

## Usporedba kvalitativnih podataka

Nora Nikolac  
Klinički zavod za kemiju  
KB "Sestre milosrdnice"

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu  
Biostatistika  
3.-5.12. 2008.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

## Kvalitativni podaci

- Nominalni - ne uspoređuju se!

Spol

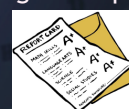


M = 1

Ž = 0

- Ordinalni - mogu se uspoređivati!

Ocjene



= 1, 2, 3, 4, 5

= A, B, C, D, E, F



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

## Statistička hipoteza - postupak testiranja

1. postavljanje hipoteze:  $H_0$ ,  $H_1$
2. odabir statističkog testa
3. određivanje razine značajnosti:  $\alpha$
4. izračunavanje
5. zaključivanje: P



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

## 1. Postavljanje hipoteze

Nulta hipoteza  $H_0$  - **nema** statistički značajne razlike među skupinama

Ukoliko se odbaci nulta hipoteza, prihvaća se alternativna hipoteza

Alternativna hipoteza  $H_1$  - **postoji** statistički značajna razlika među skupinama



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

## 2. Odabir statističkog testa

- Za pravilan odabir testa moramo poznavati podatke koje testiramo:
  - Parni/Neparni uzorci
  - Broj skupina
  - Veličina uzorka



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 3. Određivanje razine značajnosti

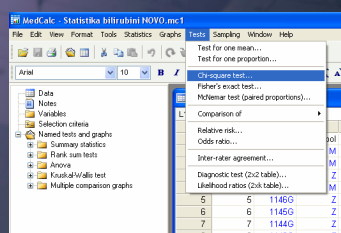
- Najčešće:  $\alpha$  - razina značajnosti: 0,05
- Kod preliminarnih istraživanja može se odabrati veća razina značajnosti: 0,1 ili 0,2
- Kod pojedinih istraživanja razina značajnosti može biti i manja: 0,01 ili 0,001

Značenje razine značajnosti: vjerojatnost odbacivanja istinite  $H_0$ .



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 4. Izračunavanje



Statistički programi



Ručno



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 5. Zaključivanje

$P > 0,05$  (ili druga postavljena razina značajnosti)



**Prihvaćamo** nultu hipotezu i zaključujemo:  
**Ne postoji** statistički značajna razlika među skupinama



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 5. Zaključivanje

$P < 0,05$  (ili druga postavljena razina značajnosti)



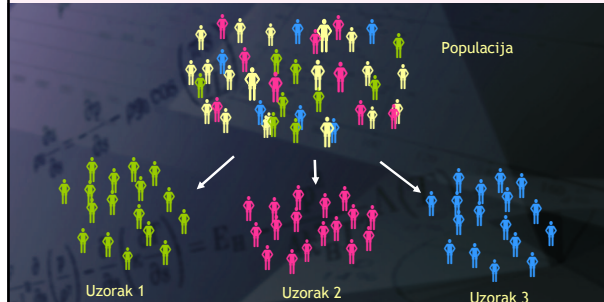
**Odbacujemo** nultu hipotezu, prihvaćamo alternativnu hipotezu i zaključujemo:

**Postoji** statistički značajna razlika među skupinama



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Nezavisni uzorci



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## $\chi^2$ test

Ispitivanje razine značajnosti razlike između:

- Očekivane razdiobe podataka i dobivenih podataka (npr. genetika);
- Dviju ili više skupina nezavisnih kvalitativnih podataka.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Ručni izračun parametara $\chi^2$ testa

1. Izračun očekivanih frekvencija
2. Izračun  $\chi^2$  vrijednosti
3. Određivanje stupnjeva slobode (DF, engl. *degree of freedom*)
4. Iščitavanje  $\chi^2$  vrijednosti za određenu razinu značajnosti
5. Zaključivanje



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Kontingencijska tablica

Ispitali smo razliku prema spolu između dvije skupine ispitanika. U prvoj je skupini od 120 ispitanika bilo 30% muškaraca, a u drugoj od 140 ispitanika 60% žena.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Kontingencijska tablica

|          | Skupina 1 | Skupina 2 | Ukupno |
|----------|-----------|-----------|--------|
| Muškarci | 36        | 56        | 92     |
| Žene     | 84        | 84        | 168    |
| Ukupno   | 120       | 140       | 260    |



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 1. Izračun očekivanih frekvencija

Stvarne frekvencije

|          | Skupina 1 | Skupina 2 | Ukupno |
|----------|-----------|-----------|--------|
| Muškarci | 36        | 56        | 92     |
| Žene     | 84        | 84        | 168    |
| Ukupno   | 120       | 140       | 260    |

|  |  |
|--|--|
| M skupina 1<br>$0,354 \times 120 = 42$ | M skupina 2<br>$0,354 \times 140 = 50$ |
| Ž skupina 1<br>$0,646 \times 120 = 78$ | Ž skupina 2<br>$0,646 \times 140 = 90$ |

|                               |                                |
|-------------------------------|--------------------------------|
| M ukupno<br>$92/260$<br>0,354 | Ž ukupno<br>$168/260$<br>0,646 |
|-------------------------------|--------------------------------|

Očekivane frekvencije

|          | Skupina 1 | Skupina 2 | Ukupno |
|----------|-----------|-----------|--------|
| Muškarci | 42        | 50        | 92     |
| Žene     | 78        | 90        | 168    |
| Ukupno   | 120       | 140       | 260    |



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## 2. Izračun $\chi^2$ vrijednosti

$$\chi^2 = \sum [(f_{\text{stvarna}} - f_{\text{očekivana}})^2 / f_{\text{očekivana}}]$$

|               | f(stv) | f(oček) | f(stv) - f(oček) | [f(stv) - f(oček)] <sup>2</sup> | [f(stv) - f(oček)] <sup>2</sup> / f(oček) |
|---------------|--------|---------|------------------|---------------------------------|---|
| M (Skupina 1) | 36     | 42      | -6               | 36                              | 0,86                                      |
| M (Skupina 2) | 56     | 50      | 6                | 36                              | 0,72                                      |
| Ž (Skupina 1) | 84     | 78      | 6                | 36                              | 0,46                                      |
| Ž (Skupina 2) | 84     | 90      | -6               | 36                              | 0,40                                      |

$$\chi^2 = 2,4$$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

### 3. Stupnjevi slobode

$$DF = (r-1) \times (s-1)$$

R = broj redaka; s = broj stupaca

$$DF = (2-1) \times (2-1) = 1$$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

### 4. Iščitanje $\chi^2$ vrijednosti iz tablice

Iz tablice iščitamo  $\chi^2$  vrijednost za odabranu razinu značajnosti i odgovarajući broj stupnjeva slobode.

Table of Chi-square statistics

| df | 0.05  | 0.01  | 0.001 |
|----|-------|-------|-------|
| 1  | 3.84  | 6.64  | 10.83 |
| 2  | 5.99  | 9.21  | 13.82 |
| 3  | 7.82  | 11.35 | 16.27 |
| 4  | 9.49  | 13.28 | 18.47 |
| 5  | 11.07 | 15.09 | 20.52 |
| 6  | 12.59 | 16.81 | 22.46 |

$$\alpha = 0,05$$

$$DF = 1$$

$$\chi^2(\alpha) = 3,84$$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

### 5. Zaključivanje

Ako je  $\chi^2 < \chi^2(\alpha)$

Prihvaćamo  $H_0$

Ako je  $\chi^2 > \chi^2(\alpha)$

Odbacujemo  $H_0$   
Prihvaćamo  $H_1$

Ne postoji statistički značajna razlika u raspodjeli muškaraca i žena između skupina 1 i 2.

$$\chi^2 = 2,4 < \chi^2(\alpha) = 3,84$$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

### A može i jednostavnije...

$$P = 0,1209 (P > 0,05)$$

Prihvaćamo  $H_0$  i zaključujemo da nema statistički značajne razlike među skupinama.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Dodatni uvjeti $\chi^2$ testa

Sve očekivane frekvencije moraju biti  $> 1$   
 Najviše 20% očekivanih frekvencija  $< 5$   
 Ukupan broj uzoraka  $> 100$

|    |    |
|----|----|
| 28 | 39 |
| 2  | 45 |

2 x 2 tablica  
 4 očekivane frekvencije  
 $20\% = 0,2 \times 4 = 0,8 < 1$

|    |    |    |
|----|----|----|
| 58 | 12 | 24 |
| 23 | 4  | 14 |
| 12 | 2  | 18 |

3 x 3 tablica  
 9 očekivanih frekvencija  
 $20\% = 0,2 \times 9 = 1,8 > 2$

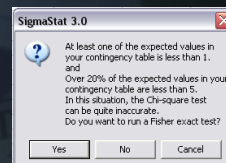


Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Primjer pogrešne upotrebe testa

Postoji li razlika u učestalosti pojave bolesti između ispitanika koji su primili cjepivo i onih koji nisu?

|             | Cijepljenje - | Cijepljenje + |
|-------------|---------------|---------------|
| Bez bolesti | 65            | 120           |
| Bolest      | 7             | 3             |



$\chi^2$  test:  $P = 0,059$

Fisherov egzaktni test:  $P = 0,040$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Fisherov egzaktni test

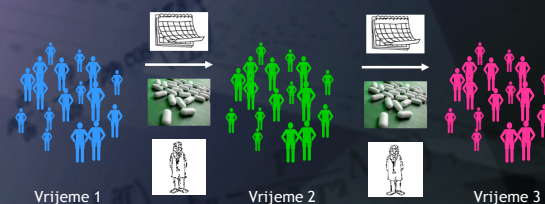
Test za ispitivanje razine značajnosti razlike između dviju skupina nezavisnih kategoričkih podataka.

Koristi se kod malih uzoraka ( $N < 100$ ).  
 Očekivane frekvencije mogu biti  $\leq 1$ .



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Zavisni uzorci



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika



## Test

**Parni  $\chi^2$  test (McNemarov test):** ispitivanje razine značajnosti razlike između dvije skupine zavisnih kategoričkih podataka.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## McNemarov test

Utječe li kava na stanje pospanosti studenata 1. godine FBF-a na kolegiju "Biostatistika"?

|                      |    | Pospanost nakon kave |     | Ukupno |
|----------------------|----|----------------------|-----|--------|
|                      |    | DA                   | NE  |        |
| Pospanost prije kave | DA | 30                   | 80  | 110    |
|                      | NE | 12                   | 50  | 62     |
| Ukupno               |    | 42                   | 130 | 172    |



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## McNemarov test

$$\chi^2 = \frac{(|B-C| - 1)^2}{B + C} = \frac{(|80-12| - 1)^2}{80 + 12} = \frac{67^2}{92} = 48,8$$

**B** = prije intervencije **POZ**,  
nakon intervencije **NEG**

**C** = prije intervencije **NEG**,  
nakon intervencije **POZ**



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Izračunavanje

Table of Chi-square statistics

| df | 0.1   | 0.05  | 0.01  | 0.001 |
|----|-------|-------|-------|-------|
| 1  | 3.84  | 3.84  | 6.63  | 10.83 |
| 2  | 4.61  | 5.99  | 9.21  | 13.82 |
| 3  | 6.25  | 7.88  | 11.35 | 16.27 |
| 4  | 7.78  | 9.49  | 13.28 | 18.47 |
| 5  | 9.49  | 11.07 | 15.09 | 20.52 |
| 6  | 11.07 | 12.59 | 16.81 | 22.46 |

$\alpha = 0,05$

DF = 1

$$\chi^2(\alpha) = 3,84 < \chi^2 = 48,8$$

Odbacujemo  $H_0$   
Prihvaćamo  $H_1$

Postoji statistički značajna razlika u pospanosti prije i poslije ispijanja kave.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Zadaci



1. Ispitali smo postoji li statistički značajna razlika u učestalosti pohađanja predavanja između studenata prve i druge godine studija FBF-a. Ishod smo klasificirali kao **redovito pohađa** ili **ne pohađa redovito**. Kojim ćemo testom ispitati značajnost razlike smo:

a) Ispitali status redovitog pohađanja predavanja na 100 studenata prve i 100 studenata druge godine?

Nezavisni kvalitativni podaci (Hi-kvadrat test, Fisherov test)

b) Ispitali status redovitog pohađanja predavanja na 100 studenata prve godine, te nakon godinu dana ponovno ispitati status redovitosti?

Zavisni kvalitativni podaci (Parni hi-kvadrat test)



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Zadaci



2. Prikazane su tablice očekivanih frekvencija za nezavisne kvalitativne podatke. Možemo li ove podatke usporediti hi-kvadrat testom?

|    |    |    |
|----|----|----|
| 56 | 37 | 60 |
| 24 | 45 | 12 |
| 1  | 12 | 45 |

Očekivane frekvencije ne smiju biti  $\leq 1$

|    |    |
|----|----|
| 16 | 50 |
| 12 | 21 |

Ukupan broj mora biti  $> 100$

|    |    |    |
|----|----|----|
| 56 | 37 | 60 |
| 24 | 45 | 12 |
| 4  | 3  | 2  |

20% očekivanih frekvencija smije biti  $< 5$

|    |    |
|----|----|
| 16 | 50 |
| 12 | 21 |
| 2  | 24 |
| 12 | 45 |



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

## Pitanja

1. Postupak testiranja statističke hipoteze.
2. O čemu ovisi izbor statističkog testa?
3. Koje je značenje razine značajnosti  $\alpha$ ?
4. Kako interpretiramo rezultat testa obzirom na dobivenu P vrijednost?
5. Kada upotrebljavamo hi-kvadrat test (uvjeti, ograničenja)?
6. Kako se izračunava broj stupnjeva slobode u kontingencijskim tablicama?
7. Kako na osnovu hi-kvadrat vrijednosti za odabranu razinu značajnosti interpretiramo rezultat hi-kvadrat testa?
8. Kada upotrebljavamo Fisherov egzaktni test?
9. Kada upotrebljavamo parni hi-kvadrat test (McNamerov test)?
10. Koje frekvencije uzimamo u obzir kod izračuna parametara parnog hi-kvadrat testa?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika