

NEPARAMETRIČNI TESTI

pregledovanje tabel
hi-kvadrat test

as. dr. Nino RODE

Parametrični in neparametrični testi

- S pomočjo z-testa in t-testa preizkušamo **domneve o parametrih**
 - na vzorcih izračunamo **statistike**, ki nas zanimajo
 - preverjamo **verjetnost**, da veljajo med **parametri** populacije odnosi, ki jih smo jih predpostavili
- Z **neparametričnimi** testi preizkušamo domneve o **celotnih porazdelitvah**
 - v vzorcu ugotovimo določeno **porazdelitev** podatkov
 - preverjamo, kakšna je **verjetnost**, da je v populaciji porazdelitev, ki jo **predpostavljamo**

Neparametrini testi

primer luninih men

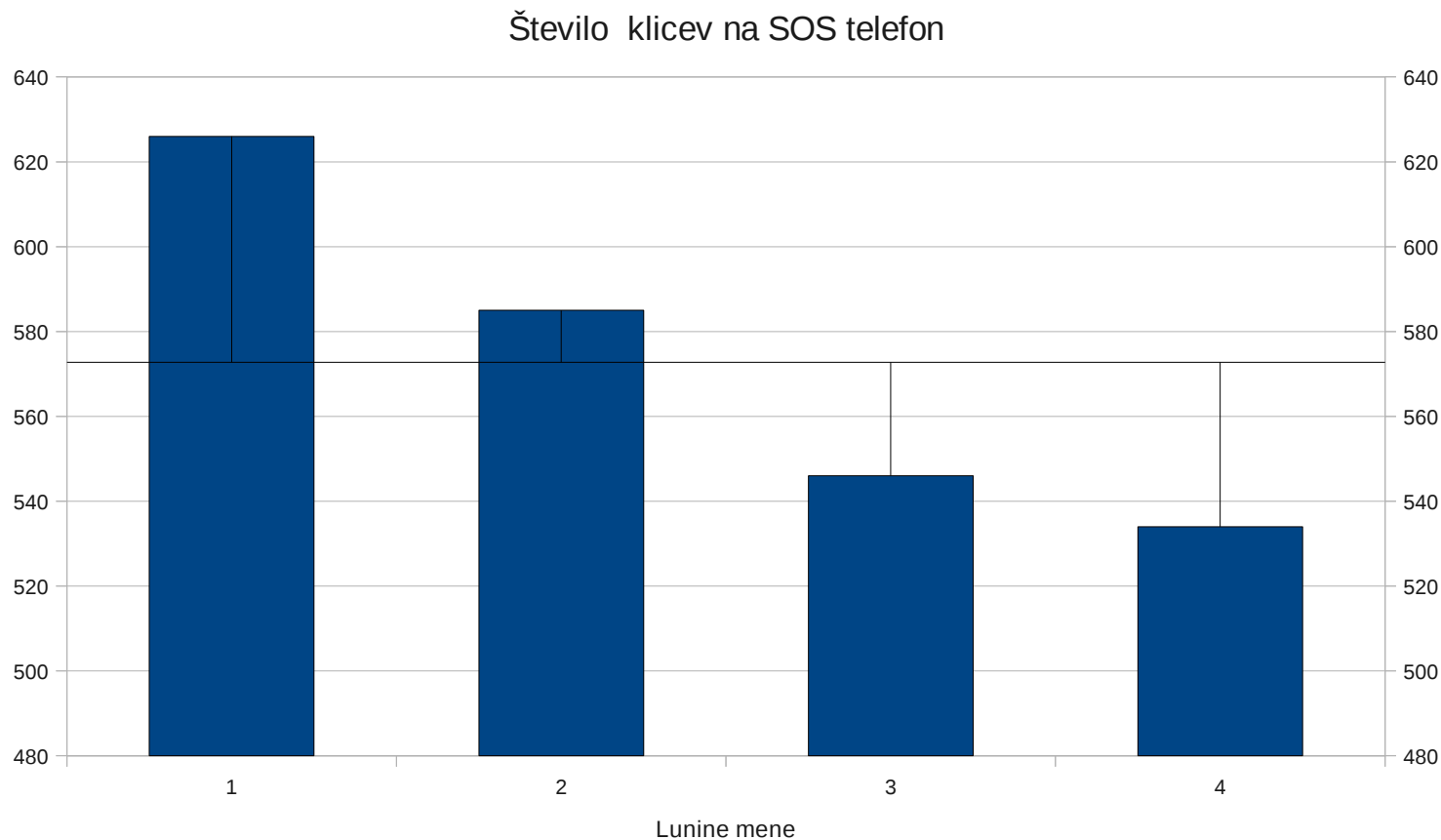
- Polna luna vpliva na obnašanje in počutje ljudi
- Ali **lunine mene** vplivajo na **število klicev** na SOS telefon?

