Kapitola 4.: Úlohy o jednom náhodném výběru z normálního rozložení

4.1. Motivace

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.

Normální rozložení je charakterizováno dvěma parametry – střední hodnotou μ a rozptylem σ^2. Budeme tedy řešit úlohy, které se týkají těchto parametrů. Jedná se především o jednovýběrový t-test či test o rozptylu. Seznámíme se rovněž se situací, kdy máme k dispozici jeden náhodný výběr z dvourozměrného normálního rozložení a posuzujeme rozdílnost středních hodnot obou náhodných veličin. K řešení tohoto problému slouží párový t-test.

4.2. Rozložení statistik odvozených z výběrového průměru a výběrového rozptylu

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

a) Výběrový průměr M a výběrový rozptyl S^2 jsou stochasticky nezávislé.

b) $M \sim N(\mu, \frac{\sigma^2}{n})$, 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.)

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

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

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

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

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

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

4.2.1. Příklad

Na výrobní lince jsou automaticky baleny balíčky rýže o deklarované hmotnosti 1000 g. Působením náhodných vlivů hmotnost balíčků kolísá. Lze ji považovat za náhodnou veličinu, která se řídí normálním rozložením se střední hodnotou 996 g a směrodatnou odchylkou 18 g. Jaká je pravděpodobnost, že náhodně vybraný balíček rýže neprojde výstupní kontrolou, jestliže je povolená tolerance ± 30 g od deklarované hmotnosti 1000 g?

Řešení: $X \sim N(996, 18^2)$, $U = \frac{X-996}{18} \sim N(0,1)$

$$P\left(X \notin \langle 970{,}1030\rangle\right) = 1 - P(970 < X < 1030) = 1 - P\left(\frac{970-996}{18} < U < \frac{1030-996}{18}\right) =$$

$$= 1 - \Phi(1{,}89) + \Phi(-1{,}44) = 2 - 0{,}971 - 0{,}925 = 0{,}104$$

4.3. Intervaly spolehlivosti pro parametry μ, σ^2

V kapitole 1 jsme se seznámili s pojmem intervalu spolehlivosti pro parametrickou funkci $h(\vartheta)$. Nyní se budeme zabývat speciálními případy, kdy za parametrickou funkci $h(\vartheta)$ považujeme střední hodnotu μ nebo rozptyl σ^2 normálního rozložení. V příkladu 1.3.5. jsme si ukázali způsob, jak zkonstruovat interval spolehlivosti pro střední hodnotu μ, když rozptyl σ^2 známe. Odvození intervalu spolehlivosti pro další tři situace (tj. pro μ, když σ^2 neznáme, pro σ^2, když μ neznáme a konečně pro σ^2, když μ známe) provádět nebudeme, uvedeme jen přehled vzorců pro meze $100(1-\alpha)\%$ empirických intervalů spolehlivosti pro tyto parametry .

4.3.1. Přehled vzorců

a) Interval spolehlivosti pro μ, když σ^2 známe

(využití pivotové statistiky $U = \dfrac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$)

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

(Odvození: viz příklad 1.3.5.)

b) Interval spolehlivosti pro μ, když σ^2 neznáme

(využití pivotové statistiky $T = \dfrac{M-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$)

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

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

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

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

d) Interval spolehlivosti pro σ^2, když μ známe

(využití pivotové statistiky $\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$)

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

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

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

4.3.2. 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 pro μ, a to

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

Řešení:
Vypočteme realizaci výběrového průměru: m = 2,06, výběrového rozptylu: $s^2 = 0{,}0404$ a výběrové směrodatné odchylky: s = 0,2011. Riziko α je 0,05. Jde o situaci popsanou v bodě (b), kde využíváme pivotovou statistiku T, která se řídí Studentovým rozložením t(9). V tabulkách najdeme kvantil $t_{0,975}(9) = 2{,}2622$ pro oboustranný interval spolehlivosti a kvantil $t_{0,95}(9) = 1{,}8331$ pro jednostranné intervaly spolehlivosti.

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

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

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

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

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

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

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

4.4. Testování hypotéz o parametrech μ, σ^2

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 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.

