

VI Simpozijum MATEMATIKA I PRIMENE

Višestrukosti u biomedicini: Problem inflacije greške I vrste i neka rešenja

Dušica Gavrilović, mast.mat.



Institut za onkologiju i radiologiju Srbije

Beograd, 17.oktobar 2015.

Višestrukosti u biomedicini

Višestrukosti (3+) u biomedicini često potiču od:

- više tretmana (komparativnih grupa)
- više podgrupa (planiranih; dobijenih stratifikacijom)
- više promenljivih (u službi primarnog cilja istraživanja)
- više merenja vrednosti (jedne ili više) varijabli
- više statističkih analiza (*eng. Interim analysis*)
koje su u funkciji kontrole toka istraživanja, itd.

Mnoštvo, samim svojim prisustvom, omogućava višestruku primenu statističkih testova nad istim setom podataka

Višestrukosti u biomedicini

Na taj način nastaje

problem višestruke komparacije
(Multiple comparisons problem)

jer dolazi do

inflacije greške I vrste
(Alpha inflation problem)

Višestrukosti u biomedicini

Kako su ove pojave

nedovoljno poznate mnogim istraživačima,

statistička analiza se kod višestrukosti
sprovodi rutinski (bez planiranih pristupa),

dobijeni rezultati tumače uobičajeno
(bez neophodnih ograničenja)

što često rezultira

veoma pogrešnim zaključcima.

Višestrukosti u biomedicini

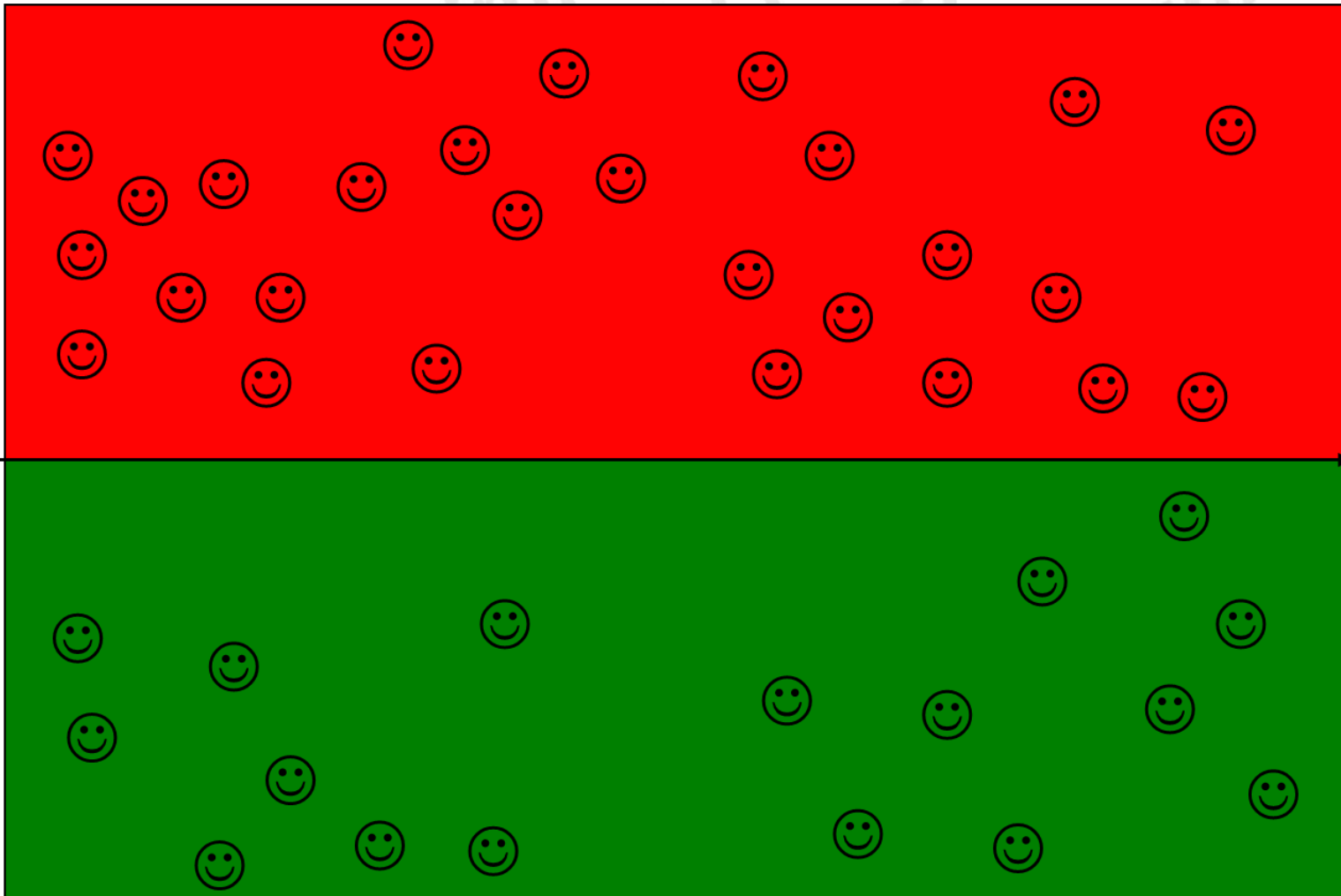
Čak i kada nema planiranih višestrukosti

- u potrazi za odgovorima
 - statistički značajnim rezultatima
 - ili usled neiskustva, neznanja
- može doći do pojave višestrukosti

U biomedicini, tipičan primer ovakve situacije jeste analiza velikog broja podgrupa nastalih stratifikacijom uzorka

