ΕΚΤΙΜΗΣΗ, Δ.Ε. και
ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ για την $\mathrm{E}(\mathrm{Y}) = \beta_0 + \beta_1 \cdot \mathrm{X}$

Εκτιμήτρια της $\mathrm{E}(\mathrm{Y})$ είναι η $\hat{\mathrm{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \mathrm{X}$.

Ιδιότητες της εκτιμήτριας: (Υποθέτω ότι έχω κανονικά και ανεξάρτητα σφάλματα)

$$\mathrm{E}(\hat{\mathrm{Y}}) = \mathrm{E}(\hat{\beta}_0 + \hat{\beta}_1 \cdot \mathrm{X}) = \mathrm{E}(\hat{\beta}_0) + \mathrm{E}(\hat{\beta}_1) \cdot \mathrm{X} = \beta_0 + \beta_1 \cdot \mathrm{X}$$

$$Var(\hat{\mathrm{Y}}) = Var(\hat{\beta}_0 + \hat{\beta}_1 \cdot \mathrm{X}) = Var(\hat{\beta}_0) + \mathrm{X}^2 \cdot Var(\hat{\beta}_1) + 2 \cdot \mathrm{X} \cdot Cov(\hat{\beta}_0, \hat{\beta}_1) =$$

$$= \sigma^2 \cdot \frac{\sum_{i=1}^{n} \mathrm{X}_i^{\,2}}{n \cdot \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} + \mathrm{X}^2 \cdot \frac{\sigma^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} - 2 \cdot \mathrm{X} \cdot \sigma^2 \cdot \frac{\overline{\mathrm{X}}}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} =$$

$$= \frac{\sigma^2}{n \cdot \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \left(\sum_{i=1}^{n} \mathrm{X}_i^{\,2} + n \cdot \mathrm{X}^2 - 2 \cdot n \cdot \mathrm{X} \cdot \overline{\mathrm{X}} \right) =$$

$$= \frac{\sigma^2}{n \cdot \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \left(\sum_{i=1}^{n} \mathrm{X}_i^{\,2} + n \cdot \mathrm{X}^2 - 2 \cdot n \cdot \mathrm{X} \cdot \overline{\mathrm{X}} + n \cdot \overline{\mathrm{X}}^2 - n \cdot \overline{\mathrm{X}}^2 \right) =$$

$$= \frac{\sigma^2}{n \cdot \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \left[\left(\sum_{i=1}^{n} \mathrm{X}_i^{\,2} - n \cdot \overline{\mathrm{X}}^2 \right) + \left(n \cdot \mathrm{X}^2 - 2 \cdot n \cdot \mathrm{X} \cdot \overline{\mathrm{X}} + n \cdot \overline{\mathrm{X}}^2 \right) \right] =$$

$$= \frac{\sigma^2}{n \cdot \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \left[\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2 + n \cdot (\mathrm{X} - \overline{\mathrm{X}})^2 \right] = \sigma^2 \cdot \left[\frac{1}{n} + \frac{(\mathrm{X} - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right]$$

Αφού λόγω της υπόθεσης της κανονικότητας, τα $\hat{\beta}_0, \hat{\beta}_1$ είναι γραμμικοί συνδυασμοί κανονικών, τότε και η $\hat{\mathrm{Y}}$ θα ακολουθεί την κανονική κατανομή σαν γραμμικός συνδυασμός των $\hat{\beta}_0, \hat{\beta}_1$, δηλαδή,

$$\hat{\mathrm{Y}} \sim \mathrm{N}\left(\beta_0 + \beta_1 \cdot \mathrm{X}, \left[\frac{1}{n} + \frac{(\mathrm{X} - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 \right)$$

Οπότε, έχουμε σαν αντιστρεπτή ποσότητα την $\dfrac{\hat{\mathrm{Y}}-(\beta_0+\beta_1\cdot \mathrm{X})}{\sqrt{\hat{Var}(\hat{\mathrm{Y}})}}\sim t_{n-2}$,

όπου $\hat{Var}(\hat{\mathrm{Y}})=\left[\dfrac{1}{n}+\dfrac{(\mathrm{X}-\overline{\mathrm{X}})^2}{\sum_{i=1}^{n}(\mathrm{X}_i-\overline{\mathrm{X}})^2}\right]\cdot\hat{\sigma}^2$.

Άρα ένα $(1-\alpha)\%$ Δ.Ε. για την $\mathrm{E}(\mathrm{Y})$ είναι:

$$\left(\hat{\mathrm{Y}}-t_{n-2,\alpha/2}\cdot\sqrt{\hat{Var}(\hat{\mathrm{Y}})},\hat{\mathrm{Y}}+t_{n-2,\alpha/2}\cdot\sqrt{\hat{Var}(\hat{\mathrm{Y}})}\right)$$

Για τον έλεγχο της υπόθεσης $\mathrm{H}_0:\beta_0+\beta_1\cdot\mathrm{X}=\nu$ προς $\mathrm{H}_1:\beta_0+\beta_1\cdot\mathrm{X}\neq\nu$, σε επίπεδο σημαντικότητας α, απορρίπτουμε την $\mathrm{H}_0:\beta_0+\beta_1\cdot\mathrm{X}=\nu$ όταν $|t|>t_{n-2,\alpha/2}$,

όπου $t=\dfrac{\hat{\mathrm{Y}}-\nu}{\sqrt{\hat{Var}(\hat{\mathrm{Y}})}}$.

ΔΙΑΣΤΗΜΑ ΠΡΟΒΛΕΨΗΣ
για την $\mathrm{E}(\mathrm{Y})=\beta_0+\beta_1\cdot\mathrm{X}$

Σε αυτό το σημείο θέλουμε να κάνουμε πόβλεψη της τιμής Y για κάποιο μελλοντικό X.

Χρησιμοποιούμε το εκτιμώμενο μοντέλο $\hat{\mathrm{Y}}=\hat{\beta}_0+\hat{\beta}_1\cdot\mathrm{X}$.

Με βάση αυτό το μοντέλο, έχουμε ότι $\hat{\mathrm{Y}}\sim\mathrm{N}\left(\beta_0+\beta_1\cdot\mathrm{X},\left[\dfrac{1}{n}+\dfrac{(\mathrm{X}-\overline{\mathrm{X}})^2}{\sum_{i=1}^{n}(\mathrm{X}_i-\overline{\mathrm{X}})^2}\right]\cdot\sigma^2\right)$.

Σκοπός μας είναι να ελαχιστοποιήσουμε το σφάλμα στην πρόβλεψη.

Από τη διακύμανση του εκτιμητή, μπορεί να φανεί και το μέγεθος του σφάλματος:

(α) Αν η τιμή της X είναι κοντά στο $\overline{\mathrm{X}}$, τότε αναμένουμε να έχουμε μικρό σφάλμα στην πρόβλεψη.

(β) Αν το εύρος των X_i είναι μεγάλο, τότε και το $\sum_{i=1}^{n}(X_i-\overline{X})^2$ θα είναι μεγάλο, άρα αναμένεται μικρό σφάλμα στην πρόβλεψη.

(γ) Όταν το πλήθος των παρατηρήσεων είναι μεγάλο, τότε πάλι αναμένουμε μικρό σφάλμα στην πρόβλεψη.

Βασικός μας στόχος είναι να δημιουργήσουμε ένα διάστημα πρόβλεψης (Δ.Π.) για την πραγματική τιμή της Y.

Έχουμε ότι $Y-\hat{Y}=Y-\hat{\beta}_0-\hat{\beta}_1\cdot X$.

