



Hi-kvadrat test nezavisnosti

Statistički metodi 2018/2019

dr Aleksandar Tomašević

18.12.2019.

Filozofski fakultet u Novom Sadu

1. Nezavisnost kvalitativnih varijabli
2. Teorijske i empirijske frekvencije
3. Hi-kvadrat statistika
4. Hi-kvadrat test nezavisnosti
5. Intenzitet asocijacije/zavisnosti

Nezavisnost kvalitativnih varijabli

- Kod **kvalitativnih** varijabli poznata nam je samo **učestalost** odgovora (*modaliteta*)

- Kod **kvalitativnih** varijabli poznata nam je samo **učestalost** odgovora (*modaliteta*)
- Ako su varijble **nezavisne**, učestalost modaliteta jedne varijble nije povezana sa učestalošću modaliteta druge!

Primer dve kvalitativne varijable

Tabela 1: Pol ispitanika

Muškarci	Žene
50	50

Tabela 2: Da li planirate da glasate na narednim parlamentarnim izborima?

Da	Ne
60	40

	Da	Ne
Muškarci	35	15
Žene	25	25

Teorijske i empirijske frekvencije

- Teorijske frekvencije su one koje bismo **očekivali** kada bi varijable bile nezavisne.

- Teorijske frekvencije su one koje bismo **očekivali** kada bi varijable bile nezavisne.
- Frekvencije modaliteta jedne varijable nije uslovljena (ograničena) frekvencijama modaliteta druge varijable.

- Teorijske frekvencije su one koje bismo **očekivali** kada bi varijable bile nezavisne.
- Frekvencije modaliteta jedne varijable nije uslovljena (ograničena) frekvencijama modaliteta druge varijable.
- Ravnomerna distribucija modaliteta po poljima tabele krostabulacije

Razlike između distribucije frekvencija

Tabela 3: Empirijske frekvencije

	Da	Ne
Muškarci	35	15
Žene	25	25

Tabela 4: Teorijske frekvencije

	Da	Ne
Muškarci	30	20
Žene	30	20

Hi-kvadrat statistika

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

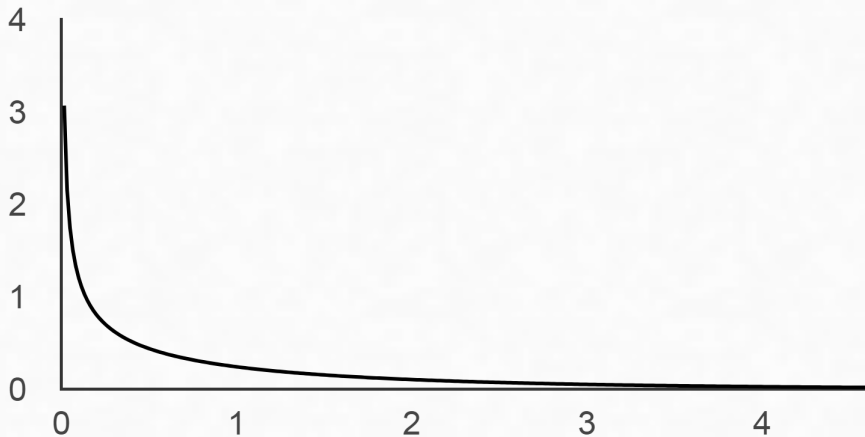
$$\chi^2 = \frac{(35 - 30)^2}{30} + \frac{(15 - 20)^2}{20} + \frac{(25 - 30)^2}{30} + \frac{(25 - 20)^2}{20}$$

Hi-kvadrat test nezavisnosti

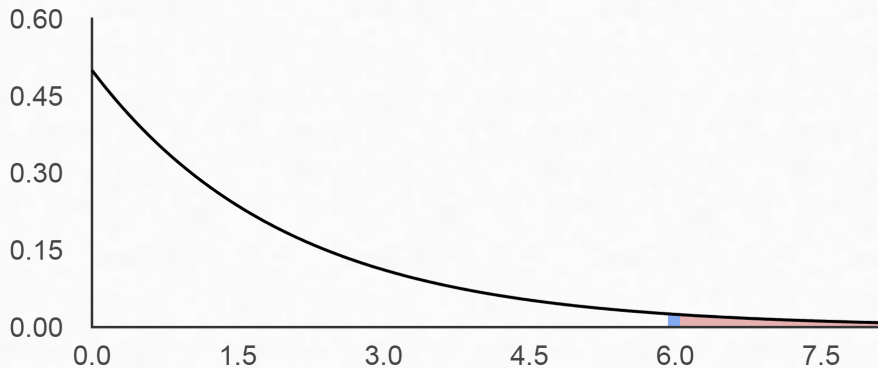
H_0 : *Variable su nezavisne*

H_1 : *Variable su zavisne*

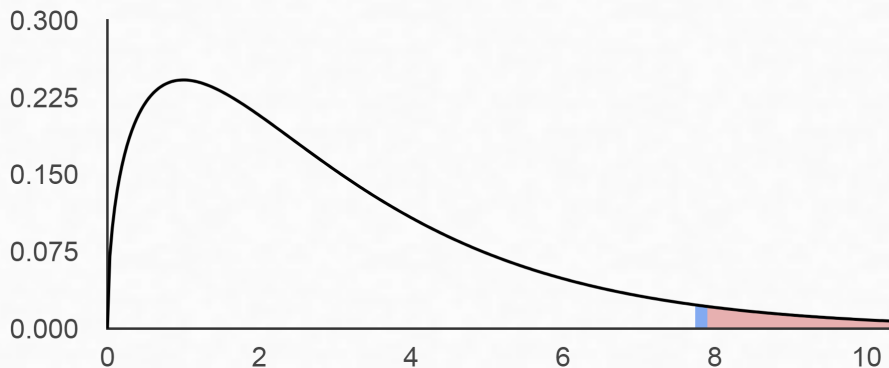
Hi-kvadrat distribucija — $d_f = 1$



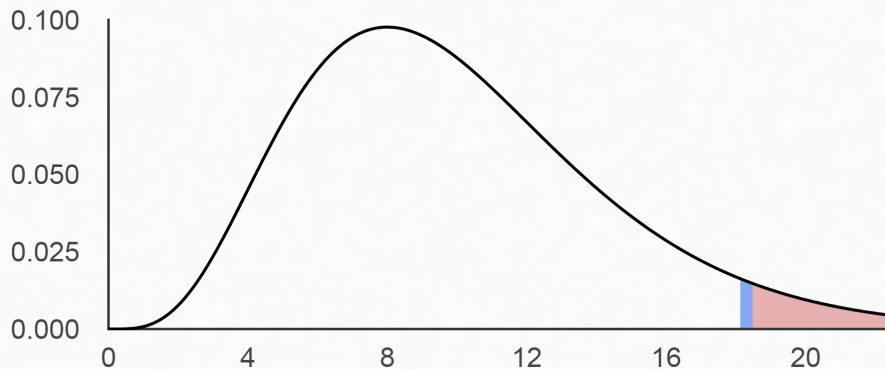
Hi-kvadrat distribucija — $d_f = 2$



Hi-kvadrat distribucija — $d_f = 3$



Hi-kvadrat distribucija — $d_f = 10$



$$\chi^2 > \chi_{d_f; \alpha}^2$$

$$d_f = (k - 1)(r - 1)$$

Intenzitet asocijacije/zavisnosti

$$V = \sqrt{\frac{\chi^2}{N(k-1)}}$$