

# 1 Εκτιμητής Ελαχίστων Τετραγώνων στην απλή γραμμική παλινδρόμηση

Θεωρούμε το απλό γραμμικό μοντέλο (απλή γραμμική παλινδρόμηση) :

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

$Y_i$  είναι η απόκριση της  $i$ -οστής παρατήρησης,  $x_i$  η τιμή της συμμεταβλητής της  $i$ -οστής παρατήρησης και  $\epsilon_i$  είναι το σφάλμα (δηλαδή η απόκλιση της παρατήρησης  $Y_i$  από την μέση τιμή της,  $E(Y_i) = \mu_i = \beta_0 + \beta_1 x_i$ ). Θεωρούμε τα  $x$  ως σταθερές τιμές κι όχι ως τ.μ.. Για τα σφάλματα υποθέτουμε πως είναι ομοσχεδαστικά ( $var(\epsilon_i) = \sigma^2$ ) και ασυσχέτιστα ( $cov(\epsilon_i, \epsilon_j) = 0$  για  $i \neq j$ ).

Εναλλακτικά το μοντέλο μπορεί να γραφτεί ως:

$$\begin{aligned} Y_i | x_i &\sim (\mu_i, \sigma^2), \quad cov(Y_i, Y_j | x_i, x_j) = 0, \quad i \neq j \\ \mu_i &= \beta_0 + \beta_1 x_i, \quad i = 1, \dots, n, \end{aligned} \quad (1.2)$$

Οι εκτιμητές ελαχίστων τετραγώνων είναι:

$$\begin{aligned} \hat{\beta}_1 &= \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2} \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \end{aligned} \quad (1.3)$$

Σημείωση: Από τα παραπάνω προκύπτει πως  $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$ , κι άρα η ευθεία ελαχίστων τετραγώνων περνάει από το σημείο  $(\bar{x}, \bar{y})$ .

Οι εκτιμητές ελαχίστων τετραγώνων ελαχιστοποιούν το άθροισμα των τετραγωνικών σφαλμάτων, ήτοι είναι οι λύσεις του μαθηματικού προβλήματος

$$\min_{\beta_0, \beta_1} \sum \epsilon_i^2(\beta_0, \beta_1) \Leftrightarrow \min_{\beta_0, \beta_1} \sum (y_i - \beta_0 - \beta_1 x_i)^2 \quad (1.4)$$

Παραγωγίζοντας καταλήγουμε πως οι εκτιμητές ελαχίστων τετραγώνων προκύπτουν ως λύσεις των εκτιμητικών εξισώσεων

$$\begin{aligned} \sum \epsilon_i(\beta_0, \beta_1) &= \sum y_i - \beta_0 - \beta_1 x_i = 0 \\ \sum x_i \epsilon_i(\beta_0, \beta_1) &= \sum x_i (y_i - \beta_0 - \beta_1 x_i) = 0 \end{aligned} \quad (1.5)$$

Οι εκτιμητές ελαχίστων τετραγώνων μπορούν να εκφραστούν ως γραμμικοί συνδυασμοί των  $Y$ :

$$\begin{aligned} \hat{\beta}_1 &= \sum c_i Y_i, \quad \text{με } c_i = \frac{x_i - \bar{x}}{S_{xx}} \\ \hat{\beta}_0 &= \sum d_i Y_i, \quad \text{με } d_i = \frac{1}{n} - \frac{(x_i - \bar{x})\bar{x}}{S_{xx}}, \end{aligned} \quad (1.6)$$

όπου

$$S_{xx} = \sum (x_i - \bar{x})^2$$

Για τους συντελεστες  $c_i, d_i$  ισχύουν τα παρακάτω:

$$\begin{aligned} \sum c_i &= 0, \quad \sum c_i^2 = \frac{1}{S_{xx}}, \quad \sum c_i x_i = 1 \\ \sum d_i &= 1, \quad \sum d_i^2 = \frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}, \quad \sum d_i x_i = 0 \\ \sum c_i d_i &= -\frac{\bar{x}}{S_{xx}} \end{aligned} \quad (1.7)$$

Χρησιμοποιώντας το γεγονός πως

$$\begin{aligned} E(\sum \alpha_r Y_r) &= \sum \alpha_r E(Y_r) \\ Cov(\sum \alpha_r Y_r, \sum \gamma_r Y_r) &= \sigma^2 \sum \alpha_r \gamma_r \end{aligned} \quad (1.8)$$

μπορούμε να αποδείξουμε πως

$$\begin{aligned} E(\hat{\beta}_0) &= \beta_0, \quad Var(\hat{\beta}_0) = \sigma^2 \left[ \frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right] \\ E(\hat{\beta}_1) &= \beta_1, \quad Var(\hat{\beta}_1) = \sigma^2 \frac{1}{S_{xx}} \\ Cov(\hat{\beta}_0, \hat{\beta}_1) &= -\sigma^2 \frac{\bar{x}}{S_{xx}}. \end{aligned} \quad (1.9)$$

Οι προσαρμοσμένες τιμές (fits) και τα υπόλοιπα ή κατάλοιπα (residuals) ισούνται με

$$\begin{aligned} \hat{Y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i \\ \hat{\epsilon}_i &= Y_i - \hat{Y}_i. \end{aligned} \quad (1.10)$$

Η προσαρμοσμένη τιμή  $\hat{Y}_i$  είναι απλά η εκτίμηση της μέσης τιμής  $\mu_i = \beta_0 + \beta_1 x_i$ . Σύγχρονως είναι και πρόβλεψη της τιμής της απόκρισης  $Y$  για μια νέα παρατήρηση με τιμή της συμμεταβλητής ίση με  $x_i$ . Καθότι η ευθεία ελαχίστων τετραγώνων περνάει από το σημείο  $(\bar{x}, \bar{y})$ , ισχύει πως  $\sum \hat{Y}_i = \sum Y_i$ . Επίσης από τις εκτιμητικές εξισώσεις (1.5) προκύπτει πως  $\sum \hat{\epsilon}_i = 0$  και  $\sum x_i \hat{\epsilon}_i = 0$

