Úvod do testování hypotéz

Motivace: 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.

Nulová a alternativní hypotéza

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.

Chyba 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 $\alpha = 0{,}05$, 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.

skutečnost	rozhodnutí	
	zdravý	nemocný
jsem zdravý	zdravý a neléčený	zdravý a léčený
jsem nemocný	nemocný a neléčený	nemocný a léčený

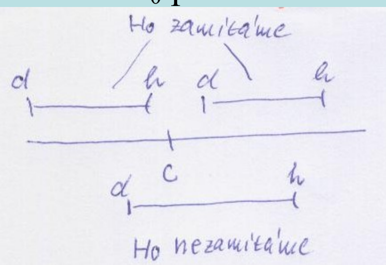
Testování pomocí kritického oboru

Najdeme statistiku $T_0 = T_0(X_1, ..., X_n)$, kterou nazveme 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).
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.
Pravděpodobnosti chyb 1. a 2. druhu nyní zapíšeme takto:
$P(T_0 \in W / H_0 \text{ platí}) = \alpha$, $P(T_0 \in V / H_1 \text{ platí}) = \beta$.

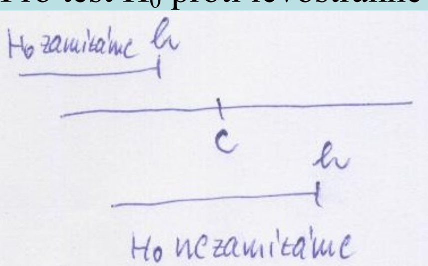
Stanovení kritického oboru pro danou hladinu významnosti α:
Označme t_{min} (resp. t_{max}) nejmenší (resp. největší) hodnotu testového kritéria.
Kritický obor v případě oboustranné alternativy má tvar
$W = \langle t_{min}, K_{\alpha/2}(T) \rangle \cup \langle K_{1-\alpha/2}(T), t_{max} \rangle$, kde $K_{\alpha/2}(T)$ 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:
$W = \langle t_{min}, K_{\alpha}(T) \rangle$.
Kritický obor v případě pravostranné alternativy má tvar:
$W = \langle K_{1-\alpha}(T), t_{max} \rangle$.

Testování pomocí intervalu spolehlivosti

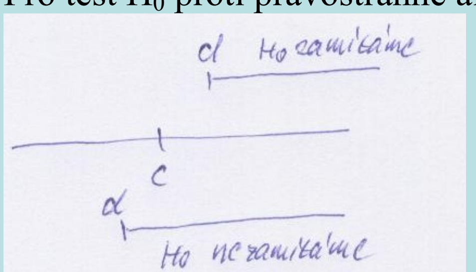
Sestrojíme 100(1-α)% empirický interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokryje-li tento interval hodnotu c, pak H_0 nezamítáme na hladině významnosti α, v opačném případě H_0 zamítáme na hladině významnosti α.
Pro test H_0 proti oboustranné alternativě sestrojíme oboustranný interval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



Pro test H_0 proti levostranné alternativě sestrojíme pravostranný interval spolehlivosti.



Pro test H_0 proti pravostranné alternativě sestrojíme levostranný interval spolehlivosti.



Testování 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 $\leq \alpha$, pak H_0 zamítáme na hladině významnosti α, je-li p-hodnota $> \alpha$, pak H_0 nezamítáme na hladině významnosti α.

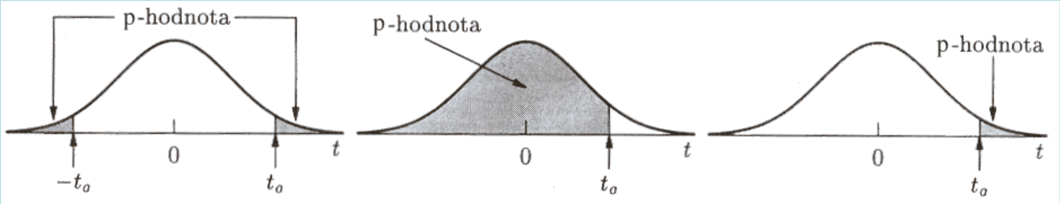
Způsob výpočtu p-hodnoty:

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 $p = P(T_0 \geq t_0)$.

Ilustrace významu p-hodnoty pro test nulové hypotézy proti oboustranné, levostranné a pravostranné alternativě:



(Zvonovitá křivka reprezentuje hustotu rozložení, kterým se řídí testové kritérium, je-li nulová hypotéza pravdivá.)

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.

Doporučený postup při testování hypotéz

1. 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.

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

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

4.
a) Testujeme-li pomocí kritického oboru, pak ho stanovíme. 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.
b) Testujeme-li pomocí intervalu spolehlivosti, vypočteme empirický $100(1-\alpha)\%$ interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokud číslo c padne do tohoto intervalu, nulovou hypotézu nezamítáme na hladině významnosti α. V opačném případě nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu.
c) Testujeme-li pomocí p-hodnoty, vypočteme ji a porovnáme ji s hladinou významnosti α. Jestliže $p \leq \alpha$, pak nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu. Je-li $p > \alpha$, pak nulovou hypotézu nezamítáme na hladině významnosti α.

