

## ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΣΤΑΤΙΣΤΙΚΗ

Κάθε στιγμή της ζωής μας κάνουμε επιλογές. Αυτές οι επιλογές βασίζονται σε ελλιπή πληροφορία. Συνεπώς κάθε επιλογή μας γίνεται κάτω από συνθήκες αβεβαιότητας.

Το αντικείμενο της σύγχρονης στατιστικής είναι η συλλογή, παρουσίαση και ανάλυση – ερμηνεία παρατηρήσεων με σκοπό την εξαγωγή βásiμων συμπερασμάτων και την λήψη βέλτιστων αποφάσεων κάτω από συνθήκες αβεβαιότητας.

Μαζί με την θεωρία πιθανοτήτων η στατιστική σήμερα χρησιμοποιείται σχεδόν σε όλους τους κλάδους της ανθρώπινης δραστηριότητας και γνώσης όπως για παράδειγμα στην οικονομία, τη γενετική, τη ψυχολογία, τις επιχειρήσεις και τη δημόσια διοίκηση.

Οι παρατηρήσεις, συχνά αναφερόμαστε σε αυτές με την ονομασία **στατιστικές μεταβλητές**, αφορούν ένα ή περισσότερα χαρακτηριστικά μονάδων ενός συνόλου που ονομάζεται **πληθυσμός**. Στις περισσότερες περιπτώσεις η εξέταση όλων των μονάδων του πληθυσμού είναι δύσκολη, δαπανηρή ή ακόμα και αδύνατη. Ένας υποψήφιος π.χ. πριν από τις εκλογές είναι δύσκολο να εξετάσει όλους τους ψηφοφόρους για να προσδιορίσει τι αντίληψη έχουν για τις θέσεις του. Σε αυτές τις περιπτώσεις, ο ερευνητής αναγκάζεται να συλλέξει την απαιτούμενη πληροφορία από κάποια μικρή ομάδα του πληθυσμού, η οποία ονομάζεται **δείγμα**. Είναι προφανές ότι τα οποιαδήποτε συμπεράσματα στα οποία θα μας οδηγήσει η ανάλυση του δείγματος θα πρέπει να μπορούν να επεκταθούν σ' ολόκληρο τον πληθυσμό. Για να γίνει όμως αυτό θα πρέπει το δείγμα μας να είναι **αντιπροσωπευτικό** του πληθυσμού, και αυτό επιτυγχάνεται μόνο όταν όλες οι μονάδες του πληθυσμού έχουν την ίδια πιθανότητα να περιληφθούν στο δείγμα. Ένα τέτοιο δείγμα θα λέγεται **τυχαίο δείγμα**. Η επιλογή λανθασμένου δείγματος μπορεί να μας οδηγήσει σε μη αληθή συμπεράσματα και για τον λόγο αυτόν η σωστή εκλογή δείγματος είναι ένα από τα πιο σημαντικά στάδια της στατιστικής μας ανάλυσης.

Ενδεικτικό είναι το παράδειγμα των προεδρικών εκλογών στις ΗΠΑ το 1936, όπου υπήρξε πλήρης αποτυχία της πρόγνωσης των εκλογικών αποτελεσμάτων. Η ανάλυση που διενήργησε η εταιρεία “Literary Digest” στηρίχτηκε σε αρχικό δείγμα 10.000.000 ατόμων του πληθυσμού από λίστα τηλεφωνικών καταλόγων των διαφόρων πόλεων και από τηλεφωνικούς καταλόγους ιδιοκτητών αυτοκινήτων. Στο ερωτηματολόγιο που ταχυδρομήθηκε απάντησε περίπου το 1/4 των επιλεγέντων και από την ανάλυση που πραγματοποιήθηκε φάνηκε ότι το εκλογικό σώμα θα έδινε μεγάλη πλειοψηφία στους Ρεπουμπλικάνους με υποψήφιο πρόεδρο τον Landon. Η μέρα των εκλογών όμως επιφύλαξε μια δυσάρεστη έκπληξη στους στατιστικούς αναλυτές της δημοσκοπήσης. Ο Δημοκρατικός υποψήφιος Roosevelt εκλέχτηκε με την ιστορική πλειοψηφία 60.7%, σε αντίθεση με το 40.9% που είχε προβλέψει η δημοσκόπηση της “Literary Digest”. Το λάθος ασφαλώς της πρόβλεψης (πάνω από 19 ποσοστιαίες μονάδες) οφείλετο στο γεγονός ότι οι τηλεφωνικοί κατάλογοι εκείνη την εποχή δεν

περιλάμβαναν τις κατώτερες κοινωνικο-οικονομικές τάξεις, με αποτέλεσμα το δείγμα να μην ήταν αντιπροσωπευτικό του πληθυσμού.

Με την αντιμετώπιση του προβλήματος επιλογής κατάλληλου δείγματος ασχολείται ένας ιδιαίτερος κλάδος της Στατιστικής γνωστός ως “**Θεωρία και Τεχνικές Δειγματοληψίας**”.

Η στατιστική χωρίζεται σε τρεις μεγάλες ενότητες:

- **Την Θεωρία και Τεχνικές Δειγματοληψίας (Sampling Theory).** Αντιμετώπιση του προβλήματος επιλογής κατάλληλου δείγματος.

- **Την Περιγραφική Στατιστική (Descriptive Statistics)** : Χρησιμοποιείται για την ταξινόμηση και επεξεργασία των δεδομένων του δείγματος μας. Ο σκοπός μας είναι ο σχηματισμός μιας συνοπτικής παρουσίασης του δείγματος μας με πίνακες και γραφήματα, από την οποία μπορούν να προκύψουν εύκολα διάφορα συμπεράσματα πάνω στο δείγμα μας.

- **Την Στατιστική Συμπερασματολογία (Inferential Statistics):** Στην Στατιστική Συμπερασματολογία σκοπός μας είναι όχι μόνο να περιγράψουμε ότι παρατηρούμε και να εξάγουμε συμπεράσματα πάνω στο δείγμα μας, αλλά να επεκτείνουμε αυτά τα συμπεράσματα σ’ ολόκληρο τον πληθυσμό μας. Θέλουμε δηλαδή από το δείγμα να επιστρέψουμε πίσω στον πληθυσμό, γενικεύοντας με τον μικρότερο δυνατό κίνδυνο λάθους τα αποτελέσματα που προέκυψαν από το δείγμα μας, έτσι ώστε να εξάγουμε αποτελέσματα για όλο τον πληθυσμό.

Το παρακάτω σχήμα μας βοηθάει στην κατανόηση των στατιστικών μεθοδολογιών:

