4. Metody hledání bodových odhadů parametrů. Úvod do testování hypotéz.

4.1. Motivace: Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $L(\vartheta)$, které závisí na parametru ϑ. Úkolem je najít statistiku $T = T(X_1, ..., X_n)$, která nabývá hodnot blízkých parametru ϑ resp. parametrické funkci $h(\vartheta)$, ať je hodnota parametru ϑ jakákoliv. Seznámíme se se dvěma metodami hledání bodových odhadů, a to metodou maximální věrohodnosti a metodou momentů.

4.2. Definice: Definice maximálně věrohodného odhadu.
Nechť $X_1, ..., X_n$ je náhodný výběr z diskrétního rozložení (je popsáno pravděpodobnostní funkcí $\pi(x;\ \vartheta)$ resp. ze spojitého rozložení (je popsáno hustotou $\varphi(x;\ \vartheta)$).
Simultánní pravděpodobnostní funkce resp. simultánní hustota náhodného vektoru $(X_1, ..., X_n)$ je $\pi(x_1\ ;\ \vartheta)\ldots\pi(x_n\ ;\ \vartheta)$ resp. $\varphi\ (x_1;\ \vartheta)\ldots\varphi\ (x_n;\ \vartheta)$. Pro pevně dané $\mathbf{x} = (x_1, ..., x_n)$ zavedeme věrohodnostní funkci $L(\vartheta)=\prod_{i=1}^{n}\pi(x_i;\vartheta)$ v diskrétním případě resp. $L(\vartheta)=\prod_{i=1}^{n}\varphi(x_i;\vartheta)$ ve spojitém případě.
Statistika $\hat{\vartheta}(\mathbf{X})$, která má tu vlastnost, že $\forall\vartheta\in\Xi: L(\hat{\vartheta})\geq L(\vartheta)$, se nazývá maximálně věrohodný odhad parametru ϑ.
(Místo věrohodnostní funkce $L(\vartheta)$ používáme logaritmickou věrohodnostní funkci $\ln L(\vartheta)$.)

4.3. Definice: Definice věrohodnostních rovnic.
Nechť $\vartheta = (\vartheta_1,\ldots,\vartheta_k)$. Logaritmickou věrohodnostní funkci $\ln L(\vartheta)$ parciálně derivujeme podle $\vartheta_1,\ldots,\vartheta_k$ a derivace položíme rovny 0:

$$\frac{\partial \ln L(\vartheta_1,\ldots,\vartheta_k)}{\partial\vartheta_i}=0,\ i = 1, 2, \ldots, k.$$

Dostaneme systém věrohodnostních rovnic. Jeho řešením je maximálně věrohodný odhad parametru ϑ: $\hat{\vartheta}(X)=\left(\hat{\vartheta}_1(X),\ldots,\hat{\vartheta}_k(X)\right)$.

4.4. Příklad: (Maximálně věrohodný odhad v diskrétním skalárním případě)
Nechť $X_1, ..., X_n$ je náhodný výběr z alternativního rozložení $A(\vartheta)$. Metodou maximální věrohodnosti najděte odhad parametru ϑ.
Řešení:

$$X \sim A(\vartheta) \Rightarrow \pi(x) = \begin{cases}\vartheta^{x_i}(1-\vartheta)^{1-x_i} & \text{pro } x = 0, 1\\ 0 \text{ jinak}\end{cases}$$

Věrohodnostní funkce: $L(\vartheta)=\prod_{i=1}^{n}\vartheta^{x_i}(1-\vartheta)^{1-x_i}=\vartheta^{\sum_{i=1}^{n}x_i}(1-\vartheta)^{n-\sum_{i=1}^{n}x_i}$ pro $x_i = 0, 1$, $= 0$ jinak.

