

Statistični praktikum

Opisna statistika in grafično predstavljanje podatkov

Asistent dr. Kristina Veljković

Fakulteta za računalništvo in informatiko

TIPI SLUČAJNIH SPREMENLJIVK

- Slučajna spremenljivka X je količina, katere vrednosti so rezultat slučajnega poskusa.

TIPI SLUČAJNIH SPREMENLJIVK

- Slučajna spremenljivka X je količina, katere vrednosti so rezultat slučajnega poskusa.
- Na osnovi zaloge vrednosti razlikujemo:
 1. **numerične slučajne spremenljivke:** *diskretne* in *zvezne*,
 2. **kategorične slučajne spremenljivke:** *nominalne* (imenske) in *ordinalne* (urejenostne).

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti.

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.
- **spol**: 0=moški, 1=ženski,

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.
 - **spol**: 0=moški, 1=ženski,
 - **barva las**: 1=črna, 2=rjava, 3=rdeča, 4=blondna, 5=bela.

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.
 - **spol**: 0=moški, 1=ženski,
 - **barva las**: 1=črna, 2=rjava, 3=rdeča, 4=blondna, 5=bela.
- Ordinalne slučajne spremenljivke: kategorije so urejene, obstaja hierarhija med njimi.

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.
 - **spol**: 0=moški, 1=ženski,
 - **barva las**: 1=črna, 2=rjava, 3=rdeča, 4=blondna, 5=bela.
- Ordinalne slučajne spremenljivke: kategorije so urejene, obstaja hierarhija med njimi.
 - **stopnja bolečin**: 0=ni, 1=blage, 2=srednje, 3=močne,

KATEGORIČNE SLUČAJNE SPREMENLJIVKE

- Kategorične slučajne spremenljivke imajo kategorije za vrednosti. Kategorije se lahko numerično kodirajo.
- Nominalne slučajne spremenljivke: med kategorijami ni hierarhije.
 - **spol**: 0=moški, 1=ženski,
 - **barva las**: 1=črna, 2=rjava, 3=rdeča, 4=blondna, 5=bela.
- Ordinalne slučajne spremenljivke: kategorije so urejene, obstaja hierarhija med njimi.
 - **stopnja bolečin**: 0=ni, 1=blage, 2=srednje, 3=močne,
 - **stopnja izobrazbe**: 1= brez, 2=osnovna šola, 3=srednja šola, 4=fakulteta, 5=postdiplomski študij.

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Primer:* Zabeležili smo krvno skupino naključnega vzorca 100 Slovencev enega majhnega mesta. V naslednjo tabelo smo vpisali dobljene frekvence

krvna skupina	O	A	B	AB
frekvenca f	38	40	15	7

Grafično prikaži dobljene podatke.

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Primer:* Zabeležili smo krvno skupino naključnega vzorca 100 Slovencev enega majhnega mesta. V naslednjo tabelo smo vpisali dobljene frekvence

krvna skupina	O	A	B	AB
frekvenca f	38	40	15	7

Grafično prikaži dobljene podatke.

- Ustvarimo vektor karakterjev v R-ju

```
krv.skupina<-c(rep("O",38),rep("A",40),  
               rep("B",15), rep("AB",7))
```

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Frekvenčna tabela v R-ju

```
table(krv.skupina)
```

```
## krv.skupina  
##   A AB  B  O  
##  40  7 15 38
```

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Frekvenčna tabela v R-ju

```
table(krv.skupina)
```

```
## krv.skupina  
##   A AB  B  O  
##  40  7 15 38
```

- Podatke grafično prikažemo s stolpčnim diagramom (ang. *barplot*): na *x*-osi so kategorije, višina stolpca je frekvenca kategorije (absolutna frekvenca).

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Frekvenčna tabela v R-ju

```
table(krv.skupina)
```

```
## krv.skupina  
##   A  AB   B   O  
##  40   7  15  38
```

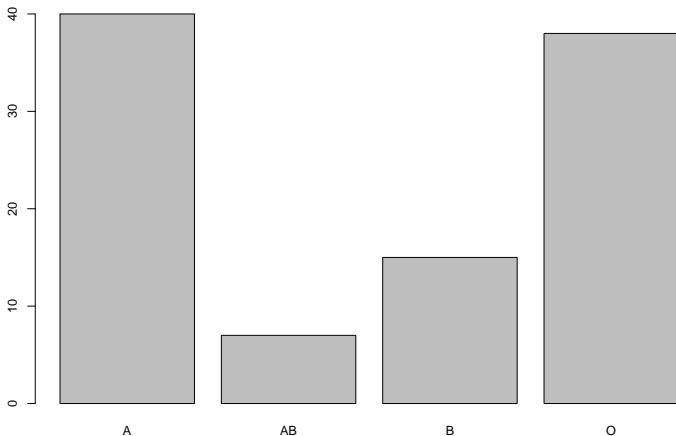
- Podatke grafično prikažemo s stolpčnim diagramom (ang. *barplot*): na *x*-osi so kategorije, višina stolpca je frekvenca kategorije (absolutna frekvenca).
- Stolpčni diagram v R-ju

```
barplot(table(krv.skupina))
```

- Na osnovi tabele in grafa opazimo, katera kategorija se javlja najbolj pogosto (**modus**), katera pa je najbolj redka.

NOMINALNA SLUČAJNA SPREMENLJIVKA - KRVNA SKUPINA

- Stolpčni diagram krvne skupine



ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.
- **Vzorčna mediana**

$$M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho} \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo} \end{cases}$$

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.
- **Vzorčna mediana**

$$M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho} \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo} \end{cases}$$

- Vzorčni prvi in tretji kvartil (po Tukey-jevi metodi):

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.
- **Vzorčna mediana**

$$M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho} \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo} \end{cases}$$

- Vzorčni **prvi** in **tretji kvartil** (po Tukey-jevi metodi):
 - Prvi kvartil Q_1 je mediana prve polovice sortiranih podatkov,

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.
- **Vzorčna mediana**

$$M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho} \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo} \end{cases}$$

- Vzorčni **prvi** in **tretji kvartil** (po Tukey-jevi metodi):
 - Prvi kvartil Q_1 je mediana prve polovice sortiranih podatkov,
 - Tretji kvartil Q_3 je mediana druge polovice sortiranih podatkov.

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Za ordinalne podatke lahko izračunamo, poleg modusa, vzorčno mediano in vzorčne kvartile.
- Podatki se razvrstijo v naraščajočem vrstnem redu $Y_1 \leq Y_2 \leq \dots \leq Y_n$.
- **Vzorčna mediana**

$$M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho} \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo} \end{cases}$$

- Vzorčni **prvi** in **tretji kvartil** (po Tukey-jevi metodi):
 - Prvi kvartil Q_1 je mediana prve polovice sortiranih podatkov,
 - Tretji kvartil Q_3 je mediana druge polovice sortiranih podatkov.

Za liho število podatkov, pri računanju kvartilov vključimo mediano v obe polovici sortiranih podatkov.

ORDINALNA SLUČAJNA SPREMENLJIVKA

- Primer:* Zabeležili smo stopnjo bolečin naključnega vzorca 20 oseb z revmatološkimi težavami (0 = brez bolečin, 1 = blage bolečine, 2 = srednje bolečine, 3 = močne bolečine). Podatke so razvrstili v naraščajočem vrstnem redu:

0 0 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3

Izračunaj opisno statistiko in grafično prikaži dobljene podatke.

ORDINALNA SPREMENLJIVKA

- Ustvarimo vektor z numeričnimi kodami stopnje bolečin

```
bol<-c(rep(0,2),rep(1,7),rep(2,8),rep(3,3))
```

ORDINALNA SPREMENLJIVKA

- Ustvarimo vektor z numeričnimi kodami stopnje bolečin

```
bol<-c(rep(0,2),rep(1,7),rep(2,8),rep(3,3))
```

- Modus (najbolj pogosta vrednost ali kategorija) je 2 (srednje bolečine).

