

Παλινδρόμηση και Ανάλυση Διακύμανσης.

ΑΣΚΗΣΕΙΣ ΑΠΟ ΦΥΛΛΑΔΙΟ.

ΤΕΛΕΥΤΑΙΟ
ΜΑΘΗΜΑΑσκ 4

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

$$\text{NB0} \quad \alpha) r_{x,y}^2 = R^2$$

$$\beta) r_{\hat{y}, \hat{y}}^2 = R^2 = r_{x,y}^2.$$

$$R^2 = \frac{\text{SS}_{\text{reg}}}{\text{SS}_{\text{tot}}} \stackrel{\text{Num}}{=} \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{\sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 X_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

$$= \frac{\sum (\bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 X_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

$$\Rightarrow R^2 \stackrel{*}{=} \frac{\hat{\beta}_1^2 \sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2} \quad (1)$$

(*) Έχουμε αποδ. στην θεωρία
ότι $\text{SS}_{\text{reg}} = \hat{\beta}_1^2 \sum (X_i - \bar{X})^2$ άρα
μπορούσα απ' ευθείας

$$r_{x,y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}, \quad \hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\hat{\beta}_1 = \left[\frac{\sum (Y_i - \bar{Y})^2}{\sum (X_i - \bar{X})^2} \right]^{1/2} r_{x,y} \quad (2)$$

Από (1) και (2)

$$R^2 = \frac{\sum (Y_i - \bar{Y})^2}{\sum (X_i - \bar{X})^2} r_{x,y}^2 \frac{\sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2} \Rightarrow R^2 = r_{x,y}^2.$$

$$\beta) r_{y, \hat{y}} = \frac{\sum (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{\hat{y}})^2}} \quad (3)$$

$$\begin{aligned} \text{Αλλά: } \bar{\hat{y}} &= \frac{1}{n} \sum \hat{y}_i = \frac{1}{n} \sum (\hat{\beta}_0 + \hat{\beta}_1 x_i) = \frac{1}{n} n \hat{\beta}_0 + \frac{1}{n} \hat{\beta}_1 \sum x_i \\ &= \hat{\beta}_0 + \hat{\beta}_1 \bar{x} = \bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 \bar{x} = \bar{y} \\ \Rightarrow \bar{\hat{y}} &= \bar{y} \quad (4) \end{aligned}$$

Από (3) και (4)

$$r_{y, \hat{y}} = \frac{\sum (y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sqrt{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{y})^2}} = \frac{\left[\sum (y_i - \bar{y})(\hat{y}_i - \bar{y}) \right]^2}{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{y})^2} \quad (5)$$

$$\begin{aligned} \triangleright \left[\sum (y_i - \bar{y})(\hat{y}_i - \bar{y}) \right]^2 &= \left[\sum (y_i - \bar{y})(\hat{\beta}_0 + \hat{\beta}_1 x_i - \bar{y}) \right]^2 \\ &= \left[\sum (y_i - \bar{y})(\bar{x} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_i - \bar{y}) \right]^2 = \hat{\beta}_1^2 \left[\sum (y_i - \bar{y})(x_i - \bar{x}) \right]^2 \quad (6) \end{aligned}$$

$$\triangleright \sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{y})^2 = \sum (y_i - \bar{y})^2 \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 \quad (7)$$

Από (5), (6), (7)

$$r_{y, \hat{y}}^2 = \frac{\hat{\beta}_1^2 \left[\sum (y_i - \bar{y})(x_i - \bar{x}) \right]^2}{\sum (y_i - \bar{y})^2 \hat{\beta}_1^2 \sum (x_i - \bar{x})^2} = \frac{\left[\sum (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} = r_{x, y}^2 \stackrel{\text{α}}{=} R^2$$

Ασκ 7

Y_1, Y_2 ανεξ. τ.μ.

$$E(Y_1) = \theta \quad \theta \in \mathbb{R}$$

$$E(Y_2) = 2\theta$$

α) Να βρεθεί ΕΕΤ της θ .

