ΣΤΑΤΙΣΤΙΚΗ
ΠΡΟΒΛΗΜΑΤΑ ΓΡΑΜΜΙΚΗΣ ΠΑΛΙΝΔΡΟΜΙΣΗΣ

των Γ. Τσικαλουδάκη – Κ. Βακαλόπουλου

ΠΡΟΒΛΗΜΑ 1

Για να προσδιορίσουμε μια γραμμική σχέση μεταξύ της θερμοκρασίας και της διαστολής του χάλυβα κάναμε 10 μετρήσεις ακριβείας σε ράβδο 1m στους 0^oC και είχαμε τα παρακάτω αποτελέσματα:

Θερμοκρασίας θ	**5^o**	**10^o**	**15^o**	**20^o**	**25^o**	**30^o**	**35^o**	**40^o**	**45^o**	**50^o**
Μήκος $1m = x \cdot mm$	**0,3**	**0,5**	**0,8**	**0,9**	**1**	**1,1**	**1,2**	**1,3**	**1,4**	**1,5**

i) Να προσδιορίσετε την ευθεία παλινδρόμησης της αύξησης x του μήκους ως προς τη θερμοκρασία θ.

ii) Για τη ζεύξη Ρίου – Αντιρρίου με μεταλλική γέφυρα από χάλυβα θα χρησιμοποιήσουμε κομμάτια χάλυβα με διάκενο μεταξύ τους (για τις διαστολές) μήκους 100m το καθένα. Να προσδιορίσετε το ελάχιστο πλάτος που πρέπει να έχει το κάθε διάκενο, δεδομένου ότι στην θέση αυτή η μέγιστη θερμοκρασία που μπορεί να υπάρξει είναι 44^oC.

Λύση

Είναι $\nu = 10$

Ακόμη: $\hat{\beta} = \dfrac{\nu \cdot \sum_{i=1}^{10} x_i\theta_i - \left(\sum_{i=1}^{10}\theta_i\right)\left(\sum_{i=1}^{10}x_i\right)}{\nu \cdot \sum_{i=1}^{10}\theta_i^2 - \left(\sum_{i=1}^{10}\theta_i\right)^2} = 0,025$, $\bar{\theta} = \dfrac{1}{10}\sum_{i=1}^{10}\theta_i = \dfrac{275}{10} = 27,5$, $\bar{x} = \dfrac{1}{10}\sum_{i=1}^{10}x_i = \dfrac{10}{10} = 1$

											Σύνολα
θ	5	10	15	20	25	30	35	40	45	50	275
X	0,3	0,5	0,8	0,9	1	1,1	1,2	1,3	1,4	1,5	10
$\theta \cdot X$	1,5	5	12	18	25	33	42	52	63	75	326,5
θ^2	25	100	225	400	625	900	1225	1600	2025	2500	9625

Οπότε $\hat{\alpha} = \bar{x} - \hat{\beta}\cdot\bar{\theta} = 1 - 0,025 \cdot 27,5 = 0,3$ και η ευθεία παλινδρόμησης της x πάνω στην $\hat{\theta}$ είναι: $\hat{x} = 0,3 + 0,025 \cdot \theta$

i) Μια ράβδος μήκους 1m στους 44^oC αναμένεται να επιμηκυνθεί κατά $\hat{x} = 0,3 + 0,025 \cdot 44 = 1,40\ mm$.

ii) Όταν η ράβδος είναι 100m αναμένουμε να επιμηκυνθεί κατά $100 \cdot 1,4 = 140\ mm = 14\ cm$. Οπότε το ελάχιστο διάκενο θα πρέπει να είναι 14 cm.

ΠΡΟΒΛΗΜΑ 2

Στον παρακάτω πίνακα φαίνονται τα εβδομαδιαία κέρδη (σε εκ. δρχ.) μιας νέας βιοτεχνίας, κατά τις 5 πρώτες εβδομάδες λειτουργίας.

εβδομάδα (x)	**1**	**2**	**3**	**4**	**5**
κέρδη (Υ)	**2**	**2**	**3**	**2**	**1**

i) Να βρεθεί (με τη μέθοδο ελαχίστων τετραγώνων η ευθεία παλινδρόμησης των κερδών Υ πάνω στη x.

ii) Με βάση την παραπάνω ευθεία να εκτιμήσετε ποια εβδομάδα, η βιοτεχνία θα παρουσιάσει για πρώτη φορά ζημιά.

Λύση

X	Y	X^2	XY
1	2	1	2
2	2	4	4
3	3	9	9
4	2	16	8
5	1	25	5
$\sum_{i=1}^{5} x = 15$	$\sum_{i=1}^{5} Y = 10$	$\sum_{i=1}^{5} X^2 = 55$	$\sum_{i=1}^{5} XY = 28$

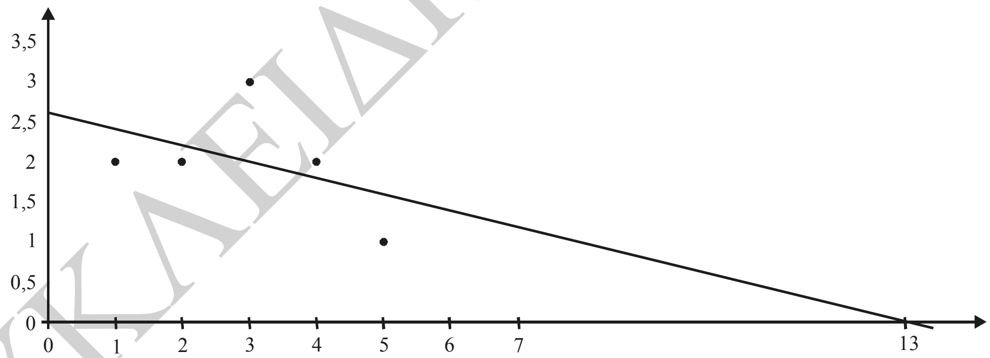
i) Είναι $v=5$, $\bar{X}=\dfrac{15}{5}=3$, $\bar{Y}=\dfrac{10}{5}=2$

$$\hat{\beta}=\frac{v\cdot\sum_{i=1}^{5}x_i y_i-\left(\sum_{i=1}^{5}x_i\right)\left(\sum_{i=1}^{5}y_i\right)}{v\cdot\sum_{i=1}^{5}x_i^2-\left(\sum_{i=1}^{5}x_i\right)^2}=\frac{5\cdot 28-15\cdot 10}{5\cdot 5-15^2}=-0,2$$

Οπότε: $\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}=2+0,2\cdot 3=2,6$.

Συνεπώς η ευθεία ελαχίστων τετραγώνων θα είναι: $\hat{y}=2,6-0,2x$ (1)

ii) Η βιοτεχνία θα παρουσιάσει για πρώτη φορά ζημιά όταν τα κέρδη της μηδενιστούν, όταν δηλαδή $\hat{y}=0$. Όπως φαίνεται από την ακόλουθη ευθεία ελαχίστων τετραγώνων, μετά από το σημείο $(x_0, 0)$ τα κέρδη παρουσιάζουν αρνητικές τιμές, έχουμε δηλαδή ζημιά μετά την τιμή $x_0=13$.



Η τιμή $\hat{y}=0$ προκύπτει από την (1) για: $x=13$ εβδομάδες.

ΠΡΟΒΛΗΜΑ 3

Στον παρακάτω πίνακα φαίνεται η αύξηση Y των κερδών (σε εκατ. δρχ) από τις πωλήσεις ενός προϊόντος, η οποία προκύπτει από την αντίστοιχη διαφήμιση κόστους X (εκατ. δρχ).

X	**1**	**3**	**4**	**5**	**7**
Y	**1**	**2**	**3**	**4**	**5**

i) Να βρεθεί, με την μέθοδο ελαχίστων τετραγώνων, η ευθεία παλινδρόμησης της μεταβολής Y πάνω στη X.

ii) Αν τα κέρδη, χωρίς διαφήμιση είναι $Y_0=10$ εκ. δρχ., να προσδιοριστεί σε πόσα διαφημιστικά έξοδα, εκτιμάται ότι θα έχουμε διπλασιασμό των κερδών Y_0.

