Задатак 1. На основу узорка обима n тестирати хипотезу H_0(обележје X има нормалну $\mathcal{N}(0,1)$ расподелу) против хипотезе H_1(обележје X има густину расподеле $f(x) = \frac{1}{2}e^{-|x|}$). Одредити најбољу критичну област ако је обим узорка 100, а праг значајности 0.05.

Решење. Хипотезе H_0 и H_1 су непараметарске, па ћемо од њих да направимо параметарске увођењем фиктивног параметра. Означимо $f(x;\theta_0) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, а $f(x;\theta_1) = \frac{1}{2}e^{-|x|}$. Сада можемо тестирати $H_0(\theta = \theta_0)$ против $H_1(\theta = \theta_1)$.

Овако дефинисане хипотезе су параметарске и просте, па ћемо критичну област одредити помоћу леме Нејман-Пирсона:

$$
\begin{aligned}
W = \left\{\frac{L(\theta_1)}{L(\theta_0)} \geq k\right\} &= \left\{\frac{\frac{1}{2^n}e^{-\sum\limits_{i=1}^{n}|x_i|}}{\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}\sum\limits_{i=1}^{n}x_i^2}} \geq k\right\} \\
&= \left\{e^{\frac{1}{2}\sum\limits_{i=1}^{n}x_i^2 - \sum\limits_{i=1}^{n}|x_i|} \geq k_1\right\} = \left\{\sum_{i=1}^{n}(|x_i|-1)^2 \geq c\right\}.
\end{aligned}
$$

Праг значајности је 0.05, па важи једнакост:

$$
P_{H_0}\left\{\sum_{i=1}^{100}(|X_i|-1)^2 \geq c\right\} = 0.05
$$

Уведимо ознаку $Y_i = (|X_i|-1)^2$. Расподела случајних величина Y_i нам није позната, али не морамо ни да је одређујемо, јер је $n = 100$, па можемо применити централну граничну теорему:

$$
\frac{\sum\limits_{i=1}^{100}Y_i - E\left(\sum\limits_{i=1}^{100}Y_i\right)}{\sqrt{D\left(\sum\limits_{i=1}^{100}Y_i\right)}} \xrightarrow{D} Y^\star \in \mathcal{N}(0,1).
$$

Одредимо очекивање и дисперзију случајне величине Y_i при H_0.

$$
\begin{aligned}
EY_i &= E(|X_i|-1)^2 = EX_i^2 - 2E|X_i| + 1 = 1 - 2\int_{-\infty}^{+\infty}|x|\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx + 1 \\
&= 2 - \frac{4}{\sqrt{2\pi}}\int_0^{+\infty}xe^{-\frac{x^2}{2}}dx = \left|\frac{x^2}{2} = t,\ xdx = dt\right| \\
&= 2 - \frac{4}{\sqrt{2\pi}}\int_0^{+\infty}e^{-t}dt = 2 - \frac{4}{\sqrt{2\pi}}, \\
EY_i^2 &= EX_i^4 - 4E|X_i|^3 + 6E|X_i|^2 - 4E|X_i| + 1 \\
&= \int_{-\infty}^{+\infty}x^4\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx - 8\int_0^{+\infty}x^3\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx + 6 - 4\frac{2}{\sqrt{2\pi}} + 1 \\
&= \frac{2}{\sqrt{2\pi}}\int_0^{+\infty}(\sqrt{2t})^3e^{-t}dt - \frac{1}{\sqrt{2\pi}}\int_0^{+\infty}2te^{-t}dt + 7 - 4\frac{2}{\sqrt{2\pi}} \\
&= \frac{2}{\sqrt{2\pi}}2\sqrt{2}\Gamma\left(\frac{5}{2}\right) - \frac{16}{\sqrt{2\pi}}\Gamma(2) + 7 - \frac{8}{\sqrt{2\pi}} \\
&= \frac{4}{\sqrt{\pi}}\frac{3}{4}\sqrt{\pi} - \frac{16}{\sqrt{2\pi}} + 7 - \frac{8}{\sqrt{2\pi}} = 10 - \frac{24}{\sqrt{2\pi}},
\end{aligned}
$$

