

ΘΕΜΑΤΙΚΗ
ΕΝΟΤΗΤΑ
ΔΕΟ 13



Eclass4U

The best Choice for you

ΑΣΚΗΣΕΙΣ
ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

ΗΜΕΡΟΜΗΝΙΑ ΑΝΑΡΤΗΣΗΣ: 15-01-22

Επιμέλεια Ύλης: ΣΠΥΡΟΣ ΒΛΑΧΟΠΟΥΛΟΣ



ΘΕΡΜΟΠΥΛΩΝ 17
ΠΕΡΙΣΤΕΡΙ

100M ΑΠΟ ΤΗ ΣΤΑΣΗ
ΜΕΤΡΟ «ΠΕΡΙΣΤΕΡΙ»

ΤΗΛΕΦΩΝΟ: 210-5711484

ΚΙΝΗΤΟ: 6970401981

EMAIL: grammateia.eclass4u@gmail.com

ΤΟΠΟΘΕΣΙΑ WEB : www.eclass4u.gr

SOCIAL MEDIA:



1^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Από ένα τυχαίο δείγμα 5 οικογενειών προέκυψαν τα παρακάτω δεδομένα που παρουσιάζονται στον παρακάτω πίνακα και εκφράζουν το ετήσιο εισόδημα τους (X) και τις ετήσιες δαπάνες τους για τρόφιμα (Y). Οι μεταβλητές X και Y είναι εκφρασμένες σε χιλιάδες ευρώ.

Ετήσιο εισόδημα X	20	35	30	40	15
Ετήσιες δαπάνες για τρόφιμα Y	7	9	8	11	5

Με βάση τα στοιχεία αυτά και αν επιπλέον γνωρίζουμε ότι οι μεταβλητές X και Y συνδέονται με γραμμική σχέση:

- A) Να εκτιμηθεί η ευθεία παλινδρόμησης της Y πάνω στη X
- B) Να ερμηνευτούν οι συντελεστές της ευθείας παλινδρόμησης
- Γ) Να υπολογισθεί ο συντελεστής συσχέτισης και να ερμηνευτεί.
- Δ) Ποιες προβλέπετε να είναι οι ετήσιες δαπάνες για τρόφιμα μιας οικογένειας της οποίας το ετήσιο εισόδημα είναι 50.000€.

Eclass4U

The best Choice for you

Α τρόπος (με χρήση των τύπων με απόκλιση από τους μέσους)

Δημιουργούμε όλες τις απαραίτητες στήλες που χρειάζονται οι τύποι των β_0 , β_1 και r . Προφανώς χρειαζόμαστε τις τιμές των \bar{X} και \bar{Y} .

Όμως, όπως γνωρίζουμε από τα μη ομαδοποιημένα δεδομένα,

$$\bar{X} = \frac{\sum X_i}{n} = \frac{140}{5} = 28$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{40}{5} = 8$$

Μπορούμε, πλέον, να δημιουργήσουμε τον παρακάτω πίνακα με όλες τις απαραίτητες στήλες για την μελέτη του προβλήματος.

i	X_i	Y_i	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(Y_i - \bar{Y}) \cdot (X_i - \bar{X})$
1	20	7	-8	-1	64	1	8
2	35	9	7	1	49	1	7
3	30	8	2	0	4	0	0
4	40	11	12	3	144	9	36
5	15	5	-13	-3	169	9	39
Σύνολα	140	40			430	20	90

Α. Για τους συντελεστές της ευθείας παλινδρόμησης (εκτιμήτριες) έχουμε :

$$\hat{\beta}_1 = \frac{\sum[(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sum[(X_i - \bar{X})^2]} = \frac{90}{430} = 0,2093$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{X} = 8 - 0,2093 \cdot 28 = 8 - 5,8604 = 2,1396$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = 2,1396 + 0,2093 \cdot X$$

Β. Ερμηνεία των β_0 και β_1 (συντελεστών της ευθείας παλινδρόμησης):

Για την $\hat{\beta}_1 = 0,2093$:

Αν το ετήσιο οικογενειακό εισόδημα αυξηθεί κατά 1 χιλιάδα ευρώ (1.000€) τότε οι ετήσιες δαπάνες για τρόφιμα θα αυξηθούν κατά 0,2093 χιλιάδες ευρώ (209,30€)

Για την $\hat{\beta}_0 = 2,1396$:

Αν το ετήσιο οικογενειακό εισόδημα γίνει ίσο με μηδέν, τότε οι ετήσιες δαπάνες για τρόφιμα θα γίνουν ίσες με 2,1396 χιλιάδες ευρώ (2.139,60€)

Γ. Συντελεστής συσχέτισης

$$r = \frac{\sum[(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sqrt{\sum(Y_i - \bar{Y})^2 \cdot \sum(X_i - \bar{X})^2}} = \frac{90}{\sqrt{20 \cdot 430}} = \frac{90}{\sqrt{8600}} = \frac{90}{92,7362} = 0,97$$

Συμπεραίνουμε ότι οι μεταβλητές X και Y (δηλαδή το ετήσιο οικογενειακό εισόδημα και οι ετήσιες δαπάνες για τρόφιμα) **παρουσιάζουν ισχυρή θετική γραμμική συσχέτιση.**

Συντελεστής προσδιορισμού (δεν ζητείται από την άσκηση)

Ξέρουμε ότι $R^2 = r^2 = 0,97^2 = 0,9409$ ή 94,09%

Ερμηνεία συντελεστή προσδιορισμού : Το 94,09% της μεταβλητότητας των ετήσιων δαπανών για τρόφιμα ερμηνεύονται από το ετήσιο οικογενειακό εισόδημα.

