

## 12. TESTIRANJE HIPOTEZA

### 12.1. Uvod i osnovni pojmovi

Pr.)  $\rightarrow$  printer  $\rightarrow$  50 str./min (proizvođač tvrdi specifikacijama)  
 $\rightarrow$  isprobali 30 ispisa  $\rightarrow$  47 str./min (mi dobili)

$H_0$ ..... NULTA hipoteza

$H_1$ ..... ALTERNATIVNA HIPOTEZA

$\rightarrow$  za naš primjer:  $H_0$ ..... proizvođač govori istinu

$H_1$ ..... proizvođač LAŽE

def.

**SNAGA TESTA** se definira sa

$$S(w) = P(\text{prihvati } H_1)$$

$\rightarrow$  idealno bi bilo ako je  $w \in H_0$  i  $S(w) = 0$  (proizvođač rekao istinu),  
ako je  $w \in H_1 \Rightarrow S(w) = 1$ .



def.

**POGREŠKA PRVE VRSTE** se definiira kao  $\alpha = \sup S(\alpha)$  ako je  $\alpha \in H_0$ . Vjerojatnost da prihvatimo  $H_1$ , a istina je  $H_0$ !

**POGREŠKA DRUGE VRSTE** se definiira kao  $\beta = \sup (1 - S(\alpha))$  ako je  $\alpha \in H_1$ . Vjerojatnost da prihvatimo  $H_0$ , a istina je  $H_1$ !

→ Pogreška prve vrste je OZBIJNJA POGREŠKA! Slučaj da smo nevinog strpali u zatvor (nego kriveca pustili iz zatvora!)

→ Svi su NEVINI, DOK SE NE DOKAŽE SUPROTNO!

## 12.2. TESTIRANJE PARAMETARSKIH HIPOTEZA

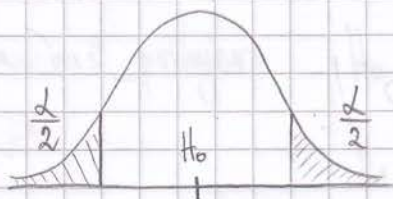
$$H_0 \dots \sigma = \sigma_0$$

$$H_1 \dots \sigma \neq \sigma_0 \quad (111) \quad \sigma > \sigma_0 \quad (111) \quad \sigma < \sigma_0 \quad \rightsquigarrow \text{SKUŽITI U ZADATKU!}$$

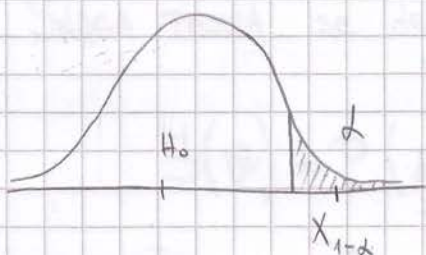
DVOSTRANI TEST

LIJEVI TEST

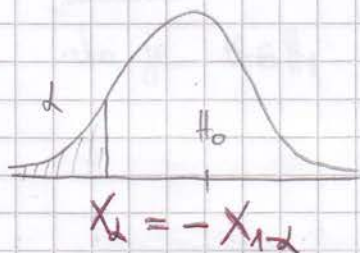
DESNI TEST



DVOSTRANI TEST



DESNI TEST



LIJEVI TEST



$$t = \frac{\bar{X} - a_0}{\frac{s}{\sqrt{n}}}$$

SLUŽBENE FORMULE

21-09-4.) Rafinerija proizvodi benzin<sup>za</sup> koji se očekuje da ima oktanski broj 98. Uzeto je 12 uzoraka benzina i dobiveni brojevi 99,1... (zadano) Možemo li na temelju ovih podataka<sup>prosjeka</sup> tvrditi da rafinerija proizvodi benzin<sup>prosjeka</sup> oktanskog broja 98 uz nivo značajnosti 95%?

$$H_0 \dots a = 98$$

$$H_1 \dots a \neq 98 \quad (1 \text{ bod})$$

BILOŠTO RAZLIČITO OD 98 NAM NE ODGOVARA

→ službene formule:  
(nepoznata disp.)

$$t = \frac{\bar{X} - a_0}{\frac{s}{\sqrt{n}}}$$

$$\bar{X} = \frac{99,1 + \dots}{12} = 97,6583$$

$$n = 12$$

$$s^2 = 0,6079 \quad (\text{točkaste procjene})$$

$$s = \sqrt{0,6079}$$

$$\Rightarrow t = -1,518 \quad (1 \text{ bod})$$

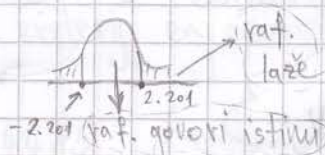
→ moramo vidjeti gdje se  $t$  nalazi

KRITERIJ ZA ODBACIVANJE  $H_0$ :

$$|t| > t_{n-1, 1-\frac{\alpha}{2}} = \{ \text{TABLICA STUDENTOVE RAZDIOBE} \} = 2,201 \quad (1 \text{ bod})$$

tablica 3.

1,518 < 2,201 NIJE → Rafinerija govori ISTINU!  
→ prihvaćamo  $H_0$ !



**MORATE TO  
NAPISATI !**  
NEMA BODA BEZ !!



7.DZ-3.) Proizvođač tvrdi da je vrijeme rada uređaja 200 dana.  
Zabrali smo par uzoraka, 165, 170..., 203, 210.

Provjeri proizvođača uz nivo značajnosti 95%.

$$H_0 \dots \mu = 200 \quad \alpha = 5\%$$

$H_1 \dots \mu < 200 \rightarrow$  Hoće li zvati proizvođača i poslati mu bogove ako je radio 210 dana??  
Pa ne.

$$\bar{X} = \frac{165 + 170 + \dots + 210}{n} = 188.5$$

zadano točno u zad.

$$s^2 = 253.43$$

$$n = 8$$

hipoteza... uz nepoznatu  $\sigma^2$ :

$$\hat{t} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = -2.04$$

kriterij za odbacivanje  $H_0$ :

$$\hat{t} < -t_{n-1, \frac{\alpha}{2}} = -t_{7, 1-0.05} = -1.895$$

$\hookrightarrow$  PAZI! U TABLICI IMAS  $1 - \frac{\alpha}{2}$  KVANTIL!

(GLEDATI) UZETI DUPLO VEĆI KVANTIL!! 0.1 U TABLICI

$\Rightarrow$  Kod JEDNOSTRANOG t-TESTA IZ TABLICE GLEDAMO  $\alpha' = 2\alpha$ !!

NAJVEĆI TRIK  
U STATISTICI

$\rightarrow$  SAMO KOD STUDENTOVE!

II



$$-2.04 < -1.895 \quad \text{JEST!}$$

⇒ Prihvaćamo  $H_1$ , odnosno odbacujemo  $H_0$ ! Proizvođač laže.

"BEZ TOG ZAKLJUČKA NEMA BODA !!!"

7.D2-6.) Pri proizvodnji u normalnim uvjetima stroj daje 2% škartu. Na uzorku od 500 proizvoda primjećeno je 16 škartnih. Proveri hipotezu o ispravnosti deklaracije uz nivo značajnosti 5%.

$$\alpha = 5\%$$

$H_0 \dots p = 0.02$   
 $H_1 \dots p > 0.02$  → ako dobijemo manje od 2% škartu to nam je super, nećemo se žaliti!  $\pi$

HIPOTEZA O PROPORCIJI:

$$\hat{u} = \left( \frac{\frac{m}{n} - p_0 \right) \sqrt{\frac{p_0 q_0}{n}}$$

$\downarrow$                        $\downarrow$                        $\downarrow$   
 $\frac{16}{500}$                        $0.02$                        $0.98$

$$\hat{p} = \frac{16}{500} = 0.032 = 3.2\%$$

PAZI!  $p_0 = 0.02$  ovaj koji ispitujemo!

$$\hat{u} = 1.917$$

$$\alpha = 0.05 \Rightarrow \hat{u} > u_{1-\alpha} \quad \{ \text{"kriterij za odbacivanje } H_0" \}$$

$$u_{1-\alpha} = u_{0.95} = 1.64485 \quad (\text{zadnja tablica u podsjetniku})$$

$$1.917 > 1.64485 \quad ? \text{ DA!}$$

⇒ Deklaracija NIJE ISPRAVNA, tj. odbacujemo  $H_0$ !

