

Παρουσίαση

βασισμένη στο βιβλίο

**Στατιστική
Τόμος Α**

ΤΩΝ

Βασίλη Αγγελή & Κατερίνας Δημάκη

Αθήνα, 2011

Περιεχόμενα

- Εισαγωγικές Έννοιες
- Γραφικές Μέθοδοι Σύνοψης Δεδομένων

1. Εισαγωγικές Έννοιες

Πρωτεύοντα ρόλο στη μελέτη ενός φαινομένου παίζει η κατανόηση της μεταβλητότητας η οποία είναι αναπόφευκτη.

Αν, για παράδειγμα, ζυγίσουμε με ακρίβεια δοχεία καφέ των 500 γραμμαρίων θα διαπιστώσουμε ότι τα παρατηρούμενα βάρη αποκλίνουν, άλλα περισσότερο και άλλα λιγότερο, από το αναγραφόμενο βάρος των 500 γραμμαρίων. Κατά συνέπεια, για να μελετήσουμε ένα χαρακτηριστικό θα πρέπει να διαθέτουμε περισσότερες από μία μετρήσεις και να κατανοήσουμε την παρατηρούμενη μεταβλητότητα και τα αίτια αυτής.

Η γνώση που θα αποκτηθεί θα αποτελέσει τη βάση για τη λήψη απόφασης σχετικής με το υπό μελέτη φαινόμενο.

Η στατιστική ανάλυση είναι ο τρόπος με τον οποίο επιτυγχάνεται η κατανόηση και η ερμηνεία της μεταβλητότητας ενός συνόλου παρατηρήσεων.

Ορισμός της Στατιστικής

Η Στατιστική **ορίζεται** ως η επιστήμη που ασχολείται με τη **συλλογή**, την **παρουσίαση**, την **επεξεργασία**, την **ανάλυση** και την **ερμηνεία** δεδομένων που προέρχονται από παρατηρήσεις ή μετρήσεις, οι οποίες αφορούν ιδιότητες φυσικών, οικονομικών ή κοινωνικών φαινομένων.

Αντικείμενό της είναι να αναπτύξει τις αρχές και τις μεθόδους που υποστηρίζουν τη λήψη αποφάσεων κάτω από συνθήκες αβεβαιότητας.

Είδη δεδομένων

Τα δεδομένα διακρίνονται σε δύο κατηγορίες.

Ποσοτικά

Ποιοτικά

Τα **ποσοτικά δεδομένα** είναι αριθμητικές παρατηρήσεις, δηλαδή πραγματικοί αριθμοί που αναφέρονται στην τιμή του χαρακτηριστικού που μας ενδιαφέρει (π.χ. ύψος, βάρος, μισθός).

Τα **ποιοτικά δεδομένα** είναι κατηγορικές παρατηρήσεις, δηλαδή παρατηρήσεις που μπορούν να ενταχθούν σε κατηγορίες με βάση κάποιο ποιοτικό χαρακτηριστικό (π.χ. φύλο, οικογενειακή κατάσταση, γεωγραφική περιφέρεια).

Κλίμακες Μέτρησης

Για τη μέτρηση των δεδομένων χρησιμοποιούνται οι παρακάτω τέσσερις κλίμακες μέτρησης:

- ονομαστική κλίμακα
- κλίμακα διάταξης
- κλίμακα διαστήματος
- κλίμακα λόγου

Βασικοί Ορισμοί

Πληθυσμός

Είναι το σύνολο των υπό μελέτη μονάδων.

Παράμετρος

Είναι η ποσότητα η οποία περιγράφει και συνοψίζει κάποιο χαρακτηριστικό του πληθυσμού .

Στατιστική Μεταβλητή

Είναι μία κατάλληλη έκφραση του υπό μελέτη χαρακτηριστικού.

Δείγμα

Είναι ένα υποσύνολο του πληθυσμού το οποίο επιλέγεται για τη μελέτη.

Στατιστική Συνάρτηση

Είναι μία συνάρτηση των στοιχείων του δείγματος.

Παράδειγμα Ο πρύτανης ενός πανεπιστημίου θέλει να διερευνήσει αν οι φοιτητές συμφωνούν με τις αλλαγές που προωθούνται με το νέο νομοσχέδιο για την παιδεία. Για το σκοπό αυτόν επιλέγει τυχαίο δείγμα 150 φοιτητών και τους ζητά να δηλώσουν αν είναι «υπέρ» ή «κατά» των αλλαγών. Από αυτούς, 85 δήλωσαν ότι είναι «υπέρ» των αλλαγών και 65 «κατά». Με βάση τα στοιχεία αυτά να καθορίσετε:

- τον πληθυσμό
- το δείγμα
- τη στατιστική μεταβλητή
- τις παραμέτρους
- τη στατιστική συνάρτηση

Λύση

- Ο πληθυσμός είναι το σύνολο των φοιτητών του πανεπιστημίου.
- Το δείγμα αποτελείται από τους 150 φοιτητές που ρωτήθηκαν από τον πρύτανη.
- Η στατιστική μεταβλητή είναι η άποψη καθενός από τους φοιτητές.
- Οι παράμετροι είναι τα ποσοστά αυτών που θα ήταν «υπέρ» ή «κατά» των αλλαγών, στην περίπτωση που είχαν ερωτηθεί όλοι οι φοιτητές του πανεπιστημίου.
- Η στατιστική συνάρτηση είναι τα ποσοστά των φοιτητών που είναι «υπέρ» ή «κατά» των αλλαγών στο δείγμα. Οι τιμές της στατιστικής συνάρτησης για το συγκεκριμένο δείγμα είναι 56,7% «υπέρ» και 43,3% «κατά».

Τρόποι Συλλογής Δεδομένων

Απογραφική μέθοδος

- Πληροφορίες από κάθε στοιχείο του πληθυσμού.
- Η διαδικασία είναι εξαιρετικά χρονοβόρα και πολυδάπανη και για αυτό χρησιμοποιείται σε ειδικές μόνο περιπτώσεις στην πράξη.

Δειγματοληπτική μέθοδος

- Πληροφορίες από ένα δείγμα το οποίο περιέχει ορισμένα μόνο από στοιχεία του πληθυσμού.
- Η δειγματοληπτική μέθοδος είναι ταχύτερη και οικονομικότερη της απογραφικής.

Βασική Επιδίωξη της Στατιστικής

Τα συμπεράσματα που προκύπτουν από τα στοιχεία του δείγματος να επαχθούν με επιστημονικό και αξιόπιστο τρόπο

στον υπό μελέτη πληθυσμό

Είδη Σφαλμάτων

- **δειγματοληπτικά σφάλματα** παρατηρούνται στην περίπτωση που οι πληροφορίες συλλέγονται με τη δειγματοληπτική μέθοδο και οφείλονται ακριβώς στο ότι οι πληροφορίες αυτές δεν προέρχονται από όλες τις μονάδες του πληθυσμού.
- **Μη δειγματοληπτικά σφάλματα** παρατηρούνται στα στοιχεία που συλλέγονται και με τις δύο μεθόδους και οφείλονται σε πολλούς παράγοντες, όπως η αδύνατη μνήμη του ερωτώμενου, η συστηματική απόκρυψη της αλήθειας από αυτούς και άλλες ειδικές συνθήκες.

Δειγματοληπτικά Σχέδια

1. Απλή Τυχαία Δειγματοληψία

Είναι δειγματοληπτικό σχέδιο σύμφωνα με το οποίο καθένα από τα **ισοπληθή** δείγματα που είναι δυνατό να σχηματισθούν από τον πληθυσμό έχει την **ίδια πιθανότητα επιλογής**.

2. Στρωματοποιημένη Δειγματοληψία

Είναι το δειγματοληπτικό σχέδιο σύμφωνα με το οποίο τα στοιχεία του πληθυσμού διαιρούνται σε **ομοιογενείς υποπληθυσμούς** ή στρώματα **στη βάση κάποιου σημαντικού χαρακτηριστικού** και στη συνέχεια **από κάθε ένα** από τα στρώματα του πληθυσμού επιλέγεται ένα **απλό τυχαίο δείγμα** μεγέθους $n_i, i=1,2,\dots,k$ ανεξάρτητα από τα άλλα. Το ενιαίο δείγμα μεγέθους $n = n_1 + n_2 + \dots + n_k$ που προκύπτει από **την ένωση των k ανεξαρτήτων απλών τυχαίων δειγμάτων** ονομάζεται στρωματοποιημένο τυχαίο δείγμα.

3. Δειγματοληψία Κατά Ομάδες

Είναι η δειγματοληπτική τεχνική η οποία διαιρεί τις στοιχειώδεις μονάδες του πληθυσμού σε ομάδες και στη συνέχεια με απλή τυχαία δειγματοληψία επιλέγει ένα δείγμα από τις ομάδες αυτές. Όταν όλα τα στοιχεία των επιλεγμένων ομάδων περιέχονται στο τελικό δείγμα έχουμε δειγματοληψία κατά ομάδες σε ένα στάδιο. Αντίθετα όταν περιλάβουμε στο ενιαίο δείγμα ένα δείγμα μόνο από τις μονάδες των επιλεγμένων ομάδων έχουμε δειγματοληψία κατά ομάδες σε δύο στάδια. Η δειγματοληψία με το σχέδιο αυτό μπορεί να προχωρήσει σε τρία ή και περισσότερα στάδια.

Παράδειγμα Έστω ότι θέλουμε να εκτιμήσουμε τις μέσες μηνιαίες αποδοχές των 3000 εργαζομένων (αντρών και γυναικών) σε μια μεγάλη βιομηχανική μονάδα. Για το σκοπό αυτόν πρέπει να επιλέξουμε δείγμα 150 εργαζομένων. Να περιγραφεί πώς θα επιλεγεί το δείγμα αυτό με:

- απλή τυχαία δειγματοληψία
- στρωματοποιημένη δειγματοληψία

Λύση

Για να επιλέξουμε δείγμα 150 εργαζομένων με απλή τυχαία δειγματοληψία, ακολουθούμε τα παρακάτω βήματα:

- κωδικοποιούμε τους 3000 εργαζομένους χρησιμοποιώντας διαδοχικά τους αριθμούς από το 0000 για τον πρώτο, μέχρι το 2999 για τον τελευταίο.
- επιλέγουμε τυχαία μια σελίδα του Πίνακα 1 των τυχαίων αριθμών και ένα σημείο εκκίνησης μέσα σε αυτήν και σημειώνουμε τους 150 πρώτους τετραψήφιους αριθμούς που είναι μικρότεροι από το 3000. (απόσπασμα του πίνακα υπάρχει στην τελευταία διαφάνεια).
- το δείγμα αποτελείται από εκείνους τους 150 εργαζομένους των οποίων οι κωδικοί αντιστοιχούν στους 150 τυχαίους αριθμούς που επιλέξαμε.

Για να επιλέξουμε δείγμα 150 εργαζομένων με στρωματοποιημένη δειγματοληψία, ακολουθούμε τα παρακάτω βήματα:

- χωρίζουμε τους εργαζομένους σε άντρες και γυναίκες, δηλαδή σε δύο στρώματα με κριτήριο το φύλο.
- επιλέγουμε απλό τυχαίο δείγμα μεγέθους n_1 από τους άντρες και απλό τυχαίο δείγμα μεγέθους n_2 από τις γυναίκες.
- το ενιαίο δείγμα θα περιλαμβάνει $n = n_1 + n_2 = 150$ εργαζομένους.

Ο προσδιορισμός των n_1 και n_2 αποτελεί αντικείμενο της θεωρίας της δειγματοληψίας και μπορεί να γίνει με πολλούς τρόπους. Ενδεικτικά αναφέρεται ότι τα n_1 και n_2 μπορούν να προσδιοριστούν με αναλογικό επιμερισμό του $n = n_1 + n_2 = 150$ στα δύο στρώματα, δηλαδή με επιμερισμό του $n = 150$ με βάση το αντίστοιχο ποσοστό αντρών/γυναικών στον πληθυσμό.

2. Γραφικές μέθοδοι σύνοψης δεδομένων

Προκειμένου να αντλήσουμε πολύτιμες πληροφορίες από ένα πολυπληθές σύνολο δεδομένων και παράλληλα να αποκτήσουμε μία συνολική εικόνα του φαινομένου στο οποίο αναφέρονται τα δεδομένα αυτά, θα πρέπει αυτά να συνοψιστούν με κάποιον τρόπο. Οι βασικές μέθοδοι σύνοψης των δεδομένων διακρίνονται σε γραφικές και αριθμητικές μεθόδους.

Στην ενότητα αυτή θα παρουσιάσουμε τις βασικές γραφικές μεθόδους σύνοψης δεδομένων που περιλαμβάνουν τους πίνακες και τα διαγράμματα.

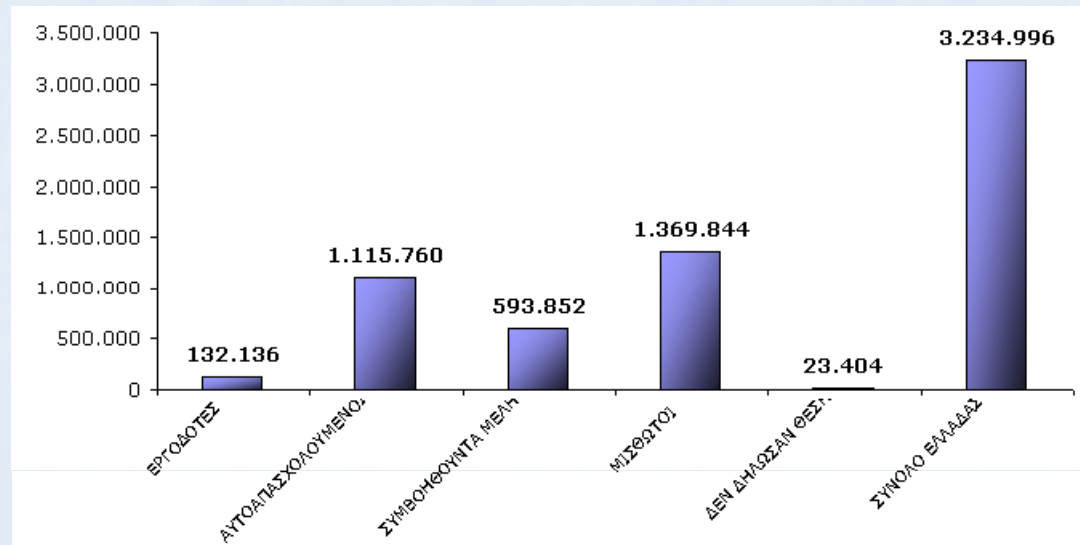
Ραβδόγραμμα

Το ραβδόγραμμα απεικονίζει τη συχνότητα των κατηγοριών στις οποίες είναι ταξινομημένα τα δεδομένα ως ένα ορθογώνιο κάθετο στον οριζόντιο ή τον κατακόρυφο άξονα.

Όλα τα ορθογώνια έχουν το ίδιο (αυθαίρετο πλάτος) ενώ το μήκος κάθε ορθογωνίου είναι ανάλογο της συχνότητας της αντίστοιχης κατηγορίας δεδομένων.

Παράδειγμα

ΘΕΣΗ ΣΤΟ ΕΠΑΓΓΕΛΜΑ	ΣΥΝΟΛΟ
ΕΡΓΟΔΟΤΕΣ	132.136
ΑΥΤΟΑΠΑΣΧΟΛΟΥΜΕΝΟΙ	1.115.760
ΣΥΜΒΟΗΘΟΥΝΤΑ ΜΕΛΗ	593.852
ΜΙΣΘΩΤΟΙ	1.369.844
ΔΕΝ ΔΗΛΩΣΑΝ ΘΕΣΗ	23.404
ΣΥΝΟΛΟ ΕΛΛΑΔΑΣ	3.234.996

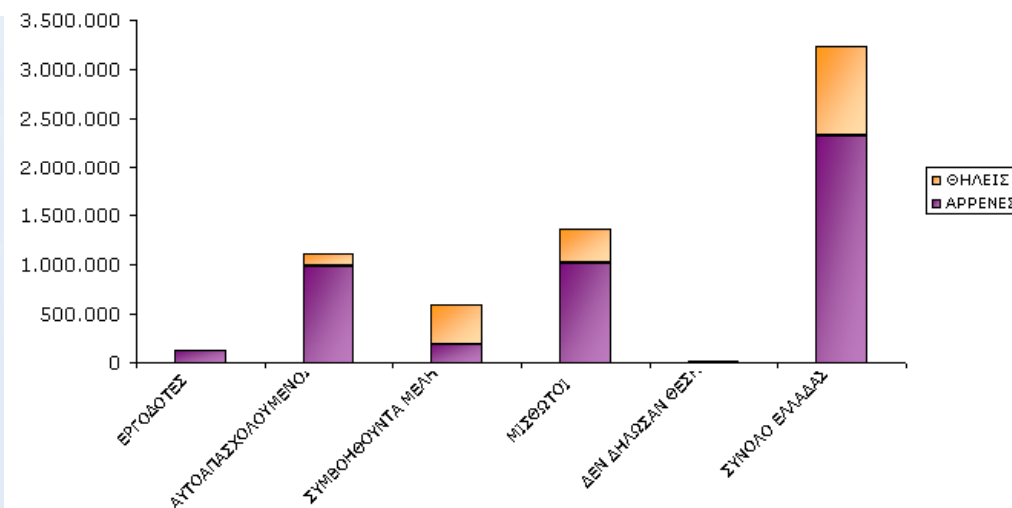


Ραβδόγραμμα Συνιστωσών

Στο ραβδόγραμμα συνιστωσών **κάθε** ορθογώνιο χωρίζεται σε τμήματα (συνιστώσες) που το καθένα αντιστοιχεί σε μία υποδιαίρεση της κατηγορίας που αντιπροσωπεύει το ορθογώνιο

Το **μήκος** κάθε συνιστώσας είναι ανάλογο της συχνότητας (ή σχετικής συχνότητας) της αντίστοιχης υποδιαίρεσης.

ΘΕΣΗ ΣΤΟ ΕΠΑΓΓΕΛΜΑ	ΣΥΝΟΛΟ	ΑΡΡΕΝΕΣ	ΘΗΛΕΙΣ
ΕΡΓΟΔΟΤΕΣ	132.136	120.832	11.304
ΑΥΤΟΑΠΑΣΧΟΛΟΥΜΕΝΟΙ	1.115.760	982.280	133.480
ΣΥΜΒΟΗΘΟΥΝΤΑ ΜΕΛΗ	593.852	183.696	410.156
ΜΙΣΘΩΤΟΙ	1.369.844	1.026.236	343.608
ΔΕΝ ΔΗΛΩΣΑΝ ΘΕΣΗ	23.404	16.544	6.860
ΣΥΝΟΛΟ ΕΛΛΑΔΑΣ	3.234.996	2.329.588	905.408



Γραμμογράφημα

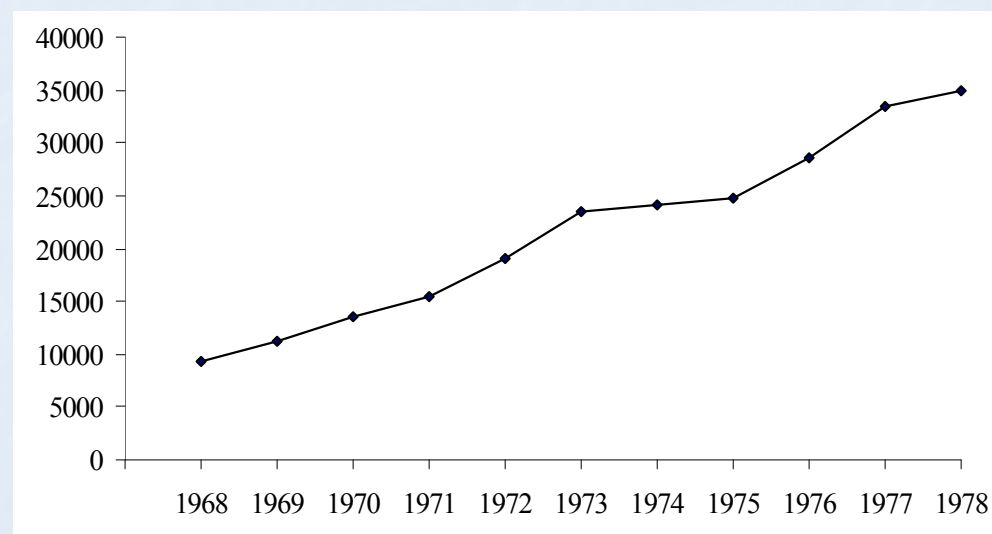
- Τα γραμμογραφήματα χρησιμοποιούνται κυρίως όταν τα δεδομένα μας είναι χρονολογικές σειρές
- Ο οριζόντιος άξονας εκφράζει τις χρονικές στιγμές στις οποίες έχουμε κατατάξει τα δεδομένα και ο κατακόρυφος τις αντίστοιχες συχνότητες
- Για να κατασκευάσουμε ένα γραμμογράφημα προσδιορίζουμε τις συχνότητες για κάθε κατηγορία (χρονική στιγμή) και τις ενώνουμε με ευθύγραμμα τμήματα.

Γραμμογράφημα

(Παράδειγμα)

Η εξέλιξη της δύναμης του ελληνικού στόλου για την περίοδο 1968-1979 φαίνεται στον πίνακα

ΤΕΛΟΣ ΕΤΟΥΣ	ΧΩΡΗΤΙΚΟΤΗΤΑ (σε χιλ. ΚΟΧ)
1968	9216
1969	11139
1970	13539
1971	15441
1972	19093
1973	23400
1974	24080
1975	24820
1976	28573
1977	33475
1978	34987



Κυκλικό Διάγραμμα

- Τα κυκλικά διαγράμματα χρησιμοποιούνται για τη γραφική απεικόνιση δεδομένων που αποτελούν υποδιαίρεσεις ενός συνόλου.
 - Αν οι υποδιαίρεσεις εκφράζονται σε ποσοστά το διάγραμμα αναφέρεται ως κυκλικό διάγραμμα ποσοστών
- Για να κατασκευάσουμε ένα κυκλικό διάγραμμα
 - χαράσσουμε καταρχήν ένα κύκλο με αυθαίρετη ακτίνα
 - Τον χωρίζουμε σε κυκλικούς τομείς οι οποίοι αντιστοιχούν στις διάφορες κατηγορίες των παρατηρήσεων και το εμβαδόν τους είναι ανάλογο με τη συχνότητα της κάθε κατηγορίας.

Παράδειγμα Οι δαπάνες ενός μεγάλου νοσοκομείου κατά το έτος 2008 κατανέμονται ως εξής:

73% για μισθούς

13% για προμήθεια ιατρικού υλικού

8% για συντήρηση μηχανημάτων

6% για γενικά έξοδα διοίκησης.

Να αποδοθούν γραφικά οι πληροφορίες αυτές με το κατάλληλο διάγραμμα.

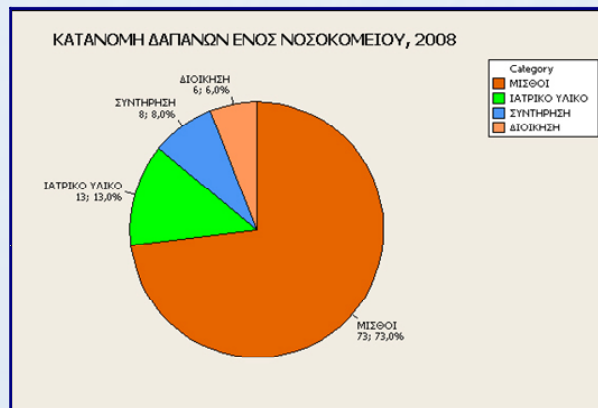
Λύση. Για να απεικονίσουμε γραφικά τις πληροφορίες αυτές, επιλέγουμε ένα κυκλικό διάγραμμα. Όπως ήδη αναφέρθηκε, κάθε 1% των παρατηρήσεων θα αντιστοιχεί σε κυκλικό τομέα γωνίας 3.6° . Κατά συνέπεια, θα έχουμε τις παρακάτω αντιστοιχίες μεταξύ της ποσοστιαίας κατανομής των δαπανών του νοσοκομείου και των κεντρικών γωνιών των κυκλικών τομέων.

$$\text{για μισθούς} \quad 73\% \rightarrow (3.6^\circ)(73) = 262.8^\circ$$

$$\text{για προμήθεια ιατρικού υλικού} \quad 13\% \rightarrow (3.6^\circ)(13) = 46.8^\circ$$

$$\text{για συντήρηση μηχανημάτων} \quad 8\% \rightarrow (3.6^\circ)(8) = 28.8^\circ$$

$$\text{γενικά έξοδα διοίκησης} \quad 6\% \rightarrow (3.6^\circ)(6) = 21.6^\circ$$



Κατανομή Συχνότητας

- Για την ταξινόμηση σε κατανομή συχνοτήτων ακολουθούμε τρία βήματα:
 - Δημιουργία Τάξεων ή Κλάσεων για τη μεταβλητή που εξετάζουμε
 - Ένταξη των δεδομένων στις τάξεις
 - Αναγραφή του πλήθους των τιμών της μεταβλητής που εμπίπτουν στην κάθε τάξη
- Για το πρώτο βήμα πρέπει να προσδιορίσουμε κάθε φορά
 - το πλήθος των τάξεων και
 - το πλάτος της κάθε τάξης.

Για να το πετύχουμε αυτό λαμβάνουμε υπόψη μας τα ακόλουθα:

- ✓ Οι κατανομές έχουν συνήθως περισσότερες από 6 και λιγότερες από 16 τάξεις.
- ✓ Η κάθε τιμή της μεταβλητής εντάσσεται σε μία μόνο τάξη
- ✓ Επιδιώκουμε οι τάξεις να έχουν όλες το ίδιο πλάτος

Κατανομή Συχνότητας

Ο Πίνακας που ακολουθεί περιέχει τις τελικές βαθμολογίες (με άριστα το 100) ενός τυχαίου δείγματος 150 φοιτητών ενός Πανεπιστημιακού Τμήματος, στη θεματική ενότητα «Ποσοτικές μέθοδοι», κατά το ακαδημαϊκό έτος 2009-10.

27	79	69	40	51	88	55	48	36	61
53	44	94	51	65	42	58	55	69	63
70	48	61	55	60	25	47	78	61	54
57	76	73	62	36	67	40	51	59	68
27	46	62	43	54	83	59	13	72	57
82	45	54	52	71	53	82	69	60	35
41	65	62	75	60	42	55	34	49	45
49	64	40	61	73	44	59	46	71	86
43	69	54	31	56	51	75	44	66	53
80	71	53	56	91	60	41	29	56	57
35	54	43	39	56	27	62	44	85	61
59	89	60	51	71	53	58	26	77	68
62	57	48	69	76	52	49	45	54	41
33	61	80	57	42	45	59	44	68	73
55	70	39	58	69	51	85	46	55	67

Κατανομή Συχνότητας

- Οι βαθμολογίες κυμαίνονται από 13 ως και 94, επιλέγουμε τις εννέα τάξεις 10-20, 20-30,...,90-100
- Εντάσσουμε τις τιμές της μεταβλητής (βαθμολογία υποψηφίων) στις δημιουργηθείσες τάξεις.

Κάθε τάξη ορίζεται από δύο αριθμούς.
Ο μικρότερος από αυτούς ονομάζεται
κατώτερο όριο και ο μεγαλύτερος **ανώτερο όριο**

Ανώτατο όριο=20
Κατώτατο όριο=10

Πλάτος Τάξεως (δ) =
Η διαφορά μεταξύ δύο διαδοχικών
κάτω (ή πάνω) ορίων
 $\delta = 40 - 30 = 10$

ΤΑΞΕΙΣ	ΠΕΡΙΠΤΩΣΕΙΣ	ΣΥΧΝΟΤΗΤΑ
10-20	I	1
20-30	IIII I	6
30-40	IIII III	9
40-50	IIII IIII IIII IIII IIII IIII I	31
50-60	IIII IIII IIII IIII IIII IIII IIII	42
60-70	IIII IIII IIII IIII IIII IIII II	32
70-80	IIII IIII IIII II	17
80-90	IIII IIII	10
90-100	II	2
		150

Κεντρικός όρος= ημίαθροισμα
του κατώτερου και ανώτερου ορίου κάθε τάξης

$$\frac{50+60}{2} = 55$$

Οι τάξεις είναι **διαστήματα ανοικτά**
προς τα πάνω.
π.χ. [10,20)

Συχνότητα:
το πλήθος των τιμών της μεταβλητής
που εντοπίζονται σε μια δεδομένη τάξη

Κατανομή Σχετικής Συχνότητας

- Το ποσοστό όλων των τιμών της μεταβλητής που βρίσκονται σε μία συγκεκριμένη τάξη ονομάζεται **σχετική συχνότητα**.

ΤΑΞΕΙΣ	ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ	ΣΧΕΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΩΣ ΠΟΣΟΣΤΟ
10-20	$1/150=0,007$	$0,007*100= 0,7$
20-30	$6/150=0,040$	4,0
30-40	$9/150=0,060$	6,0
40-50	$31/150=0,207$	20,7
50-60	$42/150=0,280$	28,0
60-70	$32/150=0,213$	21,3
70-80	$17/150=0,113$	11,3
80-90	$10/150=0,067$	6,7
90-100	$2/150=0,013$	1,3
	1,000	100,0

Μετά την κατασκευή της κατανομής συχνοτήτων δίνεται και μία γραφική απεικόνιση της.

Η πιο δημοφιλής είναι το ιστόγραμμα

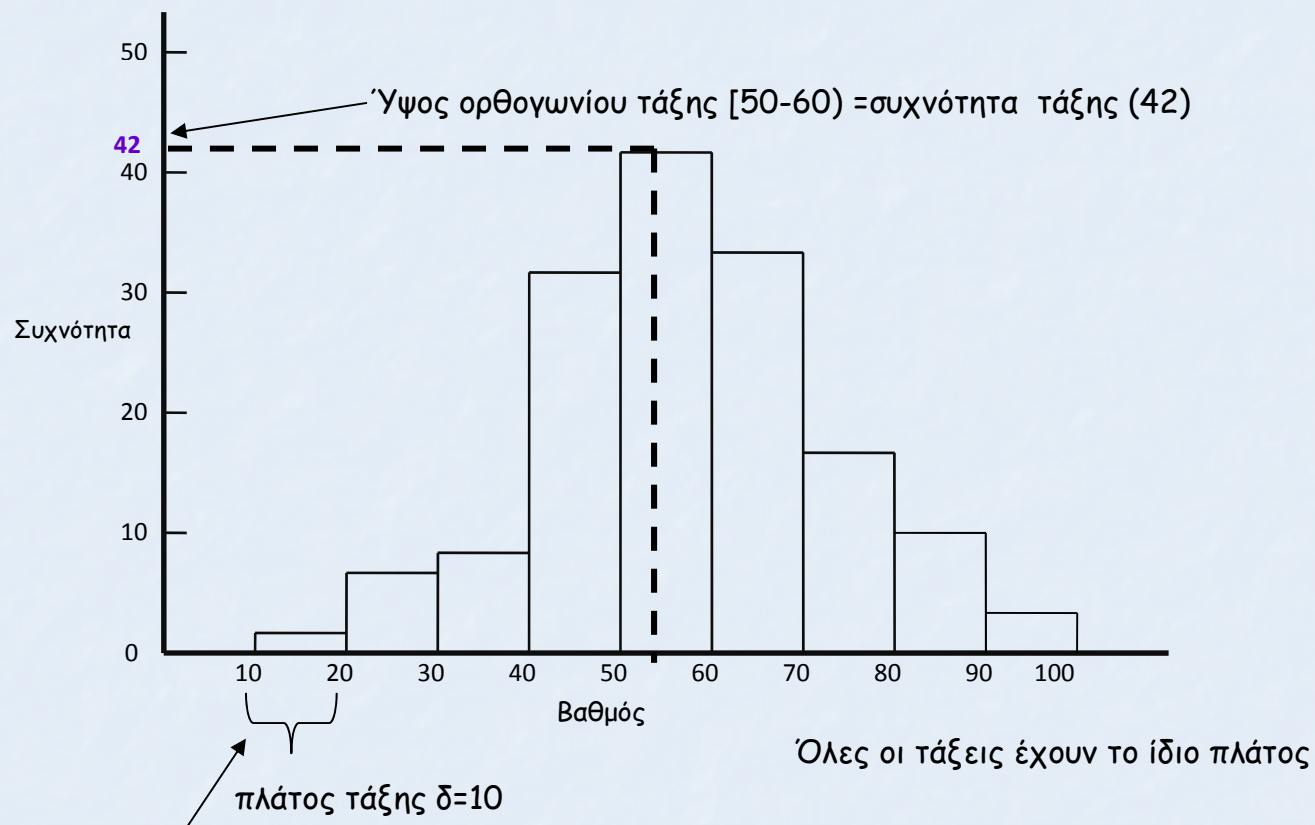
Ιστόγραμμα

ΚΑΤΑΣΚΕΥΗ ΙΣΤΟΓΡΑΜΜΑΤΟΣ:

1. Στον οριζόντιο άξονα παίρνουμε **ίσα** διαδοχικά ευθύγραμμα τμήματα τα οποία *αντιστοιχούν* στις **τάξεις** της κατανομής.
2. Οι **συχνότητες** της κατανομής τοποθετούνται στον κατακόρυφο άξονα.
3. Στη συνέχεια κατασκευάζουμε ορθογώνια παραλληλόγραμμα με **βάσεις** τα ευθύγραμμα τμήματα που αντιστοιχούν στις **τάξεις** της κατανομής και **ύψη** ανάλογα των **συχνοτήτων** της κάθε τάξης.

Ιστόγραμμα

(Παράδειγμα συν.)



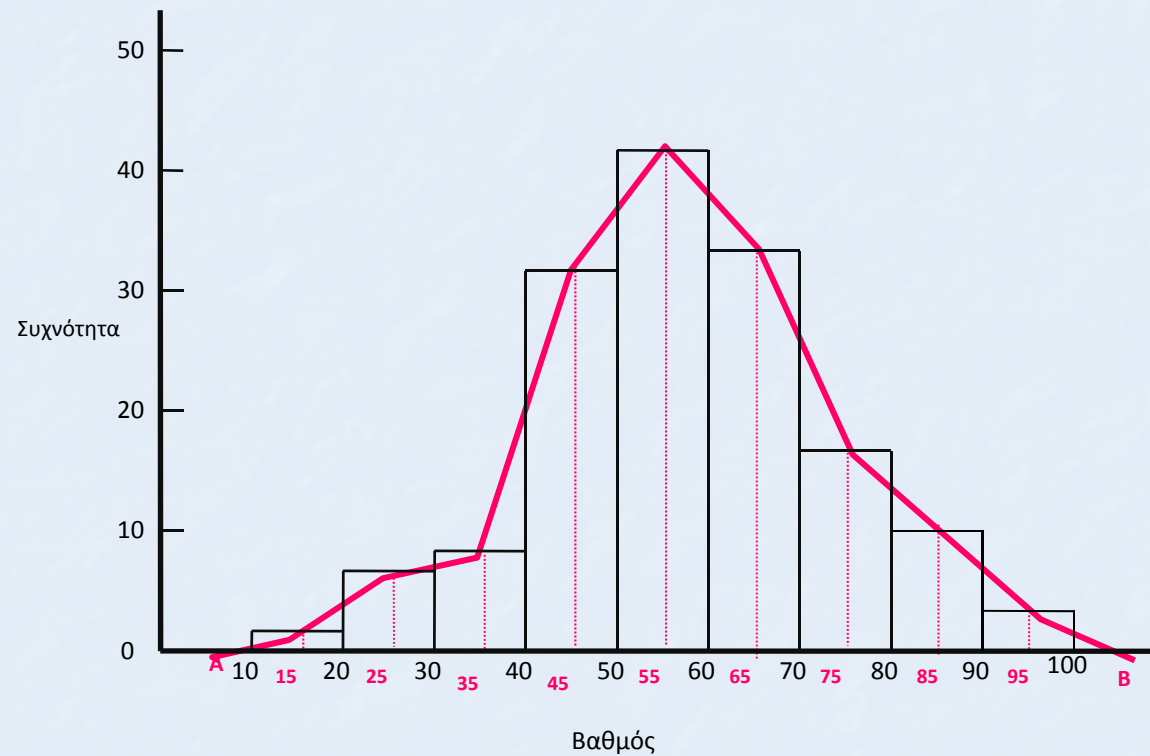
Πολύγωνο Συχνοτήτων

ΚΑΤΑΣΚΕΥΗ ΠΟΛΥΓΩΝΙΚΗΣ ΓΡΑΜΜΗΣ

- Τοποθετούμε στον οριζόντιο άξονα τους κεντρικούς όρους των τάξεων της κατανομής
- Υψώνουμε στα σημεία αυτά κάθετα ευθύγραμμα τμήματα ίσα προς τις αντίστοιχες συχνότητες των τάξεων.
- Ενώνουμε με ευθύγραμμα τμήματα τα διαδοχικά σημεία που έχουν προκύψει.
- Η πολυγωνική γραμμή είναι μία κλειστή τεθλασμένη

Για να το πετύχουμε αυτό θεωρούμε δύο υποθετικές τάξεις με συχνότητα μηδέν στα δύο άκρα της κατανομής και συνδέουμε τα άκρα A και B της τεθλασμένης που έχει ήδη προκύψει με τους κεντρικούς όρους των δύο αυτών υποθετικών τάξεων.

Πολυγωνική Γραμμή



Διάγραμμα Σημείων

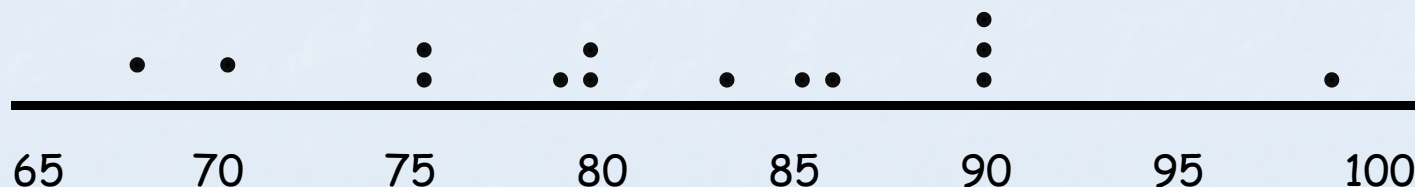
- Χρησιμοποιείται όταν έχουμε μικρό αριθμό παρατηρήσεων
- Πρόκειται για την τοποθέτηση όλων των διαθέσιμων τιμών πάνω σε ένα ευθύγραμμο τμήμα.

Αν υπάρχουν περισσότερες από μια τιμές οι οποίες συμπίπτουν τοποθετούνται η μία πάνω στην άλλη

Παράδειγμα: Η βαθμολογία 15 φοιτητών στο μάθημα της Στατιστικής δίνεται στον πιο κάτω πίνακα

70 90 82 90 75 85 86 79 84 98 75 80 90 80 68

Να κατασκευαστεί το διάγραμμα σημείων των βαθμολογιών.



Να κατασκευαστεί το διάγραμμα σημείων των τελικών βαθμολογιών (με άριστα το 100) ενός τυχαίου δείγματος 150 φοιτητών ενός Πανεπιστημιακού Τμήματος, στη θεματική ενότητα «Ποσοτικές μέθοδοι», κατά το ακαδημαϊκό έτος 2009-10.



Διάγραμμα Μίσχου-Φύλλου

- Είναι ένας απλός αλλά περιγραφικός τρόπος παρουσίασης των δεδομένων, που υποδηλώνει την κατανομή τους.
- Κατασκευή διαγράμματος
 - Χωρίζουμε κάθε παρατήρηση σε δύο μέρη, το μίσχο και το φύλλο.
 - Κατατάσσουμε τις τιμές του μίσχου και τοποθετούμε τις τιμές των φύλλων στα δεξιά του αντίστοιχου μίσχου.
- Σημειώνεται ότι ο καθορισμός του φύλλου και μίσχου είναι αυθαίρετος.

Παράδειγμα: Η βαθμολογία 15 φοιτητών στο μάθημα της Στατιστικής είναι:

70 90 82 90 75 85 86 79 84 98 75 80 90 80 68

Να κατασκευαστεί το διάγραμμα μίσχου φύλλου των βαθμολογιών.

Μίσχος	Φύλλο
6	8
7	0 5 9 5
8	2 5 6 4 0 0
9	0 0 8 0

Να κατασκευαστεί το διάγραμμα μίσχου φύλλου των τελικών βαθμολογιών (με άριστα το 100) ενός τυχαίου δείγματος 150 φοιτητών ενός Πανεπιστημιακού Τμήματος, στη θεματική ενότητα «Ποσοτικές μέθοδοι», κατά το ακαδημαϊκό έτος 2009-10.

Stem-and-Leaf Display: Βαθμολογία 150 Φοιτητών

Stem-and-leaf of C1 N = 150

Leaf Unit = 1.0

1	1	3
7	2	567779
16	3	134556699
47	4	0001112223334444455556667888999
(42)	5	111111223333344444455555566667777788899999
61	6	00000111111222223455677888999999
29	7	00111123335566789
12	8	0022355689
2	9	14

Αθροιστική Κατανομή Συχνοτήτων

Τα δεδομένα πολλές φορές συνοψίζονται με
τις αθροιστικές κατανομές συχνοτήτων

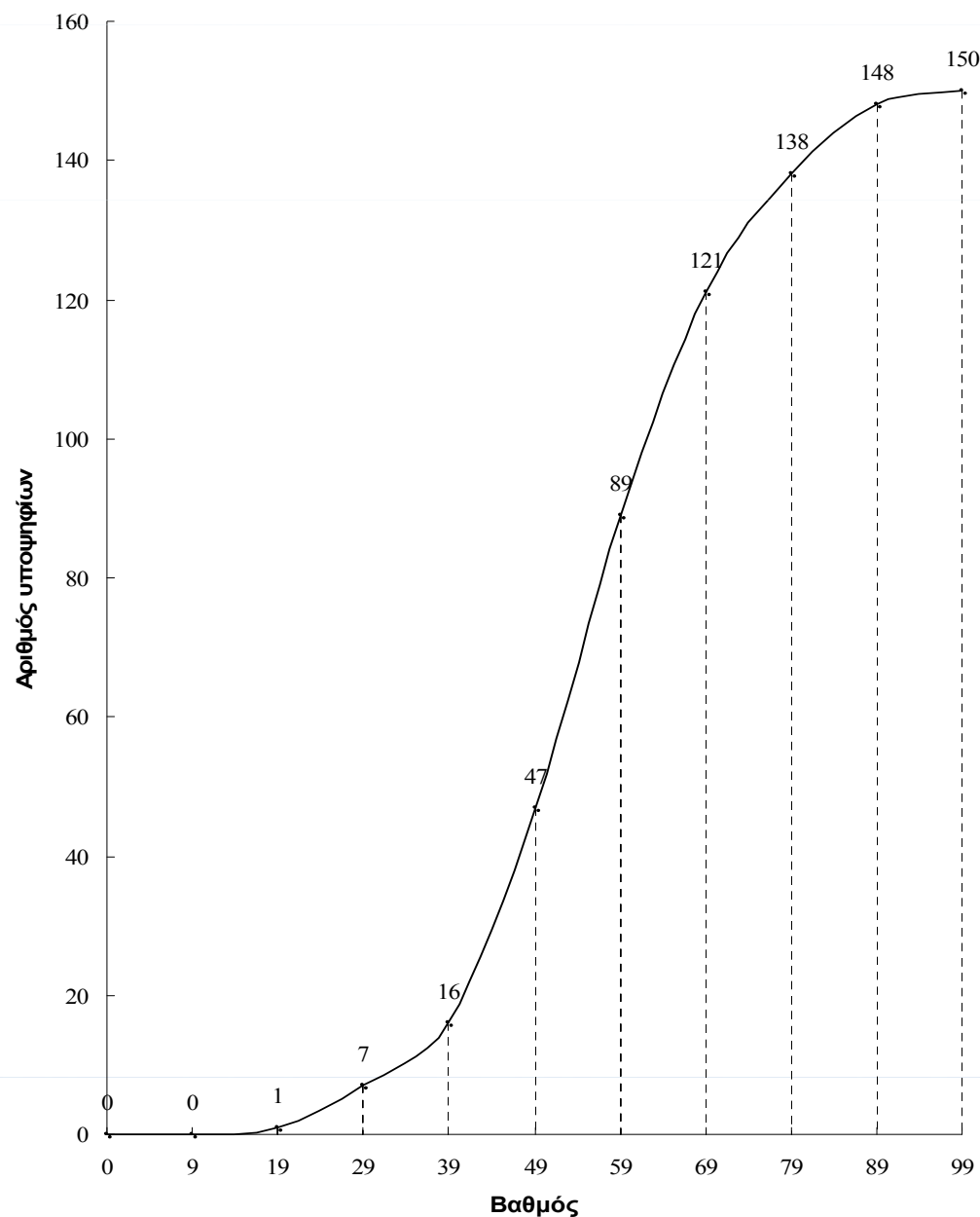
Στη συνέχεια παρουσιάζεται η διαδικασία σχηματισμού της
δεξιόστροφης αθροιστικής κατανομής συχνοτήτων

ΤΑΞΕΙΣ	ΣΥΧΝΟΤΗΤΑ	ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ	ΣΧΕΤΙΚΗ ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ	ΣΧΕΤΙΚΗ ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ ΩΣ %
10-20	1	1	$1/150=0,007$	0,7
20-30	6	$1+6=7$	$7/150=0,047$	4,7
30-40	9	$7+9=16$	$16/150=0,106$	10,6
40-50	31	$16+31=47$	$47/150=0,313$	31,3
50-60	42	$47+42=89$	$89/150=0,593$	59,3
60-70	32	$89+32=121$	$121/150=0,806$	80,6
70-80	17	$121+17=138$	$138/150=0,92$	92
80-90	10	$138+10=148$	$148/150=0,986$	98,6
90-100	2	$148+2=150$	$150/150=1$	100
	150			

Δεξιόστροφη Αθροιστική Κατανομή Συχνοτήτων

Αν ενδιαφερόμαστε
για το πόσες τιμές της
μεταβλητής είναι
μικρότερες
από κάποια τιμή
χρησιμοποιούμε
**δεξιόστροφη
αθροιστική κατανομή**

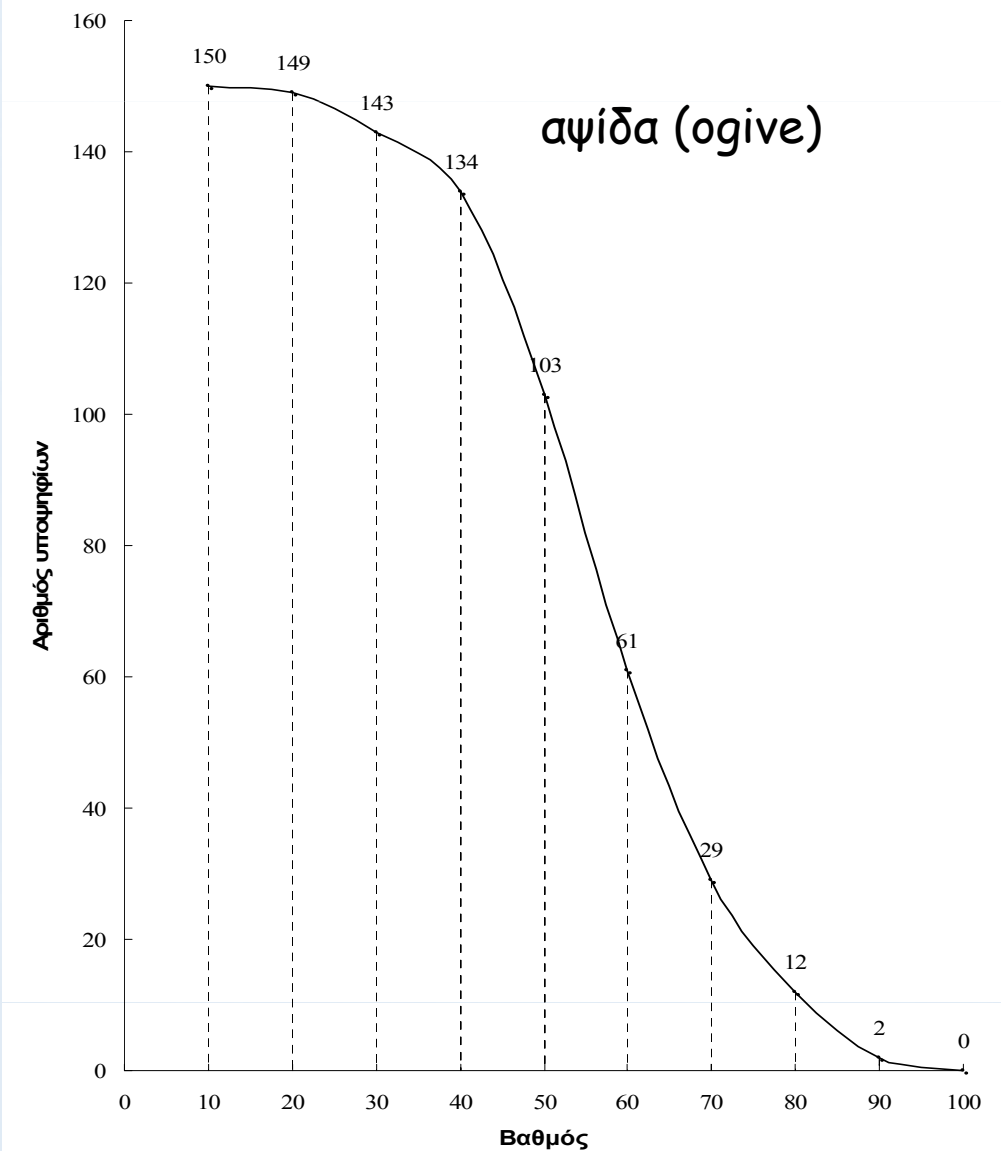
ΔΕΞΙΟΣΤΡΟΦΗ ΑΘΡΟΙΣΤΙΚΗ ΚΑΤΑΝΟΜΗ		
ΒΑΘΜΟΛΟΓΙΑ	ΥΠΟΨΗΦΙΟΙ	
	ΑΡΙΘΜΟΣ	%
ΚΑΤΩ ΤΟΥ 20	1	0,67
ΚΑΤΩ ΤΟΥ 30	7	4,67
ΚΑΤΩ ΤΟΥ 40	16	10,67
ΚΑΤΩ ΤΟΥ 50	47	31,33
ΚΑΤΩ ΤΟΥ 60	89	59,33
ΚΑΤΩ ΤΟΥ 70	121	80,67
ΚΑΤΩ ΤΟΥ 80	138	92,00
ΚΑΤΩ ΤΟΥ 90	148	98,67
ΚΑΤΩ ΤΟΥ 100	150	100,00



Αριστερόστροφη Αθροιστική Κατανομή Συχνοτήτων

Αν ενδιαφερόμαστε
για το πόσες τιμές της
μεταβλητής είναι
μεγαλύτερες
από κάποια τιμή
χρησιμοποιούμε
αριστερόστροφη
αθροιστική κατανομή

ΑΡΙΣΤΕΡΟΣΤΡΟΦΗ ΑΘΡΟΙΣΤΙΚΗ ΚΑΤΑΝΟΜΗ		
ΒΑΘΜΟΛΟΓΙΑ	ΥΠΟΨΗΦΙΟΙ	
	ΑΡΙΘΜΟΣ	%
10 ΚΑΙ ΠΑΝΩ	150	100,00
20 ΚΑΙ ΠΑΝΩ	149	99,53
30 ΚΑΙ ΠΑΝΩ	143	95,33
40 ΚΑΙ ΠΑΝΩ	134	89,33
50 ΚΑΙ ΠΑΝΩ	103	68,66
60 ΚΑΙ ΠΑΝΩ	61	40,66
70 ΚΑΙ ΠΑΝΩ	29	19,33
80 ΚΑΙ ΠΑΝΩ	12	8,00
90 ΚΑΙ ΠΑΝΩ	2	1,33



Πίνακας Διπλής Εισόδου

Υπάρχουν περιπτώσεις που οι συχνότητες αναφέρονται σε *ζευγάρια μεταβλητών*. Ο πίνακας με τον οποίο επιτυχαίνομε τη σύνοψη των δεδομένων μας ονομάζεται **Πίνακας Διπλής Εισόδου**.

Παράδειγμα: Έστω ότι διαθέτουμε 149 ζευγάρια τιμών που αφορούν τα ύψη σπουδαστριών και τα αντίστοιχα βάρη τους

Πίνακας Διπλής Εισόδου

Έστω ότι διαθέτουμε 149 ζευγάρια τιμών που αφορούν τα ύψη σπουδαστριών και τα αντίστοιχα βάρη τους

ΥΨΟΣ	ΒΑΡΟΣ	ΥΨΟΣ	ΒΑΡΟΣ	ΥΨΟΣ	ΒΑΡΟΣ	ΥΨΟΣ	ΒΑΡΟΣ	ΥΨΟΣ	ΒΑΡΟΣ
173,8	49,7	162,2	60,4	168,7	57,3	162,1	65,5	165,5	55,1
161,4	47,3	164,8	58,3	165,0	62,1	156,2	58,3	164,2	59,9
156,9	57,7	168,0	64,8	168,9	51,3	152,7	59,1	170,0	53,1
172,5	60,0	161,5	53,8	161,8	54,9	157,5	54,8	159,1	53,2
169,4	62,6	169,9	64,4	164,0	56,2	156,2	60,9	157,6	52,9
170,6	63,3	171,7	59,7	162,8	48,8	164,5	60,5	150,5	48,5
173,7	61,1	160,0	50,4	164,0	61,2	172,7	56,9	166,1	66,6
160,7	68,8	158,9	52,0	170,4	64,0	164,5	54,2	171,0	68,1
164,4	55,0	168,4	53,8	159,8	50,7	168,0	73,4	165,5	56,9
165,7	56,5	164,3	62,8	164,0	55,7	169,4	55,0	168,4	64,9
157,1	71,3	158,5	55,6	159,0	52,7	161,4	68,9	166,6	58,7
152,7	43,4	151,8	49,4	158,2	56,7	164,0	55,1	161,7	49,4
157,5	52,2	164,3	61,0	160,5	47,1	167,9	69,9	157,8	54,6
169,6	64,9	170,4	52,9	158,3	57,1	169,0	70,2	168,6	62,2
165,5	58,5	159,9	52,1	160,5	50,7	167,8	63,8	169,3	54,2
164,9	62,0	174,5	68,6	154,0	53,3	159,8	51,6	154,8	45,0
155,6	60,4	172,0	69,2	167,5	70,8	167,9	58,9	177,5	56,4
169,3	58,7	168,6	64,9	165,2	53,3	171,0	64,1	167,0	71,7
162,2	52,4	163,4	62,1	167,2	61,6	156,6	54,0	151,0	43,7
162,5	62,4	167,1	56,8	165,7	53,9	163,2	68,1	164,0	52,7
160,2	49,6	171,1	58,4	165,8	68,6	166,1	54,8	152,0	56,2
154,1	59,8	167,1	64,5	172,4	64,6	159,4	55,8	159,1	49,4
159,4	50,5	160,4	62,0	156,5	46,0	163,7	60,0	160,5	59,3
175,9	71,8	159,8	60,6	162,5	69,5	162,5	65,2	158,7	57,9
158,6	49,6	157,7	58,0	165,4	53,4	165,1	50,3	159,4	54,5
164,5	71,0	172,7	63,4	160,3	51,3	159,7	58,4	159,7	50,2
164,0	56,4	169,7	64,1	171,5	66,1	165,5	63,4	157,4	56,5
161,3	49,7	167,0	61,4	165,4	55,5	158,7	51,4	162,5	58,9
164,1	63,2	166,8	64,8	160,4	49,2	155,0	55,7	159,0	69,7
168,5	62,7	159,1	51,5	166,7	59,5	164,1	69,5		

Πίνακας Διπλής Εισόδου

ΥΨΟΣ ΒΑΡΟΣ	150-154	154-158	158-162	162-166	166-170	170-174	174-178	ΣΥΝΟΛΟ
39-47	2	2	-	-	-	-	-	4
47-55	2	6	24	8	4	3	-	47
55-63	2	8	9	23	12	5	1	60
63-71	-	-	3	8	13	8	1	33
71-79	-	1	-	1	2	-	1	5
ΣΥΝΟΛΟ	6	17	36	40	31	16	3	149

Από τα σύνολα των συχνοτήτων της τελευταίας **γραμμής** προκύπτουν οι αντίστοιχες κατανομές συχνοτήτων για τα **ύψη** των 149 σπουδαστών.

ΤΑΞΕΙΣ	ΣΥΧΝΟΤΗΤΑ
150-154	6
154-158	17
158-162	36
162-166	40
166-170	31
170-174	16
174-178	3
	149

Από τα σύνολα των συχνοτήτων της τελευταίας **στήλης** προκύπτουν οι αντίστοιχες κατανομές συχνοτήτων για τα **βάρη** των 149 σπουδαστών.

ΤΑΞΕΙΣ	ΣΥΧΝΟΤΗΤΑ
39-47	4
47-55	47
55-63	60
63-71	33
71-79	5
	149

Προσάρτημα

Πίνακας 1. Τυχαίοι αριθμοί (απόσπασμα)

Ο πίνακας περιέχει 2600 ψηφία τυχαίων αριθμών τοποθετημένα σε σύνολα 4 ψηφίων το καθένα. Για την επιλογή τυχαίων αριθμών από το 1 ως το 10^k αρκεί από μια τυχαία επιλεγμένη στήλη ή γραμμή να επιλέγουμε διαδοχικά ψηφία σε ομάδες των $(k-1)$ ψηφίων.

Στήλη	Γραμμή									
	1	2	3	4	5	6	7	8	9	10
1	6126	4874	3421	0964	9162	0413	9518	4391	3937	6147
2	0248	4421	4994	2853	1025		7081	6431	2691	9270
3	1230	6438	8888	2561	5148	7325	2444	6602	5112	7658
4	2408	2078	4239	8453	8769	7513	7005	8760	8589	6254
5	1409	5374	9214	5878	9049	9410	7060	7015	4879	4218
6	2515	8535	6370	3437	6603	5063	3718	9732	5741	1266
7	7260	7680	0643	0401	0909	9452	1138	9930	0263	9366
8	7081	4579	9973	0986	5680	6877	7053	3050	7630	3872
9	3603	5948	3575	1059	6892	0861	0051	8050	5941	0928
10	3741	2855	2086	7078	6976	1937	4900	9626	3653	8032
11	6141	1921	1328	7548	8762	2382	0075	9768	3086	7804
12	6894	6295	2947	4894	2339	6846	9947	8253	9785	9066
13	8969	1124	9210	6187	4215	6381	2054	1418	5181	3270
14	7768	9362	1010	2201	7693	3793	4968	4799	6117	0389
15	6391	1258	8255	8540	4980	9103	7955	2637	7512	8361
16	2344	8230	5303	5187	6799	5875	2158	5118	2958	6707
17	7280	7699	1736	4276	4129	7253	3531	7566	9909	9359
18	2756	2579	8231	6859	1844	3114	9282	4063	9043	9856
19	3086	2524	5835	9728	7513	4830	2393	0743	3206	4582
20	0231	6420	8827	5488	7486	0882	6213	7367	4780	4555