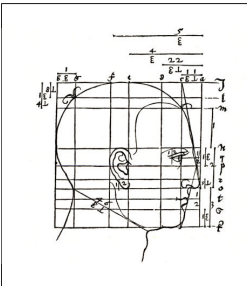


## Metodologija psiholoških istraživanja

### Obrada frekvencijski 1



5. novembar 2018.

#### IV. Obrada podataka

##### A. Frekvencijski nacrti

- Univarijantni frekvencijski nacrti
  - organizacija podataka
  - deskriptivne statističke mere
    - frekvencija
    - proporcija
    - procent
    - šansa

verovatnoća

## IV. Obrada podataka

- sadržaj glave:
- prikaz obrade podataka za tri vrste nacrti:
  - frekvencijski, faktorijalni, korelaciono-regresioni
- u okviru svake od tri vrste nacrti, biće posebno razmotreno nekoliko *tipova* nacrti, zavisno od broja varijabli u nacrtu:
  - frekvencijski nacrti*: 3 tipa: nacrti sa jednom, sa dve, i sa tri varijable
  - faktorijalni nacrti*: 3 tipa: nacrti sa jednom, sa dve, i sa tri nezavisne varijable, i jednom zavisnom varijablom
  - korelaciono-regresioni nacrti*: 2 tipa: nacrti sa dve i sa tri varijable
- u okviru svakog tipa, biće obrađeno pet osnovnih *tema*:
  - organizacija podataka
  - deskriptivne statističke mere
  - prikaz rezultata
  - struktura rezultata
  - statistička značajnost rezultata

• ovih 5 tema biće obrađeno kod svakog od gore nabrojanih 8 tipova nacrti

## A. Frekvencijski nacrti (FN)

- FN: nacrti koji sadrže samo *kategoričke* varijable
- broj *varijabli* u frekvencijskim nacrtima (*tip* nacrti):
  - univarijantni FN (UFN)**: FN sa jednom varijablom
  - bivarijantni FN (BFN)**: FN sa dve varijable
  - trivarijantni FN (TFN)**: FN sa tri varijable
- broj *kategorija* varijabli FN:
  - varijabla sa dve kategorije: **dihotomija**
  - varijabla sa tri kategorije: **trihotomija**
  - varijabla sa više od tri kategorije: **poli(ho)tomija**
- osnovna obrada podataka u FN
  - utvrđivanje **frekvenc(ij)e** (broja članova) kategorija varijabli
    - takođe se nazva: učestalost, učestanost
- u daljoj obradi koriste se i dodatne mere

## A. Frekvencijski nacrti (FN)

- Univarijantni frekvencijski nacrti (UFN)
  - frekvencijski nacrti sa samo jednom varijablom
  - PRIMERI**:
    - dihotomije: utvrđivanje broja članova dve kategorije
      - odgovori 'da' i 'ne' u anketi
      - za koga će se glasati (od dva kandidata)
      - pušači, nepušači
      - levoruke i desnoruke osobe
      - novi primer: preferencije beba za boje
        - da li bebe više vole crvenu ili plavu boju?
    - trihotomije: utvrđivanje broja članova tri kategorije
      - imenice muškog, ženskog i srednjeg roda
    - politomije: utvrđivanje broja članova za više kategorija
      - 'koju boju (od navedenih) najviše volite?'; 'koje novine najradije čitate?'; 'za koju partiju ćete glasati?' ...

## 1. Univarijantni frekvencijski nacrti (UFN)

### a. Organizacija podataka

- matrica podataka**: tabela u koju se unose podaci
  - nije isto što i matrica *rezultata* (ona se dobija *obradom* podataka)
- postoje različiti *formati* (načini prikaza) matrica podataka
- format matrice u frekvencijskim nacrtima: *objekti x varijable*
- primeri matrica podataka u UFN:
 

konkretni prikaz	apstraktni prikaz
BOJA	A
1. crvena	O1 a1
2. plava	O2 a2
3. crvena	O3 a1
...	...

redovi: objekti istraživanja  
kolone: varijable (u UFN: samo jedna)  
ćelije: kategorije (a1, a2, a3, ...), često numerički kodirane (1, 2, 3, ...)  
margine: oznake objekata (npr. O1, O2, ...) i varijabli (npr. A)

## 1. Univarijantni frekvencijski nacrti (UFN)

### b. Deskriptivne statističke mere

u frekvencijskim nacrtima koriste se **mere prebrojavanja**:

- apsolutna** mera prebrojavanja: **frekvencija**
- relativne** mere prebrojavanja: **proporcija, procent, šansa**

