Fakultet Elektrotehnike i računarstva

Seminar

Statisticko testiranje

Mentor:
Prof. dr. sc. Bojana Dalbelo-Bašić
Dr. sc. Goran Glavaš

Student:
Neven Miculinić

20. svibnja 2015.

Sadržaj

Poglavlje 1

Uvod

Ovaj seminar se bavi statističkim testovima, tj. prihvaćanjem ili odbacivanjem hipoteza. Započinje kratim opisom postupaka testiranja, objašnjavanjem ključnih pojmova te zatim demonstrira postupak testiranja na statističkom testiranju očekivanja.

Početna je premisa jednostavna, imamo hipotezu koju želimo provjeriti. Ona može biti svakojaka:

- na FERu je 99% muškaraca
- dnevno 1000 studenata FFZG idu u Cassandru
- Varijanca bodova na SISu je 15

Kao što vidimo u primjerima one testiraju vjerojatnost za neki parameter populacije Θ. On može biti različit: postotak pripadnika jedne subpopulacije, broj, očekivanje, varijanca neke statistike i mnoge druge oblike. Za daljnje primjere slučajna varijabla populacije ce biti X.

Uzorak nazivamo n-torku $(x_1, \ldots, x_n)$ koji su nezavisne realizacije slučajne varijable X. One imaju identičnu razdiobu kao i ona.[6]

Statistikom nazivamo svaku funkciju koja ovisi o uzorku $X_1, X_2, \ldots X_n$, a ne ovisi (eksplicitno) o nepoznatom parametru kojeg dobivamo iz tog uzorka. [6] Vrijednost te statistike naziva se procjenom parametra. [6]

Na temelju uzorka iz populacije donosimo zaključke. Budući da sami uzorci podliježu šansi (npr. nas uzorak FERovaca sastoji se od 20 zena no prezentira li to sliku populacije?) želimo znati koliko su sigurne naše pretpostavke.

Također prilikom uzorkovanja populacije moramo paziti na razne sistematske greške te kako ih izbjegavati.

Poglavlje 2

Hipoteza

Statistička je hipoteza tvrdnja o parametru jedne ili vise populacija.[3] Npr. tvrdnja da prosjecan broj jabuka pojedenih na dan iznosi 2 je validna hipoteza.

- Nul hipoteza *(eng. null hypothesis)* H_0
- Alternativna hipoteza *(eng. alternative hypothesis)* H_a

2.1 Nul hipoteza

Nul je hipoteza *(eng. null hypothesis)* teza o populaciji.

Ona može biti donesena na temelju modela tog sustava (npr. kroz biološke, fizikalne, matematičke ili druge zakonitosti), empirijska mjerenja sustava te na mnoge druge načina.

U većini literature [3][6] se označava s H_0. Primjeri nul hipoteza mogu biti razne, od hipoteza očekivanja populacije, hipoteza udjela subpopulaciju u populaciji, i mnoge druge primjene.

2.2 Alternativna hipoteza

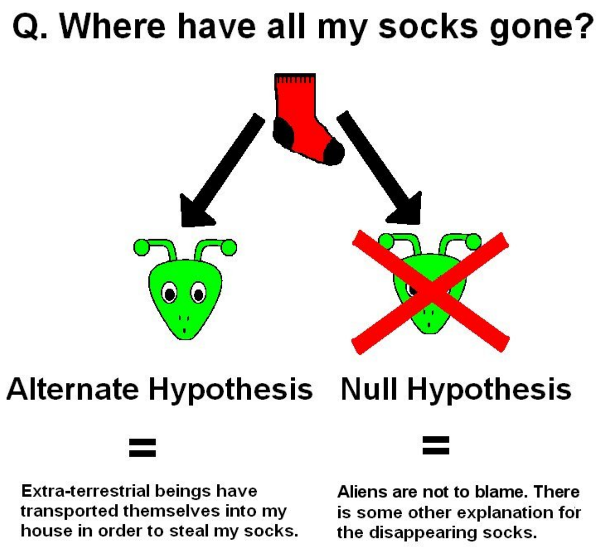
No sto ako postupkom testiranja odbacujemo nul hipotezu? Onda prihvaćamo alternativnu hipotezu. To naravno ne implicira u apsolutnu točnost jer na temelju rezultata dobivenih u uzorku ne možemo nikad biti sasvim sigurni je li odabrana hipoteza ispravna ili ne. [6]

Neka je H_0 : *Varijanca bodova na SISu je 15.* tj. $H_0 : \sigma^2 = 15$. Sto bi bila alternativna hipoteza?

Logično negacija H_0 te je time $H_a : \sigma^2 \neq 15$. Ovo je dvostrana alternativna hipoteza jer parametar σ^2 u alternative može biti i veći i manji od nul hipoteze.

Postoji još jedna mogućnost alternativne hipoteze, a to je jednostrana. Glasi ovako: $H_a : \sigma^2 > 15$ ili $H_a : \sigma^2 < 15$. Kao sto vidimo ova alternativna hipoteza gleda samo jednu stranu toga parametra te se zato zove jednostavna. [1]

Kada rabimo koju? Jednostranu rabimo samo ako smo iznimno uvjereni da je suprotan slučaj nemoguć, zbog fizikalnih, matematičkih ili ostalih zakona. Stoga najčešće uzimamo dvostranu alternativnu hipotezu.



Poglavlje 3

Struktura testa

3.1 Greške

Nakon uvoda u osnovne pojmove vezane uz hipotezu pogledajmo s kolikom sigurnošću možemo tvrditi istinost nase prosudbe u tesu. Imamo 4 moguca ishoda testa:

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Nul hipoteza je</td></tr>
<tr><td>Točna</td><td>Netočna</td></tr>
<tr><td rowspan="2">Presuda testa je:</td><td>Odbaci</td><td>Greška 1. vrste
Lažno pozitivni
$P = \alpha$</td><td>Točno</td></tr>
<tr><td>Prihvati</td><td>Točno</td><td>Greška 2. vrste
Lažno negativni
$P = \beta$</td></tr>
</table>

Kakav god test odabrali da potvrdimo ili odbacimo H_0 on može rezultirati u greskama 1. ili 2. vrste. Evo jednog ilustrativnog primjera za gresku 1. vrste.

3.1.1 Greske 1. vrste

Odbacivanje nul hipoteze kada je ona tocna naziva se greskom 1. vrste. [1]. Radi daljnjeg pojasnjenja pogledati primjer u nastavku:

Situacija je sljedeca. Vi ste cuvar nekoga sela i vasa je duznost oglasiti uzbunu ukoliko se vuk priblizava. Time je H_0 vuka nema. E ukoliko vuka stvarno nema, a vi ste oglasili uzbunu vi ste nacinili pogresku prve vrste ilitiga lazna pozitivnost.

Slicnu scenarij mozete vidjeti i s testom za trudnocu. Ukoliko krecete od hipoteze H_0 : *Nema trudnoce* te H_a : *trudnoca* te stvarno niste trudni, ali test pokazuje trudnocu to je jos jedan primjer Type I greske. Ona se oznacava s grckim slovom α te se naziva nivo znacajnost testa *(eng. significance level)*.

Pri samoj konstrukcji statistickog testa ukoliko je H_0 tocna mozemo lijepo ustimati α na prihvatljivu granicu te ga mozemo lijepo ustimati jer cesto pretpostavljamo kako funkcija razdiobe izgleda.

Ovdje dodati jos par slika i lijepse pojasniti

3.1.2 Greske 2. vrste

No dobro, ovo je sve super, mozemo podesiti test da nam je $\alpha \approx 0$ no sto time dobivamo? Tu u pricu ulaze greske 2. vrste koje imaju vjerojatnost β. Ne odbacivanjem nul hipoteze, H_0, kada je ona netocna se naziva pogreska 2. vrste. [1]

Koristeci primjere iz prethodne sekcije, vas test presudi da vuka nema dok on stvarno dolazi pred vasa vrata. Takoder vi ste trudni dok test za trudnocu to ne pokazuje dok nije prekasno. Faktor β je povezan s pojmom *snaga testa (eng. power)* koja iznosti $1 - \beta$. Ona je definirana kao vjerojatnost da ce test odbiti H_0 kada je ona lazna.

No taj faktor β je cesto nemoguce odrediti bez nekih pretpostavki. Npr. ((ovdje ubaciti neku normalnu distribuciju i koliko je beta ako je $\mu + \delta$ te kako on ovisi. Malo matematike i formulu upisati))

Kod izrade svakog statistickog testa dolazimo do biranja omjera snage i znacajnosti istoga. Oni se ponasaju kao na klackalici, sto je veca snaga testa to je veca znacajnost i obratno. Idealni test bi imao snagu 1 te znacajnost 0, no to u praksi nije moguce, te time treba pazljivo odabrati ta dva parametra tjekom izrade samoga testa.

3.2 P vrijednosti

Sljedeci vazan pojam u ovom kontekstu je p vrijednost. Ona se definira kao najmanja znacajnost test, α, za koju bi ovaj uzorak presudili odbacivanjem nul hipoteze, H_0 [1]

3.3 Podrucje prihvacanja

Potom zelimo konstruirati podrucje prihvacanja za nas test. [6] Ponovimo imamo H_0 hipotezu koja pretpostavlja neke parametre o distribuciju koju zelimo testirati.

Neka je slucajna varijabla $X \sim \mathcal{D}$. Prvo definirajmo oznaku kao u vecini literatura [6] [1]

x_α je $1 - \alpha$ percentil te zadane distribucije.

Za zadani nivo znacajnosti α to zapravo znaci $P(x \in \langle l, u \rangle) = 1 - \alpha$. Tj. da sansa da ukoliko je H_0 tocna vjerojatnost tocne presude iznosti $1 - \alpha$

Najcesce se rabe tri tipa podrucja prihvacanja:

- Gornji: $x \in \langle x_\alpha, +\infty \rangle$
- Doljni: $x \in \langle -\infty, x_{1-\alpha} \rangle$
- Dvostrani: $x \in \left\langle x_{1-\frac{\alpha}{2}}, x_{\frac{\alpha}{2}} \right\rangle$

Poglavlje 4

Statističko testiranje ocekivanja

U ovom poglavlju cu prikazati kako se testiranje vrsi na prosjek uzoraka. Prvo pretpostavljamo da je populacija normalno distribuirana.

Zatim zelimo odrediti koliki nam mora biti uzorak iz te populacije da bi znacajnost testa te njegova snaga bila zadovoljavajuca i uz to sve izbalansirati. Konkretni postupci odredivanj velicine uzorka izlaze izvan okvira ovog seminara.

Treba naravno napomenti da ce uzorak biti distribuiran po t distribuciji jer nam varijanca populacije nije poznata te pretpostavljamo normalnu distribuciju populacije.

Ukoliko je H_0 tocna nasa statisika:

$$\frac{\sqrt{n}(\bar{x}-\mu)}{s} \sim t_{n-1}$$

slijedi t distribuciju s $n-1$ stupnja slobode. Nakon malo matematike uz pretpostavku H_0:

$$\bar{x} \sim \frac{s}{\sqrt{n}} \cdot t_{n-1} + \mu$$

Uz znacajnost testa α dobivamo podrucje prihvacanja H_0 kao [3]:

$$\bar{x} \in \left\langle \mu - \frac{s}{\sqrt{n}} \cdot t_{n-1,\frac{\alpha}{2}}, \mu + \frac{s}{\sqrt{n}} \cdot t_{n-1,\frac{\alpha}{2}} \right\rangle$$

gdje $t_{n,\alpha}$ gdje pretstavlja $1-\alpha$ percentila t distribucije s n stupnjeva slobode, prikazano u prethodnom poglavlju.

Potom uz odredeni α napravimo podrucje prihvacanja za ovaj test. Ukoliko je prosjek uzorka unutar tog intervala nas test nije odbio H_0 te je presuda nul hipoteza, dok u suprotnom biramo H_a.

U nastavku je dan primjer programskog koda koju uz pythona te biblioteka NumPy odraduje cijeli postupak testiranja.

Treba ovdje uzeti neki primjer podataka i lijepo ih obraditi... ako stignem

```
from pylab import *
from numpy import *
from scipy.stats import t, norm
from numpy.random import normal

#Ovo je primjer ulaznih podata za testiranje
data = normal(3.1, 2, 15)
mu_test = 3
alpha = 0.05



n = data.size        # Broj uzoraka
s = data.std(ddof=1) # Procjena standardne devijacije populaije (ne uzoraka)
mu_s = data.mean()   # Procjena prosjeka iz uzoraka
mu_s_s = s/sqrt(n)   # Standardna devijacija prosjeka uzoraka

#Plottanje cisto da vidimo ulazne podatke
hist(data, bins=7);

(l,u) = t.interval(1-alpha, n - 1, mu_test,  mu_s_s)
print (l,u) #Podrucje prihvacanja

if l < mu_s and  mu_s < u:
    print ("H_0")
else:
    print ("H_a")
```

Poglavlje 5

Zakljucak

U ovako kratkom seminarskom radu je nemoguce prekriti svu raskos statisckih testova te on pokriva samo najelementarnije testiranje statistike. No ono moze biti preduvjet kao pretpostavke za neke naprednije testove koji pretpostavljaju mnogo vise nego normalnu distribuciju populacije te naravno puno vise i zakljucuju.

Bibliografija

[1] N. F. Hubele D. C. Montgomery, G. C. Runger. *Engineering statistics.* London: John Wiley & Sons, Inc., 2001.

[2] V. A. Clark O. J. Dunn. *Applied statistics: analysis of variance and regression.* London: John Wiley & Sons, Inc., 1974.

[3] Željko Pauše. *Uvod u matematičku statistiku.* Školska knjiga, 1993.

[4] Neven Elezović. *Diskretna vjerojatnost.* Element, Zagreb, 2007.

[5] Neven Elezović. *Slučajne variable.* Element, Zagreb, 2007.

[6] Neven Elezović. *Matematička statistika i stohastički procesi.* Element, Zagreb, 2007.