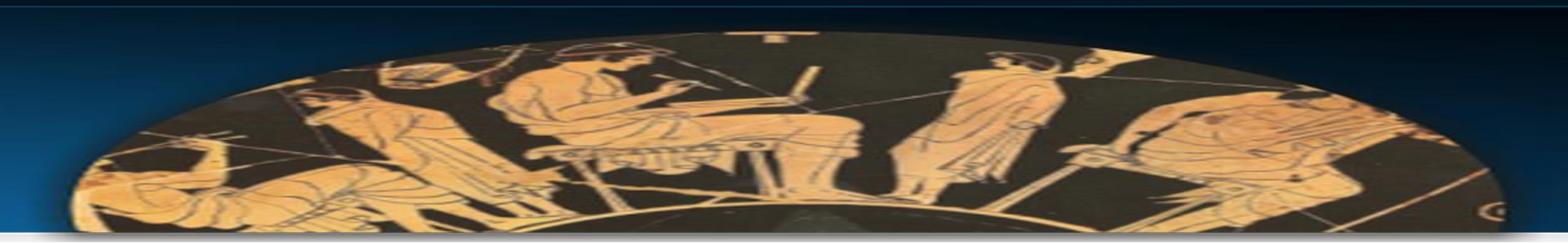
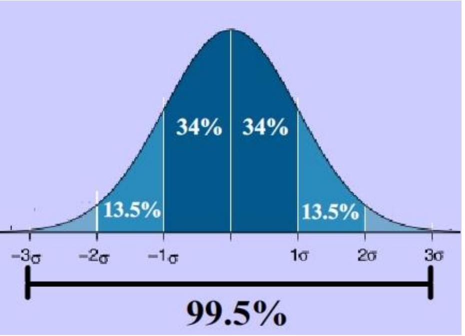


# Κανονική κατανομή


Δημήτρης Μαυρίδης

Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης

Πανεπιστήμιο Ιωαννίνων


# Διακριτές κατανομές

- Κατανομή Bernoulli: ένα πείραμα τύχης με 2 πιθανά αποτελέσματα (πείραμα Bernoulli)
- Γεωμετρική κατανομή: ο αριθμός των πειραμάτων Bernoulli μέχρι την πρώτη επιτυχία
- Διωνυμική κατανομή: ο αριθμός επιτυχιών σε n πειράματα Bernoulli

# Κανονική κατανομή (Normal distribution)

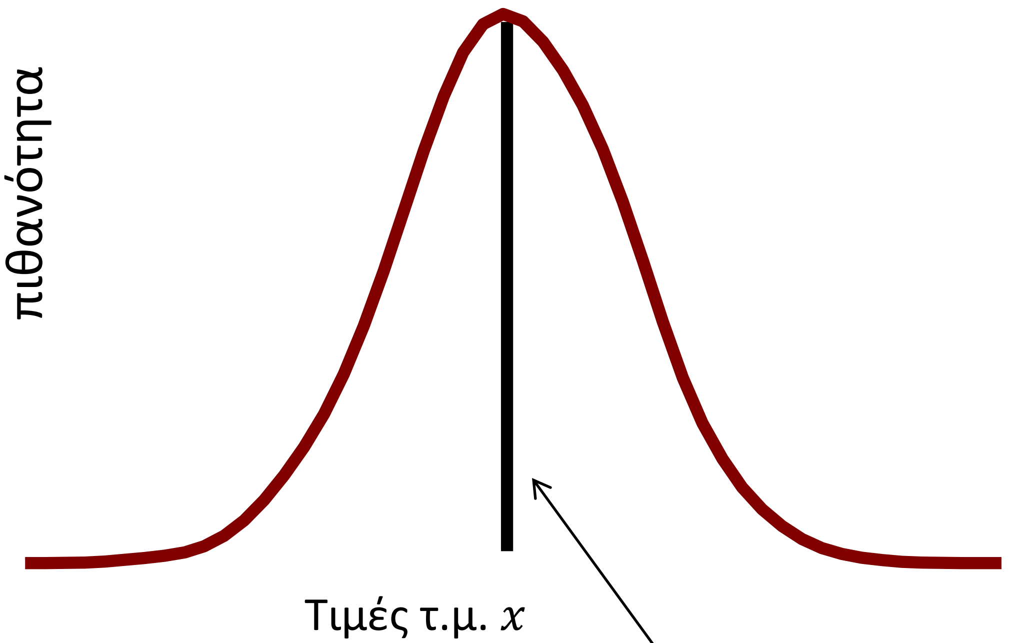
- Πολλές τυχαίες μεταβλητές χαρακτηρίζονται ή προσεγγίζονται ικανοποιητικά από την κανονική κατανομή
- Χαρακτηρίζεται από 2 παραμέτρους, το μέσο μ και την τυπική απόκλιση σ
- Συμβολίζεται ως $x \sim N(\mu, \sigma^2)$
- Είναι συμμετρική ως προς τον μέσο όρο

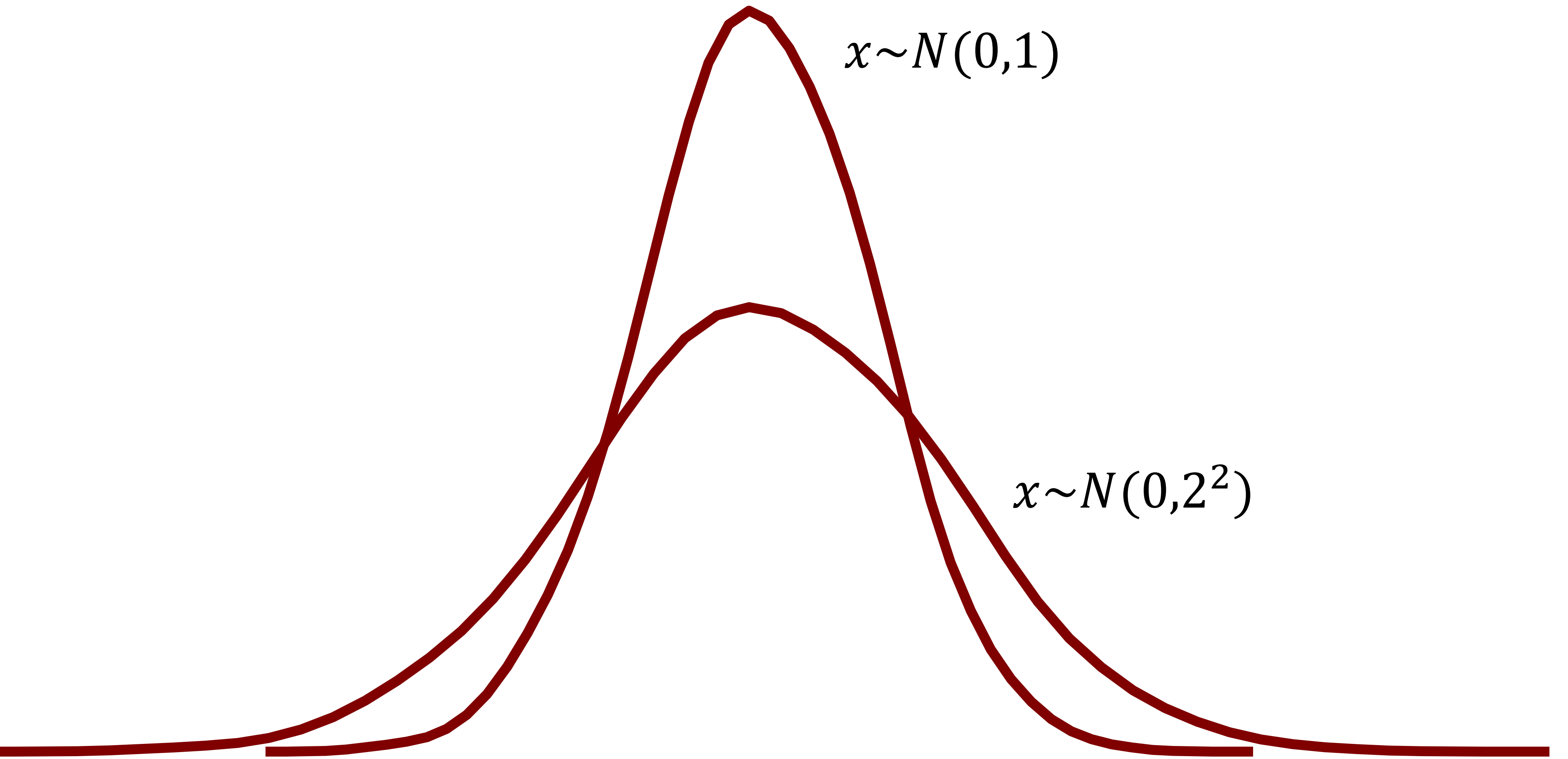
# Κανονική κατανομή

$$x \sim N(\mu, \sigma^2)$$

μέσος

Τυπική απόκλιση



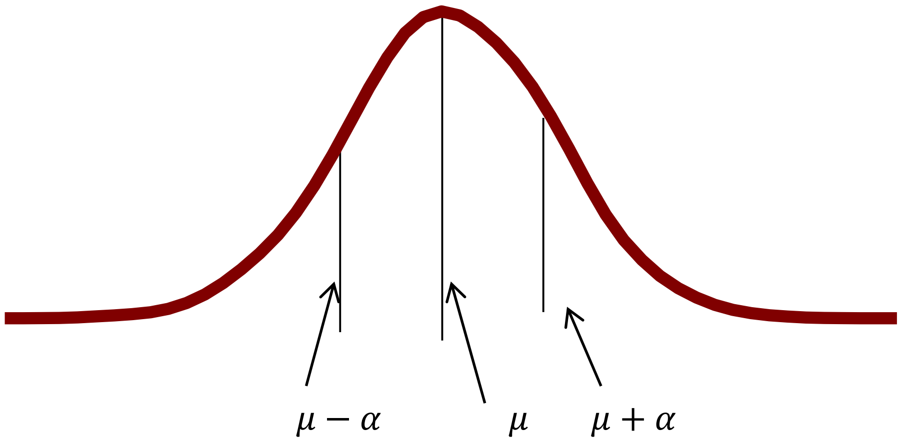

$x \sim N(0,1)$
$x \sim N(0,2^2)$

# Συμμετρία

Η κανονική κατανομή είναι συμμετρική προς το μέσο όρο. Ο μέσος όρος και η διάμεσος ταυτίζονται

$$P(X < \mu) = P(x > \mu) = 0.5$$
$$P(X < \mu - \alpha) = P(X > \mu + \alpha)$$

Η περιοχή κάτω από την καμπύλη εκφράζει πιθανότητα. Επομένως, το συνολικό εμβαδό της περιοχής κάτω από την καμπύλη ισούται με τη μονάδα. Στις συνεχείς κατανομές έχει νόημα μόνο η πιθανότητα ενός διαστήματος τιμών και όχι η πιθανότητα συγκεκριμένων τιμών



# Εμπειρικός κανόνας

- το 68% των παρατηρήσεων βρίσκονται στο διάστημα που τα άκρα του απέχουν μια τυπική απόκλιση από το μέσο των παρατηρήσεων (μ-σ, μ+σ)

