

Kolokvij 2 - Osnove teoretične statistike

6. april, 2018

1. Zanima nas, ali opazovanje X prihaja iz enakomerne porazdelitve med 0 in 1. Alternativna domneva je, da opazovanje prihaja iz porazdelitve s porazdelitveno funkcijo $F(x) = \sqrt{x}$, $0 \leq x \leq 1$.
 - (a) Zapišite ničelno in alternativno domnevo. Ali sta enostavni, sestavljeni?
 - (b) Utemeljite, za kakšne vrednosti X bo test razmerja verjetij zavračal ničelno domnevo.
 - (c) Izračunajte mejo zavrnitve za primer vzorca velikosti 1 pri stopnji značilnosti α . Na tem istem primeru izpeljite še moč testa.
2. Na televiziji predvajajo zabavno oddajo, kjer se osebe spopadajo z različnimi izzivi. Vsaka oseba dobi tri naloge. Pred vsako nalogo oseba pritisne gumb, ki 'izžreba', ali bo predvideno nalogo sploh reševala. Pri gledanju ene oddaje zabeležimo naslednje rezultate pritiskov na gumb

število reševanj (k)	0	1	2	3
število oseb	1	2	6	1

Naj bo X spremenljivka z vrednostjo 1, če oseba nalogo rešuje in 0 sicer. Spremenljivka Y naj predstavlja število reševanj za eno osebo (od 0 do 3). Zanima nas, kakšna je verjetnost, da po pritisku na gumb oseba lahko rešuje nalogo.

- (a) Kako je porazdeljena spremenljivka X in kako Y ? Na vzorcu ocenite vrednost neznanega parametra po metodi največjega verjetja (izpeljava ni potrebna).
- (b) Kateri statistični test boste uporabili, da ugotovite, ali se porazdelitev Y statistično značilno dobro prilega podatkom.
- (c) Nastavite izračun za vrednost testne statistike iz točke (b). Vse vrednosti, ki jih zahteva enačba testne statistike, morajo biti izračunane (oz. vsaj utemeljene). Povejte/utemeljite, kako je testna statistika porazdeljena.

OBRNITE LIST

3. Zbrali smo naključen vzorec dečkov v starosti 10-12 let, za vsakega smo zabeležili podatek o tedenskem številu ur gledanja televizije in igranja igrice, radi bi primerjali povprečno število ur obeh aktivnosti (ničelna domneva pravi, da je povprečno število ur gledanja televizije enako povprečnemu številu ur igranja igrice). Predpostavimo, da je standardni odklon pri obeh aktivnostih enak. Odločamo se med uporabo testa t za odvisne in neodvisne vzorce.

- (a) Zapišite varianco razlike $\bar{X} - \bar{Y}$ v obeh primerih.
- (b) Predpostavimo, da število ur ukvarjanja z aktivnostima med seboj pozitivno korelira. Povejte, kateri test naj bi uporabili in razložite, kako se bosta spremenila velikost in moč, če uporabite napačen test. Skicirajte porazdelitev vrednosti p za vsak primer posebej.
- (c) Denimo, da bi primerjali povprečno število ur gledanja televizije s povprečnim številom ur branja, pričakujemo, da ti dve aktivnosti negativno korelirata. Ponovite razmislek iz točke (b).

4. Pri operaciji kolena prihaja do specifičnega neželenega učinka, ki ga kirurg zna predvideti sredi operacije (ko vidi izgled kolena). Zadnjih nekaj operacij je kirurg pri takih (rizičnih) pacientih nadaljnji postopek operacije malenkost spremenil in od takrat naprej ni pri nobenem pacientu po operaciji več prišlo do neželenega učinka. Kirurga zanima, kako velik vzorec pacientov mora zbrati, da bo lahko statistično dokazal, da je z novim postopkom verjetnost uspešnega zdravljenja (brez neželenega učinka) več kot 90%.

- (a) Kaj je v tem primeru vzorec?
- (b) Zapišite ničelno domnevo.
- (c) Zapišite testno statistiko. Zapišite, kako je testna statistika porazdeljena pod H_0 in povejte, za kakšne vrednosti (velike/majhne/v obe smeri) bi zavrnili H_0 .
- (d) Zapišite območje zavrnitve pri $\alpha = 0,05$. Narišite skico in ob njej razložite, kako bi izračunali mejo/meji z besedami. Nato zapišite še ustrezno (psevdo) kodo v R.
- (e) Najmanj kako velik vzorec mora zdravnik zbrati, da bo lahko zavrnil H_0 ($\alpha = 0,05$), če predpostavimo, da pri njegovih pacientih ne bo nikoli več prišlo do neželenega učinka? Zapišite psevdo kodo, s katero bi to izračunali.
- (f) Kirurg najbrž ne bo nezmotljiv (naj bo verjetnost neželenega učinka π_1). Za objavo odkritja v znanstveni reviji mora biti velikost vzorca izbrana tako, da je moč testa za alternativno domnevo vsaj 80%. Zapišite, kako bi to v tem primeru lahko izvedli. Razložite, kaj pri tem testu vpliva na določitev velikosti vzorca in na kakšen način (povečuje, zmanjšuje)...