Testiranje statističkih hipoteza

Primjer

Želimo provjeriti pretpostavku da beton s čeličnim vlaknima ima veću tlačnu čvrstoću od betona bez čeličnih vlakana.
Pretpostavimo da smo dobili sljedeće vrijednosti tlačne čvrstoće (u N/mm^2) za beton s čeličnim vlaknima:

$$68, 70, 62, 65, 70, 66, 67,$$

a za beton bez čeličnih vlakana

$$63, 61, 59, 64, 64, 62, 66.$$

Vidi se da su vrijednosti tlačne čvrstoće za beton bez čeličnih vlakana uglavnom niže od onih za beton s čeličnim vlaknima.
Ako izračunamo aritmetičke sredine, dobijemo

$$\overline{x}_7^{(1)} = 66.8571 \quad \text{i} \quad \overline{x}_7^{(2)} = 62.1428.$$

No, ni ovo nije dovoljno da dođemo do željenog zaključka. Osim prosjeka, bitno je i koliko se podaci raspršuju (variraju)!

Primjer (nastavak)

- Razliku između tih prosjeka ćemo usporediti (podijeliti) s nekom mjerom raspršenja tih mjerenja, uzimajući u obzir i broj mjerenja.
- Uz dodatne pretpostavke, to će nam omogućiti da dobiveni rezultat usporedimo s nekim poznatim vrijednostima (koje se nalaze u statističkim tablicama) i zaključimo koliko je (mala) vjerojatnost da se naša razlika, na broju mjerenja te veličine, pojavi slučajno, iako su stvarni prosjeci jednaki.
- Ako je ta vjerojatnost dovoljno mala (npr. manja od 5%), odbacit ćemo hipotezu da su stvarna očekivanja jednaka i doći do željenog zaključka da beton s čeličnim vlaknima ima veću tlačnu čvrstoću.

U prethodnom primjeru zapravo testiramo statističku hipotezu, tj. provodimo statistički test!

Statističke hipoteze

- Neka je X slučajna varijabla koja modelira ishode slučajnog eksperimenta.
- **Statistička hipoteza** je bilo koja pretpostavka o raspodjeli od X.
- Statističke hipoteze označavamo s H_0 i H_1 te ih zovemo **nul-hipoteza** i **alternativna hipoteza**, redom. Hipoteza H_1 je alternativna (u nekom smislu suprotna) hipotezi H_0.
- **Neparametarska hipoteza** je pretpostavka o raspodjeli od X. Npr. H_0: X ima normalnu raspodjelu.
- **Parametarska hipoteza** je pretpostavka o parametrima od X. Npr. $H_0 : \mu = \mu_0$, gdje je μ očekivanje od X, a μ_0 fiksan broj.

Testiranje statističkih hipoteza

- **Testiranje statističkih hipoteza** je postupak donošenja odluke o odbacivanju ili neodbacivanju H_0 na osnovu informacije dobivene iz opažanja slučajnog uzorka.
- U slučaju neodbacivanja H_0 ne možemo tvrditi da je H_0 točna, dok u slučaju odbacivanja H_0 tvrdimo da je H_1 točna.
- Statistika pomoću koje se donosi odluka o odbacivanju ili neodbacivanju H_0 zove se **test-statistika**.
- Niti jedna odluka bazirana na uzorcima nije 100% pouzdana, pa ni zaključci statističkog testa nisu 100% pouzdani.
- Može se dogoditi da je odluka donesena temeljem statističkog testa pogrešna.

Pogreške prilikom testiranja

- Test-statistikom želimo odrediti granice područja odbacivanja i neodbacivanja H_0, tj. **kritičnog područja testa** (c_α).
- Odluku o odbacivanju ili neodbacivanju H_0 donosimo na temelju pripadanja ili nepripadanja vrijednosti test-statistike kritičnom području.
- U tom postupku javljaju se moguće pogreške koje su prikazane u tablici.

