Predavanje 4

TESTIRANJE PARAMETARSKIH HIPOTEZA

Formulisanje hipoteze

Primer:

Proizvođač baterija tvrdi da njegove punjive baterije imaju <u>prosečan vek trajanja od najmanje 1000 punjenja.</u>

$\mu \geq 1000$

H_0: $\mu \geq 1000$

Nulta hipoteza je izraz koji sadrži jednakost (=, ≤, ≥).

H_1: $\mu < 1000$

Radna hipoteza je komplement nulte.

Formulisanje hipoteze

Primer:
Određeni fakultet tvrdi da se 94% diplomiranih studenata zaposli najkasnije 6 meseci od diplomiranja.

$p = 0.94$

H_0: $p = 0.94$

H_1: $p \neq 0.94$

Vrste grešaka

Bez obzira koju hipotezu u stvari želimo da potvrdimo (H_0 ili H_1), prilikom testiranja statističkih hipoteza uvek počinjemo uz **pretpostavku da je H_0 tačna.**

Na kraju testa, donosimo jednu od sledeće dve odluke:

1. Odbacujemo nultu hipotezu, ili
2. Ne odbacujemo nultu hipotezu.

Greška I vrste se dešava ako na osnovu testa odbacimo H_0, a ona je zapravo tačna.

Greška II vrste se dešava ako na osnovu testa ne odbacimo H_0, a ona je zapravo netačna.

Nivo značajnosti

Prilikom testiranja hipoteza, **nivo značajnosti (α)** je najveća dozvoljena verovatnoća greške prve vrste.

Najčešće korišćeni nivoi značajnosti:

$\alpha = 0.10$ $\alpha = 0.05$ $\alpha = 0.01$

Na primer, ako je $\alpha = 0.05$, to znači da će u našem testu verovatnoća da odbacimo tačnu nultu hipotezu biti najviše 5%.

Statistički test

Nakon formulisanja nulte i radne hipoteze, i zadavanja nivoa značajnosti, računamo **test statistiku** na osnovu realizovanog uzorka.

Cilj test statistike je da utvrdi koliko se odgovarajuća statistika uzorka razlikuje od pretpostavljene vrednosti parametra populacije.

Parametar populacije	Statistika uzorka	Standardizovana test statistika
μ	$\overline{X}$	$z \ (n \geq 30)$ $t \ (n < 30)$
p	$\hat{p}$	z
σ^2	s^2	X^2

P-vrednosti

Ako je nulta hipoteza tačna, ***p*-vrednost** testa je verovatnoća dobijanja vrednosti ekstremnije od one dobijene iz uzorka. Ovo posredno odgovara verovatnoći da uzorak podržava nultu hipotezu.

P-vrednost testa zavisi od vrste testa, tj. od vrste hipoteza.

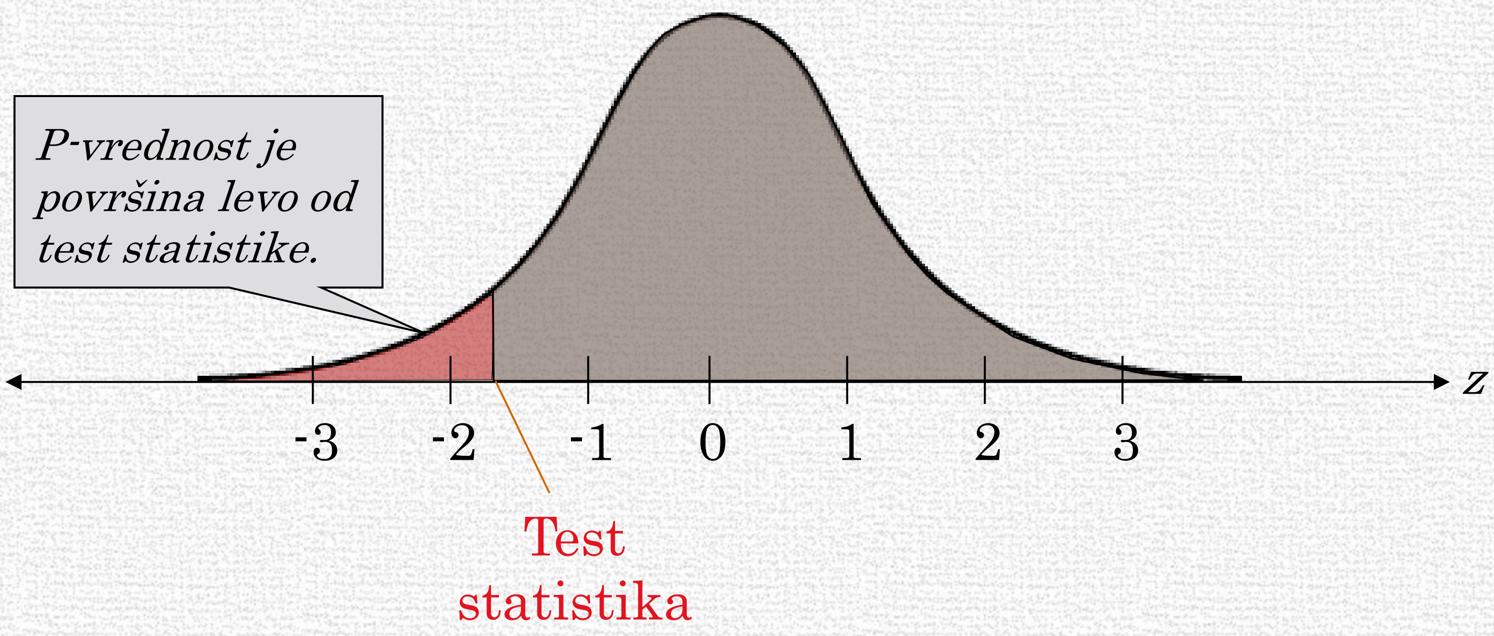
Postoje tri vrste testova – levi, desni, i dvostrani test. Izbor vrste zavisi od formulacije nulte i alternativne hipoteze.

