

PÍ SOMKA 2 - UKÁŽKA

1. **Príklad 1:** [spolu 3,5 b.]

Uvažujme model $Y = X\beta + \varepsilon$, kde vektor ε má normálne rozdelenie $N(0, \sigma^2\Omega)$, kde Ω je známa symetrická a kladne definitná matica.

- (a) [2 b.] Predpokladajme, že model odhadneme metódou najmenších štvorcov a štandardným spôsobom odhadneme štandardne odchýlky $\hat{\beta}_i$ ako odmocniny z diagonálnych prvkov matice $s^2(X^T X)^{-1}$. Z nich vypočítame t-štatistiky. Ktoré z nasledujúcich tvrdení pre tieto odhady sú pravdivé?
- Odhady $\hat{\beta}_i$ sú nevychýlené.
 - Odhady $\hat{\beta}_i$ majú normálne rozdelenie.
 - Odhad $\hat{\beta}$ má podľa Gauss-Markovovej vety najmenšiu varianciu spomedzi všetkých lineárnych nevychýlených odhadov parametra β .
 - Takto vypočítané t-štatistiky dajú korektné výsledky o signifikancii parametrov.
- (b) [0,5 b.] Napíšte vzorec na výpočet odhadu parametra β zovšeobecnenou metódou najmenších štvorcov.
- (c) [1 b.] Odvoďte (nestačí výsledok) kovariančnú maticu odhadu z predchádzajúcej úlohy.

2. **Príklad 2:** [spolu 4,5 b.]

Uvažujme model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon,$$

pričom máme k dispozícii 150 dát. O premenných ε predpokladáme, že majú nulovú strednú hodnotu a že sú nezávislé.

V modeli budeme testovať heteroskedasticitu Whitovym testom. Použijeme jednu z jeho verzií - "no cross terms".

- (a) [0,5 b.] Akú hypotézu testujeme?
- (b) [0,5 b.] Akú pomocnú regresiu odhadneme? Napíšte vysvetľovanú premennú a vysvetľujúce premenné.
- (c) [0,5 b.] Koeficient determinácie z tejto pomocnej regresie vyšiel 0.1254. Vypočítajte hodnotu testovacej štatistiky.
- (d) [0,5 b.] Aké je pravdepodobnostné rozdelenie štatistiky, ak platí nulová hypotéza? Pre aké hodnoty štatistiky nulovú hypotézu zamietame?
- (e) [0,5 b.] Rozhodnite o zamietnutí, resp. nezamietnutí nulovej hypotézy v našom konkrétnom prípade. Je v modeli heteroskedastita?

[2 b.] Chceme testovať hypotézu $\beta_1 + \beta_2 = 10$, $\beta_3 = 15$. Zapište túto hypotézu maticovo v tvare $R\beta = r$ a napíšte podrobný postup, ako ju testovať. (Ako vypočítame štatistiku? Aké pravdepodobnostné rozdelenie má štatistika, ak platí nulová hypotéza? Kedy hypotézu zamietame?)

3. **Príklad 3:** [3 b.]

Uvažujme lineárne regresné modely tvaru $Y = X\beta + \varepsilon$ pre dáta Y . Jedným z kritérií na výber vhodného modelu, ktoré berú do úvahy aj počet premenných v modeli, je Akaikeho informačné kritérium. Vzťah, ktorým sa definuje, je

$$AIC = \ln \frac{e^T e}{n} + \frac{2k}{n},$$

kde n je počet dát, k je počet premenných v modeli a e je vektor rezíduí. Pri porovnávaní je lepší ten model, ktorý má menšiu hodnotu Akaikeho kritéria.

V helpe programu EViews sa dočítame, že Akaikeho informačné kritérium, ktoré dostaneme vo výstupe z regresie, sa počíta ako

$$AIC_{EV} = -\frac{2\ell}{n} + \frac{2k}{n},$$

kde ℓ je logaritmus funkcie vierohodnosti modelu za predpokladu $\varepsilon \sim N(0, \sigma^2 I)$, vyčíslený v bode $\beta = \hat{\beta}_{MLE}$, $\sigma^2 = \hat{\sigma}_{MLE}^2$.

Dokážte, že tieto dve hodnoty, AIC a AIC_{EV} , sa líšia o konštantu, ktorá nezávisí od modelu ani od dát.

4. **Príklad 4:** [4 b.]

Bude sa týkať binárnych premenných, ktoré preberieme na cvičení 16.5.