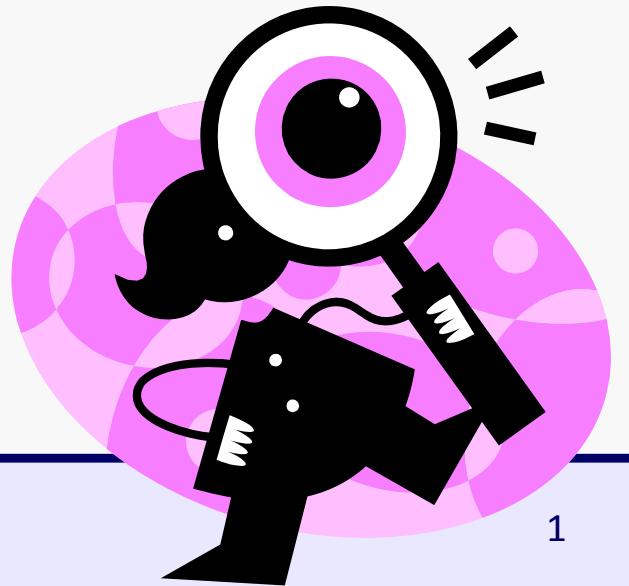


UVODNA RAZMATRANJA

izv. prof. dr. sc. Vesna Ilakovac

vesna.ilakovac@mefos.hr



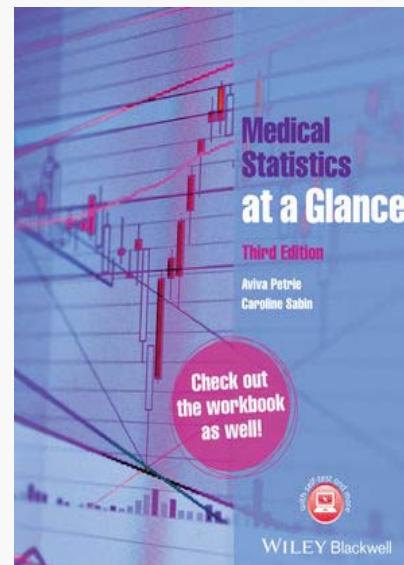
Literatura

- Dawson-Saunders B, Trapp RG. Basic & Clinical Biostatistics. Prentice-Hall Int. Inc., London
- Petrie A, Sabin C. Medical statistics at a glance. Blackwell Science, Oxford.
- Marušić M., Petrak J., Petrovečki M., Marušić A. Uvod u znanstveni rad u medicini. Medicinska naklada Zagreb.

2004.



2009.



2013.



MedCalc Manual

[https://www.medcalc.org/manual/index.php.](https://www.medcalc.org/manual/index.php)

MedCalc

easy-to-use statistical software

Home Features Download Order Contact FAQ **Manual**

Contents

Introduction

File menu

Edit menu

View menu

Format menu

Tools menu

Statistics menu

MedCalc manual - contents

Introduction

[Program installation](#)

[Auto-update](#)

[Regional settings support](#)

[Selection of display language](#)

Statistics menu (continued)

[Serial measurements](#)

[Reference intervals ...](#)

[Reference interval](#)

[Age-related reference interval](#)

PDF:

<http://www.medcalc.eu/download/medcalcmanual.pdf>



Medicinski fakultet Osijek

Katedra za medicinsku statistiku i medicinsku informatiku

Literatura

StatSoft, Inc. (2013). Electronic Statistics Textbook. Tulsa, OK:
StatSoft. WEB: <http://www.statsoft.com/textbook/>.

The screenshot shows the homepage of the StatSoft Electronic Statistics Textbook. At the top, there's a navigation bar with links for Software, Products, Solutions, Buy, Trials, Support, a search bar, and a 'Search' button. Below the navigation bar, a large banner features a blue grid pattern and a white box containing the text: 'Looking for info about statistics? We wrote the book on it. And you can read it for free!'. To the left of the main content area is a sidebar with various statistical analysis icons and their corresponding names: Elementary Concepts, Statistics Glossary, Basic Statistics, ANOVA / MANOVA, Association Rules, Boosting Trees, Canonical Analysis, CHAID Analysis, C & R Trees, Classification Trees, Cluster Analysis, Correspondence Analysis, Data Mining Techniques, Discriminant Analysis, Distribution Fitting, Experimental Design, Factor Analysis, and General Discrim. Analysis.

**Now Available
Printed Copy of this Textbook**

A printed version of this textbook is now available!
STATISTICS: Methods and Applications is \$80
which includes a coupon for a free CD version.
[click here to order or go to amazon.com.](#)

STATISTICS
Methods and Applications

StatSoft has freely provided the *Electronic Statistics Textbook* as a public service since 1995.

This textbook offers training in the understanding and application of statistics and data mining. The material was developed at the StatSoft R&D department based on many years of teaching undergraduate and graduate statistics courses. It covers a wide variety of applications, including laboratory research (biomedical, agricultural, etc.), business statistics, credit scoring, forecasting, social science statistics and survey research.

Testimonials

"Thank you and thank you again for providing a complete, well-structured, and easy-to-understand online resource. Every other website or snobbish research paper has not deigned to explain things in words consisting of less than four syllables. I was tossed to and fro like a man holding on to a frail plank that he calls his determination until I came across your electronic textbook... You have cleared the air for me. You have enlightened. You have illuminated. You have educated me."

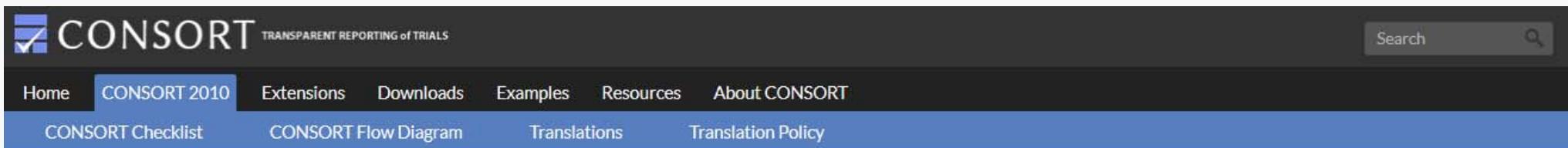
— Mr. Daniel Chen Kwok Lee
October 2013

Literatura

Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. [Ann Int Med 2010;152](#).

(<http://annals.org/article.aspx?articleid=745807>)

<http://www.consort-statement.org/consort-2010>



CONSORT 2010

The CONSORT (CONsolidated Standards of Reporting Trials) 2010 guideline is intended to improve the reporting of parallel-group randomized controlled trial (RCT), enabling readers to understand a trial's design, conduct, analysis and interpretation, and to assess the validity of its results. This can only be achieved through complete adherence and transparency by authors.

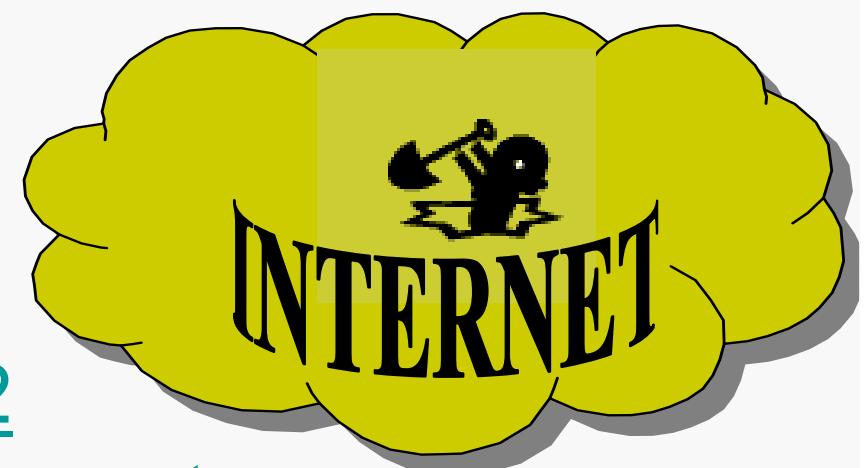
CONSORT 2010 was developed through collaboration and consensus between clinical trial methodologists, guideline developers, knowledge translation specialists, and journal editors (see [CONSORT group](#)). CONSORT 2010 is the current version of the guideline and supersedes the 2001 and 1996 versions. It contains a 25-item [checklist](#) and [flow diagram](#), freely available for viewing and [downloading](#) through this website.

[Extensions of the CONSORT Statement](#) have been developed for different types of trial designs, different interventions, and different types of data.

