

Statistični testi

Nataša Kejžar

Povzetek

Postopek preizkušanja domnev

1. določi H_0 (navadno bolj preprosta od domnev) in stopnjo značilnosti α
2. definiraj smiselno testno statistiko
 - a. določi njeno porazdelitev
 - b. kaj vse so predpostavke
3. določi meje zavrnitve (kritične vrednosti) in območje zavrnitve / sprejema (prostor parametrov Θ pri zavrnitvi / pri sprejemu)

Napake pri statističnem sklepanju

- α – napaka I. vrste (zavrնemo H_0 , čeprav drži), velikost testa
- β – napaka II. vrste (ne zavrնemo H_0 , čeprav ne drži)
- $1 - \beta$ – moč testa (zavrնemo H_0 , ki ne drži)

Moč statističnega testa je odvisna od:

- α
- razlike med H_0 in H_A (ki mora biti natančno postavljena!)
- n
- variabilnosti

Pomembnost rezultatov statističnega testa

- statistična značilnost ($p < \alpha$)
- strokovna pomembnost (razlika, ki je pomembna za stroko; pomemben podatek pri določanju H_A za izračun moči testa)

Naloge

1. Radi bi ugotovili, ali se povprečna starost mladih v Sloveniji, ki začnejo kaditi, kaj razlikuje od tiste, ki jo po rezultatih Harvard School of Public Health upoštevajo pri označevanju primernosti filmov (13 let). Predpostavimo, da je starost ob začetku kajenja v populaciji normalno porazdeljena in da 80% mladih, ki začnejo kaditi, začne kaditi do 16. leta starosti.
 - a. Definirajte populacijo.
 - b. Kaj je ničelna domneva? Je enostavna/sestavljena?
 - c. Kaj je alternativna domneva? Je enostavna/sestavljena? Je enostranska/dvostranska?
 - d. Določite testno statistiko in njeno porazdelitev. (Kaj morate še določiti?) Zapišite vse predpostavke.
 - e. Določite meje zavrnitve (kritične vrednosti) pri $\alpha = 0.05$ za vzorec velikosti 100.
 - f. Izračunajte, kako verjetno je, da bomo zavrnil H_0 , če je populacijsko povprečje 14. Kaj smo s tem izračunali?
 - g. Simulirajte podatke:
 - i. Generirajte en vzorec ($n = 100$) iz porazdelitve pod ničelno domnevo, izračunajte vrednost testne statistike, priredite ji pripadajočo vrednost p in interpretirajte rezultate.
 - ii. Generirajte en vzorec z drugačnim povprečjem (npr. 14), izračunajte vrednost testne statistike, priredite ji pripadajočo vrednost p in interpretirajte rezultate.
 - iii. Generirajte 1000 vzorcev iz porazdelitve pod ničelno domnevo, preštejte, kolikokrat zavrnemo ničelno domnevo. Narišite porazdelitev vrednosti p pod ničelno domnevo (histogram in ecdf). Komentirajte, zakaj je histogram take oblike.
 - iv. Generirajte 1000 vzorcev z drugačnim povprečjem, preštejte, kolikokrat zavrnemo H_0 . Narišite porazdelitev vrednosti p pod H_0 (histogram in ecdf).
2. Naj bo populacija iz naloge o mladih kadilcih porazdeljena po gama porazdelitvi.

Pri porazdelitvi gama velja: $X \sim \Gamma(\alpha, \lambda)$, $E(X) = \alpha/\lambda$, $var(X) = \alpha/\lambda^2$.

Za generiranje vzorcev vzemite porazdelitev $\Gamma(39, 3)$ pod ničelno domnevo in alternativno porazdelitev $\Gamma(42, 3)$.

 - a. Na graf narišite gostoti porazdelitev za osnovne podatke pod H_0 iz naloge o mladih kadilcih in iz te naloge ($\Gamma(39, 3)$).
 - b. Razmislite o testni statistiki za ta primer. Kaj bi uporabili in kako je porazdeljeno?
 - c. S pomočjo grafa gostot **utemeljite**, ali bo velikost testa (verjetnost, da napačno zavrnemo H_0) v tem primeru večja/enaka/manjša kot prej.
 - d. S pomočjo grafa gostot **utemeljite**, ali bo napaka II.vrste (verjetnost, da napačno obdržimo H_0) pri isti razliki med povprečjema v tem primeru večja/enaka/manjša kot prej.
 - e. Izračunajte kritični vrednosti za velike n (npr. $n = 100$). Kako bi poiskali kritični vrednosti (meje zavrnitve) pri $n = 4$? Izračunajte jih s pomočjo R.
 - f. Ponovite simulacije iz prejšnje točke.
 - i. Generirajte 1000 vzorcev iz porazdelitve $\Gamma(39, 3)$, preštejte, kolikokrat zavrnemo ničelno domnevo. Narišite porazdelitev vrednosti p pod ničelno domnevo.
 - ii. Generirajte 1000 vzorcev iz alternativne porazdelitve, preštejte, kolikokrat zavrnemo H_0 . Narišite porazdelitev vrednosti p pod H_0 .
 - g. Spreminjajte velikost vzorca, da vidite, kakšna je dejanska velikost testa (napaka I.vrste) pri majhnih vzorcih. S simulacijo za $n = 4$ pokažite, da velikost testa v povprečju odstopa od vrednosti α .
3. Vrnimo se nazaj na osnovni primer mladih kadilcev, na nalogo 1. Kaj smo poleg $\mu_0 = 13$ pod H_0 še predpostavili in lahko v primeru, ko velja H_A , ne drži? Naredite nekaj simulacij na to temo.
4. Raziskovalci želijo ugotoviti, ali so otroci z nizkim samospoštovanjem bolj nagnjeni k depresijam, kot otroci v splošnem. Povprečen rezultat pri vprašalniku, ki meri nagnjenost k depresiji, je 90 z odklonom 14 (manjša vrednost pomeni večjo nagnjenost k depresiji). Podatki so v datoteki **data_depression.csv**.
 - a. Kaj je populacija?
 - b. Kaj je ničelna domneva?

