Testiranje hipoteza

Proveravanje tvrđenja o nepoznatim parametrima populacije zasnovano na informaciji iz uzorka

Elementi testiranja hipoteza:

- **Nulta hipoteza-** Tvrđenje o nepoznatim parametrima osnovnog skupa
- **Alternativna hipoteza**- Tvrđenje suprotno tvrđenju nulte hipoteze
- **Test kriterijum**-Funkcija uzorka koja pod nultom hipotezom ima poznatu raspodelu
- **Kritična oblast-**Oblast odbacivanja nulte hipoteze

Postupak testiranja hipoteze se izvodi u nekoliko koraka:

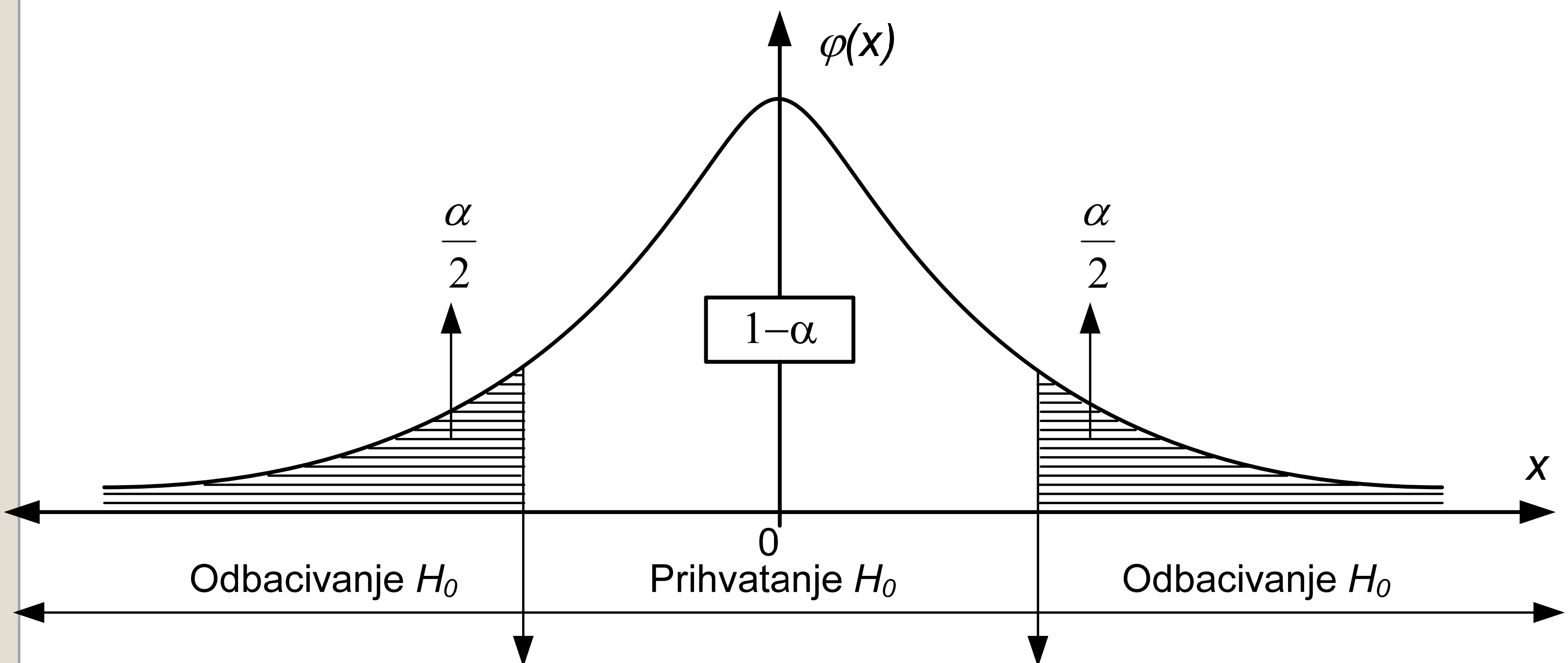
1. Definišu se nulta i alternativna hipoteza.
2. Izbor modela teorijskog rasporeda-test kriterijuma.
3. Određuje se nivo značajnosti testa α odnosno verovatnoća $(1-\alpha)$.
4. Izbor uzorka.
5. Izračunavanje statistike testa na osnovu uzorka.
6. Iz tablice teorijskog rasporeda očitava se tablična vrednost (kriterijum).
7. Upoređivanje statistike testa sa tabličnom vrednošću.
8. Odluka o prihvatanju ili odbacivanju formulisane hipoteze.

