

Metodologija psiholoških istraživanja

1

obrada faktorijalni 3



B. Faktorijalni nacrti

1. Jednofaktorski nacrti

- organizacija podataka
- deskriptivne statističke mere
- standardne mere
- transformacije skupova mera
- prikaz rezultata
- struktura rezultata
- značajnost rezultata

1. bivalentni nacrti

- neponovljeni nacrti
- ponovljeni nacrti

3. decembar 2018

1. Jednofaktorski nacrti (JFN)

2

- **Faza IV: p-vrednost**
 - p-vrednost je verovatnoća da se u istraživanju dobije veličina F-količnika koja je zaista dobijena (ili još veća), ako H0 važi u populaciji
 - PRIMER: H0: nema razlike u matematičkoj sposobnosti pušača i nepušača
 - u datom primeru: $p = 0.106$
 - kompjuterski programi izračunavaju p-vrednost preko tzv. **F-distribucije**
 - distribucija vrednosti $F = MS_A / MS_e = (SS_A / df_A) / (SS_e / df_e) = (\Sigma e^2 / 1) / (\Sigma g^2 / 2(N-1))$
 - uočiti: elementi F-količnika su dve vrste devijacija (e i g) i broj podataka (N)
- **Faza V: odluka o statističkoj značajnosti**
 - H0 se odbacuje ako je $p < 0.05$, a ne odbacuje se ako je $p > 0.05$
 - PRIMER: budući da je $p > 0.05$, H0 se ne može odbaciti
 - zaključak: na osnovu dobijenih (izmišljenih) rezultata, nema osnove da se tvrdi da matematička sposobnost zavisi od pušenja
 - uočiti: procent objašnjene varijabilnosti ($r^2 = 80\%$), i koeficijent korelacije NV i ZV ($r = 0.89$) su veliki, ali rezultat ipak nije statistički značajan
 - u zapis rezultata statističke analize treba uneti i vrednosti df_A i df_e
 - PRIMER: piše se $F(1, 2) = 8$, $p > 0.05$, budući da je $df_A = 1$, i $df_e = 2$
 - opisani postupak analize rezultata naziva se **F-test**

1. Jednofaktorski nacrti (JFN)

3

odnos veličine F-količnika i statističke značajnosti rezultata

- slično kao kod hi-kvadrata, što je veći F-količnik, manja je p-vrednost, i više ima izgleda da je rezultat statistički značajan
- važno pitanje za istraživača: kada će F-količnik biti veliki?
- kao i svaki količnik dva broja, F-količnik će biti utoliko veći ukoliko je:
 - veći brojilac B (gornji član razlomka), a manji imenilac I (donji član razlomka)
- kako je $F = MS_A / MS_e$, može se zaključiti da će F će biti utoliko veće ukoliko je veći $B = MS_A = SS_A / df_A = \Sigma e^2 / 1$, a manji $I = MS_e = SS_e / df_e = \Sigma g^2 / 2(N-1)$
- dakle, B će biti utoliko veći ukoliko je zbir kvadrata efekata Σe^2 veći
 - Σe^2 će biti utoliko veći što su veći efekti $e = M - My$, tj. ukoliko se proseci grupa više razlikuju od opšteg proseka, odn. ukoliko je razlika $R = M1 - M2$ veća
- kako je I razlomak, I će biti utoliko manji ukoliko je Σg^2 manje a $2(N-1)$ veće
 - Σg^2 će biti utoliko manje ukoliko su greške g manje, tj. ukoliko je varijabilnost pojedinačnih mera oko svojih proseka manja
 - $df_e = 2(N-1)$ će biti utoliko veće ukoliko je N veće, tj. kada je uzorak veći
- zaključak: ova matematička razmatranja se slažu sa ranijom analizom:
 - rezultati su utoliko pouzdaniji ukoliko su: veće veličina razlike proseka (R) i veličina uzorka (N), a manja veličina varijabilnosti podataka unutar grupa (SD)

