

Μέτρα θέσης

Τα μέτρα θέσης χρησιμοποιούνται για να περιγράψουν με τη χρήση ενός μόνο αριθμού το κέντρο της κατανομής των τιμών μίας τυχαίας μεταβλητής. Με απλά λόγια, τα μέτρα θέσης περιγράφουν γύρω από ποια τιμή συγκεντρώνονται οι τιμές μίας τυχαίας μεταβλητής που μετρήθηκε κατά την εκτέλεση μίας μελέτης. Δυστυχώς, δεν υπάρχει ένας μοναδικός τρόπος που να μας δίνει με τον καλύτερο τρόπο τη θέση των τιμών μίας μεταβλητής. Στη συνέχεια θα δούμε τρεις μεθόδους για τον υπολογισμό της θέσης: τη μέση τιμή, τη διάμεσο και την επικρατούσα τιμή.

Μέση τιμή

Η μέση τιμή για την κατανομή των τιμών μίας τυχαίας μεταβλητής υπολογίζεται ως μέσος όρος των τιμών της, δηλαδή αθροίζοντας όλες τις τιμές και διαιρώντας με το πλήθος τους:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Καθώς η μέση τιμή υπολογίζεται με τη χρήση $\sum_{i=1}^n x_i$ και με το n μπορούμε να την υπολογίσουμε ακόμα κι αν δεν έχουμε τις αρχικές μετρήσεις αλλά τον πίνακα συχνοτήτων των μετρήσεων για μία μεταβλητή. Για παράδειγμα, έστω ότι σε μία μελέτη ο/η ερευνητής/τρια θέλησε να αξιολογήσει τις κινητικές δεξιότητες των μαθητών μίας τάξης νηπιαγωγείου δίνοντάς τους 4 τουβλάκια και ζητώντας τους να τα τοποθετήσουν το ένα πάνω στο άλλο φτιάχνοντας έναν πύργο. Έστω X η μεταβλητή «το πλήθος από τουβλάκια που χρησιμοποίησε ο μαθητής για τον πύργο του». Ο πίνακας συχνοτήτων των μετρήσεων της μεταβλητής X φαίνεται στον παρακάτω πίνακα.

x	f
1	1
2	4
3	9
4	6

Για τον υπολογισμό της μέσης τιμής της X θα χρειαστούμε το άθροισμα

$$\sum X = \sum_{i=1}^4 f(x_i)x_i = 1 \times 1 + 4 \times 2 + 9 \times 3 + 6 \times 4 = 60$$

Επίσης θα χρειαστούμε το μέγεθος του δείγματος N το οποίο μπορούμε να βρούμε αθροίζοντας όλες τις συχνότητες

$$N = \sum_{i=1}^4 f(i) = 1 + 4 + 9 + 6 = 20$$

Επομένως, η μέση τιμή μ μπορεί να υπολογιστεί από

$$\mu = \frac{\sum X}{N} = \frac{\sum_{i=1}^4 f(x_i)x_i}{\sum_{i=1}^4 f(i)} = \frac{60}{20} = 3$$

Διάμεσος

Αν και η μέση τιμή είναι η πιο εύληπτη μέθοδος για τον υπολογισμό της θέσης των μετρήσεων μιας μεταβλητής, έχει ένα σημαντικό μειονέκτημα. Στην περίπτωση που υπάρχουν ακραίες τιμές η μέση τιμή μπορεί να επηρεαστεί τόσο πολύ, ώστε το αποτέλεσμα να μην είναι αντιπροσωπευτικό του δείγματος. Για παράδειγμα, έστω ότι μετράμε τη μέση ηλικία σε μια αίθουσα νηπιαγωγείου. Η αίθουσα έχει 4 μαθητές ηλικίας 5 ετών και έναν δάσκαλο ηλικίας 60 ετών. Τότε, η μέση ηλικία στην αίθουσα είναι

$$\mu = \frac{4 \times 5 + 60}{5} = \frac{80}{5} = 16$$

Στη συγκεκριμένη περίπτωση, η τιμή 16 δεν αντιπροσωπεύει κανέναν. Το νούμερο είναι πολύ μεγάλο για όλους τους μαθητές της αίθουσας και πολύ μικρό για τον δάσκαλο. Είναι σημαντικό να θυμόμαστε ότι σκοπός των μέτρων θέσης είναι να μας δώσουν την τιμή εκείνη γύρω από την οποία συγκεντρώνεται η πλειοψηφία των τιμών μίας μεταβλητής, ώστε να μπορέσουμε να αντιπροσωπεύσουμε όλες τις μετρήσεις με έναν αριθμό. Συμβαίνει αυτό στο συγκεκριμένο παράδειγμα;

