ΕΠΑΝΑΛΗΠΤΙΚΑ ΘΕΜΑΤΑ

ΚΕΦΑΛΑΙΟ 2ο: ΣΤΑΤΙΣΤΙΚΗ

ΘΕΜΑ Α

Ερώτηση θεωρίας 1

Τι λέγεται ιστόγραμμα αθροιστικών απολύτων – σχετικών συχνοτήτων;

<u>Λύση</u>

Ιστόγραμμα αθροιστικών απολύτων ή σχετικών συχνοτήτων είναι μια σειρά από διαδοχικά ορθογώνια παραλληλόγραμμα που έχουν βάσεις ίσες με το πλάτος των κλάσεων c και ύψη ίσα με την αθροιστική απόλυτη ή σχετική συχνότητα της κλάσης που αντιπροσωπεύουν.

Ερώτηση θεωρίας 2

Τι ονομάζεται συχνότητα κλάσης.

Λύση

Συχνότητα μιας κλάσης ονομάζεται το πλήθος των παρατηρήσεων ν_i που προκύπτουν από τη διαλογή για την κλάση i.

Ερώτηση θεωρίας 3

α) Έστω $x_1, x_2, ..., x_k$ οι τιμές μιας μεταβλητής X και $f_1, f_2, ..., f_k$ οι αντίστοιχες σχετικές συχνότητες των προηγούμενων τιμών. Να αποδείξετε ότι:

i. $0 \le f_i \le 1$ για κάθε $i = 1, 2, ..., k$.

ii. $f_1 + f_2 + ... + f_k = 1$.

β) Πως ορίζεται ο συντελεστής μεταβολής ή μεταβλητότητας ενός δείγματος τιμών;

<u>Λύση</u>

α)

i. Ισχύει $f_i = \dfrac{\nu_i}{\nu}$, όπου ν_i η (απόλυτη) συχνότητα της τιμής x_i.
Επίσης για $i = 1, 2, ..., k$ έχουμε
$0 \le \nu_i \le \nu \Leftrightarrow 0 \le \dfrac{\nu_i}{\nu} \le \dfrac{\nu}{\nu} = 1$. Συνεπώς $0 \le f_i \le 1$ για κάθε $i = 1, 2, ..., k$.

ii. Ισχύει $f_1 + f_2 + ... + f_k = \dfrac{\nu_1}{\nu} + \dfrac{\nu_2}{\nu} + \ldots + \dfrac{\nu_k}{\nu} = \dfrac{\nu_1 + \nu_2 + \ldots \nu_k}{\nu} = \dfrac{\nu}{\nu} = 1$, άρα αποδείχτηκε.

β) Αν $\overline{x}$ η μέση τιμή ενός δείγματος τιμών μιας μεταβλητής και s η τυπική απόκλιση των τιμών αυτών, τότε ο συντελεστής μεταβολής ή συντελεστής μεταβλητότητας του δείγματος για $\overline{x} \ne 0$ ορίζεται από το λόγο:

$$CV = \frac{s}{\overline{x}}.$$

Αν $\overline{x} < 0$, τότε αντί της $\overline{x}$ χρησιμοποιούμε την $|\overline{x}|$.

Ερώτηση θεωρίας 4

α) Έστω $t_1, t_2, ..., t_\nu$ οι τιμές ενός δείγματος και $\overline{x}$ η μέση τιμή αυτών. Θεωρούμε τις τιμές y_i με $y_i = t_i - \overline{x}$, $i = 1, 2, ..., \nu$. Να δείξετε ότι η μέση τιμή των τιμών y_i είναι $\overline{y} = 0$.

β) Δίνεται ένα δείγμα ν παρατηρήσεων. Πως ορίζεται η διάμεσος του δείγματος;

Λύση

α) Έχουμε

$$\overline{y} = \frac{(t_1 - \overline{x}) + (t_2 - \overline{x}) + \ldots + (t_\nu - \overline{x})}{\nu} = \frac{t_1 + t_2 + \ldots t_\nu - \nu \cdot \overline{x}}{\nu} =$$

$$\frac{t_1 + t_2 + \ldots t_\nu}{\nu} - \frac{\nu \cdot \overline{x}}{\nu} = \overline{x} - \overline{x} = 0 .$$

β) Διάμεσος δ ενός δείγματος ν παρατηρήσεων οι οποίες έχουν διαταχθεί σε αύξουσα σειρά ορίζεται η μεσαία παρατήρηση, όταν το ν είναι περιττός αριθμός, ή ο μέσος όρος (ημιάθροισμα) των δυο μεσαίων παρατηρήσεων όταν το ν είναι άρτιος αριθμός.

Ερώτηση θεωρίας 5

Αν $x_1, x_2, ..., x_\kappa$ οι τιμές μιας μεταβλητής Χ σ' ένα δείγμα μεγέθους ν, τότε:

i. Τι λέγεται συχνότητα ν_i της τιμής $x_i, i = 1, 2, ..., \kappa$;

ii. Τι λέγεται σχετική συχνότητα f_i της τιμής $x_i, i = 1, 2, ..., \kappa$;

iii. Τι εκφράζει η αθροιστική συχνότητα N_i της τιμής $x_i, i = 1, 2, ..., \kappa$;

iv. Να δείξετε ότι $f_1 + f_2 + ... + f_\kappa = 1$

<u>Λύση</u>

i. Συχνότητα της τιμής $x_i, i = 1, 2, ..., \kappa$ λέγεται ο φυσικός αριθμός ν_i που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της εξεταζόμενης μεταβλητής στο σύνολο των παρατηρήσεων.

ii. Σχετική συχνότητα f_i της τιμής $x_i, i = 1, 2, ..., \kappa$ λέγεται ο αριθμός $f_i = \frac{\nu_i}{\nu}, i = 1, 2, ..., \kappa$.

iii. Η αθροιστική συχνότητα N_i της τιμής $x_i, i = 1, 2, ..., \kappa$ εκφράζει το πλήθος των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i , όταν οι τιμές $x_1, x_2, ..., x_\kappa$ έχουν τοποθετηθεί σε αύξουσα σειρά.

iv. Είναι $f_1 + f_2 + ... + f_\kappa = \frac{\nu_1}{\nu} + \frac{\nu_2}{\nu} + ... + \frac{\nu_\kappa}{\nu} = \frac{\nu_1 + \nu_2 + ... + \nu_\kappa}{\nu} = \frac{\nu}{\nu} = 1$.

Ερώτηση θεωρίας 6

Έστω $t_1, t_2, ..., t_\nu$ οι παρατηρήσεις μιας ποσοτικής μεταβλητής X ενός δείγματος μεγέθους ν, που έχουν μέση τιμή $\overline{x}$ και τυπική απόκλιση s.

i. Να αποδείξετε ότι ο αριθμητικός μέσος των διαφορών $t_1 - \overline{x}, t_2 - \overline{x}, ..., t_\nu - \overline{x}$ είναι ίσος με μηδέν.

ii. Πώς ορίζεται ο συντελεστής μεταβολής CV του δείγματος;

<u>Λύση</u>

i. Ο αριθμητικός μέσος των διαφορών $t_1 - \overline{x}, t_2 - \overline{x}, ..., t_\nu - \overline{x}$ είναι ίσος με

$$\frac{(t_1 - \overline{x}) + (t_2 - \overline{x}) + ... + (t_\nu - \overline{x})}{\nu} = \frac{t_1 + t_2 + ... + t_\nu}{\nu} - \frac{\nu \cdot \overline{x}}{\nu} = \overline{x} - \overline{x} = 0 .$$

ii. Ο συντελεστής μεταβολής είναι ο αριθμός:

$CV = \frac{s}{\overline{x}}$, αν $\overline{x} > 0$ ή $CV = \frac{s}{|\overline{x}|}$, αν $\overline{x} < 0$.

Ερώτηση θεωρίας 7

α) Πότε μία ποσοτική μεταβλητή ονομάζεται διακριτή και πότε συνεχής;

β) Αν $x_1, x_2, \ldots\ldots x_\nu$ είναι οι παρατηρήσεις μιας ποσοτικής μεταβλητής X ενός δείγματος μεγέθους ν και $w_1, w_2, \ldots\ldots w_\nu$ οι αντίστοιχοι συντελεστές στάθμισης (βαρύτητας), να ορίσετε το σταθμικό μέσο της μεταβλητής X.

<u>Λύση</u>

α) Διακριτή ονομάζεται η ποσοτική μεταβλητή που παίρνει μόνο "μεμονωμένες" τιμές.
Συνεχής ονομάζεται η ποσοτική μεταβλητή που μπορεί να πάρει οποιαδήποτε τιμή σε ένα διάστημα πραγματικών αριθμών (α,β).

β) Ο σταθμικός μέσος ορίζεται από τον τύπο:

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_\nu w_\nu}{w_1 + w_2 + \ldots. + w_\nu} = \frac{\sum_{i=1}^{\nu} x_i w_i}{\sum_{i=1}^{\nu} w_i}.$$

Ερώτηση θεωρίας 8

α) Να δώσετε τον ορισμό της διαμέσου δ ενός δείγματος, ν παρατηρήσεων περιττού πλήθους.

β) Αν $w_1, w_2, ..., w_\nu$, είναι οι συντελεστές βαρύτητας, ενός δείγματος ν παρατηρήσεων $x_1, x_2, ..., x_\nu$, μιας ποσοτικής μεταβλητής Χ, να δώσετε τον ορισμό του σταθμικού μέσου της μεταβλητής Χ.

<u>Λύση</u>

α) Διάμεσος δ ενός δείγματος, ν παρατηρήσεων περιττού πλήθους, που έχουν διαταχθεί σε αύξουσα σειρά, ορίζεται ως η μεσαία τους παρατήρηση.

β) Αν $w_1, w_2, ..., w_\nu$, είναι οι συντελεστές βαρύτητας, ενός δείγματος ν παρατηρήσεων $x_1, x_2, ..., x_\nu$, μιας ποσοτικής μεταβλητής Χ, τότε ο σταθμικός μέσος, ορίζεται ως εξής:

$$\overline{x} = \frac{x_1 w_1 + x_2 w_2 + ... + x_\nu w_\nu}{w_1 + w_2 + ... + w_\nu} = \frac{\sum_{i=1}^{\nu} x_i w_i}{\sum_{i=1}^{\nu} w_i}$$

Ερώτηση θεωρίας 9

α) Σε ένα δείγμα μεγέθους ν , αν $x_1, x_2, ..., x_\kappa$, είναι οι τιμές μιας μεταβλητής X, με αντίστοιχες απόλυτες συχνότητες $\nu_1, \nu_2, ..., \nu_\kappa$, όπου $\kappa \le \nu$, ενώ $f_1, f_2, ..., f_\kappa$, είναι οι αντίστοιχες σχετικές τους συχνότητες, τότε να αποδείξετε ότι ισχύουν οι ιδιότητες:

i. $0 \le f_i \le 1$, για $i = 1, 2, ..., \kappa$

ii. $f_1 + f_2 + ... + f_\kappa = 1$.

β) Να δώσετε τον ορισμό της διαμέσου δ ενός δείγματος $,\nu$ παρατηρήσεων άρτιου πλήθους.

<u>Λύση</u>

α)

i. Για κάθε $\nu_1, \nu_2, ..., \nu_\kappa$, $i = 1, 2, ..., \kappa$, ισχύουν :

$$0 \le \nu_1 \le \nu , \quad 0 \le \nu_2 \le \nu, ..., 0 \le \nu_\kappa \le \nu .$$

Διαιρώντας κάθε μια από τις πιο πάνω σχέσεις δια ν (που είναι θετικός), έχουμε:

$$\frac{0}{\nu} \le \frac{\nu_1}{\nu} \le \frac{\nu}{\nu} \Leftrightarrow 0 \le f_1 \le 1 , \ \frac{0}{\nu} \le \frac{\nu_2}{\nu} \le \frac{\nu}{\nu} \Leftrightarrow 0 \le f_2 \le 1 , \dots ,$$

$$\frac{0}{\nu} \le \frac{\nu_\kappa}{\nu} \le \frac{\nu}{\nu} \Leftrightarrow 0 \le f_\kappa \le 1 ,$$

Αφού είναι: $f_i = \dfrac{\nu_i}{\nu}$, $i = 1, 2, ..., \kappa$

Άρα, ισχύει:

$$0 \le f_i \le 1 , \text{ για } i = 1, 2, ..., \kappa .$$

ii. Έχουμε:

$$f_1 + f_2 + ... + f_\kappa = \frac{\nu_1}{\nu} + \frac{\nu_2}{\nu} + ... + \frac{\nu_\kappa}{\nu} = \frac{\nu_1 + \nu_2 + ... + \nu_\kappa}{\nu} = \frac{\nu}{\nu} = 1 ,$$

Αφού είναι, $f_i = \dfrac{\nu_i}{\nu}$, $i = 1, 2, ..., \kappa$ και $\nu_1 + \nu_2 + ... + \nu_\kappa = \nu$, που είναι το μέγεθος του δείγματος.

β) Διάμεσος δ ενός δείγματος ν παρατηρήσεων, άρτιου πλήθους, που έχουν διαταχθεί σε αύξουσα σειρά, ορίζεται ως ο μέσος όρος (ημιάθροισμα) των δύο μεσαίων παρατηρήσεων.

ΘΕΜΑ Β

Άσκηση 1

Ένα μεσιτικό γραφείο κατέταξε σε τέσσερις κλάσεις ίσου πλάτους ένα δείγμα 60 οικοπέδων ανάλογα με την τιμή πώλησης σε ευρώ του τ.μ.

α) Να συμπληρώσετε τον παρακάτω πίνακα

[-)	x_i	ν_i	N_i	f_i	$f_i\%$	$F_i\%$
[1250,)				0,2		
			30			
[, 2750)						
					10	
	Σύνολο	60				

β) Να βρεθούν τα ποσοστά των οικοπέδων που έχουν τιμή
- i. το πολύ 2000 ευρώ
- ii. τουλάχιστον 2500 ευρώ

γ) Να βρεθεί η μέση τιμή και η διάμεσος του δείγματος.

<u>Λύση</u>

α) Η διαφορά του κατώτερου από το ανώτερο όριο μιας κλάσης ονομάζεται πλάτος της κλάσης και συμβολίζεται με c.
Επομένως η διαφορά του κατώτερου ορίου της πρώτης κλάσης από το ανώτερο όριο της τρίτης κλάσης θα είναι 3c, δηλαδή, $3c = 2750 - 1250 = 1500$, οπότε c=500 και οι κλάσεις διαμορφώνονται
ως εξής:

$[1250,1750),[1750,2250),[2250,2750)$ και $[2750,3250)$.

Η κεντρική τιμή x_i μιας κλάσης $[\alpha,\beta)$ είναι ίση με το ημιάθροισμα των άκρων της άρα $x_i = \frac{\alpha+\beta}{2}$ οπότε $x_1 = \frac{1250+1750}{2} = 1500,\ x_2 = 2000,\ x_3 = 2500$ και $x_4 = 3000$

- Επειδή $\nu = 60$ και $f_1 = 0,2$ από τον τύπο $f_i = \frac{\nu_i}{\nu}$ έχουμε:

  $f_1 = \frac{\nu_1}{\nu} \Leftrightarrow \nu_1 = f_1 \cdot \nu \Leftrightarrow \nu_1 = 0,2 \cdot 60 \Leftrightarrow \nu_1 = 12$,
  οπότε $N_1 = 12$.
  Επιπλέον έχουμε $f_1\% = 20$ και $F_1\% = 20$.

- Επειδή $N_2 = 30$ και $\nu_1 = 12$ από τον τύπο $N_i = N_{i-1} + \nu_i$ έχουμε:

$N_2 = N_1 + \nu_2 \Leftrightarrow 30 = 12 + \nu_2 \Leftrightarrow \nu_2 = 18$, οπότε $f_2 = \frac{\nu_2}{\nu} \Leftrightarrow$

$f_2 = \frac{18}{60} \Leftrightarrow f_2 = 0,30, f_2\% = 30$ και

$F_2\% = F_1\% + f_2\% = 20 + 30 = 50$.

- Επειδή $\nu = 60$ και $f_4\% = 10$ ή $f_4 = 0,10$ από τον τύπο $f_i = \frac{\nu_i}{\nu}$ έχουμε:

  $f_4 = \frac{\nu_4}{\nu} \Leftrightarrow \nu_4 = f_4 \cdot \nu \Leftrightarrow \nu_4 = 0,10 \cdot 60 \Leftrightarrow \nu_4 = 6$.

- Επειδή $\nu = 60, \nu_3 = 60 - (12 + 18 + 6) = 60 - 36, \nu_3 = 24$, άρα
  $N_3 = N_2 + \nu_3 = 30 + 24 = 54$,
  $f_3 = \frac{\nu_3}{\nu} \Leftrightarrow f_3 = \frac{24}{60} \Leftrightarrow f_3 = 0,40$, οπότε $F_3\% = F_2\% + f_3\% = 50 + 40 = 90$ και
  $F_4\% = F_3\% + f_4\% = 90 + 10 = 100$.

  Ακόμα $N_4 = N_3 + \nu_4 = 54 + 6 = 60$.

Ο πίνακας συμπληρωμένος διαμορφώνεται ως εξής.

[-)	x_i	ν_i	N_i	f_i	$f_i\%$	$F_i\%$
[1250, 1750)	1500	12	12	0,2	20	20
[1750, 2250)	2000	18	30	0,3	30	50
[2250, 2750)	2500	24	54	0,4	40	90
[2750, 3250)	3000	6	60	0,1	10	100
Σύνολο		60		1	100	

β)

i. Τα οικόπεδα με τιμή ανά τ.μ το πολύ 2000 ευρώ είναι όσα κατανέμονται στην πρώτη κλάση $[1250,1750)$ και τα μισά από αυτά που ανήκουν στην 2η κλάση $[1750, 2250)$ αφού το 2000 είναι το κέντρο της 2ης κλάσης. Το ποσοστό των οικοπέδων με τιμή ανά τ.μ το πολύ 2000 ευρώ είναι επομένως $f_1\% + f_2\% / 2 = 20 + (30/2) = 35$

ii. Τα οικόπεδα με τιμή ανά τ.μ τουλάχιστον 2500 ευρώ είναι τα μισά από όσα κατανέμονται στην τρίτη κλάση $[2250, 2750)$ αφού το 2500 είναι το κέντρο της κλάσης αυτής και όσα ανήκουν στην τελευταία κλάση $[2750,3250)$. Το ποσοστό των οικοπέδων με τιμή ανά τ.μ τουλάχιστον 2500 ευρώ είναι επομένως $f_3\% / 2 + f_4\% = 20\% + 10\% = 30\%$

γ) Η μέση τιμή δίνεται από τον τύπο $\overline{x} = \dfrac{x_1\nu_1 + x_2\nu_2 + \cdots + x_\kappa\nu_\kappa}{\nu_1 + \nu_2 + \cdots \nu_\kappa} = \dfrac{\sum_{i=1}^{\kappa} x_i\nu_i}{\nu}$,

άρα η μέση τιμή πώλησης ανά τ.μ. των 60 οικοπέδων του δείγματος είναι

$$\overline{x} = \frac{12 \cdot 1500 + 18 \cdot 2000 + 24 \cdot 2500 + 6 \cdot 3000}{60} = 2200$$

Παρατηρούμε από τον πίνακα ότι το 50% των παρατηρήσεων έχουν τιμή κάτω από 2250 ευρώ. Άρα η διάμεσος είναι 2250 ευρώ.

Άσκηση 2

Ο πίνακας αναφέρεται στη βαθμολογία 200 φοιτητών στο μάθημα της Στατιστικής.

Κλάσεις	x_i	x_i^2	v_i	N_i	$x_i \cdot v_i$	$x_i^2 \cdot v_i$
[0, 2)			35			
[2, 4)			40			
[4, 6)			45			
[6, 8)			50			
(8, 10)			30			
Σύνολο			200			

α) Να συμπληρώσετε τον πίνακα

β) Να υπολογιστεί η μέση τιμή, η τυπική απόκλιση και ο συντελεστή μεταβλητότητας

γ) Ποιο βαθμό δεν υπερβαίνει το 85% των φοιτητών;

Λύση

α) Ο πίνακας συμπληρωμένος είναι ο εξής :

Κλάσεις	x_i	x_i^2	v_i	N_i	$x_i \cdot v_i$	$x_i^2 \cdot v_i$
[0, 2)	1	1	35	35	35	35
[2, 4)	3	9	40	75	120	360
[4, 6)	5	25	45	120	225	1125
[6, 8)	7	49	50	170	350	2450
(8, 10)	9	81	30	200	270	2430
Σύνολο			200		1000	6400

β) Η μέση τιμή της βαθμολογία του δείγματος των 200 φοιτητών είναι:

$$\overline{x} = \frac{\sum_{i=1}^{5} x_i \cdot \nu_i}{\nu} = \frac{1000}{200} = 5$$

Για τον υπολογισμό της τυπικής απόκλισης εφαρμόζουμε τον τύπο:

$$s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{5} x_i^2 \cdot \nu_i - \frac{\left(\sum_{i=1}^{5} x_i \nu_i\right)^2}{\nu}\right\}$$

Με βάση τον πίνακα η διακύμανση της μεταβλητής x είναι:

$$s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{5} x_i^2 \cdot \nu_i - \frac{\left(\sum_{i=1}^{5} x_i \cdot \nu_i\right)^2}{\nu}\right\} = \frac{1}{200}\left[6400 - \frac{1000^2}{200}\right] = \frac{1}{200}\left[6400 - \frac{1000000}{200}\right] =$$

$$\frac{1}{200}\left[6400 - 5000\right] = \frac{1}{200}1400 = 7$$

και η τυπική απόκλιση $s = \sqrt{7} \approx 2,65$

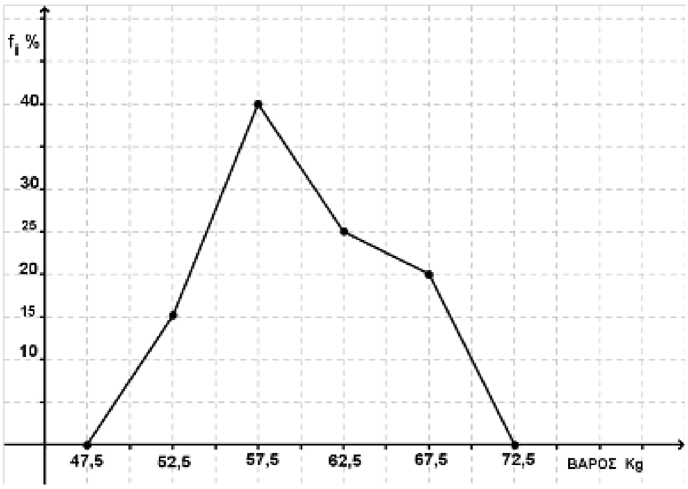
Ο συντελεστής μεταβλητότητας δίνεται από τον τύπο $CV = \frac{s}{\overline{x}} \Leftrightarrow CV = \frac{2,65}{5} = 0,53$

γ) Το 85% των φοιτητών είναι $85\% \cdot 200 = 170$ φοιτητές$= N_4$.

Από τον πίνακα παρατηρούμε ότι ο βαθμός των φοιτητών αυτών δεν υπερβαίνει το 8.

Άσκηση 3

Μια ομάδα μαθητών ενός Γυμνασίου μετρήθηκε ως προς το βάρος και προέκυψε το παρακάτω πολύγωνο σχετικών συχνοτήτων επί τοις εκατό.



Αν επιπλέον γνωρίζουμε ότι 8 μαθητές ζυγίζουν πάνω από 65 Kg, τότε:

α. Να συμπληρώσετε τον παρακάτω πίνακα

Κλάσεις $[\ ,\ )$	x_i	v_i	$f_i\%$	N_i	$F_i\%$
Σύνολο					

β. Να βρείτε τη μέση τιμή και τη διάμεσο των τιμών του δείγματος.

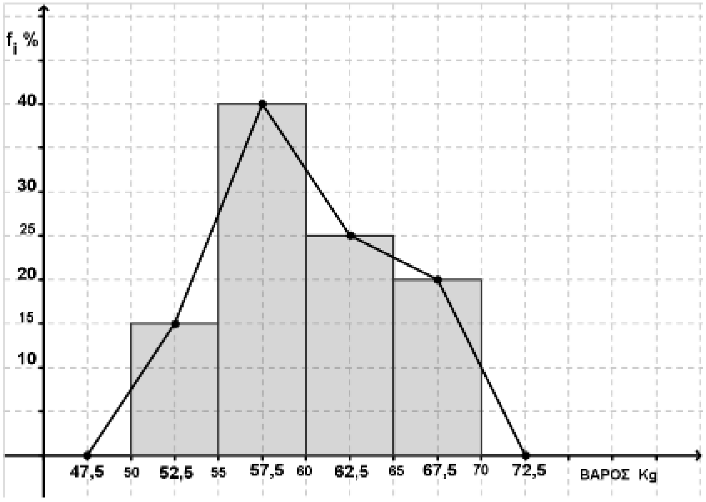
γ. Να βρείτε το ποσοστό των μαθητών που έχουν βάρος μέχρι 68 Kg.

<u>Λύση</u>

α. Η οριζόντια απόσταση των κορυφών του πολυγώνου είναι 5, άρα και το πλάτος των κλάσεων είναι $c=5$. Τα άκρα λοιπόν των κλάσεων θα είναι οι αριθμοί:

50, 55, 60, 65, 70. ($47{,}5+\frac{5}{2}=50$, $50+5=55$, $55+5=60$, $60+5=65$, $65+5=70$.)

Έτσι κατασκευάζουμε το ιστόγραμμα σχετικών συχνοτήτων επί τοις εκατό:



Έτσι έχουμε:

$f_1 = 0,15$,

$f_2 = 0,4$,

$f_3 = 0,25$ και

$f_4 = 0,2$.

Επίσης $\nu_4 = 8$, οπότε αν ν ο αριθμός των μαθητών έχουμε

$$f_4 = 0,2 \Leftrightarrow \frac{\nu_4}{\nu} = 0,2 \Leftrightarrow \nu = \frac{8}{0,2} = 40.$$

Από τον τύπο $\nu_i = f_i \cdot \nu$ βρίσκουμε

$\nu_1 = 0,15 \cdot 40 = 6$,

$\nu_2 = 0,4 \cdot 40 = 16$ και

$\nu_3 = 0,25 \cdot 40 = 10$.

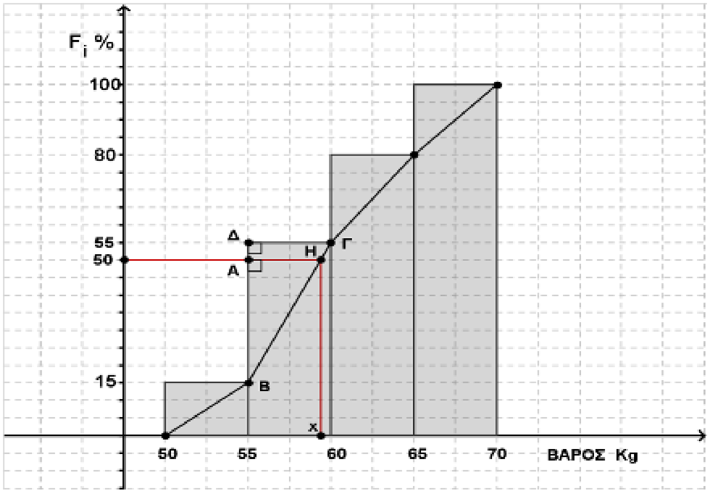
Από τα παραπάνω συμπληρώνουμε τον πίνακα:

Κλάσεις $[\ ,\ )$	x_i	v_i	f_i%	N_i	F_i%
$[50,55)$	52,5	6	15	6	15
$[55,60)$	57,5	16	40	22	55
$[60,65)$	62,5	10	25	32	80
$[65,70)$	67,5	8	20	40	100
Σύνολο		40	100		

β. Η μέση τιμή δίνεται από τον τύπο $\overline{x}=\sum_{i=1}^{k} x_i \cdot f_i$, οπότε

$$\overline{x}=52,5\cdot 0,15+57,5\cdot 0,4+62,5\cdot 0,25+67,5\cdot 0,2=60$$

και για τη διάμεσο σχεδιάζουμε το πολύγωνο σχετικών αθροιστικών συχνοτήτων επί τοις εκατό:



Φέρνουμε την οριζόντια ευθεία που τέμνει τον κατακόρυφο άξονα στο 50% και το πολύγωνο στο σημείο Η, και από το σημείο Η την κατακόρυφη ευθεία που τέμνει τον οριζόντιο άξονα στο σημείο με τετμημένη x .

Εφαρμόζοντας το θεώρημα του Θαλή για τις κατακόρυφες (άρα παράλληλες) ευθείες που διέρχονται από τα σημεία Β, Η και Γ έχουμε:

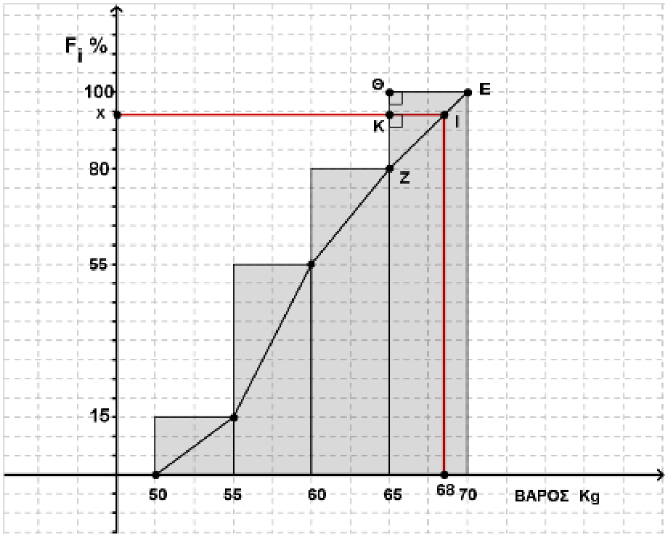
$$\frac{x-55}{60-55}=\frac{ΒΗ}{ΒΓ} .$$

Επίσης τα ορθογώνια τρίγωνα $ΒΑΗ$ και $ΒΔΓ$ είναι όμοια αφού έχουν κοινή την οξεία γωνία $\hat{Β}$, οπότε:

$$\frac{ΒΗ}{ΒΓ}=\frac{ΒΑ}{ΒΔ} \Leftrightarrow \frac{x-55}{60-55}=\frac{50-15}{55-15} \Leftrightarrow \frac{x-55}{5}=\frac{35}{40} \Leftrightarrow x=59,375$$

Άρα η διάμεσος είναι $\delta=59,375$.

γ.



Φέρνουμε την κατακόρυφη ευθεία που τέμνει τον οριζόντιο άξονα στο 68 και το πολύγωνο στο σημείο Ι, και από το σημείο Ι την οριζόντια ευθεία που τέμνει τον κατακόρυφο άξονα στο σημείο με τεταγμένη x.

Εφαρμόζοντας το θεώρημα του Θαλή για τις κατακόρυφες (άρα παράλληλες) ευθείες που διέρχονται από τα σημεία Ζ, Ι και Ε έχουμε:

$$\frac{68-65}{70-65}=\frac{\mathrm{ZI}}{\mathrm{EZ}}.$$

Επίσης τα ορθογώνια τρίγωνα $\mathrm{E\Theta Z}$ και IKZ είναι όμοια αφού έχουν κοινή την οξεία γωνία $\hat{\mathrm{Z}}$, οπότε:

$$\frac{\mathrm{ZI}}{\mathrm{EZ}}=\frac{\mathrm{ZK}}{\mathrm{\Theta Z}}\Leftrightarrow\frac{68-65}{70-65}=\frac{x-80}{100-80}\Leftrightarrow\frac{3}{5}=\frac{x-80}{20}\Leftrightarrow x=92.$$

Άρα το 92% των μαθητών έχει βάρος κάτω από 68 kg.

Άσκηση 4

Μια μεταβλητή X παίρνει τις τιμές $x_1 = 3$, $x_2 = 5$, $x_3 = 6$ και $x_4 = 7$. Αν οι συχνότητες των τιμών x_1, x_2, και x_4 είναι $\nu_1 = 10$, $\nu_2 = 3$, και $\nu_4 = 6$ και η γωνία στο κυκλικό διάγραμμα που αντιστοιχεί στην τιμή x_2 είναι $\omega_2 = 40^{\circ}$, τότε

α. Να δείξετε ότι $\nu_3 = 8$.
β. Να βρείτε τη μέση τιμή και τη διάμεσο των τιμών του δείγματος.
γ. Να δείξετε ότι η διασπορά του δείγματος είναι $s^2 = \dfrac{8}{3}$.
δ. Να εξετάσετε αν το δείγμα είναι ομοιογενές.

<u>Λύση</u>

α. Από τον τύπο $\omega_2 = 360^{\circ} \cdot \dfrac{\nu_2}{\nu}$ παίρνουμε $\nu = \dfrac{360^{\circ} \cdot \nu_2}{\omega_2} = \dfrac{360^{\circ} \cdot 3}{40^{\circ}} = 27$, άρα το πλήθος των τιμών είναι 27, οπότε $\nu_3 = \nu - \nu_1 - \nu_2 - \nu_4 = 27 - 10 - 3 - 6 = 8$.

β. Η μέση τιμή είναι $\bar{x} = \dfrac{3 \cdot 10 + 5 \cdot 3 + 6 \cdot 8 + 7 \cdot 6}{27} = \dfrac{135}{27} = 5$.

Επειδή το πλήθος των παρατηρήσεων είναι περιττό, η διάμεσος θα είναι η μεσαία παρατήρηση, δηλαδή η 14η. Επειδή $\nu_1 + \nu_2 = 10 + 3 = 13$, έπεται ότι η 14η παρατήρηση είναι $x_3 = 6$, άρα η διάμεσος είναι $\delta = x_3 = 6$.

γ. Έχουμε

$$s^2 = \frac{(3-5)^2 \cdot 10 + (5-5)^2 \cdot 3 + (6-5)^2 \cdot 8 + (7-5)^2 \cdot 6}{27} = \frac{72}{27} = \frac{8}{3}.$$

δ. Η τυπική απόκλιση είναι $s = \sqrt{s^2} = \sqrt{\dfrac{8}{3}}$, οπότε ο συντελεστής μεταβολής ισούται με:

$$CV = \frac{\sqrt{\frac{8}{3}}}{5}.$$

Υψώνουμε στο τετράγωνο και έχουμε $CV^2 = \dfrac{\frac{8}{3}}{25} = \dfrac{8}{75} > \dfrac{1}{100} = \left(\dfrac{1}{10}\right)^2$ άρα $CV > \dfrac{1}{10} = 10\%$ (αφού CV>0), το οποίο σημαίνει ότι το δείγμα δεν είναι ομοιογενές.

Άσκηση 5

Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα από μια έρευνα σ' ένα δείγμα 50 ατόμων.

Τιμές (x_i)	**Συχνότητα (ν_i)**
0	11
1	α
2	12
3	8
4	11
Σύνολο	

α) Να δείξετε ότι $\alpha = 8$.

β) Να συμπληρώσετε τον πίνακα με τις στήλες $f_i\%, N_i, F_i\%$.

γ) Να υπολογίσετε τη μέση τιμή $\bar{x}$ και τη διάμεσο δ του δείγματος.

δ) Να εξετάσετε αν το δείγμα είναι ομοιογενές. (Δίνεται : $\sqrt{2,08} \simeq 1,4$)

<u>Λύση</u>

α) Είναι $\nu_1 + \nu_2 + \nu_3 + \nu_4 + \nu_5 = 50 \Rightarrow 11 + \alpha + 12 + 8 + 11 = 50 \Leftrightarrow \alpha = 8$.

β) Για τις σχετικές συχνότητες $f_i\%$, έχουμε:

$f_1\% = \frac{\nu_1}{\nu} \cdot 100 = \frac{11}{50} \cdot 100 = 22$, $f_2\% = \frac{\nu_2}{\nu} \cdot 100 = \frac{8}{50} \cdot 100 = 16$,

$f_3\% = \frac{\nu_3}{\nu} \cdot 100 = \frac{12}{50} \cdot 100 = 24$, $f_4\% = \frac{\nu_4}{\nu} \cdot 100 = \frac{8}{50} \cdot 100 = 16$,

$f_5\% = \frac{\nu_5}{\nu} \cdot 100 = \frac{11}{50} \cdot 100 = 22$.

Για τις αθροιστικές συχνότητες N_i έχουμε:

$N_1 = \nu_1 = 11$, $N_2 = N_1 + \nu_2 = 19$, $N_3 = N_2 + \nu_3 = 31$,

$N_4 = N_3 + \nu_4 = 39$, $N_5 = \nu = 50$.

Για τις αθροιστικές σχετικές συχνότητες $F_i\%$ έχουμε:

$F_1\% = f_1\% = 22$, $F_2\% = F_1\% + f_2\% = 38\%$, $F_3\% = F_2\% + f_3\% = 62$,

$F_4\% = F_3\% + f_4\% = 78$, $F_5\% = 100$.

Έτσι προκύπτει ο παρακάτω πίνακας

Τιμές (x_i)	**Συχνότητα (ν_i)**	**Σχετική συχνότητα ($f_i\%$)**	**Αθροιστική συχνότητα (N_i)**	**Αθροιστική σχετική συχνότητα ($F_i\%$)**
0	11	22	11	22
1	8	16	19	38
2	12	24	31	62
3	8	16	39	78
4	11	22	50	100
Σύνολο	50	100		

γ) Για τον υπολογισμό της μέσης τιμής συμπληρώνουμε τον πίνακα (πίνακας 3) με τη στήλη $x_i \cdot \nu_i$ και έχουμε: $\bar{x} = \dfrac{\sum_{i=1}^{5} x_i \cdot \nu_i}{\nu} = \dfrac{100}{50} = 2$

Επειδή το μέγεθος του δείγματος είναι $\nu = 50$ άρτιος, η διάμεσος θα είναι ίση με το ημιάθροισμα των δύο μεσαίων παρατηρήσεων. Όπως διαπιστώνουμε από τη στήλη N_i (πίνακας 3), αυτές είναι $t_{25} = t_{26} = 2$. Άρα έχουμε $\delta = \dfrac{t_{25} + t_{26}}{2} = \dfrac{2+2}{2} = 2$.

δ) Πρώτα υπολογίζουμε τη διακύμανση s^2 και στη συνέχεια την τυπική απόκλιση s.

Έχουμε (πίνακας 3), $s^2 = \dfrac{\sum_{i=1}^{5} (x_i - \bar{x})^2 \cdot \nu_i}{\nu} = \dfrac{104}{50} = 2,08$ και $s = \sqrt{s^2} = \sqrt{2,08} \simeq 1,4$.

Ο συντελεστής μεταβολής του δείγματος είναι $CV = \dfrac{s}{\bar{x}} = \dfrac{1,4}{2} = 0,7 > 0,1$. Άρα το δείγμα δεν είναι ομοιογενές.