1. Jednofaktorski nacrti (JFN)

4

- analiza strukture F-količnika pruža važnu informaciju istraživaču na šta treba da obrati pažnju prilikom planiranja istraživanja i analize rezultata
- PRIMER: u datom istraživanju podaci Y su bili {4, 2, 8, 6}, i dobijeno je $p > 0.05$
- rezultati opisanog istraživanja mogli su biti statistički značajni da je bio:
 - veći uzorak: na pr. da je N bilo veće, npr. N=4, i recimo $Y = \{4, 2, 4, 2; 8, 6, 8, 6\}$
 - tada bi bili isti proseci, ali $df_e = 2(N-1) = 6$, pa bi bilo $F(1, 6) = 24$, $p < 0.05$
 - veće R: na pr. da su nep. rešili više zadataka, npr. još po 4, tj. $Y = \{4, 2; 12, 10\}$
 - tada bi bilo $M1=3$, $M2=11$, $R=M2-M1=11-3=8$, pa bi bilo $F(1, 2) = 32$, $p < 0.05$
- manja varijabilnost: na pr. da su pušači imali ne 4 i 2 već 3.5 i 2.5 poena, a nepušači ne 8 i 6 već 7.5 i 6.5 poena, tj. $Y = \{3.5, 2.5; 7.5, 6.5\}$
 - tada bi bili isti proseci (3 i 5), ali greške ne bi bile po ± 1 već po ± 0.5
 - tada bi bilo $F(1, 2) = 32$, $p < 0.05$
- napomena: kao i za hi-kvadrat, i za F-količnik postoje kritične vrednosti
 - najmanje vrednosti koje F može da ima da bi rezultat bio stat. znač., za dato α
 - F-količnik i kritične vrednosti zavisi od veličine df_A i df_e u datom istraživanju
 - ranije su se koristile opsežne tablice kritičnih vrednosti F-količnika za razne df
- PRIMER: za $df_A = 1$ i $df_e = 2$, krit. vrednost za $F(1,2)$ je 18.51 (za $\alpha = 0.05$)
 - u primeru je $F=8$, što je manje od kritične vrednosti, pa rezultat nije stat. značajan

1. Jednofaktorski nacrti (JFN)

5

analogija sa signalom i šumom

- dejstvo NV na ZV je analogno signalu koji istraživač želi da detektuje
- dejstvo drugih varijabli na ZV je analogno šumu koji ometa prijem signala
- kako takav šum utiče na MS_A i MS_e ?
 - $MS_e = \Sigma g^2 / 2(N-1)$ predstavlja čisto dejstvo šuma
 - PRIMER: to je varijabilnost koja se ne može pripisati pušenju odn. nepušenju, već potiče od razlika u okviru pušača i razlika u okviru nepušača
 - $MS_A = \Sigma e^2 / 1$ predstavlja kombinaciju dejstva signala i šuma
 - naime, razlika u prosecima dve grupe može poticati od dejstva NV (signal), ali može biti i slučajna (šum)
 - PRIMER: čak i ako pušenje ne utiče na matematičku sposobnost, može se očekivati da razlika u postignuću pušača i nepušača neće biti nulta
- usled ovoga, F-količnik se simbolički može napisati na sledeće načine:

$$\text{ako postoji signal: } F = \frac{MS_A}{MS_e} = \frac{\text{signal} + \text{šum}}{\text{šum}} \quad \text{ako ne postoji signal: } F = \frac{MS_A}{MS_e} = \frac{\text{šum}}{\text{šum}}$$
- kada nema signala, očekuje se da F-količnik bude relat. blizak jedinici (šum/šum)
- ako signal postoji, očekuje se da MS_A nadmaši MS_e , pa da F-količnik, tj. izraz (signal+šum)/šum, bude veći od 1
 - što je jači signal, tj. uticaj NV na podatke, utoliko će F-količnik biti veći

1. Jednofaktorski nacrti (JFN)

6

uslovi korišćenja F-testa u bivalentnim neponovljenim nacrtima

- kao i za hi-kvadrat test, i za F-test postoje izvesni uslovi koje podaci treba da ispune, da bi primena testa bila statistički opravdana
 - psiholozi istraživači ponekad zanemaruju važnost ovih uslova
- dva uslova za korišćenje F-testa su isti kao kod hi-kvadrat testa:
 - slučajnost: slučajni izbor objekata iz populacije
 - ovaj uslov je posebno važan kod nehomogenih populacija
 - nezavisnost: sve mere ZV su međusobno nezavisne
- dva uslova za korišćenje F-testa su novi:
 - normalnost: normalna raspodela mera ZV u populaciji za oba nivoa NV
 - važno: uslov nije ključan, jer je F-test robustan s obzirom na normalnost
 - pokazano je da on daje pouzdane rezultate čak i ako raspodela mera ZV odstupa od normalne, naročito ako su uzorci veliki
 - napomena: statistički testovi koji, kao F-test, zahtevaju normalnost, se nazivaju parametarski testovi; ako nemaju taj uslov, oni su neparametarski
 - homogenost: varijabilnost mera ZV na oba nivoa u populac. nije različita
 - postoje posebni testovi za homogenost
 - ovaj uslov je posebno važan za nebalansirane nacrti

1. Jednofaktorski nacrti (JFN)

7

alternativa F-testu: t-test za NBJFN (engl.: 'independent samples t-test')

