καλύτερο από το 4

$|\bar{Y}_{\cdot 3} - \bar{Y}_{\cdot 4}| > ΕΣΔ$. Απορ. $β_3 = β_4$, Επειδή $\bar{Y}_{\cdot 3} - \bar{Y}_{\cdot 4} > 0 \Rightarrow 3 > 4$

Συγκεντρωτικά: $1 \equiv 4 < 2 \equiv 3 \Rightarrow$ τρόπος εντόξευσης: όποιο θέλω καύσιμο.

Μάθημα 12 ... 5, 6 SPSS

ΑΣΚΗΣΗ 7:
Y_1, Y_2 ανεξάρτητες τυχαίες μεταβλητές
$E(Y_1) = θ$, $E(Y_2) = 2θ$, $θ \in \mathbb{R}$.
ΕΕΤ θ? SSres = ?

ΛΥΣΗ:
Μοντέλο π.γ.π.: $\underset{\sim}{Y} = X\underset{\sim}{β} + \underset{\sim}{ε} \xRightarrow{E(\underset{\sim}{ε})=0} E(\underset{\sim}{Y}) = X\underset{\sim}{β}$

$$\left.\begin{array}{l} E(Y_1) = θ \\ E(Y_2) = 2θ \end{array}\right\}$$

Αν $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ τότε $E(\underset{\sim}{Y}) \overset{\text{ορ.}}{=} \begin{pmatrix} E(Y_1) \\ E(Y_2) \end{pmatrix} = \begin{pmatrix} θ \\ 2θ \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} θ$

Θεωρώ το μοντέλο πολλαπλής γραμμικής παλινδρόμησης $\underset{\sim}{Y} = X\underset{\sim}{β} + \underset{\sim}{ε}$ όπου

$\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, $X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\underset{\sim}{β} = θ$, $\underset{\sim}{ε} = \begin{pmatrix} ε_1 \\ ε_2 \end{pmatrix}$ με $E(\underset{\sim}{ε}) = 0$.

Οι ΕΕΤ στο $\underset{\sim}{Y} = X\underset{\sim}{β} + \underset{\sim}{ε}$ είναι $\hat{\underset{\sim}{β}} = (X'X)^{-1}X'\underset{\sim}{Y}$

Άρα, ο ΕΕΤ της θ είναι: $\hat{θ} = (X'X)^{-1}X'\underset{\sim}{Y} = \left[(1 \;\; 2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right]^{-1}(1 \;\; 2)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$

$$\hat{θ} = \frac{1}{5}(Y_1 + 2Y_2)$$

$SSres = Y'Y - \hat{\underset{\sim}{β}}'X'\underset{\sim}{Y} = (Y_1 \;\; Y_2)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} - \frac{1}{5}(Y_1 + 2Y_2)\left[(1 \;\; 2)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}\right] \Rightarrow$

$SSres = Y_1^2 + Y_2^2 - \frac{1}{5}(Y_1 + 2Y_2)(Y_1 + 2Y_2) = Y_1^2 + Y_2^2 - \frac{1}{5}(Y_1 + 2Y_2)^2$

ΑΣΚΗΣΗ 9: Αν σε γραμμικό μοντέλο υπάρχει σταθερός όρος τότε $\sum_{i=1}^{n} e_i = 0$.

ΛΥΣΗ:

Θεωρώ το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης:

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

$\beta_0 \to \exists$ σταθερός όρος $\beta_0 \neq 0$ $\Updownarrow$

αν δεν ξέρω πως να ξεκινήσω, πρέπει να σκεφτώ κανονικές εξισώσεις

$$\underset{\sim}{Y} = X \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

όπου $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \\ 1 & X_{n1} & \cdots & X_{np} \end{pmatrix}$, $\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$, $\underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

$e_i = Y_i - \hat{Y}_i$, γνωρίζουμε όμως $\hat{\underset{\sim}{Y}} = X \hat{\underset{\sim}{\beta}}$

Α.ο κανονικές εξισώσεις: $(X'X)\underset{\sim}{\beta} = X'\underset{\sim}{Y}$

Αφού $\hat{\underset{\sim}{\beta}}$ αποτελούν λύση των κανονικών εξισώσεων τις ικανοποιούν:

$(X'X)\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{Y}$

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

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

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

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

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

Αποδείχθηκε.