Πίνακας 3

Τιμές (x_i)	**Συχνότητα (ν_i)**	**Σχετική συχνότητα ($f_i\%$)**	**Αθροιστική συχνότητα (N_i)**	**Αθροιστική σχετική συχνότητα ($F_i\%$)**	$x_i \cdot \nu_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot \nu_i$
0	11	22	11	22	0	4	44
1	8	16	19	38	8	1	8
2	12	24	31	62	24	0	0
3	8	16	39	78	24	1	8
4	11	22	50	100	44	4	44
Σύνολο	50	100			100		104

Άσκηση 6

Στον παρακάτω πίνακα φαίνονται οι πόντοι που πέτυχε ένας παίκτης του μπάσκετ σε 30 αγώνες στους οποίους συμμετείχε την περασμένη αγωνιστική περίοδο.

1	15	13	8	10	5	12	7	7	8
14	2	9	13	4	12	6	11	14	15
15	14	2	11	12	10	10	7	11	13

α) Να ομαδοποιήσετε τα δεδομένα σε 5 κλάσεις ίσου πλάτους.

β) Να κατασκευάσετε πίνακα συχνοτήτων που να περιέχει επίσης τις στήλες $f_i\%, N_i, F_i\%$.

γ) Σε πόσους αγώνες ο παίκτης πέτυχε:
i. τουλάχιστον 7 πόντους
ii. λιγότερο από 13 πόντους
iii. το πολύ 8 πόντους;

δ) Να κατασκευάσετε το πολύγωνο $F_i\%$ και να υπολογίσετε τη διάμεσο του δείγματος.

<u>Λύση</u>

α) Το εύρος του δείγματος είναι $R = 15 - 1 = 14$ και ο αριθμός των κλάσεων είναι $\kappa = 5$. Άρα το πλάτος κάθε κλάσης είναι $c = \frac{R}{\kappa} = \frac{14}{5} = 2,8 \simeq 3$.
Στον παρακάτω πίνακα φαίνονται οι κλάσεις, οι κεντρικές τιμές καθώς και η συχνότητα κάθε κλάσης όπως προκύπτει από τη διαλογή.

β)

Κλάσεις $[\alpha,\beta)$	**Κεντρικές τιμές (**x_i**)**	**Διαλογή**	**Συχνότητα (**ν_i**)**	**Σχετική συχνότητα (**$f_i\%$**)**	**Αθροιστική συχνότητα (**N_i**)**	**Αθροιστική σχετική συχνότητα (**$F_i\%$**)**
$[1,4)$	2,5	|||	3	10	3	10
$[4,7)$	5,5	|||	3	10	6	20
$[7,10)$	8,5	~~||||~~ |	6	20	12	40
$[10,13)$	11,5	~~||||~~ ||||	9	30	21	70
$[13,16)$	14,5	~~||||~~ ||||	9	30	30	100
Σύνολο			30	100		

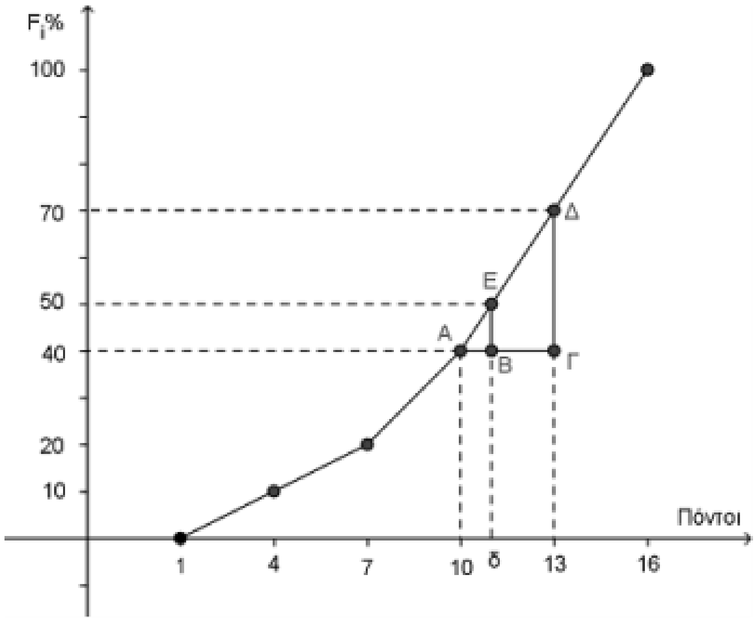
γ) Ο παίκτης πέτυχε:

i. τουλάχιστον 7 πόντους σε $\nu_3+\nu_4+\nu_5=6+9+9=24$ αγώνες

ii. λιγότερο από 13 πόντους σε $N_4=21$ αγώνες

iii. το πολύ 8 πόντους σε $N_2+\nu_3'=6+2=8$ αγώνες. Επειδή οι παρατηρήσεις θεωρούνται ομοιόμορφα κατανεμημένες μέσα στις κλάσεις θεωρούμε ότι το πλήθος των αγώνων στους οποίους ο παίκτης πέτυχε από 7 έως και 8 πόντους είναι $\nu_3'=\frac{1}{3}\nu_3=\frac{1}{3}\cdot 6=2$.

δ) Οι κορυφές του πολυγώνου αθροιστικών σχετικών συχνοτήτων $F_i\%$ είναι τα σημεία $(1,0),(4,10),(7,20),(10,40),(13,70),(16,100)$ και το πολύγωνο είναι:



Το σημείο Ε του πολυγώνου έχει τεταγμένη 50% και η τετμημένη του θα είναι ίση με τη διάμεσο δ. Τα ορθογώνια τρίγωνα ΑΒΕ και ΑΓΔ είναι όμοια, άρα οι πλευρές τους είναι ανάλογες. Έτσι έχουμε:

$$\frac{AB}{A\Gamma}=\frac{BE}{\Gamma\Delta}\Rightarrow\frac{\delta-10}{13-10}=\frac{50-40}{70-40}\Leftrightarrow\frac{\delta-10}{3}=\frac{10}{30}\Leftrightarrow$$

$$\delta-10=1\Leftrightarrow\delta=11.$$

Άσκηση 7

Ο παρακάτω πίνακας μας δίνει τις ημέρες άδειας που δικαιούνται 51 υπάλληλοι μιας επιχείρησης για ένα έτος εργασίας τους, ανάλογα με τα χρόνια υπηρεσίας τους.

Ημέρες άδειας $[\,,\,)$	**Κεντρική τιμή** x_i	**Συχνότητα** ν_i	$x_i \cdot \nu_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \nu_i$.
$[0,6)$		11				
$[6,12)$		20				
$[12,18)$		5				
$[18,24)$		5				
$[24,30)$		10				
Σύνολο		51				

α) Να μεταφέρετε τον παραπάνω πίνακα στο τετράδιό σας και να τον συμπληρώσετε.

β) Βρείτε τη μέση τιμή των ημερών αδείας των υπαλλήλων.

γ) Πόσοι υπάλληλοι της επιχείρησης δικαιούνται άδεια λιγότερο από δεκαοκτώ (18) ημέρες τον χρόνο;

δ) Βρείτε τη διακύμανση της παραπάνω κατανομής.

ε) Βρείτε την τυπική απόκλιση.

στ) Βρείτε τον συντελεστή μεταβολής.

<u>Λύση</u>

α)

Ημέρες άδειας $[,)$	**Κεντρική τιμή** x_i	**Συχνότητα** ν_i	$x_i \nu_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \nu_i$
$[0,6)$	3	11	33	-10	100	1100
$[6,12)$	9	20	180	-4	16	320
$[12,18)$	15	5	75	2	4	20
$[18,24)$	21	5	105	8	64	320
$[24,30)$	27	10	270	14	196	1960
Σύνολο		51	663			3720

β) Για τον υπολογισμό της μέσης τιμής από τον συμπληρωμένο πίνακα και τη στήλη $x_i \cdot \nu_i$ έχουμε: $\bar{x} = \dfrac{\sum_{i=1}^{5} x_i \nu_i}{\nu} = \dfrac{663}{51} = 13$

γ) Οι υπάλληλοι που δικαιούνται άδεια λιγότερο από 18 ημέρες ανήκουν στις 3 πρώτες κλάσεις, οπότε θα είναι 11+20+5=36 υπάλληλοι.

δ) Για τον υπολογισμό της διακύμανσης από τον συμπληρωμένο πίνακα και τη στήλη $(x_i - \bar{x})^2 \nu_i$ έχουμε:

$$s^2 = \frac{\sum_{i=1}^{5} \left(x_i - \bar{x}\right)^2 \cdot \nu_i}{\nu} = \frac{3720}{51} = 72,9$$

ε) Η τυπική απόκλιση είναι:

$$s = \sqrt{s^2} = \sqrt{72,9} \approx 8,5$$

στ) Ο συντελεστής μεταβολής είναι:

$$CV = \frac{s}{\bar{x}} = \frac{8,5}{13} \approx 0,654 = 65,4\%$$

Άσκηση 8

Α. Να γράψετε στο τετράδιό σας τον πίνακα των τιμών της μεταβλητής Χ σωστά συμπληρωμένο.

Τιμές Μεταβλητής x_i	**Συχνότητα** ν_i	**Σχετ. Συχνότητα** f_i	**Σχετ. Συχνότητα** $f_i\%$	**Αθροιστική Συχνότητα** N_i	$x_i \cdot \nu_i$	x_i^2	$x_i^2 \cdot \nu_i$
1	10				10	1	10
2				35		4	
3						9	
Σύνολο	$\nu = 50$	1	100				

Β. Να υπολογίσετε τη μέση τιμή και τη διάμεσο.

Γ. Να δείξετε ότι η διακύμανση είναι $s^2 = 0,49$. Δίνεται ότι: $s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{\kappa} x_i^2\nu_i - \frac{\left(\sum_{i=1}^{\kappa} x_i\nu_i\right)^2}{\nu}\right\}$

(Απολυτήριες εξετάσεις Γ΄ Τάξης 2000)

<u>Λύση</u>

Α.

Τιμές Μεταβλητής x_i	**Συχνότητα** ν_i	**Σχετ. Συχνότητα** f_i	**Σχετ. Συχνότητα** $f_i\%$	**Αθροιστική Συχνότητα** N_i	$x_i \cdot \nu_i$	x_i^2	$x_i^2 \cdot \nu_i$
1	10	0,2	20	10	10	1	10
2	25	0,5	50	35	50	4	100
3	15	0,3	30	50	45	9	135
Σύνολο	$\nu = 50$	1	100		105		245

Β. Για τον υπολογισμό της μέσης τιμής από τον συμπληρωμένο πίνακα και τη στήλη $x_i \cdot \nu_i$

έχουμε: $\bar{x} = \frac{\sum_{i=1}^{3} x_i \cdot \nu_i}{\nu} = \frac{105}{50} = 2,1$

Επειδή $\nu = 50$ άρτιος, οι δύο μεσαίες παρατηρήσεις θα είναι η 25η και η 26η παρατήρηση, που και οι δύο αντιστοιχούν στην τιμή 2. Άρα η διάμεσος είναι ο μέσος όρος (ημιάθροισμα) των δύο μεσαίων παρατηρήσεων άρα $\delta = \frac{2+2}{2} = 2$.

Γ. Είναι $\sum_{i=1}^{3} x_i^2 \nu_i = 245$ και $(\sum_{i=1}^{3} x_i \nu_i)^2 = 105^2$

Άρα $s^2 = \dfrac{1}{\nu}\left\{\sum_{i=1}^{3} x_i^2 \nu_i - \dfrac{\left(\sum_{i=1}^{3} x_i \nu_i\right)^2}{\nu}\right\} = \dfrac{1}{50}\left\{245 - \dfrac{105^2}{50}\right\} =$

$\dfrac{1}{50}\{245 - 220,5\} = \dfrac{24,5}{50} = 0,49$

Άσκηση 9

Έστω $\overline{x}$ η μέση τιμή των τριάντα παρατηρήσεων $t_1, t_2, \ldots, t_{30}$, ενός δείγματος. Αν οι είκοσι πρώτες παρατηρήσεις, έχουν μέση τιμή $\overline{x_1}$, ενώ οι υπόλοιπες έχουν μέση τιμή $\overline{x_2}$, τότε να αποδείξετε ότι ισχύει η σχέση:

$$\overline{x} = \frac{2 \cdot \overline{x_1} + \overline{x_2}}{3}$$

<u>Λύση</u>

Γνωρίζουμε ότι ισχύει: $\overline{x} = \frac{1}{\nu}\sum_{i=1}^{\nu} t_i = \frac{t_1 + t_2 + \ldots + t_\nu}{\nu}$

Οπότε, για $\nu = 30$ έχουμε:

$$\overline{x} = \frac{1}{30} \cdot \sum_{i=1}^{30} t_i = \frac{t_1 + t_2 + \ldots + t_{20} + t_{21} + \ldots + t_{30}}{30} =$$

$$\frac{t_1 + t_2 + \ldots + t_{20}}{30} + \frac{t_{21} + t_{22} + \ldots + t_{30}}{30} = \frac{\sum_{i=1}^{20} t_i}{30} + \frac{\sum_{i=21}^{30} t_i}{30} \quad \textbf{(1)}$$

Εξάλλου, ισχύουν: $\overline{x_1} = \frac{1}{20}\sum_{i=1}^{20} t_i \Leftrightarrow 20 \cdot \overline{x_1} = \sum_{i=1}^{20} t_i$ **(2)**

και $\overline{x_2} = \frac{1}{10}\sum_{i=21}^{30} t_1 \Leftrightarrow 10 \cdot \overline{x_2} = \sum_{i=21}^{30} t_i$ **(3)**.

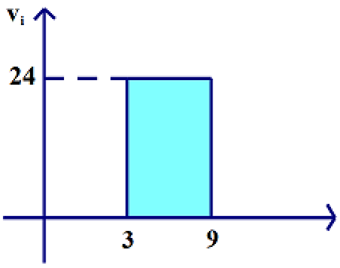
Επομένως η σχέση **(1)**, λόγω των **(2)** και **(3)**, γίνεται:

$$\overline{x} = \frac{1}{30} \cdot 20 \cdot \overline{x_1} + \frac{1}{30} \cdot 10 \cdot \overline{x_2} = \frac{2 \cdot \overline{x_1} + \overline{x_2}}{3}$$

ΘΕΜΑ Γ

Άσκηση 1

Στο ιστόγραμμα συχνοτήτων κλάσεων ίσου πλάτους ενός δείγματος μεγέθους 200 δίνεται μόνο το ορθογώνιο της 1ης κλάσης $[3,9)$.



Να βρείτε:

α) Το πλήθος των παρατηρήσεων της κλάσης $[3,9)$ που είναι μικρότερες του 5.

β) Το ποσοστό των παρατηρήσεων της κλάσης $[3,9)$ οι οποίες έχουν τιμή μεγαλύτερη ή ίση του 4 και μικρότερη του 8.

γ) Να βρείτε την τιμή του πραγματικού αριθμού α ώστε στο διάστημα $[3,\alpha)$

i. να ανήκουν 16 παρατηρήσεις.

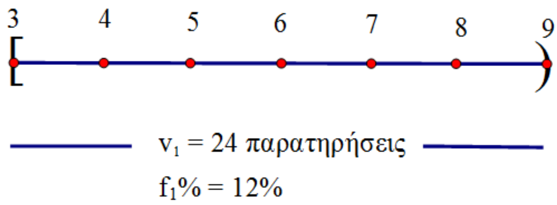
ii. να ανήκει το 3% των παρατηρήσεων του δείγματος

Λύση

α) Από το ιστόγραμμα συχνοτήτων προκύπτει ότι η συχνότητα της 1ης κλάσης $[3,9)$ είναι $\nu = 24$ όσο και το ύψος του ορθογωνίου στο ιστόγραμμα συχνοτήτων.

Επιπλέον η σχετική συχνότητα της κλάσης αυτής είναι $f_1 = \frac{\nu_1}{\nu} = \frac{24}{200} = 0{,}12$, οπότε $f_1 = 12\%$.

Οι 24 παρατηρήσεις κατανέμονται ομοιόμορφα στην κλάση $[3,9)$ άρα από 3 έως 4 έχουμε 4 παρατηρήσεις, από 4 έως 5 άλλες 4 κ.τ.λ.



Έτσι οι παρατηρήσεις που είναι μικρότερες από 5 είναι 8.
Στο ίδιο αποτέλεσμα μπορούμε να καταλήξουμε και με άλλο τρόπο. Αφού σε κάθε κλάση οι παρατηρήσεις κατανέμονται ομοιόμορφα τα ποσά πλάτος κλάση και συχνότητα κλάσης είναι ανάλογα, οπότε:

$$\frac{5-3}{9-3} = \frac{x}{24} \Leftrightarrow \frac{2}{6} = \frac{x}{24} \Leftrightarrow x = 8$$

β) Επειδή και το ποσοστό % όταν πρόκειται για συχνότητα $f_i\%$ κατανέμεται ομοιόμορφα σε μια κλάση αν θεωρήσουμε ότι οι παρατηρήσεις που είναι στο διάστημα από 4 έως 6 είναι ν_κ και έχουν σχετική συχνότητα $f_\kappa\%$ τότε

$$\frac{9-3}{8-4}=\frac{f_1\%}{f_\kappa\%} \Leftrightarrow \frac{6}{4}=\frac{12}{f_\kappa\%} \Leftrightarrow f_\kappa\%=8$$

Άρα το ποσοστό των παρατηρήσεων της κλάσης $[3,9)$ οι οποίες έχουν τιμή μεγαλύτερη ή ίση του 4 και μικρότερη του 8 είναι 8%.

γ)

i. Έχουμε $\frac{\alpha-3}{9-3}=\frac{16}{24} \Leftrightarrow \frac{\alpha-3}{6}=\frac{16}{24} \Leftrightarrow \alpha-3=\frac{6\cdot16}{24}=\frac{96}{24} \Leftrightarrow \alpha-3=4 \Leftrightarrow \alpha=7$

Επομένως στο διάστημα $[3,7)$ ανήκουν 16 παρατηρήσεις.

ii. Έχουμε $\frac{\alpha-3}{9-3}=\frac{3}{12} \Leftrightarrow \alpha-3=\frac{6\cdot3}{12} \Leftrightarrow \alpha-3=\frac{18}{12} \Leftrightarrow \alpha-3=1,5 \Leftrightarrow \alpha=4,5$

Επομένως στο διάστημα $[3,4,5)$ ανήκει το 3% των παρατηρήσεων του δείγματος.

Άσκηση 2

Οι καθαρές μηνιαίες αποδοχές των εργαζομένων σε μια επιχείρηση χωρισμένες σε πέντε κλάσεις ίσου πλάτους, είναι από 700 έως 1200 ευρώ. Αν γνωρίζουμε ότι:

i. Το πολύγωνο συχνοτήτων του δείγματος έχει εμβαδόν 125.
ii. Οι εργαζόμενοι που έχουν καθαρές μηνιαίες αποδοχές τουλάχιστον 900 ευρώ είναι 80.
iii. Η γωνία του κυκλικού τομέα στο κυκλικό διάγραμμα που αντιστοιχεί στην κλάση $[800,900)$ είναι 72°.
iv. Το ύψος του ορθογωνίου της κλάσης $[900,1000)$ στο ιστόγραμμα σχετικών αθροιστικών συχνοτήτων είναι 0,64.
v. Οι εργαζόμενοι με καθαρές μηνιαίες αποδοχές από 1000 έως 1100 ευρώ είναι διπλάσιοι από αυτούς που έχουν καθαρές μηνιαίες αποδοχές από 1100 έως 1200 ευρώ.

α) Να κάνετε πίνακα συχνοτήτων (απόλυτων, σχετικών και αθροιστικών)
β) Να υπολογίσετε τη μέση τιμή και τη διάμεσο.

<u>Λύση</u>

α) Αφού το εύρος του δείγματος είναι $R = 1200 - 700 = 500$ τότε το πλάτος των κλάσεων είναι $c = 500/5 = 100$. Αν θεωρήσουμε ως κατώτερο όριο της $1^{ης}$ κλάσης το 700 τότε οι κλάσεις διαμορφώνονται ως εξής:
$[700,800), [800,900), [900,1000), [1000,1100), [1100,1200)$

- Στο ιστόγραμμα συχνοτήτων κάθε ορθογώνιο του έχει εμβαδόν ίσο με το ύψος του αφού η βάση του έχει μήκος ίσο με το πλάτος των κλάσεων c που θεωρείται μονάδα μέτρησης. Άρα το ιστόγραμμα συχνοτήτων έχει εμβαδόν ίσο με το άθροισμα των συχνοτήτων του δείγματος, ίσο δηλαδή με το μέγεθος του δείγματος. Επομένως αφού το πολύγωνο συχνοτήτων του δείγματος έχει εμβαδόν 125, τότε $\boxed{\nu = 125}$ **(1)**

- Αφού οι εργαζόμενοι που έχουν καθαρές μηνιαίες αποδοχές τουλάχιστον 900 ευρώ, είναι 80, τότε $\boxed{\nu_3 + \nu_4 + \nu_5 = 80}$ **(2)**

- Αφού η γωνία του κυκλικού τομέα στο κυκλικό διάγραμμα που αντιστοιχεί στην $2^{η}$ κλάση $[800,900)$ είναι 72^0, τότε $\alpha_2 = 72^{\circ}$ αλλά $\alpha_2 = f_2 \cdot 360^{\circ}$ άρα $f_2 = 72^0/360^0 = 0,20$, $\boxed{f_2 = 0,20}$ **(3)**

- Επειδή το ύψος του ορθογωνίου της κλάσης $[900,1000)$ στο ιστόγραμμα σχετικών αθροιστικών συχνοτήτων είναι 0,64 έχουμε $\boxed{F_3 = 0,64}$ **(4)**

- Αφού οι εργαζόμενοι με καθαρές μηνιαίες αποδοχές από 1000 έως 1100 ευρώ είναι διπλάσιοι από αυτούς που έχουν καθαρές μηνιαίες αποδοχές από 1100 έως 1200 ευρώ, τότε $\boxed{\nu_4 = 2\nu_5}$ **(5)**

- Τέλος γνωρίζουμε ότι $\nu_1+\nu_2+\nu_3+\nu_4+\nu_5=\nu$ **(6)**

  - Από τις σχέσεις (1) και (3) έχουμε: $\nu_2=f_2\cdot\nu=0,20\cdot125$, $\boxed{\nu_2=25}$

  - Από τις σχέσεις (2) και (6) έχουμε: $\nu_1+25+80=125$, $\boxed{\nu_1=20}$. Οπότε από τον τύπο $f_1=\nu_1/\nu$ έχουμε: $f_1=20/125=0,16$, άρα $\boxed{f_1\%=16}$ και $\boxed{F_1\%=16}$

  - Από τη σχέση (3) έχουμε $f_2\%=20$, οπότε από τον τύπο $F_2\%=F_1\%+f_2\%$ έχουμε:

    $F_2\%=16\%+20\%$, $\boxed{F_2\%=36\%}$.

  - Από τη σχέση (4) και από τον τύπο $F_3\%=F_2\%+f_3\%$ έχουμε: $f_3\%=64-36$, $\boxed{f_3\%=28}$.

    Από τον τύπο $\nu_3=f_3\cdot\nu$ έχουμε: $\nu_3=0,28\cdot125$, $\boxed{\nu_3=35}$ **(7)**

  - Από τις σχέσεις (2), (5) και (7) έχουμε:
    $35+2\nu_5+\nu_5=80 \Leftrightarrow 3\nu_5=45 \Leftrightarrow \nu_5=15$.
    Επομένως $\boxed{\nu_4=30}$.

Ο πίνακας συμπληρωμένος είναι ο εξής

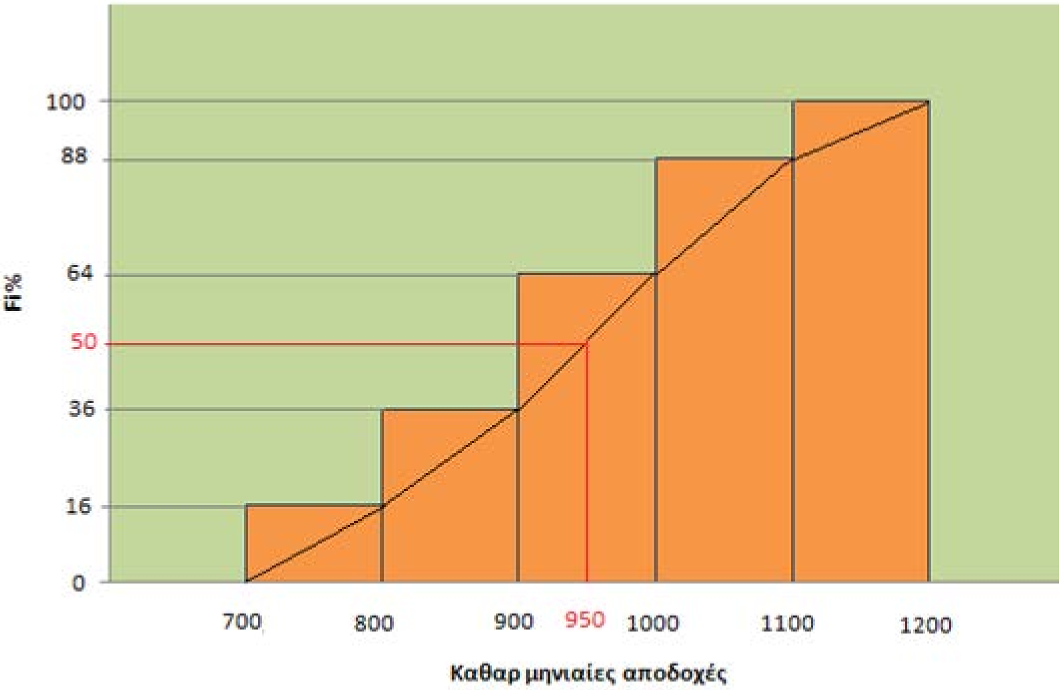
Κλάσεις	x_i	v_i	N_i	f_i	$f_i\%$	$F_i\%$
[700, 800)	750	20	20	0, 16	16	16
[800, 900)	850	25	45	0,20	20	36
[900, 1000)	950	35	80	0,28	28	64
[1000, 1100)	1050	30	110	0,24	24	88
[1100, 1200)	1150	15	125	0,12	12	100
Σύνολο		125		1	100	

Β) Η μέση τιμή είναι:

$$\bar{x}=\frac{\sum_{i=1}^{5}x_i\nu_i}{\nu}=\frac{750\cdot20+850\cdot25+950\cdot35+1050\cdot30+1150\cdot15}{125}=946$$

Άρα οι μέσες καθαρές μηνιαίες αποδοχές είναι 946 ευρώ.
Η διάμεσος μιας κατανομής είναι η μοναδική τιμή που χωρίζει τις παρατηρήσεις σε δύο ισοπληθείς ομάδες. Επειδή το πλήθος των παρατηρήσεων είναι μονός αριθμός η διάμεσος του δείγματος των 125 παρατηρήσεων είναι η μεσαία παρατήρηση δηλαδή η $63^η$. Από τον πίνακα

παρατηρούμε ότι η $63^{η}$ παρατήρηση βρίσκεται στην κλάση [900, 1000). Άρα η διάμεσος είναι 950 ευρώ.
Στο ίδιο αποτέλεσμα θα καταλήγαμε αν κατασκευάζαμε το πολύγωνο των σχετικών αθροιστικών συχνοτήτων και από το σημείο 50% του άξονα y΄y φέρναμε παράλληλη μέχρι το πολύγωνο των σχετ. αθροιστικών συχνοτήτων και από εκεί κάθετη στον άξονα x΄x. Τότε επειδή το 50 είναι στο μέσον του διαστήματος [36,64] η κάθετη στον άξονα x΄x θα κόβει το διάστημα [900,1000] στο μέσον, άρα η διάμεσος είναι 950 ευρώ.



Άσκηση 3

Ο μέσος μισθός 50 υπαλλήλων σε έναν οργανισμό είναι $\overline{x}=1500$€. Η τυπική απόκλιση των μισθών είναι $s=120$€.

α. Αν οι 20 εργαζόμενοι με τον υψηλότερο μισθό έχουν μέσο μισθό 1800€, να βρείτε το μέσο μισθό των υπολοίπων υπαλλήλων.

β. Αν οι μισθοί όλων των υπαλλήλων αυξηθούν κατά 5%, να βρείτε τη μέση τιμή, την τυπική απόκλιση και τον συντελεστή μεταβολής των νέων μισθών.

γ. Ομοίως αν οι αρχικοί μισθοί όλων των υπαλλήλων αυξηθούν κατά 50€, να βρείτε τη μέση τιμή, την τυπική απόκλιση και τον συντελεστή μεταβολής των νέων μισθών.

<u>Λύση</u>

α. Αν οι 20 εργαζόμενοι με τον υψηλότερο μισθό έχουν μισθούς $t_{31},t_{32},\ldots,t_{50}$ και οι μισθοί των υπολοίπων είναι $t_1,t_2,\ldots,t_{30}$, τότε έχουμε:

$$\frac{t_{31}+t_{32}+\ldots+t_{50}}{20}=1800 \Leftrightarrow t_{31}+t_{32}+\ldots+t_{50}=36000 .$$

Επίσης ο μέσος μισθός των 50 υπαλλήλων είναι $\overline{x}=1500$€, άρα

$$\frac{\left(t_1+t_2+\ldots+t_{30}\right)+\left(t_{31}+t_{32}+\ldots+t_{50}\right)}{50}=1500 \Leftrightarrow$$

$$\left(t_1+t_2+\ldots+t_{30}\right)+36000=75000 \Leftrightarrow$$

$$t_1+t_2+\ldots+t_{30}=39000 \Leftrightarrow \frac{t_1+t_2+\ldots+t_{30}}{30}=1300$$. Άρα ο μέσος μισθός των 30 χαμηλόμισθων υπαλλήλων είναι 1300€.

β. Κάθε μισθός γίνεται $t_i{}'=t_i+\frac{5}{100}\cdot t_i=1,05\cdot t_i$. Οπότε σύμφωνα με γνωστή εφαρμογή του βιβλίου η νέα μέση τιμή $\overline{x}'$ θα είναι $\overline{x}'=1,05\cdot\overline{x}=1,05\cdot 1500=1575$€ και η νέα τυπική απόκλιση s' είναι $s'=1,05\cdot s=1,05\cdot 120=126$€.

Ο συντελεστής μεταβολής είναι $CV'=\frac{s'}{\overline{x}'}=\frac{1,05\cdot 120}{1,05\cdot 1500}=\frac{120}{1500}=0,08=8\%$.

γ. Κάθε μισθός γίνεται $t_i{}''=t_i+50$. Οπότε σύμφωνα με γνωστή εφαρμογή του βιβλίου η νέα μέση τιμή $\overline{x}'$ θα είναι $\overline{x}''=\overline{x}+50=1550$€ και η νέα τυπική απόκλιση s'' είναι $s''=s=120$€.

Ο συντελεστής μεταβολής είναι $CV''=\frac{s''}{\overline{x}''}=\frac{120}{1550}=0,0774=7,74\%$.

Άσκηση 4

Οι τιμές μιας μεταβλητής Χ είναι ομαδοποιημένες σε κλάσεις ίσου πλάτους c, όπως δίνονται στον παρακάτω ελλιπή πίνακα:

Κλάσεις [,)	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
$[3\ ,\ )$		0,1		
$[\ ,\ )$				
$[\ ,12)$				
$[\ ,\ )$				
Σύνολο				

Γνωρίζουμε επίσης ότι οι συχνότητες f_2, f_3, f_4 είναι ανάλογες προς τους αριθμούς 3, 2, 4 αντίστοιχα.
α. Να δείξετε ότι $c=3$ και να συμπληρώσετε τον παραπάνω πίνακα.
β. Να βρείτε τη μέση τιμή και τη διασπορά των τιμών του δείγματος.
γ. Να σχεδιάσετε το ιστόγραμμα σχετικών συχνοτήτων.
δ. Να βρείτε το συντελεστή μεταβολής των τιμών του δείγματος και να εξετάσετε αν το δείγμα είναι ομοιογενές.

Λύση

α. Η πρώτη κλάση είναι η $[3,3+c)$, η δεύτερη $[3+c,3+2\cdot c)$ και η τρίτη $[3+2\cdot c,3+3\cdot c)$.
Άρα $3+3\cdot c=12 \Leftrightarrow c=3$.

Επίσης

$$\frac{f_2}{3}=\frac{f_3}{2}=\frac{f_4}{4} \Leftrightarrow \frac{f_2}{3}=\frac{f_3}{2}=\frac{f_4}{4}=\frac{f_2+f_3+f_4}{3+2+4}\overset{\sum_{i=1}^{4} f_i=1}{=}\frac{1-f_1}{9}=$$

$\frac{0,9}{9}=0,1$, άρα
$f_2=3\cdot 0,1=0,3$
$f_3=2\cdot 0,1=0,2$ και
$f_4=4\cdot 0,1=0,4$.

Έτσι συμπληρώνουμε τον πίνακα

Κλάσεις [,)	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
$[3\ ,6)$	4,5	0,1	0,45	2,025
$[6\ ,9)$	7,5	0,3	2,25	16,875
$[9\ ,12)$	10,5	0,2	2,1	22,05
$[12\ ,15)$	13,5	0,4	5,4	72,9
Σύνολο		1	10,2	113,85

β. Η μέση τιμή είναι $\overline{x}=\sum_{i=1}^{4} x_i \cdot f_i = 10,2$

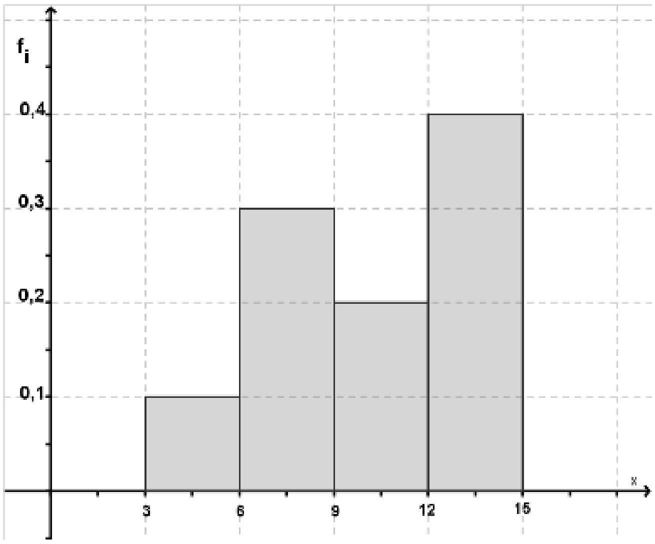
και η διασπορά υπολογίζεται από τον τύπο

$$s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{k} x_i{}^2 \cdot \nu_i - \frac{\left(\sum_{i=1}^{k} x_i \cdot \nu_i\right)^2}{\nu}\right\} = \sum_{i=1}^{k} x_i{}^2 \cdot \frac{\nu_i}{\nu} - \frac{\left(\sum_{i=1}^{k} x_i \cdot \nu_i\right)^2}{\nu^2} =$$

$$\sum_{i=1}^{k} x_i{}^2 \cdot f_i - \left(\overline{x}\right)^2$$

οπότε $s^2 = 113,85 - 10,2^2 = 9,81$.

γ. Το ιστόγραμμα σχετικών συχνοτήτων είναι το παρακάτω



δ. Ο συντελεστής μεταβολής είναι

$$CV = \frac{s}{\overline{x}} = \frac{\sqrt{9,81}}{10,2} .$$

Υψώνουμε στο τετράγωνο και έχουμε

$$CV^2 = \frac{9,81}{10,2^2} = \frac{9,81}{104,04} = 0,0943 > 0,01 = \frac{1}{100} ,$$

άρα $CV^2 > \left(\frac{1}{10}\right)^2 \overset{CV>0}{\Leftrightarrow} CV > 10\%$, επομένως το δείγμα δεν είναι ομοιογενές.

Άσκηση 5

Σε δύο τμήματα Γ1, Γ2 της Γ΄ τάξης ενός Λυκείου ο μέσος όρος της βαθμολογίας στο Α΄ τετράμηνο ήταν $\overline{x}_1 = \overline{x}_2 = 12$ με τυπική απόκλιση $s_1 = s_2 = 2$. Στο 2^{o} τετράμηνο όλοι οι μαθητές του Γ1 αύξησαν τη βαθμολογία τους κατά 1 μονάδα, ενώ οι μαθητές του Γ2 αύξησαν τη βαθμολογία τους κατά 10%.

α) Σε ποιο τμήμα η βαθμολογία παρουσιάζει μεγαλύτερη ομοιογένεια, μετά τις αυξήσεις του Β΄ τετραμήνου;

β) Να βρεθεί η μικρότερη τιμή της θετικής σταθεράς c που πρέπει να προστεθεί στις βαθμολογίες των μαθητών του Γ2 μετά το τέλος του Β΄ τετραμήνου, ώστε το δείγμα της βαθμολογίας τους να γίνει ομοιογενές.

γ) Αν οι βαθμολογίες των μαθητών του Γ1 στο Β΄ τετράμηνο αποτελούν κανονική κατανομή, να βρεθεί το ποσοστό των μαθητών που είχαν βαθμολογία από 11 έως 19.

δ) Να υπολογίσετε τη μέση τιμή των τετραγώνων των βαθμολογιών των μαθητών του Γ1 στο Α΄ τετράμηνο.

(Δίνεται: $s^2 = \dfrac{1}{\nu}\left[\sum_{i=1}^{\nu} x_i^2 - \dfrac{\left(\sum_{i=1}^{\nu} x_i\right)^2}{\nu}\right]$)

Λύση

α) Έστω x_i οι βαθμολογίες των μαθητών του Γ1 στο Α΄ τετράμηνο για τις οποίες ισχύει $\overline{x}_1 = 12$ και $s_1 = 2$. Οι βαθμολογίες αυτών στο Β΄ τετράμηνο, θα είναι $y_i = x_i + 1$, με $\overline{y} = \overline{x_1} + 1 = 12 + 1 = 13$ και $s_y = s_1 = 2$.

Άρα ο συντελεστής μεταβολής των βαθμολογιών του Γ1 στο Β΄ τετράμηνο είναι $CV_1 = \dfrac{s_y}{\overline{y}} = \dfrac{2}{13}$.

Έστω t_i οι βαθμολογίες των μαθητών του Γ2 στο Α΄ τετράμηνο για τις οποίες ισχύει $\overline{x}_2 = 12$ και $s_2 = 2$. Οι βαθμολογίες αυτών στο Β΄ τετράμηνο, θα είναι $\phi_i = t_i + 0{,}1 \cdot t_i = 1{,}1 \cdot t_i$, με $\overline{\phi} = 1{,}1 \cdot \overline{x_2} = 1{,}1 \cdot 12$ και $s_\phi = 1{,}1 \cdot s_2 = 1{,}1 \cdot 2$.

Άρα ο συντελεστής μεταβολής των βαθμολογιών του Γ2 στο Β΄ τετράμηνο είναι $CV_2 = \dfrac{s_\phi}{\overline{\phi}} = \dfrac{1{,}1 \cdot 2}{1{,}1 \cdot 12} = \dfrac{2}{12}$.

Ισχύει $\dfrac{2}{13} < \dfrac{2}{12}$, άρα $CV_1 < CV_2$. Οπότε η βαθμολογία του Β΄ τετραμήνου στο Γ1 παρουσιάζει μεγαλύτερη ομοιογένεια σε σχέση με τη βαθμολογία στο Γ2.

