

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Statističko testiranje očekivanja

Neven MICULINIĆ

Voditelj: *Prof. dr. sc. Bojana DALBELO-BAŠIĆ Dr. sc. Goran GLAVAŠ*

Zagreb, svibanj 2015.

SADRŽAJ

1. Uvod	1
2. Hipoteza	3
2.1. Nul hipoteza	3
2.2. Alternativna hipoteza	3
3. Struktura testa	5
3.1. Greške	5
3.1.1. Greške 1. vrste	5
3.1.2. Greške 2. vrste	6
3.2. P vrijednosti	7
3.3. Područje prihvatanja	7
4. Statističko testiranje očekivanja	9
5. Zaključak	11
6. Literatura	12
7. Sažetak	13

1. Uvod

Ovaj seminar se bavi statističkim testovima, tj. prihvatanjem 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 će biti X .

Uzorak nazivamo n -torku (x_1, \dots, x_n) koji su nezavisne realizacije slučajne varijable X . One imaju identičnu razdiobu kao i ona.[5]

Statistikom nazivamo svaku funkciju koja ovisi o uzorku X_1, X_2, \dots, X_n , a ne ovisi (eksplicitno) o nepoznatom parametru kojeg dobivamo iz tog uzorka. [5] Vrijednost te statistike naziva se procjenom parametra. [5]

Na temelju uzorka iz populacije donosimo zaključke. Budući da sami uzorci podliježu šansi (npr. nas uzorak FERovaca sastoji se od 20 žena 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.



2. Hipoteza

Statistička je hipoteza tvrdnja o parametru jedne ili više populacija.[7] Npr. tvrdnja da prosječan 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 [7][5] 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. [5]

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

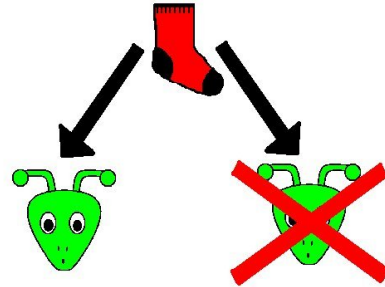
Logično negacija H_0 te je time $H_a : \sigma^2 \neq 15$. Ovo je dvostrana alternativna hipoteza jer parametar σ^2 u alternative može biti i veći i manji od nul hipoteze.

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. [2]

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.

Q. Where have all my socks gone?



Alternate Hypothesis

Null Hypothesis

=

=

Extra-terrestrial beings have transported themselves into my house in order to steal my socks.

Aliens are not to blame. There is some other explanation for the disappearing socks.

3. Struktura testa

3.1. Greške

Nakon uvoda u osnovne pojmove vezane uz hipotezu pogledajmo s kolikom sigurnošću možemo tvrditi istinitost naše prosudbe u testu. Imamo 4 moguća ishoda testa:

		Nul hipoteza je	
		Točna	Netočna
Presuda testa je:	Odbaci	Greška 1. vrste Lažno pozitivni $P = \alpha$	Točno
	Prihvati	Točno	Greška 2. vrste Lažno negativni $P = \beta$

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

3.1.1. Greške 1. vrste

Odbacivanje nul hipoteze kada je ona točna naziva se greškom 1. vrste. [2]. Radi daljnjeg pojašnjenja pogledati primjer u nastavku:

Situacija je sljedeća. Vi ste čuvar nekoga sela i vaša je dužnost oglasiti uzbunu ukoliko se vuk približava. Time je H_0 vuka nema. E ukoliko vuka stvarno nema, a vi ste oglasili uzbunu vi ste načinili pogrešku prve vrste ilitiga lažna pozitivnost.

Slični scenarij možete vidjeti i s testom za trudnoću. Ukoliko krećete od hipoteze H_0 : Nema trudnoće te H_a : trudnoća te stvarno niste trudni, ali test pokazuje trudnoću

to je još jedan primjer greške 1. vrste. Njena vjerojatnost se označava grčkim slovom α te se naziva nivo značajnosti testa (*eng. significance level*).

Pri samoj konstrukciji statističkog testa ukoliko je H_0 točna možemo lijepo uštimati α na prihvatljivu granicu te ga možemo lijepo uštimati jer često pretpostavljamo kako funkcija razdiobe izgleda.

Ovdje
dodati
još par
slika i
lijepke
pojasniti

3.1.2. Greške 2. vrste

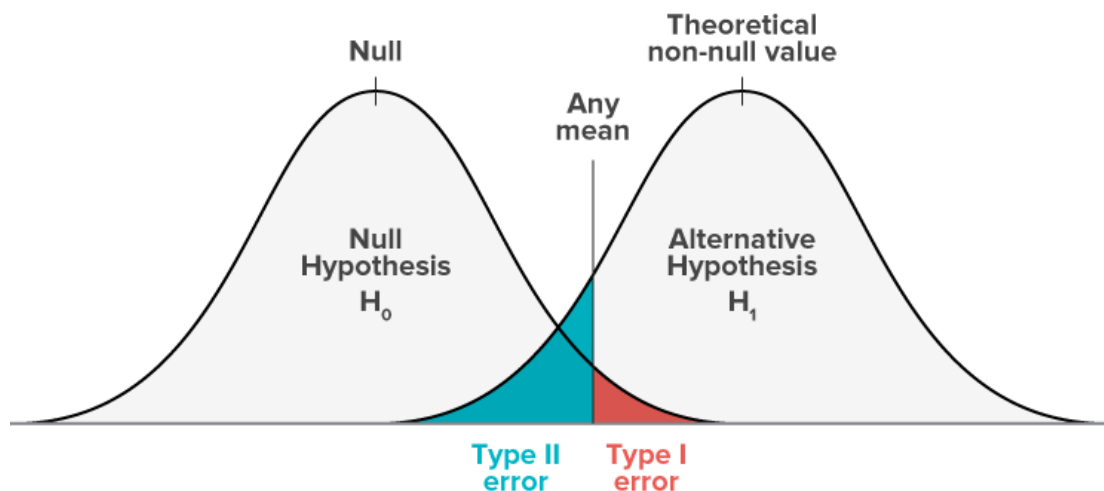
No dobro, ovo je sve super, možemo podesiti test da nam je $\alpha \approx 0$ no što time dobivamo? Tu u priču ulaze greške 2. vrste koje imaju vjerojatnost β . Ne odbacivanjem nul hipoteze, H_0 , kada je ona netočna se naziva pogreška 2. vrste. [2]

Koristeći primjere iz prethodne sekcije, vas test presudi da vuka nema dok on stvarno dolazi pred vaša vrata. Također vi ste trudni dok test za trudnoću to ne pokazuje dok nije prekasno.

Vjerojatnost pojave greške 2. vrste označava se grčkim slovom β te je povezan s pojmom *snaga testa* (*eng. power*) koja iznosi $1 - \beta$. Ona je definirana kao vjerojatnost da će test odbiti H_0 kada je ona lažna.

Međutim β je često nemoguće odrediti bez nekih pretpostavki. Npr. u testiranju očekivanja pretpostavljajući da je naša populaciju normalno distribuirana samo da se očekivanje razlikuje za neki δ tj. $\mu = \mu_0 + \delta$ uz $H_0 : \mu = \mu_0$.

Kod izrade svakog statističkog testa dolazimo do biranja omjera snage i značajnosti istoga. Oni se ponašaju kao na klackalici, što je veća snaga testa to je veća značajnost i obratno. Idealni test bi imao snagu 1 te značajnost 0, no to u praksi nije moguće, te time treba pažljivo odabrati ta dva parametra tijekom izrade samoga testa.

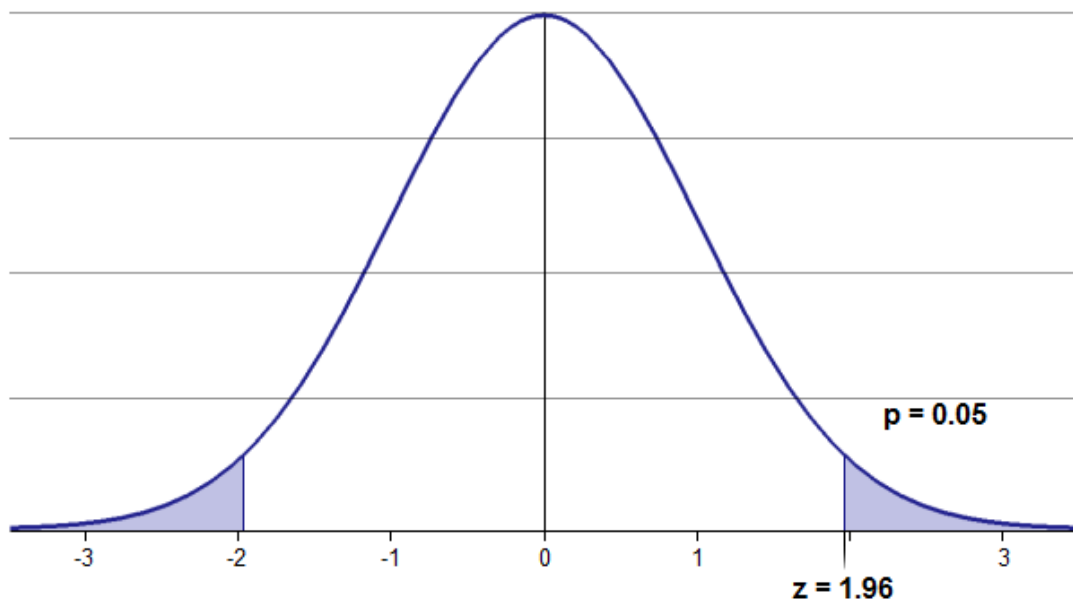


3.2. P vrijednosti

Sljedeći važan pojam u ovom kontekstu je p vrijednost. Ona se definira kao najmanja značajnost test, α , za koju bi ovaj uzorak presudili odbacivanjem nul hipoteze, H_0 [2]

Na slici 3.1 vidimo primjer. Neka je $H_0 : \mu = 0$ te se uzima dvostrana alternativna hipoteza $H_a : \mu \neq 0$. Pretpostavlja se normalna distribucija s jediničnom varijancom.

Za opažanje iz uzorka $z = 1.96$ p vrijednost našeg testa iznosi 0.05 što znači da za bilo koji $\alpha < 0.05$ presuda testa iznosi prihvatanje nul hipoteze.



Slika 3.1: P vrijednost statistike $\mu = 0$ s distribucijom $X \sim \mathcal{N}(0, 1)$

3.3. Područje prihvatanja

Područje prihvatanje *eng. region of acceptance* se definira kao interval vrijednosti statistike uzorka za koji presuđujemo nul hipotezu, H_0 . [1]

Ponovimo imamo H_0 hipotezu koja pretpostavlja neke parametre o distribuciju koju želimo testirati. [5]

Neka je slučajna varijabla $X \sim \mathcal{D}$.

Prvo definirajmo oznaku kao u većini literature: x_α je $1 - \alpha$ percentil te zadane distribucije. [5] [2]

Za zadani nivo značajnosti α to zapravo znači $P(x \in \langle l, u \rangle) = 1 - \alpha$. Tj. da šansa da ukoliko je H_0 točna, vjerojatnost točne presude iznosi $1 - \alpha$

Najčešće se rabe tri tipa područja prihvatanja:

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

4. Statističko testiranje očekivanja

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

Zatim želimo odrediti koliki nam mora biti uzorak iz te populacije da bi značajnost testa te njegova snaga bila zadovoljavajuća i uz to sve izbalansirati. Konkretni postupci određivanja veličine uzorka izlaze izvan okvira ovog seminara.

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

Ukoliko je H_0 točna naša statistika:

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

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 značajnost testa α dobivamo područje prihvatanja H_0 kao [7]:

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

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

Potom uz određeni α napravimo područje prihvatanja za ovaj test. Ukoliko je prosjek uzorka unutar tog intervala naš 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.

```

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 u
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 prihvatanja

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

```

5. Zakljucak

U ovako kratkom seminarskom radu je nemoguće prekriti svu raskoš statističkih testova te on pokriva samo najelementarnije testiranje statistike. No ono može biti preduvjet kao pretpostavke za neke naprednije testove koji pretpostavljaju mnogo više nego normalnu distribuciju populacije te naravno puno više i zaključuju.

6. Literatura

- [1] Region of acceptance. http://stattrek.com/statistics/dictionary.aspx?definition=Region_of_acceptance, 2015. 20.5.2015.
- [2] N. F. Hubele D. C. Montgomery, G. C. Runger. *Engineering statistics*. London: John Wiley & Sons, Inc., 2001.
- [3] Neven Elezović. *Diskretna vjerojatnost*. Element, Zagreb, 2007.
- [4] Neven Elezović. *Slučajne variable*. Element, Zagreb, 2007.
- [5] Neven Elezović. *Matematička statistika i stohastički procesi*. Element, Zagreb, 2007.
- [6] V. A. Clark O. J. Dunn. *Applied statistics: analysis of variance and regression*. London: John Wiley & Sons, Inc., 1974.
- [7] Željko Pauše. *Uvod u matematičku statistiku*. Školska knjiga, 1993.

7. Sažetak