

## Βασικές στατιστικές έννοιες

Η έρευνα στην εκπαίδευση κατά ένα πολύ σημαντικό μέρος της περιλαμβάνει τη συλλογή πληροφοριών. Συνήθως, με τη λήξη συλλογής της πληροφορίας οι ερευνητές/τριες καταλήγουν με έναν τεράστιο όγκο μετρήσεων είτε, παλιότερα, σε χαρτί είτε σήμερα σε ηλεκτρονική μορφή. Για την επεξεργασία και αξιοποίηση αυτής της πληροφορίας χρησιμοποιούνται μέθοδοι που συνοπτικά αποκαλούμε «στατιστική». Η στατιστική μπορεί να χρησιμοποιηθεί τόσο για την οργάνωση και τη συνόψιση της συγκεντρωμένης πληροφορίας όσο και για την συστηματική κατασκευή μίας ερμηνείας της κατάστασης από την οποία η πληροφορία αυτή προέκυψε.

## Πληθυσμός και δείγμα

Κατά την εκτέλεση μίας μελέτης ο/η ερευνητής/τρια ξεκινά διατυπώνοντας ένα γενικό ερώτημα για μία συγκεκριμένη ομάδα ανθρώπων που τον/την ενδιαφέρουν. Με στατιστικούς όρους, το σύνολο όλων των ατόμων ενδιαφέροντος μίας μελέτης ονομάζεται *πληθυσμός*. Για παράδειγμα, πληθυσμός σε μία εκπαιδευτική έρευνα θα μπορούσε να είναι όλοι οι μαθητές νηπιαγωγείων της Δυτικής Μακεδονίας.

Προφανώς, σε μία μελέτη είναι πολύ δύσκολο να συλλεχθούν οι απαιτούμενες πληροφορίες για το σύνολο του πληθυσμού ενδιαφέροντος, καθώς, τις περισσότερες φορές, αυτός μπορεί να είναι απαγορευτικά μεγάλος. Οπότε, η αμέσως καλύτερη προσέγγιση είναι να χρησιμοποιηθεί ένα υποσύνολο του πληθυσμού, με την ελπίδα ότι αυτό θα «μοιάζει» με τον πληθυσμό. Ένα τέτοιο υποσύνολο ονομάζεται *δείγμα*. Με στατιστικούς όρους, *δείγμα* ονομάζεται ένα σύνολο ατόμων το οποίο διαλέχθηκε από τον πληθυσμό με στόχο να τον αντιπροσωπεύσει σε μία ανάλυση.

Μία μελέτη εκπαιδευτικής έρευνας στοχεύει να περιγράψει ή να βγάλει συμπεράσματα για συγκεκριμένα χαρακτηριστικά του πληθυσμού. Τα χαρακτηριστικά αυτά πρέπει να είναι επακριβώς καθορισμένα και μετρήσιμα. Στη στατιστική, τα χαρακτηριστικά αυτά αποκαλούνται *παραμέτροι*. Επομένως, *παραμέτρος* είναι ένα χαρακτηριστικό του πληθυσμού, άγνωστο αλλά μετρήσιμο.

Καθώς η γνώση της τιμής μιας παραμέτρου συνήθως απαιτεί τη λήψη μετρήσεων από το σύνολο του πληθυσμού — πράγμα συνήθως αδύνατο — κατά την εκτέλεση στατιστικών αναλύσεων αρκούμαστε στην εξαγωγή εκτιμήσεων για αυτήν την τιμή με την αξιολόγηση του δείγματος. Η *εκτίμηση* της τιμής μιας παραμέτρου υπολογίζεται συνήθως με τον συνδυασμό των μετρήσεων του δείγματος (π.χ. άθροισμα, μέσος όρος). Με τη χρήση στατιστικής ορολογίας, ένας μαθηματικός συνδυασμός των μετρήσεων ενός δείγματος από τον πληθυσμό ενδιαφέροντος ονομάζεται *στατιστικό* ή *στατιστική συνάρτηση* ή *δειγματοσυνάρτηση*.

## Μεταβλητές

*Μεταβλητή* είναι οποιοδήποτε μετρήσιμο χαρακτηριστικό ατόμων ή πραγμάτων το οποίο μπορεί να παίρνει διαφορετικές τιμές. Σε αντίθεση, μία *σταθερά* είναι ένα χαρακτηριστικό το οποίο παίρνει μία και μοναδική τιμή σε ολόκληρο τον πληθυσμό. Οι σταθερές, αν και φαίνεται ότι δεν έχουν ιδιαίτερο ενδιαφέρον, στην πραγματικότητα αντικατοπτρίζουν τις αποφάσεις του ερευνητή/τριας για περιορισμό του πληθυσμού ενδιαφέροντος. Για παράδειγμα, σε μία μελέτη

με στόχο το σύνολο των μαθητριών στα νηπιαγωγεία της Δυτικής Μακεδονίας το φύλο είναι μία σταθερά.

