

**PRIPREMA: 4. deo**

G. Podaci

- a. kategorički podaci
- b. numerički podaci

H. Merenje

16. oktobar 2018.

## 2. Numerički podaci

- **2. Numerički podaci**
  - takođe nazvani: mere
  - to su: vrednosti numeričkih varijabli za pojedinačne objekte
  - **PRIMERI:**
    - varijabla: *visina*; mera (za određenu osobu): *160 cm*
    - varijabla: *težina*; mera (za određenu osobu): *60 kg*
    - ...
- *merenje*: pripisivanje (dodeljivanje) brojeva osobinama objekata, određenim postupcima
- važan metodološko-statistički problem:
  - nemaju svi tako utvrđeni brojevi jednake osobine
  - tj., kod različitih varijabli brojevi se mogu koristiti na različite načine
- da bi se ove razlike objasnile treba razlikovati:
  - s jedne strane: osobine *brojeva*
  - s druge strane: osobine *objekata* i njihovih *mera*

## 2. Numerički podaci

- osobine *brojeva*: spadaju u matematiku
  - brojevi su apstraktni pojmovi sa određenim karakteristikama
- osobine *objekata* i njihovih *mera*: spadaju u empiriju
  - brojevi su vrednosti numeričkih varijabli, utvrđene u istraživanjima
- *merenje*: povezivanje matematičkog i empirijskog domena
  - objektima *pripisuјемо* brojeve kao mere
- *problem*: ta dva domena mogu biti *u neskladu*
  - različite vrste *empiriski* dobijenih numeričkih podataka u različitoj meri ispunjavaju određene *matematičke uslove*
  - neke *matematičke* operacije ne mogu se na smislen način primeniti na brojeve dobijene *empiriski* kao mere varijabli
- razmotrićemo **pet** tipova numeričkih podataka odn. mera
  - *nominalni*, *ordinalni*, *intervalni*, *racio* i *apsolutni* podaci odn. mere
  - skup mera istog tipa čini *menu skalu*
    - dakle, postoje nominalne skale, ordinalne skale, intervalne skale, racio skale, i absolutne skale

## 2. Numerički podaci

4

### (1) Nominalni podaci (nominalne mere)

- brojevi koji služe kao *imena* odn. oznake kategorija
  - različitim kategorijama dodeljuju se različiti brojevi, kao imena
  - poreklo naziva: *nomen* (latinski): ime
  - PRIMER:*
    - dresova sportista ('1': golman, '2': levi bek, ...; )
    - odeljenja u školi (prvo, drugo,...)
    - grupe subjekata u istraživanju ('1': pušači, '2': nepušači)
- ponekad takve kategorije imaju samo jedan član
  - broj indeksa, broj pasoša, JMBG, telefonski broj, broj auta
- takvi brojevi su ustvari kategorički podaci
  - jedina razlika: kategorije su označene brojevima
  - pogodnost: brojevi su neiscrpni izvor oznaka
- nominalni podaci se veoma razlikuju od pravih brojeva
  - i najosnovnije računske operacije uglavnom su matematički besmislene:
    - poređenje po veličini, vršenje aritmetičkih operacija (+ - \* /, prosek, ...)

## 2. Numerički podaci

5

### (2) Ordinalni podaci (ordinalne mere)

- brojevi koji izražavaju *poređenje* po nekoj osobini
  - objekti mogu po toj osobini biti: prvi, drugi, treći, ...
  - poreklo naziva: *ordo* (latinski): red, poredak
  - PRIMERI:*
    - poredak na prijemnom ispitу (ne broj poena već mesto na listi)
      - važniji za prijem nego broj poena
    - plasiranost u trci (ne vreme trčanja nego redosled)
      - važnija za medalje nego vreme
    - subjektivni stavovi i procene
- terminologija: ključni pojam je *rang*
  - rangiranje:** uspostavljanje *poretka* niza objekata po nekoj osobini
  - rang** datog objekta: njegovo mesto u odnosu na ostale objekte iz niza
  - rang-lista:** rezultat rangiranja, tj. niz objekata uređen prema rangovima
- matematičke osobine ordinalnih podataka
  - bliži** su pravim, matematičkim brojevima nego nominalni podaci
  - inap postoe važne razlike, i dve vrste nedostataka, matematičkih 'mana':

