

CVIČENÍ 3

6. březen 2017

Cvičení 1. Ze stránky <http://www.statsci.org/data/general/fullmoon.txt> získejte data `fulmoon`, zopakujte si, co znamenají jednotlivé proměnné a jaké jsou mezi nimi vztahy. Nafitujte v R model `model.0` ze cvičení z minulého týdne (model \mathcal{F}_{H_0} z přednášky), t.j.

$$Y_i = \mu + \varepsilon_i, \quad i = 1, \dots, N \quad (1)$$

kde $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$, a model `model.1` ze cvičení z minulého týdne (model \mathcal{F}_{H_1} z přednášky), t.j.

$$Y_{ji} = \mu + \alpha_j + \varepsilon_{ji}, \quad j = 1, \dots, J; \quad i = 1, \dots, n_j, \quad (2)$$

kde $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$.

Cvičení 2. Zopakujte si následující definice a vztahy z podzimního semestru:

$$\begin{aligned} TSS &= \sum_{i=1}^n (Y_i - \bar{Y})^2 = \|\mathbf{Y} - \bar{Y} \cdot \mathbf{1}\|^2 \\ ESS &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \|\hat{\mathbf{Y}} - \bar{Y} \cdot \mathbf{1}\|^2 \\ RSS &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 = \|\mathbf{e}\|^2 \\ TSS &= ESS + RSS \\ R^2 &= 1 - \frac{RSS}{TSS} \\ R_{adj}^2 &= 1 - \frac{RSS/(n-p)}{TSS/(n-1)} \end{aligned}$$

Vzorce jsou zavedeny a odvozeny v přednášce z týdne 5 z podzimního semestru.

$$F = \frac{(\|\mathbf{e}_s\|^2 - \|\mathbf{e}_b\|^2)/r}{\|\mathbf{e}_b\|^2/(n-p)} \sim F_{r, n-p}$$

Značení je zavedeno a vzorec je odvozen v přednášce z týdne 9 z podzimního semestru.

Cvičení 3. Uvažujte model \mathcal{F}_{H_0} z přednášky, t.j. model (1).

- Odvoďte vzorec pro $\hat{\beta}$ (t.j. pro $\hat{\mu}$) v modelu \mathcal{F}_{H_0} .
- Odvoďte vzorec pro ESS a RSS v modelu \mathcal{F}_{H_0} .
- Odvoďte vzorec pro R^2 a R_{adj}^2 v modelu \mathcal{F}_{H_0} .
- Dejte odvozená fakta do souvislosti se výstupem `summary(model.0)`.

Cvičení 4. Uvažujte model \mathcal{F}_{H_1} z přednášky, t.j. model (2).

- Odvoďte vzorec pro $\hat{\beta}$ (t.j. pro $(\hat{\mu}, \hat{\alpha}_1, \dots, \hat{\alpha}_J)^\top$ za podmínky, že $\sum_{j=1}^J n_j \alpha_j = 0$) v modelu \mathcal{F}_{H_1} .

Můžeme si pomoci znalostmi z přednášky z týdne 11 z podzimního semestru.

- Odvoďte vzorec pro ESS a RSS v modelu \mathcal{F}_{H_1} .
- Odvoďte vzorec pro F statistiku na testování modelu \mathcal{F}_{H_0} proti modelu \mathcal{F}_{H_1} .

Cvičení 5. Spočtete kvantily ze Cvičení 3 a 4 pro data `fullmoon` a najděte spočtené hodnoty ve výstupech z funkcí

- (a) `summary(model.0)`;
- (b) `summary(model.1)`;
- (c) `anova(model.1)`;
- (d) `anova(model.0, model.1)`;
- (e) `aov`;
- (f) `oneway.test`.

Domácí úloha (12 bodů) Uvažujte model \mathcal{F}_{H_1} pro data `fullmoon`. Následující tabulka udává odhad koeficientu, příslušnou směrodatnou odchylku, pozorovanou hodnotu t statistiky pro test hypotézy o nulovosti koeficientu proti oboustranné alternativě a příslušnou p -hodnotu pro μ_{Before} , $\alpha_{\text{During}} - \alpha_{\text{Before}}$ a $\alpha_{\text{After}} - \alpha_{\text{Before}}$.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10.9167	1.2138	8.994	2.15e-10 ***
MoonDuring	2.5000	1.7165	1.456	0.155
MoonAfter	0.5417	1.7165	0.316	0.754

Spočtete stejnou tabulku pro μ_{Before} , μ_{During} a μ_{After} . Přidejte také sloupce s p -hodnotami pro testy hypotéz o nulovosti koeficientu proti jednostranným alternativám (t.j. testujte $H_0 : \beta = 0$ proti $H_1 : \beta > 0$ a $H_0 : \beta = 0$ proti $H_1 : \beta < 0$). Přiložte zdrojový kód v R.