Δ. Οι 50.000€, σε μονάδες μέτρησης της εκφώνησης είναι 50 χιλ €

Με αντικατάσταση στην εξίσωση της ευθείας παλινδρόμησης, έχουμε

$$Y = 2,1396 + 0,2093 \cdot 50 = 2,1396 + 10,465 = 12,6046 \text{ χιλιάδες ευρώ ή } 12.604,60\text{€}$$

Αν το ετήσιο εισόδημα γίνει 50.000 €, η **εκτίμηση** είναι ότι οι ετήσιες δαπάνες για τρόφιμα θα διαμορφωθούν σε 12.604,60€.

Β τρόπος (με χρήση των τύπων χωρίς απόκλιση από τους μέσους)

Δημιουργούμε, και πάλι, όλες τις απαραίτητες στήλες που χρειάζονται οι τύποι των β_0 , β_1 και r .

i	X_i	Y_i	X_i^2	Y_i^2	$X_i \cdot Y_i$
1	20	7	400	49	140
2	35	9	1225	81	315
3	30	8	900	64	240
4	40	11	1600	121	440
5	15	5	225	25	75
Σύνολα	140	40	4350	340	1210

Α. Για τους συντελεστές της ευθείας παλινδρόμησης (εκτιμήτριες) έχουμε :

$$\hat{\beta}_1 = \frac{\sum X_i \cdot Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{1210 - \frac{140 \cdot 40}{5}}{4350 - \frac{(140)^2}{5}} = \frac{1210 - 1120}{4350 - 3920} = \frac{90}{430} = 0,2093$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{X} = 8 - 0,2093 \cdot 28 = 8 - 5,8604 = 2,1396$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = 2,1396 + 0,2093 \cdot X$$

(Καταλήγουμε στο ίδιο ακριβώς αποτέλεσμα με αυτό του πρώτου τρόπου)

Β. Οι ερμηνείες των συντελεστών είναι ίδιες με του πρώτου τρόπου

Γ. Συντελεστής συσχέτισης

$$r = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n} \right] \cdot \left[\sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \right]}} = \frac{1210 - \frac{140 \cdot 40}{5}}{\sqrt{\left[4350 - \frac{(140)^2}{5} \right] \cdot \left[340 - \frac{(40)^2}{5} \right]}} =$$
$$= \frac{1210 - 1120}{\sqrt{[4350 - 3920] \cdot [340 - 320]}} = \frac{90}{\sqrt{430 \cdot 20}} = \frac{90}{92,7362} = 0,97$$

Συμπεραίνουμε και με τον δεύτερο τρόπο ότι οι μεταβλητές X και Y **παρουσιάζουν ισχυρή θετική γραμμική συσχέτιση.**

Δ. ομοίως με Α τρόπο

Συνολικά ο δεύτερος τρόπος είναι γρηγορότερος, ειδικά στο στάδιο κατασκευής των απαιτούμενων στηλών του πίνακα δεδομένων, για αυτό και προτιμάται (εφόσον δεν προσδιορίζει η άσκηση αν θα χρησιμοποιηθούν οι τύποι με απόκλιση από τους μέσους ή όχι)



Eclass4U

The best Choice for you

2^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Ένα μέλος ΣΕΠ του ΕΑΠ ενδιαφέρεται να μελετήσει τη συσχέτιση που υφίσταται μεταξύ του μέσου όρου των βαθμολογιών των 4 εργασιών στην ΔΕΟ 13 και του βαθμού του γραπτού στις τελικές εξετάσεις. Για το σκοπό αυτό βασίζεται σε ένα δείγμα 7, τυχαία επιλεγμένων, φοιτητών του. Τα δεδομένα για το εν λόγω δείγμα φοιτητών, συνοψίζονται στον πίνακα που ακολουθεί.

Μέσος όρος στις εργασίες (X)	Βαθμός γραπτού (Y)
9	4
7	7
8	4
9	4
8	5
10	2
7	7

- 1 . Να εκτιμήσετε τους συντελεστές α και β του γραμμικού υποδείγματος $Y = \alpha + \beta x + u$.
- 2 . Να ερμηνεύσετε την εκτίμηση του β .
- 3 . Να υπολογιστούν ο συντελεστής συσχέτισης και ο συντελεστής προσδιορισμού και να ερμηνευθούν.
- 4 . Με δεδομένο ότι κάποιος φοιτητής έχει στις εργασίες μέσο όρο βαθμολογίας 9, να προβλέψετε το ύψος της βαθμολογίας του γραπτού.