AKO NISMO SIGURNI  
 ŠTO ODBACUJEMO  
 PIŠE U PODSJETNIKU



21-11-5.) Svi studenti 2. godine Fera polažu jedan ispit. Na uzorku od 97, 67 ih je prošlo. Uz koji nivo značajnosti možemo tvrditi da vrijedi bolonjski prolaz od 75%?

$$H_0 \dots p = 0.75$$

$$\hat{p} = \frac{m}{n} = \frac{67}{97} = 0.711$$

$$H_1 \dots p < 0.75$$

→ ako je više od 75% onda je to o.k. i vrijedi bolonjska prolaznost!!

testiranje proporcije:

$$z = \frac{(\hat{p} - p_0) \sqrt{\frac{n}{p_0 q_0}}}{1} \rightarrow 97$$

$\downarrow$        $\downarrow$        $\downarrow$   
 0.711    0.75    0.25

$$\hat{u} = -0.879$$

→ kriterij za odbacivanje lijevog testa:

**PAZI!**

$$\hat{u} < -u_{1-\alpha}$$

→ NAS ZANIMA KADA CE BITI BOLONJSKA PROLAZOST (PRIHVATITI  $H_0$ )!

Dakle,

$$\hat{u} > -u_{1-\alpha}$$

$$-0.879 > -u_{1-\alpha}$$

$$0.879 < u_{1-\alpha}$$

→ PRVI VEĆI BROJ (OBRNUTO U TABLICI TRAJIMO)

$$\Rightarrow 1 - \alpha = 0.82$$

$$\alpha = 18\%$$



## 12.3. Usporedba Dviju Populacija

21-09-5.) 90 od 200 anketiranih muškaraca i 69 od 150 anketiranih žena se rekreativno bavi sportom. Uz nivo značajnosti od 5% provjeriti da li se jednak postotak muških i ženskih bave sportom.

→ hipoteza o JEDNAKOSTI PROPORCIJA:

$$H_0 \dots p_1 = p_2$$

$$H_1 \dots p_1 \neq p_2$$

formule. 
$$\hat{u} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}^2}{n_1} + \frac{\hat{p}^2}{n_2}}}$$

$$\Rightarrow \hat{u} = 0.186$$

→ kriterij za odbacivanje:

$$|\hat{u}| > u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$$

$$0.186 > 1.96 ? \text{ (NE) !}$$

⇒ Prihvataemo  $H_0$ , tj. nema razlike između muškaraca i žena u postotku bavljenja sportom.

$$n_1 = 200$$

$$n_2 = 150$$

$$\hat{p}_1 = \frac{90}{200} = 0.45$$

$$\hat{p}_2 = \frac{69}{150} = 0.46$$

$$\hat{p}^2 = \hat{p}(1-\hat{p}) = 0.248$$

$$\hat{p} = \frac{90 + 69}{200 + 150} = 0.454$$

FORMULE

II



21-08-5.) "VEČNI SUKOB E vs R"

E-grupe su na M1 imale prosjek 20.625 s devijacijom 8.07, od njih 254, a R-ovci od njih 299 prosjek je bio 19.697 s devijacijom 8.36.

Uz koji nivo značajnosti možemo tvrditi da NE postoji razlika između ta dva svijera?  $\Pi$

$$E(254), \bar{x} = 20.625 \quad \sigma_x = 8.07$$

$$R(299), \bar{y} = 19.697 \quad \sigma_y = 8.36$$

→ s.f. POZNATA DISPERZIJA:

$$H_0 \dots \mu = \bar{d} \rightarrow \bar{x} = \bar{y}$$

ZADANO  
DA JE BILO "NETKO BOJI" ONDA

BI BILO  $<>$

$$H_1 \dots \mu \neq \bar{d}$$

$$\hat{u} = \frac{\bar{x} - \bar{y}}{\sigma_z} \quad \sigma_z^2 = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m} = 0.49 \rightarrow \text{KORJENOVATI!}$$

$$\Rightarrow \sigma_z = \sqrt{0.49}$$

$$\Rightarrow \hat{u} = 1.32$$

→ PRIHVACAMO  $H_0$  → "da NE postoji razlika → ISTI SU →  $\mu = \bar{d}$  →  $H_0$  !

$$|\hat{u}| < u_{1-\frac{\alpha}{2}}$$

kriterij za ODBACIVANJE  $H_0$ :

$$|\hat{u}| > u_{1-\frac{\alpha}{2}}$$

$$1.32 < u_{1-\frac{\alpha}{2}} = 1.34$$

KADA PRIHVACAMO  $H_0$

ZADNJA TABLICA,

$$1 - \frac{\alpha}{2} = 0.91$$

(ZADANO) **MJENJAMO**

BR0J OD 1.32

$$\alpha = 18\%$$

**ZNAM** !



21-11-6.) U dvije trgovine <sup>istog lanca</sup> prodaje se čokolada  $\pi$ . Prodaja u jednom tjednu bila je:

A	12	9	10	8	7	13	11
B	10	8	9	7	12	10	7

Može li se zaključiti uz nivo značajnosti 10% da se prosječna dnevna prodaja čokolade u ove dvije trgovine bitno NE razlikuje?