5. Na základě rozhodnutí, které jsme učinili o nulové hypotéze, provedeme 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.)

Příklad: 10 x nezávisle na sobě byla změřena jistá konstanta μ. Výsledky měření byly: 2 1,8 2,1 2,4 1,9 2,1 2 1,8 2,3 2,2. Tyto výsledky považujeme za číselné realizace náhodného výběru $X_1, \ldots, X_{10}$ z rozložení $N(\mu, 0{,}04)$. Nějaká teorie tvrdí, že $\mu = 1{,}95$.

1. Oboustranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme oboustrannou alternativu H_1: $\mu \neq 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

$m = \frac{1}{10}(2 + \ldots + 2{,}2) = 2{,}06$, $\sigma^2 = 0{,}04$, $n = 10$, $\alpha = 0{,}05$, $c = 1{,}95$

a) Test provedeme pomocí kritického oboru.

Pro úlohy o střední hodnotě normálního rozložení při známém rozptylu používáme pivotovou statistiku $U = \frac{M - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

Testové kritérium tedy bude

$T_0 = \frac{M - c}{\frac{\sigma}{\sqrt{n}}}$ a bude mít rozložení N(0, 1), pokud je nulová hypotéza pravdivá. Vypočítáme realizaci testového kritéria: $t_0 = \frac{2{,}06 - 1{,}95}{\frac{0{,}2}{\sqrt{10}}} = 1{,}74$. Stanovíme kritický obor:

$$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right) = \left(-\infty, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) = \left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) = \left(-\infty, -u_{0{,}975}\right\rangle \cup \left\langle u_{0{,}975}, \infty\right) = \left(-\infty, -1{,}96\right\rangle \cup \left\langle 1{,}96, \infty\right).$$

Protože $1{,}74 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze 100(1-α)% empirického intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

$$(d, h) = \left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2},\ m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right).$$

V našem případě dostáváme:

$$d = 2{,}06 - \frac{0{,}2}{\sqrt{10}} u_{0{,}975} = 2{,}06 - \frac{0{,}2}{\sqrt{10}} .1{,}96 = 1{,}936,$$

$$h = 2{,}06 + \frac{0{,}2}{\sqrt{10}} u_{0{,}975} = 2{,}06 + \frac{0{,}2}{\sqrt{10}} .1{,}96 = 2{,}184.$$

Protože 1,95 $\in$ (1,936; 2,184), H_0 nezamítáme na hladině významnosti 0,05.

c) Test provedeme pomocí p-hodnoty.

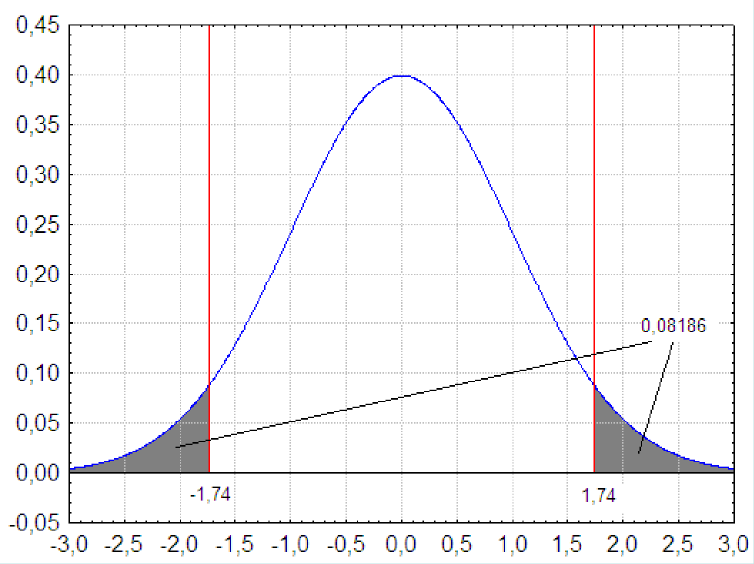
Protože proti nulové hypotéze stavíme oboustrannou alternativu, použijeme vzorec

$p = 2 \min\{P(T_0 \leq t_0), P(T_0 \geq t_0)\} = 2 \min \{P(T_0 \leq 1{,}74), P(T_0 \geq 1{,}74)\} =$

$= 2 \min \{ \Phi(1{,}74), 1 - \Phi(1{,}74) \} = 2 \min \{ 0{,}95907, 1 - 0{,}95907 \} = 0{,}08186.$

Jelikož 0,08186 > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Ilustrace významu p-hodnoty pro oboustranný test



2. Levostranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme levostrannou alternativu H_1: $\mu < 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

a) Test provedeme pomocí kritického oboru.