## 2. Numerički podaci

- **(a) neodređenost razmaka:** problem poređenja *intervala*
- za razliku od pravih brojeva, kod ordinalnih podataka ne postoji informacija o veličini *intervala* (razmaka rangova)
- *kod pravih brojeva:*
  - uvek se može utvrditi *interval* odn. *razmak* među njima
  - takođe: intervali se mogu poređati po veličini
    - PRIMER: vreme (a ne plasiranost) u trci; npr. A=10, B=11, C=13, D=15 sek.
    - važi: razmak između A i B je 1
    - takođe: taj razmak manji je nego razmak između C i D (jer je  $1 < 2$ )
- *kod ordinalnih podataka:*
  - razmak među merama nije poznat ili je u principu neodređen
  - ako znamo da je  $A > B$ , ne znamo za koliko je A veće od B
  - takođe: ne možemo da *poredimo* razmak A i B sa razmakom C i D
    - PRIMER: *plasiranost* (ne vremelj) u trci; npr. A prvi, B drugi, C treći, D četvrti
    - ne znamo za koliko je A brži od B, niti da li je *razmak* između A i B manji ili veći ili jednaka razmaku između C i D (ako ne znamo njihova vremena)
    - PRIMER: subjektivni stavovi i procene: razmaci nisu precizno određeni

## 2. Numerički podaci

7

- (b) **neodređenost umnoška:** problem poređenja mera
- za razliku od pravih brojeva, kod ordinalnih podataka *ne postoji* informacija o tome koliko *puta* je jedan objekt veći ili manji od drugog
  - kod pravih brojeva:*
    - za data dva broja A i B uvek se može utvrditi njihov *odnos*, tj. *koliko puta* je A veće ili manje od B (za koji *umnožak*)
      - naime, umnožak tj. odnos je njihov *količnik*  $k = A/B$ , pa je  $A = kB$
      - to znači da za A važi da je  $k$  puta veće (ili manje) od B
      - PRIMER:* 6 je 3 puta veće od 2, jer je  $6/3 = 2$ , odn.  $6 = 2 \times 3$
  - kod ordinalnih podataka (rangova):*
    - umnožak je nepoznat ili neodređen
    - PRIMER:* neka su rangovi 1, 2, 3, itd.
      - za prave brojeve važi:  $2 = 2 \times 1$ ,  $3 = 3 \times 1$ ,  $3 = 1.5 \times 2$ , itd
      - za rangove to ne važi: *ne znači* da je drugi u rangu dva puta veći (ili bolji ili lošiji) od prvog, niti da je treći tri puta veći od prvog, itd
    - PRIMER:* plasiranost (a ne vreme!) u trci
      - ako ne znamo vremena, ne znamo koliko je *puta* razmak između prvog i drugog takmičara manji ili veći od razmaka između trećeg i četvrtog

## 2. Numerički podaci

8

- uočimo razliku nominalnih i ordinalnih mera:
  - nominalne mere izražavaju postojanje *razlike*
  - ordinalne mere takođe izražavaju razlike, ali povrh toga izražavaju *poređenje*, čega nema kod nominalnih mera
- uočimo razliku između dve vrste nedostataka ordinalnih mera:
  - neodređenost razmaka (poređenje intervala):
    - ne znamo za *koliko* je jedan razmak veći od drugog
  - neodređenost umnoška (poređenje mera):
    - ne znamo *koliko puta* je jedan razmak veći od drugog
- posledice ovih nedostataka:**
  - iako uvek ima smisla *porediti* same rangove, vršiti računske operacije sa njima može da bude problematično

## 2. Numerički podaci

9

### (3) Intervalni podaci (intervalne mere)

- brojevi za koje je rešen problem poređenja intervala
  - naime, veličine razmaka susednih mera su *jednake*
  - poreklo naziva: *interval* (latinski): razmak
  - PRIMER:* neke mere u fizici; bisekcija u psihofizici; količnik inteligencije
- međutim: *nije* rešen problem poređenja mera
- računske operacije sa intervalnim podacima:
  - ima smisla porediti mere i poređiti *intervale* mera
  - međutim, utvrđivati *umnoške* mera je matematički problematično
    - tj., koliko puta je jedna mera veća odn. manja od druge
- PRIMER:* merenje temperature u stepenima Celzijusa ( $^{\circ}\text{C}$ ) i Farenhajta ( $^{\circ}\text{F}$ )
  - Celzijusova i Farenhajtova skala su *ekvivalentne* (jednako valjane)
    - u suštini je svejedno koja skala se koristi, to je stvar navike i dogovor
  - postoje *formule* za preračunavanje jednih stepeni u druge
    - za svaku temperaturu izraženu u stepenima Celzijusa može se izračunati odgovarajuća temperatura izražena stepenima Farenhaita, i obrnuto

