Μάθημα 5ο 03/11/17

Π.Γ.Π. (Πολλαπλή Γραμμική Παλινδρόμηση)

$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ $(n\times 1) = n\times(p+1)\cdot(p+1)\times 1 + (n\times 1)$

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

ΕΕΤ: $\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'Y$

Μελέτη μοντέλου π.γ.π. $\to$ Ιδιότητες ΕΕΤ $\hat{\underset{\sim}{\beta}}$ $\to$ Υποθέσεις για τα σφάλματα $\underset{\sim}{\varepsilon}$

(1) $E(\underset{\sim}{\varepsilon}) = 0$
(2) $Var(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$ (I_n: ταυτοτικός $n\times n$) $\xrightarrow{\text{συνέπειες}}$
(3) $\underset{\sim}{\varepsilon} \sim N_n(0, \sigma^2 I_n)$

(1) $E(\underset{\sim}{Y}) = X\underset{\sim}{\beta}$
(2) $Var(\underset{\sim}{Y}) = \sigma^2 I_n$
(3) $\underset{\sim}{Y} \sim N_n(X\underset{\sim}{\beta}, \sigma^2 I_n)$

Πολυδιάστατη Κανονική κατανομή

Η τ.μ. W έχει μονοδιάστατη κανονική κατανομή αν $f_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(w-\mu)^2}$

Το τυχαίο διάνυσμα: $\underset{\sim}{W}$ έχει την n-διάσταση κανονική κατανομή και θα συμβολίζουμε $\underset{\sim}{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix} \sim N_n\left(\underset{\sim}{\mu} = E(\underset{\sim}{W}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}, \Sigma = Var(\underset{\sim}{W}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & & \sigma_{nn} \end{pmatrix}\right)$

αν η σ.π.π. (ή από κοινού κατανομή των $W_1, \ldots, W_n$) είναι:

$$f_{\underset{\sim}{W}}(w) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\underset{\sim}{w}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{w}-\underset{\sim}{\mu})} \quad , \underset{\sim}{w} \in \mathbb{R}^n \ (\Sigma > 0) \text{ (θετικά ορισμένος)}$$

όπου $\sigma_{ii} = Var(W_i) = \sigma_i^2$, $\sigma_{ij} = Cov(W_i, W_j)$

Δύο χαρακτηριστικές ιδιότητες της $N_n(\underset{\sim}{\mu}, \Sigma)$

1) Αν $\underset{\sim}{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix} \sim N_n(\underset{\sim}{\mu}, \Sigma)$ τότε $W_i \sim N(\mu_i, \sigma_{ii} = \sigma_i^2) \ \forall i = 1, \ldots, n$

(Το αντίστροφο δεν ισχύει πάντα, ισχύει μόνο αν οι W_i ασυσχέτιστες, δηλαδή $Cov(W_i, W_j) = 0$)

2) **Αναλλοίωτο της $N_n(\underset{\sim}{\mu}, \Sigma)$**

Αν $\underset{\sim}{W} \sim N_n(\underset{\sim}{\mu}, \Sigma)$ και A ένας $q \times n$ πίνακας τότε $A\underset{\sim}{W} \sim N_q(A\underset{\sim}{\mu}, A\Sigma A')$ ανάστροφο

(ή $A\underset{\sim}{W} + \underset{\sim}{b} \sim N_q(A\underset{\sim}{\mu} + \underset{\sim}{b}, A\Sigma A')$, $\underset{\sim}{b} \in \mathbb{R}^q$

Περαιτέρω Ιδιότητες των ΕΕΤ $\hat{\underset{\sim}{\beta}}$

1) Οι ΕΕΤ $\hat{\underset{\sim}{\beta}} \sim N_{p+1}(\underset{\sim}{\beta}, \sigma^2 (X'X)^{-1})$

ΑΠΟΔΕΙΞΗ: Δείξαμε ότι $E(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}$, $Var(\hat{\underset{\sim}{\beta}}) = \sigma^2 (X'X)^{-1}$

Αρκεί να δείξω ότι $\hat{\underset{\sim}{\beta}} \sim N_{p+1}$.

Πράγματι: $\hat{\underset{\sim}{\beta}} = \underbrace{(X'X)^{-1} X'}_{A \to (p+1)\times n} \underset{\sim}{Y} \sim N_{p+1}\left([(X'X)^{-1}X'] X\underset{\sim}{\beta}, [(X'X)^{-1}X'] \sigma^2 I_n [(X'X)^{-1}X']'\right)$

$\equiv N_{p+1}(\underset{\sim}{\beta}, \sigma^2 (X'X)^{-1} X'X (X'X)^{-1})$

$\equiv N_{p+1}(\underset{\sim}{\beta}, \sigma^2 (X'X)^{-1})$

Άρα, $\hat{\underset{\sim}{\beta}}$ έχει $(p+1)$-διάστατη κανονική ως γραμμικός μετασχηματισμός του $\underset{\sim}{Y}$ και επειδή $\underset{\sim}{Y} \sim N_n$ από την 3η υπόθεση για σφάλματα.

2) Οι ΕΕΤ $\hat{\underset{\sim}{\beta}}$ είναι ΑΟΕΔ

ΘΕΩΡΗΜΑ Gauss-Markov: Οι ΕΕΤ $\hat{\underset{\sim}{\beta}}$ έχουν τη μικρότερη διακύμανση μεταξύ όλων των αμερόληπτων εκτιμητών της παραμέτρου $\underset{\sim}{\beta}$ που είναι γραμμικές συναρτήσεις των εξαρτημένων μεταβλητών $Y_1, \dots, Y_n$.

ΑΠΟΔΕΙΞΗ: Απαιτητική!

3) Οι ΕΕΤ $\hat{\underset{\sim}{\beta}}$ ταυτίζονται με τον ΕΜΠ της παραμέτρου $\underset{\sim}{\beta}$. Περαιτέρω ο ΕΜΠ της σ^2 είναι $\hat{\sigma}^2 = \frac{SSres}{n} = \frac{n-p-1}{n} MSres$

ΑΠΟΔΕΙΞΗ: Οι ΕΕΤ προέρχονται από την ελαχιστοποίηση $\rightarrow S = \sum_{i=1}^{n} \varepsilon_i^2 = \underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = (\underset{\sim}{Y} - X\underset{\sim}{\beta})'(\underset{\sim}{Y} - X\underset{\sim}{\beta})$ (1)

$L(\underset{\sim}{\beta})$ = Από κοινού κατανομή δεδομένων = Από κοινού κατανομή $\underset{\sim}{Y} = f_{\underset{\sim}{Y}}(\underset{\sim}{y}) \overset{Y \sim N_n(X\beta, \sigma^2 I_n)}{=}$

$$= \frac{1}{(2\pi)^{n/2} \cdot |\sigma^2 I_n|^{1/2}} e^{-\frac{1}{2}(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\sigma^2 I_n)^{-1}(\underset{\sim}{y} - X\underset{\sim}{\beta})} \overset{|aA| = a^n \cdot A}{=}$$

$$\frac{1}{(2\pi)^{n/2} \cdot (\sigma^2)^{n/2} \cdot \underbrace{|I_n|}_{1}} e^{-\frac{1}{2}(\underset{\sim}{y} - X\underset{\sim}{\beta})' \cdot \frac{1}{\sigma^2} I_n^{-1}(\underset{\sim}{y} - X\underset{\sim}{\beta})}$$

$$\Rightarrow L = L(\underset{\sim}{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{Y} - X\underset{\sim}{\beta})}$$

Βολεύει να πάρω τον λογάριθμο:

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})$$

Για την εύρεση των ΕΜΠ του $\underset{\sim}{\beta}$ μεγιστοποιώ ως προς $\underset{\sim}{\beta}$ το $\log L$

ή μεγιστοποιώ ως προς $\underset{\sim}{\beta}$ το $-\frac{1}{2\sigma^2}(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})$

μεγιστοποίηση $-(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})$

ή ελαχιστοποίηση $(\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})$ (2)

Από (1), (2) $\Rightarrow$ ΕΕΤ $\equiv$ ΕΜΠ

Ο ΕΜΠ της σ^2 προκύπτει από μεγιστοποίηση ως προς σ^2 του $\log L$.

