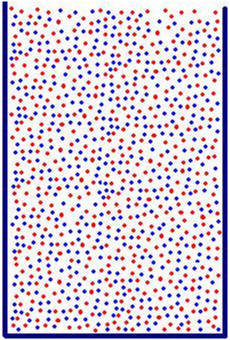
TESTIRANJE HIPOTEZA O SREDNJIM VRIJEDNOSTIMA

Testiranje hipoteze.

Problem. U kutiji se nalaze crvene i plave kuglice. Je li udio crvenih kuglica jednak udjelu plavih?



Je li ukupan broj crvenih kuglica jednak ukupnom broju plavih?

Je li udio crvenih kuglica jednak 0.5?

Rješenje. Slučajno izaberemo (izvučemo) određen broj kuglica i prebrojimo crvene i plave te izračunamo udio crvenih.

Cilj. Provjeriti da li su vrijednosti iz uzorka u skladu s pretpostavkom.

Idealno. Polovica kuglica u uzorku crvena a druga polovica plava.

STATISTIČKO ZAKLJUČIVANJE. Ukoliko je udio crvenih kuglica u uzorku blizu 0.5, zaključujemo da je udio crvenih kuglica u **kutiji** jednak 0.5. Ako je udio crvenih kuglica **daleko** od 0.5 zaključujemoda udio crvenih kuglica u **kutiji** nije jednak 0.5.

Kako definirati da je udio crvenih kuglica u uzorku blizu, tj. daleko od 0.5?

Ukoliko je udio crvenih kuglica u kutiji jednak 0.5 i slučajno smo izvukli 100 kuglica, izračunajte

- standardnu pogrešku procjene udjela crvenih kuglica;
- vjerojatnost da se procjena udjela crvenih kuglica nalazi u intervalu
$$\left[0.5 - 1.96 \cdot \frac{\sigma}{\sqrt{n}},\ 0.5 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Rješenje.

$$\begin{aligned}
p &= 0.5 \\
\sigma^2 &= p \cdot (1 - p) = 0.25 \\
\sigma &= \sqrt{\sigma^2} = \sqrt{0.25} = 0.5 \\
\frac{\sigma}{\sqrt{n}} &= \frac{0.5}{\sqrt{100}} = 0.05
\end{aligned}$$

$\hat{p}$ - udio crvenih kuglica u 100 izvučenih kuglica

Za veliki n ($n = 100$), $\hat{p}$ je približno distribuiran prema normalnoj distribuciji, $N(0.5, (\sigma/\sqrt{n})^2)$.

$$P\left(0.5 - 1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq \hat{p} \leq 0.5 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right) =$$

$$P\left(-1.96 \leq \frac{\hat{p} - p_0}{\sigma/\sqrt{n}} \leq 1.96\right)$$

Standardizirana varijabla

$$Y = \frac{\hat{p} - p_0}{\sigma/\sqrt{n}}$$

distribuirana je prema jediničnoj normalnoj distribuciji ($N(0, 1)$) pa je

$$P(-1.96 \leq Y \leq 1.96) = \mathbf{0.95}.$$

Vjerojatnost smo mogli izračunati i direktno.

$$\left[0.5 - 1.96 \cdot \frac{\sigma}{\sqrt{n}},\ 0.5 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right] = [0.402,\ 0.598]$$

$\hat{p}$ je normalna slučajna varijabla s $\mu = 0.5$ i standardnom devijacijom 0.05.

$$P(0.402 \leq \hat{p} \leq 0.598) = F(0.598) - F(0.402)$$

```
> pnorm(0.598,mean=0.5,sd=0.05)
0.9750021
> pnorm(0.402,mean=0.5,sd=0.05)
0.0249979
```

$$P(0.402 \leq \hat{p} \leq 0.598) = 0.975002 - 0.024998 = 0.950004$$

Ukoliko je udio crvenih kuglica u uzorku unutar intervala

$$\left[0.5 - 1.96 \cdot \frac{\sigma}{\sqrt{n}},\ 0.5 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right]. = [0.402,\ 0.598]$$

tada prihvaćamo pretpostavku da je udio crvenih kuglica u kutiji jednak 0.5.

U suprotnom (ako nije unutar intervala) pretpostavku ne prihvaćamo

Ako je udio crvenih kuglica u kutiji jednak 0.5 (tj. hipoteza je točna), tada je vjerojatnost da ćemo prihvatiti pretpostavku jednaka 0.95.

Vjerojatnost da nećemo prihvatiti točnu pretpostavku (tj. da ćemo pogriješiti) je 0.05.

Što ako pretpostavka nije točna?

Kolika je vjerojatnost da ćemo pogriješiti (prihvatiti krivu pretpostavku)?

Primjer. Kolika je vjerojatnost da ćemo prihvatiti pretpostavku (tj. da će udio crvenih kuglica u uzorku biti intervalu $[0.402,\ 0.598]$) ako je udio crvenih kuglica u kutiji jednak

a) 0.4;
b) 0.2?

Rješenje. a)

$$\begin{aligned} p &= 0.4 \\ \sigma^2 &= p \cdot (1-p) = 0.4 \cdot 0.6 = 0.24 \\ \sigma &= \sqrt{\sigma^2} = \sqrt{0.24} = 0.4899 \\ \frac{\sigma}{\sqrt{n}} &= \frac{0.4899}{\sqrt{100}} = 0.04899 \end{aligned}$$

$\hat{p}$ je normalna slučajna varijabla s $\mu = 0.4$ i standardnom devijacijom 0.04899.

$$P(0.402 \leq \hat{p} \leq 0.598) = F(0.598) - F(0.402)$$

```
> pnorm(0.598,mean=0.4,sd=0.04899)
0.9999735
> pnorm(0.402,mean=0.4,sd=0.04899)
0.5162822
```

$$P(0.402 \leq \hat{p} \leq 0.598) = 0.999973 - 0.516282 = 0.483691$$

b)

$$\begin{aligned} p &= 0.2 \\ \sigma^2 &= p \cdot (1-p) = 0.2 \cdot 0.8 = 0.16 \\ \sigma &= \sqrt{\sigma^2} = \sqrt{0.16} = 0.4 \\ \frac{\sigma}{\sqrt{n}} &= \frac{0.4}{\sqrt{100}} = 0.04 \end{aligned}$$

$\hat{p}$ je normalna slučajna varijabla s $\mu = 0.2$ i standardnom devijacijom 0.04.

$$P(0.402 \leq \hat{p} \leq 0.598) = F(0.598) - F(0.402)$$

```
> pnorm(0.598,mean=0.2,sd=0.04)
1
> pnorm(0.402,mean=0.2,sd=0.04)
0.9999998
```

$$P(0.402 \leq \hat{p} \leq 0.598) = 1 - 0.99998 = 0.00002$$

- Ukoliko je $p = 0.4$ tada je vjerojatnost prihvaćanja hipoteze $p = 0.5$ jednaka 0.483691.
- Ukoliko je $p = 0.2$ tada je vjerojatnost prihvaćanja hipoteze $p = 0.5$ jednaka 0.000019.

Testiranje hipoteze

Hipoteza - pretpostavka o populaciji

Statistička hipoteza - pretpostavka o parametrima populacije

Mora postojati mogućnost provjere statističkim metodama.

Statistički test - formalni postupak kojim određujemo hoćemo li odbaciti statističku hipotezu.

Nul hipoteza (H_0) - hipoteza koju testiramo. Ovisno o rezultatu testa nul hipotezu odbacujemo ili ne odbacujemo.

Alternativna hipoteza (H_a, H_1) - hipoteza koju prihvaćamo ukoliko odbacimo nul hipotezu.

Primjer s kuglicama u kutiji.

Hipoteza - Broj crvenih i plavih kuglica je jednak.

Statistička hipoteza = nul hipoteza - udio crvenih kuglica je 0.5.

Kraće: $H_0 : \quad p = 0.5$

Alternativna hipoteza - udio crvenih kuglica nije 0.5.

Kraće: $H_1 : \quad p \neq 0.5$

Testiranje hipoteze.

1. Izbor statistike. Broj crvenih kuglica u uzorku je distribuiran prema binomnoj distribuciji.

Ukoliko je hipoteza točna ($p = 0.5$), za uzorak veličine n radi se o $B(n, p)$.

Za dovoljno veliki n, proporcija crvenih kuglica u uzorku distribuirana je približno prema normalnoj distribuciji ($N(0.5, (\sigma/\sqrt{n}))^2$).

Standardizirana varijabla

$$Z = \frac{\hat{p} - 0.5}{\sigma/\sqrt{n}}$$

približno je distribuirana prema jediničnoj normalnoj distribuciji.

2. Izbor razine značajnosti.

Određujemo α, takav da, ako je hipoteza točna, vjerojatnost odbacivanja nulte hipoteze bude jednaka α.

U tom slučaju je vjerojatnost prihvaćanja (točne) nulte hipoteze jednaka $1 - \alpha$.

Broj α se naziva **razina značajnosti** testa.

U našem primjeru je bilo $\alpha = 0.05$.

3. Definiranje područja odbacivanja/prihvaćanja nul hipoteze.

Na osnovu distribucije statistike i odabrane razine značajnosti definiramo interval u kojem će se statistika nalaziti uz vjerojatnost jednaku $1-\alpha$.

Ovaj interval nazivamo **područje prihvaćanja** nul hipoteze.

Područje izvan ovog intervala nazivamo **područje odbacivanja** nul hipoteze ili **kritično područje**.

Jer statistika Z ima jediničnu normalnu razdiobu, vjerojatnost da Z poprimi vrijednost iz intervala

$$\left[z_{\alpha/2}, z_{1-\alpha/2}\right]$$

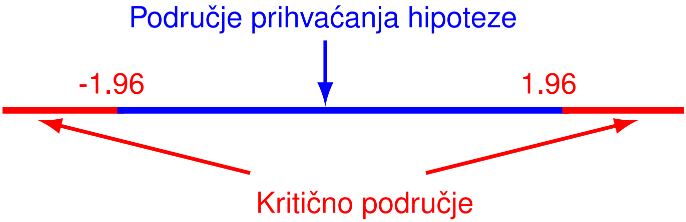
iznosi točno $1-\alpha$.

Za $\alpha = 0.05$, ovaj interval glasi $\quad [-1.96,\ 1.96]$

Područje prihvaćanja hipoteze je interval $[-1.96,\ 1.96]$.

Hipotezu odbacujemo ukoliko je z, izračunata vrijednost statistike Z, izvan tog intervala, tj. $z < -1.96$ ili $z > 1.96$.

Kritično područje su sve vrijednosti manje od -1.96 ili veće od 1.96.



4. Računanje statistike. $z = \frac{\hat{p}-0.5}{\sigma/\sqrt{n}}$

5. Odbacivanje ili prihvaćanje hipoteze.

Testiranje hipoteze o proporciji - jedna populacija

Hipoteza. Proporcija obilježja populacije jednake je zadanom broju p_0.

$$H_0 : \quad p = p_0.$$

Alternativna hipoteza.

$$H_1 : \quad p \neq p_0.$$

Test.

- Odredimo razinu značajnosti α.
- Izaberemo uzorak veličine n.
- Procjenimo proporciju obilježja u populaciji (tj. izračuna se proporcija obilježja u uzorku) $\hat{p}$

- Izračunamo statistiku

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 \cdot (1 - p_0)/n}}$$

- Ukoliko je $z < z_{\alpha/2}$ ili $z > z_{1-\alpha/2}$ hipotezu odbacujemo uz razinu značajnosti α.

  U suprotnom, hipotezu ne odbacujemo.

Primjer. Novootvoreni fitness centar naručio je promotivnu kampanju. Markentiška agencija garantirala je da će nakon kampanje 25% domaćeg tržišta znati za ovaj novi resort. Nakon kampanje provedena je telefonska anketa koja je obuhvatila 160 ispitanika. Za fitness centar je čulo njih 32. Da li centar ima razloga tvrditi da marketinška agencija nije ostvarila ciljeve kampanje?

Rješenje.

Nul hipoteza: $H_0: \quad p = 0.25$

Alternativna hipoteza: $H_1: \quad p \neq 0.25$

Razina značajnosti nije zadana. $\rightarrow \ \alpha = 0.05$

$$\begin{aligned} p_0 &= 0.25 \qquad n = 160 \\ \sigma^2 &= p \cdot (1 - p) = 0.25 \cdot 0.75 = 0.1875 \end{aligned}$$

$$\begin{aligned}
\sigma &= \sqrt{\sigma^2} = \sqrt{0.1875} = 0.4330 \\
\hat{p} &= \frac{32}{160} = 0.2 \\
z &= \frac{\hat{p} - p_0}{\sigma/\sqrt{n}} = \frac{0.2 - 0.25}{0.4330/\sqrt{160}} = -1.46
\end{aligned}$$

Jer je $z \in [-1.96, 1.96]$, nul hipotezu **ne odbacujemo**.

Napomena. U ovom testu se koristi pretpostavka da je za veliki n statistika $\hat{p}$ distribuirana približno normalno.

Budući da $n \cdot \hat{p}$ predstavlja broj pojavljivanja obilježja (uspjeha) u uzorku, $n \cdot \hat{p}$ je distribuiran prema binomnoj razdiobi.

Ukoliko je hipoteza točna, parametri razdiobe su n i p_0.

U prošlom zadatku je uvjet $z \in [-1.96, 1.96]$ zbog

$$z = \frac{\hat{p} - p_0}{\sigma/\sqrt{n}} = \frac{\hat{p} - 0.25}{0.4330/\sqrt{160}}$$

ekvivalentan uvjetu

$$\hat{p} \in [0.182905955, 0.317094045]$$

odonosno

$$n \cdot \hat{p} \in [29.26495278, 50.73504722].$$

Jer je $n \cdot \hat{p}$ cijeli broj, uvjet zapravo glasi $n \cdot \hat{p} \in [30, 50]$.

Za binomnu distribuciju $B(160, 0.25)$ je

$$F(29) = 0.024703 \quad \text{i} \quad F(50) = 0.970074$$

pa je

$$P\left(30 \leq n \cdot \hat{p} \leq 50\right) = F(50) - F(29) = 0.970074 - 0.024703 = 0.945371.$$

Jer je n velik, normalna distribucija je slična distribuciji statistike $\hat{p}$.

Ukoliko je n malen, treba koristiti binomnu distribuciju u testu!

Testiranje hipoteze o srednjoj vrijednosti - jedna populacija - standardna devijacija poznata

Pretpostavke:

- obilježje je distribuirano normalno ili je uzorak velik (veći od 30)
- poznata je standardna devijacija populacije (σ)

Zbog prve pretpostavke, aritmetička sredina ($\bar{X}$)uzorka je distribuirana normalno (ili približno normalno).

Ukoliko je točna hipoteza da je srednja vrijednost populacije jednaka nekom zadanom broju μ_0, tada je statistika

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

distribuirana prema jediničnoj normalnoj distribuciji.

Testiranje hipoteze o srednjoj vrijednosti - jedna populacija - standardna devijacija poznata

Hipoteza. Srednja vrijednost obilježja populacije jednake je zadanom broju μ_0.

$$H_0 : \quad \mu = \mu_0.$$

Alternativna hipoteza.

$$H_1 : \quad \mu \neq \mu_0.$$

Test.

- Odredimo razinu značajnosti α.
- Izaberemo uzorak veličine n.
- Procjenimo srednju vrijednost obilježja u populaciji (tj. izračuna se aritmetička sredina uzorka) $\bar{X}$

- Izračunamo statistiku
$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$
- Ukoliko je $z < z_{\alpha/2}$ ili $z > z_{1-\alpha/2}$ hipotezu odbacujemo uz razinu značajnosti α.

