

6 Testovanie hypotéz

Príklad 6.1. Botanik sa vybral na štyri objaviteľské výpravy s cieľom nájsť nové exotické rastliny. Do tabuľky nižšie si zapísal, koľko kolegov so sebou zobral na jednotlivé výpravy (účastníci) a koľko nových rastlín na výpravách objavili (objavy).

účastníci	3	1	2	4
objavy	15	3	8	25

Botanik modeluje počet objavov nasledovne:

$$objavy_i = \beta_0 + \beta_1 \text{účastníci}_i + \beta_2 \text{účastníci}_i^2 + \varepsilon_i.$$

- a) Testujte signifikantnosť kvadratického vplyvu počtu účastníkov, teda hypotézu $H_0 : \beta_2 = 0$.
- b) Určte, či z pohľadu Akaikeho informačného kritéria je lepší model bez kvadratického vplyvu alebo model s kvadratickým vplyvom počtu účastníkov.
- c) Modely z (b) porovnajte pomocou \bar{R}^2 .

Pomôcka: dá sa vypočítať, že

$$(\mathbf{X}^T \mathbf{X})^{-1} = \begin{pmatrix} 7,75 & -6,75 & 1,25 \\ -6,75 & 6,45 & -1,25 \\ 1,25 & -1,25 & 0,25 \end{pmatrix}.$$

Príklad 6.2. Agentúra sa v prieskume spýtala 50 firiem na nasledovné údaje: ročný príjem, počet zamestnancov (*zam*), počiatočný kapitál, priemerný vek zamestnancov (*vek*), počet počítačov (*počítače*). Ekonóm sformuloval model

$$príjem = \beta_0 + \beta_1 \ln(zam) + \beta_2 \text{kapitál} + \beta_3 \ln(vek) + \beta_4 \ln(počítače) + \varepsilon.$$

Predpokladajme, že chyby pozorovaní sú nezávislé a z $N(0, \sigma^2)$, kde σ^2 je neznáme. Ekonóm chcel testovať hypotézu, že priemerný vek nie je signifikantný a že počet počítačov má rovnaký vplyv ako počet zamestnancov (teda $\beta_3 = 0$ a $\beta_1 = \beta_4$).

- a) Sformulujte submodel zodpovedajúci testovanej hypotéze.
- b) Funkcia *lm* v Rku ekonómovi vrátila výsledky: $R^2 = 0,584$ a $R_{sub}^2 = 0,316$ v príslušnom submodeli. Ekonóm nedával pozor na ekonometrii a úlohu nevie dopočítať. Dá sa z dostupných údajov hypotéza otestovať? Ak áno, aký je výsledok?

Príklad 6.3. Aký je vzťah medzi testovaním hypotézy o kontraste ($H_0 : \mathbf{a}^T \beta = r$) pomocou \mathcal{F} a pomocou t -rozdelenia?