### 1. frekvenc(ij)a (učestalost): broj članova neke kategorije objekata

- oznake: varijabla: A
 

	a1	a2	T
A	f1	f2	N

  - kategorije varijable A: a1, a2, ...
  - broj članova kategorija a1, a2 ...: f1, f2, ...
  - ukupni broj članova uzorka: N
- PRIMERI**:
 

	mušk.	žene	TOTAL
POL	25	75	100

  - utvrđivanje frekvence muškaraca (a1) i žena (a2)
    - f1 = 25 muškaraca, f2 = 75 žena, N = 25 + 75 = 100
  - utvrđivanje frekvence rukosti u jednom istraživanju (stvarni podaci)
    - f1 = 11196 desnorukih, f2 = 888 levorukih, N = 11196 + 888 = 12084
- za dihotomije (A ima dve kategorije, a1 i a2) važi: **f1 + f2 = N**
- za trihotomije (A ima tri kategorije, a1, a2, a3) važi: **f1 + f2 + f3 = N, ...**

## 1. Univarijantni frekvencijski nacrti (UFN) 7

- **relativne mere** prebrojavanja: **proporcija, procent, šansa**
  - odražavaju *relacije* (odnose) apsolutnih mera
  - matematički se definišu *količnicima* apsolutnih mera
    - izražavaju se *razlomcima* ili *decimalnim brojevima* (češće)
  - **proporcija i procent**: prikazuju odnos pojedinih kategorija i celog uzorka
    - odnos dela i celine, *udeo* kategorije u uzorku
  - **šansa**: prikazuje međusobni odnos pojedinih kategorija (odnos delova)

### 2. proporcija (p): udeo kategorije u uzorku

- oznake: f: frekvencija kategorije; N: totalna frekvencija; proporcija:  $p = \frac{f}{N}$

#### PRIMERI:

- proporcije polova: udeo muškaraca i žena u celom uzorku
 

	mušk.	žene	TOT.
POL	0.25	0.75	1
- proporcija muškaraca:  $p_1 = f_1/N = 25/100 = 1/4 = 0.25$   
proporcija žena:  $p_2 = f_2/N = 75/100 = 3/4 = 0.75$
- proporcije rukosti: udeo desnorukih i levorukih osoba u celom uzorku
  - **frekvence**: desnoruki:  $f_1 = 11196$ , levoruki:  $f_2 = 888$ , uzorak:  $N = f_1 + f_2 = 12084$
  - **proporcije**: desnor.:  $p_1 = 11196/12084 = 0.926$ ; levor.:  $p_2 = 888/12084 = 0.074$

## 1. Univarijantni frekvencijski nacrti (UFN) 8

- za dihotomije važi:  $p_1 + p_2 = 1$ 
  - to sledi iz jednačine  $f_1 + f_2 = N$ , kada se obe strane podele sa N
  - naime  $f_1/N + f_2/N = p_1 + p_2 = N/N = 1$
- za trihotomije važi:  $p_1 + p_2 + p_3 = 1$ 
  - to sledi iz jednačine  $f_1 + f_2 + f_3 = N$ , kada se obe strane podele sa N
- **koliko** mogu biti frekvence i proporcije?
- **frekvence** mogu biti 0, 1, 2, 3 itd., tj. bilo koji celi pozitivni broj ili 0
- **proporcije** moraju biti između 0 i 1, jer je  $p=f/N$ , a f ne može biti veće od N

### 3. procent (P): broj sto puta veći od proporcije

#### PRIMERI:

- procenti muškaraca (P1) i žena (P2)
  - muškarci:  $P_1 = 100 \cdot p_1 = 100 \cdot 0.25 = 25\%$ ; žene:  $P_2 = 100 \cdot p_2 = 100 \cdot 0.75 = 75\%$
- procenti levorukih (P1) i desnorukih (P2) osoba
  - desnoruki:  $P_1 = 0.926 \cdot 100 = 92.6\%$ ; levoruki:  $P_2 = 0.074 \cdot 100 = 7.4\%$
- za dihotomije važi  $P_1 + P_2 = 100$ 
  - to sledi iz jednačine  $p_1 + p_2 = 1$ , kada se obe strane pomnože sa 100
- za trihotomije važi  $P_1 + P_2 + P_3 = 100$ , itd

## 1. Univarijantni frekvencijski nacrti (UFN) 9

- **uočiti**: procenti i proporcije nose *identičnu* informaciju
  - razlika je samo u *opsegu*: p je broj između 0 i 1, P je broj između 0 i 100
  - proporcije su matematički jednostavnije, jer nema množenja sa 100
  - ali, u svakodnevnom životu koriste se skoro uvek samo procenti
- jedan način razumevanja značenja procenta: kao frekvence!
  - uočiti: ako u uzorku ima  $N=100$  članova, procent je isto što i frekvencija
  - npr., ako je procent levorukih ljudi  $P = 7.4\%$ , to znači da na svakih 100 osoba možemo očekivati 7.4 tj. 7 do 8 levorukih (a 92-93 desnorukih)
- jedna razlika frekvenci i procenata
  - neka je uzorak  $N=1000$  i neka se smanji na 500, tj. za 50%
  - ako se uzorak zatim poveća za 500, onda je opet  $N=1000$
  - ali, ako se uzorak poveća za 50%, onda je  $N=750$  ( $500+0.5 \cdot 500$ )
- ponekad se procenti koriste za veličine koje se ne odnose na **frekvence**, već za izražavanje **umnožaka**
  - uočiti: u takvoj primeni procenti mogu biti i *veći* brojevi od 100 (ne u ovom kursu!)
  - npr. 'proizvod je poskupeo za 300% ili 200% ili 100%': tj. 3x odn. 2x odn. 1x?