ORDINALNA SPREMENLJIVKA

- Ustvarimo vektor z numeričnimi kodami stopnje bolečin

```
bol<-c(rep(0,2),rep(1,7),rep(2,8),rep(3,3))
```

- Modus (najbolj pogosta vrednost ali kategorija) je 2 (srednje bolečine).
- Mediana, prvi in tretji kvartil vzorca

```
quantile(bol,c(0.25,0.5,0.75))
```

```
## 25% 50% 75%  
##   1   2   2
```

ORDINALNA SPREMENLJIVKA

- Ustvarimo vektor z numeričnimi kodami stopnje bolečin

```
bol<-c(rep(0,2),rep(1,7),rep(2,8),rep(3,3))
```

- Modus (najbolj pogosta vrednost ali kategorija) je 2 (srednje bolečine).
- Mediana, prvi in tretji kvartil vzorca

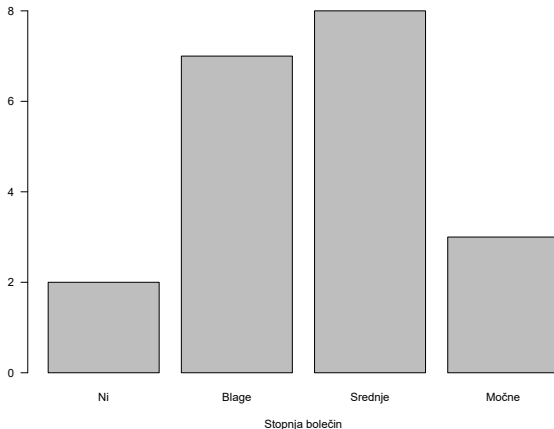
```
quantile(bol,c(0.25,0.5,0.75))
```

```
## 25% 50% 75%  
##   1   2   2
```

- Stolpčni diagram v R-ju

ORDINALNA SLUČAJNA SPREMENLJIVKA

```
par(las=1,mar=c(4,4,1,1),cex=1.1)
barplot(table(bol),xlab="Stopnja bolečin",
        names.arg=c("Ni","Blage","Srednje","Močne"))
```



ZVEZNE SLUČAJNE SPREMENLJIVKE

- Opisna statistika vzorca: minimalna in maksimalna vrednost, povprečje, mediana, standardni odklon, kvartili, itn.

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Opisna statistika vzorca: minimalna in maksimalna vrednost, povprečje, mediana, standardni odklon, kvartili, itn.
- **Vzorčno povprečje** $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Opisna statistika vzorca: minimalna in maksimalna vrednost, povprečje, mediana, standardni odklon, kvartili, itn.
- **Vzorčno povprečje** $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.
- **Popravljeni vzorčni standardni odklon** $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$.

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Primer:* Izmerili smo težo (v kg) slučajnega vzorca 50 študentov prvega letnika.

```
teza<-c(60.9, 72.7, 69.2, 74.6, 64.6, 69.4, 79.4,  
        71.1, 75.6, 68.5, 63.0, 69.9, 61.7, 62.3,  
        73.0, 62.6, 58.6, 66.8, 69.9, 65.9, 61.8,  
        67.1, 76.8, 74.9, 57.3, 72.4, 80.9, 64.5,  
        68.5, 68.4, 83.5, 75.4, 58.4, 59.7, 73.3,  
        68.8, 76.5, 81.0, 60.6, 71.2, 71.1, 61.4,  
        86.0, 60.4, 69.9, 79.3, 68.9, 74.4, 78.7,  
        66.6)
```

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Vzorčno povprečje \bar{x}

```
mean(teza)
```

```
## [1] 69.548
```

- Popravljeni vzorčni standardni odklon s

```
sd(teza)
```

```
## [1] 7.07722
```

ZVEZNE SLUČAJNE SPREMENLJIVKE

- **Povzetek s petimi števili:** minimum, maksimum, mediana, prvi in tretji kvartil (in še dodatno vzorčno povprečje)

```
summary(teza)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	57.30	63.38	69.30	69.55	74.55	86.00

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Osamelec (ang. *outlier*) lahko ima veliki vpliv na vzorčno povprečje, ne pa tudi na mediano.

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Osamelec (ang. *outlier*) lahko ima veliki vpliv na vzorčno povprečje, ne pa tudi na mediano.
- *Primer*: Dodajmo podatkom visoko težo 140 kg.

ZVEZNE SLUČAJNE SPREMENLJIVKE

- Osamelec (ang. *outlier*) lahko ima veliki vpliv na vzorčno povprečje, ne pa tudi na mediano.
- *Primer:* Dodajmo podatkom visoko težo 140 kg.