Logaritmická funkce věrohodnosti: $\ln L(\vartheta) = \ln\vartheta \cdot \sum_{i=1}^{n} x_i + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-\vartheta)$

Věrohodnostní rovnice:

$$\frac{d\ln L(\vartheta)}{d\vartheta} = \frac{\sum_{i=1}^{n} x_i}{\vartheta} - \frac{n - \sum_{i=1}^{n} x_i}{1-\vartheta} = 0 \Rightarrow$$

$$\Rightarrow (1-\vartheta)\sum_{i=1}^{n} x_i - \vartheta\left(n - \sum_{i=1}^{n} x_i\right) = 0 \Rightarrow$$

$$\Rightarrow \sum_{i=1}^{n} x_i - \vartheta\sum_{i=1}^{n} x_i - n\vartheta + \vartheta\sum_{i=1}^{n} x_i = 0 \Rightarrow \hat{\vartheta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Maximálně věrohodným odhadem parametru ϑ alternativního rozložení $A(\vartheta)$ je tedy statistika $\hat{\vartheta}(\mathbf{X}) = M$, tj. výběrový průměr.

4.5. Příklad: (Maximálně věrohodný odhad ve spojitém vektorovém případě)
Nechť $X_1, \ldots, X_n$ je náhodný výběr z normálního rozložení $N(\mu, \sigma^2)$. Metodou maximální věrohodnosti najděte odhad vektorového parametru $\boldsymbol{\vartheta} = (\mu, \sigma^2)$.

Řešení: $X \sim N(\mu, \sigma^2) \Rightarrow \varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Věrohodnostní funkce: $L(\boldsymbol{\vartheta}) = \prod_{i=1}^{n} \varphi(x_i;\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)}{\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}}$.

Logaritmická funkce věrohodnosti: $\ln L(\boldsymbol{\vartheta}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$.

Věrohodnostní rovnice:

$$\frac{\partial \ln L(\vartheta)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\frac{\partial \ln L(\vartheta)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

Z první rovnice plyne $\sum_{i=1}^{n} x_i - n\mu = 0 \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Maximálně věrohodným odhadem parametru μ je tedy statistika $\hat{\mu}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i = M$, tj. výběrový průměr.

Z druhé rovnice plyne $-n\sigma^2 + \sum_{i=1}^{n}(x_i - \mu)^2 = 0$. Za μ dosadíme odhad $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = m$ a získáme $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2$. Maximálně věrohodným odhadem parametru σ^2 je tedy statistika $\hat{\sigma}^2(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - M)^2$.

4.6. Definice: Definice momentového odhadu
Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozložení $L(\vartheta), \vartheta \in \Xi$. Předpokládáme, že existuje prvních k počátečních momentů $\mu_r' = E(X^r)$ rozložení $L(\vartheta)$, r = 1, 2, …, k. Označme $M_r' = \frac{1}{n}\sum_{i=1}^{n} X_i^r$ výběrové počáteční momenty, r = 1, 2, …, k. Statistika $\hat{\vartheta}(\mathbf{X})$, která je řešením systému momentových rovnic $\mu_r' = M_r'$, r = 1, 2, …, k, se nazývá momentový odhad parametru ϑ.

4.7. Příklad: Nechť $X_1, \ldots, X_n$ je náhodný výběr z geometrického rozložení $Ge(\vartheta)$. Metodou momentů najděte odhad parametru ϑ.

Řešení: X ~ $Ge(\vartheta)$ => $\pi(x)$ = $\begin{cases} (1-\vartheta)^x \vartheta \text{ pro } x = 0, 1, \ldots \\ 0 \text{ jinak} \end{cases}$. Lze odvodit, že $E(X) = \frac{1-\vartheta}{\vartheta}$.

Momentová rovnice: $\mu_1' = M_1'$, tj. $\frac{1-\vartheta}{\vartheta} = M \Rightarrow 1 - \vartheta = \vartheta M \Rightarrow \hat{\vartheta}(\mathbf{X}) = \frac{1}{1+M}$.