  U suprotnom, hipotezu ne odbacujemo.

Primjer. U istraživanju je mjerena frekvencija otkucaja srca prije treninga. Na 57 ispitanika izmjerena je srednja vrijednost od 70.42 otkucaja srca u minuti. Razlikuje li se frekvencija otkucaja srca populacije od standardnih 72 otkucaja u minuti? Standardna devijacija otkucaja srca u populaciji je 10.

Rješenje.

Nul hipoteza: $\mu = 72$

Alternativna hipoteza: $\mu \neq 72$

$$
\begin{aligned}
n &= 57 \\
\sigma &= 10 \\
\bar{X} &= 70.42 \\
\alpha &= 0.05
\end{aligned}
$$

Kritično područje: $z < -1.96$ ili $z > 1.96$

Statistika

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{70.42 - 72}{10/\sqrt{57}} = -1.192873841$$

Jer je $-1.96 \leq z \leq 1.96$ nul hipotezu ne odbacujemo.

Srednja vrijednost frekvencije otkucaja srca u populaciji je standardnih 72 otkucaja u minuti.

Testiranje hipoteze o srednjoj vrijednosti - jedna populacija - standardna devijacija nepoznata

Ukoliko je standardna devijacija σ nepoznata, tada ne možemo izračunati statistiku

$$z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Standardnu devijaciju procjenjujemo pomoću uzorka:

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1} \sum_i \left(X_i - \bar{X}\right)^2}$$

Ako σ zamijenimo sa S, statistika

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

je distribuirana prema Studentovoj (t) distribuciji s $n - 1$ stupnjeva slobode.

Hipoteza. Srednja vrijednost obilježja populacije jednake je zadanom broju μ_0.

$$H_0 : \quad \mu = \mu_0.$$

Alternativna hipoteza.

$$H_1 : \quad \mu \neq \mu_0.$$

Test.

- Odredimo razinu značajnosti α.
- Izaberemo uzorak veličine n.
- Procjenimo srednju vrijednost obilježja u populaciji (tj. izračuna se aritmetička sredina uzorka) $\bar{X}$

- Izračunamo statistiku
$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$
- Ukoliko je $t < t_{\alpha/2}(n-1)$ ili $t > t_{1-\alpha/2}(n-1)$ hipotezu odbacujemo uz razinu značajnosti α.

  U suprotnom, hipotezu ne odbacujemo.

Primjer. U istraživanju je mjerena frekvencija otkucaja srca prije treninga. Na 57 ispitanika izmjerena je srednja vrijednost od 70.42 otkucaja srca u minuti uz standardnu devijaciju 9.9480. Razlikuje li se frekvencija otkucaja srca populacije od standardnih 72 otkucaja u minuti?

Rješenje.

Nul hipoteza: $\mu = 72$

Alternativna hipoteza: $\mu \neq 72$

$$
\begin{aligned}
n &= 57 \\
S &= 9.9480 \\
\bar{X} &= 70.42 \\
\alpha &= 0.05
\end{aligned}
$$

Kritično područje:

$t < t_{\alpha/2}(56) = -2.003241$ ili $t > t_{1-\alpha/2}(56) = 2.003241$.

Statistika

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{70.42 - 72}{9.9480/\sqrt{57}} = -1.199109209$$

Jer je $-2.003241 \leq z \leq 2.003241$ nul hipotezu ne odbacujemo.

Srednja vrijednost frekvencije otkucaja srca u populaciji je standardnih 72 otkucaja u minuti.

Jednostrani i dvostrani test

Kod testiranja hipoteze o srednjoj vrijednosti

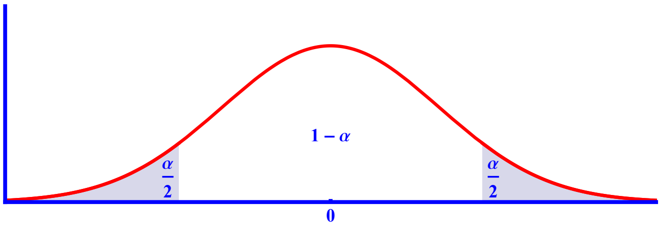
$$H_0 : \quad \mu = \mu_0$$

za alternativnu hipotezu koristili smo

$$H_1 : \quad \mu \neq \mu_0.$$

U slučaju odbacivanja nul hipoteze, zaključujemo da je srednja vrijednost populacije različito od μ_0.

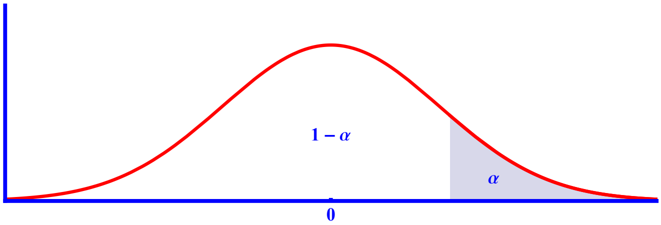
Kritično područje za dvostrani test



Kritično područje (područje odbacivanje hipoteze) se sastoji od dva rubna područja.

Stoga se ovakav test zove **dvostrani test**.

- Često nas zanima utjecaj nekog tretmana.
- U takvim situacijama želimo zaključiti da tretman povećava (ili smanjuje) srednju vrijednost.
- Alternativnu hipotezu prihvaćamo ukoliko je srednja vrijednost uzorka značajno veća od referentne vrijednosti.
- Ne prihvaćamo alternativnu hipotezu ukoliko je srednja vrijednost uzorka manja od referentne vrijednosti (makar i značajno).



Ovakav test se zove **jednostrani test**.

Kod jednostranog testa je nul hipoteza:

$$H_0 : \quad \mu = \mu_0$$

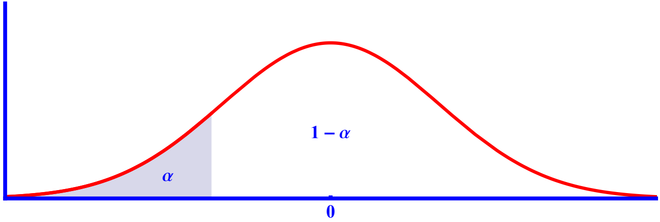
dok alternativnu hipoteza glasi

$$H_1 : \quad \mu > \mu_0.$$

Ukoliko tretman smanjuje srednju vrijednost, tada nul hipoteza i alternativna hipoteza glase

$$H_0 : \quad \mu = \mu_0$$
$$H_1 : \quad \mu < \mu_0.$$

Kritično područje je



Kod jednostranog testa bi bilo korektnije nul hipotezu i alternativnu hipotezu zapisati u obliku

$$H_0 : \quad \mu \geq \mu_0$$

$$H_1 : \quad \mu < \mu_0.$$

Alternativna hipoteza je sada suprotna nul hipotezi.

Problem je u tome što test provodimo uz pretpostavku da je nul hipoteza istinita.

Koji μ koristiti u testu?

Izbor $\mu = \mu_0$ je najnepovoljniji za odbacivanje hipoteze.

Ukoliko je $\bar{X}$ značajno manji od μ_0, tada će biti i značajno manji od bilo kojeg μ za koji je $\mu > \mu_0$.

Ovaj zapis nul hipoteze se također često koristi u literaturi.

Pogreške I. i II. vrste

- Razina značajnosti α - definiramo vjerojatnost odbacivanja nul hipoteze
- Ako odbacimo točnu nul hipotezu $\rightarrow$ radimo pogrešku
- Ako je nul hipoteza pogrešna i mi je prihvatimo $\rightarrow$ radimo pogrešku

Nul hipoteza		
	Točna	**Netočna**
Prihvaćena	**Ispravno**	**Pogrešno**
Odbaćena	**Pogrešno**	**Ispravno**

Pogrešku radimo ukoliko

- odbacimo točnu hipotezu
- prihvatimo netočnu hipotezu

- **Pogreška I. vrste** - odbacivanje točne hipoteze
- **Pogreška II. vrste** - prihvaćanje netočne hipoteze

- Pogreška I. vrste - α pogreška
- Pogreška II. vrste - β pogreška

- β - vjerojatnost prihvaćanja nul hipoteze uz uvjet da je ona netočna

- Problem: ukoliko nul hipoteza nije točna ne znamo o kojoj se točno distribuciji radi
- $\rightarrow$ ne možemo izračunati vjerojatnost.

  Snaga testa = $1-\beta$ - vjerojatnost odbacivanja nul hipoteze ukoliko je ona netočna (tj. ukoliko je alternativna hipoteza točna)

Problem određivanja vjerojatnosti pogreške II. vrste, odnosno snage testa, ilustrirat ćemo na primjeru izvlačenja kuglica iz kutije.

Testiramo hipotezu $H_0 : \quad p = 0.5$.

Veličina uzorka: $\quad n = 100$.

Područje prihvaćanja hipoteze: $\hat{p} \in [0.402,\ 0.598]$.

$\alpha = 0.05 \quad \rightarrow \quad P(0.402 \leq \hat{p} \leq 0.598) = 0.95$

Izračunali smo da ukoliko je hipoteza netočna i

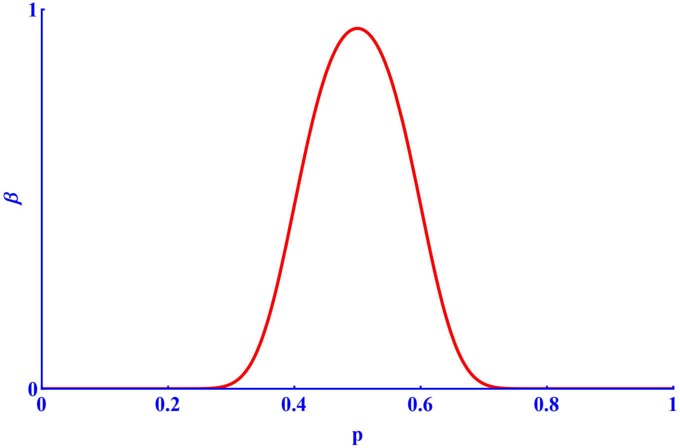
- $p = 0.4 \quad \rightarrow \quad P(0.402 \leq \hat{p} \leq 0.598) = 0.483691$
- $p = 0.2 \quad \rightarrow \quad P(0.402 \leq \hat{p} \leq 0.598) = 0.000019$

Vjerojatnost prihvaćanja nul hipoteze ukoliko je ona netočna ovisi o stvarnom udjelu crvenih kuglica u kutiji (stvarnoj vrijednosti parametra p).

Različitte vrijednosti parametra $p \quad \rightarrow \quad$ različite vrijednosti β.

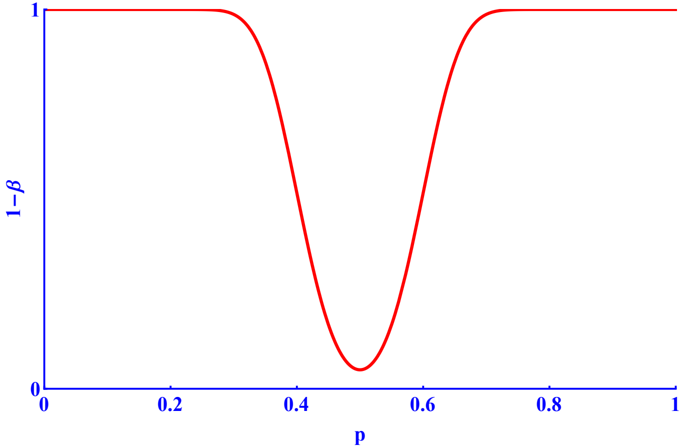
Ovisnost pogreške II. vrste o stvarnoj proporciji

$n = 100, \quad p_0 = 0.5$



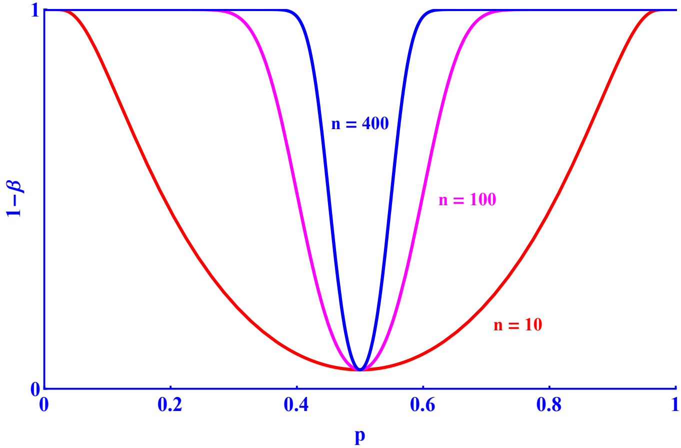
Krivulja snage testa

$n = 100, \quad p_0 = 0.5$



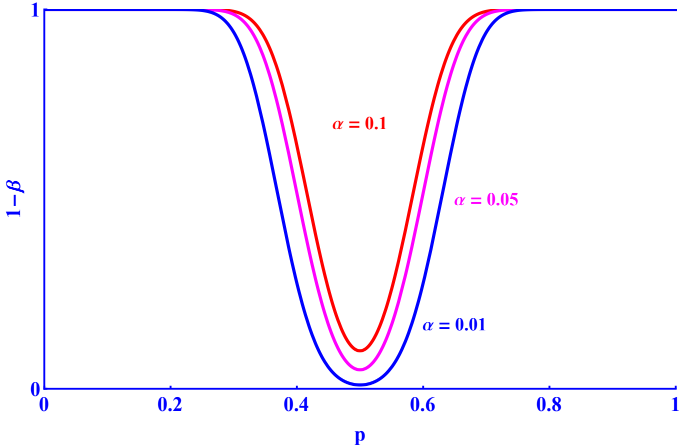
Ovisnost snage testa o veličini uzorka

$p_0 = 0.5$



Ovisnost snage testa o razini značajnosti α

$n = 100, \quad p_0 = 0.5$



- povećanje veličine uzorka povećava snagu testa
- manja razina značajnosti

  $\rightarrow$ veće područje prihvaćanja hipoteze

  $\rightarrow$ veća vjerojatnost prihvaćanja nul hipoteze

  $\rightarrow$ manja vjerojatnost prihvaćanja alternativne hipoteze (makar ona bila točna)

  $\rightarrow$ manja snaga testa

Usporedba srednjih vrijednosti dviju populacija

Ovo je vjerojatno najčešće korištena usporedba

- Uspoređujemo dvije populacije.
- Testiramo hipotezu o jednakosti srednjih vrijednosti dviju populacija.
- Najčešće se ispituje efekt nekog tretmana.

  Jedna populacija su osobe na koje se primjenjuje tretman.

  Druga populacija su osobe na koje se ne primjenjuje tretman.

  Ovdje se radi o hipotetskim populacijama. Fizički to može biti ista skupina osoba.

  Jednu populaciju čine te osobe kada bi se na njih primijenio tretman a drugu čine iste te osobe ali kada se na njih ne bi primijenio tretman.

- Iz svake populacije izaberemo uzorak veličine n_X za prvu i n_Y za drugu populaciju. (Uzorci ne trebaju biti iste veličine.)
- Uzorak iz prve populacije označimo s

$$X_1, X_2, \ldots, X_{n_X}$$

  a uzorak iz druge populacije označimo s

$$Y_1, Y_2, \ldots, Y_{n_Y}.$$

- Ukoliko se radi o ispitivanju efikasnosti tretmana, tada uzorak na kojem se provodi tretman nazivamo **eksperimentalna skupina** a uzorak na kojem se ne provodi tretman nazivamo **kontrolna skupina**
- Pretpostavka je da je obilježje u obje populacije **normalno** distribuirano.
- Srednje vrijednosti populacije označimo s μ_X i μ_Y a standardne devijacije s σ_X i σ_Y.

