Παράδειγμα Υπολογισμού

Συντελεστού Συσχέτισης
και
Ευθείας της Παλινδρόμησης

Χαράλαμπος Γναρδέλλης

Εφαρμογές Πληροφορικής στην Αλιεία και τις Υδατοκαλλιεργειες

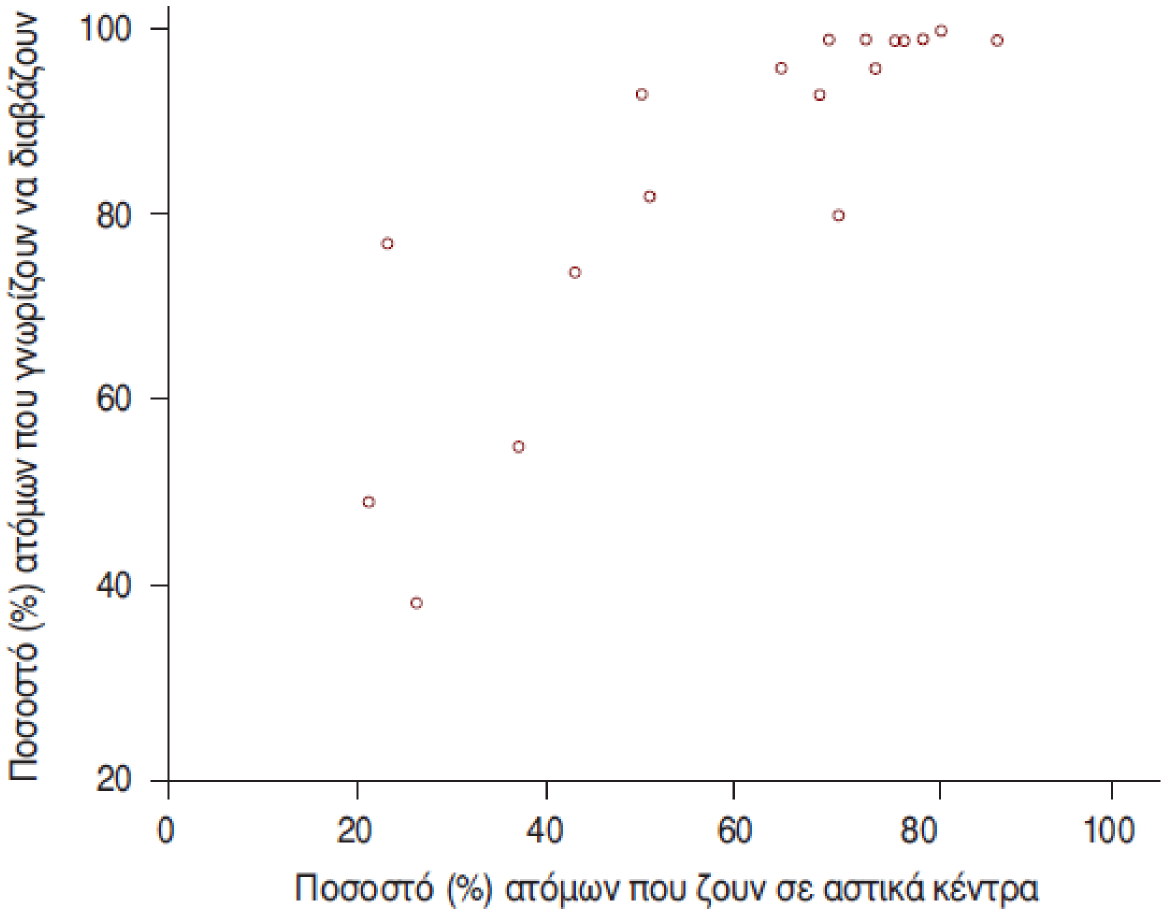
Υπολογισμός συντελεστού συσχέτισης

Ποσοστό ατόμων που ζουν σε αστικά κέντρα και ποσοστό ατόμων που γνωρίζουν να διαβάζουν σε ένα τυχαίο δείγμα χωρών του Οργανισμού Ηνωμένων Εθνών

	Ποσοστό (%) ατόμων που ζουν σε αστικά κέντρα	Ποσοστό (%) ατόμων που γνωρίζουν να διαβάζουν
Αίγυπτος	43	74
Αιθιοπία	21	49
Αφγανιστάν	26	38
Βολιβία	69	93
Γαλλία	80	99
Γερμανία	77	99
Γουατεμάλα	51	82
Δανία	88	99
Ελβετία	74	99
Ιράκ	71	80
Ιταλία	70	99
Καμπότζη	23	77
Νορβηγία	82	100
Πακιστάν	37	55
Πορτογαλία	65	96
Ταϊβάν	78	99
Ταϊλάνδη	50	93
Τουρκία	75	96