Stanje \ Odluka	H_0 se ne odbacuje	H_0 se odbacuje
H_0 istinita	ispravna odluka	α – pogreška 1. tipa
H_0 lažna	β – pogreška 2. tipa	ispravna odluka

Tablica: Mogući zaključci testiranja

- Broj $\alpha \in (0, 1)$ predstavlja maksimalnu pogrešku koju činimo kada odbacujemo istinitu H_0 - **pogreška 1. tipa**:

$$\alpha = \mathbb{P}(H_0 \text{ se odbacuje} | H_0 \text{ je istinita}).$$

  Navedenu pogrešku zovemo **nivo značajnosti**.

- Broj $\beta \in (0, 1)$ predstavlja maksimalnu pogrešku koju činimo kada ne odbacujemo H_0, a istinita je H_1 - **pogreška 2. tipa**:

$$\beta = \mathbb{P}(H_0 \text{ se ne odbacuje} | H_0 \text{ nije istinita}).$$

  Broj $1 - \beta$ naziva se **jakost testa**.

- Razumno je zahtjevati test kojim se mogu kontrolirati (smanjiti) obje pogreške. Međutim, to nije moguće jer ako se α smanji, onda se smanjuje kritično područje testa, što rezultira povećanjem β.

- U praksi se prije provođenja testa izabere značajnost (najčešće $\alpha = 0.01$ ili 0.05), odredi se kritično područje, pa se naknadno izračuna β. Ako je β prevelik, test ponavljamo s većim uzorkom.

Postupak testiranja

(1) Postaviti test koji želimo provesti i odrediti nivo značajnosti α.

(2) Ovisno o testu i nekim svojstima uzorka (poznata/nepoznata varijanca σ^2, mali/veliki uzorak) odrediti kritično područje c_α.

(3) Odrediti vrijednost test-statistike t.

(4) Zaključak se izvodi na sljedeći način:

- $t \in c_\alpha \implies$ odbacujemo nul-hipotezu H_0
- $t \notin c_\alpha \implies$ ne odbacujemo nul-hipotezu H_0

Testiranje statističkih hipoteza za μ

Uzorak iz normalne raspodjele

Neka je $(X_1, \ldots, X_n)$ slučajni uzorak iz $N(\mu, \sigma^2)$. Kao i za intervale pouzdanosti razlikujemo dva slučaja:

(1) varijanca populacije σ^2 je poznata.

(2) varijanca populacije σ^2 nije poznata.

U praksi je slučaj (2) puno češći.

Provodit ćemo tri tipa testova:

(1) dvostrani test

(2) lijevi jednostrani test

(3) desni jednostrani test

σ^2 je poznata - dvostrani test

- **Dvostrani test** uz nivo značajnosti $\alpha \in (0,1)$. Testiramo:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu \neq \mu_0. \end{aligned}$$

- U ovom slučaju test-statistika je

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma} \sim N(0,1)},$$

a kritično područje je

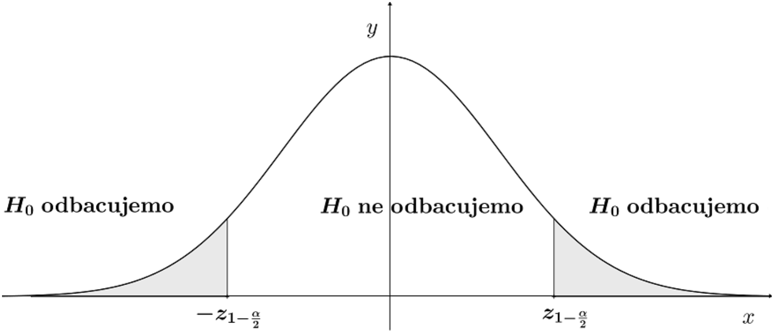
$$c_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty),$$

gdje je $z_{1-\frac{\alpha}{2}}$, $(1-\alpha/2)$-ti kvantil od $N(0,1)$.

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma}$$

za realizaciju $(x_1, \ldots, x_n)$ pripada kritičnom području c_α.