## 1. Univarijantni frekvencijski nacrti (UFN) 10

- **prednosti** relativnih mera (procenta i proporc.) nad frekvencom
  - **informativnost**: podatak o *udelu* nosi drugačiju informaciju
    - ta informacija je često korisnija nego podatak o apsolutnom broju
    - '7.4% levorukih' je informativnije nego '888 levorukih'
  - **razumljivost**: podatak o udelu se lakše razume, jasniji je
    - lakše je poimati odnos neke brojke prema 100 (ili prema 1, kod proporcija) nego prema nekom 'neokruglom' broju
    - '7.4% levorukih' je jasnije nego 'u 12056 osoba ima 888 levorukih'
  - **uporedivost**: mogu se uporediti rezultati različitih istraživanja
    - razlog: različite brojke se svode na iste *srazmere*
    - drugo istraživanje rukosti: 1933 desnorukih, 150 levorukih
      - frekvence su znatno manje, ali da li su rezultati zaista različiti?
      - u procentima: ovde je nađeno je 7.2% levorukih, a u prvom istr. 7.4%
- **mana** relativnih mera u odnosu na frekvencu
  - **informativnost**: nema informacije o apsolutnim brojkama
  - **koliko je 60%**: može biti 60 od 100, ali i 3 od 5, ili 1200 od 2000, itd

## 1. Univarijantni frekvencijski nacrti (UFN) 11

### 4. šansa: $f_1 / f_2$ ili $f_1 : f_2$ (čita se: $f_1$ sa $f_2$ , ili $f_1$ prema $f_2$ )

- iskazana je količnikom frekvenci pojedinih kategorija
- drugačije izražena informacija nego preko procenata i proporcije
- **PRIMERI**: frekvence muškaraca ( $f_1=25$ ) i žena ( $f_2=75$ )
  - šansa da je osoba u uzorku muškarac je  $f_1:f_2 = 25:75 = 1:3 = 1/3$ 
    - odn.: na svaku mušku osobu u uzorku dolaze tri ženske osobe
  - šansa da je osoba u uzorku žena je  $f_2:f_1 = 3:1$ 
    - odn.: u uzorku ima tri puta više žena nego muškaraca
  - uporediti: **proporcija** muškaraca je  $25/100 = 1/4$ , a žena  $75/100 = 3/4$ 
    - odn.: na svake četiri osobe dolaze jedna muška i tri ženske osobe
- frekvence desnorukih ( $f_1=11196$ ) i levorukih ( $f_2=888$ )
  - šansa da je osoba desnoruka je  $f_1 : f_2 = 11196 : 888 = 12.6 : 1$ 
    - u uzorku ima 12.6 puta više desnorukih nego levorukih osoba
    - na svaku levoruku osobu u proseku dolazi 12.6 desnorukih
    - obrnuto: na svaku desnoruku osobu dolazi  $1/12.6 = 0.079$  levorukih
    - uporediti procenata: levorukih ima 7.4%, desnorukih ima 92.6%

## 1. Univarijantni frekvencijski nacrti (UFN) 12

### verovatnoća: veoma važan pojam, srodan proporciji

- verov. izražava koliko ima *izgleda* (izvesnosti) da se neki događaj desi
- u matematici, verovatnoća se izražava brojem između 0 i 1
- oznaka za verovatnoću: p (lat.: 'probabilitas'), isto kao za proporciju
  - događaj koji se *sigurno (izvesno) dešava*: ima verovatnoću  $p=1$
  - događaj koji se sigurno *ne dešava*: ima verovatnoću  $p=0$
  - ostali događaji imaju verovatnoću između 0 i 1, tj.  $0 < p < 1$
  - npr. događaj koji ima istu verovat. da se desi i da se ne desi:  $p=0.5$
- teorija verovatnoće je grana matematike
  - teorija verovatnoće je teoretska *osnova* matematičke statistike
  - analiza rezultata istraživanja je *primena* matematičke statistike
- teorija verovatnoće se bavi tzv. **slučajnim** pojavama
  - one se definišu kao procesi, zbivanja, događaji:
    - koji se mogu više puta *ponavljati*, ali
    - čiji ishod *nije* unapred određen, već se menja od slučaja do slučaja
    - tj., koji nisu *deterministički* nego *stohastički*