## 2. Numerički podaci

10

celzijus	$0^{\circ}\text{C}$	$10^{\circ}\text{C}$	$20^{\circ}\text{C}$	$30^{\circ}\text{C}$	...	$X^{\circ}\text{C}$	$(5/9)(Y^{\circ}\text{F}-32)$
farenhajt	$32^{\circ}\text{F}$	$50^{\circ}\text{F}$	$68^{\circ}\text{F}$	$86^{\circ}\text{F}$	...	$(9/5)X^{\circ}\text{C}+32$	$Y^{\circ}\text{F}$

- fizički *iste temperature* su izražene različitim brojevima: npr.  $10 = 50, 20 = 68$
- isti intervali* su izraženi različitim brojevima: npr.  $20-10 = 10, 68-50 = 18$
- ali: *jednakim* intervalima jedne skale odgovaraju *jednaki* intervali druge skale
  - Celzijusi: interval od  $10^{\circ}\text{C}$  do  $20^{\circ}\text{C}$  je isti kao od  $20^{\circ}\text{C}$  do  $30^{\circ}\text{C}$  (jednak je  $10^{\circ}\text{C}$ )
  - Farenhajti: interval od  $50^{\circ}\text{F}$  do  $68^{\circ}\text{F}$  je isti kao od  $68^{\circ}\text{F}$  do  $86^{\circ}\text{F}$  (jednak je  $18^{\circ}\text{F}$ )
- iznenadjujuća osobina intervalnih podataka:** nema smisla tvrditi da je neka temperatura za određen umnožak veća ili manja od neke druge temperature!
  - na pr.: *ne može* se reći da je temperatura od  $20^{\circ}\text{C}$  dva puta veća od  $10^{\circ}\text{C}$ 
    - iako je tačno da je, u domenu brojeva, broj 20 dva puta veći od broja 10
    - naime:  $10^{\circ}\text{C} = 50^{\circ}\text{F}, 20^{\circ}\text{C} = 68^{\circ}\text{F}$ ; važi  $20 = 2 \times 10$ , ali ne važi da je  $68 = 2 \times 50$ !?
- nedopustivo je isto stanje stvari iskazati na dva *protivrečna* načina!
  - ne može ista temperatura i biti i ne biti dva puta veća od druge temperature
- uzrok problema *nije* postojanje dve različite merne skale, već to što *nula* (nulla mera) nije *prirodna* već *arbitrarna* (konvencionalna, dogovarena)
  - npr.  $0^{\circ}\text{C}$  je temperatura mržnjenja vode (a mogla se odabrat i neka druga nula)

## 2. Numerički podaci

11

### (4) Racio podaci (ratio mere)

- brojevi kod kojih je rešen ne samo problem poređenja *intervala*, već i problem poređenja *mera*
  - smisleno* je tvrditi da neki objekt ima  $2x$  veću ili  $3x$  manju meru od drugog
  - poreklo naziva: *racio* (latinski: ratio): odnos, količnik, razum
  - PRIMER:* dužina, težina, Kelvinova temperaturna skala, intenzitet svetlosti, vremenski intervali, ...
- u racio mere spada većina mera u prirodnim naukama
- sa racio podacima dozvoljene sve matematičke operacije
  - na pr. poređenje mera po veličini, merenje veličine razmaka između mera, izračunavanje umnožaka, računanje proseka, ...
- racio skale imaju *prirodnu* a ne *proizvoljnu* nulu
  - npr., nula se odnosi na odsustvo dužine, težine, kretanja molekula itd
- ali: racio skale imaju *arbitrарne odn. konvencijom* (dogovorom) utvrđene merne jedinice
  - santimetar (cm), gram (g), stepen Kelvina ( $^{\circ}\text{K}$ ), itd.

