



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ
DEMOCRITUS UNIVERSITY OF THRACE

Δημοκρίτειο Πανεπιστήμιο Θράκης

**Μεταπτυχιακό Πρόγραμμα:
Διοίκηση Τουριστικών Επιχειρήσεων
και
Οργανισμών για Στελέχη
(Executive MBA in Tourism)**

Dr. Efstathios Dimitriadis
Mathematic
Ph.D in Applied Statistics
M.Sc in Statistics and Demography
M.Sc in Quality Assurance

Καβάλα, 2024

Ανάλυση Δεδομένων- Αποτελέσματα

- **Βασικά Στατιστικά Μέτρα**
- **Ανάλυση Συσχέτισης**
- **Ανάλυση Παλινδρόμησης**
- **χ^2 έλεγχος Ανεξαρτησίας**
- **Σύγκριση Μέσων Τιμών (ANOVA, t-test)**
- **Μοντέλα Δομικών Εξισώσεων**

Μεταβλητή Variable

Το χαρακτηριστικό γνώρισμα ως προς το οποίο μελετάμε κάθε άτομο του δείγματος ή του πληθυσμού

Μεταβλητές

```
graph TD; A[Μεταβλητές] --> B[Ποσοτικές]; A --> C[Ποιοτικές];
```

Ποσοτικές

Βάρος

Ύψος

Εισόδημα

Απόδοση

Θερμοκρασία

Ποιοτικές

Φύλο

Θρήσκευμα

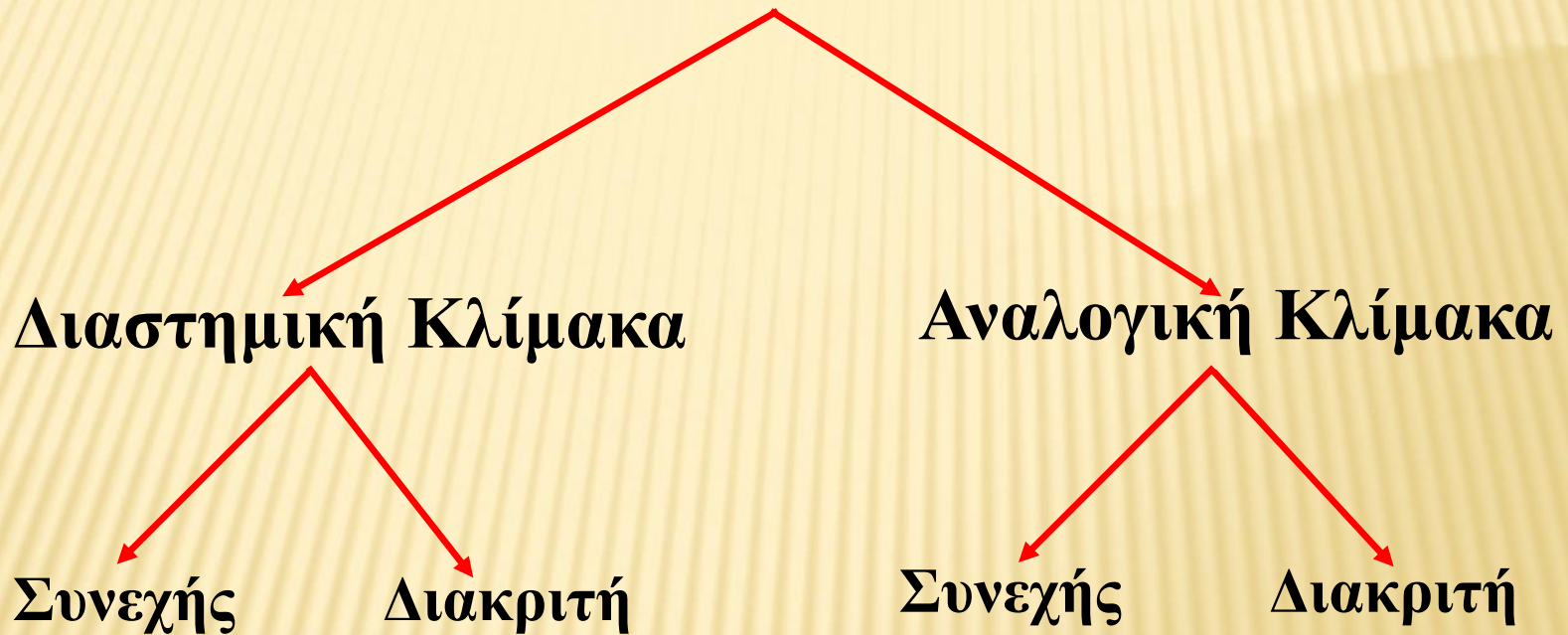
Εθνικότητα

Απόδοση

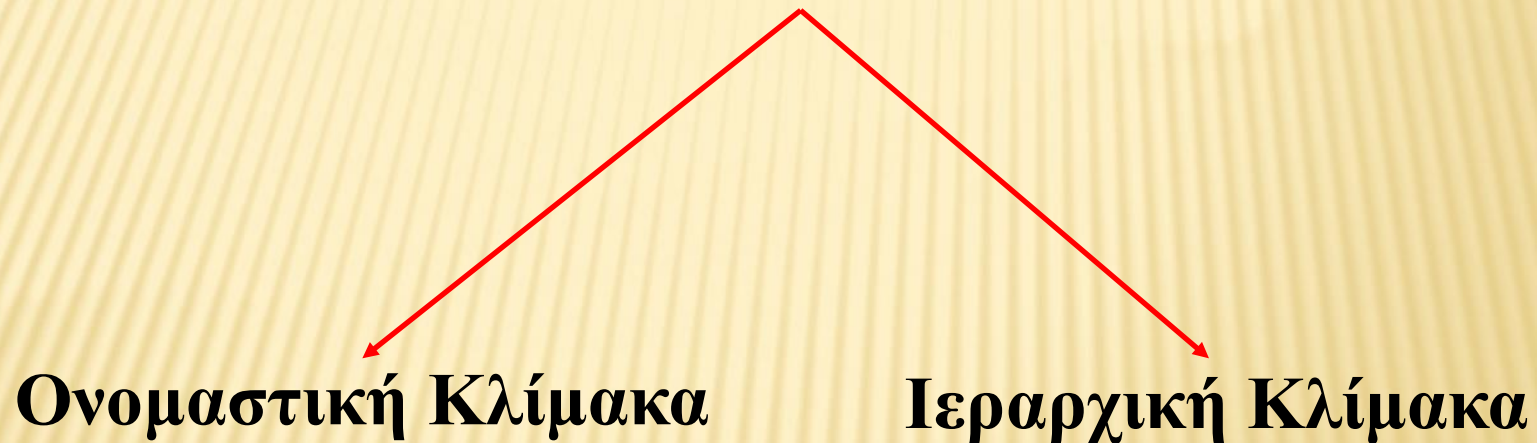
Άποψη

Επίπεδο Εκπαίδευσης

Ποσοτική Μεταβλητή



Ποιοτική Μεταβλητή



Βασικά Στατιστικά Μέτρα

Basic Statistical Measures

- ❑ **Μέτρα Κεντρικής Τάσης -Measures of Central Tendency**
- ❑ **Μέτρα Θέσης –Measures of Position**
- ❑ **Μέτρα Διασποράς -Measures of Dispersion**
- ❑ **Ασυμμετρία- Skewness**
- ❑ **Κύρτωση- Kurtosis**

Μέτρα Κεντρικής Τάσης - Measures of Central Tendency

- × **Αριθμητικός Μέσος - Arithmetic Mean**
- × **Γεωμετρικός Μέσος- Geometric Mean**
- × **Αρμονικός Μέσος- Harmonic Mean**
- × **Διάμεσος- Median**
- × **Σημείο Μέγιστης Συχνότητας -Mode**

- **Αριθμητικός Μέσος**
 - **Αστάθμητος (Αταξινόμητων Παρατηρήσεων)**

$$\bar{X} = \frac{\chi_1 + \chi_2 + \dots + \chi_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

\bar{X} : Αριθμητικός Μέσος

n : Πλήθος στοιχείων του δείγματος

X_i : Οι τιμές της μεταβλητής

!!!! Ο Αριθμητικός Μέσος δεν μπορεί να είναι μικρότερος από την μικρότερη τιμή των παρατηρήσεων ή μεγαλύτερος από τη μεγαλύτερη τιμή των παρατηρήσεων.

■ Αριθμητικός Μέσος

➤ Αστάθμητος (Ταξινομημένων και Ομαδοποιημένων Παρατηρήσεων)

$$\bar{X} = \frac{F_1X_1 + F_2X_2 + \dots + F_vX_v}{F_1 + F_2 + \dots + F_v} = \frac{\sum_{i=1}^v F_i X_i}{\sum_{i=1}^v F_i}$$

F_i : Συχνότητα (πλήθος παρατηρήσεων) κάθε τιμής

Αστάθμητος Ομαδοποιημένων Παρατηρήσεων

$$\bar{X} = X_0 + \frac{\delta \cdot \sum_{i=1}^v F_i \xi_i}{\sum_{i=1}^v F_i}$$

Επιλέγουμε όποιο διάστημα επιθυμούμε και στη στήλη ξ_i βάζουμε την τιμή 0.

X_0 : Το κέντρο του διαστήματος στο οποίο βάλουμε 0

δ : Το πλάτος του διαστήματος στο οποίο βάλουμε 0 και

$$\xi_i = \frac{\chi_i - \chi_0}{\delta}$$

Παράδειγμα Ταξινομημένων Παρατηρήσεων

<i>Αριθμός παιδιών X_i</i>	<i>Αριθμός Οικογενειών F_i</i>	<i>$X_i F_i$</i>
0	4	0*4=0
1	7	1*7=7
2	9	2*9=18
3	5	3*5=15
4	2	4*2=8
5	3	5*3=15
ΣΥΝΟΛΟ	30	63

$$\bar{X} = \frac{63}{30} = 2,1 \cong 2$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Ετήσιες μικτές Αποδοχές (σε Euro)	Αριθμός στελεχών F_i	Κέντρο Ομάδας X_i	$F_i X_i$
35.000-45.000	2	40.000	80.000
45.000-55.000	2	50.000	100.000
55.000-65.000	7	60.000	420.000
65.000-75.000	13	70.000	910.000
75.000-85.000	11	80.000	880.000
85.000-95.000	11	90.000	990.000
95.000-105.000	4	100.000	400.000
Σύνολο	50		3.780.000

$$\bar{X} = 3.780.000/50 = 75.600 \text{ Euro ετησίως}$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Ετήσιες μικτές αποδοχές (σε Euro)	Αριθμός στελεχών F_i	Κέντρο Ομάδας X_i	ξ_i	$F_i \xi_i$
35.000-45.000	2	40.000	-3	-6
45.000-55.000	2	50.000	-2	-4
55.000-65.000	7	60.000	-1	-7
65.000-75.000	13	70.000	0	0
75.000-85.000	11	80.000	1	11
85.000-95000	11	90.000	2	22
95.000-105.000	4	100.000	3	12
Σύνολο	50			28

$$\bar{X} = 70.000 + (10.000 * 28) / 50 = 75.600$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Μηνιαίοι μισθοί 500 υπαλλήλων συγκεκριμένης επιχείρησης

Μηνιαίοι μισθοί σε Euro X_i	Συχνότητα F_i	Κέντρο ομάδας	ξ_i	$F_i \xi_i$
500-650	50	575	-2	-100
650-800	100	725	-1	-100
800-950	250	875	0	0
950-1100	80	1025	1	80
1100-1250	20	1175	2	40
Σύνολο	500			-80

$$\bar{X} = X_0 + \frac{\delta \cdot \sum_{i=1}^v F_i \xi_i}{\sum_{i=1}^v F_i} = 875 + [150(-80)/500] = 851$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Ομάδες ημερομισθίων X_i (σε \$)	Αριθμός υπαλλήλων F_i	Κέντρο ομάδας X_i	ξ_i	$F_i \xi_i$
15-35	338	25	0	0
35-55	79	45	1	79
55-75	40	65	2	80
75-95	11	85	3	33
95-115	5	105	4	20
115-135	1	125	5	5
Σύνολο	474			217

$$\bar{X} = 25 + [(20 * 217) / 474] = 34,156 \$$$

■ Αριθμητικός Μέσος

➤ Σταθμικός (Ταξινομημένων και Ομαδοποιημένων Παρατηρήσεων)

$$\bar{X} = \frac{\chi_1 c_1 + \chi_2 c_2 + \dots + \chi_v c_v}{c_1 + c_2 + \dots + c_v} = \frac{\sum_{i=1}^v X_i c_i}{\sum_{i=1}^v c_i}$$

C_i : Συντελεστής βαρύτητας.

Παράδειγμα Σταθμικού Αριθμητικού Μέσου

<i>Μάθημα</i>	<i>Βαθμός X_i</i>	<i>Συντελεστής C_i</i>	<i>$X_i C_i$</i>
Μαθηματικά	13,8	1,15	15,870
Έκθεση	15,5	0,95	14,725
Ιστορία	18,0	0,95	17,100
Πολιτική Οικονομία	17,5	0,95	16,625
ΣΥΝΟΛΟ	64,8	4,00	64,320

$$\bar{X} = \frac{64,32}{4} = 16,08$$

$$\bar{X} = \frac{64,8}{4} = 16,2$$

■ Αριθμητικός Μέσος

Δύο ή περισσότερων κατανομών

$$\bar{X} = \frac{\bar{X}_1 N_1 + \bar{X}_2 N_2 + \dots + \bar{X}_k N_k}{N_1 + N_2 + \dots + N_k}$$

$\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$ οι μέσοι των k κατανομών και

N_1, N_2, \dots, N_k οι αντίστοιχες συχνότητες αυτών

Παράδειγμα Αριθμητικού Μέσου Δυο ή Περισσοτέρων Κατανομών

Κατηγορία Εργαζομένων	Μέσος Μισθός Κατηγορίας	Πλήθος ατόμων ανά κατηγορία	$\bar{X}_i N_i$
Κατώτερα Στελέχη	1.100	363	399.300
Μεσαία Στελέχη	1.700	84	142.800
Ανώτερα Στελέχη	2.300	27	62.100
ΣΥΝΟΛΟ	5.100	474	604.200

$$\bar{X} = \frac{604.200}{474} = 1.274,7 \quad \bar{X} = \frac{5.100}{3} = 1.700$$

Διάμεσος

$$M_{1/2} = X_i + \frac{\delta}{F_i} \left(\frac{N}{2} - \Phi_i \right)$$

$\frac{N}{2}$ = **Θέση Διαμέσου**. Την προσδιορίζουμε στη στήλη της αθροιστικής συχνότητας Φ_i μεταξύ της προηγούμενης και της επόμενης και τραβάμε μια γραμμή.

X_i = Η μικρότερη τιμή του διαστήματος κάτω από τη γραμμή.

δ = Το εύρος του διαστήματος κάτω από τη γραμμή.

F_i = Η συχνότητα του διαστήματος κάτω από τη γραμμή

Φ_i = Η αθροιστική συχνότητα πάνω από τη γραμμή

!!!! Το 50% των παρατηρήσεων έχει τιμές το πολύ μέχρι την τιμή της διαμέσου και το υπόλοιπο 50% τιμές μεγαλύτερες
!!! Η τιμή της διαμέσου πρέπει να βρίσκεται στο διάστημα κάτω από τη γραμμή

Παράδειγμα:

Ετήσιες μικτές αποδοχές (σε Euro)	Αριθμός στελεχών F_i	Φ_i
35.000-45.000	2	2
45.000-55.000	2	4
55.000-65.000	7	11
65.000-75.000	13	24
75.000-85.000	11	35
85.000-95000	11	46
95.000-105.000	4	50
Σύνολο	50	

$$\Theta \acute{\epsilon}\sigma\eta = 50/2 = 25^{\eta}$$

$$M_{1/2} = X_i + \frac{\delta}{F_i} \left(\frac{N}{2} - \Phi_i \right) =$$

$$= 75.000 + \frac{10.000}{11} (25 - 24) =$$

$$= 75.909$$

25^η

Σημείο Μέγιστης Συχνότητας ή Επικρατούσα Τιμή

$$M_0 = X_i + \frac{\delta \cdot \Delta_1}{\Delta_1 + \Delta_2}$$

X_i . = Η μικρότερη τιμή του διαστήματος με τη μεγαλύτερη συχνότητα.

δ = Το εύρος του διαστήματος με τη μεγαλύτερη συχνότητα.

Δ_1 = Μέγιστη- Προηγούμενη συχνότητα

Δ_2 = Μέγιστη- Επόμενη συχνότητα

!!!! Η επικρατούσα τιμή πρέπει να βρίσκεται οπωσδήποτε στο διάστημα με τη μεγαλύτερη συχνότητα

Παράδειγμα:

Ετήσιες μικτές αποδοχές (σε Euro)	Αριθμός στελεχών F_i	Φ_i
35.000-45.000	2	2
45.000-55.000	2	4
55.000-65.000	7	11
65.000-75.000	13	24
75.000-85.000	11	35
85.000-95000	11	46
95.000-105.000	4	50
Σύνολο	50	

$$M_0 = X_i + \frac{\delta \cdot \Delta_1}{\Delta_1 + \Delta_2}$$

$$M_0 = 65.000 + [(10.000 \cdot 6) / (6 + 2)] = 72.500 \text{€}$$

Μέτρα Θέσης

- **Τεταρτημόρια -Quartiles**
- **Δεκατημόρια**
- **Εκατοστημόρια- Percentiles**

Μέτρα Θέσης

Τεταρτημόρια

Πρώτο:
$$M_{1/4} = X_i + \frac{\delta}{F_i} \left(\frac{N}{4} - \Phi_i \right)$$

Τρίτο:
$$M_{3/4} = X_i + \frac{\delta}{F_i} \left(\frac{3N}{4} - \Phi_i \right)$$

Μέτρα Θέσης

Δεκατημόρια

$$M_{j/10} = X_i + \frac{\delta}{F_i} \left(\frac{j \cdot N}{10} - \Phi_i \right) \quad j = 1, \dots, 9$$

Εκατοστημόρια

$$M_{j/100} = X_i + \frac{\delta}{F_i} \left(\frac{j \cdot N}{100} - \Phi_i \right) \quad j = 1, \dots, 99$$

Γενικός Τύπος

$$M = X_i + \frac{\delta}{F_i} (a - \Phi_i)$$

Μέτρα Διασποράς

× Εύρος: $R = Max - Min$

× Ενδοτεταρτομοριακό Εύρος: $Q = M_{3/4} - M_{1/4}$

× Μέση Απόλυτη Απόκλιση: $M.A.A = \frac{\sum |X_i - \bar{X}|}{N}$

× Συντελεστής Μεταβλητότητας: $C.V = \frac{S}{\bar{X}} * 100$

Διακύμανση Αταξινόμητων παρατηρήσεων:

$$V = S^2 = \frac{\sum (X_i - \bar{X})^2}{N} \quad S^2 = \frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2$$

Διακύμανση Ταξινομημένων ή και ομαδοποιημένων παρατηρήσεων:

$$S^2 = \frac{\sum_{i=1}^N F_i \cdot (X_i - \bar{X})^2}{N} \quad S^2 = \frac{\sum_{i=1}^v F_i X_i^2}{\sum_{i=1}^v F_i} - \bar{X}^2 \quad S^2 = \delta^2 \cdot \left[\frac{\sum_{i=1}^v F_i \xi_i^2}{\sum_{i=1}^v F_i} - \left(\frac{\sum_{i=1}^v F_i \xi_i}{\sum_{i=1}^v F_i} \right)^2 \right]$$

Τυπική Απόκλιση Αταξινόμητων παρατηρήσεων:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N}} \quad s = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$

Τυπική Απόκλιση Ταξινομημένων ή και ομαδοποιημένων παρατηρήσεων:

$$S = \sqrt{\frac{\sum_{i=1}^N F_i \cdot (X_i - \bar{X})^2}{N}} \quad S = \sqrt{\frac{\sum_{i=1}^v F_i X_i^2}{\sum_{i=1}^v F_i} - \bar{X}^2} \quad S = \delta^* \cdot \sqrt{\frac{\sum_{i=1}^v F_i \xi_i^2}{\sum_{i=1}^v F_i} - \left(\frac{\sum_{i=1}^v F_i \xi_i}{\sum_{i=1}^v F_i} \right)^2}$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Μηνιαίοι μισθοί σε Euro X_i	Συχνότητα F_i	Κέντρο ομάδας	ξ_i	$F_i \xi_i$	$F_i \xi_i^2$
500-650	50	575	-2	-100	200
650-800	100	725	-1	-100	100
800-950	250	875	0	0	0
950-1100	80	1025	1	80	80
1100-1250	20	1175	2	40	80
Σύνολο	500			-80	460

Μέση Τιμή=851

$$S = \delta^* \sqrt{\frac{\sum_{i=1}^v F_i \xi_i^2}{\sum F_i} - \left(\frac{\sum_{i=1}^v F_i \xi_i}{\sum F_i} \right)^2}$$

$$S = 150^* \sqrt{\frac{460}{500} - \left(\frac{-80}{500} \right)^2} = 141.859$$

Παράδειγμα Ομαδοποιημένων Παρατηρήσεων

Μηνιαίοι μισθοί σε Euro X_i	Συχνότητα F_i	Κέντρο ομάδας	$F_i X_i$	$F_i X_i^2$
500-650	50	575	28750	16531250
650-800	100	725	72500	52562500
800-950	250	875	218750	191406250
950-1100	80	1025	82000	84050000
1100-1250	20	1175	23500	27612500
Σύνολο	500		425500	372162500

Μέση Τιμή = **851**

$$S = \sqrt{\frac{\sum_{i=1}^v F_i X_i^2}{\sum_{i=1}^v F_i} - \bar{X}^2} \quad S = \sqrt{\frac{372162500}{500} - 851^2} = 141.859$$

Ασυμμετρία

$$S_k = \frac{\bar{X} - M_0}{S} \quad S_k = \frac{3(\bar{X} - M_{1/2})}{S}$$

Αν $S_k=0$ η κατανομή είναι συμμετρική

Αν $S_k<0$ η κατανομή είναι αρνητικά ασύμμετρη

Αν $S_k>0$ η κατανομή είναι θετικά ασύμμετρη

Κύρτωση

$$b_2 = \frac{m_4}{S^4} \quad m_4 = \frac{\sum F_i (X_i - \bar{X})^4}{N}$$

Αν $b_2=3$ η κατανομή είναι μεσόκυρτη

Αν $b_2<3$ η κατανομή είναι πλατύκυρτη

Αν $b_2>3$ η κατανομή είναι λεπτόκυρτη

Μηνιαίοι μισθοί σε Euro X_i	Συχνότητα F_i	Κέντρο ομάδας	ξ_i	$F_i \xi_i$	$F_i \xi_i^2$	Φ_i
500-650	50	575	-2	-100	200	50
650-800	100	725	-1	-100	100	150
800-950	250	875	0	0	0	400
950-1100	80	1025	1	80	80	480
1100-1250	20	1175	2	40	80	500
Σύνολο	500			-80	460	

Παράδειγμα:

Ο πίνακας παρουσιάζει την κατανομή των μηνιαίων μισθών των 500 υπαλλήλων μιας επιχείρησης, σε €

Να υπολογιστούν:

1. Ο μέσος μηνιαίος μισθός
2. Η τυπική απόκλιση των μισθών
3. Μέχρι ποιο ύψος ανέρχονται οι μισθοί του 50% των υπαλλήλων;
4. Ποιος είναι ο συνηθέστερος μισθός;
5. Υπάρχει ομοιογένεια στην κατανομή των μισθών;
6. Ποιο είναι το ύψος των απολαβών του 5% των πλέον υψηλόμισθων υπαλλήλων;
7. Πόσοι υπάλληλοι έχουν μισθό μεταξύ 680€ και 1150€;
8. Ποιος είναι ο 300^{ος} σε τάξη μεγέθους μισθός;
9. Να μελετηθεί ως προς την ασυμμετρία η κατανομή των μισθών

1. Μέση Τιμή = 851€

2. $S = 141,859€$

3. $M_{1/2}$: Θέση = $500/2 = 250^{\text{η}}$

$$M_{1/2} = 800 + (150/250) * (250 - 150) = 860$$

4. M_0 : Μέγιστη συχνότητα = 250

$$M_0 = 800 + [150 * (250 - 100)] / [(250 - 100) + (250 - 80)] = 870$$

5. Συντελεστής Μεταβλητότητας

$$C.V = \frac{S}{\bar{X}} * 100 = (141,859/851) * 100 = 16,7\%$$

Μηνιαίοι μισθοί σε Euro X_i	Συχνότητα F_i	Κέντρο ομάδας	ξ_i	$F_i \xi_i$	Φ_i
500-650	50	575	-2	-100	50
650-800	100	725	-1	-100	150
800-950	250	875	0	0	400
950-1100	80	1025	1	80	480
1100-1250	20	1175	2	40	500
Σύνολο	500			-80	

Παράδειγμα:

Ο πίνακας παρουσιάζει την κατανομή των μηνιαίων μισθών των 500 υπαλλήλων μιας επιχείρησης, σε €

Να υπολογιστούν:

6. Ποιο είναι το ύψος των απολαβών του 5% των πλέον υψηλόμισθων υπαλλήλων;
7. Πόσοι υπάλληλοι έχουν μισθό μεταξύ 680€ και 1150€;
8. Ποιος είναι ο 300^{ος} σε τάξη μεγέθους μισθός;
9. Να μελετηθεί ως προς την ασυμμετρία η κατανομή των μισθών

$$9. S_k = \frac{3(\bar{X} - M_{1/2})}{S}$$

$$S_k = 3(851 - 860) / 141,859 = -0,19$$

Ελαφρά αρνητική ασυμμετρία

6. Βρίσκουμε πρώτα τη θέση του 95^{ου} εκατοστημορίου.

Θέση = $(95 * 500) / 100 = 475^{\eta}$ και στη συνέχεια υπολογίζουμε την τιμή του.

$M_{95/100} = 950 + (150/80) * (475 - 400) = 1090,625$ Συνεπώς το ύψος των απολαβών του 5% των πλέον υψηλόμισθων είναι μεγαλύτερο των **1090,625€**

7. Θα βρούμε πρώτα σε ποιες θέσεις αντιστοιχούν οι τιμές 680 και 1150€

$$680 = 650 + (150/100) * (\alpha_1 - 50) \quad \alpha_1 = 70^{\eta} \text{ θέση}$$

$$1150 = 1100 + (150/20) * (\alpha_2 - 480) \quad \alpha_2 = 487^{\eta} \text{ θέση. Συνεπώς: } 487 - 70 = 417 \text{ υπάλληλοι}$$

8. Η θέση είναι μεταξύ 150^{ης} και 400^{ης}. Συνεπώς $M = 800 + (150/250) * (300 - 150) = 890€$