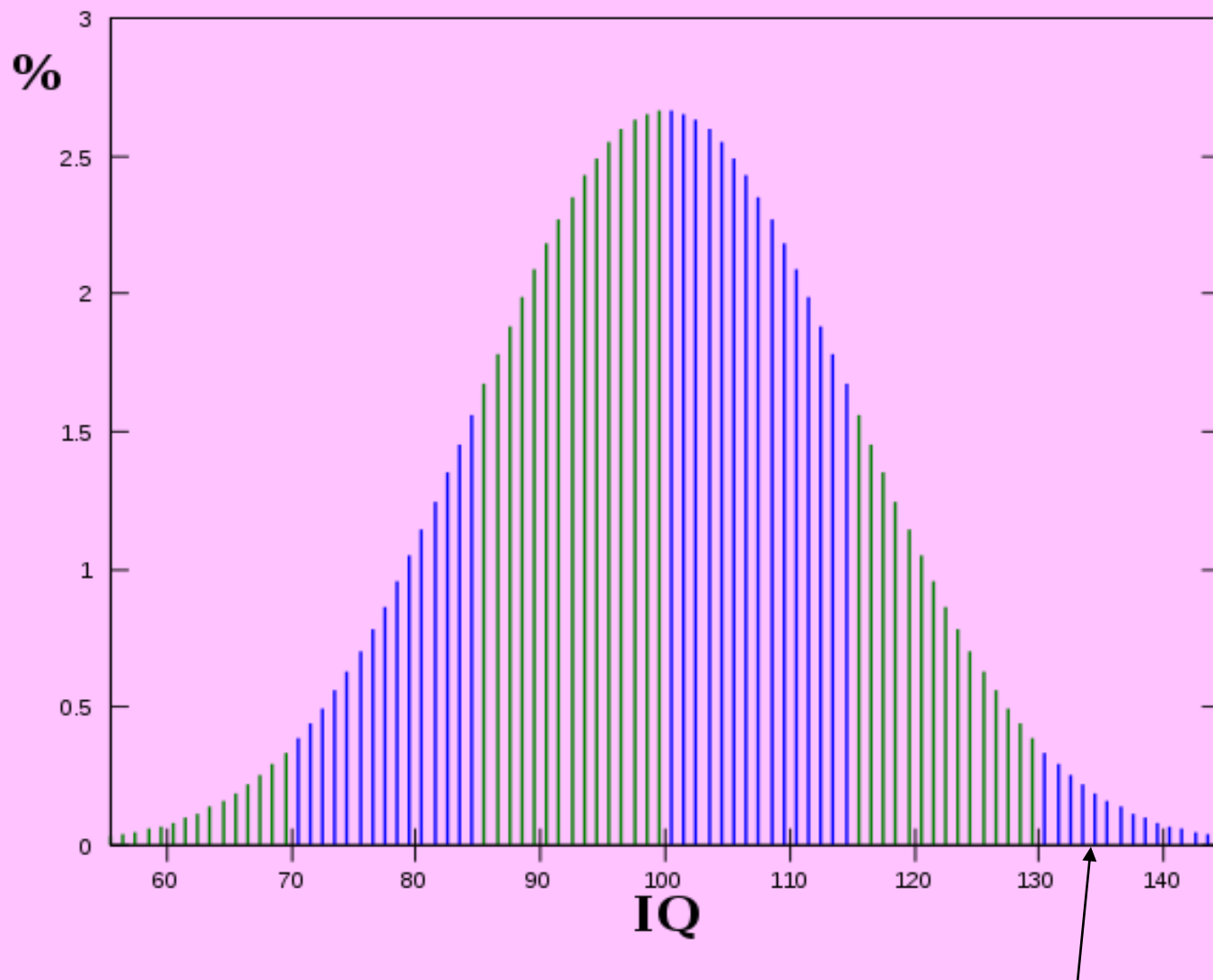


Κανονική Κατανομή



Αυτή η περιοχή αποτελεί το 5% των ατόμων με το υψηλότερο IQ

Γιατί χρειαζόμαστε την μορφή μιας κατανομής;

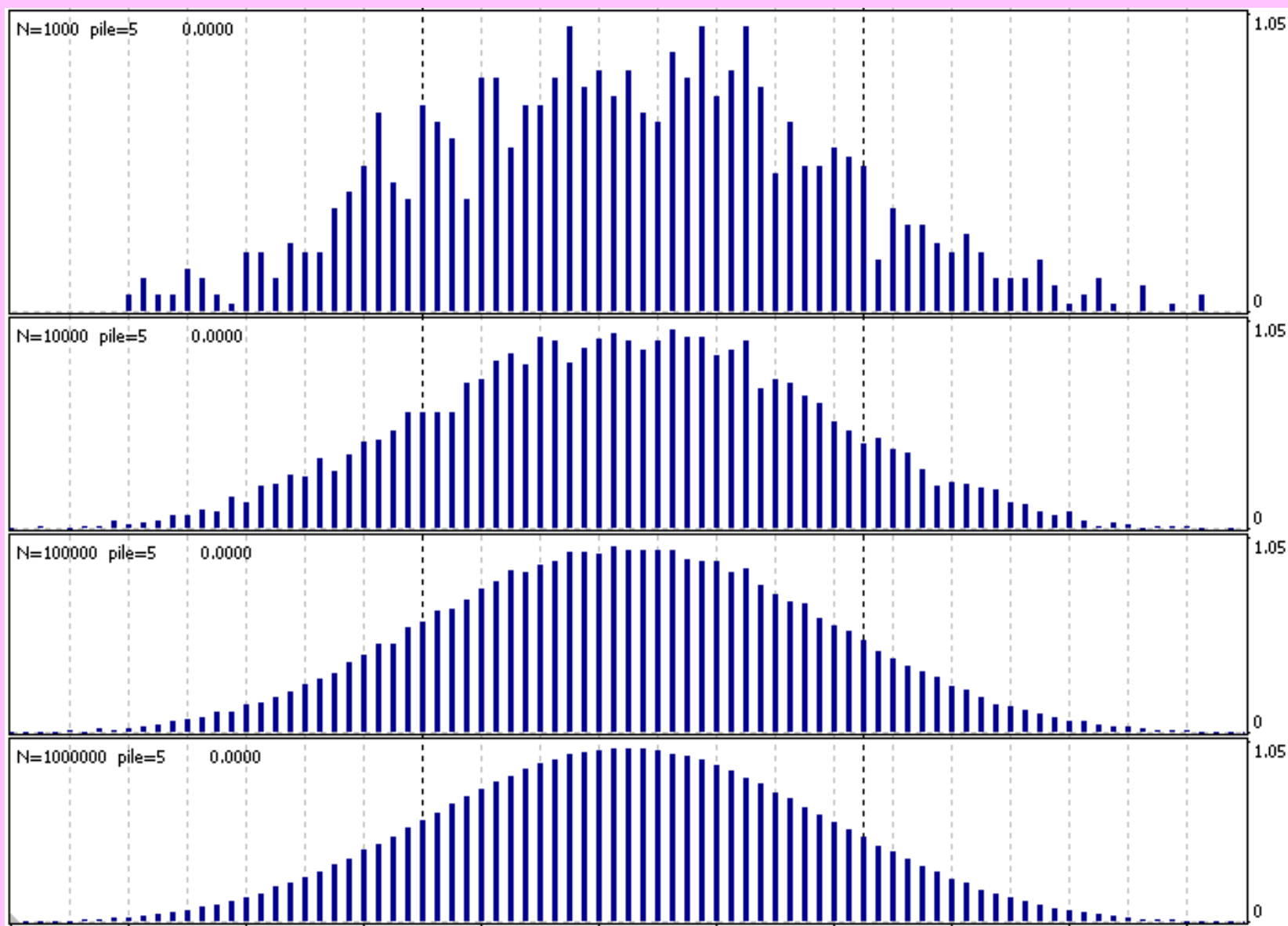
Για να βρούμε ΠΟΣΟΣΤΑ (%).