## 1. Univarijantni frekvencijski nacrti (UFN) 13

- **PRIMERI:** postoje tri standardna, osnovna primera slučajnih pojava
- 1. **bacanje novčića (BN):** postoje dva događaja: 'glava' (G), i 'pismo' (P)
  - skup svih mogućih događaja je skup  $S = \{G, P\}$
  - koje su **verovatnoće** da se prilikom BN *desi* jedan ova dva događaja?
    - uočiti: ima ih dva, jednako su verovatni, i jedan od njih se mora desiti
    - verovatnoće događaja G i P su:  $p(G) = 1/2 = 0.5$ ,  $p(P) = 1/2 = 0.5$
- 2. **bacanje kocke (BK):** postoji 6 događaja: da padne 1, 2, 3, 4, 5 ili 6
  - skup svih mogućih događaja je:  $S = \{1, 2, 3, 4, 5, 6\}$
  - verovatnoće:  $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6 = 0.166... \approx 0.17$
- 3. **izvlačenje karte iz špila (IK):** postoji 52 događaja: kralj karo, dama pik, ...
  - verovatnoća svakog takvog događaja je:  $p = 1/52 \approx 0.019$
  - uočiti: *svaki* događaj je *malo* verovatan, ali se *jedan* od njih *mora* desiti ( $p=1$ )
- u opštem slučaju važi 'pravilo jednakih verovatnoća':
  - za bilo koji skup S koji ima N *jednako* verovatnih događaja važi:  $p = 1/N$
  - važi: zbir verovatnoća svih N događaja skupa S je  $N \cdot p = N \cdot (1/N) = 1$ 
    - u gornjim primerima: BN:  $N=2$ ,  $p=1/2$ ; BK:  $N=6$ ,  $p=1/6$ ; IK:  $N=52$ ,  $p=1/52$
  - uočimo: svi primeri (BN, BK, IK) odnose se na *isključive* i *iscrpne* klasifikacije

## 1. Univarijantni frekvencijski nacrti (UFN) 14

- **prosti i složeni događaji**
- događaji opisani u dosadašnjim primerima slučaj. fenomena (BN, BK, IK) nazivaju se **prosti (elementarni) događaji (PD)**
- osim PD postoje i **složeni događaji (SD)**
  - SD su *skupovi* PD, definisani određenim pravilima odn. *uslovima*
- razmotrićemo prvo *primere* različitih SD i njihove *uslove*, a kasnije i njihove *verovatnoće*
- **PRIMERI:** neki SD koji se javljaju prilikom bacanja kocke (BK)
  - uslov: 'da padne paran broj'
    - SD:  $S = \{2, 4, 6\}$ ; broj članova skupa S je  $n=3$  od ukupno 6
  - uslov: 'da padne neparan broj'
    - SD:  $S = \{1, 3, 5\}$ ;  $n=3$
  - uslov: 'da padne šestica'
    - SD:  $S = \{6\}$ ;  $n=1$  (uočiti: skup S može da ima i samo jedan član)
- uslov: 'da ne padne broj manji od 6'
  - SD:  $S = \{6\}$ ;  $n=1$  (uočiti: isti skup S se može definisati na razne načine)

## 1. Univarijantni frekvencijski nacrti (UFN) 15

- uslov: 'da padne broj manji od 7'
  - SD:  $S = \{1, 2, 3, 4, 5, 6\}$ ;  $n=6$
  - S je skup svih mogućih događaja, tj. 'puni skup' (iscrpna kategorizacija)
- uslov: da padne broj veći od 6
  - SD:  $S = \{7\}$ ; broj članova skupa S je  $n=0$
  - uočiti: skup S nema *nijedan* član, tj. S je 'prazan skup' (nije isto što i nulat)
- **PRIMERI:** neki SD koji se javljaju prilikom izvlačenja karte iz špila (IK)
  - uslov: 'da bude izvučena dama';
    - SD:  $S = \{\text{dama pik, dama herc, dama tref, dama karo}\}$ ;  $n=4$
  - uslov: 'da bude izvučena crvena karta'
    - SD: skup svih crvenih karata;  $n=26$
- terminologija:
  - za PD koji ispunjava neki postavljeni *uslov* za SD kaže se da *pripada* dotičnom SD, odn. da je *povoljni* PD za dotični SD
  - za SD se kaže da se *desio*, ako se desio bilo koji za njega povolj. PD
- **PRIMER:** SD je dat uslovom 'da padne paran broj', tj.  $S = \{2, 4, 6\}$ 
  - njemu *pripadajući* odn. *povoljni* PDi su: kad padne ili 2 ili 4 ili 6
  - događaj 'paran broj' se *desio*: ako se *desilo* da padne ili 2 ili 4 ili 6