β) Ισχύει $CV_2=\frac{2}{12}=\frac{1}{6}>\frac{1}{10}$, άρα το δείγμα των βαθμολογιών του Γ2 στο Β΄ τετράμηνο δεν είναι ομοιογενές. Αν στις βαθμολογίες των μαθητών του Γ2 στο Β΄ τετράμηνο προσθέσουμε τη σταθερά $c>0$, τότε γίνονται $\omega_i=\phi_i+c$ και θα έχουν μέση τιμή $\overline{\omega}=\overline{\phi}+c=13,2+c$ και τυπική απόκλιση $s_\omega=s_\phi=2,2$. Αφού το δείγμα των βαθμολογιών ω_i είναι ομοιογενές, ισχύει $CV_\omega\le\frac{1}{10}\Leftrightarrow\frac{s_\omega}{\overline{\omega}}\le\frac{1}{10}\Leftrightarrow\frac{2,2}{13,2+c}\le\frac{1}{10}\Leftrightarrow 13,2+c\ge 22\Leftrightarrow c\ge 8,8$. Άρα η μικρότερη τιμή της σταθεράς c είναι 8,8.

γ) Οι βαθμολογίες των μαθητών του Γ1 στο Β΄ τετράμηνο αποτελούν κανονική κατανομή με $\overline{y}=13$ και $s_y=2$. Το ποσοστό των μαθητών του Γ1 που είχαν βαθμολογία από 11 έως 19 είναι το ίδιο με το ποσοστό των παρατηρήσεων που βρίσκονται στο διάστημα $(\overline{y}-s,\overline{y}+3s)$ μιας κανονικής κατανομής, αφού είναι $\overline{y}-s=13-2=11$ και $\overline{y}+3s=13+6=19$. Το ποσοστό αυτών των μαθητών είναι $\frac{68\%}{2}+\frac{99,7\%}{2}=83,85\%$.

δ) Έστω $x_i, i=1,2,...,\nu$ οι βαθμολογίες των μαθητών του Γ1 στο Α΄ τετράμηνο για τις οποίες ισχύει $\overline{x_1}=12$ και $s_1=2$. Τα τετράγωνα των βαθμολογιών δίνονται από τον τύπο $\tau_i={x_i}^2, i=1,2,...,\nu$.

Η μέση τιμή αυτών των βαθμολογιών είναι $\overline{\tau}=\frac{\sum_{i=1}^{\nu}{x_i}^2}{\nu}$. Από τον τύπο που δίνεται έχουμε:

$$ {s_1}^2=\frac{\sum_{i=1}^{\nu}{x_i}^2}{\nu}-\frac{\left(\sum_{i=1}^{\nu}x_i\right)^2}{\nu^2}\Leftrightarrow {s_1}^2=\overline{\tau}-\left(\frac{\sum_{i=1}^{\nu}x_i}{\nu}\right)^2\Leftrightarrow $$

$$ {s_1}^2=\overline{\tau}-\overline{x_1}^2\Leftrightarrow\overline{\tau}={s_1}^2+\overline{x_1}^2. $$

Άρα $\overline{\tau}=4+144=148$.

Άσκηση 6

Δίνεται η συνάρτηση $f(x)=x^3-\frac{11}{2}x^2-4x+41, x\in\mathbb{R}$ και μια ποσοστική μεταβλητή Χ ως προς την οποία εξετάσαμε ένα δείγμα ν παρατηρήσεων $x_1,x_2,...,x_\nu$ οι οποίες παίρνουν μόνο θετικές τιμές. Αν η εφαπτομένη της γραφικής παράστασης της f στο σημείο της $A\left(\overline{x},s^2\right)$ είναι παράλληλη στον άξονα $x'x$, τότε:

α) Να δείξετε ότι $\overline{x}=4$ και $s=1$.

β) Αν το δείγμα των παρατηρήσεων ακολουθεί κανονική κατανομή και γνωρίζουμε ότι 3 παρατηρήσεις έχουν τιμή μικρότερη του 1, τότε να υπολογίσετε:

i. το μέγεθος του δείγματος.
ii. το πλήθος των παρατηρήσεων που έχουν τιμή στο διάστημα $(2,3)$.

<u>Λύση</u>

α) Αρχικά βρίσκουμε την παράγωγο της f που είναι $f'(x)=3x^2-11x-4, x\in\mathbb{R}$.
Επειδή η εφαπτομένη της γραφικής παράστασης της f στο σημείο της $A\left(\overline{x},s^2\right)$ είναι παράλληλη στον άξονα $x'x$, θα ισχύει $f'\left(\overline{x}\right)=0\Leftrightarrow 3\overline{x}^2-11\overline{x}-4=0$ (1).
Η εξίσωση (1) έχει διακρίνουσα $\Delta=121+48=169$ και ρίζες $\overline{x}=\frac{11\pm13}{6}$, δηλαδή $\overline{x}=4$ ή $\overline{x}=-\frac{1}{3}$.
Όμως οι παρατηρήσεις $x_1,x_2,...,x_\nu$ παίρνουν μόνο θετικές τιμές, άρα θα είναι και $\overline{x}>0$. Έτσι προκύπτει ότι $\overline{x}=4$.
Επειδή το σημείο $A\left(\overline{x},s^2\right)$ ανήκει στη γραφική παράσταση της f, ισχύει $f\left(\overline{x}\right)=s^2$. Άρα $s^2=f\left(4\right)=64-88-16+41=1\Rightarrow s=1$.

β) Στο διπλανό άξονα βλέπουμε την αντιστοιχία των άκρων των διαστημάτων της κανονικής κατανομής με τις τιμές τους στο συγκεκριμένο δείγμα, για το οποίο έχουμε $\overline{x}=4$ και $s=1$.
Έτσι είναι:

1	2	3	4	5	6	7
$\overline{x}-3s$	$\overline{x}-2s$	$\overline{x}-s$	$\overline{x}$	$\overline{x}+s$	$\overline{x}+2s$	$\overline{x}+3s$

i. Στην κανονική κατανομή το 99,7% των παρατηρήσεων βρίσκονται στο διάστημα $\left(\overline{x}-3s,\overline{x}+3s\right)$. Άρα το 0,3% είναι μικρότερες του $\overline{x}-3s$ ή μεγαλύτερες του $\overline{x}+3s$. Συγκεκριμένα το 0,15% έχουν τιμή μικρότερη του $\overline{x}-3s$. Εδώ γνωρίζουμε ότι 3 παρατηρήσεις έχουν τιμή μικρότερη από 1, δηλαδή μικρότερη από $\overline{x}-3s$.

Άρα έχουμε $\frac{0{,}15}{100}=\frac{3}{\nu}\Leftrightarrow 0{,}15\nu=300\Leftrightarrow \nu=2000$.

ii. Στο διάστημα $\left(2,3\right)$, δηλαδή στο $\left(\overline{x}-2s,\overline{x}-s\right)$ βρίσκεται το $\frac{95\%}{2}-\frac{68\%}{2}=13{,}5\%$ των παρατηρήσεων. Άρα το πλήθος των παρατηρήσεων που βρίσκονται σ’ αυτό το διάστημα είναι $13{,}5\%\cdot\nu=\frac{13{,}5}{100}\cdot 2000=270$.

Άσκηση 7

Οι ώρες παρακολούθησης τηλεοπτικών προγραμμάτων από 50 άτομα σε διάστημα μιας εβδομάδας αναγράφονται στον παρακάτω (ελλιπή) πίνακα:

Ώρες παρακολούθησης (x_i)	**Συχνότητα (ν_i)**	**($x_i \nu_i$)**	**($x_i^2 \nu_i$)**
0			
1	18		
2	12		
3			
4	8		
Σύνολο	50		

Στο κυκλικό διάγραμμα συχνοτήτων του παραπάνω πίνακα δίνεται ότι η γωνία του κυκλικού τομέα που αντιστοιχεί στην παρατήρηση $x_1 = 0$ ώρες, είναι $\alpha_1 = 36^0$.

Α) Να μεταφέρετε στο τετράδιό σας τον παραπάνω πίνακα και να τον συμπληρώσετε.

Β) Γνωρίζοντας ότι για την διακύμανση ισχύει ο τύπος $s^2 = \frac{1}{\nu}\left\{ \sum_{i=1}^{\kappa} x_i^2 \nu_i - \frac{\left(\sum_{i=1}^{\kappa} x_i \nu_i \right)^2}{\nu} \right\}$, να υπολογίσετε την τυπική απόκλιση.

Γ) Υπολογίστε τον συντελεστή μεταβολής. Είναι το δείγμα ομοιογενές;

Λύση

Α) Από τον τύπο $\alpha_i = \frac{\nu_i}{\nu} \cdot 360^0$ υπολογίζουμε $\alpha_1 = \frac{\nu_1}{\nu} \cdot 360^0 \Leftrightarrow 36^0 = \frac{\nu_1}{50} \cdot 360^0 \Leftrightarrow \nu_1 = 5$.
Έχουμε
$\nu_1 + \nu_2 + \nu_3 + \nu_4 + \nu_5 = \nu \Leftrightarrow 5 + 18 + 12 + \nu_4 + 8 = 50 \Leftrightarrow \nu_4 = 7$.

Ώρες παρακολούθησης (x_i)	Συχνότητα (ν_i)	($x_i\nu_i$)	($x_i^2\nu_i$)
0	5	0	0
1	18	18	18
2	12	24	48
3	7	21	63
4	8	32	128
Σύνολο	50	95	257

Β) Πρώτα υπολογίζουμε τη διακύμανση s^2 και στη συνέχεια την τυπική απόκλιση s.

Έχουμε (πίνακας), $s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{\kappa} x_i^2\nu_i - \frac{\left(\sum_{i=1}^{\kappa} x_i\nu_i\right)^2}{\nu}\right\} =$

$$\frac{1}{50}\left\{257 - \frac{95^2}{50}\right\} = \frac{1}{50}\{257 - 180,5\} = 1,53$$

και $s = \sqrt{s^2} = \sqrt{1,53} \simeq 1,24$

Γ) Για τον υπολογισμό της μέσης τιμής από τον συμπληρωμένο πίνακα και τη στήλη $x_i \cdot \nu_i$

έχουμε: $\bar{x} = \frac{\sum_{i=1}^{5} x_i \cdot \nu_i}{\nu} = \frac{95}{50} = 1,9$.

Ο συντελεστής μεταβολής του δείγματος είναι:

$$CV = \frac{s}{\bar{x}} = \frac{1,24}{1,9} = 0,653 > 0,1$$

Άρα το δείγμα δεν είναι ομοιογενές.

Άσκηση 8

Α. Δίνεται η συνάρτηση f με τύπο $f(x)=x^2\ln x+x+1$. Να βρεθεί η εξίσωση της εφαπτόμενης (ε), της γραφικής παράστασης της συνάρτησης, f στο σημείο της $A(1,f(1))$.
Αν $M(x_i,y_i)$, είναι 50 σημεία της εφαπτομένης της καμπύλης της f, με μέση τιμή των τετμημένων τους, $\bar{x}=15$, να βρεθεί η μέση τιμή $\bar{y}$ των τεταγμένων των σημείων αυτών.

Β. Έστω ένα δείγμα ν παρατηρήσεων $x_1,x_2,...,x_\nu$ με μέση τιμή $\bar{x}$. Αν η συνάρτηση f, με:

$$f(x)=(x+x_1)^2+(x+x_2)^2+...+(x+x_\nu)^2, \qquad x\in\mathbb{R}$$

τότε να αποδείξετε ότι η f παρουσιάζει ελάχιστη τιμή, όταν $x=-\bar{x}$.

Γ. Εκατό μαθητές ενός Γενικού Λυκείου των Αθηνών, ερωτήθηκαν για το πόσες ώρες μελετούν εβδομαδιαίως. Όμως το 20% των μαθητών δήλωσε 5 ώρες λιγότερες από τις πραγματικές, ενώ το 50% των μαθητών, δήλωσε 6 ώρες περισσότερες. Από τις δηλώσεις των μαθητών, προέκυψε ότι ο μέσος χρόνος μελέτης είναι $\bar{x}$ ώρες. Αν $\bar{y}$, είναι ο πραγματικός μέσος των ωρών της εβδομαδιαίας μελέτης των μαθητών αυτών, τότε να αποδείξετε ότι ισχύει η σχέση:

$$\bar{x}=\bar{y}+2$$

Λύση

Α. Για να έχει νόημα πραγματικού αριθμού ο τύπος της συνάρτησης f, πρέπει $x>0$.
Η εξίσωση της εφαπτομένης της καμπύλης της f στο σημείο $A(1,f(1))$, δίνεται από τον τύπο:

$(\varepsilon): y=\alpha\cdot x+\beta$, όπου ο συντελεστής $\alpha=f'(1)$.

Είναι: $f'(x)=\left(x^2\ln x+x+1\right)'=2x\ln x+x^2\dfrac{1}{x}+1=2x\ln x+x+1,\ x>0;$

Οπότε είναι: $f'(1)=2\cdot 1\cdot\ln 1+1+1=2\cdot 1\cdot 0+2=2$

Επομένως, η εξίσωση της εφαπτομένης θα είναι:

(ε): $y=f'(1)\cdot x+\beta=2\cdot x+\beta$. **(1)**.

Εξάλλου είναι, $f(1)=1^2\cdot\ln 1+1+1=1\cdot 0+2=2$
Επειδή η εφαπτομένη διέρχεται από το σημείο $A(1,f(1))$
δηλαδή το $A(1,2)$, οι συντεταγμένες του θα επαληθεύουν την **(1)**.
Έχουμε:

$$2=2\cdot 1+\beta\Leftrightarrow\beta=2-2\Leftrightarrow\beta=0.$$

Οπότε η εξίσωση **(1)** της εφαπτομένης, γίνεται:

(ε): $y=2\cdot x$ **(2)**.

Τώρα, αν x_i , $i = 1,2,...,50$, είναι οι τετμημένες των 50 σημείων της εφαπτομένης της καμπύλης της f , με μέση τιμή $\overline{x} = 15$, τότε για τη μέση τιμή $\overline{y}$ των τεταγμένων y_i λόγω της (2) θα ισχύει:

$$\overline{y} = 2\overline{x} \quad \textbf{(3)}.$$

Όμως μας δίνεται ότι $\overline{x} = 15$. Οπότε η (3), γίνεται:

$$\overline{y} = 2\overline{x} \overset{\overline{x}=15}{=} 2 \cdot 15 = 30 \Leftrightarrow \overline{y} = 30$$

Β. Έστω η συνάρτηση f , με:

$$f(x) = (x + x_1)^2 + (x + x_2)^2 + ... + (x + x_\nu)^2$$, όπου $x \in \mathbb{R}$<

Παίρνουμε την πρώτη παράγωγο της f ως προς $x \in \mathbb{R}$:

$$f'(x) = \left[(x + x_1)^2\right]' + \left[(x + x_2)^2\right]' + ... + \left[(x + x_\nu)^2\right]' =$$

$$= 2(x + x_1)(x + x_1)' + 2(x + x_2)(x + x_2)' + ... + 2(x + x_\nu)(x + x_\nu)' =$$

$$= 2(x + x_1) + 2(x + x_2) + ... + 2(x + x_\nu) = 2(\nu x + x_1 + x_2 + ... + x_\nu) =$$

$$= 2(\nu x + \nu \cdot \frac{x_1 + x_2 + ... + x_\nu}{\nu}) = 2(\nu x + \nu \overline{x}) = 2\nu(x + \overline{x})$$

Στη συνέχεια βρίσκουμε τις ρίζες της $f'(x) = 0$, δηλαδή:

$$2\nu(x + \overline{x}) = 0 \Leftrightarrow x + \overline{x} = 0 \Leftrightarrow x = -\overline{x}$$

Ελέγχουμε πότε είναι $f'(x) > 0$ και πότε $f'(x) < 0$, δηλαδή:

$f'(x) > 0 \Leftrightarrow x + \overline{x} > 0 \Leftrightarrow x > -\overline{x}$ και

$f'(x) < 0 \Leftrightarrow x + \overline{x} < 0 \Leftrightarrow x < -\overline{x}$

Τέλος κατασκευάζουμε πίνακα πρόσημου της $f'(x)$, με τα αντίστοιχα συμπεράσματα για την f . Έχουμε:

x	$-\infty$	$-\overline{x}$	$+\infty$
$f'(x)$	-	0	+
$f(x)$	$\searrow$	Ολικό Ελάχιστο	$\nearrow$

Παρατηρούμε ότι για $x = -\overline{x}$, η $f\left(x\right)$ γίνεται ελάχιστη.

Γ. Από τους 100 μαθητές, 20 δήλωσαν 5 ώρες λιγότερες από τις πραγματικές και 50 δήλωσαν 6 ώρες περισσότερες, ενώ οι υπόλοιποι 30 είπαν αλήθεια για το χρόνο της εβδομαδιαίας μελέτης τους.

Έστω $\kappa_1,\kappa_2,...,\kappa_{20}$, $\lambda_1,\lambda_2,...,\lambda_{50}$ και $\mu_1,\mu_2,...,\mu_{30}$ οι αντίστοιχες ώρες μελέτης που δήλωσαν αντίστοιχα οι 20, οι 50 και οι 30 μαθητές.

Τότε θα ισχύει:

$$\bar{x}=\frac{(\kappa_1+\kappa_2+...+\kappa_{20})+(\lambda_1+\lambda_2+...+\lambda_{50})+(\mu_1+\mu_2+...+\mu_{30})}{100}$$

Οι πραγματικές όμως ώρες μελέτης τους είναι αντίστοιχα:

$\kappa_1+5,\kappa_2+5,...,\kappa_{20}+5$, $\lambda_1-6,\lambda_2-6,...,\lambda_{50}-6$, $\mu_1,\mu_2,...,\mu_{30}$.

Επομένως ο πραγματικός μέσος όρος των ωρών μελέτης των μαθητών, θα είναι:

$$\bar{y}=\frac{[(\kappa_1+5)+(\kappa_2+5)+...+(\kappa_{20}+5)]+[(\lambda_1-6)+(\lambda_2-6)+...+(\lambda_{50}-6)]+[(\mu_1+\mu_2+...+\mu_{30})]}{100}=$$

$$=\frac{(\kappa_1+\kappa_2+...+\kappa_{20})+(\lambda_1+\lambda_2+...+\lambda_{50})+(\mu_1+\mu_2+...+\mu_{30})}{100}+\frac{20\cdot 5-50\cdot 6}{100}=$$

$$\bar{x}-\frac{300-100}{100}=\bar{x}-\frac{200}{100}=\bar{x}-2$$

Άρα, $\bar{x}=\bar{y}+2$

Άσκηση 9

Α. Δίνονται πέντε αριθμοί με μέση τιμή $\bar{x}=\delta$, όπου δ είναι η διάμεσος των παρατηρήσεων 5, 3, 8, 9, 7 ενός δείγματος. Αν οι πέντε αυτοί αριθμοί, αποτελούν διαδοχικούς όρους μιας αριθμητικής προόδου και η μέση τιμή των τετραγώνων τους είναι ίση με 67, τότε να βρείτε τους πέντε αυτούς αριθμούς.

Β. Μια μεταβλητή Χ , παίρνει επτά τιμές με αντίστοιχες συχνότητες:

$$5,\ 8,\ \kappa^2-4,\ 12,\ 11,\ \kappa+2,10.$$

Αν γνωρίζουμε ότι το πλήθος των παρατηρήσεων είναι εκατό, τότε να προσδιορίσετε τον αριθμό κ.