```
teza1<-c(teza,140)  
c(staro.povp=mean(teza),novo.povp=mean(teza1))
```

```
## staro.povp  novo.povp  
##    69.54800    70.92941
```

```
c(stara.med=median(teza),nova.med=median(teza1))
```

```
## stara.med  nova.med  
##    69.3      69.4
```

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Pri risanju histograma razvrstimo teže v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Pri risanju histograma razvrstimo težo v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.
- Za težo študentov je minimalna vrednost 57.3kg, maksimalna vrednost pa 86kg.
- Koliko razredov teže potrebujemo?

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Pri risanju histograma razvrstimo težo v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.
- Za težo študentov je minimalna vrednost 57.3kg, maksimalna vrednost pa 86kg.
- Koliko razredov teže potrebujemo?
- Sturges-ovo pravilo: za n podatkov je optimalno število razredov
$$k = \text{ceiling}(\log_2(n)) + 1$$

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Pri risanju histograma razvrstimo težo v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.
- Za težo študentov je minimalna vrednost 57.3kg, maksimalna vrednost pa 86kg.
- Koliko razredov teže potrebujemo?
- Sturges-ovo pravilo: za n podatkov je optimalno število razredov
$$k = \text{ceiling}(\log_2(n)) + 1$$
- Podatki o teži: $n = 50$, $k = \text{ceiling}(\log_2(50)) + 1 = 7$ (začetno število razredov).

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Pri risanju histograma razvrstimo težo v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.
- Za težo študentov je minimalna vrednost 57.3kg, maksimalna vrednost pa 86kg.
- Koliko razredov teže potrebujemo?
- Sturges-ovo pravilo: za n podatkov je optimalno število razredov
$$k = \text{ceiling}(\log_2(n)) + 1$$
- Podatki o teži: $n = 50$, $k = \text{ceiling}(\log_2(50)) + 1 = 7$ (začetno število razredov).
- Dolžina razreda

$$h = \frac{\max - \min}{k} = \frac{86 - 57.3}{7} = 4.1.$$

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

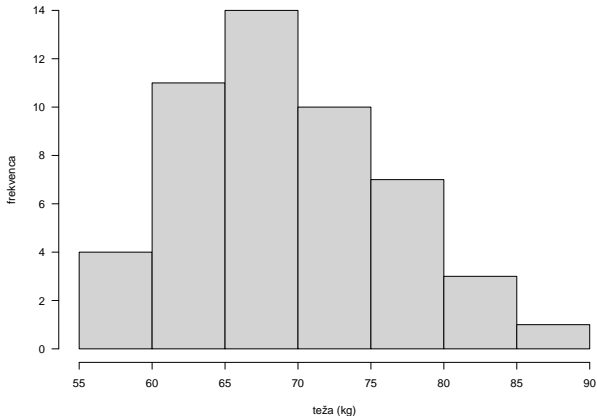
- Pri risanju histograma razvrstimo težo v željene razrede (intervale) in preštajemo število podatkov v vsakem razredu.
- Za težo študentov je minimalna vrednost 57.3kg, maksimalna vrednost pa 86kg.
- Koliko razredov teže potrebujemo?
- Sturges-ovo pravilo: za n podatkov je optimalno število razredov
$$k = \text{ceiling}(\log_2(n)) + 1$$
- Podatki o teži: $n = 50$, $k = \text{ceiling}(\log_2(50)) + 1 = 7$ (začetno število razredov).
- Dolžina razreda

$$h = \frac{\max - \min}{k} = \frac{86 - 57.3}{7} = 4.1.$$

- Vzamemo $k = 7$ razredov dolžine $h = 5$.

ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