Κατασκευή διαγράμματος διασποράς

Διάγραμμα διασποράς του βαθμού αστικότητας και του ποσοστού των ατόμων που γνωρίζουν να διαβάζουν σε ένα τυχαίο δείγμα χωρών του Οργανισμού Ηνωμένων Εθνών



Συντελεστής Συσχέτισης

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right]\left[\sum_{i=1}^{n} (y_i - \bar{y})^2\right]}},$$

Υπολογισμός μέσων τιμών

- για τη μεταβλητή X, που αναπαριστά το βαθμό αστικότητας των χωρών

$$\bar{x} = \frac{1}{18}\sum_{i=1}^{18} x_i = \frac{(43 + 21 + \ldots + 75)}{18} = 60{,}0$$

- για τη μεταβλητή Y, που αναπαριστά το ποσοστό των ατόμων που γνωρίζουν να διαβάζουν

$$\bar{y} = \frac{1}{18}\sum_{i=1}^{18} x_i = \frac{(74 + 49 + \ldots + 96)}{18} = 84{,}8$$

Υπολογισμός αθροίσματος γινομένων

- για τις δύο μεταβλητές X και Y

$$\sum_{i=1}^{18} (x_i - \bar{x})(y_i - \bar{y}) =$$

$$= (43 - 60)(74 - 84{,}8) + (21 - 60)(49 - 84{,}8) + \\ + \ldots + (75 - 60)(96 - 84{,}8) = 6187{,}0$$

- και

$$\sum_{i=1}^{18} (x_i - \bar{x})^2 =$$

$$= (43 - 60)^2 + (21 - 60)^2 + \ldots + (75 - 60)^2 = 8074{,}0,$$

$$\sum_{i=1}^{18} (y_i - \bar{y})^2 =$$

$$= (74 - 84{,}8)^2 + (49 - 84{,}8)^2 + \ldots + (96 - 84{,}8)^2 = 6394{,}5.$$

Υπολογισμός συντελεστού συσχέτισης

Άρα, ο συντελεστής συσχέτισης του Pearson ισούται με

$$r = \frac{\sum_{i=1}^{18} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{18} (x_i - \bar{x})^2\right]\left[\sum_{i=1}^{18} (y_i - \bar{y})^2\right]}} =$$

$$= \frac{6187}{\sqrt{(8074)(6394{,}5)}} = 0{,}86.$$

Επαγωγικός έλεγχος στο συντελεστή συσχέτισης

Προκειμένου να ελεγχθεί, αν η έντονη θετική συσχέτιση, που προσδιορίστηκε μεταξύ του βαθμού αστικότητας και του ποσοστού των ατόμων που γνωρίζουν να διαβάζουν ισχύει για όλες τις χώρες του Οργανισμού Ηνωμένων εθνών, μπορεί να πραγματοποιηθεί ένας αμφίπλευρος έλεγχος με μηδενική υπόθεση

$$H_0 : \varrho = 0$$

και εναλλακτική την

$$H_A : \varrho \neq 0,$$

όπου ϱ ο πληθυσμιακός συντελεστής συσχέτισης για όλες τις χώρες του Οργανισμού Ηνωμένων Εθνών.

Ο έλεγχος θα γίνει σε επίπεδο σημαντικότητας $\alpha = 0{,}05$ με τη βοήθεια του κριτηρίου

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

όπου $r = 0{,}86$, ο δειγματικός συντελεστής συσχέτισης για το τυχαίο δείγμα των 18 χωρών.

Έχουμε, επομένως,

$$t = 0{,}86\sqrt{\frac{18-2}{1-(0{,}86)^2}} = 6{,}74.$$

Ανατρέχοντας στον Πίνακα Α.5, παρατηρούμε ότι για μια κατανομή t με 16 βαθμούς ελευθερίας, η πιθανότητα να προκύψει μια τιμή τόσο ακραία όσο η 6,74 ή η −6,74, είναι $p < 2(0{,}0005) = 0{,}001$. Επειδή $p < \alpha$, η μηδενική υπόθεση της ισότητας του πληθυσμιακού συντελεστή συσχέτισης με το 0 απορρίπτεται.

Επομένως, υπάρχει θετική συσχέτιση μεταξύ του βαθμού αστικότητας και του ποσοστού των ατόμων που γνωρίζουν να διαβάζουν, για το σύνολο των χωρών του Οργανισμού Ηνωμένων Εθνών.

