12° online μάθημα

6/5/2020

Υποθέτοντας ότι $\varepsilon_i \sim N(0, \sigma^2)$ τότε προκύπτει ότι:

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i-\bar{x})^2}\right)\right)$$

και

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}\right)$$

Πώς προκύπτει;

Θυμηθείτε

$$\hat{\beta}_0 = \sum_{i=1}^{n}\left\{\frac{1}{n} - \frac{\bar{x}(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2}\right\} y_i$$

$$\hat{\beta}_1 = \frac{1}{\Sigma(x_i-\bar{x})^2}\Sigma(x_i-\bar{x})y_i \quad \text{ή} \quad \hat{\beta}_1 = \sum_{i=1}^{n}\frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} y_i$$

Αν ήθελε κάποιος να προχωρήσει σε έλεγχο υποθέσεων για τις παραμέτρους β_0, β_1 ή σε κατασκευή διαστημάτων εμπιστοσύνης θα αντιμετώπιζε το πρόβλημα ότι σ^2 άγνωστη. Επομένως χρειαζόμαστε μια εκτίμησή της. Ένας αμερόληπτος εκτιμητής του σ^2 δίνεται από τη σχέση:

$$S^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

ή $$S^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

το λεγόμενο μεσοτετραγωνικό σφάλμα

Επιπρόσθετα αποδεικνύεται ότι $\frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{n-2}$

και S^2 ανεξάρτητα από $\hat{\beta}_0, \hat{\beta}_1$.

Συνδυάζοντας τα αποτελέσματα έχουμε

$$\frac{\hat{\beta}_0 - \beta_0}{S\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i-\bar{x})^2}}} \sim t_{n-2}$$

$$\frac{\hat{\beta}_1 - \beta_1}{S/\sqrt{\Sigma(x_i-\bar{x})^2}} \sim t_{n-2}$$

Με αυτή κατασκευάζονται Δ.Ε. και έλεγχοι. Πώς;
Δείτε t-test

Άσκηση - Θεωρία
Έστω $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 X_0$
Ν.δ.ο. $E\hat{Y}_0 = \beta_0 + \beta_1 X_0 = Y_0$ (απλά $E\hat{\beta}_0 = \beta_0$, $E\hat{\beta}_1 = \beta_1$)

$$Var\,\hat{Y}_0 = \sigma^2\left(\frac{1}{n} + \frac{(X_0-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}\right)$$

Θυμηθείτε: $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{x}$ άρα

$$\hat{Y}_0 = \bar{Y} - \hat{\beta}_1\bar{x} + \hat{\beta}_1 X_0 \Rightarrow$$

$$\hat{Y}_0 = \sum_{i=1}^{n}\frac{1}{n}Y_i + \hat{\beta}_1(X_0-\bar{x}) \Rightarrow$$

$$\hat{Y}_0 = \sum_{i=1}^{n}\frac{1}{n}Y_i + \sum_{i=1}^{n}\frac{(x_i-\bar{x})(x_0-\bar{x})}{\Sigma(x_i-\bar{x})^2}Y_i$$

$$\hat{Y}_0 = \sum_{i=1}^{n}\left\{\frac{1}{n} + \frac{(X_0-\bar{X})}{\Sigma(X_i-\bar{X})^2}(X_i-\bar{X})\right\}Y_i$$

οπότε προκύπτει το ζητούμενο. Αν τώρα υποθέσουμε και κανονική κατανομή για τα σφάλματα δηλαδή ότι $\varepsilon_i \sim N(0,\sigma^2)$ τότε

$$\hat{Y}_0 \sim N(\underbrace{\beta_0 + \beta_1 X_0}_{= EY_0},\ Var(\hat{Y}_0))$$

και με παρόμοιο τρόπο αντικαθιστώντας την άγνωστη σ^2 προκύπτει ότι

$$\frac{\hat{Y}_0 - EY_0}{\sqrt{\widehat{Var}(\hat{Y}_0)}} \sim t_{n-2}$$