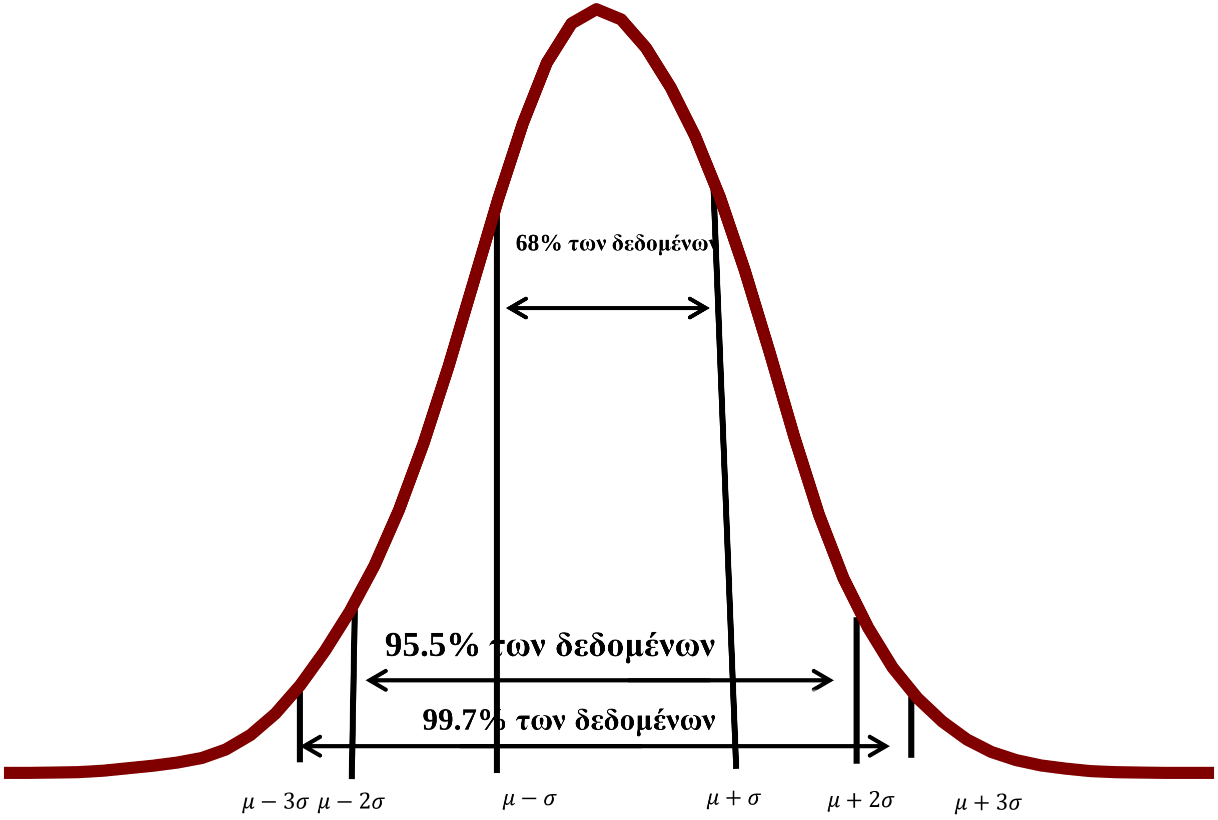
$$P(\mu - \sigma < x < \mu + \sigma) = 0.68$$

- το 95.5% των παρατηρήσεων βρίσκονται στο διάστημα που τα άκρα του απέχουν δυο τυπικές αποκλίσεις από το μέσο των παρατηρήσεων (μ-2σ,μ+2σ)

