

Ανάλυση δεδομένων στο περιβάλλον του SPSS

Λαβίδας Κωνσταντίνος
Μαθηματικός

lavidas@upatras.gr

Κανονική Κατανομή

Normal Distribution

Κανονική Κατανομή

- Πολύ σημαντική κατανομή:
 - Περιγράφει πολλά φαινόμενα
 - Αποτελεί τη θεωρητική βάση για την επαγωγική στατιστική

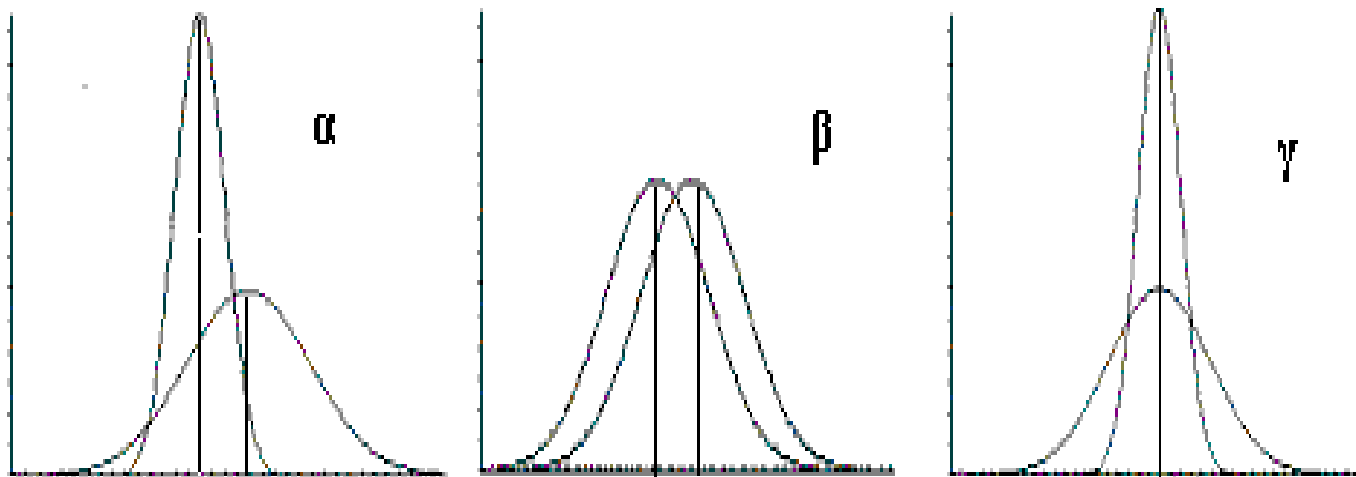
Η συνάρτηση της κανονικής κατανομής

- Η συνάρτηση που αντιστοιχεί στις καμπύλες της κανονικής κατανομής είναι:

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

- Με παραμέτρους:
 - σ : πληθυσμιακή τυπική απόκλιση και
 - μ : πληθυσμιακή μέση τιμή των τιμών της μεταβλητής.
- Η συνάρτηση αυτή και επομένως η κανονική κατανομή είναι μια **οικογένεια κατανομών**: Οι διαφορετικές τιμές των παραμέτρων μ και σ , όπως αυτές προσδιορίζουν μοναδικά την κάθε μεταβλητή, διαφοροποιούν τα μέλη της οικογένειας των κατανομών.

Κανονική κατανομή και παράμετροι



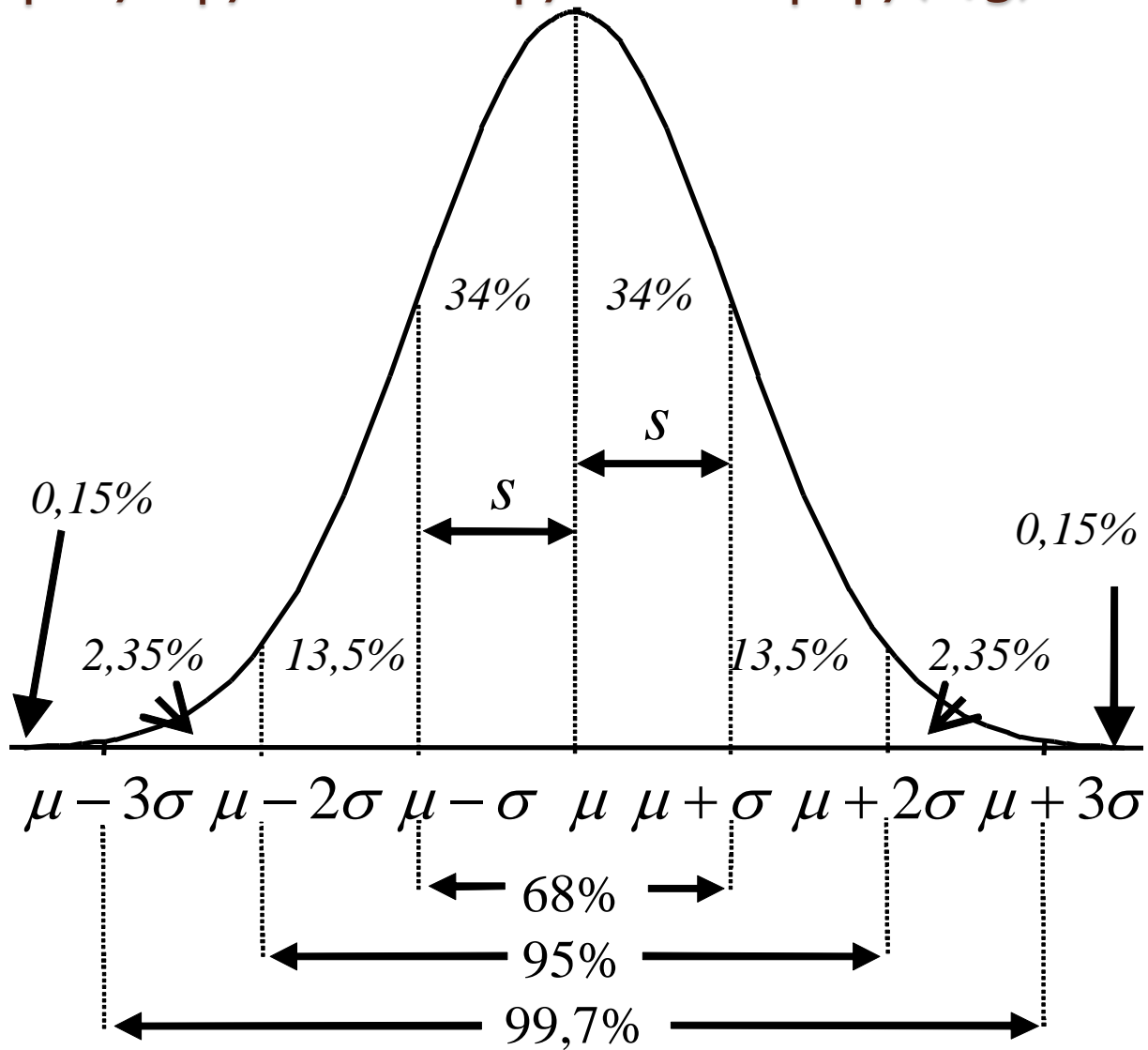
Περιπτώσεις διαφορών δύο κανονικών κατανομών

- (α) Διαφορετικές μέσες τιμές και διακυμάνσεις,
- (β) Διαφορετικές μέσες τιμές και ίσες διακυμάνσεις. Μεγαλώνοντας (μικραίνοντας) τη μέση τιμή η καμπύλη μετακινείται δεξιά (αριστερά).
- (γ) Ίσες μέσες τιμές και διαφορετικές διακυμάνσεις. Μεγαλώνοντας (μικραίνοντας) τη διακύμανση η καμπύλη απλώνεται (μαζεύεται).

Ιδιότητες της κανονικής κατανομής (1/3)

- Το εμβαδόν της περιοχής που περικλείεται από τον οριζόντιο άξονα και από την καμπύλη είναι ίσο με 1.
- Η κατανομή είναι συμμετρική ως προς τη μέση τιμή μ .
- Η μέση τιμή μ ταυτίζεται με τη διάμεσο και την επικρατούσα τιμή
- Συντελεστές: ασυμμετρίας και κύρτωσης σχεδόν μηδέν.

Ιδιότητες της κανονικής κατανομής (2/3)

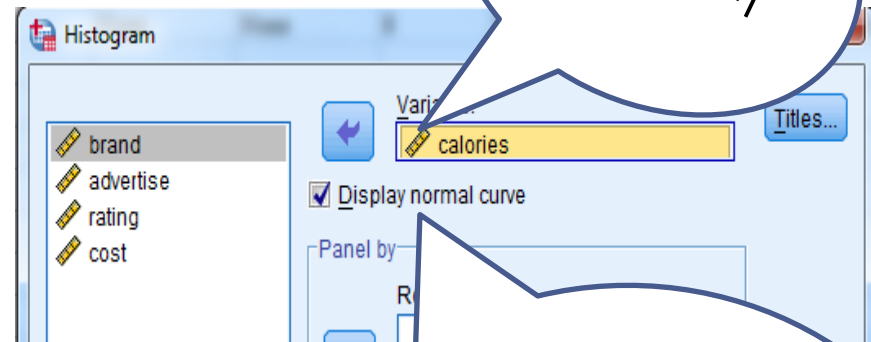
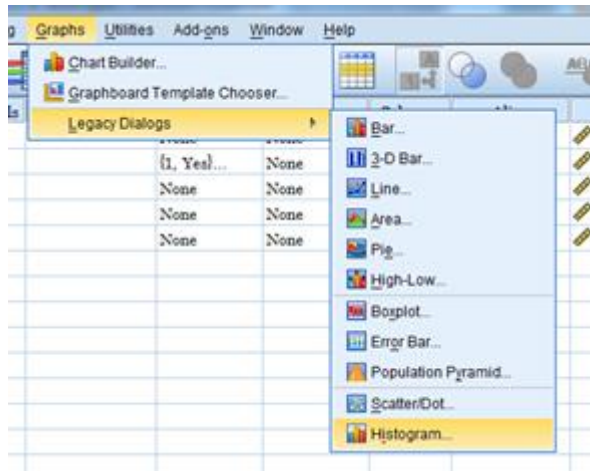


Ιδιότητες της κανονικής κατανομής (3/3)

- το 68% περίπου των παρατηρήσεων βρίσκεται στο διάστημα, $\mu - \sigma$ και $\mu + \sigma$.
- το 95% περίπου των παρατηρήσεων βρίσκεται στο διάστημα, $\mu - 2\sigma$ και $\mu + 2\sigma$.
- το 99,7% περίπου των παρατηρήσεων βρίσκεται στο διάστημα, $\mu - 3\sigma$ και $\mu + 3\sigma$.

Κατασκευή ιστογράμματος και καμπύλης (curve) συχνοτήτων με το SPSS

- Graphs- Legacy Dialogs- Histogram



Μεταβλητή

Για να εμφανιστεί η καμπύλη που αντιστοιχεί στην κατανομή των τιμών της μεταβλητής, όπως αυτή προσδιορίζεται από την μέση τιμή και την τυπική απόκλιση

Τυποποιημένες ή τυπικές τιμές (Z-Score)

- Οι τυπικές τιμές επιτρέπουν την καλύτερη κατανόηση των ατομικών περιπτώσεων σε μια κατανομή.
- Μας δίνουν τη δυνατότητα να συγκρίνουμε τιμές διαφόρων μετρήσεων που ακολουθούν την κανονική κατανομή.
- Οι τυποποιημένες τιμές υποδεικνύουν πόσες τυπικές αποκλίσεις είναι μία τιμή μακριά (πάνω ή κάτω) από την μέση τιμή.
 - Μετασχηματίζουμε τις τιμές της κατανομής σε νέες τιμές με μονάδα μέτρησης την τυπική απόκλιση της κατανομής.
 - Ακατέργαστα δεδομένα (X_i) → τυπικές τιμές (Z_i)

$$Z_i = \frac{X_i - \mu}{\sigma}$$

- Τυπικές τιμές (Z_i) → Ακατέργαστα δεδομένα (X_i)

$$X_i = Z_i \cdot (\sigma) + \mu$$

Ιδιότητες των τυπικών τιμών

- Θετική τιμή δείχνει τιμή πάνω (μετά) από τη μέση τιμή.
- Αρνητική τιμή δείχνει τιμή κάτω (πριν) από τη μέση τιμή
- Οι περισσότερες από τις τυπικές τιμές (περίπου 99%) βρίσκονται μεταξύ -3 και 3 .

Τυπικές τιμές

- Παράδειγμα: Αν η επίδοση σε ένα μάθημα έχει μέση τιμή ($\mu=50$) και τυπική απόκλιση ($\sigma=10$) να βρεθεί:
 - Η τυπική τιμή του μαθητή με σκορ (60).
 - Αν η τυπική τιμή ενός μαθητή είναι $z=1,8$, ποια είναι η επίδοση του μαθητή.

Τυπικές τιμές

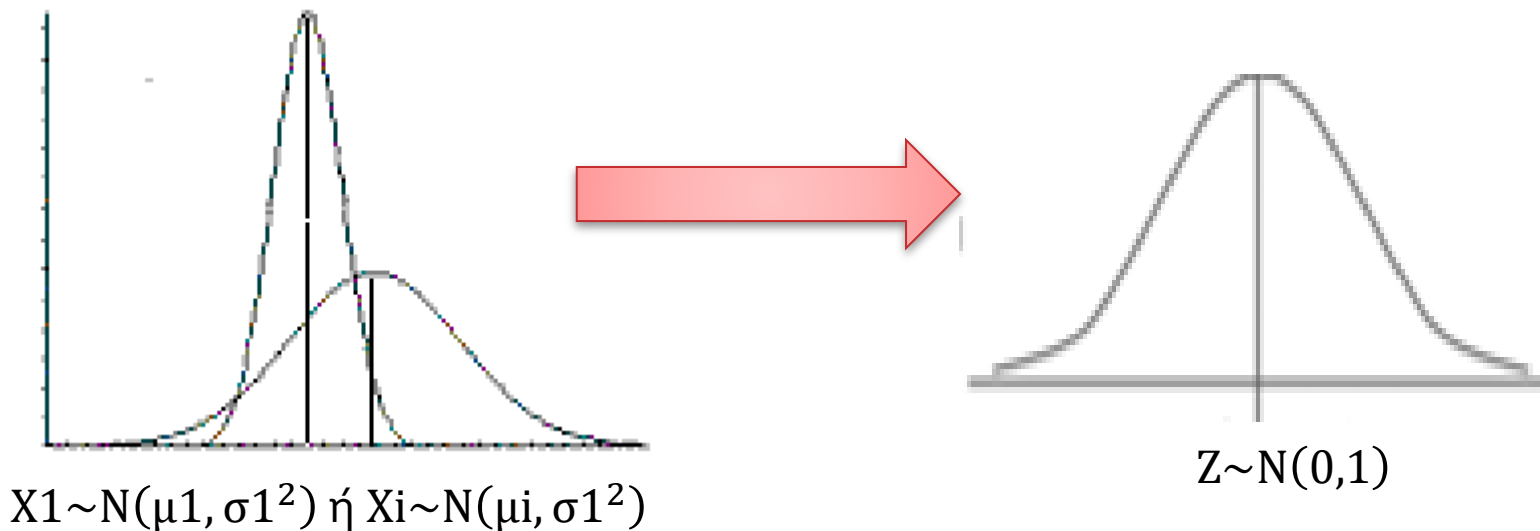
- Απάντηση: Αν η επίδοση σε ένα μάθημα έχει μέση τιμή ($\mu=50$) και τυπική απόκλιση ($\sigma=10$) να βρεθεί:
 - Η τυπική τιμή του μαθητή με σκορ 60.
 - $Z = \frac{60-50}{10} = 1$
 - Αν η τυπική τιμή ενός άλλου μαθητή είναι $z=1,8$, ποια είναι η επίδοση x του μαθητή.
 - $z = 1,8 = \frac{x-50}{10}$, δηλαδή $x=68$

Η τυπική κανονική κατανομή

- Η κανονική κατανομή μπορεί να έχει οποιαδήποτε μέση τιμή και τυπική απόκλιση και έτσι υπάρχει άπειρος αριθμός τέτοιων κατανομών.
 - Για διευκόλυνση στη χρήση της κανονικής κατανομής με έναν ενιαίο τρόπο, γίνεται μετατροπή της κανονικής κατανομής σε τυπική κανονική κατανομή.
- Έστω οι τιμές μιας τυχαίας μεταβλητής X που ακολουθούν την κανονική κατανομή με μέση τιμή μ και τυπική απόκλιση σ .
 - $X \sim N(\mu, \sigma^2)$
- Ο μετασχηματισμός $Z_i = \frac{x_i - \mu}{\sigma}$ ορίζει μια καινούργια τυχαία μεταβλητή που ακολουθεί την κανονική κατανομή με μέση τιμή $z_\mu = 0$ και τυπική απόκλιση $Z_\sigma = 1$.
 - $Z = \frac{x - \mu}{\sigma} \sim N(0, 1)$
 - Η κατανομή αυτή καλείται **τυπική κανονική κατανομή**

Η τυπική κανονική κατανομή

- Ο μετασχηματισμός αυτός μας επιτρέπει να περάσουμε σε συγκεκριμένη καμπύλη (τυπική κανονική) για όλες τις μεταβλητές X_i που ακολουθούν την κανονική κατανομή



Διάφορες Κανονικές
κατανομές για διαφορετικά
 μ και σ

Τυπική κανονική κατανομή για
 $\mu=0$ και $\sigma=1$