## 1. Univarijantni frekvencijski nacrti (UFN) 16

- **veza verovatnoće (p), proporcije (označene: pr), i frekvence (f)**
- pretpostavimo da više puta bacamo novčić, na pr. 10 puta
  - niz događaja može biti: G, G, P, G, P, P, G, P, P, P
  - frekvence padanja glave i pisma su:  $f(G) = 4$ ,  $f(P) = 6$
  - proporcije padanja glave i pisma su:  $pr(G) = 4/10 = 0.40$ ;  $pr(P) = 6/10 = 0.60$
- pretpostavimo da 100 puta bacamo novčić
  - ako je  $f(G) = 45$ ,  $f(P) = 55$ , onda je  $pr(G) = 0.45$ ;  $pr(P) = 0.55$
  - uočiti: iako je razlika *frekvenci* ( $55-45=10$ ) veća nego ranije kada je novčić bacan 10 puta ( $6-4=2$ ), razlika *proporcija* je manja (0.11) nego ranije (0.20)
- pretpostavimo da **sve duže i duže** bacamo novčić
  - tada bi frekvence i njihove razlike bile sve *veće*, ali bi *razlika* proporcija  $pr(G)$  i  $pr(P)$  bivala sve *manja*, i one bi postajale sve bliže vrednosti 0.5 odn. 1/2
  - ovakav proces se može simulirati na kompjuteru
- pretpostavimo da **beskonačno** dugo bacamo novčić
  - tada bi  $pr(G)$  i  $pr(P)$  bi bile tačno po 0.5
  - te vrednosti su *verovatnoće* događaja G i P
    - $p(G) = 0.5 = 1/2$ ;  $p(P) = 0.5 = 1/2$
- **dakle:** verovatnoće se mogu shvatiti kao *dugoročne proporcije*

## 1. Univarijantni frekvencijski nacrti (UFN) 17

- **PRIMER:** izbor objekta iz skupa objekata
- pretpostavimo:
  - da imamo skup od  $f(M)=25$  muškaraca i  $f(\bar{Z})=75$  žena
  - da se iz tog skupa slučajno odabere jedna osoba i zabeleži njen pol
  - da se taj postupak ponavlja mnogo puta
    - pritom ista osoba može biti i više puta izabrana
- na pr., ako se postupak vrši 10 puta, izbori bi mogli biti:
  - $\bar{Z}, \bar{Z}, M, \bar{Z}, \bar{Z}, \bar{Z}, M, \bar{Z}, \bar{Z}, M$  (sedam žena, tri muškarca)
  - proporcije izbora su  $pr(M)=0.3$ ,  $pr(\bar{Z})=0.7$
- ako se postupak vrši 100 puta, proporcije bi mogle biti:
  - $pr(M)=0.22$ ,  $pr(\bar{Z})=0.78$
- ako se postupak vrši sve *duže*, proporcije bi bile sve *bliže* vrednostima
  - $pr(M)=0.25$ ,  $pr(\bar{Z})=0.75$
- kaže se (budući da je verovatnoća dugoročna proporcija):
  - *verovatnoća* da je osoba iz tog skupa muškarac je  $p(M) = 0.25$
  - *verovatnoća* da je osoba iz tog skupa žena je  $p(\bar{Z}) = 0.75$

## 1. Univarijantni frekvencijski nacrti (UFN) 18

- SD se može definisati ne samo kao skup PD, već i pomoću drugih SD, na tri osnovna načina: *unija, presek, komplement*
- **unija** dva složena događaja A i B; oznaka:  $Un(A, B)$
- to je događaj C, kome pripadaju svi PD koji su povoljni *bilo* za A *bilo* za B
- kažemo da se događaj C *desio* ako se desio *ili* događaj A *ili* događaj B
- **PRIMERI:** bacanje kocke (BK)
  - $A = \{1, 2\}$ ,  $B = \{3, 4\}$ ; šta je događaj  $C = Un(A, B)$ ?
    - $Un(A, B) = C = \{1, 2, 3, 4\}$
    - C se *desio*: ako je se desilo ili A ili B, tj. ako je pala ili 1 ili 2 (A) ili 3 ili 4 (B)
  - $A = \{1, 3, 5\}$ ,  $B = \{2, 4, 6\}$ ; šta je  $Un(A, B)$ ?
    - $Un(A, B) = C = \{1, 2, 3, 4, 5, 6\}$
    - uočiti: skup neparnih i parnih brojeva čine *pun* skup
  - $A = \{1, 2\}$ ,  $B = \{2, 3\}$ ; šta je  $Un(A, B)$ ?
    - $Un(A, B) = C = \{1, 2, 3\}$ ; tj. događaj C se desio ako je palo ili 1 ili 2 ili 3
    - uočiti: A i B imaju *zajednički* član, 2, koji se *ne broji dva puta* u uniji
  - $A = \{1, 3, 5\}$ ,  $B = \{3, 5\}$ ; šta je  $Un(A, B)$ ?
    - $Un(A, B) = C = \{1, 3, 5\}$
    - uočiti: skup A *sadrži* skup B, na je njihova unija C jednaka skupu A

