klesá, přičemž t test je o něco málo úspěšnější než Wilcoxonův test. Výsledky všech testů jsou znázorněny na obrázku 5.3.



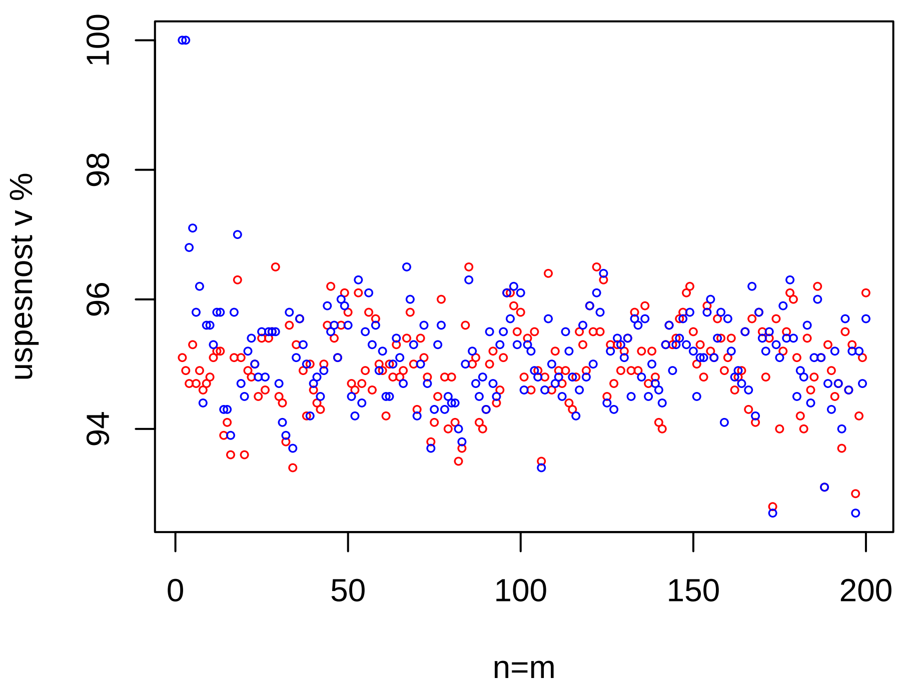
Obrázek 5.3: Porušen předpoklad shody rozptylů, H_0 neplatí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(5,\ 0{,}1)$	100	100	-5	100,0	100,0
83	$N(0,1)$	$N(5,\ 8{,}3)$	100	100	-5	100,0	99,9
88	$N(0,1)$	$N(5,\ 8{,}8)$	100	100	-5	99,9	99,7
124	$N(0,1)$	$N(5,\ 12{,}4)$	100	100	-5	97,9	94,6
136	$N(0,1)$	$N(5,\ 13{,}6)$	100	100	-5	94,3	92,1
199	$N(0,1)$	$N(5,\ 19{,}9)$	100	100	-5	69,9	63,6

5.3 Stejné rozsahy výběrů

V této části práce jsem se zabývala náhodným výběrem $X_1, \ldots, X_n$ výběr z $N(0,1)$, který jsem porovnávala nejprve s $Y_1, \ldots, Y_n$ výběr z $N(0,1)$ a potom s $Y_1, \ldots, Y_n$ výběr z $N(1,1)$, $n = 2, \ldots, 200$.

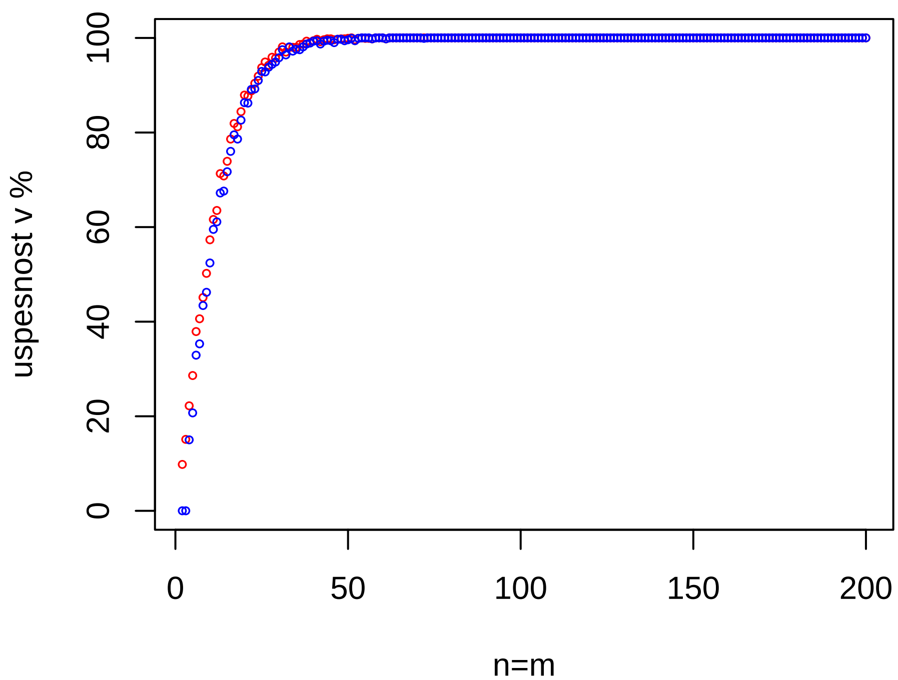
Jak je vidět na obrázku 5.4, úspěšnost t testu i Wilcoxonova testu se v prvním případě pohybovala kolem hodnoty 95%, což odpovídá tomu, že testy byly prováděny na hladině $\alpha = 0{,}05$.