Aritmetičke sredine uzoraka

$$\bar{X} = \frac{1}{n_X}\sum_i X_i \quad \text{i} \quad \bar{Y} = \frac{1}{n_Y}\sum_i Y_i$$

su normalno distribuirani i vrijedi

$$\mathrm{E}\,(\bar{X}) = \mu_X, \quad \mathrm{E}\,(\bar{Y}) = \mu_Y, \quad \mathrm{Var}\,(\bar{X}) = \frac{\sigma_X^2}{n_X}, \quad \mathrm{Var}\,(\bar{Y}) = \frac{\sigma_Y^2}{n_Y}.$$

Razlika $D = \bar{X} - \bar{Y}$ je također normalno distribuirana i

$$\mathrm{E}\,(D) = \mathrm{E}\,(\bar{X} - \bar{Y}) = \mathrm{E}\,(\bar{X}) - \mathrm{E}\,(\bar{Y}) = \mu_X - \mu_Y$$

$$\mathrm{Var}\,(D) = \mathrm{Var}\,(\bar{X} - \bar{Y}) = \mathrm{Var}\,(\bar{X}) + \mathrm{Var}\,(\bar{Y}) = \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}.$$

Ukoliko je hipoteza o jednakosti srednjih vrijednosti točna ($\mu_X = \mu_X$) tada je

$$\mathrm{E}\,(\bar{X}) - \mathrm{E}\,(\bar{Y}) = 0$$

i standardizirana varijabla

$$\frac{D - \mathrm{E}\,(D)}{\sqrt{\mathrm{Var}\,(D)}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

distribuirana je prema jediničnoj normalnoj razdiobi.

Standardne devijacije populacije su najčešće nepoznate.

σ_X^2 i σ_Y^2 procijenimo pomoću uzorka s S_X^2 i S_Y^2.

Time dolazimo do statistike

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}.$$

t nije distribuiran prema Studentovoj distribuciji!

U nazivniku se ne nalazi procjenitelj varijance S^2.

Distribucija od t je slična Studentovoj razdiobi.

Broj stupnjeva slobode (ν) je dan s

$$\nu = \frac{\left[S_X^2/n_X + S_Y^2/n_Y\right]^2}{\frac{\left(S_X^2/n_X\right)^2}{n_X} + \frac{\left(S_Y^2/n_Y\right)^2}{n_Y}} - 2.$$

Gornja veličina nije cijeli broj te se koristi najbliža cjelobrojna vrijednost.

Ukoliko je standardna devijacija u obje populacije jednaka:

$$\sigma_X = \sigma_Y = \sigma$$

tada je

$$\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} = \sigma\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}.$$

Varijancu σ^2 možemo procijeniti s

$$S_P^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}.$$

U ovom slučaju statistika je oblika

$$t = \frac{\bar{X} - \bar{Y}}{S_P \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}.$$

t je distribuiran prema Studentovoj razdiobi s $n_X + n_Y - 2$ stupnjeva slobode.

Razlikujemo slučajeve kada su varijance populacija jednake i kada su različite.

Praktični pristup je da se varijance smatraju istim ukoliko je omjer procjena standardne devijacije (S_X i S_Y) između 0.5 i 2.

Hipoteza o jednakosti varijanci može se testirati. (Kasnije.)

Testiranje hipoteze o jednakim srednjim vrijednostima dvije populacije - standardne devijacije populacija jednake

Hipoteza. Srednja vrijednost obilježja je jednako za dvije populacije.

Pretpostavke: Standardna devijacija obilježja jednako je za obje populacije i nepoznato. Obilježje distribuirano normalno.

Nul hipoteza: $H_0: \quad \mu_X = \mu_Y$

Alternativna hipoteza: $H_1: \quad \mu_X \neq \mu_Y$

Odaberemo razinu značajnosti α.

Statistika:

$$t = \frac{\bar{X} - \bar{Y}}{S_P \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}.$$

gdje je

$$S_P^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}.$$

Distribucija statistike: Studentova (t) distribucija s $n_X + n_Y - 2$ stupnjeva slobode.

Kritično područje: $t \leq t_{\alpha/2}(n_X + n_Y - 2)$ ili $t \geq t_{1-\alpha/2}(n_X + n_Y - 2)$.

Testiranje hipoteze o jednakim srednjim vrijednostima dvije populacije - standardne devijacije populacija različite

Hipoteza. Srednja vrijednost obilježja je jednako za dvije populacije.

Pretpostavke: Standardna devijacija obilježja različito je obje populacije i nepoznato. Obilježje distribuirano normalno.

Nul hipoteza: $H_0: \quad \mu_X = \mu_Y$

Alternativna hipoteza: $H_1: \quad \mu_X \neq \mu_Y$

Odaberemo razinu značajnosti α.

Statistika:

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}.$$

Distribucija statistike: Studentova (t) distribucija s ν stupnjeva slobode gdje je ν najbliži cijeli broj veličini

$$\nu = \frac{\left[S_X^2/n_X + S_Y^2/n_Y\right]^2}{\frac{\left(S_X^2/n_X\right)^2}{n_X} + \frac{\left(S_Y^2/n_Y\right)^2}{n_Y}} - 2.$$

Kritično područje: $t \leq t_{\alpha/2}(\nu)$ ili $t \geq t_{1-\alpha/2}(\nu)$.

Usporedba srednjih vrijednosti dviju populacija - sparene vrijednosti

U prethodnim testovima važna je pretpostavka da su uzorci iz dvije populacije međusobno nezavisni.

Narušavanje ove pretpostavke onemogućava primjenu prethodnih testova.

Česti ndizajn eksperimenta je da se efekt tretmana mjeri usporedbom vrijednosti obilježja prije i poslije primjene tretmana.

Na svakoj jedinki obavljaju se dva mjerenja.

Dizajn prije i poslije.

Promatramo da li tretman povečava (smanjuje) srednju vrijednost obilježja.

Dva mjerenja na istoj jedinki (prije i poslije) su **zavisna**.

X obilježje prije tretmana

Y obilježje poslije tretmana

Za svaku jedinku postoje dva mjerenja: (X_i, Y_i).

Uzorci $X_1, X_2, \ldots, X_n$ i $Y_1, Y_2, \ldots, Y_n$ su zavisni.

Mjerenja unutar svakog uzorka su nezavisna (nezavisne slučajne varijable).

Promatramo učinak tretmana na pojedinu jedinku. Definiramo promjenu obilježja:

$$D_i = Y_i - X_i.$$

$D = Y - X$ je nova varijabla.

Hipoteza

$$\mu_X = \mu_Y$$

je ekvivalentna hipotez

$$\mu_D = 0.$$

Novi uzorak je $D_1, D_2, \ldots, D_n$.

Iako imamo $2 \cdot n$ mjerenja (podataka), veličina uzorka je n.

Hipotezu $\mu_D = 0$ testiramo pomoću testa o srednjoj vrijednosti za jednu populaciju.

Hipoteza. Tretman ne utječe na povećanje srednja vrijednosti obilježja.

$$H_0 : \quad \mu_X \leq \mu_Y.$$

Alternativna hipoteza. $H_1 : \quad \mu_X < \mu_Y.$

Test.

- Odredimo razinu značajnosti α.
- Izaberemo uzorak veličine n.
- Izračunamo aritmetičku sredinu obilježja prije tretmana ($\bar{X}$) i poslije tretmana ($\bar{Y}$).
- Definiramo efekt tretmana: $D = Y - X$.

- Izračunamo statistiku

$$t = \frac{\bar{Y} - \bar{X}}{S_D/\sqrt{n}} = \frac{\bar{D}}{S_D/\sqrt{n}}$$

- Ukoliko je $t > t_{1-\alpha}(n-1)$ hipotezu odbacujemo uz razinu značajnosti α.

  U suprotnom, hipotezu ne odbacujemo.

Primjer. U istraživanju utjecaja stimulansa na krvni tlak, istraživači su dvanaestorici pacijenata dali stimulans. Svakom pacijentu je krvni tplak izmjeren prije i poslije davanja stimulansa.

Postoji li opravdanje za tvrdnju da stimulans povećava krvni tlak?

Pacijent	Krvni tlak	
	Prije (X)	Poslije (Y)
1	120	128
2	124	131
3	130	131
4	118	127
5	140	132
6	128	125

Pacijent	Krvni tlak	
	Prije (X)	Poslije (Y)
7	140	141
8	135	137
9	126	118
10	130	132
11	126	129
12	127	135

Rješenje.

Hipoteza: $H_0 : \quad \mu_X \geq \mu_Y$

Alternativna hipoteza: $H_1 : \quad \mu_X < \mu_Y$

Razina značajnosti: $\alpha = 0.05$

Uzorak: $n = 12$

Definiramo razliku D:

Pacijent	Krvni tlak		
	Prije (X)	Poslije (Y)	Razlika (D)
1	120	128	8
2	124	131	7
3	130	131	1
4	118	127	9
5	140	132	-8
6	128	125	-3
7	140	141	1
8	135	137	2
9	126	118	-8
0	130	132	2
11	126	129	3
12	127	135	8

Statistika:

$$t = \frac{\bar{Y} - \bar{X}}{S_D/\sqrt{n}} = \frac{\bar{D}}{S_D/\sqrt{n}}$$

Distribucija: Studentova razdioba $t(11)$

Kritično područje: $t > t_{0.95}(11) = 1.796$

$$\bar{D} = 1.833333333 \quad S_D = 5.828352852$$

$$t = \frac{1.833333333}{5.828352852/\sqrt{12}} = 1.09$$

Ne odbacujemo hipotezu.

Nemamo dovoljno pokazatelja za zaključak da stimulans povećava krvni tlak.

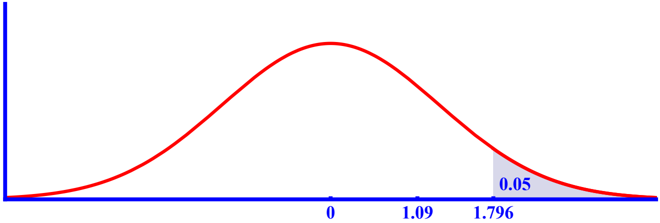
p-vrijednost

U prošlom primjeru izračunali smo vrijednost t statistike:

$$t = 1.09.$$

Dobivenu vrijednost uspoređujemo s kritičnom vrijednošću jednostranog t-testa za razinu značajnosti $\alpha = 0.05$:

$$t_{0.95}(11) = 1.796.$$

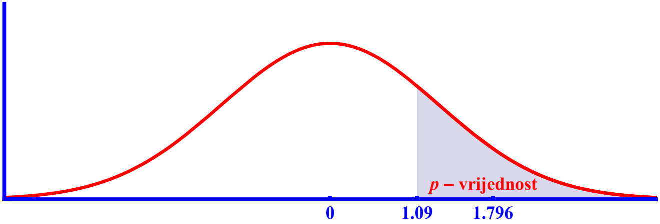


Umjesto usporedbe vrijednosti statistike t s kritičnom vrijednošću, danas je uobičajeno računati vjerojatnost $P(X \geq t)$ gdje je X slučajna varijabla distribuirana prema Studentovoj distribuciji s 11 stupnjeva slobode.

Ovu vjerojatnost nazivamo **p-vrijednost**.

Koristeći R lagano izračunamo da je

$$P(X \geq 1.09) = 0.149508.$$



Ako koristimo p-vrijednost nije potrebno računati kritično područje.

Hipotezu odbacujemo ukoliko p-vrijednost zadovoljava $p^* < \alpha$.

U suprotnom hipotezu ne odbacujemo.

Rješenje primjera u R-u.

Podaci.

```
> prije
[1] 120 124 130 118 140 128 140 135 126 130 126 127
> poslije
[1] 128 131 131 127 132 125 141 137 118 132 129 135
```

T-test.

```
> t.test(prije,poslije, paired = TRUE)

Paired t-test

data:  prije and poslije

t = -1.0896, df = 11, p-value = 0.2992

alternative hypothesis:  true difference in means is
not equal to 0

95 percent confidence interval:

-5.536492 1.869825

sample estimates:

mean of the differences

-1.833333
```

```
Paired t-test
data:  prije and poslije
t = -1.0896, df = 11, p-value = 0.2992
alternative hypothesis:  true difference in means is
not equal to 0
95 percent confidence interval:
-5.536492 1.869825
sample estimates:
mean of the differences
-1.833333
```

Razlika = -1.83333

Razlika=Prije - Poslije

$t = -1.08965$ - t je negativan zbog obratnog računanja razlike

```
Paired t-test
data:  prije and poslije
t = -1.0896, df = 11, p-value = 0.2992
alternative hypothesis:  true difference in means is
not equal to 0
95 percent confidence interval:
-5.536492 1.869825
sample estimates:
mean of the differences
-1.833333
```

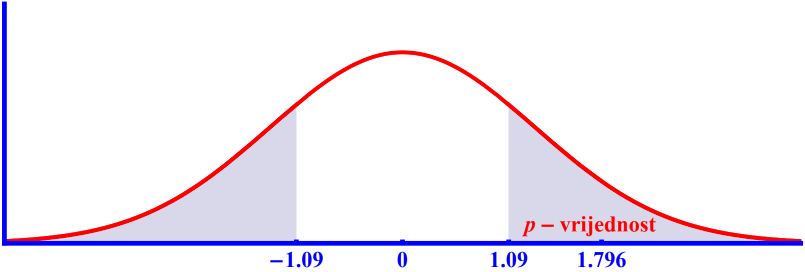
$t = -1.08965$ - t je negativan zbog obratnog računanja razlike

Broj stupnjeva slobode: df $= 11$

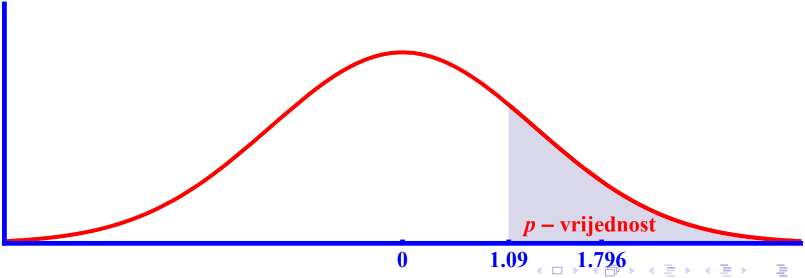
p-vrijednost: 0.299163

Različito od našeg primjera. **Test u Statistici je dvostran!**

p-vrijednost za dvostrani test.



p-vrijednost za jednostrani test.



p-vrijednost za dvostrani test je dvostruko veća od p-vrijednosti za jednostrani test.

Računanje p-vrijednosti u primjeru:

```
t = -1.0896, df = 11, p-value = 0.2992
```

$$p^* = 0.299163/2 = 0.1495815$$

To možemo i direktno u R-u:

```
> t.test(prije,poslije, paired = TRUE,
alternative=c("less"))

Paired t-test

data:  prije and poslije

t = -1.0896, df = 11, p-value = 0.1496

alternative hypothesis:  true difference in means is
less than 0

95 percent confidence interval:

-Inf 1.188244

sample estimates:

mean of the differences

-1.833333
```

Pretpostavke o uzorku

- Glavna pretpostavka je **nezavisnost** uzorka.

  $\rightarrow$ Izbor jedinki u uzorak je nezavisan.

- U slučaju dva uzorka to vrijedi za jedinje u oba uzorka.
- Izuzetak je testiranje između sparenih uzoraka.

  $\rightarrow$ Mjerenja trebaju biti sparena.

- **Normalna distribucija** populacije.
- Kada je standardna devijacija poznata uvjet normalnosti zamjenjuje uvjet dovoljno velikog uzorka (centralni granični teorem, $n \geq 30$)
- Kod testova o proporciji ne javlja se normalna već binomna distribucija. Uvjet je dovoljno velika velečina uzorka.