Slika: Kritično područje kod dvostranog testa

- Odbacivanjem H_0 na nivou značajnosti α zaključujemo da je razlika između realizacije od $\overline{X}_n$ i μ_0 statistički značajna.
- Neodbacivanjem H_0 zaključujemo da razlika nije statistički značajna.
- $\alpha \cdot 100\%$ realizacija od $(X_1, \ldots, X_n)$ uzrokuje odbacivanje H_0 ako je H_0 istinita. (U $\alpha \cdot 100\%$ slučajeva radimo pogrešku 1. tipa.)

σ^2 je poznata - lijevi jednostrani test

- **Lijevi jednostrani test** uz nivo značajnosti $\alpha \in (0,1)$. Testiramo:

$$\begin{aligned} H_0 &: \mu \geqslant \mu_0 \\ H_1 &: \mu < \mu_0. \end{aligned}$$

- Test-statistika je ponovno

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma} \sim N(0,1)},$$

a kritično područje je

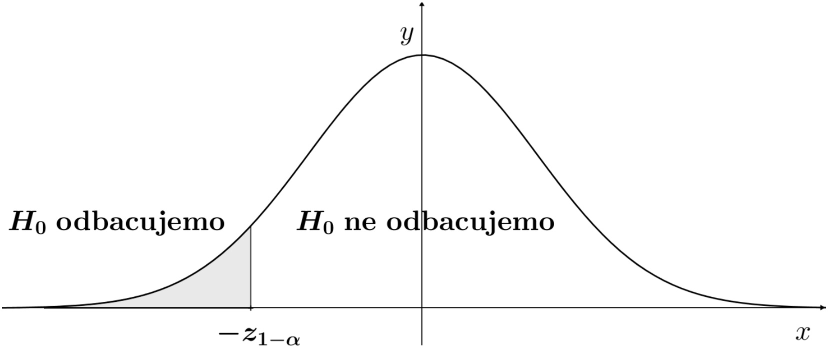
$$c_\alpha = (-\infty, -z_{1-\alpha}],$$

gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma}$$

za realizaciju $(x_1, \ldots, x_n)$ pripada kritičnom području c_α.



Slika: Kritično područje kod lijevog jednostranog testa

- Odbacivanjem H_0 na nivou značajnosti α zaključujemo da je realizacija od $\overline{X}_n$ (statistički značajno) manja od μ_0.
- Opet $\alpha \cdot 100\%$ realizacija od $(X_1, \ldots, X_n)$ uzrokuje odbacivanje H_0 ako je H_0 istinita.

σ^2 je poznata - desni jednostrani test

- **Desni jednostrani test** uz nivo značajnosti $\alpha \in (0, 1)$. Testiramo:

$$H_0 \; : \; \mu \leqslant \mu_0$$
$$H_1 \; : \; \mu > \mu_0.$$

- Test-statistika je ponovno

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma} \sim N(0, 1)},$$

  a kritično područje je

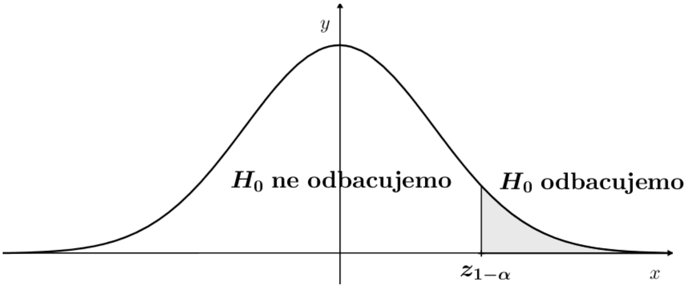
$$c_\alpha = [z_{1-\alpha}, \infty),$$

  gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0, 1)$.

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma}$$

  za realizaciju $(x_1, \ldots, x_n)$ pripada kritičnom području c_α.



Slika: Kritično područje kod desnog jednostranog testa

- Odbacivanjem H_0 na nivou značajnosti α zaključujemo da je realizacija od $\overline{X}_n$ (statistički značajno) veća od μ_0.
- Opet $\alpha \cdot 100\%$ realizacija od $(X_1, \ldots, X_n)$ uzrokuje odbacivanje H_0 ako je H_0 istinita.

σ^2 nije poznata

- Kad σ^2 nije poznata procedura je analogna kao u slučaju poznate varijance, samo test-statistiku zamijenimo s

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{S_n} \sim t(n-1)},$$

a vrijednost test-statistike s

$$t = \sqrt{n}\frac{\overline{x}_n - \mu_0}{s_n}$$

za realizaciju $(x_1, \ldots, x_n)$ slučajnog uzorka $(X_1, \ldots, X_n)$.

- Kvantile koji se javljaju za određivanje kritičnog područja određujemo iz $t(n-1)$ raspodjele. Dakle, $z_{1-\frac{\alpha}{2}}$ i $z_{1-\alpha}$ zamijenjujemo s $t_{1-\frac{\alpha}{2}(n-1)}$ i $t_{1-\alpha}(n-1)$.

σ^2 nije poznata

Kritična područja su

(1) za dvostrani test:

$$c_\alpha = (-\infty, -t_{1-\frac{\alpha}{2}}(n-1)] \cup [t_{1-\frac{\alpha}{2}}(n-1), \infty)$$

(2) za lijevi jednostrani test:

$$c_\alpha = (-\infty, -t_{1-\alpha}(n-1)]$$

(3) za desni jednostrani test:

$$c_\alpha = [t_{1-\alpha}(n-1), \infty)$$

Veliki uzorak ($n > 30$)

Imamo veliki (u praksi $n > 30$) slučajni uzorak $(X_1, \ldots, X_n)$ iz raspodjele koja **ne mora biti normalna**.

Opet ćemo promatrati dva slučaja:

(1) varijanca populacije σ^2 je poznata.

(2) varijanca populacije σ^2 nije poznata.

σ^2 je poznata - dvostrani test

- **Dvostrani test** uz nivo značajnosti $\alpha \in (0,1)$. Testiramo:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu \neq \mu_0. \end{aligned}$$

- U ovom slučaju test-statistika je

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma}},$$

  a kritično područje je dano s

$$c_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty),$$

  gdje je $z_{1-\frac{\alpha}{2}}$ $(1-\alpha/2)$-ti kvantil od $N(0,1)$.

- Koristimo činjenicu da za velike n raspodjelu od T možemo aproksimirati s $N(0,1)$ (centralni granični teorem)!

- Vrijednost test-statistike je $t = \sqrt{n}(\overline{x}_n - \mu_0)/\sigma$.

σ^2 je poznata - lijevi jednostrani test

- **Lijevi jednostrani test** uz nivo značajnosti $\alpha \in (0,1)$. Testiramo:

$$\begin{aligned} H_0 &: \mu \geqslant \mu_0 \\ H_1 &: \mu < \mu_0 \end{aligned}$$

- Test-statistika je ponovno

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma}},$$

a kritično područje je dano s

$$c_\alpha = (-\infty, -z_{1-\alpha}],$$

gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

- Vrijednost test-statistike je $t = \sqrt{n}(\overline{x}_n - \mu_0)/\sigma$.

σ^2 je poznata - desni jednostrani test

- **Desni jednostrani test** uz nivo značajnosti $\alpha \in (0,1)$. Testiramo:

$$\begin{aligned} H_0 &: \mu \leqslant \mu_0 \\ H_1 &: \mu > \mu_0 \end{aligned}$$

- Test-statistika je ponovno

$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma}},$$

a kritično područje je dano s

$$c_\alpha = [z_{1-\alpha}, \infty),$$

gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

- Vrijednost test-statistike je $t = \sqrt{n}(\overline{x}_n - \mu_0)/\sigma$.

σ^2 nije poznata

- I kad varijanca nije poznata, opet je procedura analogna kao u slučaju poznate varijance samo test-statistiku zamijenimo s

$$T = \sqrt{n}(\overline{X}_n - \mu_0)/S_n$$

  i vrijednost test-statistike s

$$t = \sqrt{n}(\overline{x}_n - \mu_0)/s_n.$$

- Primjenom centralnog graničnog teorema može se pokazati da za velike n raspodjelu od T možemo aproksimirati s $N(0,1)$. Dakle, kvantile koji određuju kritično područje ponovno određujemo iz $N(0,1)$.

Testiranje parametra Bernoullijeve slučajne varijable

- Koristeći navedene rezultate, primjenom centralnog graničnog teorema, možemo testirati hipoteze o parametru Bernoullijeve slučajne varijable
$$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

- Neka je $(X_1, \ldots, X_n)$ veliki slučajni uzorak za X. Kako je $\mathbb{E}(X) = p$ i $\mathrm{Var}(X) = p(1-p)$, za test-statistiku uzimamo
$$\boxed{T = \sqrt{n}\frac{\overline{X}_n - p_0}{S_n}}.$$

- Uočimo da vrijedi
$$S_n^2 = \frac{n\overline{X}_n(1-\overline{X}_n)}{n-1} \qquad \text{i} \qquad s_n^2 = \frac{n\overline{x}_n(1-\overline{x}_n)}{n-1}$$
pa je
$$T = \sqrt{n-1}\frac{\overline{X}_n - p_0}{\sqrt{\overline{X}_n(1-\overline{X}_n)}} \qquad \text{i} \qquad t = \sqrt{n-1}\frac{\overline{x}_n - p_0}{\sqrt{\overline{x}_n(1-\overline{x}_n)}}.$$

Sada možemo testirati hipoteze:

(i) Dvostrani test

$$\begin{aligned} H_0 &: p = p_0 \\ H_1 &: p \neq p_0. \end{aligned}$$

(ii) Lijevi jednostrani test

$$\begin{aligned} H_0 &: p \geqslant p_0 \\ H_1 &: p < p_0. \end{aligned}$$

(iii) Desni jednostrani test

$$\begin{aligned} H_0 &: p \leqslant p_0 \\ H_1 &: p > p_0. \end{aligned}$$

Primjer (8.5.)

Bacamo novčić 100 puta i bilježimo rezultat. Kao rezultat dobivamo 60 glava i 40 pisama. Sumnjamo da je novčić neispravan i želimo testirati tu pretpostavku.

Danu situaciju modeliramo Bernoullijevom raspodjelom

$$\begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix},$$

gdje 0 označava da je prilikom bacanja novčića palo pismo, a 1 da je pala glava. Testiramo

$$H_0 \;:\; p = \frac{1}{2}$$
$$H_1 \;:\; p \neq \frac{1}{2}$$

uz nivo značajnosti $\alpha = 0.05$. Na osnovu rezultata (opažanja) imamo $\overline{x}_{100} = 0.6$.

Primjer (nastavak)

Test-statistika je

$$T = \sqrt{99}\frac{\overline{X}_{100} - 1/2}{\sqrt{\overline{X}_{100}(1 - \overline{X}_{100})}},$$

a kritično područje

$$(-\infty, -z_{0.975}] \cup [z_{0.975}, \infty) = (-\infty, -1.96] \cup [1.96, \infty),$$

gdje je $z_{0.975}$ 0.975-ti kvantil od $N(0, 1)$.

Vrijednost test-statistike je $t = 2.031$.

Očito t pripada kritičnom području, pa odbacujemo H_0 s mogućnošću pogreške od 5%.

Drugim riječima, na osnovu gornjeg testa zaključujemo da novčić nije simetričan, s mogućnošću pogreške od 5%.

Uspoređivanje očekivanja dviju slučajnih varijabli

- Želimo ispitati razlikuje li se neko statističko obilježje u dvije različite populacije.
- Preciznije, imamo dvije slučajne varijable $X^{(1)}$ i $X^{(2)}$, njihova očekivanja μ_1 i μ_2 te varijance σ_1^2 i σ_2^2. Zanima nas postoji li razlika u očekivanjima ove dvije slučajne varijable.
- Koristit ćemo **t-test** ili **z-test za dva uzorka**.

Uzorci iz normalne raspodjele

Neka su $(X_1^{(1)}, \ldots, X_{n_1}^{(1)})$ i $(X_1^{(2)}, \ldots, X_{n_2}^{(2)})$ slučajni uzorci (koji su međusobno nezavisni) iz, redom, $N(\mu_1, \sigma_1^2)$ i $N(\mu_2, \sigma_2^2)$.

Promatramo 2 slučaja:

(1) varijance populacije σ_1^2 i σ_2^2 su **poznate**.

(2) vrijedi $\sigma_1^2 = \sigma_2^2 = \sigma^2$ i da σ^2 **nije poznata**.

σ_1^2 i σ_2^2 su poznate

- Promatramo **dvostrani test** uz nivo značajnost $\alpha \in (0,1)$. Testiramo:

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

- U ovom slučaju test-statistika je

$$\boxed{T = \frac{\overline{X}_{n_1}^{(1)} - \overline{X}_{n_2}^{(2)}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)}.$$

- Kritično područje je dano s

$$(-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty),$$

gdje je $z_{1-\frac{\alpha}{2}}$ $(1-\alpha/2)$-ti kvantili od $N(0,1)$.

σ_1^2 i σ_2^2 su poznate

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \frac{\overline{x}_{n_1}^{(1)} - \overline{x}_{n_2}^{(2)}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

  za realizacije $(x_1^{(1)}, \ldots, x_{n_1}^{(1)})$ i $(x_1^{(2)}, \ldots, x_{n_2}^{(2)})$ pripada kritičnom području.
  Ako t ne pripada kritičnom području, onda ne odbacujemo H_0.
- Možemo provoditi i lijevi i desni jednostrani test, pri čemu kvantil $z_{1-\frac{\alpha}{2}}$ zamijenjujemo s $z_{1-\alpha}$, gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ i σ^2 nije poznata

- Promatramo **dvostrani test** uz nivo značajnost $\alpha \in (0,1)$. Testiramo:

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

- U ovom slučaju test-statistika je

$$\boxed{T = \frac{\overline{X}_{n_1}^{(1)} - \overline{X}_{n_2}^{(2)}}{S_D\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2),}$$

gdje je

$$S_D = \sqrt{\frac{(n_1 - 1)(S_{n_1}^{(1)})^2 + (n_2 - 1)(S_{n_2}^{(2)})^2}{n_1 + n_2 - 2}}.$$

- Kritično područje je dano s

$$(-\infty, -t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)] \cup [t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2), \infty),$$

gdje je $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$ $(1 - \alpha/2)$-ti kvantili od $t(n_1 + n_2 - 2)$.

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ i σ^2 nije poznata

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \frac{\overline{x}_{n_1}^{(1)} - \overline{x}_{n_2}^{(2)}}{s_D\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

pripada kritičnom području. Ovdje je

$$s_D = \sqrt{\frac{(n_1 - 1)(s_{n_1}^{(1)})^2 + (n_2 - 1)(s_{n_2}^{(2)})^2}{n_1 + n_2 - 2}}.$$

- Za lijevi i desni jednostrani test $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$ zamijenjujemo s $t_{1-\alpha}(n_1 + n_2 - 2)$, gdje je $t_{1-\alpha}(n_1 + n_2 - 2)$ $(1 - \alpha)$-ti kvantil od $t(n_1 + n_2 - 2)$.

Veliki uzorci ($n_1, n_2 > 30$)

Imamo velike (u praksi $n_1, n_2 > 30$) slučajne uzorke $(X_1^{(1)}, \ldots, X_{n_1}^{(1)})$ i $(X_1^{(2)}, \ldots, X_{n_2}^{(2)})$ (koji su međusobno nezavisni) iz raspodjela koje **ne moraju biti normalne**.

Opet ćemo promatrati 2 slučaja:

(1) varijance populacije σ_1^2 i σ_2^2 su **poznate**.

(2) varijance populacije σ_1^2 i σ_2^2 **nisu poznate**.

σ_1^2 i σ_2^2 su poznate

- Radimo **dvostrani test** uz nivo značajnost $\alpha \in (0, 1)$. Testiramo:

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

- Test-statistika je sad

$$\boxed{T = \frac{\overline{X}_{n_1}^{(1)} - \overline{X}_{n_2}^{(2)}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}.$$

- Iz centralnog graničnog teorema možemo zaključiti da zbog veličine uzoraka raspodjelu od T možemo aproksimirati $N(0, 1)$.

- Kritično područje je dano s

$$(-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty),$$

gdje je $z_{1-\frac{\alpha}{2}}$ $(1 - \alpha/2)$-ti kvantili od $N(0, 1)$.

σ_1^2 i σ_2^2 su poznate

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \frac{\overline{x}_{n_1}^{(1)} - \overline{x}_{n_2}^{(2)}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

  za realizacije $(x_1^{(1)}, \ldots, x_{n_1}^{(1)})$ i $(x_1^{(2)}, \ldots, x_{n_2}^{(2)})$ pripada kritičnom području.
  Ako t ne pripada kritičnom području, onda ne odbacujemo H_0.

- Možemo provoditi i lijevi i desni jednostrani test, pri čemu kvantil $z_{1-\frac{\alpha}{2}}$ zamijenjujemo s $z_{1-\alpha}$, gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

σ_1^2 i σ_2^2 nisu poznate

- Radimo **dvostrani test** uz nivo značajnost $\alpha \in (0,1)$. Testiramo:

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

- U ovom slučaju test-statistika je

$$\boxed{T = \frac{\overline{X}_{n_1}^{(1)} - \overline{X}_{n_2}^{(2)}}{\sqrt{\frac{(S_{n_1}^{(1)})^2}{n_1} + \frac{(S_{n_2}^{(2)})^2}{n_2}}}}.$$

- Ponovno, zbog veličine uzoraka, iz centralnog graničnog teorema možemo zaključiti da raspodjelu od T možemo aproksimirati s $N(0,1)$.

- Kritično područje je dano s

$$(-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty),$$

gdje je $z_{1-\frac{\alpha}{2}}$ $(1-\alpha/2)$-ti kvantil od $N(0,1)$.

σ_1^2 i σ_2^2 nisu poznate

- Nul-hipotezu H_0 odbacujemo ako vrijednost test-statistike

$$t = \frac{\overline{x}_{n_1}^{(1)} - \overline{x}_{n_2}^{(2)}}{\sqrt{\frac{(s_{n_1}^{(1)})^2}{n_1} + \frac{(s_{n_2}^{(2)})^2}{n_2}}}$$

  za realizacije $(x_1^{(1)}, \ldots, x_{n_1}^{(1)})$ i $(x_1^{(2)}, \ldots, x_{n_2}^{(2)})$ pripada kritičnom području.
  Ako t ne pripada kritičnom području, onda ne odbacujemo H_0.
- Možemo provoditi i lijevi i desni jednostrani test, pri čemu kvantil $z_{1-\frac{\alpha}{2}}$ zamijenjujemo s $z_{1-\alpha}$, gdje je $z_{1-\alpha}$ $(1-\alpha)$-ti kvantil od $N(0,1)$.

Razlika parametara Bernoullijevih raspodjela

- Koristeći navedene rezultate, primjenom centralnog graničnog teorema, možemo testirati hipoteze o razlici parametara Bernoullijevih raspodjela

$$\begin{pmatrix} 0 & 1 \\ 1-p_1 & p_1 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 0 & 1 \\ 1-p_2 & p_2 \end{pmatrix}.$$

- Neka su $(X_1^{(1)}, \ldots, X_{n_1}^{(1)})$ i $(X_1^{(2)}, \ldots, X_{n_2}^{(2)})$, redom, veliki slučajni uzorci iz gornjih raspodjela.
- Kako je $\mathbb{E}(X_i^{(1)}) = p_1$, $\mathrm{Var}(X_i^{(1)}) = p_1(1-p_1)$, $\mathbb{E}(X_i^{(2)}) = p_2$ i $\mathrm{Var}(X_i^{(2)}) = p_2(1-p_2)$, za test-statistiku uzimamo

$$T = \frac{\overline{X}_{n_1}^{(1)} - \overline{X}_{n_2}^{(2)}}{\sqrt{\frac{(S_{n_1}^{(1)})^2}{n_1} + \frac{(S_{n_1}^{(2)})^2}{n_2}}}.$$

Razlika parametara Bernoullijevih raspodjela

Sada možemo testirati hipoteze:

(i) Dvostrani test

$$\begin{aligned} H_0 &: p_1 = p_2 \\ H_1 &: p_1 \neq p_2. \end{aligned}$$

(ii) Lijevi jednostrani test

$$\begin{aligned} H_0 &: p_1 \geqslant p_2 \\ H_1 &: p_1 < p_2. \end{aligned}$$

(iii) Desni jednostrani test

$$\begin{aligned} H_0 &: p_1 \leqslant p_2 \\ H_1 &: p_1 > p_2. \end{aligned}$$

Primjer (8.7.)

Vratimo se na primjer u kojem su izmjerene tlačne čvrstoće betona s čeličnim vlaknima i betona bez čeličnih vlakana. Podaci su prikazani u tablici.

beton s čeličnim vlaknima	68	70	62	65	70	66	67
beton bez čeličnih vlakana	63	61	59	64	62	66	60

Tablica: Podaci o tlačnoj čvrstoći različitih vrsta betona

Pretpostavimo da su tlačne čvrstoće koje mjerimo normalno distribuirane s raspodjelama $N(\mu_1, \sigma^2)$ (za beton s čeličnim vlaknima) i $N(\mu_2, \sigma^2)$ (za beton bez čeličnih vlakana).

Na osnovu gornjih mjerenja slijedi

$$\overline{x}_7^{(1)} = 66.8571,\ s_7^{(1)} = 2.8535,\ \overline{x}_7^{(2)} = 62.1428 \quad \text{i} \quad s_7^{(2)} = 2.4103.$$

Primjer (nastavak)

Imamo male uzorke iz normalnih raspodjela s jednakom varijancom koja nije poznata, pa je test-statistika

$$T = \sqrt{7}\frac{\overline{X}_7^{(1)} - \overline{X}_7^{(2)}}{S_D\sqrt{2}}, \quad \text{gdje je} \quad S_D = \sqrt{\frac{(S_7^{(1)})^2 + (S_7^{(2)})^2}{2}}.$$

Vrijednost test-statistike za gornje realizacije iznosi $t = 3.3391$.

Kritično područje, uz nivo značajnosti $\alpha = 0.05$, je

$$(-\infty, -t_{0.975}] \cup [t_{0.975}, \infty) = (-\infty, -2.18] \cup [2.18, +\infty),$$

gdje je $t_{0.975}$ 0.975-ti kvantili od $t(12)$.

t pripada kritičnom području, pa odbacujemo nul-hipotezu da su očekivane vrijednosti jednake s mogućnošću pogreške od 5%.

Praktično interpretirajući, postoji statistički značajna razlika u tlačnoj čvrstoći betona s čeličnim vlaknima i betona bez tih vlakana.

Napomene

(1) U slučajevima kad su varijance populacije poznate (σ^2 ili σ_1^2 i σ_2^2) u definiciji test-statistike pojavljuju se **standardne devijacije populacije** (σ, tj. σ_1 i σ_2).

(2) U slučajevima kad varijance populacije nisu poznate, u definiciji test-statistike pojavljuju se **standardne devijacije uzorka** (S_n, tj. $S_{n_1}^{(1)}$ i $S_{n_2}^{(2)}$).

(3) Kod testiranja malih uzorka iz normalne raspodjele za određivanje kritičnog područja koristimo kvantile

   (a) normalne raspodjele (z-kvantile) ako su varijance populacije poznate.

   (b) t-raspodjele ili Studentove raspodjele (t-kvantile) ako varijance populacije nisu poznate.

(4) Kod testiranja velikih uzorka (koji ne moraju biti iz normalne raspodjele) za određivanje kritičnog područja uvijek koristimo kvantile jedinične normalne raspodjele.