



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΔΙΟΙΚΗΣΗΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
“ΜΕΤΑΠΤΥΧΙΑΚΟ ΔΙΠΛΩΜΑ ΔΙΟΙΚΗΣΗΣ
ΕΠΙΧΕΙΡΗΣΕΩΝ - MBA”

ΣΤΑΤΙΣΤΙΚΗ
ΚΑΙ
ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

*Επιλεγμένες Ασκήσεις με ενδεικτική λύση
στην Παλινδρόμηση*

ΑΣΚΗΣΗ 1^η

Ο Πίνακας 1 περιέχει την βαθμολογία πέντε φοιτητών στα μαθήματα των Μαθηματικών (X) και της Στατιστικής (Y).

Πίνακας 1

Μαθηματικά (X)	Στατιστική (Y)
2	1
3	2
5	5
6	6
8	7

- α. Να εκτιμηθεί η ευθεία γραμμικής παλινδρόμησης της Y πάνω στη X και να ερμηνευτούν οι συντελεστές της.
 β. Να εκτιμηθεί το τυπικό σφάλμα της Y και να ερμηνευτεί.
 γ. Να εκτιμηθεί ο συντελεστής προσδιορισμού και να ερμηνευτεί.
 δ. Να εκτιμηθεί ο συντελεστής συσχέτισης, να ερμηνευτεί και να γίνει ο παρακάτω έλεγχος :
 $H_0: \rho=0$
 $H_1: \rho \neq 0$
 ε. Να εκτιμηθεί ο βαθμός ενός φοιτητή στο μάθημα Στατιστικής όταν ο βαθμός του στα Μαθηματικά είναι 4.
 στ. Να υπολογιστεί το 95% διάστημα πρόβλεψης του βαθμού της Στατιστικής όταν ο βαθμός των Μαθηματικών είναι 7.

Σημείωση: Δίνεται ότι το επίπεδο σημαντικότητας είναι $\alpha=0.05$.

ΛΥΣΗ

Με βάση τα δεδομένα της άσκησης δημιουργείται ο Πίνακας 1.1 ο οποίος επιπλέον περιέχει τα σύνολα και τους μέσους τόσο των αγωγών όσο και των blocks.

Πίνακας 1.1

(1)	(2)	(3)	(4)	(5)	(6)
X_i	Y_i	X_i^2	Y_i^2	$(X_i - \bar{X})$	$(X_i - \bar{X})Y_i$
2	1	4	1	-2.8	-2.8
3	2	9	4	-1.8	-3.6
5	5	25	25	0.2	1.0
6	6	36	36	1.2	7.2
8	7	64	49	3.2	22.4
ΣΥΝΟΛΑ	24	21	138	115	24,2

Υπολογισμός βοηθητικών στοιχείων

$$\bar{X} = \frac{24}{5} = 4.8, \quad \bar{Y} = \frac{21}{5} = 4.2$$

$$S_{XX} = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n} = 138 - \frac{24^2}{5} = 22.8$$

$$S_{YY} = \sum Y_i - \frac{\left(\sum Y_i\right)^2}{n} = 115 - \frac{21^2}{5} = 26.8$$

$$S_{YX} = \sum (X_i - \bar{X})Y_i = 24.2$$

α) Ευθεία Παλινδρόμησης

$$\alpha = \bar{Y} - \frac{S_{XY}}{S_{XX}} \bar{X} = 4.2 - \frac{24.2}{22.8} 4.8 \Rightarrow a = -0.89$$

$$b = \frac{S_{XY}}{S_{XX}} = \frac{24.2}{22.8} \Rightarrow b = 1.06$$

Άρα η ευθεία παλινδρόμησης είναι: $\hat{Y} = -0.89 + 1.06X$

β) Τυπικό Σφάλμα της Y

$$S_{y/x}^2 = \frac{1}{n-2} \left(S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right) = \frac{1}{5-2} \left(26.8 - \frac{(24.2)^2}{22.8} \right) \Rightarrow S_{y/x}^2 \cong 0.4$$

$$\text{Άρα } S_{y/x} = \sqrt{S_{y/x}^2} = 0.6$$

γ) Συντελεστής Προσδιορισμού

$$R^2 = b \frac{S_{xy}}{S_{yy}} = 1.06 \frac{24.2}{26.8} \Rightarrow R^2 = 0.96$$

Συμπεραίνουμε ότι η μεταβολή του βαθμού στη Στατιστική ερμηνεύεται κατά 96% από τη μεταβολή του βαθμού στα Μαθηματικά.