The best Choice for you

1. Δημιουργούμε όλες τις απαραίτητες στήλες, κι ο αρχικός πίνακας διαμορφώνεται ως εξής

i	Μέσος όρος στις εργασίες X_i	Βαθμός γραπτού Y_i	X_i^2	Y_i^2	$X_i \cdot Y_i$
1	9	4	81	16	36
2	7	7	49	49	49
3	8	4	64	16	32
4	9	4	81	16	36
5	8	5	64	25	40
6	10	2	100	4	20
7	7	7	49	49	49
Σύνολα	58	33	488	175	262

$$\bar{X} = \frac{\sum X_i}{n} = \frac{58}{7} = 8,29$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{33}{7} = 4,71$$

Άρα, για τους συντελεστές του γραμμικού υποδείγματος, έχουμε :

$$\hat{\beta} = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{262 - \frac{58 \cdot 33}{7}}{488 - \frac{(58)^2}{7}} = \frac{262 - 273,43}{488 - 480,57} = \frac{-11,43}{7,43} = -1,54$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = 4,71 - (-1,54) \cdot 8,29 = 4,71 + 12,77 = 17,48$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = 17,48 - 1,54 \cdot X$$

2 . Ερμηνεία του $\hat{\beta} = -1,54$

Αν ο μέσος όρος των εργασιών αυξηθεί κατά 1 μονάδα τότε ο βαθμός του γραπτού θα μειωθεί κατά 1,54 μονάδες.

3 . Συντελεστής συσχέτισης

$$r = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right] \cdot \left[\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}\right]}} = \frac{262 - \frac{58 \cdot 33}{7}}{\sqrt{\left[488 - \frac{(58)^2}{7}\right] \left[175 - \frac{(33)^2}{7}\right]}} =$$
$$= \frac{262 - 273,43}{\sqrt{[488 - 480,57] \cdot [175 - 155,57]}} = \frac{-11,43}{\sqrt{7,43 \cdot 19,43}} = -0,9509$$

Οι μεταβλητές X και Y (μέσος όρος στις εργασίες και βαθμός του γραπτού στις εξετάσεις) **παρουσιάζουν ισχυρή αρνητική γραμμική συσχέτιση.**

Συντελεστής προσδιορισμού

$$\Xi\acute{\epsilon}\rho\omicron\upsilon\mu\epsilon \acute{\omicron}\tau\iota R^2 = r^2 = (-0,9509)^2 = 0,9042 \text{ ή } 90,42\%$$

Ερμηνεία συντελεστή προσδιορισμού : Το 90,42% της μεταβλητότητας του βαθμού του γραπτού στις τελικές εξετάσεις ερμηνεύεται από τον μέσο όρο των εργασιών.

4 . Με αντικατάσταση στην εξίσωση της ευθείας παλινδρόμησης, έχουμε

$$Y = 17,48 - 1,54 \cdot 9 = 17,48 - 13,86 = 3,62$$

Αν ένας φοιτητής έχει μέσο όρο εργασιών ίσο με 9, η **εκτίμηση** είναι ότι στις τελικές εξετάσεις θα γράψει 3,62.

3^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Μια τράπεζα ενδιαφέρεται να μελετήσει την αποταμιευτική συμπεριφορά των πελατών της. Θεωρείται σαν δεδομένο ότι η ετήσια αποταμίευση των πελατών της εξαρτάται από το ετήσιο εισόδημά τους. Σε δείγμα 10 πελατών είχαμε τα ακόλουθα στοιχεία.

Ετήσια αποταμίευση σε χιλ € (Y)	Ετήσια εισοδήματα σε χιλ € (X)
5	50
3	31
6	28
4	45
6	50
1	32
2	36
8	55
2	26
3	47

1. Να εκτιμήσετε τους συντελεστές α και β του γραμμικού υποδείγματος $Y = \alpha + \beta x + u$.
2. Να ερμηνεύσετε την εκτίμηση του β .
3. Να υπολογιστούν ο συντελεστής συσχέτισης και ο συντελεστής προσδιορισμού R^2 και να ερμηνευθούν.
4. Με δεδομένο ότι κάποιος έχει εισόδημα 30 χιλ € να προβλέψετε το ύψος της αποταμίευσής του.

1. Δημιουργούμε όλες τις απαραίτητες στήλες, κι ο αρχικός πίνακας διαμορφώνεται ως εξής

i	Ετήσια αποταμίευση σε χιλ € Y_i	Ετήσια εισοδήματα σε χιλ € X_i	X_i^2	Y_i^2	$X_i \cdot Y_i$
1	5	50	2500	25	250
2	3	31	961	9	93
3	6	28	784	36	168
4	4	45	2025	16	180
5	6	50	2500	36	300
6	1	32	1024	1	32
7	2	36	1296	4	72
8	8	55	3025	64	440
9	2	26	676	4	52
10	3	47	2209	9	141
Σύνολα	40	400	17000	204	1728

$$\bar{X} = \frac{\sum X_i}{n} = \frac{400}{10} = 40$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{40}{10} = 4$$

Άρα, για τους συντελεστές του γραμμικού υποδείγματος, έχουμε :

$$\hat{\beta} = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{1728 - \frac{400 \cdot 40}{10}}{17000 - \frac{(400)^2}{10}} = \frac{1728 - 1600}{17000 - 16000} = \frac{128}{1000} = 0,128$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = 4 - 0,128 \cdot 40 = 4 - 5,12 = -1,12$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = -1,12 + 0,128 \cdot X$$

2. Ερμηνεία του $\hat{\beta} = 0,128$

Αν το ετήσιο εισόδημα αυξηθεί κατά 1 χιλιάδα ευρώ (1.000€) τότε η ετήσια αποταμίευση θα αυξηθεί κατά 0,128 χιλιάδες ευρώ (128€)