при чему је коришћено да је $\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{4}\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi}$. Дакле,

$$DY_i = E(|X_i| - 1)^4 - (E(|X_i| - 1))^2 = 6 - \frac{8}{\sqrt{2\pi}} - \frac{8}{\pi}$$

Уврстимо вредности за очекивање и дисперзију и добијамо да је:

$$P_{H_0}\left\{\frac{\sum_{i=1}^{100}(|X_i| - 1)^2 - E\left(\sum_{i=1}^{100}(|X_i| - 1)^2\right)}{\sqrt{D\left(\sum_{i=1}^{100}(|X_i| - 1)^2\right)}} \geq \frac{c - 100\left(2 - \frac{4}{\sqrt{2\pi}}\right)}{\sqrt{100\left(6 - \frac{8}{\sqrt{2\pi}} - \frac{8}{\pi}\right)}}\right\} = 0.05.$$

Из таблица за нормалну расподелу добијамо да је $\frac{c-40.42}{5.12} = \Phi^{-1}(0.95) = 1.645$, па је $c = 48.8424$.

Дакле, тражена критична област је

$$W = \left\{\sum_{i=1}^{100}(|x_i| - 1)^2 \geq 48.8424\right\}.$$

✓

p–вредност теста

Још један начин да донесемо одлуку да ли прихватити или одбацити нулту хипотезу је коришћење p–вредности теста.

Нека је $p(\mathbf{X})$ тест статистика која задовољава $0 \leq p(\mathbf{x}) \leq 1$ за свако $\mathbf{x} \in \mathcal{X}$. Мале вредности ове статистике указују на одбацивање нулте хипотезе. Уколико за свако $\theta \in \Theta_0$ и свако $\alpha \in [0, 1]$ важи

$$P_\theta\{p(\mathbf{X}) \leq \alpha\} \leq \alpha, \tag{**}$$

статистика $p(\mathbf{X})$ назива се p–вредност.

Ако разматрамо тестове одређене тест функцијом облика $\varphi_\alpha = I\{T(\mathbf{X}) > k_\alpha\}$, где је T тест статистика, тада је p–вредност теста

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} P_\theta\{T(\mathbf{X}) \geq T(\mathbf{x})\}.$$

Аналогно се формирају за остале облике тест функције, односно критичне области. Што је p–вредност теста мања, то је јачи доказ против нулте хипотезе.

Задатак 2. Нека је X_n узорак из експоненцијалне расподеле са непознатим параметром λ. За тестирање $H_0(\lambda = 1)$ против $H_1(\lambda > 1)$ користи се једнострани тест с тест функцијом

$$\varphi(x) = I\left\{\sum_{i=1}^{n} x_i \leq k\right\}.$$

а) Одредити k тако да тест буде мере α.

б) Одредити статистику $p(X_n)$ која представља p–вредност овог теста (фамилије тестова за различите α) и доказати да је стварно p–вредност.

в) За узорак обима 5 у коме је $\sum x_i = 4.2$ и праг значајности 0.05, донети закључак на основу реализоване p–вредности теста.

Решење. а) Да би тест са наведеном тест функцијом био мере α треба да важи

$$\alpha = P_{H_0}\left\{\varphi(\mathbf{X}) = 1\right\} = P_{H_0}\left\{\sum_{i=1}^{n} X_i \leq k\right\}.$$

Пошто $X_i \in \mathcal{E}(\lambda)$, онда $\sum_{i=1}^{n} X_i \in \gamma(n, \lambda)$. Под H_0, $\lambda = 1$, па $\sum_{i=1}^{n} X_i \in \gamma(n, 1)$, односно $2\sum_{i=1}^{n} X_i \in \chi^2_{2n}$.

Како је

$$\alpha = P_{H_0}\left\{2\sum_{i=1}^{n} X_i \leq 2k\right\},$$

следи да је $k = \frac{1}{2}F^{-1}_{\chi^2_{2n}}(\alpha)$.