- t-test je prostiji, istorijski prethodi F-testu, i predaje se obično pre F-testa
- F-test je opštiji: može se koristiti i u multivalentnim i višefaktorskim nacrtima
- t-test je obično *ekvivalentan* F-testu: daje istu odluku o značajnosti
 - uslovi ekvivalentnosti: balansiranost nacrtu, homogenost varijabilnosti
 - tada je t-količnik jednak korenu F-količnika, a stepen slobode je isti kao df_e
 - ako ovi uslovi nisu ispunjeni, koriste se složenije formule za t-količnik i df
- sprovođenje t-testa prolazi kroz istih 5 faza kao kod F-testa
 - krajnji zapis o ishodu je sličan kao za F-test: na pr.: $t(98) = 1.58, p > 0.05$
 - t-test nema dve vrste stepena slobode kao F-test, već jednu; u primeru $df=98$
- prednost t-testa nad F-testom: mogu se testirati ne samo dvosmerne alternativne hipoteze (kao kod F-testa) već i *jednosmerne*
 - na pr. da su pušači *bolji* ili da su pušači *lošiji* od nepušača
 - a ne samo da se pušači *razlikuju* od nepušača (dvosmerna hipoteza)
 - jednosmerne hipoteze se lakše statistički dokazuju (u smislu da su kritične vrednosti t-količnika manje), ali moraju biti teorijski opravdane
- uslovi korišćenja t-testa: slični kao za F-test (na pr. traži se normalnost)
- postoji i *neparametarski* test razlike proseka dva nivoa: *U-test*

1. Jednofaktorski nacrti (JFN)

8

(b) testiranje značajnosti u ponovljenim bivalentnim JFN (PBJFN)

- setimo se: postoje dve vrste PBJFN
 - 1. dve *sparene* grupe
 - 2. *repeticija*: ista grupa na oba nivoa
 - obe vrste se obrađuju na isti način, a razmatraćemo samo repeticiju
 - PRIMER**: vreme reakcije na crvenu i zelenu boju
 - uočiti: podaci (izmišljeni) su drugačiji nego u ranijem primeru sa crvenom i zelenom bojom
 - da bi uporedili neponovljene i ponovljene nacrtu, interpretiraćemo *iste* podatke na dva načina:
- | # | Crv. (Y1) | Zel. (Y2) |
|----|-----------|-----------|
| 1. | 300 | 380 |
| 2. | 600 | 720 |
| 3. | 200 | 300 |
| 4. | 400 | 490 |
| 5. | 500 | 610 |
| M | 400 | 500 |
- 1. neponovljeni nacrt
 - zamislilićemo da podaci potiču od dve *različite* grupe subjekata, 5 u jednoj i 5 u drugoj grupi
 - jedna grupa reaguje samo na crvenu, a druga samo na zelene draži
 - 2. ponovljeni nacrt
 - zamislilićemo da podaci potiču od *iste* grupe subjekata, ukupno 5
 - svi subjekti reaguju i na crvene i na zelene draži

1. Jednofaktorski nacrti (JFN)

9

- 1. analiza rezultata po *neponovljenom* nacrtu
 - važi da je: $SS_A=25000, df_A=1, SS_e=215000, df_e=8$
 - sledi da je: $MS_A=25000, MS_e=215000/8=26875$
 - tako da je: $F(1,8) = 25000/26875 = 0.93, p > 0.05$
 - zaključak: nema osnova za tvrdnju da se crvena i zelena boja obrađuju različito
 - uočiti: velika varijabilnost i preklapanje rezultata
 - iako su reakcije su u *proseku* sporije na zelenu, neki subjekti brže reaguju na zelenu nego većina subjekata na crvenu
 - 2. analiza rezultata po *ponovljenom* nacrtu
 - podaci su isti, ali su neki aspekti drugačiji:
 - nema 10 već 5 subjekata, a od svakog nema samo 1 već 2 podatka
 - podaci u istom redu ne odnose se na dva različita subjekta već na istog
 - ima smisla računati *razliku* rezultata u istom redu ($D = Y_2 - Y_1$), što ne bi imalo smisla da se radi o neponovljenom nacrtu
 - uočiti: u ovim (izmišljenim) podacima razlika D je uvek dosledno u *istom* smeru, tj. *svaki* subjekt reaguje brže na crvenu nego na zelenu boju
 - analiza ista kao za *neponovljeni* nacrt ne bi uzela u obzir ovu činjenicu!
- | # | Crv. (Y1) | Zel. (Y2) | D (Y2-Y1) |
|----|-----------|-----------|-----------|
| 1. | 300 | 380 | 80 |
| 2. | 600 | 720 | 120 |
| 3. | 200 | 300 | 100 |
| 4. | 400 | 490 | 90 |
| 5. | 500 | 610 | 110 |
| M | 400 | 500 | 100 |

1. Jednofaktorski nacrti (JFN)

10

- analiza rezultata kod ponovljenih nacrtu se vrši pomoću F-testa, u istih 5 faza kao kod neponovljenih nacrtu, uz određene promene
 - te promene uzimaju u obzir ponovljenost nacrtu
- brojilac (B) F-količnika je MS_A , tj. isti je kao kod neponovljenih nacrtu
- imenilac (I) F-količnika nije MS_e , tj. nije isti kao kod neponovljenih nacrtu, već se računa drugačije, i označava se sa MS_{rezidual}
 - odgovarajući df_{rezidual} nije $2(N-1)$ već $N-1$ (jer ima samo jedna grupa subjekata)
 - dodatni detalji statističke analize opisani su udžbeniku (str. 333-334)
- $F = MS_A / MS_{\text{rezidual}}$, gde je $MS_A = SS_A / df_A$, i $MS_{\text{rezidual}} = SS_{\text{rezidual}} / df_{\text{rezidual}}$
- ključna razlika je da je MS_{rezidual} često znatno *manji* od MS_e
- zato je F-količnik u ponovljenom nacrtu obično *veći* nego u neponovljenom nacrtu (iako je df manji), pa su veći izgledi da bude statistički značajan
 - npr., u analizi po nep. nacrtu bilo je $F(1,8) = 0.93, p > 0.05$, dok je u analizi po pon. nacrtu $F(1,4) = 200, p < 0.05$, tj. reakcija na crvenu je stat. značajno brža
- setimo se: postupak repeticije ima metodološki dobru osobinu da je svaki subjekt 'sam sebi kontrola'
- analiza rezultata po ponovljenom nacrtu na statistički način uzima u obzir ovu važnu metodološku činjenicu

1. Jednofaktorski nacrti (JFN)

11

- uslovi korišćenja F-testa u ponovljenim BJFN
 - kao i kod neponovljenih BJFN, traže se *slučajnost* i *nezavisnost*
 - traži se i *normalnost*, ali ne samih mera Y_1 i Y_2 (kao kod NBJFN), već se traži da su njihove *razlike* $D = Y_1 - Y_2$ normalno distribuirane
 - za razliku od NBJFN, ne traži se *homogenost* varijabilnosti mera ZV
- alternativa F-testu u ponovljenim BJFN: t-test
 - engl.: 'paired samples t-test' tj. 't-test za sparene uzorke'
 - zaista se i koristi kod sparenih grupa, ali mnogo češće kod repeticija
 - u tom slučaju su 'grupe' takode 'sparene', naime radi se o jednoj istoj grupi
 - uslovi korišćenja: slični kao kod F-testa
- neparametarske varijante testova kod ponovljenih BJFN
 - test predznaka* (engl.: 'sign test')
 - Wilcoxonov test* (engl.: 'Wilcoxon's test')
- postoji i 't-test za jedan uzorak' (engl.: 'one-sample t-test')
 - služi za poređenje *jedne* grupe sa nekom unapred poznatom merom
 - PRIMER**: da li se količnik inteligencije kod Srba (grupa koja se testira) razlikuje od evropskog proseka (unapred poznata veličina)?

1. Jednofaktorski nacrti (JFN)

12

razmatranje ishoda testova značajnosti

- dva su moguća ishoda:
 - pozitivan*: $p < 0.05$, nađena je statistički *značajna* razlika, odbacuje se H_0
 - negativan*: $p > 0.05$, nije nađena statistički značajna razlika, ne odbacuje se H_0
- (a) razmatranja nakon *negativnog* ishoda: ne odbacuje se H_0
 - ovakav rezultat je često *suprotan* ličnom očekivanju istraživača, izraženom u radnoj hipotezi
 - ovakav ishod NE znači: *dokazano* je da razlika u populaciji *ne postoji*
 - već znači: *nije dokazano* (sa zadovoljavajućim stepenom verovatnoće) da razlika u populaciji *postoji*
 - razlog ovakvog ishoda može biti: razlika zaista *ne postoji* u populaciji
 - ali, moguće su i alternativne (drugačije) interpretacije:
 - razlika u populaciji *postoji*, ali se nije pokazala kao statistički značajna na *uzorku*, tj. dogodila se greška tipa 2 (neprihvatanje postojećeg efekta)
- istraživač *mora* prihvatiti dobijeni ishod testa: H_0 se ne može odbaciti
 - ali: istraživač *može* da uzme u obzir i alternativna objašnjenja, da razmotri moguće uzroke negativnog ishoda, da pokuša da ih otkloni, i da možda *ponovi istraživanje sa odgovarajućim promenama*

