

## Κεφάλαιο 1 Εισαγωγή

### 1.1. Τι είναι στατιστική

Στατιστική είναι η επιστήμη που ασχολείται με τη συστηματική:

- ⇒ συλλογή στοιχείων ή δεδομένων
- ⇒ οργάνωση και επεξεργασία,
- ⇒ παρουσίαση αριθμητικών δεδομένων ή στοιχείων και
- ⇒ την ανάλυσή τους

τα οποία προέρχονται από πολλές μετρήσεις ή παρατηρήσεις.

Οι παρατηρήσεις αυτές ή οι μετρήσεις αναφέρονται σε συγκεκριμένο γεγονός ή αντικείμενο. Ανάλογα με το αντικείμενο ή το γεγονός στο οποίο αναφέρονται τα αριθμητικά δεδομένα τα κατατάσσουμε σε:

- Οικονομικά
- Δημογραφικά
- Φυσικά
- Πολιτικά
- Άλλα

φαινόμενα.

Σκοπός πάντα είναι να καταλήξουμε σε συμπεράσματα για τη λήψη ορθών αποφάσεων.

Στην Επιστημονική γλώσσα η λέξη στατιστική έχει ευρύτερη σημασία, σημαίνει την επιστήμη που έχει σαν αντικείμενο όχι μόνο τη συγκέντρωση και παρουσίαση, αλλά και τη μελέτη και ανάλυση των παρατηρήσεων ή μετρήσεων που αναφέρονται σ' ένα συγκεκριμένο αντικείμενο ή γεγονός.

Έτσι η Στατιστική περιλαμβάνει τόσο τις μεθόδους συλλογής και επεξεργασίας στοιχείων, όσο και τις μεθόδους ανάλυσης και μελέτης τους, ανακαλύπτοντας έτσι σχέσεις που υπάρχουν ανάμεσα στα διάφορα φαινόμενα και διατυπώνοντας συμπεράσματα που είναι χρήσιμα για τη λήψη ορθών αποφάσεων.

Η Στατιστική χωρίζεται σε δυο κλάδους την περιγραφική και την επαγωγική ή Συμπερασματική Στατιστική.

- Η περιγραφική Στατιστική έχει σαν σκοπό τη γενική διερεύνηση, σύντομη παρουσίαση, περιγραφή και ερμηνεία του στατιστικού υλικού, καθώς και την εξαγωγή συμπερασμάτων τα οποία όμως αναφέρονται μονό σ' ένα τμήμα του στατιστικού πληθυσμού, χωρίς να γενικεύονται για το σύνολο των στοιχείων του πληθυσμού.

- Η επαγωγική ή Συμπερασματική Στατιστική προσφέρει μεθόδους με τις οποίες είναι δυνατόν από τα δεδομένα ενός αντιπροσωπευτικού δείγματος να εξαρθούν, επαγωγικά, γενικότερα συμπεράσματα που αφορούν το σύνολο των στοιχείων του πληθυσμού από το οποίο έχει λειφθεί το δείγμα.

## 1.2. Ιστορία της Στατιστικής

(Στατίζω = τοποθετώ, ταξινομώ, στήνω, συμπεραίνω, ίσταμαι = στέκομαι. Αναφέρονται από τον Σωκράτη στο έργο του Ξενοφώντος «Απομνημονεύματα και από τον Αριστοτέλη στην «Πολιτεία»).

(Status = κατάσταση, θέση, πολιτική θέση, κράτος).

Η στατιστική ως αυτοτελής επιστήμη εμφανίζεται από τον 17ο αιώνα, εποχή που για πρώτη φορά καθιερώθηκε στην Ανωτάτη Παιδεία.

Μέχρι τότε υπάρχουν αποδείξεις χρήσης στατιστικών μεθόδων από τους Αιγυπτίους 3.500 π.χ., τους Κινέζους 2.300 π.χ., τους Ασσύριους και άλλους που έκαναν απογραφές και συγκεντρώνουν δεδομένα για τον πληθυσμό, το εισόδημα, τις στρατιωτικές δυνάμεις, τη φορολογία καθώς και για της γεννήσεις, τους θανάτους και άλλες μεταβολές του πληθυσμού.

Την εποχή που γεννήθηκε ο Χριστός αναφέρεται ότι ο Ιωσήφ και η Μαρία πήγαν στην Βηθλεέμ για την Γενική Απογραφή του πληθυσμού από τους Ρωμαίους.

