(1) ΣΥΝΤΕΛΕΣΤΗΣ ΠΡΟΣΔΙΟΡΙΣΜΟΥ

ΚΡΙΤΗΡΙΟ ΠΟΥ ΜΑΣ ΒΟΗΘΑ ΝΑ ΕΞΕΤΑΣΟΥΜΕ ΤΟ "ΠΟΣΟ ΚΑΛΑ" Η ΕΥΘΕΙΑ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΕΡΜΗΝΕΥ-ΕΙ ΤΑ ΔΕΔΟΜΕΝΑ.

ΠΟΙΟ ΤΟ ΠΟΣΟΣΤΟ ΤΗΣ ΔΙΑΣΠΟΡΑΣ ΤΟΥ y (ΓΥΡΩ ΑΠΟ $\bar{y}$) ΠΟΥ ΜΠΟΡΕΙ ΝΑ ΑΠΟΔΟΘΕΙ ΣΤΗΝ ΓΝΩΣΗ ΤΟΥ x; (ΔΗΛΑΔΗ ΠΟΥ ΜΠΟΡΕΙ ΝΑ ΕΡΜΗΝΕΥΤΕΙ ΑΠΟ ΤΟ x).

ΈΣΤΩ ΤΥΧΑΙΟ ΔΕΙΓΜΑ ΑΠΟ n ΖΕΥΓΗ ΠΑΡΑΤΗΡΗΣΕΩΝ $\{(x_i, y_i), i=1,\dots,n\}$ ΣΤΙΣ ΟΠΟΙΕΣ ΠΡΟΣΑΡΜΟΣΑΜΕ ΤΗΝ ΕΥΘΕΙΑ $\hat{y} = a + bx$ ΜΕ ΜΕΘΟΔΟ OLS.

ΕΣΤΩ $\bar{y}$ Ο ΑΡΙΘΜ. ΜΕΣΟΣ ΤΩΝ ΤΙΜΩΝ y_i. Η ΣΥΝΟΛΙΚΗ ΤΕΤΡΑΓΩΝΙΚΗ ΑΠΟΚΛΙΣΗ ΤΩΝ y_i ΑΠΟ ΤΟ $\bar{y}$ ΕΚΦΡΑΖΕΤΑΙ ΑΠΟ $\sum_{i=1}^{n}(y_i-\bar{y})^2 = S_{yy} = SST$ → ΣΥΜΒΟΛΙΣΜΟΣ

ΟΝΟΜΑΖΕΤΑΙ ΣΥΝΟΛΙΚΟ ΑΘΡΟΙΣΜΑ ΤΕΤΡΑΓΩΝΩΝ

②

ΜΠΟΡΟΥΜΕ ΝΑ ΓΡΑΨΟΥΜΕ:

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

$$\Rightarrow \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} \left[(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}) \right]^2$$

$$= \sum_{i=1}^{n} \left[(y_i - \hat{y}_i)^2 + 2(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) + (\hat{y}_i - \bar{y})^2 \right]$$

ΜΠΟΡΟΥΜΕ ΝΑ ΥΠΟΛΟΓΙΣΟΥΜΕ

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})$$

ΓΝΩΡΙΖΟΝΤΑΣ ΑΠΟ ΤΑ ΠΡΟΗΓΟΥΜΕΝΑ (ΒΛΕΠΕ OLS) ΟΤΙ

$$\hat{y}_i = \bar{y} + b(x_i - \bar{x})$$

ΣΥΝΕΠΩΣ

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})$$

$$= \sum_{i=1}^{n} (y_i - \bar{y} - b(x_i - \bar{x}))(b(x_i - \bar{x}))$$

$$= b \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) - b^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= b S_{xy} - b^2 S_{xx} = b(S_{xy} - b S_{xx})$$

ΟΜΩΣ $b = \frac{S_{xy}}{S_{xx}} \Rightarrow \frac{S_{xy}}{S_{xx}} \left(S_{xy} - \frac{S_{xy}}{\cancel{S_{xx}}} \cancel{S_{xx}} \right) = 0$

3

ΔΗΛΑΔΗ $\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y})=0$

ΚΑΙ ΣΥΝΕΠΩΣ

$$SST=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$$

ΟΝΟΜΑΖΟΥΜΕ

$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = SSE$ (ΑΘΡΟΙΣΜΑ ΤΕΤΡΑΓΩΝΩΝ ΚΑΤΑΛΟΙΠΩΝ)

$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2 = SSR$ (ΑΘΡΟΙΣΜΑ ΤΕΤΡΑΓΩΝΩΝ ΠΑΛΙΝΔΡΟΜΗΣΗΣ)

