6 Bodové a intervalové odhady

Dataset: 21-goldman-tigara.csv
Datový soubor 21-goldman-tigara.csv obsahuje osteometrické údaje o délce stehenní kosti (v mm) a acetabulární výšce (v mm) z pravé a levé strany u mužů a žen z aljašské populace z kmene Tigara. Data pochází ze souboru dokumentovaných skeletů (Goldman, 2006).

Popis proměnných v datasetu:

- sex . . . pohlaví jedince (m - muž, f - žena);
- pop . . . populace (Tigara = aljašská populace z kmene Tigara);
- femur.LR . . . délka stehenní kosti z pravé strany (v mm);
- femur.LL . . . délka stehenní kosti z levé strany (v mm);
- acetab.HR . . . acetabulární výška z pravé strany (v mm);
- acetab.HL . . . acetabulární výška z levé strany (v mm).

Příklad 6.1. Bodové odhady parametrů μ a σ^2 normálního rozdělení
Načtěte datový soubor 21-goldman-tigara.csv. Nechť náhodná veličina X popisuje *délku stehenní kosti* (v mm) z pravé strany u mužů z kmene Tigara. Za předpokladu, že náhodná veličina $X \sim N(\mu, \sigma^2)$, stanovte nestranný (bodový) odhad (a) střední hodnoty μ; (b) rozptylu σ^2; (c) směrodatné odchylky σ; (d) koeficientu variace $\frac{\sigma}{\mu}$. Všechny vypočítané hodnoty řádně interpretujte.

Řešení příkladu 6.1

```
         m       s2       s      v
1 427.8958 539.1735 23.2201 0.0543
```

Interpretace výsledků: Nestranný odhad střední hodnoty μ je .... mm. Nestranný odhad rozptylu σ^2 (resp. směrodatné odchylky σ) je .... mm^2 (resp. .... mm). Nestranný odhad koeficientu variace je .... Délka stehenní kosti z pravé strany u mužů z kmene Tigara se pohybuje okolo hodnoty .... mm se směrodatnou odchylkou .... mm. Směrodatná odchylka představuje .... % aritmetického průměru. ★

Příklad 6.2. Bodové odhady parametrů $\boldsymbol{\mu}$ a $\boldsymbol{\Sigma}$ normálního rozdělení
Načtěte datový soubor 21-goldman-tigara.csv. Nechť náhodná veličina X popisuje *délku stehenní kosti* (v mm) z pravé strany a náhodná veličina Y popisuje *acetabulární výšku* (v mm) z pravé strany u mužů z kmene Tigara. Za předpokladu, že náhodný vektor $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, stanovte (a) nestranný (bodový) odhad vektoru středních hodnot $\boldsymbol{\mu}$; (b) nestranný (bodový) odhad kovariance σ_{12}; (c) asymptoticky nestranný (bodový) odhad korelačního koeficientu ρ; (d) nestranný (bodový) odhad varianční matice $\boldsymbol{\Sigma}$.

Řešení příkladu 6.2

```
   m.LR  m.HR  s2.LR s2.HR   s.LR s.HR   s12  r12
1 427.9 51.93 539.17  9.77 23.22 3.13 42.11 0.58
```

Interpretace výsledků: Nestranný odhad vektoru středních hodnot $\boldsymbol{\mu}$ =(......................... , ...)^T mm. Nestranný odhad kovariance σ_{12} je .... Asymptoticky nestranný odhad korelačního koeficientu ρ je .... Nestranný odhad varianční matice $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$ je matice $\begin{pmatrix} s_1^2 & s_{12} \\ s_{12} & s_2^2 \end{pmatrix}$, kde s_1^2 =........................... mm^2, s_2^2 = ... mm^2 a $s_{12} = r_{12}s_1s_2$ = ... Délka stehenní kosti z pravé strany mužů z kmene Tigara se pohybuje okolo hodnoty .... mm se směrodatnou odchylkou .... mm. Acetabulární výška z pravé strany se pohybuje okolo hodnoty .... mm se směrodatnou odchylkou .... mm. Mezi délkou stehenní kosti a acetabulární výškou z pravé strany existuje .... stupeň .... závislosti (r_{12} = ...). ★

Dataset: 21-goldman-shells.csv
Datový soubor 21-goldman-shells.csv obsahuje osteometrické údaje o délce kyčelní kosti z pravé a levé strany u mužů a žen ze tří japonských populací(Tsugumo Shell Mound, Yoshigo Shell Mound a Yasaki Shell Mound). Data pochází ze souboru dokumentovaných skeletů (Goldman, 2006).

Popis proměnných v datasetu:

- sex ... pohlaví jedince (m - muž, f - žena);
- pop ... populace (tsg = Tsugumo Shell Mound, yos = Yoshigo Shell Mound, yas = Yasaki Shell Mound);
- iblade.LR ... délka kyčelní kosti z pravé strany (v mm);
- iblade.LL ... délka kyčelní kosti z levé strany (v mm).

Příklad 6.3. Dvouvýběrové statistiky
Načtěte datový soubor 21-goldman-shells.csv. Vypočítejte vážený průměr výběrových rozptylů délek kyčelních kostí z levé strany u mužů (a) z populací Yoshigo Shell Mound a Yasaki Shell Mound; (b) ze všech tří uvedených populací.

Řešení příkladu 6.3

```
  sh.YoYa sh.TgYoYa
1    16.5  24.34722
```

Interpretace výsledků: Vážený průměr výběrových rozptylů délek kyčelních kostí z levé strany mužů z populací Yoshigo Shell Mound a Yasaki Shell Mound s^2_{YoYa*} = ... mm^2. Vážený průměr výběrových rozptylů délek kyčelních kostí z levé strany mužů všech tří japonských populací s^2_* = ... mm^2. ★

Příklad 6.4. Intervalové odhady parametrů normálního rozdělení
Načtěte datový soubor 21-goldman-tigara.csv. Nechť náhodná veličina X popisuje *délku stehenní kosti* (v mm) z pravé strany u mužů z kmene Tigara. Za předpokladu, že náhodná veličina $X \sim N(\mu, \sigma^2)$, stanovte (a) 95% intervalový odhad střední hodnoty μ; (b) 99% levostranný intervalový odhad rozptylu σ^2; (c) 90% pravostranný intervalový odhad směrodatné odchylky σ.

Řešení příkladu 6.4

```
   dh.mu hh.mu D.sig2 H.sig
1 418.09 437.7 297.83  28.9
```

Interpretace výsledků: 95% empirický interval spolehlivosti pro střední hodnotu μ má tvar .... To znamená, že .... $< \mu <$ .... s pravděpodobností 95 %. V 95 případech ze sta bude střední hodnota délky stehenní kosti z pravé strany u mužů z kmene Tigara nabývat hodnoty z intervalu .... mm.

99% levostranný empirický interval spolehlivosti pro rozptyl σ^2 má tvar .... To znamená, že $\sigma^2 >$ .... s pravděpodobností 99 %. V 99 případech ze sta bude rozptyl délky stehenní kosti z pravé strany u mužů z kmene Tigara větší / menší než .... mm.

90% pravostranný empirický interval spolehlivosti pro směrodatnou odchylku σ má tvar .... To znamená, že $\sigma <$ .... s pravděpodobností 90 %. V 90 případech ze sta bude směrodatná odchylka délky stehenní kosti z pravé strany u mužů z kmene Tigara větší / menší než .... mm.
★

Příklad 6.5. Bodový a intervalový odhad parametru p alternativního rozdělení
Načtěte datový soubor 17-anova-newborns-2.txt. Mějme náhodnou veličinu X popisující *ženské pohlaví* novorozenců. Za předpokladu, že náhodná veličina $X \sim \mathrm{Alt}(p)$, kde p je pravděpodobnost narození holčičky, stanovte (a) bodový odhad parametru p; (b) 95% intervalový odhad parametru p.

Řešení příkladu 6.5

```
       p  dh.p   hh.p
1 0.4794 0.453 0.5057
```

Interpretace výsledků: Bodový odhad pravděpodobnosti narození holčičky je .... K narození holčičky dojde s pravděpodobností .... %. 95% empirický IS pro pravděpodobnost narození holčičky p má tvar .... To znamená, že .... $< p <$ .... s pravděpodobností 95 %. Pravděpodobnost narození holčičky se pohybuje v rozmezí ....% – ....% s pravděpodobností 95 %. ★