б) Нека је $T(\mathbf{X}) = \sum_{i=1}^{n} X_i$. Тада је p–вредност теста одређена са

$$\begin{aligned} p(x) &= \sup_{H_0} P_\lambda\left\{T(\mathbf{X}) \leq T(x)\right\} = \sup_{\lambda=1} P_\lambda\left\{\sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} x_i\right\} \\ &= P_{\lambda=1}\left\{2\sum_{i=1}^{n} X_i \leq 2\sum_{i=1}^{n} x_i\right\} = F_{\chi^2_{2n}}\left(2\sum_{i=1}^{n} x_i\right). \end{aligned}$$

Докажимо још да је то заиста p–вредност теста. Треба да покажемо да за свако λ из параметарског простора нулте хипотезе (заправо само $\lambda = 1$) и свако α важи (**). Приметимо да:

$$\begin{aligned} P_\lambda\left\{p(\mathbf{X}) \leq \alpha\right\} &= P_\lambda\left\{F_{\chi^2_{2n}}\left(2\sum_{i=1}^{n} X_i\right) \leq \alpha\right\} = P_\lambda\left\{2\sum_{i=1}^{n} X_i \leq F^{-1}_{\chi^2_{2n}}(\alpha)\right\} \\ &= P_\lambda\left\{2\sum_{i=1}^{n} X_i \leq 2k\right\} = \alpha. \end{aligned}$$

Дакле, $p(\mathbf{X})$ заиста јесте p–вредност теста.

в) Из дела (б), следи да је p–вредност теста

$$p(x) = F_{\chi^2_{2n}}\left(2\sum_{i=1}^{n} x_i\right) = F_{\chi^2_{10}}(2 \cdot 4.2) = 0.41.$$

✓

Дакле, $p(x) > \alpha$, па се не одбацује нулта хипотеза.

Пошто за обележје из нормалне расподеле постоје статистике које се могу повезати са параметром тако да се добије случајна величина за коју знамо расподелу, често (као што је био случај и код интервала поверења) користимо њихова својства да бисмо формирали тестове. У наредним задацима користићемо критичне области које се добијају на основу тих случајних величина чије расподеле су познате када узорак потиче из нормалне расподеле и тада важи $W = \{T_n < c\}$ или $W = \{T_n > c\}$ или $W = \{|T_n| > c\}$, а знак у критичној области зависи од алтернативне хипотезе и слагаће се са знаком у њој.

Задатак 3. Из популације чије обележје X има нормалну $\mathcal{N}(m, \sigma^2)$ расподелу извучен је узорак обима 20 и констатовано је да је узорачка средина $\bar{x}_{20} = 53.5$, а узорачка дисперзија $\bar{s}_{20}^2 = 45.85$. Са прагом значајности 0.05 тестирати:

а) хипотезу $H_0(m = 60)$;

б) хипотезу $H_0(\sigma^2 = 50)$ против хипотезе $H_1(\sigma^2 < 50)$.

Решење. а) Тестирамо хипотезу $H_0(m = 60)$ против алтернативне $H_1(m \neq 60)$. У задатку није наведена алтернативна хипотеза, па се подразумева она која обухвата све оне вредности параметарског простора које се не налазе у нултој хипотези. За овако дефинисану алтернативну хипотезу, с обзиром да дисперзија није позната, критична област је облика

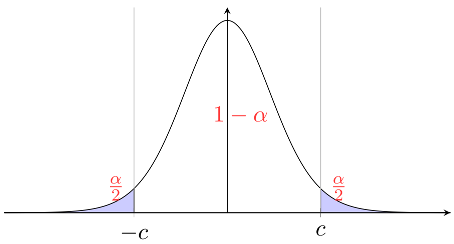
$$W = \left\{ \left| \frac{\overline{x}_n - m_0}{\overline{s}_n} \sqrt{n-1} \right| \geq c \right\},$$

где је m_0 вредност из нулте хипотезе, а константу c одређујемо тако да важи $P_{H_0}\{\mathbf{X}_n \in W\} = \alpha$.

