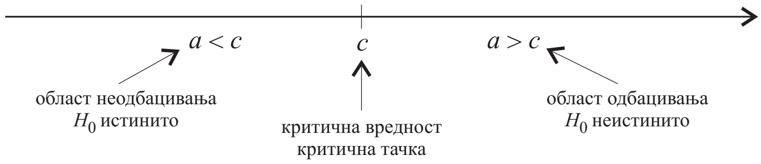
Глава 7
Тестирање хипотеза

1. Основни појмови

Нулта хипотеза је тврђење о неком параметру основног скупа, која се сматра истинитом све док се не докаже супротно. Обично је означавамо са H_0.

Алтернативна хипотеза је тврђење о неком параметру основног скупа, које ће бити истинито ако је нулта хипотеза неистинита. Обично је означавамо са H_1.



Постоје грешке два типа које се могу појавити при тестирању хипотеза.

Грешка I *врсте* се јавља када се истинита нулта хипотеза одбацује. Вредност α представља вероватноћу јављања грешке ове врсте, односно,

$$\alpha = P(H_0 \text{ се одбацује} \,|\, H_0 \text{ је истинита})$$

Вредност α се зове *ниво значајности теста*.

Грешка II *врсте* се јавља када се неистинита нулта хипотеза не одбаци. Вредност β представља вероватноћу јављања грешке ове врсте, односно,

$$\beta = P(H_0 \text{ се не одбацује} \,|\, H_0 \text{ је неистинита})$$

Вредност $1 - \beta$ се зове *јачина теста*.

Деоба укупне области на области одбацивања и неодбацивања зависи од изабране вредности нивоа значајности, α.

Област одбацивања може се у односу на област неодбацивања налазити са обе стране, лево од ње или десно од ње. Одговарајући тестови се редом зову *двострани*, *левострани* и *деснострани*. Левострани и деснострани тестови чине групу тзв. *једностраних* (или *једносмерних*) тестова, а двострани се зове још и *двосмерним* тестом.

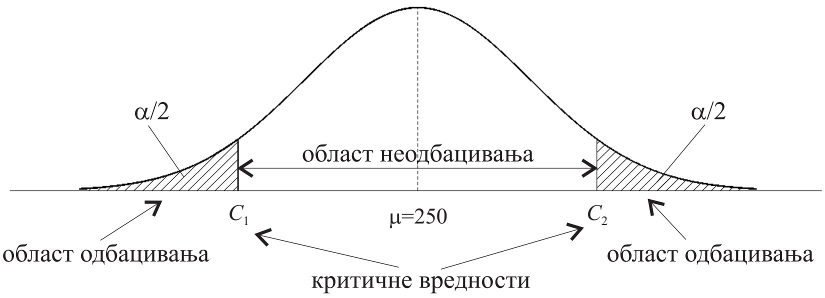
Илустујмо уведене облике тестова следећим примерима.

Пример 1. На основу анкете утврђено је да деца од 8–10 година проводе дневно 250 минута испред телевизијског екрана. Дечији психолог жели да провери да ли је дошло до промене у просечној дужини гледања телевизијског програма у датом тренутку.

У датом случају можемо узети да је:

H_0: просечна дужина гледања се није изменила, $\mu = 250$ минута
H_1: просечна дужина гледања се променила, $\mu \neq 250$ минута

Под претпоставком да је расподела статистике узорка $\overline{X}$ нормална овај случај можемо илустровати следећом сликом.



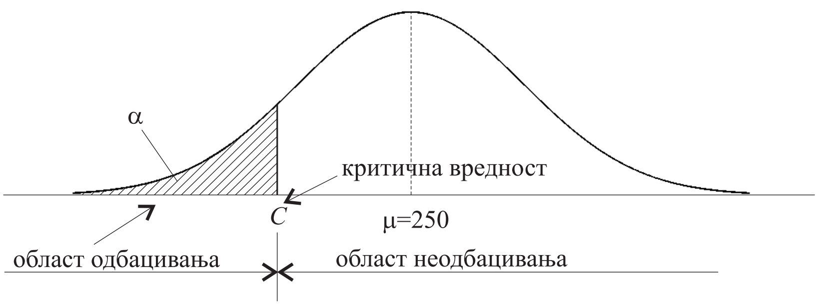
Дакле, овде се ради о двостраном тесту.