ή του $-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})$

ή του $-\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2} SSres$

Το οποίο μεγιστοποιείται για: $\hat{\sigma}^2 = \frac{SSres}{n}$

4) $\frac{SSres}{\sigma^2} \sim \chi^2_{n-p-1}$

ΑΠΟΔΕΙΞΗ: Απαιτεί γνώση θεωρίας κατανομών τετραγωνικών μορφών.

5) Οι ΕΕΤ $\hat{\underset{\sim}{\beta}}$ είναι ανεξάρτητοι του MSres ή του SSres

6) Για τον έλεγχο της $b_i = b_i^*$ (b_i^* γνωστό), $i=1,\dots,p$ χρησιμοποιείται η ΣΣΤ: $t_i = \dfrac{\hat{b}_i - b_i^*}{\sqrt{MS_{res}}\sqrt{C_{i+1,i+1}}}$, με κατανομή t_{n-p-1} υπό H_0 και κ.π.

$|t_i| \geq t_{n-p-1,\frac{a}{2}}$, $i=1,\dots,p$, όπου $C_{i+1,i+1}$ το $(i+1)$-διαγώνιο στοιχείο του πίνακα $(X'X)^{-1}$.

ΠΑΡΑΤΗΡΗΣΗ: Έχει ενδιαφέρον πρακτικό η περίπτωση $b_i^*=0$. Γιατί αν δεν μπορώ να απορρίψω την $H_0: b_i=0$ αυτό σημαίνει ότι η αντίστοιχη ανεξάρτητη μεταβλητή X_i, $i=1,\dots,p$ δεν συνδέεται γραμμικά με την Y.

ΑΠΟΔΕΙΞΗ: Επειδή $\underset{\sim}{\hat{b}} \sim N_{p+1}(\underset{\sim}{b}, \sigma^2(X'X)^{-1})$, $\underset{\sim}{\hat{b}} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_p \end{pmatrix}$, κάθε συνιστώσα έχει μονοδιάστατη κανονική Κ.

$\hat{b}_i \sim N(b_i, C_{i+1,i+1})$ $i=1,\dots,p$

και υπό την H_0: $b_i = b_i^*$, $\hat{b}_i \sim N(b_i^*, \sigma^2 C_{i+1,i+1})$, $i=1,\dots,p$

Υπό την H_0: $b_i = b_i^*$, $\dfrac{\hat{b}_i - b_i^*}{\sigma\sqrt{C_{i+1,i+1}}} \sim N(0,1)$.

Θεωρώ
$$\dfrac{\dfrac{\hat{b}_i - b_i^*}{\sigma\sqrt{C_{i+1,i+1}}}}{\sqrt{\dfrac{SS_{res}}{\sigma^2}/(n-p-1)}} \cong \dfrac{N(0,1)}{\sqrt{\dfrac{X^2_{n-p-1}}{(n-p-1)}}} \sim t_{n-p-1}$$

{SS_{res} ανεξ. $\underset{\sim}{\hat{b}}$}

$$= \dfrac{\hat{b}_i - b_i^*}{\sqrt{MS_{res}}\sqrt{C_{i+1,i+1}}} = t_i \sim t_{n-p-1}.$$

Μορφή κ.π.: είναι μεγάλες τιμές του t_i δηλ. $|t_i| > c$

Υπολογισμός του c: $a = P(\text{Απορ. } H_0 \mid H_0 \text{ αληθ.}) = P(|t_i| > c \mid t_i \sim t_{n-p-1})$
$= P(t_i > c \text{ ή } t_i < -c \mid t_i \sim t_{n-p-1})$
$= 2P(t_{n-p-1} > c) \Rightarrow$
$c = t_{n-p-1,\frac{a}{2}}$.