Greške I i II vrste

	Stvarno stanje	
Odluka:	H_0 je tačna	H_0 je pogrešna
Odbacivanje H_0	**POGREŠAN ZAKLJUČAK** **Greška I vrste** $P(H_1/H_0)=\alpha$	**TAČAN ZAKLJUČAK** $P(H_1/H_1)=1-\beta$ **MOĆ TESTA [POWER]**
Ne odbacivanje H_0	**TAČAN ZAKLJUČAK** $P(H_0/H_0)=1-\alpha$ **POVERENJE [CONFIDENCE]**	**POGREŠAN ZAKLJUČAK** **Greška II vrste** $P(H_0/H_1)=\beta$

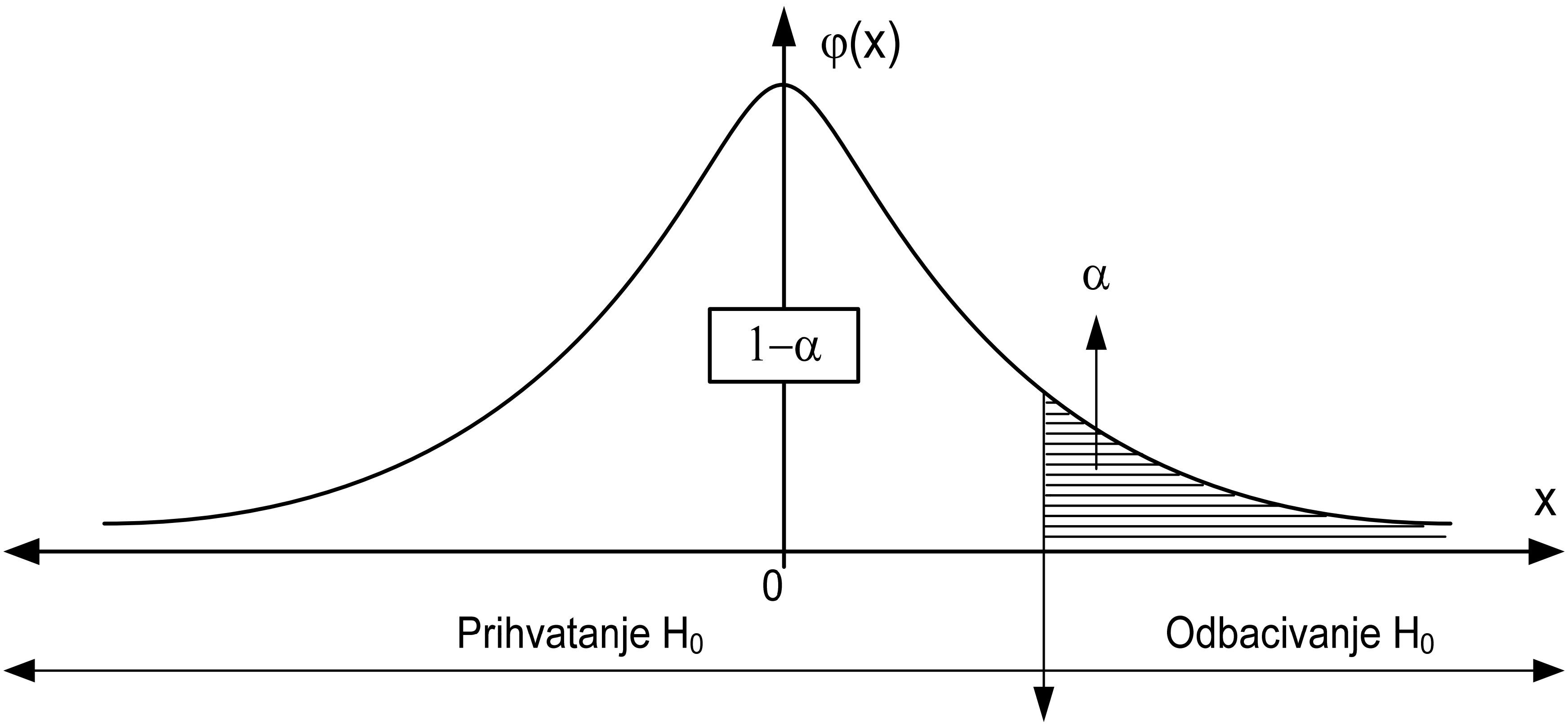
Područja prihvatanja i odbacivanja hipoteza

1.Dvosmerni test: H_0: $\mu = \mu_0$; H_1: $\mu \neq \mu_0$.