Obrázek 5.4: Stejné rozsahy výběrů, H_0 platí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(0,1)$	2	2	0	95,1	100,0
28	$N(0,1)$	$N(0,1)$	29	29	0	96,5	95,5
50	$N(0,1)$	$N(0,1)$	51	51	0	94,7	94,5
100	$N(0,1)$	$N(0,1)$	101	101	0	94,8	94,6
150	$N(0,1)$	$N(0,1)$	151	151	0	95,0	94,5
172	$N(0,1)$	$N(0,1)$	173	173	0	92,8	92,7
199	$N(0,1)$	$N(0,1)$	200	200	0	96,1	95,7

V testech, v nichž H_0 neplatila, byla úspěšnost t testu i Wilcoxonova testu závislá na rozsazích náhodných výběrů, pro malé hodnoty n byla velmi nízká, se zvyšujícím se n roste. Pro malé rozsahy výběrů ($n \leq 30$) funguje t test o něco lépe, neplatnou H_0 zamítá ve větším počtu případů než test Wilcoxonův. Při velkých rozsazích náhodných výběrů není mezi úspěšností obou testů větší rozdíl, což je ostatně dobře patrné i na obrázku 5.5.



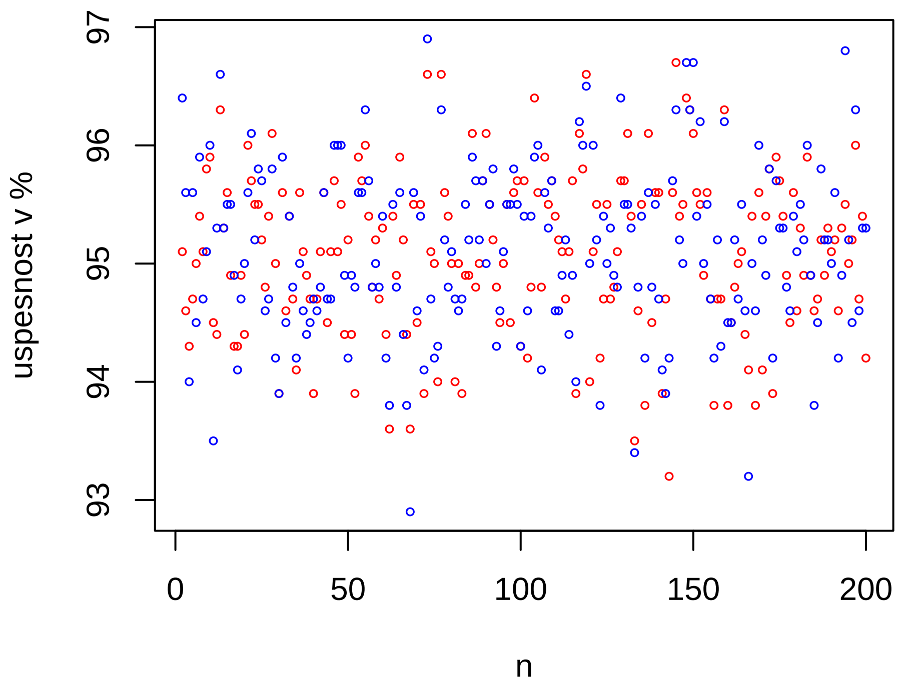
Obrázek 5.5: Stejné rozsahy výběrů, H_0 neplatí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(1,1)$	2	2	-1	9,8	0,0
8	$N(0,1)$	$N(1,1)$	9	9	-1	50,2	46,2
9	$N(0,1)$	$N(1,1)$	10	10	-1	57,3	52,4
27	$N(0,1)$	$N(1,1)$	28	28	-1	95,9	94,4
29	$N(0,1)$	$N(1,1)$	30	30	-1	97,0	95,8
50	$N(0,1)$	$N(1,1)$	51	51	-1	100,0	99,9
53	$N(0,1)$	$N(1,1)$	54	54	-1	100,0	100,0

5.4 Nestejné rozsahy výběrů

V této části práce jsem nejprve pracovala s náhodnými výběry $Y_1, \dots, Y_{100}$ výběr z $N(0,1)$ a $X_1, \dots, X_n$ výběr z $N(0,1)$, $n = 2, \dots, 200$ a poté s náhodnými výběry $Y_1, \dots, Y_{100}$ výběr z $N(1,1)$ a $X_1, \dots, X_n$ výběr z $N(0,1)$, $n = 2, \dots, 200$.

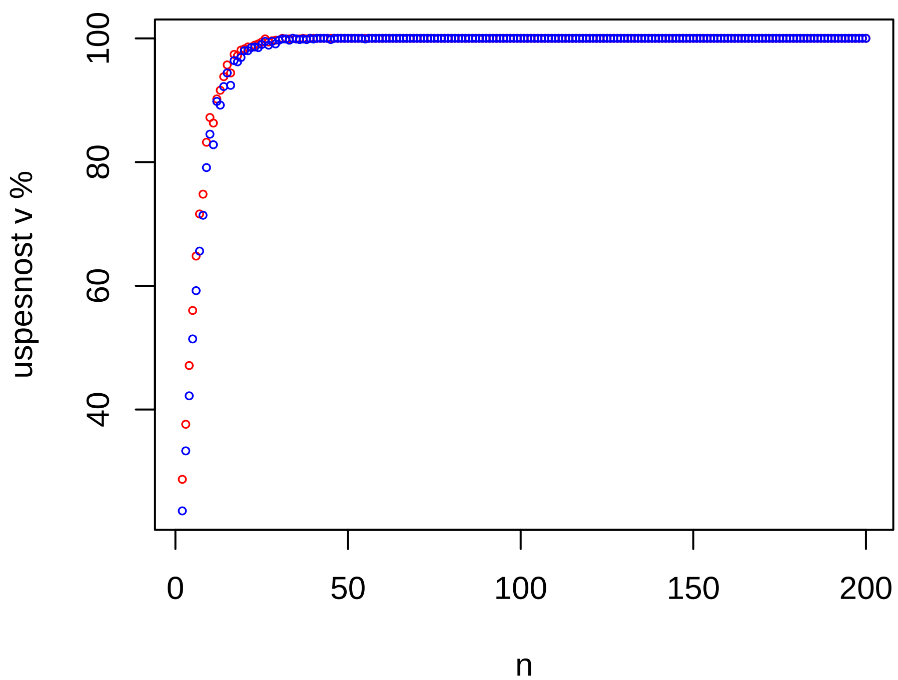
Jak je vidět v tabulce i na obrázku 5.6, za platnosti H_0 úspěšnost dvouvýběrového t testu ani dvouvýběrového Wilcoxonova testu na rozsahu náhodného výběru $X_1, \dots, X_n$ nezávisí, pohybuje se v rozmezí 92,9 − 96,9% u Wilcoxonova a 93,2 − 96,7% u t testu.