- Testovi normalnosti $\rightarrow$ testiranje hipoteze o normalnosti populacije (kasnije)
- Ako distribucija populacija odstupa od normalne, ali obje distribucije su sličnog oblika, *t* test je pouzdaniji ukoliko je veličina uzorka za obje populacije jednaka.
- Neki autori dozvoljavaju da je distribucija populacije **približno normalna** ukoliko je ispunjen jedan od sljedećih uvjeta:
    - Distribucija podataka je simetrična, unimodalna, bez ekstremnih vrijednosti (outliera) i veličina uzorka je 15 ili manja.
    - Podaci su blago zakošeni, unimodalni, bez ekstremnih vrijednosti (outliera) i veličina uzorka je između 16 i 40
    - veličina uzorka je preko 40 i nema ekstremnih vrijednosti.

- p-vrijednost je vjerojatnost da je nul hipoteza netočna. **Krivo.**

  Ispravna interpretacija. p-vrijednost je vjerojatnost da statistika ima vrijednost veliku (ili veću) od dobivene vrijednosti ukoliko je nul hipoteza točna.

- Neznačajan rezultat znači da je nul hipoteza vjerojatno točna. **Krivo.**

  Ispravna interpretacija. Neznačajan rezultat znači da podaci ne pokazuju da je nul hipoteza netočna.

- Ukoliko odbacimo nul hipotezu uz $\alpha = 0.05$ tada smo 95% sigurni da je alternativna hipoteza točna. **Krivo.**

  Ispravna interpretacija. α određuje vjerojatnost da odbacimo nul hipotezu ukoliko je ona točna.

  $1 - \alpha$ je vjerojatnost ne odbacivanja nul hipoteze ukoliko je ona točna.

  Vjerojatnosti da su nul hipoteza ili alternativna hipoteza točne su nam nepoznate. Znamo samo za uvjetne vjerojatnosti prihvaćanja ili odbacivanja hipoteze.

- Odbacivanje nul hipoteze uz $\alpha = 0.05$ znači da vjerojatnost da smo napravili pogrešku I. vrste iznosi 5%. **Krivo.**

  Ispravna interpretacija. To je uvjetna vjerojatnost. Vjerojatnost da smo napravili pogrešku I. vrste ukoliko je nul hipoteza točna.

- Neodbacivanje nul hipoteze znači da je ona točna. **Krivo.**

  Ispravna interpretacija. Neodbacivanje nul hipoteze znači da podaci ne podupiru zaključak da je ona netočna.

- Ako je snaga testa 80% i mi ne odbacimo nul hipotezu to znači da vjerojatnost da je nul hipoteza istinita iznosi 20%. **Krivo.**

  Ispravna interpretacija. 20% je vjerojatnost prihvaćanja nul hipoteze ukoliko je ona netočna.

  Bez obzira na rezultat testa mi ne znamo je li nul hipoteza točna ili nije.

- Ako je snaga testa 80% i mi odbacimo nul hipotezu to znači da vjerojatnost da je alternativna hipoteza istinita iznosi 80%. **Krivo.**

  Ispravna interpretacija. 80% je vjerojatnost prihvaćanja alternativne hipoteze ukoliko je ona točna.

Određivanje veličine uzorka
Test o srednjoj vrijednosti - jedna populacija

Želimo testirati hipotezu

$$H_0: \quad \mu = \mu_0$$

uz alternativnu hipotezu

$$H_1: \quad \mu \neq \mu_0$$

Ukoliko je standardna devijacija poznata, statistika je dana s

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

Jer se radi o dvostranom testu, područje prihvaćanja hipoteze je

$$\left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right].$$

Ovdje smo iskoristili simetričnost jedinične normalne razdiobe-

Znači, hipotezu prihvaćamo ukoliko je

$$\bar{X} - \mu_0 \in \left[-z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}},\ z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

U istraživanju je važno kolika je razlika relevantna.

Definiramo d, veličinu efekta koju smatramo značajnom.

Cilj nam je da ukoliko je uočena razlika između $\bar{X}$ i μ_0 veća od d odbaciti hipotezu o jednakosi uz razinu značajnosti α.

Hipotezu odbacujemo ukoliko je

$$\left\| \bar{X} - \mu_0 \right\| > d$$

pa treba vrijediti

$$z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} < d.$$

Ovo nam daje uvjet na veličinu uzorka

$$n > \left(z_{1-\alpha/2}\frac{\sigma}{d}\right)^2.$$

Za određivanje veličine uzorka potrebno je poznavati:

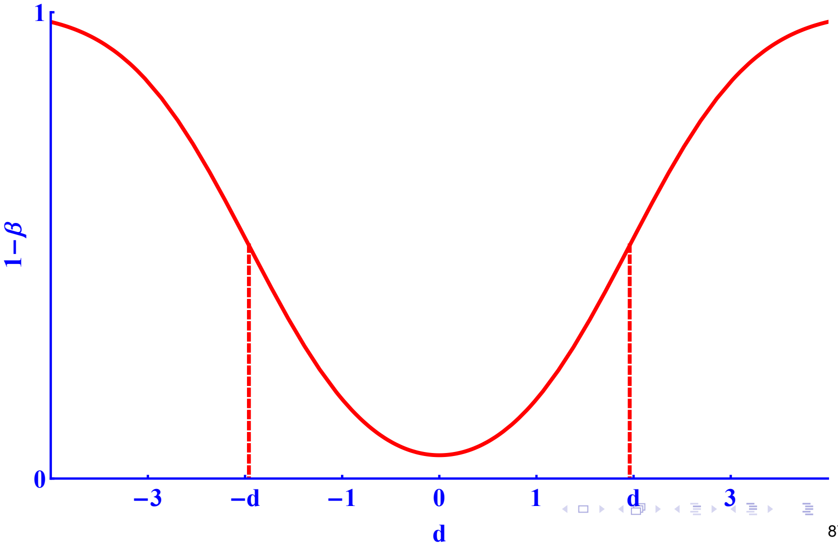
- razinu značajnosti α
- standardnu devijaciju populacije (σ)
- veličinu efekta (d)

U obzir treba uzeti i snagu testa.

Standardni zahtjev na snagu testa je da je on veći ili jednak od **0.8**. ($\beta \leq 0.2$).

β je vjerojatnost pogreške II. vrste: prihvaćanja nul hipoteze ukoliko ona nije točna.

Krivulja snage testa u ovisnosti o efektu:



Snaga testa raste s povečanjem veličine uzorka.

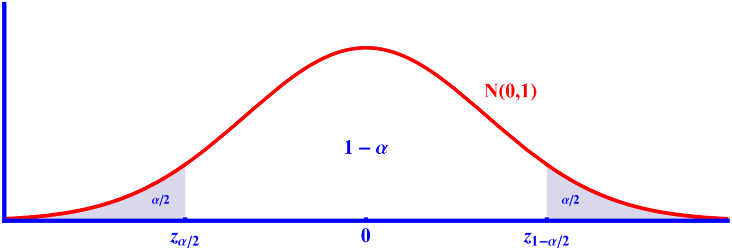
Koliki treba biti uzorak da vjerojatnost prihvaćanja alternativne hipoteze ukoliko je ona točna bude veća od $1 - \beta$?

Želimo da ukoliko je uočena razlika između $\bar{X}$ i μ_0 manja od d da prihvatimo nul hipotezu uz pogrešku II. vrste manju od β.

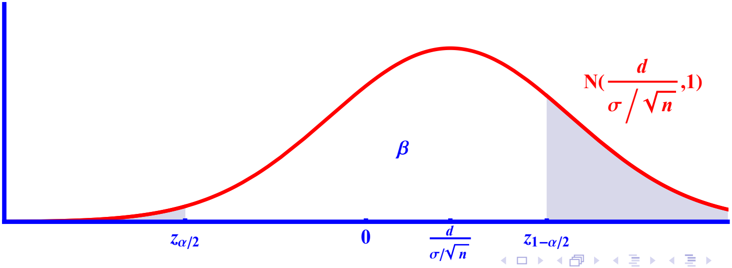
Snaga je najmanja ukoliko je razlika točno d.

Računamo snagu uz pretpostavku $\mu = \mu_0 + d$.

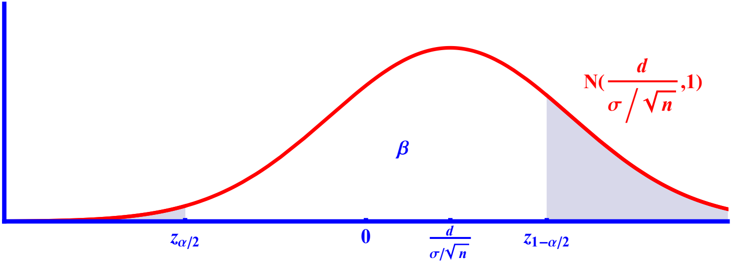
Ako je nul hipoteza točna: $\mu = 0$ Radi jednostavnosti smo stavili da je $\mu_0 = 0$.

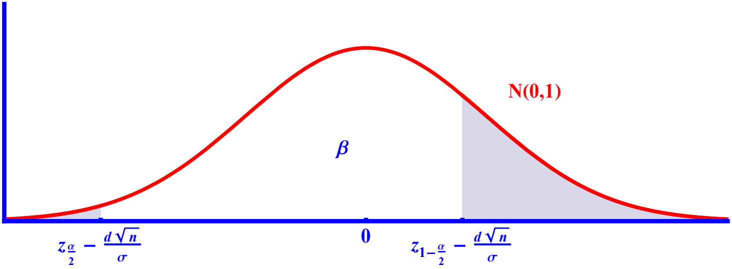


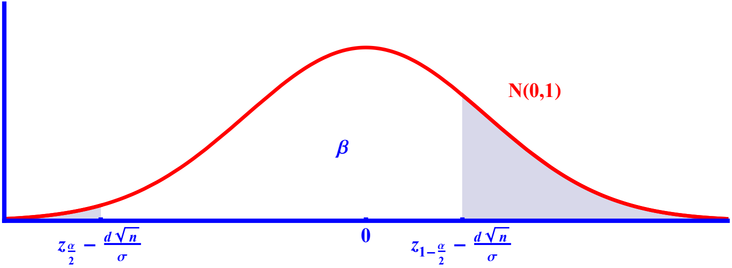
Ako nul hipoteza nije točna i $\mu = d$:



Pomaknemo da dobijemo jediničnu normalnu razdiobu:







Želimo da je za $\mu = \mu_0 + d$

$$\begin{aligned}\beta &= P\left(z_{\alpha/2} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right)\\ &= P\left(z_{\alpha/2} - \frac{d}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu_0 - d}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}}\right)\end{aligned}$$

Ako je $\mu = \mu_0 + d$ onda je

$$\frac{\bar{X} - \mu_0 - d}{\sigma/\sqrt{n}}$$

distribuiran prema jediničnoj normalnoj razdiobi.

S Φ označimo funkciju distribucije za jediničnu normalnu razdiobu.

$$\begin{aligned}
\beta &= P\left(-z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu_0 - d}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

$$\begin{aligned} 1-\beta &= 1-\Phi\left(z_{1-\alpha/2}-\frac{d}{\sigma/\sqrt{n}}\right)+\Phi\left(-z_{1-\alpha/2}-\frac{d}{\sigma/\sqrt{n}}\right) \\ &= \Phi\left(-z_{1-\alpha/2}+\frac{d}{\sigma/\sqrt{n}}\right)+\Phi\left(-z_{1-\alpha/2}-\frac{d}{\sigma/\sqrt{n}}\right) \end{aligned}$$

Zadnja jednakost vrijedi zbog simetričnosti normalne distribucije: $1-\Phi(a)=\Phi(a)$.

Uvjet na n:

$$1-\beta=\Phi\left(-z_{1-\alpha/2}+\frac{d}{\sigma/\sqrt{n}}\right)+\Phi\left(-z_{1-\alpha/2}-\frac{d}{\sigma/\sqrt{n}}\right)$$

Iz ove jednadžbe ne možemo dobiti jednostavni izraz za n, ali se n može odrediti.

U izrazu se d i σ javljaju u omjeru.

Omjer d/σ se naziva **standardizirani efekt**.

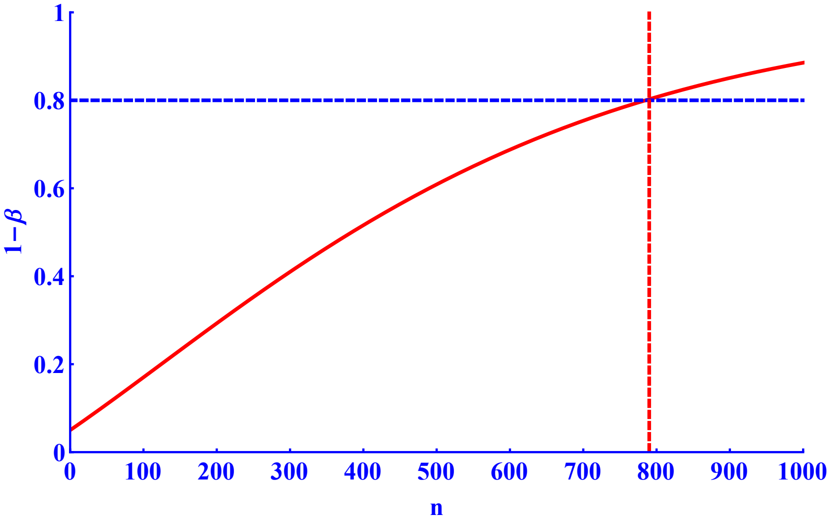
Primjer. Odredite veličinu uzorka ukoliko je standardna devijacija populacije 10, traženi efekt 1 te $\alpha = 0.05$ i $\beta = 0.2$.

$$\begin{aligned} d &= 1 \\ \sigma &= 10 \\ \frac{d}{\sigma} &= 0.1 \\ \alpha &= 0.05 \\ z_{1-\alpha/2} &= z_{0.975} = 1.96 \\ \beta &= 0.2 \\ 1-\beta &= 0.8 \end{aligned}$$

$$1-\beta = \Phi\left(-z_{1-\alpha/2} + \frac{d}{\sigma/\sqrt{n}}\right) + \Phi\left(-z_{1-\alpha/2} - \frac{d}{\sigma/\sqrt{n}}\right)$$

$$0.8 = \Phi\left(-1.96 + 0.1 \cdot \sqrt{n}\right) + \Phi\left(-1.96 - 0.1 \cdot \sqrt{n}\right)$$

Ovisnost snage testa ($1-\beta$) o veličini uzorka (n)



$n \approx 800$

Da smo veličinu uzorka određivali samo iz α pogreške:

$$n > \left(z_{1-\alpha/2}\frac{\sigma}{d}\right)^2 = \left(1.96 \cdot \frac{10}{1}\right)^2 = 384.16$$

Rješenje u R-u:

Za određivanje veličine uzorka nam treba paket 'pwr':

```
> install.packages(c("pwr"))

> library(pwr)

> pwr.t.test(d=0.1, sig.level=0.05, power=0.80,
type="one.sample", alternative="two.sided")

One-sample t test power calculation

n = 786.8089

d = 0.1

sig.level = 0.05

power = 0.8

alternative = two.sided
```

Važno:

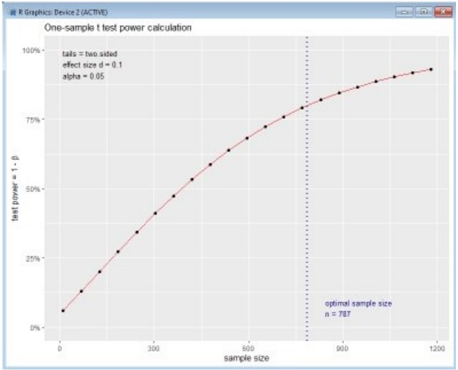
```
> pwr.t.test(d=0.1, sig.level=0.05, power=0.80,
type="one.sample", alternative="two.sided")
```

Parametar `d` je standardizirani efekt:

$$\mathtt{d} = \frac{d}{\sigma} = \frac{1}{10} = 0.1.$$

Graf:

```
> pl = pwr.t.test(d=0.1, sig.level=0.05, power=0.80,
type="one.sample", alternative="two.sided")
> plot(pl)
```



Test o srednjoj vrijednosti - dvije populacije

Želimo testirati hipotezu

$$H_0 : \quad \mu_X = \mu_Y$$

uz alternativnu hipotezu

$$H_1 : \quad \mu_X \neq \mu_Y$$

Ukoliko je standardna devijacija poznata, statistika dana s

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

distribuirana je prema jediničnoj normalnoj razdiobi.