Οι προσαρμοσμένες τιμές και τα υπόλοιπα είναι γραμμικοί συνδυασμοί των παρατηρήσεων  $Y_1, \dots, Y_n$ :

$$\begin{aligned}\hat{Y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i = \sum_{j=1}^n (d_j + x_i c_j) Y_j \\ \hat{\epsilon}_i &= Y_i - \hat{Y}_i = \sum_{j=1}^n [\delta_{[j=i]} - (d_j + x_i c_j)] Y_j,\end{aligned}\quad (1.11)$$

όπου  $\delta_{[j=i]} = 1$  όταν  $j = i$  και ισούται με 0 όταν  $j \neq i$ .

Χρησιμοποιώντας τα αποτελέσματα στις εξισώσεις (1.8) μπορούμε να δείξουμε πώς

$$\begin{aligned}E(\hat{Y}_i) &= \mu_i = \beta_0 + \beta_1 x_i, \quad \text{var}(\hat{Y}_i) = \sigma^2 \left( \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \\ E(\hat{\epsilon}_i) &= 0, \quad \text{var}(\hat{\epsilon}_i) = \sigma^2 \left( 1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \\ \text{cov}(\hat{\epsilon}_i, \hat{\epsilon}_j) &= -\sigma^2 \left( \frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{S_{xx}} \right), \quad \text{cov}(\hat{\epsilon}_i, \hat{\beta}_0) = 0 \\ \text{cov}(\hat{\epsilon}_i, \hat{\beta}_1) &= 0, \quad \text{cov}(\hat{\epsilon}_i, \hat{Y}_j) = 0 \quad \forall i, j \\ \text{cov}(\hat{\epsilon}_i, \bar{Y}) &= 0, \quad \text{cov}(\hat{\beta}_i, \bar{Y}) = 0\end{aligned}\quad (1.12)$$

Μια απερόληπτη εκτίμηση της διασποράς των σφαλμάτων δίνεται από την  $s^2 = \frac{\sum \hat{\epsilon}_i^2}{n-2} = \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2}$ . Δηλαδή, ισχύει πως

$$E(s^2) = E\left(\frac{\sum \hat{\epsilon}_i^2}{n-2}\right) = E\left(\frac{\sum (Y_i - \hat{Y}_i)^2}{n-2}\right) = \sigma^2. \quad (1.13)$$

## 2 Κανονική Απλή Γραμμική Παλινδρόμηση

Το κανονικό απλό γραμμικό μοντέλο γράφεται ως

$$\begin{aligned}Y_i &= \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \dots, n, \\ \epsilon_1, \epsilon_2, \dots, \epsilon_n &\sim^{iid} N(0, \sigma^2),\end{aligned}\quad (2.1)$$

ή εναλλακτικά

$$\begin{aligned}Y_i | x_i &\sim^{ind} N(\mu_i, \sigma^2), \quad i = 1, \dots, n, \\ \mu_i &= \beta_0 + \beta_1 x_i\end{aligned}\quad (2.2)$$

Η υπόθεση κανονικότητας μαζί με τα αποτελέσματα στο προηγούμενο κεφάλαιο συνεπάγονται τα εξής:

$$\begin{aligned} \hat{\beta} &\sim MVN(\beta, \Sigma) \quad , \\ \hat{\beta}_0 &\sim N(\beta_0, \sigma^2[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}]) \quad , \quad \hat{\beta}_1 \sim N(\beta_1, \sigma^2 \frac{1}{S_{xx}}), \end{aligned} \quad (2.3)$$

όπου η 1η πρόταση παραπάνω απλά αναφέρει πως το  $2 \times 1$  διάνυσμα των εκτιμητριών

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$$

ακολουθεί πολυμεταβλητή κανονική κατανομή με μέση τιμή

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

και πίνακα διακυμάνσεων-συνδιακυμάνσεων  $\Sigma$ , όπου

$$\Sigma = \begin{bmatrix} var(\hat{\beta}_0) & cov(\hat{\beta}_0, \hat{\beta}_1) \\ cov(\hat{\beta}_0, \hat{\beta}_1) & var(\hat{\beta}_1) \end{bmatrix} = \begin{bmatrix} \sigma_{\hat{\beta}_0}^2 & \sigma_{\hat{\beta}_0, \hat{\beta}_1} \\ \sigma_{\hat{\beta}_0, \hat{\beta}_1} & \sigma_{\hat{\beta}_1}^2 \end{bmatrix} = \begin{bmatrix} \sigma^2[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}] & -\sigma^2 \frac{\bar{x}}{S_{xx}} \\ -\sigma^2 \frac{\bar{x}}{S_{xx}} & \sigma^2 \frac{1}{S_{xx}} \end{bmatrix}$$

Μία άμεση συνέπεια του αποτελέσματος στην (2.3) είναι πως κάθε γραμμικός συνδυασμός των εκτιμητριών  $\hat{\beta}_0$  και  $\hat{\beta}_1$  ακολουθούν κανονική κατανομή. Συγκεκριμένα, έστω  $l_1, l_2$  πραγματικοί αριθμοί (όχι και τα 2 ίσα μηδέν). Τότε

$$l_1 \hat{\beta}_0 + l_2 \hat{\beta}_1 \sim N(l_1 \beta_0 + l_2 \beta_1, l_1^2 \sigma_{\hat{\beta}_0}^2 + l_2^2 \sigma_{\hat{\beta}_1}^2 + 2l_1 l_2 \sigma_{\hat{\beta}_0, \hat{\beta}_1}), \quad (2.4)$$

