Fakultet Elektrotehnike i računarstva

Seminar

Statisticko testiranje

Student:
Neven Miculinić

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

4. svibnja 2015.

Sadržaj

Poglavlje 1

Uvod

Ovaj seminar se bavi statistickim testovima, tj. provjeravanjem ili pobijanjem hipoteza. Pocetna premisa je jednostavna, imamo neku hipotezu koju zelimo provjeriti. Ona moze biti svakojaka:

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

Kao sto vidimo u primjerima one testiraju vjerojatnost za neki paramater u populaciji Θ. On moze biti svakojak kao sto vidimo u primjerima: postotak pripadnika jedne subpopulacije u populaciji, broj, varijanca neke statistike i mnoge druge oblike.

Na temelju uzorka iz populacije donosimo zakljucke. Buduci da sami uzorci podlijezu sansi (npr. nas uzorak FERovaca sastoji se od 20 zena no prezentira li to sliku populacije?) zelimo znati koliko su sigurne nase pretpostavke.

Dodati neku sliku koja prezentira pogresku u uzorkovanju kroz statitski bias... npr. radimo anketu o bolestima pluca na studentima dok time netocno inferiramo za cijelu populaciju

Poglavlje 2

Hipoteza

Sto sve moze biti hipoteza? U uvodu smo vidjeli neke primjere, dok cu se ovdje baviti s dvaja pojma:

- Nul hipoteza *(eng. Null hypothesis)* H_0
- Altrenativna hipoteza *(eng. Altrenative hypothesis)* H_a

2.1 Nul hipoteza

Nul hipoteza je nasa osnova pretpostavka o populaciji. Ona kao takva podlijeze statistici uzorka te populaciju. Zasto uzorka? Evo jedan primjer. Zelimo provjeriti hipotezu iz uvoda *na FERu je 99% muskaraca.* Nerealno je i skupo ici od svakog pojedinca na fakultetu, provjeriti s kojeg je on faksa te zapravo analizirati svakog pripadnika populacije.

Time se bavi deskriptivna statstika, dok statisticka testiranja ulaze u inferencijsku statistiku.

Zato radimo uzorak od n primjeraka iz populacije i na temelju njih zakljcujemo za cijelu populaciju. Naravno zanima nas kolika je greska toga naseg suda.

2.2 Altrenativna hipoteza

No sto ako nasa hipoteza nije tocna? Onda vrijedi altrentivna hipoteza. Evo dat cu primjer.

Neka je H_0 : *Varianca bodova na SISu je 15.* tj. $H_0 : \ \Theta = 15$. Sto bi bila altrenativna hipoteza?

Logicno negacija H_0 te je time $H_a : \ \Theta \neq 15$. Ovo je dvostrana altrenativna hipoteza jer parametar Θ u altrenative moze biti i veci i manji od nulla.

Postoji jos jedana mogucnost altrenativne hipoteze, a to je jednostrana. Glasi ovako: $H_a : \ \Theta > 15$ ili $H_a : \ \Theta < 15$. Kao sto vidimo ova altrenativna hipoteza gleda samo jednu stranu toga parametra te se zato zove jednostavna.

Kada rabimo koju? Ako iz uzorka dobijemo $\hat{\Theta} = 16$ logicnije je uzeti jednostranu koja kaze $H_a : \ \Theta > 15$

Najcesce uzimamo dvostranu altrenativnu hipotezu osim ukoliko nam nije bitno... la la.

Sada kada malo razmislim, nije mi bas najjasnije tocno kada korstimo jednostranu altrenativu, a kada dvostranu

2.3 Primjeri

Imamo tvornicu igracaka i bitno nam je da su igracke u projeku vece od 15 cm. Tj. $\mu > 15cm$. Proizvodni proces je nastiman tako da prosjek bude 16 cm s nepoznatom varijancom. Naravno, samo zato jer je proces tako nastima, to ne znaci da se on u stvarnosti tako i ponasa. Stoga je ovdje prirodno uzeti $H_0 : \ \mu = 16cm$ te $H_a : \mu < 16cm$ buduci da nam je dulje od 15 cm nego da je tocan prosjek.

Mozda naci neki bolji primjer? Takoder tu bi bilo dobro neku sliku ubaciti

Poglavlje 3

Odluke Odluke

3.1 Greske

E sada kada smo prezentirali hipoteze trebamo se za jednu odluciti te odrediti koliko smo sigurni u nasu pretpostavku. Imamo 4 moguca ishoda:

		Null hypothesis is	
		True	False
Presuda testa je:	Odbaci	Type I error False Positive Sansa je significance level α	Tocno
	Prihvati (Fail to accept)	Tocno	Type II error β False Negative

Kakav god test odabrali da potvrdimo ili odbacimo H_0 on moze rezultirati u Type I ili Type II greskama. Evo jednog ilustrativnog primjera za Type I gresku.

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 Type I pogresku ilitiga False positive.

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

No dobro, ovo je sve super, mozemo nastimati test da nam je $\alpha \approx 0$ no sto time dobivamo? Tu u pricu ulaze Type II greske koje imaju vjerojatnost β. Ona se dogada kada je H_0 netocno, no test neuspijeje pobiti H_0 nego presudi tocnosti nul hipoteze.

Kod svih ovih testova koliko mi pretpostavljamo o funkciji razdiobe? Da li i u altrenativi pretpostavljamo normalnu npr. samo s drugim parametrima ili kako?

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 svakog statistickog testa dolazimo do balansacije snage i znacajnosti istoga. 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 balansirati ta dva parametra tjekom izrade samoga testa.

Koliko se parametri slucajne varijable mijenjaju kod linearnih transformacija? Konkretno pdf, μ, σ i σ^2. Koliko sam skuzio isto linearno osim varijance koja se kvadratno mijenja

3.2 Intervali pouzdanosti

Potom zelimo konstruirati intervale pouzdanosti za nas test. Ponovimo imamo H_0 hipotezu koja pretpostavlja neke parametre o distribuciju koju zelimo testirati. Na primjer imamo $H_0 : X \sim \mathcal{N}(\mu, \sigma^2)$ tj. po normalnoj distribuciji. Onda dvostrani $(1-\alpha)$ti interval pouzdanosti se dobije kao $\left\langle \mathcal{N}(\mu, \sigma^2)_{\frac{\alpha}{2}}, \mathcal{N}(\mu, \sigma^2)_{1-\frac{\alpha}{2}} \right\rangle$ gdje $\mathcal{N}(\mu, \sigma^2)_\alpha$ predstavlja α precentilu distribucije.

Takoder opcenito ako $H_0 : X \sim \mathcal{D}$ onda je dvostrani interval pouzdanosti: $\left\langle \mathcal{D}_{\frac{\alpha}{2}}, \mathcal{D}_{1-\frac{\alpha}{2}} \right\rangle$. Naravno mozemo slicno definirati i gornji interval pouzdanosti kao $\langle \mathcal{D}_\alpha, +\infty \rangle$ te doljni kao $\langle -\infty, \mathcal{D}_{1-\alpha} \rangle$

3.3 P vrijednosti

p-vrijednost odgovara vjerojatnosti da je nul-hipoteza tocna (tj. vjerojatnosti dobivanja uocene ili još vece razlike na slucajnim uzorcima). Takoder je najmanja α za koju ce nas test ne odbaciti H_0

Kako lijepo ovdje napisati fail to rejecet H_0 jer to ovo zapravo i radi. Ne dokazuje H_0 nego ju samo ne odbacuje.

3.3.1 Opasnosti p vrijednosti

Treba naravno uzeti u obzir sto p vrijednosti objasnjavaju. Po uzoru na clanak[1] ukoliko puno razlicitih populacija testiramo s relativno velikom p vrijednoscu od 5% vjerojatnost laznog otkrica nije zanemariva.

Koliko o ovome trebam detaljnije pisati u seminaru? Clanak super objasnjava sto je problematika bolje nego sto mogu ja u nekoliko paragrafa

[1] `http://rsos.royalsocietypublishing.org/content/1/3/140216`

Poglavlje 4

Statistiko testiranje na prosjek uzorka

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

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

ok, kako odrediti velicinu uzorka? Ako znamo varijancu mozemo pretpostaviti da ce biti distribuiran s $\mathcal{N}(\mu + \delta, \sigma^2)$ te onda odrediti preciznost δ, β i onda iz toga n.

Ukoliko je H_0 tocna nasa statisika:

$$\frac{\sqrt{n}(\hat{\mu} - \mu)}{s} \sim T_{n-1}$$

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

$$\hat{\mu} \sim \frac{s}{\sqrt{n}} \cdot T_{n-1} + \mu$$

Uz znacajnost testa α dobivamo gornji i doljni interval pouzdanosti kao:

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

gdje $T_{n,\alpha}$ gdje pretstavlja α percentila T distribucije s n stupnjeva slobode.

Potom uz odredeni α napravimo intervale pouzdanosti za ovaj test. Ukoliko je prosjek uzorka unutar tog intervala nas test nije odbio H_0 te je presuda null hipoteza, dok u suprotnom biramo H_a tj. altrenativnu hipotezu.

Kako bi bili u mogucnosti efokasno i bezbolno raditi ovaj tip testova sljedeci programski kod pisan u pythonu te koristeci numpy biblioteku radi za nas:

Treba ovdje uzeti neki primjer podataka i lijepo ih obraditi

```
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) #Interval pouzdanosti

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.