Na rozdíl od oboustranné alternativy bude mít kritický obor tvar

$W = \langle -\infty, u_\alpha) = \langle -\infty, u_{0,05}) = \langle -\infty, -1{,}645)$.

Protože $1{,}74 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze $100(1-\alpha)\%$ empirického pravostranného intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

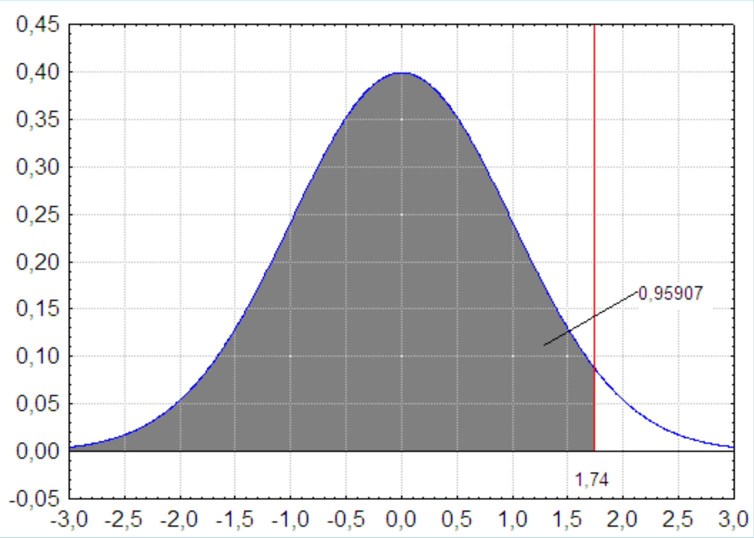
$(-\infty, h) = (-\infty, m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha})$.

V našem případě dostáváme: $h = 2{,}06 + \frac{0{,}2}{\sqrt{10}} u_{0{,}95} = 2{,}06 + \frac{0{,}2}{\sqrt{10}} .1{,}645 = 2{,}164$.

Protože $1{,}95 \in (-\infty; 2{,}164)$, H_0 nezamítáme na hladině významnosti 0,05.

c) Test provedeme pomocí p-hodnoty.
Protože proti nulové hypotéze stavíme levostrannou alternativu, použijeme vzorec
$p = P(T_0 \leq t_0) = \Phi(1{,}74) = 0{,}95907$.
Jelikož 0,95907 > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Ilustrace významu p-hodnoty pro levostranný test



3. Pravostranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme pravostrannou alternativu H_1: $\mu > 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

a) Test provedeme pomocí kritického oboru.

Na rozdíl od oboustranné alternativy bude mít kritický obor tvar

$W = \langle u_{1-\alpha}, \infty) = \langle u_{0,95}, \infty) = \langle 1{,}645, \infty)$.

Protože $1{,}74 \in W$, H_0 zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze $100(1-\alpha)\%$ empirického levostranného intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

$(d, \infty) = (m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha}, \infty)$.

V našem případě dostáváme: $d = 2{,}06 - \frac{0{,}2}{\sqrt{10}} u_{0,95} = 2{,}06 - \frac{0{,}2}{\sqrt{10}} .1{,}645 = 1{,}956$.

Protože $1{,}95 \notin (1{,}956, \infty)$, H_0 zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

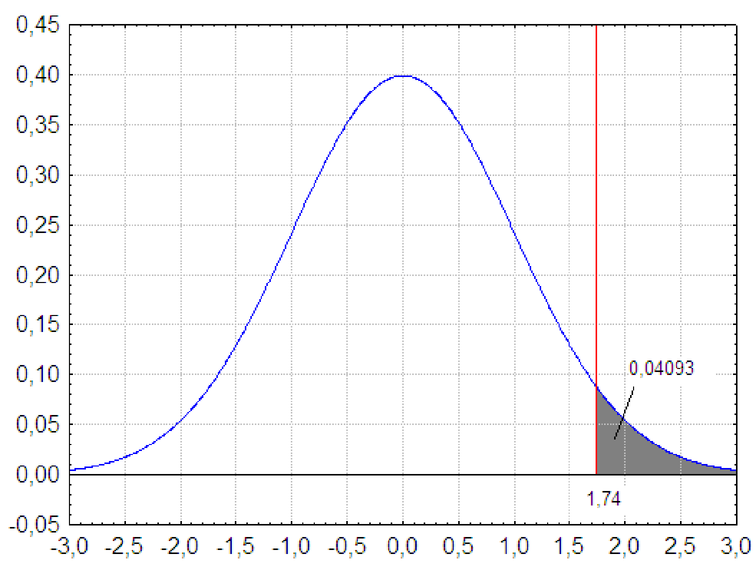
c) Test provedeme pomocí p-hodnoty.

