Statistika i osnovna mjerenja

Testiranje hipoteza

M. Makek
2017/2018

Postavljanje hipoteza

- Postavljamo neku teoriju i želimo provjeriti da li je ona istinita
- Primjeri:
  1. Kocka (za bacanje) nije poštena
  2. Novi lijek je bolji od starog
  3. Teorijski model dobro opisuje izmjerene podatke
  4. Optuženi nije kriv
- Da bi se izvršila provjera postavljamo dvije hipoteze:
  a) H_0 – nul-hipoteza
  b) H_1 – alternativna hipoteza

  → Samo jedna od njih je točna!
- H_0 se smatra ispravnom dok se ne dokaže suprotno → H_0 i H_1 se ne tretiraju ravnopravno, već se konzervativnija postavlja kao istinita dok se ne dokaže suprotno.

Ako ne odbacimo H_0, to ne znači da je ona ispravna, nego samo da nemamo dovoljno dokaza da ju odbacimo.

Postavljanje problema

1. Prvo određujemo:
   - A - područje prihvaćanja hipoteze H_0
   - B - Područje odbacivanja hipoteze H_0 (kritično područje)
2. Uzimamo uzorak iz populacije koju istražujemo
3. Donosimo odluku:

		ODLUKA	
		Odbaci H_0	Prihvati H_0
ISTINA	H_0	Pogreška 1. vrste	Ispravan zaključak
	H_1	Ispravan zaključak	Pogreška 2. vrste

4. Određujemo vjerojatnost pogreške:
   - $\alpha = P(B|H_0)$, vjerojatnost da odbacimo H_0 kada je istinita
   - $\beta = P(A|H_1)$, vjerojatnost da prihvatimo H_0 kada je neistinita

→ Smatra se da je pogreška 1. vrste teža pogreška (npr. krivnja optuženika)

Signifikantnost testa

- Provjera (test) ima **signifikantnost** α, ako je vjerojatnost za pogrešku prve vrste **<** α. (Test razine α)
- Konvencionalno se za razinu signifikantnosti uzima:
  - α=0,05 (signifikantan)
  - α=0,01 (vrlo signifikantan)
- **Snaga testa** je vjerojatnost da test uputi na ispravnu odluku, tj. da odbacimo H_0 kad je uistinu neistinita: **P=1-β**

Primjer – bacanje novčića

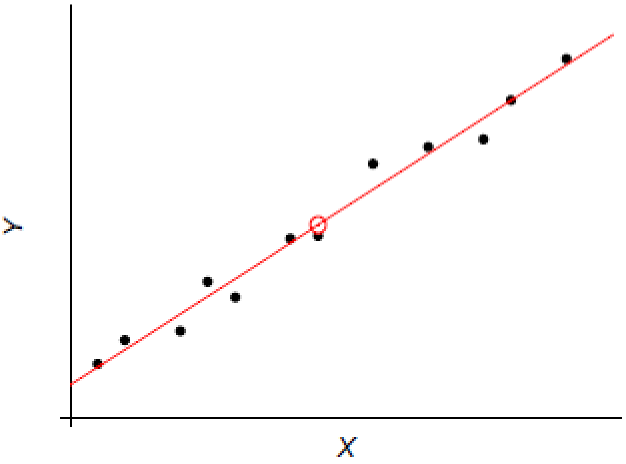
- Hipoteze:
  - H_0 = novčić je pošten
  - H_1 = novčić nije pošten
- X = broj 'pisama' u 6 bacanja → ako je novčić pošten onda je vjerojatnost $P(X|H_0) \sim Bin(6; 0.5)$
- Mogući su sljedeći ishodi:

X	0	1	2	3	4	5	6
$P(X\|H_0)$	0,0156	0,0938	0,2344	0,3125	0,2344	0,0938	0,0156

- Odaberemo: **A** za X=1,2,3,4,5 i **B** za X=0,6
- Vjerojatnost pogreške 1. vrste $\boldsymbol{\alpha = P(B|H_0)}$=0,0156 x 2
- Dakle ako u 6 bacanja kocke dobijemo 0 ili 6 puta pismo i na temelju toga zaključimo da je novčić nepošten, vjerojatnost da smo donjeli pogrešnu odluku je 3,12%

Prilagodba krivulje mjerenim podacima

- Pretpostavimo da smo izmjerili neku raspodjelu događaja (raspodjelu frekvencija)
- Prema obliku raspodjele i teoriji (ako je poznata) možemo mjerenim podacima prilagoditi krivulju regre-sijskim metodama (pravac, polinom, Gaussijan, Poissonova raspodijela, binomna raspodjela, itd.)



- Obzirom da svako mjerenje ima ograničenu preciznost javit će se odstupanja od krivulje
- Pitamo se da li su ta odstupanja slučajne prirode i da li pretpostavljena krivulja dobro opisuje izmjerene podatke

Dobrota prilagodbe

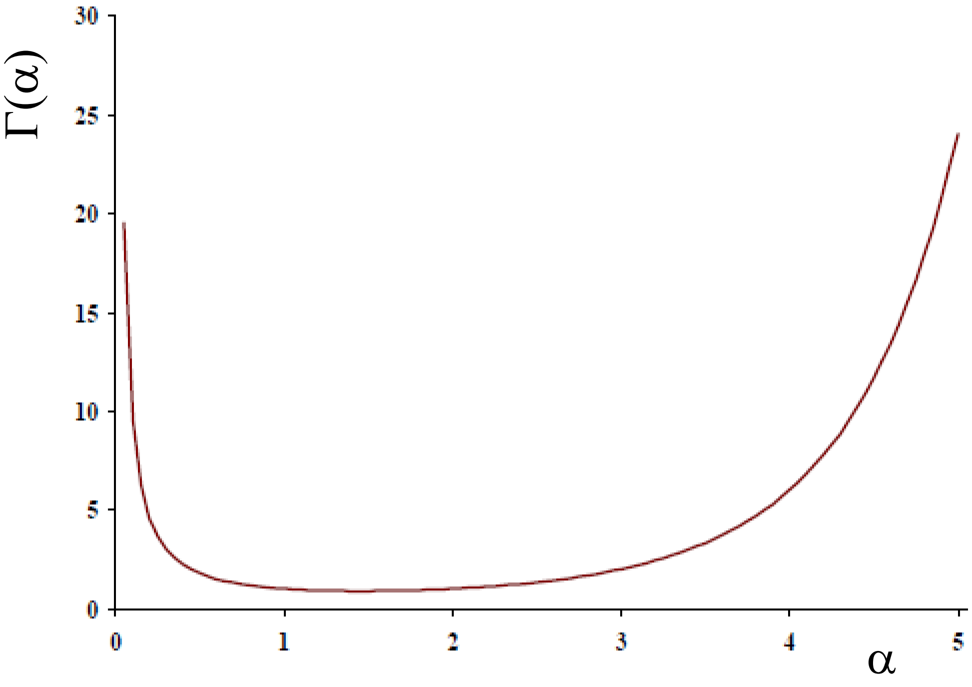
- Da bi ispitali kvalitetu (dobrotu) prilagodbe izmjerenim podacima izvodimo statistički test
- Jedan od testova koji se najčešće koriste je **χ^2-test**
- Podsjetimo se $\Gamma-$ funkcije, općenite $\Gamma-$ raspodjele i specifično χ^2 raspodjele

Gama funkcija

Definiramo Γ-funkciju:

$$\boxed{\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \qquad \alpha > 0}$$

- Svojstva:
  1. $\forall \alpha > 1 \rightarrow \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ (rekurzija)
  2. $\forall n \in \mathbb{N} \rightarrow \Gamma(n) = (n-1)!$ **(poopćenje faktorijela!)**
  3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$



Općenita Γ-raspodjela

- Funkcija gustoće vjerojatnosti:

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & x\geq 0\\ 0, & x<0\end{cases}$$

gdje su α, β >0

- Očekivanje: $E(X)=\alpha\beta$
- Varijanca: $V(X)=\alpha\beta^2$

χ^2-raspodjela

- Specijalni slučaj općenite Γ raspodjele za β=2 i α=ν/2
- Funkcija gustoće vjerojatnosti:

$$f(x;\nu) = \begin{cases} \dfrac{x^{\nu/2-1}e^{-x/2}}{2^{\nu/2}\Gamma(\nu/2)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

gdje je ν prirodan broj i naziva se “broj stupnjeva slobode”

- Veza s normalnom raspodjelom:

  Tm. Ako je ***u*** varijabla koja ima normalnu raspodjelu s očekivanjem i standardnom devijacijom μ i σ, onda je varijabla

$$x = \left(\frac{u-\mu}{\sigma}\right)^2$$

  raspodijeljena po χ^2-raspodjeli. Dokaz preskačemo.

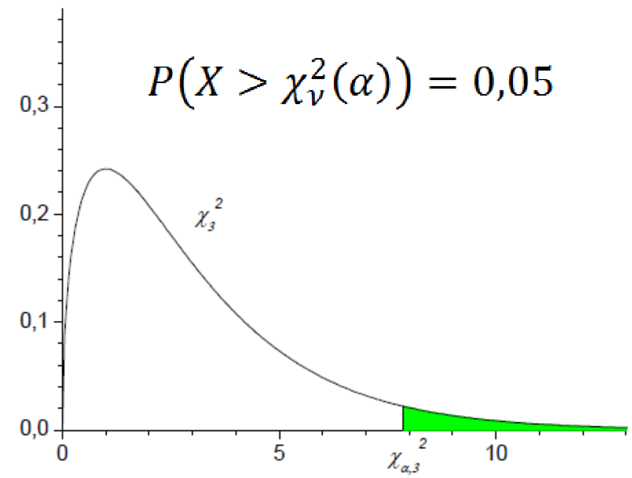
χ^2-test

1. Uzmimo da eksperiment daje raspodjelu frekvencija f_i, koje grupiramo u ***n*** razreda (i=1…n).
2. Izmjerenim podacima pomoću regresije prilagodimo krivulju, koja daje teorijske frekvencije za pojedini razred f_{ti}
3. Definiramo veličinu:
$$\chi^2 = \sum_{i=1}^{n} \frac{(f_i - f_{ti})^2}{f_{ti}}$$
4. ***Paersonov teorem***: veličina χ^2 je približno raspodijeljena prema χ^2-raspodjeli sa ν stupnjeva slobode koji ovise o broju opažanja (**n**) i broju ograničenja (**k**): $\nu = n - k$
5. Pri tome diskretnu raspodjelu aproksimiramo kontinuiranom. Aproksimacija ne vrijedi ako je $f_{ti} < 5$ → u tom slučaju treba grupirati razrede tako da je frekvencija pojedinog razreda uvijek veća od 4

Signifikantnost χ^2-testa

- Ako su razlike $|f_i - f_{ti}|$ relativno male, možemo zaključiti da imaju slučajan karakter
- Ako su razlike $|f_i - f_{ti}|$ relativno velike (signifikantne), onda će i χ^2 biti velik (signifikantan), tj. možemo zaključiti da je prevelik da bi bio slučajan. Tada je razumno odbaciti hipotezu.
- Vjerojatnost da varijabla ***X*** poprimi vrijednost veću od neke zadane $\boldsymbol{\chi_\nu^2(\alpha)}$, jednaka je površini repa ispod krivulje i dana je integralom:

$$P\left(X > \chi_\nu^2(\alpha)\right) = \int_{\chi_\nu^2(\alpha)}^{\infty} f(x;\nu)\,dx$$



- Po dogovoru se uzima da je χ^2 signifikantan ako padne u kritičnih 5% raspodjele i u tom slučaju odbacujemo hipotezu
- U prosječno 5% slučajeva će odbacivanje biti pogrešno jer je u toliko slučajeva moguće da je χ^2 signifikantan, a hipoteza je istinita.

Primjer 1

- Neki uređaj daje određeni postotak neispravnih proizvoda. Proizvodi se pakiraju u kutije od 20 komada. Za probu se otvara 50 kutija i određuje se broj neispravnih proizvoda X
- Opaženo je sljedeće:

X	0	1	2	3	>=4
f(x)	5	12	15	11	7

→ ***srednji x=1,8***

- ***Koja je vjerojatnost da se u kutiji nađe X defektnih komada?***
  - teorijski bi trebala biti dana binomnom raspodjelom: za $\bar{x} = np$ i n=20, dobivamo da je p=0.09

$$b(x) = \binom{20}{x} p^x q^{20-x}$$

  - Teorijske frekvencije:

X	0	1	2	3	>=4
f_t(x)	7,6	15,0	14,1	8,3	5,0

Primjer 1

- Nul-hipoteza: binomna raspodjela uz pretpostavku od 9% defektnih proizvoda dobro opisuje podatke
- Želimo potvrditi/odbaciti $\mathbf{H_0}$ uz signifikantnost **0.05**
- Broj stupnjeva slobode:
  - Broj razreda n=5
  - Broj ograničenja: 2 (određen **p** i **N**=50)
  - ν=5-2=3
- Izračunamo $\chi^2 = 3{,}08$
- Iz tablica vidimo da je za 3 stupnja slobode i 5% signifikantnosti $\chi^2_{0,05;3}=7{,}82$
- Obzirom da dobivena vrijednost ne spada u kritično područje tj. $\boldsymbol{\chi^2 < \chi^2_{0,05;3}}$ zadržavamo nul hipotezu.

Primjer 2

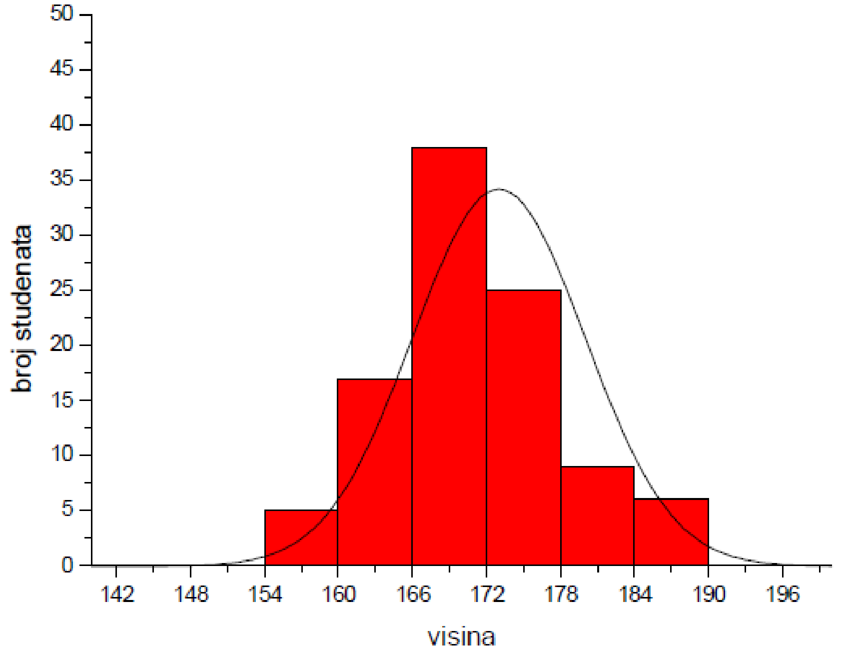
Dugogodišnje statistike pokazuju da je visina studenata na nekom sveučilištu normalno raspodijeljena s očekivanjem $\mu = 173$ cm i standardnom devijacijom $\sigma = 7$ cm.
Iz jedne generacije studenata na tom sveučilištu izdvojeno je 100 studenata, mjerene su njihove visine i svrstane u razrede.
Opažene su frekvencije:

x	154-160	160-166	166-172	172-178	178-184	184-190	Total
$f(x)$	5	17	38	25	9	6	100

Provjeri s 5% signifikantnosti je li raspodjela normalna s s očekivanjem $\mu = 173$ cm i standardnom devijacijom $\sigma = 7$ cm !

Primjer 2

H$_0$: *Raspodjela je Gaussova s očekivanjem* $\mu = 173$ cm *i standardnom devijacijom* $\sigma = 7$ cm.



Imamo jedno ograničenje: $\Sigma f_i = \Sigma f_{ti} = N = 100$

Primjer 2

- Izmjerene i teorijske frekvencije su dane tablicom:

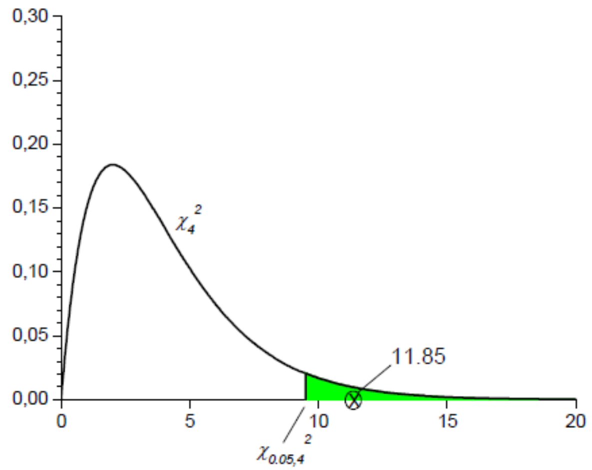
razred i	f_i	f_{ti}	$(f_i - f_{ti})^2 / f_{ti}$
154-160	5	3	
160-166	17	13	2,25
166-172	38	28	3,57
172-178	25	32	1,53
178-184	9	18	4,5
184-190	6	6	0
zbroj	100	100	11,85

- Tablica ima 5 razreda te imamo jedno ograničenje N=100 →broj stupnjeva slobode je 4

Primjer 2

Za 4 stupnja slobode i 5% signifikantnosti, kritična vrijednost (iz tablica) je

$$\chi^2_{0.05,4} = 9.49$$



Izračunata vrijednost $\chi^2_{\mathrm{op}} = 11.85$ pada u kritično područje.

Stoga hipotezu H_0 **odbacujemo**.

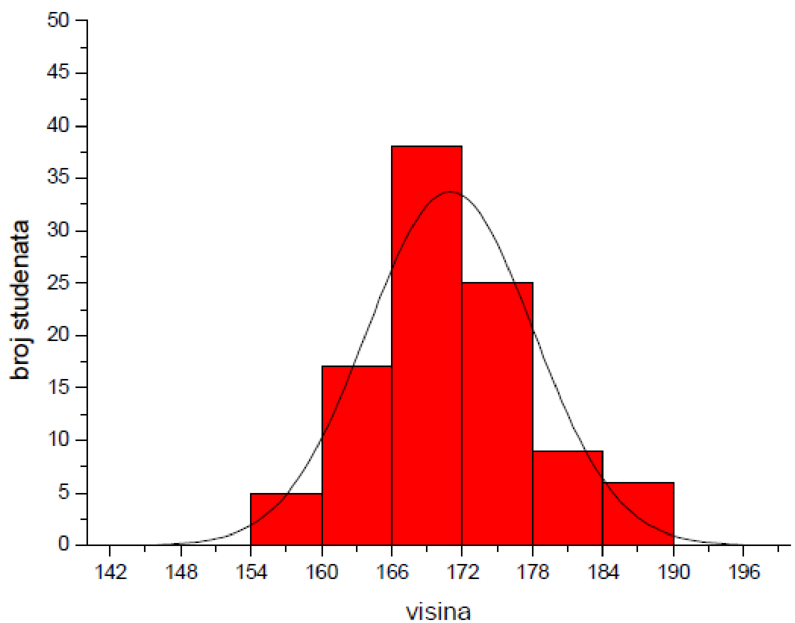
Primjer 2

- Nova hipoteza:

H$_0$: *Raspodjela je Gaussova s očekivanjem* $\mu = \bar{x}_{\text{uzorka}}$
i standardnom devijacijom $\sigma = \sigma_{\text{uzorka}}$.

Izračunamo:

$$\bar{x}_{\text{uzorka}} = \frac{1}{N}\sum f_i x_i = 171\,\text{cm} \qquad \sigma_{\text{uzorka}} = \sqrt{\frac{1}{N}\sum f_i (x_i - \bar{x}_{\text{uzorka}})^2} = 7{,}1\ \text{cm}$$



Primjer 2

- Izmjerene i teorijske frekvencije su dane tablicom:

razred i	f_i	f_{ti}	$(f_i-f_{ti})^2/f_{ti}$
154-160	5	6	0,167
160-166	17	18	0,056
166-172	38	32	1,125
172-178	25	28	0,321
178-184	9	13	0,063
184-190	6	3	
zbroj	100	100	1,732

Dakle, imamo pet razreda i tri ograničenja: $N = 100$

$$\mu = \bar{x}_{\text{uzorka}}$$

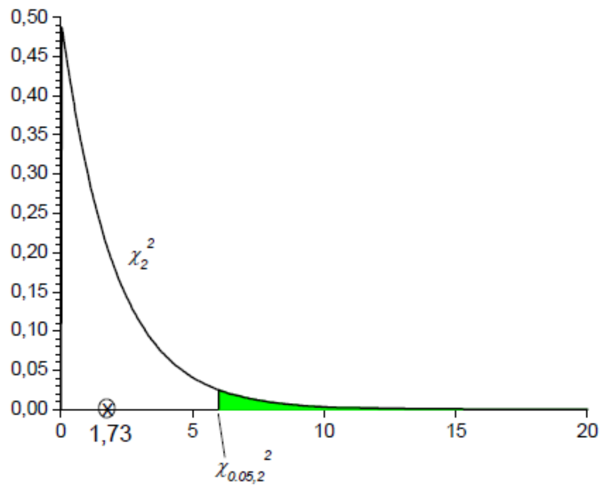
$$\sigma = \sigma_{\text{uzorka}}$$

pa je $\nu = 5\text{-}3 = 2$

Primjer 2

Za 2 stupnja slobode i 5% signifikantnosti, kritična vrijednost je

$$\chi^2_{0.05,2} = 5{,}99$$



Izračunata vrijednost $\chi^2_{\mathrm{op}} = 1{,}73$ ne pada u kritično područje.
Stoga hipotezu H_0 **zadržavamo**.

Dodatak: χ^2-tablica

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997