Пример 2. Фирма која производи сокове тврди да њени паковања сокова садрже тачно 1000 ml сока. Желимо да проверимо ову тврдњу.

У овом случају природно је узети (за купца је битно да није преварен — може за исте паре да добије и већу количину сока, али она не сме да буде мања од декларисане) да је:

H_0: $\mu = 1000\,\text{ml}$
H_1: $\mu < 1000\,\text{ml}$

Под претпоставком да је расподела статистике узорка $\overline{X}$ нормална овај случај можемо илустровати следећом сликом.



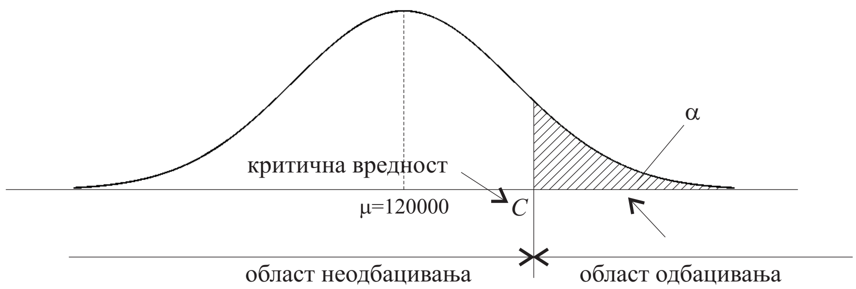
Дакле, овде се ради о левостраном тесту.

Пример 3. Просечна цена куће у неком граду је $\mu = 120.000$ евра. Желимо да проверимо да ли је она већа од наведене.

У датом случају природно је узети да је:

H_0: $\mu = 120.000$
H_1: $\mu > 120.000$

Под претпоставком да је расподела статистике узорка $\overline{X}$ нормална овај случај можемо илустровати следећом сликом.



Дакле, овде се ради о десностраном тесту.

Из управо наведених примера намеће се закључак да облик теста зависи од знака у алтернативној хипотези. Следећа таблица нам помаже да се оријентишемо у избору врсте теста.

	двострани тест	левострани тест	деснострани тест
знак у H_0	$=$	$=$ или $\geq$	$=$ или $\leq$
знак у H_1	$\neq$	$<$	$>$
област одбацивања	на оба краја	на левом крају	на десном крају

Користићемо два поступка тестирања хипотеза

1. Приступ заснован на p-вредности. Прво израчунавамо тзв. p-вредност за реализовану вредност статистике узорка. Ако је ниво значајности α унапред изабран, поредимо ову p-вредност са α и доносимо одговарајућу одлуку. p-вредност представља вероватноћу ...

2. Приступ заснован на критичној вредности.

2. Тестирање хипотеза о μ: σ је познато

У случају када је стандардна девијација σ позната, могућа су следећа три случаја.

Случај I. Ако је величина узорка мала ($n < 30$) и основни скуп има нормалну расподелу, онда, како смо видели у претходном поглављу, узорачка расподела од $\overline{X}$ је нормална са аритметичком средином μ и стандардном девијацијом $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$, при услову да је $\dfrac{n}{N} \leq 0{,}05$, па тестирање хипотезе о μ спроводимо коришћењем нормалне расподеле.

Случај II. Ако је величина узорка велика ($n \geq 30$), онда, што је последица централне граничне теореме, за узорачку расподела од $\overline{X}$ користимо нормалну расподелу са аритметичком средином μ и стандардном девијацијом $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$, при услову да је $\dfrac{n}{N} \leq 0{,}05$, па тестирање хипотезе о μ спроводимо коришћењем нормалне расподеле.

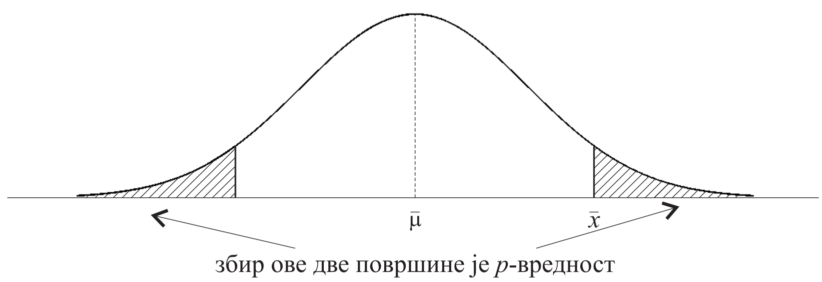
Случај III. Ако је величина узорка мала ($n < 30$) и основни скуп нема нормалну расподелу или је његова расподела непозната, онда се користи тзв. непараметарски метод за одређивање интервала поверења за μ. Ми нећемо разматрати овај случај.

Као што смо рекли при тестирању хипотеза примењујемо два приступа. Објаснимо, сада, сваки од ових приступа подробније.

А) **Приступ заснован на p-вредности.** Када примењујемо овај приступ прво израчунавамо z вредност за $\overline{x}$ (ако користимо нормалну расподелу):

$$z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}}, \quad \text{где је} \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}.$$

Израчунату вредност z зовемо *реализованом вредношћу статистике z теста*. За израчунату вредност z из таблице тражимо одговарајућу p-вредност. Ако је p-вредност $< \alpha$ — одбацујемо нулту хипотезу, а ако је p-вредност $\geq \alpha$ — не одбацујемо нулту хипотезу. Смисао p-вредности (у случају двостраног теста) можемо појаснити следећом сликом:



Етапе у тестирању хипотезе применом приступа заснованог на p-вредности су:

1. Формулисање нулте и алтернативне хипотезе.
2. Избор расподеле која ће се користити.
3. Израчунавање p-вредности.
4. Доношење одлуке.

Пример 4. Управа једног клуба за мршављење тврди да чланови клуба губе у просеку најмање 10 фунти током првог месеца од учлањења. У намери да провери ову тврдњу инспекција је изабрала 36 чланова клуба и утврдила да су они изгубили у просеку 9,2 фунте у току првог месеца од учлањења. Познато је да је стандардна девијација основног скупа 2,4 фунте. Одредите p-вредност у овом тесту. Шта ћете закључити, ако је $\alpha = 0{,}01$, а шта ако је $\alpha = 0{,}05$?

Решење. У нашем примеру је

$$n = 36, \quad \overline{x} = 9{,}2 \text{ фунте}, \quad \sigma = 2{,}4 \text{ фунте}, \quad \mu = 10$$

Задатак решимо поетапно по поступку изложеном горе:

Етапа 1. Формулисање хипотеза H_0 и H_1:

$$H_0 \colon \mu \geq 10$$
$$H_1 \colon \mu < 10$$

Етапа 2. Користимо нормалну расподелу, јер је у питању други случај.

Етапа 3. Израчунајмо сада p-вредност. Како је у H_1 знак $<$, то је тест левострани. Имамо да је

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2{,}4}{\sqrt{36}} = 0{,}4,$$

па је

$$z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{9{,}2 - 10}{0{,}4} = -2{,}00.$$

За израчунато $z = -2{,}00$ налазимо из таблице да је

$$p\text{-вредност} = 0{,}0228.$$

Етапа 4. Доношење одлуке:

$$\begin{array}{ll} \alpha = 0{,}01 < p = 0{,}0228 & \text{не одбацујемо} \\ \alpha = 0{,}05 > p = 0{,}0228 & \text{одбацујемо} \end{array}$$

Б) **Приступ заснован на критичној вредности.** Приликом тестирања хипотезе о μ применом нормалне расподеле случајна променљива

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}, \quad \text{где је} \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}},$$

се зове *статистика теста*. Статистика теста може да се дефинише као правило или критеријум који користимо при доношењу одлуке о одбацивању или неодбацивању нулте хипотезе H_0. Вредност променљиве Z која се добија за дату вредност од $\overline{X}$ зове се *реализована вредност z статистике теста*.

Етапе у тестирању хипотезе применом приступа заснованог на критичној вредности су:

1. Формулисање нулте и алтернативне хипотезе.
2. Избор расподеле која ће се користити.
3. Одређивање области одбацивања и неодбацивања.
4. Израчунавање вредности статистике теста
5. Доношење одлуке.

Пример 5. Телефонска компанија је утврдила да је просечна дужина међународних разговора прошле године била 12,44 минуте. Управа компаније хоће да утврди да ли се просечна дужина ових разговора променила. Узорак од 150 таквих разговора даје просечну дужину од 13,71 минут. Стандардна девијација оваквих разговора је 2,65 минута. Да ли на нивоу значајности 2 % можете да закључите да се просечна дужина разликује, сада, од 12,44?

Решење. У нашем примеру је

$$n = 150, \quad \overline{x} = 13{,}71 \text{ минут}, \quad \sigma = 2{,}65 \text{ минута}, \quad \mu = 12{,}44 \text{ минута}.$$

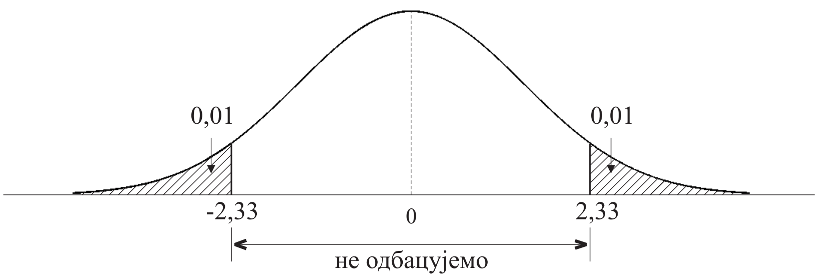
Задатак решимо поетапно по поступку изложеном горе:

Етапа 1. Формулисање хипотеза H_0 и H_1:

$$H_0: \ \mu = 12{,}44, \quad H_1: \ \mu \neq 12{,}44\,.$$

Етапа 2. Користимо нормалну расподелу, јер је у питању други случај.

Етапа 3. Одређивање области одбацивања и неодбацивања. Због знака $\neq$ следи да је у питању двострани тест. Како је $\alpha/2 = 0{,}02/2 = 0{,}01$, то су критичне тачке 0,0100 и 0,9900, а одговарајуће вредности су $-2{,}33$ и 2,33.



Етапа 4. Израчунавајмо вредности статистике теста z за $\overline{x} = 13{,}71$:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2{,}65}{\sqrt{150}} = 0{,}2163759, \quad z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{13{,}71 - 12{,}44}{0{,}2163759} = 5{,}87.$$

Етапа 5. Доношење одлуке. Како је $5{,}87 > 2{,}33$, то нулту хипотезу одбацујемо.

Пример 6. Група која заступа права потрошача верује да паковање сира чедар у локалном супермаркету уместо декларисане тежине од 10 унци заправо тежи мање од 10 унци. У циљу провере изабран је случајан узорак од 20 паковања, чија просечна тежина износи 9,995 унци. Основни скуп има нормалну расподелу са стандардном девијацујом од 0,15 унци.

А) Ако алтернативна хипотеза гласи да је просечна тежина мања од 10 унци, одредите, p-вредност у овом тесту. Да ли ћете одбацити нулту хипотезу ако је $\alpha = 0,01$?

Б) Тестирајте хипотезу у задатку под А), користећи приступ заснован на критичној вредности са $\alpha = 0,01$.

Решење. У нашем примеру је

$$n = 20, \quad \overline{x} = 9{,}995 \text{ унци}, \quad \sigma = 0{,}15 \text{ унци}, \quad \mu = 10 \text{ унци}.$$

Задатак решимо поетапно по поступку изложеном горе:

Етапа 1. Формулисање хипотеза H_0 и H_1:

$$H_0: \ \mu \geq 10, \quad H_1: \ \mu < 10\,.$$

Етапа 2. Користимо нормалну расподелу, јер је у питању први случај.

Етапа 3. Одредимо p-вредност. Због знака $<$ у алтернативној хипотези у питању је левострани тест. Имамо да је

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{0{,}15}{\sqrt{20}} = 0{,}0335, \quad z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{9{,}995 - 10}{0{,}0335} = -0{,}1492.$$

Сада из таблице налазимо да је p-вредност једнака 0,4404.

Етапа 5. Доношење одлуке. Како је $0{,}01 < 0{,}4404$, то нулту хипотезу не одбацујемо.

3. Тестирање хипотеза о μ: σ није познато

У случају када стандардна девијација σ није позната, могућа су следећа три случаја:

Случај I. Ако је величина узорка мала ($n < 30$) и основни скуп има нормалну расподелу, онда користимо t-расподелу за тестирање хипотеза о μ.

Случај II. Ако је величина узорка велика ($n \geq 30$), онда, као и у претходном случају, користимо t-расподелу за тестирање хипотеза о μ.

Случај III. Ако је величина узорка мала ($n < 30$) и основни скуп нема нормалну расподелу или је његова расподела непозната, онда се користи тзв. непараметарски метод за тестирање хипотеза о μ.

У прва два од три наведена случаја промењлива

$$t = \frac{\overline{x} - \mu}{S_{\overline{X}}}, \quad \text{где је} \quad S_{\overline{X}} = \frac{S}{\sqrt{n}},$$

има t-расподелу, промењива t се зове *статистика теста* и она се користи при тестирању хипотезе о аритметичкој средини основног скупа. Овако добијена вредност t се зове *реализована вредност статистике t теста*.

Као и у случају када је σ познато, при тестирању хипотеза примењујемо два приступа. Објаснимо, сада, сваки од ових приступа подробније.

А) **Приступ заснован на p-вредности.** Етапе у тестирању хипотезе применом приступа заснованог на p-вредности су:

1. Формулисање нулте и алтернативне хипотезе.
2. Избор расподеле која ће се користити.
3. Израчунавање p-вредности.
4. Доношење одлуке.

Пример 7. Психолози тврде да је просек старости деце у тренутку када проходају 12,5 месеци. Хоћемо да проверимо ово тврђење. На узорку од 18 деце утврдили смо да је за њих аритметичка средина старости у тренутку када су проходала 12,9 месеци са дисперзијом 0,8 месеци. Познато је да старост деце у тренутку када проходају има приближно нормалну расподелу. Одредити p-вредност за тестирање хипотезе да се аритметичка средина старости деце у тренутку када проходају разликује од 12,5 месеци. Шта закључујемо за ниво значајности од 1 %?

Решење. У нашем примеру је

$$n = 18, \quad \overline{x} = 12{,}9 \text{ месеци}, \quad s = 0{,}8 \text{ месеци}, \quad \mu = 12{,}5 \text{ месеци}.$$

Задатак решимо поетапно по поступку изложеном горе:

Етапа 1. Формулисање хипотеза H_0 и H_1:

$$H_0 : \mu = 12{,}9 \text{ месеци}$$
$$H_1 : \mu \neq 12{,}9 \text{ месеци}$$

Етапа 2. Избор расподеле која се користи. Како је први случај, то користимо t-расподелу.

