Интервали поверења

Нека је θ непознати параметар у расподели обележја X и нека је $(X_1, ..., X_n)$ прост случајан узорак обима n за посматрано обележје. Нека су U_n и V_n статистике дефинисане на основу узорка такве да је $P\{U_n \leq V_n\} = 1$ за које важи

$$P\{U_n \leq \theta \leq V_n\} = \beta.$$

Интервал $[U_n, V_n]$ се назива $\beta\%$ двострани интервал поверења параметра θ, а β је ниво поверења.

Задатак 1. Из популације чије обележје X има нормалну $\mathcal{N}(m, \sigma^2)$ расподелу извучен је узорак обима 10, чија је узорачка средина $\bar{x}_{10} = 5.5$, а узорачка дисперзија $\bar{s}_{10}^2 = 36$.

а) Одредити 90% интервал поверења за непознати параметар m.

б) Одредити 90% једнострани (доњи, горњи) и двострани интервал поверења за непознати параметар σ^2, као и за непознати параметар σ.

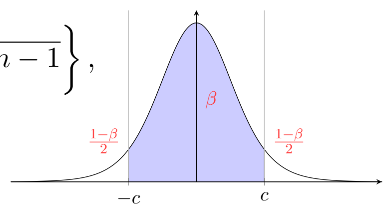
Решење. а) Приметимо прво да је σ^2 непознато, па ћемо користити случајну величину $T_n = \frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1}$, која има Студентову t_{n-1} расподелу.

Потребно је да одредимо U_n и V_n такве да важи

$$\begin{aligned} & P\{U_n \leq m \leq V_n\} = \beta \\ \iff & P\left\{\overline{X}_n - V_n \leq \overline{X}_n - m \leq \overline{X}_n - U_n\right\} = \beta \\ \iff & P\left\{\frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1} \leq \frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1} \leq \frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}\right\} = \beta \end{aligned}$$

Једино што знамо у овом случају је да је вероватноћа између $\frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1}$ и $\frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}$ једнака β, међутим њихов положај може бити произвољан. Због тога се, у случају да није другачије наглашено, увек узимају интервали такви да

$$P\left\{T_n < \frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1}\right\} = P\left\{T_n > \frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}\right\},$$



а пошто је Студентова расподела симетрична око 0 то важи да су ове вредности једнаке по апсолутној вредности.

Користећи да T_n има Студентову расподелу добијамо да је

$$\begin{aligned} \frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1} &= -t_{n-1,1-\beta}, \\ \frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1} &= t_{n-1,1-\beta}. \end{aligned}$$

Према томе,

$$U_n = \overline{X}_n - \frac{\overline{S}_n}{\sqrt{n-1}} t_{n-1,1-\beta},$$

$$V_n = \overline{X}_n + \frac{\overline{S}_n}{\sqrt{n-1}} t_{n-1,1-\beta}.$$

Односно интервал поверења је

$$I_m = \left[\overline{X}_n - \frac{\overline{S}_n}{\sqrt{n-1}} t_{n-1,1-\beta}, \overline{X}_n + \frac{\overline{S}_n}{\sqrt{n-1}} t_{n-1,1-\beta} \right].$$

У конкретном случају, $\beta = 0.9$ и $t_{9,0.1} = 1.833$. Тражени интервал је

$$I_m = \left[5.5 - 1.833 \cdot \frac{6}{3}, 5.5 + 1.833 \cdot \frac{6}{3} \right] = [1.834, 9.166].$$

б) Прво приметимо да је m непознато, па ћемо користити случајну величину $\frac{n\overline{S}_n^2}{\sigma^2}$, која има χ^2_{n-1} расподелу.

Једнострани доњи интервал $(0, V_n]$: Потребно је да одредимо V_n такво да важи

$$P\{\sigma^2 \le V_n\} = \beta \iff P\left\{ \frac{1}{V_n} \le \frac{1}{\sigma^2} \right\} = \beta$$
$$\iff P\left\{ \frac{n\overline{S}_n^2}{V_n} \le \frac{n\overline{S}_n^2}{\sigma^2} \right\} = \beta.$$

Из таблица за χ^2 расподелу добијамо да је $\frac{n\overline{S}_n^2}{V_n} = \chi^2_{n-1;\beta}$. Према томе $V_n = \frac{n\overline{S}_n^2}{\chi^2_{n-1;\beta}}$. Интервал поверења непознатог параметра σ^2 је

$$I_{\sigma^2} = \left(0, \frac{n\overline{S}_n^2}{\chi^2_{n-1;\beta}} \right].$$

У конкретном случају, $\beta = 0.9$ и $\chi^2_{9;0.9} = 4.168$, па је тражени интервал $I_{\sigma^2} = (0, 86.3724]$. Интервал поверења непознатог параметра σ се добија узимањем корена из леве и десне границе претходног интервала па се добија $I_\sigma = (0, 9.2937]$.

