# TESTIRANJE HIPOTEZA

Testiranje podrazumeva postupak provere određene pretpostavke koju zovemo nulta hipoteza.

Postupak testiranja obuhvata tri faze:

1. Formulisanje polazne pretpostavke – nulte hipoteze

2. Postupak provere postavljene hipoteze

3. Zaključak o postavljenoj hipotezi

***Testovi aritmetičkih sredina***

Postoje sledeći vidovi testiranja aritmetičkih sredina:

1. Upoređivanje aritmetičke sredine uzorka sa aritmetičkom sredinom osnovnog skupa ili sa nekom hipotetičkom vrednošću – test značajnosti jedne sredine

2. Upoređivanje dve aritmetičke sredine iz dva nezavisna uzorka – test značajnosti razlike dve sredine

3. Upoređivanje više od dve sredine iz više od dva uzorka – metod analize varijanse

*Test značajnosti jedne sredine*

Test značajnosti jedne sredine podrazumeva upoređivanje aritmetičke sredine iz uzorka sa aritmetičkom sredinom osnovnok skupa (ili nekom pretpostavljenom vrednošću). Kao i kod ocene i kod testiranja razlikujemo dva slučaja, kada jepoznat varijabilitet osnovnog skupa i kada varijabilitet osnovnog skupa nije poznat.

1. Slučaj: poznata $\boldsymbol{\sigma}^2$

2.Slučaj: nije poznata $\boldsymbol{\sigma}^2$

1. **Slučaj: poznata** $\sigma^2$

Prvi korak (definisanje nulte hipoteze):

$$H_0 : \bar{x} = \mu$$

$$H_1 : \bar{x} \neq \mu$$

Drugi korak (definisanje odgovarajuće test statistike):

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

$$\boldsymbol{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N - n}{N - 1}}}$$

Treći korak (donošenje zaključka):

U slučaju da je poznat varijabilitet osnovnog skupa apsolutna vrednost Z-količnika se poredi isključivo sa kritičnim vrednostima iz tablice Normalne raspodele. Ako je apsolutna vrednost Z-količnika manja od kritične vrednosti prihvata se nulta hipoteza. S druge strane, ako je apsolutna vrednost Z-količnika manja od kritične vrednosti nulta hipoteza se odbacuje i prihvata se alternativna.

**Istu hipotezu možemo proveriti i izračunavanjem intervala poverenja:**

$$\boldsymbol{\bar{x} - z_\alpha \sigma_{\bar{X}} < \mu < \bar{x} + z_\alpha \sigma_{\bar{X}}}$$

$$\boldsymbol{\mu \in (L_1, L_2) \Rightarrow H_o}$$
$$\boldsymbol{\mu \notin (L_1, L_2) \Rightarrow H_1}$$

Ako aritmetička sredina osnovnog skupa (ili neka pretpostavljena vrednost) ulazi u prethodno izračunati interval poverenja, prihvata se nulta hipoteza, u obrnutom slučaju nulta hipoteza se odbacuje.

**Primer 1.** Prosečna visina grebena kod jedne rase goveda je 165cm a varijabilitet iskazan standardnom devijacijom 2cm. Na osnovu ispitivanja 80 grla iste rase konstatovana je prosečna visina od 165,5cm. Može li se smatrati da je odabrani uzorak reprezentativan, odnosno da potvrđuje karakteristike posmatrane rase**.**

$$\mu = 165cm$$
$$\sigma = 2cm$$
$$n = 80$$
$$\bar{X} = 165{,}5cm$$

Prvi korak:

$$\boxed{H_0 : \overline{X} = \mu}$$

Drugi korak:

$$\boxed{Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}}$$

$$\boxed{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{80}} = 0{,}224 cm}$$

$$\boxed{Z = \frac{165{,}5 - 165}{0{,}224} = 2{,}232^{*}}$$

Treći korak:

$$\boxed{\begin{array}{l} |Z|^{*} \rangle 1{,}96 \Rightarrow H_1 \\ |Z| \langle 2{,}58 \Rightarrow H_0 \end{array}}$$

Istu hipotezu možemo proveriti i putem intervala poverenja:

$$\boxed{\begin{array}{l} \alpha = 0{,}05 \\ 165{,}5 - 1{,}96 \cdot 0{,}224 < \mu < 165{,}5 + 1{,}96 \cdot 0{,}224 \\ 165{,}061 \langle \mu \langle 165{,}939 cm \\ \mu \notin (L_1, L_2) \Rightarrow H_1 \end{array}} \quad \boxed{\begin{array}{l} \alpha = 0{,}01 \\ 165{,}5 - 2{,}58 \cdot 0{,}224 < \mu < 165{,}5 + 2{,}58 \cdot 0{,}224 \\ 164{,}922 \langle \mu \langle 166{,}078 cm \\ \mu \in (L_1, L_2) \Rightarrow H_0 \end{array}}$$

2. **Slučaj: nije poznata** $\boxed{\sigma^2}$

**U slučaju kada nisu poznate vrednosti standardne devijacije ili varijanse osnovnog skupa, hipoteza se proverava izračunavanjem količnika t .**

$$\boxed{t = \frac{\overline{X} - \mu}{S_{\overline{X}}}}$$

Apsolutna vrednost izračunatog količnika *t* upoređuje se sa kritičnim vrednostima Normalne raspodele ili t-distribucije, što zavisi od veličine uzorka na osnovu kog se izvodi provera hipoteze.

Ako se provera hipoteze izvodi na osnovu velikog uzorka (n› 30) izračunati količnik t upoređuje se sa kritičnim vrednostima iz tablica ND.

$$
\begin{aligned}
&|t| \langle Z_\alpha \Rightarrow H_0 \\
&|t|^{**} \rangle Z_\alpha \Rightarrow H_1 \\
&|t|^{*} \rangle Z_{0,05} \Rightarrow H_1
\end{aligned}
$$

Ako se provera hipoteze izvodi na osnovu malog uzorka (n‹ 30) izračunati količnik t upoređuje se sa kritičnim vrednostima t – distribucije.

$$
\begin{aligned}
&|t| \langle t_{\alpha(n-1)} \Rightarrow H_0 \\
&|t|^{**} \rangle t_{\alpha(n-1)} \Rightarrow H_1 \\
&|t|^{*} \rangle t_{0,05(n-1)} \Rightarrow H_1
\end{aligned}
$$

Testiranje hipoteze o značajnosti jedne sredine u slučaju nepoznatih parametara osnovnog skupa, može da se izvede i izračunavanjem odgovarajućeg intervala poverenja.

$$
\begin{aligned}
&\alpha = 0,05 \\
&\alpha = 0,01 \\
&n \langle 30 \\
&\overline{X} - t_{\alpha(n-1)} \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + t_{\alpha(n-1)} \cdot S_{\overline{X}} \\
&n \rangle 30 \\
&\overline{X} - Z_\alpha \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + Z_\alpha \cdot S_{\overline{X}}
\end{aligned}
$$

$$
\begin{aligned}
&\mu \in (L_1, L_2) \Rightarrow H_0 \\
&\mu \notin (L_1, L_2) \Rightarrow H_1
\end{aligned}
$$

**Primer 1.** Na sedam hektara ostvaren je sledeći prinos zelenog zrna graška: 6,8 6,9 7,1 7,2 7,2 7,4 7,5 tona. Da li je postignuti prosečan prinos na nivou očekivanog prosečnog prinosa od 7,6 tona po hektaru?

$$\boldsymbol{n} = 7$$
$$\boldsymbol{\mu} = 7{,}6\boldsymbol{t}$$

| **Prinos graška (t) Xi** | **Xi2** |
|---|---|
| **6,8** | **46,24** |
| **6,9** | **47,61** |
| **7,1** | **50,41** |
| **7,2** | **51,84** |
| **7,2** | **51,84** |
| **7,4** | **54,76** |
| **7,5** | **56,25** |
| **50,1** | **358,95** |

**Prvi korak:**

$$H_0 : \overline{\mathrm{X}} = \mu$$

**Drugi korak:**

$$\boldsymbol{t} = \frac{\overline{\mathrm{X}} - \boldsymbol{\mu}}{\boldsymbol{S}_{\overline{\mathrm{X}}}}$$

$$\overline{\mathrm{X}} = \frac{\Sigma \mathrm{X}_{\boldsymbol{i}}}{\boldsymbol{n}} = \frac{50{,}1}{7} = 7{,}16\boldsymbol{t}$$

$$\boldsymbol{S}_{\overline{\mathrm{X}}} = \sqrt{\frac{\Sigma \mathrm{X}_{\boldsymbol{i}}^2 - \frac{(\Sigma \mathrm{X}_{\boldsymbol{i}})^2}{\boldsymbol{n}}}{\boldsymbol{n}(\boldsymbol{n}-1)}} = \sqrt{\frac{358{,}95 - \frac{(50{,}1)^2}{7}}{7(7-1)}} = 0{,}095\boldsymbol{t}$$

$$\boldsymbol{t} = \frac{7{,}16 - 7{,}6}{0{,}095} = -4{,}63^{**}$$

**Treći korak:**

$$\boldsymbol{t}_{0,05(6)} = 2{,}447$$
$$\boldsymbol{t}_{0,01(6)} = 3{,}707$$

$$|\boldsymbol{t}| \rangle \boldsymbol{t}_{\alpha(\boldsymbol{n}-1)} \Rightarrow \boldsymbol{H}_1$$

Kako je apsolutna vrednost t-količnika veća od obe kritične vrednosti iz tablice Studentove raspodele altrenativna hipoteza se prihvata za oba praga značajnosti.