δ) Συντελεστής Συσχέτισης

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{24.2}{\sqrt{22.8 * 26.8}} \Rightarrow r = 0.98$$

Συμπεραίνουμε ότι υπάρχει έντονη θετική συσχέτιση μεταξύ του βαθμού στη Στατιστική και στα Μαθηματικά

Υποθέσεις:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Στατιστική συνάρτηση ελέγχου:

$$T_0 = r \sqrt{\frac{n-2}{1-r^2}}$$

Κριτήριο απόρριψης:

Απορρίπτουμε την H_0 αν $T_0 > t_{n-2, 1-\alpha/2}$

Στη συγκεκριμένη περίπτωση:

$$T_0 = 0.98 \sqrt{\frac{5-2}{1-(0.98)^2}} \Rightarrow T_0 = 8.5$$

$$\alpha = 0.05$$

$$n=5$$

$$t_{n-2, 1-\alpha/2} = t_{3, 0.975} = 3.1824$$

Συμπέρασμα:

Επειδή $T_0 = 8.5 > 3.1824 = t_{3, 0.975}$ απορρίπτουμε την H_0 και δεχόμαστε την H_1

ε. Εκτίμηση της τιμής του y

Η ευθεία παλινδρόμησης είναι:

$$\hat{Y}_i = -0.89 + 1.06 * x$$

$$\text{Για } x = 4 \text{ έχουμε } \hat{Y}_i = -0.89 + 1.06 * 4 \Rightarrow \hat{Y}_i = 3.35$$

στ. Διάστημα πρόβλεψης της τιμής του y

Ο γενικός τύπος του διαστήματος πρόβλεψης είναι:

$$\hat{Y}_k \pm t_{n-2, 1-\alpha/2} S_{\hat{Y}_k / X_k} \quad (1)$$

Στη συγκεκριμένη περίπτωση

$$x_k = 7 \rightarrow \hat{Y}_k = -0.89 + 1.06 * 7 = 6.53$$

$$n = 5, \alpha = 0.05$$

$$t_{n-2, 1-\alpha/2} = t_{3, 0.975} = 3.1824$$

$$S_{\hat{Y}_k / X_k} = S_{Y/X} \sqrt{\frac{n+1}{n} + \frac{(X_k - \bar{X})^2}{S_{xx}}} = 0.06 \sqrt{\frac{6}{5} + \frac{4.84}{22.8}} = 0.72$$

Κατά συνέπεια το διάστημα πρόβλεψης είναι: $6.5 \pm 3.1824 * 0.72$ δηλαδή 6.53 ± 2.29 ή $(4.24, 8.82)$

ΑΣΚΗΣΗ 2^η

Στον Πίνακα 2 με X συμβολίζεται ο αριθμός της γραπτής εξέτασης φοιτητών στα Μαθηματικά και με Y ο αντίστοιχος βαθμός στο Marketing (κλίμακα βαθμολογίας 0 - 4).

Πίνακας 2

X_i	0	1	2	1	0	3	4	2	2	1
Y_i	0	2	1	3	1	3	4	2	1	2

Οι τιμές (X_i, Y_i) παρατηρήθηκαν σε τυχαίο δείγμα 10 φοιτητών ενός Πανεπιστημιακού Τμήματος Διοίκησης Επιχειρήσεων. Να εκτιμηθεί η συνάρτηση $Y = a + bX^2 + \varepsilon$ και να συγκριθεί η ερμηνευτική της ικανότητα με την αντίστοιχη της απλής γραμμικής παλινδρόμησης.