3 . Συντελεστής συσχέτισης

$$r = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right] \cdot \left[\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}\right]}} = \frac{1728 - \frac{400 \cdot 40}{10}}{\sqrt{\left[17000 - \frac{(400)^2}{10}\right] \left[204 - \frac{(40)^2}{10}\right]}} =$$
$$= \frac{1728 - 1600}{\sqrt{[17000 - 16000] \cdot [204 - 160]}} = \frac{128}{\sqrt{1000 \cdot 44}} = \frac{128}{209,76} = 0,6102$$

Οι μεταβλητές X και Y (ετήσιο εισόδημα και ετήσιες αποταμιεύσεις) **παρουσιάζουν θετική γραμμική συσχέτιση.**

Συντελεστής προσδιορισμού

Ξέρουμε ότι $R^2 = r^2 = 0,6102^2 = 0,3723$ ή 37,23%

Ερμηνεία συντελεστή προσδιορισμού : Το 37,23% της μεταβλητότητας των ετήσιων αποταμιεύσεων ερμηνεύεται από το ετήσιο εισόδημα.

4 . Με αντικατάσταση στην εξίσωση της ευθείας παλινδρόμησης, έχουμε

$$Y = -1,12 + 0,128 \cdot 30 = -1,12 + 3,84 = 2,72 \text{ χιλιάδες ευρώ ή } 2.720\text{€}$$

Αν το ετήσιο εισόδημα γίνει 30.000 €, η **εκτίμηση** είναι ότι οι ετήσιες αποταμιεύσεις θα διαμορφωθούν σε 2720€.

4^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Ένας ερευνητής εφάρμοσε τη μέθοδο της γραμμικής παλινδρόμησης για τις μεταβλητές X και Y , και υπολόγισε την εξίσωση $Y = 12,014 - 2,518 \cdot X$. Υπολόγισε επίσης ότι η διακύμανση της Y ήταν 301,6111 και αντίστοιχα της X ήταν 45,4333.

1. Να δώσετε την ερμηνεία των συντελεστών που υπολόγισε.
2. Να βρείτε την τιμή του συντελεστή συσχέτισης και να την ερμηνεύσετε.
3. Ένας άλλος ερευνητής εφαρμόζοντας τη μέθοδο της γραμμικής παλινδρόμησης βρήκε ότι η μεταβλητή Y σχετίζεται επίσης με την μεταβλητή Z και ότι η σχέση τους εκφράζεται με την εξίσωση

$Y = -7,483 + 1,386 \cdot Z$. Υπολόγισε τον συντελεστή συσχέτισης και βρήκε την τιμή 0,7087. Ποια από τις δυο εφαρμογές της μεθόδου είναι πιο χρήσιμη για τον υπολογισμό της Y και γιατί;

1. Εφόσον η εξίσωση της ευθείας παλινδρόμησης είναι η

$$Y = 12,014 - 2,518 \cdot X$$

έχουμε $\hat{\beta}_1 = -2,518$ και $\hat{\beta}_0 = 12,014$

Ερμηνεία $\hat{\beta}_1 = -2,518$: Αν η μεταβλητή X αυξηθεί κατά 1 μονάδα, τότε η μεταβλητή Y θα **μειωθεί** κατά 2,518 μονάδες.

Ερμηνεία $\hat{\beta}_0 = 12,014$: Αν η μεταβλητή X γίνει 0, τότε η μεταβλητή Y θα γίνει ίση με 12,014 μονάδες.

2. Συντελεστής συσχέτισης

Γνωρίζουμε ότι ισχύει ο τύπος

$$r = \frac{S_X}{S_Y} \cdot \beta_1$$

Συνεπώς, πρέπει να υπολογιστούν οι τυπικές αποκλίσεις S_X και S_Y δεδομένου ότι έχουμε τις διακυμάνσεις S_X^2 και S_Y^2 .

$$S_X = \sqrt{45,4333} = 6,74$$

$$S_Y = \sqrt{301,6111} = 17,367$$

Επομένως,

$$r = \frac{S_X}{S_Y} \cdot \beta_1 = \frac{6,74}{17,367} \cdot (-2,518) = -0,9772$$

Οι μεταβλητές X και Y παρουσιάζουν ισχυρή αρνητική γραμμική συσχέτιση.

3. Η σύγκριση θα γίνει μέσω του συντελεστή προσδιορισμού R^2

(Υπενθυμίζεται ότι όσο μεγαλύτερη είναι η τιμή του R^2 , τόσο καλύτερη είναι η προσαρμογή των δεδομένων στο γραμμικό μοντέλο ή αλλιώς τόσο καλύτερα το γραμμικό μοντέλο εκφράζει τα δεδομένα.)

Στην αρχική περίπτωση έχουμε

$$R^2 = r^2 = (-0,9772)^2 = 0,9549 \text{ ή } 95,49\%$$

Ενώ στην δεύτερη περίπτωση

$$R^2 = r_2^2 = (0,7087)^2 = 0,5023 \text{ ή } 50,23\%$$

Συμπεραίνουμε ότι στην πρώτη περίπτωση η μεταβλητή Y ερμηνεύεται πολύ πιο ικανοποιητικά από την X σε σχέση με την δεύτερη περίπτωση και την μεταβλητή Z. (η μεταβλητή X ερμηνεύει το 95,49% της μεταβλητότητας της Y, ενώ η Z ερμηνεύει μόλις το 50,23%)

Η πρώτη περίπτωση (μέσω της μεταβλητής X) είναι πιο χρήσιμη για τον υπολογισμό της Y.