Επίσης, $E(Y-\hat{Y})=E(Y)-E(\hat{Y})=\beta_0+\beta_1\cdot X-\hat{\beta}_0-\hat{\beta}_1\cdot X=0$ και

$$Var(Y-\hat{Y})=Var(Y)+Var(\hat{Y})-2\cdot Cov(Y,\hat{Y})=\sigma^2+\sigma^2\cdot\left[\frac{1}{n}+\frac{(X-\overline{X})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}\right]-0=$$

$$=\sigma^2\cdot\left[1+\frac{1}{n}+\frac{(X-\overline{X})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}\right]$$

<u>Σημείωση:</u> Γνωρίζουμε ότι $Cov(Y_i-Y_j)=0, i\neq j$. Η $\hat{Y}$ εκτιμήθηκε βάσει τις $Y_i, i=1,2,...,n$ παρατηρήσεις και η Y είναι μια καινούρια παρατήρηση. Άρα η Y είναι ανεξάρτητη από τις $Y_i, i=1,2,...,n$ παρατηρήσεις, άρα $Cov(Y,\hat{Y})=0$.

Επομένως, $Y-\hat{Y}\sim N\left(0,\sigma^2\cdot\left[1+\frac{1}{n}+\frac{(X-\overline{X})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}\right]\right)$, σαν γραμμικός συνδυασμός κανονικών.

Άρα έχουμε σαν αντιστρεπτή ποσότητα την $t=\dfrac{Y-\hat{Y}}{\hat{\sigma}\cdot\sqrt{1+\frac{1}{n}+\frac{(X-\overline{X})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}}}\sim t_{n-2}$.

Οπότε ισχύει ότι, $P\left(-t_{n-2,\alpha/2}<\dfrac{Y-\hat{Y}}{\hat{\sigma}\cdot\sqrt{1+\frac{1}{n}+\frac{(X-\overline{X})^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}}}<t_{n-2,\alpha/2}\right)=1-\alpha$

και άρα ένα $(1-\alpha)\%$ Δ.Π. για την τιμή της Y είναι:

$$\left(\hat{\mathrm{Y}} - t_{n-2,\alpha/2} \cdot \hat{\sigma} \cdot \sqrt{1+\frac{1}{n}+\frac{(\mathrm{X}-\overline{\mathrm{X}})^2}{\sum_{i=1}^{n}(\mathrm{X}_i-\overline{\mathrm{X}})^2}}, \hat{\mathrm{Y}} + t_{n-2,\alpha/2} \cdot \hat{\sigma} \cdot \sqrt{1+\frac{1}{n}+\frac{(\mathrm{X}-\overline{\mathrm{X}})^2}{\sum_{i=1}^{n}(\mathrm{X}_i-\overline{\mathrm{X}})^2}} \right)$$

<u>Έλεγχος έλλειψης προσαρμογής:</u>

Όταν υιοθετούμε το μοντέλο $\mathrm{Y}_i = \beta_0 + \beta_1 \cdot \mathrm{X}_i + \varepsilon_i, i = 1,2,...,n$, δεν σημαίνει ότι η σχέση μεταξύ X και Y είναι γραμμική.

Για παράδειγμα, θα μπορούσε η πραγματική σχέση μεταξύ X και Y να ήταν $\mathrm{Y}_i = \beta_0^* + \beta_1^* \cdot \mathrm{X}_i + \beta_2^* \cdot \mathrm{X}_i^2 + \varepsilon_i, i = 1,2,...,n$.

Σε ορισμένες περιπτώσεις μπορούμε να ελέγξουμε κατά πόσο το μοντέλο είναι σωστό.

<u>Συνέπειες από την υιοθέτηση λανθασμένου μοντέλου:</u>

(α) Οι υποθέσεις για τα σφάλματα παραμένουν οι ίδιες.

(β) Οι υποθέσεις για τα Y_i γίνονται ως εξής:

$\mathrm{E}(\mathrm{Y}_i) \neq \beta_0 + \beta_1 \cdot \mathrm{X}_i$

$Cov(\mathrm{Y}_i, \mathrm{Y}_j) = 0, i \neq j$

$Var(\mathrm{Y}_i) = \sigma^2$

$\mathrm{Y}_i \sim \mathrm{N}(\mathrm{E}(\mathrm{Y}_i), \sigma^2)$

(γ) Το εκτιμώμενο μοντέλο θα παραμείνει το ίδιο: $\hat{\mathrm{Y}}_i = \hat{\beta}_0 + \hat{\beta}_1 \cdot \mathrm{X}_i$

(δ) Γνωρίζουμε ότι για το σωστό μοντέλο ισχύει: $\mathrm{E}(SS_{res}) = (n-2) \cdot \sigma^2$

Για το λανθασμένο μοντέλο ισχύει:

$\mathrm{Y}_i - \hat{\mathrm{Y}}_i = [(\mathrm{Y}_i - \hat{\mathrm{Y}}_i) - \mathrm{E}(\mathrm{Y}_i - \hat{\mathrm{Y}}_i)] + \mathrm{E}(\mathrm{Y}_i - \hat{\mathrm{Y}}_i) = q_i + \mathrm{B}_i$, όπου

$q_i = [(\mathrm{Y}_i - \hat{\mathrm{Y}}_i) - \mathrm{E}(\mathrm{Y}_i - \hat{\mathrm{Y}}_i)]$ και $\mathrm{B}_i = \mathrm{E}(\mathrm{Y}_i - \hat{\mathrm{Y}}_i)$.

Άρα, $SS_{res} = \sum_{i=1}^{n}(\mathrm{Y}_i - \hat{\mathrm{Y}}_i)^2 = \sum_{i=1}^{n}(q_i + \mathrm{B}_i)^2 = \sum_{i=1}^{n} q_i^2 + \sum_{i=1}^{n} \mathrm{B}_i^2 + 2 \cdot \sum_{i=1}^{n} q_i \cdot \mathrm{B}_i$.

Άρα, $\mathrm{E}(SS_{res}) = \sum_{i=1}^{n} \mathrm{E}(q_i^2) + \sum_{i=1}^{n} \mathrm{B}_i^2 + 2 \cdot \sum_{i=1}^{n} \mathrm{B}_i \cdot \mathrm{E}(q_i) = \sum_{i=1}^{n} \mathrm{E}(q_i^2) + \sum_{i=1}^{n} \mathrm{B}_i^2$,

αφού $\mathrm{E}(q_i) = 0$.

Αλλά, $\mathrm{E}(q_i^2) = Var(\mathrm{Y}_i - \hat{\mathrm{Y}}_i) = Var(\mathrm{Y}_i) + Var(\hat{\mathrm{Y}}_i) - 2 \cdot Cov(\mathrm{Y}_i, \hat{\mathrm{Y}}_i) =$

$$= \sigma^2 + \left[\frac{1}{n} + \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 - 2 \cdot Cov\left(\mathrm{Y}_i, \overline{\mathrm{Y}} + \hat{\beta}_1 \cdot (\mathrm{X}_i - \overline{\mathrm{X}})\right) =$$

$$= \sigma^2 + \left[\frac{1}{n} + \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 - 2 \cdot Cov\left(\mathrm{Y}_i, \overline{\mathrm{Y}}\right) - 2 \cdot (\mathrm{X}_i - \overline{\mathrm{X}}) \cdot Cov\left(\mathrm{Y}_i, \hat{\beta}_1\right) =$$

$$= \sigma^2 + \left[\frac{1}{n} + \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 - 2 \cdot \frac{\sigma^2}{n} - 2 \cdot (\mathrm{X}_i - \overline{\mathrm{X}}) \cdot Cov\left(\mathrm{Y}_i, \sum_{i=1}^{n} \kappa_i \cdot \mathrm{Y}_i\right) =$$

$$= \sigma^2 + \left[\frac{1}{n} + \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 - 2 \cdot \frac{\sigma^2}{n} - 2 \cdot \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \cdot \sigma^2 =$$

$$= \left[1 - \frac{1}{n} - \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2$$

Άρα τελικά,

$$\mathrm{E}(SS_{res}) = \sum_{i=1}^{n} \left[1 - \frac{1}{n} - \frac{(\mathrm{X}_i - \overline{\mathrm{X}})^2}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2} \right] \cdot \sigma^2 + \sum_{i=1}^{n} \mathrm{B}_i^{\,2} = (n - 1 - 1) \cdot \sigma^2 + \sum_{i=1}^{n} \mathrm{B}_i^{\,2} \Rightarrow$$

$$\Rightarrow \mathrm{E}(MS_{res}) = \sigma^2 + \frac{1}{n-2} \cdot \sum_{i=1}^{n} \mathrm{B}_i^{\,2}$$

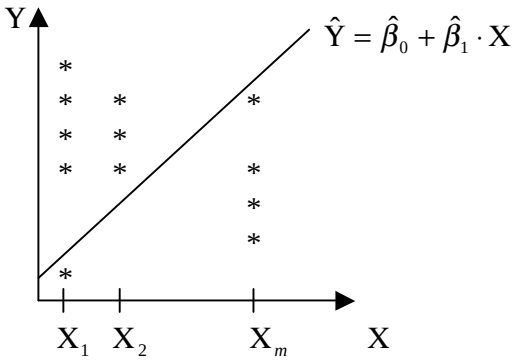
Άρα για το λανθασμένο μοντέλο, η ποσότητα MS_{res} δεν εκτιμά την σ^2, αλλά μια ποσότητα μεγαλύτερη.

Αυτό επηρεάζει τις διακυμάνσεις των $\hat{\beta}_0, \hat{\beta}_1$, αφού $\hat{Var}\left(\hat{\beta}_1\right) = \dfrac{MS_{res}}{\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2}$.

Γι' αυτό το λόγο χρησιμοποιούμε τον παρακάτω έλεγχο ορθότητας υποθέτοντας ότι έχουμε επαναλαμβανόμενες παρατηρήσεις:

$\mathrm{X} = \mathrm{X}_1$	$\mathrm{Y}_{11}, \mathrm{Y}_{12}, ..., \mathrm{Y}_{1n_1}$	(τ.δ.)	$\overline{\mathrm{Y}}_1$
$\mathrm{X} = \mathrm{X}_2$	$\mathrm{Y}_{21}, \mathrm{Y}_{22}, ..., \mathrm{Y}_{2n_2}$	(τ.δ.)	$\overline{\mathrm{Y}}_2$

.	.		.
.	.		.
.	.		.
$X = X_m$	$Y_{m1}, Y_{m2}, ..., Y_{mn_m}$	(τ.δ.)	$\overline{Y}_m$



Θεωρώ το άθροισμα των υπολοίπων:

$$SS_{res} = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \hat{Y}_i\right)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i + \overline{Y}_i - \hat{Y}_i\right)^2 =$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2 + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(\overline{Y}_i - \hat{Y}_i\right)^2 + 2\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)\left(\overline{Y}_i - \hat{Y}_i\right) =$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2 + \sum_{i=1}^{m} n_i \cdot \left(\overline{Y}_i - \hat{Y}_i\right)^2 + 2\cdot\sum_{i=1}^{m}\left(\overline{Y}_i - \hat{Y}_i\right)\cdot\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right) =$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2 + \sum_{i=1}^{m} n_i \cdot \left(\overline{Y}_i - \hat{Y}_i\right)^2$$

Παρατηρώντας ότι $\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2$ εκφράζει την μεταβλητότητα του i δείγματος, έχουμε ότι $\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2$ είναι το άθροισμα των τετραγώνων που οφείλεται σε γνήσιο σφάλμα (SS of pure error, SS_{pe}).

Για τον δεύτερο όρο στο παραπάνω άθροισμα, έχω ότι $E\left(\overline{Y}_i\right) = \beta_0 + \beta_1 \cdot X$ όταν ισχύει το μοντέλο, αλλά $E\left(\hat{Y}_i\right) = E\left(\hat{\beta}_0 + \hat{\beta}_1 \cdot X_i\right) = \beta_0 + \beta_1 \cdot X_i$.

Άρα για το σωστό μοντέλο τα $\overline{Y}_i$ είναι κοντά στα $\hat{Y}_i$.

Επομένως για το σωστό μοντέλο, αν το $\sum_{i=1}^{m} n_i \cdot \left(\overline{Y}_i - \hat{Y}_i\right)^2$ είναι μικρό, τότε έχω ένδειξη για την ορθότητα του μοντέλου. Το παραπάνω άθροισμα ονομάζεται άθροισμα τετραγώνων που οφείλεται στην έλλειψη προσαρμογής (SS of lack of fit, SS_{lof}).