Testiranje hipoteze izračunavanjem intervala poverenja:

$$\alpha = 0{,}05$$
$$\overline{X} - t_{\alpha(n-1)} \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + t_{\alpha(n-1)} \cdot S_{\overline{X}}$$
$$7{,}16 - 2{,}447 \cdot 0{,}095 \langle \mu \langle 7{,}16 + 2{,}447 \cdot 0{,}095$$
$$6{,}93 \langle \mu \langle 7{,}39t$$

$$\mu \notin (L_1, L_2) \Rightarrow H_1$$

$$\alpha = 0{,}01$$
$$\overline{X} - t_{\alpha(n-1)} \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + t_{\alpha(n-1)} \cdot S_{\overline{X}}$$
$$7{,}16 - 3{,}707 \cdot 0{,}095 \langle \mu \langle 7{,}16 + 3{,}707 \cdot 0{,}095$$
$$6{,}81 \langle \mu \langle 7{,}51t$$

$$\mu \notin (L_1, L_2) \Rightarrow H_1$$

**Primer 2**: Prinos pšenice sa odabranih 40 ha parcele čija je površina 400 ha dat je u tabeli. Da li se može očekivati da će prosečan prinos pšenice biti na nivou od 3,5 tona?

| **Prinos (t/ha)** | **Površina (ha)** | | |
|---|---|---|---|
| **X** | **f** | **fx** | $fX^2$ |
| **4,9** | **4** | **19,6** | **96,04** |
| **4,5** | **6** | **27,0** | **121,50** |
| **4,3** | **10** | **43,0** | **184,90** |
| **3,8** | **7** | **26,6** | **101,08** |
| **3,4** | **5** | **17,0** | **57,80** |
| **3,0** | **8** | **24,0** | **72,00** |
| | **40** | **157,2** | **633,32** |

$$N = 400$$
$$n = 40$$
$$\mu = 3{,}5t/ha$$

$$H_0 : \overline{X} = \mu$$

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}}$$

$$\overline{X} = \frac{\sum fX}{n} = \frac{157{,}2}{40} = 3{,}93t/ha$$

$$s_{\overline{x}} = \sqrt{\frac{\sum fX^2 - \frac{(\sum fX)^2}{n}}{n(n-1)} \cdot \frac{N-n}{N}}$$

$$S_{\overline{X}} = \sqrt{\frac{633{,}32 - \frac{157{,}2^2}{40}}{40 \cdot 39} \cdot \frac{400-40}{400}} = 0{,}095t/ha$$

$$t = \frac{3{,}93 - 3{,}5}{0{,}095} = 4{,}53^{**}$$

$$z_{0,05} = 1{,}96$$
$$z_{0,01} = 2{,}58$$

Testiranje hipoteze izračunavanjem intervala poverenja:

$$\alpha = 0{,}05$$
$$\overline{X} - Z_\alpha \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + Z_\alpha \cdot S_{\overline{X}}$$
$$3{,}93 - 1{,}96 \cdot 0{,}095 \langle \mu \langle 3{,}93 + 1{,}96 \cdot 0{,}095$$
$$3{,}74 \langle \mu \langle 4{,}12t$$

$$\alpha = 0{,}01$$
$$\overline{X} - Z_\alpha \cdot S_{\overline{X}} \langle \mu \langle \overline{X} + Z_\alpha \cdot S_{\overline{X}}$$
$$3{,}93 - 2{,}58 \cdot 0{,}095 \langle \mu \langle 3{,}93 + 2{,}58 \cdot 0{,}095$$
$$3{,}69 \langle \mu \langle 4{,}18t$$

$$\mu \notin (L_1, L_2) \Rightarrow H_1$$

$$\mu \notin (L_1, L_2) \Rightarrow H_1$$

*Upoređivanje dve aritmetičke sredine iz dva nezavisna uzorka – test značajnosti razlike dve sredine*

Kod ovog vida testiranja polazna pretpostavka glasi:

$$\boldsymbol{H}_0 : \overline{X}_1 = \overline{X}_2 \qquad \mu_1 = \mu_2$$

Za proveru ove hipoteze takođe postoje dva slučaja:

1. kada su poznate standardne devijacije ili varijanse osnovnih skupova

1. kada nisu poznate standardne devijacije ili varijanse osnovnih skupova već ih ocenjujemo na osnovu uzoraka

Kod prvog slučaja kada su poznati pokazatelji osnovnih skupova nulta hipoteza se proverava izračunavanjem količnika Z.

$$\boldsymbol{H}_0 : \overline{X}_1 = \overline{X}_2 \qquad \mu_1 = \mu_2$$

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{(\overline{X}_1 - \overline{X}_2)}}$$

$$|Z| \langle Z_\alpha \Rightarrow \boldsymbol{H}_0$$
$$|Z|^{**} \rangle Z_\alpha \Rightarrow \boldsymbol{H}_1$$
$$|Z|^{*} \rangle Z_{0,05} \Rightarrow \boldsymbol{H}_1$$

$\overline{X}_1, \overline{X}_2 \Rightarrow$ aritmetičke sredine uzoraka ( proste ili ponderisane)