5^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Σε μια μελέτη για την επίδραση των επιτοκίων πάνω στις επενδύσεις, χρησιμοποιήθηκε ένα δείγμα 12 παρατηρήσεων σε ζεύγη τιμών (X_i, Y_i) ($i = 1, 2, 3, \dots, 12$) όπου X η τιμή του μακροχρόνιου επιτοκίου επί τοις % και Y τα επενδυτικά κεφάλαια (σε δις ευρώ). Από τις τιμές του δείγματος προέκυψαν οι ακόλουθοι υπολογισμοί :

Δειγματική μέση τιμή των X_i	5,5
Δειγματική μέση τιμή των Y_i	34
$\sum X_i^2$	407
$\sum Y_i^2$	14150
$\sum X_i Y_i$	2143

Με χρήση του υποδείγματος γραμμικής παλινδρόμησης $Y = \beta_0 + \beta_1 X + u$

ζητούνται :

- 1 . Να εκτιμήσετε τους συντελεστές β_0 και β_1 της παλινδρόμησης με τη μέθοδο ελαχίστων τετραγώνων.
- 2 . Να υπολογίσετε το συντελεστή προσδιορισμού R^2 και να τον ερμηνεύσετε.

Eclass4U

The best Choice for you

1 . Παρατηρούμε δίνονται όλα τα απαραίτητα αθροίσματα του τύπου

$$\widehat{\beta}_1 = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

εκτός των $\sum X_i$ και $\sum Y_i$. Οι τιμές τους μπορούν να υπολογιστούν από τους αριθμητικούς μέσους, ως εξής :

$$\bar{X} = \frac{\sum X_i}{n} \Leftrightarrow \sum X_i = \bar{X} \cdot n = 5,5 \cdot 12 \Leftrightarrow \sum X_i = 66$$

$$\bar{Y} = \frac{\sum Y_i}{n} \Leftrightarrow \sum Y_i = \bar{Y} \cdot n = 34 \cdot 12 \Leftrightarrow \sum Y_i = 408$$

Πλέον, γνωρίζουμε όλα τα απαραίτητα αθροίσματα για την εύρεση των συντελεστών της ζητούμενης ευθείας παλινδρόμησης. Έτσι, έχουμε :

$$\widehat{\beta}_1 = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{2143 - \frac{66 \cdot 408}{12}}{407 - \frac{(66)^2}{12}} = \frac{2143 - 2244}{407 - 363} = \frac{-101}{44} = -2,30$$

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \cdot \bar{X} = 34 - (-2,30) \cdot 5,5 = 34 + 12,65 = 46,65$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = 46,65 - 2,3 \cdot X$$

Ερμηνεία των β_0 και β_1 (δε ζητούνται από την άσκηση):

Για την $\widehat{\beta}_1 = -2,3$: Αν το μακροχρόνιο επιτόκιο αυξηθεί κατά 1% τότε τα επενδυτικά κεφάλαια θα μειωθούν κατά 2,3 δις ευρώ

Για την $\widehat{\beta}_0 = 46,65$: Αν το μακροχρόνιο επιτόκιο γίνει ίσο με μηδέν, τότε τα επενδυτικά κεφάλαια θα γίνουν ίσα με 46,65 δις ευρώ

2 . Συντελεστής προσδιορισμού

$$R^2 = \frac{\left(\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}\right)^2}{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right] \cdot \left[\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}\right]} = \frac{\left(2143 - \frac{66 \cdot 408}{12}\right)^2}{\left[407 - \frac{(66)^2}{12}\right] \cdot \left[14150 - \frac{(408)^2}{12}\right]} =$$
$$= \frac{(2143 - 2244)^2}{[407 - 363] \cdot [14150 - 13872]} = \frac{(-101)^2}{44 \cdot 278} = \frac{10201}{12232} = 0,8340 \text{ ή } 83,40\%$$

Ερμηνεία συντελεστή προσδιορισμού R^2 : Το 83,40% της μεταβλητότητας των επενδύσεων ερμηνεύεται από τις μεταβολές των επιτοκίων



Eclass4U

The best Choice for you

6^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Σε μια μελέτη για το πως η απόδοση των αγορών σε κάποιο προϊόν επηρεάζεται από τη χρήση συγκεκριμένου λιπάσματος, έγιναν ορισμένα πειράματα. Έτσι X είναι η ποσότητα λιπάσματος (κιλά/10 στρέμματα) και Y είναι η ποσότητα (τόνοι/10 στρέμματα) του προϊόντος. Ένα μέρος των αποτελεσμάτων της γραμμικής παλινδρόμησης του X στο Y για το δείγμα των παρατηρήσεων του πειράματος ακολουθούν :

<i>Regression Statistics</i>	
Multiple R	0.90
R Square	0.80
Adjusted R Square	0.74
Standard Error	6.48
Observations	5

	<i>Coefficient</i>	<i>Stand. Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Low. 95%</i>	<i>Upp. 95%</i>
Intercept	36.07	6.21	5.81	0.01	16.30	55.84
Variable X	0.60	0.17	3.52	0.04	0.06	1.14

1 . Να εκτιμήσετε την αύξηση στο Y εάν το X αυξηθεί κατά 5 μονάδες. Επίσης, να εκτιμηθεί το Y όταν το X είναι 20.

