Узорачке оцјене параметара расподјеле

Расподјелу случајне промјенљиве X карактерише више параметара. Међу њима су и нумерички: математичко очекивање, дисперзија, модус, медијана. Њихове вриједности одређујемо из закона расподјеле када је познат. Ако се ради о случајном узорку на основу кога одређујемо непознате вриједности параметара говоримо о статистичким оцјенама или просто о оцјенама. Неке од оцјена добијамо на следећи начин:

Узорачка средина $\tilde{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Код груписаних података користимо облик

$\tilde{x} = \frac{1}{n}\sum_{i=1}^{n} n_i^* x_i^*$.

Оцјена модуса $\tilde{d}_X$ је елемент који се јавља највише пута. У случају груписаних података када је дужина свих интервала једнака, оцјена модуса добија се по формули $\tilde{d}_X = a_d + b\left(\frac{n_d^* - n_{d-1}^*}{2n_d^* - n_{d-1}^* - n_{d+1}^*}\right)$ гдје је a_d доња граница интервала који садржи највише вриједности, n_d^* број елемената у том интервалу, n_{d-1}^* и n_{d+1}^* број елемената у сусједним интервалима, b дужина сваког интервала.

Оцјена медијане $\tilde{h}_X$ је број који дијели све вриједности на двије групе са једнаким бројем елемената. Ако је обим узорка непаран $n = 2k+1$ онда је то $\tilde{h}_X = x^{(k+1)}$, а ако је обим узорка паран $n = 2k$ онда је $\tilde{h}_X = (x^{(k)} + x^{(k+1)})/2$. Оцјена медијане у случају података груписаних у интервале једнаке дужине b добија се формулом $\tilde{h}_X = a_h + b\left(\frac{n/2 - (n_1^* + n_2^* + \ldots + n_{h-1}^*)}{n_h^*}\right)$ гдје је a_h доња граница интервала који садржи медијану, n_h^* број елемената у том интервалу, $n_1^* + \ldots + n_{h-1}^*$ број елемената у интервалима који се налазе лијево од интервала који садржи медијану, b дужина сваког интервала.

Примјер 1

Одредити узорачку средину, модус и медијану за узорак: 5, 6, 8, 2, 3, 1, 1, 4.

Представимо податке у облику 1, 1, 2, 3, 4, 5, 6, 8. Тада је $\tilde{x} = \frac{1}{8}(1+1+2+3+4+5+6+8) = 3.75$. Све вриједности се јављају по једном осим 1 па је оцјена модуса $\tilde{d}_X = 1$. Медијана је $\tilde{h}_X = (3+4)/2 = 3.5$

Примјер 2

Одредити узорачку средину, модус и медијану за узорак:

Границе	34-36	36-38	38-40	40-42	42-44	44-46
Учестаност	2	3	30	40	20	5

Обим узорка је 100. $\tilde{x} = \frac{35*2 + 37*3 + 39*30 + 41*40 + 43*20 + 45*5}{100} = 40.76$. Модус се налази у интервалу 40-42 који садржи највише елемената.

$\tilde{d}_X = 40 + 2\left(\frac{40-30}{2*40-30-20}\right) \approx 40.67$. Медијана се налази у интервалу 40-42 јер прва три интервала садрже 2+3+30=35 елемената а прва четири интервала 2+3+30+40=75 што је више од половине елемената. $\tilde{h}_X = 40 + 2\left(\frac{100/2-(2+3+30)}{40}\right) = 40.75$.

Примјери за вјежбу:

<u>Примјер 1</u>

Одредити узорачку средину, модус и медијану за узорак: 7, 3, 3, 6, 4, 5, 1, 2, 1, 3.
Решење: $\tilde{x} = 3.5$, $\tilde{d}_X = \tilde{h}_X = 3$.

<u>Примјер 2</u>

Одредити узорачку средину, модус и медијану за узорак: 3.1, 3.0, 1.5, 1.8, 2.5, 3.1, 2.4, 2.8, 1.3.
Решење: $\tilde{x} \approx 2.39$, $\tilde{d}_X = 3.1$, $\tilde{h}_X = 2.5$.