$\sigma_{(\overline{X}_1 - \overline{X}_2)} \Rightarrow$ standardna greška razlike dve sredine $\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Provera hipoteze o značajnosti razlike dve sredine može da se izvede i izračunavanjem odgovarajućeg intervala poverenja.

$$\alpha = 0{,}05$$
$$\alpha = 0{,}01$$
$$\left(\overline{X}_1 - \overline{X}_2\right) - Z_\alpha \cdot \sigma_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \rangle \left(\overline{X}_1 - \overline{X}_2\right) + Z_\alpha \cdot \sigma_{(\overline{X}_1 - \overline{X}_2)}$$

$$0 \in (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_0$$
$$0 \notin (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_1$$

**Primer 1.** Na površini od 35 ha ostvaren je prosečan prinos jedne sorte boranije od 15 t/ha. Prosečan prinos druge sorte je 16,2 t/ha i postignut je na površini od 38 ha. Varijabilitet prosečnih prinosa ovih sorata u direktnoj proizvodnji iskazan vrednostima stndardnih devijacija iznosi 1,6 t/ha i 1,9 t/ha. Da li su proizvodni rezultati ovih sorata jednaki?

$$\boldsymbol{n}_1 = 35\boldsymbol{ha} \qquad \overline{X}_1 = 15\boldsymbol{t}/\boldsymbol{ha} \qquad \boldsymbol{\sigma}_1 = 1{,}6\boldsymbol{t}/\boldsymbol{ha}$$

$$\boldsymbol{n}_2 = 38\boldsymbol{ha} \qquad \overline{X}_2 = 16{,}2\boldsymbol{t}/\boldsymbol{ha} \qquad \boldsymbol{\sigma}_2 = 1{,}9\boldsymbol{t}/\boldsymbol{ha}$$

$$\boldsymbol{H}_0 : \overline{X}_1 = \overline{X}_2 \qquad \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$$

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\boldsymbol{\sigma}_{(\overline{X}_1 - \overline{X}_2)}}$$

$$\boldsymbol{\sigma}_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{\boldsymbol{\sigma}_1^2}{\boldsymbol{n}_1} + \frac{\boldsymbol{\sigma}_2^2}{\boldsymbol{n}_2}} = \sqrt{\frac{(1{,}6)^2}{35} + \frac{(1{,}9)^2}{38}} = 0{,}41\boldsymbol{t}/\boldsymbol{ha}$$

$$Z = \frac{15 - 16{,}2}{0{,}41} = -2{,}927$$

$$|Z|^{**} \rangle Z_{\boldsymbol{\alpha}} \Rightarrow \boldsymbol{H}_1 \qquad Z_{0,05} = 1{,}96 \qquad Z_{0,01} = 2{,}58$$

Testiranje preko intervala poverenja:

$$\begin{aligned}
&\boldsymbol{\alpha} = 0{,}05\\
&(\overline{X}_1 - \overline{X}_2) - Z_{\boldsymbol{\alpha}} \cdot \boldsymbol{\sigma}_{(\overline{X}_1 - \overline{X}_2)} \langle (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \rangle (\overline{X}_1 - \overline{X}_2) + Z_{\boldsymbol{\alpha}} \cdot \boldsymbol{\sigma}_{(\overline{X}_1 - \overline{X}_2)}\\
&(15 - 16{,}2) - 1{,}96 \cdot 0{,}41 \langle (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \rangle (15 - 16{,}2) + 1{,}96 \cdot 0{,}41\\
&-2{,}004 \langle (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \rangle -0{,}396
\end{aligned}$$

$$0 \notin (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_1$$

$$\begin{aligned}
&\alpha = 0{,}01\\
&(\overline{X}_1 - \overline{X}_2) - Z_{\alpha} \cdot \sigma_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \rangle (\overline{X}_1 - \overline{X}_2) + Z_{\alpha} \cdot \sigma_{(\overline{X}_1 - \overline{X}_2)}\\
&(15 - 16{,}2) - 2{,}58 \cdot 0{,}41 \langle (\mu_1 - \mu_2) \rangle (15 - 16{,}2) + 2{,}58 \cdot 0{,}41\\
&-2{,}258 \langle (\mu_1 - \mu_2) \rangle -0{,}142
\end{aligned}$$

$$0 \notin (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_1$$

*Testiranje značajnosti razlike dve sredine u slučaju kada nisu poznati pokazatelji varijabiliteta osnovnih skupova*

U ovom slučaju nepoznate pokazatelje varijabiliteta osnovnog skupa zamenjujemo njihovim ocenama na osnovu uzoraka, a polazna pretpostavka se proverava izračunavanjem količnika t.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{(\overline{X}_1 - \overline{X}_2)}}$$

$S_{(\overline{X}_1 - \overline{X}_{21})} \Rightarrow$ ocenjena standardna greška razlike dve sredine

Ako su uzorci sa jednakim brojem jedinica koristi se sledeći izraz:

$$n_1 = n_2 = n \Rightarrow S_{(\overline{X}_1 - \overline{X}_2)} = \sqrt{\frac{2 \cdot S_{1+2}^2}{n}}$$