β) Να βρεθεί SSres

Λύση

α) Επειδή $E(Y_1) = \theta$
 $E(Y_2) = 2\theta$ ΤΟΤΕ:

ΣΚΕΨΗ

$$E \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \theta$$

$$\rightarrow \underline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \theta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

$$\mu\epsilon \quad E(\varepsilon_1) = E(\varepsilon_2) = 0$$

Άρα θεωρούμε το γραμμικό μοντέλο

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \theta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \quad \text{όπου } E(\varepsilon_1) = E(\varepsilon_2) = 0$$

ή

$$\underline{Y} = X\theta + \underline{\varepsilon} \quad \text{όπου } \underline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \theta = \beta, \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

Για το μοντέλο $\underline{Y} = X\beta + \varepsilon$ γνωρίζω ότι οι ΕΕΤ του β
είναι $\hat{\beta} = (X'X)^{-1} X'Y$

Άρα οι ΕΕΤ της θ είναι $\hat{\theta} = (X'X)^{-1} X'Y$ όπου

$$X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \underline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

$$\text{Έτσι } \hat{\theta} = \left((1 \ 2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right)^{-1} (1 \ 2) \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \Rightarrow$$

$$\hat{\theta} = (5)^{-1} (Y_1 + 2Y_2)$$

Άρα ο ΕΕΤ της θ είναι $\hat{\theta} = \frac{1}{5} (Y_1 + 2Y_2)$

$$b) SS_{res} = \underline{y}' \underline{y} - \underline{\beta}' \underline{x}' \underline{y} \Rightarrow$$

$$SS_{res} = y_1^2 + y_2^2 - \frac{1}{5} (y_1 + 2y_2) (1 \ 2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$= y_1^2 + y_2^2 - \frac{1}{5} (y_1 + 2y_2) (y_1 + 2y_2) \Rightarrow$$

$$\Rightarrow SS_{res} = y_1^2 + y_2^2 - \frac{1}{5} (y_1 + 2y_2)^2.$$

Άσκηση 5 και 6 εκτός

Άσκ 9 $\sum_{i=1}^n e_i = 0$

Θεωρούμε το γενικό μοντέλο με σταθερό όρο

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \quad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

\Downarrow

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i, i=1, \dots, n$$

Θεωρώ τις κανονικές εξισώσεις που είναι:

$$(X'X)\underline{\beta} = X'\underline{Y}$$

Ο ΕΕΤ $\hat{\underline{\beta}}$ τις ικανοποιεί (τις κανονικές εξισώσεις)

$$(X'X)\hat{\underline{\beta}} = X'\underline{Y} \Rightarrow X'\underline{Y} - (X'X)\hat{\underline{\beta}} = \underline{0}$$

$$\Rightarrow X'(\underline{Y} - X\hat{\underline{\beta}}) = \underline{0}$$

$$\Rightarrow (\underline{Y} - X\hat{\underline{\beta}})'X = \underline{0}'$$

$$\Rightarrow (\underline{Y} - \hat{\underline{Y}})'X = \underline{0}$$

$$\begin{pmatrix} Y_1 - \hat{Y}_1 \\ Y_2 - \hat{Y}_2 \\ \vdots \\ Y_n - \hat{Y}_n \end{pmatrix}' X = \underline{0}$$

$$\Rightarrow (Y_1 - \hat{Y}_1, Y_2 - \hat{Y}_2, \dots, Y_n - \hat{Y}_n) \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow (Y_1 - \hat{Y}_1) \cdot 1 + (Y_2 - \hat{Y}_2) \cdot 1 + \dots + (Y_n - \hat{Y}_n) \cdot 1 = 0$$

$$\Rightarrow \sum_{i=1}^n (Y_i - \hat{Y}_i) = 0$$

$$\Rightarrow \sum_{i=1}^n e_i = 0$$

Ασκ 10 $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$

α) $\sum_{i=1}^n \hat{Y}_i (Y_i - \hat{Y}_i) = 0$ β) $\sum_{i=1}^n \text{Var}(\hat{Y}_i) = \sigma^2(p+1)$

α) $\left[\Rightarrow \sum_{i=1}^n \hat{Y}_i e_i = 0 \Leftrightarrow \underline{\hat{Y}} \perp \underline{e} \right]$

Παίρνει όπως και στην Ασκ 9 από κανονικές εξισώσεις.

$(X'X)\hat{\underline{\beta}} = X'\underline{Y} \Rightarrow X'\underline{Y} - (X'X)\hat{\underline{\beta}} = \underline{0}$.

$\Rightarrow X'(\underline{Y} - X\hat{\underline{\beta}}) = \underline{0} \Rightarrow (\underline{Y} - X\hat{\underline{\beta}})'X = \underline{0}'$

$\Rightarrow (\underline{Y} - X\hat{\underline{\beta}})'X\hat{\underline{\beta}} = \underline{0}'\hat{\underline{\beta}} = 0$

$\Rightarrow (\underline{Y} - X\hat{\underline{\beta}})'\hat{\underline{Y}} = 0$

$\Rightarrow (\underline{Y} - \hat{\underline{Y}})'\hat{\underline{Y}} = 0$

$\Rightarrow (Y_1 - \hat{Y}_1, Y_2 - \hat{Y}_2, \dots, Y_n - \hat{Y}_n) \begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{pmatrix} = 0 \Rightarrow \sum_{i=1}^n (Y_i - \hat{Y}_i)\hat{Y}_i = 0$.

β) Από θεωρία $\text{Var}(\hat{\underline{Y}}) = \text{Var}(X\hat{\underline{\beta}}) = X \text{Var}(\hat{\underline{\beta}})X'$
 $= X(\sigma^2(X'X)^{-1})X'$

$\Rightarrow \text{Var}(\hat{\underline{Y}}) = \sigma^2 X(X'X)^{-1}X'$

$\sum_{i=1}^n \text{Var}(\hat{Y}_i) = \text{tr}[\text{Var}(\hat{\underline{Y}})]$ όπου $\text{Var}(\hat{\underline{Y}}) = \begin{pmatrix} \text{Var}(\hat{Y}_1) & & & \\ & \text{Var}(\hat{Y}_2) & & \\ & & \dots & \\ & & & \text{Var}(\hat{Y}_n) \end{pmatrix}$

$= \text{tr}[\sigma^2 X(X'X)^{-1}X']$
 $= \sigma^2 \text{tr}[X(X'X)^{-1}X']$

$n \times (p+1)$ $(p+1) \times n$ $n \times (p+1)$ $(p+1) \times n$

Θα χρησιμοποιήσω ΠΙΟΤΗΤΕΣ ΙΧΝΟΥΣ

Για το ~~ix~~ ισχύει: $\text{tr}(AB) = \text{tr}(BA)$
 $(AB^{-1})' = (B^{-1})'A'$

$$= \sigma^2 \text{tr}(X'((X'X)^{-1}X'))$$

$$= \sigma^2 \text{tr}(X'(X(X'X)^{-1}))$$

$$= \sigma^2 \text{tr}(X'X(X'X)^{-1})$$

$$= \sigma^2 \text{tr}(I_{(p+1) \times (p+1)})$$

$$= \sigma^2 \sum_{i=1}^{p+1} 1$$

$$= \sigma^2(p+1).$$

* SOS την συμπληρώσει. (μπορεί να πάρει)
Άσκ 11

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i, \quad i=1, \dots, n \quad \otimes$$

Με ε_i ικανοποιούν τις υποθέσεις.

Οι κανονικές εξισώσεις είναι:

$$\begin{cases} 10\hat{\beta}_0 + 2\hat{\beta}_1 - 6\hat{\beta}_2 = 4 \\ 2\hat{\beta}_0 + 2\hat{\beta}_1 = 6 \\ -6\hat{\beta}_0 + \dots + 5\hat{\beta}_2 = 7 \end{cases}$$

α) Αν $n=10$ και $\sum_{i=1}^{10} Y_i^2 = 107$

Να υπολογιστεί $\hat{\beta} = ?$ και MSres.

β) F-TEST. $H_0: \beta_1 = 2\beta_2$

γ) Κατασκευή t-TEST: $H_0: \beta_1 = 2\beta_2$.

Λύση

$$\otimes \Leftrightarrow \underline{Y} = X\underline{\beta} + \underline{\varepsilon} \quad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & X_{11} & X_{12} \\ \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} \end{pmatrix}_{n \times 3}$$

α) Οι κανονικές εξισώσεις είναι της μορφής

$$(X'X)\hat{\beta} = X'Y, \quad X'X = \begin{pmatrix} 10 & 2 & -6 \\ 2 & 2 & 0 \\ -6 & 0 & 5 \end{pmatrix}$$

$$\underline{X}'\underline{Y} = \begin{pmatrix} 4 \\ 6 \\ 7 \end{pmatrix}$$

$$(\underline{X}'\underline{X})^{-1} = \frac{1}{8} \begin{pmatrix} 10 & -10 & 12 \\ -10 & 14 & -12 \\ 12 & -12 & 16 \end{pmatrix}$$

(Θα μου δίνει τον αντίστροφο αν μη τι και τέτοιο στην εξέταση.)

Άρα $\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y}$

$$= \begin{pmatrix} 8 \\ -5 \\ 11 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \quad \begin{array}{l} \beta_0 = 8 \\ \beta_1 = -5 \\ \beta_2 = 11 \end{array}$$

$$\begin{aligned} MS_{res} &= \frac{SS_{res}}{n-p-1} = \frac{\underline{Y}'\underline{Y} - \hat{\underline{\beta}}'\underline{X}\underline{Y}}{10-2-1} \\ &= \frac{\sum_{i=1}^{10} Y_i^2 - (8-5 \ 11) \cdot \begin{pmatrix} 4 \\ 6 \\ 7 \end{pmatrix}}{7} \\ &= \frac{107 - (8 \cdot 4 - 5 \cdot 6 + 11 \cdot 7)}{7} \\ &= \dots = 4. \end{aligned}$$