$$P(\mu - 2\sigma < x < \mu + 2\sigma) = 0.955$$

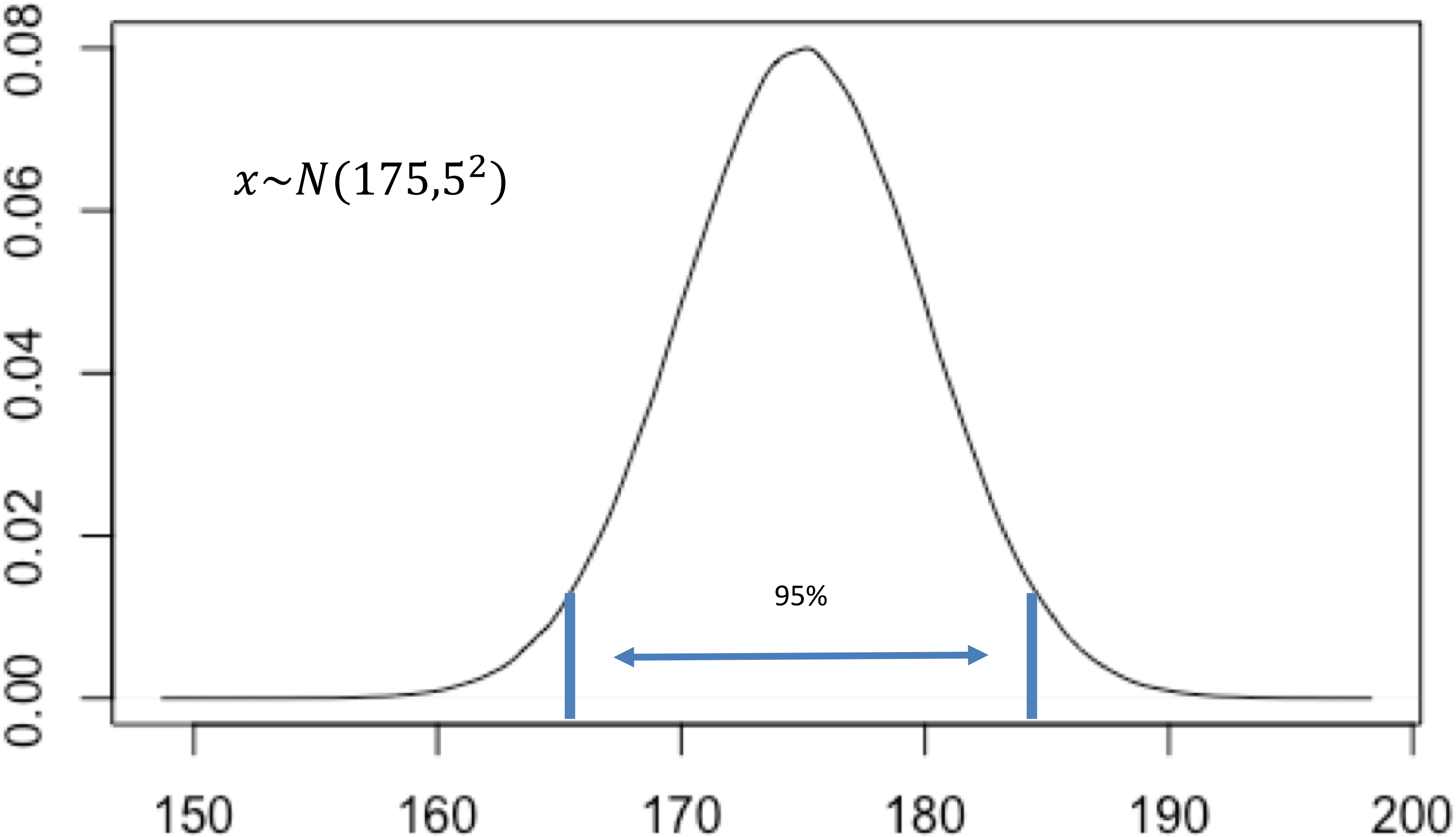
- το 99.7% των παρατηρήσεων βρίσκονται στο διάστημα που τα άκρα του απέχουν τρεις τυπικές αποκλίσεις από το μέσο των παρατηρήσεων (μ-3σ,μ+3σ)

$$P(\mu - 3\sigma < x < \mu + 3\sigma) = 0.997$$


68% των δεδομένων
95.5% των δεδομένων
99.7% των δεδομένων
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$