Ako su uzorci sa nejednakim brojem jedinica izračunavanje se izvodi na sledeći način:

$$n_1 \neq n_2 \Rightarrow S_{(\overline{X}_1 - \dot{X}_2)} = \sqrt{S_{1+2}^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Združena varijansa za negrupisane podatke:

$$S_{1+2}^2 = \frac{\Sigma\left(X_{1i} - \overline{X}_1\right)^2 + \Sigma\left(X_{2i} - \overline{X}_2\right)^2}{n_1 + n_2 - 2}$$

$$S_{1+2}^2 = \frac{\Sigma X_{1i}^2 - \frac{\left(\Sigma X_{1i}\right)^2}{n_1} + \Sigma X_{2i}^2 - \frac{\left(\Sigma X_{2i}\right)^2}{n_2}}{n_1 + n_2 - 2}$$

Združena varijansa za grupisane podatke:

$$S_{1+2}^2 = \frac{\Sigma f_1\left(X_{1i} - \overline{X}_1\right)^2 + \Sigma f_2\left(X_{2i} - \overline{X}_2\right)^2}{n_1 + n_2 - 2}$$

$$S_{1+2}^2 = \frac{\Sigma f_1 X_{1i}^2 - \frac{\left(\Sigma f_1 X_{1i}\right)^2}{n_1} + \Sigma f_2 X_{2i}^2 - \frac{\left(\Sigma f_2 X_{2i}\right)^2}{n_2}}{n_1 + n_2 - 2}$$

Izračunata vrednost količnika t upoređuje se sa kritičnim vrednostima iz tablica Studentove distribucije ako su $n_1$ i $n_2$ mali uzorci, odnosno ako njihov zbir nije veći od 30 jedinica, ili sa kritičnim vrednostima iz tablica Normalne raspodele ako su uzorci ili njihov zbir veći od 30 jedinica.

$$|t| \langle t_{\alpha(n_1+n_2-2)}; Z_\alpha \Rightarrow H_0$$

$$|t| \rangle t_{\alpha(n_1+n_2-2)}; Z_\alpha \Rightarrow H_1$$

Testiranje postavljene nulte hipoteze možemo izvesti i izračunavanjem odgovarajućeg intervala poverenja:

$$
\begin{aligned}
&\boldsymbol{n}_1, \boldsymbol{n}_2 \rangle 30 \\
&\boldsymbol{n}_1 + \boldsymbol{n}_2 \rangle 30 \\
&\boldsymbol{\alpha} = 0{,}05 \\
&\boldsymbol{\alpha} = 0{,}01 \\
&\left(\overline{X}_1 - \overline{X}_2\right) - Z_{\alpha} \cdot \boldsymbol{S}_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \langle \left(\overline{X}_1 - \overline{X}_2\right) + Z_{\alpha} \cdot \boldsymbol{S}_{(\overline{X}_1 - \overline{X}_2)}
\end{aligned}
$$

$$
\begin{aligned}
&\boldsymbol{n}_1, \boldsymbol{n}_2 \langle 30 \\
&\boldsymbol{n}_1 + \boldsymbol{n}_2 \langle 30 \\
&\boldsymbol{\alpha} = 0{,}05 \\
&\boldsymbol{\alpha} = 0{,}01 \\
&\left(\overline{X}_1 - \overline{X}_2\right) - \boldsymbol{t}_{\alpha(\boldsymbol{n}_1 + \boldsymbol{n}_2 - 2)} \cdot \boldsymbol{S}_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \langle \left(\overline{X}_1 - \overline{X}_2\right) + \boldsymbol{t}_{\alpha(\boldsymbol{n}_1 + \boldsymbol{n}_2 - 2)} \cdot \boldsymbol{S}_{(\overline{X}_1 - \overline{X}_2)}
\end{aligned}
$$

$$
\begin{aligned}
&0 \in (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_0 \\
&0 \notin (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_1
\end{aligned}
$$

**Primer 1.** Utvrditi da li postoje razlike u prosečnim prinosima dve posmatrane sorte šećerne repe, na osnovu prinosa po jedinici površine datih u tabeli.

| **Sorta A** | | **Sorta B** | | **f1X1** | **f2X2** | **f1X12** | **f2X22** |
|---|---|---|---|---|---|---|---|
| **Prinos** | **Površina** | **Prinos** | **Površina** | | | | |
| **X1** | **f1** | **X2** | **f2** | | | | |
| **4,2** | **3** | **4,6** | **2** | **12,6** | **9,2** | **52,92** | **42,32** |
| **4,6** | **4** | **4,9** | **4** | **18,4** | **19,6** | **84,64** | **96,04** |
| **5,1** | **6** | **5,3** | **5** | **30,6** | **26,5** | **156,06** | **140,45** |
| **5,3** | **5** | **5,9** | **4** | **26,5** | **23,6** | **140,45** | **139,24** |
| **5,8** | **2** | **6,2** | **3** | **11,6** | **18,6** | **67,28** | **115,32** |
| | **20** | | **18** | **99,7** | **97,5** | **501,35** | **533,37** |

