Statistični praktikum

Statistični testi

Asistent dr. Kristina Veljković

Fakulteta za računalništvo in informatiko

DELI STATISTIČNEGA TESTA

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.

DELI STATISTIČNEGA TESTA

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.
- Ničelno domnevo H_0 preverjamo s testno statistiko
  - Testna statistika ima določeno porazdelitev pri pravilni H_0.

DELI STATISTIČNEGA TESTA

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.
- Ničelno domnevo H_0 preverjamo s testno statistiko
  - Testna statistika ima določeno porazdelitev pri pravilni H_0.
  - Izberemo majhno verjetnost lažnega pozitivnega rezultata α (verjetnost napake 1. vrste ali stopnja značilnosti).

DELI STATISTIČNEGA TESTA

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.
- Ničelno domnevo H_0 preverjamo s testno statistiko
  - Testna statistika ima določeno porazdelitev pri pravilni H_0.
  - Izberemo majhno verjetnost lažnega pozitivnega rezultata α (verjetnost napake 1. vrste ali stopnja značilnosti).
  - Na osnovi porazdelitve testne statistike, stopnje značilnosti α in oblike alternativne domneve, določimo kritično območje, kjer se nahajajo vrednosti, kritične za H_0.

Deli statističnega testa

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.
- Ničelno domnevo H_0 preverjamo s testno statistiko
  - Testna statistika ima določeno porazdelitev pri pravilni H_0.
  - Izberemo majhno verjetnost lažnega pozitivnega rezultata α (verjetnost napake 1. vrste ali stopnja značilnosti).
  - Na osnovi porazdelitve testne statistike, stopnje značilnosti α in oblike alternativne domneve, določimo kritično območje, kjer se nahajajo vrednosti, kritične za H_0.
  - Če vrednost testne statistike pade v kritično območje, zavrnemo H_0 in sprejmemo H_1.

DELI STATISTIČNEGA TESTA

- Določimo ničelno domnevo H_0 (ni razlike ali spremembe) in alternativno domnevo H_1 (obstaja pomembna razlika).
- Privzamemo, da je H_0 pravilna, dokler se ne dokaže drugače.
- Ničelno domnevo H_0 preverjamo s testno statistiko
  - Testna statistika ima določeno porazdelitev pri pravilni H_0.
  - Izberemo majhno verjetnost lažnega pozitivnega rezultata α (verjetnost napake 1. vrste ali stopnja značilnosti).
  - Na osnovi porazdelitve testne statistike, stopnje značilnosti α in oblike alternativne domneve, določimo kritično območje, kjer se nahajajo vrednosti, kritične za H_0.
  - Če vrednost testne statistike pade v kritično območje, zavrnemo H_0 in sprejmemo H_1.
  - Če vrednost testne statistike ne pade v kritično območje, ne zavrnemo H_0.

UVOD

- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec.
- Testiramo $H_0 : \mu = \mu_0$ (μ_0 je predpostavljena vrednost) proti eni od alternativnih domnev
  - $H_1 : \mu \neq \mu_0$,
  - $H_1 : \mu < \mu_0$,
  - $H_1 : \mu > \mu_0$.

UVOD

- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec.
- Testiramo $H_0 : \mu = \mu_0$ (μ_0 je predpostavljena vrednost) proti eni od alternativnih domnev
  - $H_1 : \mu \neq \mu_0$,
  - $H_1 : \mu < \mu_0$,
  - $H_1 : \mu > \mu_0$.
- Predpostavka: slučajna spremenljivka X je normalno porazdeljena ali je vzorec dovolj veliki (za uporabo CLI).

STATISTIČNI TEST ZA μ (σ JE ZNAN)

- Za testiranje H_0 uporabljamo testno statistiko

$$Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}.$$

- $Z \sim \mathcal{N}(0, 1)$.

STATISTIČNI TEST ZA μ (σ JE ZNAN)

- Za testiranje H_0 uporabljamo testno statistiko

$$Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}.$$

- $Z \sim \mathcal{N}(0, 1)$.
- Kritično območje K

H_1	K	c
$\mu \neq \mu_0$	$\lvert Z\rvert \geq c$	$F^{-1}\left(1 - \frac{\alpha}{2}\right)$
$\mu < \mu_0$	$Z \leq c$	$F^{-1}(\alpha) = -F^{-1}(1 - \alpha)$
$\mu > \mu_0$	$Z \geq c$	$F^{-1}(1 - \alpha)$

STATISTIČNI TEST ZA μ (σ JE ZNAN)

- *Primer:* Teža lososov, gojenih v komercialni ribogojnici $X \sim \mathcal{N}(\mu, 0.5)$. Ribogojnica trdi, da je povprečna teža letošnjih lososov večja kot 3.5 kg. Da bi testirali njihovo trditev, smo zbrali naključni vzorec 16 lososov. Dobili smo vzorčno povprečno težo 3.8 kg. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?

STATISTIČNI TEST ZA μ (σ JE ZNAN)

- *Primer:* Teža lososov, gojenih v komercialni ribogojnici $X \sim \mathcal{N}(\mu, 0.5)$. Ribogojnica trdi, da je povprečna teža letošnjih lososov večja kot 3.5 kg. Da bi testirali njihovo trditev, smo zbrali naključni vzorec 16 lososov. Dobili smo vzorčno povprečno težo 3.8 kg. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Povprečna teža letošnjih lososov je večja kot 3.5 kg.

STATISTIČNI TEST ZA μ (σ JE ZNAN)

- *Primer:* Teža lososov, gojenih v komercialni ribogojnici $X \sim \mathcal{N}(\mu, 0.5)$. Ribogojnica trdi, da je povprečna teža letošnjih lososov večja kot 3.5 kg. Da bi testirali njihovo trditev, smo zbrali naključni vzorec 16 lososov. Dobili smo vzorčno povprečno težo 3.8 kg. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Povprečna teža letošnjih lososov je večja kot 3.5 kg.
- Ničelna domneva $H_0 : \mu = 3.5$ proti alternativni domnevi $H_1 : \mu > 3.5$.

Statistični test za μ (σ je znan)

- *Primer:* Teža lososov, gojenih v komercialni ribogojnici $X \sim \mathcal{N}(\mu, 0.5)$. Ribogojnica trdi, da je povprečna teža letošnjih lososov večja kot 3.5 kg. Da bi testirali njihovo trditev, smo zbrali naključni vzorec 16 lososov. Dobili smo vzorčno povprečno težo 3.8 kg. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Povprečna teža letošnjih lososov je večja kot 3.5 kg.
- Ničelna domneva $H_0 : \mu = 3.5$ proti alternativni domnevi $H_1 : \mu > 3.5$.
- Statistični test: Z test za en vzorec.

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- Standardni odklon σ ocenjujemo s popravljenim vzorčnim standardnim odklonom S.

Statistični test za μ (σ ni znan)

- Standardni odklon σ ocenjujemo s popravljenim vzorčnim standardnim odklonom S.
- Za testiranje H_0 uporabljamo testno statistiko

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n}.$$

- $T \sim t_{n-1}$ (t_{n-1} je Studentova porazdelitev z $n-1$ prostnostnimi stopnjami).

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- Standardni odklon σ ocenjujemo s popravljenim vzorčnim standardnim odklonom S.
- Za testiranje H_0 uporabljamo testno statistiko

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n}.$$

- $T \sim t_{n-1}$ (t_{n-1} je Studentova porazdelitev z $n-1$ prostnostnimi stopnjami).
- Kritično območje K

H_1	K	c
$\mu \neq \mu_0$	$\lvert T\rvert \geq c$	$t_{n-1;1-\frac{\alpha}{2}}$
$\mu < \mu_0$	$T \leq c$	$t_{n-1;\alpha} = -t_{n-1;1-\alpha}$
$\mu > \mu_0$	$T \geq c$	$t_{n-1;1-\alpha}$

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- *Primer:* Dvajset let nazaj so dijaki prvega letnika srednje šole v povprečju naredili 24 sklec v minuti. Raziskovalec trdi, da današnji dijaki prvega letnika ne morejo narediti toliko sklec v minuti. Da bi preveril svojo trditev, je zbral slučajni vzorec 40 učencev prvega letnika. Dobil je vzorčno povprečje 22.5 s popravljenim vzorčnim standardnim odklonom 3.1. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- *Primer:* Dvajset let nazaj so dijaki prvega letnika srednje šole v povprečju naredili 24 sklec v minuti. Raziskovalec trdi, da današnji dijaki prvega letnika ne morejo narediti toliko sklec v minuti. Da bi preveril svojo trditev, je zbral slučajni vzorec 40 učencev prvega letnika. Dobil je vzorčno povprečje 22.5 s popravljenim vzorčnim standardnim odklonom 3.1. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Današnji dijaki prvega letnika srednje šole ne morejo, v povprečju, narediti 24 sklec v minuti.