Η λέξη "Στατιστική" προέρχεται από την λατινική λέξη status (που σημαίνει κράτος) και δηλώνει αρχικά συλλογή στοιχείων για της κρατικές ανάγκες.

Στην αρχή λοιπόν ο ρόλος της στατιστικής ήταν καθαρά περιγραφικός και χρησιμοποιούσε μόνο στοιχειώδη αριθμητική.

Στην συνέχεια όμως αναπτύσσεται ένας νέος κλάδος: ο λογισμός των πιθανοτήτων ο οποίος προήλθε από τη μελέτη των τυχερών παιχνιδιών. Έτσι βοηθούμενη από αυτόν η Στατιστική ξεφεύγει από τον περιγραφικό χαρακτήρα, σημειώνεται αυξημένη χρήση των μαθηματικών και αναπτύσσεται η Επαγωγική Στατιστική ή Συμπερασματική Στατιστική.

Από τους θεμελιωτές του λογισμού των πιθανοτήτων αναφέρουμε τον Bernoulli ο οποίος στο βιβλίο του "Η τέχνη των προβλέψεων" διατυπώνει το περίφημο Νόμο των Μεγάλων Αριθμών.

Όμως τις βάσεις έχουν δώσει οι Γάλλοι μαθηματικοί Pascal και Fermat. Ο Βέλγος αστρονόμος Quetelet επεκτείνει την εφαρμογή της στατιστικής στη σπουδή των φυσικών, διανοητικών και ηθικών ιδιοτήτων του ανθρώπου και παίρνει την πρωτοβουλία για την σύγκληση του πρώτου διεθνούς συνεδρίου στατιστικής που έγινε στις Βρυξέλλες το 1853.

Αργότερα ο E. Galton εφαρμόζει την στατιστική στην βιολογία, προσπάθεια που συνεχίστηκε από τον Άγγλο μαθηματικό Pearson.

Στην συνέχεια ο Gallor πάει στις ΗΠΑ και ιδρύει κατ' την εντολή του Τσόρτσιλ εταιρεία δημοσκοπήσεων.

Στην διάρκεια του Δευτέρου Παγκοσμίου πολέμου η Αγγλία δίνει μεγάλη ώθηση στην ανάπτυξη της Στατιστικής με τις μελέτες για την αποκρυπτογράφηση.

Την ίδια περίοδο οι ΗΠΑ και η ΕΣΣΔ σε συνεργασία αναπτύσσουν τις μεθόδους Προσομοίωσης και Ελέγχου Ποιότητας και Διαδικασιών.

Τέλος στις αρχές της δεκαετίας του 1960 οι αμερικανοί επιστήμονες πάνε στην Ιαπωνία.

### 1.3.Χρησιμότητα και πεδία εφαρμογής της Στατιστικής

Η στατιστική είναι χρήσιμη και πολλές φορές απαραίτητη για πολλούς τομείς δράσης της ανθρωπίνης δραστηριότητας.

Η στατιστική είναι απαραίτητη στη διοίκηση γενικά όπου η λήψη ορθών αποφάσεων έχει μεγάλη σημασία για την πρόοδο ενός οργανισμού, μιας βιομηχανίας ή μιας επιχείρησης.

Γι αυτό και δεν υπάρχει σήμερα στις σύγχρονες επιχειρήσεις κανένας τομέας που να μην χρησιμοποιεί τις στατιστικές μεθόδους στη λήψη επιχειρηματικών αποφάσεων.

Επιγραμματικά θα λέγαμε ότι η Στατιστική εφαρμόζεται στη:

- Λήψη αποφάσεων για την Διοίκηση κράτους, Οργανισμού, βιομηχανίας ή επιχείρησης
- Δημογραφία για τη μελέτη των γεννήσεων – θανάτων, γαμηλιότητας, μετανάστευσης κλπ.
- Επιστήμες όπως Ιατρική, Φυσική, Γενετική, Αστυνομία, Βιολογία, Μετεωρολογία, Γεωργία, Βιομηχανία, μελετών των ανθρώπινων ιδεών και προθέσεων (σφυγμομετρήσεις / δημοσκοπήσεις), Θεωρεία Αποφάσεων, Έλεγχος Ποιότητας ...
- Οικονομικό Τομέα όπως Επίπεδο Τιμών, Εθνικό Εισόδημα, Νομισματική ισοτιμία, απασχόληση, παραγωγικότητα, ρυθμό ανάπτυξης, οικονομικοί δείκτες, Εθνικοί Πόροι...

Πολύ μεγάλη εφαρμογή βρίσκει στον οικονομικό τομέα, στην ιατρική, φυσική, γενετική, αστρονομία κλπ.

### 1.4. Το κράτος και η Στατιστική

Ένα καλά οργανωμένο κράτος οφείλει να γνωρίζει κάθε στιγμή

- τον πληθυσμό της χώρας,
- την κατανομή του πληθυσμού κατά φύλο, ηλικία, επάγγελμα κλπ
- καθώς και την κίνηση και την πιθανή εξέλιξη του
- τα οικονομικά φαινόμενα της χώρας (παράγωγή, εισαγωγές, εξαγωγές, κίνηση αγαθών κλπ)
- διοικητικά και κοινωνικά φαινόμενα της χώρας, διοίκηση, εργασία, δημοσία υγεία, πρόνοια, κοινωνικές ασφάλισεις, κατάρτιση τιμάρθμου κλπ.

Για τον σκοπό αυτό κάθε κράτος έχει μια στατιστική υπηρεσία η οποία συγκεντρώνει τα απαραίτητα στοιχεία και παρακολουθεί την εξέλιξη τους.

Μια τέτοια υπηρεσία πρέπει να είναι καλά οργανωμένη και να διαθέτει πλούσιο κεντρικό αρχείο στατιστικών στοιχείων από το οποίο θα παίρνει πληροφορίες κάθε ερευνητής ή διοικητικός παράγοντας του κράτους.

## 1.5. Βασικές έννοιες

- Χαρακτηριστική ιδιότητα: είναι το αντικείμενο της εκάστοτε Στατιστικής μελέτης και καλείται **στατιστική μεταβλητή**
- Πληθυσμός: είναι το σύνολο του οποίου τα στοιχεία (άτομα, ή αντικείμενα, ή φαινόμενα, ή κτλ) μελετώνται ως προς τουλάχιστο μία χαρακτηριστική τους ιδιότητα, δηλαδή ένα χαρακτηριστικό τους.
- Κάθε υποσύνολο του πληθυσμού αποτελεί ένα **δείγμα** (sample) από τον πληθυσμό.
- Στατιστικές μονάδες ή άτομα του πληθυσμού: είναι τα στοιχεία του συνόλου αυτού. Οποιαδήποτε πρόσωπα, πράγματα, γεγονός ή φαινόμενο που ανήκει στον ορισμένο πληθυσμό.
- Παρατηρήσεις: είναι οι διάφορες τιμές της χαρακτηριστικής ιδιότητας κάθε στατιστικής μονάδας.

Ονομάζεται λοιπόν, **στατιστική μεταβλητή** το χαρακτηριστικό ως προς το οποίο μελετώνται τα στοιχεία ενός πληθυσμού.

### Παράδειγμα:

Μελετούμε τη βαθμολογία στα Μαθηματικά των μαθητών μιας τάξης.

Μεταβλητή ή χαρακτηριστική ιδιότητα: η βαθμολογία των μαθητών στα Μαθηματικά

Πληθυσμός: όλοι οι μαθητές της τάξης αυτής

Στατιστική μονάδα: ο κάθε μαθητής ξεχωριστά

Παρατηρήσεις: οι κάθε ένας από τους βαθμούς. Οι παρατηρήσεις μας παίρνουν τιμές 10, 18, 17, 12,...

### Παράδειγμα:

Τροχαία ατυχήματα στην τάδε διασταύρωση.

Μεταβλητή ή χαρακτηριστικό: τα τροχαία ατυχήματα

Πληθυσμός: όλα τα αυτοκίνητα που κινούνται στην τάδε διασταύρωση

Στατιστική μονάδα: το κάθε ένα αυτοκίνητο που πέρασε από την τάδε διασταύρωση

Παρατηρήσεις: Ναι και Όχι ή 0 και 1 αν έγινε ή δεν έγινε ατύχημα.

### Άσκηση:

Το ποσό των χρημάτων που έχουν οι μαθητές της παρούσας τάξης πάνω τους.

## 1.6. Έννοια στατιστικής μεταβλητής – Διακρίσεις αυτής

Στατιστικές μεταβλητές είναι οι χαρακτηριστικές ιδιότητες των στατιστικών μονάδων ενός πληθυσμού με την μελέτη των οποίων ασχολείται η στατιστική.

### Παραδείγματα

Το σύνολο των φοιτητών μιας σχολής, το σύνολο των κατοίκων μιας χώρας, το σύνολο των εργαζομένων μιας επιχείρησης αποτελούν παραδείγματα πληθυσμών.

το ανάστημα, το επάγγελμα, το εισόδημα, η μόρφωση αποτελούν χαρακτηριστικές ιδιότητες των στατιστικών μονάδων που ονομάσαμε μεταβλητές.

## Διακρίσεις των μεταβλητών

- Α. Στις ποιοτικές ή κατηγορικές: που δεν επιδέχονται μέτρηση και οι τιμές τους εκφράζονται με λέξεις ή σύμβολα. Είναι εκείνες οι μεταβλητές που ταξινομούν τον πληθυσμό σε κατηγορίες και δεν εκφράζουν απαραίτητα κάτι μετρήσιμο.

**Παραδείγματα:** η οικογενειακή κατάσταση (π.χ. έγγαμος/η, ή άγαμος/η, ή χήρος/α), το επάγγελμα, η κατάσταση υγείας, οι προτιμήσεις σε ένα χρώμα, η διαγωγή κλπ.

Αυτές οι μεταβλητές αν επιδέχονται στις τιμές τους κάποια διάταξη ή ιεράρχηση λέγονται διατάξιμες μεταβλητές, ενώ αν δεν μπορεί να γίνει αυτό, λέγονται μη διατάξιμες.

- Β. Στις ποσοτικές: που επιδέχονται μέτρηση και διάταξη και οι τιμές τους είναι αριθμοί εκφρασμένοι σε συγκριμένες μονάδες.

**Παραδείγματα:** το βάρος, η ηλικία, το εισόδημα, η θερμοκρασία κλπ.

Οι ποσοτικές μεταβλητές διακρίνονται σε:

- Συνεχείς: ονομάζονται οι μεταβλητές εκείνες που μπορεί να πάρουν οποιαδήποτε τιμή ενός διαστήματος πραγματικών αριθμών  $[α, β]$

Παραδείγματα: το βάρος, το ύψος, η ταχύτητα ενός κινητού, η θερμοκρασία κλπ.

- Ασυνεχείς ή Διακριτές: ονομάζονται οι μεταβλητές εκείνες που παίρνουν μονό πεπερασμένο ή αριθμήσιμο πλήθος τιμών ή ακέραιες τιμές.

Παραδείγματα: ένδειξη ζαριού, αγόρια σε μια οικογένεια, ρίψεις νομίσματος μέχρι να έρθει «κεφαλή» κλπ.

Οι μεταβλητές συμβολίζονται με κεφαλαία γράμματα π.χ. X, Y, Z, W, ...

Τιμές της μεταβλητής είναι οι αριθμοί (π.χ. το ποσό των χρημάτων σε € ή το ύψος σε cm κλπ) ή άλλα σύμβολα και άλλες συμβολικές εκφράσεις (π.χ. ΝΑΙ – ΟΧΙ ή κόκκινο, κίτρινο, μπλε ή Κακή, Μέτρια, Καλή, Άριστη κλπ.) που αντιπροσωπεύουν τις διάφορες καταστάσεις.

Οι αντίστοιχες τιμές της μεταβλητής συμβολίζονται με:

Για τις ποσοτικές με μικρά π.χ.  $x_1, x_2, \dots, x_n$ , ενώ αν είναι ποιοτικές (ή κατηγορικές) εκφράζεται με τις λέξεις ή άλλα σύμβολα.

| ΜΕΤΑΒΛΗΤΕΣ                  | ΤΙΜΕΣ                              |
|-----------------------------|------------------------------------|
| Τα αγόρια σε μια οικογένεια | 0, 1, 2, ...                       |
| Ένδειξη ενός ζαριού         | 1,2,3,4,5,6                        |
| Ένδειξη ενός νομίσματος     | «κεφαλή» ή «γράμματα               |
| Φυλή ατόμων                 | λευκή, μαύρη, κίτρινη κλπ          |
| Ύψος ατόμων σε cm           | 160, 175, 147, ...                 |
| Ηλικία ατόμων               | 5,15,40,...                        |
| Υγεία ατόμων                | άριστη, καλή, μέτρια, κακή θάνατος |
| Φύλο ατόμων                 | άνδρας, γυναίκα                    |

## Κλίμακες Μέτρησης

Ευρέως χρησιμοποιούνται οι εξής τέσσερις κλίμακες: κατηγορίας, διάταξης, διαστήματος και αναλογίας. Οι δυο πρώτες κλίμακες μέτρησης αφορούν τις ποιοτικές μεταβλητές ενώ οι δυο τελευταίες τις ποσοτικές.

**Κατηγορίας** (nominal) ή **ονομαστικές** είναι οι μεταβλητές των οποίων το σύνολο των τιμών δεν έχει καμία ιδιότητα. Για τη μεταβλητή αυτή, μοναδική σημασία έχουν οι διαφορετικές τιμές (το πλήθος των κατηγοριών της) που μπορεί να πάρει. Η μοναδική σχέση που μπορεί να προσδιοριστεί μεταξύ των κατηγοριών αυτών είναι απλά η ύπαρξη διαφοράς. Παραδείγματα τέτοιων μεταβλητών είναι το φύλο, το επάγγελμα, η οικογενειακή κατάσταση, η φυλή κλπ.

**Διάταξης** (ordinal) είναι οι μεταβλητές που για το σύνολο τιμών τους μπορούμε να ορίσουμε μια σχέση διάταξης, δηλαδή να τοποθετηθούν στη σειρά. Η διάταξη μπορεί να είναι από τη μικρότερη τιμή προς τη μεγαλύτερη ή και αντίστροφα. Οι ίσες διαφορές μεταξύ των τιμών μιας τέτοιας μεταβλητής δεν συνεπάγονται και ίσες διαφορές για το χαρακτηριστικό που μετράει η μεταβλητή. Δηλαδή, δεν υπάρχει αντιστοίχιση σε υποδιαίρεσεις ή πολλαπλάσια κάποιας μονάδας. Η διάταξη δηλαδή, το μόνο που εξασφαλίζει είναι τον προσδιορισμό της μεγαλύτερης, καλύτερης, προτιμότερης κατηγορίας αλλά όχι πόσο μεγαλύτερη, καλύτερη, προτιμότερη είναι σε σχέση με κάποια από τις υπόλοιπες. Δεν υπάρχει δηλαδή η ιδιότητα της απόστασης οπότε δεν μπορούμε να υπολογίσουμε το βαθμό διαφοράς ή το σχετικό μέγεθος μεταξύ δύο κατηγοριών. Παραδείγματα τέτοιων μεταβλητών είναι το επίπεδο γνώσης κάποιου θέματος, το επίπεδο εκπαίδευσης, η θέση σε μία εταιρεία.

**Διαστήματος** (interval) είναι οι μεταβλητές των οποίων οι ίσες διαφορές μεταξύ των τιμών τους συνεπάγονται και ίσες διαφορές για το χαρακτηριστικό που μετράει η μεταβλητή. Η κλίμακα αυτή δεν επιτρέπει μόνο την ιεράρχηση των υποκειμένων αλλά προσδιορίζει επίσης και την ακριβή διαφορά τους. Η απόσταση μεταξύ δυο οποιονδήποτε διαδοχικών τιμών της μεταβλητής αυτής είναι ίση με την απόσταση δυο άλλων τυχαίων διαδοχικών τιμών της. Επίσης, δεν έχει νόημα ο υπολογισμός αναλογιών. Βασικό χαρακτηριστικό των μεταβλητών αυτής της διάταξης είναι ο αυθαίρετος ορισμός του μηδενός, που δεν υποδηλώνει παντελή έλλειψη του μετρήσιμου χαρακτηριστικού. Χαρακτηριστικότερο παράδειγμα εδώ είναι η θερμοκρασία.

**Αναλογίας** (rate) είναι οι μεταβλητές των οποίων οι τιμές αντιστοιχούν αναλογικά στην ποσότητα του χαρακτηριστικού που μετρούν. Εδώ το μηδέν ανήκει στο διάστημα τιμών της μεταβλητής και δηλώνει την πλήρη απουσία. Και επίσης για τις τιμές των μεταβλητών αυτών έχει έννοια ο υπολογισμός των αναλογιών. Παραδείγματα είναι η ηλικία, το σωματικό βάρος, εμβαδό οικοπέδων, κλπ.