Protože proti nulové hypotéze stavíme pravostrannou alternativu, použijeme vzorec

$p = P(T_0 \geq t_0) = 1 - \Phi(1{,}74) = 1 - 0{,}95907 = 0{,}04093$.

Jelikož $0{,}04093 \leq 0{,}05$, nulovou hypotézu zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

Ilustrace významu p-hodnoty pro pravostranný test



Parametrické úlohy o jednom náhodném výběru z normálního rozložení

Mnoho náhodných veličin, s nimiž se setkáváme ve výzkumu i praxi, se řídí normálním rozložením. Za jistých předpokladů obsažených v centrální limitní větě se dá rozložení jiných náhodných veličin aproximovat normálním rozložením. Proto je zapotřebí věnovat velkou pozornost právě náhodným výběrům
z normálního rozložení.

Rozložení statistik odvozených z výběrového průměru a rozptylu

Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $N(\mu, \sigma^2)$. Pak platí

a) $M \sim N(\mu, \frac{\sigma^2}{n})$, tedy $U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

(Pivotová statistika U slouží k řešení úloh o μ, když σ^2 známe.)

b) $K = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

(Pivotová statistika K slouží k řešení úloh o σ^2, když μ neznáme.)

c) $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$.

(Tato pivotová statistika slouží k řešení úloh o σ^2, když μ známe.)

d) $T = \frac{M-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$.

(Pivotová statistika T slouží k řešení úloh o μ, když σ^2 neznáme.)

Vysvětlení

ad a) Výběrový průměr M je lineární kombinace náhodných veličin s normálním rozložením, má tedy normální rozložení s parametry $E(M) = \mu$, $D(M) = \sigma^2/n$. Statistika U se získá standardizací M.

ad b) Vhodnou úpravou výběrového rozptylu S^2, kde použijeme obrat $X_i - M = (X_i - \mu) - (M - \mu)$, lze statistiku K vyjádřit jako součet kvadrátů n - 1 stochasticky nezávislých náhodných veličin se standardizovaným normálním rozložením. Tento součet se řídí rozložením $\chi^2(n-1)$.

ad c) Tato statistika je součet kvadrátů n stochasticky nezávislých náhodných veličin se standardizovaným normálním rozložením, řídí se tedy rozložením $\chi^2(n)$.

ad d) $U \sim N(0, 1)$, $K \sim \chi^2(n-1)$ jsou stochasticky nezávislé, protože M a S^2 jsou stochasticky nezávislé, tudíž statistika

$$T = \frac{U}{\sqrt{\frac{K}{n-1}}} = \frac{M - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

Příklad: Hmotnost balíčku krystalového cukru baleného na automatické lince se řídí normálním rozložením se střední hodnotou 1002 g a směrodatnou odchylkou 8 g. Kontrolor náhodně vybírá 9 balíčků z jedné série a zjišťuje, zda jejich průměrná hmotnost je alespoň 999 g. Pokud ne, podnik musí zaplatit pokutu 20 000 Kč. Jaká je pravděpodobnost, že podnik bude muset zaplatit pokutu?

Řešení:

$X \sim N(1002, 64)$, $M \sim N\left(1002, \frac{64}{9}\right)$

$$P(M \le 999) = P\left(\frac{M - 1002}{\sqrt{\frac{64}{9}}} \le \frac{999 - 1002}{\sqrt{\frac{64}{9}}}\right) = P\left(U \le -\frac{9}{8}\right) = \Phi\left(-\frac{9}{8}\right) = 1 - \Phi\left(\frac{9}{8}\right) = 1 - \Phi(1{,}125) = 1 - 0{,}87076 = 0{,}12924$$

Pravděpodobnost, že podnik bude platit pokutu, je asi 12,9%.

Řešení pomocí systému STATISTICA:

Využijeme toho, že STATISTICA pomocí funkce INormal(x;mu;sigma) umí vypočítat hodnotu distribuční funkce normálního rozložení se střední hodnotou mu a směrodatnou odchylkou sigma. Tedy $P(M \le 999) = \Phi(999)$, kde Φ je distribuční funkce rozložení N(1002, 64/9).

Otevřeme nový datový soubor o jedné proměnné a jednom případu. Dvakrát klikneme na název proměnné Prom1. Do Dlouhého jména této proměnné napíšeme = INormal(999;1002;8/3).

V proměnné Prom1 se objeví hodnota 0,130295.