Једнострани горњи интервал $[U_n, +\infty)$: Потребно је да одредимо U_n такво да важи $P\{\sigma^2 \ge U_n\} = \beta$, односно

$$P\left\{ \frac{n\overline{S}_n^2}{U_n} \ge \frac{n\overline{S}_n^2}{\sigma^2} \right\} = \beta.$$

Из таблица за χ^2 расподелу добијамо да је $\frac{n\overline{S}_n^2}{U_n} = \chi^2_{n-1;1-\beta}$. Према томе $U_n = \frac{n\overline{S}_n^2}{\chi^2_{n-1;1-\beta}}$. Интервал поверења непознатог параметра σ^2 је

$$I_{\sigma^2} = \left[\frac{n\overline{S}_n^2}{\chi^2_{n-1;1-\beta}}, \infty \right).$$

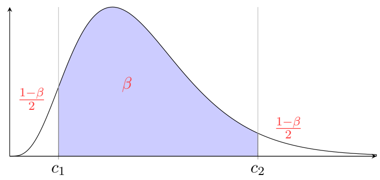
У конкретном случају, $\beta = 0.9$ и $\chi^2_{n-1;1-\beta} = 14.684$, па је тражени интервал поверења $I_{\sigma^2} = [24.5165, +\infty)$. Интервал поверења непознатог параметра σ је $I_\sigma = [4.9514, +\infty)$.

Двострани интервал $[U'_n, V'_n]$: Потребно је да одредимо U'_n и V'_n такве да важи $P\{U'_n \leq \sigma^2 \leq V'_n\} = \beta$, односно

$$P\left\{\frac{n\overline{S}_n^2}{V'_n} \leq \frac{n\overline{S}_n^2}{\sigma^2} \leq \frac{n\overline{S}_n^2}{U'_n}\right\} = \beta.$$

Као и у делу а), интервал ћемо бирати тако да

$$P\left\{\frac{n\overline{S}_n^2}{\sigma^2} > \frac{n\overline{S}_n^2}{U'_n}\right\} = P\left\{\frac{n\overline{S}_n^2}{\sigma^2} < \frac{n\overline{S}_n^2}{V'_n}\right\} = \frac{1-\beta}{2}.$$



Из таблица за χ^2 расподелу добијамо да је $\frac{n\overline{S}_n^2}{U'_n} = \chi^2_{n-1;\frac{1-\beta}{2}}$ и $\frac{n\overline{S}_n^2}{V'_n} = \chi^2_{n-1;\frac{1+\beta}{2}}$. Према томе $U'_n = \frac{n\overline{S}_n^2}{\chi^2_{n-1;\frac{1-\beta}{2}}}$ и $V'_n = \frac{n\overline{S}_n^2}{\chi^2_{n-1;\frac{1+\beta}{2}}}$. Интервал поверења непознатог параметра σ^2 је

$$I_{\sigma^2} = \left[\frac{n\overline{S}_n^2}{\chi^2_{n-1;\frac{1-\beta}{2}}}, \frac{n\overline{S}_n^2}{\chi^2_{n-1;\frac{1+\beta}{2}}}\right].$$

У конкретном случају, $\beta = 0.9$ и $\chi^2_{9;0.05} = 16.919$ и $\chi^2_{9;0.95} = 3.325$. Интервал поверења је $I_{\sigma^2} = [21.2778, 108.2707]$, а интервал поверења непознатог параметра σ је $I_\sigma = [4.6128, 10.4053]$.

