

## Metodologija psiholoških istraživanja

### Obrada frekvencijski 3



13. novembar 2018.

#### A. Frekvencijski nacrti

- Univarijantni frekvencijski nacrti
  - organizacija podataka
  - deskriptivne statističke mere verovatnoća
  - prikaz rezultata
  - struktura i značajnost rezultata

UFN sa dve kategorije

## 1. Univarijantni frekvencijski nacrti (UFN)

- oblici nulte hipoteze (H0) u UFN sa dve kategorije
  - najčešći oblik H0: proporcije kategorija p1 i p2 su *jednake* (nulta razlika)
    - formalno:  $H_0: p_1 = p_2 = 0.5$ , odn.  $p_1 - p_2 = 0$
    - PRIMER: u populaciji, bebe bi *jednako* često birale cucle obe boje
  - znatno ređi oblik H0: p1 i p2 imaju *različite* date numeričke vrednosti
    - npr., to mogu biti 0.6 i 0.4, dakle njihova razlika *nije* nulta
    - PRIMER: predpostavljamo da u populaciji 60% beba preferira crvenu a 40% plavu boju (npr. na osnovu ranijih istraživanja)
- oblici alternativne hipoteze (Ha) u UFN sa dve kategorije
  - najčešći oblik Ha je *negacija* najčešće H0: proporcije p1 i p2 *nisu* jednake
    - formalno:  $H_a: p_1 \neq p_2$
    - to je tzv. *dvosmerna* forma Ha
    - PRIMER: preferencije beba za boje u populaciji *nisu* jednake
  - nešto ređa (i nešto preciznija) *jednosmerna* forma Ha:
    - $H_a: p_1 > p_2$  ili  $H_a: p_1 < p_2$
    - PRIMER: u populaciji bebe više vole crvenu od plave boje, ili obrnuto
    - postavlja se ako ima osnova da se pretpostavi *smer* očekivane razlike

## 1. Univarijantni frekvencijski nacrti (UFN)

- razlike između H0 i Ha:
  - H0 uvek specifikuje *numeričke* vrednosti proporcija p1 i p2
  - Ha najčešće *ne* specifikuje numeričke vrednosti proporcija
    - tvrdi se uglavnom samo da razlike *ima* (dvosmerna, nominalna forma), ili da je razlika u određenom *smeru* (jednosmerna, ordinalna forma)
  - numerička *preciznost* H0 omogućava njeno testiranje statističkim metodama, dok nepreciznost Ha to otežava
- kako se vrši testiranje H0 u UFN sa dve kategorije?
  - problem:** znamo da ako  $H_0: p_1 = p_2$  važi u *populaciji*, i ako je uzorak reprezentativan, to još ne znači da će u *uzorku* proporcije biti jednake
  - pitanje:** da li to znači da na osnovu uzorka ne možemo baš ništa zaključiti o populaciji?
  - odgovor:** *ne!*
    - naime, neki ishodi su više verovatni, a neki manje!*
      - npr., ako važi H0, više je verovatan ishod 51:49 nego 80:20
  - ključna činjenica:** verovatnoće različitih ishoda se mogu *izračunati* na osnovu teorije verovatnoće, ako su ispunjeni određeni uslovi

## 1. Univarijantni frekvencijski nacrti (UFN)

- osnovna ideja testiranja H0 (i veza sa teorijom verovatnoće)
  - pretpostavimo da H0 *važi* u populaciji, tj. da je  $p_1 = p_2$
  - zamislimo da se istraživanje *ponavlja* mnogo puta, u teoriji beskonačno
  - teoretski izračunamo verovatnoće svih *moću* ishoda istraživanja
  - utvrdimo koliko je verovatan ishod *našeg* konkretnog istraživanja
- ilustrativno poređenje: istraživanje i analogni primer

istraživanje	analogni primer
populacija sa objektima (na pr. bebama)	(ogromna) vreća sa klikerima
kategorička varijabla sa dve kategorije	u vreći su klikeri dve boje (crvena i plava)
u populaciji važi $H_0: p_1 = p_2$	u vreći ima <i>jednak broj</i> klikera obe boje
uzimanje uzorka iz populacije	slučajan izbor izvesnog broja (N) klikera
utvrđivanje frekvenci dve kategorije	prebrojavanje broja klikera svake boje
ponavljanje istraživanja mnogo puta	ponavljanje izbora klikera i prebrojavanja

- pitanje:** koliko će, u nekom konkretnom izvlačenju, od npr. 100 klikera, biti izvučeno crvenih (C), a koliko plavih (P) klikera?
- odgovor:** to ne možemo znati, jer je to stvar slučajja
- ali:** možemo da proračunamo verovatnoće različitih ishoda!

## 1. Univarijantni frekvencijski nacrti (UFN)

- setimo se analogne situacije sa bacanjem novčića:
  - ne znamo* da li će u konkretnom bacanju pasti P ili G
  - ali, možemo precizno da utvrdimo, kolika je *verovatnoća* da u N bacanja:
    - ne padne P *nijednom*, da padne P *tačno* jednom, tačno dva puta, itd.,
    - da padne P *bar* jednom, da padne P *bar* dva puta, itd.
- u primeru sa uzorcima iz vreće se takođe može izračunati verovatnoća:
  - da *nijedan* kliker nije C (tj. ishod 0:100), da je *tačno* jedan C (1:99), itd.
  - da je *bar* jedan kliker C (unija ishoda 1:99, 2:98, 3:97,...), da su *bar dva* C, itd.
- nastavimo ilustrativno poređenje: kako testirati H0?

istraživanje	analogni primer
ne znamo stanje u populaciji	ne znamo koliko ima C i P klikera u vreći
izvršimo istraživanje i utvrdimo frekvence	prebrojimo klikere u datom uzorku
pitanje: da li u populaciji važi H0 ili Ha?	da li je u vreći <i>jednak ili razl. broj</i> klikera?
kako proceniti stanje stvari u populaciji na osnovu rezultata na uzorku?	kako proceniti brojeve klikera u vreći, kada smo prebrojali samo mali deo?

- problem je naizgled nerešiv! – čak i ako u vreći ima jednak broj plavih i crvenih klikera, u principu bi mogli da se dese svi ishodi, tj. da C:P bude:
  - 0:100, 1:99, 2:98, ..., 49:51, 50:50, 51:49, ..., 99:1, 100:0 (ukupno 101 ishoda)

## 1. Univarijantni frekvencijski nacrti (UFN)

- rešenje:** iako je *sve moguće*, nije *sve jednako verovatno*
  - npr., i intuitivno je jasno da je, ako važi H0, u uzorku mnogo verovatniji ishod 51:49 nego 80:20
- postoji velika prednost statističkih testova nad intuicijom:
  - naime: verovatnoće različitih ishoda ne moramo *intuitivno* procenjivati, već ih možemo *matematički* precizno utvrditi
- PRIMERI: može se *izračunati* da ako je H0 tačna, važi sledeće:
  - ishod: izabrano je *tačno* 50 crvenih i 50 plavih klikera
    - verovatnoća:  $p(50:50) = 0.08$  (tj. u 8% slučajeva, redak ali najčešći ishod)
  - ishod: izabrano je *tačno* 60 crvenih i 40 plavih klikera
    - verovatnoća:  $p(60:40) = 0.01$  (tj. dešava se u 1% slučajeva, isto kao i 40:60)
  - ishod: izabrano je *bar* 60 crvenih klikera (60 ili više)
    - $p(60:40 \text{ ili } 61:39 \text{ ili } 62:38 \text{ ili } \dots \text{ ili } 100:0) = 0.028$  (tj. u 2.8% slučajeva)
  - ishod: izabrano je *bar* 70 crvenih klikera
    - $p(70:30 \text{ ili } 71:29 \text{ ili } 72:28 \text{ ili } \dots \text{ ili } 100:0) < 0.0001$  (tj. < 0.01%)
- drugim rečima, ako je H0 tačna, ovakav ishod je *veoma malo verovatan*, tj. dešava se manje nego jednom u 10000 slučajeva!
- uočiti: ovakvi numerički rezultati ne poklapaju se uvek sa intuicijom, ali su tačni