1. Jednofaktorski nacrti (JFN)

13

- **statistički** uzrok negativnog ishoda je premala veličina F-količnika
- stoga, budući da veličina F-količnika zavisi od tri faktora (N, SD, R), uzrok premale veličine može biti jedan ili više tih faktora:
- 1. broj objekata po nivou, odn. veličina uzorka (N)
 - mogući razlog negativnog ishoda: uzorak je bio *premal*
 - zaključak: moglo bi se ponoviti istraživanje, ali sa većim uzorkom
- 2. veličina unutargrupne varijabilnosti (SD)
 - mogući razlog negativnog ishoda: *prevelika* varijabilnost, odn. šum
 - zaključak: moglo bi se ponoviti istraživanje, ali sa boljom kontrolom varijabli šuma
- 3. veličina razlike proseka dva nivoa (R)
 - mogući razlog negativ. ishoda: razlika u uzorku *podcenjuje* razliku u populaciji
 - mogući razlozi podcenjenosti:
 - neprimerena NV (npr. izbor nivoa jačine muzike)
 - neosetljiva ZV (npr. samo jedno pitanje na testu)
 - nekontrolisana konfundirajuća varijabla (npr. veličina obojane draži)
 - zaključak: moglo bi se ponoviti istraživanje, ali sa primerenijom NV, osjetljivijom ZV, ili bolje kontrolisanom konfundirajućom varijablom
- problem: istraživači obično ne objavljuju rezultate sa negativnim ishodom

1. Jednofaktorski nacrti (JFN)

14

- (b) razmatranja nakon *pozitivnog* ishoda: odbacuje se H0
- ovakav rezultat je često u skladu sa ličnim očekivanjem istraživača izraženom u radnoj hipotezi
 - razlog ovakvog ishoda može biti: razlika zaista *postoji* u populaciji
 - ipak, moguće su još dve interpretacije ishoda:
 - (a) dogodila se greška tipa 1 (prihvatanje nepostojećeg efekta)
 - setimo se: nivo značajnosti od $\alpha=0.05$ znači:
 - ako je H0 tačna, verovatnoća da se dobije test-statistik jednak ili veći od stvarno dobijenog je manja od 0.05
 - po usvojenom kriterijumu, ova verovatnoća je previše mala, i H0 se odbacuje
 - ali: ovakav događaj nije nemoguć, već samo malo verovatan: $p < 0.05$
 - međutim, i malo verovatni događaju se ponekad događaju!
 - najbolja zaštita od ovakve greške: *ponavljanje* istraživanja (replikacija), posebno od strane drugih istraživača
 - poslednjih godina se sve više naglašava važnost replikacije
 - (b) dejtvovala je konfundirajuća varijabla čije dejstvo *oponaša* NV
 - zaštitu od ove mogućnosti daju raznovrsni postupci kontrole istraživanja

1. Jednofaktorski nacrti (JFN)

15

(2) Testiranje značajnosti u multivalentnim nacrtima (MJFN)

- ključna razlika od bivalentnih nacrti: postoje *više* od dva nivoa NV (a_1, a_2, a_3, \dots), i stoga *više* od dva proseka ZV (M_1, M_2, M_3, \dots)
- obrada je slična kao kod bivalentnih nacrti, uz odgovarajuće izmene
- **PRIMER:** brzina reakcije na imenice u različitim padežima (nominativ, genitiv, dativ, akuzativ), u balansiranom nacrtu (po 25 reči po padežu)
- dva moguća nacrti:
 - neponovljen nacrt: 4 grupe subjekata, po jedna po padežu
 - ponovljen nacrt (bolji): istoj grupi se prikazuju svi padeži u istoj seansi
- **Faza I: Podaci i deskriptivne mere**
 - nećemo detaljno razmatrati podatke pojedinih subjekata, kao u primeru bivalentnog nacrti, već samo proseke
 - **PRIMER:** neka je: $M_1=680, M_2=700, M_3=720, M_4=715$, tako da je $M_y=704$