2 . Εάν η ποσότητα λιπάσματος αυξηθεί από 20 κιλά κατά 10%, πόσο είναι η εκτιμώμενη αύξηση της παραγόμενης ποσότητας;

Παρατήρηση : Στους υπολογισμούς να χρησιμοποιηθούν 2 δεκαδικά

1 . Για να απαντήσουμε στα ερωτήματα, πρέπει πρώτα να βρούμε την εξίσωση της ευθείας παλινδρόμησης. Από την αναφορά του Excel εντοπίζουμε τα εξής :

Regression Statistics	
Multiple R	0.90
R Square	0.80
Adjusted R Square	0.74
Standard Error	6.48
Observations	5

$\hat{\beta}_0$	Coefficient	Stand. Error	t Stat
Intercept	36.07	6.21	5.81
Variable X	0.60	0.17	3.52

$\hat{\beta}_1$

Συνεπώς, έχουμε έτοιμους τους συντελεστές της ευθείας παλινδρόμησης, η οποία είναι η

$$Y = 36,07 + 0,6 \cdot X$$

Από την ερμηνεία του β_1 , γνωρίζουμε ότι αν η X αυξηθεί κατά 1 μονάδα, τότε η Y θα αυξηθεί κατά 0,6 μονάδες. Επομένως, αν η X αυξηθεί κατά 5 μονάδες, τότε η Y θα αυξηθεί κατά $5 \cdot 0,6 = 3$ μονάδες (τόνοι/10 στρέμματα).

Με αντικατάσταση στην εξίσωση της ευθείας, για $X=20$ έχουμε :

$$Y = 36,07 + 0,6 \cdot X = 36,07 + 0,6 \cdot 20 = 36,07 + 12 = 48,07 \text{ τόνοι/10στρ.}$$

2 . Η αρχική ποσότητα λιπάσματος είναι 20 κιλά. Αύξηση 10% σημαίνει ότι η τελική ποσότητα λιπάσματος θα διαμορφωθεί σε

$$\text{Νέα ποσότητα} = \text{Παλιά ποσότητα} + 10\% \text{ Παλιάς ποσότητας}$$

Άρα

$$\text{Νέα ποσότητα} = 20 + 10\% \cdot 20 = 20 + 0,1 \cdot 20 = 22 \text{ κιλά}$$

Οι αντίστοιχες τιμές της Y είναι οι εξής :

Για $X=20$

$$Y = 36,07 + 0,6 \cdot 20 = 36,07 + 12 = \mathbf{48,07} \text{ τόνοι}$$

Για $X=22$

$$Y = 36,07 + 0,6 \cdot 22 = 36,07 + 13,2 = \mathbf{49,27} \text{ τόνοι}$$

Άρα η εκτιμώμενη αύξηση της παραγόμενης ποσότητας είναι

$$Y_{\text{ΤΕΛ}} - Y_{\text{ΑΡΧ}} = 49,27 - 48,07 = 1,2 \text{ τόνοι/10 στρ.}$$

Ποσοστό μεταβολής :

$$\frac{Y_{\text{ΤΕΛ}} - Y_{\text{ΑΡΧ}}}{Y_{\text{ΑΡΧ}}} \cdot 100\% = \frac{49,27 - 48,07}{48,07} \cdot 100\% = \frac{1,2}{48,07} \cdot 100\% = 2,5\%$$

Eclass4U

The best Choice for you

7^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Ένας αναλυτής θέλει να διερευνήσει αν τα έτη 2015 και 2016 ο αριθμός των αφίξεων στην χώρα μας πρόκειται να αυξηθεί. Συσχετίζει τον αριθμό των αφίξεων με την πραγματική οικονομική δραστηριότητα στην Ευρωπαϊκή Ένωση (από όπου δέχεται και την πλειοψηφία των τουριστών που επισκέπτονται την χώρα). Τα δεδομένα που έχει στην διάθεσή του είναι ετήσια, αφορούν το διάστημα 2010-2014 και δίνονται σε εκατομμύρια αφίξεις Y , και χιλιάδες ευρώ X (κατά κεφαλήν πραγματικό ακαθάριστο εθνικό εισόδημα, Α.Ε.Π. σε τιμές έτους 2010 για τις χώρες της Ευρωπαϊκής Ένωσης). Αναλυτικά,

Έτος	Αριθμός αφίξεων (εκατ.) (Y)	Κατά κεφαλήν πραγματικό Α.Ε.Π. (χιλ ευρώ) (X)
2010	15,01	29,60
2011	16,43	30,00
2012	15,52	29,70
2013	17,92	30,60
2014	22,04	30,90

1 . Θεωρώντας ότι η σχέση μεταξύ των μεταβλητών X και Y είναι γραμμική, να εκτιμηθούν οι συντελεστές α και β του γραμμικού υποδείγματος

$Y = \alpha + \beta x + u$ με τη μέθοδο των ελαχίστων τετραγώνων. Υπόδειξη : Οι υπολογισμοί θα πρέπει να γραφούν αναλυτικά και να γίνουν με τη χρήση των τύπων που περιλαμβάνουν τις μεταβλητές σε αποκλίσεις από τους μέσους τους.

2 . Δώστε την ερμηνεία της εκτίμησης του β .

3 . Σύμφωνα με ένα αισιόδοξο σενάριο, το πραγματικό κατά κεφαλήν Α.Ε.Π. στις χώρες της Ευρωπαϊκής Ένωσης θα παρατηρηθεί ετήσια αύξηση κατά 1% το 2015 και κατά 2,5% κατά το 2016. Ποια είναι η πρόβλεψή σας για τις αφίξεις του έτους 2015 και του έτους 2016;

1 . Χρειαζόμαστε τις τιμές των \bar{X} και \bar{Y} .