ΑΣΚΗΣΗ 10: $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \varepsilon$, $\text{rank}(X) = p+1$

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

ΛΥΣΗ:

(α)

$$\sum_{i=1}^{n} \hat{Y}_i \underbrace{(Y_i - \hat{Y}_i)}_{e_i} = 0 \quad \left(\Leftrightarrow \sum_{i=1}^{n} \hat{Y}_i e_i = 0 \Leftrightarrow \underset{\sim}{e} \perp \hat{\underset{\sim}{Y}} \right)$$

Από κανονικές εξισώσεις: $\underbrace{(\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})'}_{\underset{\sim}{e}} X = \underset{\sim}{0}'$ $\Longrightarrow$ (πολλαπλασιάζω από αριστερά με $\hat{\underset{\sim}{\beta}}$)

$$\Rightarrow (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})' X \hat{\underset{\sim}{\beta}} = 0 \Rightarrow (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})' X \hat{\underset{\sim}{\beta}} = 0 \Rightarrow \underset{\sim}{e}' \hat{\underset{\sim}{Y}} = 0 \Rightarrow$$

$$\Rightarrow (e_1 \dots e_n) \begin{pmatrix} \hat{Y}_1 \\ \vdots \\ \hat{Y}_n \end{pmatrix} = 0 \Rightarrow \sum_{i=1}^{n} e_i \hat{Y}_i = 0$$

(β)

$$\hat{\underset{\sim}{Y}} = \begin{pmatrix} \hat{Y}_1 \\ \vdots \\ \hat{Y}_n \end{pmatrix} , \quad Var(\hat{\underset{\sim}{Y}}) = \begin{pmatrix} Var(\hat{Y}_1) & & Cov(\hat{Y}_i, \hat{Y}_j) \\ & Var(\hat{Y}_2) & \\ Cov(\hat{Y}_j, \hat{Y}_i) & & \ddots \quad Var(\hat{Y}_n) \end{pmatrix}$$

$$Var(\hat{\underset{\sim}{Y}}) = Var(X\hat{\underset{\sim}{\beta}}) \overset{\text{ιδιοτητα}}{=\!=\!=} X\, Var(\hat{\underset{\sim}{\beta}}) X' = X[\sigma^2 (X'X)^{-1}] X'$$

$$\Rightarrow Var(\hat{\underset{\sim}{Y}}) = \sigma^2 \underbrace{X(X'X)^{-1}X'}_{P}$$

$$\sum_{i=1}^{n} Var(\hat{Y}_i) = tr[Var(\hat{\underset{\sim}{Y}})] = tr\{\sigma^2 X(X'X)^{-1}X'\} = \sigma^2\, tr\{ \overbrace{X}^{A} \overbrace{(X'X)^{-1}X'}^{B} \}$$

θα θελα να τον κάνω μοναδιαίο

$$\overset{tr(AB') = tr(A'B)}{=\!=\!=} \sigma^2\, tr[\underbrace{\overset{(p+1)\times n}{X'}}_{A} \underbrace{\overset{n\times(p+1)}{X} \overset{(p+1)\times n}{(X'X)^{-1}}}_{B}] = \sigma^2\, tr[I_{p+1}]$$

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

Αποδείχθηκε.

ΑΣΚΗΣΗ 11: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$, $i=1,\dots,n$

$\varepsilon_i \sim N(0,\sigma^2)$ και δίνονται οι κανονικές εξισώσεις:

$$\left.\begin{cases} 10\hat\beta_0 + 2\hat\beta_1 - 6\hat\beta_2 = 4 \\ 2\hat\beta_0 + 2\hat\beta_1 \qquad = 6 \\ -6\hat\beta_0 + \qquad +5\hat\beta_2 = 7 \end{cases}\right\} \text{Κ.Ε. } (*)$$

Αν $n=10$, $\sum_{i=1}^{n} Y_i^2 = 107$

ⓐ $\hat{\beta} = ?$, $MS_{res} = ?$

ⓑ F-τεστ, $H_0: \beta_1 = 2\beta_2$, $a=5\%$

ⓒ Κατασκευή t-test, $H_0: \beta_1 = 2\beta_2$, $a=5\%$

ΛΥΣΗ:

ⓐ $\hat{\beta} = (X'X)^{-1}X'Y$

$$MS_{res} = \frac{1}{n-p-1}\left(Y'Y - \hat{\beta}'X'Y\right), \quad n=10,\ p=2$$

Οι κ.ε γράφονται $(X'X)\hat{\beta} = X'Y$ (**)

Από (*) και (**) έχω: $X'X = \begin{pmatrix} 10 & 2 & -6 \\ 2 & 2 & 0 \\ -6 & 0 & 2 \end{pmatrix} \rightsquigarrow (X'X)^{-1} = \frac{1}{8}\begin{pmatrix} 10 & -10 & 12 \\ -10 & 14 & -12 \\ 12 & -12 & 16 \end{pmatrix}$

Από (*) και (**) έχω: $\underset{n\times3}{X'}\ \underset{3\times1}{Y} = \begin{pmatrix} 4 \\ 6 \\ 7 \end{pmatrix}$

Άρα $\hat{\beta} = \begin{pmatrix} \hat\beta_0 \\ \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} = (X'X)^{-1}X'Y = \text{πράξεις} = \begin{pmatrix} 8 \\ -5 \\ 11 \end{pmatrix}$

$$MS_{res} = \frac{1}{10-2-1}\left(\sum_{i=1}^{n} Y_i^2 - \hat{\beta}'X'Y\right) = \frac{1}{7}\left(107 - (8\ -5\ \ 11)\begin{pmatrix} 4 \\ 6 \\ 7 \end{pmatrix}\right) = 4.$$

→

(ς στη γενική μορφή)

(β) F-τεστ για έλεγχο της Γενικής Γραμμικής Υπόθεσης $H_0: A\underset{\sim}{\beta} = \underset{\sim}{0}$, $A \to q \times (p+1)$

Για έλεγχο $H_0: A\underset{\sim}{\beta} = \underset{\sim}{0}$ το F-τεστ είναι:

$$F = \frac{(A\hat{\underset{\sim}{\beta}} - 0)'(A(X'X)^{-1}A')^{-1}(A\hat{\underset{\sim}{\beta}} - 0)}{SS_{res}} \frac{n-p-1}{q} \sim F_{q,n-p-1} \text{ υπό την } H_0$$

και κρίσιμη περιοχή $F \geq F_{q,n-1-p,a}$

Αρκεί να το εφαρμόσω για $A=?$ το 1×3, $A = (0 \;\; 1 \;\; -2)$ γιατί

$A\underset{\sim}{\beta} = 0 \Leftrightarrow \beta_1 = 2\beta_2$.

(γ) Σκεπτόμενοι κατά Wald ξεκινώ με έναν εκτιμητή των παραμέτρων που εμφανίζονται στην H_0.

$H_0: \beta_1 = 2\beta_2 \Leftrightarrow H_0: \beta_1 - 2\beta_2 = 0 \Leftrightarrow H_0: 0\beta_0 + 1\beta_1 - 2\beta_2 = 0$

$\Leftrightarrow H_0: \underset{\sim}{c}'\underset{\sim}{\beta} = 0$, όπου $\underset{\sim}{c} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$

Θα στηριχτώ στον εκτιμητή $\hat{\underset{\sim}{\beta}}$ της $\underset{\sim}{\beta}$.

Έστω $\hat{\underset{\sim}{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ ο ΕΕΤ της $\underset{\sim}{\beta}$.

Γνωρίζουμε ότι υπό τις υποθέσεις για τα σφάλματα $\hat{\underset{\sim}{\beta}} \sim N_3\left(\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \underset{\sim}{\beta}, \sigma^2 (X'X)^{-1}\right)$

$$\xRightarrow{H_0: \underset{\sim}{c}'\underset{\sim}{\beta} = 0} \underset{\sim}{c}'\hat{\underset{\sim}{\beta}} = \hat{\beta}_1 - 2\hat{\beta}_2 \sim N\Big(\underset{\sim}{c}'\underset{\sim}{\beta} = \beta_1 - 2\beta_2, \underbrace{\underset{\sim}{c}'[\sigma^2 (X'X)^{-1}]\underset{\sim}{c}}_{=10,25\sigma^2}\Big)$$

Άρα, $\hat{\beta}_1 - 2\hat{\beta}_2 \sim N(\beta_1 - 2\beta_2, 10,25\sigma^2)$.

Υπό την $H_0: \beta_1 = 2\beta_2$ το $\hat{\beta}_1 - 2\hat{\beta}_2 \sim N(0, 10.25\sigma^2)$

$\Rightarrow$ Υπό την H_0 το $\dfrac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sigma\sqrt{10.25}} \sim N(0,1)$

Αλλά $\frac{SSres}{\sigma^2} \sim \chi^2_{n-p-1} \equiv \chi^2_{10-2-1} \equiv \chi^2_7$ και SSres ανεξ. $\hat{\underset{\sim}{\beta}}$.

Θεωρώ: $\left(\dfrac{N(0,1)}{\sqrt{\dfrac{\chi^2_{n-p-1}}{(n-p-1)}}} \right) \underset{H_0}{\overset{\text{υπό}}{\equiv}} \dfrac{\dfrac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sigma\sqrt{10.25}}}{\sqrt{\dfrac{SSres}{\sigma^2(n-p-1)}}} = \dfrac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{10.25}\sqrt{MSres}}$

Άρα, υπό την $H_0: \beta_1 = 2\beta_2$ το $t = \dfrac{\hat{\beta}_1 - 2\hat{\beta}_2}{\sqrt{10.25}\sqrt{MSres}} \sim t_7$

Μορφή κρίσιμης περιοχής: Μεγάλες τιμές του t, δηλ. $|t| \geq c$ όπου $c = t_{\frac{a}{2},7}$