## 1. Univarijantni frekvencijski nacrti (UFN) 7

- opšta šema testova statističke značajnosti rezultata UFN
  - postavimo statističku hipotezu: najčešće  $H_0: p_1 = p_2$
  - izvršimo istraživanje i utvrdimo frekvence dve kategorije
  - ustanovimo *verovatnoću* datog ishoda, pod pretpostavkom da *važi H0*
    - time *dati* ishod postavljamo u kontekst svih *moгуćih* ishoda takvih istraživ.
    - formalni opis: radi se o *uslovnoj* verovatnoći oblika:  $p(\text{ishod}|H_0)$ 
      - uočiti: ovo  $p$  je *različito* od nekih drugih, srodnih bezuslovnih i uslovnih verovatnoća, kao što su  $p(H_0)$  ili  $p(\text{ishod})$  ili  $p(H_0|\text{ishod})$
      - te verovatnoće mogu biti zanimljive za razmatranje, ali standardni testovi značajnosti se ne odnose na njih, već na  $p(\text{ishod}|H_0)$
  - na osnovu  $p(\text{ishod}|H_0)$  odlučuje se da li da se  $H_0$  odbaci ili *ne odbaci*
    - detalji i izvesne komplikacije će biti diskutovati kasnije
  - uslov za ispravnu primenu testova značajnosti: istraživanje je izvedeno tako da su ispunjeni određeni statistički i metodološki uslovi
    - odnose se na slučajnost izbora objekata, nezavisnost itd.
- opisana logika koristi se u većini statističkih analiza rezultata
  - međutim, zavisno od vrste nacrtu, postoje izvesne razlike u prirodi nulte hipoteze, matematičkih uslova testiranja i tipa testova značajnosti

## 1. Univarijantni frekvencijski nacrti (UFN) 8

- tok statističke analize u UFN sa dve kategorije biće prikazan postupno, prolazanjem kroz *pet faza*:
  - Faza I: Podaci i deskriptivne mere**
    - sakupljanje podataka i utvrđivanje frekvenci
  - Faza II: Nulta hipoteza, očekivane vrednosti, devijacije**
    - postavljanje statističkih hipoteza i izvođenje njihovih posledica
  - Faza III: Test-statistik**
    - računanje određene mere statistike zaključivanja
  - Faza IV: p-vrednost**
    - utvrđivanje verovatnoće izračunate mere
  - Faza V: Odluka o statističkoj značajnosti**
    - donošenje zaključka o ishodu istraživanja
- i naredna izlaganja o statističkoj značajnosti, u svim ostalim nacrtima, biće prikazana preko ovih istih pet faza
  - nekad su to bile faze kroz koje su, redom, zaista prolazili istraživači
  - danas sva računanja vrši kompjuter
  - mi ćemo proći kroz njih 'peške', da bi bolje razumeli zašto se one vrše

## 1. Univarijantni frekvencijski nacrti (UFN) 9

- Faza I: Podaci i deskriptivne mere**
  - vrši se utvrđivanje u koju kategoriju spada svaki podatak, i prebrojavanje koliko ima podataka u svakoj kategoriji
    - tako se dobijaju *opservirane (empirijske) frekvence*
- koristićemo *četiri* primera ishoda istraživanja sa bebama
  - u svim primerima biće veći broj izabranih *crvenih* cuclu nego plavih
  - pitajanje: da li je takav ishod pouzdan pokazatelj stanja u populaciji?
- PRIMER A:** početni primer, koji se poredi sa ostalima: C:P = 55:45
  - od 100 beba, 55 je izabralo crvenu cuclu a 45 plavu
- PRIMER B:** isti broj beba kao u A, ali više izbora crvenih cuclu: C:P = 80:20
  - od 100 beba, 80 je izabralo crvenu cuclu a 20 plavu
- PRIMER C:** isto kao u B, ali je još po 1000 beba izabralo crvenu i plavu cuclu
  - od 2100 beba, 1080 je izabralo crvenu a 1020 plavu cuclu
- PRIMER D:** isti odnos C:P kao u A, ali sa 10 puta većim uzorkom
  - od 1000 beba, 550 je izabralo crvenu cuclu a 450 plavu
- pitajanje:** šta nam govori intuicija o *pouzdanosti* rezultata u ova 4 primera?
  - da li ne samo u uzorku već u *populaciji* postoji preferencija prema crvenoj boji?