Vzorce pro meze 100(1-α)% empirických intervalů spolehlivosti pro μ a σ^2

a) Interval spolehlivosti pro μ, když σ^2 známe (využití pivotové statistiky U)

Oboustranný: $(d, h) = \left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2},\ m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right)$

Levostranný: $(d, \infty) = \left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha},\ \infty\right)$

Pravostranný: $(-\infty, h) = \left(-\infty,\ m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha}\right)$

b) Interval spolehlivosti pro μ, když σ^2 neznáme (využití pivotové statistiky T)

Oboustranný: $(d, h) = \left(m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1),\ m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\right)$

Levostranný: $(d, \infty) = \left(m - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1),\ \infty\right)$

Pravostranný: $(-\infty, h) = \left(-\infty,\ m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1)\right)$

c) Interval spolehlivosti pro σ^2, když μ neznáme (využití pivotové statistiky K)

Oboustranný: $(d, h) = \left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)} \right)$

Levostranný: $(d, \infty) = \left(\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}, \infty \right)$

Pravostranný: $(-\infty, h) = \left(-\infty, \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)} \right)$

d) Interval spolehlivosti pro σ^2, když μ známe (využití pivotové statistiky $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma}$)

Oboustranný: $(d, h) = \left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}, \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\alpha/2}(n)} \right)$

Levostranný: $(d, \infty) = \left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\alpha}(n)}, \infty \right)$

Pravostranný: $(-\infty, h) = \left(-\infty, \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\alpha}(n)} \right)$

Příklad: 10 krát nezávisle na sobě byla změřena jistá konstanta μ. Výsledky měření byly: 2 1,8 2,1 2,4 1,9 2,1 2 1,8 2,3 2,2. Tyto výsledky považujeme za číselné realizace náhodného výběru $X_1, ..., X_{10}$ z rozložení $N(\mu, \sigma^2)$, kde parametry μ, σ^2 neznáme. Najděte 95% empirický interval spolehlivosti jak pro μ, tak pro σ^2 a to

a) oboustranný,
b) levostranný,
c) pravostranný.

Řešení: $m = 2{,}06$, $s^2 = 0{,}0404$, $s = 0{,}2011$, $\alpha = 0{,}05$, $t_{0,975}(9) = 2{,}2622$, $t_{0,95}(9) = 1{,}8331$, $\chi^2_{0,975}(9) = 19{,}023$, $\chi^2_{0,025}(9) = 2{,}7$, $\chi^2_{0,95}(9) = 16{,}919$, $\chi^2_{0,05}(9) = 3{,}325$

ad a) Oboustranný interval spolehlivosti pro střední hodnotu μ

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2{,}06 - \frac{0{,}2011}{\sqrt{10}} 2{,}2622 = 1{,}92$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2{,}06 + \frac{0{,}2011}{\sqrt{10}} 2{,}2622 = 2{,}20$$

$1{,}92 < \mu < 2{,}20$ s pravděpodobností aspoň 0,95.

Oboustranný interval spolehlivosti pro rozptyl σ^2

$$d = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} = \frac{9 \cdot 0{,}0404}{19{,}023} = 0{,}0191$$

$$h = \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)} = \frac{9 \cdot 0{,}0404}{2{,}7} = 0{,}1347$$

$0{,}0191 < \sigma^2 < 0{,}1347$ s pravděpodobností aspoň 0,95.

ad b) Levostranný interval spolehlivosti pro střední hodnotu μ

$$d = m - \frac{s}{\sqrt{n}}\, t_{1-\alpha}(n-1) = 2{,}06 - \frac{0{,}2011}{\sqrt{10}}\, 1{,}8331 = 1{,}94$$

$1{,}94 < \mu$ s pravděpodobností aspoň 0,95.

Levostranný interval spolehlivosti pro rozptyl σ^2

$$d = \frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)} = \frac{9 \cdot 0{,}0404}{16{,}919} = 0{,}0215$$

$\sigma^2 > 0{,}0215$ s pravděpodobností aspoň 0,95.

ad c) Pravostranný interval spolehlivosti pro střední hodnotu μ

$$h = m + \frac{s}{\sqrt{n}}\, t_{1-\alpha}(n-1) = 2{,}06 + \frac{0{,}2011}{\sqrt{10}}\, 1{,}8331 = 2{,}18$$

$\mu < 2{,}18$ s pravděpodobností aspoň 0,95.

Pravostranný interval spolehlivosti pro rozptyl σ^2

$$h = \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)} = \frac{9 \cdot 0{,}0404}{3{,}325} = 0{,}1094$$

$\sigma^2 < 0{,}1094$ s pravděpodobností aspoň 0,95.

Řešení pomocí systému STATISTICA:
Vytvoříme nový datový soubor o jedné proměnné X a 10 případech. Do proměnné X napíšeme dané hodnoty.
Statistika – Základní statistiky a tabulky – Popisné statistiky – OK – Proměnné X – OK – Detailní výsledky – zaškrtneme Meze spolehl. prům. a Meze sp. směr. odch. (ostatní volby zrušíme) – pro oboustranný 95% interval spolehlivosti ponecháme implicitní hodnotu pro Interval 95,00, pro jednostranné intervaly změníme hodnotu na 90,00.

