

Tipovi grešaka u statističkom zaključivanju

Moć testa



PREDAVANJE BR.3

Greške u zaključivanju



- Nulta hipoteza, kao tvrđenje o parametru osnovnog skupa, je ili istinita ili neistinita.
- Informacija iz uzorka je ili saglasna sa H_0 ili joj protivrječi.

Odluka	Stvarno stanje u populaciji	
	H_0 istinita	H_0 neistinita
Odbacujemo H_0	Greška I vrste	Nema greške
Ne odbacujemo H_0	Nema greške	Greška II vrste

Greške u zaključivanju



- Prilikom testiranja hipoteze nastojimo da odredimo vjerovatnoće javljanja navedenih grešaka i da ih, ako je moguće, smanjimo.

- Rizik greške I vrste ili nivo značajnosti testa

$$\alpha = P(H_0 \text{ odbacujemo} \mid H_0 \text{ je istinita})$$

- Rizik greške II vrste

$$\beta = P(H_0 \text{ ne odbacujemo} \mid H_0 \text{ nije istinita})$$

Greške u zaključivanju

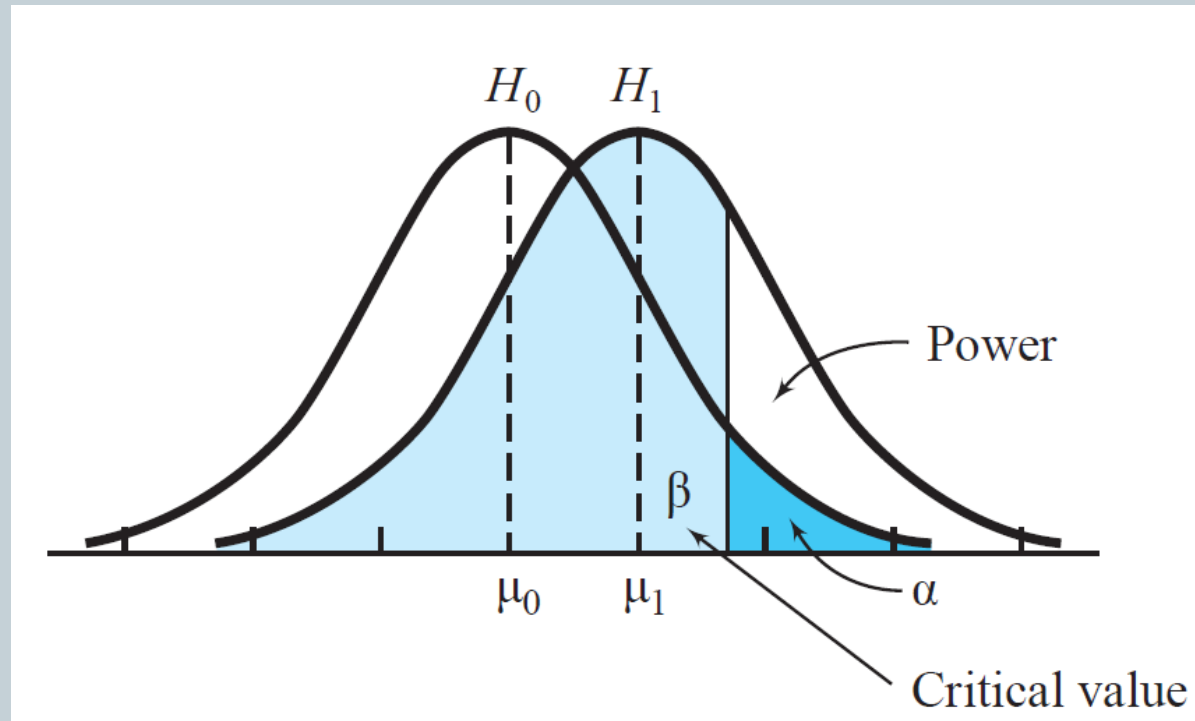


- Vrijednost rizika β je određena sljedećim faktorima:
 - 1) Stvarnom vrijednošću testiranog parametra
 - 2) Nivoom značajnosti testa
 - 3) Veličinom uzorka
 - 4) Oblikom testa (jednosmjernim ili dvosmjernim)
- Jačina testa ($1 - \beta$) je vjerovatnoća odbacivanja netačne H_0
- Za dati obim uzorka, sa porastom rizika greške I vrste smanjuje se rizik greške II vrste i obrnuto. Pritom, rizici se ne dopunjuju do 1.