✓

Задатак 2. Из популације чије обележје X има нормалну $\mathcal{N}(m, \sigma^2)$ расподелу извучен је узорак:

(-5, -3]	(-3, -1]	(-1, -0.5]	(-0.5, 0.5]	(0.5, 1.5]	(1.5, 3.5)
3	13	56	100	60	8

На основу тог узорка добијено је да је $(c, 0.13756)$ 95.8%-ни интервал поверења за непознати параметар m. Израчунати реалан број c.

Решење. Како је σ^2 непознато, користићемо случајну величину $\frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1}$, која има Студентову t_{n-1} расподелу.

Одредимо прво реализоване вредности статистика. С обзиром да немамо тачне вредности елемената узорка, већ само број елемената у одговарајућем интервалу, можемо сматрати да су елементи случајно расподељени унутар интервала и да је једнака вероватноћа да се налазе са леве и десне стране средине интервала, па можемо њом апроксимирати вредности.

На основу података из задатка можемо одредити узорачку средину и узорачку дисперзију:

$$\overline{x}_{240} = \frac{3 \cdot (-4) + 13 \cdot (-2) + 56 \cdot (-0.75) + 100 \cdot 0 + 60 \cdot 1 + 8 \cdot 2.5}{240} = 0,$$

$$\overline{s}^2_{240} = \frac{3 \cdot (-4)^2 + 13 \cdot (-2)^2 + 56 \cdot 0.75^2 + 60 \cdot 1^2 + 8 \cdot 2.5^2}{240} = 1.00625.$$

Одредимо $\beta\%-$ни интервал поверења за m у општем случају. Потребно је да одредимо U_n и V_n такве да важи

$$\begin{aligned}
&P\{U_n \leq m \leq V_n\} = \beta \\
\Longleftrightarrow\ &P\left\{\overline{X}_n - V_n \leq \overline{X}_n - m \leq \overline{X}_n - U_n\right\} = \beta \\
\Longleftrightarrow\ &P\left\{\frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1} \leq \frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1} \leq \frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}\right\} = \beta \\
\Longleftrightarrow\ &F_{t_{n-1}}\left(\frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}\right) - F_{t_{n-1}}\left(\frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1}\right) = \beta.
\end{aligned}$$

Како је n довољно велико $(n > 30)$ одговарајуће вредности се не могу прочитати из таблица Студентове расподеле. У том случају, Студентову расподелу можемо апроксимирати стандардном нормалном расподелом. Односно, горњи израз је еквивалентан са

$$\Phi\left(\frac{\overline{X}_n - U_n}{\overline{S}_n}\sqrt{n-1}\right) - \Phi\left(\frac{\overline{X}_n - V_n}{\overline{S}_n}\sqrt{n-1}\right) = \beta.$$

Како је дата лева граница полазног интеграла, не можемо претпоставити да је у питању симетричан интервал. Према томе, остале прорачуне морамо радити за конкретан узорак.

Из таблица за нормалну расподелу и за реализовану вредност узорка, имамо да је $\Phi(\frac{-0.13756}{\sqrt{1.00625}}\sqrt{239}) = \Phi(-2.12) = 0.017$. Према томе

$$\Phi\left(\frac{\overline{x}_n - c}{\overline{s}_n}\sqrt{n-1}\right) \approx 0.958 + 0.017 = 0.975.$$

Даље из таблица за нормалну расподелу добијамо да је

$$\frac{\overline{x}_n - c}{\overline{s}_n}\sqrt{n-1} = \Phi^{-1}(0.975) = 1.96,$$

па добијамо да је $c = -0.12718$. ✓

Задатак 3. Из популације чије обележје X има нормалну $\mathcal{N}(m, 16)$ расподелу извучен је узорак обима 64, чија је узорачка средина $\bar{x}_{64} = 5$. Константовано је да је $\beta\%$ интервал поверења за m једнак (4,6). Израчунати β.