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

V kapitole 1 byly uvedeny tři způsoby testování hypotéz – pomocí kritického oboru, pomocí intervalu spolehlivosti a pomocí p-hodnoty. V tomto odstavci si ukážeme, jak testovat hypotézy o střední hodnotě μ a rozptylu σ^2 pomocí kritického oboru.

a) Provedení z-testu

H_0: $\mu = c$ proti H_1: $\mu \neq c$ (resp. H_1: $\mu < c$ resp. H_1: $\mu > c$) zamítáme na hladině významnosti α,

jestliže $\left|\dfrac{m-c}{\frac{\sigma}{\sqrt{n}}}\right| \geq u_{1-\alpha/2}$ (resp. $\dfrac{m-c}{\frac{\sigma}{\sqrt{n}}} \leq -u_{1-\alpha}$ resp. $\dfrac{m-c}{\frac{\sigma}{\sqrt{n}}} \geq u_{1-\alpha}$).

b) Provedení jednovýběrového t-testu

H_0: $\mu = c$ proti H_1: $\mu \neq c$ (resp. H_1: $\mu < c$ resp. H_1: $\mu > c$) zamítáme na hladině významnosti α,

jestliže $\left|\dfrac{m-c}{\frac{s}{\sqrt{n}}}\right| \geq t_{1-\alpha/2}(n-1)$ (resp. $\dfrac{m-c}{\frac{s}{\sqrt{n}}} \leq -t_{1-\alpha}(n-1)$ resp. $\dfrac{m-c}{\frac{s}{\sqrt{n}}} \geq t_{1-\alpha}(n-1)$).

c) Provedení testu o rozptylu

H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 \neq c$ (resp. H_1: $\sigma^2 < c$ resp. H_1: $\sigma^2 > c$) zamítáme na hladině významnosti α, jestliže $\dfrac{(n-1)s^2}{c} \leq \chi^2_{\alpha/2}(n\text{-}1)$ nebo $\dfrac{(n-1)s^2}{c} \geq \chi^2_{1-\alpha/2}(n\text{-}1)$ (resp. $\dfrac{(n-1)s^2}{c} \leq \chi^2_{\alpha}(n\text{-}1)$ nebo $\dfrac{(n-1)s^2}{c} \geq \chi^2_{1-\alpha}(n\text{-}1)$).

Před provedením kteréhokoli z uvedených testů je zapotřebí ověřit normalitu dat pomocí diagnostických grafů a testů normality popsaných v kapitole 3. Zjistíme-li u jednovýběrového t-testu, že rozsah souboru je malý ($n < 30$) a porušení normality je výraznější, doporučuje se přejít k neparametrickému jednovýběrovému Wilcoxonovu testu (viz kapitola 7). Pro výběry větších rozsahů není mírné porušení normality na překážku použití uvedených testů.

4.4.2. Příklad

Podle údajů na obalu čokolády by její čistá hmotnost měla být 125g. Výrobce dostal několik stížností od kupujících, ve kterých tvrdili, že hmotnost čokolád je nižší než deklarovaných 125g. Z tohoto důvodu oddělení kontroly náhodně vybralo 50 čokolád a zjistilo, že jejich průměrná hmotnost je 122g a směrodatná odchylka 8,6g. 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 $\frac{m-c}{\frac{s}{\sqrt{n}}} = \frac{122-125}{\frac{8,6}{\sqrt{50}}} = -2,4667$. Absolutní hodnotu testového kritéria porovnáme s kvantilem $t_{0,99}(49) = 2,4049$. Jelikož $2,4667 \geq 2,4049$, zamítáme nulovou hypotézu na hladině významnosti 0,01. Stížnosti kupujících tedy lze považovat za oprávněné.

4.5. Náhodný výběr z dvourozměrného normálního rozložení

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 $n \geq 2$. Označíme $\mu = \mu_1 - \mu_2$ a zavedeme rozdílový náhodný výběr $Z_1 = X_1 - Y_1, ..., Z_n = X_n - Y_n$. Vypočteme $M = \frac{1}{n}\sum_{i=1}^{n} Z_i$, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(Z_i - M)^2$.

4.5.1. Interval spolehlivosti pro parametr μ

Pro výpočet mezí $100(1-\alpha)\%$ empirického intervalu spolehlivosti pro střední hodnotu μ použijeme vzorec uvedený v 4.3.1. (b).

4.5.2. Párový t-test

Testujeme H_0: $\mu_1 - \mu_2 = 0$ (tj. $\mu = 0$) proti H_1: $\mu_1 - \mu_2 \neq 0$ (tj $\mu \neq 0$). Přechodem k rozdílovému náhodnému výběru převedeme párový t-test na jednovýběrový t-test, jehož provedení je popsáno v 4.4.1. (b).

Před provedením párového t-testu je zapotřebí se aspoň orientačně přesvědčit o dvourozměrné normalitě dat pomocí dvourozměrného tečkového diagramu. Je-li rozsah výběru malý ($n < 30$) a porušení normality je výraznější, je zapotřebí místo párového testu použít neparametrický párový Wilcoxonův test (viz kapitola 7). Pro výběry větších rozsahů, které vykazují jen mírné porušení normality, můžeme použít párový t-test.