1. Jednofaktorski nacrti (JFN)

16

Faza II: nulta hipoteza, očekivane vrednosti, devijacije

- novi aspekt: postoji ne samo jedna već *više* nultih hipoteza (H0)
 - *omnibus* H0: nema nikakvih razlika između proseka, uzetih *svi* zajedno
 - *dodatne* H0: npr. nema razlika među *parovima* proseka (biće obrađeno kasnije)
- omnibus H0: očekivane vrednosti *grupnih* proseka M_1, M_2, M_3, \dots jednake su međusobno, i jednake su *opštem* proseku M_y
- devijacije: postoje tri vrste, potpuno isto kao kod bivalentnih nacrti
 - *totalne devijacije:* razlika individualnih mera i opšteg proseka: $y = Y - M_y$
 - *unutargrupne devijacije (greške):* razlike ind. mera i grupnih proseka: $g = Y - M$
 - *međugrupne devijacije (efekti):* razlike grupnih pros. i opšteg pros.: $e = M - M_y$
- novi aspekt: u multivalentnim nacrtima postoje više od dva efekta
 - naime, postoji više od dva grupna proseka, jer postoji više od dve grupe odn. nivoa
- **PRIMER:** nećemo prikazati y i g , već samo e
- u primeru istraživanja sa padežima, sa četiri nivoa, ima četiri efekta e :
 - nominativ: $e_1 = M_1 - M_y = 680 - 704 = -24$; genitiv: $e_2 = M_2 - M_y = 700 - 704 = -4$;
 - dativ: $e_3 = M_3 - M_y = 720 - 704 = 16$; akuzativ: $e_4 = M_4 - M_y = 715 - 704 = 11$
- važe iste jednačine kao kod bivalentnih nacrti:
 - devijaciona jednačina: $y = e + g$; struktura jednačina: $Y = M_y + e + g$

1. Jednofaktorski nacrti (JFN)

17

Faza III: test statistik

- u analizi se koriste iste jednačine i koraci kao kod bivalentnih nacrti:
 - sve tri vrste devijacija sabiraju se do nule: $\Sigma y = 0$; $\Sigma e = 0$; $\Sigma g = 0$;
 - važi jednačina zbirova kvadrata: $\Sigma y^2 = \Sigma e^2 + \Sigma g^2$, odn. $SS_T = SS_A + SS_e$
 - važi proporcionalna jednačina: $r^2 + q^2 = 1$ (objašn. i neobj. varijabilnost)
 - novi aspekt: ovde r^2 nije kvadrat koeficijenta point-biserijalne korelacije (naime, ima više od dve serije), i označava se češće sa η^2 (eta-kvadrat)
 - napomena: postoje i drugačiji eta-kvadrati, npr. tzv. parcijalni eta-kvadrat
- neponovljeni MJFN
 - važe iste jednačine za srednje kvadrate: $MS_A = SS_A/df_A$, $MS_e = SS_e/df_e$
 - novi aspekt: budući da ima više od dva nivoa, formule za df se razlikuju
 - broj nivoa varijable A označimo sa a (tako da je za bivalentni nacrt: $a=2$)
 - važe formule: $df_A = a-1$; $df_e = a(N-1)$ za bival. nacrt: $df_A = 2-1 = 1$; $df_e = 2(N-1)$
 - formula za F-količnik se blago razlikuje, zbog df : $F = \frac{MS_A}{MS_e} = \frac{SS_A/(a-1)}{SS_e/a(N-1)}$
- ponovljeni MJFN: postoje slične razlike od bival. nacrti, zbog razlika df :
 - F-količnik: $F = \frac{MS_A}{MS_{rezi}} = \frac{SS_A/(a-1)}{SS_{rezi}/(a-1)(N-1)}$

1. Jednofaktorski nacrti (JFN)

18

Faza IV: p-vrednost, Faza V: značajnost: isto kao u bival. nacrtu

- uslovi korišćenja F-testa u multivalentnim JFN
 - isti su kao kod bivalentnih nacrti (slučajnost, nezavisnost, normalnost), uz izvesne izmene u vezi uslova homogenosti varijabilnosti podataka
 - kod *neponovljenih* nacrti, traži se homogenost varijabilnosti za podatke na svakom nivou (u svakoj grupi)
 - kod *ponovljenih* nacrti, traži se homogenost varijabilnosti *razlika* parova podataka za svaka dva nivoa
 - mogu se koristiti i složeniji statistički postupci, za koje nije potrebno da bude ispunjen uslov homogenosti
- Alternative F-testu
 - neparametarski testovi za neponovljene nacrti
 - Kruskal-Volison test (engl.: 'Kruskal-Wallis test')
 - test medijane (engl.: 'median test')
 - neparametarski testovi za ponovljene nacrti
 - Fridmanov test (engl.: 'Friedman test')
 - uočiti: t-test *nije* alternativa, ne može se koristiti za više od dva nivoa