Levi (jednostrani) test

1. Ukoliko alternativna hipoteza sadrži nejednakost ”manje od” (<), reč je o **levom testu**.

H_0: $\mu \geq m$

H_1: $\mu < m$

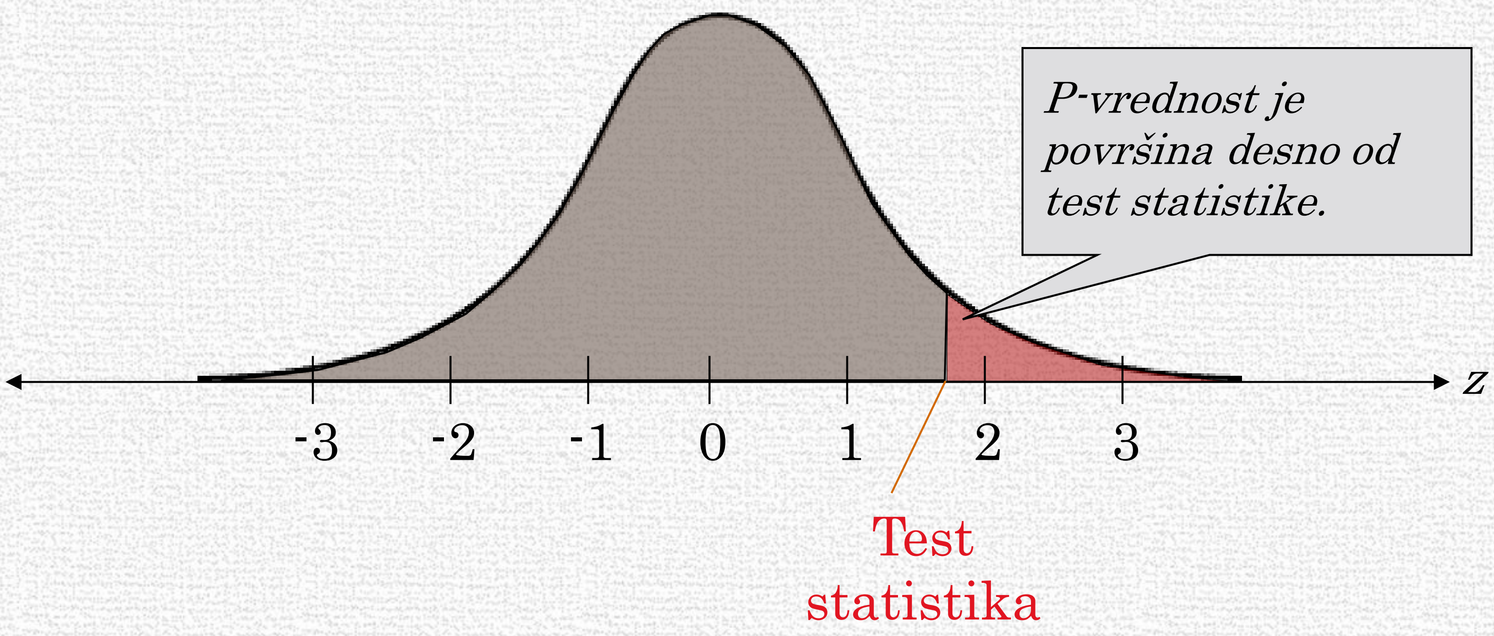


Desni (jednostrani) test

2. Ukoliko alternativna hipoteza sadrži simbol "veće od" (>), reč je o **desnom testu**.

H_0: $\mu \leq m$

H_1: $\mu > m$

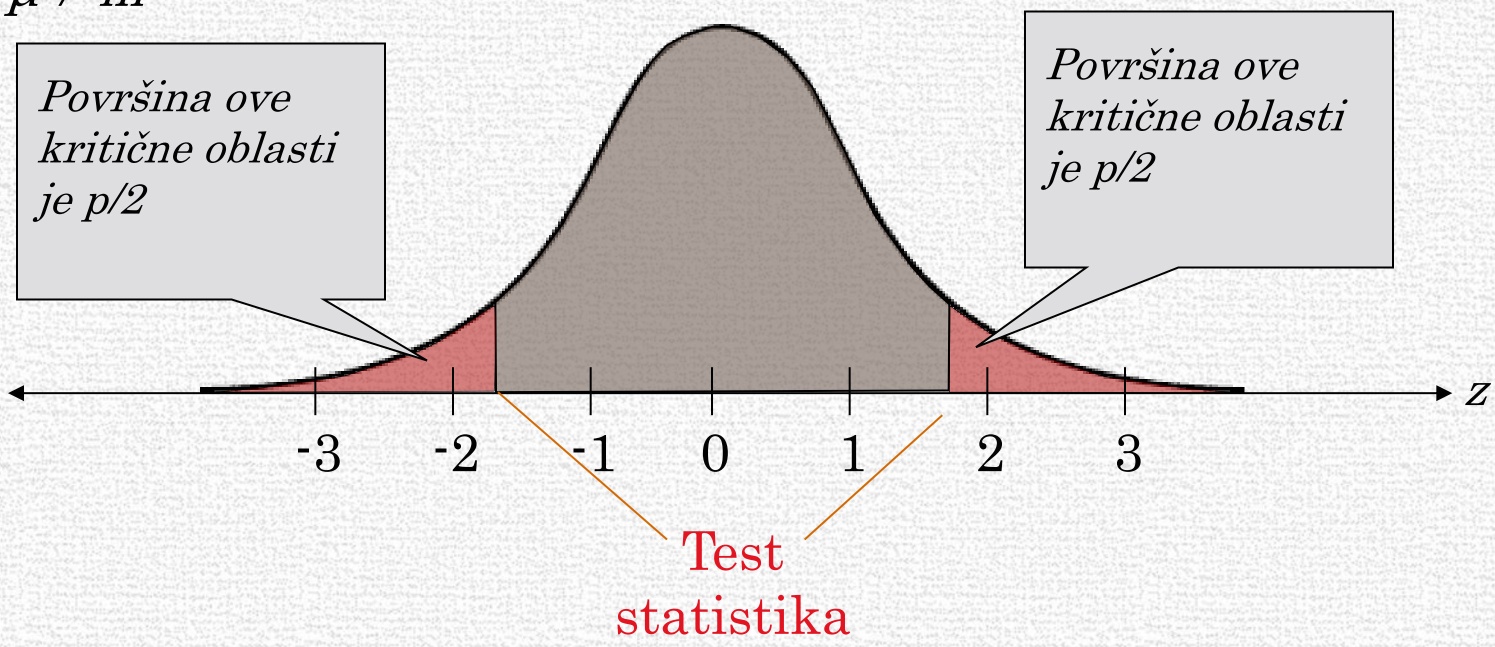


Dvostrani test

Ukoliko alternativna hipoteza sadrži simbol nejednakosti (≠), reč je o **dvostranom testu**. U tom slučaju, oba repa imaju površinu *p/2*.

H_0: $\mu = m$

H_1: $\mu \neq m$



Određivanje tipa testa

Primer:

Za svaku tvrdnju, formuliši H_0 i H_1, a zatim odredi tip testa koji treba koristiti u datom slučaju.

a.) Proizvođač cigareta tvrdi da manje od jedne osmine odraslih građana puši.

H_0: $p \geq 0.125$

H_1: $p < 0.125$ → levi test

b.) Operator mobilne telefonije tvrdi da je prosečno vreme razgovora 8 minuta.