Obrázek 5.6: Nestejné rozsahy výběrů, H_0 platí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(0,1)$	2	100	0	95,1	96,4
50	$N(0,1)$	$N(0,1)$	51	100	0	94,4	94,9
67	$N(0,1)$	$N(0,1)$	68	100	0	93,6	92,9
72	$N(0,1)$	$N(0,1)$	73	100	0	96,6	96,9
100	$N(0,1)$	$N(0,1)$	101	100	0	95,7	95,4
142	$N(0,1)$	$N(0,1)$	143	100	0	93,2	94,2
144	$N(0,1)$	$N(0,1)$	145	100	0	96,7	96,3
150	$N(0,1)$	$N(0,1)$	151	100	0	95,6	95,4
199	$N(0,1)$	$N(0,1)$	200	100	0	94,2	95,3

Pokud H_0 neplatí, tak se úspěšnost dvouvýběrového t testu i dvouvýběrového Wilcoxonova testu s rostoucím rozsahem náhodného výběru $X_1, \ldots, X_n$ zvyšuje, pro $n \geq 30$ se pohybuje v rozmezí $99,8 - 100,0\%$, pro $n \geq 56$ je dokonce u obou testů stoprocentní. Výsledky jsou zachyceny na obrázku 5.7.

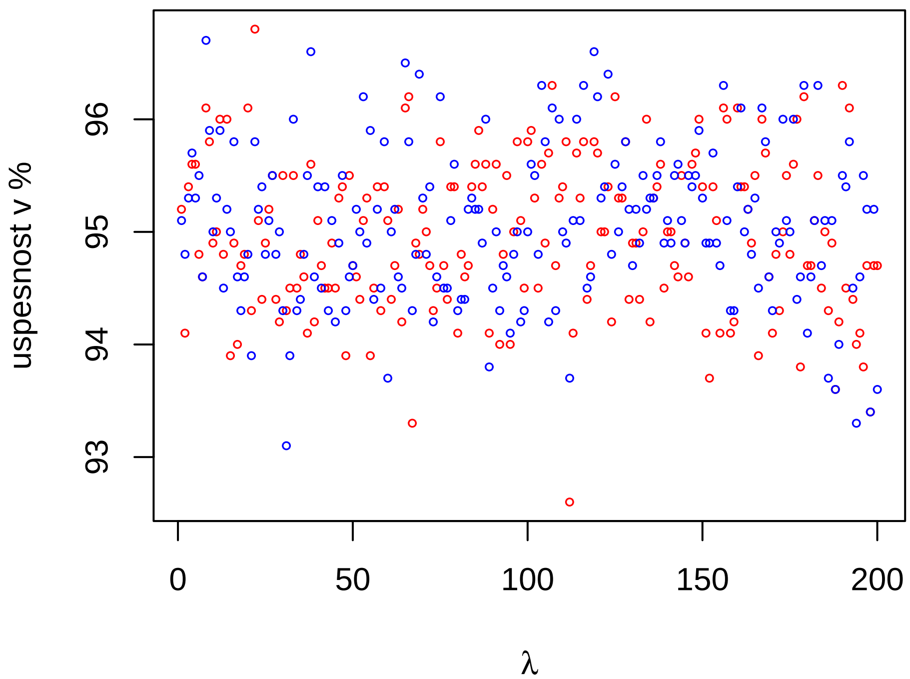


Obrázek 5.7: Nestejné rozsahy výběrů, H_0 neplatí

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$N(0,1)$	$N(1,1)$	2	100	-1	28,7	23,6
4	$N(0,1)$	$N(1,1)$	5	100	-1	56,0	51,4
14	$N(0,1)$	$N(1,1)$	15	100	-1	95,7	94,4
16	$N(0,1)$	$N(1,1)$	17	100	-1	97,4	96,4
30	$N(0,1)$	$N(1,1)$	31	100	-1	100,0	99,9
33	$N(0,1)$	$N(1,1)$	34	100	-1	100,0	100,0

5.5 Dvojitě exponenciální rozdělení

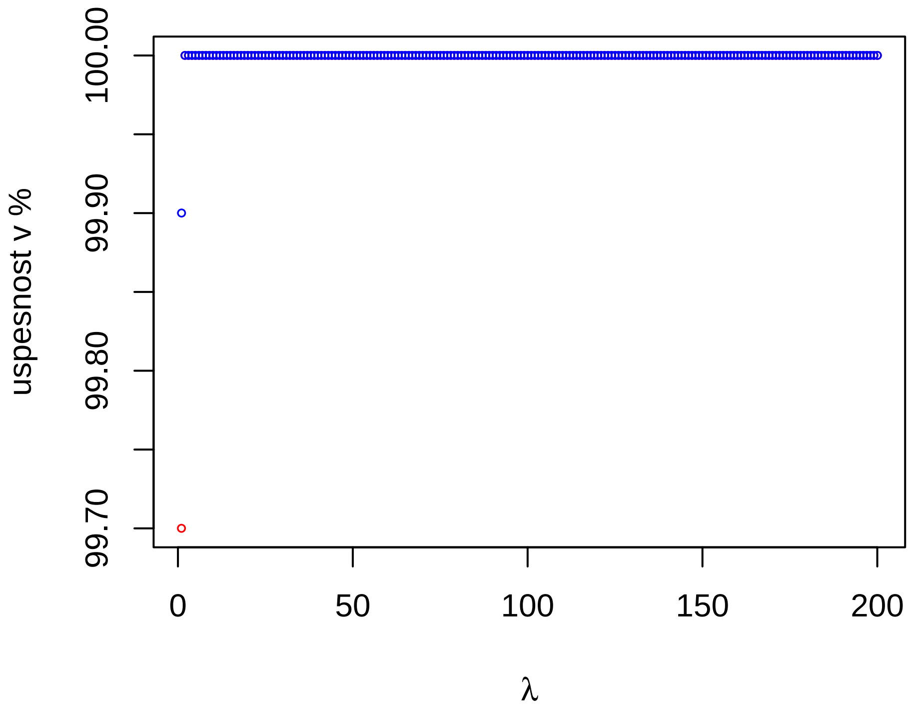
Mějme $X_1, \ldots, X_{100}$ a $Y_1, \ldots, Y_{100}$ výběry z $\mathsf{DEx}(0, \lambda)$, $\lambda = 1, \ldots, 200$. Přestože je v těchto výběrech porušen předpoklad normality, který na data klademe při používání dvouvýběrového t testu, v 95 z 200 případů byl t test lepší nebo stejně dobrý jako Wilcoxonův test. Rozdíly mezi oběma testy ale nejsou nijak veliké, jejich výsledky se liší nanejvýš o 2%. Z obrázku 5.8 není patrná žádná závislost mezi úspěšností testů a hodnotou parametru λ. Na tomto obrázku si můžeme rovněž povšimnout, že úspěšnost obou testů se pohybuje kolem hladiny $95,0\%$.