Za dvostrani test, područje prihvaćanja hipoteze je

$$\left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right].$$

Jednake varijance i jednake veličine uzorka

Za

$$\sigma = \sigma_X = \sigma_Y \quad \text{i} \quad n = n_X = n_Y$$

je

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} = \frac{\bar{X} - \bar{Y}}{\frac{\sigma}{\sqrt{n}}\sqrt{2}}$$

Zahtjev da za efekt d razina značajnosti testa bude α daje uvjet:

$$z_{1-\alpha/2} \leq \frac{d}{\frac{\sigma}{\sqrt{n}}\sqrt{2}}.$$

Uvjet na veličinu uzoraka je

$$n \geq 2\left(z_{1-\alpha/2}\frac{\sigma}{d}\right)^2$$

Snaga testa.

Slično kao i kod jedne populacije, ukoliko je zadana i snaga testa, uvjet na veličinu uzorka je dan s:

$$1-\beta=\Phi\left(-z_{1-\alpha/2}+\frac{d}{\sqrt{2}\,\sigma/\sqrt{n}}\right)+\Phi\left(-z_{1-\alpha/2}-\frac{d}{\sqrt{2}\,\sigma/\sqrt{n}}\right)$$

Pojednostavljena aproksimacija bi dala:

$$n=2\left(\frac{\sigma}{d}(z_{1-\alpha/2}+z_{1-\beta})\right)^2.$$

Jednake varijance i različite veličine uzorka

Označimo s r omjer

$$r = \frac{n_Y}{n_X}.$$

Za

$$\sigma = \sigma_X = \sigma_Y, \quad n_X = n \quad \text{i} \quad n_Y = r\,n$$

je

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} = \frac{\bar{X} - \bar{Y}}{\frac{\sigma}{\sqrt{n}}\sqrt{1 + \frac{1}{r}}}$$

distribuiran prema normalnoj distribuciji.

Zahtjev da za efekt d razina značajnosti testa bude α daje uvjet:

$$z_{1-\alpha/2} \leq \frac{d}{\frac{\sigma}{\sqrt{n}}\sqrt{1 + \frac{1}{r}}}.$$

Uvjet na veličinu uzoraka je

$$n \geq \frac{r+1}{r}\left(z_{1-\alpha/2}\frac{\sigma}{d}\right)^2$$

Snaga testa.

Slično kao i kod jedne populacije, ukoliko je zadana i snaga testa, uvjet na veličinu uzorka je dan s:

$$1-\beta = \Phi\left(-z_{1-\alpha/2} + \frac{d}{\sqrt{\frac{r+1}{r}\,\frac{\sigma}{\sqrt{n}}}}\right) + \Phi\left(-z_{1-\alpha/2} - \frac{d}{\sqrt{\frac{r+1}{r}\,\frac{\sigma}{\sqrt{n}}}}\right)$$

Pojednostavljena aproksimacija bi dala:

$$n = \frac{r+1}{r}\left(\frac{\sigma}{d}(z_{1-\alpha/2} + z_{1-\beta})\right)^2.$$

Za koji r je ukupna veličina uzorka najmanja?

Ukupna veličina uzorka:

$$n_X + n_Y = n + r \cdot n = (1 + r)n = \frac{(r+1)^2}{r} \left(\frac{\sigma}{d} (z_{1-\alpha/2} + z_{1-\beta}) \right)^2 .$$

Izraz

$$\frac{(r+1)^2}{r}$$

je najmanji za $r = 1$.

Ukoliko je standardna devijacija populacija jednaka, najefikasnije je imati uzorke iste veličine.

Usporedba proporcija dviju populacija

Testiramo hipotezu

$$H_0 : \quad p_X = p_Y$$

uz alternativnu hipotezu

$$H_1 : \quad p_X \neq p_Y$$

Na osnovu uzoraka veličina n_X i n_Y procijenimo proporcije populacija.

Ukoliko je hipoteza točna ($p_X = p_Y = p$) procjene proporcija $\hat{p}_X$ i $\hat{p}_Y$ zadovoljavaju

$$\mathrm{E}\left(\hat{p}_X\right) = \mathrm{E}\left(\hat{p}_Y\right) = p$$

i

$$\mathrm{Var}\left(\hat{p}_X\right) = \frac{\sigma^2}{n_X} = \frac{p \cdot (1-p)}{n_X}, \quad \mathrm{Var}\left(\hat{p}_Y\right) = \frac{\sigma^2}{n_Y} = \frac{p \cdot (1-p)}{n_Y}.$$

Razlika $D = \hat{p}_X - \hat{p}_Y$ zadovoljava

$$\mathrm{E}\left(D\right) = \mathrm{E}\left(\hat{p}_X - \hat{p}_Y\right) = \mathrm{E}\left(\hat{p}_X\right) - \mathrm{E}\left(\hat{p}_Y\right) = p - p = 0,$$

$$\mathrm{Var}\left(D\right) = \mathrm{Var}\left(\hat{p}_X - \hat{p}_Y\right) = \mathrm{Var}\left(\hat{p}_X\right) + \mathrm{Var}\left(\hat{p}_Y\right) = \frac{\sigma^2}{n_X} + \frac{\sigma^2}{n_Y} =$$

$$= \sigma^2 \left(\frac{1}{n_X} + \frac{1}{n_Y}\right) = p \cdot (1 - p) \left(\frac{1}{n_X} + \frac{1}{n_Y}\right)$$

Za dovoljno veliki uzorak, statistika

$$\frac{D - \mathrm{E}(D)}{\sqrt{\mathrm{Var}(D)}} = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{p \cdot (1-p)}\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}$$

distribuirana je prema jediničnoj normalnoj razdiobi.

Proporcija populacije je nepoznata pa je procijenjujemo na osnovu cijelog uzorka.

Ukoliko je m_X broj pojavljivanja uspjeha u prvoj a m_Y broj uspjeha u drugoj populaciji:

$$\hat{p}_X = \frac{m_X}{n_X} \quad \text{i} \quad \hat{p}_Y = \frac{m_Y}{n_Y}$$

tada p procjenjujemo s

$$\bar{p} = \frac{m_X + m_Y}{n_X + n_Y} = \frac{n_X \cdot \hat{p}_X + n_Y \cdot \hat{p}_Y}{n_X + n_Y}.$$

Time dolazimo do statistike

$$z = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\bar{p} \cdot (1 - \bar{p})}\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}},$$

gdje je

$$\bar{p} = \frac{m_X + m_Y}{n_X + n_Y} = \frac{n_X \cdot \hat{p}_X + n_Y \cdot \hat{p}_Y}{n_X + n_Y}.$$

z je distribuiran prema jediničnoj normalnoj distribuciji.

Primjer. U istraživanju o kvaliteti nastave Tjelesne i zdravstvene kulture na sveučilištima, istraživači su na uzorku studenata testirali ispunjavanje norme iz plivanja. Od 88 studenata s prvog sveučilišta, njih 33 je ispunilo normu dok je od 84 studenata s drugog sveučilišta njih 5 ispunilo normu? Postoji li razlika između studenata dva sveučilišta?

Rješenje.

$$n_X = 88, \quad n_Y = 84$$

$$\hat{p}_X = \frac{33}{88} = 0.375, \quad \hat{p}_Y = \frac{5}{4} = 0.060$$

$$\bar{p} = \frac{m_X + m_Y}{n_X + n_Y} = \frac{33 + 5}{88 + 84} = 0.221$$

$$\begin{aligned} z &= \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\bar{p}\cdot(1-\bar{p})}\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}} = \\ &= \frac{0.375 - 0.060}{\sqrt{0.221\cdot(1-0.221)}\sqrt{\frac{1}{88}+\frac{1}{84}}} = \\ &= 4.98 \end{aligned}$$

Razina značajnosti $\alpha = 0.05$

Kritično područje: $z < z_{0.025} = -1.96$ ili $z > z_{0.925} = 1.96$.

$z = 4.98 > 1.96 = z_{0.925}$ pa hipotezu odbacujemo.

Postoji razlika između studenata dvaju sveučilišta.

Testiranje hipoteze o varijanci

Kod usporedbe srednjih vrijednosti dvije populcije u slučaju kada standardna devijacija populacije nije poznata odabir statistike ovisio je o jednakosti varijance dvije populacije.

Procjena varijance:

$$S^2 = \frac{1}{n-1} \sum_i \left(X_i - \bar{X}\right)^2 .$$

χ^2 distribucija

Neka su $X_1, X_2, \ldots, X_k$ nezavisne slučajne varijable s jediničnom normalnom distribucijom.

Tada je slučajna varijabla

$$Y = X_1^2 + X_2^2 + X_3^2 + \ldots + X_k^2$$

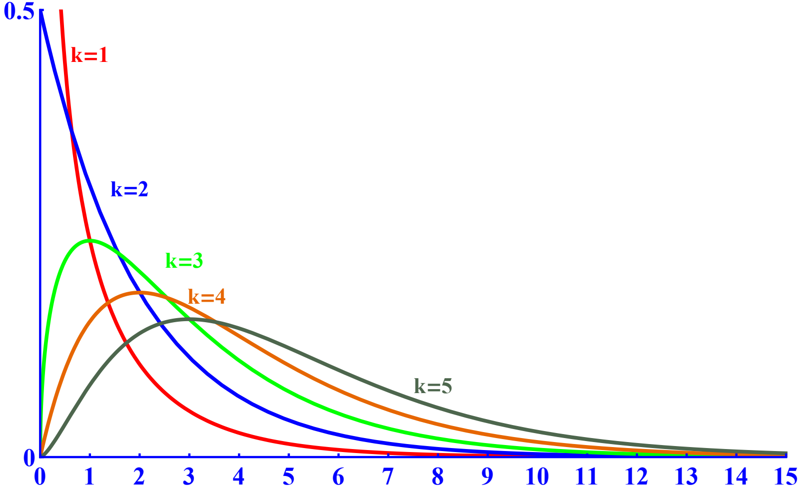
distribuirana prema χ^2 **distribuciji** s k **stupnjeva slobode**.

Očekivanje: $\mathrm{E}\left(Y\right) = k$

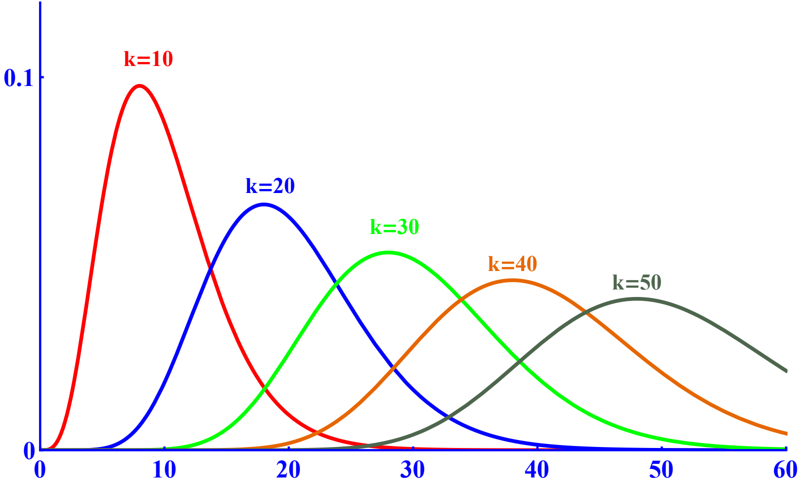
Varijanca: $\mathrm{Var}\left(Y\right) = 2 \cdot k$

Oznaka: $\chi^2(k)$

Funkcija gustoće za χ^2 distribuciju s k stupnjeva slobode ($k = 1,\ 2,\ 3,\ 4,\ 5$)



Funkcija gustoće za χ^2 distribuciju s k stupnjeva slobode ($k = 10, 20, 30, 40, 50$)



Ako je populacija normalna, varijabla

$$(n-1)\frac{S^2}{\sigma^2}$$

distribuirana je prema χ^2 distribuciji s $n-1$ stupnjeva slobode ($\chi^2(n-1)$).

Uočimo da χ^2 distribucija nije simetrična kao npr. normalna distribucija i Studentove distribucije.

Interval pouzdanosti za σ^2

$$P\left(\chi^2_{\alpha/2}(n-1) \leq (n-1)\frac{S^2}{\sigma^2} \leq \chi^2_{1-\alpha/2}(n-1)\right) = \alpha$$

$$P\left(\frac{\sigma^2}{n-1}\chi^2_{\alpha/2}(n-1) \leq S^2 \leq \frac{\sigma^2}{n-1}\chi^2_{1-\alpha/2}(n-1)\right) = \alpha$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right) = \alpha$$

$100(1-\alpha)\%$-tni interval pouzdanosti za S^2:

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right]$$

Testiranje hipoteze o varijanci - jedna populacija.

Hipoteza. Varijanca populacije jednaka je zadanom broju.

Nul hipoteza: $H_0 : \quad \sigma^2 = \sigma_0^2$

Alternativna hipoteza: $H_1 : \quad \sigma^2 \neq \sigma_0^2$

Zadamo razinu značajnosti α.

Statistika: $\chi^2 = \dfrac{S^2}{\sigma_0^2}(n-1)$

Distribucija statistike: $\chi^2(n-1)$.

Kritično područje: $\chi^2 < \chi^2_{\alpha/2}(n-1)$ ili $\chi^2 > \chi^2_{1-\alpha/2}(n-1)$

Jednostrani testovi provode se na isti način, jedino se mijenja kritično područje.

Za $H_0: \ \sigma^2 \leq \sigma_0^2$ kritično područje je $\chi^2 > \chi^2_{1-\alpha}(n-1)$

Za $H_0: \ \sigma^2 \geq \sigma_0^2$ kritično područje je $\chi^2 < \chi^2_{\alpha}(n-1)$

Primjer. Proizvođač satova želi nešto saznati o varijabilnosti svojih proizvoda. Da bi to ustanovio, odlučio je odrediti interval pouzdanosti za σ dobiven na osnovu slučajnog uzorka od 10 satova izabranih iz mnogo većeg broja satova koji su prošli završnu kontrolu. Odstupanje tih satova od standardnog sata sata je zabilježeno nakon jednog mjeseca i izračunate su sljedeće statistike:

$$\bar{X} = 0.7 \text{ s}, \qquad S = 0.4 \text{ s}.$$

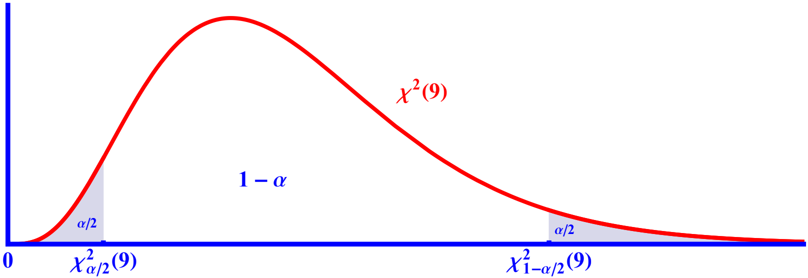
Pretpostavimo da razdioba mjerenja može biti dobro aproksimirana normalnom razdiobom.
Nađite 90%-tni interval pouzdanosti za σ.

