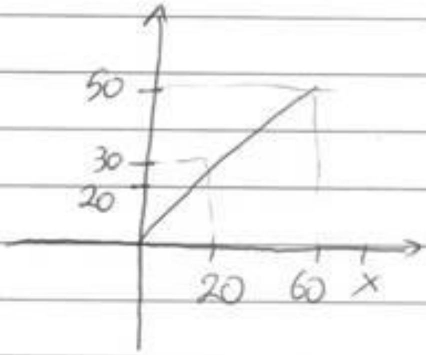
06-04-16

Ανάλυση Παλινδρόμησης

→ Γραμμική Παλινδρόμηση

$y = 20 + 0{,}5x$



y: €, x: km

(x,y): $(20,30)$, $(60,50)$

Συναρτησιακή – προσδιοριστική → {μη στοχαστική}

→ Παράδειγμα 5.2 {Βιβλίο Σ. Λουκά} ✓

• Απλή Γραμμική Παλινδρόμηση

$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad , i = 1, \dots, n$

β_0, β_1: παράμετροι, άγνωστοι {συντελεστές παλινδρόμησης}

x_i: τιμές της ελεγχόμενης μεταβλητής

Y_i: τιμές (αποκρίσεις) της εξαρτημένης μεταβλητής

ε_i: τυχαία σφάλματα, ασυσχέτιστες ανά δύο τυχαίες μεταβλητές με $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$: άγνωστο

• $E(Y_i) = \beta_0 + \beta_1 X_i$ → άγνωστη συνάρτηση παλινδρόμησης

• $\varepsilon_i = Y_i - \beta_0 - \beta_1 x_i = Y_i - E(Y_i)$

• $Var(Y_i) = \sigma^2$, $Y_i \sim (\beta_0 + \beta_1 x_i, \sigma^2)$

· Εκτίμηση της συνάρτησης παλινδρόμησης με τη μέθοδο ελαχίστων τετραγώνων.

$$Q(\beta_0,\beta_1)=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1 x_i)^2 \longrightarrow \min \hat{\beta}_0, \hat{\beta}_1$$

$$\frac{\partial Q}{\partial \beta_0}=-2\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1 x_i)(=0) \rightsquigarrow \sum_{i=1}^{n} Y_i = n\hat{\beta}_0+\hat{\beta}_1\sum_{i=1}^{n} x_i \quad (1)$$

$$\frac{\partial Q}{\partial \beta_1}=-2\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1 x_i)Y_i(=0) \rightsquigarrow \sum_{i=1}^{n} x_i Y_i = \hat{\beta}_0\sum_{i=1}^{n} x_i+\hat{\beta}_1\sum_{i=1}^{n} x_i^2 \quad (2)$$

Οι (1), (2) είναι οι κανονικές εξισώσεις

$$\overset{(1)}{\rightsquigarrow} \hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{x}$$

$$\overset{(2)}{\rightsquigarrow} \sum x_i Y_i=(\bar{Y}-\hat{\beta}_1\bar{x})\sum x_i+\hat{\beta}_1\sum x_i^2 \rightsquigarrow \sum x_i Y_i-\frac{\sum x_i\sum Y_i}{n}=$$

$$=\hat{\beta}_1\left(\sum x_i^2-\frac{(\sum x_i)^2}{n}\right)$$

· $\hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{x}$

· $$\hat{\beta}_1=\frac{\sum x_i Y_i-\frac{\sum x_i\sum Y_i}{n}}{\sum x_i^2-\frac{(\sum x_i)^2}{n}}=\frac{\sum(x_i-\bar{x})(Y_i-\bar{Y})}{\sum(x_i-\bar{x})^2}$$

$$A=\begin{pmatrix}\frac{\partial^2 Q}{\partial \beta_0^2}=2n & 2\sum x_i \\ \frac{\partial^2 Q}{\partial \beta_0\partial\beta_1}=2\sum x_i & \frac{\partial^2 Q}{\partial \beta_1^2}=2\sum x_i^2\end{pmatrix} \rightarrow$$

$$\rightarrow |A|=4n\sum x_i^2-4\left(\sum x_i\right)^2=4n\sum(x_i-\bar{x})^2>0$$

• $E(Y_i) = \beta_0 + \beta_1 x_i$

• $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$ για $x = x_0$, εκτιμώμενη συνάρτηση παλινδρόμησης

Στο παράδειγμα της ΕΒΟ

$\hat{\beta}_0 = 10.0$
$\hat{\beta}_1 = 2.0$

$\hat{Y}_0 = 10.0 + 2.0 * X_0$

• $e_i = Y_i - \hat{Y}_i \longrightarrow$ υπόλοιπα

• Ιδιότητες της εκτιμώμενης συνάρτησης παλινδρόμησης

i) Το κέντρο βάρους των παρατηρήσεων $(\bar{X}, \bar{Y})$ είναι σημείο της ευθείας γιατί αν: $X = \bar{X}$, $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X} = \bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 \bar{X} = \bar{Y}$

ii) $\sum_{1}^{n} e_i = 0$ γιατί $\sum_{1}^{n} (Y_i - \hat{Y}_i) = \sum Y_i - \sum (\hat{\beta}_0 + \hat{\beta}_1 X_i) =$

$= \sum Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum X_i = 0$, από την κανονική εξίσωση (1)

$\sum Y_i = \sum \hat{Y}_i \Rightarrow \bar{Y} = \bar{\hat{Y}}$

iii) $\sum_{i=1}^{n} e_i x_i = 0$, γιατί $\sum (Y_i - \hat{Y}_i) X_i = \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) X_i =$

$= \sum X_i Y_i - \hat{\beta}_0 \sum X_i - \hat{\beta}_1 \sum X_i^2 = 0$, από την κανονική εξίσωση (2)

Παρατήρηση: $\hat{\sigma}^2 = S^2 = \frac{1}{n-2} \sum_{i=1}^{n} e_i^2 = \frac{1}{n-2} \sum (Y_i - \hat{Y}_i)^2 =$

$= \frac{1}{n-2} \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$

Το $\hat{\sigma}^2$: μέσο τετραγωνικό σφάλμα.