# Δώστε ένα 95% διάστημα εμπιστοσύνης για το ύψος του αντρικού πληθυσμού

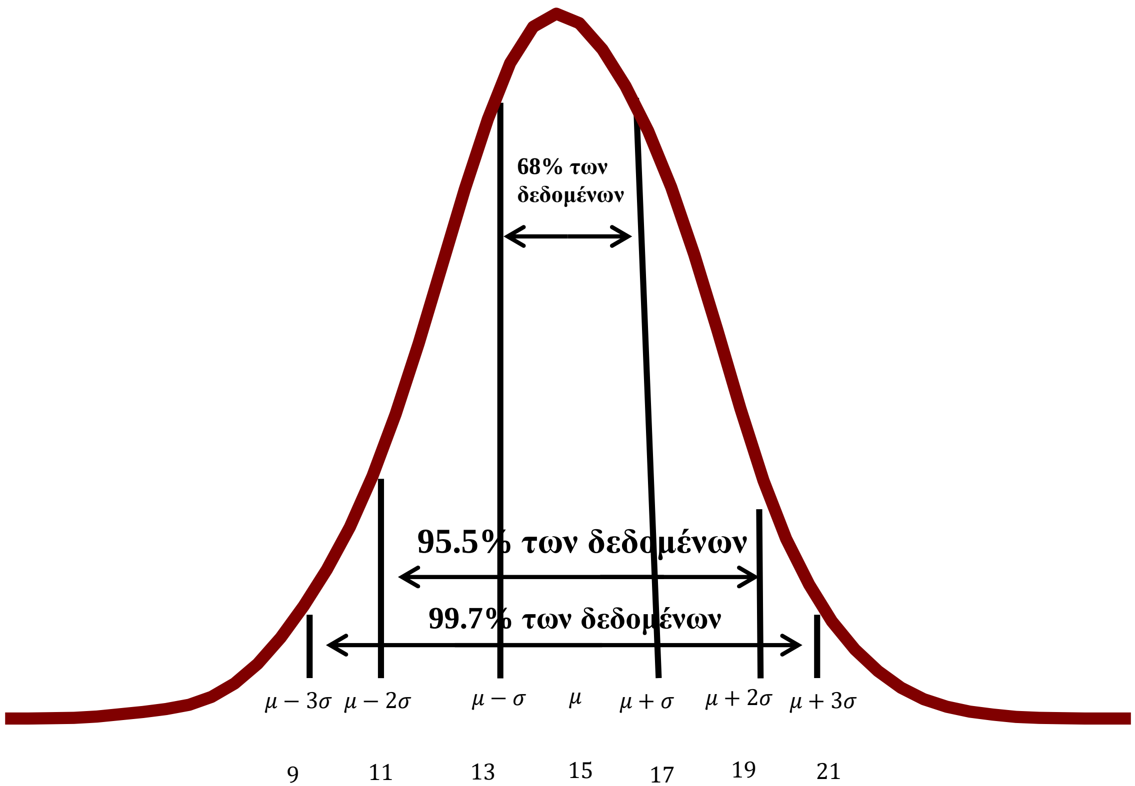
- Θεωρούμε ότι το ύψος του αντρικού πληθυσμού κατανέμεται κανονικά. Είναι συμμετρικό ως προς το μέσο ύψος και η κατανομή φθίνει με τον ίδιο τρόπο όσο απομακρυνόμαστε εκατέρωθεν του μέσου ύψους
- Πρέπει να υπολογίσουμε το μέσο ύψος και την τυπική απόκλιση
- Έστω ότι θεωρούμε ότι η πλειοψηφία των αντρών είναι από 165 έως 185 εκατοστά
- Επομένως ο μέσος όρος είναι μ=175εκ
- Επίσης $\mu + 2 \times \sigma = 185 \leftrightarrow \sigma = 5$


$x \sim N(175, 5^2)$
95%
0.00
0.02
0.04
0.06
0.08
150
160
170
180
190
200

Έστω ότι οι βαθμοί σε ένα διαγώνισμα ακολουθούν την κανονική κατανομή με μέσο όρο 15 και τυπική απόκλιση 2.

- Έστω x η τυχαία μεταβλητή με τους βαθμούς των μαθητών $$x \sim N(15,2^2)$$
- Ποιά είναι η πιθανότητα ένας μαθητής να γράψει λιγότερο από 15
- Ποιά είναι η πιθανότητα ένας μαθητής να πάρει βαθμό στο διάστημα [13,17]
- Ποιά είναι η πιθανότητα να γράψει πάνω από 17
- Ποιά είναι η πιθανότητα να γράψει από 13 έως 15

$$x \sim N(\mu = 15, \sigma^2 = 2^2)$$



Έστω ότι οι βαθμοί σε ένα διαγώνισμα ακολουθούν την κανονική κατανομή με μέσο όρο 15 και τυπική απόκλιση 2.