Biomedicina: stratifikacija uzorka

Terapija

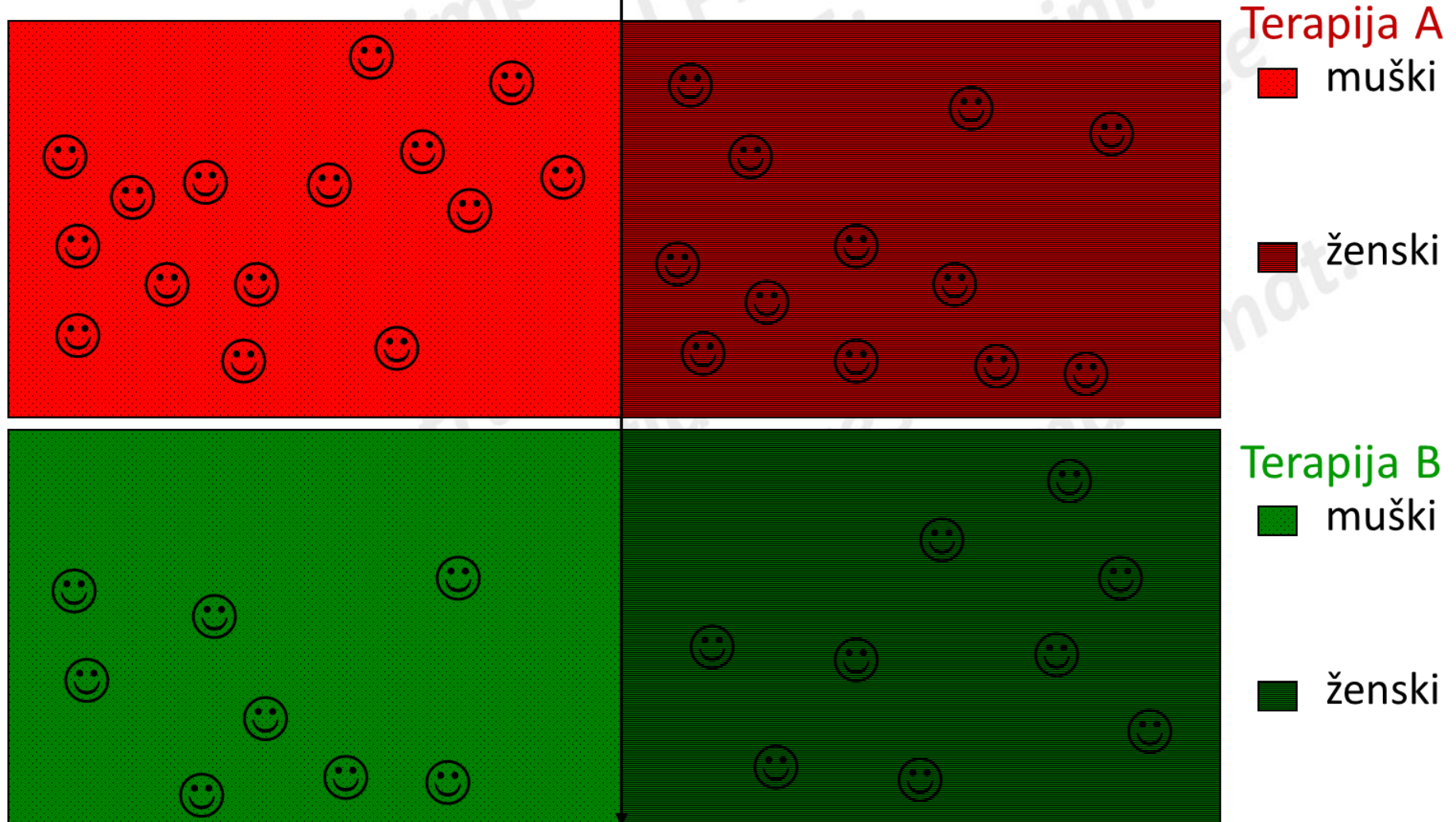


Terapija A

Terapija B

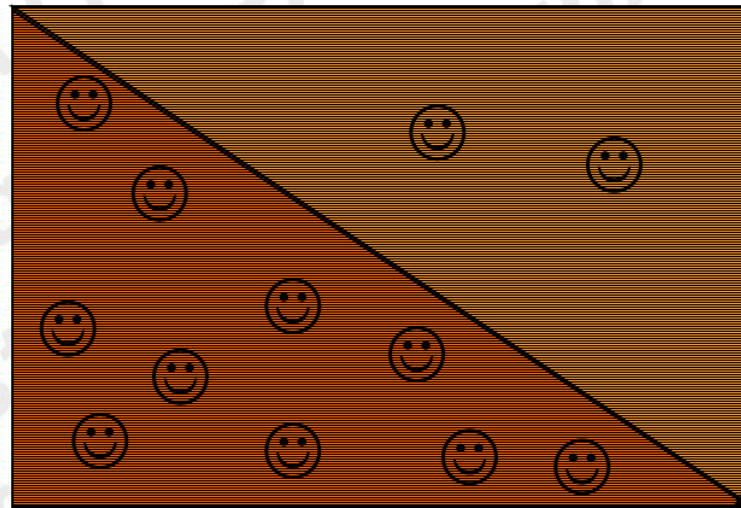
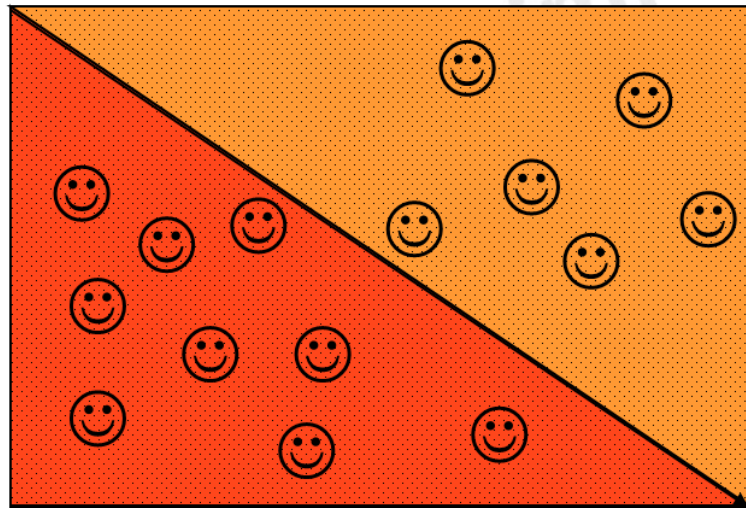
Biomedicina: stratifikacija uzorka