Решење. Пошто је дисперзија расподеле позната, можемо користити случајну величину $X^\star = \frac{\overline{X}_n - m}{\sigma}\sqrt{n}$ која има стандардну нормалну расподелу. На основу тога добија се

$$\begin{aligned}
\beta &= P\{U_n < m < V_n\} \\
\beta &= P\left\{\overline{X}_n - V_n < \overline{X}_n - m < \overline{X}_n - U_n\right\} \\
\beta &= P\left\{\frac{\overline{X}_n - V_n}{\sigma}\sqrt{n} < \frac{\overline{X}_n - m}{\sigma}\sqrt{n} < \frac{\overline{X}_n - U_n}{\sigma}\sqrt{n}\right\}.
\end{aligned}$$

За реализовани узорак, следи

$$\begin{aligned}
\beta &= P\left\{\frac{5-6}{4}\sqrt{64} < X^\star < \frac{5-4}{4}\sqrt{64}\right\} = \Phi(2) - \Phi(-2) \\
&= \Phi(2) - (1 - \Phi(2)) = 2\Phi(2) - 1 = 2(0.5 + 0.47725) - 1 = 0.9545.
\end{aligned}$$

Дакле, $\beta = 0.9545$, односно 95.45%. ✓

Задатак 4. Обележје X има униформну $\mathcal{U}[0, 1+\theta], \theta > 0$, расподелу, где је θ непознати параметар. Узет је узорак обима 200 и константовано је да је у узорку 150 елемената који су мањи од 1.

а) Одредити 95% интервал поверења за вероватноћу p, где је $p = P\{X < 1\}$.

б) На основу резултата под (а) одредити 95% интервал поверења за θ.

Решење. а) Уведимо случајну величину Y (индикатор) која узима вредност 1 ако случајна величина X узме вредност мању од један, односно

$$Y : \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

Нас интересује укупан број елемената који је задовољио наведени услов, односно сума претходних индикатора који баш говоре да ли је услов задовољен или не.

Како је узет узорак обима 200 (> 30) можемо да искористимо централну граничну теорему:

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

С обзиром да је

$$E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} E(Y_i) = np$$

и

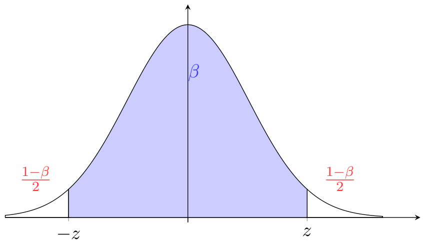
$$D\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} D(Y_i) = np(1-p),$$

интервал поверења за вероватноћу p добијамо из:

$$P\left\{\left|\frac{\sum\limits_{i=1}^{n} Y_i - 200p}{\sqrt{200p(1-p))}}\right| < z\right\} = 0.95.$$

$$P\left\{-z < \frac{\sum\limits_{i=1}^{n} Y_i - 200p}{\sqrt{200p(1-p))}} < z\right\} = 0.95.$$

Из задатка је познато да је $\sum\limits_{i=1}^{200} y_i = 150$.



Из таблица за нормалну расподелу имамо да је $z = \Phi^{-1}(0.95 + \frac{1-0.95}{2}) = 1.96$, па решавањем неједначине

$$\frac{(150 - 200p)^2}{200p(1-p)} < 1.96^2,$$

добијамо да је $I_p = (0.686, 0.805)$.

б) Користећи део а) и из једнакости $p = P\{X < 1\} = \frac{1}{1+\theta}$ имамо да је

$$\begin{aligned} 0.686 &< p < 0.805 \\ 0.686 &< \frac{1}{1+\theta} < 0.805 \\ 0.242 &< \theta < 0.458 \end{aligned}$$

тј. $I_\theta = (0.242, 0.458)$.

✓

Задатак 5. Из популације чије обележје X има експоненцијалну $\mathcal{E}(\lambda)$ расподелу извучен је узорак:

I_k	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)	[5,6)	$[6,\infty)$
n_k	493	378	298	211	171	45	4

Одредити 98% интервал поверења за непознати параметар λ.

Решење. Узет је узорак обима $n = \sum n_k = 1600$ (>30), па можемо да искористимо централну граничну теорему:

$$\frac{\sum\limits_{i=1}^{n} X_i - E\left(\sum\limits_{i=1}^{n} X_i\right)}{\sqrt{D\left(\sum\limits_{i=1}^{n} X_i\right)}} \xrightarrow{D} X^*, \quad X^* \in \mathcal{N}(0,1).$$

Користећи податке из задатка можемо одредити збир вредности из узорака

$$\sum_{k=1}^{n} x_k = 493 \cdot 0.5 + 378 \cdot 1.5 + 298 \cdot 2.5 + 211 \cdot 3.5 + 171 \cdot 4.5 + 45 \cdot 5.5 + 4 \cdot 6.5 = 3340.$$

