

ΑΣΚΗΣΕΙΣ ΚΕΦ. 3.

(1)

① Ένας πωλητής κατά την διάρκεια 5 ταξιδιών του πραγματοποιήσε τις παρακάτω πωλήσεις:

A/A Ταξιδιού	Αριθ. Ημερών	Αξία Πωλήσεων (6€ x 12.€)	Μέση ημερήσια αξία Πωλήσεων (6€ x 12.€)
1	3	500	166,67
2	5	800	160
3	7	1500	214,29
4	4	750	187,5
5	6	900	150
Σύνολο	25	4.450	878,46

Ο διωθοντής πωλήσεων χαρακτήρισε μη ικανοποιητική την απόδοση του πωλητή αφού, κατ'αυτόν, η μέση ημερήσια αξία των πωλήσεων δια την περίοδο των 25 ημερών είναι $\frac{878,46}{5} \approx 175,69$ χιλ.€.

Ο πωλητής διαφώνησε με τη δυνάμη αυτή

και χαρακτηρίσει την απόδοσή του ικανο-⁽²⁾
ποιητική δεδομένου ότι, κατ'αυτόν, η
μέση ημερήσια αξία των πωλήσεων για
την περίοδο των 25 ημερών είναι $\frac{4450}{25}$
 $= 178$ χιλ. €. Ποιος από τους δύο έχει
δίκιο;

Λύση Για να υπολογίσουμε τη μέση ημε-
ρήσια αξία των πωλήσεων θα πρέπει να
υπολογίσουμε τον βταθμισμένο αριθμητικό
μέσο. Έχουμε ότι $\bar{x} = 166,67 \cdot \frac{3}{25} + 160 \cdot \frac{5}{25}$
 $+ 214,29 \cdot \frac{7}{25} + 187,5 \cdot \frac{4}{25} + 150 \cdot \frac{6}{25} =$
 $\frac{166,67 \cdot 3 + 160 \cdot 5 + 214,29 \cdot 7 + 187,5 \cdot 4 + 150 \cdot 6}{25}$
 $= \frac{4.450,04}{25} \approx 178 \text{ χιλ. €}.$

Άρα δίκιο έχει ο πωλητής αφού ορθολογικά
υπολόγισε τον βταθμισμένο αριθμητικό μέσο
του δείγματος.

(2) Έστω μια μονοκόρυφη κατανομή βυχο-
τήτων τέτοια ώστε το 89% περίπου των παρα-
τηρήσεων βρίσκεται στο διάστημα $(\bar{x} - 3 \cdot s, \bar{x} + 3 \cdot s)$

δχέβι με την κανονική.

(4)

Σημείωση Σε μια κανονική κατανομή στο διάστημα $(\bar{x} - 6 \cdot s, \bar{x} + 6 \cdot s)$ περιέχεται και το 100% των παρατηρήσεων.

(3) Έστω ο ακόλουθος πίνακας κατανομής συχνότητων υψών ευχάλικων ανδρών ενός δείγματος (ομαδοποιημένων σε κλάσεις πλάτους 10) :

Υψι (σε cm)	n_i	f_i	N_i	F_i
[160, 170)	4			
[170, 180)				
[180, 190)		0,10		

Για την κορυφή της κατανομής αυτής θεωρούμε ότι $M_0 = 175$. Να συμπληρωθεί ο παραπάνω πίνακας και να υπολογισθεί το μέγεθος του δείγματος.

Λύση Η επικρατούσα κλάση είναι η [170, 180)
δχλ. $i=2$. $L_2 = 170$, $C = 10$, $M_0 = 175$.
Άρα $M_0 = L_2 + \frac{1}{2} \cdot C = 170 + 5 = L_2 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot C$

όπου $\Delta_1 = v_2 - v_1$, $\Delta_2 = v_2 - v_3$. Άρα (5)

$$\frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot c = \frac{1}{2} \cdot c \Rightarrow \frac{\Delta_1}{\Delta_1 + \Delta_2} = \frac{1}{2} \Rightarrow 2\Delta_1 =$$

$$\Delta_1 + \Delta_2 \Rightarrow 2\Delta_1 - \Delta_1 = \Delta_2 \Rightarrow \Delta_1 = \Delta_2 \Rightarrow$$

$$v_2 - v_1 = v_2 - v_3 \Rightarrow -v_1 = -v_3 \Rightarrow v_1 = v_3.$$

Όμως $v_1 = 4$ οπότε $v_3 = 4$. $f_3 = \frac{v_3}{4}$, όπου

η το μέγεθος του δείχματος. Άρα $\frac{4}{4} = 0,10$

$$\Rightarrow 4 = \frac{4}{0,10} = 40. \text{ Έτσι, } f_1 = \frac{4}{40} = 0,10,$$

$$N_1 = 4, F_1 = 0,10. \text{ Επιπλέον } v_1 + v_2 + v_3 = 40$$

$$\Rightarrow 4 + v_2 + 4 = 40 \Rightarrow 8 + v_2 = 40 \Rightarrow v_2 = 40 - 8$$

$$= 32. \text{ Άρα } f_2 = \frac{32}{40} = 0,80, N_2 = N_1 + v_2 =$$

$$4 + 32 = 36, F_2 = F_1 + f_2 = 0,10 + 0,80 = 0,90,$$

$$N_3 = N_2 + v_3 = 36 + 4 = 40 \text{ (αναμενόμενο),}$$

$$F_3 = F_2 + f_3 = 0,90 + 0,10 = 1 \text{ (αναμενόμενο)}$$

Ο πίνακας συμπληρώνεται ως εξής:

γ_4 (6ε cm)	v_i	f_i	N_i	F_i
--------------------	-------	-------	-------	-------

[160, 170)	4	0,10	4	0,10
------------	---	------	---	------

[170, 180)	32	0,80	36	0,90
------------	----	------	----	------

[180, 190)	4	0,10	40	1
------------	---	------	----	---

Συμείωση | Με βάση τον συμπληρωμένο πίνακα μπορούν στη συνέχεια π.χ. να υπολογισθούν τα αριθμητικά περιγραφικά μέτρα και να δίνουν τα βχετικά διαγράμματα.

④ Πήραμε τυχαία 50 φοιτητές του Τμήματος Μάρκετινγκ και Επικοινωνίας και καταγράψαμε τα ύψη τους σε cm :

170, 172, 180, 181, 174, 168, 175, 179, 176,
177, 169, 171, 177, 176, 177, 175, 174, 172,
173, 179, 172, 173, 174, 175, 177, 176, 174,
175, 179, 176, 176, 176, 177, 177, 177, 175,
175, 179, 173, 172, 177, 177, 174, 175, 173,
170, 171, 176, ~~176~~, 177.

Να ταξινομήσετε τα ύψη σε κλάσεις συχνοτήτων πρώτους $c=2$ και να προσδιορίσετε όλες τις παραμέτρους που είναι αναγκαίες για την πλήρη μελέτη της κατανομής τους.

Λύση | Αρχικά παρατηρούμε ότι $x_{\min} = 168$, $x_{\max} = 181$ οπότε $R = x_{\max} - x_{\min} = 181 - 168 = 13$. Αν q το πλήθος των κλάσεων τότε ισχύει ότι $c = \frac{R}{q} \Rightarrow q = \frac{R}{c} \Rightarrow q = \frac{13}{2} =$

6,5 και συνεπώς θα συστηματοποιήσουμε (7)
 $q=7$ κλάσεις (κοινού πλάτους $c=2$).

Στη συνέχεια με διαδοχική κατασκευή του πίνακα κατανομής συχνотήτων:

i	y_i (6εcm)	y_i	n_i	N_i	f_i	F_i
1	[168, 170)	169	2	2	4%	4%
2	[170, 172)	171	4	6	8%	12%
3	[172, 174)	173	8	14	16%	28%
4	[174, 176)	175	12	26	24%	52%
5	[176, 178)	177	18	44	36%	88%
6	[178, 180)	179	4	48	8%	96%
7	[180, 182)	181	2	50	4%	100%
Σύνολο		50	—	100%	—	—

α) Αρχικά υπολογίζουμε τα μέτρα θέσης

Αριθμητικός μέσος

$$\bar{x} = \sum_{i=1}^7 y_i \cdot f_i = 169 \cdot 0,04 + 171 \cdot 0,08 + 173 \cdot 0,16 + 175 \cdot 0,24 + 177 \cdot 0,36 + 179 \cdot 0,08 + 181 \cdot 0,04 = 6,76 + 13,68 + 27,68 + 42 + 63,72 + 14,32 + 7,24 = \underline{175,4 \text{ cm}} \approx 0 \text{ μέσος όρος των ψώνων.}$$

Σημείωση] Θα μπορούσαμε να είχαμε χρησιμοποιήσει τα πρωτεύοντα δεδομένα για τον υπολογισμό του αριθμητικού μέσου. Αυτό όμως θα απαιτούσε πολύ περισσότερο χρόνο και θα αλάσανε την πιθανότητα να δίνει κάποιο αριθμητικό λάθος στις πράξεις. Εξάλλου η διαφορά των δύο αριθμητικών μέσων είναι πολύ μικρή.

Γεωμετρικός μέσος

Δεν θα υπολογίζουμε γεωμετρικό μέσο δεδομένου ότι δεν έχουμε χρονοδεδομένα (δηλ. μεταβολή μεγεθών σε σχέση με τον χρόνο).

Τεταρτημόρια

• $\left[\frac{n}{4} \right] = \left[\frac{50}{4} \right] = [12,5] = 12$. Δεδομένου ότι έχουμε ήδη ταξινομήσει τα δεδομένα μας κατ'αύξουσα σειρά μεγεθών, από την 6η των αθροιστικών συχνοτήτων N_i παρατηρούμε ότι βρίσκουμε στην κλάση $i=3$. Το κάτω όριο της κλάσης αυτής είναι το $L_3 = 172$, $\frac{n}{4} = 12,5$, $N_{i-1} = N_2 = 6$, $v_3 = 8$, $c=2$, οπότε με εφαρμογή του τύπου έχουμε ότι

$$Q_1 = L_3 + \frac{\frac{4}{4} - N_2}{v_3} \cdot c = 172 + \frac{(12,5 - 6)}{8} \cdot 2$$

$\approx 173,63 \text{ cm}$. Αυτό σημαίνει ότι 25%

των διατεταγμένων παρατηρήσεων είναι $\leq 173,63 \text{ cm}$ ενώ το υπόλοιπο 75% αυτών είναι $> 173,63 \text{ cm}$.

Συμείωση] Επειδή εργαζόμαστε με 1446 cm θα μπορούσαμε να βτροχολοποιήσουμε με το πρώτο δεκαδικό ψηφίο (842,6 εκ. χιλιοστά του εκατοστό). Αυτό όμως δεν είναι απαραίτητο.