Obrázek 5.8: $X \sim \mathsf{DEx}(0, \lambda)$, $Y \sim \mathsf{DEx}(0, \lambda)$

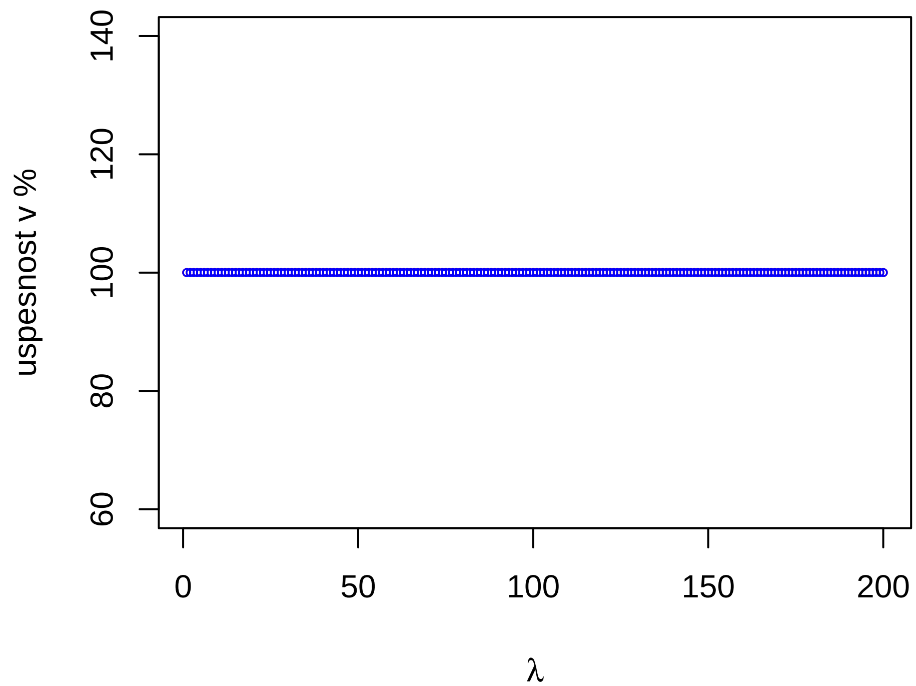
č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$\mathsf{DEx}(0,1)$	$\mathsf{DEx}(0,1)$	100	100	0	95,2	95,1
8	$\mathsf{DEx}(0,8)$	$\mathsf{DEx}(0,8)$	100	100	0	96,1	96,7
22	$\mathsf{DEx}(0,22)$	$\mathsf{DEx}(0,22)$	100	100	0	96,8	95,8
31	$\mathsf{DEx}(0,31)$	$\mathsf{DEx}(0,31)$	100	100	0	94,3	93,1
50	$\mathsf{DEx}(0,50)$	$\mathsf{DEx}(0,50)$	100	100	0	94,7	94,7
100	$\mathsf{DEx}(0,100)$	$\mathsf{DEx}(0,100)$	100	100	0	95,8	95,0
112	$\mathsf{DEx}(0,112)$	$\mathsf{DEx}(0,112)$	100	100	0	92,6	93,7
150	$\mathsf{DEx}(0,150)$	$\mathsf{DEx}(0,150)$	100	100	0	95,4	95,3
200	$\mathsf{DEx}(0,200)$	$\mathsf{DEx}(0,200)$	100	100	0	94,7	93,6

Dále jsem se zabývala následujícími dvěma situacemi. Výběr $Y_1, \ldots, Y_{100}$ z $\mathsf{DEx}(0, \lambda)$ jsem porovnávala nejprve s výběrem $X_1, \ldots, X_{100}$ z $\mathsf{DEx}(1, \lambda)$ a poté s výběrem $X_1, \ldots, X_{100}$ z $\mathsf{DEx}(5, \lambda)$, $\lambda = 1, \ldots, 200$ Domnívám se, že jakýkoli komentář je na tomto místě téměř zbytečný, neboť v 399 ze 400 uskutečněných testů zamítl jak Wilcoxonův tak t test neplatnou hypotézu H_0 ve $100,0\%$ případů. Situace jsou ilustrovány obrázky 5.9 a 5.10.



Obrázek 5.9: $X \sim \mathsf{DEx}(1, \lambda)$, $Y \sim \mathsf{DEx}(0, \lambda)$