С обзиром да је $EX_1 = \frac{1}{\lambda}$ и $DX_1 = \frac{1}{\lambda^2}$, интервал поверења за параметар λ добијамо из једнакости:

$$P\left\{\left|\frac{\sum\limits_{i=1}^{1600} X_i - \frac{1600}{\lambda}}{\sqrt{\frac{1600}{\lambda^2}}}\right| < z\right\} = 0.98.$$

Из таблица за нормалну расподелу имамо да је $z = \Phi^{-1}(0.99) = 2.33$. Решавањем неједначине

$$-2.33 < \frac{3340 - \frac{1600}{\lambda}}{\frac{40}{\lambda}} < 2.33,$$

добијамо да је $I_\lambda = (0.451, 0.507)$. ✓

Задатак 6. За густину расподеле обележја X, које представља ефикасност новог лека, важи да је $f(x;\theta) = \theta x^{\theta-1}$, $x \in (0,1)$, $\theta > 0$. Ако је на узорку од 300 пацијената код њих 250 нови лек имао ефикасност већу од 0.5, формирати 95% интервал поверења за параметар θ.

Решење. Уочимо да за $x \in (0,1)$ важи да је

$$P\{X \le x\} = \int_0^x \theta t^{\theta-1}\, dt = \theta \left(\frac{t^\theta}{\theta}\right)\Bigg|_{t=0}^{t=x} = x^\theta.$$

Следи да је $P\{X > x\} = 1 - x^\theta$, па је, специјално, $P\left\{X > \frac{1}{2}\right\} = 1 - 2^{-\theta}$.

Означимо $I_j = I\left\{X_j > \frac{1}{2}\right\}$. Број пацијената код којих је лек имао ефикасност већу од $\frac{1}{2}$ дат је као $\sum_{j=1}^{300} I_j$. Обим узорка је довољно велики, а индикатори I_j су независни и једнако расподељени, те се на њих може применити централна гранична теорема, те ћемо интервал поверења тражити из једнакости:

$$P\left\{-c < \frac{\sum\limits_{j=1}^{300} I_j - E\left(\sum\limits_{j=1}^{300} I_j\right)}{\sqrt{D\left(\sum\limits_{j=1}^{300} I_j\right)}} < c\right\} = 0.95,$$

односно

$$P\left\{-c < \frac{\sum\limits_{j=1}^{300} I_j - 300\cdot(1-2^{-\theta})}{\sqrt{300\cdot 2^{-\theta}(1-2^{-\theta})}} < c\right\} = 0.95.$$

Из таблица за нормалну расподелу имамо да је $c = \Phi^{-1}\left(0.95 + \frac{1-0.95}{2}\right) = \Phi(0.975) = 1.96$. Из задатка је познато да је $\sum_{j=1}^{300} i_j = 250$, па решавањем неједначине

$$-1.96 < \frac{250 - 300\cdot(1-2^{-\theta})}{\sqrt{300\cdot 2^{-\theta}(1-2^{-\theta})}} < 1.96,$$

односно увођењем смене $p = 1 - 2^{-\theta}$

$$\frac{(250-300p)^2}{300p(1-p)} < 1.96^2 = 3.84$$

добијамо да је $I_p = (0.787,\ 0.871)$. Сада из једнакости $p = 1 - 2^{-\theta}$ имамо да је

$$\begin{aligned} 0.787 &< p < 0.871, \\ 0.787 &< 1 - 2^{-\theta} < 0.871, \end{aligned}$$

тј. тражени интервал поверења је $I_\theta = (2.231,\ 2.954)$. ✓

Задатак 7. Број поена на пријемном испиту за упис на Математички факултет има нормалну расподелу. Изабрано је 30 ученика на случајан начин и регистрован је њихов резултат на пријемном испиту:

број поена	[0,20]	(20,40]	(40,60]	(60,80]	(80,100]
број ученика	3	2	9	11	5

На основу овог узорка, одредити 92.5% интервал поверења за просечан број поена свих ученика на пријемном испиту, такав да вероватноћа да просечан број поена свих ученика на пријемном испиту буде мањи од леве границе интервала поверења буде дупло мања од вероватноће да просечан број поена свих ученика на пријемном испиту буде већи од десне границе интервала поверења.