• $\left[\frac{4}{2} \right] = \left[\frac{50}{2} \right] = [25] = 25$ και από την βτήση των αθροιστικών συχνοτήτων συμπεραίνουμε ότι βρίσκουμε στην κλάση $i=4$. $L_4 = 174$, $\frac{4}{2} = 25$, $N_3 = 14$, $v_4 = 12$, $c=2$ οπότε εφαρμόζοντας τον τύπο έχουμε ότι $Q_2 = \delta = L_4 + \frac{\frac{4}{2} - N_3}{v_4} \cdot c = 174 + \frac{(25 - 14)}{12} \cdot 2$

$\approx 175,83 \text{ cm}$. Αυτό σημαίνει ότι το 50%

των διατεταγμένων παρατηρήσεων είναι $\leq 175,83 \text{ cm}$ ενώ το υπόλοιπο 50% αυτών είναι

7, 175, 83 cm.

(10)

$$\bullet \left[\frac{34}{4} \right] = \left[\frac{3 \cdot 50}{4} \right] = \left[\frac{150}{4} \right] = [37,5] =$$

37 οπότε, όμοια με παραπάνω, βρίσκουμε
ΜΑΒΤΕ στην κλάση $i=5$. $L_5 = 176$,

$$\frac{34}{4} = 37,5, N_4 = 26, V_5 = 44, c=2 \text{ και}$$

$$\text{άρα } Q_3 = 176 + \frac{(37,5 - 26)}{44} \cdot 2 \approx 176,52 \text{ cm}$$

Αυτό σημαίνει ότι το 75% των διατεταγμένων
Παρατηρήσεων είναι $\leq 176,52 \text{ cm}$ ενώ
το υπόλοιπο 25% αυτών είναι $> 176,52 \text{ cm}$.

Συμείωση Ορισμένοι βιοχαραφείς, κατά
τη διαδικασία υπολογισμού των Q_1, Q_2, Q_3 ,
υπολογίζουν τα $\frac{n}{4}, \frac{n}{2}, \frac{3n}{4}$ αντίστοιχα
χωρίς ακέραια μέρη, προκειμένου να εντο-
πίσουν τη σχετική κλάση και να εφαρμό-
σουν τους αντίστοιχους τύπους.

Κορυφή

Η επικρατούσα κλάση (δηλ. η κλάση
με τη μεγαλύτερη συχνότητα) είναι η
[176, 178) σύμφωνα με τη βτήλη συχνότητων

Αρα $i=5$, $v_5 = 18$, $L_5 = 176$, $\Delta_1 = v_5 - v_4 =$ (11)
 $18 - 12 = 6$, $\Delta_2 = v_5 - v_6 = 18 - 4 = 14$, $c = 2$.

$$M_0 = L_5 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot c = 176 + \frac{6}{6+14} \cdot 2 =$$

$$176 + \frac{12}{20} = 176,6 \text{ cm. Αρα η κατα-}$$

νομή των υψών είναι μονοκύρφη.
 β) Στη συνέχεια υπολογίζουμε τα μέτρα
διαφοράς.

Εύρος.

$R = \text{άνω όριο τελευταίας κλάσης} - \text{κάτω όριο}$
 πρώτης κλάσης $= 182 - 168 = 14 \text{ cm.}$

Ενδοτεταρτημοριακό εύρος και τεταρτημοριακή απόκλιση

$$IR = Q_3 - Q_1 = 176,52 - 173,63 = 2,89 \text{ cm.}$$

Αυτό σημαίνει ότι το μέγαλο 50% των δια-
 τεταγμένων (καί' αλξούρα γειρά) υψών βρίσκε-
 ται στο διάστημα $[173,63, 176,52]$ που
 έχει εύρος 2,89 cm. Για την τεταρτημοριακή
 απόκλιση ισχύει ότι $QD = \frac{IR}{2} = \frac{2,89}{2} \approx 1,45 \text{ cm.}$

$$s^2 = \frac{1}{50-1} \left(\sum_{i=1}^7 n_i \cdot y_i^2 - 50 \cdot \bar{x}^2 \right) \text{ όπου } \bar{x}$$

ο δείγματικός μέσος που έχει υπολογισθεί από τα ομαδοποιημένα δεδομένα δηλ. $\bar{x} = 175,4 \text{ cm}$. Άρα έχουμε ότι

$$s^2 = \frac{1}{49} \left(2 \cdot 169^2 + 4 \cdot 171^2 + 8 \cdot 173^2 + 12 \cdot 175^2 + 18 \cdot 177^2 + 4 \cdot 179^2 + 2 \cdot 181^2 - 50 \cdot 175,4^2 \right) =$$

$$\dots = \frac{368}{49} \approx 7,51 \text{ cm}^2 \text{ Άρα η μέση τετραγωνική απόκλιση των υψών από το μέσο όρο τους είναι } 7,51 \text{ cm}^2.$$

Τυπική απόκλιση

$s = \sqrt{s^2} = \sqrt{7,51} \approx 2,74 \text{ cm}$. Αυτό σημαίνει ότι, κατά μέσο όρο, τα ύψη απέχουν από τον αριθμητικό τους μέσο κατά 2,74 cm περίπου.

Συντελεστής μεταβλητότητας

$$CV = \frac{s}{|\bar{x}|} = \frac{s}{\bar{x}} = \frac{2,74}{175,4} \approx 0,0156 \text{ ή } 1,56\%$$

Επειδή $CV \leq 10\%$ το δείγμα των υψών είναι ομοιογενές δηλ. παρουσιάζει μικρή σχετική μεταβλητότητα. Αυτό σημαίνει ότι τα ύψη

δεν διαβπείρονται αρκετά χάρω από τον ⁽¹³⁾
αριθμητικό τους μέσο.

δ) Τέλος υπολογίζουμε τα μέτρα ασυμμετρίας

Συντελεστής ασυμμετρίας Bowley

$$S_K = \frac{Q_3 + Q_1 - 2\bar{x}}{Q_3 - Q_1} = \frac{176,52 + 173,63 - 2 \cdot 175,83}{176,52 - 173,63}$$
$$= \frac{350,15 - 351,66}{2,89} = \frac{-1,51}{2,89} \approx -0,5225 \text{ ή}$$

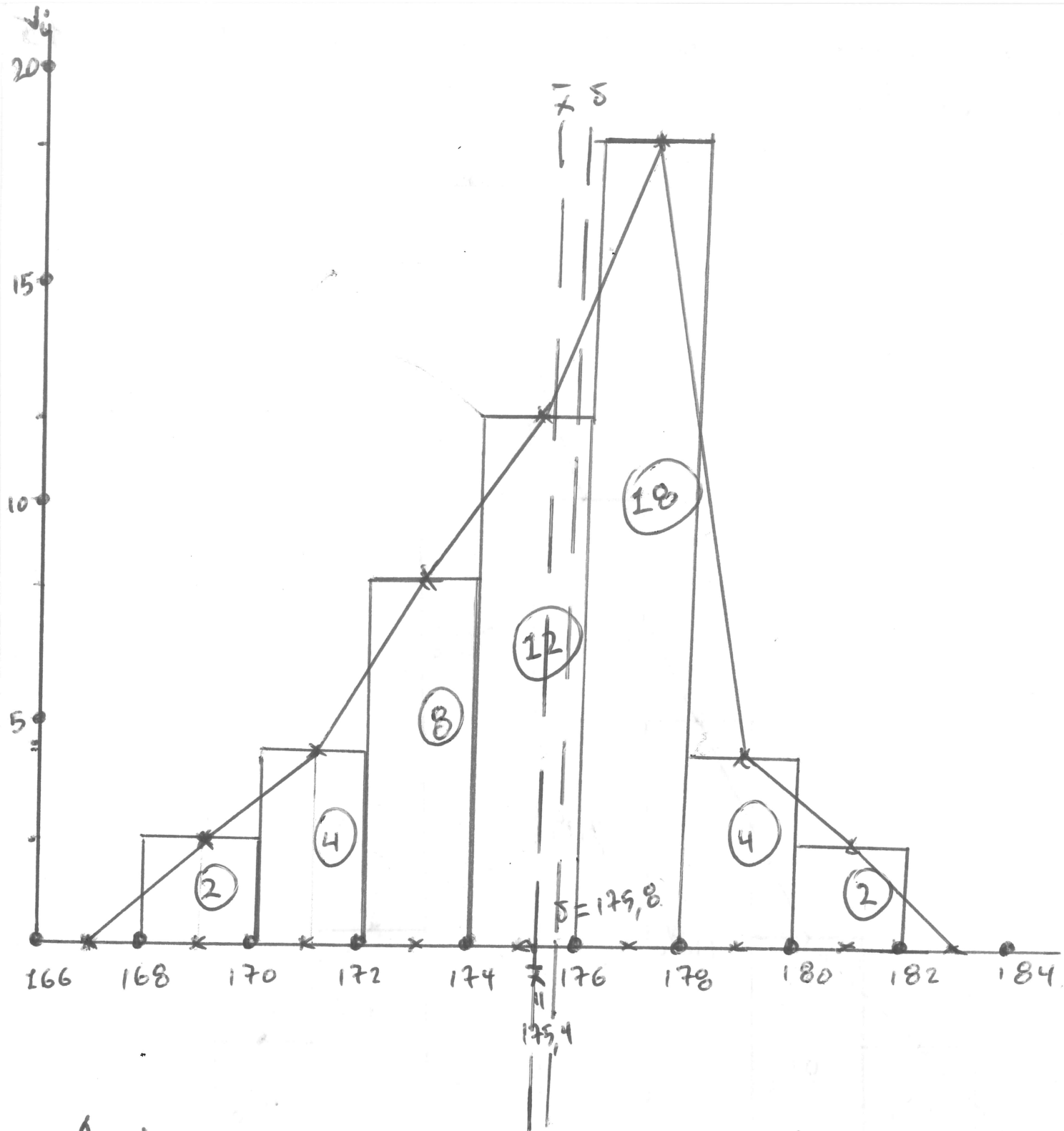
-52,25%. Αυτό σημαίνει ότι η κατανομή των υψών είναι αρνητικά ασυμμετρική

δηλ. ότι υπάρχει ουρά ή έλλειμμα τιμών στην περιοχή των μικρών υψών. Αυτό επαληθεύεται από το γεγονός ότι $Q_1 = 173,63 \text{ cm}$

$$< \bar{x} = 175,4 \text{ cm} < \bar{s} = 175,83 \text{ cm} < Q_3 = 176,52 \text{ cm}.$$

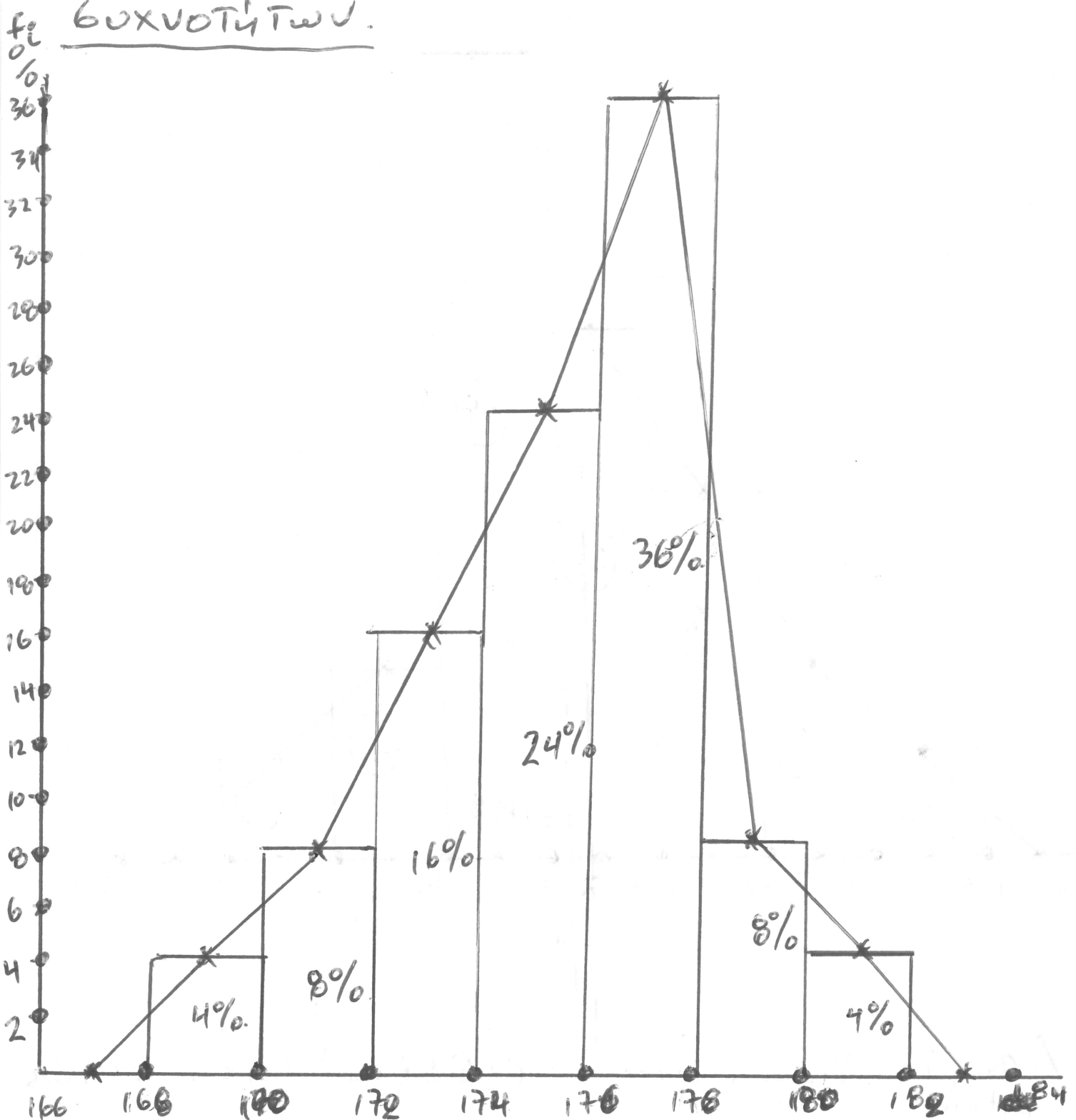
Οστόσο η μικρή διαφορά ανάμεσα στον \bar{x} και στην \bar{s} (μόλις 0,43 cm) απαιτεί επαλήθευση και μέσω του πολυχώνου συχνότητας, που θα κατασκευάσουμε παρακάτω, δεδομένου ότι η κατανομή των υψών είναι επιπλέον και μονόκορυφη.

δ) Στη συνέχεια κατασκευάζουμε τα χρυσήματα
16τόγραμμα και πολύγωνο συχνότητας



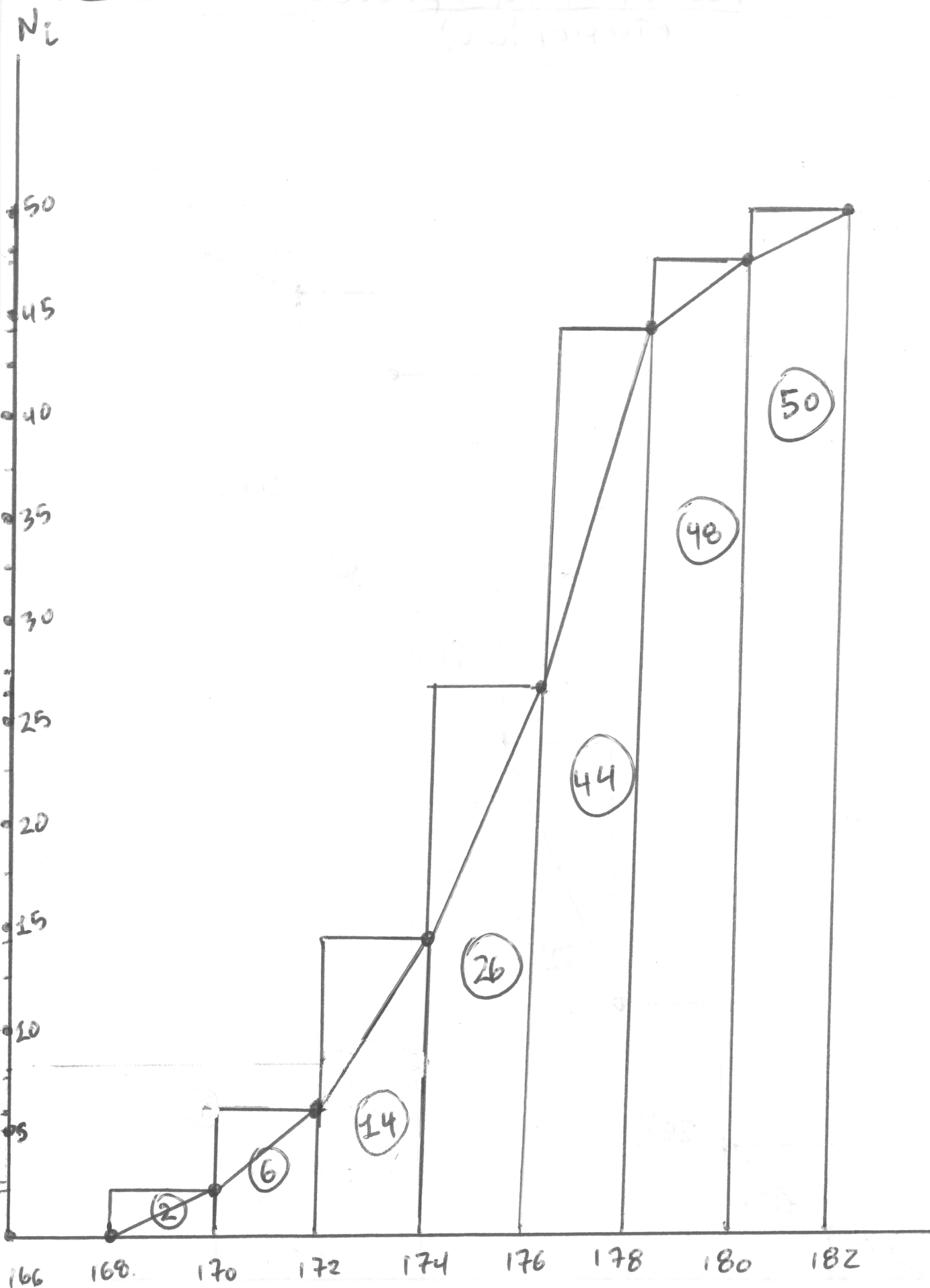
Από το διάγραμμα επιβεβαιώνεται βαρύν
και μικρή
και αριστερή ουρά δηλ. και αρνητική ασυμμε-
τρία (οπότε η κατανομή των υψών δεν
είναι κανονική).

Ιστογράμματα και πολλαπλό σχετικό (15) συχνότητες.

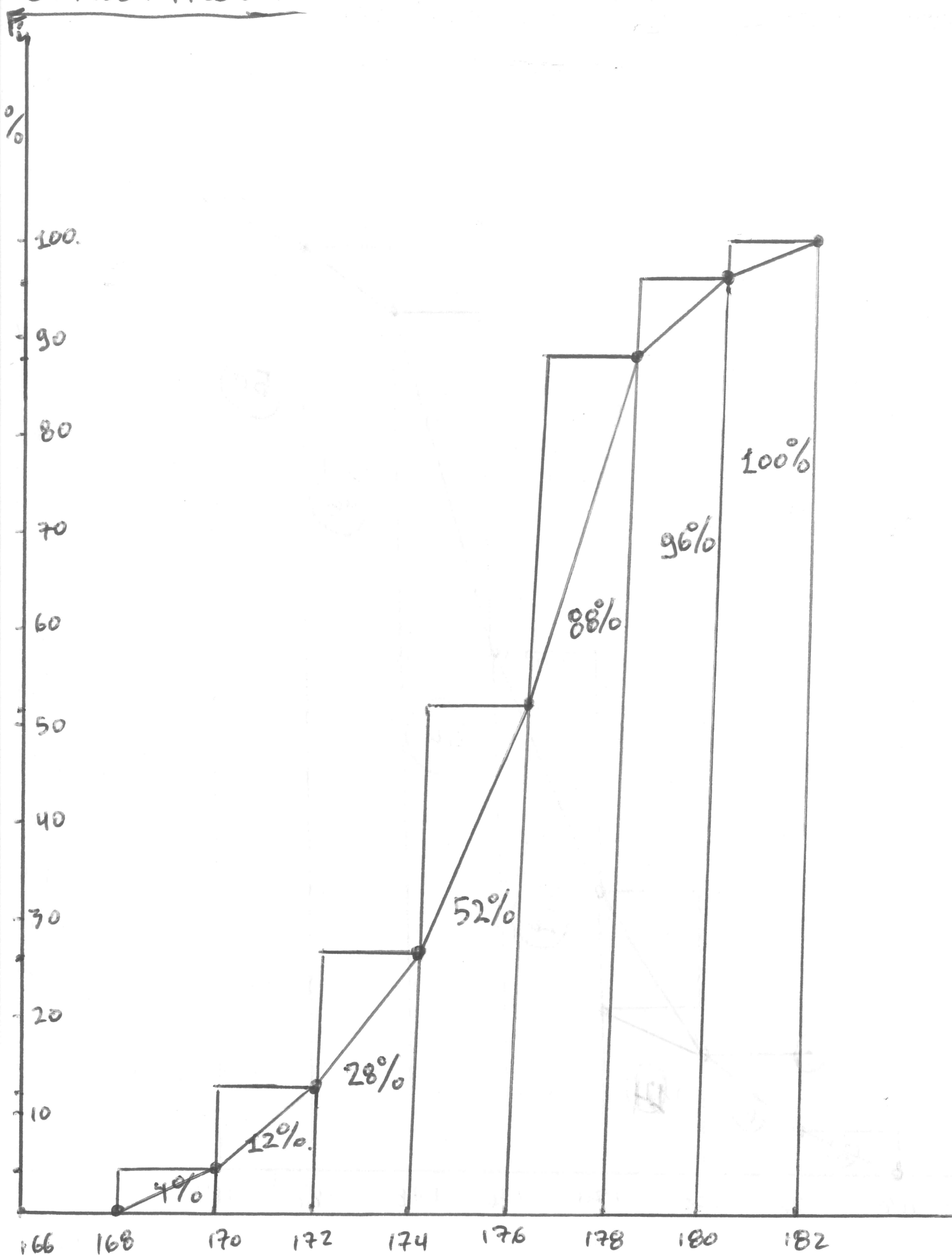


(η αρνητική ασυμμετρία μπορεί να επιβεβαιωθεί και από το πολλαπλό σχετικό συχνότητες).

Ιστογράμμο αθροιστικών συχνοτήτων (16)
και ποσοδωνο αθροιστικών συχνοτήτων



16. Το διάγραμμα σχετικών αθροιστικών συχνοτήτων (17)
τύπων και πολλαπλοο σχετικών αθροιστικών
συχνοτήτων



Διάγραμμα πλαισίου - απολύσεων

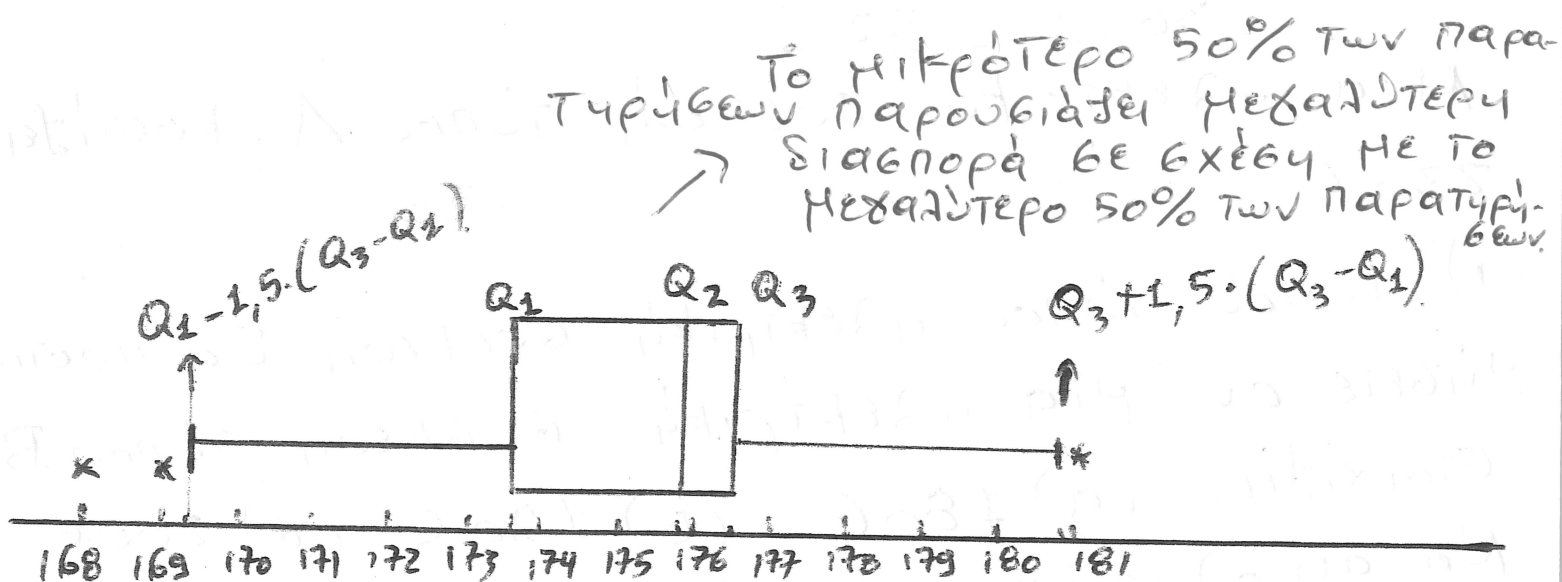
(18)

$$Q_1 - 1,5 \cdot (Q_3 - Q_1) = 173,63 - 1,5 \cdot (176,52 - 173,63) = 173,63 - 1,5 \cdot 2,89 \approx 173,63 - 4,34 = 169,29$$

$$Q_3 + 1,5 \cdot (Q_3 - Q_1) = 176,52 + 1,5 \cdot 2,89 \approx 176,52 + 4,34 = 180,86$$

Με βάση αυτά υπάρχουν οι εξής έκτρο-
πες τιμές: 168, 169, 181.