Rješenje.

$n = 10 \quad \rightarrow \quad$ broj stupnjeva slobode=9.

$\alpha = 0.1$

Odredimo kritično područje.



R:
```
> qchisq(0.05,9)
3.325113
> qchisq(0.95,9)
16.91898
```

$$\begin{aligned}\chi^2_{\alpha/2}(n-1) &= \chi^2_{0.05}(n-1) = 3.325113,\\ \chi^2_{1-\alpha/2}(n-1) &= \chi^2_{0.95}(n-1) = 16.91898\end{aligned}$$

90%-tni interval pouzdanosti za σ^2:

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right] = \left[\frac{9 \cdot 0.16}{16.918978}, \frac{9 \cdot 0.16}{3.325113}\right] = [0.085, 0.433]$$

90%-tni interval pouzdanosti za σ:

$$\left[\sqrt{0.085}, \sqrt{0.433}\right] = [0.29, 0.66]$$

Uočite. $S = 0.4$ nije u sredini intervala pouzdanosti za standardnu devijaciju (kao niti $S^2 = 0.16$)

Razlog tome je nesimetričnost funkcije gustoće za χ^2 distribuciju.

Primjer. Nastavnik je uzorku studenata dao standardizirani test. Rezultati na testu su:

39	37	57	61	50	47	48	46
69	61	65	54	53	48	44	56

Testirajte hipotezu da je varijanca populacije veća od 100 uz razinu značajnosti $\alpha = 0.05$. Pretpostavljamo da je razdioba rezultata u populaciji normalna.

Rješenje.

Nul hipoteza: $H_0 \quad \sigma^2 \geq 100$

Alternativna hipoteza: $H_1 \quad \sigma^2 < 100$ (Jednostrani test.)

Statistika: $\chi^2 = \frac{S^2}{\sigma_0^2}(n-1) = \frac{S^2}{100}15$

Distribucija statistike: $\chi^2(15)$.

Kritično područje: $\chi^2 < \chi^2_\alpha(n-1) = \chi^2_{0.05}(15) = 7.26$

Računamo: $\bar{X} = 52.19$ i $S^2 = 81.36$.

Vrijednost statistike:

$$\chi^2 = \frac{S^2}{\sigma_0^2}(n-1) = \frac{S^2}{100}15 = \frac{81.36}{100}15 = 12.20 > 7.26.$$

Hipotezu ne odbacujemo.

Napomena. Testiranje hipoteze o standardnoj devijaciji je isto što i testiranje hipoteze o varijanci.

Hipoteza $H_0: \ \sigma = \sigma_0$ je ekvivalentna hipotezi $H_0: \ \sigma^2 = \sigma_0^2$.

Usporedba varijanci dvije populacije

Uzorak:

Prva populacija: $X_1, X_2, \ldots, X_{n_X}$

Druga populacija: $Y_1, Y_2, \ldots, Y_{n_Y}$

Procjena srednje vrijednosti:

$$\bar{X} = \frac{1}{n_X} \sum_i X_i, \quad \bar{Y} = \frac{1}{n_Y} \sum_i Y_i,$$

Procjena varijanci:

$$S_X^2 = \frac{1}{n_X - 1} \sum_i \left(X_i - \bar{X}\right)^2, \qquad S_Y^2 = \frac{1}{n_Y - 1} \sum_i \left(Y_i - \bar{Y}\right)^2$$

Populacija normalna i uzorak nezavisan:

$$\frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim \chi^2(n_X - 1), \qquad \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim \chi^2(n_y - 1)$$

F-distribucija

Ukoliko su X i Y nezavisne slučajne varijable distribuirane prema χ^2 distribuciji s m i n stupnjeva slobode:

$$X \sim \chi^2(n), \qquad Y \sim \chi^2(m),$$

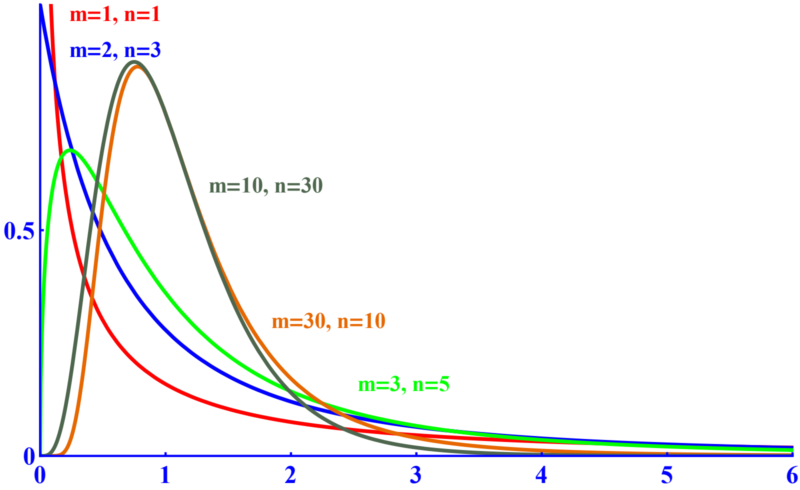
tada je slučajna varijabla

$$F = \frac{X/m}{Y/n}$$

Distribuirana prema F distribuciji s m i n stupnjeva slobode:

$$F \sim F(m, n).$$

Funkcija gustoće za F distribuciju s različitim brojem stupnjeva slobode



Za $F \sim F(m, n)$ je

$$\mathrm{E}\,(F) = \frac{n}{n-2} \quad \text{za} \quad n > 2$$

i

$$\mathrm{Var}\,(F) = \frac{2 \cdot n^2(n+m-2)}{m(n-2)^2(n-4)} \quad \text{za} \quad n > 4$$

$$\frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim \chi^2(n_X - 1), \qquad \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim \chi^2(n_y - 1)$$

$$\frac{\frac{\frac{(n_X-1)S_X^2}{\sigma_X^2}}{n_X-1}}{\frac{\frac{(n_Y-1)S_Y^2}{\sigma_Y^2}}{n_Y-1}} = \frac{S_X^2}{S_Y^2}\,\frac{\sigma_Y^2}{\sigma_X^2} \sim F(n_X - 1, n_Y - 1)$$

Ukoliko je $\sigma_X^2 = \sigma_Y^2$ tada je

$$\frac{S_X^2}{S_Y^2} \sim F(n_X - 1, n_Y - 1)$$

na osnovu ovoga definiramo statistiku za usporedbu varijanci.

Testiranje hipoteze o varijanci - dvije populacije.

Hipoteza. Varijance populacija su jednake.

Nul hipoteza: $H_0 : \quad \sigma_X^2 = \sigma_Y^2$

Alternativna hipoteza: $H_1 : \quad \sigma_X^2 \neq \sigma_Y^2$

Zadamo razinu značajnosti α.

Statistika: $f = \dfrac{S_X^2}{S_Y^2}$

Distribucija statistike: $f \sim F(n_X - 1, n_Y - 1)$ gdje je n_X veličina uzorka iz prve a n_Y veličina uzorka iz druge populacije.

Kritično područje: $f < F_{\alpha/2}(n_X - 1, n_Y - 1)$ ili $f > F_{1-\alpha/2}(n_X - 1, n_Y - 1)$.

Primjer. U studiji o razlici u apsolutnoj pogrešci kod pozicioniranja aktivne i pasivne ruke istraživači su testirali 20 studenata. Kod 10 studenata su mjerili apsolutnu pogrešku kod pozicioniranja aktivne ruke a kod preostalih 10 pogrešku kod pozicioniranja pasivne ruke. Rezultati mjerenja (u cm) prikazani su u tablici:

Ispitanik	Aktivna ruka	Ispitanik	Pasivna ruka
1	2.65	11	3.30
2	2.42	12	2.00
3	3.30	13	0.09
4	0.19	14	0.04
5	1.25	15	4.56
6	2.00	16	3.33
7	3.34	17	1.02
8	4.08	18	0.89
9	0.70	19	2.78
10	2.89	20	1.65

Postoji li razlika između aktivne i pasivne ruke u apsolutnoj pogrešci kod pozicioniranja?

Primjer. Dva nezavisna uzorka.

Razina značajnosti $\alpha = 0.05$

X - aktivna ruka

Y - pasivna ruka

Nul hipoteza: $H_0 : \quad \mu_X = \mu_Y$

Alternativna hipoteza: $H_1 : \quad \mu_X \neq \mu_Y$

$$\bar{X} = 2.282, \quad \bar{Y} = 1.966, \quad S_X^2 = 1.548, \quad S_Y^2 = 2.268$$

Test o jednakosti varijanci:

$$f = \frac{S_X^2}{S_Y^2} = \frac{1.548}{2.268} = 0.683$$

Kritično područje: $f < F_{0.025}(9,9) = 0.248$ ili $f > F_{0.925}(9,9) = 2.735$.

Hipotezu o jednakosti varijanci ne odbacujemo.

Koristimo test o usporedbi srednjih vrijednosti kada su varijance jednake.

Računanje statistike:

$$S_P^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2} = \frac{9 \cdot 1.548 + 9 \cdot 2.268}{10 + 10 - 2} = 1.908$$

$$t = \frac{\bar{X} - \bar{Y}}{S_P\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} = \frac{2.282 - 1.966}{\sqrt{1.908}\sqrt{\frac{1}{10} + \frac{1}{10}}} = 0.511$$

Distribucija statistike: Studentova (t) distribucija s $10 + 10 - 2 = 18$ stupnjeva slobode.

Kritično područje: $t < t_{0.025}(18) = -1.504$ ili $t > t_{0.975}(18) = 1.504$.

Jer je $t_{0.025}(18) = -1.504 < 0.511 < 1.504 = t_{0.975}(18)$ nul hipotezu ne odbacujemo.

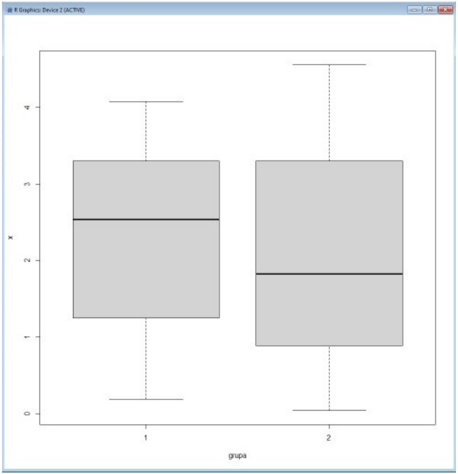
Rješavanje primjera u R-u.

Podaci:

```
> podaci
   grupa    x
1      1 2.65
2      1 2.42
3      1 3.30
4      1 0.19
5      1 1.25
6      1 2.00
7      1 3.34
8      1 4.08
9      1 0.70
10     1 2.89
11     2 3.30
12     2 2.00
13     2 0.09
14     2 0.04
15     2 4.56
16     2 3.33
17     2 1.02
18     2 0.89
19     2 2.78
20     2 1.65
```

Uvijek je dobro početi s grafom:

```
> boxplot(x ~ grupa, data = podaci)
```



Testiranje hipoteze o jednakosti varijanci:

```
> var.test(x ~ grupa, podaci,
alternative="two.sided")
```

```
F test to compare two variances

data:  x by grupa

F = 0.68269, num df = 9, denom df = 9,
```
p-value = 0.5787

```
alternative hypothesis:  true ratio of variances is
not equal to 1

95 percent confidence interval:

0.1695693 2.7484860

sample estimates:

ratio of variances

0.6826851
```

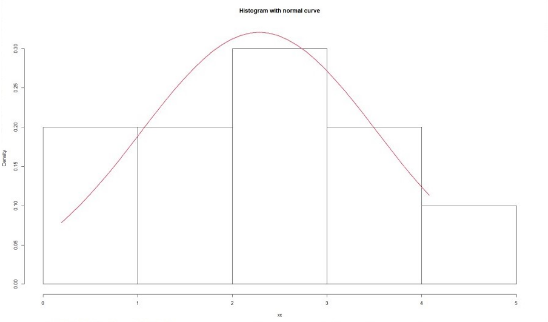
Testiranje hipoteze o jednakosti srednjih vrijednosti:

```
> t.test(x ~ grupa, podaci, alternative =
"two.sided", var.equal = TRUE)

Two Sample t-test

data:  x by grupa

t = 0.5115, df = 18, p-value = 0.6152

alternative hypothesis:  true difference in means
between group 1 and group 2 is not equal to 0

95 percent confidence interval:

-0.9819399 1.6139399

sample estimates:

mean in group 1 mean in group 2

2.282 1.966
```

Da su varijance bile različite:

```
> t.test(x ~ grupa, podaci, alternative =
"two.sided", var.equal = FALSE)

Welch Two Sample t-test

data:  x by grupa

t = 0.5115, df = 17.382, p-value = 0.6154

alternative hypothesis:  true difference in means
between group 1 and group 2 is not equal to 0

95 percent confidence interval:

-0.9852563 1.6172563

sample estimates:

mean in group 1 mean in group 2

2.282 1.966
```

Normalna distribucija - provjera.
Histogram



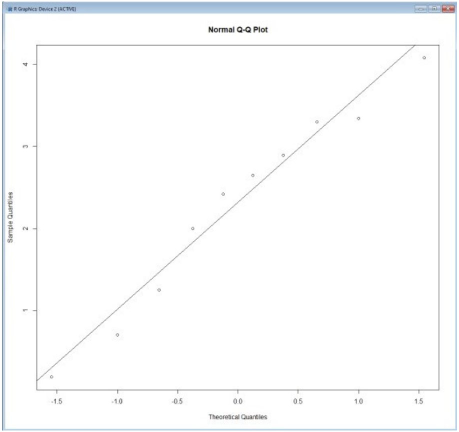
Usporedba distribucije s normalnom distribucijom (crvena krivulja).

R:

```
> xx=x[1:10,1]
> x2 <- seq(min(xx), max(xx), length = 40)
> fun <- dnorm(x2, mean = mean(xx), sd = sd(xx))
> hist(xx, prob = TRUE, col = "white",
> + ylim = c(0, max(fun)),
> + main = "Histogram with normal curve")
> lines(x2, fun, col = 2, lwd = 2)
```

Normalni vjerojatnosni graf

'Normal probabilistic plot', 'Q-Q plot'



Usporedba distribucije s normalnom distribucijom (pravac).