Π.χ.

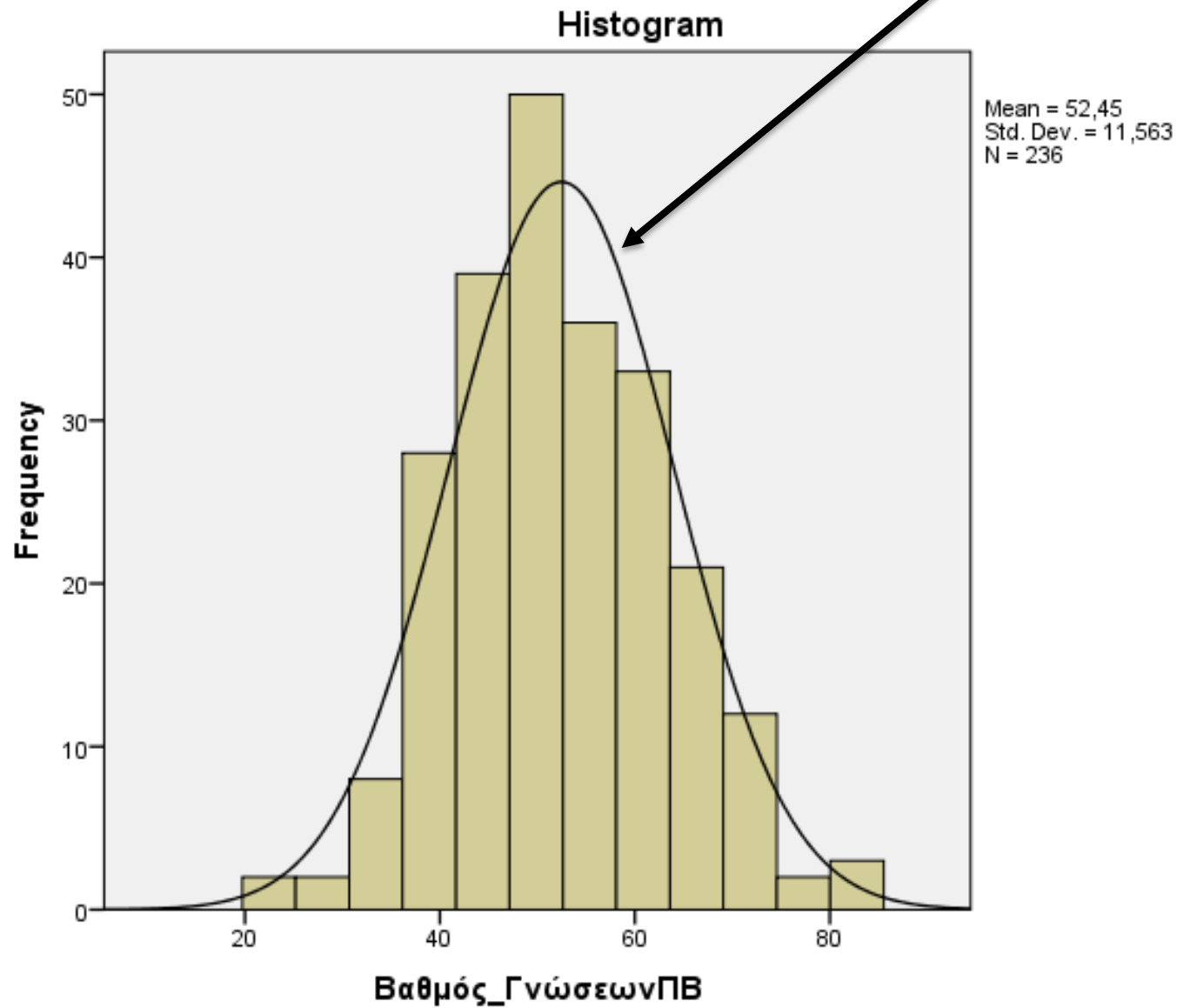
Πόσα άτομα (%) έχουν IQ μεγαλύτερο από 120;