Lunina mena	Število klicev
1	626
2	585
3	546
4	534
Skupaj	2291

Neparametrini testi

primer luninih men

- Če ne ni vpliva luninih men, bi (**teoretično**) pričakovali, da so **frekvence** pri vseh menah enake, skupne razlike pa le **slučajne** (**majhne**)



Neparametrini testi

primer luninih men

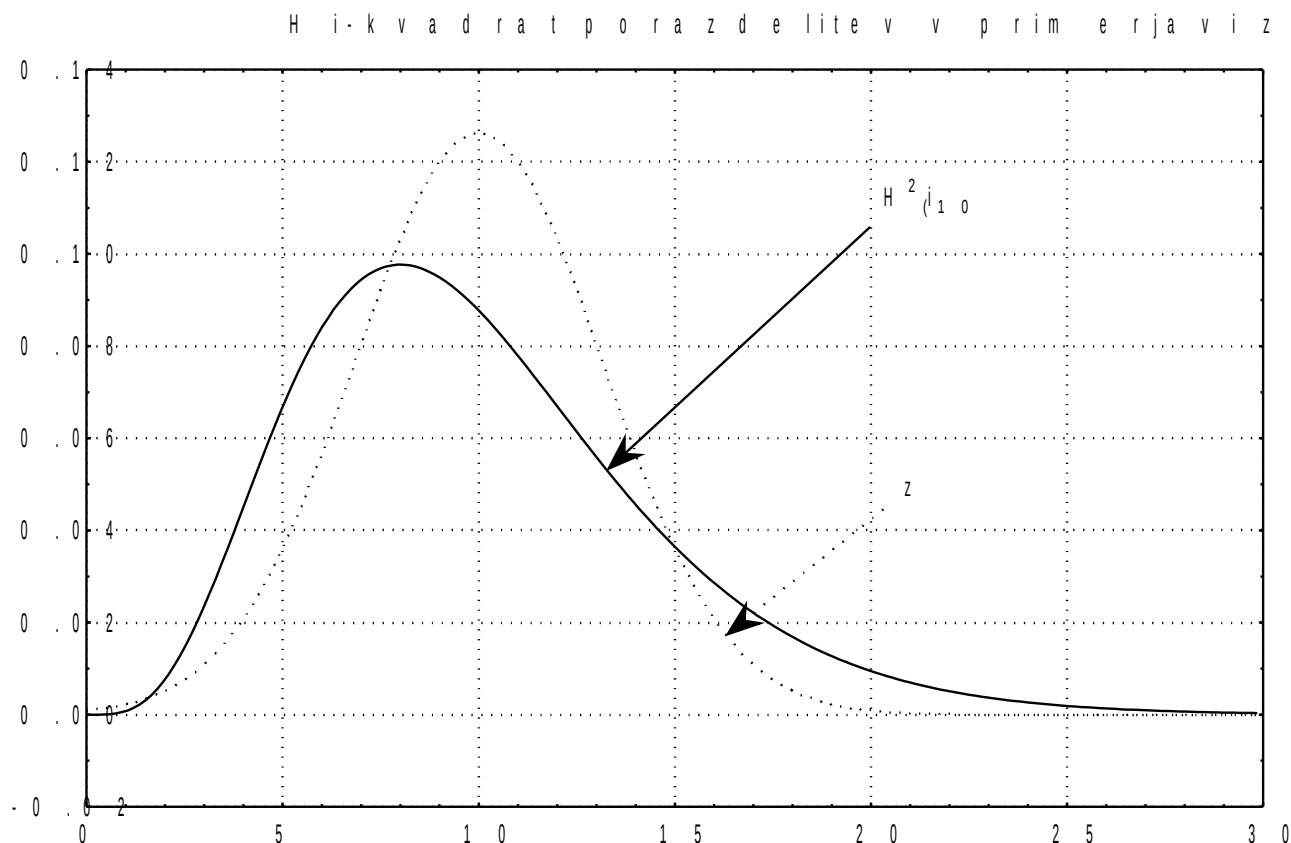
- Odkloni se med sabo uničijo, zato (kot ponavadi) uporabimo **kvadrate odklonov**

	Odkloni	Kvadrati odklonov
1	53,25	2835,56
2	12,25	150,06
3	-26,75	715,56
4	-38,75	1501,56
Skupaj	0	5202,75

Neparametrini testi: χ^2 porazdelitev

χ^2 porazdelitev izhaja iz normalne porazdelitve
seštevek kvadratov z-vrednosti:

$$\chi_{sp}^2 = \sum_{i=1}^{sp} z_i^2$$



Neparametrini testi: χ^2 test

Pri χ^2 testu praviloma **primerjamo opaženo porazdelitev** podatkov s **teoretično porazdelitvijo**

Preverjamo, ali sta si dve porazdelitvi **podobni**:

Če sta si porazdelitvi podobni, bodo **razlike** med njima praviloma **majhne**

Seštevek kvadratov razlik se bo porazdeljeval po **χ^2 porazdelitvi**

Manj ko sta si dve porazdelitvi podobni, večje so razlike med njima, večja je **χ^2 statistika**

Neparametrini testi

primer luninih men

- Večje teoretične frekvence dovoljujejo večje napake.

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

	Odkloni	Kvadrati odklonov	Elementi X2
1	53,25	2835,56	4,95
2	12,25	150,06	0,26
3	-26,75	715,56	1,25
4	-38,75	1501,56	2,62
Skupaj	0	5202,75	9,08

Neparametrini testi

primer luninih men

- **Izračunano** χ^2 vrednost primerjamo s **kritičnimi vrednostmi** iz tabele χ^2 porazdelitve
- Stopinje prostosti izračunamo kot število vseh celic tabele minus število seštevkov (v našem primeru $4 - 1 = 3$)
- Kritična vrednost za pri tveganju 0,05 in 3 stopinjah prostosti je 7,815

$$\chi^2_{izracun} = 9,08 > 7,815 = \chi^2_{0,05;3}$$

Neparametrini testi

primer luninih men

- Z manj kot 0,05 (5%) tveganja lahko trdimo, da obstajajo razlike v številu klicev glede na lunine mene (razlike med številom klicev po luninih menah so statistično različne na nivoju 0,05)

TODA

- 1 Prvi krajec
- 2 Polna luna
- 3 Mlaj
- 4 Zadnji krajec

χ^2 test: primer dvodimenzionalne tabele

- Na Stigmi smo dobili podatke o tem ali so obiskovalci redni, ali pa so se oglasili le enkrat

	Večkrat	Enkrat	Skupaj
Moški	257	536	793
Ženske	120	117	237
Skupaj	377	653	1030

- S χ^2 testom bomo preverili, ali so razlike med spoloma glede rednosti obiskovanja Stigme

χ^2 test: primer dvodimenzionalne tabele

- V tabeli so izračunani odstotki za moške in ženske. Med moškimi je več enkratnih obiskovalcev

	Večkrat	Enkrat	Skupaj
Moški	32,4%	67,6%	100,0%
Ženske	50,6%	49,4%	100,0%
Skupaj	36,6%	63,4%	100,0%

- Odstotki kažejo razliko, vendar je vprašanje, če je ta razlika dovolj velika

χ^2 test: primer dvodimenzionalne tabele

- Pri χ^2 testu praviloma primerjamo **empirično tabelo**, ki smo jo dobili iz podatkov s **teoretično tabelo**, ki bi jih dobili, če bi veljali določeni pogoji
- V našem primeru nas zanima tabela, v kateri **med spoloma ni razlik**
- Teoretične frekvence v takšni tabeli izračunamo iz **robnih seštevков** (robnih frekvenc) tabele

$$f' = \frac{f_{\text{vrstice}} \cdot f_{\text{stolpca}}}{n}$$

χ^2 test: primer dvodimenzionalne tabele

- Teoretične frekvence in odstotki izračunani iz njih so prikazani v spodnjih tabelah

	Večkrat	Enkrat	Skupaj
Moški	290,25	502,75	793
Ženske	86,75	150,25	237
Skupaj	377	653	1030

	Večkrat	Enkrat	Skupaj
Moški	36,6%	63,4%	100,0%
Ženske	36,6%	63,4%	100,0%
Skupaj	36,6%	63,4%	100,0%

χ^2 test: primer dvodimenzionalne tabele

- Formula za χ^2 je takšna:
$$\chi^2 = \sum \frac{(f_o - f')^2}{f'}$$
- Pri χ^2 testu praviloma tabelo razvijemo v stolpec, da lažje računamo

	f_o	f'	$f_o - f'$	$(f_o - f')^2$	$(f_o - f')^2 / f'$
A	257	290,25	-33,25	1105,79	3,81
B	536	502,75	33,25	1105,79	2,2
C	120	86,75	33,25	1105,79	12,75
D	117	150,25	-33,25	1105,79	7,36
				$\chi^2 =$	26,12

χ^2 test: primer dvodimenzionalne tabele

- Kot pri vsakem testu, primerjamo iz podatkov izračunani χ^2 s tabelirano **kritično vrednostjo** pri danih **stopinjah prostosti**

$$sp = (\text{št.vrstic} - 1) \cdot (\text{št.stolpcev} - 1) = (2 - 1) \cdot (2 - 1) = 1 \cdot 1 = 1$$

$$\chi^2 = 26,12 > 3,84 = \chi^2_{1;0,05}$$

- Odgovor: Rednost obiskovanja Stigme se med spoloma statistično značilno razlikuje pri stopnji tveganja 0,05. Med moškimi je več enkratnih obiskovalcev, kot med ženskami