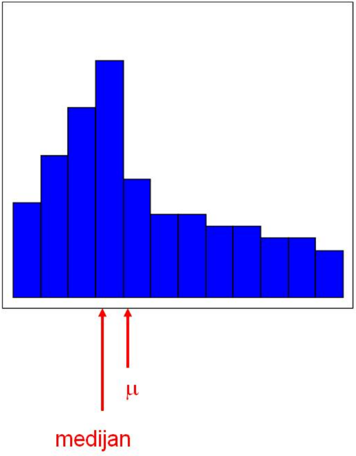
R:

```
> x<-subset(podaci,grupa==1,select="x")
> with(x,qqnorm(x);qqline(x))
```

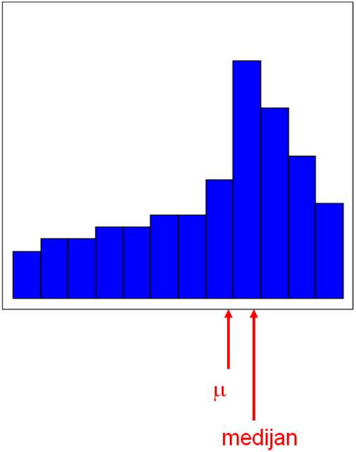
- Izabrali smo podatke za jednu grupu
- nacrtali graf

Zakošenost

Asimetrične razdiobe:



Pozitivno zakošena



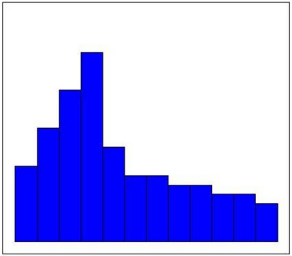
Negativno zakošena

Zakošenost

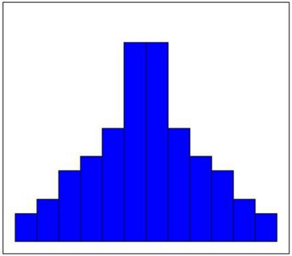
$$skew(X) = \sum_i \left(\frac{x_i - \mu}{\sigma} \right)^3$$

gdje je
μ - srednja vrijednost
σ - standardna devijacija

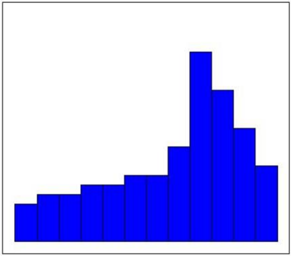
Primjer.



Zakošenost = 0.64



Zakošenost = 0



Zakošenost = -0.64

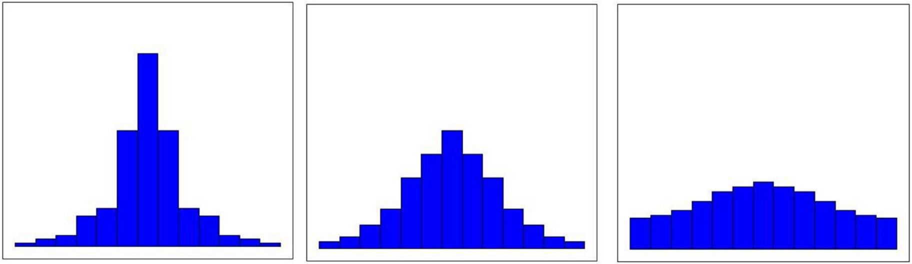
Oprez!

Zakošenost je dobar pokazatelj asimetrije za unimodalne distribucije.

- mala zakošenost: -0.5 do 0.5
- srednja zakošenost: između -1 i -0.5 te između 0.5 i 1
- velika zakošenost: manja od -1 ili veća od 1

Normalna distribucija: zakošenost = 0.

Spljoštenost

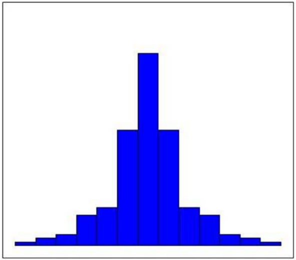


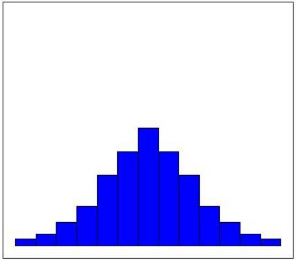
Koeficijent spljoštenosti

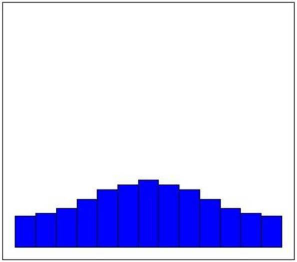
$$kurt(X) = \sum_i \left(\frac{x_i - \mu}{\sigma} \right)^4$$

gdje je

μ - srednja vrijednost
σ - standardna devijacija







$kurt(X) = 4.31$

$kurt(X) > 3$

Leptokurtična (izbočena) distribucija

$kurt(X) = 3$

$kurt(X) = 3$

Mezokurtična distribucija

$kurt(X) = 2.13$

$kurt(X) < 3$

Platikurtična (spljoštena) distribucija

Normalna distribucija: koeficijent spljoštenosti = 3.

Statistički paketi prikazuju korigirani koeficijent spljoštenosti
koeficijent spljoštenosti - 3

Normalna distribucija: korigirani koeficijent spljoštenosti = 0.

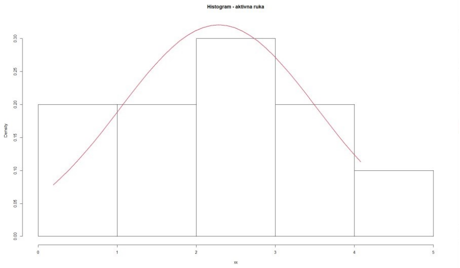
Distribucija podataka je slična normalnoj distribuciji ako je:

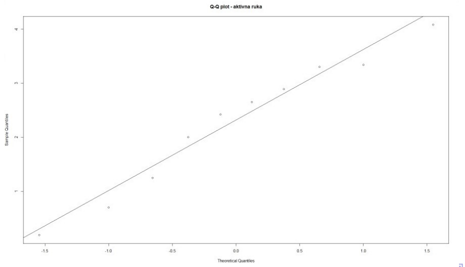
- koeficijens zakrivljenosti 'blizu' 0
- korigirani koeficijent spljoštenosti 'blizu' 0

Primjer. Aktivna i pasivna ruka.

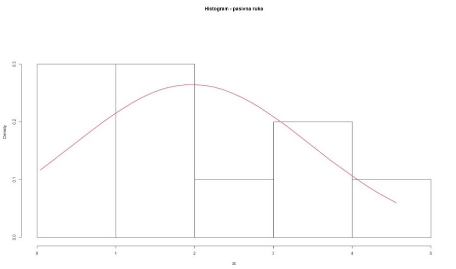
Analizu treba napraviti za svaku grupu posebno.
Posebno za aktivnu ruku i posebno za pasivnu ruku.

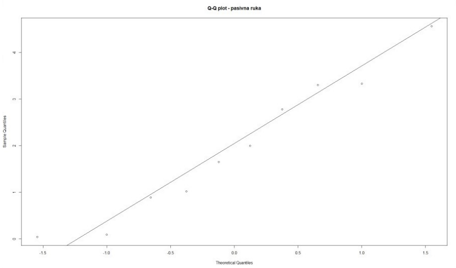
Aktivna ruka:





Pasivna ruka:





Asimetrija i spljoštenost.

```
> xaktivna = aktivna[1:10,1]
> xpasivna = pasivna [1:10,1]
> skewness(xaktivna, type = 2)
-0.4257934
> skewness(xpasivna, type = 2)
0.2860206
> kurtosis(xaktivna, type = 2)
-0.7447202
> kurtosis(xpasivna , type = 2)
-0.9183038
```

Usporedba više srednjih vrijednosti - ANOVA

Za testiranje hipoteze o jednakosti srednjih vrijednosti dvije populacije:

$$H_0 \quad \mu_X = \mu_Y$$

koristili smo t-test.

Kako usporediti više srednjih vrijednosti? Npr.

$$H_0 \quad \mu_1 = \mu_2 = \mu_3 \quad ?$$

Usporedba više srednjih vrijednosti se javlja u situaciji kada uspoređujemo više tretmana (metoda) i želimo ustanoviti jesu li one različite.

Možemo li hipotezu H_0 testirati tako da testiramo hipoteze

$$\begin{aligned} H_{0,1} &: \quad \mu_1 = \mu_2 \\ H_{0,2} &: \quad \mu_1 = \mu_3 \\ H_{0,3} &: \quad \mu_2 = \mu_3 \end{aligned}$$

Višestruko testiranje

Pretpostavimo da hipotezu

$$H_0 : \quad \mu_1 = \mu_2$$

nezavisno testiramo 5 puta.

(5 puta biramo uzorak i testiramo hipotezu.)

Nakon 5 testiranja hipotezu odbacujemo ukoliko smo je odbacili u barem jednom pojedinačnom testiranju.

Razina značajnosti u pojedinom testiranju je $\bar{\alpha}$.

Kolika je vjerojatnost da ćemo nakon 5 testiranja odbaciti hipotezu ukoliko je ona istinita?

Kolika je pogreška I. vrste za test koji se sastoji od 5 pojedinačnih testiranja?

Primjer. Na uzorku veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$.

Uzorak je generiran pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 42 puta (**4.2%**).

Ovo je u skladu s razinom značajnosti od **5%**.

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 222 puta (**22.2%**).

Ovo nije u skladu s razinom značajnosti od **5%**.

Označimo:

A = Prihvaćanje nul hipoteze

A_1 = Prihvaćanje nul hipoteze u 1. testiranju

A_2 = Prihvaćanje nul hipoteze u 2. testiranju

A_3 = Prihvaćanje nul hipoteze u 3. testiranju

A_4 = Prihvaćanje nul hipoteze u 4. testiranju

A_5 = Prihvaćanje nul hipoteze u 5. testiranju

Sada je

$$A = A_1 \textbf{ i } A_2 \textbf{ i } A_3 \textbf{ i } A_4 \textbf{ i } A_5$$

Uvjetna vjerojatnost uz uvjet da je H_0 istinita za ove događaje je

$$\begin{aligned} P(A \mid H_0) &= P(A_1 \text{ i } A_2 \text{ i } A_3 \text{ i } A_4 \text{ i } A_5 \mid H_0) = \\ &\textbf{nezavisni testovi} \\ &= P(A_1 \mid H_0) \cdot P(A_2 \mid H_0) \cdot P(A_3 \mid H_0) \cdot P(A_4 \mid H_0) \cdot P(A_5 \mid H_0) \end{aligned}$$

Oznake:

$\bar{\alpha}$ - Razina značajnosti u pojedinom testiranju.

α - Razina značajnosti ukupnog testa.

Dakle:

$$1 - \alpha = (1 - \bar{\alpha})^5$$

odnosno

$$\bar{\alpha} = 1 - \sqrt[5]{1 - \alpha}$$

Općenito, kod ponavljanja *m* testova,

$$1 - \alpha = (1 - \bar{\alpha})^m$$

i

$$\bar{\alpha} = 1 - \sqrt[m]{1 - \alpha}.$$

Jer je

$$\sqrt[m]{1 - \alpha} \approx 1 - \frac{\alpha}{m}$$

slijedi da je

$$\bar{\alpha} \approx \frac{\alpha}{m}.$$

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 222 puta (**22.2%**).

Ovo nije u skladu s razinom značajnosti od **5%**.

$$\alpha = 1 - (1 - \bar{\alpha})^m = 1 - (1 - 0.05)^5 = 0.226219 = 22.6\%$$

Ako su testovi potpuno zavisni, rezultat je drugačiji.

Npr., srednje vrijednosti uspoređujemo više varijabli na dva uzorka i varijable su zavisne. Uspoređujemo visinu danu u:

- metrima (X)
- centimetrima (Y)
- inčima (Z)

$$\begin{aligned} H_{0,1} &: \quad X_1 = X_2 \\ H_{0,2} &: \quad Y_1 = Y_2 \\ H_{0,3} &: \quad Z_1 = Z_2 \end{aligned}$$

Označimo:

A = Prihvaćanje nul hipoteze

A_1 = Prihvaćanje nul hipoteze u 1. testiranju ($H_{0,1}$)

A_2 = Prihvaćanje nul hipoteze u 2. testiranju ($H_{0,2}$)

A_3 = Prihvaćanje nul hipoteze u 3. testiranju ($H_{0,3}$)

Sada je

$$A = A_1 \text{ i } A_2 \text{ i } A_3 = A_1 = A_2 = A_3$$

i

$$P(A \mid H_0) = P(A_1 \text{ i } A_2 \text{ i } A_3 \mid H_0) = P(A_1 \mid H_0) = P(A_2 \mid H_0) = P(A_3 \mid H_0)$$

odnosno

$$1 - \alpha = 1 - \bar{\alpha} \quad \Leftrightarrow \quad \alpha = \bar{\alpha}$$

Korekcija treba biti između

- $1/m$ za nezavisne testove i
- 1 za potpuno zavisne.

Testiramo m hipoteza: $H_1,\ H_2,\ H_3,\dots,H_m$.

Generalna hipoteza:

H_0: Sve hipoteze $H_1,\ H_2,\ H_3,\dots,H_m$ su istinite.

H_a: Barem jedna od hipoteza $H_1,\ H_2,\ H_3,\dots,H_m$ nije istinita.

Oznake:

$\bar{\alpha}$ - Razina značajnosti u pojedinom testiranju hipoteza $H_1,\ H_2,\dots,H_m$.

α - Razina značajnosti ukupnog testa, testiranja hipoteze H_0.

Kako odrediti $\bar{\alpha}$ tako da vjerojatnost pogreške I. vrste bude manja od α?

'Family-wise error rate (FWER)' je vjerojatnost pojave jednog ili više lažnih otkrića, odnosno, pogrešaka I. vrste među svim hipotezama pri testiranju više hipoteza,

Bonferronijeva korekcija $\bar{\alpha} = \frac{\alpha}{m}$.

Šidákova korekcija $\bar{\alpha} = 1 - \sqrt[m]{1-\alpha}$.

Holmova korekcija

- $p_{(1)}, \ldots, p_{(m)}$ sortirane p-vrijednosti individualnih testova ($p_{(1)} \leq \ldots \leq p_{(m)}$).
  Neka su $H_1, \ldots, H_m$ odgovarajuće nul hipoteze.
- Holmova procedura je definirana korak po korak (engl. 'stepwise').
- 1. korak. Ako je $p_{(1)} \geq \alpha/m$, prihvati $H_1, \ldots, H_m$ i stani.
  Ako je $p_{(1)} < \alpha/m$, odbaci H_1 i testiraj ostalih $m-1$ hipoteza na razini $\alpha/(m-1)$.
- 2. korak. Ako je $p_{(1)} < \alpha/m$, a $p_{(2)} \geq \alpha/(m-1)$, prihvati $H_2, \ldots, H_m$ i stani. Ako je $p_{(1)} < \alpha/m$ i $p_{(2)} < \alpha/(m-1)$, odbaci H_2 uz H_1 i testiraj ostalih $m-2$ hipoteza na razini $\alpha/(m-2)$.
- ...

- Bonferronijeva (i Šidákova) korekcija je konzervativna.
- Premala pogreška I. vrste.
- Povećani broj prihvaćanja nul hipoteze.
- Veća pogreška II. vrste.
- Mala snaga testa.

Druge metode.

Tukeyeva procedura, Duncanova procedura. - Usporedba svih parova ($m(m-1)/2$ usporedbi).

Dunnettova procedura. - Usporedba kontrolne skupine s preostalim skupinama ($m-1$ usporedbi).

Umjesto FWER-a možemo kontrolirati:

FDR ('False Discovery Rate') - kontrolira se udio krivo odbačenih ('false positive') hipoteza unutar skupa odbačenih hipoteza.

- Benjamini-Hochberg procedura

- Benjamini-Hochberg-Yekutieli procedura

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu uz razinu značajnosti $\bar{\alpha} = 0.05/5 = 0.01$.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 52 puta (**5.2%**).

Ovo je u skladu s razinom značajnosti od **5%**.

ANOVA test

Za k **normalnih** populacija želimo provjerilti jesu li srednje vrijednosti obilježja u tim populacijama jednake.

Testiramo hipotezu

$$H_0 \quad \mu_1 = \mu_2 = \ldots = \mu_k.$$

Alternativna hipoteza:

$$H_a \quad \text{barem jedan par srednjih vrijednosti je različit}$$

Želimo izbjeći pristup pomoću viěstrukog uspoređivanja srednjih vrijednosti ($k(k-1)/2$ usporedbi).