Terapija + Pol



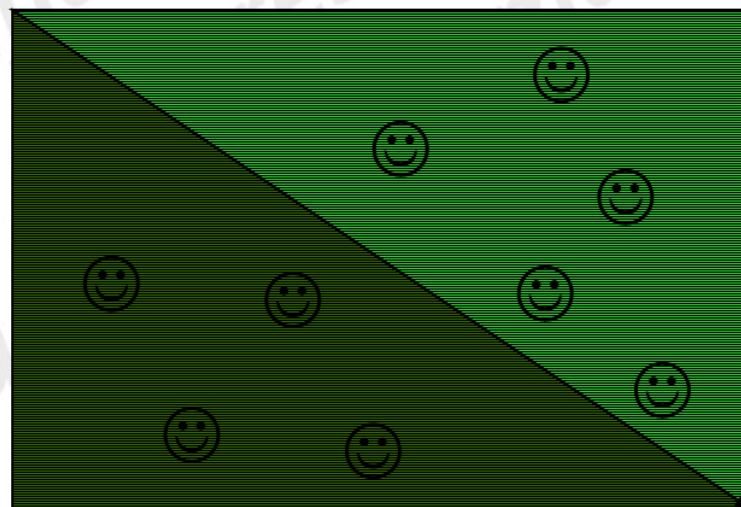
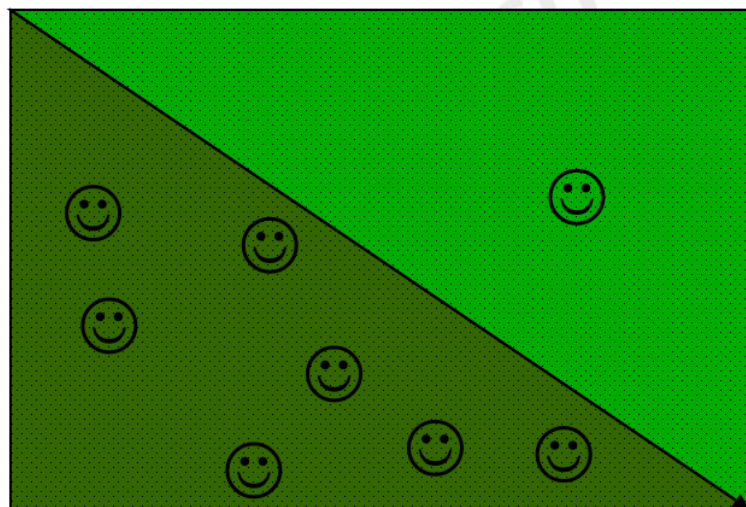
Biomedicina: stratifikacija uzorka

Terapija + Pol + Starost



Terapija A

- muški
- <45
- ≥ 45
- ženski
- <45
- ≥ 45



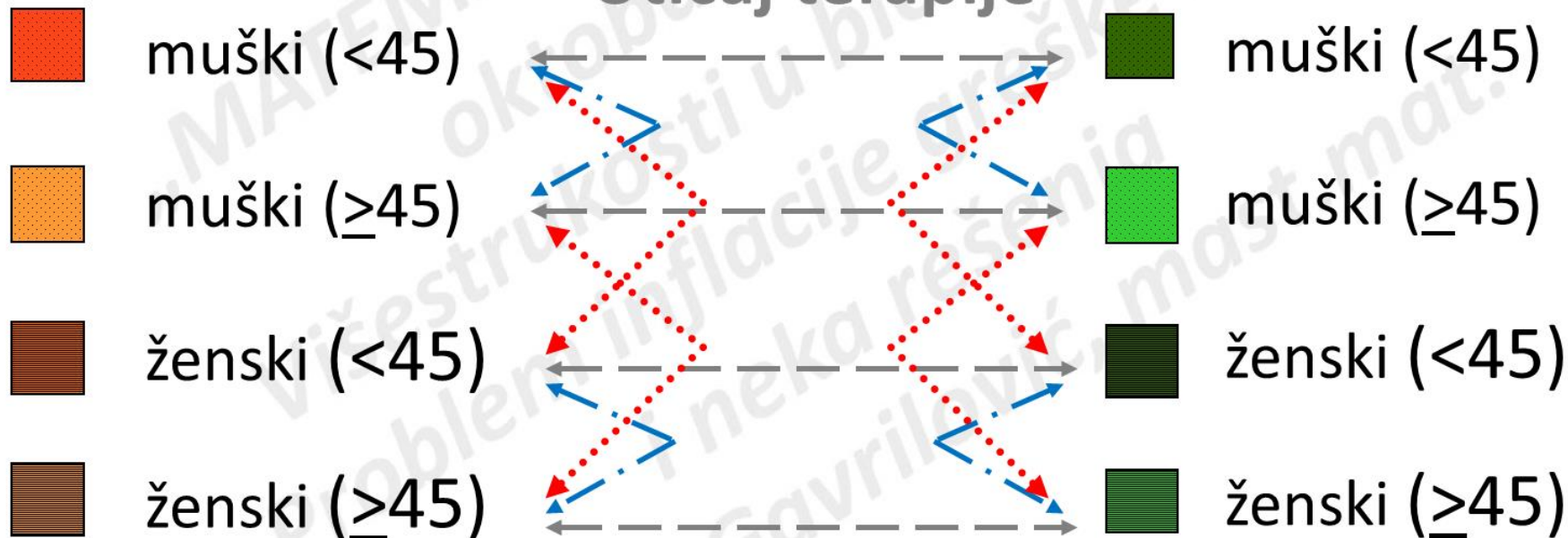
Terapija B

- muški
- <45
- ≥ 45
- ženski
- <45
- ≥ 45

Testiranje:

Terapija A

Terapija B



Uticaoaj pola

Uticaoaj starosti

Problem višestruke komparacije

(Multiple comparisons problem)



Previše testova nad istim setom podataka :

- uzrokuju „slučajne“ rezultate, otkrića, zaključke tj. vode u greške u odlučivanju o H_0
- nose rizik od lažno pozitivnih rezultata usled inflacije greške I vrste (α)

Problem višestruke komparacije

(Multiple comparisons problem)

- k - broj uzoračkih podgrupa ($k \in \mathbb{N}$, $k > 1$)
- $m = k \cdot (k - 1) / 2$ - broj testova (različiti parovi)
- α - nivo značajnosti testa
- verovatnoća slučajne greške I vrste:

$$p = 1 - (1 - \alpha)^m$$

Za $\alpha=0.05$								
m testova	1	2	3	4	5	10	15	20
$p = 1 - (1 - \alpha)^m$	0.05	0.10	0.14	0.18	0.23	0.40	0.54	0.64

Inflacija greške I vrste

„Istraživačka p -vrednost“

Svako “istraživačko p ”, bez unapred preciziranih:

- ciljeva (primarnog, sekundarnog i ostalih)
- svih oblika višestrukosti (podgrupe, ponovljena merenja...)
- statističkih metoda i pristupa višestrukostima
- minimuma prihvatljivog nivoa statističke značajnosti za rezultate višestrukih testiranja, itd.

naziva se nominalna p -vrednost

Nominalna p -vrednost obično uzrokuje grešku u tumačenju jer ukazuje na vezu koja je često rezultat slučajnosti!