<u>Примјер 3</u>

Одредити узорачку средину, модус и медијану за узорак:

Границе	5-7	7-9	9-11	11-13	13-15	15-17
Учестаност	8	14	40	26	6	4

Решење: $\tilde{x} \approx 10.4$, $\tilde{d}_X = 10.3$, $\tilde{h}_X = 10.35$.

<u>Примјер 4</u>

Одредити узорачку средину, модус и медијану за узорак:

Границе	10-14	14-18	18-22	22-26	26-30	30-34
Учестаност	1	5	10	20	18	3

Решење: $\tilde{x} \approx 24.24$, $\tilde{d}_X \approx 25.33$, $\tilde{h}_X = 24.5$.

Интервалне оцјене параметара расподјеле

Нека је θ параметар расподјеле обележја X са функцијом расподјеле $F(x;\theta)$ за који је узет прост случајни узорак $(X_1,...X_n)$. Дефинишимо двије статистике $\theta_1(X_1,...X_n)$ и $\theta_2(X_1,...X_n)$ гдје је $p[\theta_1 \le \theta_2] = 1$. Ако θ_1 и θ_2 могу бити изабрани тако да за свако $\alpha \in [0,1]$ $p(\theta_1 < \theta < \theta_2) = \alpha$ тада је (θ_1, θ_2) интервална оцјена (интервал повјерења) параметра θ са нивоом повјерења (нивоом поузданости) α.
У примјенама се за вриједности α обично узима $\alpha = 0.9$, $\alpha = 0.95$, $\alpha = 0.99$.
<u>Примјер:</u>
Приликом 7 кварова једне машине измјерени су следећи бројеви часова исправног рада: 53, 48, 50, 54, 51, 50, 51. Претпостављајући да број часова исправног рада има нормалну расподјелу наћи 99% интервал повјерења за непознато математичко очекивање броја часова исправног рада машине.
<u>Решење:</u>
Интервал повјерења за m у $N(m,\sigma^2)$ ако σ^2 није познато налазимо преко $p\left[\left|\frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1}\right| \le t_{n-1,\beta}\right] = \beta$ гдје случајна промјенљива $\frac{\overline{X}_n - m}{\overline{S}_n}\sqrt{n-1}$ има Студентову t расподјелу са $n-1$ степена слободе. Читањем таблице и узимањем вриједности за n, $\overline{X}_n$ и $\overline{S}_n$ из узорка добијамо интервал $\left(\overline{X}_n - \frac{t_{n-1,\beta}\overline{S}_n}{\sqrt{n-1}}; \overline{X}_n + \frac{t_{n-1,\beta}\overline{S}_n}{\sqrt{n-1}}\right)$ за m.

Из узорка добијамо $\overline{X}_7 = \frac{1}{7}\sum_{k=1}^{7} X_k = 51$, $\overline{S}_7^{\,2} = \frac{1}{7}\sum_{k=1}^{7}\left(X_k - \overline{X}_n\right)^2 = 3.4286$, $\overline{S}_7 = 1.85$.
Полазећи од $p[|t_6| \le t_{6;0.99}] = 0.99$ из таблица за Студентову расподјелу налазимо $t_{6;0.99} = 3.707$. Тражени интервал је $\left(\overline{X}_7 - \frac{t_{6;0.99}\overline{S}_7}{\sqrt{7-1}}; \overline{X}_7 + \frac{t_{6;0.99}\overline{S}_7}{\sqrt{7-1}}\right) = (48.2;53.8)$.

<u>Примјер:</u>
Испитује се нека бројна карактеристика X популације са нормалном расподјелом. Узет је узорак обима 30 и добијени су следећи резултати:

X_k	-6	-2.5	-1	1.5	4	6	8.5
m_k	1	3	4	10	6	4	2

Одредити 90% двострани и једнострани интервал повјерења за непознату дисперзију.