## 1. Univarijantni frekvencijski nacrti (UFN) 19

- **presek** dva složena događaja A i B; oznaka:  $\Pr(A, B)$
- to je događaj C, kome pripadaju svi PD koji su povoljni i za A i za B
- kaže se da se događaj C *desio*: ako se desio i događaj A i događaj B
- **PRIMERI**:  $A = \{2, 4, 6\}$  (tj. broj deljiv sa 2),  $B = \{3, 6\}$  (tj. broj deljiv sa 3)
  - $\Pr(A, B) = C = \{6\}$  (tj. broj deljiv i sa 2 i sa 3)
  - uočimo: ako je palo 6, desio se i događaj A i događaj B
  - $A = \{1, 3\}$ ,  $B = \{2, 4\}$ ; šta je  $\Pr(A, B)$ ?
  - $\Pr(A, B) = C = \{\}$  tj. prazan skup
- događaji A i B čiji je presek **prazan** skup se nazivaju međusobno **isključivi** događaji, naime, ako se A desio, B se nije desio, i obrnuto; takvi su svi PDI
- **komplement** događaja A; oznaka:  $Ko(A)$ 
  - to je događaj C, kome pripadaju svi *moćni* PD koji za A nisu povoljni
  - C se desio ako se A nije desio; A i  $Ko(A)$  su međusobno **isključivi**
- **PRIMERI**: ako je  $A = \{1, 3, 5\}$  (tj. neparni), onda je  $Ko(A) = \{2, 4, 6\}$  (tj. parni)
  - za  $S = \{1, 2, 3, 4, 5, 6\}$  (tj. pun skup),  $Ko(S) = \{\}$  (tj. prazan skup)
  - važi:  $\Pr(A, Ko(A)) = \{\}$  (npr.: nijedan broj nije i paran i neparan; isključiv.)
  - $Un(A, Ko(A)) = S$  (npr.: svaki broj je ili paran ili neparan);  $Ko(Ko(A)) = A$

## 1. Univarijantni frekvencijski nacrti (UFN) 20

- za složeni događaj SD, kako se izračunava **verovatnoća**  $p(SD)$ ?
- imali smo 'pravilo jednako verovatnih događaja':  $p = 1/N$
- postoji još nekoliko načina izračunavanja verovatnoća SD:
  - (a) kao **zbir** verovatnoća za SD povoljnih PD ('pravilo zbira')
  - (b) kao **količnik** broja povoljnih PD i svih mogućih PD, tj. kao proporcija svih povoljnih događaja ('pravilo količnika')
    - ali to važi **samo** pod uslovom da su svi prosti događaji **jednako verovatni**
  - (c) pomoću kasnije navedenih pravila koje se odnose verovatnoće unije, preseka i komplementa
- **PRIMERI**: za događaj 'pao je neparan broj', tj.  $A = \{1, 3, 5\}$ , koliko je  $p(A)$ ?
  - intuicija kaže da je  $p(A) = 0.5$ ; šta kažu pravila?
  - **pravilo zbira**: kako je  $p(1)=p(3)=p(5)=1/6$ , važi  $p(1)+p(3)+p(5)=3/6=1/2$
  - **pravilo količnika**: kako od ukupno 6 jednako verovatnih prostih događaja ima 3 povoljna, **proporcija** povoljnih događaja data je količnikom  $3/6 = 1/2$
- za događaj  $A = \{3, 6\}$ , tj. 'pao je broj deljiv sa 3',  $p(A) = 2/6 = 1/3 = 0.333...$
- za događaj 'pao je broj manji od 7' (pun skup, tj. nužan događaj):  $p = 1$
- za događaj 'pao je broj veći od 6' (prazan skup, nemogući događaj):  $p = 0$

## 1. Univarijantni frekvencijski nacrti (UFN) 21

- pravila računanja verovatnoće složenih događaja (SD) koji su definisani kao unija, presek ili komplement drugih SD
  - korišćenjem ovakvih pravila ponekad se verovatnoća može izračunati jednostavnije nego pomoću drugih načina
- (a) **pravilo komplementa**:  $p(Ko(A)) = 1 - p(A)$ 
  - naime, važi:  $p(A) + p(Ko(A)) = 1$
- **PRIMER**: za  $A = \{3, 6\}$ , koliko je  $p(Ko(A))$ ?
  - prema pravilu zbira ili količ.:  $Ko(A) = \{1, 2, 4, 5\}$ ,  $p(Ko(A)) = 4/6 = 2/3 = 0.666...$
  - prema pravilu komplementa: kako je  $p(A) = 1/3$ , biće  $p(Ko(A)) = 1 - 1/3 = 2/3$ 
    - uočiti: ne mora se utvrđivati šta je skup  $Ko(A)$ , već samo šta je  $p(A)$
- (b) **pravilo unije**:  $p(Un(A, B)) = p(A) + p(B)$ 
  - verovatnoća unije je **zbir** verovatnoća članova unije
  - pravilo važi **samo** ako su A i B međusobno **isključivi** (presek je prazan)
- **PRIMER**: za  $A = \{1, 3\}$ ,  $B = \{2, 4\}$ , koliko je  $p(Un(A, B))$ ?
  - prema pravilu zbira ili količnika:  $Un(A, B) = \{1, 2, 3, 4\}$ ,  $p(Un(A, B)) = 4/6 = 2/3$
  - prema pravilu unije:  $p(Un(A, B)) = p(A) + p(B) = 1/3 + 1/3 = 2/3$ 
    - uočiti: ne mora se utvrđivati šta je  $Un(A, B)$ , već samo šta su  $p(A)$  i  $p(B)$

