ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΠΕΡΙΓΡΑΦΙΚΗ και ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Εισήγηση 3Α: Η Κανονική Κατανομή

Διδάσκων: Δαφέρμος Βασίλειος
ΤΜΗΜΑ ΠΟΛΙΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
ΣΧΟΛΗΣ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

BY NC ND
Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται στην άδεια χρήσης **Creative Commons** και ειδικότερα

Αναφορά – Μη εμπορική Χρήση – Όχι Παράγωγο Έργο 3.0 Ελλάδα

(Attribution – Non Commercial – Non-derivatives 3.0 Greece)

[ή επιλογή ενός άλλου από τους έξι συνδυασμούς]

[και αντικατάσταση λογότυπου άδειας όπου αυτό έχει μπει (σελ. 1, σελ. 2 και τελευταία)]

- Εξαιρείται από την ως άνω άδεια υλικό που περιλαμβάνεται στις διαφάνειες του μαθήματος, και υπόκειται σε άλλου τύπου άδεια χρήσης. Η άδεια χρήσης στην οποία υπόκειται το υλικό αυτό αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Δευτέρα 7-4-14
ΠΕΡΙΓΡΑΦΙΚΗ ΚΑΙ ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ
Εισηγητής: Βασίλης Δαφέρμος, Αναπληρωτής Καθηγητής

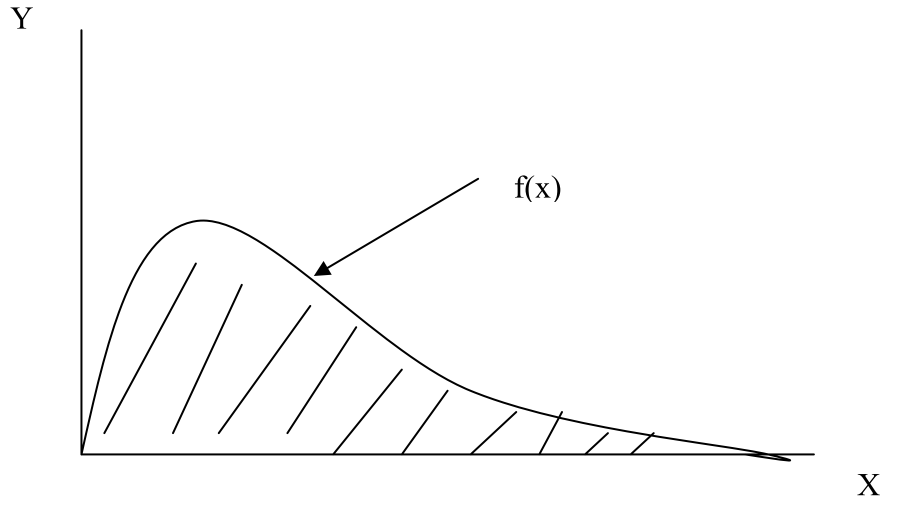
Η ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ (Normal Distribution)

Προ-έννοιες
- **Συνεχής τυχαία μεταβλητή**
- **Συνάρτηση πυκνότητας πιθανότητας**
- **Συνάρτηση κατανομής**

Ορισμός συνεχούς τυχαίας μεταβλητής
▷ Θα λέμε συνεχή τυχαία μεταβλητή X, τη μεταβλητή η οποία μπορεί να πάρει τιμές από ένα ανοικτό διάστημα της μορφής (α,β) του συνόλου των πραγματικών αριθμών.

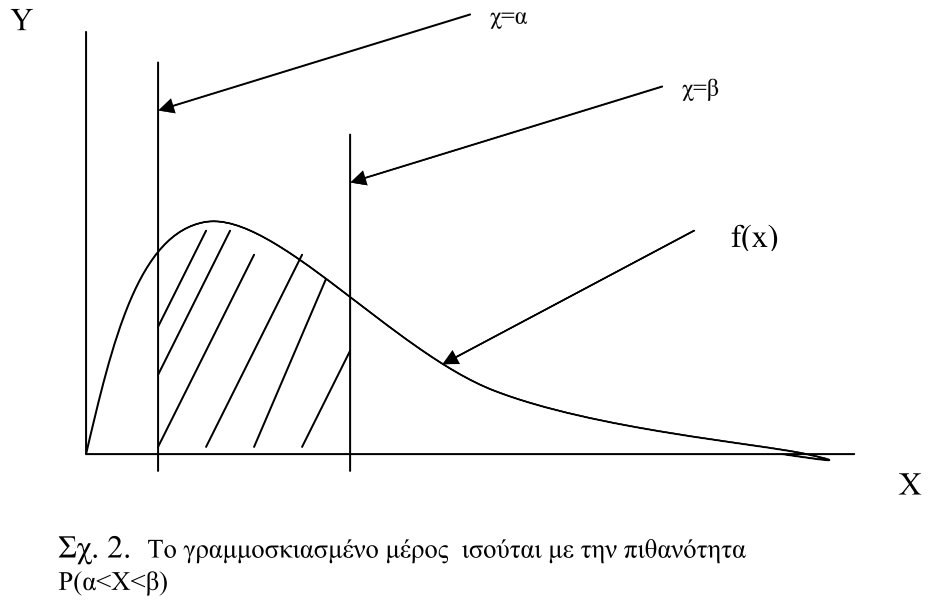
Παραδείγματα
1. Η τυχαία μεταβλητή A που συμβολίζει το χρόνο που χρειάζονται διάφοροι τύποι αεροπλάνων για να καλύψουν την απόσταση Κρήτη- Θεσσαλονίκη, αν λαμβάνει τιμές από το διάστημα (15, 60) λεπτών της ώρας, είναι μια συνεχής τυχαία μεταβλητή.
2. Η τυχαία μεταβλητή V που συμβολίζει το βαθμό που λαμβάνουν οι φοιτητές στις εξετάσεις, αν παίρνει τιμές από ένα διάστημα (0,10) μονάδων, είναι μία συνεχής μεταβλητή.

Συναρτήσεις πυκνότητας πιθανότητας
Σε κάθε τυχαία μεταβλητή X μπορούμε να αντιστοιχίσουμε κατά κάποιο τρόπο, ή να ορίσουμε μια συνάρτηση πυκνότητας πιθανότητας.
▷ Μια συνάρτηση f(x) θα λέμε πως είναι συνάρτηση πυκνότητας πιθανότητας για μια τυχαία μεταβλητή X, αν ισχύουν υποχρεωτικά και ταυτόχρονα δύο προϋποθέσεις:
1. Η f(x) λαμβάνει θετικές τιμές ή μηδέν, δηλ. $f(x) \geq 0$ για κάθε x.
2. Το εμβαδόν που περικλείεται από την καμπύλη της f(x) και τον οριζόντιο άξονα των τιμών της τυχαίας μεταβλητής X, θα πρέπει να είναι ίσο με τη μονάδα. (βλ Σχ.1).



Σχ. 1. Το γραμμοσκιασμένο μέρος ισούται με 1 τετραγωνική μονάδα

Άμεση συνέπεια του παραπάνω ορισμού είναι το γεγονός ότι, αν θέλουμε να βρούμε την πιθανότητα $P(\alpha<X<\beta)$ αυτή προφανώς θα είναι το εμβαδόν που περιορίζεται από τις κατακόρυφες ευθείες x=α και x=β, τον άξονα των τιμών της X , και την καμπύλη της συναρτήσεως f(x). (Βλ. Σχ. 5.2)



Σχ. 2. Το γραμμοσκιασμένο μέρος ισούται με την πιθανότητα $P(\alpha<X<\beta)$

Παράδειγμα

Έστω ότι η συνεχής τυχαία μεταβλητή Χ έχει συνάρτηση πυκνότητας πιθανότητας την f(x) με :

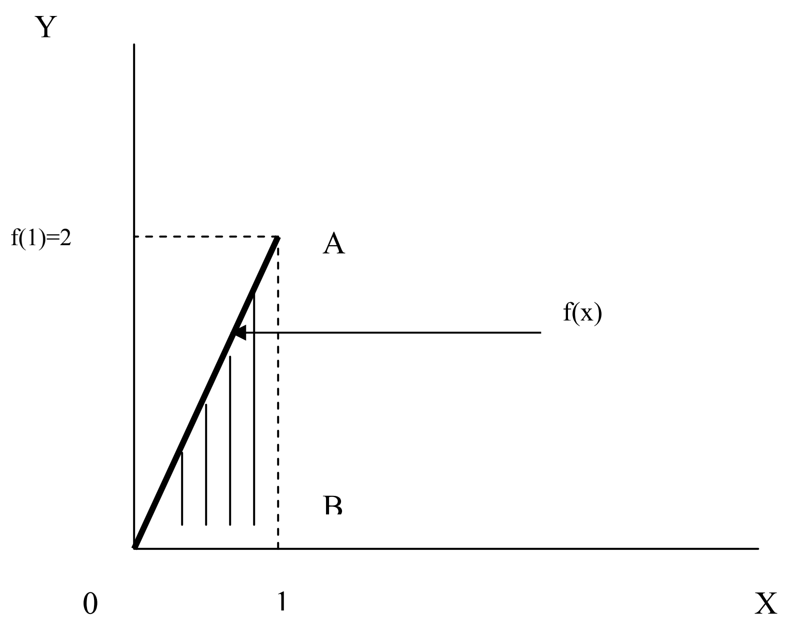
$$f(x) = \begin{cases} 2x & \alpha\nu \quad 0<x \le 1 \\ \\ 0 & o\pi o\upsilon\delta\eta\pi o\tau\varepsilon \quad \alpha\lambda\lambda o\upsilon \end{cases}$$

i. Να εξετάσετε αν όντως μπορεί η f(x) να παίξει το ρόλο της συνάρτησης πυκνότητας πιθανότητας για την τυχαία μεταβλητή Χ.

ii. Να βρείτε τις πιθανότητες $P(\frac{1}{2} < X < 1)$ και $P(X > \frac{2}{3})$

Λύση

i. Προφανώς όλες οι τιμές της f(x) είναι θετικές ή μηδέν για κάθε x. Η γραφική



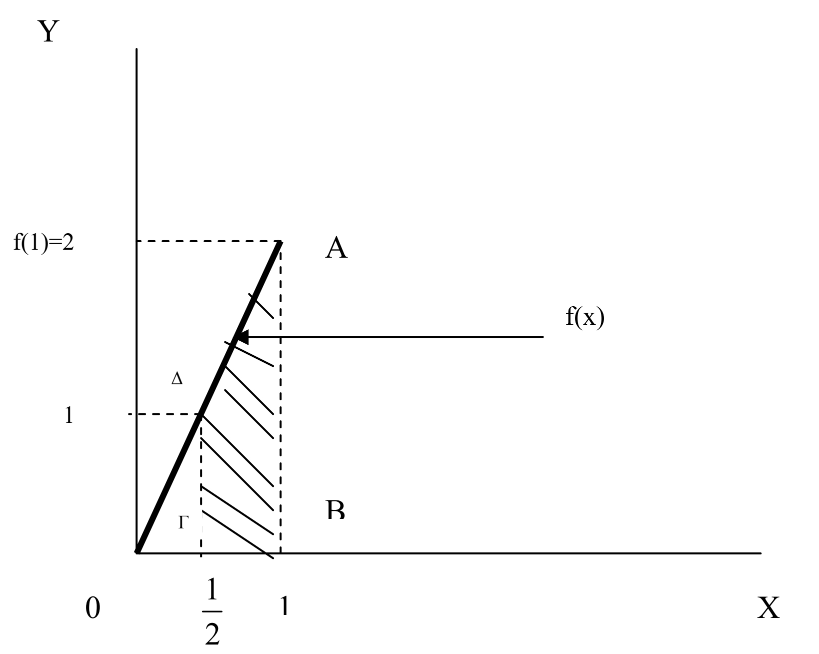
Σχ.3

αναπαράσταση της f(x) φαίνεται στο Σχ.3. Αυτό είναι φανερό από τον ορισμό της. Έχουμε δηλ. $f(x) \geq 0 \quad \forall$ x. Επίσης, από το ίδιο σχήμα, για το γραμμοσκιασμένο εμβαδόν έχουμε:

$$E_{OAB} = \frac{1}{2} \cdot 1 \cdot 2 = 1$$

Επομένως, η f(x) πληροί τις προϋποθέσεις που αναφέραμε και κατά συνέπεια είναι μια συνάρτηση πυκνότητας πιθανότητας για την τυχαία μεταβλητή X.

iii. Την πιθανότητα $P(\frac{1}{2} < X < 1)$ προφανώς θα μας την δώσει το εμβαδόν του γραμμοσκιασμένου χωρίου στο Σχ.4.



Σχ. 4

Γι'αυτό το χωρίο, το οποίο είναι ένα τραπέζιο, έχουμε:

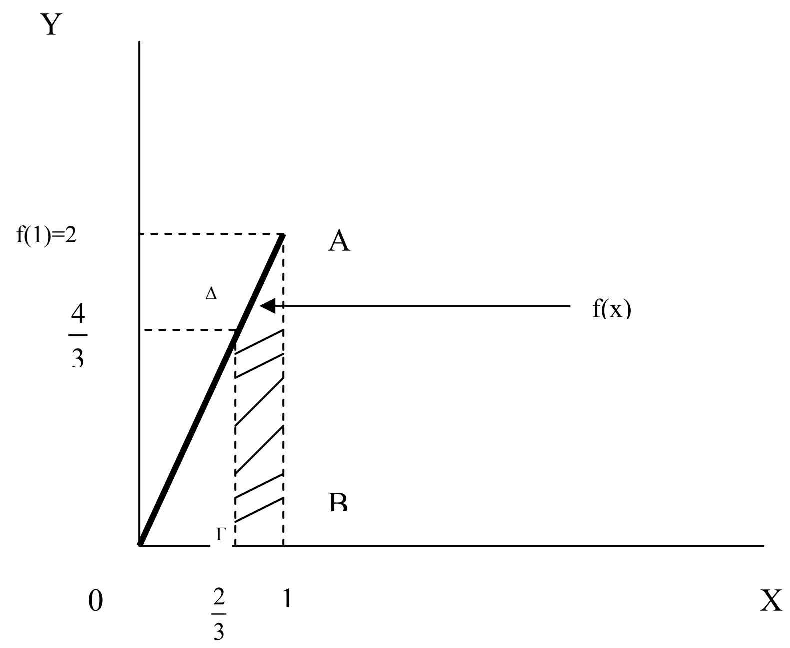
$$E_{\text{ΑΒΓΔ}} = \frac{\text{ΔΓ} + \text{ΑΒ}}{2} \cdot \text{ΓΒ} = \frac{1+2}{2} \cdot \frac{1}{2} = \frac{3}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

Δηλ. τελικά είναι:

$$P(\frac{1}{2} < X < 1) = E_{\text{ΑΒΓΔ}} = \frac{3}{4}$$

Τέλος, για τον υπολογισμό της πιθανότητας $P(X > \frac{2}{3})$ έχουμε από το Σχ. 5 τα εξής:

$$P(X > \frac{2}{3}) = E_{\mathrm{AB\Gamma\Delta}} = \frac{\Delta\Gamma + \mathrm{AB}}{2} \cdot \Gamma\mathrm{B} = \frac{\frac{4}{3} + 2}{2} \cdot \frac{1}{3} = \frac{10}{18} = \frac{5}{9}$$



Σχ. 5

Συναρτήσεις κατανομής

Όπως αντιστοιχίσαμε σε μια τυχαία συνεχή μεταβλητή X μια συνάρτηση πυκνότητας πιθανότητας f(x), με ανάλογο τρόπο θα μπορούσαμε να ορίσουμε για την ίδια τυχαία συνεχή μεταβλητή και μια συνάρτηση κατανομής.

- ▷ Θα λέμε συνάρτηση κατανομής, ή συνάρτηση αθροιστικής κατανομής για μια τυχαία συνεχή μεταβλητή X, και θα τη συμβολίζουμε με F(x), την πιθανότητα $P(X \le x)$ ή την πιθανότητα $P(X < x)$. Δηλ. ισχύει:

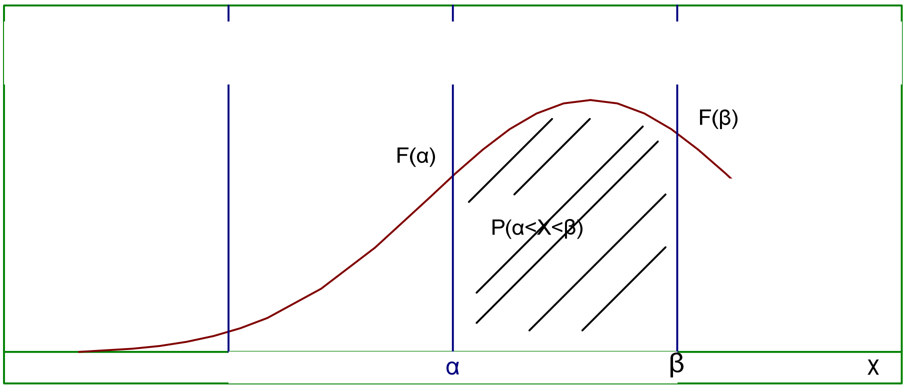
  $F(x) = P(X \le x) = P(X < x)$

  Αλλά τότε αν ζητάμε την πιθανότητα $P(\alpha < X < \beta)$ ασφαλώς αυτή η πιθανότητα θα είναι ίση με $F(\beta) - F(\alpha)$.

  Συμβολικά δηλ. θα έχουμε:

  $P[a < x < \beta] = P[a \le x \le \beta] = P[a < x \le \beta] = P[a \le x < \beta] = F(\beta) - F(a)$ (1)

  Η σχέση (1) είναι δυνατόν να παρασταθεί και γραφικά (βλ. Σχ.6).



Σχ.6. Εάν F(x) η συνάρτηση κατανομής της τυχαίας μεταβλητής X, τότε ισχύει:
P(α<X<β)=F(β)-F(α)

Ορισμός της κανονικής κατανομής

▹ Θα λέμε κανονική την κατανομή που έχει συνάρτηση πυκνότητας πιθανότητας την f(x) με:

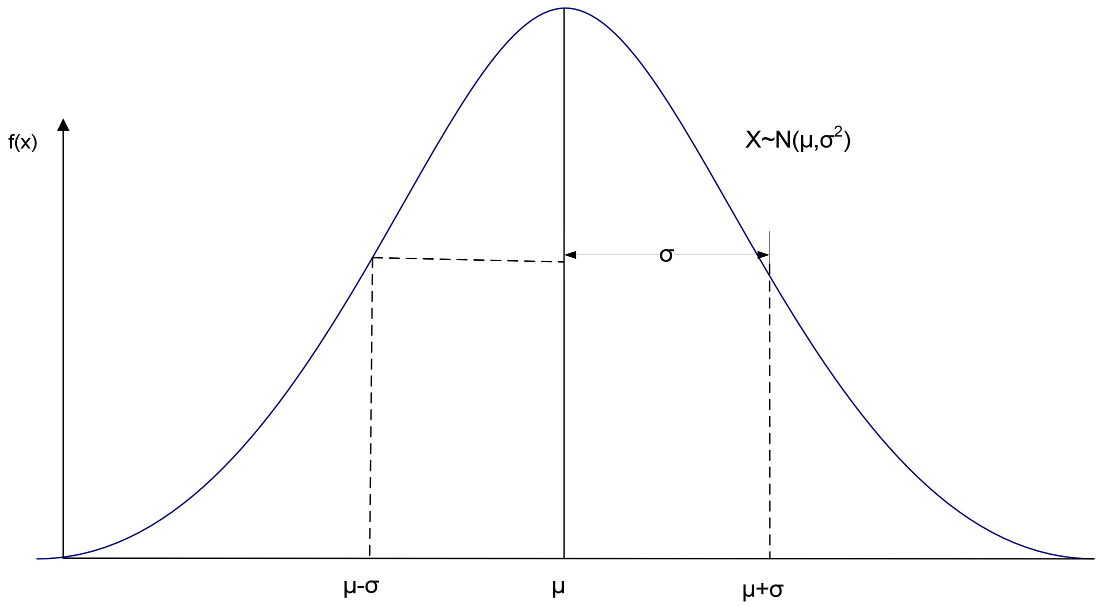
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad (2)$$

Όπου, π=3,14159 σ>0 και

$-\infty < x < +\infty$

$-\infty < \mu < +\infty$

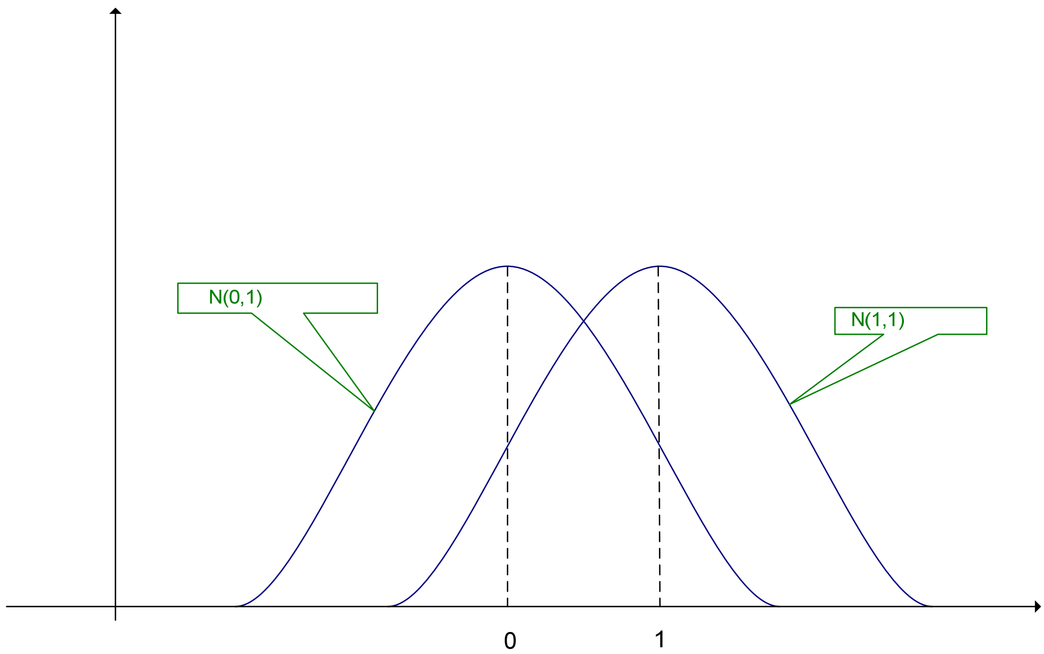
Η σχέση (2) είναι δυνατόν να αναπαρασταθεί και γραφικά (βλ. Σχ.7).



Σχ.7 Η κανονική κατανομή με μέση τιμή μ και τυπική απόκλιση σ

Όταν δε, θέλουμε να δηλώσουμε ότι μια τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή, με μέση τιμή μ και τυπική απόκλιση σ, τότε γράφουμε συμβολικά: $X \sim N(\mu,\sigma^2)$.

Εύκολα αντιλαμβάνεται κανείς, ότι σύμβολο $N(\mu,\sigma^2)$ υπαινίσσεται μια οικογένεια κατανομών, κάθε μέλος της οποίας ορίζεται από το ζεύγος των παραμέτρων μ και σ. Για παράδειγμα, αν έχουμε τα ζεύγη των κανονικών κατανομών ($\mu_1=0$, $\sigma_1^2=1$) και ($\mu_2=1$, $\sigma_2^2=1$), τότε η γραφική τους αναπαράσταση μπορεί να γίνει όπως στο Σχ. 8.



Σχ.8 Δύο κανονικές καμπύλες με διαφορετική μέση τιμή και ίδια τυπική απόκλιση

Η αξία της κανονικής κατανομής

Ανακαλύφθηκε το 1720 από το Γάλλο μαθηματικό De Moivre, ο οποίος 13 χρόνια αργότερα προέβη σε σχετική δημοσίευση. Με την κανονική κατανομή ασχολήθηκαν, εργαζόμενοι ο ένας ανεξάρτητα από τον άλλο, και δύο άλλοι μαθηματικοί: Ο επίσης Γάλλος Laplace και ο Γερμανός Gauss. Πολλές φορές δε η κανονική κατανομή ονομάζεται και κατανομή Gauss ή κατανομή Laplace.

Η κανονική κατανομή (normal distribution) πιθανότατα είναι η σημαντικότερη από όλες τις συνεχείς κατανομές και εν είναι δυνατόν να γίνει λόγος για *Παραμετρική Στατιστική*, χωρίς άμεση αναφορά στην κανονική κατανομή. Επίσης δεν είναι δυνατή η θεμελίωση και η απόδειξη του κεντρικού οριακού θεωρήματος, στο οποίο θα αναφερθούμε παρακάτω, χωρίς τη συμμετοχή της κανονικής κατανομής. Γενικά, στο χώρο της Κοινωνικής έρευνας, υπάρχουν αμέτρητες περιπτώσεις μεταβλητών που ακολουθούν, άλλοτε με **μεγαλύτερη**, και άλλοτε με **μικρότερη** προσέγγιση, την κανονική κατανομή.

Η κανονική κατανομή είναι ένα 'πρότυπο' που γεννήθηκε **μέσα από την Ιστορία** του 18ου αιώνα.

Οι μαθηματικοί της εποχής αυτής **παρατήρησαν** αργά αλλά σταθερά, **τη γέννηση** αυτού του προτύπου. Παρατήρησαν δηλ. πως τα σφάλματα των μετρήσεων είχαν μια εκπληκτική ομοιομορφία, ας την πούμε 'κανονικότητα'. Τα σφάλματα των μετρήσεων ακολουθούσαν, με άλλα λόγια, ένα μαθηματικό νόμο, ο οποίος κάποια στιγμή ονομάστηκε 'νόμος των σφαλμάτων'.

Για παράδειγμα, αν μετράμε ξανά και ξανά την ίδιο χαρακτηριστικό, ας πούμε

- το ανθρώπινο βάρος,
- το ύψος,
- την πίεση του αίματος ή
- τις τιμές της χοληστερόλης ανά μονάδα φυσιολογικού ορού,

και απεικονίσουμε γραφικά τα αποτελέσματα, κάποια στιγμή θα διαπιστώσουμε ότι οι μετρήσεις αυτές ακολουθούν, κατά προσέγγιση, την κανονική κατανομή.

Στο Σχ.7, βλέπουμε την κανονική κατανομή να έχει σχήμα καμπάνας (bell-shaped) με τα άκρα της να προσπαθούν να προσεγγίσουν το άπειρο από δεξιά και αριστερά, πάνω στον οριζόντιο άξονα, αλλά ουδέποτε να το κατορθώνουν. Το Σχ.7 μας δίνει την ευκαιρία, να ορίσουμε με γραφικό τρόπο το μήκος μιας ιδιαίτερα χρήσιμης στη στατιστική μονάδας μέτρησης, την τυπική απόκλιση. Η οριζόντια απόσταση που συνδέει το σημείο που η συνάρτηση πυκνότητας πιθανότητας της κανονικής κατανομής αλλάζει τα κοίλα της, με την κάθετο που διέρχεται από την κορυφή της καμπύλης, είναι ίση με μια τυπική απόκλιση σ.

Στο Σχ.7 παρατηρούμε επίσης ότι οι τιμές της κανονικής κατανομής συσσωρεύονται, κατά κύριο λόγο, γύρω από τη μέση τιμή, ενώ όσο προχωρούμε προς τα άκρα, οι τιμές ολοένα και αραιώνουν.

Οι ιδιότητες της κανονικής κατανομής

- Η κανονική κατανομή είναι συμμετρική, ως προς άξονα συμμετρίας την κάθετο που διέρχεται από την κορυφή της καμπύλης της και από το σημείο μ, το οποίο είναι η μέση τιμή της.
- Η μέση τιμή, η διάμεσος και η δεσπόζουσα τιμή στην κανονική κατανομή συμπίπτουν. Αυτό είναι αποτέλεσμα της συμμετρίας της κανονικής κατανομής. Ακόμη, αποτέλεσμα της συμμετρίας είναι και το γεγονός ότι τα εμβαδά δεξιά και αριστερά του άξονα συμμετρίας είναι ίσα.
- Το εμβαδόν της περιοχής που ορίζεται από την καμπύλη της κανονικής κατανομής και τον οριζόντιο άξονα είναι ίσο με 1 τετραγωνική μονάδα
- Η καμπύλη της κανονικής κατανομής τείνει να προσεγγίσει τον οριζόντιο άξονα **ασυμπτωτικά**.
- Το πεδίο ορισμού της συνάρτησης της κανονικής κατανομής είναι ολόκληρο το σύνολο των πραγματικών αριθμών R.

Η μετατροπή των τιμών της κανονικής κατανομής σε z-τιμές

Κάθε κανονική κατανομή είναι δυνατόν να μετατραπεί σε τυπική κανονική κατανομή, με βάση τον τύπο:

$$Z_i = \frac{X_i - \bar{X}}{S} \qquad (3)$$

Όπου,
Zi είναι η λεγόμενη z-τιμή,
$\bar{X}$ η μέση τιμή του δείγματος,
και S η τυπική του απόκλιση.

Παράδειγμα

Ας υποθέσουμε ότι έχουμε και πάλι το δείγμα των 15 φοιτητών και φοιτητριών του Καθηγητή Δεληβοριά:

5,6,6,7,7,8, 4,5,6,7,7,7,8,9,10

Αν θέσουμε X_1=1, X_2=6, …, X_{14}=9, X_{15}=10, τότε με βάση τον τύπο (3), και αφού $\bar{X} = 6,8$ και S=1,568 , θα έχουμε τις αντίστοιχες z-τιμές:

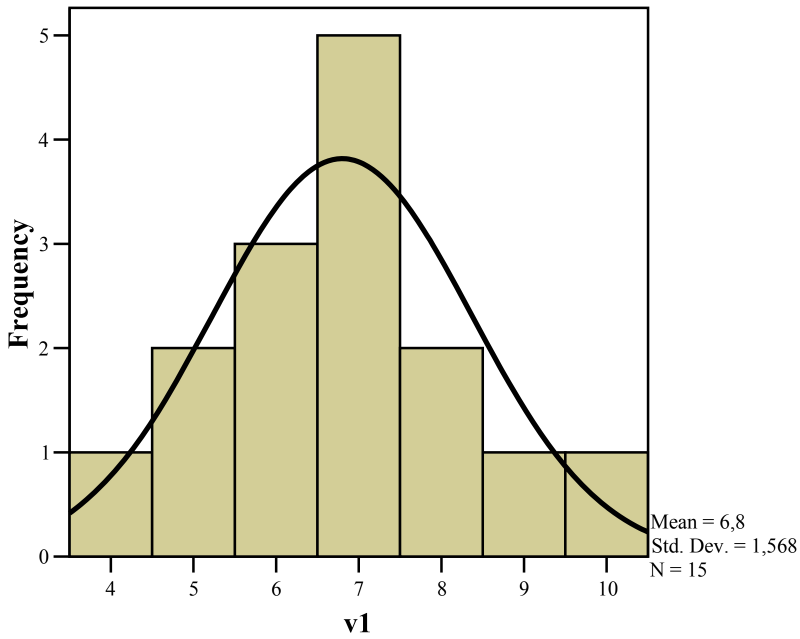
$$Z_1 = \frac{X_1 - \bar{X}}{S} = \frac{5 - 6,8}{1,568} = -1,14831$$

$$Z_2 = \frac{X_2 - \bar{X}}{S} = \frac{6 - 6,8}{1,568} = -0,51036$$

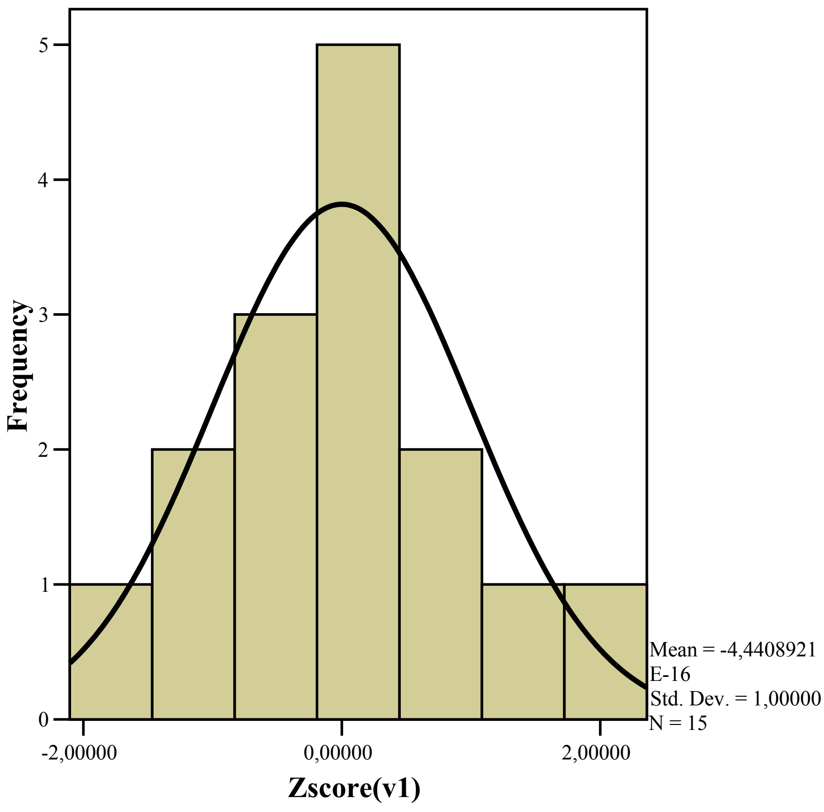
….

$$Z_{14} = \frac{X_{14} - \bar{X}}{S} = \frac{9 - 6,8}{1,568} = 1,40348$$

$$Z_{15} = \frac{X_{15} - \bar{X}}{S} = \frac{10 - 6,8}{1,568} = 2,04143$$



Σχ.11



Σχ. 12

Στο Σχ.11 παρατηρούμε ότι η γραφική αναπαράσταση της μεταβλητής V1 ακολουθεί περίπου την κανονική κατανομή με μέσο όρο μ=6,8 και τυπική απόκλιση σ=1,568.

Πως γεννιέται η τυπική κανονική κατανομή από την κανονική κατανομή….
Δηλ. πώς πάμε από το Σχ. 11 στο Σχ. 12…
Απλά μετατρέπουμε τους βαθμούς σε αντίστοιχες z-τιμές με βάση τη σχέση (3)

Στο Σχ.12 επίσης παρατηρούμε ότι περίπου κανονική είναι και η κατανομή της μεταβλητής Zv1, με μέσο όρο περίπου ίσο με μηδέν και τυπική απόκλιση ίση με ένα.

Ιδιότητες της τυπικής κανονικής κατανομής (Standardized Normal Distribution)

1. Η τυπική κανονική κατανομή είναι συμμετρική με άξονα συμμετρίας την κατακόρυφη ευθεία που διέρχεται από την κορυφή της και τέμνει τον οριζόντιο άξονα στο σημείο μ (μέση τιμή του πληθυσμού).
2. Το εμβαδόν που περικλείεται από την καμπύλη της κατανομής και τον οριζόντιο άξονα έχει εμβαδόν 1 τετραγωνική μονάδα, και αντιστοιχεί σε πιθανότητα 100%.
3. Η μέση τιμή, η διάμεσος και η δεσπόζουσα τιμή της τυπικής κατανομής συμπίπτουν και είναι ίσες με μηδέν.
4. Η τυπική της απόκλιση είναι ίση με 1.
5. Η στρεβλότητα είναι ίση με μηδέν.
6. Η κύρτωσή της είναι ίση με μηδέν, με βάση τον τύπο του Fisher που χρησιμοποιεί και το SPSS.

Συμβολικά η τυπική κανονική κατανομή αναπαρίσταται με N(0,1).

Η συνάρτηση αθροιστικής κατανομής της τυπικής κανονικής κατανομής

Η συνάρτηση αθροιστικής κατανομής της τυπικής κανονικής κατανομής συμβολίζεται με F(z) και ορίζεται από την πιθανότητα $P(Z < z) = P(-\infty < Z < z)$.

Η συνάρτηση F(z) αναπαρίσταται γραφικά όπως στο Σχ. 13. Είναι δε φανερό ότι για την F(z) ισχύει:

F(-z)=1-F(z) (4).

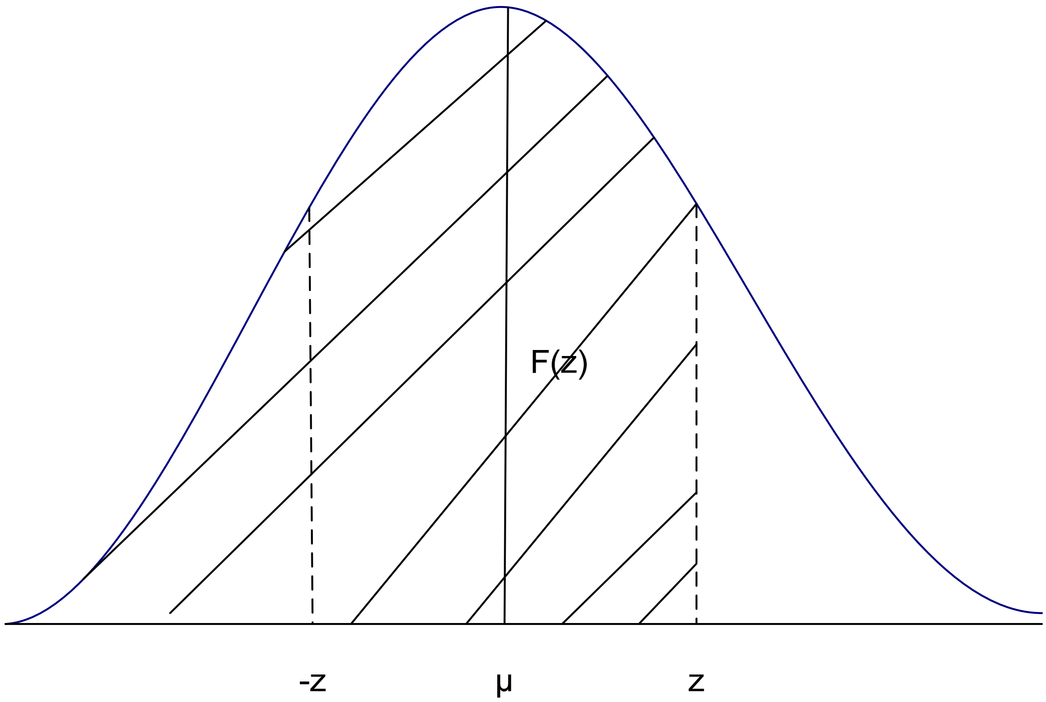
Ας υποθέσουμε τώρα πως αναζητάμε εμβαδά-πιθανότητες κάτω από την καμπύλη της F(z). Κάτι τέτοιο είναι εύκολο να υπολογιστεί με τη βοήθεια του τύπου (3), διότι μέσω αυτού του τύπου είναι δυνατός ο μετασχηματισμός των αρχικών μας δεδομένων, τα οποία μπορεί να μετρούνται σε διάφορες μονάδες, σε z-τιμές, οι οποίες στη συνέχεια με τη βοήθεια του Πίνακα της τυποποιημένης κανονικής κατανομής (βλ. φυλλάδιο που σας δόθηκε), ανάγονται σε πιθανότητες. Αλλά ας μιλήσουμε πιο συγκεκριμένα:

Ας υποθέσουμε ότι αναζητάμε την πιθανότητα, η τυχαία μας μεταβλητή X να λαμβάνει τιμές ανάμεσα στο α και το β. Συμβολικά, ας πούμε ότι ζητάμε την πιθανότητα P(α<X<β). Τότε αυτό που πρέπει να κάνουμε είναι να μετατρέψουμε αυτή την πιθανότητα, σε πιθανότητα της μορφής $P(z_\alpha < Z < z_\beta)$, ως εξής:

$$P(\alpha < X < \beta) = P(\frac{\alpha-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{\beta-\mu}{\sigma}) = P(z_\alpha < Z < Z_\beta) = F(z_\beta) - F(z_\alpha) \qquad (5)$$

Όπου,

$z_\alpha = \dfrac{\alpha-\mu}{\sigma}$ και $z_\beta = \dfrac{\beta-\mu}{\sigma}$ είναι οι τυποποιημένες τιμές της τυχαίας μας μεταβλητής X η οποία υποθέσαμε ότι ακολουθεί την κανονική κατανομή.



Σχ.13. Για το γραμμοσκιασμένο μέρος F(z) ισχύει:
F(z)=P(Z<z)

Παράδειγμα

Σε κάποια ευρωπαϊκή χώρα είναι γνωστό από μελέτες ότι το ύψος των παιδιών της προσχολικής ηλικίας ακολουθεί την κανονική κατανομή με μέση τιμή 110 cm και τυπική απόκλιση 10 cm. Συναντάμε στην τύχη ένα από τα παιδιά αυτής της ηλικίας και αυτής της χώρας. Ποια η πιθανότητα:

Α) Να έχει ύψος μεγαλύτερο από 120 cm

Β) Να έχει ύψος κάτω από 90 cm.

Γ) Το ύψος του να κυμαίνεται μεταξύ 90 και 110 cm.

Λύση

Α) Προφανώς αναζητούμε την πιθανότητα P(X>120). Αλλά γι' αυτήν την πιθανότητα έχουμε:

$$P(X>120)=P(\frac{X-110}{10}>\frac{120-110}{10})=P(Z>1)=1-P(Z<1)=1-F(1)$$

$$=1-0,8413=0,1587$$

Με άλλα λόγια η ζητούμενη πιθανότητα είναι 15,87 %.

Β) Εδώ η ζητούμενη πιθανότητα είναι P(X<90), για την οποία έχουμε:

$$P(X<90)=P(\frac{X-110}{10}<\frac{90-110}{10})=P(Z<-2)$$

$$=F(-2)=1-F(2)=1-0,9772=0,0228$$

Δηλ. η ζητούμενη πιθανότητα είναι 2,28 %.

Γ) Εδώ η ζητούμενη πιθανότητα είναι P(90<X<110), για την οποία έχουμε:

$$P(90 < X < 110) = P(\frac{90-110}{10} < \frac{X-110}{10} < \frac{110-110}{10})$$
$$= P(-2 < Z < 0) = F(0) - F(-2) = F(0) - (1 - F(2))$$
$$= F(0) - 1 + F(2) = F(2) + F(0) - 1 = 0,9772 + 0,5000 - 1$$
$$= 0,4772$$

Προφανώς, η ζητούμενη πιθανότητα είναι 47,72%.

Παράδειγμα Σύγκρισης τιμών που ανήκουν σε διαφορετικές κανονικές κατανομές

Η βαθμολογία των φοιτητών του πρώτου έτους σε κάποιο μάθημα Α ήταν κανονική,

- με μέση τιμή μ_A=70 μονάδες,
- και τυπική απόκλιση σ_A=10 μονάδες,

ενώ σε κάποιο άλλο μάθημα Β η βαθμολογία ήταν πάλι κανονική, αλλά

- με μέση τιμή μ_B=60 μονάδες
- και τυπική απόκλιση σ_B=4.

===== ================== =========== === ===== === ==

Εάν ένας φοιτητής πήρε
βαθμό X_A=80 στο μάθημα Α και
βαθμό X_B=65 στο μάθημα Β,

- **πού είναι καλύτερος, στο μάθημα Α ή στο μάθημα Β;**

============== =================== == === ========

Λύση

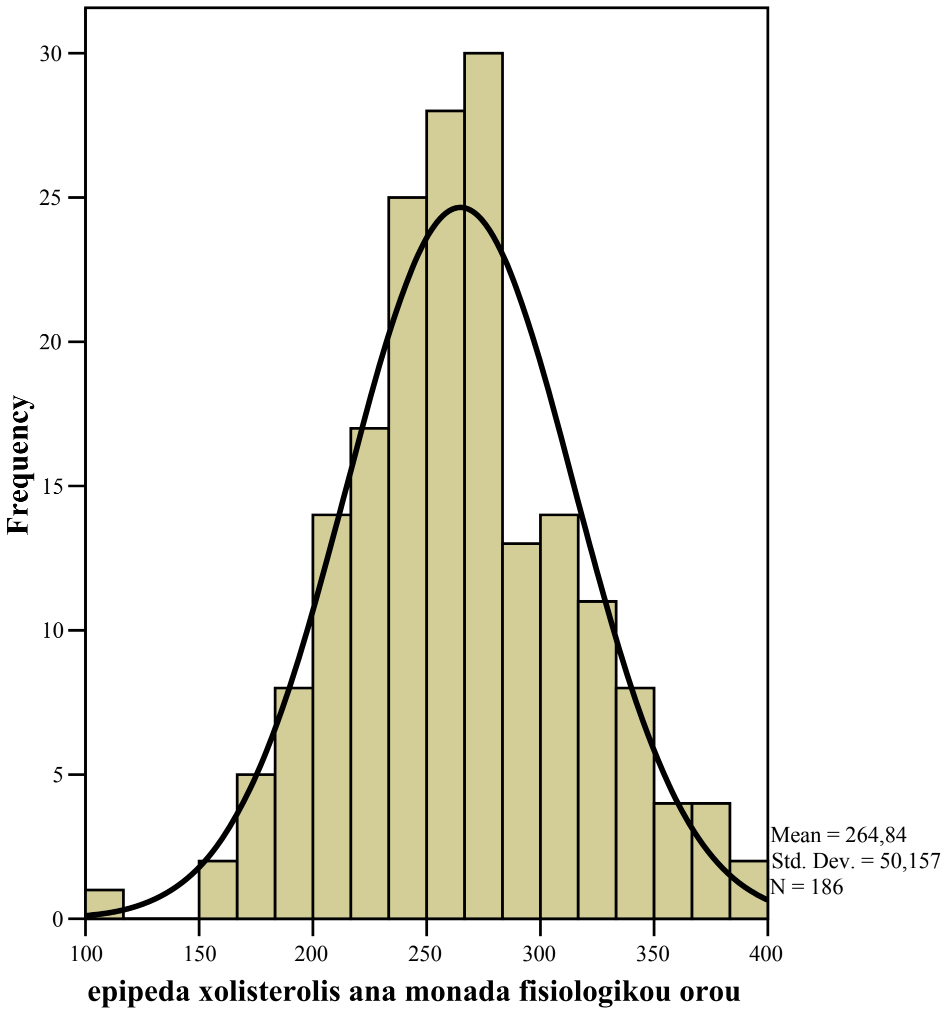
Αρχικά να επισημάνουμε ότι είναι ανάγκη να μετατραπούν οι δύο βαθμολογίες σε z-τιμές, ώστε στη συνέχεια να βρεθεί ένα κοινό μέτρο σύγκρισης. Ακόμη να επισημάνουμε ότι εάν η μία από τις δύο ή και οι δύο κατανομές, δεν ήταν κανονικές, δεν θα ήταν δυνατή η μετατροπή των τιμών τους σε τιμές τυπικής κανονικής κατανομής.

Έτσι, για το μάθημα Α έχουμε: $Z_A = \dfrac{80-70}{10} = 1$

Αντίστοιχα, για το μάθημα Β έχουμε: $Z_B = \dfrac{65-60}{4} = \dfrac{5}{4} = 1,25$

Συμπέρασμα

Αφού $Z_B > Z_B \Rightarrow$ ο φοιτητής ήταν καλύτερος στο μάθημα Β.


30
25
20
15
10
5
0
Frequency
100
150
200
250
300
350
400
epipeda xolisterolis ana monada fisiologikou orou
Mean = 264,84
Std. Dev. = 50,157
N = 186

Σχ.14

==================== =========================== ===== ====

ΓΕΝΙΚΕΣ ΑΣΚΗΣΕΙΣ

Άσκηση 1

Ο χρόνος που χρειάζεται ένας φοιτητής για να προετοιμαστεί στο μάθημα της Μεθοδολογίας των Κοινωνικών Επιστημών, στις εξετάσεις του Ιουνίου, βρέθηκε ότι προσεγγιστικά ακολουθεί την κανονική κατανομή με μέση τιμή **μ=25** ώρες και τυπική απόκλιση **σ=5** ώρες.

Α) Να υπολογιστεί το ποσοστό των φοιτητών οι οποίοι δαπανούν στην επανάληψη **περισσότερες από 30 ώρες**.

Β) Να υπολογιστεί το ποσοστό των φοιτητών οι οποίοι δαπανούν στην επανάληψη **λιγότερο από 15 ώρες**.

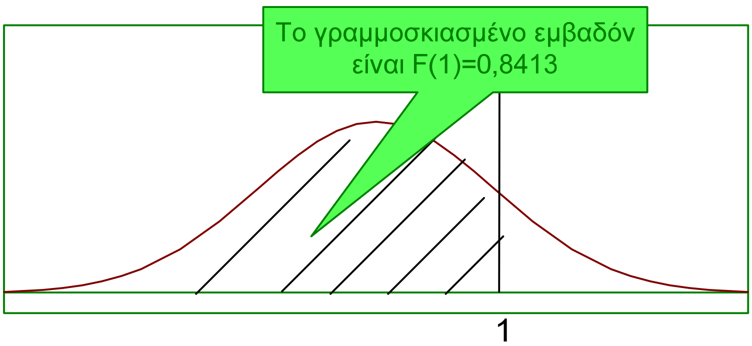
Γ) Να υπολογιστεί το ποσοστό των φοιτητών οι οποίοι δαπανούν στην επανάληψη **από 15 μέχρι 25 ώρες**.

=============== =================== ======== ======= =====

Λύση

Α) Αν συμβολίσουμε με την τυχαία μεταβλητή Χ, το χρόνο που δαπανούν οι φοιτητές για την επανάληψη του μαθήματος, προφανώς αναζητούμε την πιθανότητα $P(X > 30)$. Επίσης από τα δεδομένα του προβλήματος έχουμε ότι **μ=25 και σ=5**. Έτσι, για τη ζητούμενη πιθανότητα έχουμε:

$$P(X > 30) = P(\frac{X-25}{5} > \frac{30-25}{5}) \quad \text{[μετασχηματισμός σύμφωνα με τη σχέση 5]}$$
$$= P(Z > 1) = 1 - P(Z < 1) = 1 - F(1) \quad \text{[βλ. Σχ. 15]}$$
$$= 1 - 0,8413 = 0,1587$$



Σχ.15

Επομένως, η απάντηση είναι ότι περίπου το 15,87 % των φοιτητών δαπανά περισσότερες από 30 ώρες στην επανάληψη του μαθήματος της Μεθοδολογίας των Κοινωνικών Επιστημών.

+++++++++++++++++++++++++++ +++++++++++++++ ++++++ +++++

+++++++++++++++++++ +++++++++++++++++ +++++++++++++ ++++

Β) Εδώ, προφανώς αναζητούμε την πιθανότητα $P(X<15)$ για την οποία έχουμε:

$$P(X<15)=P(\frac{X-25}{5}<\frac{15-25}{5})$$

$$=P(Z<\frac{-10}{5}=-2)=P(Z<-2)$$

$$=F(-2)=1-F(2)=1-0,9772=0,0228$$

Δηλ. περίπου το 2,28 % των φοιτητών δαπανά στην επανάληψη του μαθήματος αυτού, χρόνο μικρότερο, των 15 ωρών.

Σημείωση

Από τον Πίνακα της τυπικής κανονικής κατανομής που σας δόθηκε βρίσκουμε ότι F(2)=0,9772.

+++++++++++++++++++++++++ +++++++++++ ++++++++++++ +++++++

Γ) Εδώ, προφανώς αναζητούμε την πιθανότητα $P(15<X<25)$ για την οποία έχουμε:

$$P(15<X<25)$$

$$=P(\frac{15-25}{5}<\frac{X-25}{5}<\frac{25-25}{5})$$

$$=P(-2<Z<0)$$

$$=F(0)-F(-2)=F(0)-[1-(F(2)]$$

$$=F(0)-1+F(2)=F(2)+F(0)-1$$

$$=0,9772+0,5000-1=0,4772$$

Συμπέρασμα

Περίπου το 47,72% των φοιτητών, δαπανά για την επανάληψη, χρόνο που κυμαίνεται από 15 μέχρι 25 ώρες.

======================= ================= ========== ==

Άσκηση 2

Από προηγούμενη έρευνα, για τις συνθήκες διαβίωσης των φοιτητών στο Ρέθυμνο Κρήτης, βρέθηκε ότι τα μηνιαία έξοδα των φοιτητών του εκεί Πανεπιστημίου, ακολουθούν προσεγγιστικά την κανονική κατανομή με μέση τιμή μ=800 ευρώ και τυπική απόκλιση σ=80 ευρώ.

Παίρνουμε στην τύχη ένα φοιτητή.

Ποια είναι η πιθανότητα ο φοιτητής:

Α) να ξοδεύει το μήνα περισσότερο από 1000 ευρώ;

Β) να ξοδεύει λιγότερο από 600 ευρώ;

Γ) να κυμαίνονται τα έξοδά του από 400 μέχρι 600 ευρώ;

Λύση

Α) Αν συμβολίσουμε με την τυχαία μεταβλητή Χ, τα μηνιαία έξοδα των φοιτητών, προφανώς

- **αναζητούμε την πιθανότητα** $P(X > 1000)$**.**
- έχουμε ότι **μ=800 και σ=80**.
- Έτσι, έχουμε:

$$P(X > 1000) = P(\frac{X-800}{80} > \frac{1000-800}{80})$$
$$= P(Z > \frac{200}{80}) = P(Z > 2,5)$$
$$= 1 - P(Z < 2,5)$$
$$= 1 - F(2,5)$$
$$= 1 - 0,9798 = 0,0202$$

Επομένως, η πιθανότητα να ξοδεύει ο φοιτητής αυτός, κατά μήνα, περισσότερο από 1000 ευρώ είναι 2,02 %.

Β) Εδώ, προφανώς αναζητούμε την πιθανότητα $P(X < 600)$ για την οποία έχουμε:

$$P(X < 600) = P(\frac{X-800}{80} < \frac{600-800}{80})$$
$$= P(Z < -2,5) = F(-2,5) = 1 - F(2,5)$$
$$= 1 - 0,9938 = 0,0062$$

Επομένως, η πιθανότητα να ξοδεύει ο φοιτητής αυτός, κατά μήνα, λιγότερο από 600 ευρώ είναι 0,62%.

Γ) Εδώ, προφανώς αναζητούμε την πιθανότητα $P(400 < X < 600)$ για την οποία έχουμε:

$$
\begin{aligned}
& P(400 < X < 600) \\
& = P(\frac{400-800}{80} < \frac{X-800}{80} < \frac{600-800}{80}) \\
& = P(-5 < Z < -2,5) \\
& = F(-2,5) - F(-5) = 1 - F(2,5) - [1 - F(5)] \\
& = 1 - F(2,5) - 1 + F(5) = F(5) - F(2,5) \\
& = 0,9999997 - 0,9938 = 0,0061997
\end{aligned}
$$

Επομένως, η πιθανότητα να ξοδεύει ο φοιτητής αυτός, κατά μήνα, περισσότερα από 400 ευρώ, αλλά όχι περισσότερο από 600 ευρώ, είναι 0,61997 %.

=============== ================= =================== ===

Τέλος Α΄ εισήγησης ….

Τέλος Ενότητας

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης