Příklady z matematické statistiky

Příklad 1.: Kontrola zvážila 5 tabulek čokolády. Výsledky vážení (v gramech) byly: 198, 199, 197, 202, 200. Předpokládáme, že tyto výsledky představují realizace náhodného výběru rozsahu 5 z normálního rozložení $N(\mu, \sigma^2)$.

a) Vypočtěte výběrový průměr a výběrový rozptyl. (definice 11.2.)
b) Sestrojte 95% interval spolehlivosti pro střední hodnotu μ. (věta 12.9. (b))
c) Sestrojte 95% interval spolehlivosti pro rozptyl σ^2. (věta 12.9. (c))

Řešení:

ad a) $m = 199{,}2$ g, $s^2 = 3{,}7$ g^2

ad b) $d = m - \frac{s}{\sqrt{n}}\, t_{1-\alpha/2}(n-1) = 199{,}2 - \frac{\sqrt{3{,}7}}{\sqrt{5}}\, t_{0{,}975}(4) = 199{,}2 - \frac{\sqrt{3{,}7}}{\sqrt{5}}\, 2{,}7764 = 196{,}8$

$h = m + \frac{s}{\sqrt{n}}\, t_{1-\alpha/2}(n-1) = 199{,}2 + \frac{\sqrt{3{,}7}}{\sqrt{5}}\, t_{0{,}975}(4) = 199{,}2 + \frac{\sqrt{3{,}7}}{\sqrt{5}}\, 2{,}7764 = 201{,}6$

$196{,}8$ g $< \mu <$ $201{,}6$ g s pravděpodobností aspoň 0,95.

ad c) $d = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} = \frac{4 \cdot 3{,}7}{\chi^2_{0{,}975}(4)} = \frac{4 \cdot 3{,}7}{11{,}143} = 1{,}33$

$h = \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)} = \frac{4 \cdot 3{,}7}{\chi^2_{0{,}025}(4)} = \frac{4 \cdot 3{,}7}{0{,}484} = 30{,}58$

$1{,}33$ g$^2 < \sigma^2 <$ $30{,}58$ g^2 s pravděpodobností aspoň 0,95.

Příklad 2.: Měřením délky deseti válečků byly získány hodnoty (v mm): 5,38 5,36 5,35 5,40 5,41 5,34 5,29 5,43 5,42 5,32. Pro úsporu času máte uveden aritmetický průměr $m = 5{,}37$ mm a směrodatnou odchylku $s = 0{,}044$ mm. Těchto deset hodnot považujeme za realizace náhodného výběru rozsahu 10 z normálního rozložení $N(\mu, \sigma^2)$.

a) Sestrojte 99% interval spolehlivosti pro neznámou střední hodnotu μ.(Návod: věta 12.9. (b))
b) Sestrojte 99% interval spolehlivosti pro neznámou směrodatnou odchylku σ. (Návod: věta 12.9. (c))
c) Na hladině významnosti 0,01 testujte hypotézu, že střední hodnota délky válečků je 5,3 mm proti oboustranné alternativě. (Návod: poznámka 13.5. (b))

Řešení:

ad a)

$d = m - \frac{s}{\sqrt{n}}\, t_{1-\alpha/2}(n-1) = 5{,}37 - \frac{0{,}044}{\sqrt{10}}\, t_{0{,}995}(9) = 5{,}37 - \frac{0{,}044}{\sqrt{10}}\, 3{,}25 = 5{,}3248$

$h = m + \frac{s}{\sqrt{n}}\, t_{1-\alpha/2}(n-1) = 5{,}37 + \frac{0{,}044}{\sqrt{10}}\, t_{0{,}995}(9) = 5{,}37 + \frac{0{,}044}{\sqrt{10}}\, 3{,}25 = 5{,}4152$

$5{,}3248$ mm $< \mu <$ $5{,}4152$ mm s pravděpodobností aspoň 0,99

ad b)

$$d = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}} = \sqrt{\frac{9 \cdot 0{,}044^2}{\chi^2_{0{,}995}(9)}} = \frac{3 \cdot 0{,}044}{\sqrt{23{,}589}} = 0{,}0272$$

$$h = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}} = \sqrt{\frac{9 \cdot 0{,}044^2}{\chi^2_{0{,}005}(9)}} = \frac{3 \cdot 0{,}044}{\sqrt{1{,}735}} = 0{,}1002$$

0,0272 mm < σ < 0,1002 mm s pravděpodobností aspoň 0,99.
ad c) Testujeme H_0: $\mu = 5{,}3$ proti H_1: $\mu \neq 5{,}3$ na hladině významnosti 0,01. Protože 99% empirický interval spolehlivosti vypočtený v bodě (a) neobsahuje hodnotu 5,3, zamítáme nulovou hypotézu na hladině významnosti 0,01 a přijímáme alternativní hypotézu.

Příklad 3.: Při kontrolních zkouškách životnosti 16 žárovek byl stanoven odhad m = 3000 h střední hodnoty jejich životnosti a odhad s = 20 h směrodatné odchylky jejich životnosti. Za předpokladu, že životnost žárovky se řídí normálním rozložením, vypočtěte

a) 99% empirický interval spolehlivosti pro střední hodnotu životnosti
b) 99% empirický interval spolehlivosti pro směrodatnou odchylku životnosti
c) 90% levostranný empirický interval spolehlivosti pro střední hodnotu životnosti
d) 95% pravostranný empirický interval spolehlivosti pro rozptyl životnosti.

Návod: věta 12.9., příklad 12.8. – konkrétní aplikace
Upozornění: Výsledek zaokrouhlete na jedno desetinné místo a vyjádřete v hodinách a minutách.

Řešení:
ad a)

$$d = m - \frac{s}{\sqrt{n}} t_{0,995}(n-1) = 3000 - \frac{20}{\sqrt{16}} 2{,}9467 = 2985{,}3,\ h = 3014{,}7$$

2985 h a 18 min < μ < 3014 h a 42 min s pravděpodobností aspoň 0,99
ad b)

$$d = \sqrt{\frac{(n-1)s^2}{\chi^2_{0,995}(n-1)}} = \frac{\sqrt{15}\cdot 20}{\sqrt{32{,}801}} = 13{,}5,\ h = \sqrt{\frac{(n-1)s^2}{\chi^2_{0,005}(n-1)}} = \frac{\sqrt{15}\cdot 20}{\sqrt{4{,}601}} = 36{,}1$$

13 h a 30 min < σ < 36 h a 6 min s pravděpodobností aspoň 0,99
ad c)

$$d = m - \frac{s}{\sqrt{n}} t_{0,9}(n-1) = 3000 - \frac{20}{\sqrt{16}} 1{,}3406 = 2993{,}3$$

2993 h a 18 min < μ s pravděpodobností aspoň 0,9
ad d)

$$h = \frac{(n-1)s^2}{\chi^2_{0,05}(n-1)} = \frac{15\cdot 400}{7{,}261} = 826{,}3$$

826 h^2 a 18 min^2 > σ^2 s pravděpodobností aspoň 0,95

Příklad 4.: Nechť $X_1, ..., X_{10}$ je náhodný výběr z $N(\mu, \sigma^2)$, kde parametry μ, σ^2 neznáme. Realizace výběrového průměru M je m = 21,2 a realizace výběrového rozptylu S^2 je s^2 = 30,25. Na hladině významnosti α = 0,01 testujte nulovou hypotézu H_0: μ = 25 proti alternativní hypotéze H_1: μ < 25. Test proveďte pomocí

a) intervalu spolehlivosti,
b) kritického oboru.

Řešení:
ad a)
Při testování nulové hypotézy proti levostranné alternativě konstruujeme pravostranný interval spolehlivosti.

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) = 21{,}2 + \frac{\sqrt{30{,}25}}{\sqrt{10}} t_{0,99}(9) = 21{,}2 + 1{,}7392527.2{,}8214 = 26{,}11$$

Protože číslo c = 25 leží v intervalu (-∞; 26,11), hypotézu H_0: μ = 25 nezamítáme na hladině významnosti 0,01.
ad b)

Vypočteme realizaci testového kritéria $t_0 = \dfrac{m-c}{\dfrac{s}{\sqrt{n}}} = \dfrac{21{,}2-25}{\dfrac{\sqrt{30{,}25}}{\sqrt{10}}} = -2{,}1848$

Číslo t_0 porovnáme s opačnou hodnotou kvantilu $t_{0,99}(9) = 2{,}8214$. Protože -2,1848 je větší než -2,8214, hypotézu H_0: μ = 25 nezamítáme na hladině významnosti 0,01.

Příklad 5.: 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
a) 95% empirický interval spolehlivosti pro střední hodnotu μ.
Návod: Věta 12.9. b), příklad 9 a) z POTu
b) 95% empirický interval spolehlivosti pro směrodatnou odchylku σ.
Návod: Věta 12.9. c), příklad 9 b) z POTu

Řešení:
ad a)

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 26{,}33 - \frac{0{,}748}{\sqrt{46}} t_{0,975}(45) = 26{,}33 - \frac{0{,}748}{\sqrt{46}} 2{,}0141 = 26{,}11$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 26{,}33 + \frac{0{,}748}{\sqrt{46}} t_{0,975}(45) = 26{,}33 + \frac{0{,}748}{\sqrt{46}} 2{,}0141 = 26{,}55$$

26,11°C < μ < 26,55°C s pravděpodobností aspoň 0,95.

ad b)

$$d = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}} = \sqrt{\frac{(n-1)s^2}{\chi^2_{0,975}(n-1)}} = \frac{0{,}748\sqrt{45}}{\sqrt{65{,}4102}} = 0{,}62$$

$$h = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}} = \sqrt{\frac{(n-1)s^2}{\chi^2_{0,025}(n-1)}} = \frac{0{,}748\sqrt{45}}{\sqrt{28{,}3662}} = 0{,}94$$

0,62°C < σ < 0,94°C s pravděpodobností aspoň 0,95.

Příklad 6.: 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)
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 (poznámka 13.5. (b), věta 12.9.)
b) pomocí kritického oboru. (poznámka 13.5. (b), věta 13.9.)

(Pro úsporu času máte uvedeny realizace výběrového průměru m = $-1,\bar{3}$ 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:

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

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

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 $t_0 = \dfrac{m-c}{\dfrac{s}{\sqrt{n}}} = \dfrac{-1,\bar{3}}{\dfrac{\sqrt{4,\overline{78}}}{\sqrt{12}}} = -2,11085$

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.