→ NEPOZNATE DISPERZIJE:

$$H_0: \mu = 0$$

$$\bar{x} = \frac{12+9+\dots+11}{7} = 10 \quad (A)$$

$$H_1: \mu \neq 0$$

$$\bar{y} = 9 \quad (B)$$

→ službene formule:

$$\left. \begin{aligned} S_x^2 &= \frac{28}{6} \\ S_y^2 &= \frac{20}{6} \end{aligned} \right\} S_z^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2} = 4$$

PRI ČEMU JE  $n=m=7$

$$\hat{t} = \frac{\bar{x} - \bar{y}}{\underbrace{S_z}_{2}} \sqrt{\frac{nm}{n+m}} = 0.935 \Rightarrow S_z = \sqrt{4} = 2 !$$

$$|\hat{t}| > t_{n+m-2, 1-\frac{\alpha}{2}}$$

$$t_{12, 0.1} = 1.782$$

$0.935 > 1.782$  ? (NE)!  $\Rightarrow$  NE ODBACUJEMO  $\Rightarrow$  PRIHVATAMO  $H_0$  !



→ Uz nivo značajnosti 10% ne možemo tvrditi da postoji razlika u prodaji.

→ KADA BURIC POSTANE NOSITEJ NA VISU TRAZIT CE TU RECENICU. 11

## 12.4. TEST PRILAGODBE RAZDIOBAMA ( $\chi^2$ -test)

→ službeni podsjetnik - SVE! 11

21-10.-6.) Bacamo 3 kocke 100 puta i bilježimo broj šestica.  
Dobili smo:

NIJEDNA 6 → 0	61
1	29
2	8
~ SVE 3 ŠESTICE → 3	2

→ TAKO ZADANO

Provjerite ispravnost kocka uz nivo značajnosti 5%.



TEORIJSKI KOLIKO SMO TREBALI DOBITI → USPOREDIVATI

$x_i$	$n_i$	$p_i$	$n \cdot p_i$
0	61	$(\frac{5}{6})^3$	57.87
1	29	$\binom{3}{1} \frac{1}{6} \cdot (\frac{5}{6})^2$	34.72
2	8	$\binom{3}{2} (\frac{1}{6})^2 (\frac{5}{6})$	6.94
3	2	$(\frac{1}{6})^3$	0.46
$\Sigma$	$n=100$	1	$\approx 100$

ZAPRAVO BINOMINALNA RAZDIOBA

$61 \cdot (\frac{5}{6})^3$

teorijski je trebalo toliko puta pasti 0 šestica (a palo je 61 → CLOSE ENOUGH!)

empirijski

7.4

SLUŽBENI PODSJETNIK VELI:

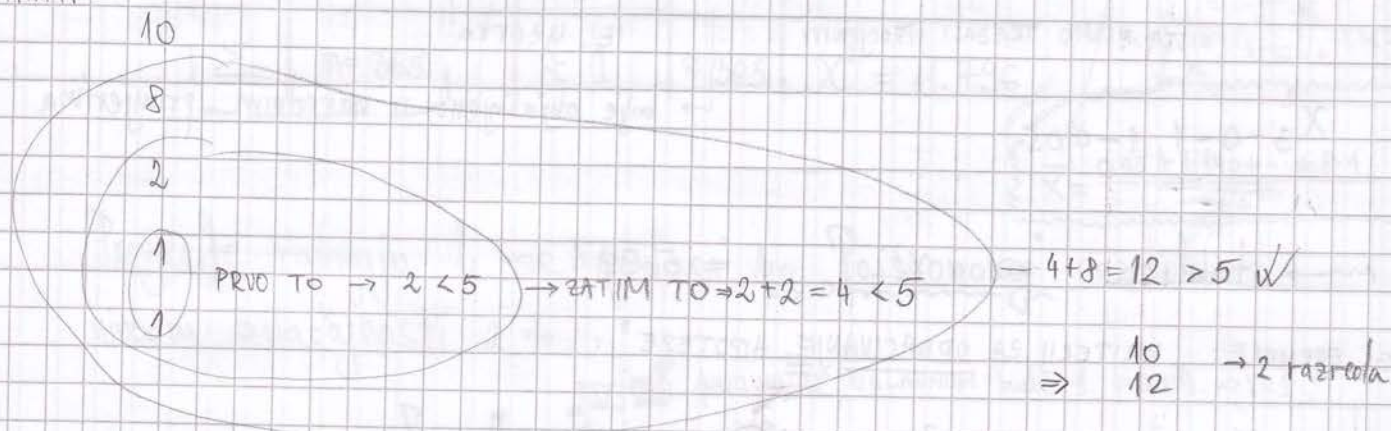
"razrede za koje je  $n_i < 5$  spajamo s njima susjednim"