4.8. Motivace k testování hypotéz.
Častým úkolem statistika je na základě dat ověřit předpoklady o parametrech nebo typu rozložení, z něhož pochází náhodný výběr. Takovému předpokladu se říká nulová hypotéza. Nulová hypotéza vyjadřuje nějaký teoretický předpoklad, často skeptického rázu a uživatel ji musí stanovit předem, bez přihlédnutí k datovému souboru. Proti nulové hypotéze stavíme alternativní hypotézu, která říká, co platí, když neplatí nulová hypotéza. Alternativní hypotéza je formulována tak, aby mohla platit jenom jedna z těchto dvou hypotéz. Pravdivost alternativní hypotézy by znamenala objevení nějakých nových skutečností nebo zásadnější změnu v dosavadních představách.
Např. výzkumník by chtěl na základě dat prověřit tezi (nový objev), že pasivní kouření škodí zdraví. Jako nulovou hypotézu tedy položí tvrzení, že pasivní kouření neškodí zdraví a proti nulové hypotéze postaví alternativní, že pasivní kouření škodí zdraví.
Testováním hypotéz se myslí rozhodovací postup, který je založen na daném náhodném výběru a s jehož pomocí rozhodneme o zamítnutí či nezamítnutí nulové hypotézy.

4.9. Definice: Definice nulové a alternativní hypotézy.

Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $L(\vartheta)$, kde parametr $\vartheta \in \Xi$ neznáme. Nechť $h(\vartheta)$ je parametrická funkce a c daná reálná konstanta.

a) Oboustranná alternativa: Tvrzení H_0: $h(\vartheta) = c$ se nazývá jednoduchá nulová hypotéza. Proti nulové hypotéze postavíme složenou oboustrannou alternativní hypotézu H_1: $h(\vartheta) \neq c$.

b) Levostranná alternativa: Tvrzení H_0: $h(\vartheta) \geq c$ se nazývá složená pravostranná nulová hypotéza. Proti jednoduché nebo složené pravostranné nulové hypotéze postavíme složenou levostrannou alternativní hypotézu H_1: $h(\vartheta) < c$.

c) Pravostranná alternativa: Tvrzení H_0: $h(\vartheta) \leq c$ se nazývá složená levostranná nulová hypotéza. Proti jednoduché nebo složené levostranné nulové hypotéze postavíme složenou pravostrannou alternativní hypotézu H_1: $h(\vartheta) > c$.

Testováním H_0 proti H_1 rozumíme rozhodovací postup založený na náhodném výběru
$X_1, ..., X_n$, s jehož pomocí zamítneme či nezamítneme platnost nulové hypotézy.

(Volba alternativní hypotézy není libovolná, ale vyplývá z konkrétní situace. Např. při současné technologii je pravděpodobnost vyrobení zmetku $\vartheta = 0{,}01$.

a) Po rekonstrukci výrobní linky byla obnovena výroba, přičemž technologie zůstala stejná. Chceme ověřit, zda se změnila kvalita výrobků. Testujeme H_0: $\vartheta = 0{,}01$ proti H_1: $\vartheta \neq 0{,}01$.

b) Byly provedeny změny v technologii výroby s cílem zvýšit kvalitu. V tomto případě tedy testujeme H_0: $\vartheta = 0{,}01$ proti H_1: $\vartheta < 0{,}01$.

c) Byly provedeny změny v technologii výroby s cílem snížit náklady. V této situaci testujeme H_0: $\vartheta = 0{,}01$ proti H_1: $\vartheta > 0{,}01$.)

4.10. Definice: Definice chyby 1. a 2. druhu.

Při testování H_0 proti H_1 se můžeme dopustit jedné ze dvou chyb: chyba 1. druhu spočívá v tom, že H_0 zamítneme, ač ve skutečnosti platí a chyba 2. druhu spočívá v tom, že H_0 nezamítneme, ač ve skutečnosti neplatí. Situaci přehledně znázorňuje tabulka:

skutečnost	rozhodnutí	
	H_0 nezamítáme	H_0 zamítáme
H_0 platí	správné rozhodnutí	chyba 1. druhu
H_0 neplatí	chyba 2. druhu	správné rozhodnutí

Pravděpodobnost chyby 1. druhu se značí α a nazývá se hladina významnosti testu (většinou bývá α = 0,05, méně často 0,1 či 0,01). Pravděpodobnost chyby 2. druhu se značí β. Číslo $1-\beta$ se nazývá síla testu a vyjadřuje pravděpodobnost, že bude H_0 zamítnuta za předpokladu, že neplatí. Obvykle se snažíme, aby síla testu byla aspoň 0,8. Obě hodnoty, α i $1-\beta$, závisí na velikosti efektu, který se snažíme detekovat. Čím drobnější efekt, tím musí být větší rozsah náhodného výběru.

4.11. Poznámka: Testování nulové hypotézy proti alternativní hypotéze třemi způsoby.

Testování nulové hypotézy proti alternativní hypotéze lze provést pomocí kritického oboru, pomocí intervalu spolehlivosti nebo pomocí p-hodnoty.

4.12. Definice: Definice testového kritéria, oboru nezamítnutí, kritického oboru a kritických hodnot.

Statistika $T_0 = T_0(X_1, \ldots, X_n)$ se nazývá testovým kritériem. Množina všech hodnot, jichž může testové kritérium nabýt, se rozpadá na obor nezamítnutí nulové hypotézy (značí se V) a obor zamítnutí nulové hypotézy (značí se W a nazývá se též kritický obor). Tyto dva obory jsou odděleny kritickými hodnotami (pro danou hladinu významnosti α je lze najít ve statistických tabulkách).

4.13. Věta: Rozhodnutí o nulové hypotéze pomocí realizace testového kritéria v oboru nezamítnutí či v kritickém oboru.

Jestliže číselná realizace t_0 testového kritéria T_0 padne do kritického oboru W, pak nulovou hypotézu zamítáme na hladině významnosti α a znamená to skutečné vyvrácení testované hypotézy. Jestliže t_0 padne do oboru nezamítnutí V, pak jde o pouhé mlčení, které platnost nulové hypotézy jenom připouští.

4.14. Věta: Stanovení kritického oboru v případě oboustranné alternativy, levostranné alternativy, pravostranné alternativy.

Kritický obor v případě oboustranné alternativy má tvar

$$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right),$$

kde $K_{\alpha/2}(T)$ a $K_{1-\alpha/2}(T)$ jsou kvantily rozložení, jímž se řídí testové kritérium T_0, je-li nulová hypotéza pravdivá.
Kritický obor v případě levostranné alternativy má tvar:

$$W = \left(t_{min}, K_{\alpha}(T)\right\rangle.$$

Kritický obor v případě pravostranné alternativy má tvar:

$$W = \left\langle K_{1-\alpha}(T), t_{max}\right).$$

4.15. Poznámka: Doporučený postup při testování nulové hypotézy proti alternativní hypotéze pomocí kritického oboru.

- Stanovíme nulovou hypotézu a alternativní hypotézu. Přitom je vhodné zvolit jako alternativní hypotézu ten předpoklad, jehož přijetí znamená závažné opatření a mělo by k němu dojít jen s malým rizikem omylu.

- Zvolíme hladinu významnosti α. Zpravidla volíme $\alpha = 0{,}05$, méně často 0,1 nebo 0,01.

- Najdeme vhodné testové kritérium a na základě zjištěných dat vypočítáme jeho realizaci.

- Jestliže realizace testového kritéria padla do kritického oboru, nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu. V opačném případě nulovou hypotézu nezamítáme na hladině významnosti α.

- Na základě rozhodnutí, které jsme učinili o nulové hypotéze, učiníme nějaké konkrétní opatření, např. seřídíme obráběcí stroj.

(Při testování hypotéz musíme mít k dispozici odpovídající nástroje, nejlépe vhodný statistický software. Nemáme-li ho k dispozici, musíme znát příslušné vzorce. Dále potřebujeme statistické tabulky a kalkulačku.)

4.16. Věta: Testování nulové hypotézy proti alternativní hypotéze pomocí 100(1-α)% empirického intervalu spolehlivosti pro parametrickou funkci $h(\vartheta)$.

