

Kolokvij 2 - Osnove teoretične statistike

25. april, 2017

1. So spodnje trditve resnične? Kratko utemeljite vsako.
 - (a) Moč testa je pri večjem številu statističnih enot manjša/večja.
 - (b) Velikosti dopustnih testov na podlagi diskretnih spremenljivk so lahko večje/manjše od α .
 - (c) Za izračun velikosti testa potrebujemo/ne potrebujemo predpostavk(e) o porazdelitvi testne statistike pod alternativno domnevo.
 - (d) V primeru zvezne testne statistike in enostavne domneve je testna statistika pod ničelno domnevo porazdeljena po uniformni porazdelitvi.
2. Razložite, zakaj je testna statistika pri testu t za dva neodvisna vzorca ob predpostavki enakih varianc pod ničelno domnevo porazdeljena kot t_{n-2} . Pri tem lahko uporabite dejstvo, da je testna statistika pri testu t za en vzorec porazdeljena kot t_{n-1} .
3. Raziskovalci prek največjega spletnega mesta z novicami v državi izvajajo anketo glede prehranjevanja psov. V podatkovno bazo se zabeležijo vse izpolnjene ankete, zanimive pa so samo tiste, ki so jih izpolnili lastniki psov. Zanima nas, ali je verjetnost, da bo anketo izpolnil človek, ki ima doma psa, enaka 0,2. Verjetnost za dogodek, da smo po X poskusih dobili v vzorec lastnika psov, zapišemo kot

$$P(X = k) = (1 - p)^k p,$$

kjer je p verjetnost za osebo s psom. Iz izpolnjenih anket dobimo naslednjo frekvenčno tabelo o številu poskusov:

število poskusov (X)	0	1	2	3	4	5	6	7	8	9	10
frekvenca	11	2	4	2	0	6	1	1	1	0	2

- (a) Zapišite ničelno domnevo.
- (b) Natanko zapišite testno statistiko za posplošeni test razmerja verjetij za našo ničelno domnevo. Zapišite, kako je taka testna statistika porazdeljena.
- (c) Kako bi ničelno domnevo preverili s testom χ^2 ? Kaj pri tem predpostavimo? Zapišite pričakovane frekvence za gornji primer.
- (d) Zapišite psevdo kodo, s katero bi simulirali velikost testa iz prejšnje točke. Osnovne podatke v R generirate s pomočjo funkcije `rgeom(n,p)`. Posebej pozorno zapišite (in komentirajte), kako boste določili število kategorij.

OBRNITE LIST

4. Učitelj je študentom pri predmetu Statistika pripravil kratek izpit, ki sestoji iz petih vprašanj, na vsako vprašanje je možen eden izmed odgovorov A, B, C in D. Vprašanja so bila zastavljena v bengalskem jeziku, ki ga ne zna nobeden od študentov, učitelj se torej zaveda, da je pravilnost odgovorov povsem naključna. Da bi učitelj vseeno spodbudil študente k odgovarjanju, je obljubil nagrado tistim študentom, ki izpit opravijo pozitivno (pravilni vsaj trije od petih odgovorov). Ko so študenti končali z odgovarjanjem, je učitelj na tablo napisal vseh pet pravilnih odgovorov, študente pa pozval, naj se ocenijo in dvignejo roke tisti, ki so izpit opravili. Opozoril jih je, da bo nagrade podelil le v primeru, če ne bo mogel statistično dokazati, da so ga pri samoocenjevanju skupinsko goljufali (pri $\alpha = 0,05$).
- V učilnici je bilo prisotnih 60 študentov.

- a) Kakšno je največje število študentov, ki lahko dvignejo roko, ne da bi jih učitelj obtožil goljufije?
- Odgovor utemeljite s pomočjo spodnjih podvprašanj (pri izračunih lahko uporabite aproksimacije, kadar to lahko smiselno utemeljite):
- Kaj je učiteljeva ničelna domneva (zapišite jo z besedami in s formulo)?
 - Predlagajte testno statistiko, skicirajte njeno porazdelitev pod ničelno domnevo.
 - Razložite, kako bi kritično vrednost natanko izračunali z R-om. Kritične vrednosti ni treba izračunati, kljub temu zapišite, kakšno vrednost pričakujete.
- b) Učitelj pričakuje, da bi, če študentov ne bi opozoril, študenti dvignili roko z verjetnostjo 50% ne glede na njihov dejanski rezultat. Kakšno moč za odkrivanje goljufije bi imel na danem vzorcu? Približno izračunajte rezultat s pomočjo aproksimacije (pri tem za kritično vrednost vzemite številko, ki ste jo ocenili v prejšnji točki).