από όπου προκύπτει πως

$$\frac{l_1 \hat{\beta}_0 + l_2 \hat{\beta}_1 - (l_1 \beta_0 + l_2 \beta_1)}{\sigma \sqrt{l_1^2[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}] + l_2^2 \frac{1}{S_{xx}} - 2l_1 l_2 \frac{\bar{x}}{S_{xx}}}} \sim N(0, 1). \quad (2.5)$$

Από τα αποτελέσματα στη (1.12) σε συνδυασμό με την υπόθεση κανονικότητας προκύπτει πως τα υπόλοιπα,  $\hat{\epsilon}_i$ , είναι στατιστικά ανεξάρτητα των εκτιμητριών,  $\hat{\beta}_0$  και  $\hat{\beta}_1$ . Άρα τα υπόλοιπα,  $\hat{\epsilon}_i$ , είναι ανεξάρτητα και του γραμμικού συνδυασμού  $l_1 \hat{\beta}_0 + l_2 \hat{\beta}_1$ . Τέλος, καθότι  $s^2 = \frac{\sum \hat{\epsilon}_i^2}{n-2}$  είναι συνάρτηση των υπολοίπων,  $\hat{\epsilon}_i$ , έχουμε πως  $l_1 \hat{\beta}_0 + l_2 \hat{\beta}_1$  και  $s^2$  είναι στατιστικά ανεξάρτητα.

Αργότερα θα δείξουμε πως, κάτω από την υπόθεση κανονικότητας των σφαλμάτων, έχουμε πως

$$(n-2) \frac{s^2}{\sigma^2} \sim \chi_{n-2}^2. \quad (2.6)$$

Από τα αποτελέσματα στις ( 2.5) και (2.6) έχουμε

$$\begin{aligned} & \frac{l_1\hat{\beta}_0+l_2\hat{\beta}_1-(l_1\beta_0+l_2\beta_1)}{\sigma\sqrt{l_1^2[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}]+l_2^2\frac{1}{S_{xx}}-2l_1l_2\frac{\bar{x}}{S_{xx}}}} \\ & \frac{\sqrt{(n-2)\frac{s^2}{\sigma^2}/(n-2)}}{\sqrt{(n-2)\frac{s^2}{\sigma^2}/(n-2)}} \\ & = \frac{l_1\hat{\beta}_0+l_2\hat{\beta}_1-(l_1\beta_0+l_2\beta_1)}{s\sqrt{l_1^2[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}]+l_2^2\frac{1}{S_{xx}}-2l_1l_2\frac{\bar{x}}{S_{xx}}}} \sim t_{n-2}. \end{aligned} \quad (2.7)$$

## 2.1 Συμπερασματολογία για παραμέτρους, Διάστημα εμπιστοσύνης, Διάστημα Πρόβλεψης

Βάσει του αποτελέσματος στην (2.7) ένα  $100(1-a)\%$  διάστημα εμπιστοσύνης για τον γραμμικό συνδυασμό  $l_1\beta_0+l_2\beta_1$  δίνεται από:

$$l_1\hat{\beta}_0+l_2\hat{\beta}_1 \pm t_{n-2; \frac{\alpha}{2}} s \sqrt{l_1^2[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}]+l_2^2\frac{1}{S_{xx}}-2l_1l_2\frac{\bar{x}}{S_{xx}}},$$

όπου  $t_{n-2; \alpha/2}$  είναι τό άνω  $\alpha/2$  ποσοστιαίο σημείο της  $t_{n-2}$  κατανομής.

Επίσης ο έλεγχος της μηδενικής υπόθεσης  $H_0 : l_1\beta_0+l_2\beta_1 = r$  έναντι της δίπλευρης εναλλακτικής  $H_a : l_1\beta_0+l_2\beta_1 \neq r$ , σε επίπεδο σημαντικότητας  $\alpha$  απορρίπτει την μηδενική όταν

$$\frac{|l_1\hat{\beta}_0+l_2\hat{\beta}_1-r|}{s\sqrt{l_1^2[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}]+l_2^2\frac{1}{S_{xx}}-2l_1l_2\frac{\bar{x}}{S_{xx}}}} > t_{n-2; \frac{\alpha}{2}}.$$

Ειδικές περιπτώσεις ενδιαφέροντος:

1) Για την κλίση της ευθείας παλινδρόμησης παίρνουμε  $l_1 = 0$  και  $l_2 = 1$ . Ειδικά, για να ελέγξουμε τη στατιστική σημαντικότητα της παλινδρόμησης κάνουμε τον έλεγχο  $H_0 : \beta_1 = 0$  έναντι της δίπλευρης εναλλακτικής. Η μηδενική υπόθεση απορρίπτεται όταν

$$\frac{|\hat{\beta}_1|}{s\sqrt{\frac{1}{S_{xx}}}} > t_{n-2; \frac{\alpha}{2}}. \quad (2.8)$$

2) Για την μέση τιμή της απόκρισης σε καθορισμένη τιμή της συμμεταβλητής  $x_0$ , δηλαδή για την  $\mu_0 = \beta_0 + \beta_1 x_0$ , παίρνουμε  $l_1 = 1$  και  $l_2 = x_0$ . Έτσι για

παράδειγμα, ένα 95% Διάστημα Εμπιστοσύνης για την μέση τιμή  $\mu_0 = \beta_0 + \beta_1 x_0$  δίνεται από

$$\hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t_{n-2;0.025} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}$$

$$\hat{Y}_0 \pm t_{n-2;0.025} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}$$

Διάστημα Πρόβλεψης μιας καινούργιας τιμής της απόκρισης για συγκεκριμένη τιμή,  $x_0$ , της συμεταβλητής.

Έστω πως θέλουμε να προβλέψουμε μια νέα τιμη  $Y_0$  δοθέντος συγκεκριμένης τιμης της συμεταβλητής,  $x_0$ . Π.χ. έστω πως θέλουμε να προβλεψουμε την τιμή της FEV για ένα παιδί με ύψος 65 ίντσες (εννοείται πως δεν αναφερόμαστε σε παιδί που ήδη παρατηρήσαμε την τιμή του, υπάρχει δηλαδή η παρατήρησή του στην βάση δεδομένων). Θέλουμε δηλαδή να προβλέψουμε την τυχαία μεταβλητή  $Y_0 = \beta_0 + \beta_1 x_0 + \epsilon_0$ . Μπορεί να αποδειχθεί πως η βέλτιστη γραμμική πρόβλεψη της  $Y_0$  (με κριτήριο το μέσο τετραγωνικό σφάλμα πρόβλεψης) είναι η  $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$ . Η πρόβλεψη  $\hat{Y}_0$  είναι προφανώς ανεξάρτητη της τ.μ. που θέλουμε να προβλέψουμε,  $Y_0$  (γιατί;). Παρατηρούμε πως η πρόβλεψη της  $Y_0$  είναι ίδια με την αμερόληπτη εκτίμηση της μέσης τιμής  $\mu_0 = \beta_0 + \beta_1 x_0$ .

Για το σφάλμα της της πρόβλεψης  $Y_0 - \hat{Y}_0$  ισχύουν

$$E(Y_0 - \hat{Y}_0) = 0$$

$$Var(Y_0 - \hat{Y}_0) = Var(Y_0) + Var(\hat{Y}_0) = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}\right) \quad (2.9)$$

που μαζί με την υπόθεση κανονικότητας μας δίνει

$$\frac{Y_0 - \hat{Y}_0}{\sigma \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}} \sim N(0, 1). \quad (2.10)$$

Ισχύει πως  $Y_0 - \hat{Y}_0$  και  $s^2$  είναι ανεξάρτητα (καθώς  $Y_0, \hat{\beta}_0, \hat{\beta}_1$  είναι ανεξάρτητα της  $s^2$ ). Άρα, το αποτέλεσμα στην (2.10) σε συνδυασμό με το αποτέλεσμα στην (2.6) μας δίνει

$$\frac{Y_0 - \hat{Y}_0}{s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}} \sim t_{n-2}, \quad (2.11)$$

από όπου προκύπτει πως ισχύει ότι

$$P(\hat{Y}_0 - t_{n-2; \frac{\alpha}{2}} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}} < Y_0 < \hat{Y}_0 + t_{n-2; \frac{\alpha}{2}} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}) = 1 - \alpha.$$

Το διάστημα

$$\hat{Y}_0 \pm t_{n-2; \frac{\alpha}{2}} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Sxx}}$$

λέγεται ένα  $100(1 - \alpha)\%$  Διάστημα Πρόβλεψης της  $Y_0$ .

## 2.2 Διάσπαση $\chi^2$ , ANOVA

**Λήμμα 1.** Έστω  $U = U_1 + U_2$ . Αν  $U \sim \chi_k^2$ ,  $U_1 \sim \chi_l^2$ , με  $k > l$  και  $U_1$  είναι ανεξάρτητη της  $U_2$ , τότε  $U_2 \sim \chi_{k-l}^2$

Απόδειξη. Για της ροπογεννήτριες των τ.μ. ισχύει, λόγω ανεξαρτησίας των  $U_1$  και  $U_2$

$$M_U(t) = M_{U_1}(t)M_{U_2}(t)$$

Η ροπογεννήτρια μιας τ.μ. με  $\chi_\nu^2$  κατανομή ισούται με

$$M(t) = \frac{1}{(1 - 2t)^{\nu/2}}, \quad t, 1/2$$

και συνεπώς ισχύει

$$\frac{1}{(1 - 2t)^{k/2}} = \frac{1}{(1 - 2t)^{l/2}}(t)M_{U_2}(t),$$

κι άρα

$$M_{U_2}(t) = \frac{1}{(1 - 2t)^{(k-l)/2}}(t).$$

Καθώς η ροπογεννήτρια (όταν υπάρχει) καθορίζει την κατανομή, συμπεραίνουμε πως  $U_2 \sim \chi_{k-l}^2$ .  $\square$

**Θεώρημα 2.** Σε μία απλή γραμμική παλινδρόμηση με ομοσκεδαστικά και ασυσχέτιστα σφάλματα που ακολουθούν κανονική κατανομή ισχύουν:

$$\begin{aligned} (n - 2) \frac{s^2}{\sigma^2} &\sim \chi_{n-2}^2 \Leftrightarrow \frac{SSE}{\sigma^2} \sim \chi_{n-2}^2 \\ \frac{\hat{\beta}_1^2 S_{XX}}{\sigma^2} &\sim_{H_0} \chi_1^2 \Leftrightarrow \frac{SSR}{\sigma^2} \sim_{H_0} \chi_1^2 \\ SSE &\perp SSR, \end{aligned}$$

όπου  $H_0 : \beta_1 = 0$ .

Απόδειξη. Γράφουμε

$$\begin{aligned}
 \epsilon_i &= \hat{\epsilon}_i + (\epsilon_i - \hat{\epsilon}_i) \\
 &= \hat{\epsilon}_i + (Y_i - \beta_0 - \beta_1 x_i - Y_i + \hat{\beta}_0 + \hat{\beta}_1 x_i) \\
 &= \hat{\epsilon}_i + (\hat{\beta}_0 - \beta_0) + (\hat{\beta}_1 - \beta_1)x_i.
 \end{aligned} \tag{2.12}$$

Γνωρίζουμε πως

$$\begin{aligned}
 \bar{Y} &= \beta_0 + \beta_1 \bar{x} + \bar{\epsilon} \\
 \bar{Y} &= \hat{\beta}_0 + \hat{\beta}_1 \bar{x} \\
 \implies & (\hat{\beta}_0 - \beta_0) = -(\hat{\beta}_1 - \beta_1)\bar{x} - \bar{\epsilon}.
 \end{aligned} \tag{2.13}$$

Από τις (2.13) και (;) προκύπτει

$$\begin{aligned}
 \epsilon_i &= \hat{\epsilon}_i + (\hat{\beta}_1 - \beta_1)(x_i - \bar{x}) - \bar{\epsilon} \implies \\
 \epsilon_i^2 &= \hat{\epsilon}_i^2 + (\hat{\beta}_1 - \beta_1)^2(x_i - \bar{x})^2 + \bar{\epsilon}^2 \\
 &+ 2(\hat{\epsilon}_i(\hat{\beta}_1 - \beta_1)(x_i - \bar{x}) - \hat{\epsilon}_i\bar{\epsilon} - (\hat{\beta}_1 - \beta_1)(x_i - \bar{x})\bar{\epsilon})
 \end{aligned} \tag{2.14}$$

. Από τις εκτιμητικές εξισώσεις στη (1.5) έχουμε πως

$$\begin{aligned}
 \sum \hat{\epsilon}_i &= 0 \\
 \sum x_i \hat{\epsilon}_i &= 0 \\
 \implies \sum (x_i - \bar{x})\hat{\epsilon}_i &= 0
 \end{aligned} \tag{2.15}$$

Από τις (2.14) και (2.15) παίρνουμε

$$\begin{aligned}
 \sum \epsilon_i^2 &= \sum \hat{\epsilon}_i^2 + (\hat{\beta}_1 - \beta_1)^2 \sum (x_i - \bar{x})^2 + n\bar{\epsilon}^2 \implies \\
 \frac{\sum \epsilon_i^2}{\sigma^2} &= \frac{\sum \hat{\epsilon}_i^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2 S_{xx}}{\sigma^2} + \frac{n\bar{\epsilon}^2}{\sigma^2}
 \end{aligned} \tag{2.16}$$

. Καθότι τα  $\epsilon_i$  είναι ένα τυχαίο δείγμα από την  $N(0, \sigma^2)$  κατανομή έχουμε

$$\begin{aligned}
 \frac{1}{\sigma^2} \sum \epsilon_i^2 &\sim \chi_n^2 \\
 \bar{\epsilon} &\sim N(0, \frac{\sigma^2}{n}) \implies \frac{1}{\sigma^2} n\bar{\epsilon}^2 \sim \chi_1^2.
 \end{aligned} \tag{2.17}$$

. Από την (2.3) έχουμε

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2 \frac{1}{S_{xx}}) \implies \frac{(\hat{\beta}_1 - \beta_1)^2 S_{xx}}{\sigma^2} \sim \chi_1^2. \tag{2.18}$$



Τα αποτελέσματα στη (1.12) μαζί με την υπόθεση κανονικότητας συνεπάγονται πως

$$\begin{aligned} \text{A)} & \frac{\sum \hat{\epsilon}_i^2}{\sigma^2} \perp\!\!\!\perp \frac{(\hat{\beta}_1 - \beta_1)^2 S_{xx}}{\sigma^2} \\ \text{B)} & \frac{\sum \hat{\epsilon}_i^2}{\sigma^2} \perp\!\!\!\perp \frac{n\bar{\epsilon}^2}{\sigma^2} \text{ (καθότι } \bar{\epsilon} = \bar{Y} - \beta_0 - \beta_1\bar{x}) \\ \text{Γ)} & \frac{(\hat{\beta}_1 - \beta_1)^2 S_{xx}}{\sigma^2} \perp\!\!\!\perp \frac{n\bar{\epsilon}^2}{\sigma^2}. \end{aligned}$$

Από τις τρεις παραπάνω προτάσεις προκύπτει πως  $\frac{(\hat{\beta}_1 - \beta_1)^2 S_{xx}}{\sigma^2} + \frac{n\bar{\epsilon}^2}{\sigma^2} \sim \chi_2^2$  και χρησιμοποιώντας το Λήμμα παραπάνω παίρνουμε το πρώτο αποτέλεσμα του θεωρήματος. Το δεύτερο αποτέλεσμα προκύπτει από την (2.18) και το γεγονός πως η μηδενική υπόθεση λέει πως  $\beta_1 = 0$ .  $\square$

Από το παραπάνω αποτέλεσμα προκύπτει ο  $F$  έλεγχος της στατιστικής σημαντικότητας της παλινδρόμησης, δηλαδή της  $H_0 : \beta_1 = 0$  έναντι της δίπλευρης εναλλακτικής. Καθότι  $\frac{SSR}{\sigma^2} \sim^{H_0} \chi_1^2$ ,  $\frac{SSE}{\sigma^2} \sim \chi_{n-2}^2$ , και  $SSR \perp\!\!\!\perp SSE$  έχουμε πως

$$\frac{MSR}{MSE} = \frac{SSR/1}{SSE/n-2} \sim^{H_0} F_{1,n-2}, \quad (2.19)$$

με αποτέλεσμα η  $H_0 : \beta_1 = 0$  να απορρίπτεται σε επίπεδο σημαντικότητας  $\alpha$  όταν η παρατηρηθείσα τιμή της σ.σ.ε.  $F = \frac{MSR}{MSE} > F_{1,n-2;\alpha}$ . Μπορεί να δειχθεί πως ο παραπάνω έλεγχος είναι ισοδύναμος του ελέγχου στην (2.8).

### 3 Πολλαπλή Γραμμική Παλινδρόμηση

Θεωρούμε το πολλαπλό γραμμικό μοντέλο (πολλαπλή γραμμική παλινδρόμηση) με  $p-1$  συμμεταβλητές :

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{p-1} x_{i,p-1} + \epsilon_i, \quad i = 1, \dots, n, \quad (3.1)$$

$Y_i$  είναι η απόκριση της  $i$ -οστής παρατήρησης,  $x_{i,k}$  η τιμή της  $k$  συμμεταβλητής της  $i$ -οστής παρατήρησης και  $\epsilon_i$  είναι το σφάλμα (δηλαδή η απόκλιση της παρατήρησης  $Y_i$  από την μέση τιμή της). Θεωρούμε τα  $x$  ως σταθερές τιμές κι όχι ως τ.μ.. Για τα σφάλματα υποθέτουμε πως είναι ομοσκεδαστικά ( $var(\epsilon_i) = \sigma^2$ ) και ασυσχέτιστα ( $cov(\epsilon_i, \epsilon_j) = 0$  για  $i \neq j$ ).

Εναλλακτικά το μοντέλο μπορεί να γραφτεί ως:

$$\begin{aligned} Y_i | x_{i1}, x_{i2}, \dots, x_{i,p-1} & \sim (\mu_i, \sigma^2), \quad cov(Y_i, Y_j | x_{i1}, x_{i2}, \dots, x_{i,p-1}, x_{j1}, x_{j2}, \dots, x_{j,p-1}) = 0, \quad i \neq j \\ \mu_i & = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{p-1} x_{i,p-1}, \quad i = 1, \dots, n, \end{aligned} \quad (3.2)$$

Μπορούμε να γράψουμε το μοντέλο με χρήση πινάκων και διανυσμάτων, ορίζοντας

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1,p-1} \\ 1 & x_{21} & x_{22} & \dots & x_{2,p-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{n,p-1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Στις παραπάνω εκφράσεις το διάνυσμα

$$\mathbf{x}_i = \begin{bmatrix} 1 \\ x_{i1} \\ x_{i2} \\ \vdots \\ x_{i,p-1} \end{bmatrix}$$

είναι το διάνυσμα με τις τιμές των συμμεταβλητών για την  $i$  παρατήρηση. Το μοντέλο για την  $i$  απόκριση γράφεται ως

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n, \quad (3.3)$$

ή εναλλακτικά

$$\begin{aligned} Y_i | \mathbf{x}_i &\sim (\mu_i, \sigma^2), \quad \text{cov}(Y_i, Y_j | \mathbf{x}_i, \mathbf{x}_j) = 0, \quad i \neq j \\ \mu_i &= \mathbf{x}_i^T \boldsymbol{\beta}, \quad i = 1, \dots, n. \end{aligned} \quad (3.4)$$

Το μοντέλο για όλες τις παρατηρήσεις γράφεται ως

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &\sim (\mathbf{0}_n, \sigma^2 \mathbf{I}_n), \end{aligned} \quad (3.5)$$

όπου για τον  $n \times p$  πίνακα συμμεταβλητών  $\mathbf{X}$  ισχύει  $\text{rank}(X) = p$ ,  $\mathbf{0}_n = E(\boldsymbol{\epsilon})$  είναι ένα  $n \times 1$  διάνυσμα με μηδενικά,  $\mathbf{I}_n$  είναι ένας  $n \times n$  ταυτοτικός πίνακας με μοναδες στην διαγώνιο και μηδενικά εκτος διαγωνίου. Καθότι υποθέτουμε ασυσχέτιστα και ομοσχεδαστικά σφάλματα ισχύει

$$\text{Var}(\boldsymbol{\epsilon}) = \begin{bmatrix} \text{var}(\epsilon_1) & \text{cov}(\epsilon_1, \epsilon_2) & \dots & \text{cov}(\epsilon_1, \epsilon_{n-1}) & \text{cov}(\epsilon_1, \epsilon_n) \\ \text{cov}(\epsilon_2, \epsilon_1) & \text{var}(\epsilon_2) & \dots & \text{cov}(\epsilon_2, \epsilon_{n-1}) & \text{cov}(\epsilon_2, \epsilon_n) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \text{cov}(\epsilon_{n-1}, \epsilon_1) & \text{cov}(\epsilon_{n-1}, \epsilon_2) & \dots & \text{var}(\epsilon_{n-1}) & \text{cov}(\epsilon_{n-1}, \epsilon_n) \\ \text{cov}(\epsilon_n, \epsilon_1) & \text{cov}(\epsilon_n, \epsilon_2) & \dots & \text{cov}(\epsilon_n, \epsilon_{n-1}) & \text{var}(\epsilon_n) \end{bmatrix} = \sigma^2 \mathbf{I}_n.$$

Ο εκτιμητής ελαχίστων τετραγώνων του διανύσματος συντελεστών  $\beta$  ελαχιστοποιεί το άθροισμα των τετραγωνικών σφαλμάτων, ήτοι είναι οι λύσεις του μαθηματικού προβλήματος

$$\min_{\beta} \sum \epsilon_i^2(\beta) \Leftrightarrow \min_{\beta} \epsilon^T(\beta)\epsilon(\beta) \Leftrightarrow \min_{\beta} (\mathbf{Y} - \mathbf{X}\beta)^T(\mathbf{Y} - \mathbf{X}\beta) \quad (3.6)$$

Παραγωγίζοντας καταλήγουμε πως οι εκτιμητές ελαχίστων τετραγώνων προκύπτουν ως λύσεις των εκτιμητικών εξισώσεων

$$\mathbf{X}^T(\mathbf{Y} - \mathbf{X}\beta) = \mathbf{0}_p \quad (3.7)$$

Οι εκτιμητές ελαχίστων τετραγώνων προκύπτει ως λύση του παραπάνω γραμμικού συστήματος  $p$  εξισώσεων:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (3.8)$$

Προφανώς για τα υπόλοιπα  $\hat{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta}$  της προσαρμογής ισχύει

$$\mathbf{X}^T \hat{\epsilon} = \mathbf{0}_p, \quad (3.9)$$

δηλαδή τα υπόλοιπα υπόκεινται σε  $p$  γραμμικούς περιορισμούς (και άρα δεν μπορεί να είναι στατιστικά συσχετίιστα). Προκύπτει άμεσα πως

$$\begin{aligned} E(\hat{\beta}) &= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\mathbf{Y}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \beta = \beta \\ \text{Var}(\hat{\beta}) &= \text{Var}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{Var}(\mathbf{Y}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \sigma^2 \mathbf{I}_n \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}. \end{aligned} \quad (3.10)$$

Στα παραπάνω αποτελέσματα καταλήξαμε χρησιμοποιώντας το γεγονός πως αν  $\mathbf{U}$  είναι ένα  $m \times 1$  τυχαίο διάνυσμα και  $\mathbf{C}$  ένας  $k \times m$  πίνακας σταθερών, τότε  $\mathbf{C}\mathbf{U}$  είναι ένα  $k \times 1$  τυχαίο διάνυσμα με  $E(\mathbf{C}\mathbf{U}) = \mathbf{C}E(\mathbf{U})$  και  $\text{Var}(\mathbf{C}\mathbf{U}) = \mathbf{C}\text{Var}(\mathbf{U})\mathbf{C}^T$ , ένας  $k \times k$  πίνακας.

Ο  $p \times p$  πίνακας διακυμάνσεων-συνδιακυμάνσεων του  $\hat{\beta}$

$$\text{Var}(\hat{\beta}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} = \begin{bmatrix} \text{var}(\hat{\beta}_0) & \text{cov}(\hat{\beta}_0, \hat{\beta}_1) & \text{cov}(\hat{\beta}_0, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_0, \hat{\beta}_{p-1}) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_0) & \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_1, \hat{\beta}_{p-1}) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_0) & \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{var}(\hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_2, \hat{\beta}_{p-1}) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \text{cov}(\hat{\beta}_{p-1}, \hat{\beta}_0) & \text{cov}(\hat{\beta}_{p-1}, \hat{\beta}_1) & \text{cov}(\hat{\beta}_{p-1}, \hat{\beta}_2) & \dots & \text{var}(\hat{\beta}_{p-1}) \end{bmatrix}$$

έχει στην διαγώνιο τις διασπορές των εκτιμητριών ( $\text{var}(\hat{\beta}_i) = \sigma_{\hat{\beta}_i}^2$ ) και εκτός διαγώνιου τις συνδιασπορές μεταξύ των εκτιμητριών ( $\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma_{\hat{\beta}_i, \hat{\beta}_j}$ ).

Οι προσαρμοσμένες τιμές είναι

$$\begin{aligned}\hat{\mathbf{Y}} &= \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\ &= \mathbf{H}\mathbf{Y}.\end{aligned}\quad (3.11)$$

Ο πίνακας  $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$  είναι συμμετρικός και ταυτοδύναμος, δηλαδή ισχύει  $\mathbf{H}^2 = \mathbf{H}$  (δείξτε το) και λέγεται Hat Matrix. Τα υπόλοιπα είναι

$$\begin{aligned}\hat{\boldsymbol{\epsilon}} &= \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{Y} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\ &= (\mathbf{I}_n - \mathbf{H})\mathbf{Y}.\end{aligned}\quad (3.12)$$

Ο πίνακας  $(\mathbf{I}_n - \mathbf{H})$  είναι συμμετρικός και ταυτοδύναμος, δηλαδή ισχύει  $(\mathbf{I}_n - \mathbf{H})^2 = (\mathbf{I}_n - \mathbf{H})$  (δείξτε το).

Μια αμερόληπτη εκτιμήτρια της διασποράς των σφαλμάτων  $\sigma^2$  δίνεται από

$$\begin{aligned}s^2 &= \frac{1}{n-p} \sum (Y_i - \hat{Y}_i)^2 = \frac{1}{n-p} \sum \hat{\epsilon}_i^2 \\ &= \frac{1}{n-p} \hat{\boldsymbol{\epsilon}}^T \hat{\boldsymbol{\epsilon}} = \frac{1}{n-p} \mathbf{Y}^T (\mathbf{I}_n - \mathbf{H}) \mathbf{Y}\end{aligned}\quad (3.13)$$

Αντικαθιστώντας το  $\sigma^2$  με την εκτιμήτρια  $s^2$  στον  $p \times p$  πίνακα διακυμάνσεων-συνδιακυμάνσεων του  $\hat{\boldsymbol{\beta}}$  παίρνουμε

$$\hat{Var}(\hat{\boldsymbol{\beta}}) = s^2 (\mathbf{X}^T \mathbf{X})^{-1} = \begin{bmatrix} \hat{var}(\hat{\beta}_0) & \hat{cov}(\hat{\beta}_0, \hat{\beta}_1) & \hat{cov}(\hat{\beta}_0, \hat{\beta}_2) & \dots & \hat{cov}(\hat{\beta}_0, \hat{\beta}_{p-1}) \\ \hat{cov}(\hat{\beta}_1, \hat{\beta}_0) & \hat{var}(\hat{\beta}_1) & \hat{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \hat{cov}(\hat{\beta}_1, \hat{\beta}_{p-1}) \\ \hat{cov}(\hat{\beta}_2, \hat{\beta}_0) & \hat{cov}(\hat{\beta}_2, \hat{\beta}_1) & \hat{var}(\hat{\beta}_2) & \dots & \hat{cov}(\hat{\beta}_2, \hat{\beta}_{p-1}) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \hat{cov}(\hat{\beta}_{p-1}, \hat{\beta}_0) & \hat{cov}(\hat{\beta}_{p-1}, \hat{\beta}_1) & \hat{cov}(\hat{\beta}_{p-1}, \hat{\beta}_2) & \dots & \hat{var}(\hat{\beta}_{p-1}) \end{bmatrix}$$

