

ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

ΔΙΑΛΕΞΗ: ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ II
ΠΑΡΑΔΕΙΓΜΑ

Διδάσκουσα: Ε. Γάκη, Επίκ. Καθηγήτρια

Παράδειγμα

Ο υπεύθυνος ποιότητας ενός οργανισμού παροχής υπηρεσιών στα πλαίσια της πολιτικής βελτίωσης της ποιότητας υπηρεσιών θέλει να ελέγξει τον χρόνο που δαπανούν οι πελάτες για την διεκπεραίωση των υποθέσεών τους. Σε μία έρευνα καταγράφηκε ο χρόνος εξυπηρέτησης (σε λεπτά) ενός δείγματος 50 πελατών από την στιγμή που εισέρχονται στον οργανισμό μέχρι την διεκπεραίωση της αίτησής τους. Ο παρακάτω πίνακας περιγράφει τα δεδομένα αυτά.

25	31	38	28	38	41	69	23	24	51
31	35	32	57	36	31	33	42	27	24
37	36	35	23	27	35	31	32	28	26
37	47	31	25	31	33	54	69	34	29
35	33	38	59	68	44	33	29	32	65

Ερώτημα 1^ο

Να υπολογιστούν ο αριθμητικός μέσος, η επικρατούσα τιμή, η διάμεσος, το εύρος, το ενδοτεταρτημοριακό εύρος, η διασπορά των δεδομένων καθώς και ο συντελεστής ασυμμετρίας ο οποίος και να ερμηνευθεί.

Λύση

Ζητείται ο υπολογισμός στατιστικών μέτρων από αταξινόμητα δεδομένα. Θα χρησιμοποιηθούν οι αντίστοιχοι τύποι του Τυπολογίου.

Αριθμητικός Μέσος

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{1852}{50} = 37.04$$

Λύση

Διάμεσος

Για τον υπολογισμό της διαμέσου απαιτείται ταξινόμηση σε αύξουσα σειρά.

X1	23
X2	23
X3	24
X4	24
X5	25
X6	25
X7	26
X8	27
X9	27
X10	28
X11	28
X12	29
X13	29
X14	31
X15	31
X16	31
X17	31
X18	31
X19	31
X20	32
X21	32
X22	32
X23	33
X24	33
X25	33

X26	33
X27	34
X28	35
X29	35
X30	35
X31	35
X32	36
X33	36
X34	37
X35	37
X36	38
X37	38
X38	38
X39	41
X40	42
X41	44
X42	47
X43	51
X44	54
X45	57
X46	59
X47	65
X48	68
X49	69
X50	69

Λύση

Διάμεσος

Για τον υπολογισμό της διαμέσου απαιτείται ταξινόμηση σε αύξουσα σειρά.

Διάμεσος

Πλήθος Δεδομένων $n = 50$, άρτιος αριθμός

Άρα η διάμεσος δίνεται από τον τύπο $M = \frac{1}{2} \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right)$

Επομένως:

$$M = \frac{1}{2} (X_{26} + X_{25}) = \frac{1}{2} (33 + 33) = 33$$

Λύση

Επικρατούσα Τιμή

Ορίζεται ως η τιμή με τη μεγαλύτερη συχνότητα εμφάνισης. Στην προκειμένη περίπτωση, $T_0 = 31$

Εύρος

$$R = X_{\max} - X_{\min} = 69 - 23 = 46$$

Λύση

Ενδοτεταρτημοριακό εύρος

$$IR = Q_3 - Q_1$$

Θα πρέπει να υπολογιστούν τα Q_1 και Q_3

Παρατηρώ ότι για $i = 1$ και $n = 50$

$$\frac{i(n+1)}{4} = \frac{(50+1)}{4} = 12.75. \text{ Άρα } A_Q = 12 \text{ και } \Delta_Q = 0.75.$$

Επομένως,

$$\begin{aligned} Q_1 &= X_{\frac{(n+1)}{4}} = X_{A_Q} + \Delta_Q (X_{A_Q+1} - X_{A_Q}) \\ &= X_{12} + \Delta_Q (X_{13} - X_{12}) \\ &= 29 + 0.75(29 - 29) = 29 \end{aligned}$$

Λύση

Παρατηρώ ότι για $i = 3$ και $n = 50$

$$\frac{i(n+1)}{4} = \frac{3(50+1)}{4} = 38.25. \text{ Άρα } A_Q = 38 \quad \Delta_Q = 0.25.$$

Επομένως,

$$Q_3 = X_{\frac{3(n+1)}{4}}$$

$$= X_{A_Q} + \Delta_Q (X_{A_Q+1} - X_{A_Q})$$

$$= X_{38} + \Delta_Q (X_{39} - X_{38})$$

$$= 38 + 0.25(41 - 38) = 38.75$$

$$\text{Άρα } IR = Q_3 - Q_1 = 38.75 - 29 = 9.75$$

Λύση

Διακύμανση

Για τον υπολογισμό της διακύμανσης θα χρησιμοποιήσουμε τον αριθμητικό μέσο που υπολογίσαμε πιο πάνω

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{1852}{50} = 37.04$$

Η διασπορά είναι

$$S^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1} = \frac{75964 - 50 \times 37.04^2}{49} = \frac{7365.91}{49} = 150.32$$

Λύση

Συντελεστής Ασυμμετρίας:

$$\triangleright S_p = \frac{\bar{X} - T_0}{S} = \frac{37.04 - 31}{12.2607} = 0.4926$$

Ερώτημα 2^ο

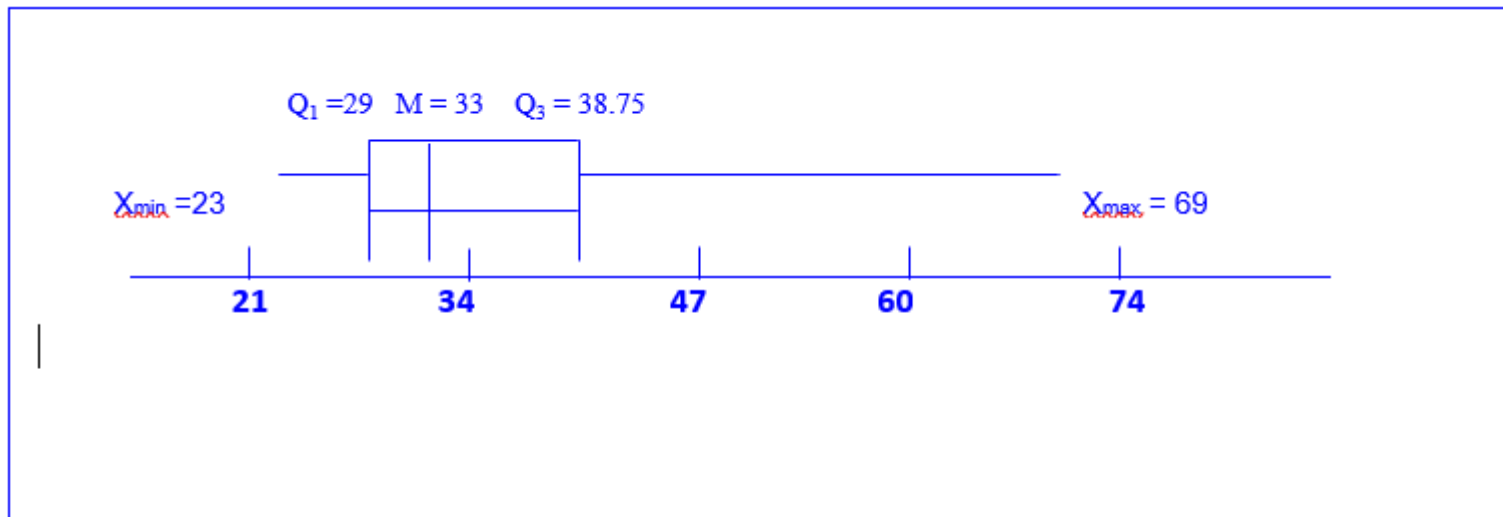
Να κατασκευασθεί το Θηκόγραμμα του χρόνου διεκπεραίωσης και με βάση αυτό να εξαχθούν συμπεράσματα για τη μορφή της κατανομής του.

Λύση

Για την κατασκευή του Θηκογράμματος απαιτούνται τα ακόλουθα στοιχεία τα οποία έχουν ήδη υπολογισθεί:

$$X_{\min} = 23 \quad X_{\max} = 69 \quad M = 33 \quad Q_1 = 29 \quad Q_3 = 38.75$$

Με βάση τα στοιχεία αυτά κατασκευάζουμε το Θηκόγραμμα των περιεκτικότητων.



Ερώτημα 3^ο

Να κατασκευασθεί Πίνακας Κατανομής Συχνοτήτων για το χρόνο διεκπεραίωσης χρησιμοποιώντας τάξεις εύρους 10, με κάτω όριο της πρώτης τάξης τα 20 λεπτά.

Λύση

Το κάτω όριο της πρώτης κλάσης – τάξης είναι 20, το άνω όριο της τελευταίας κλάσης είναι 70, και το εύρος των κλάσεων είναι 10. Επομένως ο Πίνακας κατανομής συχνοτήτων θα έχει την ακόλουθη μορφή:

Τάξη	Συχνότητα f_i
[20-30)	13
[30-40)	25
[40-50)	4
[50-60)	4
[60-70)	4
Σύνολο	50

Ερώτημα 4^ο

Χρησιμοποιώντας τα ταξινομημένα δεδομένα να υπολογιστούν ο αριθμητικός μέσος, η επικρατούσα τιμή, το τρίτο τεταρτημόριο και η διακύμανση του χρόνου διεκπεραίωσης.

Λύση

Ο πίνακας περιέχει ορισμένους βοηθητικούς υπολογισμούς

Τάξη	Κεντρική Τιμή Τάξης m_i	Συχνότητα f_i	$m_i f_i$	F_i
[20-30)	25	13	325	13
[30-40)	35	25	875	38
[40-50)	45	4	180	42
[50-60)	55	4	220	46
[60-70)	65	4	260	50
Σύνολο		50	1860	

Λύση

Αριθμητικός μέσος (\bar{X})

$$\bar{X} = \frac{\sum_{i=1}^5 f_i m_i}{\sum_{i=1}^5 f_i} = \frac{1860}{50} = 37.2 \Rightarrow \bar{X} = 37.2$$

Λύση

Επικρατούσα τιμή (T_0)

Εντοπισμός της θέσης του T_0 :

Η τάξη με τη μεγαλύτερη συχνότητα είναι η 2^η. Άρα η επικρατούσα τιμή βρίσκεται στην τάξη αυτή (διάστημα 30 – 40).

Υπολογισμός της τιμής του T_0 :

$$T_o = L_{T_o} + \delta \frac{\Delta_1}{\Delta_1 + \Delta_2} = 30 + 10 \frac{(25 - 13)}{(25 - 13) + (25 - 4)} = 30 + 10 \frac{12}{33} \Rightarrow T_o = 33.6364$$

Λύση

Τρίτο Τεταρτημόριο (Q_3)

$$\text{Εντοπισμός της θέσης του } Q_3 : \frac{n \times i}{4} = \frac{50 \times 3}{4} = 37.5$$

Άρα το τρίτο τεταρτημόριο Q_3 ανήκει στην 2^η τάξη (διάστημα 30-40)

Υπολογισμός της τιμής του Q_3 :

$$Q_3 = L_{Q_3} + \frac{\delta}{f_{Q_3}} \left(\frac{n \times i}{4} - F_{Q_3-1} \right) = 30 + \frac{10}{25} \left(\frac{50 \times 3}{4} - 13 \right) = 30 + \frac{10}{25} \times 24.5 = 39.8$$

Λύση

Διακύμανση

$$S^2 = \frac{\sum_{i=1}^k f_i (m_i - \bar{X})^2}{\sum_{i=1}^n f_i - 1} = \frac{\sum_{i=1}^5 f_i (m_i - 37.2)^2}{49} = \frac{6658}{49} = 135.877$$

Ερώτημα 5^ο

Χρησιμοποιώντας ένα μοντέλο βελτίωσης της ποιότητας των παρεχόμενων υπηρεσιών ο υπεύθυνος του οργανισμού πιστεύει ότι ο χρόνος εξυπηρέτησης του κάθε πελάτη μπορεί να μειωθεί κατά πέντε λεπτά μέσα σε χρονικό διάστημα έξι μηνών. Με βάση αυτό επανυπολογίζει ορισμένα στατιστικά μέτρα και βρίσκει ότι ενώ ο μέσος χρόνος διαπεραίωσης πράγματι μειώθηκε στα 32,04 λεπτά η τυπική απόκλιση παραμένει σταθερή. Σε ποια από τις δύο περιπτώσεις μπορεί να παρατηρηθεί μεγαλύτερη μεταβλητότητα στο χρόνο εξυπηρέτησης; Τι θα πρέπει να αποφασίσει ο υπεύθυνος ποιότητας;

Λύση

Θα υπολογίσουμε τους συντελεστές μεταβλητότητας στις δύο περιπτώσεις.

$$CV_X = \frac{S}{X} 100 = \frac{12.2607}{37.04} 100 = 33.101\%$$

και

$$CV_Y = \frac{S}{Y} 100 = \frac{12.2607}{32.04} 100 = 38.266\%$$

Παρατηρούμε ότι η μεταβλητότητα μετά την βελτίωση του χρόνου διεκπεραίωσης θα είναι μεγαλύτερη απ'ότι στην αρχική κατάσταση. Για το λόγο αυτό ο υπεύθυνος πρέπει να λάβει μέτρα για τη μείωση της διασποράς του χρόνου διεκπεραίωσης της αίτησης.

Που απευθύνομαι?????

Ελένη Γάκη,

Τηλ. 2271035161

Email: e.gaki@aegean.gr

Υλικό μαθήματος

<https://eclass.aegean.gr/>