$P(t_{n-p-1} > c) = \frac{a}{2}$
$c = t_{n-p-1,\frac{a}{2}}$

7) Για τον έλεγχο της $H_0: b_1 = b_2 = \dots = b_p = 0$. δηλ. τον έλεγχο ύπαρξης του μοντέλου της π.γ.π. χρησιμοποιούμε ως στατιστική συνάρτηση τεστ την $F = \dfrac{MS_{reg}}{MS_{res}}$ με κατανομή

$F_{p,n-p-1}$ υπό H_0 και κ.π. μεγέθους a : $F \geq F_{p,n-p-1,a}$

ΑΣΚΗΣΗ 1: $Y_i = b_0 + b_1 X_i + \varepsilon_i$, $i=1, \dots, n$.

Ισχύουν οι υποθέσεις σφάλματα: $\varepsilon_i \sim N(0, \sigma^2)$ ασυσχέτιστα

i) Ν.Δ.Ο. ΕΕΤ ≡ ΕΜΠ

ii) Να βρεθεί ΕΜΠ της σ^2 $\left(\hat{\sigma}^2 = \frac{SS_{res}}{n}\right)$

Παρένθεση: Έστω $X_1, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $f(x, \theta)$

παράμετρος άγνωστη

Εκτιμητή για το θ.

Εκτιμητή Μέγιστης Πιθανοφάνειας (ΕΜΠ) για το θ.

Πιθανοφάνεια: $L = L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$

ΕΜΠ $\hat{\theta}$ είναι εκείνος ο οποίος μεγιστοποιεί ως προς θ την L ή logL.

Για μεγιστοποίηση της L ή logL: $\frac{\partial L}{\partial \theta} = 0$ ή $\frac{\partial \log L}{\partial \theta} = 0 \xrightarrow{\text{λύνω}} \hat{\theta} = \dots$

ΛΥΣΗ:

ΕΕΤ προκύπτουν από ελαχιστοποίηση ως προς b_0, b_1: $S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$ ⊛

$Y_1, Y_2, \dots, Y_n$ ανεξάρτητα

Πιθανοφάνεια: $L = L(b_0, b_1) = \prod_{i=1}^{n} f_{Y_i}(y_i)$

$Y_i \sim N(b_0 + b_1 X_i, \sigma^2)$

$f_{Y_i}(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2\sigma^2} \cdot (Y_i - b_0 - b_1 X_i)^2\right\}$

Οπότε $L = L(b_0, b_1) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} (Y_i - b_0 - b_1 X_i)^2 \right\} =$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2 \right\}$$

$$\log L = \log L(b_0, b_1) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$$

Αρκεί να μεγιστοποιήσω ως προς b_0, b_1 το $\log L$.

ή -״- —״— —״— το $-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$

ή —״— το $-\sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$

ή Αρκεί να ελαχιστοποιήσω ως προς b_0, b_1 το $\sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$ (**)

Από τις (*) και (**) $\rightarrow$ ΕΕΤ $\equiv$ ΕΜΠ.

(HW) $\frac{\partial}{\partial \sigma^2} \log L(\hat{b}_0, \hat{b}_1) = 0 \Rightarrow \cdots \Rightarrow \hat{\sigma}^2 = \frac{SS_{res}}{n} = \frac{n-2}{n} MS_{res}$

ΠΑΡΑΤΗΡΗΣΗ: ο ΕΜΠ $\hat{\sigma}^2$ της σ^2 δεν είναι αμερόληπτος

ΑΣΚΗΣΗ 2:

α) Ν.δ.ο. $\underbrace{\hat{b}_1 \sim N(\cdots)}_{\text{έχει αποδειχθεί}}$, $\hat{b}_0 \sim N\left(b_0, \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right\} = \frac{\sigma^2 \sum X_i^2}{n \sum (X_i - \bar{X})^2} \right)$

ΛΥΣΗ:

$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$

$E(\hat{b}_0) = E(\bar{Y} - \hat{b}_1 \bar{X}) = E(\bar{Y}) - \bar{X} E(\hat{b}_1) = E\left(\frac{1}{n} \sum Y_i \right) - \bar{X} b_1 = \frac{1}{n} \sum E(Y_i) - \bar{X} b_1 =$

$= \frac{1}{n} \sum E(b_0 + b_1 X_i + \varepsilon_i) - \bar{X} b_1 = \frac{1}{n} \sum (b_0 + b_1 X_i) - \bar{X} b_1 = b_0 + b_1 \bar{X} - \bar{X} b_1 = b_0.$

$Var(\hat{\beta}_0) = Var(\bar{Y} - \hat{\beta}_1 \bar{X})$

$$\left[Var\left(\sum_{i=1}^{n} a_i W_i\right) = \sum_{i=1}^{n} a_i^2 Var(W_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} a_i a_j Cov(W_i, W_j) \right]$$

$$Var(\hat{\beta}_0) = Var(\bar{Y}) + \bar{X}^2 Var(\hat{\beta}_1) - \bar{X}\, \cancel{Cov(\bar{Y}, \hat{\beta}_1)}^{\to 0}$$

$Cov(\bar{Y}, \hat{\beta}_1) = Cov\left(\frac{1}{n}\sum_{i=1}^{n} Y_i, \sum \frac{(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} Y_i\right)$

Ισχύει:

$$Cov\left(\sum_{i=1}^{n} a_i W_i, \sum_{j=1}^{m} b_j Z_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j Cov(W_i, Z_j)$$

$$Var(\hat{\beta}_0) = Var(\bar{Y}) + \bar{X}^2 Var(\hat{\beta}_1)$$

$$Var(\bar{Y}) = \frac{\sigma^2}{n} \quad \left(Var\bar{Y} = Var\left(\frac{1}{n}\sum Y_i\right) = \frac{1}{n^2} \cdot Var\sum(Y_i) \xrightarrow{Y_i \text{ ασυσχ}} \frac{1}{n^2}\sum Var\, Y_i = \frac{1}{n^2}\sum \sigma^2 = \frac{\sigma^2}{n} \right.$$

$$Var(\hat{\beta}_0) = \frac{\sigma^2}{n} + \bar{X}^2 \cdot Var(\hat{\beta}_1) = \frac{\sigma^2}{n} + \bar{X}^2 \frac{\sigma^2}{\sum (X_i - \bar{X})^2} = \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right\}$$

Το $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ είναι γραμμικός συνδυασμός των ασυσχέτιστων κανονικών $\bar{Y}$ και $\hat{\beta}_1$
$\Rightarrow \hat{\beta}_0 \sim Normal$.

β) Ν.Δ.Ο. $\sum_{i=1}^{n} X_i e_i = 0 \quad (X \perp e)$

ΛΥΣΗ:

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

$$= \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i (\hat{\beta}_0 + \hat{\beta}_1 X_i) = \sum_{i=1}^{n} X_i Y_i - \hat{\beta}_0 \sum_{i=1}^{n} X_i - \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 =$$

Οι κανονικές εξισώσεις: $\beta_0 \sum X_i + \beta_1 \sum X_i^2 = \sum X_i Y_i$ και επαληθεύεται από τις $\hat{\beta}_0, \hat{\beta}_1$
άρα λοιπόν: $= \sum X_i Y_i - \hat{\beta}_0 \sum X_i - \hat{\beta}_1 \sum X_i^2 = 0$.

γ) $Cov(\hat{\beta}_0, \hat{\beta}_1) = Cov(\bar{Y} - \bar{X}\hat{\beta}_1, \hat{\beta}_1) = \cancel{Cov(\bar{Y}, \hat{\beta}_1)}^{\to 0} - \bar{X} \cdot Cov(\hat{\beta}_1, \hat{\beta}_1) = -\bar{X}\, Cov(\hat{\beta}_1, \hat{\beta}_1) =$

$$= -\bar{X}\, Var(\hat{\beta}_1) = -\bar{X} \cdot \frac{\sigma^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$