

## CVIČENÍ 5

20. březen 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. Nafitujte v R model  $\mathcal{F}_{H_1}$  pro závislost počtu pacientů pohotovostní služby psychiatrické kliniky na měsíci v roce (nazvěme jej `model.year`). Zopakujte si interpretaci parametrů  $\beta_0, \beta_1, \dots, \beta_{11}$ . Nafitujte stejný model s volbou `contrasts=contr.sum` a interpretujte jeho parametry.

**Cvičení 2.** Testujte v modelu `model.year` hypotézu, že střední hodnota počtu pacientů pohotovostní služby psychiatrické kliniky nezávisí na měsíci v roce.

**Cvičení 3.** Zamítneme-li hypotézu o nezávislosti počtu pacientů na měsíci v roce, zajímá nás, které měsíce nebo období se liší od jiných.

- Co o tom soudíte na základě výstupu funkce `summary`?
- Zkonstruujte konfidenční interval pro rozdíl středních hodnot ve dvou pozorovaných rocích (data jsou za období srpen 1971 až červenec 1972) a testujte významnost tohoto rozdílu.
- Zkonstruujte konfidenční intervaly pro rozdíly středních hodnot v jednotlivých ročních obdobích a celoroční střední hodnoty, a testujte významnost těchto rozdílů.
- Jakým problémem trpí tyto konfidenční intervaly a testy?

**Cvičení 4.** Zaměříme se nyní na párová porovnávání (rozdíly mezi středními hodnotami pro jednotlivé dvojice měsíců). Zkonstruujte konfidenční intervaly a testy pro rozdíly středních hodnot pomocí

- (jednorozměrných) t-testů,
- Fisherovy LSD metody,
- Scheffého metody,
- Tukeyho HSD metody.

Porovnejte výsledné intervaly graficky a uvažte výhody a nevýhody jednotlivých přístupů.

*Tip: Můžete si pomoci přednáškou z týdne 7 z podzimního semestru a přednáškami z jarního semestru.*

**Domácí úloha** (10 bodů)Uvažujte model  $\mathcal{F}_{H_1}$ :

$$Y_{ji} = \mu_j + \varepsilon_{ji}, \quad j = 1, 2, \dots, J; \quad i = 1, 2, \dots, n_j.$$

Vyjádřete parametry  $\beta_0, \beta_1, \dots, \beta_{J-1}$ , odhady kterých vidíte na výstupu funkce `summary` v R, zvolíte-li ve funkci `lm` možnost `contrasts=contr.helmert`, jako funkce skupinových středních hodnot  $\mu_1, \dots, \mu_J$ .

Jakou možnost pro `contrasts` byste museli zvolit ve funkci `lm`, pokud byste chtěli na výstupu funkce `summary` vidět

$$\begin{aligned} \beta_0 &= \mu \\ \beta_1 &= \mu_2 - \mu_1 \\ \beta_2 &= 2\mu_3 - (\mu_2 + \mu_1) \\ \beta_3 &= 3\mu_4 - (\mu_3 + \mu_2 + \mu_1) \\ &\dots \\ \beta_{J-1} &= (J-1)\mu_J - (\mu_{J-1} + \mu_{J-2} + \dots + \mu_1)? \end{aligned}$$