Ústav matematiky a statistiky
Přírodovědecká fakulta
Masarykova univerzita

Aplikovaná statistika I

Téma 3: Základní číselné charakteristiky

Veronika Bendová, Zdeňka Geršlová

Úvod a motivace

- minulá hodina → bodové/intervalové rozložení četností
  - seznámení s daty
  - výhody: široké množství informací, globální pohled na data
  - nevýhody: přemíra informací, horší interpretovatelnost, ztížené srovnávání dvou datasetů
  - Adam a Bára provádí nezávislý výzkum na stejné téma → dvě různé nemocnice → dva různé datasety (A) a (B) → dvě různé variační řady → porovnávání variačních řad (nepřehledné a neefektivní)
- → vznik *číselných charakteristik*
  - elegantní a jednoduché vystihnutí charakteristických rysů znaku zpravidla pomocí jednoho čísla
  - snadno spočítatelné i interpretovatelné
- různá data → různé charakteristiky
- typy dat
  - **Nominální**,
  - **Ordinální**,
  - **Intervalová**,
  - **Poměrová**

- typy charakteristik
  - polohy
  - variability
  - závislosti
  - $+$ nesymetrie (intervalové znaky)
- přehled číselných charakteristik podle typu znaku a sledované vlastnosti

	Poloha	**Variabilita**	**Symetrie**	**Závislost**
Nominální	modus	–	–	Cramérův koeficient
Ordinální	medián	interkvartilové rozpětí	–	Spearmanův koef. poř. korel.
Intervalový	aritmetický průměr	rozptyl směrodatná odchylka	koeficient šikmosti koeficient špičatosti	Pearsonův korel. koeficient

Nominální znaky

- varianty znaku jsou neporovnatelné
  - barva očí: modrá, zelená, hnědá
  - pohlaví: m, f
- charakteristika polohy
  - *modus* . . . nejčetnější varianta znaku
- charakteristika závislosti
  - **Cramérův koeficient** r_C - těsnost závislosti u nominálních znaků
  - $r_C \in \langle 0; 1 \rangle$.
  - stupnice míry závislosti podle Cramérova koeficientu

Cramérův koeficient	**Interpretace**
0.0 – 0.1	zanedbatelný stupeň závislosti
0.1 – 0.3	slabý stupeň závislosti
0.3 – 0.7	střední stupeň závislosti
0.7 – 1.0	silný stupeň závislosti

Příklad 3.1. Charakteristika polohy nominálního znaku
V souboru 20-more-samples-probabilities-pubis.txt jsou údaje o původu žen (european – evropský; african – africký; inuits – inuitský) a o míře změn kostního reliéfu na vnitřní straně stydké kosti v blízkosti stydké spony (absence – nepřítomnost změn; trace.to small – stopové až malé změny; moderate.to.large – střední až výrazné změny). Znaky X a Y jsou typickým příkladem znaků nominálního typu. Najděte modus pro znak $X =$ *původ ženy* i pro znak $Y =$ *míra změny kostního reliéfu*.

Tabulka 3.1: Simultánní absolutní četnosti pro znaky *původ ženy* a *míra změny kostního reliéfu*

	absence	trace.to small	moderate.to.large
European	30	20	10
African	56	37	17
Inuits	16	6	13

Řešení příkladu 3.1

```
1  (data <- data.frame(absence = c(30,    56,     16),
2                      small       = c(20, 37,     6),
3                      large =     c(10,   17, 13),
4                      row.names = c('european', 'african', 'inuits')))
```

```
          absence small large                                                   5
european       30    20    10                                                 6
african        56    37    17                                                 7
inuits         16     6    13                                                 8
```

Zaměřme se nejprve na znak $X =$ *původ ženy*. Číselná charakteristika *modus* je definována jako nejčetnější varianta sledovaného znaku.

```
(nj. <- apply(data, MARGIN = 1, FUN = sum))
```

```
european  african   inuits
      60       110       35
```

```
rowSums(data)
```

```
european  african   inuits
      60       110       35
```

Interpretace výsledků: Nejčetnější variantou znaku *původ ženy* je ..... ($n =$ .....). Nejvíce žen v datovém souboru pochází z populace ..... .

Analogicky nyní najdeme modus znaku $Y =$ *míra změny kostního reliéfu*.

```
(n.k <- apply(data, MARGIN = 2, FUN = sum))
```

```
absence   small   large
    102      63      40
```

```
colSums(data)
```

```
absence   small   large
    102      63      40
```

Interpretace výsledků: Nejčastěji se u žen v datovém souboru vyskytovala ..... míra změny kostního reliéfu (s absolutní četností $n =$ .....).

Příklad 3.2. Charakteristika závislosti mezi dvěma nominálními znaky
Zaměřte se nyní na oba znaky X = *původ ženy* a Y = *míra změny kostního reliéfu* najednou. Určete míru závislosti mezi znaky X a Y .

Řešení příkladu 3.2
Protože X a Y jsou znaky ..... typu, použijeme na určení míry závislosti mezi nimi .... Stupnice míry závislosti podle hodnoty tohoto koeficientu je uvedena v tabulce 3.2

Tabulka 3.2: Stupnice míry závislosti podle Cramérova koeficientu

Cramérův koeficient	**Interpretace**
0.0 – 0.1	Zanedbatelný stupeň závislosti
0.1 – 0.3	Slabý stupeň závislosti
0.3 – 0.7	Střední stupeň závislosti
0.7 – 1.0	Silný stupeň závislosti

```
21 lsr::cramersV(data)
```

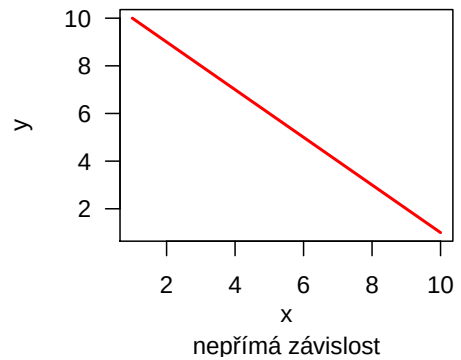
```
[1] 0.1516982                                                    22
```

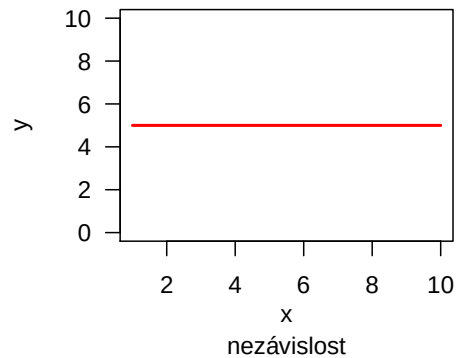
Interpretace výsledků: Hodnota Cramérova koeficientu vyšla ..... . Mezi původem ženy a mírou změny jejího kostního reliéfu existuje ..... stupeň závislosti .

Ordinální znaky

- různé hodnoty znaku můžeme porovnávat, ale nemůžeme obsahově interpretovat rozdíly mezi nimi
  - školní klasifikace, pořadí 10 pacientů podle závažnosti onemocnění, ...
- charakteristika polohy
  - α-kvantil x_α ... takové číslo, že $\alpha \times 100\%$ hodnot v datovém souboru je menších nebo rovných hodnotě x_α; $\alpha \in \langle 0; 1 \rangle$
    - medián $x_{0.5}$
    - dolní kvartil $x_{0.25}$
    - horní kvartil $x_{0.75}$
  - $n\alpha =$ celé číslo $c \rightarrow x_\alpha = \frac{x_{(c)} + x_{(c+1)}}{2}$
  - $n\alpha =$ necelé číslo $\rightarrow$ zaokrouhlíme nahoru na nejbližší vyšší celé číslo $c \rightarrow x_\alpha = x_{(c)}$
- charakteristika variability
  - (inter)kvartilové rozpětí IQR
  - $IQR = x_{0.75} - x_{0.25}$
  - v intervalu leží 50 % dat.
- charakteristika závislosti
  - dva znaky; alespoň jeden znak je ordinální
  - **Spearmanův koeficient pořadové korelace** r_S
  - určuje míru **pořadové** závislosti mezi znaky X a Y

- $r_S \in \langle -1; 1 \rangle$
- $r_S > 0$ . . . přímá závislost
- $r_S < 0$ . . . nepřímá závislost
- $r_S = 0$ . . . nezávislost

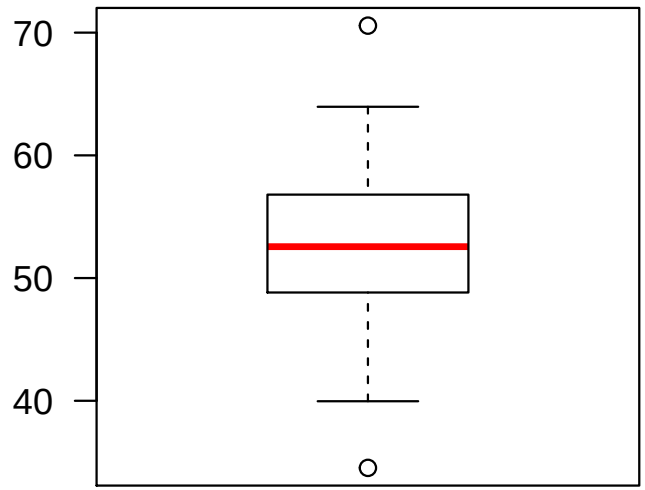




- stupnice míry závislosti podle Spearmanova koeficientu pořadové korelace

$\lvert r_S \rvert$	**Interpretace**
0.0	**pořadová** nezávislost
0.0 – 0.1	velmi nízký stupeň závislosti
0.1 – 0.3	nízký stupeň závislosti
0.3 – 0.5	mírný stupeň závislosti
0.5 – 0.7	význačný stupeň závislosti
0.7 – 0.9	vysoký stupeň závislosti
0.9 – 1.0	velmi vysoký stupeň závislosti
1.0	úplná **pořadová** závislost

- grafická viualizace ordinálních dat
  - krabicový diagram



Příklad 3.3. Základní číselné charakteristiky pro ordinální znak
Načtěte datový soubor 17-anova-newborns-2.txt a odstraňte neznámé hodnoty. Zaměřte se pouze na novorozence ženského pohlaví. Zjistěte dimenzi datové tabulky obsahující údaje o těchto novorozencích a vytvořte tabulku základních číselných charakteristik pro znak $X =$ *vzdělání matky*.

Řešení příkladu 3.3

```
23 data <- read.delim('17-anova-newborns-2.txt', sep = '\t')
24 head(data, n = 4)
```

```
  edu.M prch.N sex.C weight.C weight.K
1     2      0     m     3470        2
2     2      0     m     3240        2
3     2      0     f     2980        2
4     1      0     m     3280        2
```

```
30 data <- na.omit(data)
31 data.F <- data[data$sex == 'f', ]
32 dim(data.F) # 662 5
```

Po odstranění neznámých hodnot obsahuje datová tabulka údaje o ..... novorozencích ženského pohlaví, přičemž u každého z těchto novorozenců máme záznamy o ..... znacích.

Znak $X =$ *vzdělání matky* je příkladem ..... dat. V tabulce základních charakteristik budou obsaženy následující charakteristiky: minimální hodnota, dolní kvartil, medián, horní kvartil, maximální hodnota a interkvartilové rozpětí.

1. minimální hodnota $x_{min} = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.$

```
edu <- data.F$edu.M
edu.sort <- sort(edu) # serazena data
min.e <- min(edu)
```

2. dolní kvartil $x_{0.25}$
   - $n = \ldots\ldots\ldots\ldots\ldots\ldots$, $\alpha = \ldots\ldots\ldots\ldots\ldots\ldots$
   - $\alpha \times n = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \rightarrow$ je / není celé číslo
   - $x_{0.25} =$

```
n <- length(edu)
alpha <- 0.25
alpha * n # 165.5 ... neni cele cislo -> zaokrouhleni
n.alpha <- ceiling(alpha * n)
edu.sort[n.alpha]
x0.25 <- quantile(edu, probs = 0.25, type = 2)
```

3. medián $x_{0.50}$
   - $n = \ldots\ldots\ldots\ldots\ldots\ldots$, $\alpha = \ldots\ldots\ldots\ldots\ldots\ldots$
   - $\alpha \times n = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \rightarrow$ je / není celé číslo
   - $x_{0.50} =$

```
alpha <- 0.5
alpha * n
(edu.sort[331] + edu.sort[332]) / 2
x0.5 <- median(edu, type = 2)
x0.5 <- quantile(edu, probs = 0.50, type = 2)
```

4. horní kvartil $x_{0.75}$

   - $n =$ ....., $\alpha =$ .....
   - $\alpha \times n =$ ..... $\rightarrow$ je / není celé číslo
   - $x_{0.75} =$

```
47  alpha <- 0.75
48  alpha * n
49  edu.sort[ceiling(alpha * n) ]
50  x0.75 <- quantile(edu, probs = 0.75, type = 2)
```

5. maximální hodnota $x_{max} =$ ..... .

```
51  max.e <- max(edu)
```

6. interkvartilové rozpětí $IQR = x_{0.75} - x_{0.25} =$ ..... .

```
52  IQR.e <- x0.75 - x0.25
53  (tab <- data.frame(min = min.e, dolni.kv = x0.25, median = x0.5,
54                     horni.kv = x0.75, max = max.e, IQR = IQR.e))
```

```
    min dolni.kv median horni.kv max IQR                                    55
25%   1        1      2        3   4   2                                    56
```

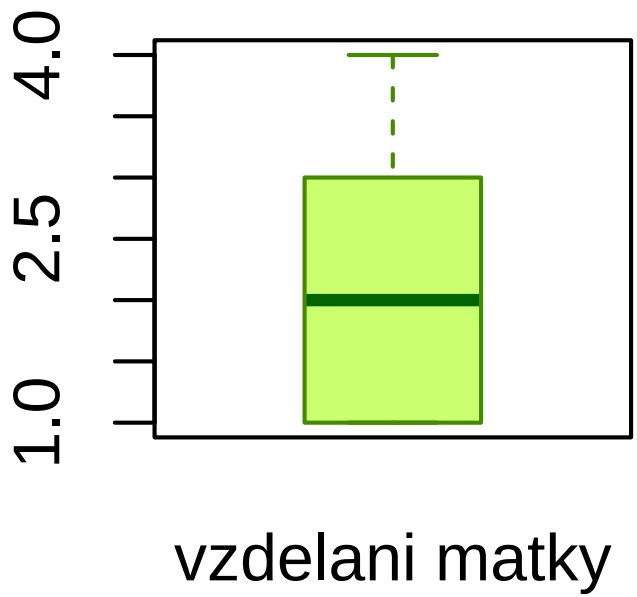
Interpretace výsledků: Vzdělání matky u novorozenců ženského pohlaví se pohybovalo v rozmezí ..... – ..... . Dolní kvartil vzdělání matky nabývá hodnoty ....., tj. ..... % novorozenců ženského pohlaví má matku se ..... vzděláním. Medián vzdělání matky nabývá hodnoty ....., tj. ..... % novorozenců ženského pohlaví má matku se ..... nebo ..... vzděláním.

Horní kvartil vzdělání matky nabývá hodnoty ....., tj. ..... % novorozenců ženského pohlaví má matku se ....., ... nebo ..... vzděláním. Hodnota interkvartilového rozpětí je rovna ..... .

Příklad 3.4. Krabicový diagram
Sestrojte krabicový diagram pro znak $X =$ *vzdělání matky* pro novorozence ženského pohlaví. Zaměřte se na vzhled krabicového diagramu a zamyslete se nad tím, kde je v krabicovém diagramu zobrazen medián, dolní kvartil, horní kvartil a mezikvartilové rozpětí.

Řešení příkladu 3.4

```
57  par(mar = c(2, 3, 1, 1))
58  boxplot(edu, col = 'darkolivegreen1',
59          border = 'chartreuse4', medcol = 'darkgreen', xlab = '')
60  mtext('vzdelani matky', side = 1, line = 1)
```



Příklad 3.5. Charakteristika závislosti mezi ordinálními znaky
Zaměřme se nyní na oba znaky X = *vzdělání matky* a Y = *porodní hmotnost novorozence* najednou. Určete míru závislosti mezi znaky X a Y u novorozenců ženského pohlaví.

Řešení příkladu 3.5
Znak X je ..... typu, zatímco znak Y je ..... typu $\rightarrow$ ke znaku Y budeme přistupovat jako ke znaku ..... typu. Ke stanovení míry závislosti použijeme ..... koeficient ..... korelace.

Tabulka 3.3: Stupnice míry závislosti podle Spearmanova koeficientu pořadové korelace

$\lvert r_S \rvert$	**Interpretace**
0.0	pořadová nezávislost
0.0 – 0.1	velmi nízký stupeň závislosti
0.1 – 0.3	nízký stupeň závislosti
0.3 – 0.5	mírný stupeň závislosti
0.5 – 0.7	význačný stupeň závislosti
0.7 – 0.9	vysoký stupeň závislosti
0.9 – 1.0	velmi vysoký stupeň závislosti
1.0	úplná pořadová závislost

```
61  data.F <- data[data$sex == 'f', ]
62  cor(data.F$edu.M, data.F$weight.C, method = 'spearman') #0.09478
```

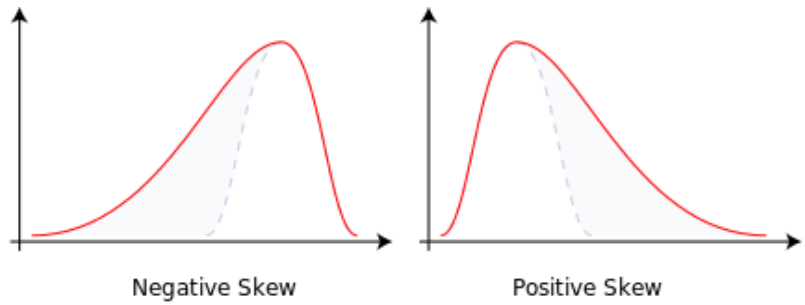
```
[1] 0.09478423                                                    63
```

Interpretace výsledku: Hodnota Spearmanova koeficientu pořadové korelace vyšla ..... . Mezi počtem starších sourozenců a porodní hmotností novorozence ženského pohlaví existuje ..... stupeň .... závislosti.

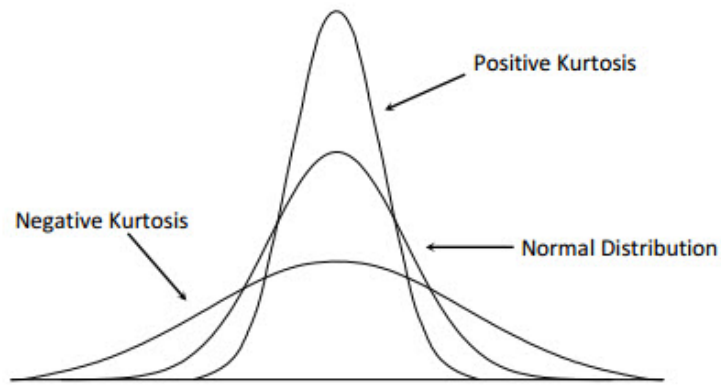
Intervalové znaky

- hodnoty znaků můžeme nejen vzájemně porovnat, ale můžeme též říci, o kolik se liší
  - porodní hmotnost novorozence (v g), největší šířka mozkovny (v mm), BMI (v kg/m^2)
- charakteristiky polohy
  - aritmetický průměr m
    - $m = \frac{1}{n} \sum_{i=1}^{n} x_i$
    - ovlivněn vybočujícími hodnotami $\rightarrow$ vhodný, máme-li symetrická data
  - α-kvantily x_α
    - medián $x_{0.5}$, dolní kvartil $x_{0.25}$, horní kvartil $x_{0.75}$, ...
    - nejsou ovlivněny vybočujícími hodnotami $\rightarrow$ vhodné, máme-li nesymetrická data
- charakteristiky variability
  - rozptyl s^2
    - $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2$
    - průměrná kvadratická odchylka hodnot od jejich aritmetického průměru.
    - $s^2 \geq 0$
    - ovlivněn vybočujícími hodnotami $\rightarrow$ vhodný, máme-li symetrická data
    - rozptyl $s^2 \rightarrow$ jednotky $^\wedge$ 2
  - směrodatná odchylka s
    - $s = \sqrt{s^2}$
    - převádí rozptyl do původních jednotek
  - interkvartilové rozpětí IQR ($x_{0.75}$ - $x_{0.25}$)

- charakteristiky nesymetrie
  - šikmost α_3
    - $\alpha_3 = 0 \rightarrow$ symetrické rozdělení dat
    - $\alpha_3 < 0 \rightarrow$ záporně zešikmené$\rightarrow$ prodloužený levý konec
    - $\alpha_3 > 0 \rightarrow$ kladně zešikmené $\rightarrow$ prodloužený pravý konec



  - špičatost α_4
    - $\alpha_4 = 0 \rightarrow$ normální rozdělení dat
    - $\alpha_4 > 0 \rightarrow$ strmé rozdělení dat
    - $\alpha_4 < 0 \rightarrow$ ploché rozdělení dat (Říp)



- charakteristika těsnosti závislosti
  - dva znaky X a Y, oba intervalového typu
  - Pearsonův korelační koeficient r_{12}
    - určuje míru **lineární** závislosti mezi znaky X a Y
    - $r_{12} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i - m_1}{s1} \frac{y_i - m_2}{s_2}$
    - $r_{12} \in \langle -1\,;\, 1 \rangle$
    - $r_{12} > 0$ . . . přímá závislost
    - $r_{12} < 0$ . . . nepřímá závislost
    - $r_{12} = 0$ . . . nezávislost
  - stupnice míry závislosti podle Pearsonova korelačního koeficientu

$\lvert r_{12} \rvert$	**Interpretace**
0.0	**lineární** nezávislost
0.0 – 0.1	velmi nízký stupeň závislosti
0.1 – 0.3	nízký stupeň závislosti
0.3 – 0.5	mírný stupeň závislosti
0.5 – 0.7	význačný stupeň závislosti
0.7 – 0.9	vysoký stupeň závislosti
0.9 – 1.0	velmi vysoký stupeň závislosti
1.0	úplná **lineární** závislost

Příklad 3.6. Základní číselné charakteristiky pro intervalový znak
Načtěte datový soubor 01-one-sample-mean-skull-mf.txt a odstraňte z načtených dat NA hodnoty. Zaměřte se pouze na znak $X =$ *největší šířka mozkovny* pro skelety mužského pohlaví. Vytvořte tabulku základních číselných charakteristik pro znak X.

Řešení příkladu 3.6

```
64 data <- read.delim('01-one-sample-mean-skull-mf.txt')
65 head(data, n = 3)
```

```
   id   pop sex skull.L skull.B
1 416 egant   m      188      145
2 417 egant   m      172      139
3 420 egant   m      176      138
```

```
70 data <- na.omit(data)
71 skull.BM <- data[data$sex == 'm', 'skull.B']
```

Po odstranění neznámých hodnot obsahuje datová tabulka údaje o ..... skeletech mužského pohlaví.

Znak $X =$ *největší šířka mozkovny* pro skelety mužského pohlaví je příkladem ..... dat. V tabulce základních číselných charakteristik budou obsaženy následující charakteristiky: aritmetický průměr, směrodatná odchylka, minimální hodnota, dolní kvartil, medián, horní kvartil, maximální hodnota, interkvartilové rozpětí, koeficient šikmosti a koeficient špičatosti.

1. aritmetický průmer m
   - $m = \frac{1}{n}\sum_{i=1}^{n} x_i =$

```
72 n <- length(skull.BM) # 216
73 m <- 1 / n * sum(skull.BM) # 137.1852
74 m <- mean(skull.BM) # 137.1852
```

2. rozptyl s^2

   - $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2 =$

```
75  s2 <- 1 / n * sum((skull.BM - m) ^ 2) # 23.16941
```

3. směrodatná odchylka s

   - $s = \sqrt{s^2} =$

```
76  s <- sqrt(s2) # 4.813461
```

4. minimální hodnota $x_{min} =$ .....

```
77  skull.BM.sort <- sort(skull.BM) # 124, 127, 127, ..., 148, 149, 149
78  min.BM <- min(skull.BM) # 124
```

5. dolní kvartil $x_{0.25}$

   - $n =$ ....., $\alpha =$ .....
   - $\alpha \times n =$ ..... $\rightarrow$ je / není celé číslo
   - $x_{0.25} =$

```
79  alpha <- 0.25
80  n * alpha # 54
81  (skull.BM.sort[54] + skull.BM.sort[55]) / 2 # 134
82  x0.25 <- quantile(skull.BM, probs = 0.25, type = 2) # 134
```

6. medián $x_{0.50}$

   - $n =$ ....., $\alpha =$ .....
   - $\alpha \times n =$ ..... $\rightarrow$ je / není celé číslo
   - $x_{0.50} =$

```
83  alpha <- 0.5
84  n * alpha # 108
85  (skull.BM.sort[108] + skull.BM.sort[109]) / 2 # 137
86  x0.50 <- quantile(skull.BM, probs = 0.50, type = 2) # 137
```

7. horní kvartil $x_{0.75}$

   - $n =$ ....., $\alpha =$ .....
   - $\alpha \times n =$ ..... $\rightarrow$ je / není celé číslo
   - $x_{0.75} =$

```
87  alpha <- 0.75
88  n * alpha # 162
89  (skull.BM.sort[162] + skull.BM.sort[163]) / 2 # 140
90  x0.75 <- quantile(skull.BM, probs = 0.75, type = 2) # 140
```

8. maximální hodnota $x_{max} =$ .....

```
91  max.BM <- max(skull.BM) # 149
```

9. interkvartilové rozpětí $IQR = x_{0.75} - x_{0.25} =$ ..... .

```
92  IQR.BM <- x0.75 - x0.25
```

6. koeficient šikmosti b_1

   - $b_1 =$ .....

```
sikmost <- e1071::skewness(skull.BM, type = 3) # 0.08410943
```

7. koeficient špičatosti b_2

   - $b_2 =$ .....

```
spicatost <- e1071::kurtosis(skull.BM, type = 3) # -0.2956831
tab <- data.frame(m, s, min = min.BM, dolni.kv = x0.25, median = x0.50,
                  horni.kv = x0.75, max = max.BM,
                  IQR = IQR.BM, sikmost, spicatost, row.names = c('muzi'))
(tab <- round(tab, digits = 2))
```

```
          m    s min dolni.kv median horni.kv max IQR sikmost spicatost
muzi 137.19 4.81 124      134    137      140 149   6    0.08      -0.3
```

Interpretace výsledků: Naměřené hodnoty největší šířky mozkovny pro skelety mužského pohlaví se pohybují v rozmezí ..... – ..... mm. Průměrná hodnota největší šířky mozkovny je ..... mm se směrodatnou odchylkou ..... mm. 25% naměřených hodnot je menších nebo rovných ..... mm, 50% naměřených hodnot je menších nebo rovných ..... mm, 75% naměřených hodnot je menších nebo rovných ..... mm. Interkvartilové rozpětí naměřených hodnot je rovno ..... mm. Hodnota koeficientu šikmosti, ..., ukazuje na ..... zešikmená data (prodloužený ..... konec). Hodnota koeficientu špičatosti, ..., ukazuje na ..... charakter dat.

Příklad 3.7. Charakteristika závislosti pro znaky intervalového typu
Zaměřme se nyní na znaky X = *největší šířka mozkovny* a Y = *největší délka mozkovny* pro skelety mužského pohlaví najednou. Určete míru závislosti mezi znaky X a Y.

Řešení příkladu 3.7
Oba znaky X a Y jsou ..... typu. Ke stanovení míry závislosti použijeme ..... korelační koeficient.

Tabulka 3.4: Stupnice míry závislosti podle Pearsonova korelačního koeficientu

$\lvert r_{12} \rvert$	**Interpretace**
0.0	**lineární** nezávislost
0.0 – 0.1	velmi nízký stupeň závislosti
0.1 – 0.3	nízký stupeň závislosti
0.3 – 0.5	mírný stupeň závislosti
0.5 – 0.7	význačný stupeň závislosti
0.7 – 0.9	vysoký stupeň závislosti
0.9 – 1.0	velmi vysoký stupeň závislosti
1.0	úplná **lineární** závislost

```
101  skull.BM <- data[data$sex == 'm', 'skull.B']
102  skull.LM <- data[data$sex == 'm', 'skull.L']
103  cor(skull.BM, skull.LM, method = 'pearson')
```

```
[1] 0.168157                                                    104
```

Interpretace výsledků: Pearsonův korelační koeficient nabývá hodnoty ..... . Mezi největší šířkou mozkovny a největší délkou mozkovny pro skelety mužského pohlaví existuje ..... stupeň .... závislosti.

Příklad 3.8. Dvourozměrný tečkový diagram
Pro znaky $X =$ *největší šířka mozkovny* a $Y =$ *největší délka mozkovny* u mužů vykreslete dvourozměrný tečkový diagram.

Řešení příkladu 3.8

```
105 par(mar = c(3, 4, 2, 2))
106 plot(skull.BM, skull.LM, pch = 21, bg = 'mintcream', col = 'darkblue',
107         xlab = '', ylab = 'nejvetsi delka mozkovny (mm)', main = '')
108 mtext('nejvetsi sirka mozkovny (mm)', side = 1, line = 2.1)
```