Односно,

$$P_{H_0} \left\{ \left| \frac{\overline{X}_n - m_0}{\overline{S}_n} \sqrt{n-1} \right| \geq c \right\} = \alpha.$$

Како, при нултој хипотези, $\frac{\overline{X}_n - m_0}{\overline{S}_n} \sqrt{n-1}$ има Студентову t_{n-1} расподелу, из таблица те расподеле добијамо да је



$$c = t_{n-1;\alpha} = t_{19;0.05} = 2.093.$$

Дакле, критична област је

$$W = \left\{ \left| \frac{\overline{x}_n - 60}{\overline{s}_n} \sqrt{n-1} \right| \geq 2.093 \right\}.$$

На основу реализованог узорка имамо да је $\frac{\overline{x}_n - 60}{\overline{s}_n} \sqrt{n-1} = 4.18428$. Приметимо да је $4.18428 \geq 2.093$, што значи да се реализовани узорак налази у критичној области, па одбацујемо H_0.

Напомена:

Уколико је алтернативна хиотеза $H_1(m > 60)$, критична област је облика

$$W = \left\{ \frac{\overline{x}_n - m_0}{\overline{s}_n} \sqrt{n-1} > c \right\}.$$

Уколико је алтернативна хипотеза $H_1(m < 60)$, критична област је облика

$$W = \left\{ \frac{\overline{x}_n - m_0}{\overline{s}_n} \sqrt{n-1} < c \right\}.$$

б) Тестирамо хипотезу $H_0(\sigma^2 = 50)$ против алтернативне $H_1(\sigma^2 < 50)$. За овако дефинисану алтернативну хипотезу критична област је облика

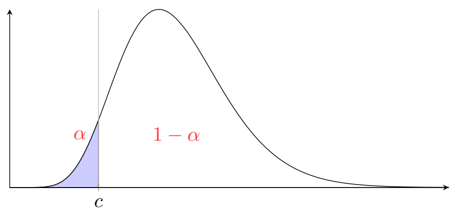
$$W = \left\{ \frac{n\overline{s}_n^2}{\sigma_0^2} \leq c \right\},$$

где је σ_0^2 вредност из нулте хипотезе, а константа c се одређује тако да

$$P_{H_0}\{\mathbf{X}_n \in W\} = \alpha.$$

Односно

$$P_{H_0}\left\{\frac{n\overline{S}_n^2}{\sigma^2} \leq c\right\} = \alpha.$$



Познато је да $\frac{n\overline{S}_n^2}{\sigma^2}$ има χ^2_{n-1} расподелу, па из таблица за χ^2 расподелу добијамо да је $c = \chi^2_{n-1;1-\alpha} = \chi^2_{19;0.95} = 10.12$.

Дакле критична област је

$$W = \left\{\frac{n\overline{s}_n^2}{\sigma_0^2} \leq 10.12\right\}.$$

На основу података из задатка имамо да је $\frac{n\overline{s}_n^2}{\sigma_0^2} = 18.34 > 10.12$, па не одбацујемо H_0, јер се реализовани узорак не налази у критичној области.

✓

Задатак 4. Из две популације чија су обележја X које има нормалну $\mathcal{N}(m_1, 6^2)$ расподелу и Y које има нормалну $\mathcal{N}(m_2, 5^2)$ расподелу извучени су независни узорци обима 12 и 10 и констатовано је да су узорачке средине $\bar{x}_{12} = 178$ и $\bar{y}_{10} = 176.6$. Са прагом значајности 0.04 тестирати хипотезу $H_0(m_1 = m_2)$ против хипотезе $H_1(m_1 > m_2)$.

Решење. Како је потребно да тестирамо једнакост очекивања два обележја $H_0(m_1 = m_2)$ против алтернативне $H_1(m_1 > m_2)$, посматраћемо критичну област облика

$$W = \{\overline{x}_{12} - \overline{y}_{10} \geq c\}.$$

Дакле, желимо да одбацимо нулту хипотезу о једнакости средњих вредности обележја X и Y, ако се узорачке средине разликују за више од неке унапред одређене вредности.

