

Η ΑΠΑΤΗ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

Το μεροληπτικό δείγμα

Στο *αμερόληπτο* ή *τυχαίο δείγμα* ή *αντιπροσωπευτικό*, πρέπει όλα τα άτομα του πληθυσμού να έχουν την ίδια πιθανότητα να συμπεριληφθούν στο δείγμα.

Παραδείγματα:

A) Στέκομαι έξω από τον σιδηροδρομικό σταθμό και ρωτάω τους περαστικούς. Αποκλείω έτσι ένα κομμάτι του πληθυσμού.

B) Το μέσο εισόδημα των αποφοίτων του 1924 του πανεπιστημίου του Γέηλ είναι 25 111 \$ (πολύ υψηλό). Ποιους βρήκαν και ρώτησαν; Πως τους βρήκαν;

Το μέγεθος του δείγματος

Το *μέγεθος του δείγματος* έχει σημασία για την αξιοπιστία μιας δημοσκόπησης.

Παράδειγμα:

Ρίχνω ένα νόμισμα 10 φορές και έρχονται 8 γράμματα. Αυτό δεν σημαίνει ότι αν ρίξω το ίδιο νόμισμα 100 φορές θα φέρει 80 γράμματα.

Μέτρα Θέσης

Μέση Τιμή - Διάμεσος - Επικρατούσα Τιμή

Το καθένα από μόνο του δεν λέει πολλά πράγματα.

Παραδείγματα:

A) Ο μέσος μισθός μιας περιοχής

B) Μια μέση οικογένεια αποτελείται από 3,6 άτομα. Αυτό σημαίνει 3 ή 4 άτομα. Οι κατασκευαστές οικιών φτιάχνουν κατοικίες με 2 ή 3 κρεβατοκάμαρες. Όμως η κατηγορία με 3 ή 4 άτομα είναι η μικρότερη αριθμητικά.

Σε δεδομένα που ακολουθούν την κανονική κατανομή τα τρία μέτρα θέσεως συμπίπτουν.

Οι μικροί αριθμοί που λείπουν

A) Επίπεδο σημαντικότητας

Παράδειγμα:

Το μέσο βάρος των μαθητών της Γ' Γυμνασίου της πόλης Α ήταν πριν 5 χρόνια 70,4 kg. Σε ένα τυχαίο δείγμα φέτος βρέθηκε μέσο βάρος 70,8. Μπορούμε να

ισχυριστούμε ότι το μέσο βάρος των μαθητών της Γ΄ τάξης αυξήθηκε; Πρέπει να απαντήσουμε αναφέροντας και την πιθανότητα σφάλματος (επίπεδο σημαντικότητας).

Π.χ αποδεικνύεται ότι αν το μέσο βάρος ενός δείγματος είναι μεγαλύτερο ή ίσο του 71,18, τότε το μέσο βάρος των μαθητών έχει αυξηθεί με πιθανότητα σφάλματος 2%.

Λόγω της στατιστικής αβεβαιότητας, δεν μπορούμε να αποδείξουμε μια υπόθεση βασισμένη σε ένα δείγμα. Εξάγουμε συμπεράσματα που συνοδεύονται από μια πιθανότητα λάθους. (Εκτός και το δείγμα μας ταυτιστεί με τον πληθυσμό).

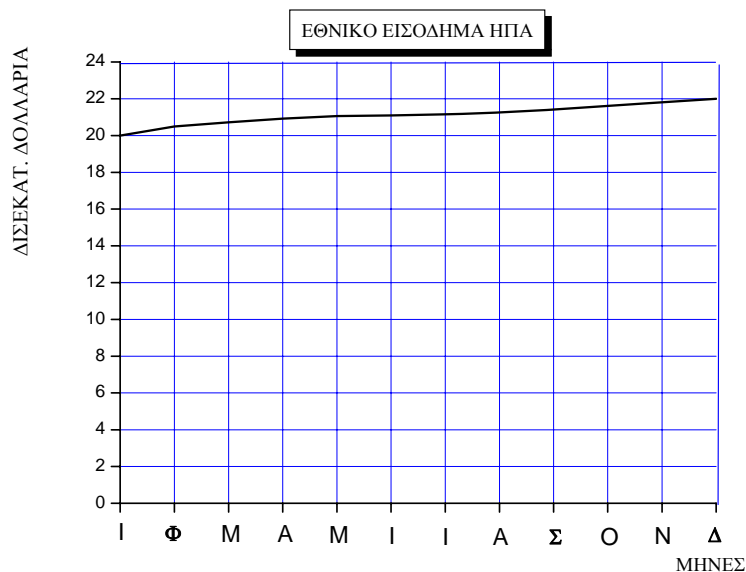
Β) Διακύμανση

Η Διακύμανση μας καθορίζει την απόκλιση των τιμών από τον μέσο όρο, ο οποίος μόνος του, ενίοτε δεν λέει πολλά πράγματα και ίσως είναι άχρηστος.

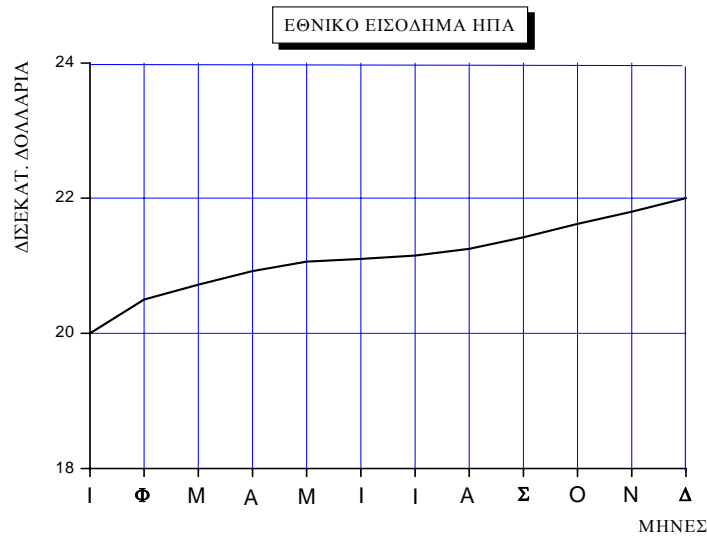
Παράδειγμα:

Ένα μικρό παιδί αρχίζει να στέκεται όρθιο στην ηλικία των x μηνών. Αν δεν συμβεί αυτό για το παιδί του ένας γονιός αρχίζει να ανησυχεί. Όμως λόγω του κεντρικού οριακού θεωρήματος, ο πραγματικός μέσος θα βρίσκεται στο διάστημα x συν πλην δύο τυπικές αποκλίσεις του δείγματος, με πιθανότητα 95%.

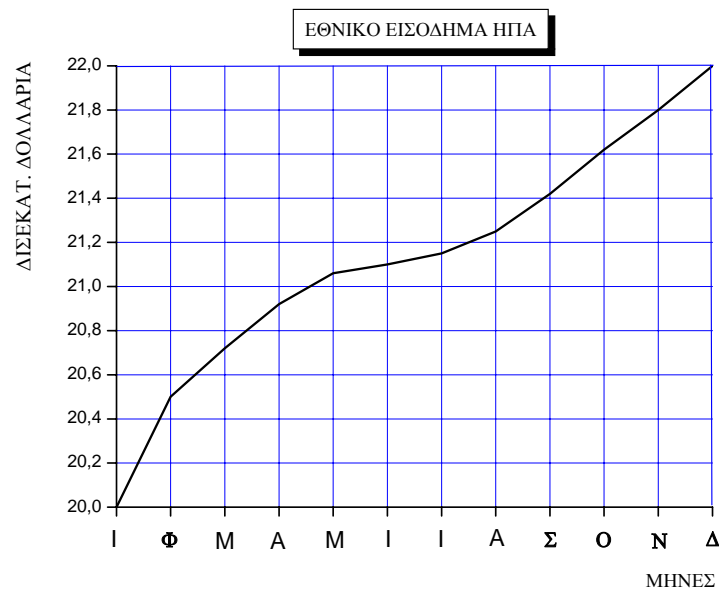
Παραπλανητικά Γραφήματα



Το εθνικό εισόδημα των ΗΠΑ μέσα σε ένα χρόνο. (Αύξηση 10%)



Η αύξηση φαίνεται μεγαλύτερη.



Η αύξηση τώρα είναι «εντυπωσιακή».

Συσχέτιση και Αιτιότητα

Συσχέτιση δεν σημαίνει και αιτιότητα.

Παράδειγμα:

α) Οι φωλιές των πελαργών στην στέγη και ο αριθμός παιδιών σε μια δανέζικη οικογένεια.

β) Ο μισθός των καθηγητών και η τιμή του καφέ.

Έλεγχος των στατιστικών εκτιμήσεων

α) Ποιος το λέει αυτό;

Έχει σημασία η αυθεντία να βρίσκεται πίσω από την πληροφορία και όχι δίπλα σε αυτή.

β) Πως το ξέρει;

Π.χ το δείγμα ήταν τυχαίο;

γ) Τι λείπει;

Π.χ οι ανεπεξέργαστοι αριθμοί.

δ) Μήπως άλλαξε κάποιος το θέμα;

Π.χ Οι Βρετανοί (ισχυρίζονται ότι) κάνουν μπάνιο συχνότερα από τις γυναίκες τους.

ε) Έχει νόημα;

Π.χ Η εξέλιξη ενός πληθυσμού μιας χώρας. Υπάρχει μια τάση, θα διατηρηθεί;