→ u konačnoj sumi

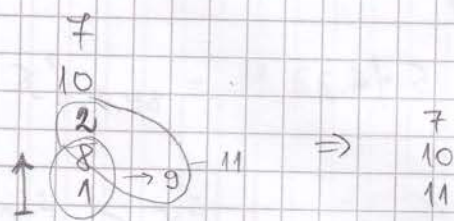
→  $p_i$ -ove računamo BEZ SPAJANJA, NA KRAJU TEK SPAJAMO!

\*DA SMO IMALI:

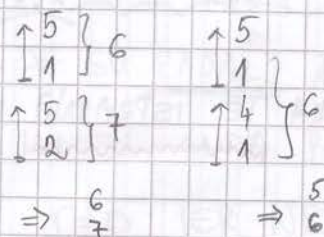
Prim 1.



Prim 2.



Prim 3.



UVJEK ODOZDOLA, PAZI NA 2 ( $< 5$ ), SVE RAZREDE  $< 5$  MORAMO SPOJITI, S NJIMA SUSJEDNIMA!



$$\frac{(n_j - n p_j)^2}{n p_j} \rightarrow \text{službeni podsjetnik}$$

$$0.169$$

$$0.942$$

$$0.913$$

suma

$$\chi^2 = 2.024 \rightarrow \text{mora biti što manji da bi prihvatili } H_0$$

$H_0$  ..... locka je ispravna (ravna se po binomnoj)

$H_1$  ..... suprotno

$$\chi^2_{m-r-1, 1-\alpha}$$

$m$  ..... broj razreda nakon spajanja

$r$  ..... broj parametara razdiobe procijenjenih iz uzorka

NIŠTA NISMO TREBALI PROCJENITI

$$\chi^2_{3-0-1, 1-0.05}$$

→ bolje objašnjeno u narednim primjerima!

→ tablica  $\Rightarrow \chi^2_{2, 0.95} = 5.991$

SL. FORMULE: "KRITERIJ ZA ODBACIVANJE HIPOTEZE":

$$\chi^2_{2, 0.95} < \chi^2 \Rightarrow \text{PRIHVATAMO } H_0 !$$

5.991      2.204      ? (NE)

⇒ ROCKA JE ISPRAVNA UZ NIVO ZNAČAJNOSTI 5%.



21-07-6.)

Tijekom godine bilježen je dnevni broj krivo spojenih poziva u nekoj centrali. Dobiveno je:

SAMI IZRACUNATI

ZADANO U  
ZADATKU

$x_j$	$n_j$	$p_j$	$n \cdot p_j$	$\frac{(n_j - n \cdot p_j)^2}{n \cdot p_j}$
0	95	0.2556	93.31	0.031
1	140	0.3486	127.25	1.277
2	73	0.2379	86.777	2.187
3	31	0.1081	39.45	1.81
4	16	0.03687	13.45	0.483
5	7	0.01005	3.67	
6	2	0.0023	0.83	6.008
7	0	0.00045	0.163	
8	1	0.00076	0.028	
$\Sigma$	$n=365$	$\approx 1$	$\approx 365$	$\chi^2 = 11.796$

\* PODSETNIK:

Poisson!