The most important documents for understanding the CONSORT 2010 statement are the following:

1. **The CONSORT 2010 Statement:** This is a declaration of the standard, and how it was developed. This declaration or "statement" has been published in many prominent journals, including the British Medical Journal, the Lancet and PLoS Medicine. You can download the CONSORT 2010 Statement documents [here](#).
2. **The CONSORT 2010 Explanation and Elaboration (E&E) Document:** This 32-page document provides an explanation of each of the CONSORT Checklist items, and how to apply them in your reporting. You can download the CONSORT 2010 E&E Document [here](#).
3. **The CONSORT 2010 Checklist:** This is a checklist of each of the 25 items that must be reported in a Randomized Clinical Trial Report in order for it to be compliant with the standard. You can see an example of how the CONSORT checklist can be applied to a well-reported study by navigating to the [Sample Study](#) in the menu above. You can download the CONSORT 2010 Checklist [here](#).
4. **The CONSORT 2010 Flow Diagram:** This is a simple flow diagram showing how your study population was recruited and handled during the course of your study. You must complete a flow diagram in order to be compliant with the CONSORT 2010 standard. You can download the CONSORT 2010 Flow Diagram template [here](#).

Dopunsko štivo



BMJ STATISTICS NOTES

www.csm-oxford.org.uk/?o=1292

www.bmj.com/specialties/statistics-notes

How to Upset the Statistical Referee

www-users.york.ac.uk/~mb55/talks/upset.htm

Lang T. Twenty statistical errors even YOU can find in biomedical research articles. Croat Med J 2004;45:361-70.

www.cmj.hr/2004/45/4/15311405.htm

.....

Lako čitljiva literatura.....

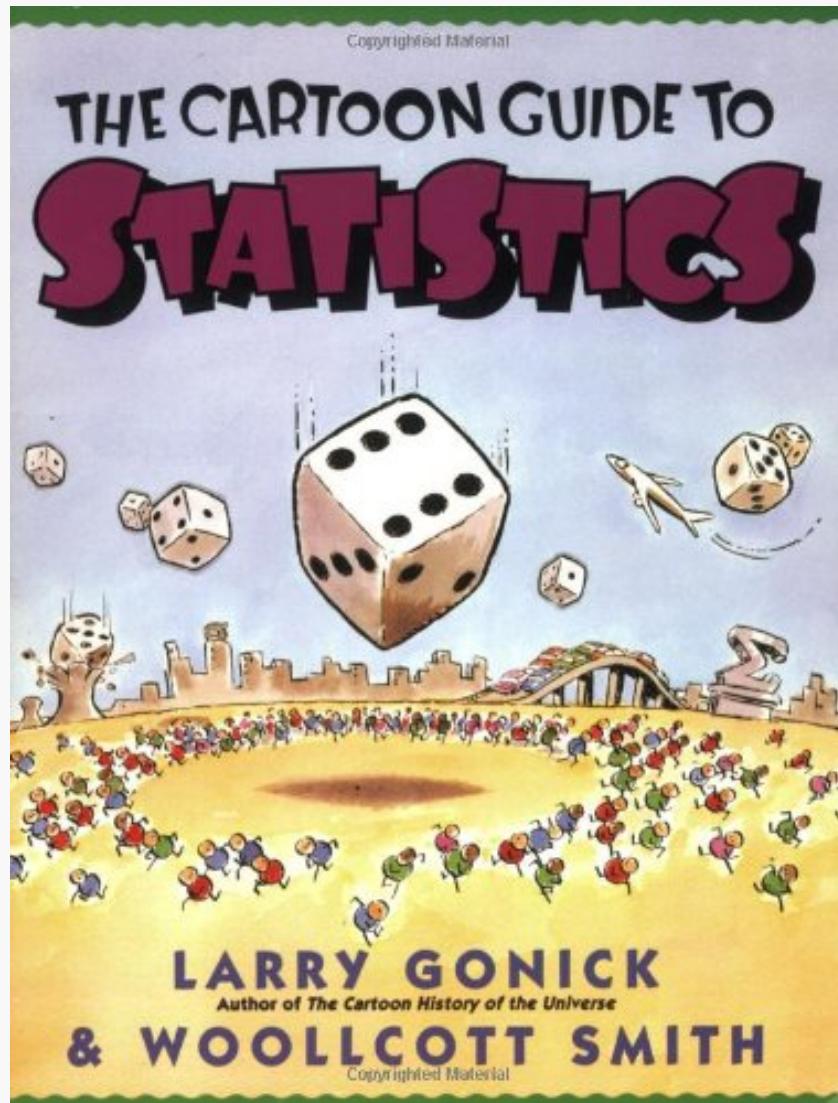
Petz B, Kolesarić V, Ivanec D.

Petzova statistika : osnovne statističke metode za nematematičare. Naklada Slap.

2012.

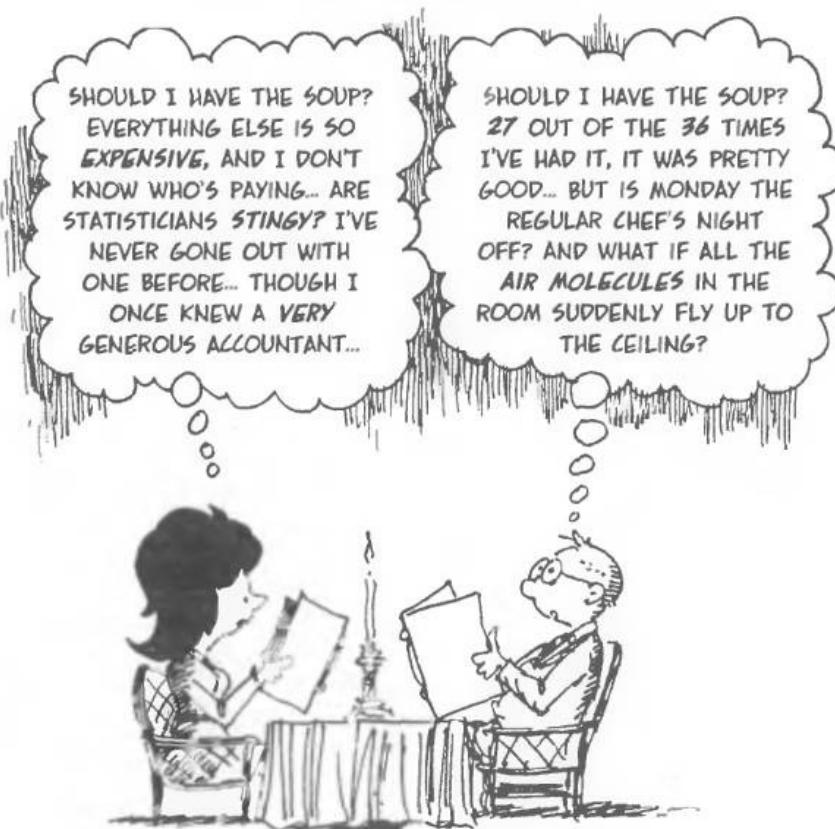


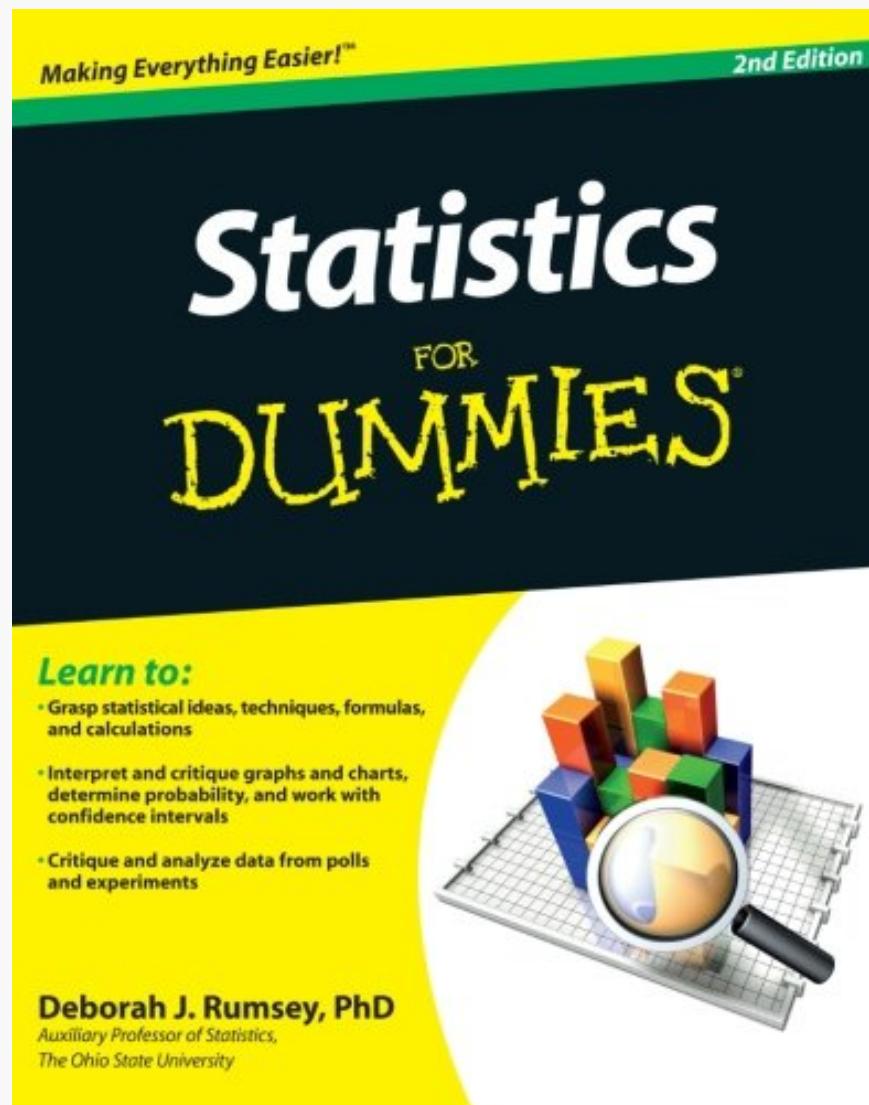
Ako ništa ne pomaže, pokušajte započeti s



WHAT IS STATISTICS?