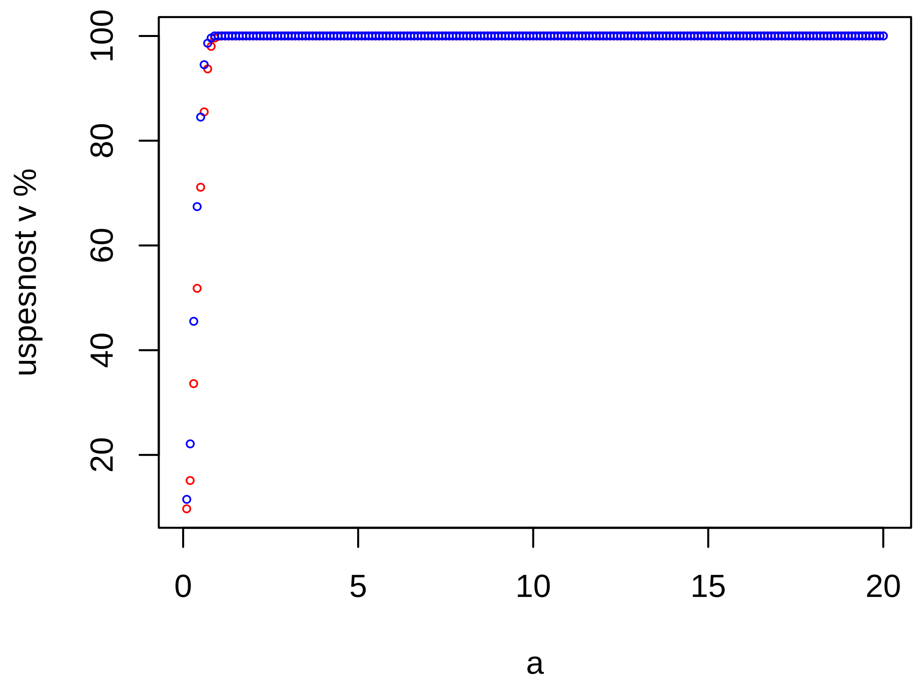
č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$\mathsf{DEx}(1, 1)$	$\mathsf{DEx}(0, 1)$	100	100	1	99,7	99,9
50	$\mathsf{DEx}(1, 50)$	$\mathsf{DEx}(0, 50)$	100	100	1	100,0	100,0
100	$\mathsf{DEx}(1, 100)$	$\mathsf{DEx}(0, 100)$	100	100	1	100,0	100,0
150	$\mathsf{DEx}(1, 150)$	$\mathsf{DEx}(0, 150)$	100	100	1	100,0	100,0
200	$\mathsf{DEx}(1, 200)$	$\mathsf{DEx}(0, 200)$	100	100	1	100,0	100,0



Obrázek 5.10: $X \sim \mathsf{DEx}(5, \lambda)$, $Y \sim \mathsf{DEx}(0, \lambda)$

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$\mathsf{DEx}(5, 1)$	$\mathsf{DEx}(0, 1)$	100	100	5	100,0	100,0
50	$\mathsf{DEx}(5, 50)$	$\mathsf{DEx}(0, 50)$	100	100	5	100,0	100,0
100	$\mathsf{DEx}(5, 100)$	$\mathsf{DEx}(0, 100)$	100	100	5	100,0	100,0
150	$\mathsf{DEx}(5, 150)$	$\mathsf{DEx}(0, 150)$	100	100	5	100,0	100,0
200	$\mathsf{DEx}(5, 200)$	$\mathsf{DEx}(0, 200)$	100	100	5	100,0	100,0

Nakonec jsem pro případ dvojitě exponenciálního rozdělení zkoumala následující situaci. Mějme $Y_1, \ldots, Y_{100}$ výběr z $\mathsf{DEx}(0, 1)$ a $X_1, \ldots, X_{100}$ výběr z $\mathsf{DEx}(a, 1)$, $a = 0{,}1,\ 0{,}2, \ldots, 20{,}0$. Na obrázku 5.11 je dobře vidět, že pro malé hodnoty a (platí $a = \mu_1 - \mu_2$) není ani jeden z testů příliš spolehlivý, ale se zvyšujícím se a úspěšnost obou testů velmi rychle roste a již pro $a = 1, 1$ je stoprocentní. V testech s menší než stoprocentní úspěšností dopadá lépe Wilcoxonův test.



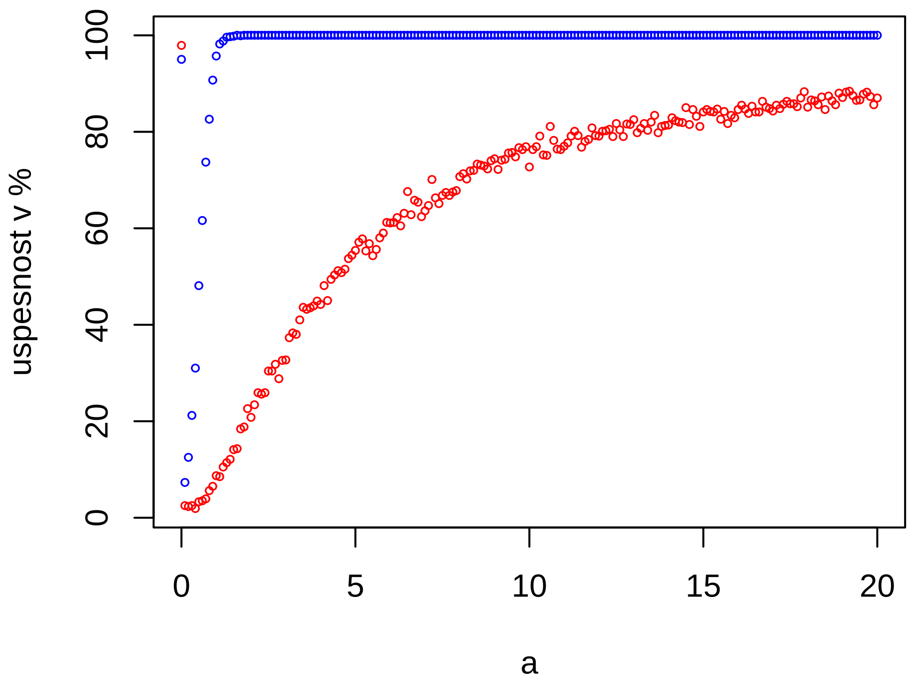
Obrázek 5.11: $X \sim \mathsf{DEx}(a, 1)$, $Y \sim \mathsf{DEx}(0, 1)$

	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
č. testu	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	DEx(0,1, 1)	DEx(0, 1)	100	100	0,1	9,7	11,5
4	DEx(0,4, 1)	DEx(0, 1)	100	100	0,1	51,8	67,4
7	DEx(0,7, 1)	DEx(0, 1)	100	100	0,7	93,7	98,6
8	DEx(0,8, 1)	DEx(0, 1)	100	100	0,8	98,0	99,6
9	DEx(0,9, 1)	DEx(0, 1)	100	100	0,9	99,6	100,0
11	DEx(1,1, 1)	DEx(0, 1)	100	100	1,1	100,0	100,0

5.6 Cauchyovo rozdělení

Na tomto místě své práce jsem se zabývala náhodnými výběry $X_1, \ldots, X_{100}$ výběr z $\mathsf{C}(0, 1)$ a $Y_1, \ldots, Y_{100}$ výběr z $\mathsf{C}(a_2, 1)$, $a_2 = 0{,}0,\ 0{,}1, \ldots, 20{,}0$. Jak víme z odstavce 2.3, Cauchyovo rozdělení $\mathsf{C}(a, b)$ nemá střední hodotu. Jeho hustota je ovšem symetrická kolem bodu a. Pro toto rozdělení přeformulujeme nulovou a alternativní hypotézu do tvaru $H_0 : a_1 = a_2$, $H_1 : a_1 \neq a_2$. V našem případě je a_1 rovno nule.

V jediném případě je pro Cauchuovo rozdělení t test lepší Wilcoxoův test, a sice v případě, kdy H_0 platí, a to o pouhá $3,1\%$. Wilcoxonův test je ve všech ostatních testech výrazně úspěšnější, pro $|a_2 - a_1| \geq 1,9$ zamítá neplatnou hypotézu H_0 ve sto procentech případů, zatímco úspěšnost t testu při posuzování neplatné hypotézy dosahuje nejvýše 88,4%. Jak vidíme v tabulce, v sedmnáctém testu, kdy Wilcoxonův test poprvé dosahuje hranice stoprocentní úspěšnosti, zamítl t test neplatnou hypotézu pouze v každém sedmém případě.

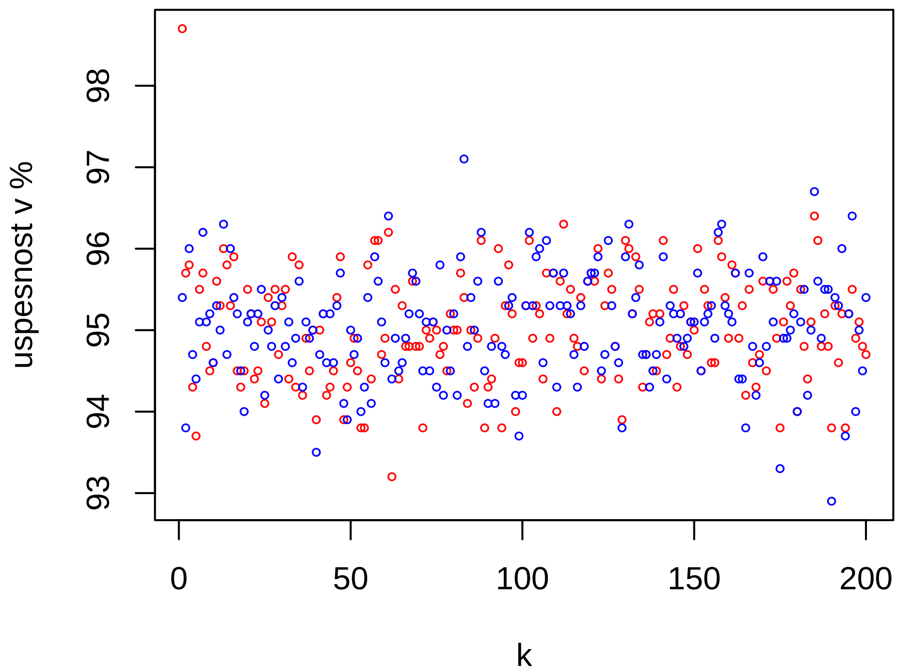


Obrázek 5.12: Cauchyovo rozdělení

č. testu	rozdělení náh. veličiny		rozsah výběru		$a_1 - a_2$	úspěšnost testů v %	
	X	Y	n	m		t	W
1	C(0, 1)	C(0,1, 1)	100	100	-0,1	97,5	94,4
2	C(0, 1)	C(0,2, 1)	100	100	-0,2	2,5	7,3
5	C(0, 1)	C(0,5, 1)	100	100	-0,5	1,9	31,0
7	C(0, 1)	C(0,7, 1)	100	100	-0,7	3,5	61,6
11	C(0, 1)	C(1,1, 1)	100	100	-1,1	8,7	95,7
17	C(0, 1)	C(1,7, 1)	100	100	-1,7	14,3	100,0
45	C(0, 1)	C(4,5, 1)	100	100	-4,5	50,3	100,0
193	C(0, 1)	C(19,3, 1)	100	100	-19,3	88,4	100,0

