Parametrické úlohy o jednom náhodném výběru z normálního rozložení

Motivace:

K nejčastěji používaným statistickým metodám patří konstrukce intervalů spolehlivosti pro parametry normálního rozložení či testování hypotéz o těchto parametrech. Normální rozložení je charakterizováno dvěma parametry – střední hodnotou μ a rozptylem σ^2. Budeme tedy řešit úlohy, které se týkají těchto dvou parametrů. K tomu slouží např. jednovýběrový z-test, t-test či test o rozptylu. Můžeme také mít k dispozici náhodný výběr z dvourozměrného rozložení s vektorem středních hodnot $\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$ a naším úkolem bude posoudit rozdílnost středních hodnot $\mu_1 - \mu_2$. K řešení tohoto problému slouží párový t-test.

Osnova:

- rozložení statistik odvozených z výběrového průměru a výběrového rozptylu

- vzorce pro meze intervalů spolehlivosti pro střední hodnotu a rozptyl

- jednotlivé typy testů pro parametry normálního rozložení

(z-test, jednovýběrový t-test, test o rozptylu, párový t-test)

Rozložení statistik odvozených z výběrového průměru a výběrového rozptylu

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

a) $M \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, 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 ~ N(1002, 64), M ~ $N\left(1002, \frac{64}{9}\right)$

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

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

Pravostranný: (-∞, 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 $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$)

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

Pravostranný: (-∞, 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

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2{,}06 - \frac{0{,}201}{\sqrt{10}} 2{,}2622 = 1{,}92$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2{,}06 + \frac{0{,}201}{\sqrt{10}} 2{,}2622 = 2{,}20$$

$1{,}92 < \mu < 2{,}20$ 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

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) = 2{,}06 - \frac{0{,}201}{\sqrt{10}} 1{,}8331 = 1{,}94$$

1,94 < μ s pravděpodobností aspoň 0,95.

Levostranný interval spolehlivosti pro rozptyl σ^2

$$d = \frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)} = \frac{9 \cdot 0{,}0404}{16{,}919} = 0{,}021$$

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

ad c) Pravostranný interval spolehlivosti pro střední hodnotu μ

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) = 2{,}06 + \frac{0{,}201}{\sqrt{10}} 1{,}8331 = 2{,}18$$

μ < 2,18 s pravděpodobností aspoň 0,95.

Pravostranný interval spolehlivosti pro rozptyl σ^2

$$h = \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)} = \frac{9 \cdot 0{,}0404}{3{,}325} = 0{,}109$$

$\sigma^2 < 0{,}1094$ 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:

Promě	Int. spol -95,000	Int. spol 95,000	Spolehliv Sm.Odc -95,000	Spolehliv Sm.Odc +95,000	NPror =v32	NPror =v42
X	1,916	2,203	0,138	0,367	0,019	0,134

Vidíme, že
1,92 < μ < 2,20 s pravděpodobností aspoň 0,95,
0,1383 < σ < 0,3671 s pravděpodobností aspoň 0,95.
0,0191 < σ^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:

Proměr	Int. spol -90,00(	Int. spol 90,00(	Spolehliv Sm.Odc -90,00(	Spolehliv Sm.Odc +90,00(	NPron =v3^	NPron =v4^
X	1,943	2,176	0,146	0,330	0,021	0,109

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 $t_0 = \frac{m -}{\frac{\sigma}{\sqrt{n}}}$. Stanovíme kritický obor W. Pokud t_0 W, H_0 zamítáme na hladině významnosti α a přijímáme H_1.

Oboustranný test: Testujeme H_0: μ = c proti H_1: μ c. Kritický obor má tvar: $W = \ldots, u_{1-\ldots/2}\rangle \cup u_{1-\ldots/2}, \infty$.

Levostranný test: Testujeme H_0: μ = c proti H_1: μ < c. Kritický obor má tvar: $W = \ldots, u_{1-\ldots}$.

Pravostranný test: Testujeme H_0: μ = c proti H_1: μ > c. Kritický obor má tvar: $W = u_{1-}, \infty$.

b) Provedení jednovýběrového t-testu

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

Oboustranný test: Testujeme H_0: μ = c proti H_1: μ c. Kritický obor má tvar: $W = \ldots, t_{1-\ldots/2}(n-1) \cup t_{1-\ldots/2}(n-1), \infty$.

Levostranný test: Testujeme H_0: μ = c proti H_1: μ < c. Kritický obor má tvar: $W = \ldots n-$.

Pravostranný test: Testujeme H_0: μ = c proti H_1: μ > c. Kritický obor má tvar: $W = t_{1-}\ n-1, \infty$.

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:.

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

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

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

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: $X_1, ..., X_{50}$ 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 $\dfrac{m-125}{\frac{s}{\sqrt{n}}} = \dfrac{122-125}{\frac{8,6}{\sqrt{50}}} = -2,466$.

Kritický obor $W = (-\infty, -t_{0,99}(n-1)\rangle = (-\infty, -t_{0,99}(49)\rangle = (-\infty, -2,404\rangle$.

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

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

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

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

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

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

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 $N(\mu, \sigma^2)$. Vypočteme meze 90% oboustranného intervalu spolehlivosti pro μ při neznámém σ:

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

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

-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 (ch	
	Int. spol	Int. spol
Proměn	-90,000	90,000
Z	-0,084	0,124

Vidíme tedy, že -0,0844 < μ < 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 = … a výběrového rozptylu $s^2 = 4,78$ 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:

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

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

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 $t_0 = \frac{m - c}{\frac{s}{\sqrt{n}}} = \frac{\ldots}{\frac{\sqrt{4,78}}{\sqrt{12}}} = \ldots 108$

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

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:

Proměı	t-test pro závislé vzorky (investování) Označ. rozdíly jsou významné na hlad. p < ,0 Průmı	Sm.od	N	Rozd	Sm.od rozdílı	t	sv	p
X	11,91	2,937						
Y	13,25	3,048	1:	-1,33:	2,188	-2,11(	1	0,058

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.