Етапа 3. Израчунајмо p-вредност. Тест је двострани, јер је знак у алтернативној хипотези H_1 знак неједнакости $\neq$. Прво нађимо реализовану вредност статистике t теста за $\overline{x} = 12{,}9$ и број степени слободе одговарајуће t расподеле:

$$\begin{aligned} s_{\overline{X}} &= \frac{s}{\sqrt{n}} = \frac{0{,}8}{\sqrt{18}} = 0{,}18856181 \\ t &= \frac{\overline{x}-\mu}{s_{\overline{X}}} = \frac{12{,}9-12{,}5}{0{,}18856181} = 2{,}121 \\ df &= n-1 = 18-1 = 17 \end{aligned}$$

Из таблице за t расподелу (17. ред) нађимо две вредности које, одоздо и одозго, апроксимирају израчунату вредност 2,121; то су вредности 2,110 и 2,567. Дакле,

$$2{,}110 < 2{,}121 < 2{,}567.$$

Овим вредностима одговарају вредности 0,025 и 0,01 променљиве t. Како је тест двостран, то добијене вредности множимо са 2, па, напокон, добијамо да је

$$0{,}02 < p\text{-вредност} < 0{,}05.$$

Да смо уместо таблице користили неки софтвер за израчунавање одговарајућих вредности за расподелу t, то би смо могли тачније да одредимо p-вредност, јер вредности 2,121 одговара приближно вредност 0,049, тј. имамо да је p-вредност $= 0{,}049$.

Етапа 4. Доношење одлуке. За било које $\alpha > 0{,}05$ одбацујемо H_0. За сваку вредност $\alpha < 0{,}02$ не одбацујемо H_0. Ако је

$$0{,}02 < \alpha < 0{,}05.$$

не можемо да донесемо никакву одлуку. Како је

$$\alpha = 0{,}01 < 0{,}02$$

нулту хипотезу не одбацујемо, тј. можемо да кажемо да је разлика између аритметичке средине основног скупа и аритметичке средине узорка мала и да се јавља због случајне грешке.

Б) **Приступ заснован на критичној вредности.** Етапе у тестирању хипотезе применом приступа заснованог на критичној вредности су:

1. Формулисање нулте и алтернативне хипотезе.
2. Избор расподеле која ће се користити.
3. Одређивање области одбацивања и неодбацивања.
4. Израчунавање вредности статистике теста
5. Доношење одлуке.

Пример 8. Недавно је на основу једног истраживања утврђено да одрасле особе у Америци у просеку проводе 18 сати недељно у слободним активностима. Један истраживач је желео да провери ово тврђење. Изабрао је узорак од 10 одраслих особа и питао их колико времена недељно проводе у слободним активностима Следе њихови одговори (у сатима):

$$14 \quad 25 \quad 22 \quad 38 \quad 16 \quad 26 \quad 19 \quad 23 \quad 41 \quad 33$$

Претпоставимо да време које сви одрасли проводе у слободним активностима има нормалну расподелу. Можете ли да закључите да је истинито тврђење истраживача (користите 5% ниво значајности)?

Решење. У нашем примеру је

$$n = 150, \quad \overline{x} = 13{,}71 \text{ минут}, \quad \sigma = 2{,}65 \text{ минута}, \quad \mu = 12{,}44 \text{ минута}.$$

Задатак решимо поетапно по поступку изложеном горе:

Етапа 1. Формулисање хипотеза H_0 и H_1:

$$H_0\colon \mu = 12{,}44$$
$$H_1\colon \mu \neq 12{,}44$$

Етапа 2. Користимо нормалну расподелу, јер је у питању други случај.

Етапа 3. Одређивање области одбацивања и неодбацивања. Због знака $\neq$ следи да је у питању двострани тест. Како је

$$\frac{\alpha}{2} = \frac{0{,}02}{2} = 0{,}01,$$

то су критичне тачке 0,0100 и 0,9900, а одговарајуће вредности су $-2{,}33$ и $2{,}33$.

Етапа 4. Израчунавајмо вредности статистике теста z за $\overline{x} = 13{,}71$:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2{,}65}{\sqrt{150}} = 0{,}2163759, \quad z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{13{,}71 - 12{,}44}{0{,}2163759} = 5{,}87.$$

Етапа 5. Доношење одлуке. Како је $5{,}87 > 2{,}33$, то нулту хипотезу одбацујемо.