Γ. Σε ένα δείγμα 50 παρατηρήσεων $t_1,t_2,\ldots,t_{50}$, με μέση τιμή $\bar{x}$, ισχύουν οι πιο κάτω σχέσεις:

$$\sum_{i=1}^{50} t_i = 2500 \quad \text{και} \quad \sum_{i=1}^{50}\left(t_i-\bar{x}\right)^2 = 8450.$$

Να εξετάσετε, εάν το πιο πάνω δείγμα, είναι ομοιογενές.

<u>Λύση</u>

Α. Αρχικά, προσδιορίζουμε τη διάμεσο των αριθμών 5, 3, 8, 9, 7. Τους διατάσσουμε σε αύξουσα σειρά:
3, 5, 7, 8, 9. Οπότε $\delta=7$, που είναι η μεσαία παρατήρηση. Οπότε και $\bar{x}=\delta=7$. Αν ω η διαφορά της αριθμητικής προόδου, τότε οι πέντε αριθμοί γράφονται κατά συμμετρικό τρόπο ως εξής:

$$x-2\omega,\ x-\omega,\ x,\ x+\omega,\ x+2\omega.$$

Οπότε έχουμε:

$$\bar{x}=\frac{(x-2\omega)+(x-\omega)+x+(x+\omega)+(x+2\omega)}{5} \text{ και } \bar{x}=7$$

Άρα προκύπτει η ισότητα:

$$\frac{(x-2\omega)+(x-\omega)+x+(x+\omega)+(x+2\omega)}{5}=7 \Leftrightarrow \frac{5x}{5}=7 \Leftrightarrow x=7 \quad \textbf{(1)}$$

Επομένως οι αριθμοί είναι:

$$7-2\omega,\ 7-\omega,\ 7,\ 7+\omega,\ 7+2\omega.$$

Οπότε έχουμε και:

$$\frac{(7-2\omega)^2+(7-\omega)^2+7^2+(7+\omega)^2+(7+2\omega)^2}{5}=67$$

Επομένως:

$$\frac{(7-2\omega)^2+(7-\omega)^2+7^2+(7+\omega)^2+(7+2\omega)^2}{5}=67\Leftrightarrow$$

$$49+4\omega^2-28\omega+49+\omega^2-14\omega+49+49+\omega^2+14\omega+49+4\omega^2+28\omega=335\Leftrightarrow$$

$$\Leftrightarrow 245+10\omega^2=335\Leftrightarrow 10\omega^2=90\Leftrightarrow \omega^2=9$$

Από όπου $\omega=-3$ ή $\omega=3$.

- Αν $\omega=-3$ και $x=7$, τότε οι ζητούμενοι αριθμοί είναι:

  13, 10, 7, 4, 1

- Αν $\omega=3$ και $x=7$, τότε οι ζητούμενοι αριθμοί είναι:

  1, 4, 7, 10, 13

Β. Αφού οι 100 παρατηρήσεις της μεταβλητής X, παίρνουν επτά μόνο τιμές με αντίστοιχες συχνότητες:

5, 8, κ^2-4, 12, 11, $\kappa+2$, 10, άρα θα ισχύει:

$5+8+(\kappa^2-4)+12+11+(\kappa+2)+10=100$ **(1)**

Όμως, όλες οι συχνότητες αυτές, είναι αριθμοί φυσικοί. Επομένως αναζητούμε φυσικούς αριθμούς ως λύσεις της εξίσωσης (1).

Επιλύουμε την (1) και έχουμε:

$$5+8+(\kappa^2-4)+12+11+(\kappa+2)+10=100\Leftrightarrow \kappa^2+\kappa+44=100\Leftrightarrow$$

$$\Leftrightarrow \kappa^2+\kappa-56=0$$

που έχει ως ρίζες τους αριθμούς:
$\kappa=-8$, που απορρίπτεται,
γιατί είναι: $\kappa+2=-8+2=-6$ που δεν είναι φυσικός αριθμός,
και $\kappa=7$ που είναι αποδεκτή ως λύση, γιατί $\kappa+2=7+2=9$.

Γ. Αφού η μέση τιμή των 50 παρατηρήσεων είναι ίση με $\overline{x}$, άρα θα ισχύει η σχέση:

$\overline{x}=\dfrac{\sum_{i=1}^{\nu} t_i}{\nu}=\dfrac{\sum_{\iota=1}^{50} t_i}{50}=\dfrac{2500}{50}=50$, όπου $\nu=50$.

Εξάλλου ισχύει και:

$$S^2 = \frac{\sum_{i=1}^{50}\left(t_i - \overline{x}\right)^2}{\nu} = \frac{8450}{50} = 169$$

Οπότε θα είναι και $S = \sqrt{S^2} = \sqrt{169} = 13$.

Επομένως:

$$CV = \frac{S}{\overline{x}} = \frac{13}{50} = 0,26 = 26\% .$$

Επειδή $CV = 26\% > 10\%$, άρα το δείγμα, δεν είναι ομοιογενές.

Άσκηση 10

Στον πιο κάτω πίνακα έχουν ταξινομηθεί σε τέσσερεις κλάσεις, τα ύψη 200 μαθητών ενός Γενικού Λυκείου της Θεσσαλονίκης, με αντίστοιχες σχετικές συχνότητες f_1, f_2, f_3, f_4. Αν γνωρίζουμε ότι ισχύει η σχέση:

$$f_1^2 + f_2^2 + f_3^2 + 0,29 = 2(0,2f_1 + 0,4f_2 + 0,3f_3)$$

Ύψη σε κλάσεις [-)	**Σχετικές συχνότητες**
[164,170)	f_1
[170,176)	f_2
[176,182)	f_3
[182,188)	f_4

Τότε:
α) να υπολογίσετε τις τιμές των σχετικών συχνοτήτων f_1, f_2, f_3, f_4.
β) να υπολογίσετε το μέσο ύψος των μαθητών.
γ) να βρείτε πόσοι μαθητές έχουν ύψος τουλάχιστον 180 εκατοστά.

Λύση

α) Από τη δοσμένη σχέση έχουμε:

$$f_1^2 + f_2^2 + f_3^2 + 0,29 = 2(0,2f_1 + 0,4f_2 + 0,3f_3) \Leftrightarrow$$

$$\Leftrightarrow f_1^2 + f_2^2 + f_3^2 - 2\cdot 0,2f_1 - 2\cdot 0,4f_2 - 2\cdot 0,3f_3 + 0,29 = 0 \Leftrightarrow$$

$$\Leftrightarrow f_1^2 - 2\cdot 0,2f_1 + 0,04 + f_2^2 - 2\cdot 0,4f_2 + 0,16 + f_3^2 - 2\cdot 0,3f_3 + 0,09 = 0 \Leftrightarrow$$

$$\Leftrightarrow (f_1 - 0,2)^2 + (f_2 - 0,4)^2 + (f_3 - 0,3)^2 = 0$$

Εξάλλου γνωρίζουμε ότι ισχύει η ισοδυναμία:

$\alpha^2 + \beta^2 + \gamma^2 = 0 \Leftrightarrow \alpha = 0$ και $\beta = 0$ και $\gamma = 0$.

Επομένως από την προηγούμενη σχέση προκύπτει:

$f_1 - 0,2 = 0$ και $f_2 - 0,4 = 0$ και $f_3 - 0,3 = 0$, οπότε έχουμε:

$f_1 = 0,2$ και $f_2 = 0,4$ και $f_3 = 0,3$. **(1)**

Επειδή το άθροισμα όλων των σχετικών συχνοτήτων είναι ίσο με 1, άρα θα έχουμε:

$$f_1+f_2+f_3+f_4=1 \Leftrightarrow f_4=1-\left(f_1+f_2+f_3\right) \Rightarrow f_4=1-\left(0,2+0,4+0,3\right) \Leftrightarrow$$

$$\Leftrightarrow f_4=1-0,9 \Leftrightarrow f_4=0,1$$

Επομένως οι σχετικές συχνότητες είναι:

$f_1=0,2$, $f_2=0,4$, $f_3=0,3$, $f_4=0,1$.

β) Από τα στοιχεία που προέκυψαν από το προηγούμενο ερώτημα, συμπληρώνουμε τον πιο κάτω πίνακα:

Ύψη σε κλάσεις [-)	**Κεντρικές Τιμές x_i**	**Σχετικές Συχνότητες f_i**	$x_i f_i$
[164,170)	167	0,2	33,4
[170,176)	173	0,4	69,2
[176,182)	179	0,3	53,7
[182,188)	185	0,1	18,5
Σύνολο			174,8

Επομένως, αν $\overline{x}$ είναι το μέσο ύψος τους και x_i οι αντίστοιχες κεντρικές τιμές των κλάσεων ύψους, τότε θα έχουμε:

$$\overline{x}=\sum_{i=1}^{4} x_i f_i = 167 \cdot 0,2+173 \cdot 0,4+179 \cdot 0,3+185 \cdot 0,1=174,8$$

Άρα το μέσο ύψος τους είναι 174,8 εκατοστά.

γ) Θεωρούμε ότι η κατανομή του ύψους στο εσωτερικό κάθε κλάσης είναι ομοιόμορφη, οπότε το πλήθος των μαθητών με ύψος τουλάχιστον 180 εκατοστά, θα είναι:

$$\frac{1}{3} \cdot \nu_3+\nu_4=\frac{1}{3} f_3 \cdot \nu+f_4 \cdot \nu=\frac{1}{3} 0,3 \cdot 200+0,1 \cdot 200=20+20=40 \,,$$

όπου $f_3=\frac{\nu_3}{\nu} \Leftrightarrow \nu_3=f_3 \cdot \nu$ και $f_4=\frac{\nu_4}{\nu} \Leftrightarrow \nu_4=f_4 \cdot \nu$

Άρα με ύψος τουλάχιστον 180 εκατοστών, αριθμούν 40 μαθητές.

Άσκηση 11

Μια μεταβλητή παίρνει τις τιμές 1, 2, 3, 4 με αντίστοιχες συχνότητες 4, $\kappa+1$, 4, $\lambda-2$ και με μέση τιμή 2,5. Αν γνωρίζουμε ότι το πλήθος των παρατηρήσεων είναι 20, τότε:

α) Να υπολογίσετε τις τιμές των κ και λ.

β) Να βρείτε την τυπική απόκλιση και το συντελεστή μεταβολής της μεταβλητής.

Λύση

α) Κατασκευάζουμε τον πιο κάτω πίνακα:

Τιμές της μεταβλητής	**Συχνότητες**
1	4
2	$\kappa+1$
3	4
4	$\lambda-2$
Σύνολο	20

Οπότε θα έχουμε:

$$2,5=\frac{1\cdot4+2\cdot(\kappa+1)+3\cdot4+4\cdot(\lambda-2)}{20}\Leftrightarrow$$

$$1\cdot4+2\cdot(\kappa+1)+3\cdot4+4\cdot(\lambda-2)=50\Leftrightarrow$$

$$\Leftrightarrow 4+2\kappa+2+12+4\lambda-8=50\Leftrightarrow 2\kappa+4\lambda=40\Leftrightarrow \kappa+2\lambda=20$$

Οπότε έχουμε:

$\kappa+2\lambda=20$ **(1)**.

Εξάλλου είναι:

$$4+(\kappa+1)+4+(\lambda-2)=20\Leftrightarrow 4+\kappa+1+4+\lambda-2=20\Leftrightarrow$$

$$\Leftrightarrow \kappa+\lambda=13.$$

Οπότε έχουμε και:

$\kappa+\lambda=13$ **(2)**

Από τις (1) και (2) προκύπτει το σύστημα:

$$\begin{cases}\kappa+2\lambda=20\\ \kappa+\lambda=13\end{cases}\Leftrightarrow\begin{cases}\kappa=6\\ \lambda=7\end{cases}.$$

Επομένως οι τιμές των κ και λ, είναι $\kappa = 6$ και $\lambda = 7$.

β) Εφαρμόζουμε τον τύπο της διακύμανσης:

$S^2 = \frac{1}{\nu}\sum_{1}^{\kappa}\left(\overline{x} - x_i\right)^2 \cdot \nu_i$, οπότε έχουμε:

$$S^2 = \frac{1}{20}\left[4\left(2,5-1\right)^2 + 7\left(2,5-2\right)^2 + 4\left(2,5-3\right)^2 + 5\left(2,5-4\right)^2\right] =$$

$$= \frac{1}{20}\left(4 \cdot 2,25 + 7 \cdot 0,25 + 4 \cdot 0,25 + 5 \cdot 2,25\right) = \frac{1}{20}\left(9 + 2,75 + 11,25\right) =$$

$$= \frac{1}{20} \cdot 23 = 1,15$$

Επομένως:

$$S = \sqrt{1,15} \cong 1,07 \text{ και } CV = \frac{S}{\overline{x}} = \frac{1,07}{2,5} = 0,428 \text{ ή } 42,8\% .$$

ΘΕΜΑ Δ

Άσκηση 1

Από μια έρευνα που έγινε σχετικά με τους μισθούς των εργατών μιας επιχείρησης προέκυψε ότι το 2,5% των εργατών έχει μηνιαίο μισθό μικρότερο από 500 ευρώ, ενώ το 84% των εργατών έχει μισθό μικρότερο από 800 ευρώ. Υποθέτουμε ότι η κατανομή των μισθών είναι περίπου κανονική.

α) Να βρείτε τη μέση τιμή και την τυπική απόκλιση των μισθών.

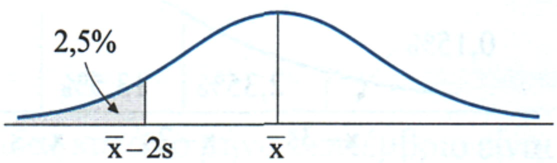
β) Να εξετάσετε αν το δείγμα είναι ομοιογενές

γ) Αν η επιχείρηση απασχολεί 400 εργάτες να βρείτε:
  i. Πόσοι εργάτες έχουν μισθό από 500 έως 800 ευρώ.
  ii. Πόσοι εργάτες έχουν μισθό μεγαλύτερο από 900 ευρώ.

Λύση

Γνωρίζουμε ότι:

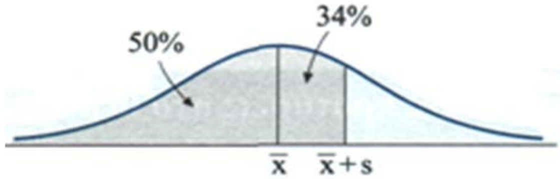
• Σε μια κανονική κατανομή το 95% περίπου των παρατηρήσεων βρίσκεται στο διάστημα $(\bar{x}-2s, \bar{x}+2s)$ άρα εκτός του διαστήματος αυτού βρίσκεται το υπόλοιπο 5%, από το οποίο το 2,5% είναι μικρότερες από $\bar{x}-2s$ ενώ το άλλο 2,5% είναι μεγαλύτερες από το $\bar{x}+2s$.



Επειδή το 2,5% των εργατών έχει μηνιαίο μισθό μικρότερο από 500 ευρώ, έχουμε:

$\bar{x}-2s=500$ **(1)**

• Σε μια κανονική κατανομή το 84% περίπου των παρατηρήσεων είναι μικρότερες από $\bar{x}+s$. Επειδή το 84% των εργατών έχει μισθό το πολύ 800 ευρώ, έχουμε:



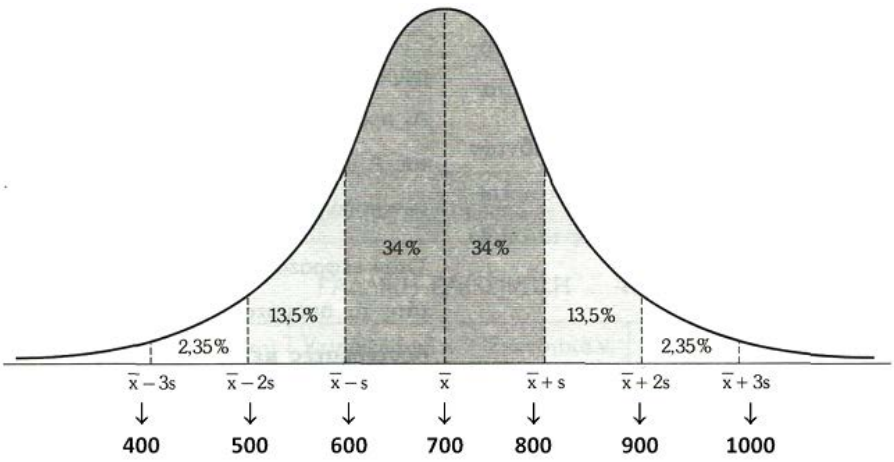
$\bar{x}+s=800$ **(2)**

Από τη σχέση (2) έχουμε $\bar{x}+s=800 \Leftrightarrow \bar{x}=800-s$ και με αντικατάσταση στη σχέση (1) έχουμε

$800-s-2s=500 \Leftrightarrow 3s=300 \Leftrightarrow \boxed{s=100}$

άρα $\bar{x}=800-100=700$, $\boxed{\bar{x}=700}$

Επειδή $\bar{x}=700$ και $s=100$ έχουμε την παρακάτω κανονική κατανομή



β) Για να βρούμε αν το δείγμα είναι ομοιογενές υπολογίζουμε τον συντελεστή μεταβλητότητας

$$CV = \frac{s}{\bar{x}} = \frac{100}{700} = 0{,}143 = 14{,}3\%$$

Εφόσον ο συντελεστής μεταβλητότητας ξεπερνά το 10%, το δείγμα δεν είναι ομοιογενές.

γ)

i. Αν η επιχείρηση απασχολεί 400 εργάτες τότε το ποσοστό των εργατών που έχουν μισθό από 500 έως 800 ευρώ όπως παρατηρούμε από το παραπάνω διάγραμμα είναι $13{,}5\% + 34\% + 34\% = 81{,}5\%$, άρα το πλήθος αυτών των εργατών είναι: $81{,}5\% \cdot 400 = 326$.

ii. Το ποσοστό των εργατών που έχουν μισθό μεγαλύτερο από 900 ευρώ, όπως παρατηρούμε από το παραπάνω διάγραμμα, είναι $\frac{2{,}5\%}{2}$, άρα το πλήθος αυτών των εργατών είναι $\left(\frac{2{,}5\%}{2}\right) \cdot 400 = 5$.

Άσκηση 2

Οι βαθμοί 100 μαθητών μιας τάξης του Λυκείου ομαδοποιημένοι σε τέσσερις κλάσεις ίσου πλάτους δίνονται στον παρακάτω πίνακα

Κλάσεις	$f_i\%$
[0, 4)	35
[4, 8)	
[8, 12)	30
[12, 16)	

Αν ο μέσος όρος της βαθμολογίας των μαθητών αυτών είναι 7

α) Να υπολογίσετε τις συχνότητες $f_2\%$ και $f_4\%$ που λείπουν
β) Να υπολογίσετε τη διάμεσο και την τυπική απόκλιση
γ) Να εξετάσετε αν το δείγμα είναι ομοιογενές
δ) Αν ο καθηγητής που διόρθωσε τα γραπτά ανεβάσει τη βαθμολογία όλων των μαθητών κατά μια μονάδα πόσο θα είναι τότε ο μέσος όρος της βαθμολογίας τους;

<u>Λύση</u>

α) Επειδή $\nu = 100$ και $f_1\% = 35, f_1 = 0,35$ οπότε $\nu_1 = f_1 \cdot \nu = 0,35 \cdot 100 \Leftrightarrow \nu_1 = 35$
Όμοια $\nu_3 = f_3 \cdot \nu = 0,30 \cdot 100 \Leftrightarrow \nu_3 = 30$

Επειδή $\nu = \nu_1 + \nu_2 + \nu_3 + \nu_4 = 100$ έχουμε $35 + \nu_2 + 30 + \nu_4 = 100 \Leftrightarrow \nu_2 + \nu_4 = 35$ **(1)**

Ακόμα επειδή $\bar{x} = 7$ έχουμε $\bar{x} = \dfrac{2 \cdot 35 + 6 \cdot \nu_2 + 10 \cdot 30 + 14 \cdot \nu_4}{100} = \dfrac{70 + 6 \cdot \nu_2 + 300 + 14 \cdot \nu_4}{100} = 7 \Leftrightarrow$

$70 + 6 \cdot \nu_2 + 300 + 14 \cdot \nu_4 = 700 \Leftrightarrow 6 \cdot \nu_2 + 14 \cdot \nu_4 = 330$ **(2)**

Από τη σχέση **(1)** έχουμε $\nu_2 = 35 - \nu_4$ οπότε από την **(2)** προκύπτει
$6(35 - \nu_4) + 14\nu_4 = 330 \Leftrightarrow$
$210 - 6\nu_4 + 14\nu_4 = 330 \Leftrightarrow 8\nu_4 = 120 \Leftrightarrow \nu_4 = 15$ άρα $\nu_2 = 35 - 15 \Leftrightarrow \nu_2 = 20$

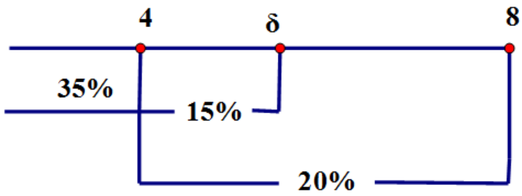
Άρα $f_4\% = 15$ και $f_2\% = 20$.