Najčešće se radi o usporedbi više tretmana.

Svaka populacija odgovara pojedinom tretmanu.

Biramo

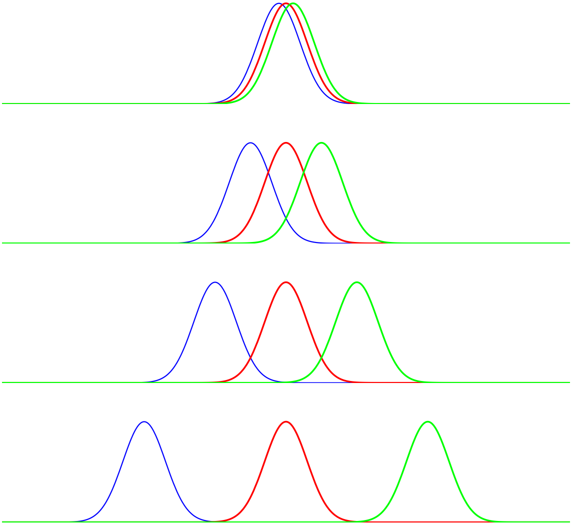
- k uzoraka iz k populacija ili
- k uzoraka iz jedne populacije i na svaki uzorak primijenimo drugi tretman.
- veličina uzoraka ne treba biti ista. Mi ćemo radi jednostavnosti pretpostaviti da su svi uzorci iste veličine (n, ukupno $k \cdot n$ podataka)

Važne pretpostavke na populacije:

- normalna distribucija obilježja u svakoj populaciji
- ista standardna devijacija (varijanca) obilježja u svih k populacija.

Distribucija za 3 (jednako velike) populacije.

Gdje je najveća ukupna varijanca?



ANOVA - Analysis of variance

U prvom koraku izračunamo srednje vrijednosti za svaki uzorak.

Tretman	Slučajni uzorak	Sr. vr. uzorka
Tretman 1	$X_{11}\ X_{12}\ X_{13}\ \ldots\ X_{1j}\ \ldots\ X_{1n}$	$\bar{X}_1$
Tretman 2	$X_{21}\ X_{22}\ X_{23}\ \ldots\ X_{2j}\ \ldots\ X_{2n}$	$\bar{X}_2$
$\vdots$	$\vdots\ \vdots\ \vdots\ \vdots\ \vdots$	$\vdots$
Tretman i	$X_{i1}\ X_{i2}\ X_{i3}\ \ldots\ X_{ij}\ \ldots\ X_{in}$	$\bar{X}_i$
$\vdots$	$\vdots\ \vdots\ \vdots\ \vdots\ \vdots$	$\vdots$
Tretman k	$X_{k1}\ X_{k2}\ X_{k3}\ \ldots\ X_{kj}\ \ldots\ X_{kn}$	$\bar{X}_k$

$$\bar{X}_i = \frac{1}{n}\sum_j X_{ij}$$ - srednja vrijednost uzorka

Ujedno izračunamo i srednju vrijednost cijelog uzorka ($k \cdot n$ podataka)

$$\bar{X} = \frac{1}{k \cdot n} \sum_i \sum_j X_{ij} = \frac{1}{k} \sum_i \bar{X}_i.$$

Suma kvadrata (sum of squares)

Osnova ANOVA testa je usporedba različitih procjena varijance σ^2.

Na osnovu svih uzoraka ($k \cdot n$ podataka) računamo ukupnu varijancu:

$$S^2 = \frac{1}{k \cdot n - 1} \sum_i \sum_j \left(X_{ij} - \bar{X}\right)^2.$$

Posebno ćemo označiti **ukupnu sumu kvadrata**

$$SS(T) = \sum_i \sum_j \left(X_{ij} - \bar{X}\right)^2.$$

$SS(T)$ - total sum of squares

Za svaki tretman (uzorak) izračunamo procjenu varijance

$$S_i^2 = \frac{1}{n-1} \sum_j \left(X_{ij} - \bar{X}_i\right)^2, \quad i = 1, \ldots, k.$$

Na osnovu ovih k procjena konstruiramo jednu zajedničku procjenu:

$$S_P^2 = \frac{1}{k} \sum_i S_i^2 = \frac{1}{k \cdot (n-1)} \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2.$$

Odgovarajuću sumu kvadrata nazivamo **suma kvadrata za pogreške**

$$SS(E) = \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2.$$

$SS(E)$ - sum of squares for errors

Na kraju računamo **sumu kvadrata za tretmane (faktore)**

$$SS(F) = \sum_i n\left(\bar{X}_i - \bar{X}\right)^2.$$

Može se pokazati da je

$$SS(T) = SS(E) + SS(F).$$

Ukoliko su uzorci iz normalne populacije, tada je

$$(k \cdot n - 1)\frac{S^2}{\sigma^2} \sim \chi^2(k \cdot n - 1)$$

tj.

$$\frac{SS(T)}{\sigma^2} \sim \chi^2(k \cdot n - 1).$$

Nadalje,

$$(n-1)\frac{S_i^2}{\sigma^2} \sim \chi^2(n-1), \quad i = 1, \ldots, k.$$

Kako su S_i^2 nezavisni (zbog nezavisnosti uzoraka), tada je

$$\begin{aligned}
(n-1)\frac{S_1^2}{\sigma^2} + (n-1)\frac{S_2^2}{\sigma^2} + \ldots + (n-1)\frac{S_k^2}{\sigma^2} &\sim \chi^2(k \cdot (n-1)) \\
k \cdot (n-1)\frac{S_P^2}{\sigma^2} &\sim \chi^2(k \cdot (n-1)) \\
\frac{SS(E)}{\sigma^2} &\sim \chi^2(k \cdot (n-1))
\end{aligned}$$

Što je sa $SS(F)$?

Može se pokazati da je

- $\dfrac{SS(F)}{\sigma^2} \sim \chi^2(k-1)$
- $SS(E)$ i $SS(F)$ su nezavisne.

Uz

$$\frac{SS(E)}{\sigma^2} \sim \chi^2(k \cdot (n-1))$$

dobijamo statistiku

$$F = \frac{\frac{SS(F)}{k-1}}{\frac{SS(E)}{k\cdot(n-1)}} = \frac{MS(F)}{MS(T)} \sim F(k-1, k \cdot (n-1))$$

Kritično područje za razinu značajnosti α je

$$F \geq F_{1-\alpha}(k-1, k \cdot (n-1)).$$

Rezultati se obično prikazuju pomoću tablice

Izvor varijacije	Suma kvadrata	Br. st. slobode	Srednja vr. sume kvadrata	Omjer F
Tretman (faktor)	$SS(F)$	$k-1$	$MS(F) = \frac{SS(F)}{k-1}$	$\frac{MS(F)}{MS(E)}$
Pogreška	$SS(E)$	$k \cdot (n-1)$	$MS(E) = \frac{SS(E)}{k \cdot (n-1)}$	
Ukupno	$SS(T)$	$k \cdot n - 1$		

Primjer. Trener želi usporediti tri različite metode treninga. Svaku od metoda primijenio je na po $n = 4$ studenta. Nakon 30 dana ocijenjena je uspješnost i ocjene su prikazane u tablici

Metoda	Observacije			
Metoda 1	3	6	4	7
Metoda 2	11	8	10	7
Metoda 3	6	9	5	8

Jesu li sve tri metode jednako uspješne? Hipotezu testirajte uz razinu značajnosti $\alpha = 0.05$.

Rješenje. 3 uzorka ($k = 3$) s po 4 ispitanika u svakom uzorku ($n = 4$).

Prvo računamo srednje vrijednosti uzoraka:

$$
\begin{aligned}
\bar{X}_1 &= \frac{3+6+4+7}{4} = \frac{20}{4} = 5 \\
\bar{X}_2 &= \frac{11+8+10+7}{4} = \frac{36}{4} = 9 \\
\bar{X}_3 &= \frac{6+9+5+8}{4} = \frac{24}{4} = 7 \\
\bar{X} &= \frac{20+36+28}{2} \left(= \frac{5+9+7}{3} \right) = \frac{84}{12} = 7
\end{aligned}
$$

Metoda	Observacije				Sr. vr.
Metoda 1	3	6	4	7	5
Metoda 2	11	8	10	7	9
Metoda 3	6	9	5	8	7
Ukupno					7

Zatim računamo sume kvadrata.

Suma kvadrata za faktore:

$$
\begin{aligned}
SS(F) &= \sum_i 4\left(\bar{X}_i - \bar{X}\right)^2 = \\
&= 4\cdot(5-7)^2 + 4\cdot(9-7)^2 + 4\cdot(7-7)^2 = \\
&= 16 + 16 + 0 = \mathbf{32}
\end{aligned}
$$

Suma kvadrata za pogreške:

$$
\begin{aligned}
SS(E) &= \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2 = \\
&= (3-5)^2 + (6-5)^2 + (4-5)^2 + (7-5)^2 + \\
&\quad (11-9)^2 + (8-9)^2 + (10-9)^2 + (7-9)^2 + \\
&\quad (6-7)^2 + (9-7)^2 + (5-7)^2 + (8-7)^2 = \\
&= 4+1+1+4+4+1+1+4+1+4+4+1 = \\
&= \mathbf{30}
\end{aligned}
$$

Za ilustraciju računamo i ukupnu sumu kvadrata:

$$
\begin{aligned}
SS(T) &= \sum_i \sum_j (X_{ij} - \bar{X})^2 = \\
&= (3-7)^2 + (6-7)^2 + (4-7)^2 + (7-7)^2 + \\
&\quad (11-7)^2 + (8-7)^2 + (10-7)^2 + (7-7)^2 + \\
&\quad (6-7)^2 + (9-7)^2 + (5-7)^2 + (8-7)^2 = \\
&= 16+1+9+0+16+1+9+0+1+4+4+1 = \\
&= \mathbf{62}
\end{aligned}
$$

Provjerimo

$$
\begin{aligned}
SS(T) &= SS(F) + SS(E) \\
62 &= 32 + 30
\end{aligned}
$$

Računamo srednje sume kvadrata:

$$
\begin{aligned}
MS(F) &= \frac{SS(F)}{k-1} = \frac{32}{2} = 16 \\
MS(E) &= \frac{SS(E)}{k \cdot (n-1)} = \frac{30}{3 \cdot (4-1)} = \frac{30}{9} = 3.333
\end{aligned}
$$

i statistiku

$$
F = \frac{MS(F)}{MS(E)} = \frac{16}{3.33} = 4.80
$$

Sve ove podatke pregledno prikazujemo u ANOVA tablici:

Izvor varijacije	**Suma kvadrata**	**Br. st. slobode**	**Srednja vr. sume kvadrata**	**Omjer *F***
Tretman	32	2	16	4.80
Pogreška	30	9	3.33	
Ukupno	62	11		

Pomoću kalkulatora izračunamo $F_{0.95}(2,9) = 4.26$.

Jer je

$$F = 4.80 > 4.26 = F_{0.95}(2,9)$$

hipotezu o jednakosti srednjih vrijednosti odbacujemo uz razinu značajnosti $\alpha = 0.95$.

Rješenje primjera pomoću R-a.

Podaci:

```
> podacianova
```

```
   ocjena  metoda
1       3      m1
2       6      m1
3       4      m1
4       7      m1
5      11      m2
6       8      m2
7      10      m2
8       7      m2
9       6      m3
10      9      m3
11      5      m3
12      8      m3
```

ANOVA test:

```
> aov(ocjena ~ metoda , data = podacianova )
Call:  aov(formula = ocjena ~ metoda, data =
podacianova)

Terms:

metoda Residuals

Sum of Squares 32 30

Deg.  of Freedom 2 9

Residual standard error:  1.825742

Estimated effects may be unbalanced
```

ANOVA tablica:

```
> rezultat <- aov(ocjena ~ metoda , data =
podacianova )
> summary(rezultat )
            Df Sum Sq Mean Sq F value  Pr(>F)
 metoda      2     32  16.000     4.8 0.0381  *
 Residuals   9     30   3.333
--
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
'.'  0.1 ' ' 1
```

ANOVA test nam je pokazao da su srednje vrijednosti različite, tj. da se metode razlikuju.

Koje od tih tri metoda se razlikuju?

Za odgovor na ovo pitanje treba napraviti međusobnu usporedbu svih metoda.

$\rightarrow$ višestruka usporedba.

Nije poželjno koristiti t-test za nezavisne uzorke s Bonferronijevom korekcijom.

Ukoliko ANOVA test pokaže razliku između tretmana, standardno se koriste sljedeći testovi za višestruku usporedbu ('post hoc' testovi):

- **Duncanov test** - kontrolira pogrešku II. vrste na štetu pogreške I. vrste.
- **Tukeyev test** - u principu se radi od t-testu uz kontrolu FWER-a.
- **Dunnettov test** - usporedba samo s kontrolnom grupom ($k-1$ usporedbi) za razliku od Tukeyovog testa gdje se sve grupe međusobno uspoređuju ($k \cdot (k-1)$ usporedbi).

Tukeyev 'post hoc' test:

```
> TukeyHSD(rezultat)

Tukey multiple comparisons of means

95% family-wise confidence level

Fit:  aov(formula = ocjena   metoda, data =
podacianova)

 $metoda  diff        lwr       upr     p adj
 m2-m1       4  0.3955363  7.604464  0.0310140
 m3-m1       2 -1.6044637  5.604464  0.3149865
 m3-m2      -2 -5.6044637  1.604464  0.3149865
```

Razlikuju se Metoda 1 i Metoda 2.

Razlikuju se Metoda 1 i Metoda 2.

Metoda	**Sr. vr.**
Metoda 1	5
Metoda 2	9
Metoda 3	7

ANOVA i 'post hoc' testovi pretpostavlja da

- uzorci su nezavisni
  to uvijek pretpostavljamo ukoliko nije drugačije napomenuto
- uzorci su iz normalne populacije
  Ova pretpostavka vrijedi za skoro sve do sada spomenute testove
- Standardna devijacija populacija je jednaka.
  Ova pretpostavka se u slučaju dvije populacije testira F-testom.
  Što u slučaju više populacija?

Levene-ov test - usporedba više varijanci

Za testiranje hipoteze o jednakosti varijanci k populacija:

$$H_0 : \quad \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$

koristi se **Levene-ov test**.

Homogenost varijanci = jednake varijance = homoscedastičnost

heteroscedastičnost - različite varijance

Paket 'car':

```
> install.packages(c("car"))
> library(car)
```

Levenov test:

```
> leveneTest(ocjena ~ metoda , data = podacianova )
Levene's Test for Homogeneity of Variance (center = median)
        Df F  value  Pr(>F)
 group     2       0        1
           9

Warning message:
In leveneTest.default(y = y, group = group, ...)  :
group coerced to factor.
```

Ukoliko se varijance razlikuju, ne možemo koristiti ANOVA test.

Rješenje: **Welchov test** - sve je isto kao u ANOVA testu jedino se koristi drugačija statistika

R:

```
> oneway.test(ocjena ~ metoda , data = podacianova,
var.equal = FALSE )

One-way analysis of means (not assuming equal
variances)

data:  ocjena and metoda

F = 4.32, num df = 2, denom df = 6, p-value =
0.06884
```

Napomena. Ako se uspoređuju dvije populacije, ANOVA i t-test daju isti rezultat.

Za dvije populacije (X_1 i X_2) je

$$F = \frac{\left(\bar{X}_1 - \bar{X}_2\right)^2}{\left(S_1^2 + S_2^2\right)/n}.$$

S druge strane, u t-testu za dvije populacije s jednakim varijancama, statistika je dana s

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(S_1^2 + S_2^2\right)/n}}.$$

Vrijedi:

$$F = t^2.$$