Како су случајне величине $\overline{X}_{12}$ и $\overline{Y}_{10}$ независне и $\overline{X}_{12} \in \mathcal{N}(m_1, \frac{6^2}{12})$, $\overline{Y}_{10} \in \mathcal{N}(m_2, \frac{5^2}{10})$, то $\overline{X}_{12} - \overline{Y}_{10} \in \mathcal{N}(m_1 - m_2, 5.5)$.

Формирамо критичну област тако да важи $P_{H_0}\{\mathbf{X}_n \in W\} = \alpha$. Праг значајности је 0.04, па важи једнакост

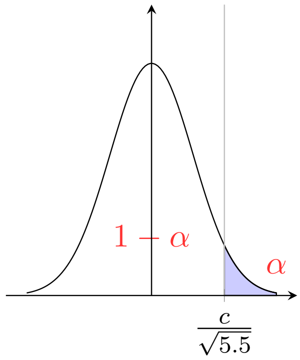
$$0.04 = P_{H_0}\left\{\overline{X}_{12} - \overline{Y}_{10} \geq c\right\} = P_{H_0}\left\{\frac{\overline{X}_{12} - \overline{Y}_{10}}{\sqrt{5.5}} \geq \frac{c}{\sqrt{5.5}}\right\}.$$

С обзиром да статистика на левој страни неједнакости има стандардну нормалну расподелу, из таблица за нормалну расподелу добијамо да је

$$\frac{c}{\sqrt{5.5}} = \Phi^{-1}(0.96) = \Phi_0^{-1}(0.46) = 1.75,$$

односно $c = 4.104$. Критична област је

$$W = \{\overline{x}_{12} - \overline{y}_{10} \geq 4.104\}.$$



На основу података из задатка имамо да је $\overline{x}_{12} - \overline{y}_{10} = 1.4 < 4.104$, па реализовани узорци нису у критичној области и не одбацујемо нулту хипотезу.

✓

Задатак 5. Из две популације чија су обележја X које има нормалну $\mathcal{N}(m_1, \sigma_1^2)$ расподелу и Y које има нормалну $\mathcal{N}(m_2, \sigma_2^2)$ расподелу извучени су независни узорци обима 8 и 10 и констатовано је да су узорачке дисперзије $\overline{s}_8^2(X) = 46$ и $\overline{s}_{10}^2(Y) = 50$. Са прагом значајности 0.1 тестирати хипотезу $H_0(\sigma_1^2 = \sigma_2^2)$.

Решење. Није наведена алтернативна хипотеза, па претпостављамо најопштији случај, односно тестирамо $H_0(\sigma_1^2 = \sigma_2^2)$ против $H_1(\sigma_1^2 \neq \sigma_2^2)$.

Добра оцена за дисперзију обележја на популацији је узорачка дисперзија, па можемо покушати да формирамо критичну област на следећи начин:

$$W = \{|\overline{s}_{n_1}^2(\mathbf{x}) - \overline{s}_{n_2}^2(\mathbf{y})| \geq c\}$$

Дакле, желимо да одбацимо нулту хипотезу о једнакости дисперзија обележја X и Y, ако се узорачке дисперзије разликују више од неке унапред одређене вредности c. За овакву критичну област, вредност c бисмо рачунали из следеће једнакости

$$P_{H_0}\{|\overline{S}_{n_1}^2(X) - \overline{S}_{n_2}^2(Y)| \geq c\} = 0.1.$$

Међутим, пошто не знамо расподелу статистике $|\overline{S}_{n_1}^2(X) - \overline{S}_{n_2}^2(Y)|$, нити било које њене трансформације која би нам омогућила лак рачун, не можемо одредити c.