Οι μεταβλητές μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες, τις *κατηγορικές μεταβλητές* και τις *αριθμητικές μεταβλητές*. Οι κατηγορικές μεταβλητές είναι ποιοτικές μεταβλητές και περιγράφουν κατηγορίες. Οι κατηγορικές μεταβλητές μπορούν να είναι είτε *ονομαστικές* είτε *διατάξιμες*.

Οι *ονομαστικές* μεταβλητές παίρνουν τιμές από ένα σύνολο κατηγοριών με διαφορετικά ονόματα. Οι μετρήσεις σε μια ονομαστική κατηγορική μεταβλητή επισημαίνουν κατηγοριοποιούν τις παρατηρήσεις, αλλά δεν κάνουν καμία ποσοτική διάκριση μεταξύ τους.

Αντίθετα, οι *διατάξιμες* μεταβλητές αναφέρονται σε ένα σύνολο κατηγοριών που είναι οργανωμένες σε μια ταξινομημένη ακολουθία. Οι μετρήσεις μία διατάξιμης κατηγορικής μεταβλητής κατατάσσουν τις παρατηρήσεις με βάση το μέγεθος ή την ένταση.

Τέλος, μία ειδική περίπτωση των κατηγορικών μεταβλητών με ιδιαίτερο ενδιαφέρον στις στατιστικές αναλύσεις είναι οι *διχοτομικές μεταβλητές*, δηλαδή οι μεταβλητές που μπορούν να πάρουν μόνο δύο πιθανές τιμές (ναι/όχι).

Οι αριθμητικές μεταβλητές είναι ποσοτικές μεταβλητές που μπορούν να μετρηθούν με αριθμούς. Μπορούν να χωριστούν περαιτέρω:

- *Διακριτές μεταβλητές*: Μεταβλητές που μπορούν να πάρουν μόνο συγκεκριμένες τιμές, π.χ. ο αριθμός των μαθητών σε μία τάξη δεν μπορεί να είναι αρνητικός ή δεκαδικός.
- *Συνεχείς μεταβλητές*: Μεταβλητές που μπορούν να πάρουν οποιαδήποτε τιμή μέσα σε ένα διάστημα με την προϋπόθεση ότι ένα αρκετά καλό μέσο μέτρησης χρησιμοποιείται, π.χ. η απόσταση του σπιτιού από το κοντινότερο νηπιαγωγείο.

## Μαθηματικοί συμβολισμοί

Συνήθως, στις στατιστικές αναλύσεις συμβολίζουμε τις τυχαίες μεταβλητές με κεφαλαία γράμματα του λατινικού αλφαβήτου. Για παράδειγμα, μπορούμε να συμβολίσουμε με  $X$  το ύψος των μαθητών των νηπιαγωγείων της περιφέρειας Δυτικής Μακεδονίας. Για συγκεκριμένες μετρήσεις της μεταβλητής  $X$  χρησιμοποιούμε το ίδιο *πεζό* γράμμα ( $x$ ). Για παράδειγμα, έστω ότι έχουμε ένα δείγμα μεγέθους  $n = 100$  από μαθητές των νηπιαγωγείων της περιφέρειας Δυτικής Μακεδονίας. Για κάθε έναν από αυτούς τους μαθητές, συμβολίζουμε την μέτρηση του ύψους με  $x_i, i = 1, \dots, 100$ .

Με τις μετρήσεις μίας μεταβλητής  $X$  συνήθως θέλουμε να κάνουμε πράξεις. Μία από τις πιο συχνές ανάγκες που πορκύπτουν κατά την εκτέλεση μίας στατιστικής ανάλυσης είναι να αθροίσουμε τις τιμές  $x_i$  που συλλέχθηκαν. Ο μαθηματικός συμβολισμός στην περίπτωση αυτή είναι το  $\sum$ . Στο προηγούμενο παράδειγμα, το άθροισμα όλων των μετρήσεων του ύψους των 100 μαθητών θα μπορούσε να συμβολιστεί με

$$\sum X = \sum_{i=1}^{100} x_i = x_1 + x_2 + \dots + x_{100}$$

Οι πράξεις αυτές μπορούν να γίνουν αρκετά πιο σύνθετες, π.χ. να αθροίσουμε τα τετράγωνα των μετρήσεων του ύψους ( $\sum X^2$ ) ή να αθροίσουμε το ύψος τους μειωμένο κατά μία μονάδα ( $\sum (X - 1)$ ).