

## **Aplikovaná statistika I**

### *Téma 7: Úvod do testování hypotéz*

**Veronika Bendová, Zdeňka Geršlová**

# Úvod do testování hypotéz

- datový soubor = náhodný výběr → stanovíme předpoklady → ověřujeme, zda platí
- předpoklady
  - o rozdělení: normální, Poissonovo, binomické, ...
  - o parametrech:  $\mu$ ,  $\sigma^2$ ,  $\sigma$ ,  $\rho$ ,  $p$  ...

## Postup testování hypotéz

- literární řešerše, formulace problému ... přesná, jednoznačná
- stanovení nulové hypotézy  $H_0$ 
  - hypotéza o níž test rozhodne, zda se zamítne, nebo ne
  - příklad: jeden náhodný výběr a publikovaná hodnota  $c$ 
    - $H_{01} : \mu = c$ ;
    - $H_{02} : \mu \leq c$ ;
    - $H_{03} : \mu \geq c$ .
- stanovení alternativní hypotézy  $H_1$ 
  - pokud  $H_0$  zamítáme, pak ve prospěch alternativy
  - příklad: jeden náhodný výběr a publikovaná hodnota  $c$ 
    - $H_{11} : \mu \neq c$  (oboustranná alt.);
    - $H_{12} : \mu > c$  (pravostranná alt.);
    - $H_{13} : \mu < c$  (levostranná alt.).

- volba hladiny významnosti  $\alpha$ 
  - pst(riziko), že  $H_0$  zamítneme, když platí - snažíme se tuto hodnotu snížit na minimum
- provedení měření; sběr dat
- testování  $H_0$  (tři různé způsoby):
  - kritický obor
  - interval spolehlivosti
  - $p$ -hodnota
- rozhodnutí o zamítnutí/nezamítnutí  $H_0$
- interpretace výsledků

### Přístupy k testování hypotézy $H_0$

- Na hladině významnosti  $\alpha$  testujeme nulovou hypotézu  $H_{01}: \theta = c$ , případně  $H_{02}: \theta \leq c$ , či  $H_{03}: \theta \geq c$  oproti alternativní hypotéze  $H_{11}: \theta \neq c$ , případně  $H_{12}: \theta > c$ , či  $H_{13}: \theta < c$ .
  - **Test kritickým oborem**
    - stanovíme hodnotu testovací statistiky  $t_0$
    - stanovíme kritický obor  $W$
    - pokud  $t_0 \in W$ ,  $H_0$  **zamítáme** na hladině významnosti  $\alpha$
  - **Test intervalem spolehlivosti**
    - sestrojíme  $100(1 - \alpha)\%$  IS
    - pokud  $c \notin IS$ ,  $H_0$  **zamítáme** na hladině významnosti  $\alpha$
  - **Test  $p$ -hodnotou**
    - vypočítáme  $p$ -hodnotu  $p$
    - pokud  $p \leq \alpha$ ,  $H_0$  **zamítáme** na hladině významnosti  $\alpha$

## Test jednorozměrné normality

- normalita = nepostradatelný předpoklad parametrických testů (jednovýběrových, párových, dvouvýběrových, ...)
- $X_1, \dots, X_n$  je náhodný výběr z jednorozměrného rozdělení  $L(\theta)$
- stanovení hypotéz
  - $H_0$ : Data pochází z normálního rozdělení.
  - $H_1$ : Data nepochází z normálního rozdělení.
- testy normality
  - Shapirův-Wilkův test: `shapiro.test()` ...  $n \leq 30$
  - Andersonův-Darlingův test: `nortest::ad.test()` ...  $30 < n \leq 100$
  - Lillieforsův test: `nortest::lillie.test()` ...  $n > 100$
- výstup testů =  $p$ -hodnota:  $p > \alpha \rightarrow H_0$  nezamítáme;  $p \leq \alpha \rightarrow H_0$  **zamítáme**
- grafické ověření normality
  - histogram + křivka hustoty normálního rozdělení ... `hist()` + `dnorm()`
  - QQ diagram ... `qqnorm()` + `qqline()`

### Dataset: 11-two-samples-means-skull.txt

Datový soubor 11-two-samples-means-skull.txt obsahuje původní kranio-metrické údaje o basion-bregmatické výšce lebky u 215 dospělých mužů a 107 dospělých žen ze starověké egyptské populace. Data pochází z archivních materiálů (Schmitd, 1888).

#### Popis proměnných v datasetu:

- id ... pořadové číslo;
- pop ... populace (egant - egyptská starověká);
- sex ... pohlaví jedince (m - muž, f - žena);
- skull.H ... basion-bregmatická výška lebky (v mm).

#### Příklad 7.1. Test o normalitě dat

Načtěte datový soubor 11-two-samples-means-skull.txt. Na hladině významnosti  $\alpha = 0.05$  testujte hypotézu, že náhodný výběr basion-bregmatických výšek lebky (skull.H) mužů starověké egyptské populace pochází z normálního rozdělení.

#### Řešení příkladu 7.1

- $H_0$  : Data pochází / nepochází z normálního rozdělení.
- $H_1$  : Data pochází / nepochází z normálního rozdělení.

Hladina významnosti  $\alpha = \dots\dots\dots$

```

1 data <- read.delim(...) # nacteni datoveho souboru
2 skull.HM <- data[... , ...] # vyber sloupce skull.H pro muze
3 skull.HM <- na.omit(...) # odstraneni NA hodnot
4 n <- length(...) # pocet udaju o delce b-b vysky lebky u muzu
5 tab <- data.frame(n, min = min(...),
6                   max = max(...)) # souhrnna tabulka: n, min, a max

```

```

      n min max
1 215 119 146

```

7  
8

Náhodný výběr obsahuje údaje o basion-bregmatické výšce lebky ..... mužů starověké egyptské populace. Naměřené hodnoty se pohybují v rozmezí .....-..... mm. Protože rozsah náhodného výběru výšek lebky  $n \leq 30$  /  $30 < n \leq 100$  /  $n > 100$  použijeme na testování hypotézy o normalitě dat Shapirův-Wilkův / Andersonův-Darlingův / Lillieforsův test.

```

9 nortest::lillie.test(...) # Lillieforsuv test

```

```

      Lilliefors (Kolmogorov-Smirnov) normality test
data:  skull.HM
D = 0.054341, p-value = 0.1263

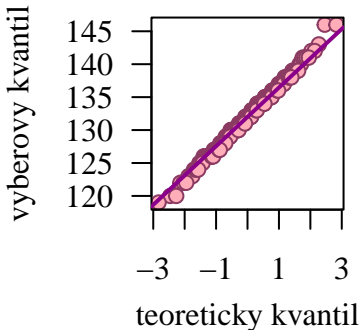
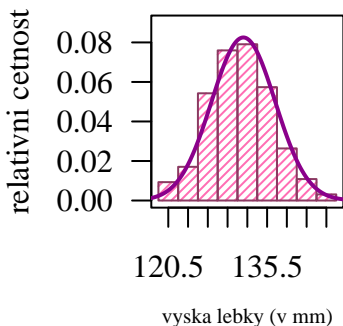
```

10  
11  
12  
13

Protože  $p$ -hodnota  $p = \dots\dots\dots$  je větší / menší než  $\alpha = 0.05$ ,  $H_0$  zamítáme / nezamítáme na hladině významnosti  $\alpha = 0.05$ .

## Grafická vizualizace rozdělení náhodného výběru

```
14 r <- ... # 9; 146 - 119 = 27 -> 27 -> seq(119, 149, by = 3)
15 hist(skull.HM, breaks = ..., prob = T, ylim = c(0, 0.09), axes = F, col = ...,
16     border = ..., density = ..., xlab = '', ylab = '', main = '') # histogram
17 axis(side = ..., at = seq(120.5, 144.5, by = 3)) # osa x (stredy trid. int.)
18 xfit <- seq(...) # posl. od 115 do 150 o delce 512
19 yfit <- dnorm(x = xfit, mean(skull.HM), sd(skull.HM)) # hustota f(x) N(m, s^2)
20 lines(xfit, yfit, col = ..., lwd = ...) # krivka hustoty f(x)
21 qqnorm(skull.HM, pch = ..., bg = ..., col = ..., main = '') # Q-Q diagram
22 qqline(skull.HM, col = ..., lwd = ...) # referencni primka
```



**Interpretace výsledků:** Náhodný výběr basion-bregmatických výšek lebky mužů starověké egyptské populace pochází / nepochází z normálního rozdělení.

### Dataset: 15-anova-means-skull.txt

Datový soubor 15-anova-means-skull.txt obsahuje původní kranioметриcké údaje o výšce horní části tváře mužů z německé, malajské, čínské, peruánské a bantuské populace. Data pochází z archivních materiálů (Schmidt, 1888).

#### Popis proměnných v datasetu:

- id ... pořadové číslo;
- pop ... populace (nem - německá, mal - malajská, cin - čínská, per - peruánská, ban - bantuská);
- sex ... pohlaví jedince (m - muž);
- upface.H ... výška horní části tváře, přímá vzdálenost mezi body *nasion* a *prosthion* (v mm).

#### Příklad 7.2. Test o normalitě dat

Načtěte datový soubor 15-anova-means-skull.txt. Na hladině významnosti  $\alpha = 0.10$  testujte hypotézu, že náhodný výběr výšek horní části tváře (upface.H) mužů německé populace pochází z normálního rozdělení.

#### Řešení příkladu 7.2

- $H_0$  : Data ..... z normálního rozdělení.
- $H_1$  : Data ..... z normálního rozdělení.

Hladina významnosti  $\alpha =$ .....



```
23 data <- read.delim(...) # nacteni datoveho souboru
24 upface.HN <- data[... , ...] # vyber sloupce upface.H pro muze nem. pop.
25 upface.HN <- ... # odstraneni NA hodnot
26 n <- ... # pocet udaju ve vektoru upface.HN
27 (tab <- data.frame(...)) # souhrnna tabulka: n, min, a max
```

```
      n min max
1 19  62  76
```

28  
29

Náhodný výběr obsahuje údaje o výšce horní části tváře ..... mužů německé populace. Naměřené hodnoty se pohybují v rozmezí .....-..... mm. Protože rozsah náhodného výběru výšek horní části tváře  $n \leq 30$  /  $30 < n \leq 100$  /  $n > 100$  použijeme na testování hypotézy o normalitě dat Shapirův-Wilkův / Andersonův-Darlingův / Lillieforsův test.

```
30 shapiro.test(...) # Shapiruv-Wilkuv test
```

```
      Shapiro-Wilk normality test
```

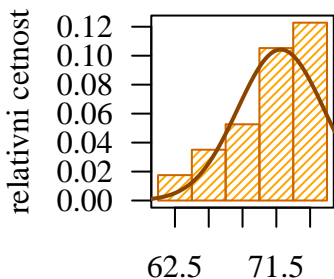
```
data:  upface.HN
W = 0.8964, p-value = 0.0419
```

31  
32  
33  
34

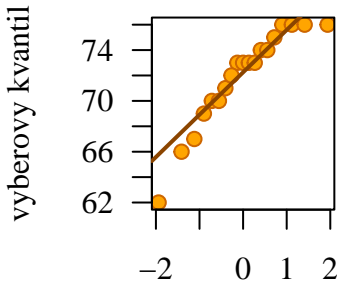
Protože  $p$ -hodnota  $p = \dots\dots\dots$  je větší / menší než  $\alpha = 0.10$ ,  $H_0 \dots\dots\dots$  na hladině významnosti  $\alpha = 0.10$ .

## Grafická vizualizace rozdělení náhodného výběru

```
35 r <- ... # 76 - 62 = 14 -> 15 -> seq(61, 76, by = 3)
36 par(mar = c(...)) # okraje 4, 4, 1, 1
37 hist(..., breaks = ..., prob = ..., ...) # histogra,
38 xfit <- seq(...) # posl. od 55 do 80 o delce 512
39 yfit <- dnorm(...) # hustota normalního rozdělení nad posl. xfit
40 lines(xfit, yfit, col = 'darkorange4', lwd = 2)
41 qqnorm(...) # Q-Q diagram
42 qqline(...) # referenční přímka
43 mtext(..., side = ..., line = ...) # popis osy x
```



vyska horni casti tvare (v mm)



teoreticky kvantil

**Interpretace výsledků:** Náhodný výběr výšek horní části tváře mužů německé populace pochází / nepochází z normálního rozdělení.

## Test dvourozměrné normality

- dvourozměrná normalita = nepostradatelný předpoklad parametrického testu o korelačním koeficientu  $\rho$
- $(X_1, Y_1)^T, \dots, (X_n, Y_n)^T$  je náhodný výběr z dvourozměrného rozdělení  $L(\theta)$
- stanovení hypotéz
  - $H_0$ : Data pochází z dvourozměrného normálního rozdělení.
  - $H_1$ : Data nepochází z dvourozměrného normálního rozdělení.
- testy normality
  - Mardiaův test: `MVN::mvn()` s argumentem `mvnTest = 'mardia'`
    - sestává ze dvou testů: (a) test šikmosti; (b) test špičatosti
  - (Henzův-Zirklerův test: `MVN::mvn()` s argumentem `mvnTest = 'hz'`)
  - (Roystonův test: `MVN::mvn()` s argumentem `mvnTest = 'royston'`)
- výstup testů =  $p$ -hodnota:  $p > \alpha \rightarrow H_0$  nezamítáme;  $p \leq \alpha \rightarrow H_0$  **zamítáme**
- grafické ověření normality
  - 3D-diagram ... `GA::persp3D()`
  - dvourozměrný tečkový diagram s  $(1 - \alpha)\%$  elipsou spolehlivosti ... `car::dataEllipse()`

### Dataset: 19-more-samples-correlations-skull.txt

Datový soubor 19-more-samples-correlations-skull.txt obsahuje údaje o šířce nosu a o interorbitální šířce mužů z německé, malajské, čínské, peruánské a bantuské populace. Data pochází z archivních materiálů (Schmid, 1888).

#### Popis proměnných v datasetu:

- pop ... populace (nem - německá, mal - malajská, cin - čínská, per - peruánská, ban - bantuská);
- sex ... pohlaví jedince (m - muž);
- nose.B ... šířka nosu (v mm);
- intorb.B ... interorbitální šířka (v mm).

#### Příklad 7.3. Test o dvourozměrné normalitě dat

Načtete datový soubor 19-more-samples-correlations-skull.txt. Necht' náhodná veličina  $X$  popisuje šířku nosu a náhodná veličina  $Y$  popisuje *interorbitální šířku* mužů peruánské populace. Na hladině významnosti  $\alpha = 0.05$  testujte hypotézu, že náhodný vektor  $(X, Y)^T$  pochází z dvourozměrného normálního rozdělení.

#### Řešení příkladu 7.3

- $H_0$  : Data pochází / nepochází z dvourozměrného normálního rozdělení.
- $H_1$  : Data pochází / nepochází z dvourozměrného normálního rozdělení.

Hladina významnosti  $\alpha = \dots\dots\dots$

Náhodný výběr obsahuje údaje o šířce nosu a interorbitální šířce ..... mužů peruánské populace. Hodnoty šířky nosu se pohybují v rozmezí .....-..... mm, hodnoty interorbitální šířky se pohybují v rozmezí .....-..... mm. Dvourozměrnou normalitu otestujeme Mardiovým testem. Ten sestává z dvou částí:

(a) z testu šikmosti

- $H_{0a}$  : Data vykazují / nevykazují kladné nebo záporné zešikmení.
- $H_{1a}$  : Data vykazují / nevykazují kladné ani záporné zešikmení.

(b) z testu špičatosti

- $H_{0b}$  : Data vykazují / nevykazují zešpičatění nebo zploštění.
- $H_{1b}$  : Data vykazují / nevykazují zešpičatění ani zploštění.

Data pochází z dvourozměrného normálního rozdělení, pokud nevykazují zešikmení ani zešpičatění.

44

```
MVN::mvn(data.P, mvnTest = 'mardia')$multivariateNormality
```

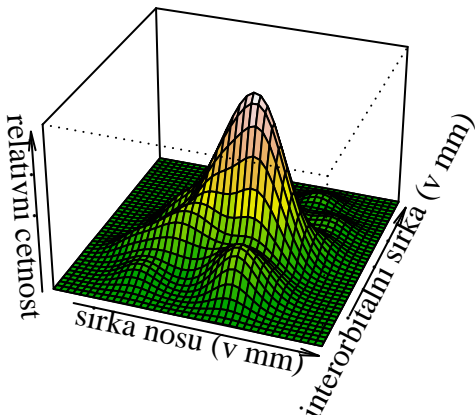
	Test	Statistic	p value	Result
1	Mardia Skewness	4.27819772855481	0.369663150730262	YES
2	Mardia Kurtosis	-0.0684871107744411	0.945397880096616	YES
3	MVN	<NA>	<NA>	YES

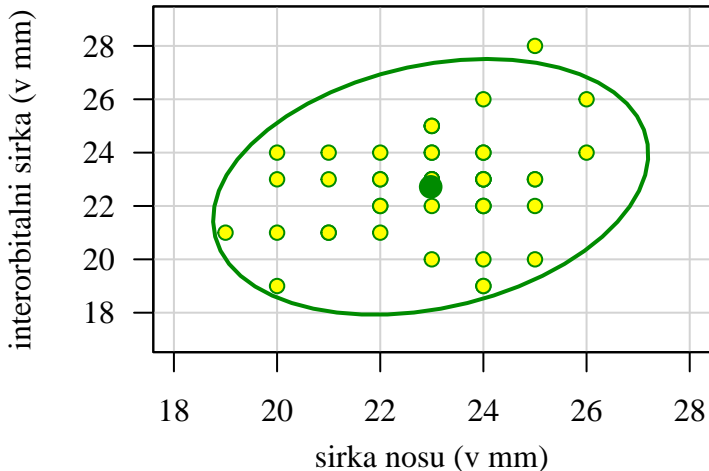
45  
46  
47  
48

Protože  $p$ -hodnota testu šikmosti  $p = \dots\dots\dots$  je menší / větší než  $\alpha = 0.05$ ,  $H_{0a}$  zamítáme / nezamítáme na hladině významnosti  $\alpha = 0.05$ . Data vykazují / nevykazují výrazné zešikmení. Protože  $p$ -hodnota testu špičatosti  $p = \dots\dots\dots$  je menší / větší než  $\alpha = 0.05$ ,  $H_{0b}$  zamítáme / nezamítáme na hladině významnosti  $\alpha = 0.05$ . Data vykazují / nevykazují výrazné zešpičatění nebo zploštění.

## Grafická vizualizace rozdělení náhodného výběru

```
49 Z <- MASS::kde2d(nose.BP, intorb.BP, n = 40,  
50                 lims = c(17, 28, 17, 30)) # jadrovy odhad hustoty  
51 GA::persp3D(Z$x, Z$y, Z$z, theta = 5, phi = 30, col.palette = terrain.colors,  
52             ticktype = ..., xlab = ..., ylab = ..., zlab = ...) # 3D-diagram  
53 car::dataEllipse(nose.BP, intorb.BP, level = 0.95, xlim = c(18, 28), las = ...,  
54                 ylim = c(17, 29), xlab = ..., ylab = ..., main = ...,  
55                 bg = ..., pch = ..., col = ...) # 95% elipsa spolehlivosti
```





95% elipsu spolehlivosti leží ..... z celkového počtu ..... pozorování, tj. .... %, tedy více / méně než 5 % dat.

**Interpretace výsledků:** Náhodný výběr šířky nosu a interorbitalní šířky mužů peruánské populace pochází / nepochází z dvourozměrného normálního rozdělení.