Покушаћемо да формирамо критичну област на другачији начин. Познато је да

$$\frac{n_1\overline{S}_{n_1}^2(X)\sigma_2^2(n_2-1)}{n_2\overline{S}_{n_2}^2(Y)\sigma_1^2(n_1-1)} \in \mathcal{F}_{n_1-1,n_2-1},$$

па при H_0

$$\frac{n_1\overline{S}_{n_1}^2(X)(n_2-1)}{n_2\overline{S}_{n_2}^2(Y)(n_1-1)} = \frac{8\cdot\overline{S}_8^2(X)\cdot 9}{10\cdot\overline{S}_{10}^2(Y)\cdot 7} = \frac{72\overline{S}_8^2(X)}{70\overline{S}_{10}^2(Y)} \in \mathcal{F}_{7,9}.$$

Искористићемо ову чињеницу да формирамо критичну област. Желимо да одбацимо нулту хипотезу о једнакости дисперзија, ако однос узорачких дисперзија није близак 1, односно ако је мањи од неке вредности c_1 или већи од неке вредности c_2. Формирамо критичну област

$$W = \left\{\frac{\overline{s}_8^2(\mathbf{x})}{\overline{s}_{10}^2(\mathbf{y})} \leq c_1\right\} \cup \left\{\frac{\overline{s}_8^2(\mathbf{x})}{\overline{s}_{10}^2(\mathbf{y})} \geq c_2\right\}.$$

Вредности константи c_1 и c_2 можемо одредити из једнакости:

$$P_{H_0}\left\{\left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \leq c_1\right\} \cup \left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \geq c_2\right\}\right\} = 0.1.$$

Ако није другачије наглашено, стандардно је да се претпостави да су вероватноће скупова у унији једнаке, тј.

$$P_{H_0}\left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \leq c_1\right\} = 0.05 \quad \text{и} \quad P_{H_0}\left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \geq c_2\right\} = 0.05.$$

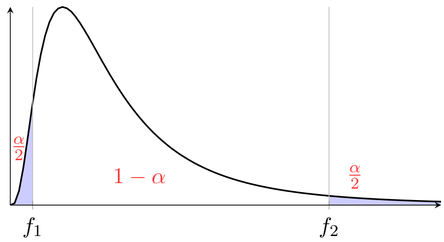
Користећи чињеницу да

$$\frac{72\overline{S}_8^2(X)}{70\overline{S}_{10}^2(Y)} \in \mathcal{F}_{7,9},$$

па

$$\frac{70\overline{S}_{10}^2(Y)}{72\overline{S}_8^2(X)} \in \mathcal{F}_{9,7},$$

коначно можемо израчунати вредности константи c_1 и c_2:



$$P_{H_0}\left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \leq c_1\right\} = P_{H_0}\left\{\frac{\overline{S}_{10}^2(Y)}{\overline{S}_8^2(X)} \geq \frac{1}{c_1}\right\} = P_{H_0}\left\{\frac{70\overline{S}_{10}^2(Y)}{72\overline{S}_8^2(X)} \geq \frac{70}{72}\frac{1}{c_1}\right\} = 0.05$$

$$\Longrightarrow \frac{70}{72}\frac{1}{c_1} = F^{-1}_{\mathcal{F}_{9,7}}(0.95) = 3.7$$

$$\Longrightarrow c_1 = \frac{70}{72}\frac{1}{3.7} = 0.26,$$

и

$$P_{H_0}\left\{\frac{\overline{S}_8^2(X)}{\overline{S}_{10}^2(Y)} \geq c_2\right\} = P_{H_0}\left\{\frac{72\overline{S}_8^2(X)}{70\overline{S}_{10}^2(Y)} \geq \frac{72}{70}c_2\right\} = 0.05$$

$$\Longrightarrow \frac{72}{70}c_2 = F^{-1}_{\mathcal{F}_{7,9}}(0.95) = 3.3$$

$$\Longrightarrow c_2 = \frac{70}{72}3.3 = 3.21.$$

Дакле, критична област је

$$W = \left\{\frac{\overline{s}_8^2(\mathbf{x})}{\overline{s}_{10}^2(\mathbf{y})} \leq 0.26\right\} \cup \left\{\frac{\overline{s}_8^2(\mathbf{x})}{\overline{s}_{10}^2(\mathbf{y})} \geq 3.21\right\}.$$

Сада можемо да проверимо да ли добијени узорци упадају у критичну област и да одлучимо да ли ћемо одбацити нулту хипотезу. На добијеним узорцима

$$\frac{\overline{s}_8^2(\mathbf{x})}{\overline{s}_{10}^2(\mathbf{y})} = \frac{46}{50} = 0.92,$$