Sestrojíme 100(1-α)% empirický interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokryje-li tento interval hodnotu c, pak H_0 nezamítáme na hladině významnosti α, v opačném případě H_0 zamítáme na hladině významnosti α.

Pro test H_0 proti oboustranné alternativě sestrojíme oboustranný interval spolehlivosti.
Pro test H_0 proti levostranné alternativě sestrojíme pravostranný interval spolehlivosti.
Pro test H_0 proti pravostranné alternativě sestrojíme levostranný interval spolehlivosti.

4.17. Věta: Testování nulové hypotézy proti alternativní hypotéze pomocí p-hodnoty.
p-hodnota udává nejnižší možnou hladinu významnosti pro zamítnutí nulové hypotézy. Je to riziko, že bude zamítnuta H_0 za předpokladu, že platí (riziko planého poplachu). Jestliže p-hodnota ≤ α, pak H_0 zamítáme na hladině významnosti α, je-li p-hodnota > α, pak H_0 nezamítáme na hladině významnosti α.

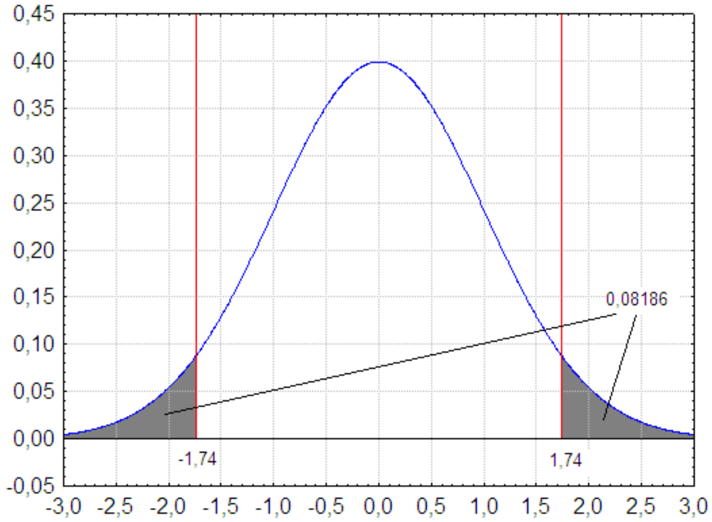
Způsob výpočtu p-hodnoty:

- Pro oboustrannou alternativu $p = 2 \min\{P(T_0 \leq t_0), P(T_0 \geq t_0)\}$.
- Pro levostrannou alternativu $p = P(T_0 \leq t_0)$.
- Pro pravostrannou alternativu $p = P(T_0 \geq t_0)$.

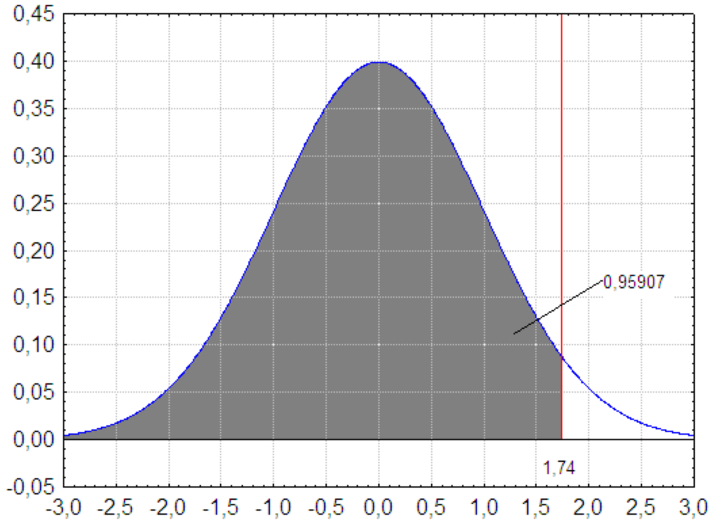
(p-hodnota vyjadřuje pravděpodobnost, s jakou číselné realizace $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ podporují H_0, je-li pravdivá. Statistické programové systémy poskytují ve svých výstupech p-hodnotu. Její výpočet vyžaduje znalost distribuční funkce rozložení, kterým se řídí testové kritérium T_0, je-li H_0 pravdivá. Vzhledem k tomu, že v běžných statistických tabulkách jsou uvedeny pouze hodnoty distribuční funkce standardizovaného normálního rozložení, bez použití speciálního software jsme schopni vypočítat p-hodnotu pouze pro test hypotézy o střední hodnotě normálního rozložení při známém rozptylu.)