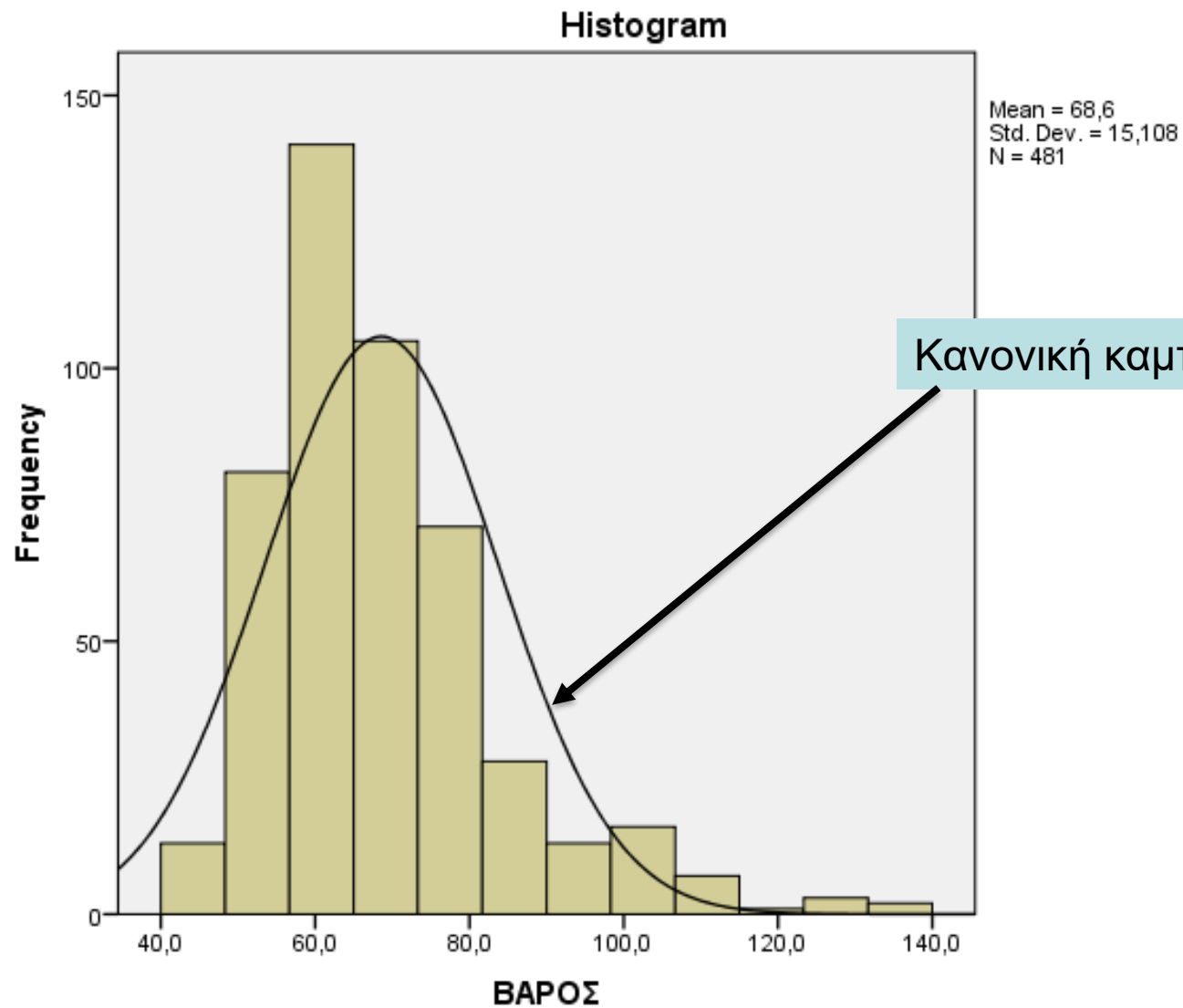
Πόσα άτομα (%) έχουν IQ μεταξύ 90 και 100;

Πόσα άτομα (%) έχουν IQ μικρότερο από 80;



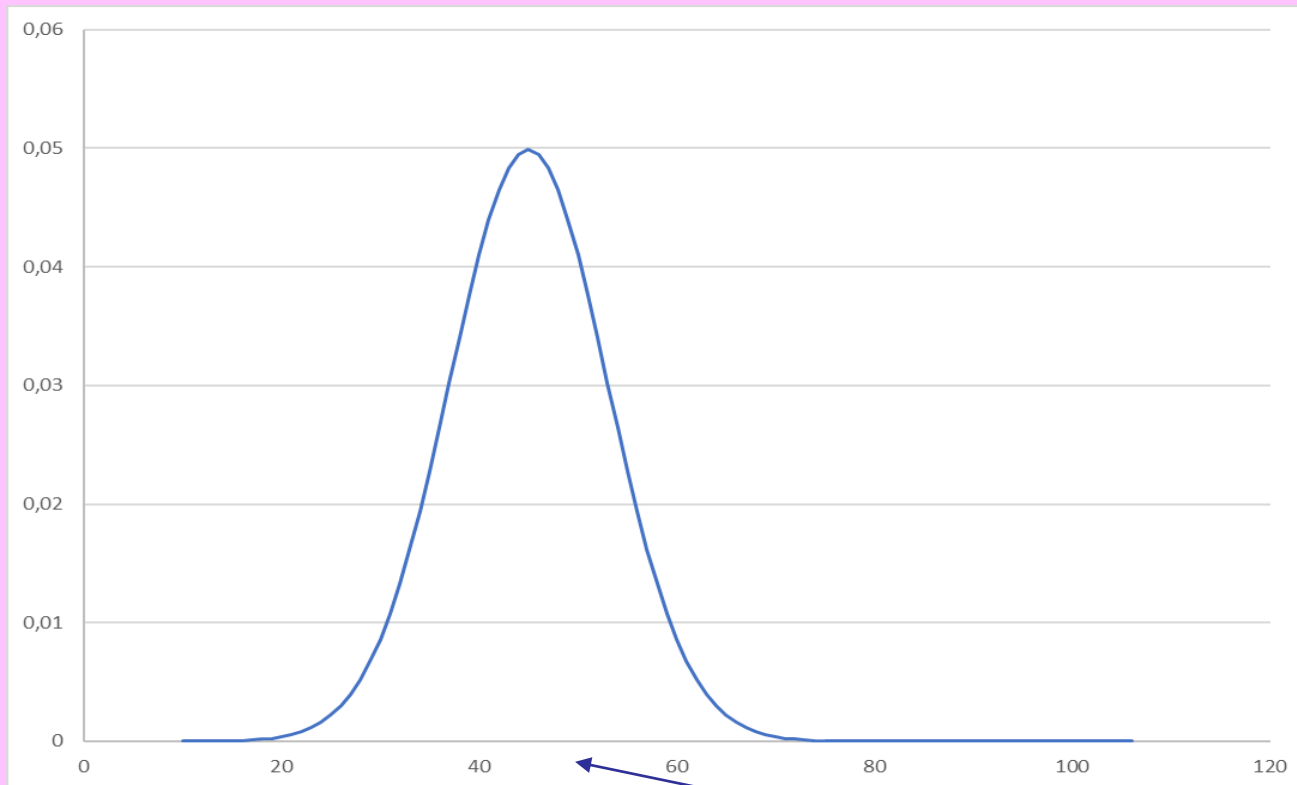
Κανονική καμπύλη





Η κανονική κατανομή είναι

- Συμμετρική
 - Μοιάζει με καμπάνα
- (μέση τιμή = διάμεσος = επικρατούσα τιμή)



Χαρακτηριστικό που μετρήσαμε

Χρειαζόμαστε ΔΥΟ στοιχεία για να περιγράψουμε ένα χαρακτηριστικό με την κανονική κατανομή:

Α) τη μέση τιμή (μ)

β) την τυπική απόκλιση (σ)

Όταν δεν τα γνωρίζουμε, τα εκτιμούμε από το δείγμα μας

Δηλαδή μέση τιμή ($\bar{X}$) και τυπική απόκλιση (S)

Για να περιγράψουμε κάποιο χαρακτηριστικό με την κανονική κατανομή, και να βρούμε κάποια ποσοστά, χρειαζόμαστε να γνωρίζουμε

1 – την μέση τιμή (μ ή $\bar{X}$)

2 – την τυπική απόκλιση (σ ή S)

Συνήθως χρησιμοποιούμε τις εκτιμήσεις που παίρνουμε από το δείγμα μας για τη μέση τιμή και την τυπική απόκλιση.

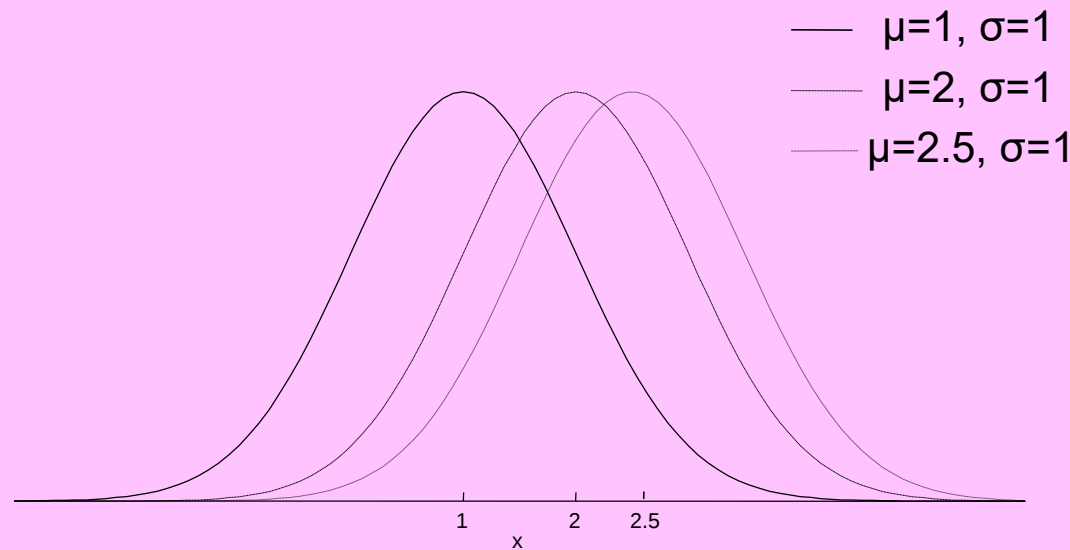
Η μορφή της συνάρτησης που εκφράζει την καμπύλη της
Κανονικής Κατανομής
(ονομάζεται Πυκνότητα Πιθανότητας)

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

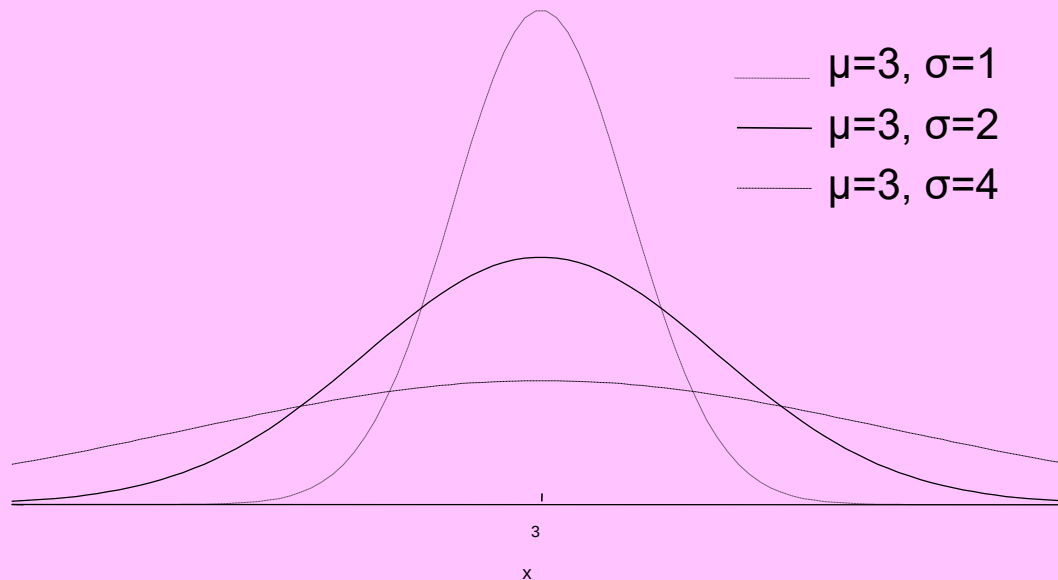
Ειδική περίπτωση ($\mu=0, \sigma=1$)

Την ονομάζουμε
**τυπική κανονική
κατανομή**
(Z – κατανομή)

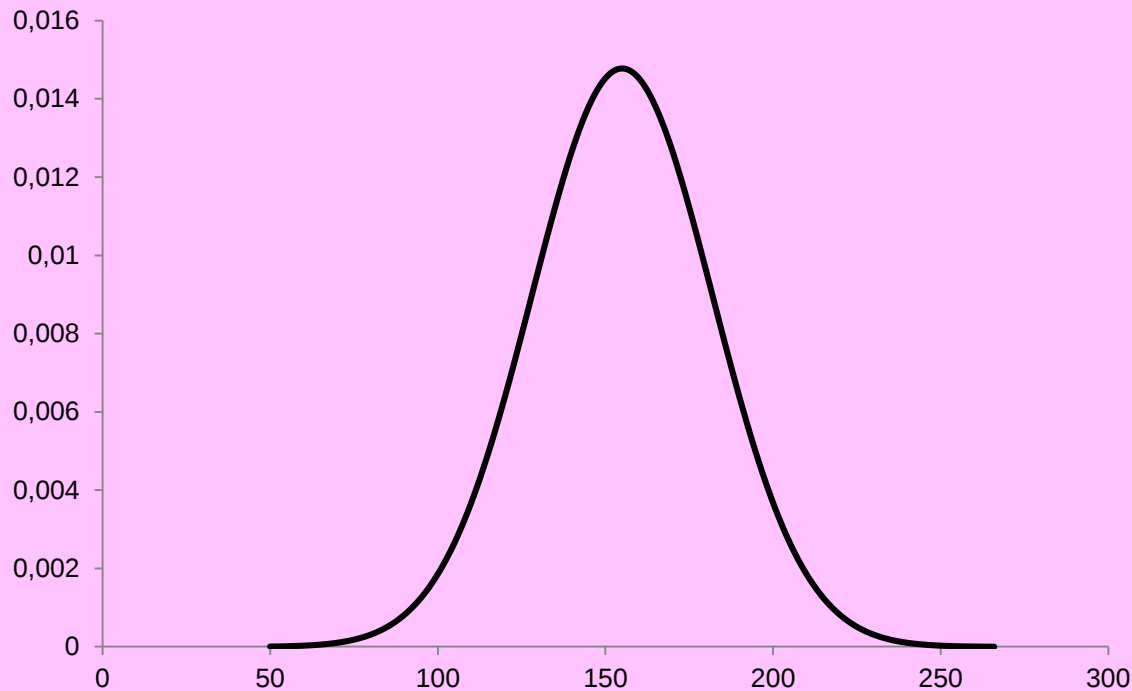
Κανονικές κατανομές με την ίδια τυπική απόκλιση (αλλά διαφορετικές μέσες τιμές)



Κανονικές κατανομές με την ίδια μέση τιμή (αλλά διαφορετικές τυπικές αποκλίσεις)

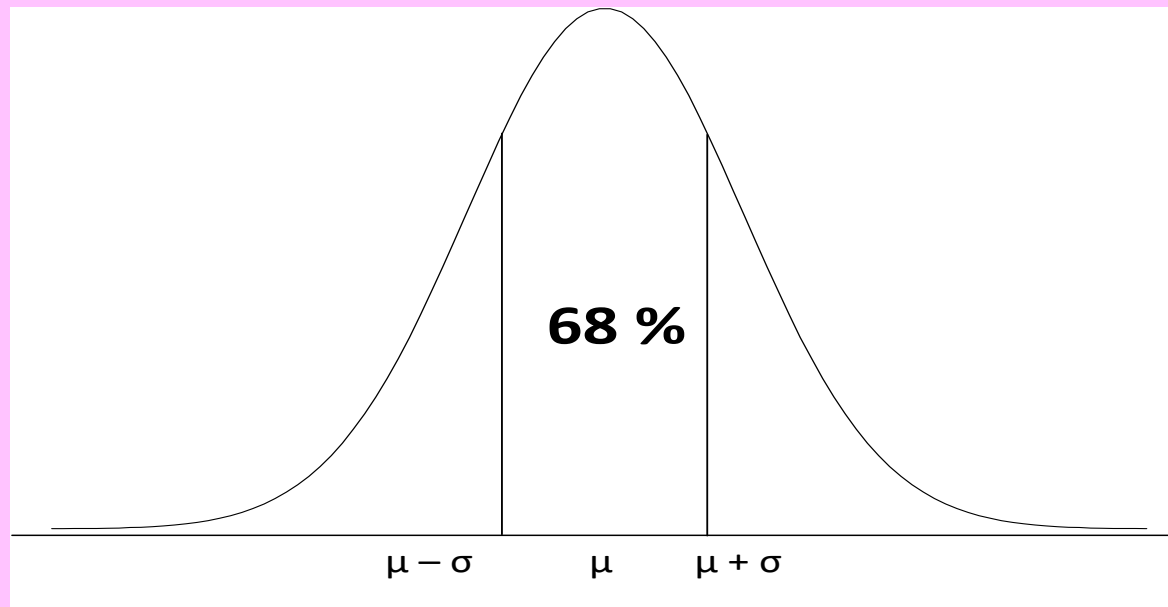


Κατανομή σκορ σε κάποιο τεστ σε κάποιο πληθυσμό



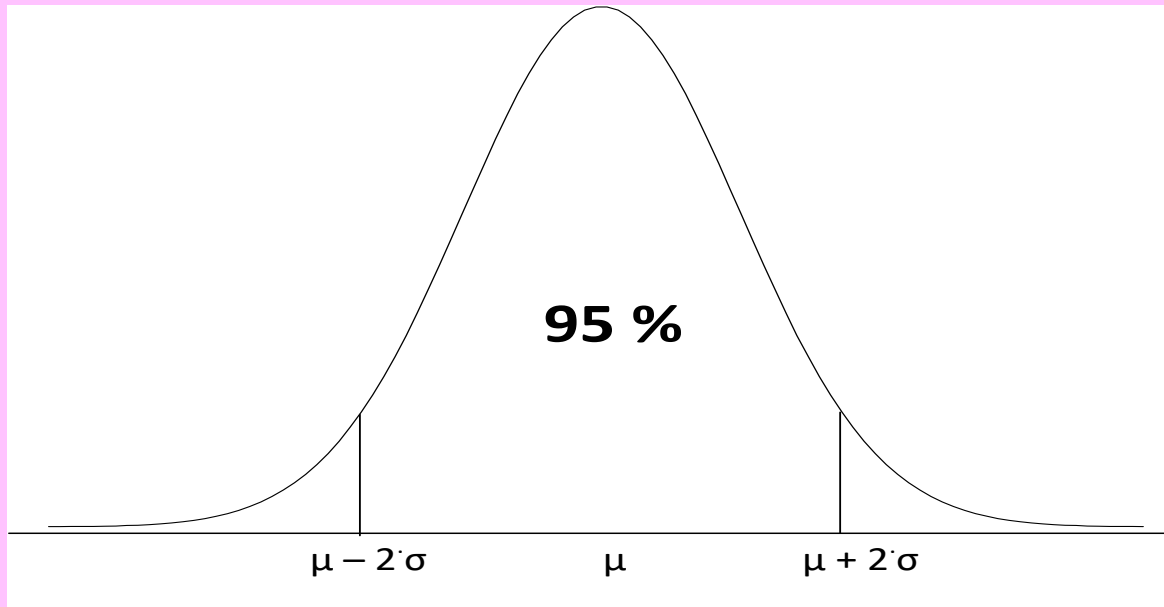
Μέση τιμή (μ) = 155
Τυπική απόκλιση (σ) = 27

Η σημασία της Τυπικής Απόκλισης (σ) – περίπτωση 1



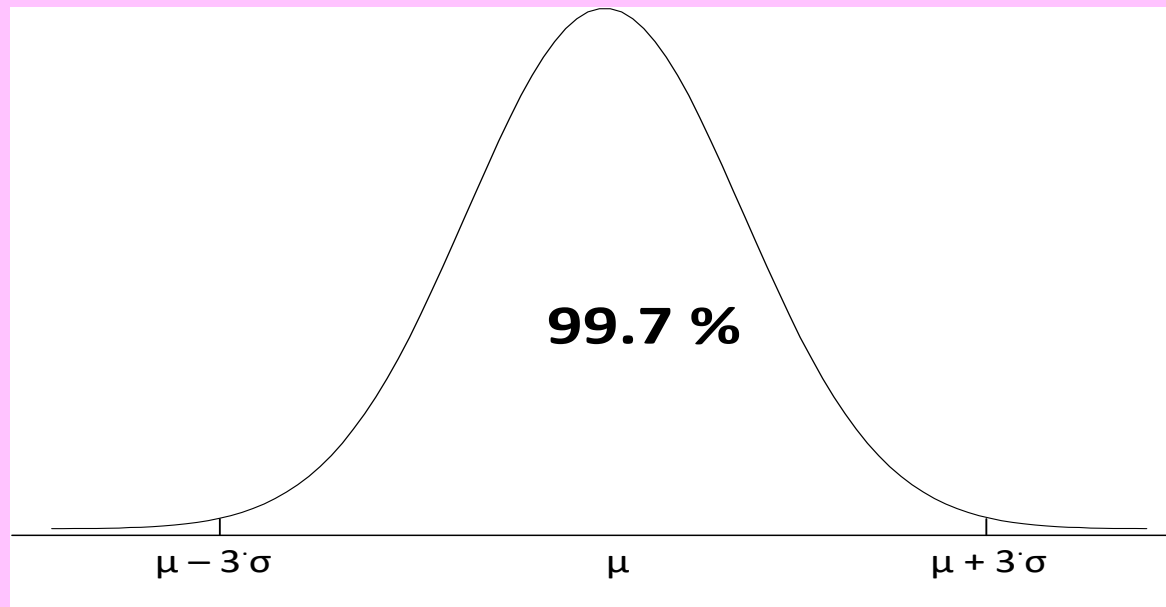
Σε απόσταση **ΜΙΑ** τυπική απόκλιση από τη μέση τιμή, υπάρχει περίπου το 68% των περιπτώσεων

Η σημασία της Τυπικής Απόκλισης (σ) – περίπτωση 2



Σε απόσταση **ΔΥΟ** τυπικών αποκλίσεων από τη μέση τιμή, υπάρχει περίπου το 95% των περιπτώσεων

Η σημασία της Τυπικής Απόκλισης (σ) – περίπτωση 3



Σε απόσταση **ΤΡΙΩΝ** τυπικών αποκλίσεων από τη μέση τιμή, υπάρχει το 99,7% των περιπτώσεων (δηλαδή σχεδόν το σύνολο)

Εφαρμογή: στο τεστ που μετρήσαμε:

Μέση τιμή = **155** και Τυπική απόκλιση = **27**

ΟΠΟΤΕ

$$155 - 27 = 128$$

$$155 + 27 = 182$$

Άρα μεταξύ 128 και 182 $\rightarrow$ 68% των ατόμων

ΟΠΟΤΕ

$$155 - 2 * 27 = 101$$

$$155 + 2 * 27 = 209$$

Άρα μεταξύ 101 και 209 $\rightarrow$ 95% των ατόμων

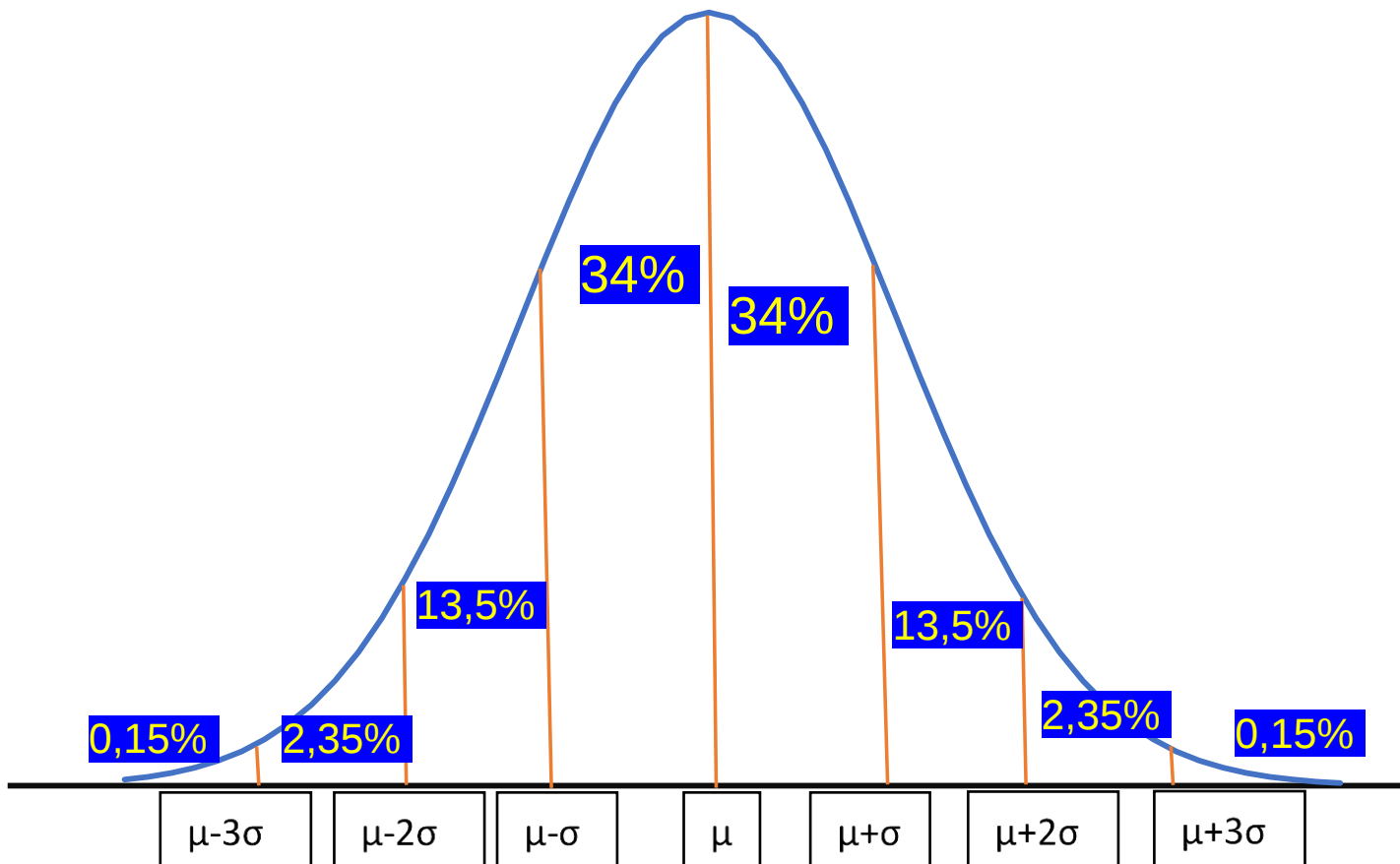
ΟΠΟΤΕ

$$155 - 3 \cdot 27 = 74$$

$$155 + 3 \cdot 27 = 236$$

Άρα μεταξύ 74 και 236 $\rightarrow$ 99,7% των
ατόμων

Ποσοστά στην Κανονική Κατανομή



Άσκηση – παράδειγμα

- Σε κάποιο τεστ για τον αλκοολισμό, δόθηκαν 15 ερωτήσεις στην κλίμακα 1-2-3-4-5 (καθόλου --- πάρα πολύ).
- Υπολογίστηκε το άθροισμα των απαντήσεων και υπολογίστηκε το συνολικό σκορ κάθε ατόμου
- Το σκόρ μπορεί να πάρει τιμές από 15 (αν σε όλες τις ερωτήσεις απαντήσουν «καθόλου») μέχρι 75 (αν σε όλες τις ερωτήσεις απαντήσουν «πάρα πολύ»)

Ερωτήσεις

Στο δείγμα μας, βρήκαμε μέση τιμή = 35 και τυπική απόκλιση = 6.

- 1) Ανάμεσα σε ποιες τιμές αλκοολισμού, βρίσκουμε το 95% των ατόμων μας;

$$35 - 2 \cdot 6 = 23 \text{ και } 35 + 2 \cdot 6 = 47$$

(μεταξύ 23 και 47 υπάρχει το 95%)

2) Πόσο ποσοστό των ατόμων θα βρούμε ανάμεσα σε σκορ 29 και 41;

Παρατηρούμε ότι

$$29 = \mu - \sigma \quad (35 - 6) \quad \text{και} \quad 41 = \mu + \sigma \quad (35 + 6),$$

οπότε το ποσοστό ανάμεσα στο 29 και το 41 είναι 68%

3) Πόσο ποσοστό των ατόμων θα βρούμε με σκορ πάνω από 47;

Το 47 είναι το $\mu+2\sigma$, οπότε το ποσοστό είναι 2,5%

4) Πόσο ποσοστό των ατόμων θα βρούμε με σκορ ανάμεσα στο 41 και στο 47;

Το 41 είναι το $\mu+\sigma$ και το 47 είναι το $\mu+2\sigma$, οπότε το ποσοστό είναι 13,5%

Τυπική Κανονική Κατανομή (Z)

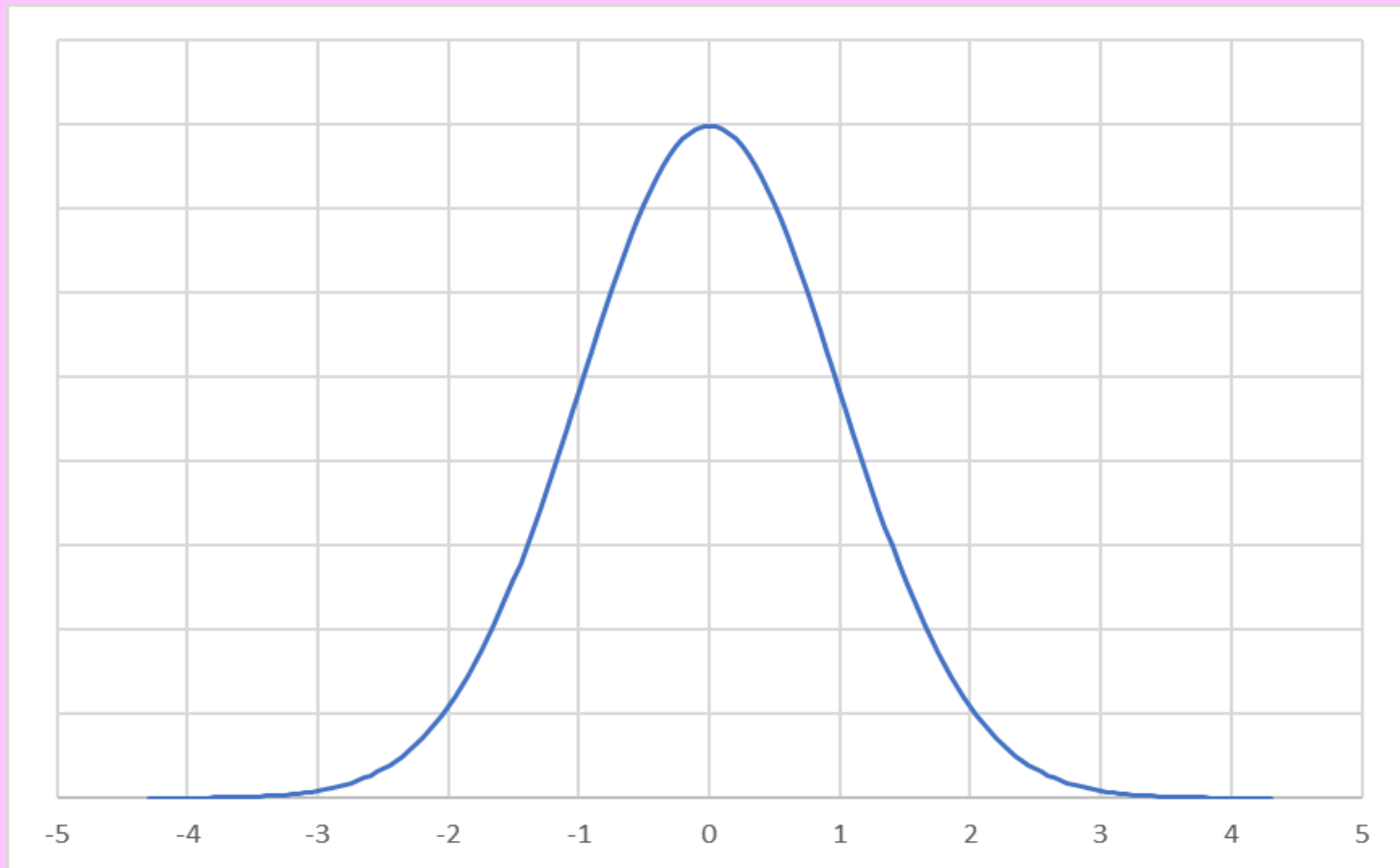
- Η **Z κατανομή** ή **τυπική κανονική κατανομή** είναι η κανονική κατανομή που έχει μέση τιμή **0** και τυπική απόκλιση **1**.
- Γιατί μας χρειάζεται;;
- Οποιοδήποτε αριθμητικό χαρακτηριστικό έχουμε, σε οποιαδήποτε κλίμακα, μπορεί να μετασχηματιστεί σε Z-σκορ. Όλες οι κλίμακες μπορούν να «τυποποιηθούν» με τον εξής απλό τρόπο:

Τυποποίηση μιας κλίμακας:

- Από την τιμή που έχουμε πάρει στην κλίμακα (π.χ. την τιμή X που μετρήσαμε) αφαιρούμε την μέση τιμή της κλίμακας (μ ή $\bar{X}$). Το αποτέλεσμα το διαιρούμε με την τυπική απόκλιση (σ ή S), και το ονομάζουμε Z-σκορ
- Δηλαδή $Z\text{-σκορ} = (X - \mu) / \sigma$

Z-κατανομή

(μέση τιμή=0 και τυπ.απόκλιση =1)



Μέση τιμή

Παράδειγμα:

Κάποιος φοιτητής εξετάστηκε σε δύο τεστ, στο Α και στο Β με διαφορετικές κλίμακες και πήρε τις εξής βαθμολογίες

A-τεστ -----βαθμολογία = 25

B-τεστ -----βαθμολογία = 84

Στο Α τεστ η μέση τιμή των συμμετεχόντων ήταν 22 και η τυπική απόκλιση 5

Το Β τεστ η μέση τιμή των συμμετεχόντων ήταν 60 και τυπική απόκλιση 12

Σε ποιο τεστ πήγε καλύτερα ο φοιτητής αυτός;

- Για να κάνουμε σύγκριση ανάμεσα σε διαφορετικές βαθμολογίες ή διαφορετικές κλίμακες, χρησιμοποιούμε το Z-μετασχηματισμό.

Για την πρώτη βαθμολογία

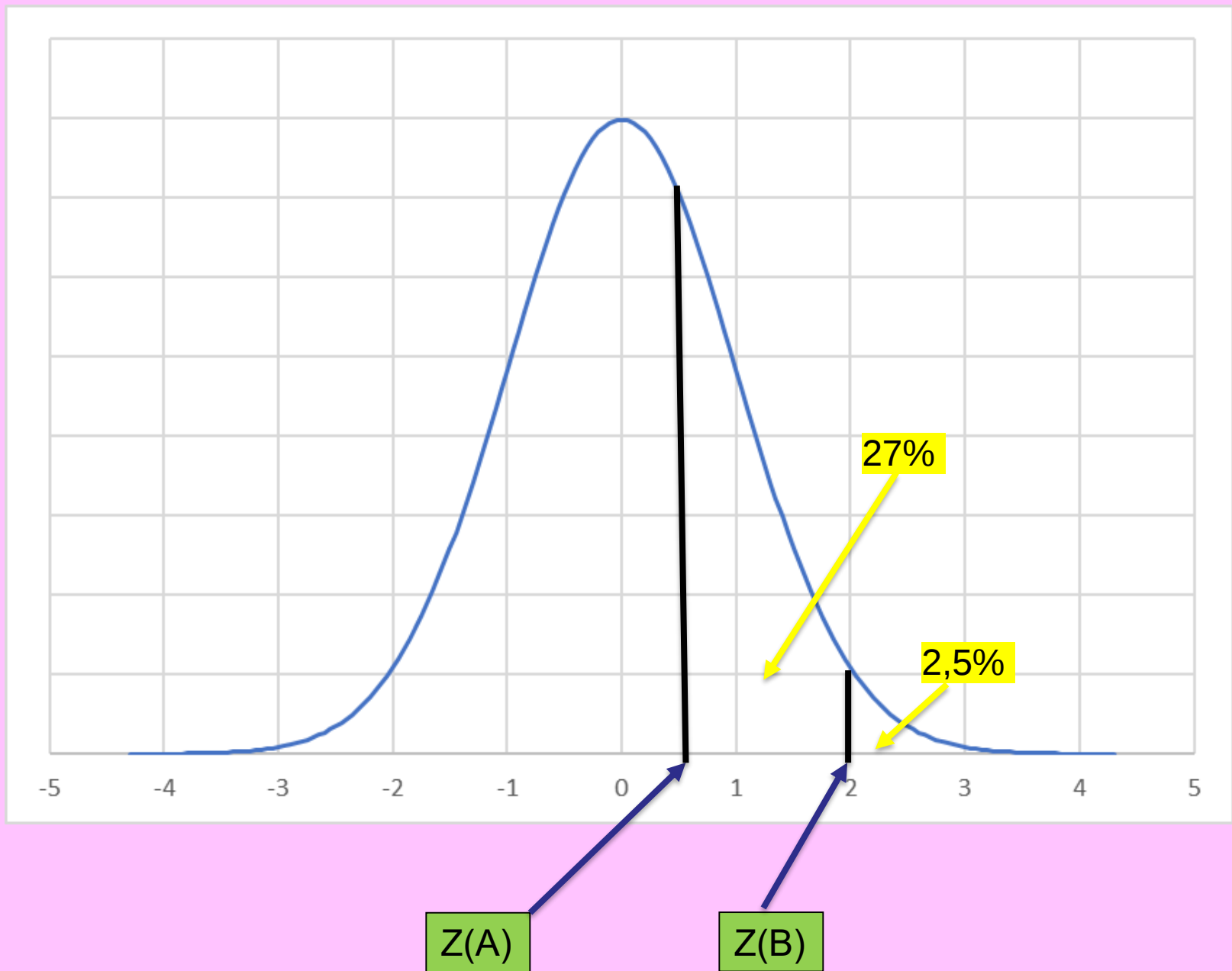
$$Z(A) = (25-22)/5 = 3/5 = 0,6$$

Για τη δεύτερη βαθμολογία

$$Z(B) = (84-60)/12 = 24/12 = 2$$

Συμπέρασμα;

- Στο B τεστ ο φοιτητής έβγαλε πιο μεγάλο Z σκορ, άρα είναι πιο κοντά στην πάνω ουρά της κανονικής κατανομής, άρα πήγε πολύ καλύτερα σε σχέση με το A τεστ.
- Μάλιστα στο B τεστ, είναι φανερό ότι μόνο 2,5% των φοιτητών είχαν καλύτερη επίδοση από τον φοιτητή αυτό.
- Στο A τέστ το ποσοστό αυτό είναι 27% (μπορεί να βρεθεί από εντολή στο Jamonι ή στο Excel)



Ποσοστά στην Z είναι αυτά που ήδη έχουμε δει