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$E(X) = \lambda = \bar{x}$$

$$\lambda = 1.364$$

$$0.2556 = \{X=0\} =$$

$$= \frac{\lambda^0}{0!} e^{-\lambda} =$$

$$= e^{-1.364}$$

$$0.3486 = \{X=1\} =$$

$$= \frac{\lambda^1}{1!} e^{-\lambda} = 1.364 e^{-1.364}$$

$$\bar{x} = \frac{0.95 + 1 \cdot 140 + \dots + 8 \cdot 1}{365}$$

Provjerite ravnaju li se podaci po Poissonovoj razdiobi s nivoom značajnosti 2%.

provjeriti parametar, uglavnom bude 1 param.  $\Rightarrow r=1!$

$$\chi^2_{m-r-1, 1-\alpha} = \chi^2_{6-1-1, 1-0.02} =$$

$$= \chi^2_{4, 0.98} = 11.668$$

$$\chi^2_{11.796} > \chi^2_{11.668, 0.98} ? \text{ (DA!) } \Rightarrow \text{ Odbacujemo } H_0 !$$

$\Rightarrow$  U nivo znač. 2% ne možemo zaključiti da se ravna po Poissonu!

PAZI!TRIKDA SU ZADALI  $\lambda = 1.364$ 

ILI SL. ONDA BI

r=0 JER NISTO

PROVJERJIVALI!

BILO PAR PUTA

NA ISPITU!



SAMI RAČUNAMO

Z1-11-10.)

ZADANO

SAMI RAČUNAMO

	$[a, b]$	$n_j$	$p_j$	$n \cdot p_j$	$\frac{(n_j - n \cdot p_j)^2}{n \cdot p_j}$
12.5	0-25	87	0.613	91.99	0.271
37.5	25-50	46	0.237	35.576	3.054
62.5	50-75	14	0.092	13.76	0.227
87.5	75-100	3	0.035	5.32	
$\Sigma$		$n=150$	$\approx 1$	$\approx 150$	$\chi^2 = 3.552$

Provjerite ravnaju li se dobiveni podaci po eksponencijalnoj razdiobi uz nivo značajnosti  $\alpha = 5\%$ .

MORAMO ZNATI!

\* PODSJETNIK: EKSPONENCIJALNA:

$$P(a < X < b) = P(b) - P(a) = 1 - e^{-b\lambda} - (1 - e^{-a\lambda}) = e^{-a\lambda} - e^{-b\lambda}$$

$$\lambda = \frac{1}{\bar{X}}$$

$$* E(X) = \frac{1}{\lambda}$$

RAČUNAMO SREDINU SVAKOG INTERVALA (LIJEVO SKROZ)

$$\text{KAO } [a, b] \rightarrow \text{sredina} = \frac{a+b}{2}$$

$$\Rightarrow \lambda = \frac{1}{26.33}$$

$$\bar{X} = 87 \cdot 12.5 + 46 \cdot 37.5 + \dots + 3 \cdot 87.5$$

$$\bar{X} = 26.33$$

$$\lambda = 0.038$$

r=1 → procijenili parametar  $\lambda$ 

$$\chi^2_{m-r-1, 1-\alpha} = \chi^2_{3-1-1, 1-0.05} = \chi^2_{1, 0.95} = 3.841$$

$$\chi^2 = 3.552 < 3.841 = \chi^2_{1, 0.95} \Rightarrow \text{Prihvaćamo } H_0!$$

Dobiveni podaci se ravnaju po eksponencijalnoj razdiobi.



12.72V-26.) U prvih 800 decimala broja  $\pi$  znamenke se pojavljuju:

ZADANO  $\rightarrow$

$x_j$	$n_j$	$p_j$	$n p_j$	$\frac{(n_j - n p_j)^2}{n p_j}$
0	74	$\frac{1}{10}$	80	0.45
1	92	$\frac{1}{10}$	80	1.8
2	83	$\frac{1}{10}$	80	0.1125
3	79	.	.	.
4	80	.	.	.
5	73	.	.	.
6	77	.	.	.
7	75	.	.	.
8	76	.	.	0.2
9	91	$\frac{1}{10}$	80	1.5152
$\Sigma$	$n=800$	1	800	$\chi^2 = 5.152$

Provjerite da li se znamenke ravnaju po jednolikoj razdiobi (pojava svake znamenke je jednako vjerojatna) uz nivo značajnosti 10%.

$r=0 \rightarrow$  NIŠTA NE PROCJENJUJEMO!

$$\alpha = 0.1$$

$$\chi^2_{10, m-r-1, 1-\alpha} = \chi^2_{9, 0.9} = 14.684$$

$$5.152 < 14.684$$

$\Rightarrow$  Prihvaćamo  $H_0$ , odnosno pojava svake znamenke jednako je vjerojatna!

LJEPI  $\pi$  ZADATAK ZA KRAJ DRUŽENJA SA ŽPIT-om

