Η ανάλυση της διακύμανσης (στην απλή γραμμική παλινδρόμηση)

$$y_i-\bar{y}=y_i-\hat{y}_i+\hat{y}_i-\bar{y}$$

$$\underbrace{\sum_{i=1}^{n}(y_i-\bar{y})^2}_{\text{ολική μεταβλητότητα, ολ. αθρ. τετραγώνων}}=\underbrace{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}_{\text{Αθρ. τετρ. υπολοίπων}}+\underbrace{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}_{\text{Αθρ. τετρ. παλινδρόμ.}}+\cancel{2\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y})}\ \ 0$$

- $\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y})=\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)(\hat{\beta}_0+\hat{\beta}_1x_i-\bar{y})=\sum_{i=1}^{n}(y_i-\bar{y}+\hat{\beta}_1\bar{x}-\hat{\beta}_1x_i)(\cancel{\bar{y}}-\hat{\beta}_1\bar{x}+\hat{\beta}_1x_i-\cancel{\bar{y}})$
$=\sum_{i=1}^{n}[(y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\,\hat{\beta}_1(x_i-\bar{x})]=\hat{\beta}_1\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})-\hat{\beta}_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2$
$=\hat{\beta}_1\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})-\hat{\beta}_1\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sum_{i=1}^{n}(x_i-\bar{x})^2=0$

- $\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{\beta}_0+\hat{\beta}_1x_i-\bar{y})^2=\sum_{i=1}^{n}(\bar{y}-\hat{\beta}_1\bar{x}+\hat{\beta}_1x_i-\bar{y})$
$=\hat{\beta}_1^2\sum(x_i-\bar{x})^2$

Για το ολ. αθρ. τετ. έχουμε: $n-1$ βαθμούς ελευθερίας γιατί $\sum(y_i-\bar{y})=0$

Για το αθρ. τετρ. παλινδρ. έχουμε: 1 βαθμό ελευθερίας

Για το αθρ. τετρ. υπολοίπων έχουμε: $n-2$ βαθμοί ελευθερίας

Προέλευση Μεταβλητότητας	Άθροισμα (SS) Τετραγώνων	Βαθμ. Ελευθ.	Μέσο (MS) τετράγωνο	F-τεστ
Παλινδρόμηση	$SS_{reg}=\sum(\hat{y}_i-\bar{y})^2$	1	$MS_{reg}=\frac{SS_{reg}}{1}$	$F=\frac{MS_{reg}}{MS_{res}}$
Υπόλοιπα	$SS_{res}=\sum(y_i-\hat{y}_i)^2$	$n-2$	$MS_{res}=\frac{SS_{res}}{n-2}$	κρ. περ. $F\geq F_{a,1,n-2}$
Ολική μεταβλητότητα	$SS_{tot}=\sum(y_i-\bar{y})^2$	$n-1$	$R^2=\frac{SS_{reg}}{SS_{tot}}$ (: Συντελεστής προσδιορισμού)	
Παλινδρόμηση (EBO):	13600	1	13600	$F=1813,3$
Υπόλοιπα (EBO):	60	8	7,5	και κρ. περ. : 5,32
Ολ. μεταβλ. (EBO):	13660	9	$\frac{13600}{13660}$	απορ. την H_0

Είναι: $E(MS_{res})=\sigma^2$ (πάντα)
Θα δείξε: $E(MS_{reg})=\sigma^2+\beta_1^2\sum(x_i-\bar{x})^2$ $\Bigg\}\Rightarrow \dfrac{MS_{reg}}{MS_{res}}\to 1$ αν $\beta_1=0$

$\dfrac{(n-2)S^2}{\sigma^2}\sim X^2_{n-2}$ δηλ. $\dfrac{SS_{res}}{\sigma^2}\sim X^2_{n-2}$

$\dfrac{SS_{res}}{\sigma^2}\overset{\beta_1=0}{\sim}X_1^2$ $\Bigg\}\Rightarrow \dfrac{MS_{reg}}{MS_{res}}=\dfrac{\frac{SS_{reg}}{\sigma^2}/1}{\frac{SS_{res}}{\sigma^2}/(n-2)}\overset{H_0}{\sim}F_{1,n-2}$ και

κρ. περ. $F\geq F_{a,1,n-2}$

$$MS_{reg} = \hat{\beta}_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$E(MS_{reg}) = \sum_{i=1}^{n} (x_i - \bar{x})^2 \left[E(\hat{\beta}_1)^2 \right] = \sum_{i=1}^{n} (x_i - \bar{x})^2 \left[Var\hat{\beta}_1 + (E\hat{\beta}_1)^2 \right] =$$

$$= \sum (x_i - \bar{x})^2 \left[\frac{\sigma^2}{\sum (x_i - \bar{x})^2} + \beta_1^2 \right] = \sigma^2 + \beta_1^2 \sum (x_i - \bar{x})^2$$

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum (x_i - \bar{x})^2}\right) \rightsquigarrow \frac{\hat{\beta}_1 - \beta_1}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}} \sim N(0,1) \rightsquigarrow \underline{\qquad\qquad}$$

Ανάλυση Υπολοίπων (Έλεγχος με καταν. των μοντέλων)

Υπ. $\varepsilon_i \sim N(0, \sigma^2)$, $\sigma^2 \ \forall x$ $\quad \frac{\varepsilon_i}{\sigma} \sim N(0,1)$

$\frac{e_i}{\sqrt{MS_{res}}} \overset{\text{προσ.}}{\sim} N(0,1)$. Το 95% των $\frac{e_i}{\sqrt{MS_{res}}}$ να είναι $[-2,2]$

Λογαριθμικός Μετασχηματισμός

$y = \gamma_0 \cdot \gamma_1^x \cdot \varepsilon$, $E(\varepsilon) = 1$, $Var(\varepsilon) = \sigma^2$

με $\beta_0 = \log_{10}\gamma_0$, $\beta_1 = \log_{10}\gamma_1$, $\varepsilon' = \log_{10}\varepsilon$, $y' = \log_{10} y$

$y' = \beta_0 + \beta_1 x + \varepsilon'$: απλό γραμ. μοντέλο $(x, y' = \log_{10} y)$

Αντίστροφος Μετασχηματισμός

$y = \beta_0 + \frac{\beta_1}{x} + \varepsilon$ με $(x' = \frac{1}{x}, y) \rightsquigarrow y = \beta_0 + \beta_1 x' + \varepsilon$