```
par(las=1,mar=c(4,4,1,1),cex=1.1)
hist(teza,right=FALSE,main="", xlab="teža (kg)",
     ylab="frekvenca")
```



ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Histogram gostote: ploščina vseh pravokotnikov histograma je enaka 1 (ploščina posameznega pravokotnika je enaka relativni frekvenci $f_{rel} = \frac{f}{n}$).

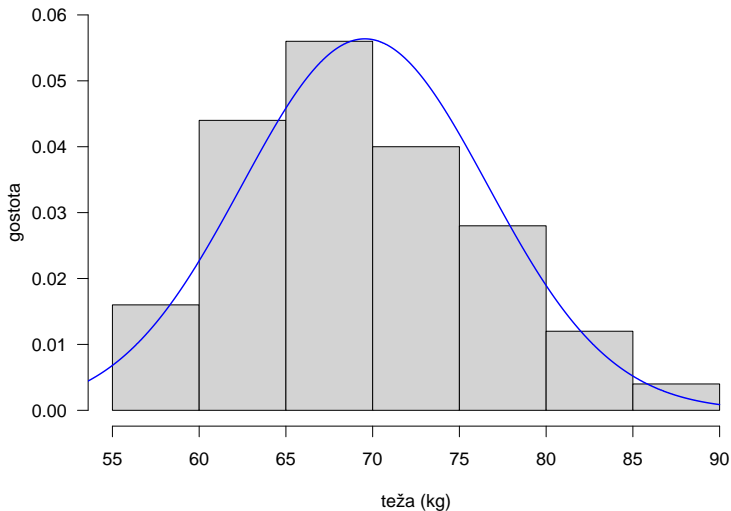
ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

- Histogram gostote: ploščina vseh pravokotnikov histograma je enaka 1 (ploščina posameznega pravokotnika je enaka relativni frekvenci $f_{rel} = \frac{f}{n}$).
- Preko histograma gostote narišemo gostoto normalne porazdelitve (μ ocenimo z $\bar{x} = 69.55$, standardni odklon σ z $s = 7.08$).

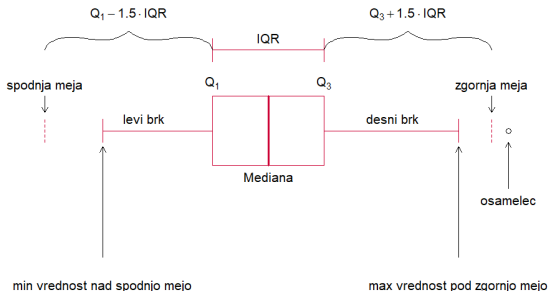
ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM

```
x<-seq(55,90,0.001)
par(las=1,cex=1.4,mar=c(4,4,1,1))
hist(teza,right=FALSE,freq=FALSE,main="",ylim=c(0,0.06),
     xlab="teža (kg)",ylab="gostota")
lines(x,dnorm(x,mean(teza),sd(teza)),col="blue",lwd=1.8)
```

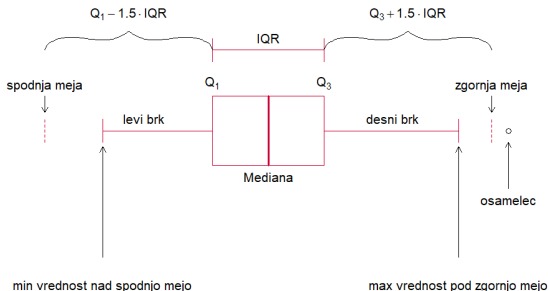
ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM



ZVEZNE SLUČAJNE SPREMENLJIVKE - ŠKATLA Z BRKI



ZVEZNE SLUČAJNE SPREMENLJIVKE - ŠKATLA Z BRKI

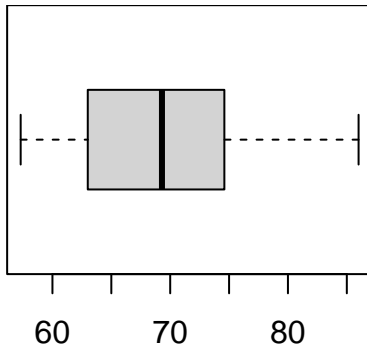


- Navpična črta v škatli je mediana, meji škatle sta Q_1 - prvi kvartil in Q_3 - tretji kvartil.
- Škatla je dolžine $IQR = Q_3 - Q_1$.
- Spodnja meja = $Q_1 - 1.5 \cdot IQR$, zgornja meja = $Q_3 + 1.5 \cdot IQR$.
- Podatki, ki so manjši od spodnje meje ali večji od zgornje meje, se imenujejo osamelci (ang. *outliers*).

ZVEZNE SLUČAJNE SPREMENLJIVKE - ŠKATLA Z BRKI

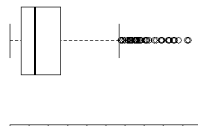
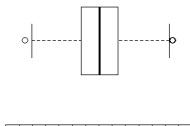
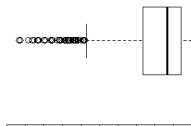
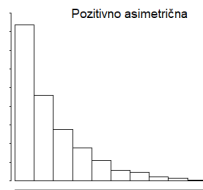
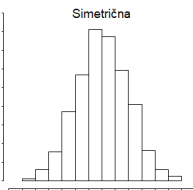
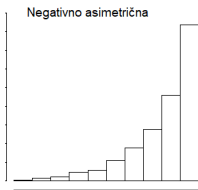
- Pri simetrični porazdelitvi je mediana na sredini škatle in sta brka enakih dolžin.
- Škatla z brki teže študentov