Στον παρακάτω πίνακα δίδονται οι τιμές μεθυλικού υδραργύρου από διατροφική πρόσληψη και οι αντίστοιχες συγκεντρώσεις υδραργύρου στο αίμα 12 ατόμων που εκτέθηκαν σε μεθυλικό υδράργυρο λόγω κατανάλωσης μολυσμένων ψαριών:

Μεθυλικός υδράργυρος από πρόσληψη (μg Hg/ημέρα)	Συγκέντρωση υδραργύρου στο αίμα (ng/g)
180	90
200	120
230	125
410	290
600	310
550	290
275	170
580	375
105	70
250	105
460	205
650	480

Να διερευνηθεί αν υπάρχει σχέση μεταξύ των δύο μεταβλητών, με τη βοήθεια του συντελεστή συσχέτισης του Spearman. Ο έλεγχος να γίνει σε επίπεδο σημαντικότητας $\alpha = 0{,}05$, αφού πρώτα κατασκευαστεί το διάγραμμα διασποράς των δύο μεταβλητών.

Υπολογισμός συντελεστών απλής παλινδρόμησης

Στο παράδειγμα του υπολογισμού του συντελεστού συσχέτισης μεταξύ βαθμού αστικότητας και βαθμού αλφαβητισμού των χωρών του ΟΗΕ θα κατασκευαστεί η ευθεία της παλινδρόμησης

$$\hat{y} = \hat{\alpha} + \hat{\beta} x$$

με

εξαρτημένη μεταβλητή y τον αλφαβητισμό, και

ανεξάρτητη μεταβλητή x το βαθμό αστικότητας.

Η εξίσωση ορίζεται από την παρακάτω σχέση

Οι συντελεστές του υποδείγματος υπολογίζονται από τις εξισώσεις

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

και

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

Υπολογισμός των συντελεστών του υποδείγματος

Αρχικά υπολογίζουμε τις μέσες τιμές των δύο μεταβλητών

- για τη μεταβλητή X, που αναπαριστά το βαθμό αστικότητας των χωρών

$$\bar{x} = \frac{1}{18} \sum_{i=1}^{18} x_i = \frac{(43 + 21 + \ldots + 75)}{18} = 60{,}0$$

- για τη μεταβλητή Y, που αναπαριστά το ποσοστό των ατόμων που γνωρίζουν να διαβάζουν

$$\bar{y} = \frac{1}{18} \sum_{i=1}^{18} x_i = \frac{(74 + 49 + \ldots + 96)}{18} = 84{,}8$$

Στη συνέχεια υπολογίζουμε το άθροισμα των γινομένων των διαφορών και το άθροισμα των τετραγώνων των διαφορών για τις δύο μεταβλητές

- για τις δύο μεταβλητές X και Y

$$\sum_{i=1}^{18} (x_i - \bar{x})(y_i - \bar{y}) =$$

$$= (43 - 60)(74 - 84{,}8) + (21 - 60)(49 - 84{,}8) + \\ + \ldots + (75 - 60)(96 - 84{,}8) = 6187{,}0$$

- και

$$\sum_{i=1}^{18} (x_i - \bar{x})^2 =$$

$$= (43 - 60)^2 + (21 - 60)^2 + \ldots + (75 - 60)^2 = 8074{,}0,$$

$$\sum_{i=1}^{18} (y_i - \bar{y})^2 =$$

$$= (74 - 84{,}8)^2 + (49 - 84{,}8)^2 + \ldots + (96 - 84{,}8)^2 = 6394{,}5.$$

Αντικαθιστούμε τα προηγούμενα αποτελέσματα στις εξισώσεις

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

και

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

Υπολογισμός του $\hat{\beta}$

$$\hat{\beta} = \frac{\sum_{i=1}^{18}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{18}(x_i - \bar{x})^2} = \frac{6187}{8074} = 0{,}77$$

Υπολογισμός του $\hat{a}$

$$\hat{a} = \bar{y} - \hat{\beta}\bar{x} = 84{,}8 - (0{,}77) \cdot (60) = 38{,}6$$

- Άρα η εξίσωση της παλινδρόμησης που εκτιμά το βαθμό αλφαβητισμού μιας χώρας με το βαθμό αστικότητας δίδεται από τη σχέση:

$$\hat{\boldsymbol{y}} = 38{,}6 + 0{,}77\,\boldsymbol{x}$$

Άσκηση

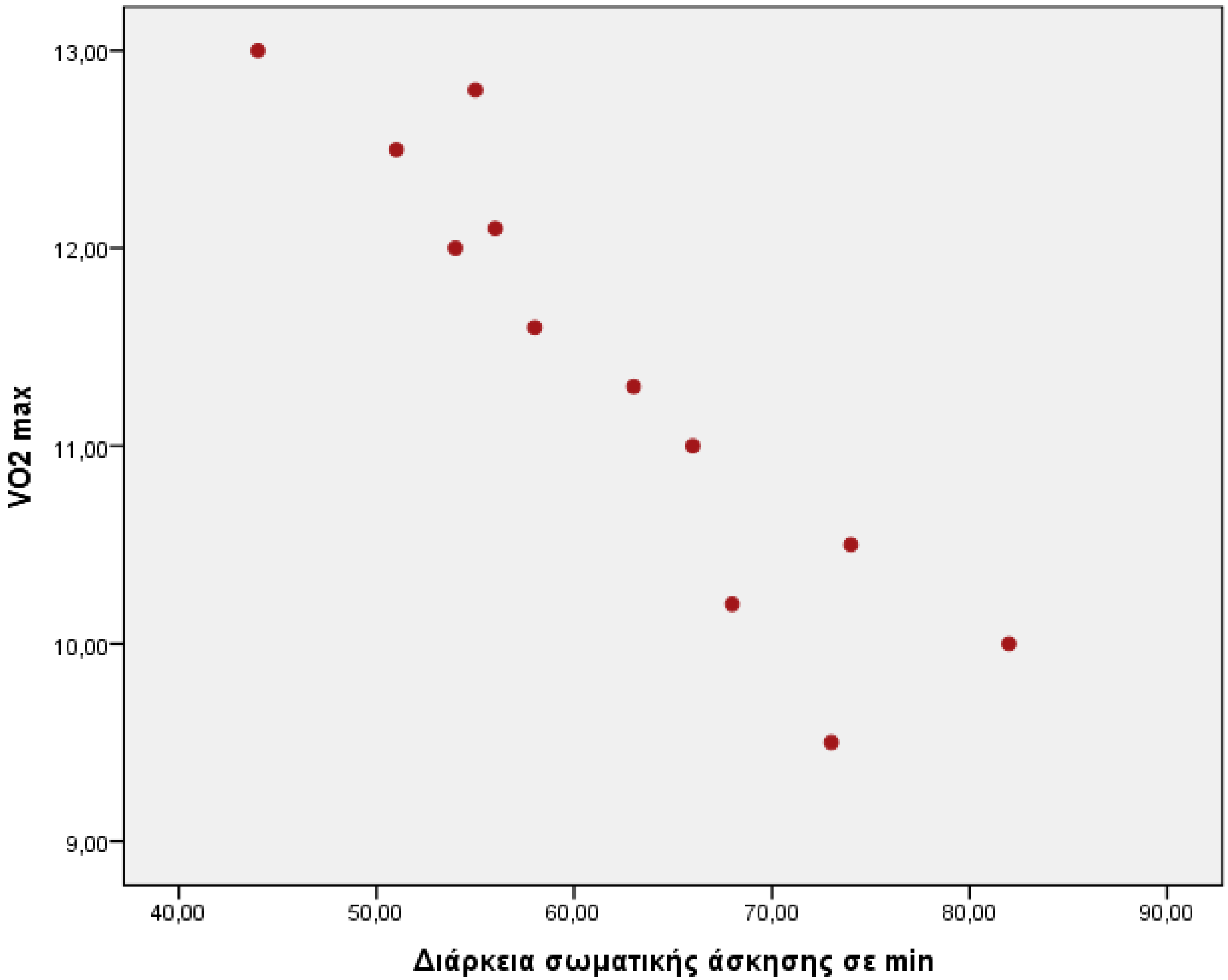
Στον παρακάτω πίνακα δίδονται οι τιμές του μέγιστου εισπνεόμενου όγκου οξυγόνου (VO_2 max) 12 ενηλίκων έπειτα από εντατική σωματική άσκηση. Γενικά, ο μέγιστος εισπνεόμενος όγκος οξυγόνου αποτελεί μέτρο της εύρυθμης καρδιακής λειτουργίας και ελαττώνεται με την αύξηση του επιπέδου της σωματικής δραστηριότητας.

Άτομο	VO_2 max	Διάρκεια σωματικής άσκησης σε min
1	82	10,0
2	73	9,5
3	68	10,2
4	74	10,5
5	66	11,0
6	63	11,3
7	58	11,6
8	54	12,0
9	56	12,1
10	51	12,5
11	55	12,8
12	44	13,0

Να γίνει το διάγραμμα σκέδασης του μέγιστου εισπνεόμενου όγκου οξυγόνου με τη διάρκεια της σωματικής άσκησης.

Να προσδιοριστεί η ευθεία των ελαχίστων τετραγώνων για τη σχέση του VO_2 max με τη διάρκεια της σωματικής άσκησης.