WE MUDDLE THROUGH LIFE MAKING CHOICES
BASED ON INCOMPLETE INFORMATION...





SVRHA ISTRAŽIVANJA

- **opis**
 - stanje u populaciji
- **usporedba**
 - novi postupak VS stari postupak
- **povezanost**
 - rizični čimbenik i bolest/stanje

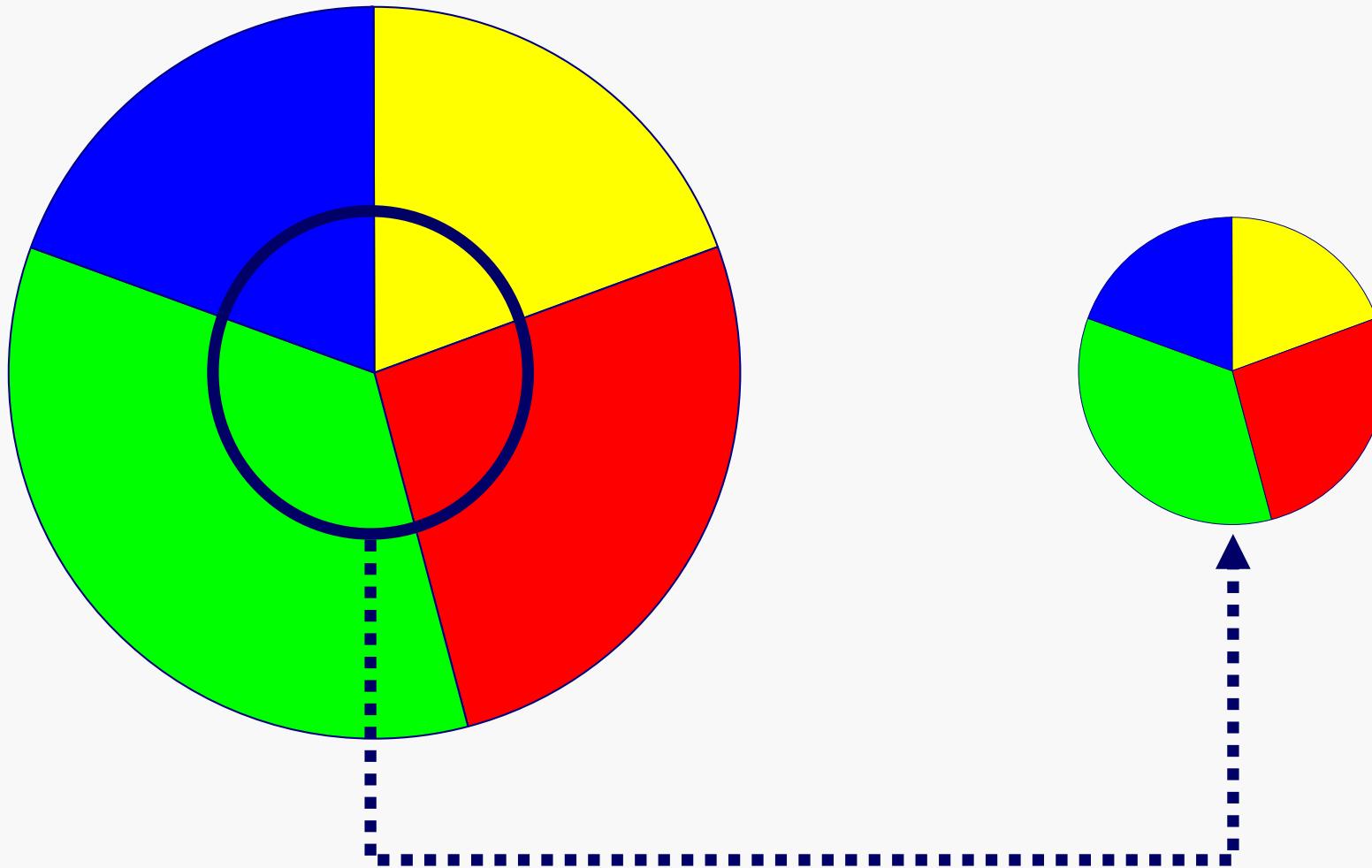
TEMELJNE VRSTE ISTRAŽIVANJA

- **opažajna istraživanja (observational studies)**
 - presječno istraživanje (cross-sectional)
 - istraživanje slučajeva i kontrola (case-control)
 - kohortno istraživanje (cohort)
- **pokusna istraživanja**
 - randomizirani kontrolirani pokus (randomised controlled trial-RCT)



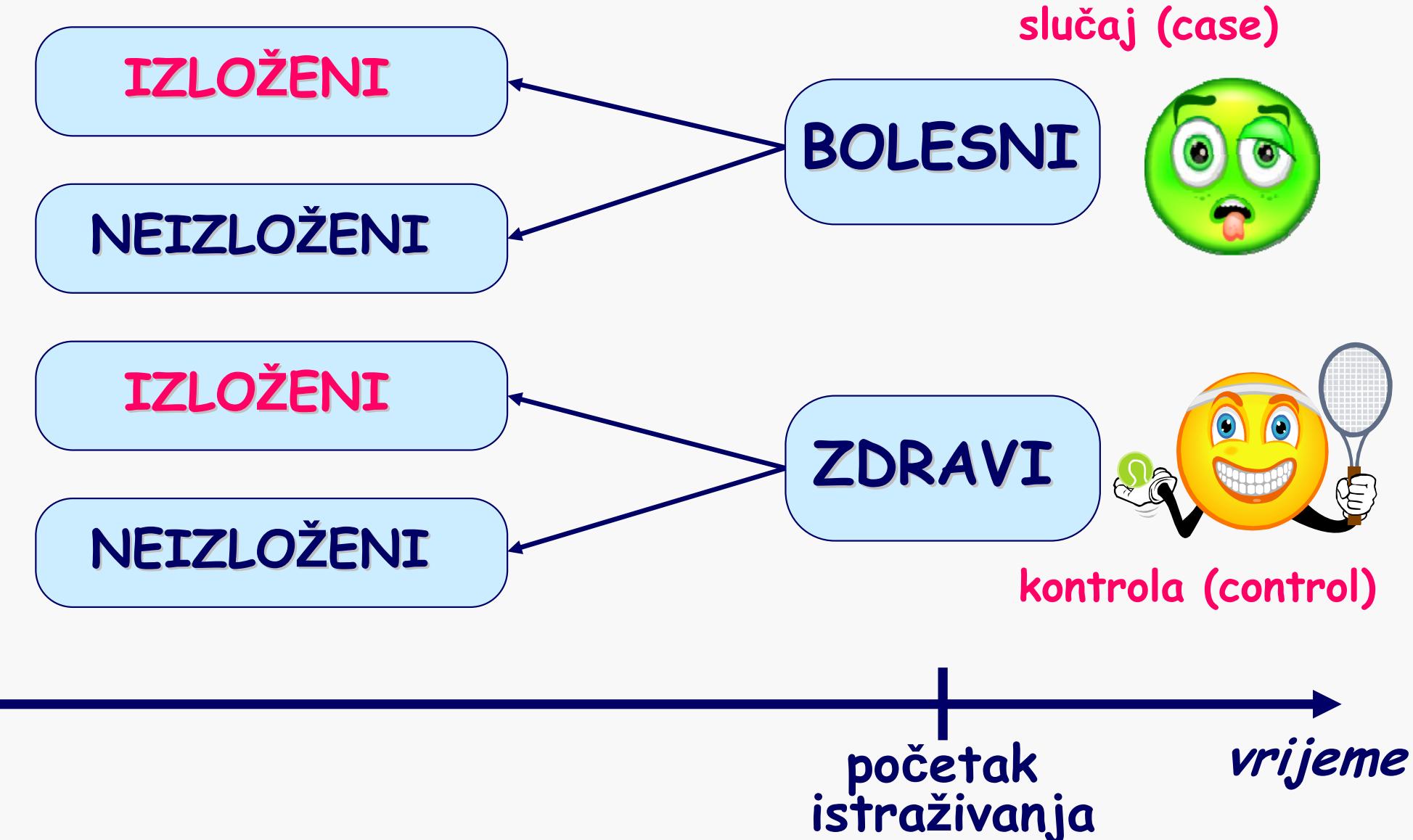
PRESJEČNO ISTRAŽIVANJE (cross-sectional)

"Što se dešava?"