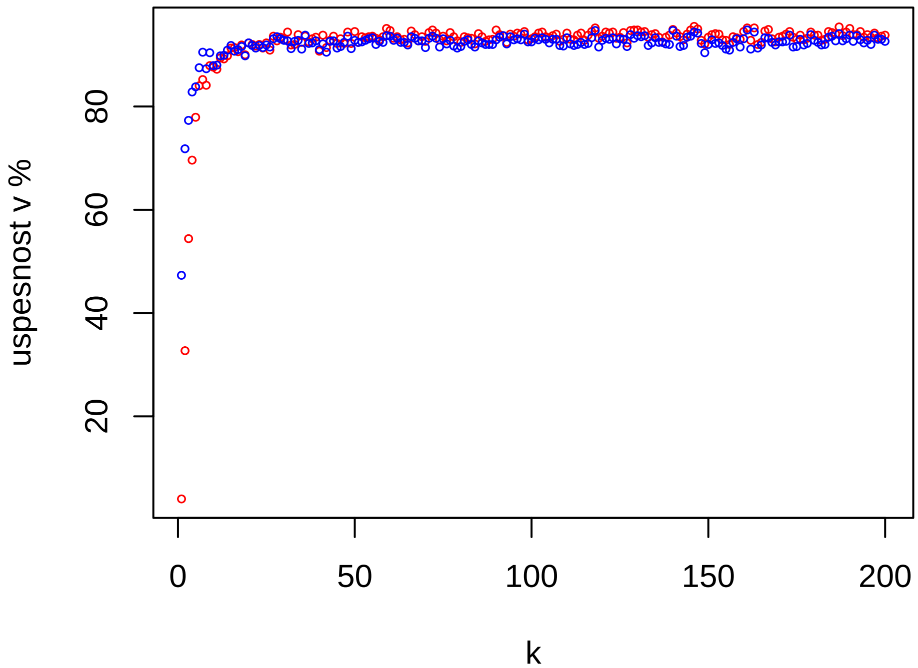
5.7 t rozdělení

V této části se zabýváme následujícími třemi situacemi. Náhodná veličina Y má ve všech třech rozdělení t_k a pro X platí nejprve $X \sim \mathsf{t}_k$, poté $X \sim \mathsf{t}_k + 0,5$ a nakonec $X \sim \mathsf{t}_k + 5$, $k = 1, \ldots, 200$. V odstavci 2.6 bylo uvedeno, že pro rozdělení t_1 neexistuje střední hodnota, jeho hustota je ale symetrická kolem bodu 0. Pro toto rozdělení přeformulujeme nulovou a alternativní hypotézu do tvaru $H_0 : \nu_1 = \nu_2$, $H_1 : \nu_1 \neq \nu_2$, kde ν_1 je bod, kolem nějž je symetrická hustota náhodných veličin $X_1, \ldots, X_{100}$, a ν_2 je bod, kolem nějž je symetrická hustota veličin $Y_1, \ldots, Y_{100}$. V části 2.6 jsme též zmínili, že rozdělení t_k konvergují k $N(0, 1)$ pro $k \to \infty$. Jelikož jedním z předpokladů dvouvýběrového t testu je normalita, očekávali bychom, že s rostoucím k se bude úspěšnost t testu zvyšovat a že pro malá k bude lepších výsledků dosahovat Wilcoxonův test. Pokud H_0 neplatí, toto očekávání je naplněno, viz obr. 5.14 a 5.15. V situaci ilustrované obrázkem 5.14 ($X \sim \mathsf{t}_k + 0,5$) je pro $k \geq 48$ t test lepší než Wilcoxonův test. Za platnosti H_0 se úspěšnost obou testů pohybuje kolem hladiny 95%, na hodnotě k nezávisí a Wilcoxonův i t test dosahují srovnatelných výsledků (viz obr. 5.13).



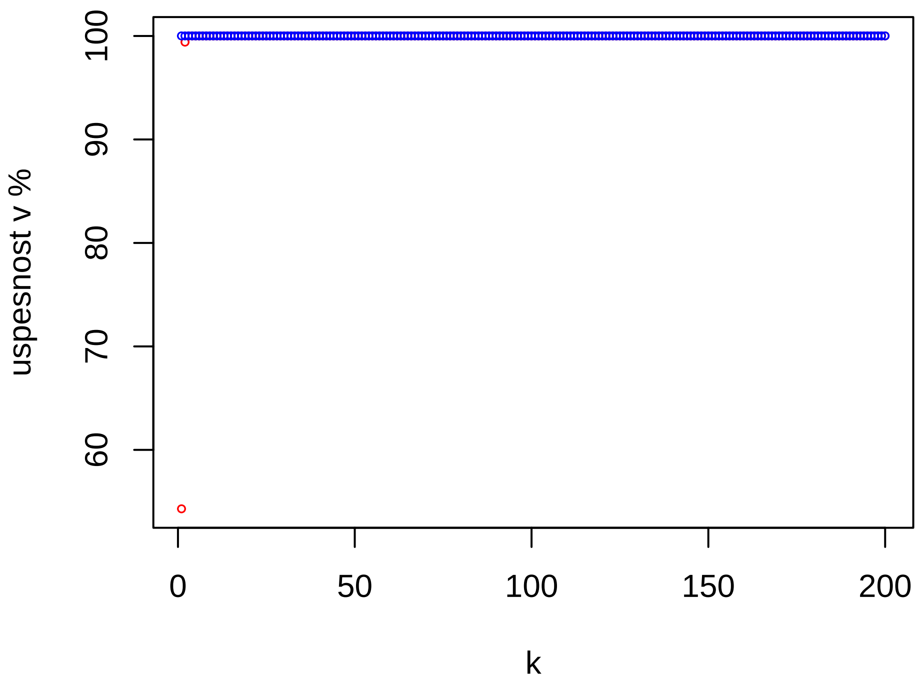
Obrázek 5.13: $X \sim \mathsf{t}_k$

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	t_1	t_1	100	100	0	98,7	95,4
50	t_{50}	t_{50}	100	100	0	94,6	95,0
62	t_{62}	t_{62}	100	100	0	93,2	94,4
83	t_{83}	t_{83}	100	100	0	95,4	97,1
100	t_{100}	t_{100}	100	100	0	94,6	94,2
150	t_{150}	t_{150}	100	100	0	95,0	95,1
190	t_{190}	t_{190}	100	100	0	93,8	92,9
200	t_{200}	t_{200}	100	100	0	94,7	95,4



Obrázek 5.14: $X \sim \mathsf{t}_k + 0,5$

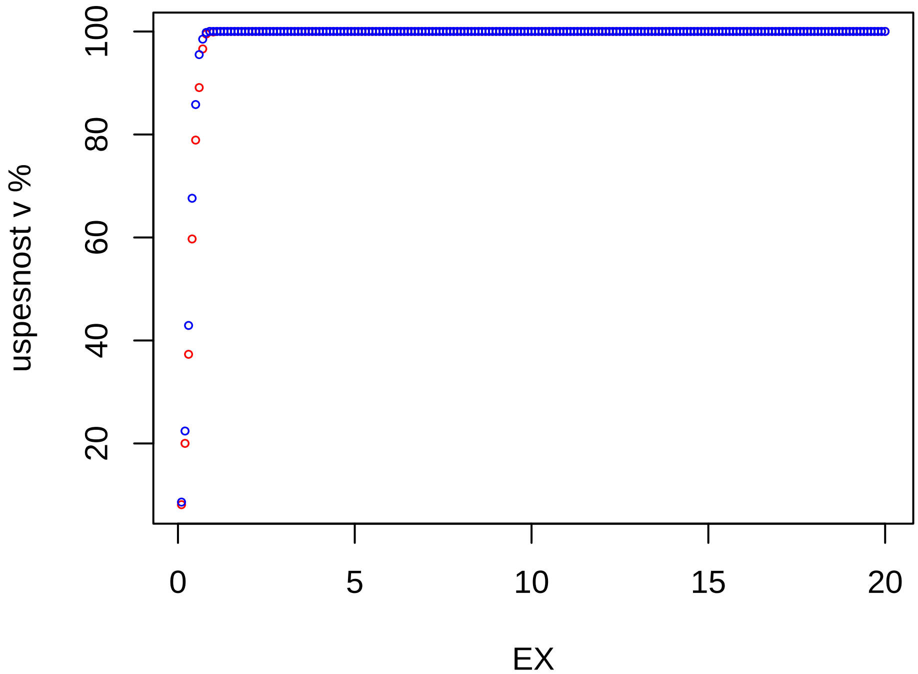
č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$\mathsf{t}_1 + 0,5$	t_1	100	100	0,5	4,0	47,3
2	$\mathsf{t}_2 + 0,5$	t_2	100	100	0,5	32,7	71,8
3	$\mathsf{t}_3 + 0,5$	t_3	100	100	0,5	54,4	77,3
146	$\mathsf{t}_{146} + 0,5$	t_{146}	100	100	0,5	95,5	94,4
161	$\mathsf{t}_{161} + 0,5$	t_{161}	100	100	0,5	95,2	94,8