Όμως, όπως γνωρίζουμε από τα μη ομαδοποιημένα δεδομένα,

$$\bar{X} = \frac{\sum X_i}{n} = \frac{150,8}{5} = 30,16$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{86,92}{5} = 17,384$$

Ο αρχικός πίνακας, με την προσθήκη των απαραίτητων στηλών για χρήση των τύπων με απόκλιση από τους μέσους, είναι ο παρακάτω

Έτος	Αρ. αφίξεων (Y_i)	Κατά κεφαλήν Α.Ε.Π. (X_i)	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y}) \cdot (X_i - \bar{X})$	$(Y_i - \bar{Y})^2$
2010	15.01	29.6	-0.56	-2.374	0.314	1.329	5.636
2011	16.43	30	-0.16	-0.954	0.026	0.153	0.91
2012	15.52	29.7	-0.46	-1.864	0.212	0.857	3.474
2013	17.92	30.6	0.44	0.536	0.194	0.236	0.287
2014	22.04	30.9	0.74	4.656	0.548	3.445	21.68
Σύνολα	86.92	150.8			1.292	6.021	31.99

$$\hat{\beta} = \frac{\sum[(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sum[(X_i - \bar{X})^2]} = \frac{6,021}{1,292} = 4,66$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = 17,384 - 4,66 \cdot 30,16 = 17,384 - 140,5456 = -123,16$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = -123,16 + 4,66 \cdot X$$

2. Ερμηνεία $\hat{\beta} = 4,66$:

Αν το κατά κεφαλήν Α.Ε.Π. αυξηθεί κατά 1 χιλιάδα ευρώ (1.000€) τότε ο αριθμός αφίξεων θα αυξηθεί κατά 4,66 εκατομμύρια (4.660.000 αφίξεις)

Υπολογισμός συντελεστή συσχέτισης (δε ζητείται από την άσκηση)

$$r = \frac{\sum[(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sqrt{\sum(Y_i - \bar{Y})^2 \cdot \sum(X_i - \bar{X})^2}} = \frac{6,021}{\sqrt{31,99 \cdot 1,292}} = \frac{6,021}{\sqrt{41,331}} = \frac{6,021}{6,429} = 0,9365$$

Οι μεταβλητές Χ και Υ (ετήσιο κατά κεφαλήν Α.Ε.Π. και ετήσιες αφίξεις) παρουσιάζουν **ισχυρή θετική γραμμική συσχέτιση**.

3. Αν το κατά κεφαλήν Α.Ε.Π. αυξηθεί κατά 1% το 2015 τότε, με δεδομένο το Α.Ε.Π. του 2014 θα διαμορφωθεί σε

$$X_{2015} = X_{2014} + 1\% \cdot X_{2014} = 30,9 + 0,01 \cdot 30,9 = 31,209 \text{ χιλ } \text{€}$$

Άρα, με αντικατάσταση στην εξίσωση της ευθείας παλινδρόμησης οι αφίξεις του 2015 **αναμένεται** να είναι :

$$Y = -123,16 + 4,66 \cdot 31,209 = 22,2739 \text{ εκατομμύρια αφίξεις}$$

Αντίστοιχα, αν το 2016 η αύξηση του Α.Ε.Π. είναι 2,5% τότε θα είναι ίσο με

$$X_{2016} = X_{2015} + 2,5\% \cdot X_{2015} = 31,209 + 0,025 \cdot 31,209 = 31,989 \text{ χιλ } \text{€}$$

Και οι αφίξεις του 2016 αναμένεται να είναι:

$$Y = -123,16 + 4,66 \cdot 31,989 = 25,9087 \text{ εκατομμύρια αφίξεις}$$

8^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Μια εταιρεία προσάρμοσε στα δεδομένα της ένα μοντέλο γραμμικής παλινδρόμησης για να μελετήσει την σχέση ανάμεσα στις πωλήσεις της (Y) σε ευρώ και του ποσού που δαπάνησε σε διαφήμιση ο βασικός ανταγωνιστής της (X). Η ευθεία παλινδρόμησης που βρήκε ήταν

$$Y = 1000 - 243,23 \cdot X$$

και ο συντελεστής προσδιορισμού 0,95. Ποια είναι η τιμή του συντελεστή συσχέτισης; (Δικαιολογήστε την απάντησή σας.)

Δίνεται ο $R^2 = 0,95$ και ζητείται η τιμή του συντελεστή συσχέτισης r

Γνωρίζουμε ότι $R^2 = r^2$ επομένως $r = \pm\sqrt{R^2} = \pm\sqrt{0,95} = \pm 0,9747$

Από τις δυο τιμές του r , αποδεκτή είναι μόνο η αρνητική αφού, επιπλέον, γνωρίζουμε ότι ο συντελεστής συσχέτισης είναι πάντα ομόσημος της εκτιμήτριας β_1 (του συντελεστή της μεταβλητής X).