$$
\begin{aligned}
&\boldsymbol{H}_0 : \overline{X}_1 = \overline{X}_2 \\
&\mu_1 = \mu_2
\end{aligned}
$$

$$
\boldsymbol{n}_1 \neq \boldsymbol{n}_2 \Rightarrow \boldsymbol{S}_{(\overline{X}_1 - \dot{X}_2)} = \sqrt{\boldsymbol{S}^2_{1+2} \cdot \left(\frac{1}{\boldsymbol{n}_1} + \frac{1}{\boldsymbol{n}_2}\right)}
$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{(\overline{X}_1 - \overline{X}_2)}} = \frac{4{,}985 - 5{,}417}{0{,}168} = -2{,}57$$

$$S_{1+2}^2 = \frac{\Sigma f_1 X_{1i}^2 - \frac{(\Sigma f_1 X_{1i})^2}{n_1} + \Sigma f_2 X_{2i}^2 - \frac{(\Sigma f_2 X_{2i})^2}{n_2}}{n_1 + n_2 - 2} =$$

$$= \frac{501{,}35 - \frac{(99{,}7)^2}{20} + 533{,}37 - \frac{(97{,}5)^2}{18}}{20 + 18 - 2} = 0{,}266$$

$$S_{(\overline{X}_1 - \dot{X}_2)} = \sqrt{0{,}266 \cdot \left(\frac{1}{20} + \frac{1}{18}\right)} = 0{,}168 t / ha$$

$$Z_{0,05} = 1{,}96$$
$$Z_{0,01} = 2{,}58$$

$$|t|^* \rangle Z_{0,05} \Rightarrow H_1$$
$$|t| \langle Z_{0,01} \Rightarrow H_0$$

Provera postavljene hipoteze izračunavanjem intervala poverenja:

$$n_1 + n_2 \rangle 30$$
$$\alpha = 0{,}05$$
$$(\overline{X}_1 - \overline{X}_2) - Z_\alpha \cdot S_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \langle (\overline{X}_1 - \overline{X}_2) + Z_\alpha \cdot S_{(\overline{X}_1 - \overline{X}_2)}$$
$$(4{,}985 - 5{,}417) - 1{,}96 \cdot 0{,}168 \langle (\mu_1 - \mu_2) \langle (4{,}985 - 5{,}417) + 1{,}96 \cdot 0{,}168$$
$$-0{,}761 \langle (\mu_1 - \mu_2) \langle -0{,}103$$
$$0 \notin (L_1, L_2) \Rightarrow H_1$$

$$n_1 + n_2 \rangle 30$$
$$\alpha = 0{,}01$$
$$(\overline{X}_1 - \overline{X}_2) - Z_\alpha \cdot S_{(\overline{X}_1 - \overline{X}_2)} \langle (\mu_1 - \mu_2) \langle (\overline{X}_1 - \overline{X}_2) + Z_\alpha \cdot S_{(\overline{X}_1 - \overline{X}_2)}$$
$$(4{,}985 - 5{,}417) - 2{,}58 \cdot 0{,}168 \langle (\mu_1 - \mu_2) \langle (4{,}985 - 5{,}417) + 2{,}58 \cdot 0{,}168$$
$$-0{,}865 \langle (\mu_1 - \mu_2) \langle 0{,}001$$
$$0 \in (L_1, L_2) \Rightarrow H_0$$

# Testiranja proporcija

Kod testiranja proporcija imamo dva vida testiranja:

1. kada upoređujemo proporciju iz uzorka sa pretpostavljenom proporcijom osnovnog skupa ili sa nekom teorijskom vrednošću

2. kada upoređujemo dve proporcije iz dva nezavisna uzorka

## *Test značajnosti jedne proporcije*

Pri testiranju jednakosti proporcije uzorka sa pretpostavljenom proporcijom osnovnog skupa polazi se od sledeće pretpostavke:

$$\boldsymbol{H}_0 : \hat{\boldsymbol{p}} = \boldsymbol{p}$$
$$\boldsymbol{H}_1 : \hat{\boldsymbol{p}} \neq \boldsymbol{p}$$

Ova hipoteza se proverava izračunavanjem količnika ***t*:**

$$\boldsymbol{t} = \frac{\hat{\boldsymbol{p}} - \boldsymbol{p}}{\boldsymbol{S}_{\hat{\boldsymbol{p}}}}$$

Izračunati količnik t upoređujemo sa vrednostima iz tablica t – distribucije ako se testiranje izvodi na osnovu malog uzorka, ili sa kritičnim vrednostima iz tablica Normalne raspodele ako se testiranje izvodi na osnovu velikog uzorka.

$$\begin{aligned} &\boldsymbol{n} \langle 30 \\ &|\boldsymbol{t}| \langle \boldsymbol{t}_{\alpha(\boldsymbol{n}-1)} \Rightarrow \boldsymbol{H}_0 \\ &|\boldsymbol{t}|^{**} \rangle \boldsymbol{t}_{\alpha(\boldsymbol{n}-1)} \Rightarrow \boldsymbol{H}_1 \\ &|\boldsymbol{t}|^{*} \rangle \boldsymbol{t}_{0,05(\boldsymbol{n}-1)} \Rightarrow \boldsymbol{H}_1 \end{aligned}$$

$$\begin{aligned} &\boldsymbol{n} \rangle 30 \\ &|\boldsymbol{t}| \langle Z_\alpha \Rightarrow \boldsymbol{H}_0 \\ &|\boldsymbol{t}|^{**} \rangle Z_\alpha \Rightarrow \boldsymbol{H}_1 \\ &|\boldsymbol{t}|^{*} \rangle Z_{0,05} \Rightarrow \boldsymbol{H}_1 \end{aligned}$$

Testiranje značajnosti jedne proporcije možemo izvesti i izračunavanjem intervala poverenja.

$$\begin{aligned} &\alpha = 0{,}05 \\ &\alpha = 0{,}01 \\ &\boldsymbol{n} \langle 30 \\ &\hat{\boldsymbol{p}} - \boldsymbol{t}_{\alpha(\boldsymbol{n}-1)} \cdot \boldsymbol{S}_{\hat{\boldsymbol{p}}} \langle \boldsymbol{p} \langle \hat{\boldsymbol{p}} + \boldsymbol{t}_{\alpha(\boldsymbol{n}-1)} \cdot \boldsymbol{S}_{\hat{\boldsymbol{p}}} \\ &\boldsymbol{n} \rangle 30 \\ &\hat{\boldsymbol{p}} - Z_\alpha \cdot \boldsymbol{S}_{\hat{\boldsymbol{p}}} \langle \boldsymbol{p} \langle \hat{\boldsymbol{p}} + Z_\alpha \cdot \boldsymbol{S}_{\hat{\boldsymbol{p}}} \end{aligned}$$

$$\begin{aligned} &\boldsymbol{p} \in (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_0 \\ &\boldsymbol{p} \notin (\boldsymbol{L}_1, \boldsymbol{L}_2) \Rightarrow \boldsymbol{H}_1 \end{aligned}$$

**Primer1.** U uzorku od 164 grla jedne rase goveda sa dužinom trupa ispod 160 cm bilo je 52 grla. Može li se doneti zaključak da je kod posmatrane rase učešće grla sa dužinom trupa ispod 160 cm 40 %.

$$n = 164 \quad a = 52 \quad p = 0{,}4$$

$$H_0 : \hat{p} = p \qquad H_1 : \hat{p} \neq p$$

$$\hat{p} = \frac{a}{n} = \frac{52}{164} = 0{,}317 \qquad \hat{q} = 1 - 0{,}317 = 0{,}683$$

$$S_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{0{,}317 \cdot 0{,}683}{164}} = 0{,}036$$

$$t = \frac{\hat{p} - p}{S_{\hat{p}}} = \frac{0{,}317 - 0{,}4}{0{,}036} = -2{,}306^{*}$$

$$Z_{0,05} = 1{,}96 \qquad Z_{0,01} = 2{,}58$$

$$|t|^{*} \rangle Z_{0,05} \Rightarrow H_1 \qquad |t| \langle Z_{0,01} \Rightarrow H_0$$

Testiranje hipoteze možemo izvesti i izračunavanjem intervala poverenja:

$$\alpha = 0{,}05$$
$$n \rangle 30$$
$$\hat{p} - Z_{\alpha} \cdot S_{\hat{p}} \langle p \langle \hat{p} + Z_{\alpha} \cdot S_{\hat{p}}$$
$$0{,}317 - 1{,}96 \cdot 0{,}036 \langle p \langle 0{,}317 + 1{,}96 \cdot 0{,}036$$
$$0{,}246 \langle p \langle 0{,}388$$
$$p \notin (L_1, L_2) \Rightarrow H_1$$

$$\alpha = 0{,}01$$
$$n \rangle 30$$
$$\hat{p} - Z_{\alpha} \cdot S_{\hat{p}} \langle p \langle \hat{p} + Z_{\alpha} \cdot S_{\hat{p}}$$
$$0{,}317 - 2{,}58 \cdot 0{,}036 \langle p \langle 0{,}317 + 2{,}58 \cdot 0{,}036$$
$$0{,}224 \langle p \langle 0{,}410$$
$$p \in (L_1, L_2) \Rightarrow H_0$$

## *Upoređivanje dve proporcije iz dva nezavisna uzorka*

Pri testitaranju značajnosti razlike dve proporcije iz dva nezavisna uzorka polazi se od sledeće nulte hipoteze:

$$H_0 : \hat{p}_1 = \hat{p}_2$$
$$p_1 = p_2$$

Polazna hipoteza proverava se izračunavanjem sledećeg količnika:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{S_{(\hat{p}_1 - \hat{p}_2)}}$$