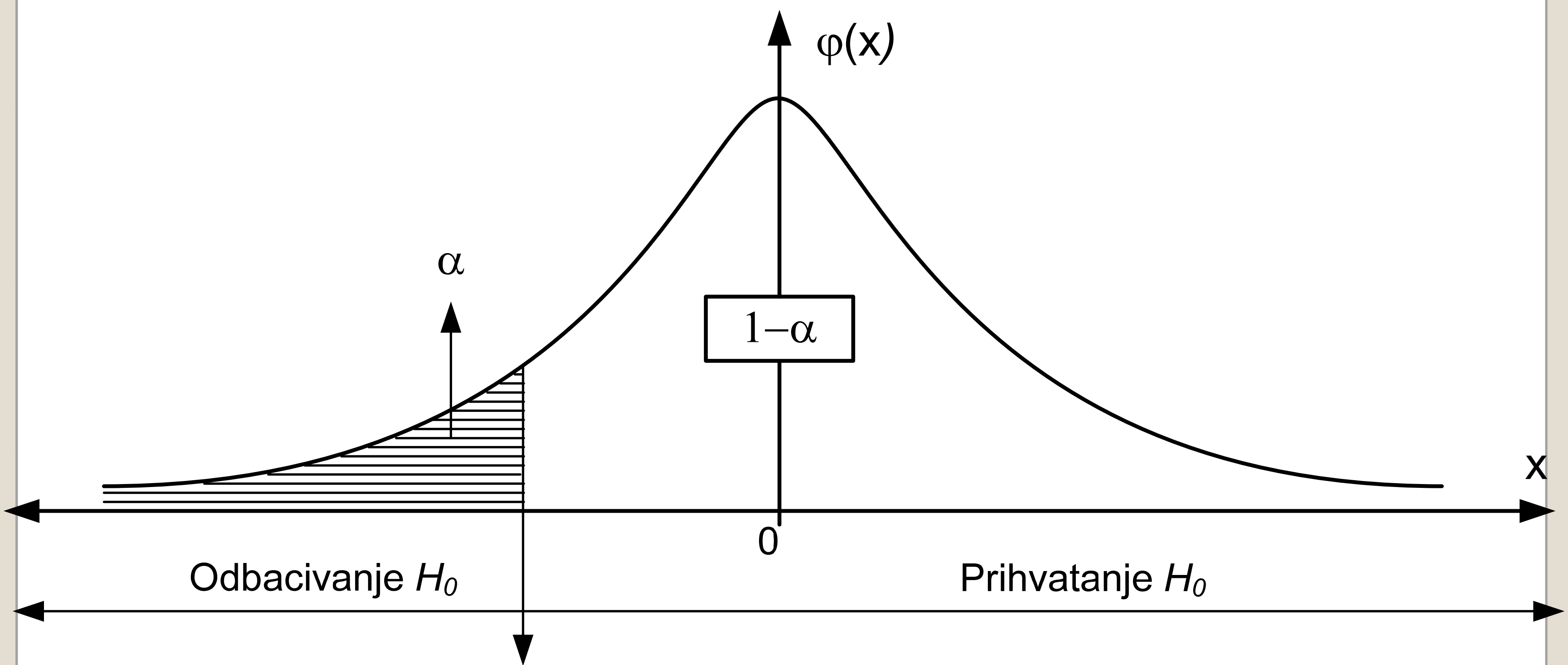
Oblasti prihvatanja i odbacivanja nulte hipoteze H_0

2.Jednosmerni test: H_0: $\mu \leq \mu_0$; H_1: $\mu > \mu_0$.



Oblasti prihvatanja i odbacivanja nulte hipoteze *H_0*

3.Jednosmerni test: H_0: $\mu \geq \mu_0$; H_1: $\mu < \mu_0$



Oblasti prihvatanja i odbacivanja nulte hipoteze H_0

TESTIRANJE HIPOTEZA U SLUČAJU JEDNOG OSNOVNOG SKUPA

Testiranje nulte hipoteze o nepoznatoj aritmetičkoj sredini osnovnog skupa μ na osnovu prostog slučajnog uzorka

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

1. Slučaj: σ^2 poznata

Primer 1: Rezultati jednog uzorka n=30, dali su da je prosečan prinos graška $\overline{X} = 3{,}6$ (t/ha). Uporediti ovaj prinos sa prinosom standardnih sorti $\mu_0 = 3{,}9$ i $\sigma^2 = 1{,}44$.

Rešenje:

n = 30

$\overline{X} = 3{,}6$ (t/ha)

$\mu_0 = 3{,}9$ (t/ha)

$\sigma^2 = 1{,}44 \Rightarrow \sigma = 1{,}2$ (t/ha)

$$\boxed{\begin{array}{l} H_0 : \mu = 3{,}9 \\ H_1 : \mu \neq 3{,}9 \end{array}}$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}}$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{1{,}2}{\sqrt{30}} = \frac{1{,}2}{5{,}4772} = 0{,}2191 \text{(t/ha)}$$

$$Z = \frac{3{,}6 - 3{,}9}{0{,}2191} = -1{,}37$$

$Z_{0,05} = 1{,}96$ $\qquad |Z| < Z_{0,05} \quad \Rightarrow \quad H_o$

Istu hipotezu možemo proveriti i putem $(1-\alpha)\%$ intervala poverenja:

$$\overline{X} - z_{\alpha}\sigma_{\overline{X}} < \mu < \overline{X} + z_{\alpha}\sigma_{\overline{X}}.$$

95% interval poverenja za parametar μ ocenjen na osnovu uzorka je:

$$3{,}6 - 1{,}96 \cdot 0{,}2191 < \mu < 3{,}6 + 1{,}96 \cdot 0{,}2191$$

$$3{,}17 < \mu < 4{,}03 \text{ (t/ha)}$$

Hipotetička vrednost μ_0 =3,9 (t/ha) pripada intervalu poverenja tako da nema razloga da se odbaci nulta hipoteza na pragu značajnosti 5%.

$$\mu_0 \in (3{,}17, 4{,}03) \quad \Rightarrow \quad H_o$$

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

Primer 2. Pri ispitivanju jedne vrste pesticida na klijavost semena hibridnog kukuruza (%) dobijeni su sledeći rezultati:

X	X^2
66	4356
67	4489
70	4900
69	4761
72	5184
75	5625
72	5184
68	4624
65	4225
73	5329
697	48677

Može li se pretpostaviti da se prosečan procenat klijavosti ne razlikuje značajno od poznate vrednosti $\mu_0 = 80\%$?

Rešenje:

μ_0 =80%   n =10

$t_{0,05}$ (9) = 2,262

$t_{0,01}$ (9) = 3,250

$$\boxed{\begin{aligned} &H_0: \mu = 80(\%) \\ &H_1: \mu \neq 80(\%) \end{aligned}}$$

$$t = \frac{\overline{X} - \mu_0}{s_{\overline{X}}}$$

$$\overline{X} = \frac{\sum X}{n} = \frac{697}{10} = 69{,}7\,(\%)$$

$$s_{\overline{X}} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n(n-1)}} =$$

$$= \sqrt{\frac{48677 - \frac{697^2}{10}}{10 \cdot 9}} = \sqrt{\frac{48677 - 48580{,}9}{90}} = 1{,}0333\,(\%)$$

$$t = \frac{69{,}7 - 80}{1{,}0333} = -9{,}97^{**}$$

Za $\alpha = 0{,}05$, $|t| = 9{,}97 > 2{,}262 \quad \Rightarrow \quad H_1$

Za $\alpha = 0{,}01$, $|t| = 9{,}97 > 3{,}25 \quad \Rightarrow \quad H_1$

Ili:

$$69{,}7 - 2{,}262 \cdot 1{,}0333 < \mu < 69{,}7 + 2{,}262 \cdot 1{,}0333$$

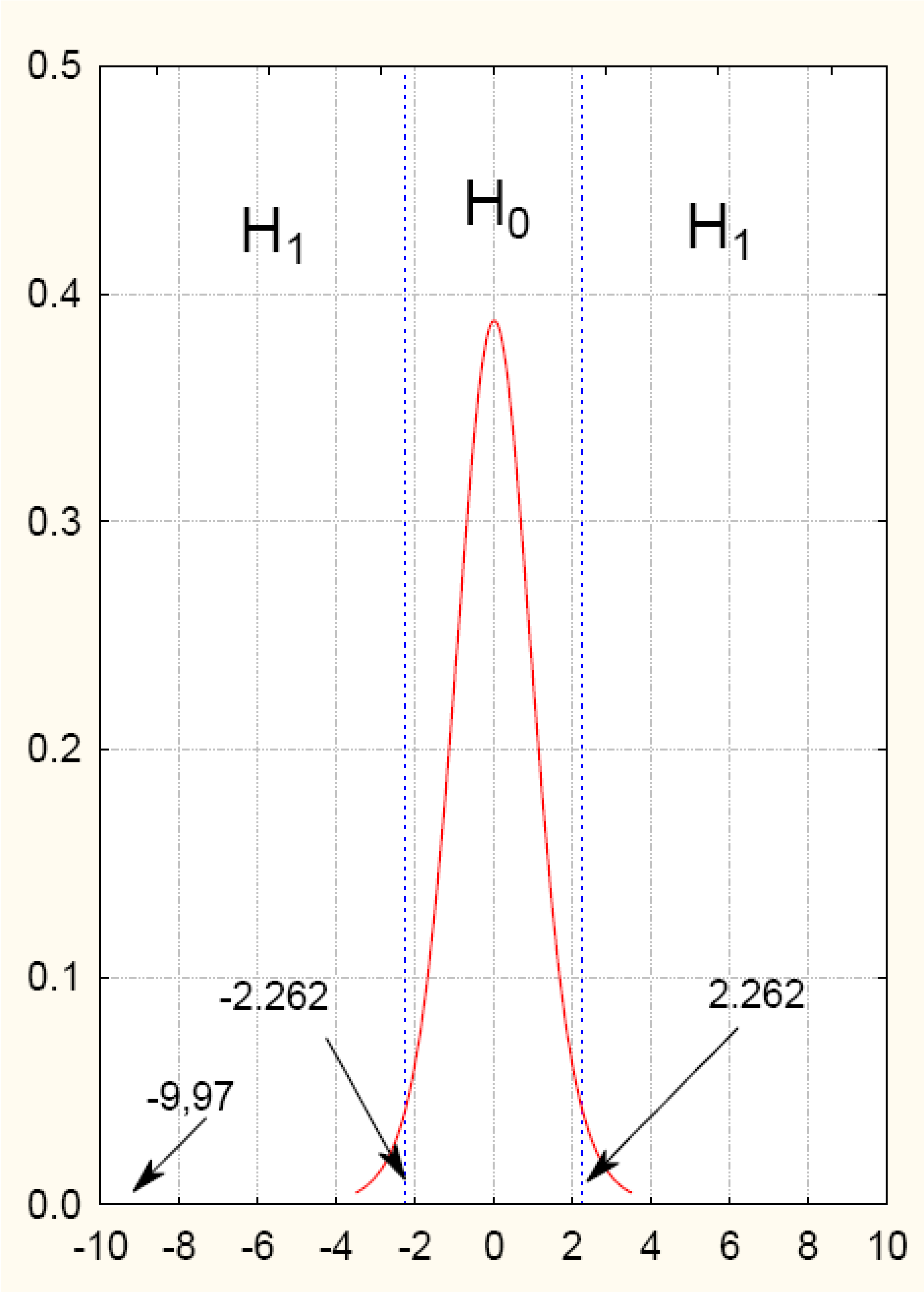
$$67{,}36 < \mu < 72{,}04(\%)$$

$$80 \notin (67{,}36,\ 72{,}04) \quad \Rightarrow \quad H_1$$

$$69{,}7 - 3{,}25 \cdot 1{,}0333 < \mu < 69{,}7 + 3{,}25 \cdot 1{,}0333$$

$$66{,}34 < \mu < 73{,}06(\%)$$

$$80 \notin (66{,}34,\ 73{,}06) \quad \Rightarrow \quad H_1$$


0.5
0.4
0.3
0.2
0.1
0.0
-10
-8
-6
-4
-2
0
2
4
6
8
10
H1
H0
H1
-2.262
2.262
-9,97

Primer 3: Da bi se ocenio prosečan prinos pšenice odabrano je 40 ha sa parcele površine N=400. U istraživanju je primenjen prost slučajan uzorak bez ponavljanja i dobijeni su sledeći rezultati. Da li može da se prihvati nulta hipoteza da je μ_0=3,5?

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

Rešenje:

$$\boxed{\begin{array}{l} H_0: \mu = 3{,}5 \\ H_1: \mu \neq 3{,}5 \end{array}}$$

μ_0=3,5 n = 40 N=400

$$t_{0,05}(39) \approx t_{0,05}(40) = 2{,}021 \text{ ili } z_{0,05} = 1{,}96$$

$$t_{0,01}(39) \approx t_{0,01}(40) = 2{,}704 \text{ ili } z_{0,01} = 2{,}58$$

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

$$s_{\overline{X}} = \sqrt{\frac{\sum fX^2 - \frac{(\sum fX)^2}{n}}{n(n-1)} \cdot \frac{N-n}{N}} = \sqrt{\frac{633{,}32 - \frac{157{,}2^2}{40}}{40 \cdot 39} \cdot \frac{400-40}{400}}$$

$$s_{\overline{X}} = \sqrt{\frac{633{,}32 - 617{,}796}{1560} \cdot 0{,}9} = 0{,}095\,(\text{t/ha})$$

$$t = \frac{\overline{X} - \mu_0}{s_{\overline{X}}} = \frac{3{,}93 - 3{,}5}{0{,}095} = 4{,}53^{**}$$

$$t > t_{0,05} \Rightarrow H_1 \qquad t > t_{0,01} \Rightarrow H_1$$

Zaključivanje na osnovu intervala poverenja

$\alpha = 0,05$

$$3,93 - 2,021 \cdot 0,095 < \mu < 3,93 + 2,021 \cdot 0,095$$

$$3,74 < \mu < 4,12 \, (\text{t/ha})$$

$$3,5 \notin (3,74,\ 4,12) \quad \Rightarrow \text{ prihvata se } H_1 \text{ (ili } H_o \text{ se odbacuje)}$$

$\alpha = 0,01$

$$3,93 - 2,704 \cdot 0,095 < \mu < 3,93 + 2,704 \cdot 0,095$$

$$3,67 < \mu < 4,19 \, (\text{t/ha})$$

$$3,5 \notin (3,67,\ 4,19) \quad \Rightarrow H_o \text{ se odbacuje ili prihvata se } H_1$$

Kako μ_0 ne pripada 95% i 99% intervalima poverenja H_0 odbacuje. Prema tome prosečan prinos pšenice se statistički visoko značajno razlikuje od pretpostavljene vrednosti 3,5 (t/ha).

Testirnje hipoteze o proporciji osnovnog skupa u slučaju velikih uzoraka

$$H_0 : p = p_0$$
$$H_1 : p \neq p_0$$

Pretpostavka za primenu testa: np>5 i nq>5

Primer 1: U uzorku veličine n=100 izniklo je 80 biljaka. Testirati značajnost razlike proporcije izniklih biljaka u uzorku i pretpostavljene proporcije u osnovnom skupu $p_0 = 0,9$.

Rešenje:

a = 80 n = 100

$p_0 = 0,9 \qquad \hat{p} = \dfrac{80}{100} = 0,80 \qquad \hat{q} = 1 - 0,8 = 0,2$

$$\boxed{\begin{array}{l} H_0 : p = 0,9 \\ H_1 : p \neq 0,9 \end{array}}$$

$$Z = \frac{\hat{p} - p_0}{s_{\hat{p}}}$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{0,80 \cdot 0,20}{100}} = 0,04$$

$$z = \frac{0{,}80 - 0{,}90}{0{,}04} = -2{,}5*$$

Na pragu značajnosti α=0,05

$|z| > z_{0,05}$ $\Rightarrow$ prihvata se H_1 ili H_o se odbacuje

Na pragu značajnosti α=0,01

$|z| < z_{0,01}$ $\Rightarrow$ H_o se prihvata

II način testiranja primenom

$\alpha = 0{,}05$

$$0{,}80 - 1{,}96 \cdot 0{,}04 < p < 0{,}80 + 1{,}96 \cdot 0{,}04$$

$$0{,}722 < p < 0{,}878$$

$p_0 = 0{,}9 \notin (0{,}722, 0{,}878)$ $\Rightarrow$ prihvata se H_1 ili H_o se odbacuje

$\alpha = 0{,}01$

$$0{,}80 - 2{,}58 \cdot 0{,}040 < p < 0{,}80 + 2{,}58 \cdot 0{,}040$$

$$0{,}697 < p < 0{,}903$$

$p_0 \in (0{,}697,\ 0{,}903)$ $\Rightarrow$ H_o se prihvata

TESTIRANJE HIPOTEZA U SLUČAJU DVA OSNOVNA SKUPA

Zaključivanje o jednakosti aritmetičkih sredina dva osnovna skupa na osnovu nezavisnih prostih slučajnih uzoraka

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

1. Slučaj: σ_1^2 i σ_2^2 poznate

Primer 1: Cilj eksperimenta je da se utvrdi da li postoji statistički značajna razlika u prosečnom prinosu dve sorte pšenice. Izabrana su dva nezavisna slučajna uzorka. Na osnovu uzorka od 25 (ha) dobijen je prosečni prinos prve sorte 6,2 (t/ha), dok je na osnovu uzorka od 36 (ha) dobijen prosečni prinos druge sorte 4,5 (t/ha). Ako su poznati varijabiliteti posmatranih sorti $\sigma_1^2 = 1$ i $\sigma_2^2 = 0,75$, testirati nultu hipotezu da ne postoji statistički značajna razlika u prosečnom prinosu posmatranih sorti.

Rešenje:

$n_1 = 25$ $n_2 = 36$ $\overline{X}_1 = 6,2$ (t/ha) $\overline{X}_2 = 4,5$ (t/ha)

$\sigma_1^2 = 1$ $\sigma_2^2 = 0,75$ $z_{0,05} = 1,96$, $z_{0,01} = 2,58$

$$\boxed{\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned}}$$

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

$$\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{1}{25} + \frac{0,75}{36}} = 0,2466$$

$$Z = \frac{6,2 - 4,5}{0,2466} = 6,89^{**}$$

$z > z_{0,01} > z_{0,05}$ $\Rightarrow H_0$ se odbacuje (razlika prosečnih prinosa ove dve sorte je visoko značajna).

II način testiranja

Testiranje primenom $(1-\alpha)\%$ intervala poverenja za razliku aritmetičkih sredina osnovnih skupova:

$$\overline{X}_1 - \overline{X}_2 - z_\alpha \cdot \sigma_{\overline{X}_1-\overline{X}_2} < \mu_1 - \mu_2 < \overline{X}_1 - \overline{X}_2 + z_\alpha \cdot \sigma_{\overline{X}_1-\overline{X}_2}$$

95% interval poverenja je:

$$6,2 - 4,5 - 1,96 \cdot 0,2466 < \mu_1 - \mu_2 < 6,2 - 4,5 + 1,96 \cdot 0,2466$$

$$1,22 < \mu_1 - \mu_2 < 2,18 \text{(t/ha)}$$

$0 \notin (1,22, 2,18) \Rightarrow$ H_o se odbacuje na pragu značajnosti 5%.

99% interval poverenja je:

$$6,2 - 4,5 - 2,58 \cdot 0,2466 < \mu_1 - \mu_2 < 6,2 - 4,5 + 2,58 \cdot 0,2466$$

$$1,06 < \mu_1 - \mu_2 < 2,34 \text{(t/ha)}$$

$0 \notin (1,06, 2,34) \Rightarrow$ H_o se odbacuje na pragu značajnosti 1%.

2. Slučaj: σ_1^2 i σ_2^2 nisu poznate

Primer 2: U ogledu sa dve sorte šećerne repe dobijeni su sledeći prinosi po jedinici površine (vag/ha):

Sorta A		Sorta B					
Prinos X_1	Površina f_1	Prinos X_2	Površina f_2	f_1X_1	f_2X_2	$f_1X_1^2$	$f_2X_2^2$
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

Rešenje:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

$$\overline{X}_1 = \frac{\sum f_1 X_1}{\sum f_1} = \frac{99,70}{20} = 4,985 \text{(vag/ha)}$$

$$\overline{X}_2 = \frac{\sum f_2 X_2}{\sum f_2} = \frac{97,50}{18} = 5,417 \text{(vag/ha)}$$

$$s_{1+2}^2 = \frac{\sum f_1X_1^2 - \frac{(\sum f_1X_1)^2}{n_1} + \sum f_2X_2^2 - \frac{(\sum f_2X_2)^2}{n_2}}{n_1 + n_2 - 2}$$

$$s_{(\overline{X}_1-\overline{X}_2)} = \sqrt{s_{1+2}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$s_{1+2}^2 = \frac{501{,}35 - \frac{(99{,}7)^2}{20} + 533{,}37 - \frac{97{,}5^2}{18}}{20+18-2} =$$

$$= \frac{4{,}3455 + 5{,}245}{36} = 0{,}2664 (\text{vag/ha})$$

$$s_{(\overline{X}_1-\overline{X}_2)} = \sqrt{0{,}2664 \cdot \left(\frac{1}{20} + \frac{1}{18}\right)} = 0{,}1677$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_{(\overline{X}_1-\overline{X}_2)}} = \frac{4{,}985 - 5{,}417}{0{,}1667} = -2{,}59*$$