Nominalna p -vrednost ne može biti osnov kauzalnih zaključaka!

Ona samo upućuje na mogući pravac povezanosti!
(zahteva dodatne analize)



Višestrukosti: rešenja

Navedeni problemi mogu se izbeći

- pažljivim planiranjem istraživanja
- planiranjem buduće statističke analize



To znači da u fazi planiranja treba izvršiti:

- **identifikaciju višestrukosti** (postojećih; potencijalnih)
- **izbeći ih, ako je moguće**
- ukoliko ih je nemoguće izbeći **izabrati, često i opravdati, optimalni pristup višestrukostima**

Optimalni pristup višestrukostima

1. Redukcija višestrukosti

- a) identifikacija “ključne” varijable
kada postoji više primarnih



Izbor primarne varijable u biomedicini

- često nije očigledan
- treba je pažljivo definisati

Hipertenzija:

- ✓ Primarna varijabla – sistolni, dijastolni pritisak

Onkologija:

- ? Primarna varijabla - odgovor na Th (CR, PR, SD, PD)
- ? Primarna varijabla - preživljavanje (OS, DFS, PFS...)
- ? Primarna varijabla - vreme do relapsa, PD...

Gerijatrija / palijativna terapija:

- ? Primarna varijabla - procena funkcionalnog stanja

Primarna varijabla: mogući pristupi

Istraživanje: efikasnost lečenja anemije

Primarna varijabla: nivo hemoglobina (Hgb)

Komparativna studija: lek 1 vs lek 2

Trajanje: 6 nedelja

Primarni cilj:

- ispitivanje razlika Hgb na početku i kraju
- ispitivanje razlika Hgb tokom više merenja
- ispitivanje porasta Hgb tokom istraživanja
- ispitivanje da li Hgb dostiže granicu normalnih vrednosti
- itd.

Optimalni pristup višestrukostima



Identifikacija “ključne”
varijable kada postoji više
primarnih

Primarni cilj:

- ispitivanje razlika Hgb na početku i kraju
- ispitivanje razlika Hgb tokom više merenja
- ispitivanje porasta Hgb tokom istraživanja

ispitivanje da li Hgb dostiže granicu
normalnih vrednosti

- itd.

Optimalni pristup višestrukostima

1. Redukcija višestrukosti

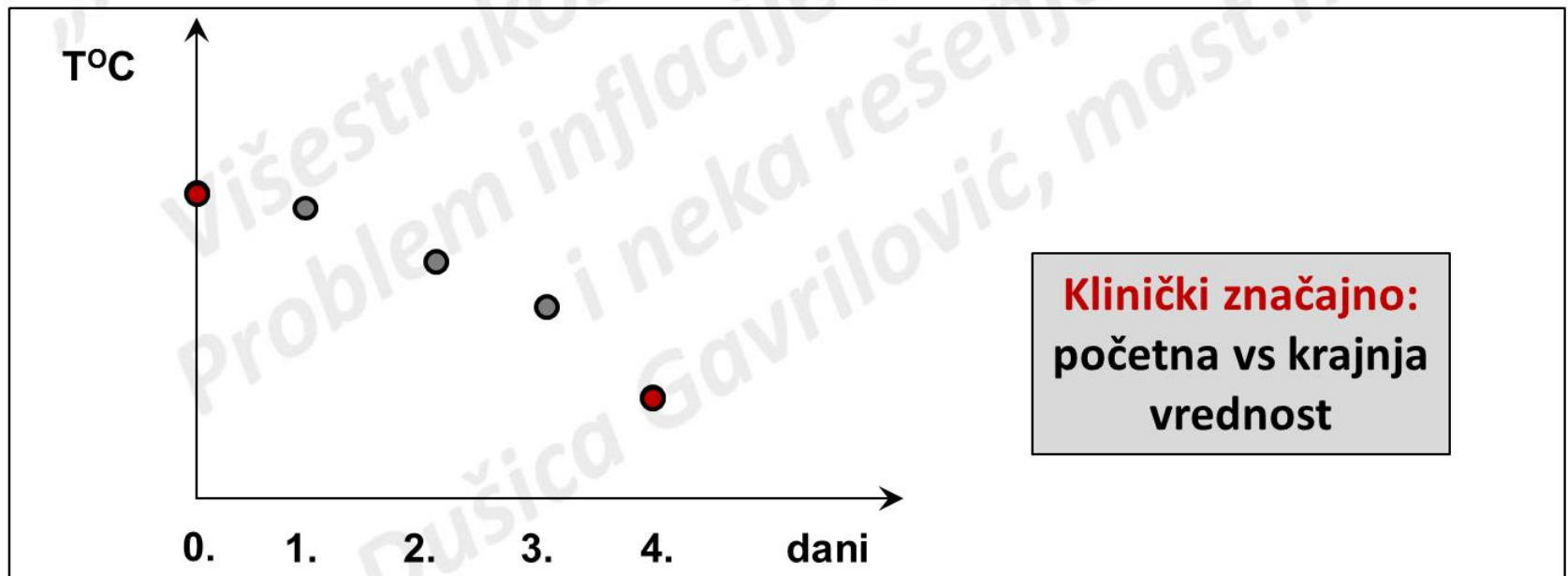
- a) identifikacija “ključne” varijable kada postoji više primarnih
- b) izbor biomedicinski najznačajnijeg ili “kritičnog” kontrasta kod višestrukih komparacija



Optimalni pristup višestrukostima



Izbor biomedicinski najznačajnijeg ili “kritičnog” kontrasta kod višestrukih komparacija



Optimalni pristup višestrukostima

1. Redukcija višestrukosti

- a) identifikacija “ključne” varijable kada postoji više primarnih
- b) izbor biomedicinski najznačajnijeg ili “kritičnog” kontrasta kod višestrukih komparacija
- c) upotreba “sumarnih” varijabli za ponovljena merenja



Optimalni pristup višestrukostima

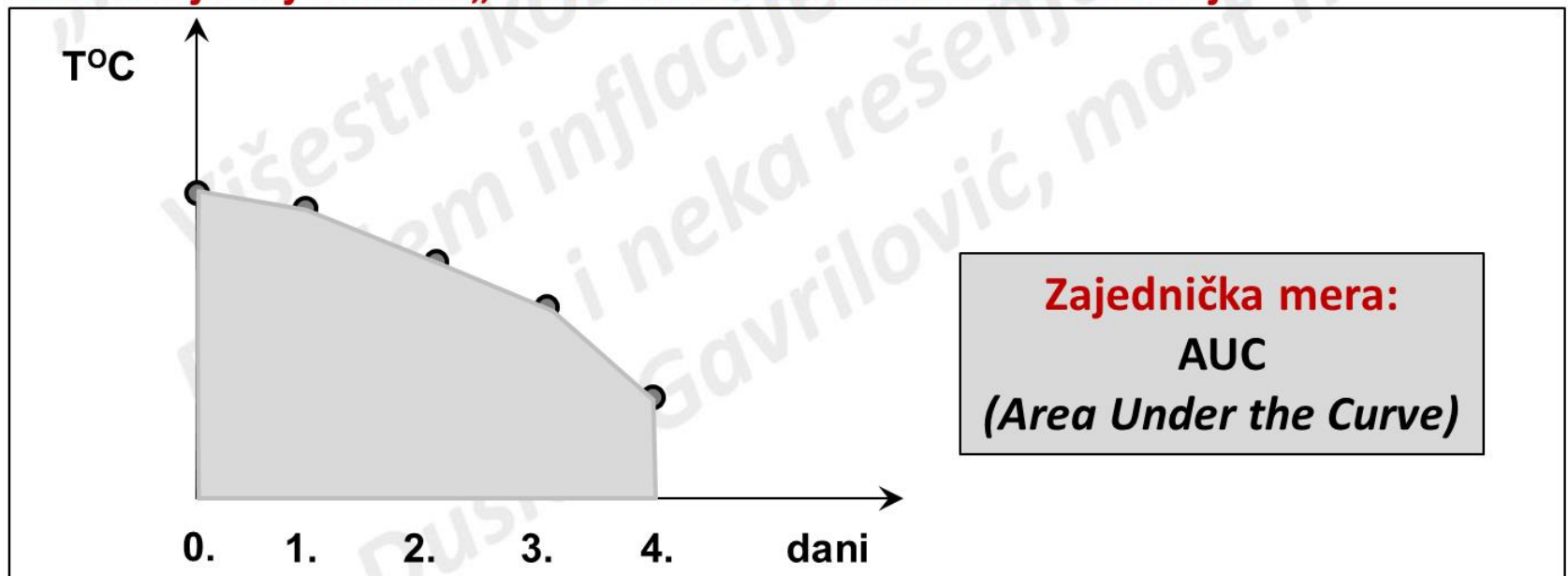


Upotreba “sumarnih” varijabli
za ponovljena merenja



AUC - površina ispod krive

Nalaženja zajedničke „mere“ za zavisna višestruka merenja



Optimalni pristup višestrukostima

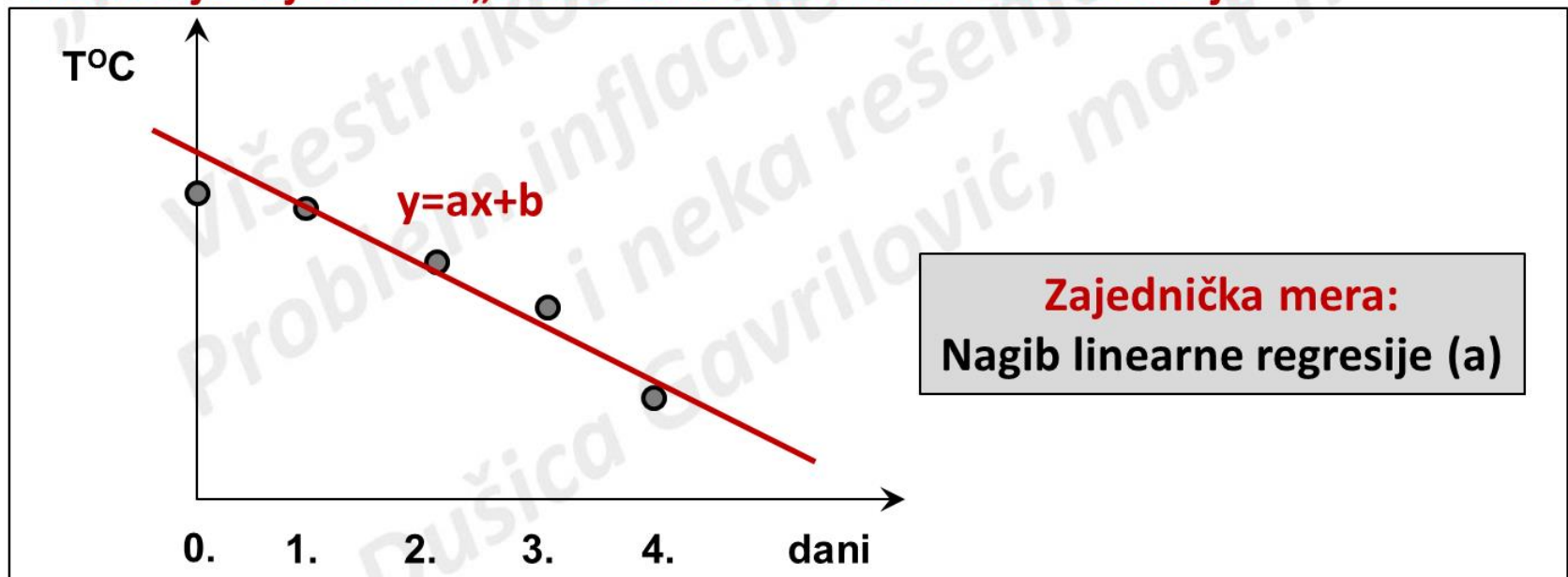


Upotreba “sumarnih” varijabli
za ponovljena merenja



Nagib linearne regresije

Nalaženja zajedničke „mere“ za zavisna višestruka merenja



Zajednička mera:
Nagib linearne regresije (a)

Optimalni pristup višestrukostima

1. Redukcija višestrukosti

- a) identifikacija “ključne” varijable kada postoji više primarnih;
- b) izbor biomedicinski najznačajnijeg ili “kritičnog” kontrasta kod višestrukih komparacija
- c) upotreba “sumarnih” varijabli za ponovljena merenja
- d) upotreba kompozitnih varijabli koje nastaju kombinovanjem više pojedinačnih varijabli



Optimalni pristup višestrukostima

Upotreba kompozitnih varijabli
koje nastaju kombinovanjem više
pojedinačnih varijabli



Grupisanje pitanja u anketi
(izračunavanje skorova)



Primer: SF-36 Health Survey

Koristi se za procenu kvaliteta života obolelih kroz sledeće aspekte:

- fizički
- psihički
- porodični
- profesionalni
- socijalni

Na osnovu opserviranih saznanja zaključuje se o potrebama obolelih i mogućnostima poboljšanja lečenja

Standardna anketa, sa metodologijom

Primer: SF-36 Health Survey

1. Uopšteno o zdravlju
2. Sadašnje zdravlje u odnosu na prošlogodisnje
3. Ograničenost u aktivnostima zbog zdravstvenog stanja
 - a. Naporna aktivnost
 - b. Umerena aktivnost
 - c. Dizanje/nosenje stvari
 - d. Stepence /nekoliko spratova
 - e. Stepence /1 sprat
 - f. Savijanje/čučenje/saginjanje
 - g. Hodanje/vise od 1 km
 - h. Hodanje/nekoliko stotina m
 - i. Hodanje/stotinak m
 - j. Samostalno kupanje/oblačenje
4. Problemi kao rezultat narušeno fizičkog zdravlje
 - a. Smanjivanje vremena potrošenog na rad ili u drugim aktivnostima

1. Opšte zdravlje pacijenta

2. Fizičko funkcionisanje

3. Uloga - fizička

Grupisanje odgovora (skorovi)

5. Problemi kao rezultat narušeno socijalno funkcionisanje
6. Fizičko zdravlje
7. Telesna bol
8. Uticaj telesnog bola na normalan posao
9. Kako su vam stvari „išle od ruke” i kako ste se osećali
 - a. Puni života
 - b. Vrlo nervozni
 - c. Potišteni
 - d. Spokojni, smireni
 - e. Imali ste dosta energije
 - f. Utučeni i umorni
 - g. Osecali ste iscrpljenosti
 - h. Srećni
 - i. Imali ste osećaj umora
10. Narušenost zdravlja kao uticaj na društvene aktivnosti
11. Koliko je za vas TAČNA ili NETAČNA svaka od sledećih tvrdnji?
 - a. Lakse se razboljevam
 - b. Zdrav sam kao i ostali koje znam
 - c. Očekujem pogoršanje zdravlja
 - d. Odlično sam zdravlje

5. Socijalno funkcionisanje

6. Mentalno zdravlje

7. Vitalnost

8. Telesna bol

Optimalni pristup višestrukostima

2. Izbor pogodnijih statističkih metoda



ANCOVA*, MANOVA*, MANCOVA*...

Kada postoje:

- faktori nehomogenosti
- poznate kovarijable
- *confounding factor*
- više zavisnih promenljivih...

Metoda	Slučajne promenljive		
	Nezavisne	Zavisne	Kovarijable
ANOVA	1+	1	-
ANCOVA	1+	1	1+
MANOVA	1+	2+	-
MANCOVA	1+	2+	1+

***Ukoliko su ispunjeni svi uslovi za njihovu primenu!**

Statističke metode: neke karakteristike

Karakteristike	Univarijantne (UV)	Multivarijantne (MV)*
Ispituju povezanost između nezavisnih varijabli i	1 zavisne varijabla	2 ili više zavisnih varijabli
Mera centr.tendencije	srednja vrednost	vektor srednjih vrednosti
Primer statističkih testova	t-test ANOVA ANCOVA	Hotelling – ov T^2 MANOVA MANCOVA
Povezanost između zavisnih varijabli	<u>ne vode računa</u>	<u>vode računa</u>
Priroda testiranja	odvojene analize za svaku zavisnu varijablu	istovremeno se analizira više zavisnih varijabli
Uticaj na grešku I vrste	zbog višestrukih testiranja uticaja nezavisnih varijabli na svaku od zavisnih varijabli, <u>povećava se inflacija greške I vrste</u>	<ul style="list-style-type: none"> • omogućavaju procenu zajedničkog modela • zato je često moguće generisati samo jednu test veličinu za testiranje MV H_0 • <u>bez inflacije greške I vrste</u>

***Ukoliko su ispunjeni svi uslovi za njihovu primenu!**

Optimalni pristup višestrukostima

3. Korigovanje granice statističke značajnosti za rezultate višestrukih testiranja



Podešavanje strožijeg praga statističke značajnosti za rezultate višestrukih testiranja

Optimalni pristup višestrukostima



U slučaju postojanja višestrukih komparacija, za njihove rezultate koriste se korigovane granice statističke značajnosti.

Postoji više metoda i algoritama

Korekcija Bonferonija (*Bonferroni correction*)

bez uslova o nezavisnosti grupa:

Ako je $m \in \mathbb{N}$ broj poređenja i α zadati nivo značajnosti, tada treba odbaciti svaku hipotezu H_{0i} , $i \in \{1, 2, 3, \dots, m\}$ ukoliko je pripadajuća verovatnoća $p_{0i} \leq \alpha_m$ gde je $\alpha_m = \alpha / m$

Optimalni pristup višestrukostima



U slučaju postojanja višestrukih komparacija, za njihove rezultate koriste se korigovane granice statističke značajnosti.

Postoji više metoda i algoritama

Šidakova procedura (*Sidak's multiple testing procedure*)
za nezavisne grupe

Ako je $m \in \mathbb{N}$ broj poređenja i α zadati nivo značajnosti, tada treba odbaciti svaku hipotezu H_{0i} , $i \in \{1, 2, 3, \dots, m\}$ ukoliko je pripadajuća verovatnoća $p_{0i} \leq \alpha_m$ gde je $\alpha_m = 1 - (1 - \alpha)^{1/m}$

Optimalni pristup višestrukostima



U slučaju postojanja višestrukih komparacija, za njihove rezultate koriste se korigovane granice statističke značajnosti.

Postoji više metoda i algoritama

Holmova procedura, Holm-Bonferoni metod

(Holm's step-up procedure; Holm-Bonferroni method)

bez uslova o nezavisnosti grupa

Neka su $p_{01} \leq \dots \leq p_{0m}$ sortirane verovatnoće dobijene testiranjem odgovarajućih hipoteza H_{01}, \dots, H_{0m} . Za zadato α neka je

$r = \min\{i \mid p_{0i} > \alpha / (m + 1 - i)\}$ gde $i \in \{1, \dots, m\}$.

Tada se odbacuju hipoteze $H_{01}, \dots, H_{0(r-1)}$.

- Ako je $r=1$, ne odbacuje se ni jedna hipoteza.
- Ako je $r \in \emptyset$, odbacuju se sve hipoteze.

Optimalni pristup višestrukostima



U slučaju postojanja višestrukih komparacija, za njihove rezultate koriste se korigovane granice statističke značajnosti.

Postoji više metoda i algoritama

Hošbergova procedura (eng. *Hochberg's step-up procedure*)

bez uslova o nezavisnosti grupa

Neka su $p_{01} \leq \dots \leq p_{0m}$ sortirane verovatnoće dobijene testiranjem odgovarajućih hipoteza H_{01}, \dots, H_{0m} . Za zadato α neka je

$r = \max\{i \mid p_{0i} \leq \alpha / (m + 1 - i)\}$ gde $i \in \{1, \dots, m\}$.

Tada se odbacuju hipoteze H_{01}, \dots, H_{0r} .

Optimalni pristup višestrukostima



U slučaju postojanja višestrukih komparacija, za njihove rezultate koriste se korigovane granice statističke značajnosti.

Postoji više metoda i algoritama

- Hošbergov metod je **moćniji** od Holmovog
- Hošbergova i Holmova metoda **moćnije** od Bonferonijeve

Postoje preporuke:

Ako je $m \in \mathbb{N}$ broj poređenja, tada se za

- ✓ $m < 10$ koristi **korekcija Bonferonija**
- ✓ $m \geq 10$ koristi **Holmova ili Hošbergova procedura**

Pristup višestrukim testiranjima: bez korekcije

Neki statističari smatraju da ne treba vršiti korekciju granice statističke značajnosti za rezultate višestrukih testiranja:

1. kada su višestruka testiranja u funkciji tumačenja rezultata

neki statističari preporučuju da se tada prikažu svi rezultati (sve p - vrednosti) i odgovarajući intervali poverenja, uz jasnu naglašenost da nema korekcije granice značajnosti za višestruka testiranja.

2. ako se sprovodi samo nekoliko planiranih poređenja

- naučno ili praktično opravdanih
- ukoliko nisu planirane, ne smeju se sprovoditi

3. kada su poređenja na komplementrnim podgrupama

tj. za testiranja koja se sprovode unutar međusobno disjunktivnih podgrupa (napr. muškarci /žene) jer se ponekad smatra da ovakva poređenja mogu pružiti podršku primarnom cilju istraživanja

Međutim, recezenti obično traže korekcije!

Drugi pristupi višestrukim testiranjima

- Poslednjih godina se vrše brojna istraživanja u genetici
- Tokom eksperimenata na DNK, RNK mogu se generisati veoma velike količine podataka koje omogućavaju procenu ukupnog stanja ćelija ili organizma
- Međutim, tako velika količina podataka može biti teška za analizu, naročito u odsustvu pogodnog zapisa za ekspresije gena
- Za potrebe takvih istraživanja sve više se korise tzv. **Microarray analysis** tehnike, namenjene tumačenju podataka dobijenih u eksperimentima
- Ove tehnike omogućavaju ispitivanje stanja (ekspresije) velikog broja gena, u nekim slučajevima čak ispitivanje u samo jednom eksperimentu celog genoma nekog organizma

U takvim istraživanjima se istovremeno testira hiljade hipoteza te je problem višestrukih testiranja jako izražen

Drugi pristupi višestrukim testiranjima:

FWER

Neka se nad istim setom podataka testira m nultih hipoteza

H_{0j} ($j = 1, \dots, m$) kojima odgovaraju H_{1j} ($j = 1, \dots, m$) alternativne hipoteze

Realnost

Rezultat testiranja H_0	H ₀ tačne	H ₁ tačne	Ukupno	
	Stat. značajno	V	S	R
	Nije stat. značajno	U	T	m-R
	Ukupno	m ₀	m-m ₀	m

- **m**: ukupan broj testiranja (tj. ukupan broj testiranih nultih hipoteza)
- **m₀**: broj stvarno tačnih nultih hipoteza
- **m-m₀**: broj stvarno tačnih alternativnih hipoteza
- **V**: broj lažno pozitivnih rezultata (*false positives* - *Type I error* - "false discoveries")
- **S**: broj potvrđenih stvarno pozitivnih rezultata (*true positives* - "true discoveries")
- **T**: broj lažno negativnih rezultata (*false negatives* - *Type II error*)
- **U**: broj potvrđenih stvarno negativnih rezultata (*true negatives*)
- **R**: broj odbačenih nultih hipoteza (*rejected null hypotheses* - "discoveries")

Drugi pristupi višestrukim testiranjima:

FWER

Neka se nad istim setom podataka testira m nulnih hipoteza

H_{0j} ($j = 1, \dots, m$) kojima odgovaraju H_{1j} ($j = 1, \dots, m$) alternativne hipoteze

Realnost

Rezultat
testiranja H_0

	H_0 tačne	H_1 tačne	Ukupno
Stat. značajno	V	S	R
Nije stat. značajno	U	T	$m-R$
Ukupno	m_0	$m-m_0$	m

Uobičajeni pristup za problem višestruke komparacije jeste podešavanje strožijeg praga statističke značajnosti za rezultate višestrukih testiranja.

Drugi pristupi višestrukim testiranjima:

FWER

Neka se nad istim setom podataka testira m nulnih hipoteza

H_{0j} ($j = 1, \dots, m$) kojima odgovaraju H_{1j} ($j = 1, \dots, m$) alternativne hipoteze

Realnost

		Realnost		
		H ₀ tačne	H ₁ tačne	Ukupno
Rezultat testiranja H ₀	Stat. značajno	V	S	R
	Nije stat. značajno	U	T	m-R
	Ukupno	m ₀	m-m ₀	m

FWER (Familywise error rate):

verovatnoća bar jednog lažno pozitivnog rezultata (greška I vrste)

$$\text{FWER} = P(V \geq 1) = 1 - P(V = 0)$$

V: broj lažno pozitivnih rezultata (*false positives* - *Type I error* - "*false discoveries*")

Drugi pristupi višestrukim testiranjima:

FWER

Neka se nad istim setom podataka testira m nulnih hipoteza

H_{0j} ($j = 1, \dots, m$) kojima odgovaraju H_{1j} ($j = 1, \dots, m$) alternativne hipoteze

Realnost

Rezultat
testiranja H_0

	H_0 tačne	H_1 tačne	Ukupno
Stat. značajno	V	S	R
Nije stat. značajno	U	T	$m-R$
Ukupno	m_0	$m-m_0$	m

Ukoliko se uvede znaka:

$$\alpha = \text{FWER} = P(V \geq 1) = 1 - P(V = 0); \alpha = 0.05$$

to znači da se možete očekivati da u svim sprovedenim komparacijama, tačne nulte hipoteze (m_0) budu slučajno statistički značajne u svega 5% slučajeva.

Drugi pristupi višestrukim testiranjima:

FWER

Neka se nad istim setom podataka testira m nulnih hipoteza

H_{0j} ($j = 1, \dots, m$) kojima odgovaraju H_{1j} ($j = 1, \dots, m$) alternativne hipoteze

Realnost

Rezultat
testiranja H_0

	H_0 tačne	H_1 tačne	Ukupno
Stat. značajno	V	S	R
Nije stat. značajno	U	T	$m-R$
Ukupno	m_0	$m-m_0$	m

Cilj **strožijeg definisanja granice** stat. značajnosti znači da, pod pretpostavkom da su sve H_0 tačne, postoji **95% šansi** da komparacije **ne dovode** do slučajno statistički značajnog rezultata

Kako se $\alpha=0.05$ odnosi na ceo eksperiment, postavljena granica FWER se naziva i *Experimentwise Error Rate (EWER)*

Drugi pristupi višestrukim testiranjima:

FDR

- **Genetaske analize** i *Microarray analysis* tehnike omogućavaju istovremeno testiranje hiljada hipoteza
- **Kontrola FWER** u takvim situacijama nije pogodna jer je **suviše konzervativna**
- Jedna od alternativa:
kontrola False Discovery Rate (FDR)

Realnost

Rezultat
testiranja H_0

	H_0 tačne	H_1 tačne	Ukupno
Stat. značajno	V	S	R
Nije stat. značajno	U	T	$m-R$
Ukupno	m_0	$m-m_0$	m

$$\text{FDR} = E(V/(V+S)) = E(V/R)$$

V: broj lažno pozitivnih rezultata (*false positives* - *Type I error* - "*false discoveries*")

R: broj odbačenih nultih hipoteza (*rejected null hypotheses* - "*discoveries*")

Drugi pristupi višestrukim testiranjima:

FDR

- **Genetaskе analize** i *Microarray analysis* tehnike omogućavaju istovremeno testiranje hiljada hipoteza
- **Kontrola FWER** u takvim situacijama nije pogodna jer je **suviše konzervativna**
- Jedna od alternativa:
kontrola *False Discovery Rate (FDR)*

Problem višestrukosti je veoma širok i veoma aktuelan.

Poslednjih godina se koristi i Bajesovski pristup.

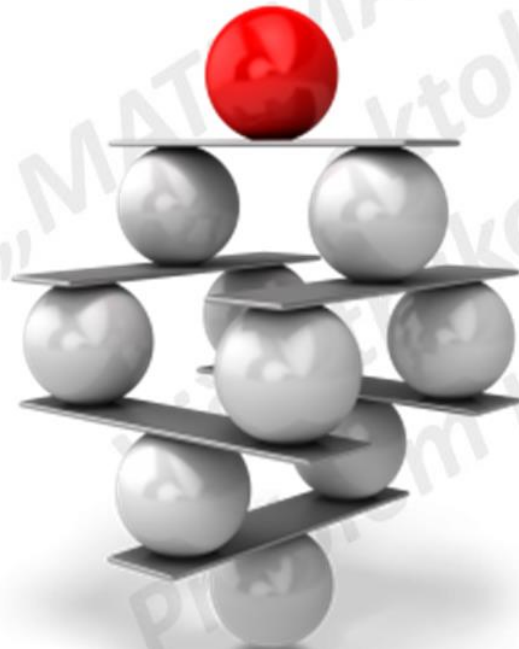
Problem višestrukosti i rešenja: zaključak



Pravilnom edukacijom istraživača i statističara

- o problemima koje izazivaju višestrukosti u istraživanjima
 - i mogućim pristupima u njihovom rešavanju
- postavljaju se dobre osnove za izbegavanje ili minimiziranje raznih grešaka**

Takva edukacija treba da bude kontinuirana i da prati savremene trendove u oblasti biostatistike, uz stalno naglašavanje važnosti faze PLANIRANJA istraživanja.



Hvala na pažnji