Στην ευθεία παλινδρόμησης που δόθηκε, $\beta_1 = -243,23 < 0$

Άρα, τελικά, $r = -0,9747$

(οι μεταβλητές X και Y παρουσιάζουν ισχυρή αρνητική γραμμική συσχέτιση)

9^Η ΑΣΚΗΣΗ ΣΤΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Ένας οικονομικός αναλυτής θέλει να διερευνήσει τη σχέση μεταξύ των πωλήσεων μιας εταιρείας και των δαπανών διαφήμισής της. Τα δεδομένα που έχει στην διάθεσή του αφορούν το διάστημα 1994-2000 και δίνονται στον παρακάτω πίνακα

Έτος	Ετήσιες πωλήσεις σε χ.μ. (Y)	Ετήσιες Δαπάνες Διαφήμισης σε χ.μ. (X)
1994	21	5
1995	23	7
1996	27	9
1997	30	10
1998	35	12
1999	50	13
2000	45	14

Η μεταβλητή X εκφράζει τις ετήσιες δαπάνες διαφήμισης της εταιρείας σε χρηματικές μονάδες και η μεταβλητή Y τις ετήσιες πωλήσεις της στις ίδιες χ.μ.

1 . Δεδομένου ότι η σχέση μεταξύ των μεταβλητών X και Y είναι γραμμική να εκτιμήσετε τις παραμέτρους α και β του γραμμικού υποδείγματος

$Y = \alpha + \beta x + u$ με τη μέθοδο των ελαχίστων τετραγώνων. Να ερμηνευθούν οι συντελεστές του υποδείγματος.

2 . Με βάση το γραμμικό υπόδειγμα $Y = \alpha + \beta x + u$ να εκτιμήσετε τις ετήσιες πωλήσεις της εταιρείας όταν οι ετήσιες δαπάνες για διαφήμιση είναι 10 χ.μ.

3 . Για το γραμμικό υπόδειγμα $Y = \alpha + \beta x + u$ να υπολογισθούν ο συντελεστής συσχέτισης καθώς και ο συντελεστής προσδιορισμού και να ερμηνευθούν.

1 . Δημιουργούμε όλες τις απαραίτητες στήλες, κι ο αρχικός πίνακας διαμορφώνεται ως εξής

Έτος	Ετήσιες πωλήσεις σε χ.μ. Y_i	Ετήσιες Δαπάνες Διαφήμισης σε χ.μ. X_i	X_i^2	Y_i^2	$X_i \cdot Y_i$
1994	21	5	25	441	105
1995	23	7	49	529	161
1996	27	9	81	729	243
1997	30	10	100	900	300
1998	35	12	144	1225	420
1999	50	13	169	2500	650
2000	45	14	196	2025	630
Σύνολα	231	70	764	8349	2509

$$\bar{X} = \frac{\sum X_i}{n} = \frac{70}{7} = 10$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{231}{7} = 33$$

Άρα, για τους συντελεστές του γραμμικού υποδείγματος, έχουμε :

$$\hat{\beta} = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{2509 - \frac{70 \cdot 231}{7}}{764 - \frac{(70)^2}{7}} = \frac{2509 - 2310}{764 - 700} = \frac{199}{64} = 3,11$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = 33 - 3,11 \cdot 10 = 33 - 31,1 = 1,9$$

Άρα, η ευθεία παλινδρόμησης της Y πάνω στη X είναι η

$$Y = 1,9 + 3,11 \cdot X$$

Ερμηνεία του $\hat{\beta} = 3,11$

Αν οι ετήσιες δαπάνες για διαφήμιση αυξηθούν κατά 1 χ.μ. τότε οι ετήσιες πωλήσεις θα αυξηθούν κατά 3,11 χ.μ.

Ερμηνεία του $\hat{\alpha} = 1,9$

Αν οι ετήσιες δαπάνες για διαφήμιση γίνουν ίσες με 0 χ.μ. τότε οι ετήσιες πωλήσεις θα γίνουν ίσες με 1,9 χ.μ.

2 . Με αντικατάσταση στην εξίσωση της ευθείας παλινδρόμησης , έχουμε

$$Y = 1,9 + 3,11 \cdot 10 = 1,9 + 31,1 = 33 \text{ χ.μ.}$$

Συνεπώς, αν οι ετήσιες δαπάνες για διαφήμιση είναι 10 χ.μ., οι ετήσιες πωλήσεις αναμένεται να γίνουν ίσες με 33 χ.μ.

3 . Συντελεστής συσχέτισης

$$r = \frac{\sum X_i Y_i - \frac{\sum X_i \cdot \sum Y_i}{n}}{\sqrt{\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n} \right] \cdot \left[\sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \right]}} = \frac{2509 - \frac{70 \cdot 231}{7}}{\sqrt{\left[764 - \frac{(70)^2}{7} \right] \left[8349 - \frac{(231)^2}{7} \right]}} =$$
$$= \frac{2509 - 2310}{\sqrt{[764 - 700] \cdot [8349 - 7623]}} = \frac{199}{\sqrt{64 \cdot 726}} = \frac{199}{215,56} = 0,9232$$

Οι μεταβλητές Χ και Υ (ετήσιες πωλήσεις και ετήσιες δαπάνες για διαφήμιση) παρουσιάζουν ισχυρή θετική γραμμική συσχέτιση.

Συντελεστής προσδιορισμού

$$\Xi\acute{\epsilon}\rho\omicron\upsilon\mu\epsilon \acute{\omicron}\tau\iota R^2 = r^2 = 0,9232^2 = 0,8523 \text{ ή } 85,23\%$$

Ερμηνεία συντελεστή προσδιορισμού : Το 85,23% της μεταβλητότητας των ετήσιων πωλήσεων ερμηνεύεται από τις ετήσιες δαπάνες για διαφήμιση.