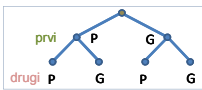
## 1. Univarijantni frekvencijski nacrti (UFN) 22

- **PRIMER**: za  $A = \{1, 2, 3, 4\}$ ,  $B = \{3, 4, 5, 6\}$ , koliko je  $p(Un(A, B))$ ?
  - važi:  $p(A) = 4/6 = 2/3$ ,  $p(B) = 4/6 = 2/3$
  - ali NE važi:  $p(Un(A, B)) = 2/3 + 2/3 = 4/3$ , što je  $>1$ , a p nikada ne može biti  $>1!$
  - pravilo unije se ne može primeniti, jer A i B nisu **isključivi**
    - njihov presek nije prazan skup:  $\Pr(A, B) = \{3, 4\}$
  - prema pravilu zbira ili količnika:  $Un(A, B) = \{1, 2, 3, 4, 5, 6\}$ , pa je  $p(Un(A, B)) = 1$
- (c) **pravilo preseka**:  $p(\Pr(A, B)) = p(A) * p(B)$ 
  - verovatnoća preseka je **proizvod** verovatnoća članova preseka
    - verovatnoća preseka se zove **zajednička verovatnoća** dva događaja
  - pravilo važi **samo** ako su A i B **nezavisni** događaji
  - **definicija**: dva događaja su **nezavisna** ako ishod jednog događaja ne utiče na verovatnoću javljanja drugog događaja
- **PRIMERI**: višestruko bacanje novčića: **svaki** put je  $p(G) = p(P) = 0.5$ 
  - bacanja su **nezavisna**, ishod prethodnog ne utiče na verovatnoću sledećeg
- višestruko bacanje kocke: **svaki** put je za svaki ishod  $p = 1/6$
- višestruko igranje ruleta: verovatnoće crnog i crvenog polja su uvek iste
  - međutim, postoji tzv. **'kockarska zabluda'**: 'ako je ranije više puta bilo crveno, veća je šansa da u sledećem krugu bude crno' – nije tačno!

## 1. Univarijantni frekvencijski nacrti (UFN) 23

analize višestrukih bacanja novčića

- (a) bacanje **jednog** novčića (B1N): skup prostih događaja je  $S = \{P, G\}$ 
  - njihove verovatnoće su:  $p(P) = p(G) = 1/2$
- (b) bacanje **dva** novčića (B2N)
  - bacanje: jedan za drugim (ili istovremeno)
  - označićemo ih kao **prvi** i **drugi**
- skup prostih događaja ima četiri člana:  $S = \{\{P, P\}, \{P, G\}, \{G, P\}, \{G, G\}\}$ 
  - oznake: npr.  $\{P, G\}$  označava da je na **prvom** palo P, a na **drugom** G
  - uočiti:  $\{P, G\}$  i  $\{G, P\}$  nisu isti događaji (iako su im krajnji ishodi isti)
- kolike su **verovatnoće** svakog od ova četiri događaja?
  - prema pravilu jednako verovatnih događaja:
    - svi su jednako verovatni, ima ih 4, pa je za sve  $p = 1/4$
    - $p(P, P) = p(P, G) = p(G, P) = p(G, G) = 1/4 = 0.25$
  - prema pravilu preseka:
    - prvo i drugo bacanje su **nezavisni** događaji, za oba je  $p = 1/2 = 0.5$
    - verovatnoća da se dese **i jedan i drugi je proizvod**:  $0.5 * 0.5 = 0.25$
    - npr.:  $p(\{P, G\}) = p(P) * p(G) = (1/2) * (1/2) = 0.5 * 0.5 = 0.25 = 1/4$