Výsledky pro oboustranné 95% intervaly spolehlivosti pro střední hodnotu μ, pro směrodatnou odchylku σ a rozptyl σ^2:

Proměnná	Int. spolehl. -95,000%	Int. spolehl. 95,000	Spolehlivost Sm.Odch. -95,000%	Spolehlivost Sm.Odch. +95,000%	NProm1 =v3 ^2	NProm2 =v4 ^2
X	1,916136	2,203864	0,138329	0,367145	0,019135	0,134795

Vidíme, že
$1{,}92 < \mu < 2{,}20$ s pravděpodobností aspoň 0,95,
$0{,}1383 < \sigma < 0{,}3671$ s pravděpodobností aspoň 0,95.
$0{,}0191 < \sigma^2 < 0{,}1348$ s pravděpodobností aspoň 0,95.

Výsledky pro jednostranné 95% intervaly spolehlivosti pro střední hodnotu μ, pro směrodatnou odchylku σ a rozptyl σ^2:

Proměnná	Int. spolehl. -90,000%	Int. spolehl. 90,000	Spolehlivost Sm.Odch. -90,000%	Spolehlivost Sm.Odch. +90,000%	NProm1 =v3^2	NProm2 =v4^2
X	1,943421	2,176579	0,146678	0,330862	0,021514	0,10947

Vidíme, že

$\mu > 1,94$ s pravděpodobností aspoň 0,95,

$\mu < 2,20$ s pravděpodobností aspoň 0,95,

$\sigma > 0,1467$ s pravděpodobností aspoň 0,95,

$\sigma < 0,3309$ s pravděpodobností aspoň 0,95,

$\sigma^2 > 0,0215$ s pravděpodobností aspoň 0,95,

$\sigma^2 < 0,1095$ s pravděpodobností aspoň 0,95,

Jednotlivé typy testů pro parametry normálního rozložení

a) Nechť $X_1, ..., X_n$ je náhodný výběr $N(\mu, \sigma^2)$, kde σ^2 známe. Nechť $n \geq 2$ a c je konstanta. Test H_0: $\mu = c$ proti H_1: $\mu \neq c$ se nazývá jednovýběrový z-test.

b) Nechť $X_1, ..., X_n$ je náhodný výběr $N(\mu, \sigma^2)$, kde σ^2 neznáme. Nechť $n \geq 2$ a c je konstanta. Test H_0: $\mu = c$ proti H_1: $\mu \neq c$ se nazývá jednovýběrový t-test.

c) Nechť $X_1, ..., X_n$ je náhodný výběr $N(\mu, \sigma^2)$, kde μ neznáme. Nechť $n \geq 2$ a c je konstanta. Test H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 \neq c$ se nazývá test o rozptylu.

Provedení testů o parametrech μ, σ^2 pomocí kritického oboru

a) Provedení jednovýběrového z-testu

Vypočteme realizaci testového kritéria $t_0 = \dfrac{m - c}{\dfrac{\sigma}{\sqrt{n}}}$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1.

Oboustranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu \neq c$. Kritický obor má tvar: $W = \left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right)$.

Levostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu < c$. Kritický obor má tvar: $W = \left(-\infty, -u_{1-\alpha}\right\rangle$.

Pravostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu > c$. Kritický obor má tvar: $W = \left\langle u_{1-\alpha}, \infty\right)$.

b) Provedení jednovýběrového t-testu

Vypočteme realizaci testového kritéria $t_0 = \dfrac{m - c}{\dfrac{s}{\sqrt{n}}}$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1.

Oboustranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu \neq c$. Kritický obor má tvar: $W = \left(-\infty, -t_{1-\alpha/2}(n-1)\right\rangle \cup \left\langle t_{1-\alpha/2}(n-1), \infty\right)$.

Levostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu < c$. Kritický obor má tvar: $W = \left(-\infty, -t_{1-\alpha}(n-1)\right\rangle$.

Pravostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu > c$. Kritický obor má tvar: $W = \left\langle t_{1-\alpha}(n-1), \infty\right)$.

c) Provedení testu o rozptylu

Vypočteme realizaci testového kritéria $t_0 = \frac{(n-1)s^2}{c}$. Stanovíme kritický obor W. Pokud $t_0 \in$ W, H_0 zamítáme na hladině významnosti α a přijímáme H_1.

Oboustranný test: Testujeme H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 \neq c$. Kritický obor má tvar:.

$$W = \left\langle 0, \chi^2_{\alpha/2}(n-1) \right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty \right)$$

Levostranný test: Testujeme H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 < c$. Kritický obor má tvar: $W = \left\langle 0, \chi^2_{\alpha}(n-1) \right\rangle$.

Pravostranný test: Testujeme H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 > c$. Kritický obor má tvar: $W = \left\langle \chi^2_{1-\alpha}(n-1), \infty \right)$.

Příklad: Podle údajů na obalu čokolády by její čistá hmotnost měla být 125 g. Výrobce dostal několik stížností od kupujících, ve kterých tvrdili, že hmotnost čokolád je nižší než deklarovaných 125 g. Z tohoto důvodu oddělení kontroly náhodně vybralo 50 čokolád a zjistilo, že jejich průměrná hmotnost je 122 g a směrodatná odchylka 8,6 g. Za předpokladu, že hmotnost čokolád se řídí normálním rozložením, můžeme na hladině významnosti 0,01 považovat stížnosti kupujících za oprávněné?

