5.2 Тестови за параметрите на нормална распределба

Задача 5.11. Една откупна станица го мери процентот на масленост во млекото за кој се претпоставува дека има нормална распределба $\mathcal{N}(m, 1)$. Во текот на еден ден собрано е млеко од 25 производители и добиени се следните резултати:

процент на масленост	3,45 - 3,75	3,75 - 4,05	4,05 - 4,35
број на прозводители	5	16	4

Со ниво на значајност $\alpha = 0,05$, провери ја хипотезата дека средната вредност на процентот на маследност на млекото е 4.

Решение. Нека X - процент на масленост во млекото, притоа $X \sim \mathcal{N}(m, 1)$, каде m е непознат параметар. Врз основа на дадените податоци, треба со ниво на значајност $\alpha = 0,05$, да се тестира хипотезата $H_0 : m = 4$, против $H_1 : m \neq 4$. При тестирање на оваа хипотеза, критичната област го има обликот

$$C = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : \Big|\frac{\overline{x} - m_0}{\sigma}\sqrt{n}\Big| \geq c_0\},$$

каде $m_0 = 4$, $\sigma = 1$, $n = 25$, аритметичката средина на дадените податоци е

$$\overline{x} = \frac{1}{25}(5 \cdot 3,6 + 16 \cdot 3,9 + 4 \cdot 4,2) = 3,888,$$

а критичната вредност c_0 ја наоѓаме од $c_0 = \Phi_0^{-1}(\frac{1-\alpha}{2}) = \Phi_0^{-1}(\frac{1-0,05}{2}) = \Phi_0^{-1}(0,475) = 1,96$, што се чита од таблицата за нормална распределба. Така, за дадените податоци имаме

$$\Big|\frac{\overline{x} - m_0}{\sigma}\sqrt{n}\Big| = \Big|\frac{3,888 - 4}{1}\sqrt{25}\Big| = 0,56 < 1,96 = c_0,$$

од каде за дадените податоци важи $x \notin C$, значи тие не ѝ противречат на нултата хипотеза $H_0 : m = 4$, па таа се прифаќа.

Задача 5.12. За да се испитаат временските услови во едно туристичко место, за 15 случајно избрани години забележан е бројот на сончеви денови во текот на годината и добиени се следните резултати:

$$220, 218, 180, 230, 245, 253, 227, 182, 194, 228, 231, 192, 260, 268, 251.$$

Со ниво на значајност $\alpha = 0,05$ и под претпоставка дека X-бројот на сончеви денови во текот на една година има нормална распределба, провери ја хипотезата дека средниот број на сончеви денови во текот на годината е 220, против алтернативната хипотеза дека е поголем од 220.

Решение. За обележјето X-број на сончеви денови во текот на една година имаме дека е нормално распределено т.е. $X \sim \mathcal{N}(m, \sigma^2)$, каде m и σ^2 се непознати параметри. Врз основа на дадените податоци, треба со ниво на значајност $\alpha = 0,05$, да се тестира хипотезата $H_0 : m = 220$, против $H_1 : m > 220$. При тестирање на оваа хипотеза, критичната област го има обликот

$$C = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : \frac{\overline{x} - m_0}{s}\sqrt{n-1} \geq c_0\},$$

каде $m_0 = 220$, $n = 15$, аритметичката средина на дадените податоци е $\overline{x} = 225,3$, дисперзијата е $s^2 = 733,7$, од каде $s = 27,1$, а критичната вредност $c_0 = t_{n-1,2\alpha} = t_{14;0,1} = 1,761$ се чита од таблицата за студентова распределба. Тогаш, за дадените податоци имаме

$$\frac{\overline{x} - m_0}{s}\sqrt{n-1} = \frac{225,3 - 220}{27,1}\sqrt{14} = 0,732 < 1,761 = c_0,$$

значи $x \notin C$, па нултата хипотеза $H_0 : m = 220$ се прифаќа.

Задача 5.13. При мерењето на масата (во kg) на една група ученици, добиени се следните резултати:

маса (во kg)	40 - 50	50 - 60	60- 70	70 - 80	80 - 90
број на ученици	6	7	10	4	3

Со ниво на значајност $\alpha = 0,10$, под претпоставка дека X - масата на еден ученик има нормална распределба, провери ја хипотезата дека средното квадратно отстапување на масата на учениците е 13 kg.

Решение. Дадено е дека обележјето X-маса на еден ученик има $\mathcal{N}(m, \sigma^2)$ распределба, каде m и σ^2 се непознати параметри. Врз основа на дадените податоци, треба со ниво на значајност $\alpha = 0,10$, да се тестира хипотезата $H_0 : \sigma = 13$, против $H_1 : \sigma \neq 13$. При тестирање на оваа хипотеза, критичната област го има обликот

$$C = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : \frac{ns^2}{\sigma_0^2} \geq b \text{ или } \frac{ns^2}{\sigma_0^2} \leq a\},$$

каде $n = 30$, $\sigma_0^2 = 13^2$, $a = \chi^2_{n-1,1-\alpha/2} = \chi^2_{29;0,95} = 17,708$ и $b = \chi^2_{n-1,\alpha/2} = \chi^2_{29;0,05} = 42,557$ се читаат од таблицата за хи-квадрат распределба, а дисперзијата на дадените податоци е $s^2 = 147,67$. Тогаш, за дадените податоци имаме

$$\frac{ns^2}{\sigma_0^2} = \frac{30 \cdot 147,67}{13^2} = 2,016 < 17,708 = a,$$

значи $x \in C$, па нултата хипотеза $H_0 : \sigma = 13$ се отфрла и се прифаќа алтернативната $H_1 : \sigma \neq 13$.

Ирена Стојковска

Задача 5.14. Од два града на случаен начин се избрани 10, односно 12 жители на кои им е измерена висината во *cm* и добиени се следните резултати:

$$\text{град А:}168, 174, 178, 182, 178, 186, 179, 180, 169, 172,$$
$$\text{град Б:}173, 182, 180, 175, 180, 178, 184, 188, 179, 181, 165, 173.$$

Ако висината на случајно избран жител има нормална распределба со средно квадратно отстапување од очекуваната вредност од 5 *cm*, со ниво на значајност $\alpha = 0,05$, провери ја хипотезата дека средните висини на жителите од градовите А и Б се еднакви, против алтернативната хипотеза дека средната висина на жителите од градот Б е поголема.

Решение. Нека X е висината на жител од градот А, а Y е висината на жител од градот Б, тогаш $X \sim \mathcal{N}(m_1, 5^2)$ и $Y \sim \mathcal{N}(m_2, 5^2)$, каде m_1 и m_2 се непознати параметри. Врз основа на дадените податоци, треба со ниво на значајност $\alpha = 0,05$, да се тестира хипотезата $H_0 : m_1 = m_2$, против $H_1 : m_1 < m_2$. При тестирање на оваа хипотеза, критичната област го има обликот

$$C = \{(x, y) \in \mathbb{R}^{n_1+n_2} : \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq -c\},$$

каде $n_1 = 10$, $n_2 = 12$, $\sigma_1^2 = \sigma_2^2 = 5^2$, $c = \Phi_0^{-1}(\frac{1-2\alpha}{2}) = \Phi_0^{-1}(\frac{1-2\cdot 0,05}{2}) = \Phi_0^{-1}(0,45) = 1,645$, а аритметичките средини на дадените податоци се $\overline{x} = 176,6$ и $\overline{y} = 178,2$. Тогаш, за дадените податоци имаме

$$\frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{176,6 - 178,2}{\sqrt{\frac{5^2}{10} + \frac{5^2}{12}}} = -0,747 > -1,645 = -c,$$

значи $(x, y) \notin C$, па нултата хипотеза $H_0 : m_1 = m_2$ се прифаќа, односно се прифаќа хипотезата дека средните висини на жителите од градовите А и Б се еднакви.

Задача 5.15. Една група од 16 студенти на еден испит ги покажала следните резултати (во број на поени од можни 100):

студент	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
писмен дел	90	90	80	90	92	88	90	63	70	54	78	86	99	84	56	85
усмен дел	84	84	82	94	90	85	89	62	65	52	72	90	98	89	58	85

Под претпоставка дека бројот на поени на еден студент добиени на писмениот, односно устниот дел од испитот имаат нормални распределби со еднакви дисперзии, врз основа на дадените податоци, со 5% ниво на значајност, провери ја хипотезата дека средните вредности на поените на писмениот и устниот дел се еднакви.

Решение. Нека X е бројот на поени добиен на писмениот испит, а Y е бројот на поени добиен на устниот испит, тогаш $X \sim \mathcal{N}(m_1, \sigma^2)$ и $Y \sim \mathcal{N}(m_2, \sigma^2)$, каде m_1, m_2 и σ^2 се непознати параметри. Врз основа на дадените податоци, треба со ниво на значајност $\alpha = 5\% = 0,05$, да се тестира хипотезата $H_0 : m_1 = m_2$, против $H_1 : m_1 \neq m_2$. При тестирање на оваа хипотеза, критичната област го има обликот

$$C = \{(x, y) \in \mathbb{R}^{n_1+n_2} : \Big| \frac{\overline{x} - \overline{y}}{\sqrt{n_1 s_x^2 + n_2 s_y^2}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}(n_1 + n_2 - 2)} \Big| \geq c\},$$

каде $n_1 = 16$, $n_2 = 16$, $c = t_{n_1+n_2-2,\alpha} = t_{16+16-2;0,05} = t_{30;0,05} = 2,042$, а аритметичките средини и дисперзиите на дадените податоци се $\overline{x} = 80,9375$, $\overline{y} = 79,9375$, $s_x^2 = 167,31$ и $s_y^2 = 178,06$. Тогаш, за дадените податоци имаме

$$\begin{aligned} &\Big| \frac{\overline{x} - \overline{y}}{\sqrt{n_1 s_x^2 + n_2 s_y^2}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}(n_1 + n_2 - 2)} \Big| = \\ = \ &\Big| \frac{80,9375 - 79,9375}{\sqrt{16 \cdot 167,31 + 16 \cdot 178,06}} \sqrt{\frac{16 \cdot 16}{16 + 16}(16 + 16 - 2)} \Big| = 0,215 < 2,042 = c, \end{aligned}$$

значи $(x, y) \notin C$, па нултата хипотеза $H_0 : m_1 = m_2$ се прифаќа, односно се прифаќа хипотезата дека средните вредности на поените на писмениот и устниот дел од испитот се еднакви.

Задача 5.16. За податоците $x_1, ..., x_{n_1}$ и $y_1, ..., y_{n_2}$ кои одговараат на обележјата X и Y соодветно, познато е дека $n_1 = 10$, $n_2 = 17$, $s_x^2 = 3,2^2$ и $s_y^2 = 3^2$. Врз основа на овие податоци, под претпоставка дека обележјата X и Y имаат нормални распределби $\mathcal{N}(a_1, \sigma_1^2)$ и $\mathcal{N}(a_2, \sigma_2^2)$ соодветно, со ниво на значајност $\alpha = 0,10$, провери ја хипотезата $H_0 : \sigma_1^2 = \sigma_2^2$, против алтернативната $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Решение. Дадено е дека обележјата $X \sim \mathcal{N}(a_1, \sigma_1^2)$ и $Y \sim \mathcal{N}(a_2, \sigma_2^2)$, каде a_1, a_2, σ_1^2 и σ_2^2 се непознати параметри. При тестирање на хипотезата $H_0 : \sigma_1^2 = \sigma_2^2$, против алтернативната $H_1 : \sigma_1^2 \neq \sigma_2^2$, со ниво на значајност $\alpha = 0,10$, врз основа на дадените податоци за кои $s_x^2 > s_y^2$, критичната област има облик

$$C = \{(x, y) \in \mathbb{R}^{n_1+n_2} : \frac{n_1(n_2 - 1)s_x^2}{n_2(n_1 - 1)s_y^2} \geq c\},,$$

каде $c = F_{n_1-1,\, n_2-1,\, \alpha/2} = F_{10-1,\, 17-1,\, 0,10/2} = F_{9,\, 16,\, 0,05} = 2,54$ е вредност од таблицата за Фишеровата распределба. Сега, за дадените податоци имаме

$$\frac{n_1(n_2 - 1)s_x^2}{n_2(n_1 - 1)s_y^2} = \frac{10 \cdot (17 - 1) \cdot 3,2^2}{17 \cdot (10 - 1) \cdot 3^2} = 1,1898 < 2,54 = c,$$

значи $(x, y) \notin C$, па нултата хипотеза $H_0 : \sigma_1^2 = \sigma_2^2$ се прифаќа.

Ирена Стојковска

Задачи за самостојна работа

Задача 5.17. Измерена е должината на стеблото (во cm) на 30 семиња после една недела од почетокот на ртењето и добиени се следните резултати:

$$4.7, 4.9, 5.2, 5.2, 5.4, 5.5, 5.5, 5.5, 5.7, 5.7, 5.8, 6.0, 6.0, 6.0, 6.0,$$
$$6.1, 6.1, 6.2, 6.4, 6.5, 6.5, 6.7, 6.9, 7.2, 7.5, 7.8, 8.0, 8.5, 8.5, 9.0.$$

Под претпоставка дека X-должината на стеблото по една недела ртење има нормална $\mathcal{N}(m, 2^2)$ распределба, врз основа на дадените податоци, со 5% ниво на значајност, тестирај ја хипотезата $H_0 : m = 6$, против алтернативната $H_1 : m > 6$.

Задача 5.18. Една физичка величина е мерена 26 пати при исти услови. Притоа средно квадратното отстапување на податоците од нивната аритметичка средина е 2,20. Под претпоставка дека физичката величина има нормална распределба, со 5% ниво на значајност, тестирај ја хипотезата дека дисперзијата на физичката величина е 5.

Задача 5.19. Извршено е мерење на масата (во g) на 15 јаболка од сортата "Златен делишес" и 25 јаболка од сортата "Црвен делишес". Добиено е дека средната вредност на масата на јаболката "Златен делишес" е 110,5 g со средно квадратно отстапувањето од 5,1 g, додека средната вредност на масата на јаболката "Црвен делишес" е 115,2 g со средно квадратно отстапувањето од 3,5 g. Под претпоставка дека масатата на двете сорти јаболка има нормална распределба, врз основа на дадените податоци, со 5% ниво на значајност, тестирај ја хипотезата за еднаквост на масите на јаболката од двете сорти.