<u>Решење:</u>
Једнострани интервал повјерења $[0,b]$ за непознату дисперзију σ^2 у $N(m,\sigma^2)$ ако m није познато добијамо следећим поступком: статистика $\frac{n\overline{S}_n^{\,2}}{\sigma^2}$ има χ^2_{n-1} расподјелу. Из таблица читамо вриједност ε за коју је $P[\chi^2_{n-1} > \varepsilon] = \beta$ а затим узимајући вриједности из узорка имамо $\chi^2_{n-1} = \frac{n\overline{S}_n^{\,2}}{\sigma^2} \ge \varepsilon$ тј. $\sigma^2 \le \frac{n\overline{S}_n^{\,2}}{\varepsilon}$ па је тражени интервал $[0, \frac{n\overline{S}_n^{\,2}}{\varepsilon}]$.
Двострани интервал повјерења $[b_1, b_2]$ се добија помоћу бројева ε_1 и ε_2 за које је $P[\chi^2_{n-1} \ge \varepsilon_1] = \frac{1+\beta}{2}$, $P[\chi^2_{n-1} \ge \varepsilon_2] = \frac{1-\beta}{2}$. Из $P[\varepsilon_1 \le \frac{n\overline{S}_n^{\,2}}{\sigma^2} \le \varepsilon_2] = \beta$ тј. $P[\frac{1}{\varepsilon_1} \ge \frac{\sigma^2}{n\overline{S}_n^{\,2}} \ge \frac{1}{\varepsilon_2}] = \beta$ добијамо двострани интервал $[\frac{n\overline{S}_n^{\,2}}{\varepsilon_2}, \frac{n\overline{S}_n^{\,2}}{\varepsilon_1}]$.

Из узорка добијамо $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{l} m_k X_k$ гдје је l број различитих реализација бројне карактеристике X па је у нашем случају $(l = 7)$:

$\overline{X}_{30} = \frac{1}{30}\sum_{k=1}^{7} m_k X_k = \frac{1}{30}(1\cdot(-6) + 3\cdot(-2.5) + ... + 2\cdot 8.5) = 2.0833$ што можемо заокружити на $\overline{X}_{30} = 2$. $\overline{S}_n^{\,2} = \frac{1}{n}\sum_{k=1}^{l} m_k \left(X_k - \overline{X}_n\right)^2 = ... = 11.19166$, $\overline{S}_{30} = 3.345$.

За једнострани интервал из таблица читамо вриједност ε за коју је $P[\chi^2_{29} > \varepsilon] = 0.9$ а то је $\varepsilon = 19.8$ па је тражени интервал $[0; \frac{n\overline{S}_n^{\,2}}{\varepsilon}] = [0; \frac{30\cdot 11.19166667}{19.8}] = [0;16.957]$.

За двострани интервал повјерења из таблица читамо вриједности ε_1 за које је $P[\chi^2_{29} \geq \varepsilon_1] = \frac{1+0.9}{2}$, и ε_2 за које је $P[\chi^2_{29} \geq \varepsilon_2] = \frac{1-0.9}{2}$ а то су $\varepsilon_1 = 17.7$ и $\varepsilon_2 = 42.6$.

Двострани интервал је:

$$[\frac{n\overline{S}_n^{\,2}}{\varepsilon_2}, \frac{n\overline{S}_n^{\,2}}{\varepsilon_1}] = [\frac{30\cdot 11.19166667}{42.6}, \frac{30\cdot 11.19166667}{17.7}] = [7.88, 18.9689]$$

<u>Примјер:</u>

На једном факултету су испитана примања 60 студената да би се утврдило просјечно примање студената тог факултета. Добијени су следећи резултати:

примања	150-250	250-350	350-450	450-550	550-650	650-750	750-850	850-950	950-1050
Број ст.	3	5	11	15	9	7	5	3	2

Одредити 98% интервал повјерења за непознато математичко очекивање примања студената тог факултета.

<u>Решење:</u>

Из узорка добијамо: $\overline{X}_{60} = \frac{1}{60}(3\cdot 200 + 5\cdot 300 + ... + 2\cdot 1000) = 520$,

$\overline{S}_{60}^{\,2} = \frac{1}{60}(3(300-520)^2 + ... + 2(1000-520)^2) = 34286.66667$, $\overline{S}_{60} = 185.1665917$. Пошто је узорак великог обима за налажење траженог интервала повјерења можемо користити или нормалну апроксимацију Студентове расподјеле или саму Студентову расподјелу. Урадимо то на други начин. Полазећи од $p\left[|t_{59}| \leq t_{59;0.98}\right] = 0.98$ из таблица за Студентову расподјелу налазимо $t_{60;0.98} = 2.39$. Тражени интервал је

$$\left(\overline{X}_{60} - \frac{t_{60;0.99}\overline{S}_{60}}{\sqrt{60-1}}; \overline{X}_{60} + \frac{t_{60;0.99}\overline{S}_{60}}{\sqrt{60-1}}\right) = (462.4; 577.6).$$

<u>Примјер:</u>

Обележје X има $N(m,(0.1)^2)$ расподјелу. Одредити најмањи обим узорка тако да дужина 95%-ног интервала повјерења за непознато математичко очекивање обележја X не буде већа од 0.05.

<u>Решење:</u>

Интервал повјерења за m у $N(m,\sigma^2)$ ако је σ^2 познато налазимо користећи то да $\overline{X}_n$ има расподјелу $N(m,\frac{\sigma^2}{n})$. Из $P[\left|\frac{\overline{X}_n - m}{\sigma}\sqrt{n}\right| \leq z_\beta] = \beta$ помоћу таблица за нормалну расподјелу и вриједности из узорка добијамо интервал повјерења

$(\overline{x}_n - \frac{z_\beta}{\sqrt{n}}\sigma; \overline{x}_n + \frac{z_\beta}{\sqrt{n}}\sigma)$. Дужина интервала је $2\frac{z_\beta}{\sqrt{n}}\sigma$. (Исти поступак користимо и када се не ради о нормалној расподјели, али у је у питању велики узорак јер према ЦГТ $\overline{X}_n$ има приближно расподјелу $N(m, \frac{\sigma^2}{n})$)

Због $P[\left|\frac{\overline{X}_n - m}{0.1}\sqrt{n}\right| \le z_\beta] = 0.95$ у таблици за нормалну расподјелу тражимо вриједност 0.95/2=0.475 и добијамо $z_\beta = 1.96$. Дакле дужина интервала је

$2\frac{z_\beta}{\sqrt{n}}\sigma = \frac{2\cdot 1.96\cdot 0.1}{\sqrt{n}} = \frac{0.392}{\sqrt{n}}$. По услову задатка та дужина не смије да буде већа од 0.05 из чега добијамо неједначину $\frac{0.392}{\sqrt{n}} \le 0.05$. Решење је $n \ge 61.46$. Најмањи обим узорка за који је 95% интервал повјерења за m не већи од 0.05 износи 62.

Тестирање параметарских хипотеза

Нека је θ параметар расподјеле обележја X са функцијом расподјеле $F(x;\theta)$. Свака хипотеза о параметру расподјеле назива се параметарска хипотеза а поступак њене провјере назива се параметарски тест. Простом хипотезом називамо хипотезу да је непознати параметар θ једнак броју θ_0 и означавамо са $H(\theta = \theta_0)$. Ако хипотеза није проста, она је сложена.

Обично радимо са двије хипотезе: H_0 и H_1 . H_0 је хипотеза коју тестирамо и називамо је нулта. H_1 је хипотеза која је у извјесном смислу супротна за H_0 и називамо је алтернативна. Она може бити проста или сложена. Одбацивањем хипотезе H_0 када је она фактички тачна чинимо грешку која се назива грешка првог типа. Прихватањем хипотезе H_0 када је фактички тачна H_1 чинимо грешку која се назива грешка другог типа.

Када тестирамо неку хипотезу не добијамо доказ да ли је она тачна или не. Добијамо само одређени степен сагласности хипотезе са подацима добијеним из узорка. Критеријум бирамо тако да вјероватноћа α одбацивања тачне хипотезе буде мала. Број α називамо праг (ниво) значајности и најчешће користимо вриједности $\alpha = 0.01$, $\alpha = 0.05$, $\alpha = 0.1$.