Jačina (snaga) testa



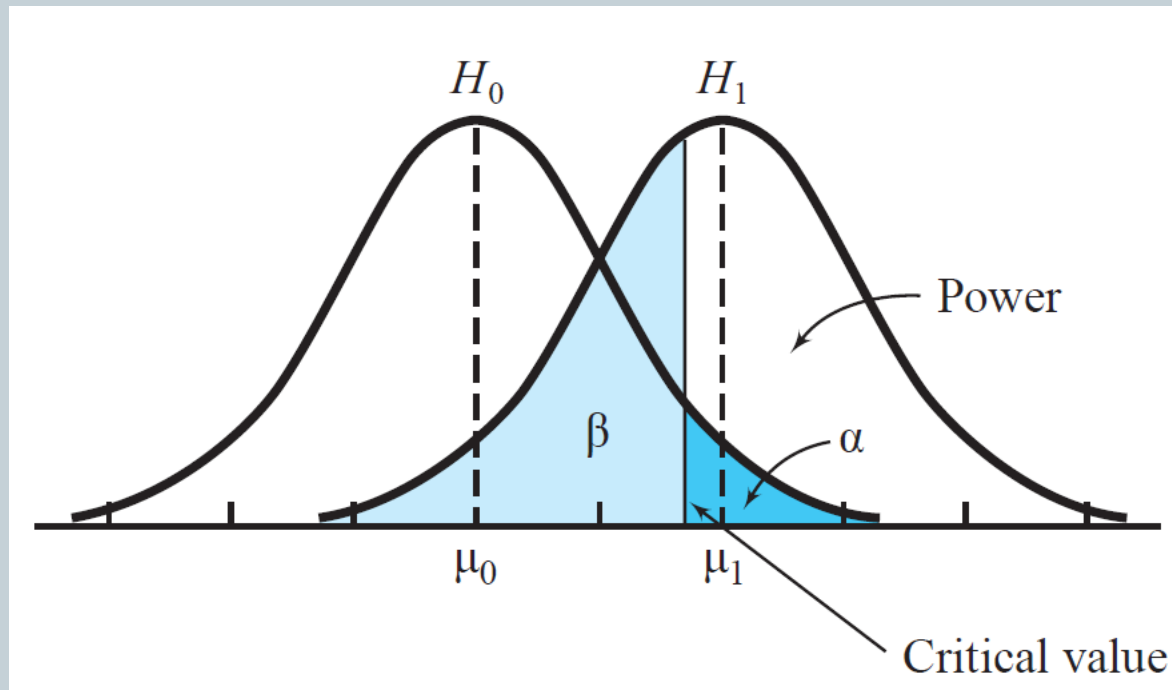
- Hipotetičke distribucije kad je nulta hipoteza tačna i kad je alternativna hipoteza tačna



Jačina (snaga) testa



- Povećanje snage povećanjem razlike između μ_0 i μ_1 .