Ο πίνακας συμπληρώνεται ως εξής:

Κλάσεις	x_i	v_i	f_i	$f_i\%$	$F_i\%$	$x_i \cdot v_i$	x_i^2	$x_i^2 \cdot v_i$
[0, 4)	2	35	0,35	35	35	70	4	140
[4, 8)	6	20	0,20	20	55	120	36	720
[8, 12)	10	30	0,30	30	85	300	100	3000
[12, 16)	14	*15*	0,15	15	100	210	196	2940
Σύνολο		100	1	100		700	336	6800

β) Η διάμεσος αντιστοιχεί στην τιμή $x=\delta$ της μεταβλητής Χ έτσι ώστε το 50% των παρατηρήσεων να είναι μικρότερες ή ίσες με δ. Από τον πίνακα παρατηρούμε ότι η διάμεσος δ θα είναι ένας αριθμός της 2ης κλάσης $[4,8)$ τέτοιος ώστε στο διάστημα από 4 έως δ να ανήκει το 15% των παρατηρήσεων ώστε $35\%+15\%=50\%$.

Με βάση το παρακάτω σχήμα έχουμε:



$$\frac{\delta-4}{8-4}=\frac{15\%}{20\%}\Leftrightarrow\frac{\delta-4}{4}=\frac{15}{20}\Leftrightarrow\delta-4=3\Leftrightarrow\delta=7$$

Για τον υπολογισμό της τυπικής απόκλισης εφαρμόζουμε τον τύπο:

$$s^2=\frac{1}{\nu}\left\{\sum_{i=1}^{\nu}x_i^2\cdot\nu_i-\frac{\left(\sum_{i=1}^{\nu}x_i\cdot\nu_i\right)^2}{\nu}\right\}$$

Με βάση τον πίνακα η διακύμανση της μεταβλητής Χ είναι:

$$s^2=\frac{1}{\nu}\left\{\sum_{i=1}^{\nu}x_i^2\cdot\nu_i-\frac{\left(\sum_{i=1}^{\nu}x_i\cdot\nu_i\right)^2}{\nu}\right\}=$$

$$\frac{1}{100}\left[6800-\frac{700^2}{100}\right]=\frac{1}{100}\left[6800-\frac{490000}{100}\right]=$$

$$\frac{1}{100}[6800-4900]=\frac{1}{100}1900=19$$

και η τυπική απόκλιση $s=\sqrt{19}=4,36$

γ) Για να εξετάσουμε την ομοιογένεια του δείγματος θα βρούμε το συντελεστή μεταβλητότητας

$$CV=\frac{s}{\bar{x}}=\frac{4,36}{7}=0,623 \text{ ή } 62,3\%$$

Επειδή ο συντελεστής μεταβλητότητας είναι μεγαλύτερος από 10% το δείγμα δεν είναι ομοιογενές.

δ) Αν ο καθηγητής που διόρθωσε τα γραπτά ανεβάσει τη βαθμολογία όλων των μαθητών κατά μια μονάδα τότε ο μέσος όρος της βαθμολογίας θα ανέβει κατά μια μονάδα και θα γίνει $\overline{x}_1 = \overline{x} + 1 = 7 + 1 = 8$.

Άσκηση 3

Οι ν τιμές $t_1, t_2, \ldots, t_\nu$ μιας μεταβλητής X ακολουθούν την κανονική κατανομή, έχουν εύρος $R \simeq 6$ και το 2,5% αυτών είναι μικρότερες από το 10. Επιπλέον, αν ισχύει $\sum_{i=1}^{\nu} t_i^2 = 58000$, τότε:

α. Να βρείτε το συντελεστή μεταβολής των τιμών του δείγματος και να δείξετε ότι το δείγμα είναι ομοιογενές.

β. Να βρείτε το πλήθος ν των παρατηρήσεων.

γ. Να βρείτε το πλήθος των παρατηρήσεων που βρίσκονται στο διάστημα $(13,14)$.

δ. Αν οι τιμές αυξηθούν κατά 12 μονάδες, να δείξετε ότι ο συντελεστής μεταβολής των νέων τιμών θα γίνει ο μισός του αρχικού.

Λύση

α. Γνωρίζουμε ότι στην κανονική κατανομή το εύρος R είναι περίπου ίσο με $6 \cdot s$, άρα $6 \cdot s = 6 \Leftrightarrow s = 1$.
Επίσης το 95% των παρατηρήσεων βρίσκονται στο διάστημα $\left(\overline{x} - 2s, \overline{x} + 2s\right)$, οπότε λόγω συμμετρίας της κατανομής θα έχουμε ότι το $2,5\%$ των παρατηρήσεων θα είναι μικρότερες από $\overline{x} - 2s$. Άρα $\overline{x} - 2s = 10 \Leftrightarrow \overline{x} - 2 = 10 \Leftrightarrow \overline{x} = 12$.
Έτσι ο συντελεστής μεταβολής του δείγματος είναι $CV = \frac{s}{\overline{x}} = \frac{1}{12} < \frac{1}{10} = 10\%$, άρα το δείγμα είναι ομοιογενές.

β. Μετασχηματίζοντας ισοδύναμα τον τύπο $s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{\nu} t_i^{\,2} - \frac{\left(\sum_{i=1}^{\nu} t_i\right)^2}{\nu}\right\}$ παίρνουμε:

$$s^2 = \frac{1}{\nu}\left\{\sum_{i=1}^{\nu} t_i^{\,2} - \frac{\left(\sum_{i=1}^{\nu} t_i\right)^2}{\nu}\right\} \Leftrightarrow s^2 = \frac{1}{\nu}\sum_{i=1}^{\nu} t_i^{\,2} - \left(\frac{\sum_{i=1}^{\nu} t_i}{\nu}\right)^2 \Leftrightarrow$$

$$s^2 = \frac{1}{\nu}\sum_{i=1}^{\nu} t_i^{\,2} - \left(\overline{x}\right)^2 \Leftrightarrow \nu = \frac{\sum_{i=1}^{\nu} t_i^{\,2}}{s^2 + \left(\overline{x}\right)^2}\text{, οπότε}$$

$$\nu = \frac{58000}{1+144} = 400\,.$$

γ. Το διάστημα $(13{,}14)$ είναι το $(\overline{x}+s, \overline{x}+2s)$. Επίσης γνωρίζουμε ότι το 95% των παρατηρήσεων βρίσκονται στο διάστημα $(\overline{x}-2s, \overline{x}+2s)$ και το 68% των παρατηρήσεων βρίσκονται στο διάστημα $(\overline{x}-s, \overline{x}+s)$, οπότε λόγω συμμετρίας της κατανομής θα έχουμε ότι το $\frac{95-68}{2}\%=13{,}5\%$ των παρατηρήσεων θα βρίσκεται στο διάστημα $(\overline{x}+s, \overline{x}+2s)$. Άρα στο διάστημα $(13{,}14)$ θα βρίσκονται $\frac{13{,}5}{100}\cdot 400=54$ παρατηρήσεις.

δ. Αν οι τιμές αυξηθούν κατά 12 μονάδες, τότε σύμφωνα με γνωστή εφαρμογή του βιβλίου θα έχουμε ότι η νέα μέση τιμή $\overline{x'}$ θα είναι $\overline{x'}=\overline{x}+12=24$ και η νέα τυπική απόκλιση s' θα είναι $s'=s=1$.

Άρα ο νέος συντελεστής μεταβολής είναι $CV'=\frac{s'}{\overline{x'}}=\frac{1}{24}=\frac{CV}{2}$.

Άσκηση 4

Δίνεται η συνάρτηση $f(x)=2x^2-\lambda\cdot x+8$, $\lambda>8$, η οποία έχει ρίζες τους αριθμούς $\frac{1}{s}$ και $\overline{x}$, όπου $s\neq 0$ η τυπική απόκλιση και $\overline{x}>1$ η μέση τιμή των τιμών ενός δείγματος μιας τυχαίας μεταβλητής X.

α. Να δείξετε ότι το δείγμα δεν είναι ομοιογενές.

β. Αν η ελάχιστη τιμή της συνάρτησης f είναι η τιμή $-\frac{225}{8}$, να δείξετε ότι $s=2$ και $\overline{x}=8$.

γ. Αν ισχύουν οι τιμές του ερωτήματος β, τότε να βρείτε πόσο πρέπει τουλάχιστον να αυξηθούν οι τιμές του δείγματος, ώστε να προκύψει ομοιογενές δείγμα.

Λύση

α. Το γινόμενο των ριζών ενός τριωνύμου γνωρίζουμε ότι δίνεται από τον τύπο $P=\rho_1\cdot\rho_2=\frac{\gamma}{\alpha}$, οπότε έχουμε $\frac{1}{s}\cdot\overline{x}=\frac{8}{2}=4\Leftrightarrow\frac{\overline{x}}{s}=4\Leftrightarrow\frac{s}{\overline{x}}=\frac{1}{4}=0,25$, άρα ο συντελεστής μεταβολής του δείγματος είναι $CV=25\%>10\%$, το οποίο σημαίνει ότι το δείγμα δεν είναι ομοιογενές.

Παρατήρηση: Η συνάρτηση είναι πολυώνυμο δευτέρου βαθμού με διακρίνουσα $\Delta=\beta^2-4\cdot\alpha\cdot\gamma=\lambda^2-64>0$, άρα έχει δύο ρίζες άνισες και μάλιστα διαφορετικές του μηδενός, αφού το γινόμενό τους είναι ίσο με 4.

β. Παραγωγίζουμε τη συνάρτηση και έχουμε:

$$f'(x)=\left(2x^2-\lambda\cdot x+8\right)'=4x-\lambda.$$

Έχουμε

$$f'(x)=0\Leftrightarrow 4x-\lambda=0\Leftrightarrow x=\frac{\lambda}{4},$$

$$f'(x)>0\Leftrightarrow 4x-\lambda>0\Leftrightarrow x>\frac{\lambda}{4} \text{ και}$$

$$f'(x)<0\Leftrightarrow 4x-\lambda<0\Leftrightarrow x<\frac{\lambda}{4}.$$

Από τα παραπάνω σχηματίζουμε τον πίνακα προσήμου:

x	$-\infty$		$\frac{\lambda}{4}$		$+\infty$
$f'(x)$		$-$	0	$+$	
$f(x)$		$\searrow$	$f\left(\frac{\lambda}{4}\right)$	$\nearrow$	

Από τα παραπάνω έχουμε ότι η συνάρτηση f είναι γνησίως φθίνουσα στο διάστημα $\left(-\infty,\frac{\lambda}{4}\right]$ και γνησίως αύξουσα στο διάστημα $\left[\frac{\lambda}{4},+\infty\right)$.

Άρα το ελάχιστο της συνάρτησης είναι το $f\left(\frac{\lambda}{4}\right)=8-\frac{\lambda^2}{8}=-\frac{225}{8}$.

Οπότε $\lambda^2=289\Leftrightarrow\lambda=17$ (η αρνητική ρίζα απορρίπτεται, αφού $\lambda>8$).

Έτσι η συνάρτηση γίνεται:

$f(x)=2x^2-17x+8$ η οποία έχει ρίζες $x_{1,2}=\frac{17\pm\sqrt{225}}{4}=\begin{cases}\frac{1}{2}\\ 8\end{cases}$.

Άρα $\frac{1}{s}=\frac{1}{2}\Leftrightarrow s=2$ και $\overline{x}=8$. (αφού $\overline{x}>1$)

γ. Αν οι τιμές αυξηθούν κατά c μονάδες, τότε σύμφωνα με γνωστή εφαρμογή του βιβλίου θα έχουμε ότι η νέα μέση τιμή $\overline{x'}$ θα είναι $\overline{x'}=\overline{x}+c=8+c$ και η νέα τυπική απόκλιση s' θα είναι $s'=s=2$.

Άρα ο νέος συντελεστής μεταβολής είναι

$$CV'=\frac{s'}{\overline{x'}}=\frac{2}{8+c}.$$

Πρέπει $CV'\leq\frac{1}{10}\Leftrightarrow\frac{2}{8+c}\leq\frac{1}{10}\Leftrightarrow 8+c\geq 20\Leftrightarrow c\geq 12$.

Άρα οι τιμές του δείγματος πρέπει να αυξηθούν τουλάχιστον κατά 12 μονάδες, ώστε να προκύψει ομοιογενές δείγμα.

Άσκηση 5

Έστω t_1, t_2, t_3, t_4, t_5 οι παρατηρήσεις ενός δείγματος, που δεν είναι όλες ίσες μεταξύ τους, $\overline{x}$ η μέση τιμή τους και s^2 η διακύμανση. Επίσης, θεωρούμε τη συνάρτηση
$f(x) = (x - t_1)^3 + (x - t_2)^3 + ... + (x - t_5)^3,\ x \in \mathbb{R}$.

α) Να δείξετε ότι η f είναι γνησίως αύξουσα.

β) Αν η συνάρτηση $f'(x)$ παρουσιάζει ελάχιστο το $f'(6) = 30$, τότε:

i. να δείξετε ότι $s = \sqrt{2}$.

ii. να υπολογίσετε το άθροισμα των τετραγώνων των παρατηρήσεων.

(Δίνεται: $s^2 = \dfrac{1}{\nu}\left[\sum_{i=1}^{\nu} t_i^{\,2} - \dfrac{\left(\sum_{i=1}^{\nu} t_i\right)^2}{\nu}\right]$)

<u>Λύση</u>

α) Η παράγωγος της f είναι

$$f'(x) = 3(x - t_1)^2 \cdot (x - t_1)' + 3(x - t_2)^2 \cdot (x - t_2)' + ... + 3(x - t_5)^2 \cdot (x - t_5)' =$$

$$= 3(x - t_1)^2 + 3(x - t_2)^2 + ... + 3(x - t_5)^2 = 3\sum_{i=1}^{5}(x - t_i)^2 =$$

$3\sum_{i=1}^{5}(t_i - x)^2 > 0$, για κάθε $x \in \mathbb{R}$.

Άρα η f είναι γνησίως αύξουσα.

β)

i. Γνωρίζουμε ότι η συνάρτηση $g(x) = \sum_{i=1}^{5}(t_i - x)^2$, έχει ελάχιστο όταν $x = \overline{x}$. Οπότε και η συνάρτηση $f'(x) = 3\sum_{i=1}^{5}(t_i - x)^2 = 3g(x)$, έχει ελάχιστο για $x = \overline{x}$.

Επειδή η $f'(x)$ έχει ελάχιστο για $x = 6$, θα είναι $\overline{x} = 6$.

Επίσης είναι $f'(6) = 30 \Leftrightarrow f'(\overline{x}) = 30 \Leftrightarrow 3\sum_{i=1}^{5}(t_i - \overline{x})^2 = 30 \Leftrightarrow \sum_{i=1}^{5}(t_i - \overline{x})^2 = 10$ **(1)**.

Όμως η διακύμανση των πέντε παρατηρήσεων δίνεται απ’ τον τύπο

$$s^2=\frac{1}{5}\sum_{i=1}^{5}\left(t_i-\overline{x}\right)^2\overset{(1)}{=}\frac{10}{5}=2$$. Άρα $s=\sqrt{2}$.

ii. Από τον τύπο που δίνεται έχουμε

$$s^2=\frac{1}{5}\left[\sum_{i=1}^{5}t_i^{\,2}-\frac{\left(\sum_{i=1}^{5}t_i\right)^2}{5}\right]\Leftrightarrow s^2=\frac{\sum_{i=1}^{5}t_i^{\,2}}{5}-\left(\frac{\sum_{i=1}^{5}t_i}{5}\right)^2\Leftrightarrow$$

$$s^2=\frac{\sum_{i=1}^{5}t_i^{\,2}}{5}-\overline{x}^2\Leftrightarrow\sum_{i=1}^{5}t_i^{\,2}=5\left(s^2+\overline{x}^2\right).$$

Άρα $\sum_{i=1}^{5}t_i^{\,2}=5\left(2+36\right)=190$.

Άσκηση 6

Δίνεται η συνάρτηση $f(x)=x\cdot\ln x,\ \ x>0$.

α) Να μελετήσετε την f ως προς τη μονοτονία και τα ακρότατα.

β) Αν τα σημεία $A_1(x_1,y_1),A_2(x_2,y_2),...,A_5(x_5,y_5)$ με $0<x_1<x_2<x_3<x_4<x_5<\frac{1}{e}$ ανήκουν στη γραφική παράσταση της f και ισχύει $x_1\cdot x_2\cdot x_3\cdot x_4\cdot x_5=e^{-22}$, τότε:

i. να βρείτε τη μέση τιμή των συντελεστών διεύθυνσης των εφαπτομένων της γραφικής παράστασης της f στα παραπάνω σημεία.

ii. να υπολογίσετε το εύρος των συντελεστών διεύθυνσης των εφαπτομένων της γραφικής παράστασης της f στα παραπάνω σημεία, αν επιπλέον ισχύει $x_5=e^5\cdot x_1$.

<u>Λύση</u>

α) Η παράγωγος της f είναι η $f'(x)=(x)'\cdot\ln x+x\cdot(\ln x)'=\ln x+x\cdot\frac{1}{x}=\ln x+1,\ x>0$.

$$f'(x)=0\Leftrightarrow \ln x+1=0\Leftrightarrow \ln x=-1\Leftrightarrow x=\frac{1}{e}$$

$$f'(x)>0\Leftrightarrow \ln x+1>0\Leftrightarrow \ln x>-1\Leftrightarrow x>\frac{1}{e}$$

Η μονοτονία και τα ακρότατα της f φαίνονται στον παρακάτω πίνακα:

x	0		$1/e$		$+\infty$
$f'(x)$		-	0	+	
$f(x)$		$\searrow$	ο.ε.	$\nearrow$	

Από τον παραπάνω πίνακα προκύπτει ότι η f:

- Είναι γνησίως αύξουσα στο $\left[\frac{1}{e},+\infty\right)$

- Είναι γνησίως φθίνουσα στο $\left(0,\frac{1}{e}\right]$.

- Έχει ελάχιστο το $f\left(\frac{1}{e}\right)=\frac{1}{e}\ln\frac{1}{e}=-\frac{1}{e}$.

