6 Číselné charakteristiky, Matematická statistika, Bodové a intervalové odhady parametrů

6.1 Výpočet číselných charakteristik náhodných veličin

α-kvantily

Příklad 6.1. Najděte medián a horní a dolní kvartil náhodné veličiny $U \sim N(0,1)$. $0; -0.6745; 0.6745$

- Pravděpodobnost, že náhodná veličina U bude nabývat hodnoty menší nebo rovné .... je ....
- Pravděpodobnost, že náhodná veličina U bude nabývat hodnoty menší nebo rovné .... je ....
- Pravděpodobnost, že náhodná veličina U bude nabývat hodnoty menší nebo rovné .... je ....

Příklad 6.2. Najděte dolní kvartil náhodné veličiny $X \sim N(3,5)$. 1.491795

- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....

Příklad 6.3. Určete kvantil $\chi^2_{0.025}(25)$. 13.11972

- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....

Příklad 6.4. Určete kvantily $t_{0.99}(30)$ a $t_{0.05}(14)$. $2.457262; -1.76131$

- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....
- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....

Příklad 6.5. Určete kvantily $F_{0.975}(5,20)$ a $F_{0.05}(2,10)$. $3.289056; 0.0515573$

- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....
- Pravděpodobnost, že náhodná veličina X bude nabývat hodnoty menší nebo rovné .... je ....

6.2 Základní pojmy matematické statistiky

- Statistiky – jednovýběrové:
  Nechť $X_1, \ldots, X_n$ je náhodný výběr, $n \geq 2$.

  1. *Výběrový průměr*
  $$M = \frac{1}{n}\sum_{i=1}^{n} X_i$$
  2. *Výběrový rozptyl*
  $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M)^2$$
  3. *Výběrová směrodatná odchylka*
  $$S = \sqrt{S^2}$$
  4. *Výběrová distribuční funkce* $F_n(x)$ ... průměrný počet těch veličin X_i, pro něž platí $X_i \geq x$.

- Statistiky – dvouvýběrové:
  Nechť $(X_1, Y_1), \ldots (X_n, Y_n)$ je náhodný výběr z dvourozměrného rozdělení. M_1, M_2 jsou výběrové průměry a S_1, S_2 jsou výběrové směrodatné odchylky.

  1. *Výběrová kovariance*

  $$S_{12} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - M_1)(Y_i - M_2)$$

  2. *Výběrový koeficient korelace*

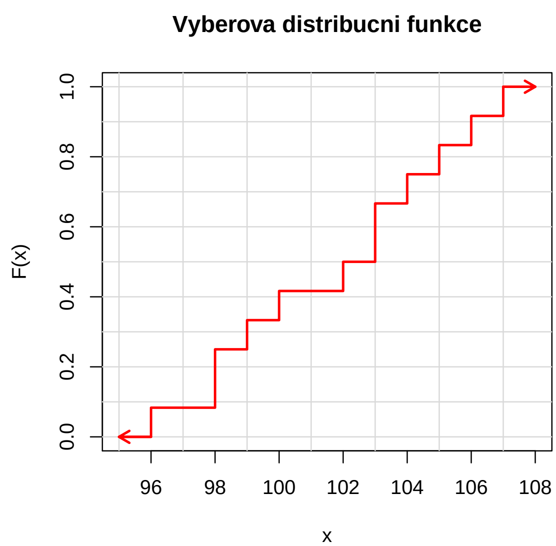
  $$R_{12} = \frac{S_{12}}{S_1 S_2}$$

6.3 Bodové a intervalové odhady parametrů

Příklad 6.6. Ve 12-ti náhodně vybraných internetových obchodech byly zjištěny následující ceny deskriptoru artefaktů (v Kč): $102, 99, 106, 103, 96, 98, 100, 105, 103, 98, 104, 107$. Těchto 12 hodnot považujeme za realizace náhodného výběru $X_1, \ldots, X_{12}$ z rozdělení, které má střední hodnotu μ a rozptyl σ^2.

a) Určete nestranné bodové odhady neznámé střední hodnoty μ a neznámého rozptylu σ^2.

b) Najděte výběrovou distribuční funkci $F_{12}(x)$ a nakreslete její graf.



Příklad 6.7. Z archivních materiálů (Schmidt, 1888) máme k dispozici původní kraniometricé údaje o výšce horní části tváře (v mm) u 13 mužů bantuské populace. Hodnoty výšky horní části tváře jsou 67, 67, 63, 68, 70, 70, 75, 74, 80, 77, 77, 67, 64.

a. Odhadněte střední hodnotu, rozptyl a směrodatnou odchylku výšky horní části tváře.

b. Odhadněte pravděpodobnost že výška tváře bantuského muže bude vyšší než 72 mm.

```
           m    s2    s
akcie 101.75 29.06 5.39
[1] 0.3846
```

ad a. Hodnota výběrového průměru výšky horní části tváře je ....mm s rozptylem ....mm a směrodatnou odchylkou ....mm.

ad b. Odhad pravděpodobnosti že výška tváře bantuského muže bude vyšší než 72 mm je ....

Příklad 6.8. Máme k dispozici antropometrické údaje mladých dospělých lidí, převážně studentů vysokých škol z Brna a Ostravy, konkrétně údaje o šířce hlavy (head.W), šířce tváře (bizyg.W) a šířce dolní čelisti (bigo.W). Dále máme u každého studenta uveden údaj o pohlaví (sex), přičemž v databázi máme celkem 75 mužů a 100 žen. Zaměřme se na údaje týkající se mužů. Najděte bodové odhady kovariance $\sigma_{1,2}$ a korelace ρ pro náhodné proměnné X_1 ... šířka hlavy a X_2 ... šířka tváře.

```
[1] 31.83279
[1] 0.6785296
```

Hodnota výběrové kovariance mezi šířkou hlavy a šířkou tváře mužů je ....mm, hodnota výběrového korelačního koeficientu je ....mm, což značí na význačný stupeň přímé lineární závislosti.

INTERVALY SPOLEHLIVOSTI

a. $X_1 \dots X_n$ ... náh.výběr z rozdělení $L(\theta)$, θ je parametr, $\alpha \in (0,1)$

b. interval (D,H) ... $100(1-\alpha)\%$ oboustranný IS pro param. θ

c. interval (D,∞)... $100(1-\alpha)\%$ levostranný IS pro param. θ

d. interval $(-\infty,H)$... $100(1-\alpha)\%$ pravostranný IS pro param. θ

e. α se nazývá *riziko*, $(1-\alpha)$ se nazývá *spolehlivost*.

Příklad 6.9. Vezměte data z příkladu 7.3. Vypočítejte

a. 95 % empirický interval spolehlivosti pro střední hodnotu délky šířky čelisti u mužů.

```
[1] 106.2321
[1] 109.3946
```

b. 90 % pravostranný empirický interval spolehlivosti pro střední hodnotu šířky dolní čelisti u mužů.

```
[1] 108.8395
```

c. 99 % levostranný empirický interval spolehlivosti pro střední hodnotu šířky dolní čelisti u mužů.

```
[1] 105.9263
```

ad a. .... % empirický IS pro střední hodnotu šířky dolní čelisti u mužů má tvar .... To znamená, že .... $< \mu <$ .... s pravděpodobností ....% .

ad b. .... % pravostranný empirický IS pro střední hodnotu šířky dolní čelisti u mužů má tvar .... To znamená, že $\mu >$ .... s pravděpodobností ....% .

ad c. .... % levostranný empirický IS pro střední hodnotu šířky dolní čelisti u mužů má tvar .... To znamená, že $\mu <$ .... s pravděpodobností ....% .

Příklad 6.10. Při kontrolních zkouškách životnosti 16 žárovek byl stanoven odhad $m = 3000\,h$ střední hodnoty jejich životnosti. Z dřívějších zkoušek je známo, že životnost žárovky se řídí normálním rozdělením se směrodatnou odchylkou $\sigma = 20\,h$. Vypočtěte

a) 99 % empirický interval spolehlivosti pro střední hodnotu životnosti (2987.1; 3012.9);

```
[1] 2987.121
[1] 3012.879
```

b) 90 % levostranný empirický interval spolehlivosti pro střední hodnotu životnosti $(29993.6; \infty)$;

```
[1] 2993.592
```

c) 95 % pravostranný empirický interval spolehlivosti pro střední hodnotu životnosti $(-\infty; 3008.2)$.

```
[1] 3008.224
```

ad a. .... % empirický IS pro střední hodnotu životnosti žárovek má tvar .... To znamená, že .... $< \mu <$ .... s pravděpodobností ....% .

ad b. .... % pravostranný empirický IS pro střední hodnotu životnosti žárovek má tvar .... To znamená, že $\mu >$ .... s pravděpodobností ....% .

ad c. .... % levostranný empirický IS pro střední hodnotu životnosti žárovek má tvar .... To znamená, že $\mu <$ .... s pravděpodobností ....% .