$\boxed{t_{0,05}(36) \approx t_{0,05}(35) = 2{,}03}$ $\boxed{t_{0,01}(36) \approx t_{0,01}(35) = 2{,}724}$

$|t| = 2{,}59 > 2{,}03 \Rightarrow H_o$ se odbacuje na pragu značajnosti 5%

$|t| = 2{,}59 < 2{,}724 \Rightarrow H_o$ se prihvata na pragu značajnosti 1%

Testiranje primenom intervala poverenja za razliku aritmetičkih sredina osnovnih skupova:

(4,985 -5,417) -2,03 · 0,1667 < μ_1- μ_2< (4,985 -5,417)+2,03 · 0,1667

-0,7704< μ_1- μ_2<-0,0936 (vag/ha)

(4,985 -5,417) -2,724 · 0,1667 < μ_1- μ_2< (4,985 -5,417)+2,724 · 0,1667

-0,8861 < μ_1- μ_2<0,0221 (vag/ha)

$0 \notin (-0{,}7704,\ -0{,}0936) \Rightarrow H_o$ se odbacuje ($\alpha = 0{,}05$).

$0 \in (-0{,}8861,\ 0{,}0221) \Rightarrow H_o$ se prihvata ($\alpha = 0{,}01$).

Zaključivanje o razlici proporcija dva osnovna skupa na osnovu velikih i nezavisnih prostih slučajnih uzoraka

Pretpostavka za primenu testa: $n_1p_1>5$, $n_1q_1>5$, $n_2p_2>5$ i $n_2q_2>5$.

Primer 1: U uzorcima od po 100 biljaka, broj izniklih biljaka je 40 i 60 u zavisnosti od toga da li je seme tretirano ili nije. Testirati nultu hipotezu da se proporcije izniklih biljaka statistički značajno ne razlikuje.

Rešenje:

$n_1 = n_2 = 100$

$$\hat{p}_1 = \frac{40}{100} = 0{,}4 \qquad \hat{p}_2 = \frac{60}{100} = 0{,}6$$

$$\boxed{\begin{array}{l} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{array}}$$

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

$$\bar{p} = \frac{a_1 + a_2}{n_1 + n_2} = \frac{100}{200} = 0{,}50 \qquad \bar{q} = 1 - \bar{p} = 0{,}50$$

$$s_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{0{,}50 \cdot 0{,}50 \cdot \left(\frac{1}{100} + \frac{1}{100}\right)} = \sqrt{0{,}25 \cdot \frac{2}{100}} = 0{,}07071$$

$$Z = \frac{0{,}4 - 0{,}6}{0{,}07071} = -2{,}83$$

$z_{0,05} = 1{,}96 \qquad |z| > z_{0,05} \Rightarrow \qquad H_o$ se odbacuje $(\alpha = 0{,}05)$

$z_{0,01} = 2{,}58 \qquad |z| > z_{0,01} \Rightarrow \qquad H_o$ se odbacuje $(\alpha = 0{,}01)$

Proporcije izniklih biljaka se statistički visoko značajno razlikuju u zavisnosti od toga da li je seme tretirano ili nije.

II način

Testiranje na osnovu intervala poverenja za razliku proporcija osnovnih skupova:

$$\hat{p}_1 - \hat{p}_2 - z_\alpha \cdot s_{(\hat{p}_1 - \hat{p}_2)} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + z_\alpha \cdot s_{(\hat{p}_1 - \hat{p}_2)}$$

$\alpha = 0{,}05$

$$0{,}4 - 0{,}6 - 1{,}96 \cdot 0{,}07071 < p_1 - p_2 < 0{,}4 - 0{,}6 + 1{,}96 \cdot 0{,}07071$$

$$-0{,}3386 < p_1 - p_2 < -0{,}06141$$

$0 \notin (-0{,}3386, -0{,}06141) \Rightarrow H_o$ se odbacuje ($\alpha = 0{,}05$).

$\alpha = 0{,}01$

(0,4-0,6 -2,58.0,07071< p_1- p_2<(0,4 -0,6)+2,58.0,07071

$$-0{,}3824 < p_1 - p_2 < -0{,}01757$$

$0 \notin (-0{,}3824, -0{,}01757) \Rightarrow H_o$ se odbacuje ($\alpha = 0{,}01$).