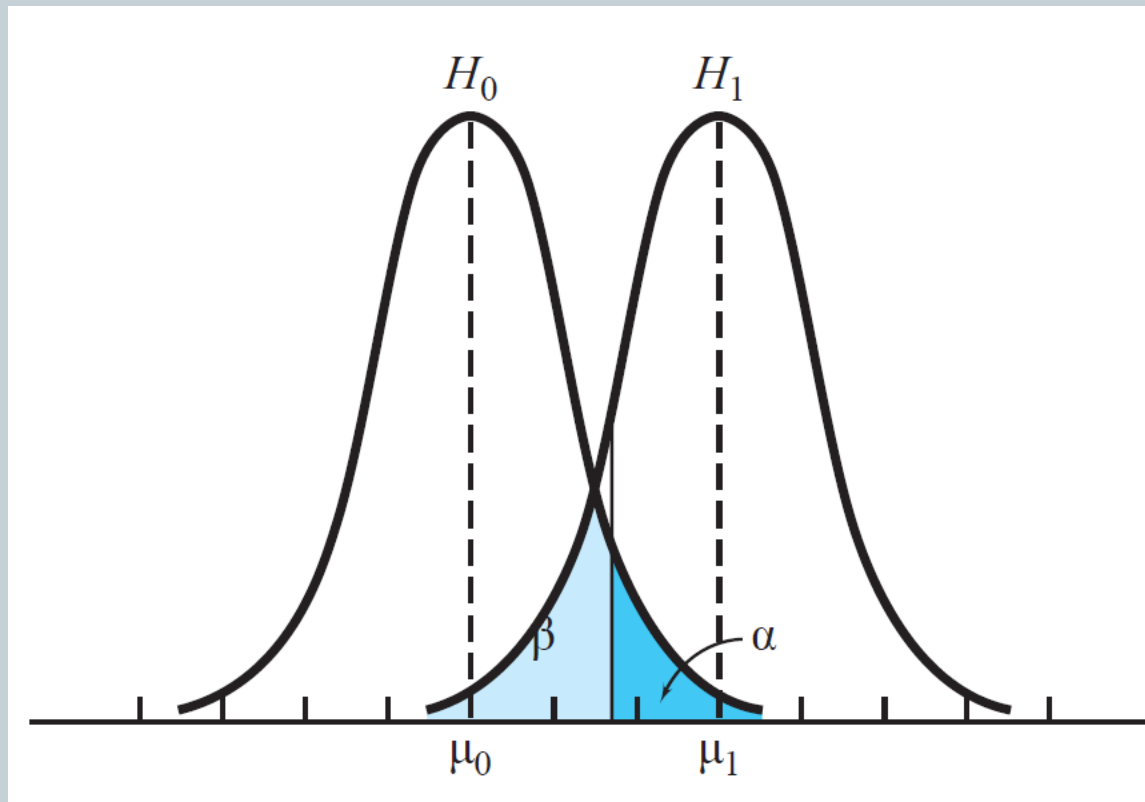


- Na veličinu razlike se može djelovati dobrim definisanjem varijabli i pažljivom kontrolom eksperimenta

Jačina (snaga) testa



- Povećanje snage smanjivanjem greške aritm. sredine



- Što je veći uzorak, veća je vjerovatnoća veće snage.

Određivanje jačine testa



- Različiti pristupi određivanja snage testa
- Cohen (1988) – minimalna snaga 0,8
- Prema Howell-u (2010), kod uzoraka jednake veličine

$$\delta = d \sqrt{\frac{n}{2}} = \frac{\mu_1 - \mu_2}{\sigma} \sqrt{\frac{n}{2}}$$

$$n = 2 \left(\frac{\delta}{d} \right)^2$$

Provjera znanja



Provjera znanja – teorijska pitanja



1. Variranje aritmetičkih sredina uzoraka oko prave aritmetičke sredine populacije mjeri

2. Greška I tipa nastaje kada:

- a) odbacimo nultu hipotezu koja je tačna
- b) ne odbacimo nultu hipotezu koja je tačna
- c) odbacimo nultu hipotezu koja je netačna
- d) ne odbacimo nultu hipotezu koja je netačna

3. Centralna granična teorema je važna u statistici zato što tvrdi da

- a) za veliki uzorak populacija ima približno normalan raspored
- b) za svaku populaciju, bez obzira na veličinu uzorka, raspored aritmetičkih sredina uzoraka je približno normalan
- c) za veliki uzorak, bez obzira na oblik rasporeda osnovnog skupa, raspored aritmetičkih sredina uzoraka teži normalnom rasporedu
- d) raspored aritmetičkih sredina iz uzoraka je približno normalan, za bilo koju veličinu uzorka

4. Nivo značajnosti testa, α je

Provjera znanja – teorijska pitanja



5. Ukoliko ljekar želi da na osnovu prikupljenih podataka zaključi da li neki lijek djeluje na prosječno povećanje pulsa za 7 udaraca u minuti, primijeniće

- a) jednosmjerni ili dvosmjerni test, sa ekvivalentnim rezultatom
- b) jednosmjerni test
- c) dvosmjerni test
- d) nijedan od navedenih testova

6. Standardna greška aritmetičke sredine je veća, što je:

- a) veći uzorak
- b) veća standardna devijacija
- c) manja standardna devijacija
- d) ništa od navedenog

7. Navesti dva preduslova za upotrebu t-testa.

_____.

8. Šta je standardna greška proporcije?

_____.

9. Relativna učestalost pojavljivanja nekog svojstva je:

- a) frekvencija
- b) proporcija
- c) standardna devijacija
- d) vjerovatnoća

10. Snaga statističkog testa je

_____.

11. Smanjenjem vjerovatnoće greške II vrste, snaga testa se _____.

Provjera znanja – zadaci



1. Jedan istraživač je želio ustanoviti da li droga ima uticaj na koordinaciju pokreta. 9 ispitanika eksperimentalne grupe i 10 ispitanika kontrolne grupe (koji su umjesto droge primili "placebo") postigli su na uređaju za mjerenje psihomotorike sljedeće rezultate (veći broj znači bolji rezultat):

E grupa	12	14	10	8	16	5	3	9	11	
K grupa	21	18	14	20	11	19	8	12	13	15

Izračunate su: $\sum x_1^2 = 996$ $\sum x_2^2 = 2445$ $F_{\alpha, n_1-1, n_2-1} = 4,36$ $t_{\frac{\alpha}{2}, n_1-1, n_2-1} = 2,11$

Sjetimo se... $s^2 = \frac{\sum x^2 - n\bar{x}^2}{(n-1)}$

Da li je razlika u prosječnom uspjehu tih grupa statistički značajna?

Provjera znanja – zadaci



$$M_1 = 9,78$$

$$M_2 = 15,10$$

$$S^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$$

$$S_E^2 = 16,895 \quad S_E = 4,11 \quad N_1 = 9$$

$$S_K^2 = 18,322 \quad S_K = 4,28 \quad N_2 = 10$$

$$F = \frac{\text{veća } SD^2}{\text{manja } SD^2} = 1,08446$$

$$F_{teor} = 4,36 = F_{\alpha, n_1-1, n_2-2} \rightarrow H_0 \rightarrow \text{iste var.}$$

$$Zaj. SD = \sqrt{\frac{SD_1^2(N_1-1) + SD_2^2(N_2-2)}{(N_1-1) + (N_2-2)}} = 4,2$$

$$SD_{M_1-M_2} = Zaj. SD \sqrt{\frac{N_1+N_2}{N_1*N_2}} = 1,9303$$

$$t = \frac{15,10 - 9,78}{1,9303} = 2,756$$

$$t_{\frac{\alpha}{2}, n_1-1+n_2-1} = 2,11 \rightarrow \text{odb. } H_0$$

Provjera znanja – zadaci



2. Od 300 studenata, 200 je položilo ispit iz Statistike, a od 500 studentkinja položilo ih je 300. Je li razlika u proporciji onih koji su položili statistički značajna?

Provjera znanja – zadaci



$$N_1 = 300$$

$$N_2 = 500$$

$$p_1 = \frac{200}{300} = 0,66 \quad p_2 = \frac{300}{500} = 0,6$$

$$t = \frac{0,66 - 0,6}{\sqrt{\frac{0,66 * 0,34}{300} + \frac{0,6 * 0,4}{500}}} = \frac{0,06666}{0,03504} = 1,9025$$

$$Z_{\frac{\alpha}{2}} = 1,96 \rightarrow H_0$$

Provjera znanja – zadaci



3. Dvije grupe učenika su izjednačene po predznanju iz psihologije. Zatim su obje grupe godinu dana učile psihologiju, svaka posebnom nastavnom metodom. Na kraju godine su svi učenici rješavali test znanja iz psihologije i dobijeni su sljedeći rezultati:

	I grupa	II grupa
N	60	60
AS	54,54	56,74
SD	6,34	7,25
r	0,50	

Da li se ove dvije grupe učenika značajno razlikuju po znanju iz psihologije?

Provjera znanja – zadaci



$$SD_{M_1-M_2} = \sqrt{SD_{M_1}^2 + SD_{M_2}^2 - 2rSD_{M_1} * SD_{M_2}} = 0,886211$$

$$SD_{M_1} = \frac{SD}{\sqrt{N}} = 0,82623$$

$$SD_{M_2} = 0,93597$$

$$t = \frac{M_1 - M_2}{SD_{M_1 - M_2}} = \frac{2,2}{0,886211} = 2,4824 \rightarrow H_1$$

Provjera znanja – zadaci



4. Od 64 djece, 32 tačno rješava jedan zadatak iz testa, a 36 tačno rješava drugi zadatak iz istog testa. Na bazi podataka iz donje tabele, odrediti da li je razlika između procenata tačnih odgovora na ova dva zadatka statistički značajna, tj. da li je jedan od ovih zadataka značajno teži?

		Zadatak 1	
		Netačno	Tačno
Zadatak 2	Tačno	10	26
	Netačno	22	6

Provjera znanja – zadaci



$$p_1 = \frac{32}{64} = 0,5 \quad p_2 = \frac{36}{64} = 0,5625$$

$$SD_{p_1-p_2} = \sqrt{\frac{d+a}{N}} = \sqrt{\frac{\frac{10}{64} + \frac{6}{64}}{64}} = 0,0625$$

$$t = \frac{0,0625}{0,0625} = 1 < 1,96 \rightarrow H_0$$