

Στοιχεία Θεωρίας (Φρ. 5°)

①

○ Στον πληθυσμό:

Συν. Παλινδρόμησης: $Y_i = \beta_0 + \beta_1 X_i + u_i, i=1, \dots, n$

○ Στο δείγμα:

Εκτιμ. Συν. Παλινδρ.: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \forall i$

όπου $\hat{Y}_i$, εκτιμήτρια της $E(Y/X_i)$ ✓

$\hat{\beta}_0, \hat{\beta}_1$ εκτιμήτριες των β_0, β_1

Κατάλοιπα:

$$u_i = Y_i - E(Y/X_i)$$

και

$$\hat{u}_i = Y_i - \hat{Y}_i$$

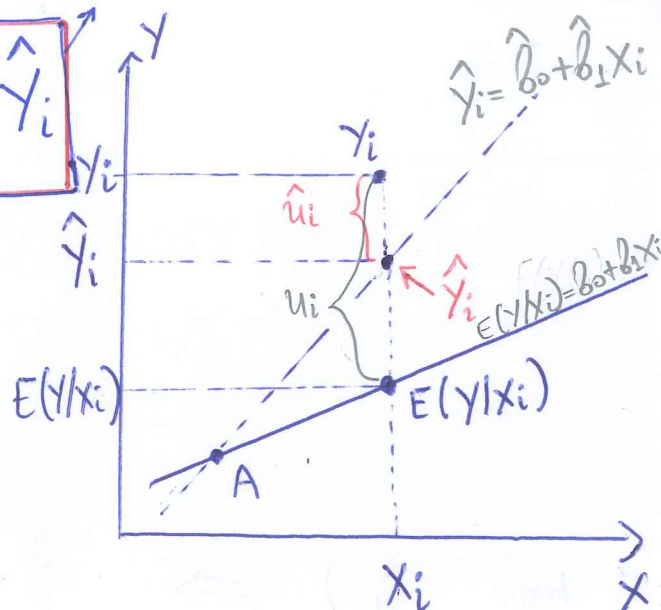
$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Ενώ

$$Y_i = \beta_0 + \beta_1 X_i + u_i \Leftrightarrow$$

$$Y_i = \beta_0 + \beta_1 X_i + Y_i - E(Y/X_i)$$

$$\Rightarrow E(Y/X_i) = \beta_0 + \beta_1 X_i$$



OLS :

$$\text{Έστω } Q(\hat{b}_0, \hat{b}_1) = \text{RSS}_{\text{unweighted}}(\hat{b}_0, \hat{b}_1)$$

Επιδιώκουμε την ελαχίστη του $\sum_{i=1}^n \hat{u}_i^2$

$$\text{δηλ. ισοδύναμα: } \min_{\hat{b}_0, \hat{b}_1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \min_{\hat{b}_0, \hat{b}_1} \sum_{i=1}^n (y_i - \hat{b}_0 - \hat{b}_1 x_i)^2$$

F.O.C.

$$\frac{\partial \sum_{i=1}^n \hat{u}_i^2}{\partial \hat{b}_0} = 0 \quad (i) \quad \text{και} \quad \frac{\partial \sum_{i=1}^n \hat{u}_i^2}{\partial \hat{b}_1} = 0 \quad (ii)$$

$\Rightarrow$ προκύπτουν
τα $\hat{b}_0, \hat{b}_1$

και για την ικανοποίηση του min:

Συνθήκες 2ης τάξης:

$$\frac{\partial^2 \sum \hat{u}_i^2}{\partial \hat{b}_0^2} > 0 \quad \text{και} \quad \frac{\partial^2 \sum \hat{u}_i^2}{\partial \hat{b}_1^2} > 0$$

$$\left. \begin{aligned} \text{Από (i)} &\Rightarrow \sum_{i=1}^n y_i = n \hat{b}_0 + \hat{b}_1 \sum_{i=1}^n x_i \\ \text{Από (ii)} &\Rightarrow \sum_{i=1}^n y_i x_i = \hat{b}_0 \sum_{i=1}^n x_i + \hat{b}_1 \sum_{i=1}^n x_i^2 \end{aligned} \right\} \text{Σ.Κ.Ε.}$$

→ Ε.Ε.Τ. με τη μορφή των Αποκλίσεων από τους Μέσους :

Σύστημα Κανονικών Εξισώσεων :

$$\begin{cases} \sum_{i=1}^n Y_i = n\hat{b}_0 + \hat{b}_1 \sum_{i=1}^n X_i & (1) \\ \sum_{i=1}^n Y_i X_i = \hat{b}_0 \sum_{i=1}^n X_i + \hat{b}_1 \sum_{i=1}^n X_i^2 & (2) \end{cases}$$

Από (1) :

$$n\hat{b}_0 = \sum_{i=1}^n Y_i - \hat{b}_1 \sum_{i=1}^n X_i \Rightarrow \hat{b}_0 = \frac{1}{n} \sum_{i=1}^n Y_i - \hat{b}_1 \frac{1}{n} \sum_{i=1}^n X_i$$
$$\Rightarrow \boxed{\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}} (*)$$

Τώρα, (2) (*) $\Rightarrow \sum_{i=1}^n Y_i X_i = (\bar{Y} - \hat{b}_1 \bar{X}) \cdot \sum_{i=1}^n X_i + \hat{b}_1 \sum_{i=1}^n X_i^2$

→ Διακρούμε με (n) και τα 2 μέλη (ως παραπάνω σχέση) :

$$\frac{1}{n} \sum_{i=1}^n Y_i X_i = (\bar{Y} - \hat{b}_1 \bar{X}) \cdot \frac{1}{n} \sum_{i=1}^n X_i + \hat{b}_1 \cdot \frac{1}{n} \sum_{i=1}^n X_i^2$$

$$\Leftrightarrow \frac{1}{n} \sum_{i=1}^n Y_i X_i = \bar{Y} \cdot \frac{1}{n} \sum_{i=1}^n X_i - \hat{b}_1 \bar{X} \cdot \frac{1}{n} \sum_{i=1}^n X_i + \hat{b}_1 \cdot \frac{1}{n} \sum_{i=1}^n X_i^2$$

$$\Leftrightarrow \frac{1}{n} \sum_{i=1}^n Y_i X_i - \bar{Y} \cdot \underbrace{\left(\frac{1}{n} \sum_{i=1}^n X_i \right)}_{\bar{X}} = \hat{b}_1 \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X} \cdot \underbrace{\left(\frac{1}{n} \sum_{i=1}^n X_i \right)}_{\bar{X}} \right)$$

$$\Leftrightarrow \frac{1}{n} \sum_{i=1}^n y_i x_i - \bar{y} \cdot \bar{x} = \hat{b}_1 \left(\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \right) \quad (4)$$

$$\Leftrightarrow \frac{1}{n} \sum_{i=1}^n (y_i x_i - \bar{y} \bar{x}) = \hat{b}_1 \cdot \frac{1}{n} \sum_{i=1}^n (x_i^2 - \bar{x}^2)$$

$$\left[\text{γιατι } \frac{1}{n} \sum_{i=1}^n \bar{y} \bar{x} = \frac{1}{n} \cdot n \cdot \bar{y} \bar{x} = \bar{y} \bar{x} \right] \quad \left[\text{γιατι } \frac{1}{n} \sum_{i=1}^n \bar{x}^2 = \frac{1}{n} \cdot n \cdot \bar{x}^2 = \bar{x}^2 \right]$$

$$\Leftrightarrow \hat{b}_1 = \frac{\frac{1}{n} \sum_{i=1}^n (y_i x_i - \bar{y} \bar{x})}{\frac{1}{n} \sum_{i=1}^n (x_i^2 - \bar{x}^2)} \quad (**)$$

$$\Leftrightarrow \hat{b}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

ο $\hat{b}_1$ δε
→ αποτελείται από
τους μέσους
($y_i - \bar{y}$), ($x_i - \bar{x}$)

Μένει να δ.ο. ισχύει η ισότητα (**), δηλ. ότι:

$$\frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (y_i x_i - \bar{y} \bar{x})}{\sum_{i=1}^n (x_i^2 - \bar{x}^2)}$$

(5)

Πράγματι, θα είναι:

$$\frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (y_i x_i - \bar{x} y_i - \bar{y} x_i + \bar{y} \bar{x})}{\sum_{i=1}^n (x_i^2 - 2\bar{x} \cdot x_i + \bar{x}^2)} =$$

$$= \frac{\sum_{i=1}^n y_i x_i - \bar{x} \sum_{i=1}^n y_i - \bar{y} \sum_{i=1}^n x_i + \sum_{i=1}^n \bar{y} \bar{x}}{\sum_{i=1}^n x_i^2 - 2\bar{x} \cdot \sum_{i=1}^n x_i + \sum_{i=1}^n \bar{x}^2} =$$

$$= \frac{\sum_{i=1}^n y_i x_i - \bar{x} \cdot n\bar{y} - \bar{y} \cdot n\bar{x} + n\bar{y}\bar{x}}{\sum_{i=1}^n x_i^2 - 2\bar{x} \cdot n\bar{x} + n\bar{x}^2}$$

$$= \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y} - \cancel{n\bar{x}\bar{y}} + \cancel{n\bar{y}\bar{x}}}{\sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2} = \frac{\sum_{i=1}^n (y_i x_i - \bar{x}\bar{y})}{\sum_{i=1}^n (x_i^2 - \bar{x}^2)} \quad \checkmark$$

$$\checkmark - n\bar{x}^2$$

$$\text{γιατί } \sum_{i=1}^n \bar{x}\bar{y} = n\bar{x}\bar{y}$$

$$\text{γιατί } \sum_{i=1}^n \bar{x}^2 = n\bar{x}^2$$

ΕΦΑΡΜΟΓΗ OLS:

(6)

Με βάση το υπόδειγμα $Y_i = b_0 + b_1 X_i + u_i$, $i=1, 2, \dots, 500$
να εκτιμηθούν με OLS οι παράμετροι b_0, b_1 και να
γραφεί η εκτιμημένη συνάρτηση παλινδρόμησης.

Δίνονται : $\sum_{i=1}^{500} X_i = 25.000$, $\sum_{i=1}^{500} Y_i = 100.000$

και $\sum_{i=1}^{500} (X_i - \bar{X})^2 = 66.000$, $\sum_{i=1}^{500} (X_i - \bar{X})(Y_i - \bar{Y}) = 200.000$

Λύση: Θα είναι $b_0 = \bar{Y} - \hat{b}_1 \bar{X}$ (1)

Γνωρίζουμε ότι σε μορφή αποκλίσεων από τους
μέσους είναι $\hat{b}_1 = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2}$ (2)

① Άρα, $\hat{b}_1 = \frac{200.000}{66.000} \approx 3,03$ (3)

Επίσης, θα είναι $\bar{X} = \frac{1}{500} \cdot 25000 = 50$ (4)

και $\bar{Y} = \frac{1}{500} \cdot 100000 = 200$ (5)

Άρα ① $\xrightarrow[\text{④, ⑤}]{\text{③}}$ $b_0 = 200 - 3,03 \cdot 50 \approx 48,5$ (6)

Άρα: Εκτιμημένη εξίσωση: $\hat{Y}_i = 48,5 + 3,03 \cdot X_i$

(7)

⊙ Θεωρείστε το υπόδειγμα: $Y_i = \alpha + u_i, i=1,2,\dots,n$

Αν $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$, τότε ν.δ.ο. $\hat{\alpha}_{LS} = \bar{Y}$

Λύση:

Ας είναι το πλήρες γραμμικό υπόδειγμα:

$$Y_i = \alpha + \beta \cdot X_i + u_i, i=1,2,\dots,n$$

Το υπόδειγμα της επιώνησης προκύπτει από το παραπάνω (πλήρες) υπόδειγμα με $X_i = 0, \forall i=1,2,\dots,n$.

Γενικά το Σ.Κ.Ε. έχει ως εξής:

$$\left\{ \begin{array}{l} \textcircled{1} \quad \sum_{i=1}^n Y_i = n \hat{\alpha} + \hat{\beta} \sum_{i=1}^n X_i \\ \textcircled{2} \quad \sum_{i=1}^n Y_i X_i = \hat{\alpha} \sum_{i=1}^n X_i + \hat{\beta} \sum_{i=1}^n X_i^2 \end{array} \right.$$

Όμως, εδώ έχουμε $\sum_{i=1}^n X_i = 0$ (αφού $\forall i=1,\dots,n, X_i = 0$)

και άρα από την $\textcircled{1}$: $\sum_{i=1}^n Y_i = n \hat{\alpha} \Rightarrow \hat{\alpha} = \frac{1}{n} \sum_{i=1}^n Y_i$

$$\Rightarrow \hat{\alpha} = \bar{Y} \quad \checkmark$$

(8)

⊙ Για το υπόδειγμα $Y_i = \beta_0 + \beta_1 X_i + u_i, i=1, 2, \dots, n$
 v.s.o. (i) $\sum_{i=1}^n \hat{u}_i = 0$ και (ii) $\sum_{i=1}^n X_i \hat{u}_i = 0$

Για το (i) :

Από την (1n) κανονική εξίσωση, έχουμε:

$$\sum_{i=1}^n Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n X_i \Leftrightarrow \sum_{i=1}^n Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^n X_i = 0$$

$$\Leftrightarrow \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

$$\Leftrightarrow \sum_{i=1}^n [Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)] = 0$$

$$\Leftrightarrow \sum_{i=1}^n (Y_i - \hat{Y}_i) = 0 \Leftrightarrow \sum_{i=1}^n \hat{u}_i = 0 \quad \checkmark$$

(9)

Για το (ii):

Από τα (2n) κανονιστικές Εξισώσεις:

$$\sum_{i=1}^n y_i x_i = \hat{b}_0 \sum_{i=1}^n x_i + \hat{b}_1 \sum_{i=1}^n x_i^2 \Leftrightarrow$$

$$\sum_{i=1}^n y_i x_i - \hat{b}_0 \sum_{i=1}^n x_i - \hat{b}_1 \sum_{i=1}^n x_i^2 = 0 \Leftrightarrow$$

$$\sum_{i=1}^n (y_i x_i - \hat{b}_0 x_i - \hat{b}_1 x_i^2) = 0 \Leftrightarrow$$

$$\sum_{i=1}^n x_i (y_i - \hat{b}_0 - \hat{b}_1 x_i) = 0 \Leftrightarrow$$

$$\sum_{i=1}^n x_i [y_i - (\hat{b}_0 + \hat{b}_1 x_i)] = 0 \Leftrightarrow$$

$$\sum_{i=1}^n x_i \cdot (y_i - \hat{y}_i) = 0 \Leftrightarrow \sum_{i=1}^n x_i \cdot \hat{u}_i = 0 \quad \checkmark$$

N.S.o. $\sum_{i=1}^n (\hat{y}_i - \bar{y}) \hat{u}_i = 0$, για το απλό γραμμικό μοντέλο $y_i = b_0 + b_1 x_i + u_i$

► Είναι $\hat{u}_i = y_i - \hat{y}_i$, άρα έχουμε:

$$\begin{aligned} \sum_{i=1}^n (\hat{y}_i - \bar{y}) \cdot \hat{u}_i &= \sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \hat{y}_i) = \\ &= \sum_{i=1}^n (\hat{y}_i y_i - \hat{y}_i^2 - \bar{y} y_i + \bar{y} \hat{y}_i) = \end{aligned}$$

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

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

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

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

Ισχύει γενικά:
 $\bar{\hat{y}} = \bar{y}$

Όμως:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (*)$$

$$1) \Rightarrow = \sum_{i=1}^n (\hat{b}_0 + \hat{b}_1 x_i) y_i - \sum_{i=1}^n (\hat{b}_0 + \hat{b}_1 x_i)^2 =$$

4)

$$= \sum_{i=1}^n \hat{b}_0 y_i + \sum_{i=1}^n \hat{b}_1 x_i y_i - \sum_{i=1}^n \hat{b}_0^2 - \sum_{i=1}^n \hat{b}_1^2 x_i^2 - \sum_{i=1}^n 2\hat{b}_0 \hat{b}_1 x_i$$

$$= \underbrace{\hat{b}_0 \sum_{i=1}^n y_i}_{+} + \underbrace{\hat{b}_1 \sum_{i=1}^n x_i y_i}_{+} - \underbrace{n \hat{b}_0^2}_{+} - \underbrace{\hat{b}_1^2 \sum_{i=1}^n x_i^2}_{+} - \underbrace{\hat{b}_0 \hat{b}_1 \sum_{i=1}^n x_i}_{+} - \underbrace{\hat{b}_0 \hat{b}_1 \sum_{i=1}^n x_i}_{+}$$

$$= \hat{b}_0 \left(\sum_{i=1}^n y_i - n \hat{b}_0 - \hat{b}_1 \sum_{i=1}^n x_i \right) +$$

↑
In κανονική
εξίσ.

$$+ \hat{b}_1 \left(\sum_{i=1}^n x_i y_i - \hat{b}_0 \sum_{i=1}^n x_i - \hat{b}_1 \sum_{i=1}^n x_i^2 \right)$$

↑
Ση κανον. εξίσ.

$$= 0 + 0 = 0 \quad \checkmark$$

↑
Κανονικές
εξισώσεις

$$\begin{cases} \sum_{i=1}^n y_i = n \hat{b}_0 + \hat{b}_1 \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i y_i = \hat{b}_0 \sum_{i=1}^n x_i + \hat{b}_1 \sum_{i=1}^n x_i^2 \end{cases}$$