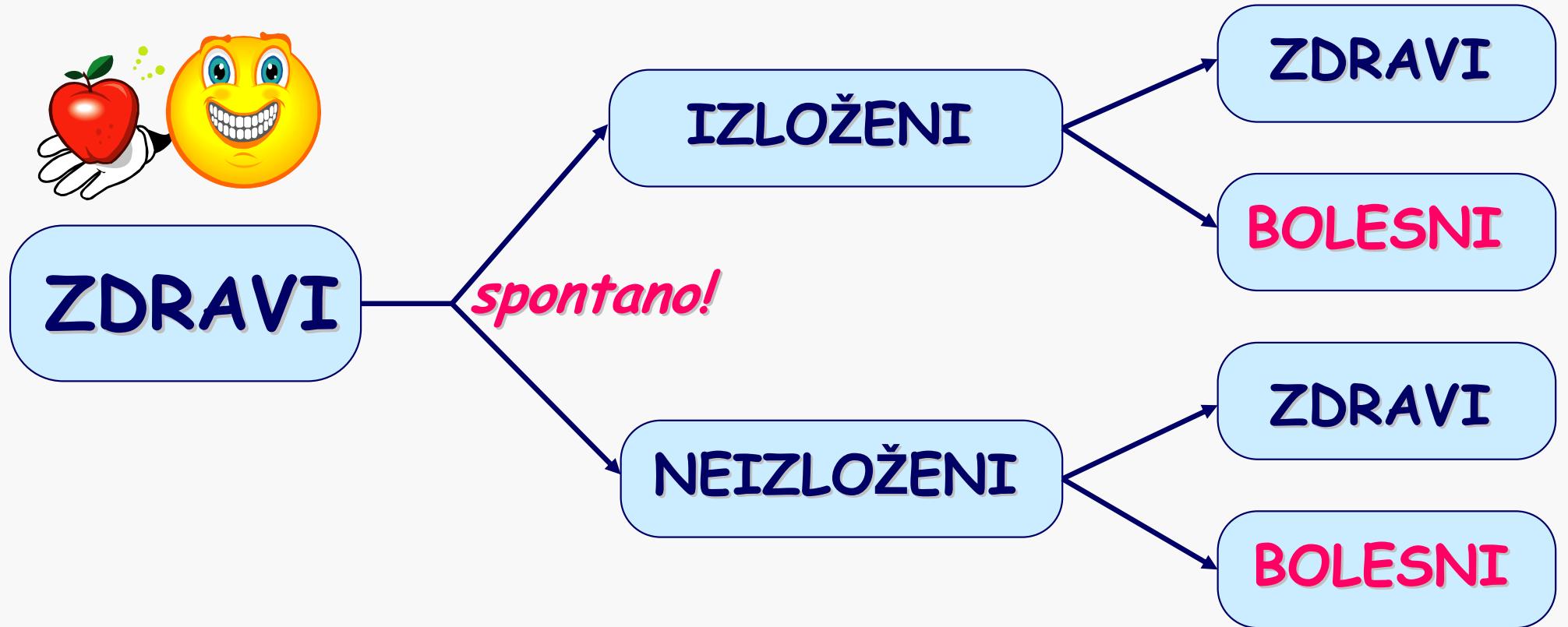
ISTRAŽIVANJE SLUČAJEVA I KONTROLA (case-control)

"Što se desilo?"



KOHORTNO ISTRAŽIVANJE

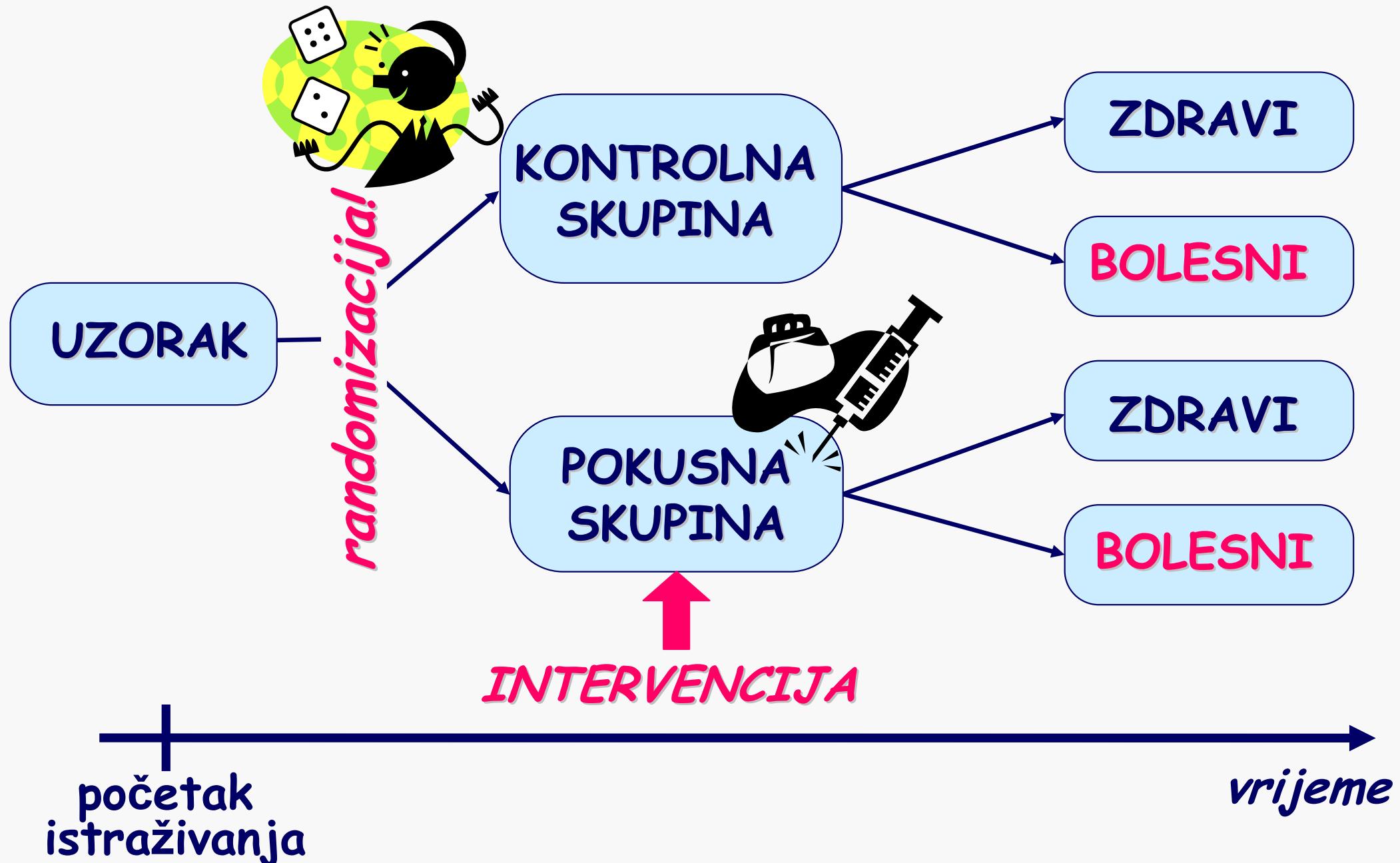
"Što će se desiti?"



početak
istraživanja

vrijeme

RANDOMIZIRANI KONTROLIRANI POKUS (RCT)



PRIKRIVANJE (*masking, blinding*)

- **jednostruko prikriveni (slijepi) pokus**
 - ispitanik ne zna kojoj skupini pripada



- **dvostruko slijepi pokus**
 - niti ispitanik niti istraživač ne znaju kojoj skupini pripada ispitanik

- **trostruko slijepi pokus**
 - niti ispitanik, niti istraživač niti osoba koja obrađuje podatke ne zna tko pripada ispitivanoj skupini



"DOKAZI" IZ ISTRAŽIVANJA

- različite vrste istraživanja donose "dokaze" različite "snage"
- dvostruko slijepi RCT smatra se "zlatnim standardom" za donošenje najsnažnijih uzročno-posljedičnih znanstvenih zaključaka

ISTRAŽIVANJE

"DOKAZNA" SNAGA ISTRAŽIVANJA

OPAŽAJNO

bez kontrole

slučajeva i kontrola

kohortno

POKUSNO

ne randomizirana kontrola

randomizirana
kontrola
(RCT)

bez prikrivanja

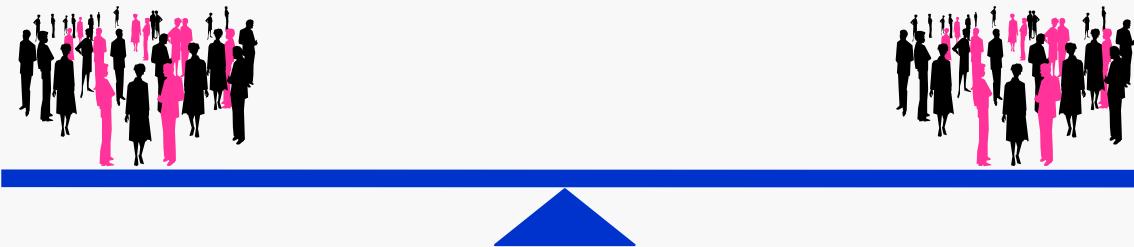
jednostruko slijepi

dvostruko slijepi



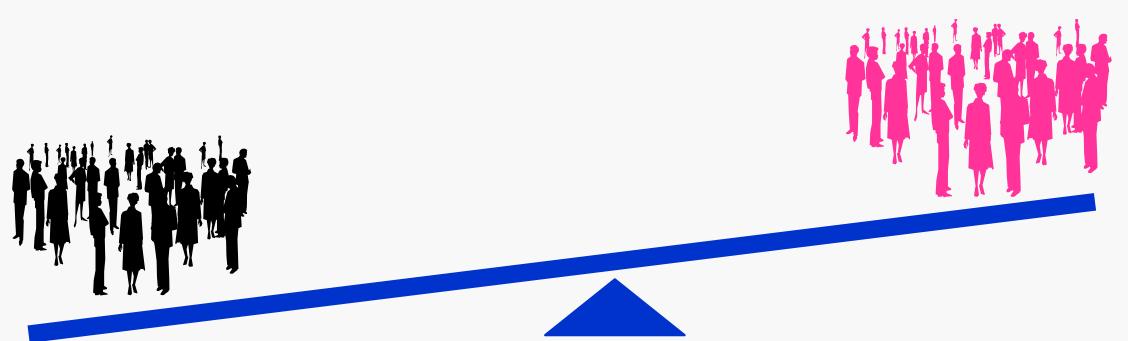
ZAŠTO JE POTREBNA RANDOMIZACIJA (nasumični odabir)?

- slučajnost izbora u postupku izbora grupa (tko dobiva koji tretman) **izjednačava grupe** s obzirom na poznate i nepoznate prognostičke faktore



"POŠTENO"!

PRISTRANO!



RANDOMIZACIJA (nasumični odabir)

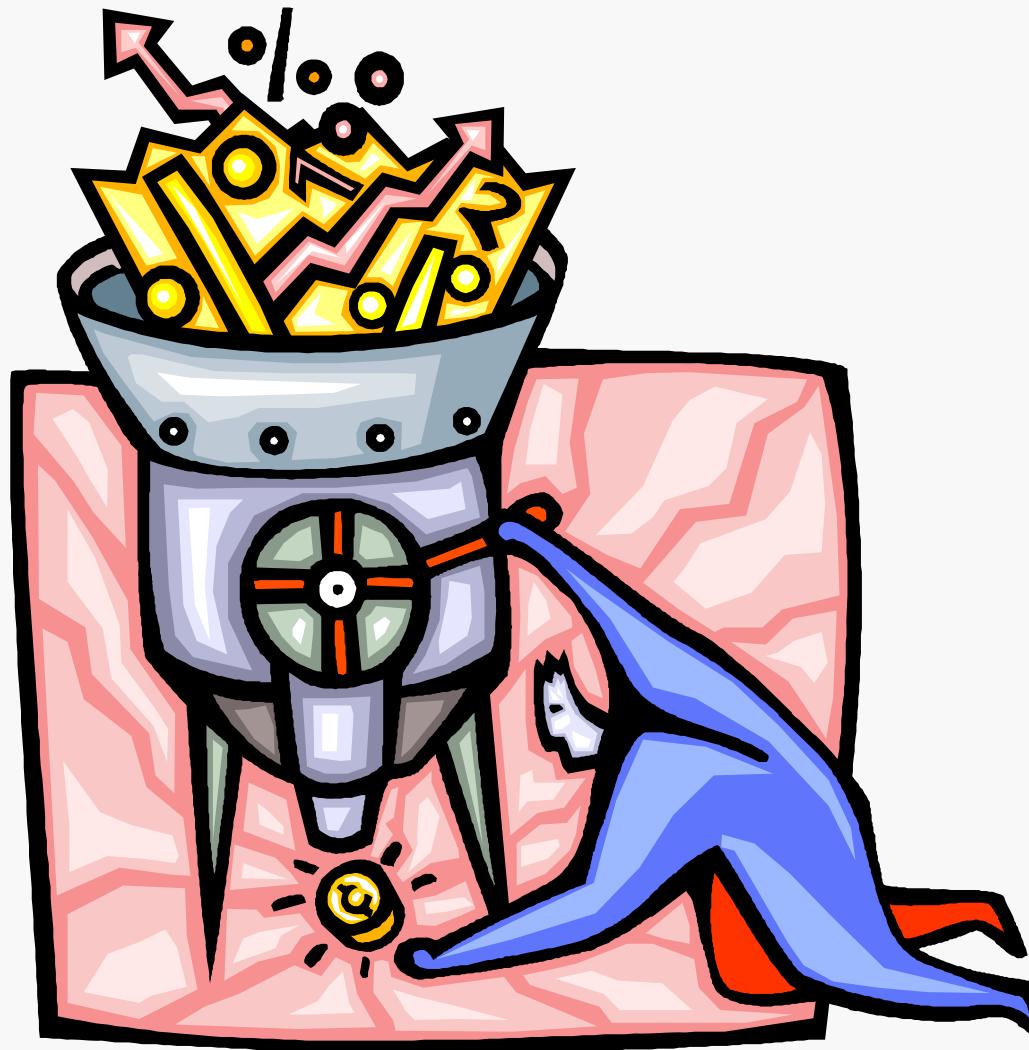
- generatori slučajnih brojeva
 - mjerenja fizikalnih pojava (atmosferski šum, terbalni šum, ...)
 - matematički algoritmi (pseudo-slučajni brojevi)
- biranje slučajnih brojeva:
 - Excel: RAND(), RANDBETWEEN()
 - <https://www.random.org/>
 - <http://www.graphpad.com/quickcalcs/randMenu/>
 - <http://www.randomizer.org/>

ZAŠTO JE POTREBNO PRIKRIVANJE?



- prikrivanje alokacije tretmana je dio dobrog nacrta studije
- izbjegavaju se moguće optužbe za pristranost ukoliko niti ispitanici niti ispitivači ne znaju pripadnost pojedinim grupama
- dvostruko slijepi pokus je najbolji način da se očuva znanstvena strogost RCT

STATISTIKA?



....neki razlozi - praćenje literature....

“ ...povećan je rizik koronarnog incidenta kod bolesnika čija je koncentracija CRP u petoj kvantili u odnosu na prve četiri kvantile...”

“ ...srednja vrijednost dobi 42 ± 8 godina...”

“ ...značajnost razlike testirana je Studentovim T-testom...”

“ ... $p < 0.05$...”

Multiple R	.96764
R Square	.93632
Adjusted R Square	.92883
Standard Error	6.54079

....neki razlozi - deskripcija i analiza rezultata....

Rezultati mjerenja visine studenata prve godine:

188	175	179	179	173	193	183	177	165	170
165	164	168	182	193	183	160	183	165	166
168	172	178	193	167	174	176	176	172	181
169	172	184	190	182	176	176	164	162	167
189	187	168	175	173	182	175	165	167	181

Tko je najviši ? Tko je najniži ?

Koju visinu ima najveći broj studenata?

Kako varira visina studenata?

Koja je visina kojoj teži najveći broj rezultata?

?????

?

neki razlozi - zaključivanje iz pojedinačnog na “opće”..

Kada i pod kojim uvjetima možemo zaključivati o populaciji iz rezultata mjerjenja provedenih na nekoj skupini ispitanika?

Kako odabratи ispitanike?

Koliko ispitanika uključiti u promatranje?

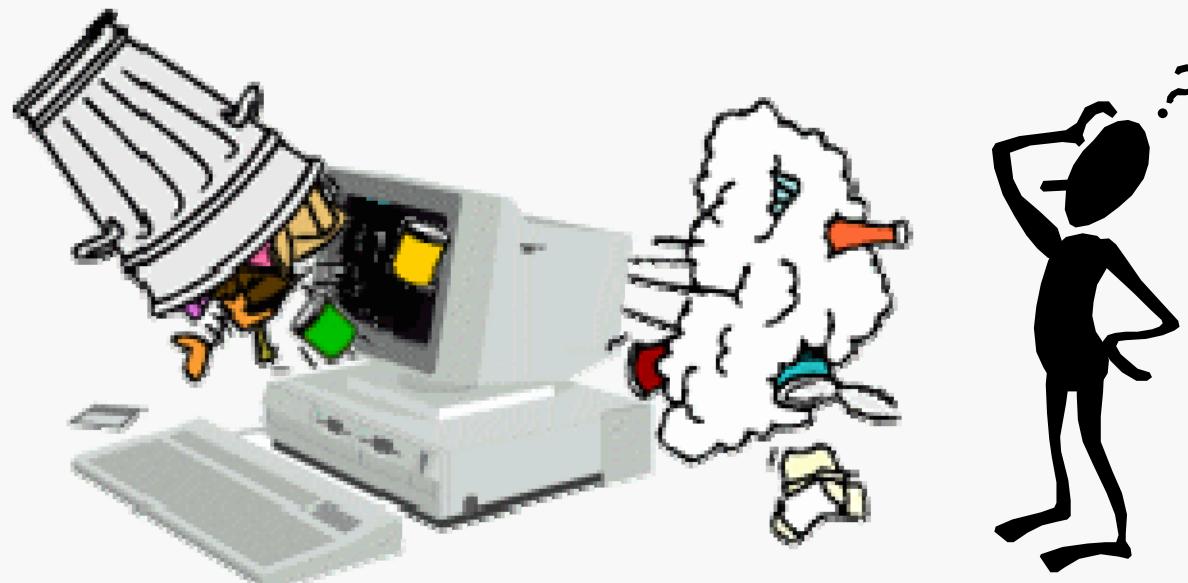
Kolika je pogreška pri zaključivanju?

...



neki razlozi - planiranje istraživanja i pokusa

- metoda "što ispadne" u istraživanju i eksperimentu može rezultirati nepouzdanim i neinterpretabilnim rezultatima





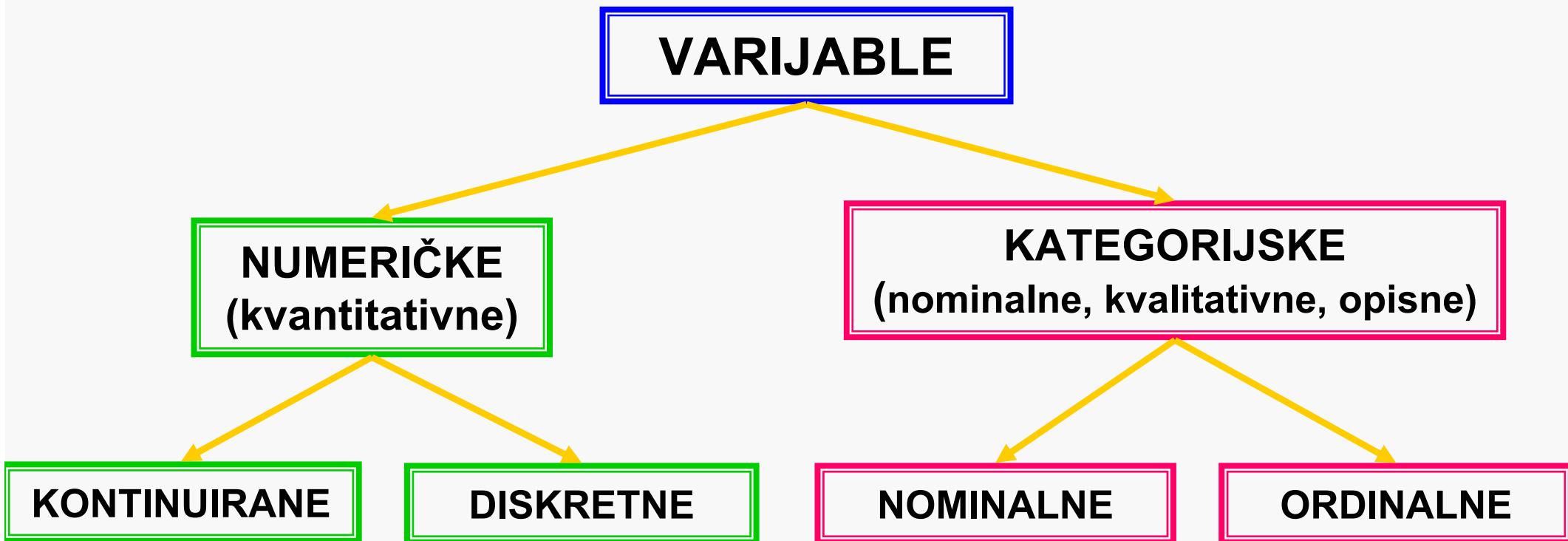
Sir Ronald Aylmer Fisher (1890. –1962.)

“To consult the statistician after an experiment is finished is often merely to ask him to conduct a *post mortem* examination.

He can perhaps say what the experiment died of.”

R. A. Fisher, 1938.

VARIJABLE



Kvantitativne (numeričke) varijable

- **kontinuirane varijable**

- varijable koje teorijski prepostavljaju postojanje beskonačnog broja vrijednosti
- **u pravilu dobivene mjerenjem**
 - visina, težina, dob, tlak,

- **diskretne varijable**

- varijable čija se vrijednost uvećava ili umanjuje za cijeli broj jedinica
- **u pravilu dobivene brojanjem**
 - broj popuštenih cigareta na dan, broj otkucaja srca u minuti, broj djece, broj proteinskih obroka tjedno,....

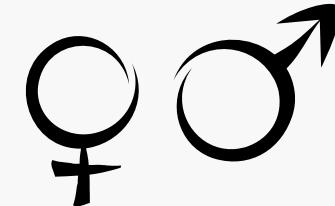
Kvalitativne (opisne) varijable

- **bez redosljeda (*nominalne*)**
 - bračno stanje, spol, krvna grupa, zanimanje, radni status,
- **s poznatim redosljedom (*ordinalne*)**
 - samostalnost: nije ovisan, ovisan u manjem stupnju, ovisan u višem stupnju, ovisan u visokom stupnju, potpuno ovisan
 - sluh: dobar, oštećen, gluhi
 -
- **dvije kategorije (*binarne, dihotomne*)**
 - spol (muški/ženski), pušač (da/ne), nesanica (da/ne),....

LJESTVICE MJERENJA

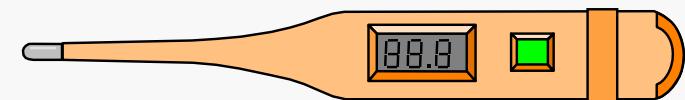
2 1 3

1. nominalna

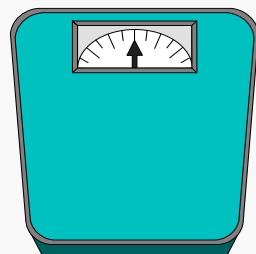


2. ordinalna

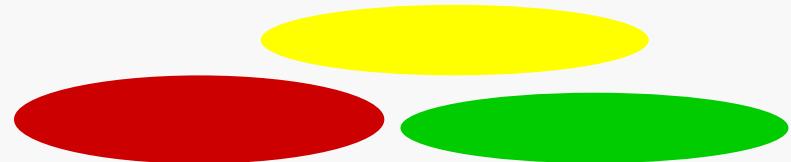
3. intervalna



4. omjerna



1. Nominalna ljestvica



- kategorizacija kojom objektima ili događajima pridružujemo riječi ili simbole
- nema informacije o veličini pojedinačnog rezultata

spol (muški/ženski)

ishod bolesti (preživljavanje/smrt)

boja očiju (plava, crna, smeđa,)

2. Ordinalna ljestvica

2

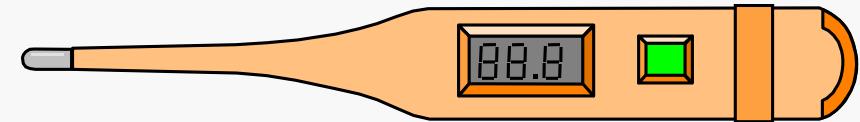
1

3

- ima sve karakteristike nominalne ljestvice i dodatno uključuje **redoslijed** (nominalna + rangiranje)
- intervali nisu jednaki, a granice među grupama nisu čvrste
- nema informacije o “jačini” razlike između pojedinih grupa

**stupanj opekotine (prvi, drugi, treći)
ocjene na ispitu**

3. Intervalna ljestvica



- ljestvica sa jednakim intervalima koji imaju definirane granice => **razlike imaju smisla**
- nema absolutnu nulu (moguće su i negativne vrijednosti) => **omjeri nemaju smisla**

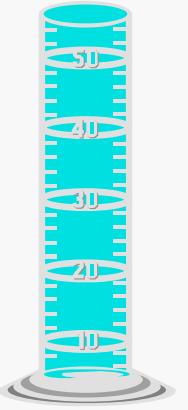
temperatura, kvocijent inteligencije

$$100^{\circ}\text{F} \neq 2 \cdot 50^{\circ}\text{F} \quad \text{jer je} \quad 100^{\circ}\text{F} = 38^{\circ}\text{C}$$

$$50^{\circ}\text{F} = 10^{\circ}\text{C}$$

$$\left(\frac{t^{\circ}\text{C}}{100} = \frac{t^{\circ}\text{F} - 32}{180} \right)$$

4. Omjerna ljestvica



- ima sva svojstva intervalne ljestvice ali i absolutnu nulu => **omjeri imaju smisla**
- nula znači totalnu odsutnost obilježja

visina, težina, dob

$$90\text{kg} = 3 \cdot 30\text{kg}$$

$$3174.653\text{oz} = 3 \cdot 1058.218\text{oz}$$

$$(1\text{kg}=35.27392\text{oz})$$

Još neke vrste podataka...

- **postotci ili proporcije** (*percentages or proportions*)
 - odnos dijela prema cjelini
 - omjer dva istovrsna podatka

ejekcijska frakcija, ...

- **omjeri** (*ratios*)
 - omjer dva raznovrsna podatka (može biti i omjer dviju varijabli)

BMI

...još neke vrste podataka...

- **stope** (rates)
 - odnos opaženog broja pojave nekog svojstva prema jedinici populacije nekog područja u nekom vremenskom periodu

morbiditet, mortalitet,...

- za izračunavanje stope trebamo:
 - **broj** pojave svojstva (npr. broj oboljelih)
 - **skup** u kojem se to obilježje pojavljuje (npr. broj stanovnika nekog područja)
 - **specifikaciju vremena i prostora** (npr. Slavonija, 1992.)

...još neke vrste podataka...

- **skorovi** (scores)
 - rezultat zbrajanja vrijednosti dodijeljenih kategorijama varijabli od interesa

Apgar, GCS, ...

...još neke vrste podataka...

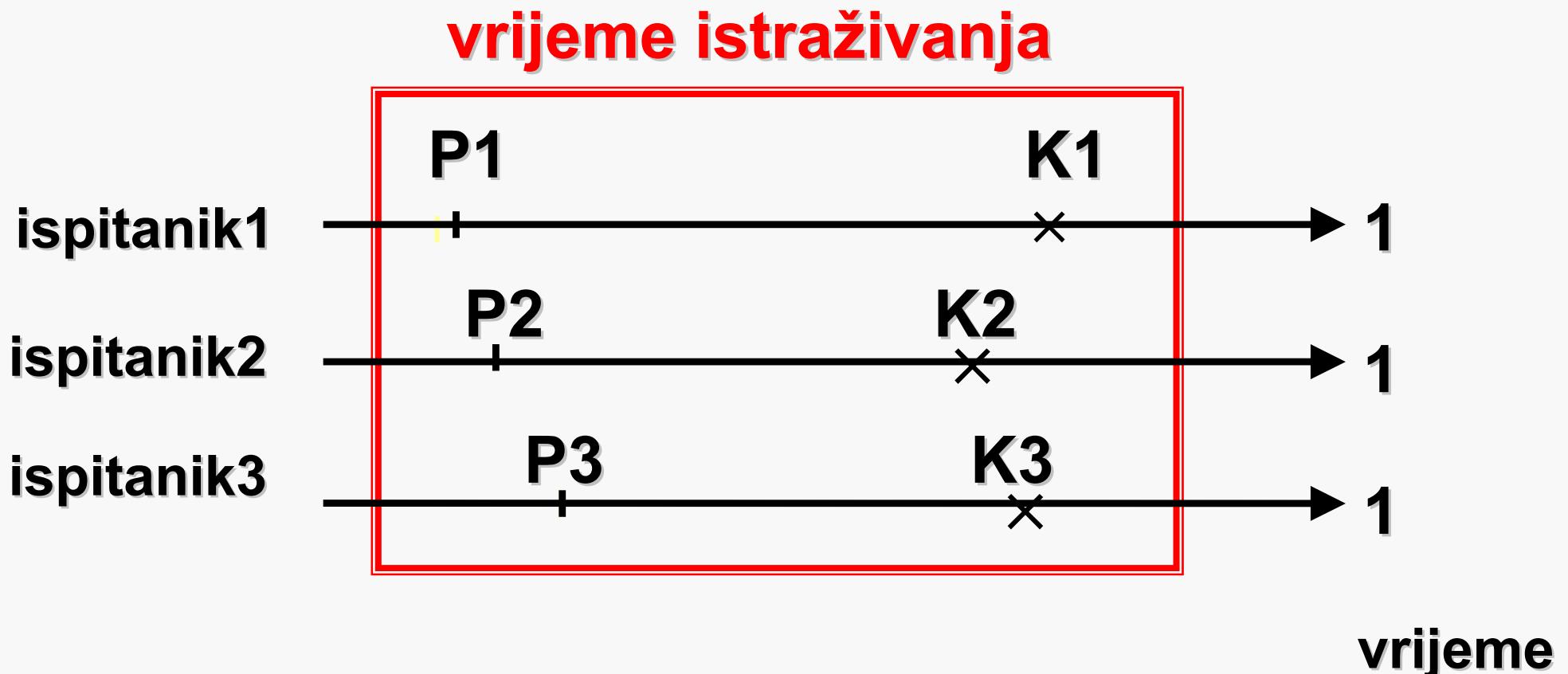
- **cenzorirani** (censored)
 - podatci koji se ne mogu točno izmjeriti, ali je poznato da prelaze neku granicu mjerenja
- najčešće se pojavljuju:
 - u laboratorijskim mjerenjima – vrijednosti ispod/iznad mogućnosti detekcije nekog uređaja
 - u studijama praćenja (follow-up):
 - očekivano svojstvo se nije pojavilo u promatranom periodu
 - ispitanik uključen u istraživanje je iz nekog razloga izuzet prije završetka istraživanja

Cenzuriranje

- događaj se ostvaruje = 1
- sve ostalo = 0 (cenzurirani podatci)

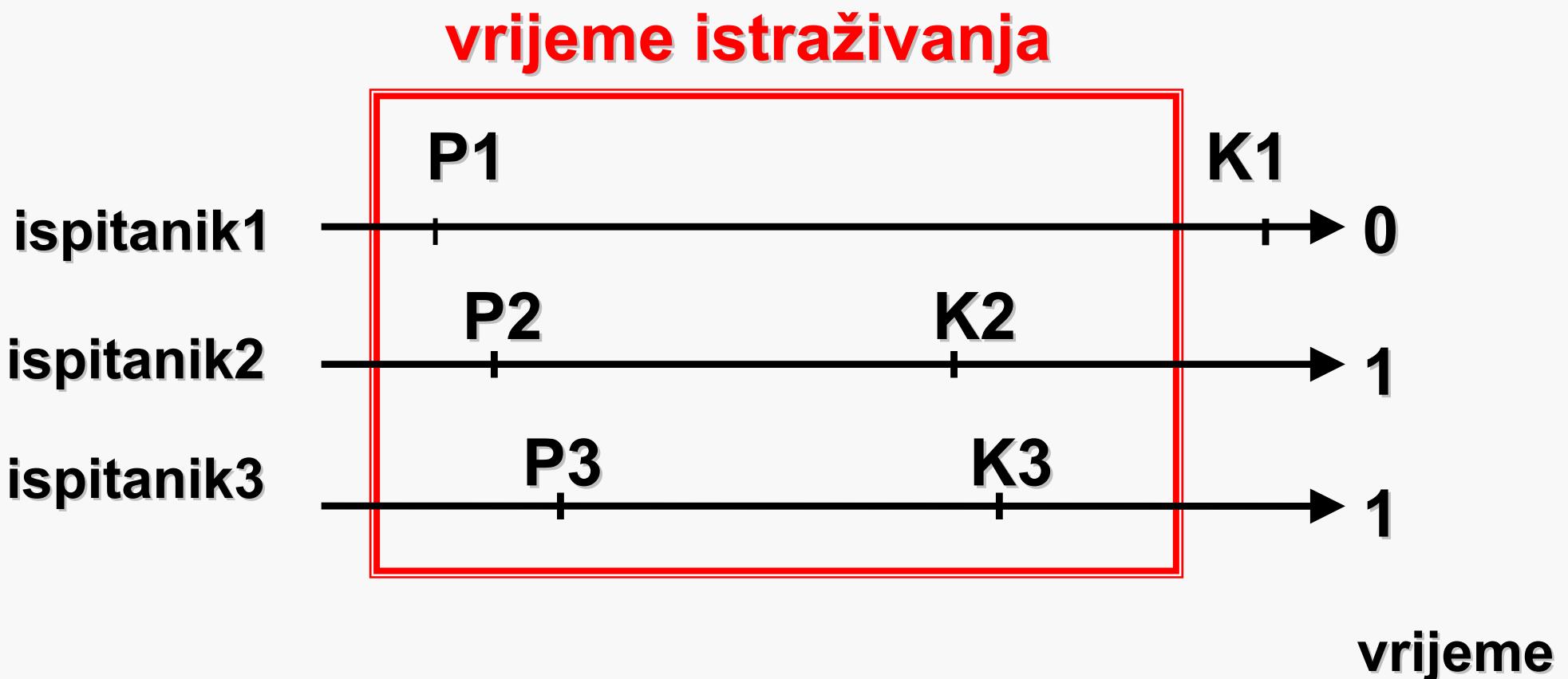
Cenzuriranje

- **potpuni podatci** (potpuno praćenje)
 - događaj se ostvario u promatranom vremenu istraživanja



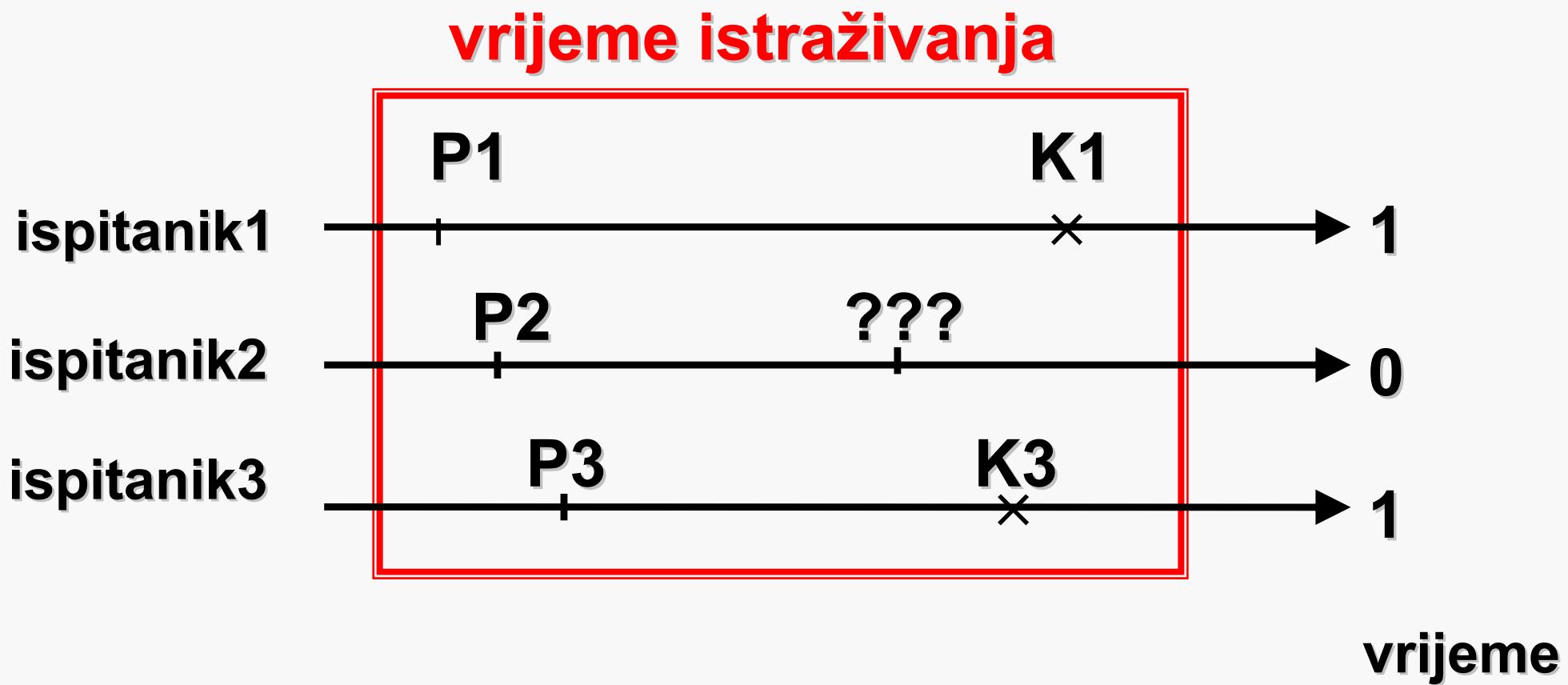
Cenzuriranje

- **cenzurirani podatci** (nepotpuni podatci)
 - za neke ispitanike događaj se ostvario IZVAN okvira promatranog vremena istraživanja



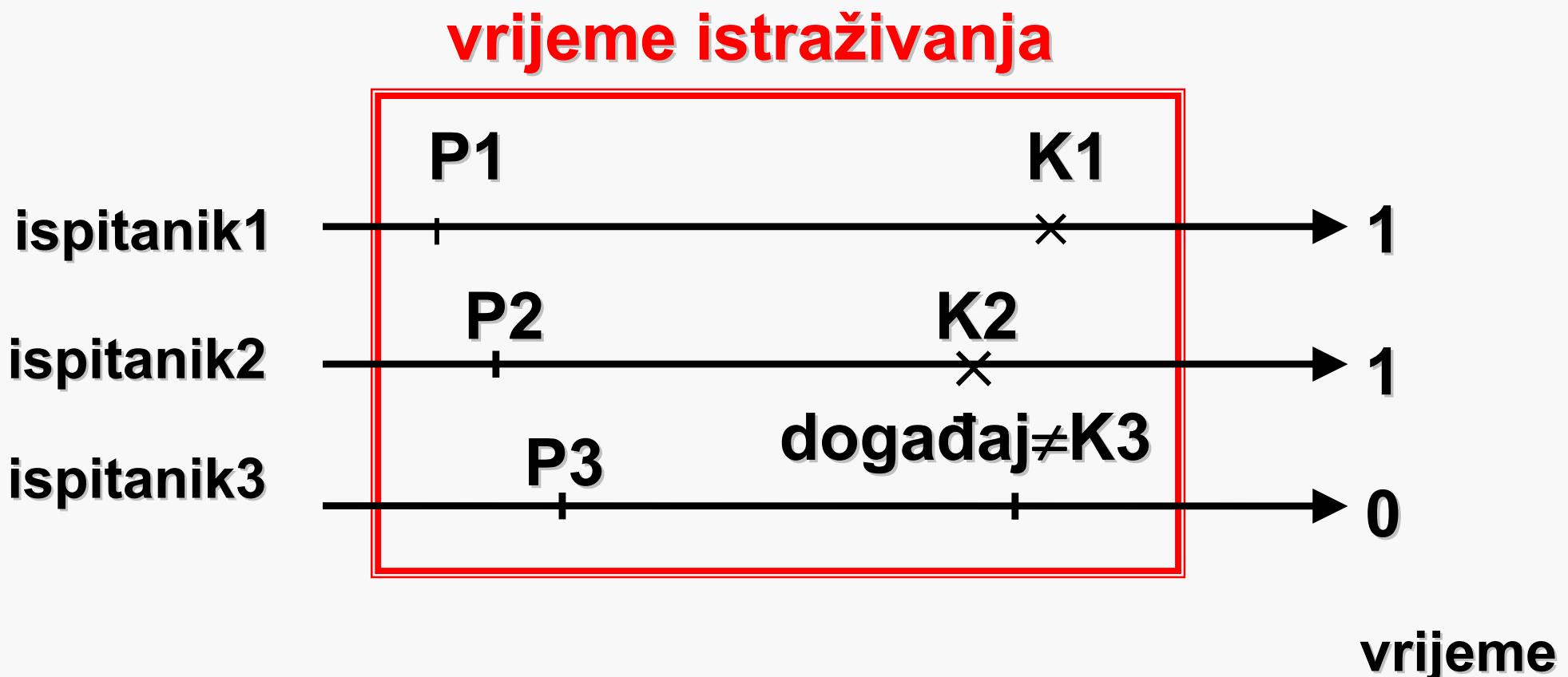
Cenzuriranje

- cenzurirani podatci (nepotpuni podatci)
 - izgubljeni ispitanici ("lost to follow-up")



Cenzuriranje

- **cenzurirani podatci** (nepotpuni podatci)
 - ispitanici koji iz nekih razloga ne mogu više biti praćeni (promjena mjesta boravka, smrt iz razloga nepovezanih s promatranim događajem)



Razlozi VIP tretmana cenzuriranih podataka

- ocjena sredine -

- **aritmetička sredina vremena preživljenja (od postavljanja dijagnoze) oboljelih od određene vrste karcinoma:**
 - srednje vrijeme preživljenja ovisi o tome KADA su podatci analizirani
 - nema smisla dok SVI ispitanici ne umru
- **medijan vremena preživljenja (od postavljanja dijagnoze) oboljelih od određene vrste karcinoma:**
 - može se procijeniti nakon što je POLOVINA ispitanika umrla
- **čekati????**

Razlozi VIP tretmana cenzuiranih podataka

- **čovjek-godine:**
 - svaki ispitanik doprinosi ukupnom vremenu preživljenja svojim vremenom provedenim u studiji (npr. 26 ispitanika promatrani su tijekom studije ukupno 336 mjeseci)
- **čovjek-godine:** suma vremena provedenih u studiji SVIH ispitanika svedena na godine:
$$336/12 = 28 \text{ čovjek-godina}$$
- **problem:** je li 1 čovjek praćen 28 godina ili 28 ljudi 1 godinu ili ???
- usporedba????