Решење. Нека просечан број поена на пијемном испиту има нормалну $\mathcal{N}(m, \sigma^2)$ расподелу. Како је σ^2 непознат параметар, користићемо случајну величину $T_n = \frac{\bar{X}_n - m}{\bar{S}_n}\sqrt{n-1}$, која има Студентову t_{n-1} расподелу.

Одредимо прво реализоване вредности статистика. С обзиром да немамо тачне вредности елемената узорка, већ само број елемената у одговарајућем интервалу, можемо сматрати да су елементи случајно расподељени унутар интервала и да је једнака вероватноћа да се налазе са леве и десне стране средине интервала, па можемо њом апроксимирати вредности.

На основу података из задатка, можемо одредити узорачку средину и узорачку дисперзију:

$$\bar{x}_{30} = \frac{3 \cdot 10 + 2 \cdot 30 + 9 \cdot 50 + 11 \cdot 70 + 5 \cdot 90}{30},$$

$$\bar{s}_{30}^2 = \frac{3 \cdot (10 - 58.67)^2 + 2 \cdot (30 - 58.67)^2 + 9 \cdot (50 - 58.67)^2 + 11 \cdot (70 - 58.67)^2 + 5 \cdot (90 - 58.67)^2}{30},$$

тј. $\bar{x}_{30} = 58.67$ и $\bar{s}_{30}^2 = 542.98$. Одредимо сада 92.5% интервал поверења за m у општем случају. Потребно је одредити статистике U_n и V_n за које је $P\{U_n \leq V_n\} = 1$ такве да важи

$$P\{U_n \leq m \leq V_n\} = 0.95,$$

$$P\{\bar{X}_n - V_n \leq \bar{X}_n - m \leq \bar{X}_n - U_n\} = 0.925,$$

$$P\left\{\frac{\bar{X}_n - V_n}{\bar{S}_n}\sqrt{n-1} \leq \frac{\bar{X}_n - m}{\bar{S}_n}\sqrt{n-1} \leq \frac{\bar{X}_n - U_n}{\bar{S}_n}\sqrt{n-1}\right\} = 0.925.$$

Из услова задатка имамо да за овакав интервал поверења треба да важи $P\{m < U_n\} = 0.025$ и $P\{m > V_n\} = 0.05$, односно

$$P\left\{T_n < \frac{\bar{X}_n - V_n}{\bar{S}_n}\sqrt{n-1}\right\} = 0.05, \quad P\left\{T_n > \frac{\bar{X}_n - U_n}{\bar{S}_n}\sqrt{n-1}\right\} = 0.025.$$

Користећи да T_n има Студентову расподелу, добијамо да је

$$\frac{\bar{X}_n - V_n}{\bar{S}_n}\sqrt{n-1} = -t_{n-1;2\cdot 0.05},$$

$$\frac{\bar{X}_n - U_n}{\bar{S}_n}\sqrt{n-1} = t_{n-1;2\cdot 0.025}.$$

Према томе,

$$U_n = \bar{X}_n - \frac{\bar{S}_n}{\sqrt{n-1}} t_{n-1;0.05},$$

$$V_n = \bar{X}_n + \frac{\bar{S}_n}{\sqrt{n-1}} t_{n-1;0.1}.$$

Односно интервал поверења је

$$I_m^{(n)} = \left[\bar{X}_n - \frac{\bar{S}_n}{\sqrt{n-1}} t_{n-1;0.05},\ \bar{X}_n + \frac{\bar{S}_n}{\sqrt{n-1}} t_{n-1;0.1}\right].$$

У конкретном случају, $\beta = 0.0.925$, $t_{29;0.05} = 2.045$ и $t_{29;0.1} = 1.699$. Тражени интервал је

$$I_m^{(30)} = \left[58.67 - 2.045 \cdot \frac{\sqrt{542.98}}{\sqrt{29}}, 58.67 + 1.699 \cdot \frac{\sqrt{542.98}}{\sqrt{29}}\right],$$

$$I_m^{(30)} = [49.821, 66.022].$$

✓

Задатак 8. Удружење банака Србије спроводи истраживање како би се испитала разлика просечне изложености банака кредитном и тржишном ризику. Претпоставља се да изложеност кредитном ризику има нормалну $\mathcal{N}(m_1, \sigma^2)$ расподелу, а изложеност тржишном ризику нормалну $\mathcal{N}(m_2, \sigma^2)$ расподелу, где су m_1, m_2 и σ непознати. Измерене су изложености кредитном и тржишном ризику у неких 7 од свих банака које послују на територији Србије. Одредити ниво поверења симетричног у односу на разлику узорачких средина интервала поверења за разлику просечне изложености кредитном и тржишном ризику, ако је дужина тог интервала мања од 2.023σ са вероватноћом 0.98.