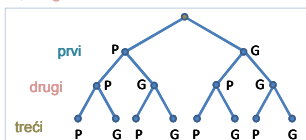


## 1. Univarijantni frekvencijski nacrti (UFN) 24

- **PRIMERI**: kolike su verovatnoće sledećih složenih događaja kod B2N:
  - da **oba** puta padne P?
    - skup povoljnih događaja  $S = \{\{P, P\}\}$ ;  $n=1$  od 4
    - verovatnoća:  $p = 1/4 = 0.25$
  - da **samo** (odn. **tačno**) **jedanput** padne P?
    - skup povoljnih događaja  $S = \{\{P, G\}, \{G, P\}\}$ ;  $n=2$  od 4
    - prema pravilu zbira:  $p = 1/4 + 1/4 = 1/2 = 0.5$
    - prema pravilu količnika: ima 2 povoljna od 4,  $p = 2/4 = 1/2 = 0.5$
  - da **bar** (odn. **najmanje**) **jedanput** padne P?
    - skup povoljnih događaja  $S = \{\{P, G\}, \{G, P\}, \{P, P\}\}$ ;  $n=3$  od 4
    - prema pravilu zbira:  $p = 1/4 + 1/4 + 1/4 = 3/4 = 0.75$
    - prema pravilu količnika: ima 3 povoljna od 4,  $p = 3/4 = 0.75$
  - da **nijednom** ne padne P?
    - skup povoljnih događaja:  $S = \{\{G, G\}\}$ ;  $n=1$
    - prema pravilu jednako verovatnih događaja:  $p = 1/4 = 0.25$
    - prema pravilu komplementa:
      - ovaj događaj je **komplement** događaja da bar jedanput padne P
      - dakle:  $p = 1 - 0.75 = 0.25$

## 1. Univarijantni frekvencijski nacrti (UFN) 25

- (c) bacanje tri novčića (B3N)
  - postoje prvi, drugi i treći novčić



- skup mogućih događaja ima  $N=8$  članova:
  - $\{(P,P,P), (P,P,G), (P,G,P), (P,G,G), (G,P,P), (G,P,G), (G,G,P), (G,G,G)\}$
- kolike su verovatnoće svakog od ovih jednako verovatnih osam događaja?
  - po pravilu jednako verovatnih događaja:  $p = 1/N = 1/8 = 0.125$
  - po pravilu preseka:
    - sva tri bacanja su *nezavisni* događaji, za svako je  $p = 1/2 = 0.5$
    - verovatnoća da se dese *sva tri* je *proizvod* njihovih verovatnoća
      - $p = (1/2) \cdot (1/2) \cdot (1/2) = (0.5) \cdot (0.5) \cdot (0.5) = 0.125 = 1/8$

## 1. Univarijantni frekvencijski nacrti (UFN) 26

- kolike su verovatnoće sledećih složenih događaja:
  - da tačno (odn. samo) jedanput padne P?
    - skup  $\{(P, G, G), (G, P, G), (G, G, P)\}$
    - P je palo ili u prvom, ili u drugom, ili u trećem bacanju
    - po pravilu jednako verovatnih događaja:  $n=3$ ;  $p = 3/8 = 0.375$
  - da tačno dvaput padne P?
    - skup  $\{(P, P, G), (P, G, P), (G, P, P)\}$ ;  $n=3$ ;  $p = 3/8 = 0.375$
    - P je palo prvo i drugo, ili prvo i treće, ili drugo i treće
  - da bar (odn. najmanje) dvaput padne P (dakle dva puta ili tri puta)?
    - skup  $\{(P, P, P), (P, P, G), (P, G, P), (G, P, P)\}$ ;  $n=4$ ;  $p = 4/8 = 1/2$
  - da ne padne P nijednom?
    - skup S:  $\{(G, G, G)\}$ ;  $n=1$ ;  $p = 1/8 = 0.125$
  - da bar jedanput padne P (dakle triput ili dvaput ili jedanput)?
    - skup:  $\{(P,P,P), (P,P,G), (P,G,P), (G,P,P), (P,G,G), (G,P,G), (G,G,P)\}$
    - po pravilu jednako verovatnih događaja:  $n=7$ ;  $p = 7/8 = 0.875$
    - po pravilu komplementa:
      - gornji skup je komplement događaja  $\{G,G,G\}$ , za koji važi  $p = 1/8$
      - prema pravilu, njegova verovatnoća je  $p = 1 - 1/8 = 7/8 = 0.875$

## 1. Univarijantni frekvencijski nacrti (UFN) 27

- (d) bacanje četiri, pet, ..., N novčića: isti principi
- ako događaji imaju *različite* verovatnoće, ne može se koristiti pravilo jednakih verovatnoća, ali može pravilo preseka
- npr.: desnorukost (D) i levorukost (L)
  - pretpostavićemo isključivost i iscrpnost ove kategorizacije
  - ako slučajno izaberemo *jednu* osobu:  $S = \{D, L\}$ ;  $p(D)=0.93$ ;  $p(L)=0.07$
  - ako slučajno izaberemo *dve* osobe, postoje 4 događaja:
    - $S = \{(D, D), (L, L), (D, L), (L, D)\}$
    - koja je verovatnoća da su obe desnoruke?
      - $p\{D, D\} = p(D) \cdot p(D) = 0.93 \cdot 0.93 = 0.865$
    - koja je verovatnoća da su obe levoruke?
      - $p\{L, L\} = p(L) \cdot p(L) = 0.07 \cdot 0.07 = 0.005$
    - koja je verovatnoća da je jedna desnoruka a druga levoruka?
      - $p\{D, L\} + p\{L, D\} = p(D) \cdot p(L) + p(L) \cdot p(D) = 0.065 + 0.065 = 0.130$
    - uočiti: verovatnoće sva četiri događaja se sabiraju do 1
      - $0.865 + 0.005 + 0.065 + 0.065 = 1$