β) F-test

Από τη θεωρία γνωρίζω το F-TEST για τον έλεγχο της Γενικής

Γραμμικής Υπόθεσης: $H_0: A\underline{\beta} = \underline{c}$ όπου γενικά A ένας $q \times (p+1)$ και \underline{c} $q \times 1$ διάνυσμα

Για τον έλεγχο της $H_0: A\underline{\beta} = \underline{c}$ η ΣΣΤ του TEST είναι:

$$F = \frac{(\underline{A}\hat{\underline{\beta}} - \underline{c})' [\underline{A}(\underline{X}'\underline{X})^{-1}\underline{A}']^{-1} (\underline{A}\hat{\underline{\beta}} - \underline{c})}{SS_{res}} \sim F_{q, n-p-1} \text{ υπο την } H_0$$

↙ θα δίνεται

και κ.π $F \geq F_{q, n-p-1, \alpha}$.

$$H_0: \beta_1 = 2\beta_2 \iff H_0 \quad \underline{A}\underline{\beta} = \underline{c} \text{ όπου } \underline{A} = \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 \\ 0 & 1 & -2 \end{pmatrix}, \quad \underline{c} = 0$$

$$(0 \ 1 \ -2) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = 0 \iff 0 \cdot \beta_0 + \beta_1 \cdot 1 - 2\beta_2 = 0$$

Αντικαθιστώ στην F και βρίσκω το ηπώκεινο

$\beta_1 = 2\beta_2$

γ) Για κατασκευή t-test ακολουθούμε την λογική του World
 δηλ. στηρίξομαστε σε ευτιμήτη των παραμέτρων που
 εμφανίζονται στην H_0 .

$$H_0: \beta_1 = 2\beta_2 \Leftrightarrow H_0: \underset{\sim}{\underline{c}}' \underset{\sim}{\underline{\beta}} = 0, \quad \underline{c} = (0 \ 1 \ -2)'$$

Θα στηρίχτω σε ένα ευτιμήτη του $\underline{\beta}$. Ποιον?



$$\hat{\underline{\beta}} = (X'X)^{-1} X'Y$$

$$\text{Ξέρω } \hat{\underline{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \sigma^2 (X'X)^{-1} \right)$$

στων ΕΕΤ
 ή ΕΜΠ

$$\text{Επομένως } \underset{\substack{\uparrow \\ \text{Τυχαία μεταβλητή}}}{\underline{c}' \hat{\underline{\beta}}} = \hat{\beta}_1 - 2\hat{\beta}_2 \sim N(\underline{c}' \underline{\beta} = \beta_1 - 2\beta_2, \sigma^2 \underbrace{\underline{c}' (X'X)^{-1} \underline{c}}_{15,75})$$

$$\text{Υπό την } H_0: \beta_1 = 2\beta_2 \text{ το } \underline{c}' \hat{\underline{\beta}} \sim N(0, \sigma^2 \cdot 15,75)$$

$$\Rightarrow \underline{c}' \hat{\underline{\beta}} = \hat{\beta}_1 - 2\hat{\beta}_2 \sim N(0, 15,75 \sigma^2)$$



$$\frac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{15,75} \sigma} \sim N(0,1) \text{ υπό την } H_0$$

$$\text{Αλλά } \frac{SS_{res}}{\sigma^2} \sim \chi_{n-p-1}^2$$

$$\text{Θεωρώ το } \frac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{15,75} \sigma}$$

$$N(0,1) \sim \frac{\frac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{15,75} \sigma}}{\frac{\sqrt{\frac{SS_{res}}{\sigma^2}} / (n-p-1)}{\sqrt{15,75} MS_{res}}}$$

ή αν εf
 $\hat{\underline{\beta}}, MS_{res}$

t_{n-p-1}

Άρα για τον έλεγχο $H_0: \beta_1 = 2\beta_2$ η ΣΖΤ είναι $t = \frac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{15,75} MS_{res}} \sim t_{n-p-1}$ υπό H_0

Κ.Π $|t| > \ln p - 1, \alpha/2$

αναμεικτώ τις τιμές και συγκρίνω.

Βλέπω την άσκηση 20.

Για εξετάσεις: Κουμπουτεράκι

3 θέματα με υποερωτήματα

(μοντελα παλινδρόμησης 6.5 με 7 μέρια)

↳ θεωρητικά (ευτυχώς ^{πχ} αμερόληπτοι)

⊕

πρακτικά

θέμα

όπως άσκηση 11

θέμα

αριθμητικό (με πίνακα ANADIA)

(προσοχή στο τι θα δίνει)