1. Jednofaktorski nacrti (JFN)

19

Vrste testova u multivalentnim nacrtima

- isto kao i kod UFN sa više od dve kategorije, i kod multivalentnih JFN postoje više vrsta testova značajnosti
- razmotrili smo test *omnibus* H0: ne postoje razlike među prosecima uopšte
 - PRIMER:** da li postoji ikakva razlika u brzini reakcije na padeže?
 - ako je omnibus test statistički *značajan*, saznajemo da postoje *neke* razlike među prosecima, ali: ne saznajemo između kojih *parova* proseka postoje razlike
 - omnibus test je *globalan* test, jer testira postojanje razlika među svim prosecima *zajedno*; detaljnije informacije dobijaju se dodatnim *lokalnim* testovima
- dodatni (naknadni) testovi** se često nazivaju '**post hoc**' testovi (lat.: 'kasnije')
- najčešći dodatni testovi su **testovi kontrasta**
 - oni testiraju da li između dva proseka postoji statički značajna razlika
 - u bivalentnim nacrtima, test kontrasta je isto što i omnibus test
 - u trivalentnim nacrtima, sa prosecima M1, M2 i M3, postoje tri testa kontrasta:
 - testiranje razlike M1 i M2, razlike M1 i M3, i razlike M2 i M3
 - u nacrtima sa *a* nivoa postoji $a(a-1)/2$ moguća testa kontrasta
 - PRIMER!** u ogledu sa padežima ima ukupno $4(4-1)/2 = 6$ testova: da li postoji razlika u reakciji na nomin. i genit? na nominativ i dativ? na genitiv i akuzativ? itd.

1. Jednofaktorski nacrti (JFN)

20

- za testove kontrasta koristićemo odgovarajuće **oznake**:
 - test razlike između proseka M1 i M2, odn. nivoa a1 i a2 var. A ima oznaku $A_{1,2}$
 - test razlike između proseka M1 i M3 označava se kao $A_{1,3}$, itd
 - omnibus test, tj. test između svih nivoa varijable A, ima oznaku A
- ponekad, mada retko, vrše se testovi koje ćemo nazvati **parcijalni testovi**:
 - da li postoji stat. znač. razlika između više od dva ali manje od svih proseka
 - PRIMER:** da li postoji razlika u brzini reakcije na kose padeže (tj. one koji nisu nominativ, a to su genitiv, dativ i akuzativ), tj. proseka M2, M3 i M4?
 - oznaka: $A_{2,3,4}$
- ponekad, vrlo retko, vrše se **višestruki (multipli) testovi kontrasta**:
 - testiraju da li postoji statistički značajna razlika između *kombinacija* proseka
 - to je kontrast (porede se dve veličine), ali višestruki (zbog kombinovanja)
 - PRIMER:** da li postoji razlika u brzini reakcije na, s jedne strane, nominativ (a1), i, s druge strane, kose padeže uzete zajedno, kao grupa (a2+a3+a4)
 - oznaka: $A_{1,2+3+4}$
- postoje i **testovi trenda**: testovi oblika funkcionalne zavisnosti ZV od NV
 - PRIMER:** da li je funkcija zavisnosti linearna ili nelinearna
- statistički principi post hoc testova su slični kao kod omnibus testova

1. Jednofaktorski nacrti (JFN)

21

Potencijalni problemi sa post hoc testovima

- uočiti: postoji samo *jedan* omnibus test, a *mnogo* post hoc testova
 - ali, uglavnom će samo *neki* od tih testova imati *smisla* u datom istraživanju
- u analizi rezultata multivalentnih nacrtu preporučuje se vršenje *samo* smislenih, unapred planiranih testova
- nikako* se ne preporučuje: vršenje *svih mogućih* testova, da bi se utvrdilo koji od njih su statistički značajni
 - to je loš metodološki pristup i može lako dovesti do pogrešnih zaključaka
- naime: kada se vrši *veliki* broj testova, oni mogu biti *nepouzdati*
- PRIMER!** višestruko bacanje jednog ili više novčića
 - (1) pretpostavimo: bacimo jedan novčić 10 puta, a 9 puta padne pismo!
 - može se izračunati da verovatnoća da se to desi kod *ispravnog* novčića iznosi približno $p = 0.01$, tj. dešava se u oko 1% slučajeva
 - zaključak: novčić je verovatno neispravan, 'namešten'
 - (2) pretpostavimo: bacimo sto novčića po 10 puta, i to *ispravnih*
 - važi: verovatnoća da *bar* na jednom od 100 padne pismo 9 puta je veća od 0.5!
 - tj., veći su izgledi da se to desi (tj. veći su od 50%) nego da se ne desi
 - zaključak: pogrešili ili kad bi smatrali da je takav novčić neispravan!

1. Jednofaktorski nacrti (JFN)

22

- važna statistička činjenica: kada se vrši *veliki* broj testova, nivo značajnosti $\alpha=0.05$ nema isto značenje kao kada se vrši *jedan* ili mali broj testova
 - ako se vrši jedan test, $p=0.05$ znači da možemo očekivati da se, iako je H0 ispravna, dati ishod ipak zbiva oko 5 puta u 100 posmatranja, tj. dosta retko
 - ali, ako se vrši sto testova, možemo očekivati da se, iako je H0 ispravna, dobije takav 'statistički značajan' ishod u oko 5 od 100 slučajeva, tj. skoro sigurno!
- PRIMER!** loša statistička analiza može dovesti do pogrešnih zaključaka
 - test ekstraverzije za grupe osoba iz svakog od 12 astroloških znakova
 - statistička obrada (lošal!): porediti *sve moguće* parove znakova ($12 \cdot 11/2 = 66!$)
 - ishod: sasvim je moguće da se nađe znač. razlika ($p < 0.05$) bar kod jednog para
 - npr.: Bikovi su u većoj meri ekstravernti nego Vage
 - međutim, takva razlika lako može biti slučajna!
 - u ponovljenom ispitivanju ova razlika verovatno ne bi bila potvrđena (ali bi zato neka druga razlika mogla biti značajna, opet slučajno)
 - uticaj mantr e ili molitve na zdravlje
 - kod jedne grupe osoba koristi se mantra, kod druge ne (ili samo relaksacija)
 - uzme se u obzir *veliki* broj pokazatelja zdravlja kao ZV (pritisak, puls, apetit, ...)
 - statistička obrada (lošal!): utvrđuje se kod kojih od njih postoji značajna razlika
 - isti problem kao u prethodnom primeru: razlika je mogla nastati slučajno

1. Jednofaktorski nacrti (JFN)

23

- problem: kako *ispravno* analizirati istraživanja sa potencijalno veoma mnogo naknadnih testova?
- jedno moguće rešenje: uvesti strožiji nivo značajnosti
 - predloženo je da se umesto $\alpha=0.05$ koristi $\alpha=0.005$
 - time se smanjuje verovatnoća greške tipa 1 (prihvatanje nepostojećeg efekta)
 - ali se povećava verovatnoća greške tipa 2 (neprihvatanje postojećeg efekta)
- danas se problem uglavnom rešava na sledeće načine:
 - korišćenje standardnog nivo značajnosti, $\alpha=0.05$
 - ali: ne vršiti *sve moguće* testove, već samo *mali* broj *smislenih* testova
 - ti testovi bi trebalo da proističu iz neke teorije iz datog područja
 - korišćenje posebno konstruisanih, složenijih i strožijih post hoc testova
 - postoji veći broj takvih testova, i još uvek se razvijaju
 - nazivaju se prema statističarima koji su ih uveli
 - Tukey, Scheffé, Bonferroni ...
 - ti testovi se mogu podeliti u dve grupe:
 - a priori testovi:** planirani pre izvođenja istraživanja i dobijanja podataka
 - a posteriori testovi:** odluka da se vrše donosi se *nakon* uvida u podatke
 - razvijanje novog pristupa testovima značajnosti (tzv. Bayes faktori)

1. Jednofaktorski nacrti (JFN)

24

- opisani skup statističkih postupaka naziva se **analiza varijanse**
- skraćena: ANOVA (engl.: ANalysis Of VAriance)
- kod jednofaktorskih nacrtu, analiza varijanse se naziva:
 - jednofaktorska* (engl.: 'single factor') analiza varijanse, ili
 - jednosmera* (engl.: 'one-way') analiza varijanse
- ako je faktor *nepovljen*, to se obično ne naglašava
- ako je faktor *ponovljen*, analiza varijanse se naziva *ponovljena*
 - engl.: 'repeated measures ANOVA', tj. 'sa ponovljenim merenjima'
- kod ponovljenih nacrtu može se koristiti i tzv. *multipla ANOVA* (MANOVA)
 - složenija je od ANOVA, ali ne zahteva uslov homogenosti
- kompjuterski statistički programi za analizu varijanse često koriste jedan opštiji tip statističke obrade, nazvan **generalni linearni model**
 - ovaj se postupak koristi i kod korelaciono-regresionih nacrtu
- kompjuterski statistički programi *automatski* vrše *omnibus* test
 - takođe se automatski daje ispis zbiora kvadrata, prosečnih kvadrata, stepena slobode, i drugih veličina (često nepotrebnih istraživaču)
- istraživač može, ako želi, da vrši i **dodatne** testove (npr. testove kontrasta)
 - programi omogućavaju primenu *većeg* broja dodatnih testova