ΔΗΛΑΔΗ: $\boxed{SST=SSR+SSE}$

ΤΟ ΣΥΝΟΛΙΚΟ ΑΘΡΟΙΣΜΑ ΤΕΤΡΑΓΩΝΩΝ ΚΑΤΑΛΟΙΠΩΝ ΜΠΟΡΕΙ ΝΑ ΑΠΟΔΟΘΕΙ ΣΤΟΥΣ ΠΑΡΑΠΑΝΩ ΔΥΟ ΠΑΡΑΓΟΝΤΕΣ

ΣΥΝΕΠΩΣ ΤΟ ΠΟΣΟΣΤΟ ΤΗΣ ΔΙΑΣΠΟΡΑΣ ΤΟΥ Υ ΠΟΥ ΜΠΟΡΕΙ ΝΑ ΑΠΟΔΟΘΕΙ ΣΤΗΝ ΓΝΩΣΗ ΤΟΥ Χ ΕΙΝΑΙ $\frac{SSR}{SST}$ ΚΑΙ ΣΥΜΒΟΛΙΖΕΤΑΙ ΜΕ R^2.

$R^2 = \frac{SSR}{SST}$ ΟΝΟΜΑΖΕΤΑΙ ΣΥΝΤΕΛΕΣΤΗΣ ΠΡΟΣΔΙΟΡΙΣΜΟΥ.

ΠΛΕΟΝΕΚΤΗΜΑ: ΔΕΝ ΕΧΕΙ ΜΟΝΑΔΕΣ ΜΕΤΡΗΣΗΣ. ΕΠΙΠΛΕΟΝ ΥΠΟΛΟΓΙΖΕΤΑΙ ΣΤΑ ΠΕΡΙΣΣΟΤΕΡΑ ΣΤΑΤΙΣΤΙΚΑ/ΟΙΚΟΝΟΜΕΤΡΙΚΑ ΛΟΓΙΣΜΙΚΑ

ΑΝ ΕΥΘΕΙΑ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΠΡΟΣΑΡΜΟΖΕΤΑΙ "ΚΑΛΑ" ΣΤΑ ΔΕΔΟΜΕΝΑ ΤΟΤΕ R^2 ΕΙΝΑΙ ΚΟΝΤΑ ΣΤΟ 1,

ΕΙΝΑΙ ΚΟΝΤΑ ΣΤΟ 0.

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}(a + bx_i - (a + b\bar{x}))^2}{S_{yy}}$$

$$= \frac{b^2 \sum_{i=1}^{n}(x_i - \bar{x})^2}{S_{yy}} = \frac{\frac{S_{xy}^2}{S_{xx}^2} S_{xx}}{S_{yy}}$$

$$= \boxed{\frac{S_{xy}^2}{S_{xx} \cdot S_{yy}}}$$ ⇐ ΣΥΝΗΘΩΣ ΑΥΤΟ ΧΡΗΣΙΜΟΠΟΙΟΥΜΕ ΓΙΑ ΠΡΑΚΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

ΓΡΑΦΕΤΑΙ ΚΑΙ ΩΣ

$$R^2 = b^2 \frac{S_{xx}}{S_{yy}} = b \frac{S_{xy}}{S_{yy}}$$

ΕΠΙΠΛΕΟΝ

$$0 \leq R^2 \leq 1$$

Γ

ΕΦΑΡΜΟΓΗ ΣΤΟ ΠΑΡΑΔΕΙΓΜΑ "ΝΕΡΟ – ΣΟΔΕΙΑ".

$$R^2 = \frac{(103.68)^2}{(1008)(11.28)} \simeq 0.946$$

ΠΟΛΥ ΚΟΝΤΑ ΣΤΟ 1

$\Rightarrow$ "ΚΑΛΗ" ΠΡΟΣΑΡΜΟΓΗ

<u>ΕΡΜΗΝΕΙΑ</u>: Η ΠΟΣΟΤΗΤΑ ΝΕΡΟΥ ΠΟΥ ΡΙΞΑΜΕ ΕΡΜΗΝΕΥΕΙ ΠΕΡΙΠΟΥ ΤΟ 95% ΤΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ ΤΗΣ ΠΑΡΑΓΟΜΕΝΗΣ ΣΟΔΕΙΑΣ (ΤΡΙΦΥΛΛΙ). ΤΟ ΥΠΟΛΟΙΠΟ (ΑΝΕΡΜΗΝΕΥΤΟ) 5% ΟΦΕΙΛΕΤΑΙ ΣΕ ΤΥΧΑΙΟΥΣ ΠΑΡΑΓΟΝΤΕΣ.