

CVIČENÍ 8

10. duben 2017

Cvičení 1. Ze stránky <http://www.statsci.org/data/general/fullmoon.txt> získejte data `fullmoon`, zopakujte si, co znamenají jednotlivé proměnné a jaké jsou mezi nimi vztahy. Vytvořte novou kategoriální proměnnou `Fullmoon` udávající, jestli se jedná o úplňk nebo ne.

Cvičení 2. Zkoumejte závislost počtu pacientů na úplňku měsíce a na měsíci v roce.

(a) Uvažujte nejprve model

```
> model.month.and.moon <- lm(Admission~Month + Moon, data=fullmoon)
```

Interpretujte jeho parametry a testujte vlivy jednotlivých faktorů pomocí vhodných tabulek ANOVA.

(b) Porovnejte výsledky získané na základě modelu `model.month.moon` s těmi, které jste získali z modelů

```
> model.moon <- lm(Admission~Moon, data=fullmoon)
>
> model.month <- lm(Admission~Month, data=fullmoon)
```

(c) Zdálo by se vám rozumné uvažovat v modelu `model.month.moon` interakční člen?

Cvičení 3. Pokračujte ve zkoumání závislost počtu pacientů na úplňku měsíce a na měsíci v roce.

(a) Uvažujte nyní model

```
> model.month.times.fullmoon <- lm(Admission~Month * Fullmoon, data=fullmoon)
```

Interpretujte jeho výsledky (také za pomoci vhodné grafiky) a testujte vlivy jednotlivých faktorů a kombinací faktorů pomocí vhodných tabulek ANOVA.

(b) Porovnejte výsledky získané na základě modelu `model.month.times.fullmoon` s těmi, které získáte z modelů

```
> model.month.and.fullmoon <- lm(Admission~Month + Fullmoon, data=fullmoon)
>
> model.fullmoon <- lm(Admission~Fullmoon, data=fullmoon)
>
> model.month <- lm(Admission~Month, data=fullmoon)
```

(c) Všimněte si (ne)závislosti výsledků v tabulkách ANOVA na volbě kontrastů a na pořadí, ve kterém jednotlivé faktory vstupují do definice lineárního modelu.

(d) Řešte část (c) pro případ, že z dat vynecháte první pozorování, t.j. místo dat `fullmoon` použijte data `fullmoon[-1,]`.

Cvičení 4.

- (a) Uvažujte náhodné jevy A_1, \dots, A_k . Omezte pravděpodobnost sjednocení $\bigcup_{i=1}^k A_i$ shora a odvoďte platnost Bonferroniho korekce p -hodnot při mnohonásobných porovnáváních.
- (b) Uvažujte nezávislé náhodné jevy A_1, \dots, A_k . Spočítejte pravděpodobnost průniku $\bigcap_{i=1}^k A_i$ a odvoďte platnost Šidákovy korekce p -hodnoty při mnohonásobných porovnáváních.

Cvičení 5.

- (a) Zkoumejte rozdíly mezi počty pacientů pro jednotlivé kombinace fáze měsíce a měsíce v roce.
- (b) Rozmyslete si využití porovnávání hlavních efektů v dvoufaktorovém ANOVA modelu s interakcemi a bez interakcí.
- (c) Řešte část (a) pomocí Bonferroniho metody.

Domácí úloha (8 bodů)

Uvažujte obecně (t.j. ne pro konkrétní data) dvoufaktorový ANOVA model s interakcemi:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i \in \{1, \dots, I\}, \quad j \in \{1, \dots, J\}, \quad k \in \{1, \dots, n_{ij}\}.$$

Zvolíte-li v R pomocí parametru `contrasts` parametrizaci pro oba hlavní efekty, tato volba implikuje parametrizaci pro interakce. Jakou parametrizaci pro interakce dostanete, vyberete-li pro oba hlavní efekty v parametru `contrasts` volbu `contr.sum`?