```
par(mar=c(3, 3, 0, 0))  
boxplot(teza, horizontal=TRUE, cex=1.2)
```



ZVEZNE SLUČAJNE SPREMENLJIVKE - HISTOGRAM IN ŠKATLA Z BRKI

- Primeri histogramov in škatle z brki negativno asimetrične, simetrične in pozitivno asimetrične vzorčne porazdelitve.



DISKRETNE SLUČAJNE SPREMENLJIVKE

- Diskretne podatke z malim številom vrednosti grafično prikažemo s stolpčnim diagramom.

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Diskretne podatke z malim številom vrednosti grafično prikažemo s stolpčnim diagramom.
- Diskretne podatke z več vrednostmi prikažemo s histogramom in škatlo z brki.

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Diskretne podatke z malim številom vrednosti grafično prikažemo s stolpčnim diagramom.
- Diskretne podatke z več vrednostmi prikažemo s histogramom in škatlo z brki.
- *Primer:* Zabeležili smo število otrok slučajnega vzorca 60 družin. Izračunaj opisno statistiko, ter nariši histogram in škatlo z brki za število otrok.

```
otroci<-c(0,4,5,2,3,2,1,2,3,3,4,5,1,8,1,2,3,2,  
          2,4,0,2,2,1,6,1,5,3,0,3,2,3,2,2,4,3,  
          3,1,2,2,1,2,4,3,1,1,0,2,2,2,3,1,2,1,  
          2,2,2,2,3,0)
```

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Za diskretne podatke izračunamo opisno statistiko na enak način kot za zvezne podatke (vzorčno povprečje, vzorčni standardni odklon, povzetek s petimi števili).

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Za diskretne podatke izračunamo opisno statistiko na enak način kot za zvezne podatke (vzorčno povprečje, vzorčni standardni odklon, povzetek s petimi števili).
- Povzetek s petimi števili in vzorčno povprečje.

```
summary(otroci)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	1.000	2.000	2.333	3.000	8.000

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Za diskretne podatke izračunamo opisno statistiko na enak način kot za zvezne podatke (vzorčno povprečje, vzorčni standardni odklon, povzetek s petimi števili).
- Povzetek s petimi števili in vzorčno povprečje.

```
summary(otroci)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	1.000	2.000	2.333	3.000	8.000

- Popravljeni vzorčni standardni odklon števila otrok.

```
sd(otroci)
```

```
## [1] 1.514525
```

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Tabela frekvenc

```
table(otroci)
```

```
## otroci
```

```
##  0  1  2  3  4  5  6  8
```

```
##  5 11 22 12  5  3  1  1
```

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Tabela frekvenc

```
table(otroci)
```

```
## otroci
##  0  1  2  3  4  5  6  8
##  5 11 22 12  5  3  1  1
```

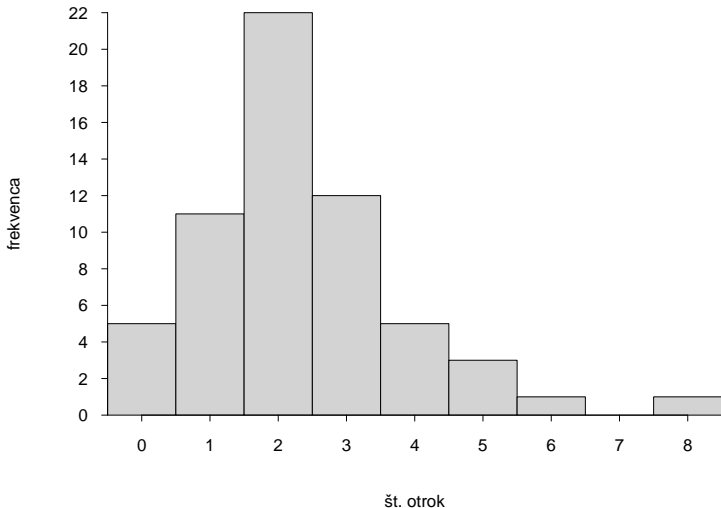
- Želimo, da je sredina vsakega razreda histograma v vrednostih 0-8. Meje razredov so: vrednost ± 0.5 .

DISKRETNE SLUČAJNE SPREMENLJIVKE

- Histogram števila otrok

```
par(las=1,cex=1.3,mar=c(4,4,1,1))
hist(otroci,breaks=seq(-0.5,8.5,by=1),ylim=c(0,22),
     main="",xlab="št. otrok",ylab="frekvenca",axes=F)
axis(1,pos=0,at=0:8,tcl=-0.3)
axis(2,pos=-0.5,at=seq(0,22,by=2),tcl=-0.3)
```

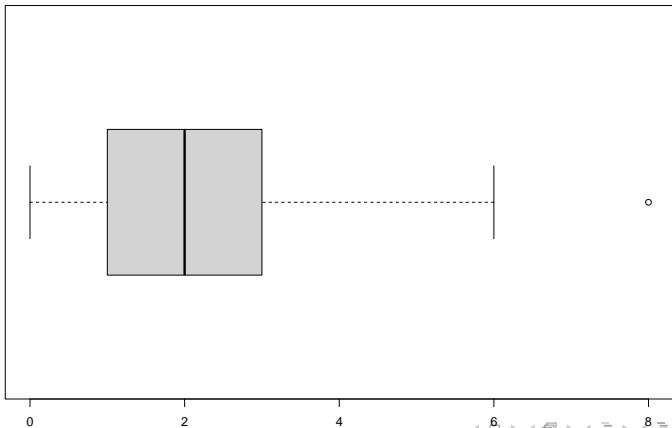
DISKRETNE SLUČAJNE SPREMENLJIVKE



DISKRETNE SLUČAJNE SPREMENLJIVKE

- Škatla z brki števila otrok

```
boxplot (otroci, horizontal=TRUE)
```



OPISNA STATISTIKA IN GRAFI

- Za vse tipe spremenljivk lahko izračunamo frekvence (absolutne in relativne).
- Tabela opisne statistike in grafov za kategorične podatke

Podatki	Opisna statistika	Grafi
<i>Nominalni</i>	Modus	Stolpčni diagram
<i>Ordinalni</i>	Modus, mediana, kvartili	Stolpčni diagram

OPISNA STATISTIKA IN GRAFI

- Tabela opisne statistike in grafov za numerične podatke

Podatki	Opisna statistika	Grafi
<i>Diskretni</i>	Minimum, maksimum, modus, mediana, kvartili, povprečje, standardni odklon	<i>malo število vrednosti:</i> stolpčni diagram; <i>večje število vrednosti:</i> histogram, škatla z brki
<i>Zvezni</i>	Minimum, maksimum, modus (za interval vrednosti), mediana, kvartili, povprečje, standardni odklon	histogram škatla z brki