## 1. Univarijantni frekvencijski nacrti (UFN) 10

- tabelarni prikaz podataka i nekih dodatnih mera**
- |                       | primer A         |      | primer B |      | primer C: B*2000   |      | primer D: A*10     |      |
|-----------------------|------------------|------|----------|------|--------------------|------|--------------------|------|
|                       | crv              | plv  | crv      | plv  | crv                | plv  | crv                | plv  |
| veličina uzorka (N)   | 100              |      | 100      |      | 2100               |      | 1000               |      |
| frekvence (f)         | 55               | 45   | 80       | 20   | 1080               | 1020 | 550                | 450  |
| proporcije (p)        | 0.55             | 0.45 | 0.80     | 0.20 | 0.51               | 0.49 | 0.55               | 0.45 |
| šansa (š)             | 55:45=11:9=1.2:1 |      | 4:1      |      | 1080:1020 = 1.06:1 |      | 550:450=11:9=1.2:1 |      |
| razlika frekvenci (R) | 10               |      | 60       |      | 60                 |      | 100                |      |
- PRIMER A:** početni primer, koji se poredi sa ostalima, u pogledu N, f, p, š i R
    - ne čini se pouzdan (tj. da stvarna razlika postoji u populaciji), šansa je blizu 1:1
  - PRIMER B:** isto kao u A: N; *različito* nego u A: f, p, š i R (veće nego u A)
    - čini se pouzdan, zbog velikih R i š
  - PRIMER C:** isto kao u B: R; *različito*: N (veće nego u B); p (skoro 0.5), š (skoro 1:1)
    - uprkos R istim kao u B, ne čini se pouzdan, p i š su 'gori' (bliže H0) nego u A
  - PRIMER D:** isto kao u A: p, š; *različito*: N (veće nego u A), R (veće nego u A)
    - čini se pouzdan zbog velikih N i R, iako su p i š iste kao u nepouzdanom A
  - zaključak: odluka se ne može doneti na osnovu samo *jednog* pokazatelja (N, f, p, š, R), već treba više njih uzeti u obzir i kombinovati – ali kako?

## 1. Univarijantni frekvencijski nacrti (UFN) 11

- Faza II: nulta hipoteza, očekivane vrednosti, devijacije**
    - postavljanje statističke hipoteze:  $H_0: p_1=p_2$
    - utvrđivanje *očekivanih (teoretskih) frekvenci f1' i f2'*
      - koje su *najverovatnije* frekvence u uzorku, kada bi H0 bila tačna?
      - najverovatnije je da su te frekvence jednake:  $f_1' = f_2'$ , tj. po N/2
    - utvrđivanje *reziduala (rezidualnih frekvenci): d = f - f'*
      - oni odražavaju odstupanja *opserviranih* od *očekivanih* frekvenci
- | faza | koraci              | primer A |       | primer B |      | primer C |       | primer D |       |
|------|---------------------|----------|-------|----------|------|----------|-------|----------|-------|
|      |                     | crv      | plv   | crv      | plv  | crv      | plv   | crv      | plv   |
| I    | empirija (f)        | 55       | 45    | 80       | 20   | 1080     | 1020  | 550      | 450   |
| II   | očekivanje (f')     | 50       | 50    | 50       | 50   | 1050     | 1050  | 500      | 500   |
|      | rezidual (d = f-f') | +5       | -5    | +30      | -30  | +30      | -30   | +50      | -50   |
|      | količnik d/f        | 1/10     | -1/10 | 3/5      | -3/5 | 1/34     | -1/34 | 1/10     | -1/10 |
- uočiti: reziduali za B i C su isti, ali se B čini pouzdanijim ishodom od C
  - od samog reziduala informativniji je njegov *odnos* sa teor. frekvencom d/f
    - ovaj odnos odražava *relativno* odstupanje opserviranih frekvenci od očekivanih
  - moguća *opšta* mera odstupanja: *prosečni* odnos d/f – ali on je nužno nula!
  - druga mera bi mogao biti *prosečni apsolutni* odnos |d/f| - ali on se ne koristi

## 1. Univarijantni frekvencijski nacrti (UFN) 12

- Faza III: Test-statistik**
    - to je broj koji je *pokazatelj* (mera) *odstupanja* dobijenih rezultata od H0
    - za UFN koristi se test-statistik nazvan *hi kvadrat*, označen sa  $\chi^2$
    - formula:  $\chi^2 = d^2/f' + d^2/f'$  (tj., za obe ćelije se računa  $d^2/f'$  i sabere se)
- | faza | koraci                                  | primer A |     | primer B |     | primer C |      | primer D |      |
|------|---|----------|-----|----------|-----|----------|------|----------|------|
|      |   | crv      | plv | crv      | plv | crv      | plv  | crv      | plv  |
| I    | empirija (f)                            | 55       | 45  | 80       | 20  | 1080     | 1020 | 550      | 450  |
| II   | očekivanje (f')                         | 50       | 50  | 50       | 50  | 1050     | 1050 | 500      | 500  |
|      | rezidual (d = f-f')                     | +5       | -5  | +30      | -30 | +30      | -30  | +50      | -50  |
|      | rezidual <sup>2</sup> (d <sup>2</sup> ) | 25       | 25  | 900      | 900 | 900      | 900  | 2500     | 2500 |
| III  | količnik (d <sup>2</sup> /f')           | 0.5      | 0.5 | 18       | 18  | 0.86     | 0.86 | 5        | 5    |
|      | zbir (χ <sup>2</sup> )                  | 1        |     | 36       |     | 1.72     |      | 10       |      |
- količnik  $d^2/f'$  je *sličan* količniku  $d/f$ , jer odražava *relativno* odstupanje
  - količnik  $d^2/f'$  je *različit* od količnika  $d/f$  utoliko što ne može da bude negativan
    - naime, kvadrati i poz. i neg. brojeva su poz. brojevi (sem 0, čije je kvadrat 0)
  - stoga  $\chi^2$  mora da bude *pozitivan* broj (ili nula, za  $f = f'$ , gde nema odstupanja)
  - uočiti:  $d^2/f'$  (a time i njihov zbir, tj.  $\chi^2$ ) će biti utoliko *veći* ukoliko su reziduali  $d$  *veći*, ali ne *apsolutno*, već u *odnosu* na veličine *očekivanih* frekvenci  $f'$

## 1. Univarijantni frekvencijski nacrti (UFN) 13

- novi pojam: **stepen slobode: df** (engleski: 'degree of freedom')
  - važna statistička veličina za  $\chi^2$  i za sve druge test-statistike
  - veličina **df** zavisi od broja mogućih ishoda istraživanja
  - veličina **df** izražava koliko 'slobode' ima prilikom odabira izvesnih vrednosti, pod uslovom da su neke druge vrednosti poznate
- koliko ima mogućih ishoda UFN sa dve kategorije?
  - za prvu kategoriju 'slobodno' je, u principu, ustanoviti bilo koju frekvencu
  - za drugu kategoriju važi isto
  - u tom smislu postojala bi **dva stepena slobode**, po jedan za svaku kategoriju
- međutim: uvek nam je poznata totalna frekvencija (N)
  - na pr., planirali smo da ispitamo, i ispitati smo, 100 beba
- u tom slučaju postoji samo **jedan stepen slobode**, tj. **df = 1**
- naime, ako je data frekvencija jedne kategorije, samim tim je precizno određena i frekvencija druge kategorije, jer se one moraju sabirati do N
  - npr., ako znamo da je 60 beba odabralo crvenu cuclu, onda samim tim znamo i da je 40 beba odabralo plavu, jer znamo da je N=100
- ovakva razmatranja su važna za statističku teoriju

## 1. Univarijantni frekvencijski nacrti (UFN) 14

- Faza IV: p-vrednost**
- opšti smisao:** to je **verovatnoća** da dobijemo ono što smo dobili ako u populaciji stvarno **nema** razlike (formalno: to je **p(ishod|H0)**)
- naime:** to je verovatnoća da se, ako je u populaciji H0 tačna, na uzorku registruje test-statistik **dobijene** ili **veće** veličine, odn. **bar** toliki  $\chi^2$ 
  - p-vrednost se izračunava na osnovu tzv. **hi-kvadrat distribucije**
  - to je **raspodela vrednosti** hi-kvadrata za sve moguće ishode, ako važi H0
- važi:** što je **veći**  $\chi^2$  (tj. odstupanje rezultata od H0), **manja** je **p-vrednost**

faze	koraci	primer A		primer B		primer C		primer D	
		crv	plv	crv	plv	crv	plv	crv	plv
I	empirija (f)	55	45	80	20	1080	1020	550	450
II	očekivanje (f')	50	50	50	50	1050	1050	500	500
	rezidual (d = f-f')	+5	-5	+30	-30	+30	-30	+50	-50
III	rezidual <sup>2</sup> (d <sup>2</sup> )	25	25	900	900	900	900	2500	2500
	količnik (d <sup>2</sup> /f')	0.5	0.5	18	18	0.86	0.86	5	5
zbir ( $\chi^2$ )		1		36		1.72		10	
IV	p-vrednost	0.32		0.000000002		0.19		0.0016	

- PRIMER A:** p da se dobije  $\chi^2 = 1$  ili  $> 1$  odn.  $\chi^2 \geq 1$  (tj. **bar** 1) iznosi **0.32**

## 1. Univarijantni frekvencijski nacrti (UFN) 15

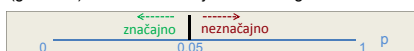
- Faza V: Odluka o statističkoj značajnosti**
- smisao:** šta možemo na osnovu uzorka da zaključimo o populaciji?
- PRIMER A: C:P = 55:45,  $\chi^2 = 1$ , p = 0.32**
  - dakle: verovatnoća **ovakvog** ishoda (ili još veće prednosti izbora jedne boje nad drugom), ako važi H0 (tj. ako u populaciji nema preferencije boja), je skoro 1/3
  - zaključak: **nema dovoljno osnova da se pouzdano odbaci H0!**
    - naime, u principu jeste moguće da ženske bebe **više** vole crvenu boju
    - međutim, sasvim je moguće i da u populaciji **nema razlike** u preferenciji boja
      - ne može se isključiti pretpostavku da je razlika dobijena na uzorku u korist crvene boje samo posledica **slučaja** (dešava se jednom u tri puta)
    - evidencija** (činjenice koje idu u prilog) radnoj hipotezi nije dovoljno uverljiva
- PRIMER B: empirijske frekvencije: C:P = 80:20,  $\chi^2 = 36$ , p = 0.000000002**
  - dakle: verovatnoća **ovakvog** ishoda (ili još veće prednosti izbora jedne boje nad drugom), ako važi H0, je 'mikroskopski' mala (dva u milijardu slučajeva)
  - zaključak: **postoje veoma dobre osnove da se odbaci H0!**
    - teoretski je **moguće** da u populaciji **ne** postoji preferencija, a da se ipak dobije **ovakav** ishod u prilog jedne boje
    - međutim, mnogo više ima osnova pretpostavka da ženske bebe zaista **više** vole crvenu boju, i da se to i pokazalo u dobijenom rezultatu

## 1. Univarijantni frekvencijski nacrti (UFN) 16

- na osnovu primera se mogu doneti opštniji zaključci:
  - ako je p-vrednost relativno **velika**, sasvim je moguće da važi H0 u populaciji, a da je dobijena razlika u uzorku nastala slučajno, tj. H0 se ne može odbaciti
  - što je **manja** p-vrednost, utoliko je manje uverljiva pretpostavka da važi H0, a da je razlika u uzorku nastala slučajno, tj. utoliko ima više osnova da se H0 odbaci
  - za veoma **mala** p-vrednosti, sasvim je razumno zaključiti da dobijena razlika nije delo slučajna, i odbaciti pretpostavku o postojanju H0 u populaciji
- mogu se postaviti dva komplementarna pitanja o p-vrednosti:
  - za koje veličine p-vrednosti je ona još previše **velika**, pa se H0 ne može odbaciti?
  - za koje veličine p-vrednosti je ona već dovoljno **mala**, pa se H0 može odbaciti?
- drugim rečima: da li postoji neka veličina p-vrednosti koja bi pouzdano **razgraničila** ove dve mogućnosti?
  - takva veličina bi uvek obezbedila potpuno ispravnu odluku o ishodu istraživanja!
- odgovor: postoji jedna **loša** i jedna (umereno) **dobra** vest
- loša vest:** nažalost ne postoji matematički opravdana granična vrednost!
  - uverljivost H0 **postepeno** opada sa opadanjem p-vrednosti, od sasvim uverljive do sasvim neuverljive, bez oštre granice uverljivosti
- kako onda doneti praktičnu odluku o ishodu istraživanja?

## 1. Univarijantni frekvencijski nacrti (UFN) 17

- dobra vest:** postoje dogovorom usvojena i u istraživačkoj praksi sada široko prihvaćena i primenjivana **pravila** za donošenje odluka o ishodu istraživanja
- u tim pravilima se koriste tzv. **nivoi značajnosti ( $\alpha$ -nivoi)**
  - to su unapred utvrđene **granične p-vrednosti**, sa kojima se **porede** p-vrednosti dobijene u istraživanju, radi donošenja odluke o ishodu istraživanja
  - u psihološkoj literaturi koriste se većinom **tri** nivoa značajnosti: 0.01, 0.05, i 0.10
- pravilo za (danas najčešće korišćen) nivo značajnosti  $\alpha = 0.05$  glasi:
  - ako je u istraživanju dobijena p-vrednost koja je:
    - manja** od  $\alpha$ -nivoa, tj. ako je **p < 0.05**
      - onda se H0 **odbacuje**, i kaže se da je rezultat **statistički značajan**
    - veća** od  $\alpha$ -nivoa, tj. ako je **p > 0.05**
      - onda se H0 **ne odbacuje**, kaže se da rezultat **nije** statistički značajan
    - jednaka ili bliska**  $\alpha$ -nivou, tj. ako je **p  $\approx$  0.05**
      - onda se zaključuje sa oprezom, i kaže se da je rezultat **marginalno** (granično) statistički značajan odn. marginalno statistički neznačajan



## 1. Univarijantni frekvencijski nacrti (UFN) 18

- uočiti: kriterijum SZ je **konvencionalan** i **arbitraran**, ali je **za sada** široko prihvaćen, kao uslov koji nalaz mora da ispuni da bi se smatrao pouzdanim
  - u praksi se ovakav kriterijum uglavnom pokazao kao koristan **vodić** u odlučivanju istraživača i naučne zajednice prilikom interpretacije rezultata istraživanja
- ali: o testiranju značajnosti se još vode diskusije u metodološkoj literaturi!
- primeri primene nivoa značajnosti  $\alpha = 0.05$ :
  - PRIMER A:** p = 0.32, tj. p > 0.05, što nije statistički značajno ('nije SZ')
  - PRIMER B:** p = 0.000000002, tj. p < 0.05, što je statistički značajno ('jeste SZ')
- druga dva kriterijuma,  $\alpha = 0.10$ , i  $\alpha = 0.01$ , ređe se primenjuju
  - u poređenju sa  $\alpha = 0.05$ , nivo  $\alpha = 0.10$  je **blaži**, liberalniji kriterijum
    - po njemu, neke p-vrednosti su SZ, koje nisu SZ prema kriterijumu  $\alpha = 0.05$ 
      - to su p-vrednosti između 0.05 i 0.10, tj. **0.05 < p < 0.10**
      - na pr., p = 0.08 nije SZ za  $\alpha = 0.05$ , ali jeste SZ za  $\alpha = 0.10$
  - u poređenju sa  $\alpha = 0.05$ , nivo  $\alpha = 0.01$  je **stroži**, konzervativniji kriterijum
    - po njemu, neke p-vrednosti nisu SZ, koje jesu SZ prema  $\alpha = 0.05$ 
      - to su p-vrednosti između 0.01 i 0.05, tj. **0.01 < p < 0.05**
      - na pr., p = 0.03 nije SZ za  $\alpha = 0.01$ , ali jeste SZ za  $\alpha = 0.05$
- u fizici i nekim oblastima neuronauka koriste se još strožiji nivoi značajnosti

## 1. Univarijantni frekvencijski nacrti (UFN) 19

- kakve su posledice korišćenja različito strogih kriterijuma SZ?
  - uočiti: koji god kriterijum da se koristi, mogu se pojaviti greške, a ima ih dve vrste
- posledice korišćenja (pre)strogog kriterijuma, tj. (pre)visoke prepreke
  - *dobre*: odbacuje se veći broj rezultata koji su samo posledica slučaja
  - *loše*: odbacuje se i veći broj realno postojećih fenomena
    - i kod realno postojećih fenomena uvek deluju varijable šuma, jače ili slabije, koje ometaju uvid u pojavu
    - ako istraživač ne prihvati realno postojeći fenomen, pravi *grešku tipa 2*
- posledice korišćenja (pre)blagog kriterijuma, tj. (pre)niske prepreke
  - *dobre*: prihvatiće se veći broj realno postojećih fenomena, ometanih šumom
  - *loše*: prihvatiće se veći broj realno nepostojećih fenomena ('lažni alarm')
    - varijable šuma mogu da stvore privid realnog fenomena
    - ako istraživač prihvati realno nepostojeći fenomen, pravi *grešku tipa 1*
- kritične p-vrednosti (nivoi značajnosti) i kritične  $\chi^2$ -vrednosti:
  - može se matematički dokazati da u UFN sa dve kategorije važi sledeće:
    - uslov da bude  $p < 0.10$  je da bude  $\chi^2 > 2.71$
    - uslov da bude  $p < 0.05$  je da bude  $\chi^2 > 3.84$
    - uslov da bude  $p < 0.01$  je da bude  $\chi^2 > 6.64$

## 1. Univarijantni frekvencijski nacrti (UFN) 20

faze	koraci	primer A		primer B		primer C		primer D	
		crv	plv	crv	plv	crv	plv	crv	plv
I	empirija (f)	55	45	80	20	1080	1020	550	450
II	očekivanje (f)	50	50	50	50	1050	1050	500	500
	rezidual (d = f-f)	+5	-5	+30	-30	+30	-30	+50	-50
III	rezidual <sup>2</sup> (d <sup>2</sup> )	25	25	900	900	900	900	2500	2500
	količnik (d <sup>2</sup> /f)	0.5	0.5	18	18	0.86	0.86	5	5
	zbir ( $\chi^2$ )	1		36		1.72		10	
IV	p-vrednost	0.32		0.000000002		0.19		0.0016	
V	značajnost	nije SZ		jeste SZ		nije SZ		jeste SZ	
	zapis	$\chi^2(1)=1, p>0.05$		$\chi^2(1)=36, p<0.05$		$\chi^2(1)=1.72, p>0.05$		$\chi^2(1)=10, p<0.05$	

- zapis rezultata testa je oblika:  $\chi^2(df) = \dots$ , pri čemu u UFN važi da je  $df=1$ 
  - ranijih godina je zapis p-vrednosti uvek bio oblika  $p < \alpha$  ili  $p > \alpha$
  - poslednjih godina, usled korišćenja kompjutera, zapis je sve češće oblika  $p = \dots$
- pitanje: zašto su za različite primere *različiti* zaključci statističke analize?
- **PRIMER**: poređenje ishoda u A (nije SZ) i B (jeste SZ):
  - u A i B je isto N, i time su iste f, ali su u B veći reziduali, odn. njihovi kvadrati  $d^2$
  - stoga su u B veći količnici  $d^2/f$ , pa je veći i  $\chi^2$ , koji je zbir tih količnika

## 1. Univarijantni frekvencijski nacrti (UFN) 21

faze	koraci	primer A		primer B		primer C		primer D	
		crv	plv	crv	plv	crv	plv	crv	plv
I	empirija (f)	55	45	80	20	1080	1020	550	450
II	očekivanje (f)	50	50	50	50	1050	1050	500	500
	rezidual (d = f-f)	+5	-5	+30	-30	+30	-30	+50	-50
III	rezidual <sup>2</sup> (d <sup>2</sup> )	25	25	900	900	900	900	2500	2500
	količnik (d <sup>2</sup> /f)	0.5	0.5	18	18	0.86	0.86	5	5
	zbir ( $\chi^2$ )	1		36		1.72		10	
IV	p-vrednost	0.32		0.000000002		0.19		0.0016	
V	značajnost	nije SZ		jeste SZ		nije SZ		jeste SZ	
	zapis: $\chi^2(df)$	$\chi^2(1)=1, p>0.05$		$\chi^2(1)=36, p<0.05$		$\chi^2(1)=1.72, p>0.05$		$\chi^2(1)=10, p<0.05$	

- **PRIMER**: poređenje B (jeste SZ) i C (nije SZ):
  - u B i C su jednaki reziduali i njihovi kvadrati  $d^2$ , ali je u C veće N, i time f
  - stoga su u C manji količnici  $d^2/f$ , pa je manji i  $\chi^2$ , koji je zbir tih količnika
    - uočiti: u B proporcije veoma različite:  $p_1=0.8, p_2=0.2$ ; šansa:  $f_1:f_2 = 4:1$
    - uočiti: u C proporcije vrlo slične:  $p_1=0.51, p_2=0.49$ ; šansa  $f_1:f_2 \approx 1.06:1$
- **PRIMER**: poređenje D (jeste SZ) i A (nije SZ):
  - u D i A su jednake proporcije i šanse, ali su u D 10x veći N, f i d, i 100x veći  $d^2$
  - stoga su u D količnici  $d^2/f$  10x veći nego u A, a time je i 10x veći i njihov zbir, tj.  $\chi^2$

## 1. Univarijantni frekvencijski nacrti (UFN) 22

- da bi se ispravno primenio  $\chi^2$  – test, potrebno je da su ispunjeni neki uslovi
  - uzorak je slučajan, svi podaci su nezavisni
  - očekivane frekvence ne treba da budu manje od 5 (ovaj zahtev može se ublažiti)
- alternative  $\chi^2$  – testu kod UFN sa dve kategorije: drugi testovi značajnosti
  - pomenućemo dva: *z-test za jednu proporciju* i *binomijalni test*
  - različiti testovi uglavnom dovode do istih zaključaka o značajnosti
- napomene o terminologiji: razlike statist. žargona i svakodnevnog govora
  - 'statistički značajan' rezultat ne znači isto što i 'značajan' rezultat
    - 'značajan' u svakodnevnom govoru znači 'veliki' ili 'važan'
    - s druge strane, 'statistički značajan' znači da se može odbaciti nulta hipoteza
      - takav rezultat može a ne mora biti važan ili se odnosi na veliku razliku
  - 'očekivano' u statističkom smislu nije isto kao u svakodnevnom smislu
    - 'očekivano' u svakodnevnom govoru znači nešto što se *lično* očekuje, ili se pretpostavlja da će se dogoditi
    - 'očekivane frekvence' nisu nešto što istraživač lično očekuje da se dogodi
  - 'hipoteza' u statističkom smislu nije isto kao u svakodnevnom smislu
    - uobičajeno, 'hipoteza' je nešto što pretpostavljamo, ali što nije dokazano
    - s druge strane, 'nulta hipoteza' nije nešto što istraživač lično pretpostavlja da je tačno, već je to prosto jedan od standardnih koraka u statističkoj analizi