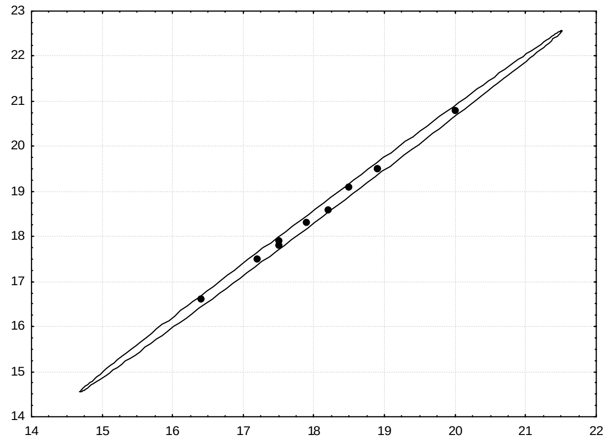
4.5.3. Příklad

Na 10 automobilech stejného typu se testovaly dva druhy benzínu lišící se oktanovým číslem. U každého automobilu se při průměrné rychlosti 90 km/h měřil dojezd (tj. dráha, kterou ujede na dané množství benzínu) při použití každého z obou druhů benzínu. Výsledky:

Číslo auta	1	2	3	4	5	6	7	8	9	10
benzín A	17,5	20,0	18,9	17,9	16,4	18,9	17,2	17,5	18,5	18,2
benzín B	17,8	20,8	19,5	18,3	16,6	19,5	17,5	17,9	19,1	18,6

Za předpokladu, že dojezd se řídí normálním rozložením, testujte na hladině významnosti 0,05 hypotézu, že rozdíl středních hodnot dojezdu při dvou druzích benzínu se neliší.

Řešení:
Pomocí dvourozměrného tečkového diagramu se zakreslenou 95% elipsou konstantní hustoty pravděpodobnosti posoudíme oprávněnost předpokladu o dvourozměrné normalitě dat.



Vidíme, že tečky se řadí do velmi úzkého elipsovitého obrazce. Data můžeme považovat za realizace náhodného výběru z dvourozměrného normálního rozložení.

Přejdeme k rozdílovému náhodnému výběru. Označíme $\mu = \mu_1 - \mu_2$. Testujeme hypotézu H_0: $\mu = 0$ proti H_1: $\mu \neq 0$ na hladině významnosti 0,05. Vypočteme m = -0,46, s = 0,1838 a testové kritérium $t_0 = -7{,}9148$. Absolutní hodnotu testového kritéria porovnáme s kvantilem $t_{0,975}(9) = 2{,}2622$. Protože $7{,}9148 \geq 2{,}2622$, zamítáme nulovou hypotézu na hladině významnosti 0,05.

Kontrolní otázky

1. Jaké pivotové statistiky odvozené z výběrového průměru M a výběrového rozptylu S^2 používáme při řešení úloh o střední hodnotě μ a rozptylu σ^2 normálního rozložení?
2. Jak vypadají meze $100(1-\alpha)\%$ empirického intervalu spolehlivosti pro neznámou střední hodnotu μ, když rozptyl σ^2 není znám?
3. Jaké testy o parametrech normálního rozložení znáte?
4. V jaké situaci a za jakých podmínek použijete jednovýběrový t-test?
5. V jaké situaci a za jakých podmínek použijete dvouvýběrový t-test?

Příklady

1. Lze předpokládat, že hmotnost pomerančů dodávaných do obchodní sítě se řídí normálním rozložením se střední hodnotou 170 g a směrodatnou odchylkou 12 g. Jaká je pravděpodobnost, že celková hmotnost devíti náhodně vybraných pomerančů balených do síťky překročí 1,5 kg?
Výsledek: Hledaná pravděpodobnost je 0,797.

2. Počet bodů v testu inteligence je náhodná veličina, která se řídí rozložením N(100,225). Jaká je pravděpodobnost, že průměr v náhodně vybrané skupině 20 osob bude větší než 105 bodů?
Výsledek: Hledaná pravděpodobnost je 0,06811.

3. Při provádění určitého pokusu bylo zapotřebí udržovat v laboratoři konstantní teplotu 26,5°C. Teplota byla v jednom pracovním týdnu 46x namátkově kontrolována v různých denních a nočních hodinách. Z výsledků měření byly vypočteny realizace výběrového průměru a výběrové směrodatné odchylky: m = 26,33°C, s = 0,748°C. Za předpokladu, že výsledky měření teploty se řídí rozložením $N(\mu,\sigma^2)$, vypočtěte 95% empirický interval spolehlivosti
a) pro střední hodnotu μ
b) pro směrodatnou odchylku σ.

Výsledek:
ad a) Dosazením do vzorce 4.3.1. (b) dostaneme 26,11°C < μ < 26,55°C s pravděpodobností aspoň 0,95.
ad b) Dosazením do vzorce 4.3.1. (d), kde meze odmocníme, dostaneme 0,62°C < σ < 0,94°C s pravděpodobností aspoň 0,95.

4. U 25 náhodně vybraných dvoulitrových lahví s nealkoholickým nápojem byl zjištěn přesný objem nápoje. Výběrový průměr činil m = 1,99 l a výběrová směrodatná odchylka s = 0,1 l. Předpokládejme, že objem nápoje v láhvi je náhodná veličina s normálním rozložením.
a) Na hladině významnosti 0,05 ověřte tvrzení výrobce, že zákazník není znevýhodněn.
b) Na hladině významnosti 0,05 ověřte tvrzení výrobce, že směrodatná odchylka je 0,08 l.
Výsledek:
ad a) Testujeme hypotézu H_0: $\mu = 2$ proti levostranné alternativě H_1: $\mu < 2$ pomocí jednovýběrového t-testu (viz 4.4.1. (b)). Jelikož hodnota testového kritéria -0,5 neleží v kritickém oboru $\left(-\infty; -2{,}064\right\rangle$, nezamítáme nulovou hypotézu na hladině významnosti 0,05.
ad b) Testujeme hypotézu H_0: $\sigma = 0{,}08$ proti oboustranné alternativě H_1: $\sigma \neq 0{,}08$ pomocí testu o rozptylu (viz 4.4.1. (c)). Jelikož hodnota testového kritéria 37,5 neleží v kritickém oboru $\left(0; 12{,}4\right\rangle \cup \left\langle 39{,}4; \infty\right)$, nejsme oprávněni na hladině významnosti 0,05 zamítnout tvrzení výrobce.

5. (S) Bylo vybráno šest nových vozů téže značky a po určité době bylo zjištěno, o kolik mm se sjely jejich levé a pravé přední pneumatiky. Výsledky: (1,8; 1,5), (1,0; 1,1), (2,2; 2,0), (0,9; 1,1), (1,5; 1,4), (1,6; 1,4). Za předpokladu, že uvedené dvojice tvoří náhodný výběr z dvourozměrného normálního rozložení s vektorem středních hodnot (μ_1, μ_2), testujte na hladině významnosti 0,05 hypotézu, že obě pneumatiky se sjíždějí stejně rychle.
Výsledek:
Vzhled dvourozměrného tečkového diagramu není v rozporu s předpokladem o dvourozměrném normálním rozložení. Přejdeme k rozdílovému náhodnému výběru a testujeme nulovouhypotézu H_0: $\mu = 0$ proti oboustranné alternativě H_1: $\mu \neq 0$ pomocí párového t-testu. Hodnota testového kritéria = 1,0512, počet stupňů volnosti = 5. Protože odpovídající p-hodnota = 0,3411 je větší než hladina významnosti 0,05, nelze na hladině významnosti 0,05 zamítnout nulovou hypotézu. Ke stejnému rozhodnutí dospějeme, pokud stanovíme kritický obor: W = $\left(-\infty; -2{,}571\right) \cup \left(2{,}571; \infty\right)$. Testové kritérium se nerealizuje v kritickém oboru, tedy nelze na hladině významnosti 0,05 zamítnout nulovou hypotézu.

6. (S) Uměle připravený vzorek minerálu obsahoval 10% křemene a byl 12 krát proměřen. Výsledky měření byly: 8,7 10,2 10,07 9,75 9,65 10,37 10,14 10,5 9,48 11,22 9,49 9,86. Na hladině významnosti 0,05 testujte hypotézu, že obsah křemene byl stanoven správně.
Výsledek:
K-S test ani S-W test nezamítají na hladině významnosti 0,05 normalitu dat. Testujeme nulovou hypotézu H_0: $\mu = 10$ proti oboustranné alternativě H_1: $\mu \neq 10$. Úloha vede na jednovýběrový t-test. Realizace testového kritéria = -0,262, počet stupňů volnosti = 1. Protože odpovídající p-hodnota = 0,7981 je větší než hladina významnosti 0,05, nelze na hladině významnosti 0,05 zamítnout nulovou hypotézu.