па вредност не упада у критичну област, те не одбацујемо хипотезу о једнакости дисперзија. ✓

Задатак 6. Мерен је систолни (горњи) притисак на узорку од 12 мушкараца и добијено је 130, 148, 122, 140, 132, 142, 124, 150, 170, 136, 146, 140, а на узорку од 13 жена добијене су следеће вредности 140, 150, 130, 132, 150, 138, 123, 124, 160, 138, 170, 144, 108. Сматра се да систолни притисак и код мушкараца и код жена има нормалну расподелу. Ако се претпостави да су дисперзије једнаке, са прагом значајности $\alpha = 0.1$, тестирати хипотезу да су средње вредности притисака мушкараца и жена једнаке против алтернативе да се разликују.

Решење. Нека је X обележје које представља систолни притисак код мушкараца, а Y обележје које представља систолни притисак код жена. Претпоставља се да обележја X и Y имају нормалну расподелу са једнаким дисперзијама. Дакле, претпостављамо да

$$X \in \mathcal{N}(m_1, \sigma^2) \qquad \text{и} \qquad Y \in \mathcal{N}(m_2, \sigma^2).$$

На основу добијених (независних) узорака обима $n_1 = 12$ и $n_2 = 13$, желимо да тестирамо хипотезу да је средња вредност обележја X једнака средњој вредности обележја Y, против алтернативе да се разликују. Односно, тестирамо $H_0(m_1 = m_2)$ против $H_1(m_1 \neq m_2)$.

Није нам позната вредност дисперзије σ^2, али знамо да:

$$\frac{(\overline{X}_{n_1} - m_1) - (\overline{Y}_{n_2} - m_2)}{\sqrt{n_1\overline{S}^2_{n_1}(X) + n_2\overline{S}^2_{n_2}(Y)}}\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}(n_1 + n_2 - 2)} \in t_{n_1+n_2-2},$$

па можемо да користимо тест статистику:

$$T = \frac{\overline{X}_{12} - \overline{Y}_{13}}{\sqrt{12\overline{S}^2_{12}(X) + 13\overline{S}^2_{13}(Y)}}\sqrt{\frac{12 \cdot 13}{12 + 13}(12 + 13 - 2)},$$

која, при H_0, има Студентову t_{23} расподелу.

Пошто је расподела статистике T симетрична око нуле, формирамо критичну област на следећи начин:

$$W = \{|T| \geq c\}.$$

Вредност константе c можемо да добијемо из задатог прага значајности и познате расподеле тест статистике при H_0 тако да важи

$$P_{H_0}\{|T| \geq c\} = 0.1.$$

Одатле је

$$c = t_{23;0.1} = 1.714.$$

Дакле, критична област је облика:

$$W = \{|T| \geq 1.714\}.$$

Када смо формирали критичну област, рачунамо реализовану вредност тест статистике за добијене узорке да бисмо одлучили да ли ћемо прихватити или одбацити хипотезу о једнакости очекивања:

$$\overline{x}_{12} = \frac{1}{12}\sum_{i=1}^{12} x_i = 140,$$

$$\overline{s}^2_{12}(x) = \frac{1}{12}\sum_{i=1}^{12}(x_i - 140)^2 = 155.34,$$

$$\overline{y}_{13} = \frac{1}{13}\sum_{i=1}^{13} y_i = 139,$$

$$\overline{s}^2_{13}(y) = \frac{1}{13}\sum_{i=1}^{13}(y_i - 139)^2 = 249.54.$$

Дакле, реализована вредност тест статистике је

$$\frac{140-139}{\sqrt{12\cdot 155.34 + 13\cdot 249.54}}\sqrt{\frac{12\cdot 13}{25}23} = \frac{11.98}{11.47} = 0.168.$$

Проверавамо да ли реализована вредност тест статистике упада у критичну област,

$$|0.168| = 0.168 \notin W.$$

Закључујемо да реализована вредност тест статистике не упада у критичну област, па на основу добијених узорака не одбацујемо хипотезу H_0 о једнакости очекивања.

✓