Řešení: $X_1, \ldots, X_{50}$ je náhodný výběr z $N(\mu, \sigma^2)$. Testujeme hypotézu
H_0: $\mu = 125$ proti levostranné alternativě H_1: $\mu < 125$. Protože neznáme rozptyl σ^2, použijeme jednovýběrový t-test.

Testové kritérium $\dfrac{m - c}{\frac{s}{\sqrt{n}}} = \dfrac{122 - 125}{\frac{8,6}{\sqrt{50}}} = -2,4667$.

Kritický obor $W = \left(-\infty, -t_{1-\alpha}(n-1)\right\rangle = \left(-\infty, -t_{0,99}(49)\right\rangle = \left(-\infty, -2,4049\right\rangle$.

Jelikož testové kritérium se realizuje v kritickém oboru, zamítáme nulovou hypotézu na hladině významnosti 0,01. Stížnosti kupujících tedy lze považovat za oprávněné.

Výpočet pomocí systému STATISTICA:

Statistiky – Základní statistiky a tabulky – Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – zaškrtneme Výběrový průměr vs. Střední hodnota a zvolíme jednostr. – do políčka Pr1 napíšeme 122, do políčka SmOd1 napíšeme 8,6, do políčka N1 napíšeme 50, do políčka Pr2 napíšeme 125 - Výpočet. Dostaneme p-hodnotu 0,0086, tedy zamítáme nulovou hypotézu na hladině významnosti 0,01

Náhodný výběr z dvourozměrného rozložení

Nechť $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ je náhodný výběr z dvourozměrného rozložení, přičemž $n \geq 2$. Označíme $\mu = \mu_1 - \mu_2$ a zavedeme rozdílový náhodný výběr $Z_1 = X_1 - Y_1, \dots, Z_n = X_n - Y_n$, o němž předpokládáme, že se řídí normálním rozložením.

Vypočteme $M = \frac{1}{n}\sum_{i=1}^{n} Z_i$, $S^2 = \frac{1}{n}\sum_{i=1}^{n} (Z_i - M)^2$.

Vzorec pro meze 100(1-α)% empirického intervalu spolehlivosti pro střední hodnotu rozdílového náhodného výběru

Oboustranný: $(d, h) = (m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1))$

Levostranný: $(d, \infty) = (m - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1), \infty)$

Pravostranný: $(-\infty, h) = (-\infty, m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1))$

Příklad: Dvěma rozdílnými laboratorními metodami se zjišťoval obsah chemické látky v roztoku (v procentech). Bylo vybráno 5 vzorků a proměřeno oběma metodami. Výsledky měření jsou obsaženy v tabulce:

číslo vzorku	1	2	3	4	5
1. metoda	2,3	1,9	2,1	2,4	2,6
2. metoda	2,4	2,0	2,0	2,3	2,5

Za předpokladu, že data mají normální rozložení, sestrojte 90% empirický interval spolehlivosti pro rozdíl středních hodnot výsledků obou metod.

Řešení:

Přejdeme k rozdílovému náhodnému výběru, jehož realizace jsou: -0,1 -0,1 0,1 0,1 0,1. Vypočteme m = 0,02, s^2 = 0,012, s = 0,109545. Předpokládáme, že tato data pocházejí z normálního rozložení N(μ, σ^2). Vypočteme meze 90% oboustranného intervalu spolehlivosti pro μ při neznámém σ:

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 0{,}02 - \frac{0{,}109545}{\sqrt{5}} t_{0,95}(4) = 0{,}02 - \frac{0{,}109545}{\sqrt{5}} 2{,}1318 = -0{,}0844$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 0{,}02 + \frac{0{,}109545}{\sqrt{5}} t_{0,95}(4) = 0{,}02 + \frac{0{,}109545}{\sqrt{5}} 2{,}1318 = 0{,}1244$$

-0,0844 < μ < 0,1244 s pravděpodobností aspoň 0,9.

Výpočet pomocí systému STATISTICA:

Vytvoříme nový datový soubor o 3 proměnných a 5 případech. Do 1. proměnné X napíšeme hodnoty pro 1. metodu, do 2. proměnné Y hodnoty pro 2. metodu a do 3. proměnné Z rozdíly mezi X a Y.

Statistiky – Základní statistiky a tabulky – Popisné statistiky, OK - Proměnné Z, Detailní výsledky – zaškrtneme Meze spolehl. Prům. – Interval 90% - Výpočet. Dostaneme tabulku:

	Popisné statistiky (chemicka latka)	
Proměnná	Int. spolehl. -90,000%	Int. spolehl. 90,000
Z	-0,084439	0,124439

Vidíme tedy, že $-0{,}0844 < \mu < 0{,}1244$ s pravděpodobností aspoň 0,9.