4.18. Poznámka: Ilustrace významu p-hodnoty.

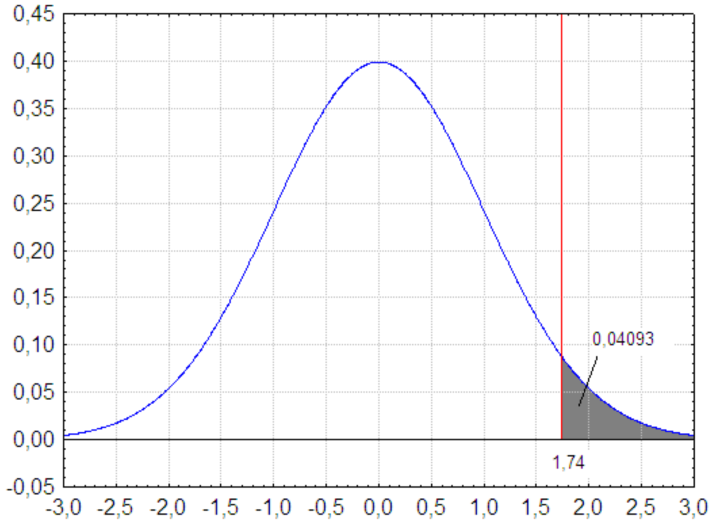
Oboustranný test



Levostranný test



Pravostranný test



4.19. Příklad: Nechť $X_1, \ldots, X_{400}$ je náhodný výběr z $N(\mu, 0{,}01)$. Je známo, že výběrový průměr se realizoval hodnotou 0,01. Na hladině významnosti 0,05 testujte hypotézu H_0: $\mu = 0$ proti pravostranné alternativě H_1: $\mu > 0$

a) pomocí intervalu spolehlivosti
b) pomocí kritického oboru
c) pomocí p-hodnoty.

Řešení:

ad a) Při testování nulové hypotézy proti pravostranné alternativě používáme levostranný interval spolehlivosti.

$$d = m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha} = 0{,}01 - \frac{0{,}1}{\sqrt{400}} u_{0{,}95} = 0{,}01 - \frac{0{,}1}{20} 1{,}64485 = 0{,}0018 .$$

Protože číslo c = 0 neleží v intervalu (0,0018; ∞), H_0 zamítáme na hladině významnosti 0,05.

ad b) Vypočteme realizaci testové statistiky:

$$t_0 = \frac{m - c}{\frac{\sigma}{\sqrt{n}}} = \frac{0{,}01 - 0}{\frac{0{,}1}{\sqrt{400}}} = \frac{0{,}01 \cdot 20}{0{,}1} = 2 .$$

Stanovíme kritický obor:

$$W = \langle u_{1-\alpha}, \infty) = \langle u_{0{,}95}, \infty) = \langle 1{,}64485, \infty)$$

Protože testová statistika se realizuje v kritickém oboru, H_0 zamítáme na hladině významnosti 0,05.

ad c) Při testování nulové hypotézy proti pravostranné alternativě se p-hodnota počítá podle vzorce: $p = P(T_0 \geq t_0)$. V našem případě: $p = P(T_0 \geq 2) = 1 - \Phi(2) = 1 - 0{,}97725 = 0{,}02275$. Protože p-hodnota je menší než hladina významnosti 0,05, H_0 zamítáme na hladině významnosti 0,05.