Διάγραμμα διασποράς των δύο μεταβλητών



Υπολογισμός των παραμέτρων του υποδείγματος

Σωματική άσκηση x_i	VO2 max y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$\hat{y}_i$	$(y_i - \hat{y})^2$
82	10	20	-1,38	-27,5	400	70,8	125,44
73	9,5	11	-1,88	-20,6	121	75,2	4,84
68	10,2	6	-1,18	-7,1	36	69,04	1,08
74	10,5	12	-0,88	-10,5	144	66,4	57,76
66	11	4	-0,38	-1,5	16	62	16
63	11,3	1	-0,07	-0,1	1	59,36	13,25
58	11,6	-4	0,23	-0,9	16	56,72	1,64
54	12	-8	0,63	-5,0	64	53,2	0,64
56	12,1	-6	0,73	-4,4	36	52,32	13,54
51	12,5	-11	1,13	-12,4	121	48,8	4,84
55	12,8	-7	1,43	-10,0	49	46,16	78,15
44	13	-18	1,63	-29,3	324	44,4	0,16
$\bar{x} = 62$	$\bar{y} = 11{,}38$			$\sum(x_i - \bar{x})(y_i - \bar{y}) =$ -129,1	$\sum(x_i - \bar{x})^2 =$ 1328		$\sum(y_i - \hat{y})^2 =$ 317,34

$$\boldsymbol{b}_1 = \frac{\sum_{\boldsymbol{i}=1}^{12} \left(\boldsymbol{x_i} - \overline{\boldsymbol{x}}\right)\left(\boldsymbol{y_i} - \overline{\boldsymbol{y}}\right)}{\sum_{\boldsymbol{i}=1}^{12} \left(\boldsymbol{x_i} - \overline{\boldsymbol{x}}\right)^2} = \frac{-129{,}1}{1328} = -0{,}097$$

$$\boldsymbol{b}_0 = \overline{\boldsymbol{y}} - \boldsymbol{b}_1 \overline{\boldsymbol{x}} = 11{,}38 - \left(-0{,}097\right) \cdot \left(62\right) = 17{,}40$$

$$s_{y|x} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{n-2} = \frac{\sqrt{\sum_{i=1}^{12}(y_i - \hat{y}_i)^2}}{12-2} = \frac{\sqrt{317{,}34}}{10} = 1{,}78$$

$$se(b_1) = \frac{s_{y|x}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{1{,}78}{\sqrt{1328}} = 0{,}0488$$

Ο έλεγχος του συντελεστή της παλινδρόμησης γίνεται με μηδενική υπόθεση

$$H_0 : \beta_1 = 0$$

και εναλλακτική

$$H_A : \beta_1 \neq 0,$$

μέσω της ποσότητας

$$t = \frac{b_1}{se(b_1)} = \frac{-0,097}{0,0488} = -1,99$$

η οποία, εφόσον ισχύει η μηδενική υπόθεση, ακολουθεί την κατανομή t με n-2 = 12-2 = 10 βαθμούς ελευθερίας.

Ανατρέχοντας στον πίνακα της κατανομής t, για επίπεδο σημαντικότητας α=0,05 (δηλ. α=0,025) και βαθμούς ελευθερίας df=10 βλέπουμε ότι

$$|t| = |-1{,}99| < t_{10} = 2{,}228$$

επομένως, η μηδενική υπόθεση ότι ο συντελεστής b_1 είναι ίσος με 0 δεν απορρίπτεται.

Άρα στον πληθυσμό των ενηλίκων από τον οποίο προέρχεται το τυχαίο δείγμα των 12 ατόμων, δεν υπάρχει σημαντική σχέση μεταξύ μέγιστου εισπνεόμενου όγκου οξυγόνου και της σωματικής άσκησης .

Στον παρακάτω πίνακα δίδονται οι τιμές μεθυλικού υδραργύρου από διατροφική πρόσληψη και οι αντίστοιχες συγκεντρώσεις υδραργύρου στο αίμα 12 ατόμων που εκτέθηκαν σε μεθυλικό υδράργυρο λόγω κατανάλωσης μολυσμένων ψαριών:

Μεθυλικός υδράργυρος από πρόσληψη (μg Hg/ημέρα)	Συγκέντρωση υδραργύρου στο αίμα (ng/g)
180	90
200	120
230	125
410	290
600	310
550	290
275	170
580	375
105	70
250	105
460	205
650	480

Να οριστεί η ευθεία της παλινδρόμησης με ανεξάρτητη μεταβλητή τις τιμές του μεθυλικού υδραργύρου από διατροφική πρόσληψη και εξαρτημένη μεταβλητή τη συγκέντρωση υδραργύρου στο αίμα.

Να ερμηνεύσετε το συντελεστή b_1 της παλινδρόμησης.

Μεθυλικός υδράργυρος x_i	Συγκέντρωση υδραργύρου y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
180	90	-194,17	-129,17	25079,86	37700,69
200	120	-174,17	-99,17	17271,53	30334,03
230	125	-144,17	-94,17	13575,69	20784,03
410	290	35,83	70,83	2538,19	1284,03
600	310	225,83	90,83	20513,19	51000,69
550	290	175,83	70,83	12454,86	30917,36
275	170	-99,17	-49,17	4875,69	9834,03
580	375	205,83	155,83	32075,69	42367,36
105	70	-269,17	-149,17	40150,69	72450,69
250	105	-124,17	-114,17	14175,69	15417,36
460	205	85,83	-14,17	-1215,97	7367,36
650	480	275,83	260,83	71946,53	76084,03
$\bar{x}$ = 374,17	$\bar{y}$ = 219,17			$\sum_{i=1}^{12}(x_i - \bar{x})(y_i - \bar{y})$ =253441,67	$\sum_{i=1}^{12}(x_i - \bar{x})^2$ =395541,67

$$b_1 = \frac{\sum_{i=1}^{12}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{12}(x_i - \bar{x})^2} = \frac{253441,67}{395541,67} = 0,64$$

$$b_0 = \bar{y} - b_1\bar{x} = 219,17 - (0,64)\cdot(374,17) = -20,58$$

- Άρα η εξίσωση της παλινδρόμησης που εκτιμά τη συγκέντρωση υδραργύρου στο αίμα σε σχέση με την πρόσληψη μεθυλικού υδραργύρου δίδεται από τη σχέση:

$$\hat{\boldsymbol{y}} = -20{,}58 + 0{,}64\,\boldsymbol{x}$$

Άρα η συγκέντρωση υδραργύρου στο αίμα αυξάνει κατά 0,64 ng/day για κάθε μg/day μεθυλικού υδραργύρου που προσλαμβάνουν.

Ο έλεγχος του συντελεστή της παλινδρόμησης γίνεται με μηδενική υπόθεση

$$H_0 : \beta_1 = 0$$

και εναλλακτική

$$H_A : \beta_1 \neq 0,$$

μέσω της ποσότητας

$$t = \frac{b_1}{se(b_1)}$$

η οποία, εφόσον ισχύει η μηδενική υπόθεση, ακολουθεί την κατανομή t με n-2 = 12-2 = 10 βαθμούς ελευθερίας.

$$s_{y|x} = \frac{\sqrt{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}}{n-2}$$

$$se\left(b_1\right) = \frac{s_{y|x}}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}}$$

Μεθυλικός υδράργυρος x_i	Συγκέντρωση υδραργύρου y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \hat{y}_i)^2$
180	90	-194,17	-129,17	25079,86	37700,69	21,34
200	120	-174,17	-99,17	17271,53	30334,03	158,26
230	125	-144,17	-94,17	13575,69	20784,03	2,62
410	290	35,83	70,83	2538,19	1284,03	2321,31
600	310	225,83	90,83	20513,19	51000,69	2853,70
550	290	175,83	70,83	12454,86	30917,36	1715,62
275	170	-99,17	-49,17	4875,69	9834,03	212,58
580	375	205,83	155,83	32075,69	42367,36	594,38
105	70	-269,17	-149,17	40150,69	72450,69	546,62
250	105	-124,17	-114,17	14175,69	15417,36	1184,74
460	205	85,83	-14,17	-1215,97	7367,36	4736,19
650	480	275,83	260,83	71946,53	76084,03	7153,78
$\bar{x}$ = 374,17	$\bar{y}$ = 219,17			$\sum_{i=1}^{12}(x_i - \bar{x})(y_i - \bar{y})$ =253441,67	$\sum_{i=1}^{12}(x_i - \bar{x})^2$ =395541,67	$\sum_{i=1}^{12}(y_i - \hat{y}_i)^2 =$ 21501,14

$$s_{y|x} = \frac{\sqrt{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}}{n-2} = \frac{\sqrt{21501{,}14}}{12-2} = 14{,}7$$

$$se\left(b_1\right) = \frac{s_{y|x}}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}} = \frac{14{,}7}{\sqrt{395541{,}67}} = 0{,}023$$

$$t = \frac{b_1}{se(b_1)} = \frac{0,64}{0,023} = 27,8$$

Για μια κατανομή *t* *με* 12 – 2 = 10 *βαθμούς ελευθερίας, η πιθανότητα να προ*κύψει μια τιμή τόσο ακραία όσο η 27,8 είναι *p* < 0,001 και, επομένως, η μηδενική υπόθεση ότι ο συντελεστής b_1 είναι ίσος με 0 απορρίπτεται.

Δηλαδή στον πληθυσμό των ενηλίκων από τον οποίο προέρχεται το τυχαίο δείγμα των 12 ατόμων, υπάρχει στατιστικά σημαντική σχέση μεταξύ κατανάλωσης μολυσμένων ψαριών και συγκέντρωσης υδραργύρου στο αίμα.

Σύμφωνα με τη σχέση αυτήν, η συγκέντρωση υδραργύρου στο αυξάνεται όσο αυξάνει η κατανάλωση μολυσμένων ψαριών.

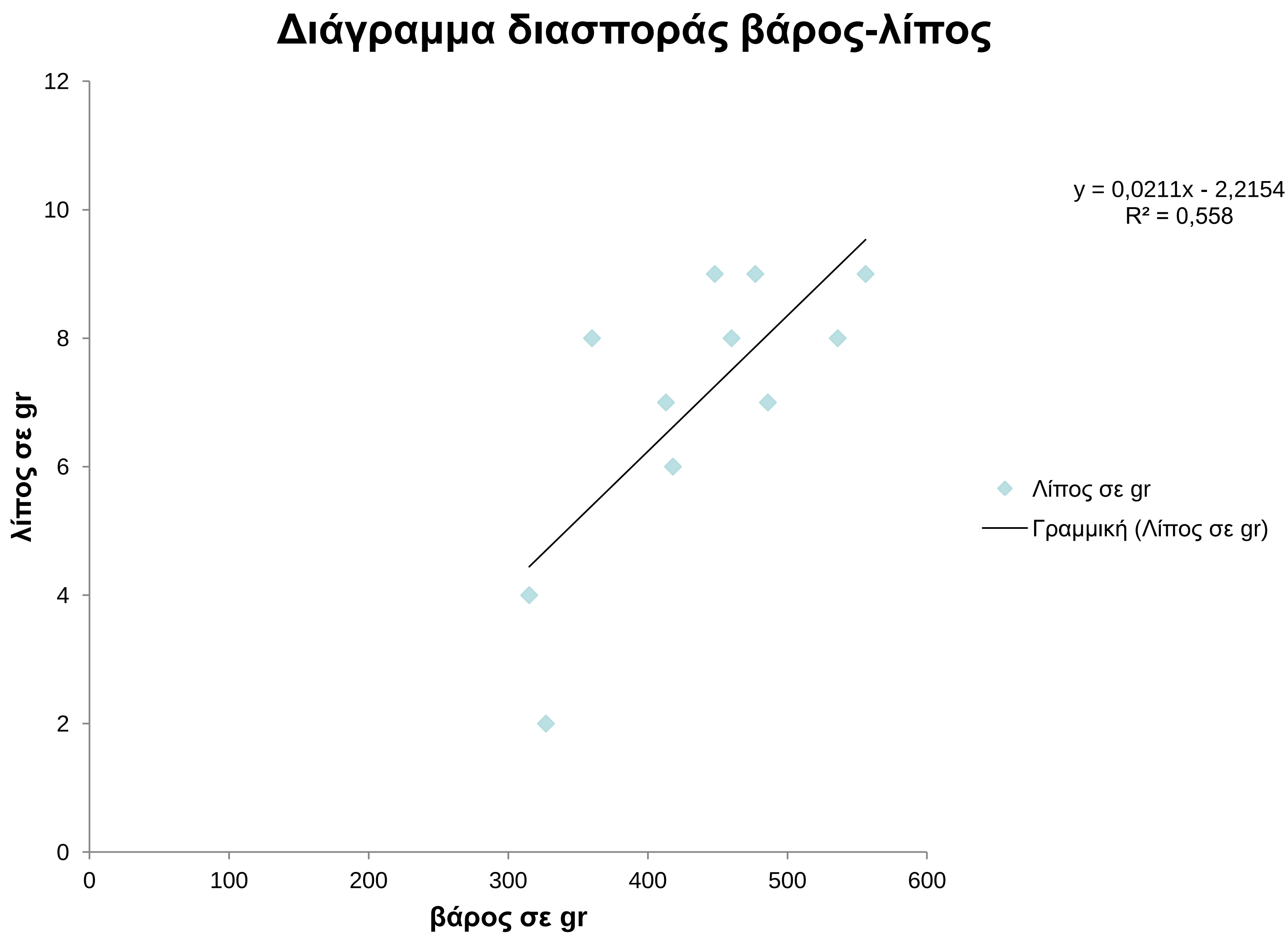
Άσκηση

Μετρήθηκε το βάρος και το λίπος 11 ψαριών (σε gr) στο τέλος μιας διαδικασίας πάχυνσης. Τα δεδομένα δίδονται στον παρακάτω πίνακα:

Βάρος σε gr	Λίπος σε gr
477	9
327	2
556	9
448	9
536	8
418	6
413	7
315	4
486	7
360	8
460	8

Να οριστεί η ευθεία της παλινδρόμησης με εξαρτημένη μεταβλητή τις τιμές του λίπους και ανεξάρτητη μεταβλητή το βάρος των ψαριών Να ερμηνεύσετε το συντελεστή b_1 της παλινδρόμησης.

Να γίνει επαγωγική αξιολόγηση του συντελεστού b1 της παλινδρόμησης και να διατυπώσετε το συμπέρασμα της αξιολόγησης.


Διάγραμμα διασποράς βάρος-λίπος
12
10
8
6
4
2
0
0
100
200
300
400
500
600
λίπος σε gr
βάρος σε gr
y = 0,0211x - 2,2154
R² = 0,558
Λίπος σε gr
Γραμμική (Λίπος σε gr)

Βάρος σε gr x_i	Λίπος σε gr y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$\hat{y}_i$	$(y_i - \hat{y})^2$
477	9	41	2	82	1681	7,86	1,30
327	2	-109	-5	545	11881	4,71	7,35
556	9	120	2	240	14400	9,52	0,27
448	9	12	2	24	144	7,25	3,06
536	8	100	1	100	10000	9,10	1,21
418	6	-18	-1	18	324	6,62	0,39
413	7	-23	0	0	529	6,52	0,23
315	4	-121	-3	363	14641	4,46	0,21
486	7	50	0	0	2500	8,05	1,10
360	8	-76	1	-76	5776	5,40	6,74
460	8	24	1	24	576	7,50	0,25
				1320	62452	77,00	22,10
$\bar{x} = 436$	$\bar{y} = 7$			$\sum(x_i - \bar{x})(y_i - \bar{y}) =$ 1320	$\sum(x_i - \bar{x})^2 =$ 62452		$\sum(y_i - \hat{y})^2 =$ 22,10

$$\boldsymbol{b}_1 = \frac{\sum_{\boldsymbol{i}=1}^{11}\left(\boldsymbol{x_i} - \overline{\boldsymbol{x}}\right)\left(\boldsymbol{y_i} - \overline{\boldsymbol{y}}\right)}{\sum_{\boldsymbol{i}=1}^{11}\left(\boldsymbol{x_i} - \overline{\boldsymbol{x}}\right)^2} = \frac{1320}{62452} = 0{,}021$$

$$\boldsymbol{b}_0 = \overline{\boldsymbol{y}} - \boldsymbol{b}_1\overline{\boldsymbol{x}} = 7 - \left(0{,}021\right)\cdot\left(436\right) = -2{,}156$$

$$s_{y|x} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{n-2} = \frac{\sqrt{\sum_{i=1}^{11}(y_i - \hat{y}_i)^2}}{11-2} = \frac{\sqrt{22{,}10}}{9} = 0{,}52$$

$$se(b_1) = \frac{s_{y|x}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{0{,}52}{\sqrt{62452}} = 0{,}002$$

Ο έλεγχος του συντελεστή της παλινδρόμησης γίνεται με μηδενική υπόθεση

$$H_0 : \beta_1 = 0$$

και εναλλακτική

$$H_A : \beta_1 \neq 0,$$

μέσω της ποσότητας

$$t = \frac{b_1}{se(b_1)} = \frac{0{,}021}{0{,}002} = 10{,}5$$

η οποία, εφόσον ισχύει η μηδενική υπόθεση, ακολουθεί την κατανομή t με n-2 = 11-2 = 9 βαθμούς ελευθερίας.

Ανατρέχοντας στον πίνακα της κατανομής t, για επίπεδο σημαντικότητας α=0,05 (δηλ. α=0,025) και βαθμούς ελευθερίας df=9 βλέπουμε ότι

$$|\boldsymbol{t}| = |10,5| > \boldsymbol{t}_9 = 2,262$$

επομένως, η μηδενική υπόθεση ότι ο συντελεστής $\boldsymbol{b_1}$ είναι ίσος με 0 απορρίπτεται.

Άρα υπάρχει σημαντική σχέση μεταξύ βάρους ψαριών και λίπους . Για κάθε γραμμάριο αύξησης του βάρους των ψαριών αυξάνει το λίπος κατά 0,021 γρ.