## 2. Numerički podaci

12

- kod racio skala za iste pojave *mogu* postojati različite merne skale i jedinice, ali su takve merne skale *ekvivalentne*
  - npr., brojčano su različite mere u cm ili inčima, u kg ili funtama, itd
    - postoje formule za *preračunavanje* jednih mera u druge
      - na pr.,  $10\text{cm} = 3.937\text{in}, 20\text{cm} = 7.874\text{in}, \dots$
    - ali, za razliku od intervalnih mera, *ima* smisla koristiti umnoške mera!
    - na pr., u cm:  $20 = 2 \times 10$ , u inčima:  $7.874 = 2 \times 3.937$
    - tj. dužina dva puta duža u santimetrima je i dva puta duža u inčima!
  - važi: intervali intervalnih mera spadaju u racio mere!
    - naime, *intervali* imaju *prirodnu* nulu (a to je odsustvo razmaka)
      - ima smisla reći:
        - interval između  $10^{\circ}\text{C}$  i  $20^{\circ}\text{C}$  je  $2x$  manji od intervala između  $10^{\circ}\text{C}$  i  $30^{\circ}\text{C}$
        - interval između  $50^{\circ}\text{F}$  i  $68^{\circ}\text{F}$  je  $2x$  manji od intervala između  $50^{\circ}\text{F}$  i  $86^{\circ}\text{F}$
      - dakle: nema protivrečnosti među iskazima sa dve skale, s obzirom na *temperaturne intervale*
      - ali: kao što smo videli, *ima* protivrečnosti s obzirom na same mere, odn. *temperature*

## 2. Numerički podaci

**(5) Apsolutni podaci (apsolutne mere)**

- brojevi koji su utvrđeni *prebrojavanjem*
  - poreklo naziva: *absolutus* (latinski): bezuslovni, savršeni, nepromenljivi
  - PRIMERI:* broj studenata prve godine; broj zapamćenih reči u eksperimentu
- kod absolutnih mera *ne* postoje ekvivalentne merne skale
  - ne mogu se isti podaci prikazati različitim brojevima
  - doduše: moguće je neke podatke dobijene prebrojavanjem izraziti ne samo absolutno već i *procentualno*, tj. drugim brojevima
    - od 30 reči zapamćeno je: 15 (apsolutno) odn. 50% (procent.)
  - ali: procenti ne nose istu informaciju kao absolutni brojevi
- apsolutni podaci imaju i prirodnu nulu i prirodnu mernu jed.
- nula: nijedan prebrojni element; jedinica: jedan prebrojni element
- računske operacije sa absolutnim podacima:
  - nema ograničenja, sve operacije su dozvoljene

## 2. Numerički podaci

**dodatane klasifikacije numeričkih podataka**

- digitalne** odn. **diskrete** mere
  - mere koje imaju samo određene, često celobrojne vrednosti, i nemaju *međuvrednosti*
  - ovde spadaju ordinalne i absolutne mere
    - PRIMERI:* prvi, drugi, treći, ...; jedna greška, dve greške, ...
- kontinuelne** odn. **analoge** mere
  - u principu *neprekidne*, tj. uvek mogu imati međuvrednosti
  - ovde spadaju intervalne i racio mere
    - PRIMERI:* temperatura, težina, procenat grešaka, ...
- kvantitativne (skalne)** mere: kod statističkih programa
  - programi razlikuju tri vrste podataka: nominalni, ordinalni i skalni
  - u skalne podatke spadaju intervalne i racio, a često i absolutne mere

**problem:** ponekad nije lako utvrditi tip podataka

- PRIMER:* u koju skalu spadaju školske ocene? skale stavova?

## H. Merenje

- postoji šira i uža definicija merenja
- šira definicija** (mi je koristimo):
  - utvrđivanje vrednosti neke varijable
    - PRIMERI:* merenje visine; utvrđivanje rukosti
  - varijabla može biti kategorička ili numerička
  - podaci mogu biti sa bilo koje merne skale
- uža definicija:**
  - utvrđivanje vrednosti kvantitativnih kontinuelnih varijabli
    - PRIMERI:* merenje dužine, mase, vremena, jačine svetlosti, itd.
  - obično se koristi u prirodnim naukama
  - odnosi se na podatke sa intervalnih i racio skala
  - ne uključuje: kategorisanje (nominalne skale), poređenje (ordinalne skale), prebrojavanje (apsolutne skale)

## H. Merenje

- 1. Tehnike merenja**
  - postupci utvrđivanja podataka
- opservacione tehnike**
  - podaci se utvrđuju opservacijom (posmatranjem)
- instrumentalne tehnike**
  - podaci se utvrđuju instrumentima
    - fizički instrumenti
    - fiziološki instrumenti
    - psihološki instrumenti
      - ankete
      - testovi