Λύση

X	Y	X^2	XY

	1	1	1	1
	3	2	9	6
	4	3	16	12
	5	4	25	20
	7	5	49	35
Σύνολο	20	15	100	74

i) Είναι: $v=5$, $\bar{x}=\frac{20}{5}=4$, $\bar{y}=\frac{15}{5}=3$

$$\hat{\beta}=\frac{v\cdot\sum_{i=1}^{5}x_iy_i-\left(\sum_{i=1}^{5}x_i\right)\left(\sum_{i=1}^{5}y_i\right)}{v\cdot\sum_{i=1}^{5}x_i^2-\left(\sum_{i=1}^{5}x_i\right)^2}=\frac{5\cdot 74-20\cdot 15}{5\cdot 100-20^2}=\frac{70}{100}=0,7$$

Οπότε: $\hat{\alpha}=\bar{y}-\hat{\beta}\cdot\bar{x}=3-0,7\cdot 4=0,2$

Συνεπώς η ευθεία γραμμικής παλινδρόμησης της ευθεία Υ πάνω στην Χ είναι: $\hat{y}=0,2+0,7x$ (1)

ii) Από την (1) προκύπτει ότι για $x=14$, έχουμε $\hat{y}=10$, (διπλασιασμό κερδών).

ΠΡΟΒΛΗΜΑ 4

Στον παρακάτω πίνακα φαίνεται το κόστος συσκευασίας Υ (σε δρχ.), ανά μονάδα προϊόντος, όταν συσκευάζονται Χ (χιλιάδες) προϊόντα ημερησίως από 10 εργάτες μιας βιομηχανίας.

X	**1**	**2**	**3**	**4**	**5**
Y	**100**	**80**	**60**	**40**	**20**

i) Να βρεθεί (με τη μέθοδο ελαχίστων τετραγώνων) η ευθεία παλινδρόμησης της Υ πάνω στη Χ.

ii) Αν υποθέσουμε ότι ο ταχύτερος εργάτης μπορεί να συσκευάσει το πολύ 550 μονάδες προϊόντος ημερησίως να εκτιμηθεί το ελάχιστο κόστος συσκευασίας, ανά μονάδα προϊόντος που μπορεί να υπάρξει.

Λύση

X	Y	X^2	XY	
1	100	1	100	
2	80	4	160	
3	60	9	180	
4	40	16	160	
5	20	25	100	
15	300	55	700	Σύνολο

i) Είναι $v=5$, $\bar{x}=3$, $\bar{y}=60$

$$\hat{\beta}=\frac{5\cdot 700-15\cdot 300}{5\cdot 55-(15)^2}=-\frac{1000}{50}=-20$$

Οπότε: $\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}=60+20\cdot 3=120$ και συνεπώς η ευθεία παλινδρόμησης της Υ πάνω στη Χ είναι λοιπόν:

$$\hat{y}=120-20x \qquad (1)$$

ii) Από την (1) για $x=5,50$ προκύπτει $\hat{y}=120-20\cdot 5,5 \Leftrightarrow \hat{y}=120-110 \Leftrightarrow \hat{y}=10$

Άρα το ελάχιστο κόστος συσκευασίας που μπορεί να υπάρξει είναι 10 δρχ. ανά μονάδα προϊόντος.

ΠΡΟΒΛΗΜΑ 5

Α. **Μελετώντας 10 βρέφη ως προς την ηλικία Χ (σε εβδομάδες) και το βάρος Υ (σε Kg) βρήκαμε:**

$$\sum x_iy_i=80 \qquad \sum x_i=20 \qquad \sum y_i=30 \qquad \sum x_i^2=100$$

i) Να βρεθεί η ευθεία παλινδρόμησης του βάρους Υ πάνω στην ηλικία Χ.

ii) Να εκτιμηθεί το βάρος ενός νεογέννητου μετά από 8 εβδομάδες.

Β. **Ένας ερευνητής διαπίστωσε με βάση τον παρακάτω πίνακα ότι υπάρχει γραμμική σχέση μεταξύ ανεργίας X και εγκληματικότητας Y, σε μια έρευνά του σε 10 χώρες.**

Ανεργία X%	**2**	**2**	**3**	**3**	**3**	**5**	**6**	**8**	**9**	**9**
Εγκληματικότητα Y%	**3**	**2**	**4**	**2**	**3**	**4**	**4**	**5**	**6**	**7**

Να βρεθεί:

i) **η** $\hat{y} = \hat{\alpha} + \hat{\beta}x$

ii) **η εγκληματικότητα με ποσοστό ανεργίας 11%.**

Λύση

Α. i) Είναι: $\nu = 10$ βρέφη. Οπότε: $\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 2$, $\bar{y} = \frac{1}{10}\sum_{i=1}^{10} y_i = 3$

$$\hat{\beta} = \frac{\nu \cdot \sum_{i=1}^{5} x_i y_i - \left(\sum_{i=1}^{5} x_i\right)\left(\sum_{i=1}^{5} y_i\right)}{\nu \cdot \sum_{i=1}^{5} x_i^2 - \left(\sum_{i=1}^{5} x_i\right)^2} = \frac{10 \cdot 80 - 20 \cdot 30}{10 \cdot 100 - (20)^2} = 0,33$$

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 3 - 0,22 \cdot 2 = 2,34$. Συνεπώς: $\hat{y} = 2,34 + 0,33x$ (1)

ii) Θέτουμε στην (1) όπου $x = 8$ και έχουμε: $\hat{y} = 2,34 + 0,33 \cdot 8 = 4,98$ Kgr

Β. i) Είναι: $\nu = 10$, $\bar{x} = \frac{50}{10} = 5$, $\bar{y} = \frac{40}{10} = 4$

$\hat{\beta} = \frac{10 \cdot 238 - 50 \cdot 40}{10 \cdot 22 - (50)^2} = \frac{1220}{720} = 1,69$, $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 4 - 1,69 \cdot = -4,45$

Οπότε: $\hat{y} = -4,45 + 1,69x$ (1)

X	Y	X^2	XY	
2	3	4	6	
2	2	4	4	
3	4	9	12	
3	2	9	6	
3	3	9	9	
5	4	25	20	
6	4	36	24	
8	5	64	40	
9	6	81	54	
9	7	81	63	
50	40	322	238	Σύνολο

ii) Από την (1) για $x = 11$, προκύπτει: $\hat{y} = -4,45 + 1,69 \cdot 11 = 14,14\%$.

ΠΡΟΒΛΗΜΑ 6

Η ΔΕΗ σκοπεύει να κατασκευάσει ευθύγραμμο αγωγό μεταφοράς ηλεκτρικής ενέργειας ο οποίος θα διέρθει μεταξύ τεσσάρων πόλεων Α, Β, Γ, Δ οι οποίες έχουν συντεταγμένες ως προς το κέντρο παραγωγής ηλεκτρικής ενέργειας: Α(2, 2) Β(3, 8) Γ(5, 12) Δ(6, 10) αντίστοιχα.

i) **Να βρεθεί η εξίσωση της ευθείας του αγωγού ώστε η συνολική κατακόρυφη απόσταση και των τεσσάρων πόλεων από τον αγωγό να είναι ελάχιστη.**

ii) **Αν η μονάδα μήκους του αγωγού, με τον οποίο θα συνδεθεί μια πόλη με τον παραπάνω κεντρικό αγωγό, κοστίζει 500 χιλιάδες δρχ. και η σύνδεση της κάθε πόλης με τον κεντρικό**

αγωγό, γίνει με ευθύγραμμο αγωγό κάθετο στον άξονα $x'x$, να υπολογιστεί το συνολικό κόστος σύνδεσης των πόλεων με τον κεντρικό αγωγό.

Λύση

i) Έχουμε τον πίνακα τιμών:

X	2	3	5	6
Y	2	8	12	10

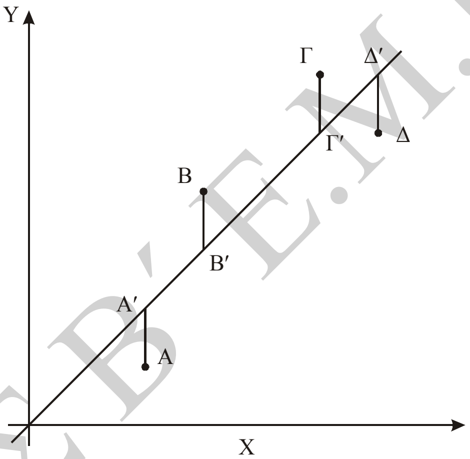
Η ζητούμενη ευθεία είναι η ευθεία ελαχίστων τετραγώνων: $\hat{y} = \hat{\alpha} + \hat{\beta}x$

Για τον υπολογισμό των $\hat{\alpha}, \hat{\beta}$ έχουμε: $\hat{\beta} = \dfrac{\nu \cdot \sum xy - \sum x \cdot \sum y}{\nu \cdot \sum x^2 - \left(\sum x\right)^2} = \dfrac{4 \cdot 148 - 16 \cdot 32}{4 \cdot 74 - 16^2} = \dfrac{80}{40} = 2$

$$\hat{\alpha} = \overline{y} - \hat{\beta}\overline{x} = 8 - 2 \cdot 4 = 0$$

Επομένως η ζητούμενη ευθεία είναι: $\hat{y} = 2x$

ii) Τα τμήματα των αγωγών με τα οποία οι πόλεις Α, Β, Γ, Δ θα συνδεθούν με τον κεντρικό αγωγό, καθέτως προς τον άξονα $x'x$ έχουν μήκος: (βλέπε σχήμα).



$$AA' = |2 \cdot 2 - 2| = 2, \quad BB' = |2 \cdot 3 - 8| = 2,$$

$$\Gamma\Gamma' = |5 \cdot 2 - 12| = 2, \quad \Delta\Delta' = |2 \cdot 6 - 10| = 2.$$

Επομένως το συνολικό μήκος είναι 8 μονάδες και κοστίζει 4 εκατ. δρχ.

Ερωτήσεις τύπου «ΣΩΣΤΟ – ΛΑΘΟΣ»

Δύο μεταβλητές Χ, Υ παρουσίασαν τις τιμές $x_i, y_i, i = 1, 2, ..., \nu$, σε ένα δείγμα μεγέθους ν.

Με τη μέθοδο ελαχίστων τετραγώνων βρήκαμε ότι η ευθεία παλινδρόμησης της y πάνω στη x είναι: $\hat{y} = 2 + x$.

Να απαντήσετε αν είναι σωστή η λάθος η πρόταση:

1. Αν $\overline{x} = 3$, τότε $\overline{y} = 5$.
2. Αν $\hat{x} = 2 + \frac{1}{2}y$, τότε $\overline{x} = 6$ και $\overline{y} = 8$.
3. Αν $x_1 = 2$, τότε $y_1 = 4$.
4. Η προβλεπόμενη τιμή της Υ για x=3 είναι 5.
5. Η ευθεία παλινδρόμησης της Χ πάνω στην Υ είναι $\hat{x} = -2 + y$.
6. Αν $\overline{x} = 3$, τότε η ευθεία παλινδρόμησης της Χ πάνω στην Υ μπορεί να είναι η $\hat{x} = \frac{1}{2} + \frac{1}{2}y$.
7. Η ευθεία παλινδρόμησης της Χ πάνω στην Υ έχει τη μορφή: $\hat{x} = 1 - 3\hat{\beta} + \hat{\beta}y$.
8. Αν η προβλεπόμενη τιμή της Χ για y=4 είναι $\hat{x} = 3$, τότε $\overline{y} = 3$.

Βιβλιογραφία: «Στατιστική» Γ΄ Λυκείου 2001 Γ. Τσικαλουδάκη