Γενικά έχω ότι: $SS_{res} = SS_{pe} + SS_{lof}$

$(n-2)$ $(n-m)$ $(m-2)$ (βαθμοί ελευθερίας)

(Όπου οι β.ε. $(n-m)=\sum_{i=1}^{m}(n_i-1)$, αφού $n=\sum_{i=1}^{m}n_i$.)

Όταν ισχύει η κανονικότητα, μπορεί εύκολα να δειχθεί ότι $\frac{SS_{lof}}{\sigma^2}\sim X^2_{m-2}$.

Άρα απορρίπτω την H_0 : "το μοντέλο είναι σωστό", αν $\frac{SS_{lof}}{\sigma^2}>c$.

Επειδή όμως η παράμετρος σ^2 είναι άγνωστη, την εκτιμούμε με $\hat{\sigma}^2=MS_{pe}=\frac{SS_{pe}}{n-m}$,

οπότε θεωρούμε την ελεγχοσυνάρτηση $F=\frac{MS_{lof}}{MS_{pe}}=\frac{SS_{lof}/m-2}{SS_{pe}/n-m}\sim F_{m-2,n-m}$.

Άρα απορρίπτεται η H_0, σε επίπεστο σημαντικότητας α , αν $F>F_{m-2,n-m,\alpha}$.

<u>Πρόταση:</u> $E(MS_{pe})=\sigma^2$, δηλαδή MS_{pe} είναι αμερόληπτη εκτιμήτρια της σ^2.

<u>Απόδειξη:</u> Παραλείπεται.

<u>ΠΙΝΑΚΑΣ ΑΝΑΔΙΑ</u>

Πηγή Μεταβλητήτητας	Άθροισμα Τετραγώνων	Βαθμοί Ελευθερίας	Μέσα Τετράτωνα	F - test
Παλινδρόμηση	SS_{reg}	1	$MS_{reg}=\frac{SS_{reg}}{1}$	$F=\frac{MS_{reg}}{MS_{res}}$
Υπόλοιπα	SS_{res}	$n-2$	$MS_{res}=\frac{SS_{res}}{n-2}$	
Έλλειψη Προσαρμογής	SS_{lof}	$m-2$	$MS_{lof}=\frac{SS_{lof}}{m-2}$	$F=\frac{MS_{lof}}{MS_{pe}}$
Γνήσιο Σφάλμα	SS_{pe}	$n-m$	$MS_{pe}=\frac{SS_{pe}}{n-m}$	
Ολική Μεταβλητότητα	SS_{tot}	$n-1$		

<u>Παρατηρήσεις:</u>

1) Πρώτα εκτελό το δεύτερο F - test του πίνακα ΑΝΑΔΙΑ, $F=\frac{MS_{lof}}{MS_{pe}}$, για τον έλεγχο ορθότητας του μοντέλου. Αν δεχτώ αυτό το F - test, τότε εκτελώ και το

πρώτο F - test του πίνακα ΑΝΑΔΙΑ, $F = \frac{MS_{reg}}{MS_{res}}$, για τον έλεγχο ύπαρξης ανεξαρτησίας μεταξύ των X και των παρατηρήσεων Y.

2) Η παραπάνω ανάλυση είναι δυνατή μόνο αν έχω επαναλαμβανόμενες μετρήσεις.

3) Η ποσότητα MS_{pe} είναι πάντα αμερόληπτη εκτιμήτρια για τη σ^2, ανεξαρτήτως του μοντέλου. Αυτό ισχύει γιατί οι επαναλαμβανόμενες μετρήσεις, επιτρέπουν την εύρεση εκτιμήτριας που να μην εξαρτάται από το μοντέλο.

4) Υπάρχει αναλογία μεταξύ: (α) $\mathrm{E}(SS_{res}) = (n-2)\cdot\sigma^2 + \sum_{i=1}^{n} \mathrm{B}_i^{\,2}$

και (β) $SS_{res} = SS_{pe} + SS_{lof}$.