

Άσκηση: Δίνονται οι τιμές 8 παρατηρήσεων των μεταβλητών Y και X.

Περίοδος:	1	2	3	4	5	6	7	8
Y:	12	13,5	14	12,4	13,2	12,9	14,5	14,1
X:	1,8	2,4	2,8	2,3	2,8	2,4	3,8	2,9

α) Να προσδιορίσετε και να ερμηνεύσετε το συντελεστή γραμμικής συσχέτισης r.

Απάντηση

	X	Y	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X}) \cdot (Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$
	1,8	12	-0,850	-1,325	1,126	0,723	1,756
	2,4	13,5	-0,250	0,175	-0,044	0,063	0,031
	2,8	14	0,150	0,675	0,101	0,023	0,456
	2,3	12,4	-0,350	-0,925	0,324	0,123	0,856
	2,8	13,2	0,150	-0,125	-0,019	0,023	0,016
	2,4	12,9	-0,250	-0,425	0,106	0,063	0,181
	3,8	14,5	1,150	1,175	1,351	1,323	1,381
	2,9	14,1	0,250	0,775	0,194	0,063	0,601
Σύνολο	21,2	106,6			3,14	2,4	5,275

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \cdot \sqrt{\sum (Y_i - \bar{Y})^2}} = \frac{3,14}{\sqrt{2,4} \cdot \sqrt{5,275}} = 0,882$$

Ο συντελεστής συσχέτισης είναι θετικός και κοντά στην μονάδα άρα θετική ισχυρή συσχέτιση.

β) Να διενεργήσετε το στατιστικό έλεγχο ότι ο συντελεστής γραμμικής συσχέτισης του πληθυσμού των δύο μεταβλητών είναι μηδέν χρησιμοποιώντας ένα επίπεδο σημαντικότητας 5%. (κ. τιμή=2,45)

$$H_0: \rho = 0, \text{ Στατιστική Ελέγχου: } t = \frac{r}{\sqrt{1-r^2}/\sqrt{n-2}} = \frac{0,882}{\sqrt{1-0,882^2}/\sqrt{8-2}} = \frac{0,882}{0,192} = 4,58$$

Με εναλλακτική υπόθεση,

• $H_1: \rho \neq 0$ η H_0 απορρίπτεται σε επίπεδο σημαντικότητας 0,05 καθώς

$$|t| = 4,58 > \text{κριτική τιμή} = t_{n-2, \alpha/2} = t_{6, 0,025} = 2,45$$

Άρα ο συντελεστής συσχέτισης είναι στατιστικά διάφορος του μηδενός.

γ) Να εκτιμήσετε και να ερμηνεύσετε τους συντελεστές του γραμμικού υποδείγματος $Y = \alpha + \beta X + \varepsilon$.

Απάντηση

$$\hat{\beta} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{3,14}{2,4} = 1,3083 \text{ και } \hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = \frac{106,6}{8} - 1,3083 \cdot \frac{21,1}{8} = 9,8579$$

$$\text{άρα } \hat{Y} = 9,8579 + 1,3083 \cdot X$$

$$\text{Παρατήρηση: Χωρίς σταθερό όρο: } \hat{\beta} = \frac{\sum X_i \cdot Y_i}{\sum X_i^2} = \frac{285,63}{58,58} = 4,876 \text{ άρα } \hat{Y} = 4,876 \cdot X$$

δ) Να προσδιορίσετε και να ερμηνεύσετε την τιμή του συντελεστή προσδιορισμού R^2 της παλινδρόμησης. Να ελεγχθεί η στατιστική σημαντικότητα της παλινδρόμησης (κριτική τιμή 5,99)

(Αντίστοιχα να ελεγχθεί η υπόθεση ότι ο συντελεστής προσδιορισμού ισούται με το μηδέν)

Ο συντελεστής προσδιορισμού ισούται με το τετράγωνο του συντελεστή συσχέτισης, άρα

$$R^2 = r^2 = (0,882)^2 = 0,778$$

Έλεγχος F

$H_0: \beta = 0$ με εναλλακτική υπόθεση

$H_1: \beta \neq 0$,

$$\text{αφού } F = \frac{R^2 / k}{(1-R^2)/(n-k-1)} = \frac{0,778/1}{(1-0,778)/8-1-1} = \frac{0,778}{0,037} = 21$$