- c. Kaj je alternativna domneva? Je enostavna/sestavljena? Je enostranska/dvostranska?
 - d. Kaj je vaša testna statistika? Kako je porazdeljena? Kaj so vaše predpostavke?
 - e. Določite meje zavrnitve in območje zavrnitve pri $\alpha = 0.05$.
 - f. Določite vrednost p za vaš primer.
 - g. Določite $(1 - \alpha)\%$ interval zaupanja. Interpretirajte interval zaupanja.
 - h. Statistično in vsebinsko interpretirajte rezultat. Kaj pomeni, da je rezultat statistično značilen in kaj pomeni, da je rezultat strokovno pomemben?
5. Dokazali bi radi, da dvotedenska terapija z nekim zdravilom uspešno znižuje sistolični tlak bolnikov z diabetesom. Vemo, da sistolični tlak precej niha, pričakujemo, da bo standardni odklon razlik enak približno 20 mmHg. želimo, da je stopnja tveganja pri našem dokazu enaka $\alpha=0.05$.
- a. Zapišite ničelno domnevo, ki jo želimo preverjati.
 - b. Kakšno standardno napako lahko pričakujemo, če vzamemo vzorec 25 bolnikov?
 - c. Kaj bomo lahko zaključili, če se bo tlak naših bolnikov v povprečju znižal za 5 mmHg?
 - d. Približno za koliko bi se moral v povprečju znižati tlak, da bi lahko ničelno domnevo zavrnil?
 - e. Na podlagi našega znanja menimo, da je dejanska povprečna razlika enaka 4 mmHg. Narišite porazdelitev povprečne razlike, ki jo lahko pričakujemo na našem vzorcu.
 - f. Kakšna je verjetnost, da bo opažena razlika dovolj velika, da bomo ničelno domnevo lahko zavrnil? Označite to tudi v grafu iz prejšnje točke.
 - g. Kaj pa, če je dejanska razlika enaka 6 mmHg? Se verjetnost, da bomo uspeli zavrniti spremeni? Kako?
 - h. Kakšna je verjetnost napake I. vrste v našem primeru?
 - i. Kakšna je verjetnost napake II. vrste v našem primeru?
 - j. Kako se bo spreminjala verjetnost zavrnitve, če bo razlika v populaciji večja? Narišite graf (v R), ki vam prikazuje verjetnost zavrnitve glede na dejansko (absolutno) razliko v populaciji (različne vrednosti μ_A).
 - k. Kaj se zgodi z napakama, če stopnjo tveganja znižamo na $\alpha = 0.01$?
 - l. Kakšna bo napaka II.vrste, če vzorec zmanjšamo? Narišite graf (v R), kaj se zgodi z napako II.vrste, če spreminjamo število bolnikov (npr. od 10 do 500)?
 - m. Raziskavo izvedemo in na podlagi podatkov ničelne domneve ne zavrnemo. Ali to pomeni, da smo jo dokazali?
 - n. Naredimo povzetek: recimo, da želimo načrtovati raziskavo podobno zgoraj opisani. Kaj vse moramo vedeti (oceniti), da bomo lahko izračunali potrebno velikost vzorca.
6. Raziskavo iz naloge o sistoličnem krvnem tlaku ponovno izvedemo na 10000 bolnikih, preverjamo torej ničelno domnevo: v populaciji je povprečna razlika sistolnega tlaka enaka 0. Standardni odklon v populaciji je enak 20 mmHg. Dobimo značilen rezultat ($p = 0.02$) in ničelno domnevo zavrnemo.
- a. Kakšna je bila standardna napaka v tem primeru?
 - b. Kakšna je bila vrednost testne statistike? Koliko se je torej tlak znižal na našem vzorcu?
 - c. Kakšen je 95% interval zaupanja za doseženo razliko?
 - d. Kaj lahko zaključimo?