

## Časté chyby v riešeniach domácej úlohy z NMSA331

M. OMEKA A S. NAGY, 2022/2023

### Základné pravidlá.

- Odovzdávaný pdf dokument musí byť pomenovaný. Na úvodnej strane musí začínať **menom** riešiteľa, nadpisom, a **kódom** použitým v analýze (vaše študentské číslo UKČO). Bez kódu nie sme schopní replikovať výsledky analýzy.
- Popis dát neznamená iba skopírovanie **summary** tabuľky z R. Celkove, R kód nie je **nikdy** výsledkom analýzy. Výsledkom analýzy je vaša správa, v ktorej slovne a matematicky popíšete to, čo ste použitím R získali. Kódy môžete pridať ako prílohu k správe, všetky výsledky ale musia byť zrozumiteľne popísané v texte.

### Popis dát (Časť 1).

- Popis dátového súboru nie je iba uvedenie základných sumárnych štatistík a obrázkov (histogram, krabicový diagram). Dôležitý je slovný popis dát, čo nemôže byť iba prepísanie údajov z tabuliek a obrázkov slovami. Popis typu „*medián našich pozorovaní je 1,54*“ je úplne zbytočný, ak toto číslo už vidíme v tabuľke. Rozumnejšie je celkové zhodnotenie dát. Napríklad:
  - Jedná sa o náhodný výber?
  - Zo spojitých, alebo diskrétnej náhodnej veličiny?
  - Nadobúda náhodná veličina všetky hodnoty v  $\mathbb{R}$ , alebo v nejakom obmedzenom intervale?
  - Existujú problémy so zaokrúhľovaním, alebo chýbajúcimi pozorovaniami?
  - Môže dátový súbor obsahovať chybné pozorovania (tj. chyby merania alebo iné podozrivé hodnoty)?
  - Môže ísť o náhodný výber z normálneho rozdelenia, alebo je rozdelenie zrejme nenormálne (je napr. zošikmené, alebo má iné vlastnosti ktoré ho odlišujú od normálneho)?

### Štatistické testy a odhady (Časti 2–5).

- Pri testoch a intervaloch spoľahlivosti vždy očakávame:
  - (i) zhodnotenie toho, či sa jedná o náhodný výber;
  - (ii) pravdepodobnostný model, ktorý uvažujete a jeho zdôvodnenie, napr. diagnostický obrázok (alebo obrázky) s ich vysvetlením;
  - (iii) špecifikáciu parametra ktorý testujeme alebo odhadujeme, určenie nulovej a alternatívnej hypotézy;
  - (iv) vzorec pre výpočet testovej štatistiky a rozdelenie štatistiky za nulovej hypotézy;
  - (v) jedná sa o presný alebo asymptotický test?
  - (vi) kritický obor a p-hodnota testu, vrátane vzťahov pre ich výpočet;
  - (vii) interval spoľahlivosti pre skúmaný parameter;
  - (viii) pozorovaná hodnota testovej štatistiky a pozorovaná hodnota intervalu spoľahlivosti;

- (ix) rozhodnutie — môžeme zamietnuť nulovú hypotézu?
- (x) interpretácia testu a záver;
- (xi) prípadná diskusia o tom, aký vplyv na výsledok testu môže mať porušenie predpokladov (prípadná nenormalita, neexistencia momentov, zaokrúhľovanie, šikmosť rozdelenia apod.).

V žiadnom prípade nestačí vypísať hodnoty získané z programu R bez ďalšieho komentára.

### Úprava textu.

- Popisky tabuliek a obrázkov sú celé vety. Musia teda začínať veľkými písmenami, a končiť bodkami. Popisky musia byť zrozumiteľné aj samotne, tj. bez toho aby čitateľ musel prechádzať text dokumentu. Popisok obrázku „*histogram*“ je teda zbytočný, omnoho lepšie je napísať „*Histogram výsledného hodnotenia filmov. Porovujeme výrazné odchýlky od normality.*“
- Pri percentách záleží na medzere. 50 % znamená 50 percent, 50% znamená 50-percentný.
- V českom texte by sa mala používať čeština aj v obrázkoch a tabuľkách. Obrázky by mali byť dostatočne veľké, a čitateľné. V R je možné veľkosť písma nastaviť voľbou `cex`, napr.

```
hist(X,cex.main=3,cex.lab=2,cex.axis=2.5).
```

- Ak niečo vykreslíte, vždy uveďte, o čo sa jedná. Napr., ak vykreslíte cez histogram krivku normálneho rozdelenia, vždy komentujte prečo túto krivku vykresľujete, o aké normálne rozdelenie sa jedná, a čo z porovnania krivky a histogramu usudzujete.

### Formátovanie textu v L<sup>A</sup>T<sub>E</sub>Xu.

- Používajte vlnku pre nezalomiteľné medzery. Takéto medzery sa nikdy nerozdelia na konci riadku. Patria napr. k výrazom typu 1 203 Kč (píšeme 1~203~Kč), tab. 9, rovnica (8).
- Ak po rovnici pokračujete v texte, nenechávajte v L<sup>A</sup>T<sub>E</sub>Xu voľný riadok. Porovnajme

$$x = 0$$

kde som nechal voľný riadok, a oproti tomu

$$x = 0.$$

Tu som riadok nevynechal. Všimnite si odsadenie textu.