H_0: $\mu = 8$

H_1: $\mu \neq 8$ → dvostrani test

Odlučivanje

Odlučivanje bazirano na p-vrednosti

1. Ako je $P \leq \alpha$, odbacujemo H_0.
2. Ako je $P > \alpha$, ne odbacujemo H_0.

	tvrdnja	
Odluka	**Tvrdnja je H_0**	**Tvrdnja je H_1**
odbaci H_0	Ima dovoljno dokaza da odba cimo tvrdnju.	Ima dovoljno dokaza da potv rdimo tvrdnju.
Ne odbacuj H_0	Nema dovoljno dokaza da od bacimo tvrdnju.	Nema dovoljno dokaza da pot vrdimo tvrdnju.

Odlučivanje

Primer:

Na osnovu testa je izračunata p-vrednost $P = 0.0256$.
Doneti odluku ako je prag značajnosti α

a.) 0.05,

b.) 0.01.

a.) Pošto je $0.0256 < 0.05$, odbacujemo nultu hipotezu sa pragom značajnosti 0.05.

b.) Pošto je $0.0256 > 0.01$, ne odbacujemo nultu hipotezu sa pragom značajnosti 0.01.

Pravilo 5%

U praksi se najčešće koristi takozvano pravilo 5%:

- Ako je **p-vrednost ≤ 0.05**, odbacujemo nultu hipotezu.
- Ako je p-vrednost > 0.05, ne odbacujemo nultu hipotezu.

Ovo pravilo potiče od toga što se smatralo da je verovatnoća 5% da napravimo grešku I vrste zlatna sredina, koja ne smanjuje previše moć testa.

U pojedinim oblastima nauke se koriste i druge granične p-vrednosti.

Interpretacija odluke

Primer:

H_1: (tvrdnja) Proizvođač cigareta tvrdi da manje od jedne osmine odraslih građana puši.

Ukoliko odbacimo H_0 ($p \leq 0.05$) zaključujemo da "postoji dovoljno dokaza da potvrdimo tvrdnju proizvođača."

Ukoliko ne odbacimo H_0 ($p > 0.05$) zaključujemo da "nema dovoljno dokaza da potvrdimo tvrdnju proizvođača."

Z-Test

Z-test za srednju vrednost je parametarski test kojim se proverava pretpostavka o srednjoj vrednosti populacije. Koristi se kada je populacija normalno distribuirana i σ *(varijansa populacije)* je poznata, ili kada je obim uzorka n bar 30.

Test statistika:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Standardna greška s.e.

Kada je $n \geq 30$, uzoračka standardna devijacija s se koristi u slučaju nepoznate varijanse populacije.

Primer z-testa

Primer:
Proizvođač tvrdi da je prosečan vek trajanja njegovih punjivih baterija više od 1000 punjenja. Slučajnim uzorkovanjem 100 baterija izračunata je AS 1002 punjenja i SD 14 punjenja. Da li sa pragom značajnosti $\alpha = 0.01$ možemo podržati tvrdnju proizvođača?

H_0: $\mu \leq 1000$ H$_1$: $\mu > 1000$ (tvrdnja)

Prag značajnosti $\alpha = 0.01$.

Standardizovana test statistika je

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{1002 - 1000}{14 / \sqrt{100}} \approx 1.43$$

Nastavak primera:

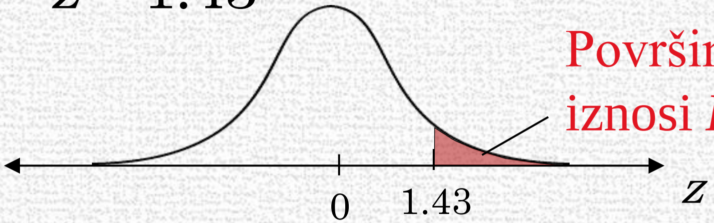
H_0: $\mu \leq 1000$

H_1: $\mu > 1000$ (tvrdnja)

Kako alternativna hipoteza sadrži znak >,

radi se o desnom jednostranom testu.

$z = 1.43$



Površina desno od $z = 1.43$ iznosi $P = 0.0764$.

P-vrednost je veća od α (0.076 > 0.01) pa ne odbacujemo H_0.

Zaključak: sa pragom značajnosti 1% zaključujemo da nema dovoljno dokaza da podrži tvrdnju proizvođača.

T -Test

T-test za srednju vrednost populacije je modifikacija z-testa koja se koristi kada je populacija skoro normalna, σ nepoznato, a obim uzorka $n < 30$.

Test statistika:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Test statistika ima Studentovu t-raspodelu sa $n - 1$ stepena slobode.

T - test

Primer:
Mobilni operater tvrdi da je prosečna dužina telefonskog razgovora 8 minuta. Na osnovu uzorka od 18 razgovora izračunata je AS 7.8 minuta i SD 0.5 minuta. Testirati tvrdnju operatera sa pragom značajnosti $\alpha = 0.05$.

H_0: $\mu = 8$ (tvrdnja) H_1: $\mu \neq 8$

Prag značajnosti $\alpha = 0.05$.

Obim uzorka je manji od 30 i varijansa je nepoznata,

pa koristimo t-test.

Broj stepeni slobode d.f. = 18 – 1 = 17.

Nastavak primera:

H_0: $\mu = 8$ (tvrdnja)

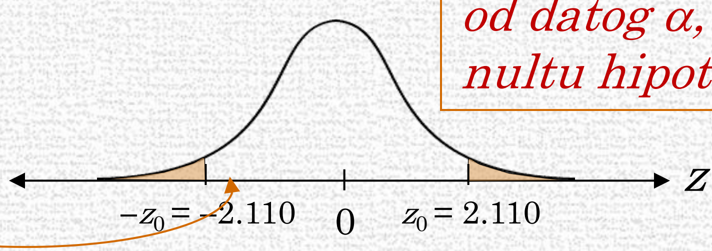
H_1: $\mu \neq 8$

Koristimo dvostrani test (radna hipoteza sadrži ≠).

Test statistika:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7.8 - 8}{0.5/\sqrt{18}}$$

$$\approx -1.70.$$



P-vrednost je 0.108, što je veće od datog α, pa ne odbacujemo nultu hipotezu.

Sa pragom značajnosti 5% zaključujemo da nema dovoljno dokaza da odbacimo tvrdnju operatera.

Z -Test za proporciju populacije

Z-test za proporciju populacije je statistički test za proveru pretpostavke o vrednosti parametra populacije p. Koristi se kod binomne raspodele kada su zadovoljeni uslovi $np \geq 5$ i $nq \geq 5$.

Test statistika:

$$z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

Standardna greška s.e.

Z-test za proporciju

Primer:
Neki fakultet tvrdi da više se od 94% diplomaca zaposli u periodu od šest meseci nakon diplomiranja. Anketiranjem 500 slučajno odabranih diplomaca utvrđeno je da se 475 zaposlilo u tom periodu. Ima li dovoljno dokaza za potvrdu tvrdnje fakulteta sa pragom značajnosti 1%?

Provera uslova:

$$np = 500*0.94 = 470,\ nq = 500*0.06 = 30$$

H_0: $p \leq 0.94$ H$_1$: $p > 0.94$ (tvrdnja)

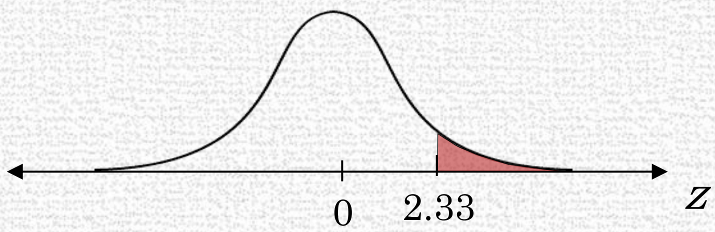
Nastavak primera:

Koristimo desni jednostrani test (radna hipoteza sadrži >).

$\alpha = 0.01$

Test statistika:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.95 - 0.94}{\sqrt{(0.94)(0.06)/500}} \approx 0.94$$



Odgovarajuća p-vrednost je 0.173, što je veće od α.

Sa pragom značajnosti 1% zaključujemo da nema dovoljno dokaza da bi potvrdili tvrdnju fakulteta.

χ^2 - Test

χ^2-test za varijansu ili standardnu devijaciju populacije se koristi se koristi za proveru tvrdnji o ovim parametrima populacije, i koristi kada je populacija normalno distribuirana.

Test statistika:

$$X^2 = \frac{(n-1)s^2}{\sigma^2}$$

Ona ima Pirsonovu hi-kvadrat raspodelu $n - 1$ stepenom slobode.

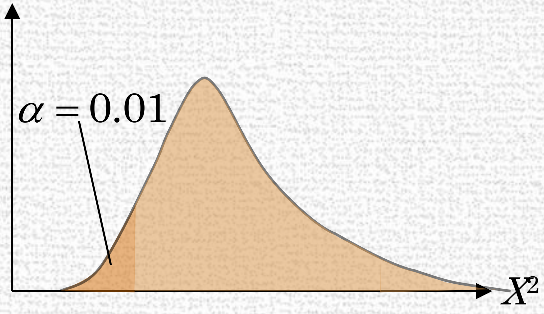
χ^2 - Test

Primer:

Jedan profesor tvrdi da je standardna devijacija bodova na testu iz statistike manja od 30. 10 testova je nasumično uzeto, i izračunata je SD 28.8. Proveriti profesorovu tvrdnju sa pragom $\alpha = 0.01$.

H_0: $\sigma \geq 30$ H$_1$: $\sigma < 30$ (tvrdnja)

U pitanju je levi jednostrani test, d.f.= 9 i $\alpha = 0.01$.

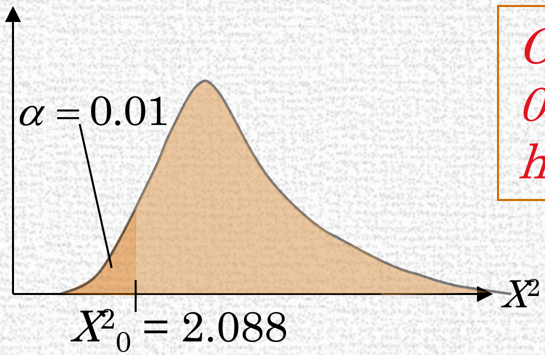


Nastavak primera:

Test statistika:

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)(28.8)^2}{30^2} \approx 8.29$$

$X^2_0 = 2.088$



Odgovarajuća p-vrednost je 0.495, pa ne odbacujemo nultu hipotezu.

Sa pragom značajnosti 1% nema dovoljno dokaza za potvrđivanje profesorove tvrdnje.

Rezime

Parametarski testovi jednog uzorka prema **vrsti parametra** na koji se odnose mogu biti:

1. Z-test za srednju vrednost populacije
2. T – test za srednju vrednost populacije
3. Z-test za proporciju populacije
4. Hi-kvadrat test za varijansu ili SD populacije

Prema **vrsti testa** mogu biti:

1. dvostrani (H_1 sadrži $\neq$)
2. levi jednostrani (H_1 sadrži $<$)
3. desni jednostrani (H_1 sadrži $>$)

Odlučivanje na osnovu **p-vrednosti**:

1. Ako je $p \leq \alpha$ odbacujemo H_0 sa pragom značajnosti α
2. Ako je $p > \alpha$ ne odbacujemo H_0 sa pragom značajnosti α

Tvrdnja koju proveravamo može biti sadržana ili u H_0 ili u H_1.

Parametarski testovi više uzoraka

- Dosadašnji testovi su se odnosili na jedan uzorak, na osnovu kog smo proveravali da li vrednost parametra populacije odgovara tvrdnji.
- Tvrdnje mogu da se odnose i na poređenje vrednosti parametara dve ili više populacija.

Parametarski testovi dva uzorka

$H_0(\theta_1 = \theta_2)$

Analiza varijanse ANOVA

$H_0(\theta_1 = \theta_2 = ... = \theta_n)$

Parametarski testovi dva uzorka

- Cilj je uporediti vrednost istog parametra kod dve grupe
- θ_1 vrednost parametra u grupi 1, θ_2 u grupi 2
- Mogu biti dvostrani i jednostrani

$H_0(\theta_1 = \theta_2)$

$H_1(\theta_1 \neq \theta_2)$

levi

$H_0(\theta_1 \geq \theta_2)$

$H_1(\theta_1 < \theta_2)$

$H_0(\theta_1 \leq \theta_2)$

$H_1(\theta_1 > \theta_2)$

desni

Z -test dva uzorka

Z-test služi za ispitivanje razlike između dve srednje vrednosti populacije μ_1 i μ_2.

Tri uslova za primenu:

1. Uzorci treba da su **nasumično odabrani**.
2. Uzorci treba da budu **nezavisni**, što znači da elementi uzorka iz prve populacije nisu povezani sa elementima uzorka iz druge populacije.
3. **Oba obima** uzoraka treba da su **bar 30**, ili, ako to nije slučaj, obe populacije treba da budu normalno distribuirane sa poznatom varijansom.

z-test dva nezavisna uzorka

Standardizovana test statistika je:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} \quad \text{where} \quad \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

- Kada su uzorci dovoljno veliki, možemo koristiti uzoračke SD s_1 i s_2 umesto σ_1 i σ_2.
- Statistika z ima Gausovu raspodelu.
- Postupak odlučivanja na osnovu p-vrednosti isti kao kod testova jednog uzorka.

Z-Test primer

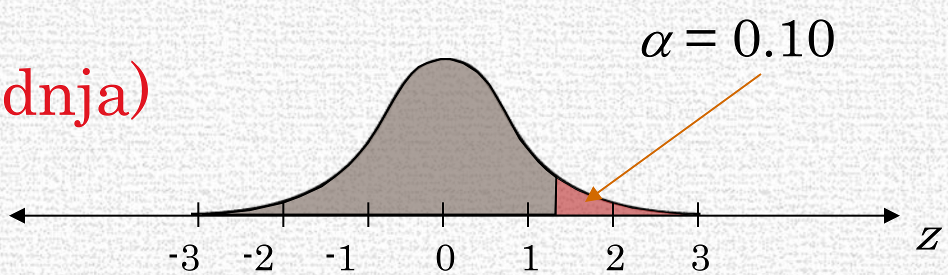
Primer:

Nastavnica u srednjoj školi tvrdi da njeni učenici postižu bolji uspeh od koleginih. Uzorak njenih učenika: 49 učenika, AS 22.1 bodova i SD 4.8. Uzorak koleginih učenika: 44 učenika, AS 19.8 bodova i SD 5.4. Za $\alpha = 0.1$, da li možemo potvrditi nastavnicinu tvrdnju?

Desni jednostrani test.

H_0: $\mu_1 \leq \mu_2$

H_a: $\mu_1 > \mu_2$ (tvrdnja)



$z_0 = 1.28$

Nastavak primera:

H_0: $\mu_1 \leq \mu_2$

H_1: $\mu_1 > \mu_2$ (tvrdnja)

Odgovarajuća p-vrednost je 0.012, $p < \alpha$, *pa odbacujemo* H_0.

Standardna greška je:

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4.8^2}{49} + \frac{5.4^2}{44}} \approx 1.0644.$$

Standardizovana test statistika:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x} - \bar{x}}} = \frac{(22.1 - 19.8) - 0}{1.0644} \approx 2.161$$

Zaključak: sa pragom značajnosti 10% zaključujemo da ima dovoljno dokaza da podržimo nastavnicinu tvrdnju.

T-test dva uzorka

Ukoliko su obimi uzoraka manji od 30, umesto z-testa koristimo ***t*-test** za ispitivanje razlike srednjih vrednosti populacije μ_1 i μ_2.

Tri uslova:

1. Nasumično odabrani uzorci.
2. Nezavisni uzorci
3. Normalno distribuirane populacije.

T –Test dva uzorka

Standardizovana test statistika:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}.$$

- Ukoliko su varijanse jednake koristimo kombinovanu ocenu

$$\hat{\sigma} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \qquad \sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- Ukoliko varijanse nisu jednake

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Z-test ili *t*-test?

Are both sample sizes at least 30?

Yes → Use the *z*-test.

No ↓

Are both populations normally distributed?

No → You cannot use the *z*-test or the *t*-test.

Yes ↓

Are both population standard deviations known?

Yes ↓ Use the *z*-test.

No → Are the population variances equal?

Yes → Use the *t*-test with

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

and d.f = $n_1 + n_2 - 2$.

No ↓ Use the *t*-test with

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

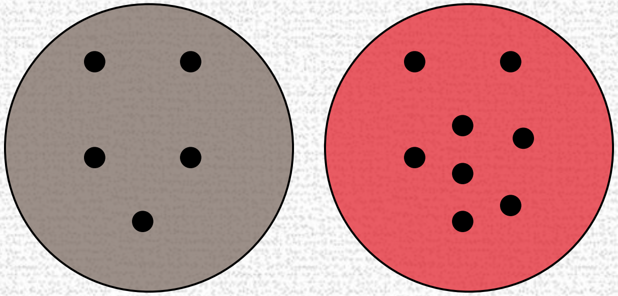
and d.f = smaller of $n_1 - 1$ or $n_2 - 1$.

U praksi se mnogo češće koristi t-test!

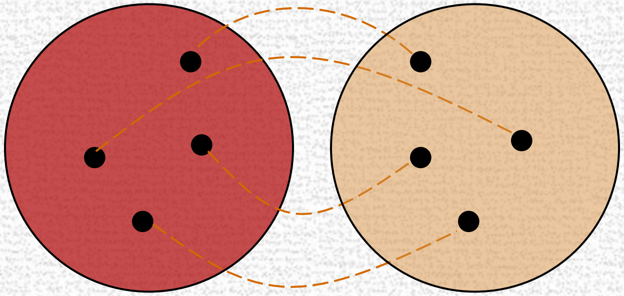
Nezavisni i zavisni uzorci

Dva uzorka su **nezavisna** ako uzorak iz jedne populacije nije povezan sa uzorkom iz druge populacije.

Dva uzorka su **zavisna (uparena)** ako svakom elementu uzorka iz jedne populacije odgovara jedan element uzorka iz druge populacije.



Nezavisni uzorci



Zavisni uzorci

Nezavisni i zavisni uzorci

Primer:
Da li su sledeći parovi uzoraka nezavisni ili spareni?

Uzorak 1: Težina 24 osobe pre dijete.
Uzorak 2: Težina iste 24 osobe nakon dijete.

spareni

Uzorak 1: Cena 150 polovnih automobila u BG
Uzorak 2: Cena 100 polovnih automobila u NS

nezavisni

t-test za sparene uzorke

Test statistika se zasniva na razlici obeležja kod svakog para:

$$d = x_1 - x_2$$

Razlika unosa za par podataka.

Potom se određuje prosečna vrednost $\bar{d}$ ovih razlika.

$$\bar{d} = \frac{\sum d}{n}.$$

AS razlika obeležja kod zavisnih uzoraka.

Nulta hipoteza je oblika H_0: $\mu_{\mathrm{d}} = 0$ (ili $\mu_{\mathrm{d}} \leq 0$ ili $\mu_{\mathrm{d}} \geq 0$) ako nas zanima samo postojanje razlike

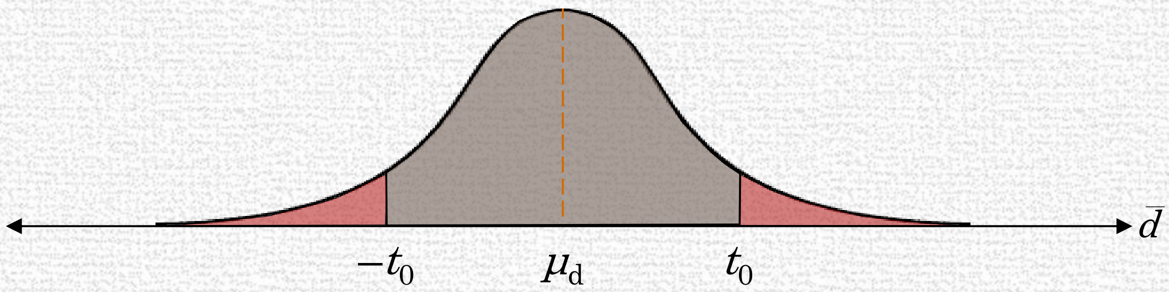
Umesto 0 može stajati neko c, ako nas zanima da li se parovi razlikuju za (tačno, bar, najviše) c.

t-test za zavisne uzorke

Tri uslova:

1. Nasumično odabrani uzorak.
2. Zavisni (upareni) uzorci.
3. Normalno distribuirane populacije.

Ukoliko uslovi važe, onda uz. raspodela prosečnih razlika za $\bar{d}$ ima Studentovu raspodelu sa $n - 1$ d.f., gde je n broj parova podataka.



z-test za proporciju dva uzorka

Koristi se za testiranje razlike proporcija kod dve populacije, p_1 and p_2. Proporcije se odnose na zastupljenost istog svojstva u te dve populacije.

Tri uslova:

1. Nasumično odabrani uzorci
2. Nezavisni uzorci
3. Dovoljno velik obim da važi

$$n_1 p_1 \geq 5, \quad n_1 q_1 \geq 5,$$

$$n_2 p_2 \geq 5, \quad n_2 q_2 \geq 5.$$

z-test za razliku proporcija

Standardizovana test statistika je

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Gde su $\overline{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ and $\overline{q} = 1 - \overline{p}.$

- Ukoliko su uslovi ispunjeni, z ima Gausovu raspodelu
- Zaključivanje na osnovu p-vrednosti isto kao kod svih.

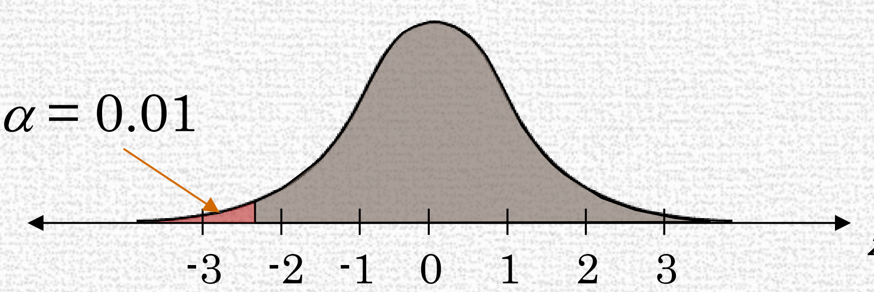
Z-test za razliku proporcija

Primer:

Cilj istraživanja je uporediti pušačke navike u zavisnosti od pola studenata. Od anketiranih1245 M studenata, 361 je izjavilo da puše više od pakle dnevno. Od 1065 Ž studenata, 341 puši više od pakle na dan. Za $\alpha = 0.01$, proveriti tvrdnju da je udeo M pušača manji od udela Ž pušača.

H_0: $p_1 \geq p_2$

H_1: $p_1 < p_2$ (tvrdnja)

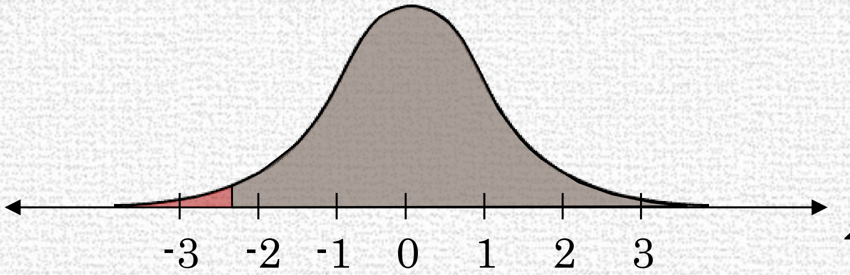


$-z_0 = -2.33$

Nastavak primera:

H_0: $p_1 \geq p_2$

H_1: $p_1 < p_2$ (tvrdnja)



$-z_0 = -2.33$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{361 + 341}{1245 + 1065} = \frac{702}{2310} \approx 0.304$$

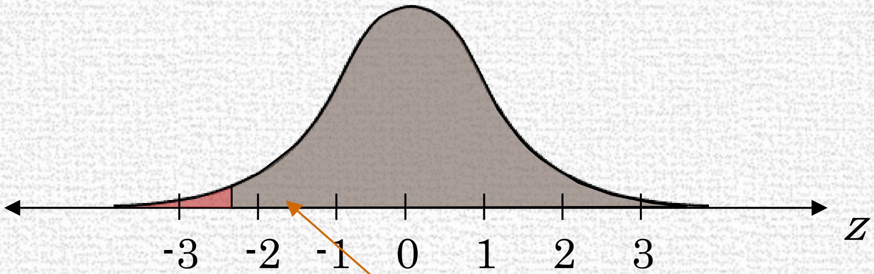
$$\bar{q} = 1 - \bar{p} = 1 - 0.304 = 0.696$$

Svi proizvodi 1245*0.304, 1245*0.696, 1065*0.304, i 1065*0.696 su bar 5, pa možemo koristiti *z*-test.

nastavak:

H_0: $p_1 \geq p_2$

H_1: $p_1 < p_2$ (tvrdnja)



$-z_0 = -2.33$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.29 - 0.32) - 0}{\sqrt{(0.304)(0.696)\left(\frac{1}{1245} + \frac{1}{1065}\right)}} \approx -1.56$$

Izračunata p-vrednost je 0,058 pa ne odbacujemo H_0.

Zaključak: sa pragom značajnosti 1% zaključujemo da nema dovoljno dokaza da je proporcija M pušača koji puše više od pakle dnevno niža od proporcije takvih Ž pušača.

Napomena: da je prag značajnosti bio 10%, odbacili bi H_0.