

4 Preverjanje domnev

4.1 Enostavni domnevi

Računalnik skuša razlikovati med dvema viroma signalov. Prvi vir oddaja signale, katerih jakost je normalno porazdeljena z $N(0,1)$, drugi vir ima enako porazdelitev, a višjo povprečno jakost - $N(2,1)$. Računalnik prejme 10 signalov in se mora odločiti iz katerega vira so prišli. Posamezni signali iz istega vira so med seboj neodvisni.

- Računalnik se odloča med domnevama

H_1 : Signal prihaja iz vira 1 in H_2 : Signal prihaja iz vira 2.

Zapišite testno statistiko, ob pomoči katere naj se odloča računalnik. (Izrazite malo bolj splošno - naj imata porazdelitvi obeh virov varianco σ^2 , povprečna jakost drugega vira naj bo a , $a > 0$, velikost vzorca naj bo n)

Namig: Pri tem, kaj je bolj verjetno, si pomagajte z gostotami

- Denimo, da želimo, da računalnik reagira le, če je precej prepričan, da signal ne prihaja iz vira 1. Domnevo H_1 proglasimo za ničelno domnevo, domnevo H_2 pa za alternativno. Odločitveno pravilo postavimo tako, da bo verjetnost zmote, kadar je ničelna domneva res, največ $\alpha = 0,05$.

– Testna statistika je slučajna spremenljivka (označite jo z Y). Kaj lahko rečemo o njeni porazdelitvi pod ničelno domnevo?

Namig: Da bo porazdelitev preprostejša, uporabite logaritem

– Izrazite mejno vrednost, pri kateri naj računalnik reagira.

- Kakšna je verjetnost, da bo računalnik reagiral, če signal v resnici prihaja iz drugega vira? (Tej verjetnosti pravimo moč testa)

Namig: Izračunajte porazdelitev testne statistike pod alternativno domnevo.

- Testno statistiko Y transformirajte tako, da bo pod ničelno domnevo standardna normalna spremenljivka.

- Povzemite: ali je mejna vrednost testne statistike odvisna od a ? Intuitivno razložite. Je vrednost a torej sploh pomembna?

Predlogi za vaje v R-u:

- Generirajte podatke v dveh korakih. Najprej z enako verjetnostjo izberite vrednost a (0 ali 2), nato generirajte 10 vrednosti iz porazdelitve $N(a,1)$. Izračunajte vrednost testne statistike iz prve točke in se odločite za eno izmed domnev glede na to ali je vrednost testne statistike večja ali manjša od 1. Postopek velikokrat ponovite in izračunajte delež poskusov, v katerih se odločite za vsako od domnev ter delež poskusov, ko je ta odločitev pravilna.
- Zamenjajte verjetnost v prvem koraku (naj bo npr. bolj verjetna izbira $a = 2$). Kako s spreminjajo deleži iz prejšnje točke?
- Generirajte podatke, tako da je a ves čas enak 0 in preverite, da je vrednost α pri izračunani mejni vrednosti c zares enaka 0,05.
- Generirajte podatke še tako, da je $a = 2$ in preverite moč testa.

4.2 Enostavni domnevi, posplošitev

Prejšnjo nalogo zapišimo v splošnem (primer izpitne naloge prof. Permana). Predpostavljamo, da so opazovane vrednosti neodvisne, enako porazdeljene slučajne spremenljivke X_1, X_2, \dots, X_n . Predpostavite, da sta samo dve možnosti: ali je gostota spremenljivk enaka $f(x)$ ali pa $g(x)$, kjer sta $f(x)$ in $g(x)$ znani pozitivni gostoti. Formalno postavimo:

$$H_0 : \text{gostota je } f(x) \text{ proti } H_1 : \text{gostota je } g(x).$$

- Predlagajte testno statistiko za preizkušanje zgornje domneve, če imate opazovane vrednosti x_1, \dots, x_n .
- Kdaj bi zavrnili ničelno domnevo pri stopnji značilnosti α ? Izrazite aproksimativno kritično mejo s količinama

$$a = \int \log \left(\frac{g(x)}{f(x)} \right) f(x) dx \text{ in } b = \int \log \left(\frac{g(x)}{f(x)} \right)^2 f(x) dx$$

4.3 Osnovni pojmi pri statističnem preverjanju domnev

Želimo preveriti ali je kovanec pošten. Naredili smo poizkus, kjer smo 10-krat vrgli kovanec in dobili, da je grb padel 7-krat.

- Zapišite ničelno domnevo za vaš primer. Ali je ničelna domneva enostavna ali sestavljena? Zapišite testno statistiko, označite jo z X - kakšna je njena porazdelitev pod ničelno domnevo?
- Denimo, da je vaša alternativna domneva $H_A : p > 0,5$. Ali je ta domneva enostavna ali sestavljena? Pri kakšnih vrednostih X boste zavrnilo ničelno domnevo v prid alternativni? Ali je domneva enostranska ali dvostranska?
- V našem primeru je $X = 7$. Kolikšna je verjetnost, da se na vzorcu zgodi ta dogodek, če ničelna domneva drži?
- Denimo, da je območje zavrnitve sestavljeno iz vrednosti $\{10\}$. Kakšna je stopnja značilnosti α v tem primeru? Kakšna je stopnja značilnosti, če je območje zavrnitve sestavljeno iz vrednosti $\{6,7,8,9,10\}$?
- Določite območje zavrnitve pri stopnji značilnosti $\alpha = 0,05$. Ali lahko na podlagi dobljenih podatkov zavrnete ničelno domnevo pri stopnji značilnosti $\alpha = 0,05$?
- Kolikšna je moč testa pri tej vrednosti α , če predpostavimo, da je prava vrednost parametra $p = 0,6$? Kaj pa pri $p = 0,7$? Kolikšna je v teh primerih napaka druge vrste?
- Predpostavite sedaj, da je vaša alternativna domneva $H_A : p \neq 0,5$. Ali je ta domneva enostavna ali sestavljena? Ali je domneva enostranska ali dvostranska?
- Kakšno bo sedaj območje zavrnitve, če želite, da je $\alpha \leq 0,05$? Kakšna natanko bo stopnja značilnosti za to območje?
- Izračunajte še moč testa v tem primeru.

Predlogi za vaje v R-u:

- 1000x ponovite poskus v katerem po 10x mečete kovanec. Oglejte verjetnost zavrnitve za posamezno območje.

p \ k	0	1	2	3	4	5	6	7	8	9	10
0,5	0,001	0,011	0,055	0,172	0,377	0,623	0,828	0,945	0,989	0,999	1
0,6	0,000	0,002	0,012	0,055	0,166	0,367	0,618	0,833	0,954	0,994	1
0,7	0,000	0,000	0,002	0,011	0,047	0,150	0,350	0,617	0,851	0,972	1

Tabela 1: Kumulativne verjetnosti za binomsko porazdelitev pri $n = 10$
 $(P(X \leq k|p))$

- Spremenite verjetnost s katero pade grb in si oglejte moč testa.
- Povečajte vzorec in si oglejte, kako se spreminja moč testa.

4.4 Moč testa

Iz literature lahko povzamemo, da se športnikovo povprečje hemoglobina ob vsaj 14-dnevnem bivanju na višini nad 1500m zviša za 2 g/l, medtem ko višinski treningi ne vplivajo na varianco njegovih vrednosti. Ob običajnih treningih se posameznikove vrednosti porazdeljujejo normalno, $X \sim N(\mu_1, 5^2)$, kjer je μ_1 športnikovo povprečje.

Športnik pogosto opravlja višinske treninge, vendar v krajših intervalih. Zanima ga, ali se njegovo povprečje hemoglobina v obdobju višinskih treningov kljub temu zviša. V sezoni opravi 12 meritev, 8 med obdobjem višinskih priprav in 4 sicer. Cilj naloge je ugotoviti, kakšna bo moč njegovega testa, če bo pri sklepanju uporabil stopnjo značilnosti $\alpha = 0,05$?

- Kaj je športnikova ničelna in kaj alternativna domneva?
- Predlagajte testno statistiko. Izračunajte njeno porazdelitev pod ničelno domnevo.
- Izračunajte moč testa, torej verjetnost, da bo ničelno domnevo uspel zavrniti, če se mu povprečje hemoglobina v obdobju višinskih priprav zares poveča za 2 g/l?
- Kako bi se moč testa spremenila, če bi imel na voljo enako število meritev v vsakem obdobju?
- Kako je moč testa odvisna od variance posameznikovih meritev in kako od dejanske velikosti razlike v populaciji?

4.5 Posplošeni test razmerja verjetij

Zanima nas ali imajo zares vsi športniki enako variabilnost hemoglobina. Primerjati želimo meritve k športnikov, naj bodo vrednosti i -tega športnika ($i = 1, \dots, k$) porazdeljene normalno, torej $X_{ij} \sim N(\mu_i, \sigma_i^2)$, kjer $j = 1, \dots, n_i$ označujejo meritve pri posamezniku. Predpostavimo, da so vse meritve med seboj neodvisne.

- Zapišite ničelno in alternativno domnevo
- Najprej vzemimo, da imamo le enega športnika in n njegovih meritev. Kako bi ocenili njegova parametra μ in σ^2 z metodo največjega verjetja?
- Vrnimo se h k športnikom. Utemeljite, da so pod alternativno domnevo ocene parametrov enake

$$\hat{\mu}_i = \frac{1}{n_i} \sum x_{ij}$$
$$\hat{\sigma}_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{ij} - \hat{\mu}_i)^2$$

- Kakšna je ocena povprečij pod ničelno domnevo?
- Kakšna je ocena variance pod ničelno domnevo?
- Kako bi ničelno domnevo preverili s testom razmerja verjetij?