Σε αυτές τις περιπτώσεις, αντί της μέσης τιμής μπορούμε να χρησιμοποιήσουμε τη διάμεσο, η οποία ορίζεται ως η μεσαία τιμή των μετρήσεων της μεταβλητής. Για τον υπολογισμό της, πρώτα αναδιατάσσουμε το δείγμα σε αύξουσα σειρά και στη συνέχεια υπολογίζουμε τη διάμεσο ως εξής:

- Αν το μέγεθος του δείγματος (n) είναι περιττός (μονός) αριθμός, η διάμεσος είναι η μεσαία τιμή. Η μεσαία τιμή είναι η $\frac{n+1}{2}$, για παράδειγμα, αν το δείγμα έχει μέγεθος 5, τότε η μεσαία τιμή είναι η $\frac{5+1}{2} = 3$.
- Αν το μέγεθος του δείγματος είναι άρτιος (ζυγός) αριθμός, τότε η διάμεσος υπολογίζεται ως ο μέσος όρος των δύο μεσαίων τιμών. Αν το δείγμα έχει μέγεθος n , τότε οι δύο μεσαίες τιμές είναι η $\frac{n}{2}$ και η επόμενη. Για παράδειγμα, αν το μέγεθος του δείγματος είναι 4, τότε η διάμεσος θα υπολογιστεί ως ο μέσος όρος της δεύτερης και της τρίτης τιμής.

Ποσοστημότητα

Ποσοστημότητα P_k είναι η τιμή εκείνη από την οποία το $k\%$ των μετρήσεων μιας μεταβλητής είναι μικρότερο. Για παράδειγμα, αν για τη μεταβλητή X (το ύψος των

μαθητών σε μία τάξη νηπιαγωγείου) βρούμε ότι το 75% ποσοστημόριο είναι 120cm ($P_{75} = 120$) τότε το 75% των μαθητών έχει ύψος μικρότερο από 120cm.

Παρατήρηση

Η διάμεσος είναι στην πραγματικότητα το 50%-ποσοστημόριο.

Για τον υπολογισμό του ποσοστημορίων ακολουθούμε την παρακάτω διαδικασία:

1. Τοποθετούμε τις μετρήσεις σε αύξουσα σειρά
2. Κατασκευάζουμε τον *πίνακα τάξεων*. Ο πίνακας αυτός έχει δύο στήλες: (1) κάθε τιμή του της μεταβλητής και (2) την τάξη της συγκεκριμένης τιμής.
3. Υπολογίζουμε την τάξη του ζητούμενου ποσοστημορίου από τον τύπο

$$r_k = \frac{k}{100}(n + 1)$$

4. Υπολογίζουμε την τιμή της μεταβλητής που αντιστοιχεί στο συγκεκριμένο ποσοστημόριο κάνοντας παρεμβολή αν η τάξη του ποσοστημορίου δεν είναι ακέραιος αριθμός (βλέπε παρακάτω).

Για παράδειγμα, έστω ότι παρατηρούμε τις τιμές για τα ύψη: 108, 107, 118, 100, 98, 103, 112, 109, 96, 101 cm.

Έστω ότι θέλουμε να βρούμε το 25%-ποσοστημόριο των μετρήσεων του ύψους. Τοποθετούμε τις μετρήσεις σε αύξουσα σειρά και κατασκευάζουμε τον πίνακα τάξεων:

Ύψος (cm)	Τάξη (R)
96	1
98	2
100	3
101	4
103	5
107	6
108	7
109	8
112	9
118	10

Υπολογίζουμε το $r(P_{25})$:

$$r(P_{25}) = \frac{25}{100}(10 + 1) = 2.75$$

Αφού η τάξη του ποσοστημορίου δεν είναι ακέραιος αριθμός και πέφτει μεταξύ 2 και 3 θα πρέπει να εξάγουμε την τιμή που αντιστοιχεί στη μεταβλητή κάνοντας παρεμβολή. Οι τιμές του X με τάξεις 2 και 3 είναι 98 και 100 αντίστοιχα. Η απόσταση μεταξύ των δύο αυτών τιμών είναι $100 - 98 = 2$. Για την παρεμβολή πολλαπλασιάζουμε το δεκαδικό μέρος του ποσοστημορίου (0.75) με το διαφορά (2) και προσθέτοντας στο μικρότερο (98) βρίσκουμε την τιμή του 25%-ποσοστημορίου ($0.75 \times 2 + 98 = 99.5$).

Τέλος, πολύ συχνά στις στατιστικές μας ενδιαφέρουν τα 25%,50%,75%-ποσοστημόρια. Τα ποσοστημόρια αυτά αποκαλούνται τεταρτημόρια. Για παράδειγμα το 25%-ποσοστημόριο είναι το πρώτο τεταρτημόριο, το 50%-ποσοστημόριο (διάμεσος) είναι το δεύτερο τεταρτημόριο κ.ο.κ.