Σημείωση: Για διευκόλυνση των υπολογισμών δίνονται:

$$S_{xx} = 14.4 \quad S_{yy} = 12.9 \quad S_{xy} = 9.6$$

ΛΥΣΗ

Ο παρακάτω πίνακας περιέχει τα αρχικά δεδομένα (στήλες 1,2) καθώς και ορισμένα βοηθητικά στοιχεία που προκύπτουν από αυτά και τα οποία είναι απαραίτητα για τη λύση της άσκησης.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
X	Y	$Y - \bar{Y}$	$X^2 = Z$	$Z - \bar{Z}$	$(Z - \bar{Z})^2$	$(Y - \bar{Y})(Z - \bar{Z})$
0	0	-1.9	0	-4	16	7.6
1	2	0.1	1	-3	9	-0.3
2	1	-0.9	4	0	0	0
1	3	1.1	1	-3	9	-3.3
0	1	-0.9	0	-4	16	3.6
3	3	1.1	9	5	25	5.5
4	4	2.1	16	12	144	25.2
2	2	0.1	4	0	0	0
2	1	-0.9	4	0	0	0
1	2	0.1	1	3	9	-0.3
16	19		40			38.0
$\sum X_i$	$\sum Y_i$		$\sum Z_i$		S_{zz}	S_{yz}

Δίνεται επίσης ότι $S_{xx} = 14.4$, $S_{yy} = 12.9$, $S_{xy} = 9.6$

Υπολογισμός βοηθητικών στοιχείων

$$\bar{X} = \frac{\sum X_i}{n} = \frac{16}{10} \rightarrow \bar{X} = 1.6$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{19}{10} \rightarrow \bar{Y} = 1.9$$

$$\bar{Z} = \frac{\sum Z_i}{n} = \frac{40}{10} \rightarrow Z = 4.0$$

$$S_{zz} = 228, \quad S_{yz} = 38$$

Επίσης δίνονται $S_{xx} = 14.4$, $S_{yy} = 12.9$, $S_{xy} = 9.6$

Γραμμικό Μοντέλο

$$b = \frac{S_{XY}}{S_{XX}} = \frac{9.6}{14.4} \Rightarrow b = 0.67,$$

$$\alpha = \bar{Y} - b\bar{X} = 1.9 - 0.67 * 1.6 \Rightarrow a \cong 0.83$$

$$\text{Άρα } \hat{Y} = 0.83 + 0.67X$$

$$R^2 = \frac{S_{XY}}{S_{YY}} b = \frac{9.6}{12.9} * 0.67 \rightarrow R^2 \cong 0.50$$

Μη Γραμμικό Μοντέλο

$$\text{Θέτω } X^2 = Z. \quad \text{Άρα } \hat{Y} = a + bZ$$

$$b = \frac{S_{YZ}}{S_{ZZ}} = \frac{38}{228} \rightarrow b = 0.17,$$

$$\alpha = \bar{Y} - b\bar{Z} = 1.9 - 0.17 * 4 \Rightarrow a = 1.22$$

$$\text{Άρα } \hat{Y} = 1.22 + 0.17Z \rightarrow \hat{Y} = 1.22 + 0.17X^2$$

$$R^2 = b \frac{S_{XY}}{S_{YY}} = 0.17 \frac{38}{12.9} \rightarrow R^2 \cong 0.50$$

Από τη σύγκριση του R^2 προκύπτει ότι τα δύο μοντέλα είναι ισοδύναμα.

ΑΣΚΗΣΗ 3^η

Μια μεγάλη δισκογραφική εταιρεία έχει αναλάβει την παραγωγή των δίσκων 6 νέων καλλιτεχνών. Για το σκοπό αυτό και στα πλαίσια της πολιτικής προώθησής τους έχει συμφωνήσει με διάφορες τηλεοπτικές εκπομπές την εμφάνιση σε αυτές των συγκεκριμένων καλλιτεχνών. Ο Διευθυντής Πωλήσεων της εταιρείας θέλει να εξετάσει εάν ο αριθμός των τηλεοπτικών εμφανίσεων επηρεάζει το ύψος των πωλήσεων των CD των συγκεκριμένων καλλιτεχνών. Για το σκοπό αυτό καταγράφει για ένα συγκεκριμένο διάστημα τις εμφανίσεις των καλλιτεχνών (x) και τις πωλήσεις των CD τους (y), σε χιλιάδες αντίτυπα, για το συγκεκριμένο διάστημα. Ο Πίνακας 3 συνοψίζει τα στοιχεία που συνέλεξε.

Πίνακας 3

Καλλιτέχνης	1	2	3	4	5	6
Αριθμός τηλεοπτικών εμφανίσεων (X)	9	3	5	7	1	2
Πωλήσεις CD (Y)	13	7	2	8	8	1

Με βάση τα παραπάνω στοιχεία:

- Να εκτιμηθεί η ευθεία παλινδρόμησης της Y πάνω στη X και να ερμηνευθούν οι συντελεστές της.
- Να εκτιμηθεί ο συντελεστής προσδιορισμού και ο συντελεστής συσχέτισης και να ερμηνευθούν.
- Μια ανταγωνιστική δισκογραφική εταιρεία υποστηρίζει ότι ο αριθμός των τηλεοπτικών εμφανίσεων ενός καλλιτέχνη δεν επηρεάζει τις πωλήσεις των CD του. Να διερευνηθεί, σε επίπεδο σημαντικότητας $\alpha=0,05$, αν ισχύει, αυτή η άποψη και για την περίπτωση της συγκεκριμένης εταιρείας.

Σημείωση: Για διευκόλυνση των υπολογισμών δίνεται $S^2_{Y/X} = 16,57$

	x	y	xi-xmes	yi-ymes	sqr(xi-xmes)	sqr(yi-ymes)	(xi-xmes)yi	sqr(xi)	sqr(yi)	xi*yi
	9	13	4,50	6,50	20,25	42,25	58,50	81	169	117
	3	7	-1,50	0,50	2,25	0,25	-10,50	9	49	21
	5	2	0,50	-4,50	0,25	20,25	1,00	25	4	10
	7	8	2,50	1,50	6,25	2,25	20,00	49	64	56
	1	8	-3,50	1,50	12,25	2,25	-28,00	1	64	8
	2	1	-2,50	-5,50	6,25	30,25	-2,50	4	1	2
Sum	27	39			47,50	97,50	38,50	169	351	214
Aver	4,50	6,50			Sxx	Syy	Sxy	Σxi ²	Σyi ²	Σxiyi

i.

$$\bar{Y} = \sum \frac{y_i}{n} = 6,5$$

$$\bar{X} = \sum \frac{x_i}{n} = 4,5$$

$$b = \frac{S_{XY}}{S_{XX}} = \frac{38,5}{47,5} = 0,81$$

$$a = \bar{Y} - b\bar{X} = 6,5 - 0,81 * 4,5 = 2,855$$

$$\hat{y} = a + bx = 2,855 + 0,81x$$

ii.

$$R^2 = b \frac{S_{xy}}{S_{yy}} = 0,81 \frac{38,5}{97,5} = 0,32$$

$$r = \frac{S_{XY}}{\sqrt{S_{XX}} \sqrt{S_{YY}}} = \frac{38,5}{\sqrt{47,5 * 97,5}} = \frac{38,5}{68,05} = 0,57$$

iii.

$$H_0 : \beta = \beta_0 = 0$$

$$H_1 : \beta \neq \beta_0$$

Στατιστική Συνάρτηση Ελέγχου:

$$T_0 = \frac{\hat{\beta} - \beta_0}{\frac{S_{Y/X}}{\sqrt{S_{XX}}}} \sim t_{n-2}$$

Κριτήριο Απόρριψης: Απορρίπτουμε την H_0 αν:

$$T_0 > t_{n-2, 1-\alpha/2}$$

ή

$$T_0 < -t_{n-2, 1-\alpha/2}$$

$$S^2_{Y/X} = \frac{1}{n-2} \left(S_{YY} - \frac{S^2_{XY}}{S_{XX}} \right) = \frac{1}{6-2} \left(97,5 - \frac{38,5^2}{47,5} \right) = 16,5725$$

$$S_{Y/X} = \sqrt{16,5725} = 4,07$$

$$T_0 = \frac{\hat{\beta} - \beta_0}{\frac{S_{Y/X}}{\sqrt{S_{XX}}}} = \frac{0,81 - 0}{\frac{4,07}{\sqrt{47,5}}} = 1,37$$

$$t_{n-2, 1-\alpha/2} = t_{4, 0,975} = 2,7764$$

$T_0 = 1,37 < 2,7764$ άρα δέχομαι την H_0