Obrázek 5.15: $X \sim \mathsf{t}_k + 5$

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$\mathsf{t}_1 + 5$	t_1	100	100	5	54,3	100,0
2	$\mathsf{t}_2 + 5$	t_2	100	100	5	99,4	100,0
50	$\mathsf{t}_{50} + 5$	t_{50}	100	100	5	100,0	100,0
100	$\mathsf{t}_{100} + 5$	t_{100}	100	100	5	100,0	100,0
150	$\mathsf{t}_{150} + 5$	t_{150}	100	100	5	100,0	100,0
200	$\mathsf{t}_{200} + 5$	t_{200}	100	100	5	100,0	100,0

Nakonec jsem porovnávala $Y_1, \ldots, Y_{100}$ výběr z t_5 s výběrem $X_1, \ldots, X_{100}$ z rozdělení $\mathsf{t}_5 + \frac{i}{10}$, $i = 1, \ldots, 100$. Jak vidíme na obrázku 5.16, s rostoucím $\mathsf{E}\,X$ se úspěšnost obou testů zvyšuje, přičemž Wilcoxonův test je vždy buď stejně dobrý jako t test, nebo je dokonce lepší. Pro $EX \geq 1,1$ je úspěšnost obou testů stoprocentní.



Obrázek 5.16: $X \sim \mathsf{t}_k + \frac{i}{10}$, $Y \sim \mathsf{t}_k$

č. testu	rozdělení náh. veličiny		rozsah výběru		rozdíl středních hodnot	úspěšnost testů v %	
	X	Y	n	m	$\mu_1 - \mu_2$	t	W
1	$t_5 + 0,1$	t_5	100	100	0,1	8,1	8,6
4	$t_5 + 0,4$	t_5	100	100	0,4	59,7	67,6
6	$t_5 + 0,6$	t_5	100	100	0,6	89,1	95,5
7	$t_5 + 0,7$	t_5	100	100	0,7	96,6	98,5
9	$t_5 + 0,9$	t_5	100	100	0,9	100,0	100,0
50	$t_5 + 5$	t_5	100	100	5	100,0	100,0
100	$t_5 + 10$	t_5	100	100	10	100,0	100,0
150	$t_5 + 15$	t_5	100	100	15	100,0	100,0
200	$t_5 + 20$	t_5	100	100	20	100,0	100,0

Závěr

Cílem této práce bylo porovnat dvouvýběrový t test a Wilcoxonův test. Na simulovaných datech jsme zkoumali, jak se úspěšnosti těchto testů mění v závislosti na rozdělení dat, rozsazích výběrů a rozdílu mezi skutečnou a hypotetickou hodnotou výrazu $\mu_1 - \mu_2$.

Zjistili jsme, že pokud výběry pocházejí z normálního rozdělení, dosahuje lepších výsledků t test, a to i při porušení předpokladu shody rozptylů, který na data při používání t testu klademe.

U výběrů z dvojitě exponenciálního rozdělení dosahoval Wilcoxonův test lepších výsledků než t test, ale rozdíl mezi nimi nebyl příliš velký. V případě Cauchyova rozdělení byl t test výrazně horší než Wilcoxonův test.

Nakonec jsme se zabývali rozdělením t_k, o němž víme, že pro $k = 1$ je Cauchyovým rozdělením a pro $k \to \infty$ konverguje k rozdělení $N(0,1)$. Zjistili jsme, že pro malá k je u tohoto rozdělení lepší Wilcoxonův test a naopak, pro velká k lepších výsledků docílíme při užití t testu.

Při zkoumání závislosti úspěšnosti testů na rozsazích výběrů m a n jsme došli k závěru, že pokud H_0 platí, zamítají oba testy hypotézu v pěti procentech případů, jinými slovy, pravděpodobnost chyby prvního druhu je rovna očekávané hodnotě bez ohledu na rozsah výběru. Ukázalo se, že pokud H_0 neplatí, je situace poněkud odlišná, pro malé rozsahy výběrů je pravděpodobnost chyby druhého druhu velká u obou testů a klesá se zvyšujícími se hodnotami m, n. Dobrých výsledků testy dosahují pro $m > 50$, $n > 50$.

Pokud jde o závislost úspěšnosti testů na výrazu $|\mu_1 - \mu_2|$ (tj. na absolutní hodnotě rozdílu mezi skutečnou a hypotetickou hodnotou $\mu_1 - \mu_2$), ukázalo se, že s rostoucí hodnotou tohoto výrazu se pravděpodobnost zamítnutí neplatné hypotézy u obou testů zvyšuje, tedy pravděpodobnost chyby druhého druhu klesá.

Literatura

[1] Anděl, J.: Základy matematické statistiky. Matfyzpress, Praha 2005.

[2] Anděl, J.: Statistické metody. Matfzypress, Praha 2003.

[3] Dupač, V., Hušková, M.: Pravděpodobnost a matematická statistika. Nakladatelství Karolinum, Praha 2001.