απόφαση: $F = 21 > F_{1,6,0.05} = 5,99$ άρα απορρίπτεται η H_0 . σημαντική η παλινδρόμηση (και $R^2 \neq 0$)

ε) Να ελεγχθεί η στατιστική σημαντικότητα του β και να κατασκευαστεί 95% διάστημα εμπιστοσύνης του συντελεστή. (κριτική τιμή 2,44)

$H_0: \beta = 0$

$H_1: \beta \neq 0$

Έλεγχος με τη t κατανομή,

$$t = \frac{\hat{\beta}}{se(\hat{\beta})} \text{ όπου } se(\hat{\beta}) = \sqrt{\frac{s^2}{\sum (X_i - \bar{X})^2}} \cdot s^2 = \frac{SSE}{n-2}$$

Εύρεση SSE:

$$SST = \sum (Y_i - \bar{Y})^2 = 5,275$$

$$\text{Καθώς } R^2 = \frac{SSR}{SST} \Rightarrow SSR = SST \cdot R^2 = 5,275 \cdot 0,778 = 4,104,$$

$$\text{άρα } SSE = SST - SSR = 5,275 - 4,104 = 1,171$$

$$s^2 = 1,171/8 = 0,146, \text{ και από πίνακα } \sum (X_i - \bar{X})^2 = 2,4$$

$$se(\hat{\beta}) = \sqrt{\frac{0,146}{2,4}} = 0,247 \text{ άρα } t = \frac{1,3083}{0,247} = 5,29$$

$|t| = 5,29 > t_{6,0.025} = 2,44$ άρα απορρίπτεται η μηδενική υπόθεση έναντι της $H_1: \beta \neq 0$.

95% Διάστημα εμπιστοσύνης

$$\hat{\beta} - t_{n-2,\alpha/2} \cdot se(\hat{\beta}) \leq \beta \leq \hat{\beta} + t_{n-2,\alpha/2} \cdot se(\hat{\beta}) = 1,3083 - 2,44 \cdot 0,247 \leq \beta \leq 1,3083 + 2,44 \cdot 0,247 = 1,3083 - 0,603 \leq \beta \leq 1,3083 + 0,603 = (0.705, 1.911)$$

στ) Εάν η τιμή της εξαρτημένης μεταβλητής X για την επόμενη περίοδο ορισθεί στις 3,1 χρηματικές μονάδες, να προσδιορίσετε την προβλεπόμενη τιμή και το 90% διάστημα εμπιστοσύνης της εξαρτημένης μεταβλητής. (κριτική τιμή 1,943)

Ατομική πρόβλεψη 90% Διάστημα εμπιστοσύνης

$$\hat{Y}_0 - t_{n-2,\alpha/2} \cdot se(\hat{\epsilon}_0) \leq Y_0 | X_0 \leq \hat{Y}_0 + t_{n-2,\alpha/2} \cdot se(\hat{\epsilon}_0)$$

$$\text{όπου } \hat{Y}_0 | X_0 = \hat{\alpha} + \hat{\beta} \cdot X_0 = 9,8579 + 1,3083 \cdot 3,1 = 13,914 \text{ και}$$

$$se(\hat{\epsilon}_0) = s \cdot \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{S_{xx}}} \stackrel{s^2=0,146}{=} \sqrt{0,146} \cdot \sqrt{1 + \frac{1}{8} + \frac{(3,1 - 21,2/8)^2}{2,4}} = 0,382 \cdot \sqrt{1 + \frac{1}{8} + 0,084} = 0,382 \cdot 1,1 = 0,42$$

$$\text{Τελικά } 13,914 - 1,943 \cdot 0,42 \leq Y_0 | X_0 \leq 13,914 + 1,943 \cdot 0,42 = (13.1, 14.73)$$

ΓΙΑ ΝΑ ΛΑΜΒΑΝΕΤΕ ΕΝΗΜΕΡΩΣΕΙΣ ΑΚΟΛΟΥΘΗΣΤΕ ΜΑΣ ΣΤΟ FACEBOOK