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- *Primer:* Dvajset let nazaj so dijaki prvega letnika srednje šole v povprečju naredili 24 sklec v minuti. Raziskovalec trdi, da današnji dijaki prvega letnika ne morejo narediti toliko sklec v minuti. Da bi preveril svojo trditev, je zbral slučajni vzorec 40 učencev prvega letnika. Dobil je vzorčno povprečje 22.5 s popravljenim vzorčnim standardnim odklonom 3.1. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Današnji dijaki prvega letnika srednje šole ne morejo, v povprečju, narediti 24 sklec v minuti.
- Ničelna domneva $H_0 : \mu = 24$ proti alternativni domnevi $H_1 : \mu < 24$.

STATISTIČNI TEST ZA μ (σ NI ZNAN)

- *Primer:* Dvajset let nazaj so dijaki prvega letnika srednje šole v povprečju naredili 24 sklec v minuti. Raziskovalec trdi, da današnji dijaki prvega letnika ne morejo narediti toliko sklec v minuti. Da bi preveril svojo trditev, je zbral slučajni vzorec 40 učencev prvega letnika. Dobil je vzorčno povprečje 22.5 s popravljenim vzorčnim standardnim odklonom 3.1. Kakšne zaključke lahko potegnemo, pri stopnji značilnosti $\alpha = 0.05$?
- Raziskovalna domneva: Današnji dijaki prvega letnika srednje šole ne morejo, v povprečju, narediti 24 sklec v minuti.
- Ničelna domneva $H_0 : \mu = 24$ proti alternativni domnevi $H_1 : \mu < 24$.
- Statistični test: t-test za en vzorec.

UVOD

- p je delež populacije z določeno lastnostjo.

UVOD

- p je delež populacije z določeno lastnostjo.
- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec , kjer je X_i indikatorska slučajna spremenljivka.

UVOD

- p je delež populacije z določeno lastnostjo.
- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec , kjer je X_i indikatorska slučajna spremenljivka.
- Neznani delež p ocenjujemo z vzorčnim deležem

$$\hat{p} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

UVOD

- p je delež populacije z določeno lastnostjo.
- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec , kjer je X_i indikatorska slučajna spremenljivka.
- Neznani delež p ocenjujemo z vzorčnim deležem

$$\hat{p} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- Testiramo $H_0 : p = p_0$ (p_0 je predpostavljena vrednost) proti eni od alternativnih domnev
  - $H_1 : p \neq p_0$,
  - $H_1 : p < p_0$,
  - $H_1 : p > p_0$.

STATISTIČNI TEST ZA DELEŽ p

- Za testiranje H_0 uporabljamo testno statistiko

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\sqrt{n}.$$

- $Z \sim \mathcal{N}(0, 1)$ (za dovolj veliko n).

STATISTIČNI TEST ZA DELEŽ p

- Za testiranje H_0 uporabljamo testno statistiko

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}.$$

- $Z \sim \mathcal{N}(0,1)$ (za dovolj veliko n).
- Kritično območje K

H_1	K	c
$p \neq p_0$	$\lvert Z\rvert \geq c$	$F^{-1}\left(1-\frac{\alpha}{2}\right)$
$p < p_0$	$Z \leq c$	$F^{-1}(\alpha) = -F^{-1}(1-\alpha)$
$p > p_0$	$Z \geq c$	$F^{-1}(1-\alpha)$

STATISTIČNI TEST ZA DELEŽ p

- *Primer:* Generator naključnih bitov je zgeneriral zaporedje 1 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1. Testiraj ničelno domnevo, da generator zgenerira bite slučajno, pri stopnji značilnosti $\alpha = 0.01$.

STATISTIČNI TEST ZA DELEŽ p

- *Primer:* Generator naključnih bitov je zgeneriral zaporedje 1 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1. Testiraj ničelno domnevo, da generator zgenerira bite slučajno, pri stopnji značilnosti $\alpha = 0.01$.
- p - delež bitov 1.

STATISTIČNI TEST ZA DELEŽ p

- *Primer:* Generator naključnih bitov je zgeneriral zaporedje 1 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1. Testiraj ničelno domnevo, da generator zgenerira bite slučajno, pri stopnji značilnosti $\alpha = 0.01$.
- p - delež bitov 1.
- Ničelna domneva $H_0 : p = 0.5$ proti alternativni domnevi $H_1 : p \neq 0.5$.

STATISTIČNI TEST ZA DELEŽ p

- *Primer:* Generator naključnih bitov je zgeneriral zaporedje 1 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1. Testiraj ničelno domnevo, da generator zgenerira bite slučajno, pri stopnji značilnosti $\alpha = 0.01$.
- p - delež bitov 1.
- Ničelna domneva $H_0 : p = 0.5$ proti alternativni domnevi $H_1 : p \neq 0.5$.
- Statistični test: Z test za en vzorec (za delež).