Τεχνικές Ανάλυσης και Πρόβλεψης Τηλεπικοινωνιακών Αγορών

ΤΕ-3011

π

Περιεχόμενα Μαθήματος

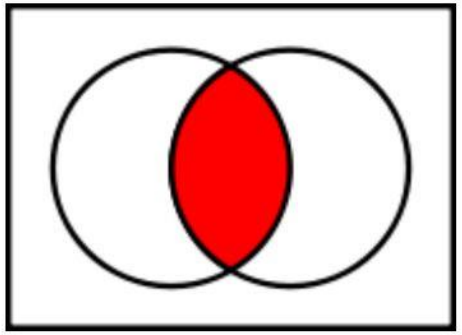
- Σύνολα
- Πιθανότητες
- Κατανομές
- Δειγματοληψία
- Παρουσίαση Στατιστικών Δεδομένων
- Εκτιμητική
- Συσχέτιση

Σύνολα

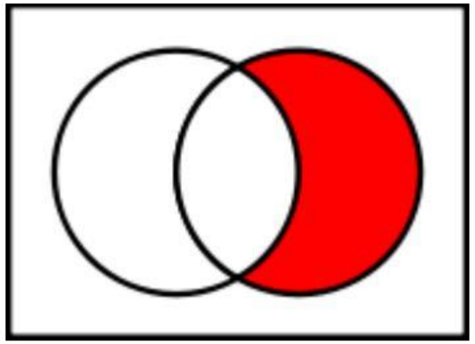
Σύνολα

- **Σύνολο** είναι κάθε συλλογή από αντικείμενα
- **Στοιχείο** ενός συνόλου ονομάζουμε κάθε αντικείμενο που περιέχεται σε ένα σύνολο
- Δύο σύνολα είναι **ίσα** όταν έχουν τα ίδια ακριβώς στοιχεία
- Ένα σύνολο A ονομάζεται **υποσύνολο** ενός συνόλου B, όταν κάθε στοιχείο του A είναι και στοιχείο του B
- **Κενό σύνολο** ονομάζεται το σύνολο που δεν περιέχει καθόλου στοιχεία

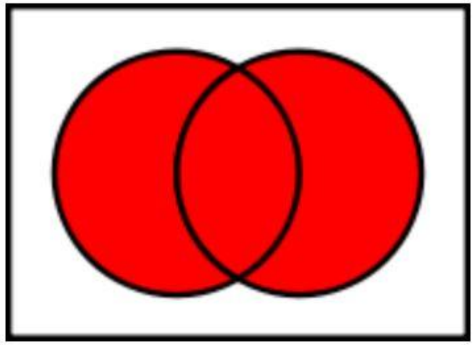
Σύνολα



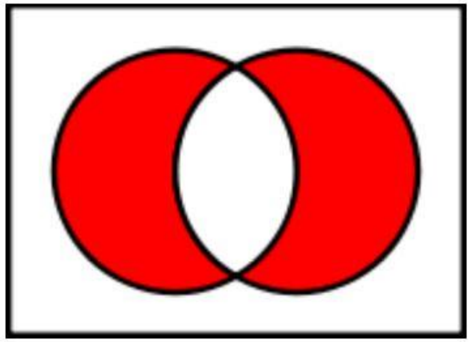
Τομή Συνόλων $A \cap B$ *(AND)*



Διαφορά Συνόλων $= A^C \cap B = B \setminus A$ $(B \text{ χωρίς } A)$



Ένωση συνόλων $A \cup B$ *(OR)*



Συμμετρική Διαφορά Συνόλων $= (A \Delta B) = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$

(OR, OR)

Πιθανότητες

Πιθανότητες

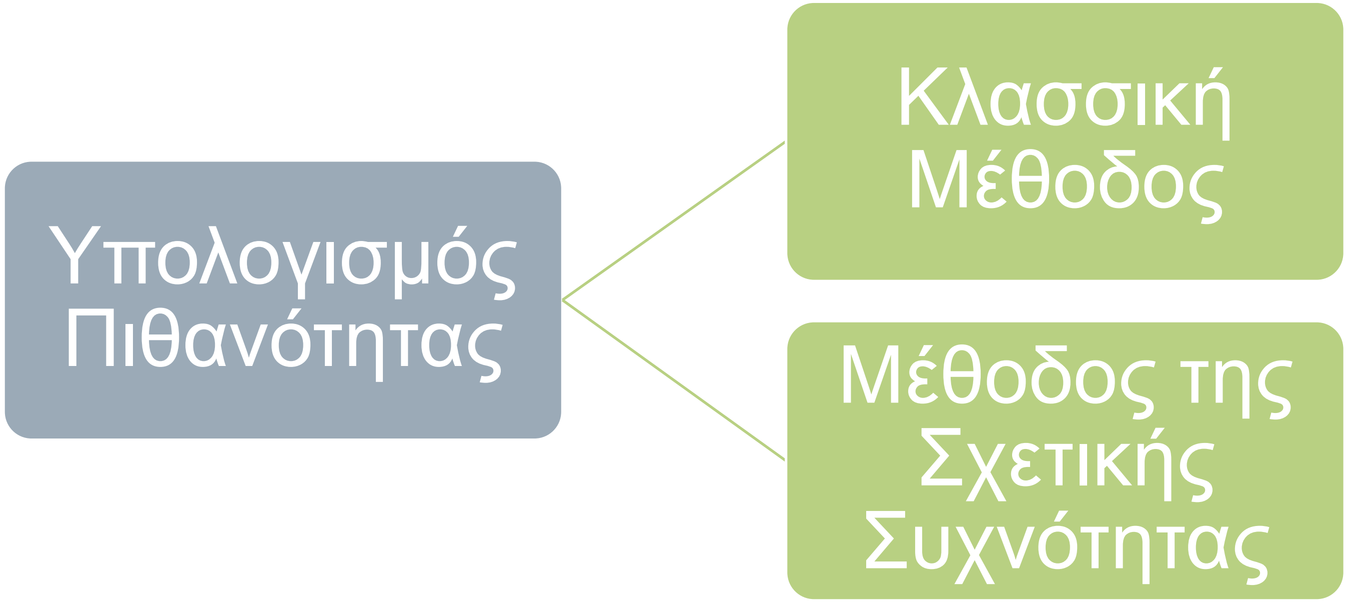
- **Δειγματοχώρος** ονομάζεται το σύνολο όλων των δυνατών αποτελεσμάτων κατά την εκτέλεση ενός πειράματος τύχης

Πιθανότητες

- **Γεγονός** ή **ενδεχόμενο** ονομάζεται ένα υποσύνολο δυνατών αποτελεσμάτων του δειγματοχώρου
- Εάν τα σύνολα που αντιστοιχούν στα γεγονότα Α και Β είναι ξένα δηλαδή $A \cap B = \emptyset$, τότε τα γεγονότα αυτά λέγονται **ασυμβίβαστα** ή **αμοιβαίως αποκλειόμενα**. Τα ενδεχόμενα αυτά δεν μπορούν να συμβούν ταυτόχρονα.
- Ο ίδιος ο δειγματοχώρος καλείται **βέβαιο γεγονός**, ενώ το κενό υποσύνολο **αδύνατο γεγονός**.

Πιθανότητες

- **Έννοια της πιθανότητας:** Σε ένα τυχαίο πείραμα υπάρχει πάντα η αβεβαιότητα, για το αν θα συμβεί ένα ορισμένο γεγονός ή όχι
- Ως **μέτρο της πιθανότητας** ορίζουμε έναν αριθμό μεταξύ 0 και 1



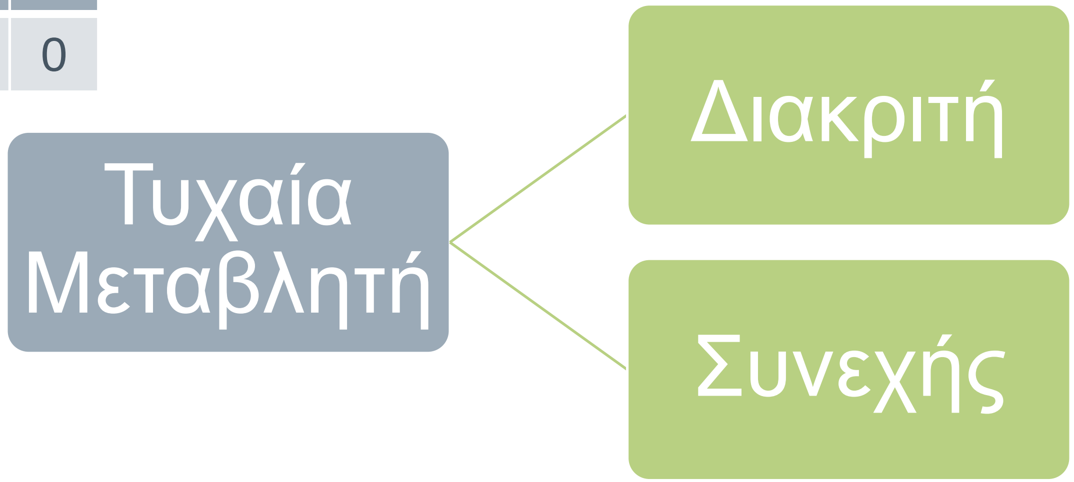
Αξιώματα Πιθανότητας (Kolmogoroff)

1. Για κάθε γεγονός Α ενός συνόλου είναι $P(A) \geq 0$.
2. Για το βέβαιο γεγονός δ είναι $P(\delta)=1$. Αν δηλαδή το δ είναι όλος ο δειγματοχώρος, τότε η πιθανότητα είναι 1.
3. Προσθετική Ιδιότητα: Αν τα γεγονότα Α και Β είναι ασυμβίβαστα, τότε $P(A \cup B)=P(A)+P(B)$.

Πιθανότητες

- **Τυχαία μεταβλητή** ή **στοχαστική συνάρτηση** καλείται μία συνάρτηση ορισμένη στο δειγματοχώρο, μέσω της οποίας κάθε σημείο του δειγματοχώρου αντιστοιχίζεται σε έναν αριθμό.
- Για δύο νομίσματα έχουμε:

Σημείο Δειγματοχώρου	ΚΚ	ΚΓ	ΓΚ	ΓΓ
X	2	1	1	0



Πιθανότητες

› Έστω Χ μία τυχαία διακριτή μεταβλητή με τιμές x_1, x_2, x_3..... Ας υποθέσουμε, ότι οι πιθανότητες να πάρει η μεταβλητή τις τιμές αυτές είναι:

$$P(X = x_k) = f(x_k)$$

› Συνάρτηση πυκνότητας πιθανότητας ή κατανομή πιθανότητας

$$P(X=x)= f(x)$$

έτσι ώστε:

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

Πιθανότητες

› Αθροιστική συνάρτηση κατανομής ή συνάρτηση κατανομής ορίζεται με τη σχέση

$$P(X \leqslant x) = F(x), \ -\infty < x < \infty$$

› Η συνάρτηση κατανομής μπορεί να ληφθεί από τη συνάρτηση πιθανότητας ως:

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$$

Κατ’ αντιστοιχία για συνεχείς τυχαίες μεταβλητές.

Πιθανότητες

- **Μέση** ή **Αναμενόμενη Τιμή** ή **Μαθηματική Ελπίδα** για μια τυχαία διακριτή μεταβλητή ορίζεται:

$$\mathrm{E(X)} = x_1 P(X = x_1) + x_2 P(X = x_2) + \ldots + x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- Η αναμενόμενη τιμή συμβολίζεται με **μ**.

Πιθανότητες

› Η **διασπορά** ή **διακύμανση** ή **μεταβλητότητα** ορίζεται από τη σχέση:

$$\mathrm{Var}(x) = \sigma^2 = E[(X - \mu)^2]$$

› Μας δίνει το ποσό της μεταβλητότητας των τιμών της μεταβλητής γύρο από το μέσο της

› Η θετική τετραγωνική ρίζα της διασποράς ονομάζεται **τυπική απόκλιση**

Πιθανότητες

- Για πολυδιάστατες κατανομές, η **συνδιακύμανση ή συνδιασπορά (Covariance)** μας δίνει το μέσο ποσό της ταυτόχρονης μεταβλητότητας των δύο μεταβλητών από τον μέσο τους.
- Η συνδιασπορά είναι το μέτρο σχέσης των δύο μεταβλητών. Δηλαδή αν οι δύο μεταβλητές συσχετίζονται τότε όπως μεταβάλλεται η μία γύρο από τον μέσο της, με τον παρόμοιο τρόπο ή τον αντίθετο θα μεταβάλλεται και η άλλη γύρο από το μέσο της

Πιθανότητες

- Η **διάμεσος** (median) ενός συνόλου n τιμών είναι η τιμή της μεσαίας παρατήρησης, όταν όλες οι παρατηρήσεις είναι ταξινομημένες σε αύξουσα ή φθίνουσα σειρά
- Μισές παρατηρήσεις έχουν τιμή μεγαλύτερη από τη διάμεσο και μισές έχουν τιμή μικρότερη από αυτήν
  - Αν n είναι περιττός αριθμός η διάμεσος είναι η τιμή της (n+1)/2 παρατήρησης
  - Εάν είναι ζυγός η διάμεσος τιμή είναι ο μέσος όρος των δύο μεσαίων τιμών

Πιθανότητες

- Η **επικρατούσα** ή **τυπική τιμή** (mode) ενός συνόλου παρατηρήσεων ορίζεται ως η τιμή με τη μεγαλύτερη συχνότητα
- Ως **εύρος** (range) των τιμών ενός συνόλου ορίζεται η διαφορά της μικρότερης από τη μεγαλύτερη τιμή τους
- **Ακραίες** (outliers) λέγονται οι τιμές, οι οποίες είναι, είτε πολύ μεγαλύτερες, είτε πολύ μικρότερες από τον κύριο όγκο των δεδομένων

Πιθανότητες

› Οι **βαθμοί ελευθερίας** είναι ο αριθμός των ελεύθερων επιλογών που έχουμε στον προσδιορισμό της τιμής των μεταβλητών

Αριθμός των βαθμών ελευθερίας = αριθμός μεταβλητών – αριθμός περιορισμών

Πιθανότητες

- Μία κατανομή που είναι συγκεντρωμένη κοντά στη μέση τιμή παρουσιάζει έντονη **κύρτωση**.
- Σαν μέτρο της κύρτωσης παίρνουμε την τιμή της αδιάστατης ποσότητας:

$$\alpha_4 = \frac{\mathrm{E}[(\mathrm{X} - \mu)^4]}{\sigma^4}$$

- Όσο πιο μεγάλη είναι αυτή η ποσότητα, τόσο μεγαλύτερη κύρτωση έχει η κατανομή, δηλαδή τόσο πιο συγκεντρωμένη είναι κοντά στη μέση τιμή

Κατανομές

Κατανομές – Δυωνυμική Κατανομή

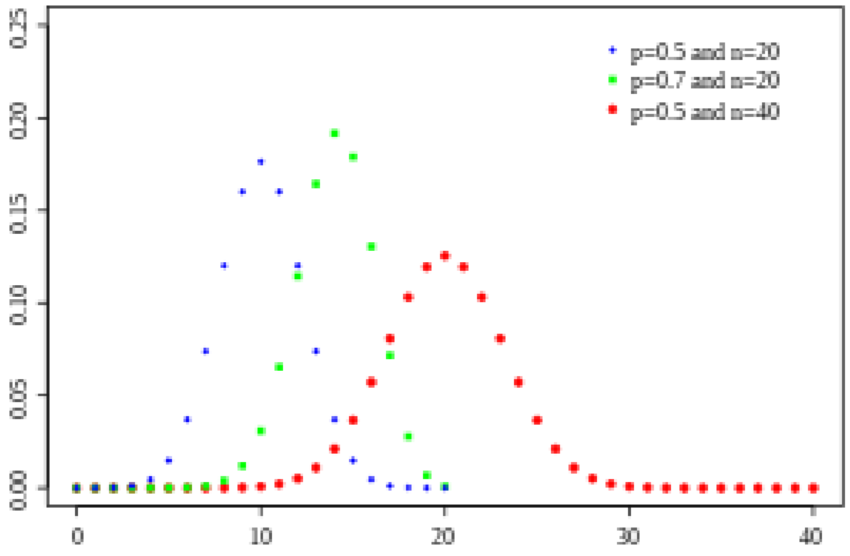
› Έστω p η πιθανότητα να πραγματοποιηθεί το γεγονός σε μία μόνο δοκιμή και 1-p η πιθανότητα να μη πραγματοποιηθεί. Η πιθανότητα να πραγματοποιηθεί το γεγονός x φορές σε n δοκιμές είναι:

$$f(x)=P(X=x)=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

όπου,

- p = η πιθανότητα επιτυχίας σε κάθε δοκιμή
- x = ο αριθμός των επιτυχιών ή εμφάνισης του γεγονότος
- n = ο αριθμός των δοκιμών

Μέση τιμή	μ=np
Διασπορά	$\sigma^2=np(1-p)$
Τυπική απόκλιση	$\sigma=\sqrt{np(1-p)}$



Κατανομές – Κανονική Κατανομή

- Η κανονική κατανομή χρησιμοποιείται ως μία πρώτη προσέγγιση για να περιγραφούν τυχαίες μεταβλητές πραγματικών τιμών, οι οποίες τείνουν να συγκεντρώνονται γύρω από μια μέση τιμή.
- Η κανονική κατανομή αποτελεί την πιο σημαντική κατανομή της στατιστικής μεθοδολογίας για τους εξής βασικούς λόγους:
  - Την κανονική κατανομή ακολουθούν είτε με ακρίβεια είτε με μεγάλη προσέγγιση τα περισσότερα συνεχή φαινόμενα.
  - Πολλές ασυνεχείς κατανομές πιθανοτήτων μπορούν να προσεγγιστούν μέσω της κανονικής κατανομής. Για παράδειγμα πολλά πληθυσμιακά χαρακτηριστικά, όπως το ύψος, το βάρος η βαθμολογία σε διαγώνισμα, κ.λπ.
  - Η κανονική κατανομή αποτελεί σύμφωνα με το κεντρικό οριακό θεώρημα τη βάση της στατιστικής συμπερασματολογίας ή επαγωγικής στατιστικής.
  - Τυχαία σφάλματα που εμφανίζονται σε διάφορες μετρήσεις έχουν κανονική κατανομή. Γι' αυτό το λόγο αναφέρεται πολλές φορές και ως κατανομή σφαλμάτων

Κατανομές – Κανονική Κατανομή

Μία από τις σημαντικότερες κατανομές είναι η κανονική κατανομή ή κατανομή του Gauss. Η συνάρτηση πυκνότητας είναι:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

Η συνάρτηση κατανομής είναι η:

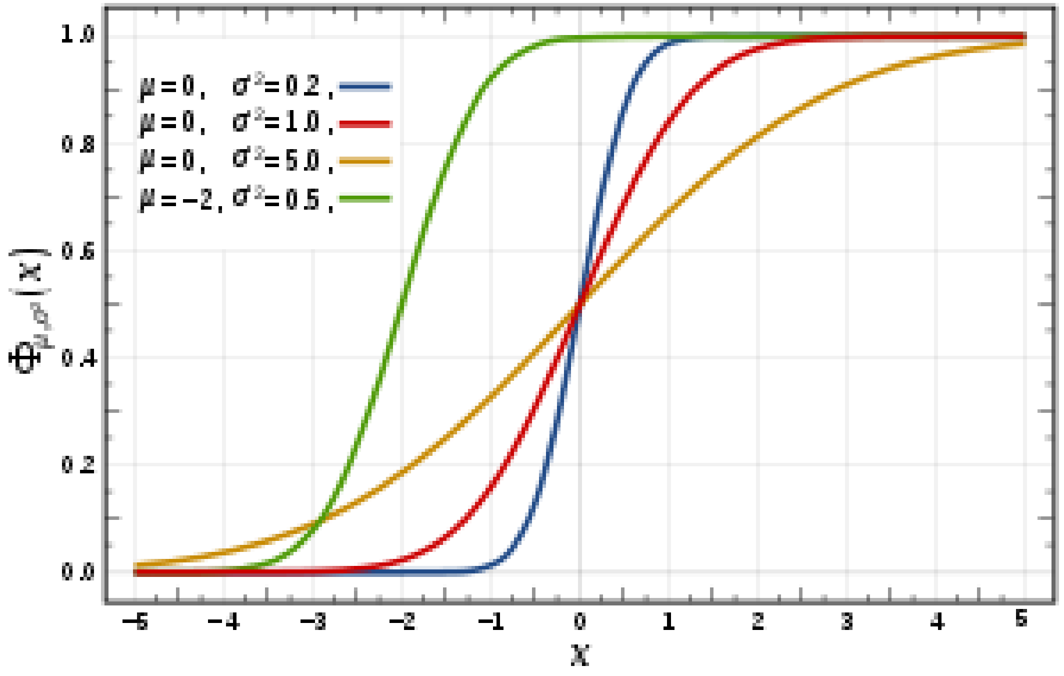
$$F(x)=P(X\leq x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{-(u-\mu)^2/2\sigma^2}du$$

Ορίζοντας $Z=\frac{X-\mu}{\sigma}$ ως την τυποποιημένη μεταβλητή, έχουμε:

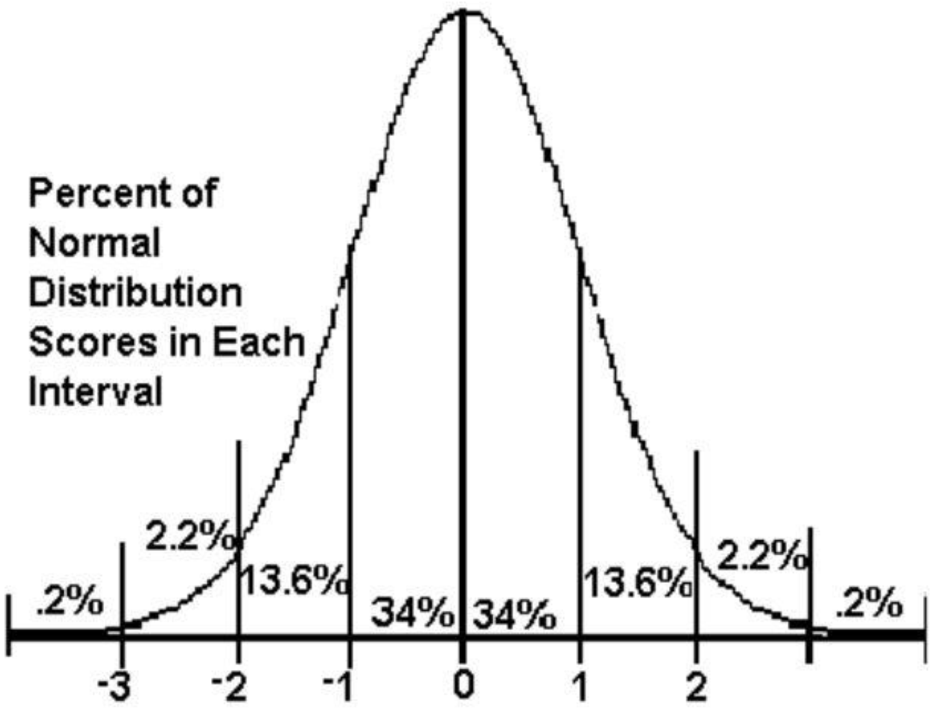
$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2} \quad \text{και}$$

$$F(z)=P(Z\leq z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-u^2/2}du=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_{0}^{z}e^{-u^2/2}du$$

όπου το δεύτερο ολοκλήρωμα είναι η error function.



Κατανομές – Κανονική Κατανομή



Μέση τιμή	μ
Διασπορά	σ^2
Τυπική απόκλιση	σ

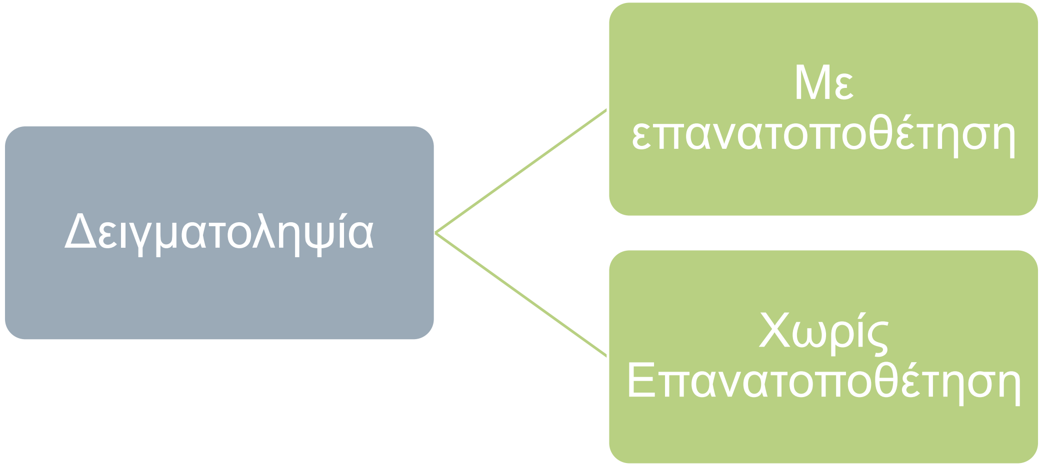
Κατανομές

- x^2
- Student
- F
- …

Δειγματοληψία

Δειγματοληψία

- Όταν θέλουμε να βγάλουμε συμπεράσματα για μια μεγάλη ομάδα ατόμων ή αντικειμένων, αντί να μελετήσουμε ολόκληρη την ομάδα, τον **πληθυσμό**, εξετάζουμε ένα μικρό εύρος αυτού του πληθυσμού, δηλαδή ένα **δείγμα**
- Η διαδικασία λήψεως ενός δείγματος λέγεται δειγματοληψία



Δειγματοληψία

- Ως **στατιστική δειγματική συνάρτηση** (που είναι συνάρτηση των τυχαίων μεταβλητών του δείγματος) ορίζουμε κάθε ποσότητα, που προκύπτει από ένα δείγμα στη διάρκεια εκτίμησης των παραμέτρων του πληθυσμού, έστω π.χ. το δειγματικό μέσο.
- Για κάθε παράμετρο θα έχουμε διαφορετική συνάρτηση.
- Η τιμή μιας στατιστικής συνάρτησης μεταβάλλεται από δείγμα σε δείγμα

Δειγματοληψία

- Η τιμή μιας στατιστικής συνάρτησης μεταβάλλεται από δείγμα σε δείγμα

- Η τιμή μιας στατιστικής συνάρτησης είναι και η ίδια τυχαία μεταβλητή και για κάθε δείγμα θα έχουμε μία τιμή αυτής της στατιστικής συνάρτησης

- Η **κατανομή δειγματοληψίας** (sampling distribution) είναι η κατανομή πιθανότητας αυτής της στατιστικής συνάρτησης

Δειγματοληψία

- Στη δειγματοληψία είναι επιθυμητό να πάρουμε ένα δείγμα, το οποίο να αποτελεί μια μικρογραφία του πληθυσμού, να είναι δηλαδή **αντιπροσωπευτικό**
- Τέλεια αντιπροσωπευτικό δείγμα δεν υπάρχει. Συμπεραίνοντας για τον πληθυσμό με βάση τις πληροφορίες του δείγματος θα κάνουμε αναπόφευκτα ένα σφάλμα, το οποίο ονομάζεται **σφάλμα δειγματοληψίας** (sampling error)
- Όταν το σφάλμα οφείλεται στην κακή μεθοδολογία μέτρησης ή στη μη τυχαιότητα του δείγματος ονομάζεται **σφάλμα μεροληψίας** (bias error) και δεν μπορεί να μετρηθεί σε καμιά περίπτωση

Δειγματοληψία

- Εάν οι X_1, X_2, X_3, ..., X_n ανεξάρτητες και ισόνομες τυχαίες μεταβλητές για ένα δείγμα (αντιπροσωπευτικό) μεγέθους n, ως **δειγματική μέση τιμή** ορίζεται η τυχαία μεταβλητή:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Δειγματοληψία

- Εάν X_1, X_2, X_3, ..., X_n ανεξάρτητες και τυχαίες μεταβλητές για ένα δείγμα μεγέθους n, τότε η τυχαία μεταβλητή, που δίνει τη **δειγματική διασπορά** ή **διακύμανση** είναι:

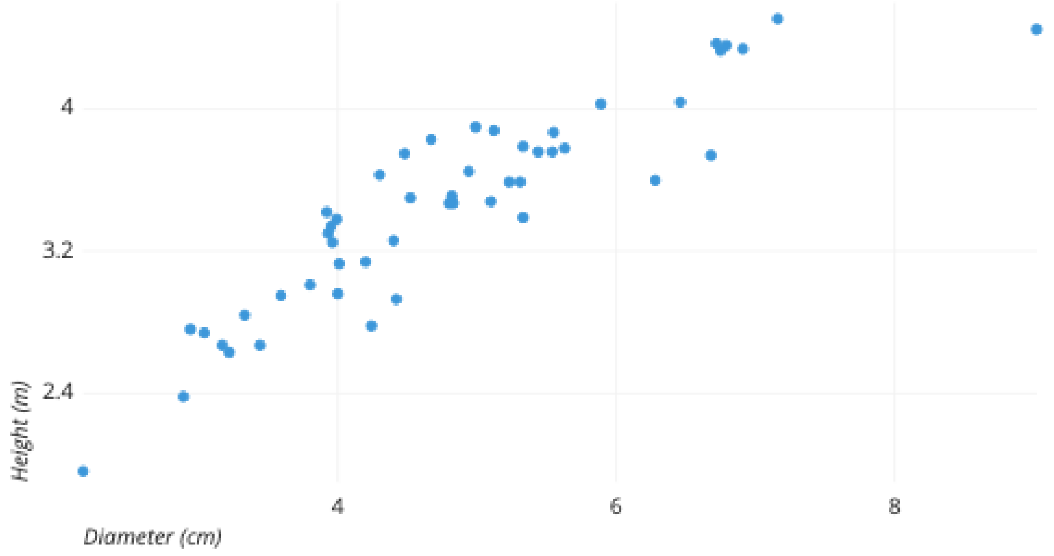
$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$$

Παρουσίαση Στατιστικών Δεδομένων

Παρουσίαση Στατιστικών Δεδομένων

Διαγράμμα διασποράς (scatter diagram)

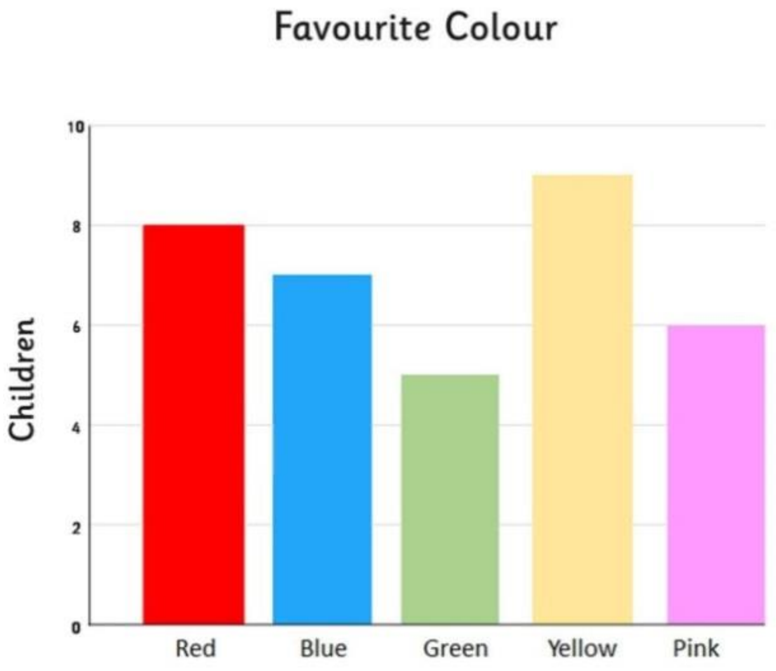
- Χρησιμοποιείται για να απεικονίσει τη σχέση μεταξύ δύο μεταβλητών



Παρουσίαση Στατιστικών Δεδομένων

Ραβδόγραμμα (bar chart)

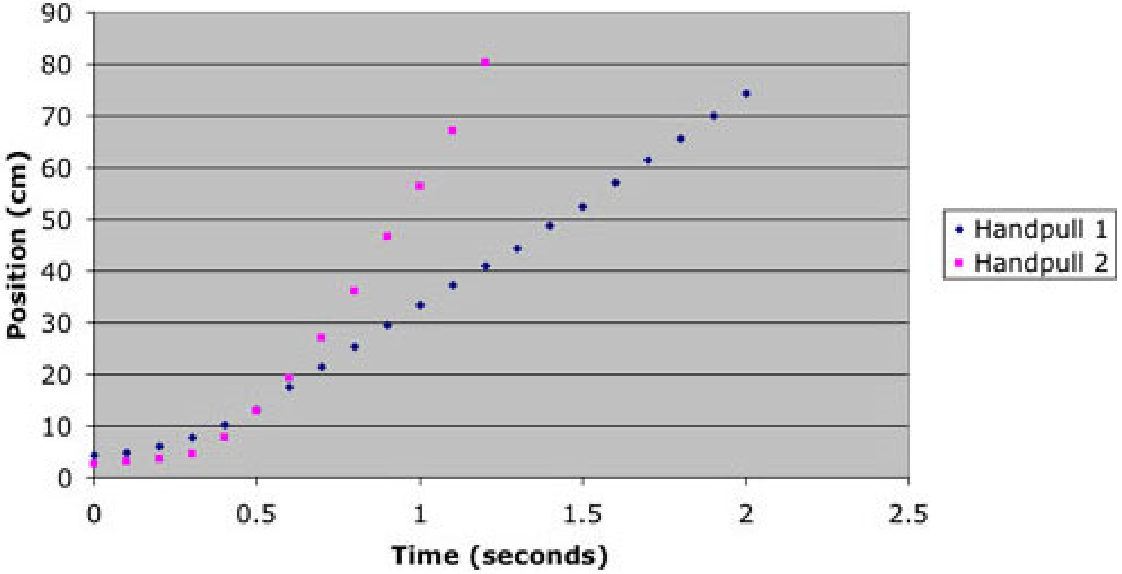
› Στον οριζόντιο άξονα υψώνονται σε ισαπέχοντα σημεία ορθογώνια παραλληλόγραμμα με ύψος ίσο με την αντίστοιχη συχνότητα



Παρουσίαση Στατιστικών Δεδομένων

Χρονολογικό διάγραμμα (time plot, run chart)

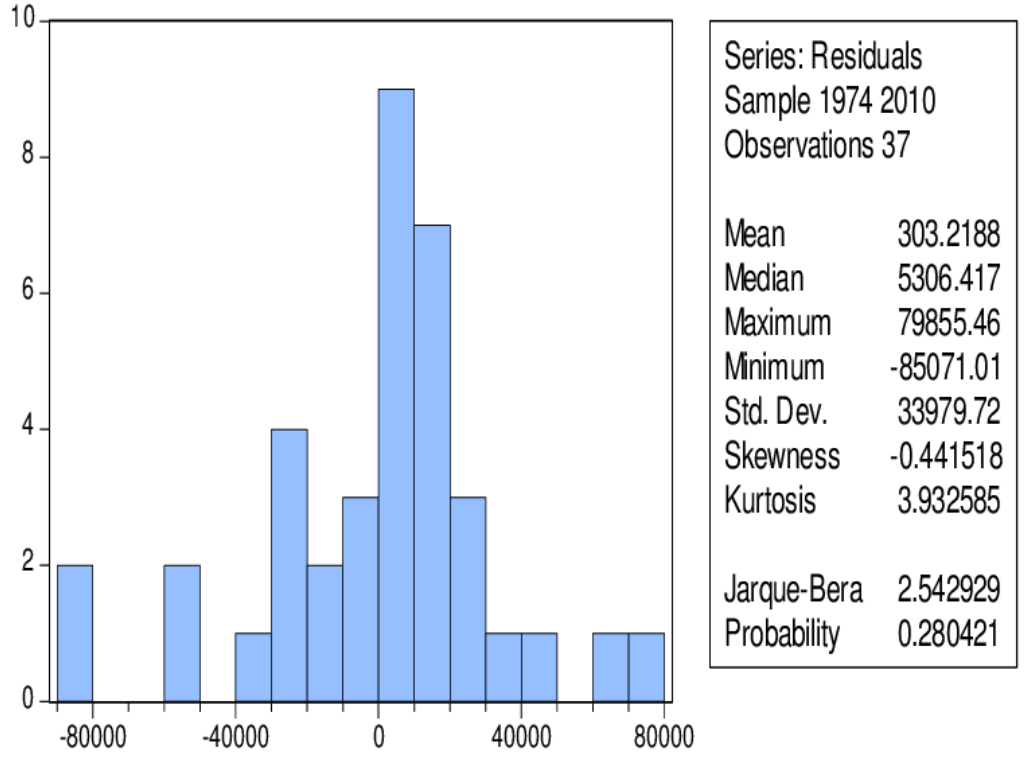
- Τα δεδομένα απεικονίζονται στην πάροδο του χρόνου.
- Στον οριζόντιο άξονα μετρούμε τον χρόνο και στον κάθετο την τιμή του μεγέθους, οπότε για κάθε παρατήρηση παίρνουμε ένα σημείο στο γράφημα



Παρουσίαση Στατιστικών Δεδομένων

Ιστόγραμμα (histogram)

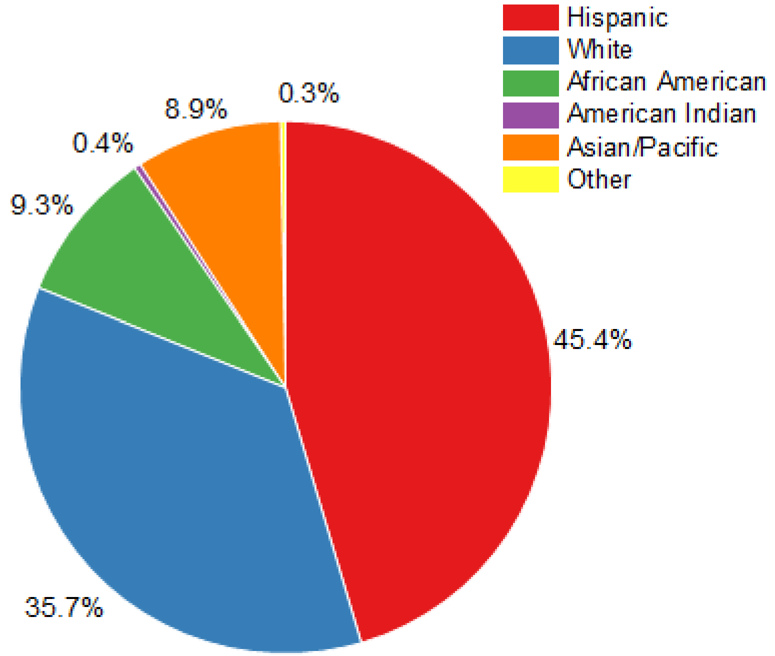
- Αποτελεί μια ακολουθία ορθογωνίων παραλληλογράμμων, όσα και οι τάξεις της κατανομής.
- Κάθε παραλληλόγραμμο έχει πλάτος και ύψος ίσα με το εύρος και τη συχνότητα της αντίστοιχης κλάσης.
- Το σχήμα του αναπαριστά την κατανομή πιθανότητας ή πυκνότητα πιθανότητας



Παρουσίαση Στατιστικών Δεδομένων

Κυκλικό διάγραμμα (pie chart)

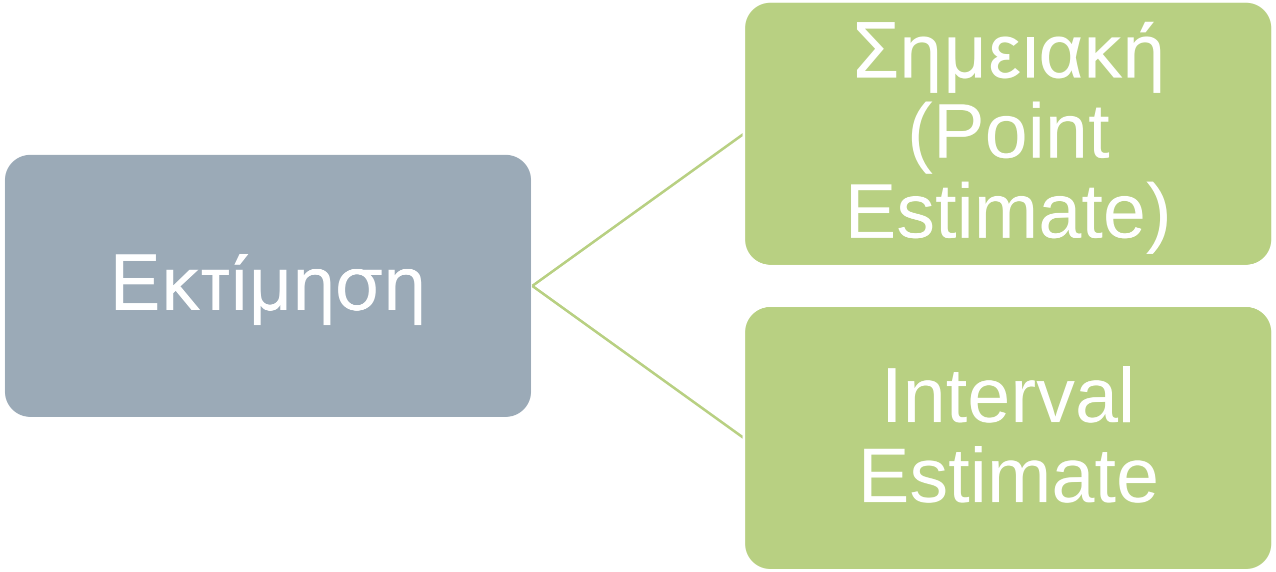
- Χρησιμοποιείται για την παρουσίαση της ποσοστιαίας αναλογίας των διαφόρων κατηγοριών ενός ολικού μεγέθους



Εκτιμητική

Εκτιμητική

- Η μέση τιμή και η διακύμανση, όταν αναφέρονται στον πληθυσμό, ονομάζονται **παράμετροι** (parameters)
- Οι παράμετροι είναι σταθερές και σχεδόν πάντα άγνωστες ποσότητες



Εκτιμητική

- Το διάστημα εμπιστοσύνης (Confidence Interval – CI) είναι ένα διάστημα εκτίμησης μιας παραμέτρου
- Αντί να εκτιμούμε την παράμετρο με μία μόνο τιμή, δίνουμε μαζί και το διάστημα πιθανότητας για την παράμετρο α

$$P(u \leq \Theta \leq v) = 1\text{-}\alpha, \text{ με } 0 \leq \alpha \leq 1$$

  – Το [u,v] ονομάζεται διάστημα εμπιστοσύνης (confidence interval) με **πιθανότητα** ή **επίπεδο εμπιστοσύνης** (level of confidence) 1-α.
  – Οι τιμές u, v ονομάζονται **όρια εμπιστοσύνης** (confidence limits) του διαστήματος και η πιθανότητα α ονομάζεται **επίπεδο σημαντικότητας** (level of significance).

Εκτιμητική

- Η **στατιστική υπόθεση** είναι μία δήλωση για μια ή περισσότερες παραμέτρους του πληθυσμού, η οποία ελέγχεται με βάση τις παρατηρήσεις ενός ή περισσότερων, αντίστοιχα, τυχαίων δειγμάτων

Εκτιμητική

- Πολλές φορές κάνουμε μία υπόθεση για να την απορρίψουμε

Ορίζουμε:

- Τη μηδενική υπόθεση (null hypothesis), η οποία είναι αυτή που περιγράφει την υπάρχουσα κατάσταση (status quo) και συμβολίζεται με H_0. Η μηδενική υπόθεση είναι αυτό που περιμένουμε
- Την εναλλακτική υπόθεση (alternative hypothesis), η οποία είναι αντίθετη, ασυμβίβαστη της μηδενικής και η οποία υπαγορεύει αλλαγή ενεργειών. Η εναλλακτική υπόθεση συμβολίζεται με H_1
- Οι δύο υποθέσεις είναι ξένες δηλαδή αμοιβαία αποκλειόμενες

Εκτιμητική

Διαδικασία Ελέγχου μιας Στατιστικής Υπόθεσης

1. Προσδιορίζουμε τις υποθέσεις, οι οποίες γίνονται δεκτές.
2. Εξειδικεύουμε την μηδενική και την εναλλακτική υπόθεση, έτσι ώστε οι δύο υποθέσεις να διαμερίζουν το σύνολο των δυνατών περιπτώσεων σε δύο υποσύνολα ξένα μεταξύ τους.
3. Ορίζουμε το κριτήριο αποφάσεως για την αποδοχή, ή την απόρριψη της μηδενικής υπόθεσης, δηλαδή ορίζουμε κανόνες απόφασης. Για τον σκοπό επιλέγουμε:
    1. Τη στατιστική του ελέγχου (test statistic), της οποίας η κατανομή πιθανοτήτων προσδιορίζεται από τις υποθέσεις, που γίνονται δεκτές και τη μηδενική υπόθεση και
    2. Το επίπεδο σημαντικότητας α του ελέγχου.

Εκτιμητική

4. Παίρνουμε την απόφαση ως εξής:
   - Αν η τιμή της στατιστικής του ελέγχου βρίσκεται στην περιοχή αποδοχής H_0, την δεχόμαστε, διαφορετικά,
   - αν τα αποτελέσματα του δείγματος είναι σημαντικά διαφορετικά, τότε θα πρέπει να απορρίψουμε τη μηδενική υπόθεση προς χάριν της εναλλακτικής.

Εκτιμητική

› Υπάρχουν 4 ξένες μεταξύ τους δυνατές περιπτώσεις:
  – Δεχόμαστε την H_0 και η H_0 είναι ορθή
  – Δεχόμαστε την H_0 και η H_0 είναι λανθασμένη
  – Απορρίπτουμε την H_0 και η H_0 να είναι ορθή
  – Απορρίπτουμε την H_0 και η H_0 να είναι λανθασμένη

- Όταν απορρίπτουμε την H_0, ενώ είναι σωστή, τότε έχουμε σφάλμα τύπου Ι. Η πιθανότητα του σφάλματος αυτού ισούται με το επίπεδο σημαντικότητας α του ελέγχου
- Όταν δεχόμαστε την H_0, ενώ είναι λανθασμένη, τότε έχουμε σφάλμα τύπου ΙΙ. Η πιθανότητα του σφάλματος αυτού συμβολίζεται με β

Εκτιμητική

- Ένα πρόβλημα ελέγχου ονομάζεται **δίπλευρο** (two – sided), αν οι αποκλίσεις προς οποιαδήποτε κατεύθυνση από τη μηδενική υπόθεση επιβάλλουν διαφορετική ενέργεια, από αυτήν που υπαγορεύεται από την μηδενική υπόθεση
- Ένα πρόβλημα ελέγχου ονομάζεται **μονόπλευρο** (one – sided), αν μόνον οι αποκλίσεις προς μια κατεύθυνση από τη μηδενική υπόθεση επιβάλλουν διαφορετική ενέργεια από ότι στην μηδενική

Εκτιμητική

- Μετά την παρατήρηση των στοιχείων του δείγματος μπορούμε να προσδιορίσουμε το ελάχιστο επίπεδο σημαντικότητας, στο οποίο μπορεί να απορριφθεί η μηδενική υπόθεση και το οποίο ονομάζεται **p-τιμή** (pvalue, prob value) ή **πιθανότητα σημαντικότητας** (significance probability) του ελέγχου.
  - Όταν απορρίπτουμε την μηδενική υπόθεση, λέμε ότι η διαφορά ανάμεσα στην παρατηρούμενη μέση τιμή του δείγματος και την μ_0 είναι στατιστικά σημαντική και με πιθανότητα σφάλματος α και δεν μπορεί να αποδοθεί στις διακυμάνσεις της δειγματοληψίας.
  - Όταν η p-τιμή του ελέγχου είναι μικρότερη από α, έχουμε μεγαλύτερη εμπιστοσύνη για την στατιστική σημαντικότητα της παρατηρούμενης διαφοράς και συνεπώς για την ισχύ της H_1.

Συσχέτιση

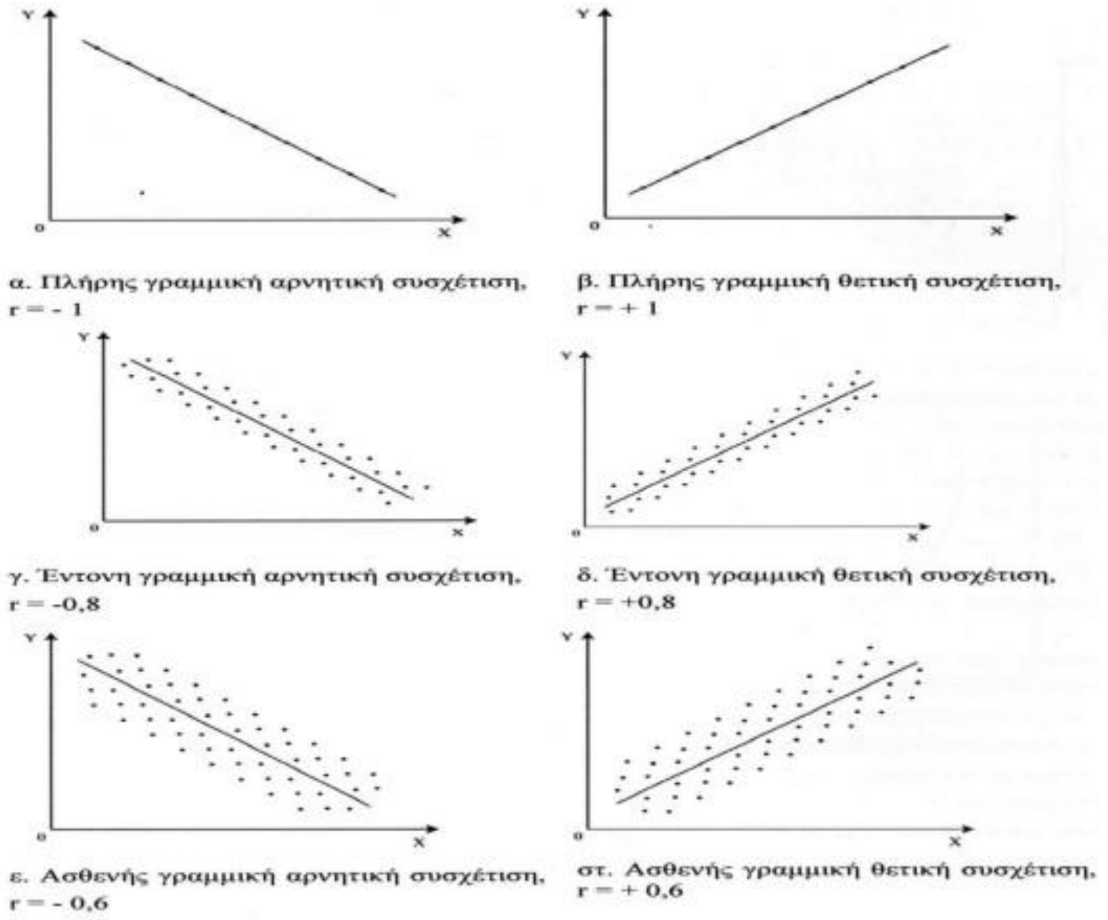
Συσχέτιση

- Ο **συντελεστής συσχέτισης** (correlation coefficient) μετρά το βαθμό, κατά τον οποίο δύο μεταβλητές σχετίζονται γραμμικά μεταξύ τους
- Δεν είναι τίποτε άλλο παρά μόνο ένα μέτρο γραμμικής συμμεταβολής δύο τυχαίων μεταβλητών

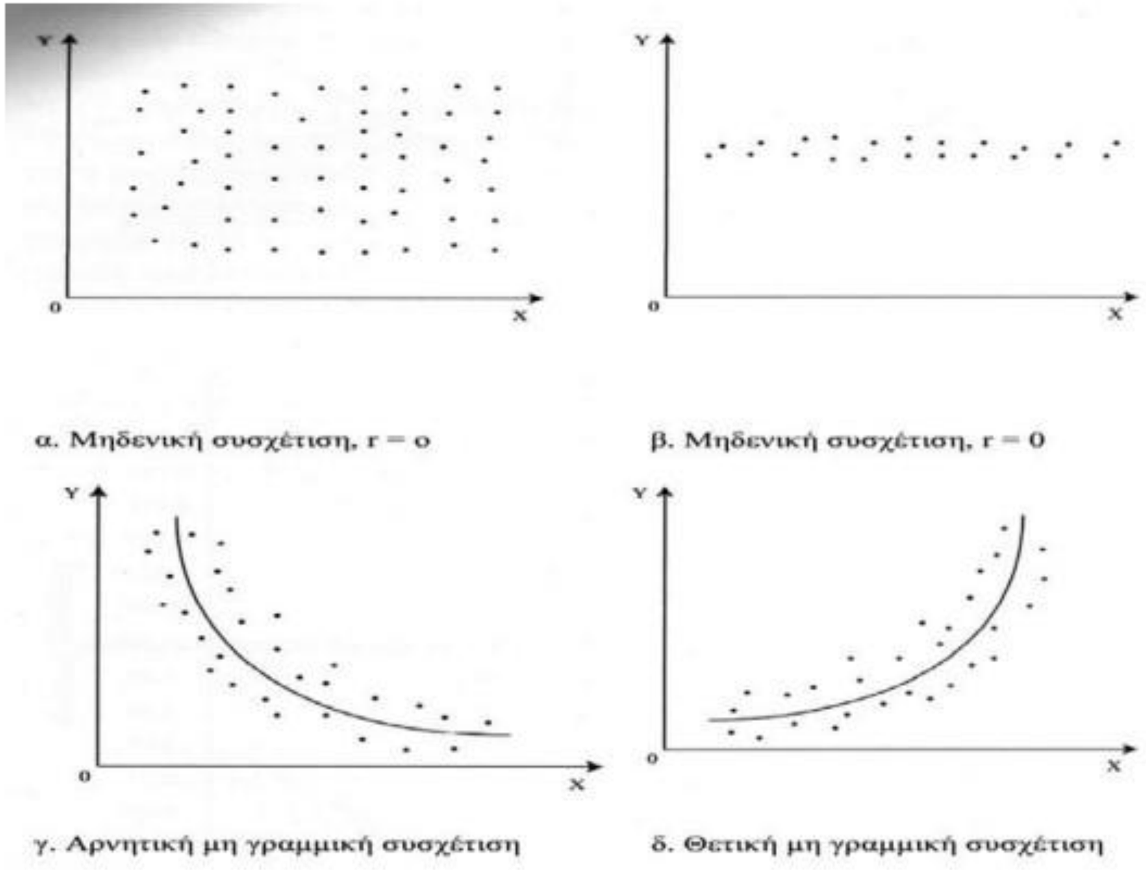
Συσχέτιση

- Η γραμμική συσχέτιση μεταξύ δύο μεταβλητών μπορεί να είναι θετική, αρνητική ή οι μεταβλητές να μη συσχετίζονται καθόλου
  - Θετική συσχέτιση υπάρχει όταν δύο μεταβλητές τείνουν να μεταβάλλονται προς την ίδια κατεύθυνση. Στην περίπτωση αυτή οι τιμές και των δύο μεταβλητών τείνουν να αυξάνονται ή να μειώνονται
  - Αρνητική συσχέτιση υφίσταται, όταν δύο μεταβλητές τείνουν να μεταβάλλονται προς αντίθετη κατεύθυνση. Στην περίπτωση αυτή όταν οι τιμές της μιας μεταβλητής αυξάνονται οι τιμές της άλλης μειώνονται και αντίστροφα
  - Μηδενική συσχέτιση παρατηρείται, όταν οι μεταβολές των τιμών της μιας μεταβλητής δεν συνοδεύονται καθόλου με τις μεταβολές των τιμών της άλλης, ή με άλλα λόγια όταν οι μεταβλητές είναι ανεξάρτητες μεταξύ τους

Συσχέτιση


α. Πλήρης γραμμική αρνητική συσχέτιση, r = - 1
β. Πλήρης γραμμική θετική συσχέτιση, r = + 1
γ. Έντονη γραμμική αρνητική συσχέτιση, r = -0,8
δ. Έντονη γραμμική θετική συσχέτιση, r = +0,8
ε. Ασθενής γραμμική αρνητική συσχέτιση, r = - 0,6
στ. Ασθενής γραμμική θετική συσχέτιση, r = + 0,6

Συσχέτιση


α. Μηδενική συσχέτιση, r = o
β. Μηδενική συσχέτιση, r = 0
γ. Αρνητική μη γραμμική συσχέτιση
δ. Θετική μη γραμμική συσχέτιση

Ερωτήσεις???