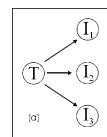
## 2. Problemi merenja

- podela varijabli: *operacionalno* i *konceptualno* definisane varij.
- operacionalno definisane** (odn. empirijske) varijable
  - za njih su utvrđene *operacije* (postupci) merenja
  - iz njihove *definicije* jasno i direktno proizilazi način njihovog *merenja*
  - PRIMERI:* visina, težina, vreme reakcije, starost
- konceptualno definisane** (odn. teoretske) varijable; **konstrukti**
  - definisane (odn. konstruisane) na osnovu *koncepta* (pojma) pojave
  - iz njihove definicije *ne* proističe jasno i direktno način merenja
  - PRIMERI:*
    - inteligencija:* 'sposobnost snalaženja u novim situacijama i shvatanja apstraktnih pojmova'
    - pamćenje:* 'trajna promena u organizmu izazvana iskustvom'
    - anksioznost:* 'spoj straha sa anticipacijom budućeg zla'
- problem: kako *meriti* konceptualno definisane varijable?
  - PRIMERI:* kako ustanoviti nečiju *ovako definisanu* inteligenciju, pamćenje, anksioznost, ...?

## 2. Problemi merenja

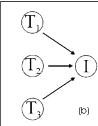
**operacionalizacija**

- postupak kojim se ustanavljava *operacionalno* definisana varijabla kojom se, posredno, meri neki *konstrukt*
- takva varijabla se naziva **indikator(ska varijabla)**
- PRIMER:*
  - test inteligencije meri inteligenciju
    - inteligencija:* konceptualno definisana varijabla; konstrukt
    - test:* operacionalno definisana varijabla; indikator (pokazatelj)
- odnos između indikatora i konstrukata je složen
  - isti konstrukt (teoretska varijabla) može imati više indikatora
  - različiti indikatori mogu da mere različite *aspekte* teoretske varijable, i da je mere različito *uspešno*
- PRIMERI:*
  - različiti testovi inteligencije
  - različiti indikatori pamćenja



## 2. Problemi merenja [19]

- validnost (valjanost)**
  - varijabla je *validna* (valjana): ako meri ono što treba da meri
    - validnost je širok pojam, ima više vrsta validnosti
  - primer: **konstruktualna validnost**
    - valjanost indikatorske varijable da meri konstrukt
- validnost je stvar stepena, može biti veća ili manja
  - na isti indikator može uticati više konstrukata
  - PRIMER:** više teoretskih varijabli može da utiče na učinak na testu inteligencije, a ne samo inteligencija kao teoretska varijabla



## 2. Problemi merenja [20]

opšti problem: kako utvrditi da li je neka varijabla valjan indikator razmotrićemo dve grupe kriterijuma:

- logičko-emprirjski kriterijumi**
  - zasnivaju se na *logici* (definiciji i osobinama konstrukta) i *empiriji* (načinu merenja indikatora)
- statistički kriterijumi**
  - zasnivaju se na korelaciji datog indikatora sa nekim već poznatim, valjanim indikatorom
- PRIMERI:**
  - indikatori pamćenja
    - reprodukacija; rekognicija; vreme reakcije
  - indikatori matematičke i jezičke sposobnosti
    - školske ocene; kvalitetni test
  - astrološki znaci kao indikatori osobina ličnosti
    - logičko-emprirjski kriterijumi
    - statistički kriterijumi

## 2. Problemi merenja [21]

### kriterijumi validnosti *istraživanja*

naime, i na istraživanja se može primeniti pojam validnosti

- eksterna** (spoljna) validnost istraživanja
  - istraživanje je valjano: ako se rezultati mogu uopštiti sa uzorka na populaciju
  - omogućeno korišćenjem statističkih *testova*
- interna** (unutrašnja) validnost istraživanja
  - istraživanje je valjano: ako omogućuje zaključivanje o *kauzalnom* odnosu nezavisne i zavisne varijable
  - omogućeno *kontrolom* varijabli u istraživanju
- pouzdanost**
  - varijabla je pouzdana: ako višestruko merenje daje iste ili slične rezultate
  - ako je varijabla valjana, mora biti i pouzdana
  - ako je varijabla pouzdana, *ne mora* biti i valjana
    - može pouzdano meriti nešto drugo, a ne konstrukt