, που είναι η εκτίμηση του πίνακα διακυμάνσεων-συνδιακυμάνσεων του  $\hat{\boldsymbol{\beta}}$  και έχει στην διαγώνιο τις εκτιμήσεις των διασπορών των εκτιμητριών ( $\hat{var}(\hat{\beta}_i) = s_{\hat{\beta}_i}^2$ ) και εκτός διαγώνιου τις εκτιμήσεις των συνδιασπορών μεταξύ των εκτιμητριών ( $\hat{cov}(\hat{\beta}_i, \hat{\beta}_j) = s_{\hat{\beta}_i, \hat{\beta}_j}$ ).

### 3.1 Πολλαπλή Κανονική Γραμμική Παλινδρόμηση

Το μοντέλο για όλες τις παρατηρήσεις γράφεται ως

$$\begin{aligned}\mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &\sim N(\mathbf{0}_n, \sigma^2 \mathbf{I}_n),\end{aligned}\quad (3.14)$$

όπου για τον  $n \times p$  πίνακα συμμεταβλητών  $\mathbf{X}$  ισχύει  $\text{rank}(X) = p$ ,  $\mathbf{0}_n = E(\boldsymbol{\epsilon})$  είναι ένα  $n \times 1$  διάνυσμα με μηδενικά,  $\mathbf{I}_n$  είναι ένας  $n \times n$  ταυτοτικός πίνακας με μοναδες στην διαγώνιο και μηδενικά εκτος διαγωνίου.

Έχουμε πως

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^T \mathbf{X})^{-1}), \quad (3.15)$$

από όπου προκύπτει πως κάθε γραμμικός συνδυασμός των εκτιμήσεων των συντελεστών στην παλινδρόμηση ακολουθεί κανονική κατανομή. Πιο συγκεκριμένα, αν πάρουμε ένα  $p \times 1$  διάνυσμα  $\mathbf{l} \neq \mathbf{0}_p$  (δηλαδή  $\mathbf{l}^T = (l_0, l_1, l_2, \dots, l_{p-1}) \neq (0, 0, 0, \dots, 0)$ ) τότε

$$\mathbf{l}^T \hat{\boldsymbol{\beta}} \sim N(\mathbf{l}^T \boldsymbol{\beta}, \text{Var}(\mathbf{l}^T \hat{\boldsymbol{\beta}}) = \mathbf{l}^T \text{Var}(\hat{\boldsymbol{\beta}}) \mathbf{l} = \sigma^2 \mathbf{l}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{l}), \quad (3.16)$$

από όπου προκύπτει πως

$$\frac{\mathbf{l}^T \hat{\boldsymbol{\beta}} - \mathbf{l}^T \boldsymbol{\beta}}{\sigma \sqrt{\mathbf{l}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{l}}} \sim N(0, 1), \quad (3.17)$$

Μπορεί να δειχθεί πως

$$\begin{aligned} (n-p) \frac{s^2}{\sigma^2} &\sim \chi_{n-p}^2 \\ s^2 &\perp \hat{\boldsymbol{\beta}}. \end{aligned} \quad (3.18)$$

Άρα, εύκολα προκύπτει

$$\frac{\mathbf{l}^T \hat{\boldsymbol{\beta}} - \mathbf{l}^T \boldsymbol{\beta}}{s \sqrt{\mathbf{l}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{l}}} \sim t_{n-p}. \quad (3.19)$$

### 3.2 Συμπερασματολογία για παραμέτρους, Διάστημα εμπιστοσύνης, Διάστημα Πρόβλεψης

Από την (3.18) προκύπτει πως ένα  $100(1 - \alpha)\%$  Δ.Ε. για τον γραμμικό συνδυασμό  $\mathbf{l}^T \boldsymbol{\beta}$  δίνεται από

$$\mathbf{l}^T \hat{\boldsymbol{\beta}} \pm t_{n-p; \frac{\alpha}{2}} \sqrt{\mathbf{l}^T \hat{\text{Var}}(\hat{\boldsymbol{\beta}}) \mathbf{l}},$$

όπου  $t_{n-p; \alpha/2}$  είναι τό άνω  $\alpha/2$  ποσοστιαίο σημείο της  $t_{n-p}$  κατανομής και  $\hat{\text{Var}}(\hat{\boldsymbol{\beta}}) = s^2(\mathbf{X}^T \mathbf{X})^{-1}$ .

Ως παράδειγμα, αν πάρουμε  $\mathbf{l}$  να έχει 1 στην  $i+1$  θέση και 0 αλλού παίρνουμε ένα  $100(1 - \alpha)\%$  Δ.Ε. για τον συντελεστή  $\beta_i$

$$\hat{\beta}_i \pm t_{n-p; \frac{\alpha}{2}} s_{\hat{\beta}_i}$$

Αντίστοιχα, ο έλεγχος της μηδενικής υπόθεση  $H_0 : \mathbf{l}^T \boldsymbol{\beta} = r$  έναντι της δίπλευρης εναλλακτικής απορρίπτει την  $H_0$  σε επίπεδο σημαντικότητας  $\alpha$  όταν

$$\frac{|\mathbf{l}^T \hat{\boldsymbol{\beta}} - r|}{\sqrt{\mathbf{l}^T \hat{Var}(\hat{\boldsymbol{\beta}}) \mathbf{l}}} > t_{n-p; \frac{\alpha}{2}},$$