Решење. Нека је X обележје које представља изложеност кредитном ризику, а Y обележје које представља изложеност тржишном ризику. Претпоставља се да обележја X и Y имају нормалну расподелу са једнаким дисперзијама. Дисперзије нам нису познате, али знамо да

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

где је $n_1 = n_2 = 7$, па случајна величина T има Студентову t_{12} расподелу. Треба да одредимо ниво поверења β тако да је

$$P\{U_7 < m_1 - m_2 < V_7\} = \beta, \tag{1}$$

где су U_7 и V_7 статистике дефинисане на основу узорка такве да је $P\{U_7 \leq V_7\} = 1$, односно на основу тога добија се

$$P\left\{\bar{X}_7 - \bar{Y}_7 - V_7 < \bar{X}_7 - \bar{Y}_7 - (m_1 - m_2) < \bar{X}_7 - \bar{Y}_7 - U_7\right\} = \beta.$$

$$P\left\{\frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} < \frac{\bar{X}_7 - \bar{Y}_7 - (m_1 - m_2)}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} < \frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6}\right\} = \beta.$$

Знамо да је вероватноћа између $\frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X)+\bar{S}^2_7(Y)}} \cdot \sqrt{6}$ и $\frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X)+\bar{S}^2_7(Y)}} \cdot \sqrt{6}$ једнака β, ме

дј утим њихов положај може бити произвољан. Узимамо вјероватносно симетричан интервал, тј. такав да

$$P\left\{T < \frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6}\right\} = P\left\{T > \frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6}\right\},$$

а пошто је Студентова расподела симетрична око нуле, то важи да су ове вредности једнаке по апсолутној вриједности. Користећи да $T \in t_{12}$ добијамо да је

$$\frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} = -t_{12;1-\beta},$$

$$\frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} = t_{12;1-\beta},$$

Према томе,

$$U_7 = \bar{X}_7 - \bar{Y}_7 - \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)},$$

$$V_7 = \bar{X}_7 - \bar{Y}_7 + \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)},$$

односно имамо да је $\beta\%$ двострани интервал поверења за разлику просечне изложености кредитном и тржишном ризику:

$$I = \left(\bar{X}_7 - \bar{Y}_7 - \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)},\, \bar{X}_7 - \bar{Y}_7 + \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)} \right),$$

па је дужина уоченог интервала поверења једнака

$$d(I) = t_{12;1-\beta} \sqrt{\frac{2}{3}} \sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)}.$$

Из услова задатка важи да је

$$P\{d(I) < 2,023\sigma\} = 0.98,$$

$$P\left\{ t_{12;1-\beta} \sqrt{\frac{2}{3}} \sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)} < 2.023\sigma \right\} = 0.98,$$

односно

$$P\left\{ \frac{\bar{S}_7^2(X) + \bar{S}_7^2(Y)}{\sigma^2} < \frac{2.023^2}{(t_{12;1-\beta})^2} \frac{3}{2} \right\} = 0.98. \tag{2}$$

Како је $\frac{n_1 \bar{S}_{n_1}^2(X)}{\sigma} \in \chi^2_{n_1-1}$ и $\frac{n_2 \bar{S}_{n_2}^2(Y)}{\sigma} \in \chi^2_{n_2-1}$, то из адитивности χ^2 расподеле добијамо да је

$$\frac{n_1 \bar{S}_{n_1}^2(X) + n_2 \bar{S}_{n_2}^2(Y)}{\sigma} \in \chi^2_{n_1+n_2-2},$$

па из (2) следи да је

$$7 \cdot \frac{2.023^2}{(t_{12;1-\beta})^2} \cdot \frac{3}{2} = \chi^2_{12;0.02} = 24.054 \quad \Rightarrow \quad t_{12;1-\beta} = 1.3365.$$

Коначно, из (1) добијамо да је тражени ниво поверења једнак

$$\beta = F_{t_{12}}(1.3365) = 0.8.$$

✓