$S_{(\hat{p}_1 - \hat{p}_2)} \Rightarrow$ ocenjena standardna greška razlike dve proporcije

Ocenjena standardna greška razlike dve proporcije može se izračunati primenom dva izraza:

$$S_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{2 \cdot \overline{p}\overline{q}}{n_1 + n_2}} \qquad S_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\overline{p}\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Da bi se izračunala standardna greška prvo se izračunava prosečna proporcija na osnovu dva uzorka.

$$\overline{p} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$$
$$\overline{p} = \frac{a_1 + a_2}{n_1 + n_2}$$

$$\overline{q} = 1 - \overline{p}$$

Izračunati količnik Z upoređuje se sa kritičnim vrednostima iz tablica Normalne raspodele.

$$|Z| \langle Z_\alpha \Rightarrow H_0$$
$$|Z|^{**} \rangle Z_\alpha \Rightarrow H_1$$
$$|Z|^{*} \rangle Z_{0,05} \Rightarrow H_1$$

Testiranje ove hipoteze možemo izvesti i izračunavanjem intervala poverenja:

$$\alpha = 0{,}05$$
$$\alpha = 0{,}01$$
$$(\hat{p}_1 - \hat{p}_2) - Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)} \langle (p_1 - p_2) \rangle (\hat{p}_1 - \hat{p}_2) + Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)}$$
$$0 \in (L_1, L_2) \Rightarrow H_0$$
$$0 \notin (L_1, L_2) \Rightarrow H_1$$

**Primer 2.** U uzorcima od po 110 grla, goveda dve rase obolela grla učestvuju sa 6 % i 13 %. Utvrditi da li je otpornost dve posmatrane rase goveda prema ispitivanoj bolesti ista.

$$\begin{aligned} n_1 &= 110 \\ \hat{p}_1 &= 0{,}06 \end{aligned} \qquad \begin{aligned} n_2 &= 110 \\ \hat{p}_2 &= 0{,}13 \end{aligned} \qquad \begin{aligned} H_0 &: \hat{p}_1 = \hat{p}_2 \\ p_1 &= p_2 \end{aligned} \qquad Z = \frac{\hat{p}_1 - \hat{p}_2}{S_{(\hat{p}_1 - \hat{p}_2)}}$$

$$\bar{p} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2} = \frac{0{,}06 \cdot 110 + 0{,}13 \cdot 110}{110 + 110} = 0{,}095 \qquad \bar{q} = 1 - \bar{p} = 0{,}905$$

**I način:**

$$S_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{0{,}095 \cdot 0{,}905\left(\frac{1}{110} + \frac{1}{110}\right)} = 0{,}0396$$

$$Z = \frac{0{,}06 - 0{,}13}{0{,}0396} = -1{,}768 \qquad \begin{aligned} Z_{0,05} &= 1{,}96 \\ Z_{0,01} &= 2{,}58 \end{aligned} \qquad |Z| \langle Z_\alpha \Rightarrow H_0$$

$$\begin{aligned}
&\alpha = 0{,}05 \\
&(\hat{p}_1 - \hat{p}_2) - Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)} \langle (p_1 - p_2) \rangle (\hat{p}_1 - \hat{p}_2) + Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)} \\
&(0{,}06 - 0{,}13) - 1{,}96 \cdot 0{,}0396 \langle (p_1 - p_2) \rangle (0{,}06 - 0{,}13) + 1{,}96 \cdot 0{,}0396 \\
&-0{,}148 \langle (p_1 - p_2) \rangle 0{,}008 \\
&0 \in (L_1, L_2) \Rightarrow H_0
\end{aligned}$$

**II način:**

$$S_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{2 \cdot \bar{p}\bar{q}}{n_1 + n_2}} = \sqrt{\frac{2 \cdot 0{,}095 \cdot 0{,}905}{110 + 110}} = 0{,}028$$

$$Z = \frac{0{,}06 - 0{,}13}{0{,}028} = -2{,}5^* \qquad \begin{aligned} Z_{0,05} &= 1{,}96 \\ Z_{0,01} &= 2{,}58 \end{aligned} \qquad \begin{aligned} |Z| &\rangle Z_{0,05} \Rightarrow H_1 \\ |Z| &\langle Z_{0,01} \Rightarrow H_0 \end{aligned}$$

$$\begin{aligned}
&\alpha = 0{,}05 \\
&(\hat{p}_1 - \hat{p}_2) - Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)} \langle (p_1 - p_2) \rangle (\hat{p}_1 - \hat{p}_2) + Z_\alpha \cdot S_{(\hat{p}_1 - \hat{p}_2)} \\
&(0{,}06 - 0{,}13) - 1{,}96 \cdot 0{,}028 \langle (p_1 - p_2) \rangle (0{,}06 - 0{,}13) + 1{,}96 \cdot 0{,}028 \\
&-0{,}125 \langle (p_1 - p_2) \rangle -0{,}015 \\
&0 \notin (L_1, L_2) \Rightarrow H_1
\end{aligned}$$