Párový t-test

Nechť $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \ldots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ je náhodný výběr z rozložení $N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$, přičemž $n \geq 2$. Testujeme H_0: $\mu_1 - \mu_2 = c$ (tj. $\mu = c$) proti H_1: $\mu_1 - \mu_2 \neq c$ (tj. $\mu \neq c$) nebo testujeme nulovou hypotézu proti jedné z jednostranných alternativ. Tento test se nazývá párový t-test.

Provedení párového t-testu

Vypočteme realizaci testového kritéria $t_0 = \dfrac{m - c}{\frac{s}{\sqrt{n}}}$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1.

Oboustranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu \neq c$. Kritický obor má tvar: $W = \left(-\infty, -t_{1-\alpha/2}(n-1)\right\rangle \cup \left\langle t_{1-\alpha/2}(n-1), \infty\right)$.

Levostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu < c$. Kritický obor má tvar: $W = \left(-\infty, -t_{1-\alpha}(n-1)\right\rangle$.

Pravostranný test: Testujeme H_0: $\mu = c$ proti H_1: $\mu > c$. Kritický obor má tvar: $W = \left\langle t_{1-\alpha}(n-1), \infty\right)$.

Příklad: V následující tabulce jsou údaje o výnosnosti dosažené 12 náhodně vybranými firmami při investování do mezinárodního podnikání (veličina X) a do domácího podnikání (veličina Y):

č.firmy	1	2	3	4	5	6	7	8	9	10	11	12
X	10	12	14	12	12	17	9	15	9	11	7	15
Y	11	14	15	11	13	16	10	13	11	17	9	19

(Výnosnost je vyjádřena v procentech a představuje podíl na zisku vložených investic za rok.)
Za předpokladu, že data pocházejí z dvourozměrného rozložení a jejich rozdíl se řídí normálním rozložením, na hladině významnosti 0,1 testujte hypotézu, že neexistuje rozdíl mezi střední hodnotou výnosnosti investic do mezinárodního a domácího podnikání proti oboustranné alternativě.
Testování proveďte
a) pomocí intervalu spolehlivosti, b) pomocí kritického oboru.
(Pro úsporu času známe realizace výběrového průměru $m = -1,\overline{3}$ a výběrového rozptylu $s^2 = 4,\overline{78}$ rozdílového náhodného výběru $Z_i = X_i - Y_i$, i = 1, …, 12.)

Řešení:

Testujeme H_0: $\mu = 0$ proti H_1: $\mu \neq 0$
ad a) 90% interval spolehlivosti pro střední hodnotu μ při neznámém rozptylu σ^2 má meze:

$$d = m - \frac{s}{\sqrt{n}} t_{0,95}(n-1) = -1,\overline{3} - \frac{\sqrt{4,\overline{78}}}{\sqrt{12}} 1,7959 = -2,4677$$

$$h = m + \frac{s}{\sqrt{n}} t_{0,95}(n-1) = -1,\overline{3} + \frac{\sqrt{4,\overline{78}}}{\sqrt{12}} 1,7959 = -0,1989$$

Protože číslo c = 0 neleží v intervalu (-2,4677; -0,1989), H_0 zamítáme na hladině významnosti 0,1.

ad b) Vypočítáme realizaci testové statistiky $t_0 = \dfrac{m - c}{\frac{s}{\sqrt{n}}} = \dfrac{-1,\overline{3}}{\frac{\sqrt{4,\overline{78}}}{\sqrt{12}}} = -2,11085$

Stanovíme kritický obor $W = (-\infty, -t_{0,95}(11)\rangle \cup \langle t_{0,95}(11), \infty) = (-\infty, -1,7959\rangle \cup \langle 1,7959, \infty)$

Protože testová statistika se realizuje v kritickém oboru, H_0 zamítáme na hladině významnosti 0,1.

Výpočet pomocí systému STATISTICA:

Vytvoříme nový datový soubor o 2 proměnných a 12 případech. Do 1. proměnné X napíšeme hodnoty pro mezinárodní podnikání, do 2. proměnné hodnoty pro domácí podnikání.

Statistiky – Základní statistiky a tabulky – t-test pro závislé vzorky, OK - Proměnné X, Y – OK – Výpočet. Dostaneme tabulku:

	t-test pro závislé vzorky (investovani) Označ. rozdíly jsou významné na hlad. p < ,05000							
Proměnná	Průměr	Sm.odch.	N	Rozdíl	Sm.odch. rozdílu	t	sv	p
X	11,91667	2,937480						
Y	13,25000	3,048845	12	-1,33333	2,188122	-2,11085	11	0,058490

Vypočtenou p-hodnotu 0,05849 porovnáme se zvolenou hladinou významnosti $\alpha = 0{,}1$. Protože $p \leq \alpha$, zamítáme nulovou hypotézu na hladině významnosti 0,1.