Хипотезу $H_0(\theta = \theta_0)$ тестирамо на следећи начин:

За узорак $(X_1,...X_n)$ формирамо статистику $\hat{\theta}_n = f(X_1,...X_n)$ и региструјемо њену вриједност θ_n^* добијену из узорка. Затим налазимо вјероватноћу $P_{H_0}[\left|\hat{\theta}_n - \theta_0\right| \ge \left|\theta_n^* - \theta_0\right|] = \alpha^*$ гдје је $P_{H_0}[\cdot]$ вјероватноћа добијена при претпоставци да је $H_0(\theta = \theta_0)$ тачна. Ако добијемо $\alpha^* \le \alpha$ одбацујемо хипотезу $H_0(\theta = \theta_0)$, у супротном прихватамо.

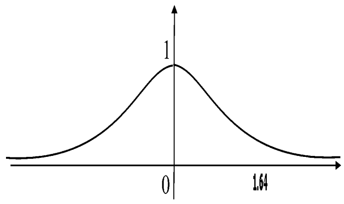
Могуће је радити и на следећи начин: Из таблица налазимо број ε из услова $P_{H_0}[\left|\hat{\theta}_n - \theta_0\right| \ge \varepsilon] = \alpha$ па ако је $\varepsilon \le \left|\theta_n^* - \theta_0\right|$ одбацујемо хипотезу $H_0(\theta = \theta_0)$, у супротном прихватамо.

<u>Примјер:</u>

У 4040 бацања новчића Бифон је добио 2048 пута грб а 1992 пута писмо. На нивоима значајности од 0.05 и 0.01 тестирати хипотезу H_0 да је новчић исправан, тј. да је вјероватноћа појављивања грба 0.5 против алтернативне хипотезе H_1 да новчић форсира појављивање грба тј. да је вјероватноћа појављивања грба већа од 0.5

Решење:

Нека је p вјероватноћа појављивања грба. Нулта хипотеза је $H_0(p=0.5)$. Алтернативна хипотеза је $H_1(p>0.5)$. Нека је случајна промјенљива S_n једнака броју појављивања грба у n бацања новчића. Очигледно S_n, ако је хипотеза



$H_0(p=0.5)$ тачна има расподјелу $S_n : B(4040{,}0.5)$, па и

$S_n^* = \dfrac{S_n - np}{\sqrt{npq}} \sim N(0,1)$.

Због тога и чињенице да је алтернативна хипотеза $H_1(p>0.5)$ услов $P_{H_0}[\left|\hat{\theta}_n - \theta_0\right| \geq \varepsilon] = \alpha$ добија облик

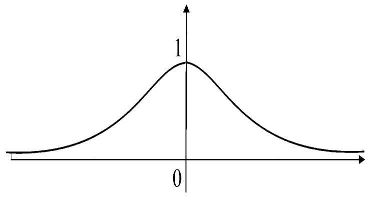
$P(\dfrac{S_n - 4040p}{\sqrt{4040pq}} > \varepsilon) = 0.05$

Са илустрације се види да у таблици за нормалну расподјелу треба да пронађемо вриједност 1-0.05-0.5 =0.95-0.5=0.45 . Добијамо $\varepsilon = 1.64$. Како је на основу узорка $S_n = 2048$ то је $\dfrac{S_n - 4040p}{\sqrt{4040pq}} = \dfrac{2048 - 4040 \cdot 0.5}{\sqrt{4040(0.5)^2}} = \dfrac{28}{32} = 0.87$. Добијена вриједност је мања од $\varepsilon = 1.64$ па нема разлога за одбацивање хипотезе $H_0(p=0.5)$ са нивоом значајности 0.05.

Коментар 1: Када у услову $P(\dfrac{S_n - 4040p}{\sqrt{4040pq}} > \varepsilon) = 0.05$ $\dfrac{S_n - 4040p}{\sqrt{4040pq}}$ и ε добију вриједности 0.87 и 1.64 не можемо више говорити о вјероватноћи јер се ту ради о константама а не више о промјенљивим величинама. Оно што можемо да закључимо је : Постављен је услов да вјероватноћа догађаја $\dfrac{S_n - 4040p}{\sqrt{4040pq}} > \varepsilon$ буде веома мала.

Резултати добијени из узорка се добро слажу са постављеним условом јер се догађај није реализовао. Да смо на основу узорка добили за $\dfrac{S_n - 4040p}{\sqrt{4040pq}}$ вриједност већу од ε то би значило да се реализовао догађај који, ако је хипотеза H_0 тачна, има веома малу вјероватноћу. То би био разлог да одбацимо хипотезу $H_0(p=0.5)$ и прихватимо $H_1(p>0.5)$.

Коментар 2: Када у услову $P(\dfrac{S_n - 4040p}{\sqrt{4040pq}} > \varepsilon) = 0.05$ замијенимо ниво значајности са 0.01 онда ће очигледно вриједност коју тражимо у таблици бити 0.5-0.01=0.49 што даје још већу вриједност за ε па тим прије нема разлога за одбацивање нулте хипотезе са нивоом значајности 0.01 .



Коментар 3: Услучају алтернативне хипотезе $H_1(p \neq 0.5)$ услов $P_{H_0}[\left|\hat{\theta}_n - \theta_0\right| \geq \varepsilon] = \alpha$ добија облик

$P(\left|\dfrac{S_n - 4040p}{\sqrt{4040pq}}\right| > \varepsilon) = \alpha$ тј. $2 \cdot P(\dfrac{S_n - 4040p}{\sqrt{4040pq}} > \varepsilon) = \alpha$

$P(\frac{S_n - 4040p}{\sqrt{4040pq}} \le \varepsilon) = 1 - \frac{\alpha}{2}$. За ниво значајности $\alpha = 0.05$ одређујемо ε из услова $P(\frac{S_n - 4040p}{\sqrt{4040pq}} \le \varepsilon) = 0.975$. У таблици тражимо вриједност 0.975-0.5=0.475. Добијамо $\varepsilon = 1.96$ што доводи до прихватања H_0. Резултат је сасвим очекиван јер за наведени узорак (2048>4040/2) хипотеза $H_1(p > 0.5)$ има веће шансе за обарање H_0 него $H_1(p \ne 0.5)$.

Коментар 4: Размотрити поступак у случају алтернативне хипотезе $H_1(p < 0.5)$.

<u>Примјер:</u>

Случајно изабрани ексери из једне кутије имају следеће дужине (у мм): 80, 81, 81, 82, 81, 82, 80, 82, 81, 81. Тестирати хипотезу $H_0(m = 80)$ против алтернативне $H_1(m \ne 80)$ на нивоу значајности 0.01 гдје је 80 номинална вриједност дужине назначена на кутији.

Тестирати $H_0(m = 80)$ против $H_1(m > 80)$.

<u>Решење:</u>

Тестирање хипотезе $H_0(m = m_0)$ у нормалној расподјели $N(m, \sigma^2)$ кад σ^2 није познато вршимо на следећи начин: Ако је H_0 тачна онда $\frac{\overline{X}_n - m_0}{\overline{S}_n}\sqrt{n-1}$ има Студентову расподјелу са $n-1$ степени слободе тј. $\frac{\overline{X}_n - m_0}{\overline{S}_n}\sqrt{n-1} = t_{n-1}$. Из таблица одређујемо ε из услова $P[\left|\overline{X}_n - m_0\right| \ge \varepsilon] = P[\frac{\left|\overline{X}_n - m_0\right|}{\overline{S}_n}\sqrt{n-1} \ge \frac{\varepsilon\sqrt{n-1}}{\overline{S}_n}] = \alpha$. Нулту хипотезу одбацујемо ако је ε мање или једнако од вриједности $\frac{\left|\overline{x}_n - m_0\right|}{\overline{S}_n}\sqrt{n-1}$ добијене из узорка.

У задатку имамо: $n = 10$, $\overline{X}_n = 81.1$, $\overline{S}_n = 0.7$ па у случају алтернативне хипотезе $H_1(m \ne 80)$ са нивоом значајности 0.01 услов који користимо има облик $P[\frac{\left|\overline{X}_n - m_0\right|}{\overline{S}_n}\sqrt{n-1} \ge \varepsilon] = 0.01$. Из узорка имамо:

$\frac{\left|\overline{x}_n - m_0\right|}{\overline{S}_n}\sqrt{n-1} = \frac{|81.1 - 80|}{0.7}\sqrt{10-1} = \frac{1.1}{0.7}3 = 4.7$. Из таблице одређујемо вриједност ε за коју је $P[|t_9| \ge \varepsilon] = 0.01$ а то је $\varepsilon = 3.25$. Како је вриједност $\frac{\left|\overline{x}_n - m_0\right|}{\overline{S}_n}\sqrt{n-1}$ из узорка једнака 4.7 тј. већа од ε то закључујемо да вриједности из узорка нису у сагласности са нултом хипотезом па је одбацујемо.

У случају алтернативне хипотезе $H_1(m > 80)$ са нивоом значајности 0.01 услов који користимо има облик $P[\frac{\overline{X}_n - m_0}{\overline{S}_n}\sqrt{n-1} \ge \varepsilon] = 0.01$. Из узорка се добија $\frac{\overline{x}_n - m_0}{\overline{S}_n}\sqrt{n-1} = 4.7$. Из таблице одређујемо вриједност ε за коју је $P[t_9 \ge \varepsilon] = 0.01$ а то је $\varepsilon = 2.82$. Како је вриједност $\frac{\overline{x}_n - m_0}{\overline{S}_n}\sqrt{n-1}$ из узорка једнака 4.7 тј. већа од ε то закључујемо да вриједности из узорка нису у сагласности са нултом хипотезом па је одбацујемо.

Примјер:

Просјечан вијек трајања сијалица произведених у једној фабрици по старом поступку био је 1120 сати и стандардно одступање било је 125 сати. Узорак од 8 сијалица произведених по новом поступку дао је средњи вијек трајања 1070 сати. Тестирати хипотезу да се вијек трајања сијалица није промијенио. За ниво значајности узети $\alpha = 0.05$.

Решење:

Када је код нормалног закона расподјеле $N(m,\sigma^2)$ познато σ^2 тестирање хипотезе $H_0(m = m_0)$ против алтернативне $H_1(m \neq m_0)$ вршимо на следећи начин: Ако је H_0 тачна тада $\overline{X}_n$ има расподјелу $N(m,\frac{\sigma^2}{n})$. Зато из таблица за нормалну расподјелу налазимо $P[\left|\overline{X}_n - m_0\right| \geq \left|\overline{x}_n - m_0\right|] = P[\frac{\left|\overline{X}_n - m_0\right|}{\sigma}\sqrt{n} \geq \frac{\left|\overline{x}_n - m_0\right|}{\sigma}\sqrt{n}] = \alpha^*$ и ако је $\alpha^* \leq \alpha$ нулту хипотезу одбацујемо. Може и: из услова $P[\left|\overline{X}_n - m_0\right| \geq \varepsilon] = P[\frac{\left|\overline{X}_n - m_0\right|}{\sigma}\sqrt{n} \geq \frac{\varepsilon\sqrt{n}}{\sigma}] = \alpha$ одређујемо $\frac{\varepsilon\sqrt{n}}{\sigma}$ па ако је то мање или једнако од $\frac{\left|\overline{X}_n - m_0\right|}{\sigma}\sqrt{n}$ одбацујемо нулту хипотезу.

У задатку је $H_0(m = 1120)$ и $H_1(m \neq 1120)$, $\alpha = 0.05$. Из $P[\left|X\right| \geq \varepsilon] = 0.05$ тј. $2P[X \geq \varepsilon] = 0.05$ добијамо $P[X < \varepsilon] = 0.975$. Из таблице очитавамо вриједност 0.475 за $\varepsilon = 1.96$. Из узорка добијамо $\frac{\left|\overline{x}_n - m_0\right|}{\sigma}\sqrt{n} = \frac{|1070 - 1120|}{125}\sqrt{8} = 1.13 < 1.96$ па нема разлога за одбацивање нулте хипотезе.