Το διάγραμμα είναι το εξής:



Το διάγραμμα είναι ενδεικτικό τόσο για
τη θέση των υψών, όσο και για την διασπορά
τους και την ασυμμετρία τους. Οι έκτρο-

πες τιμές συμβολίζονται με *. Η διαφορά των τιμών είναι σχετικά μικρή δεδομένου ότι υπάρχουν λίγες έκτροπες παρατηρήσεις.

5) Ο παρακάτω πίνακας δίνει τη διάρκεια ζωής δύο τύπων ηλεκτρικών βυβκων A και B (σε χιλ. ώρες).

A	B
12	12
14	13
23	16
30	22
36	32

Μια ηλεκτρική βυβκω τύπου A κοστίζει 230 €.

- (i) Ποιού τύπου ηλεκτρική βυβκω θα προτιμήσετε αν μια ηλεκτρική βυβκω τύπου B στοιχίζει: (α) 180 €, (β) 190 €, (γ) 200 €. Να αιτιολογώετε σε κάθε περίπτωση την απάντησή σας.
- (ii) Ποιού τύπου ηλεκτρικές βυβκωές παραγάφουν μεγαλύτερη ομοιογένεια ως προς τη διάρκεια λειτουργίας τους;

1064 (i) $\bar{X}_A = \frac{12+14+23+30+36}{5} = \frac{115}{5} =$

23.000 ώρες. $\bar{X}_B = \frac{12+13+16+22+32}{5} = 19.000$

ώρες. Άρα ο μέσος χρόνος λειτουργίας
βυβκεών τύπου A είναι 23.000 ώρες και

ο μέσος χρόνος λειτουργίας βυβκεών τύπου
B είναι 19.000 ώρες. Υπολογίζουμε τώρα

τη μέση αξία ανά ώρα λειτουργίας για
κάθε ^{τύπο} βυβκεύς. Υποθέτουμε ότι :

(α) μια βυβκεύ τύπου B βτοιχίζει 180€.
Τότε $V_{av}^A = \frac{230}{\bar{X}_A} = \frac{230}{23.000} = \frac{1}{100} = 0,01 \text{ €/ώρα}$

ώρα. $V_{av}^B = \frac{180}{19.000} \approx 0,0095 \text{ €/ώρα}$

Παρατηρούμε ότι $V_{av}^B < V_{av}^A$ και συνεπώς
προτιμούμε βυβκεύ τύπου B (αφού θα έχουμε
πληρώσει λιγότερα για την ώρα λειτουργίας της).

(β) μια βυβκεύ τύπου B βτοιχίζει 190€.

Τότε $V_{av}^B = \frac{190}{19.000} = \frac{1}{100} = 0,01 \text{ €/ώρα}$

Επειδή $V_{av}^A = V_{av}^B$ μπορούμε να επιλέξουμε
όποιο τύπο βυβκεύς θέλουμε.

(δ). Μια βούκκη τύπου Β στοιχίζει 200 €.

(21)

Τότε $V_{av}^B = \frac{200}{19.000} \approx 0,0105 \text{ € / ώρα.}$

Επειδή $V_{av}^A < V_{av}^B$ θα προτιμήσουμε βούκκη τύπου Α.

(ii). Θα υπολογίσουμε τους βυντελεστές μεταβλητότητας για κάθε δείγμα.

$$s_A^2 = \frac{1}{4} \cdot (12^2 + 14^2 + 23^2 + 30^2 + 36^2 - 5 \cdot 23^2) =$$

$$\frac{1}{4} \cdot (144 + 196 + 529 + 900 + 1296 - 2645) =$$

$$= \frac{1}{4} \cdot 420 = 105 \Rightarrow \underline{s_A \approx 10,25 \text{ κιλ. ώρες}}$$

Άρα $CV_A = \frac{s_A}{|\bar{x}_A|} = \frac{s_A}{\bar{x}_A} = \frac{10,25}{23} \approx 0,446.$

ή 44,6 %. Ομοίως $s_B^2 = \frac{1}{4} \cdot (12^2 + 13^2 + 16^2 + 22^2 + 32^2 - 5 \cdot 19^2) = \frac{1}{4} \cdot (144 + 169 + 256 + 484 + 1024 - 1805) = \frac{1}{4} \cdot 272 = 68 \Rightarrow \underline{s_B = \sqrt{68} \approx 8,25}$

κιλ. ώρες. Άρα $CV_B = \frac{s_B}{|\bar{x}_B|} = \frac{s_B}{\bar{x}_B} = \frac{8,25}{19} \approx$

0,434 ή 43,4 %. Επειδή $CV_B < CV_A$ οι τιμές

Του δείγματος B παρουσιάζουν μικρότερη
ελαστική μεταβλητότητα περί τον δείγματικό
τους μέσο όρο. Μεγαλύτερη ομοιογένεια σε
έχει με τις τιμές του δείγματος A.

Συμείωση | Η έννοια του μέσου κόστους/αξίας
ή της μέσης απόδοσης ^(π.χ. μιας επένδυσης) $\sqrt{\text{βτή μονάδα}}$ του χρόνου
(π.χ. ημέρα, μήνας, έτος κ.λπ.) έχει πολλές
εφαρμογές στην οικονομία/ χρηματοοικονομία.