β) Ο συντελεστής διεύθυνσης της εφαπτομένης της γραφικής παράστασης της f στο σημείο της $A_i\left(x_i, f(x_i)\right), i=1,2,...,5$, είναι $f'\left(x_i\right)=\ln x_i+1$.

i. Για τη μέση τιμή των αριθμών $f'\left(x_i\right), i=1,2,...,5$ έχουμε

$$\overline{y}=\frac{f'(x_1)+f'(x_2)+f'(x_3)+f'(x_4)+f'(x_5)}{5}=$$

$$\frac{\ln x_1+1+\ln x_2+1+\ln x_3+1+\ln x_4+1+\ln x_5+1}{5}=$$

$$\frac{\ln\left(x_1\cdot x_2\cdot x_3\cdot x_4\cdot x_5\right)+5}{5}=\frac{\ln e^{-22}+5}{5}=\frac{-22+5}{5}=-\frac{17}{5}.$$

ii. Ισχύει $f''(x)=\left(\ln x+1\right)'=\frac{1}{x}>0$, για κάθε $x>0$ άρα η συνάρτηση $f'(x)=\ln x+1$ είναι γνησίως αύξουσα. Έτσι, για $0<x_1<x_2<x_3<x_4<x_5<\frac{1}{e}$ θα ισχύει $f'(x_1)<f'(x_2)<f'(x_3)<f'(x_4)<f'(x_5)$. Το εύρος των συντελεστών διεύθυνσης είναι ίσο με: $R=f'(x_5)-f'(x_1)=\ln x_5+1-\ln x_1-1=\ln\frac{x_5}{x_1}=\ln\frac{x_1\cdot e^5}{x_1}=\ln e^5=5$.

Άσκηση 7

Οι παρατηρήσεις $t_1, t_2, ..., t_\nu$ ενός δείγματος μεγέθους ν έχουν μέση τιμή $\overline{x}$ και τυπική απόκλιση s. Θεωρούμε τη συνάρτηση $f(x) = x\ln x - \overline{x}\ln x + 2s,\ x > 0$. Αν η γραφική παράσταση της συνάρτησης f διέρχεται από το σημείο $A(e,e)$ και η εφαπτομένη της γραφικής παράστασης της f στο σημείο με τετμημένη $x_0 = 1$ είναι παράλληλη στον άξονα $x'x$ τότε:

Α) Να βρείτε το s και το $\overline{x}$.

Β) Αποδείξτε ότι το δείγμα δεν είναι ομοιογενές.

Γ) Αν έχουμε κανονική κατανομή και 54 παρατηρήσεις βρίσκονται στο διάστημα $(1.5, 2)$ να βρείτε:

i. το μέγεθος του δείγματος.
ii. το πλήθος των παρατηρήσεων που βρίσκονται στο διάστημα $(0, 2)$

Δ) Να βρείτε την ελάχιστη θετική τιμή του c που πρέπει να προστεθεί σε κάθε μία παρατήρηση $t_1, t_2, ..., t_\nu$ ώστε το δείγμα των παρατηρήσεων που προκύπτει να είναι ομοιογενές.

<u>Λύση</u>

Α) Αφού το $A(e,e)$ ανήκει στη γραφική παράσταση της f θα ισχύει:

$$f(e) = e \Leftrightarrow e\ln e - \overline{x}\ln e + 2s = e \Leftrightarrow e - \overline{x} + 2s = e \Leftrightarrow \overline{x} = 2s \quad \textbf{(1)}$$

Έχουμε $f'(x) = \ln x + 1 - \overline{x}\cdot\dfrac{1}{x}$, για κάθε $x > 0$.

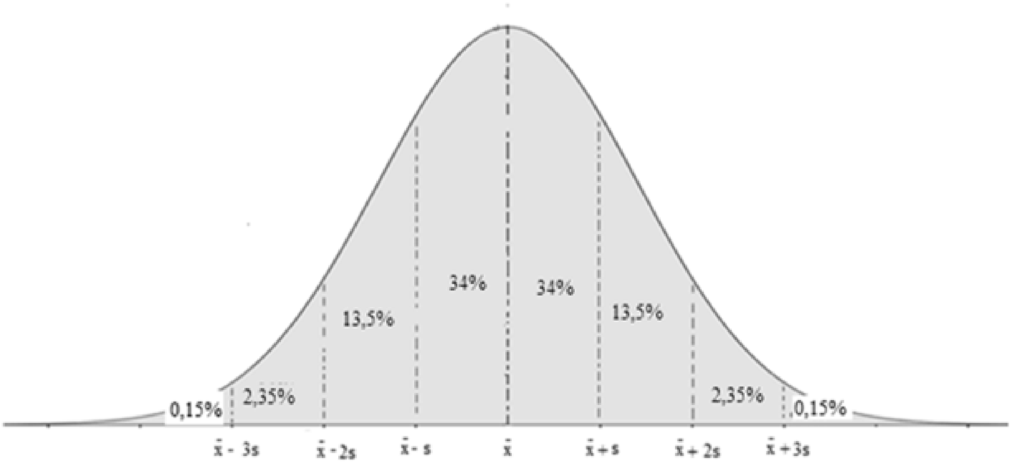
Αφού η εφαπτομένη της γραφικής παράστασης της f στο σημείο με τετμημένη $x_0 = 1$ είναι παράλληλη στον άξονα $x'x$ τότε:

$$f'(1) = 0 \Leftrightarrow \ln 1 + 1 - \overline{x} = 0 \Leftrightarrow \overline{x} = 1 \quad \textbf{(2)}$$

Τότε από την (1) έχουμε $s = \dfrac{1}{2}$.

Β) Έχουμε $CV = \dfrac{s}{\overline{x}} = \dfrac{1/2}{1} = \dfrac{1}{2} = 0{,}5 = 50\% > 10\%$. Άρα δεν είναι ομοιογενές.

Γ)



i. Στο διάστημα $\left(1.5, 2\right) = \left(\overline{x} + s, \overline{x} + 2s\right)$ αντιστοιχεί το 13,5% των παρατηρήσεων. Άρα θα έχουμε
$\frac{13,5}{100}\nu = 54 \Leftrightarrow \nu = 400$, το μέγεθος του δείγματος.

ii. Στο διάστημα $\left(0, 2\right) = \left(\overline{x} - 2s, \overline{x} + 2s\right)$ αντιστοιχεί το 95% των παρατηρήσεων. Άρα θα έχουμε $\frac{95}{100}400 = 380$ παρατηρήσεις.

Δ) Οι νέες παρατηρήσεις προκύπτουν από τη σχέση $y_i = x_i + c$, $c > 0$ και $i = 1, 2, ..., \nu$
Οπότε $\overline{y} = \overline{x} + c$ και $s_y = s_x$.

Το δείγμα είναι ομοιογενές αν και μόνο αν ισχύει η σχέση $CV' \leq 10\%$.

Δηλαδή $CV' \leq 10\% \Leftrightarrow \frac{s_y}{\overline{y}} \leq 10\% \Leftrightarrow \frac{1/2}{1+c} \leq \frac{10}{100} \Leftrightarrow 50 \leq 10 + 10c \Leftrightarrow c \geq 4$.

Άρα, η ελάχιστη τιμή του c ώστε το δείγμα να είναι ομοιογενές είναι ίση με 4.

Άσκηση 8

Έστω ένα δείγμα τριάντα παρατηρήσεων $t_1, t_2, t_3, ..., t_{30}$, με μέση τιμή $\overline{x} = 31$.

Α. Να προσδιορίσετε την τιμή του αριθμού c, που αν προστεθεί σε κάθε μία από τις παρατηρήσεις αυτές, τότε οι παρατηρήσεις $y_1, y_2, y_3, ..., y_{30}$, που θα προκύψουν, να έχουν μέση τιμή $\overline{y} = 40$.

Β. Αν πέντε από τις παρατηρήσεις $t_1, t_2, t_3, ..., t_{30}$, μειωθούν κατά 2 και δέκα παρατηρήσεις αυξηθούν κατά 4, προκύπτουν οι παρατηρήσεις $z_1, z_2, z_3, ..., z_{30}$

i. Να βρείτε τη μέση τιμή $\overline{z}$ των παρατηρήσεων $z_1, z_2, z_3, ..., z_{30}$

ii. Να βρείτε τη διακύμανση και την τυπική απόκλιση, αν ισχύει: $\sum_{i=1}^{30} z^2{}_i = 31800$.

iii. Να εξετάσετε εάν το δείγμα είναι ομοιογενές.

Γ. Δίνονται οι παρατηρήσεις $x_1, x_2, ..., x_\nu$, ενός δείγματος ως προς μια μεταβλητή X. Αν οι παρατηρήσεις y_i και t_i, όπου $i = 1, 2, ..., \nu$, που προκύπτουν από τις σχέσεις $y_i = \alpha \cdot x_i + \beta$ και $t_i = -\frac{1}{\alpha} x_i - \frac{1}{\beta}$ όπου

$\alpha > 0$, $\beta \neq 0$ και $\overline{x} = 1$, να αποδείξετε ότι ισχύει η σχέση:

$$\beta \cdot CV_y + \alpha \cdot CV_t = 0$$

όπου CV_y και CV_t, είναι οι συντελεστές μεταβολής των μεταβλητών Y και T.

<u>Λύση</u>

Α. Έχουμε: $\overline{x} = \frac{1}{30} \sum_{i=1}^{30} t_i$ **(1)**

και $y_i = t_i + c$, $i = 1, 2, 3, ..., 30$ **(2)** οπότε θα είναι:

$\overline{y} = \frac{1}{30} \sum_{i=1}^{30} y_i$ που λόγω γνωστής εφαρμογής, μας δίνει ότι:

$\overline{y} = \overline{x} + c$

Οπότε, αφού $\overline{y} = 40$ και $\overline{x} = 31$, η σχέση $\overline{y} = \overline{x} + c$, γίνεται:

$40 = 31 + c \Leftrightarrow c = 9$

Επομένως ο σταθερός αριθμός είναι $c = 9$.

Β. i. Από τα δεδομένα έχουμε ότι, πέντε παρατηρήσεις μειώνονται κατά 2 και δέκα παρατηρήσεις αυξάνονται κατά 4, η κάθε μια, ενώ οι υπόλοιπες δεκαπέντε παραμένουν αμετάβλητες. Οπότε θα έχουμε:

$$\overline{z}=\frac{\sum_{i=1}^{30}t_i-5\cdot2+10\cdot4}{30}=\frac{\sum_{i=1}^{30}t_i+30}{30}=\frac{\sum_{i=1}^{30}t_i}{30}+\frac{30}{30}=\overline{x}+1=31+1=32$$

Αφού είναι $\overline{x}=\frac{1}{30}\sum_{i=1}^{30}t_i=\frac{\sum_{i=1}^{30}t_i}{30}$.

Άρα η μέση τιμή $\overline{z}$ των παρατηρήσεων $z_1,z_2,z_3,...,z_{30}$, είναι $\overline{z}=32$.

ii. Από το προηγούμενο ερώτημα έχουμε ότι $\overline{z}=32$. Επομένως θα έχουμε:

$\overline{z}=\frac{\sum_{i=1}^{30}z_i}{30}\Rightarrow 32=\frac{\sum_{i=1}^{30}z_i}{30}\Leftrightarrow\sum_{i=1}^{30}z_i=960$, οπότε θα είναι:

$$S_z^{\,2}=\frac{1}{30}\left\{\sum_{i=1}^{30}z_i^{\,2}-\frac{\left(\sum_{i=1}^{30}z_i\right)^2}{30}\right\}=\frac{1}{30}\left(31800-\frac{960^2}{30}\right)=\frac{1}{30}\left(31800-30720\right)=$$

$$=\frac{1}{30}\cdot1080=36$$

Επομένως, η τυπική απόκλιση S_z θα είναι: $S_z=\sqrt{36}=6$

iii.Έχουμε:

$CV_z=\frac{S_z}{\overline{z}}=\frac{6}{32}=0,1875$ ή $18,75\%$,

οπότε το δείγμα δεν είναι ομοιογενές, αφού ο συντελεστής μεταβολής, είναι μεγαλύτερος του 10%.

Γ. Από γνωστή εφαρμογή, θα έχουμε:

$\overline{y}=\alpha\cdot\overline{x}+\beta$ και $\overline{t}=-\frac{1}{\alpha}\cdot\overline{x}-\frac{1}{\beta}$.

Όμως $\overline{x}=1$, οπότε έχουμε:

$\bar{y}=\alpha\cdot 1+\beta=\alpha+\beta$ **(1)** και $\bar{t}=-\frac{1}{\alpha}\cdot\bar{x}-\frac{1}{\beta}=-\frac{1}{\alpha}\cdot 1-\frac{1}{\beta}=-\frac{1}{\alpha}-\frac{1}{\beta}$ **(2)**.

Αν S_y και S_t είναι οι αντίστοιχες τυπικές αποκλίσεις, τότε θα έχουμε:

$S_y=|\alpha|\cdot S_x\overset{\alpha>0}{=}\alpha\cdot S_x$ και

$S_t=\left|-\frac{1}{\alpha}\right|\cdot S_x=\frac{1}{\alpha}\cdot S_x$.

Από τις **(1)** και **(2)**, θα έχουμε:

$CV_y=\frac{S_y}{\bar{y}}=\frac{\alpha\cdot S_x}{\alpha+\beta}$ και

$$CV_t=\frac{S_t}{\bar{t}}=\frac{\frac{1}{\alpha}\cdot S_x}{-\frac{1}{\alpha}-\frac{1}{\beta}}=$$

$$-\frac{\frac{1}{\alpha}\cdot S_x}{\frac{\alpha+\beta}{\alpha\cdot\beta}}=-\frac{\beta\cdot S_x}{\alpha+\beta}.$$

Επομένως θα έχουμε:

$$\beta\cdot CV_y+\alpha\cdot CV_t=\frac{\alpha\beta\cdot S_x}{\alpha+\beta}-\frac{\alpha\beta\cdot S_x}{\alpha+\beta}=0$$

Άσκηση 9

Α. Σε ένα δείγμα, οι τέσσερεις παρατηρήσεις μιας μεταβλητής Χ, είναι $x_1 = 2$, $x_2 = 4$, $x_3 = 6$, $x_4 = 8$. Να εξετάσετε εάν το δείγμα αυτό είναι ομοιογενές.

Β. Δίνονται x_1, x_2, x_3,..., x_ν οι παρατηρήσεις μιας μεταβλητής Χ, με μέση τιμή $\overline{x}$ και τυπική απόκλιση S.
Αν $y_i = \dfrac{2x_i - 3\overline{x}}{S}$, όπου $i = 1, 2, 3, ..., \nu$, είναι οι παρατηρήσεις μιας άλλης μεταβλητής Ψ, τότε να αποδείξετε ότι ισχύει:
$\overline{y} = -\dfrac{1}{CV_x}$.

Γ. Έστω x_1, x_2, x_3,..., x_ν, οι παρατηρήσεις μιας μεταβλητής Χ, ενός δείγματος μεγέθους ν, με τυπική απόκλιση μηδέν και μέση τιμή $\overline{x}$.
Αν οι συναρτήσεις f και g, είναι ορισμένες στο $\mathbb{R}$, και για κάθε $x \in \mathbb{R}$, ισχύει:

$$f^2(x) + g^2(x) + \overline{x}^2 = 2\overline{x}f(x) - {x_1}^2 + 2x_\nu g(x)$$

Τότε να αποδείξετε ότι είναι $f(x) = g(x)$, για κάθε $x \in \mathbb{R}$.

Λύση

Α. Αρχικά υπολογίζουμε τη μέση τιμή των τεσσάρων αυτών παρατηρήσεων. Έχουμε:
$\overline{x} = \dfrac{x_1 + x_2 + x_3 + x_4}{4} = \dfrac{20}{4} = 5$.

Υπολογίζουμε τη διακύμανση S^2:

$$S^2 = \frac{1}{\nu}\sum_{i=1}^{\nu}\left(x_i - \overline{x}\right)^2 \Rightarrow S^2 = \frac{1}{4}\sum_{i=1}^{4}\left(x_i - \overline{x}\right)^2 = \frac{1}{4}\left[(2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2\right] =$$

$$= \frac{1}{4}(9+1+1+9) = \frac{20}{4} = 5$$

Άρα, θα έχουμε: $S = \sqrt{S^2} = \sqrt{5}$, οπότε και $CV = \dfrac{S}{\overline{x}} = \dfrac{\sqrt{5}}{5} > \dfrac{1}{10}$ προφανώς

($\dfrac{\sqrt{5}}{5} > \dfrac{1}{10} \Leftrightarrow \sqrt{5} > \dfrac{1}{2} \Leftrightarrow 2\sqrt{5} > 1 \Leftrightarrow \left(2\sqrt{5}\right)^2 > 1^2 \Leftrightarrow 20 > 1$, που ισχύει)

Επομένως το δείγμα δεν είναι ομοιογενές.

Β. Έχουμε:

$$\overline{x} = \frac{\sum_{i=1}^{\nu} x_i}{\nu} \Leftrightarrow \sum_{i=1}^{\nu} x_i = \nu\overline{x} \quad \textbf{(1)}$$

Εξάλλου είναι και:

$$\overline{y}=\frac{\sum_{i=1}^{\nu}y_i}{\nu}=\frac{\frac{2x_1-3\overline{x}}{S}+\frac{2x_2-3\overline{x}}{S}+...+\frac{2x_\nu-3\overline{x}}{S}}{\nu}=$$

$$=\frac{2\left(x_1+x_2+...+x_\nu\right)-3\nu\overline{x}}{\nu S}=\frac{2\nu\overline{x}-3\nu\overline{x}}{\nu S}=\frac{-\nu\overline{x}}{\nu S}=-\frac{\overline{x}}{S}=-\frac{1}{CV_x}$$

Γ. Έχουμε:

$$S^2=0\Leftrightarrow\frac{1}{\nu}\sum_{i=1}^{\nu}\left(x_i-\overline{x}\right)^2=0\Leftrightarrow$$

$$\left(x_1-\overline{x}\right)^2+\left(x_2-\overline{x}\right)^2+...+\left(x_\nu-\overline{x}\right)^2=0\Leftrightarrow$$

$$\Leftrightarrow x_1-\overline{x}=0, x_2-\overline{x}=0,..., x_\nu-\overline{x}=0$$

από όπου προκύπτει:

$$x_1=x_2=x_3=...=x_\nu=\overline{x}$$

Δηλαδή:

$x_1=\overline{x}$ και $x_2=\overline{x}$ και ….. και $x_\nu=\overline{x}$

Επιπλέον, έχουμε ότι ισχύει για κάθε $x\in\mathbb{R}$:

$$f^2\left(x\right)+g^2\left(x\right)+\overline{x}^2=2\overline{x}f\left(x\right)-{x_1}^2+2x_\nu g\left(x\right)\Leftrightarrow$$

$$f^2\left(x\right)+g^2\left(x\right)+\overline{x}^2-2\overline{x}f\left(x\right)+{x_1}^2-2x_\nu g\left(x\right)=0\Leftrightarrow$$

$$\overset{\left(x_1=x_\nu=\overline{x}\right)}{\Leftrightarrow} f^2\left(x\right)+g^2\left(x\right)+\overline{x}^2-2\overline{x}f\left(x\right)+\overline{x}^2-2\overline{x}g\left(x\right)=0\Leftrightarrow$$

$$\Leftrightarrow f^2\left(x\right)-2\overline{x}f\left(x\right)+\overline{x}^2+g^2\left(x\right)-2\overline{x}g\left(x\right)+\overline{x}^2=0\Leftrightarrow$$

$$\Leftrightarrow\left(f\left(x\right)-\overline{x}\right)^2+\left(g\left(x\right)-\overline{x}\right)^2=0\Leftrightarrow\begin{cases}f\left(x\right)=\overline{x}\\g\left(x\right)=\overline{x}\end{cases}\Leftrightarrow f\left(x\right)=g\left(x\right)=\overline{x}\,.$$

Άρα $f\left(x\right)=g\left(x\right)$, για κάθε $x\in\mathbb{R}$.