- Έστω x η τυχαία μεταβλητή με τους βαθμούς των μαθητών
$$x \sim N(15, 2^2)$$
- Ποιά είναι η πιθανότητα ένας μαθητής να γράψει λιγότερο από 15
$$P(x < 15) = P(x < \mu) = 0.5$$
- Ποιά είναι η πιθανότητα ένας μαθητής να πάρει βαθμό στο διάστημα [13,17]
$$P(\mu - \sigma < x < \mu + \sigma) = 0.68$$
$$P(13 < x < 17) = 0.68$$
- Ποιά είναι η πιθανότητα να γράψει πάνω από 17
$$P(x > 17) = P(x < 13) = 0.16$$
- Ποιά είναι η πιθανότητα να γράψει από 13 έως 15
$$P(13 < x < 15) = \frac{P(13 < x < 17)}{2} = \frac{0.68}{2} = 0.34$$

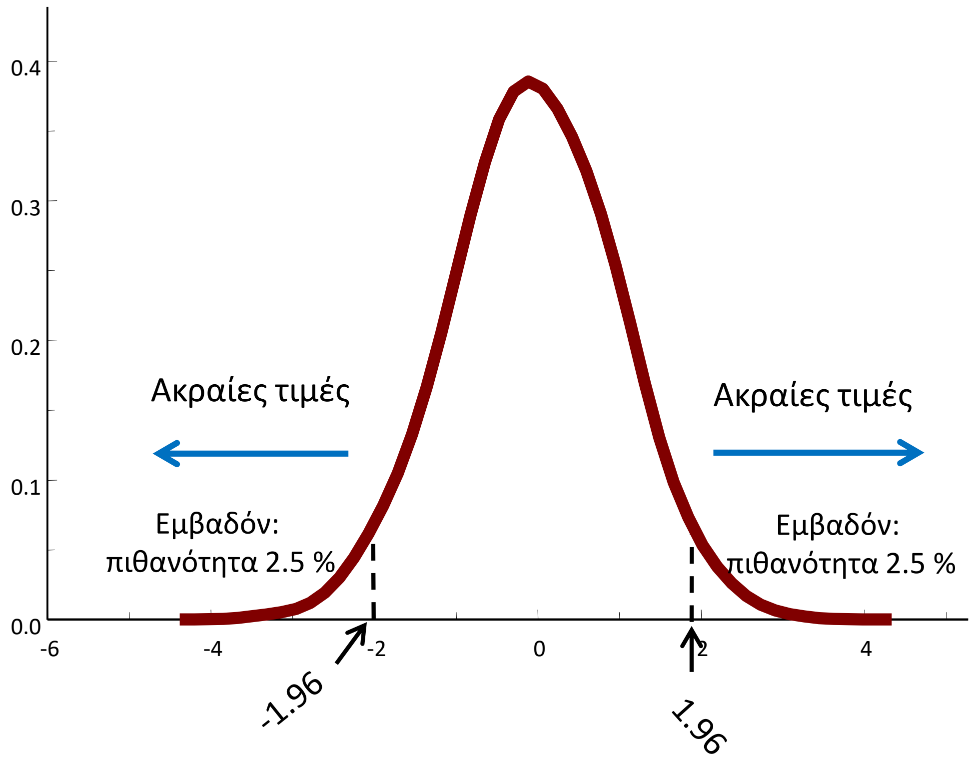
# Τυποποιημένες τιμές

$$Z_i = \frac{x_i - \mu}{\sigma} \sim \mathrm{N}(0,1)$$

- Δυο διαφορετικά διαγωνίσματα μαθηματικών δόθηκαν για τη χορήγηση υποτροφιών. Υποτροφία θα έπαιρναν όσοι φοιτητές πετύχαιναν βαθμολογία από 19 και πάνω. Στο πρώτη διαγώνισμα ο μέσος όρος των βαθμολογιών είναι 17 ενώ στο δεύτερο 13. Η τυπική απόκλιση και στις 2 τάξεις είναι 2. Ο υψηλότερος βαθμός της πρώτης τάξης πήρε 19 ενώ ο υψηλότερος βαθμός στη δεύτερη τάξη είναι 17.
- Για να δούμε τις τυποποιημένες βαθμολογίες των 2 μαθητών. Του πρώτου είναι (19-17)/2=1 ενώ του δεύτερου είναι (17-13)/2=2
- επομένως ο δεύτερος μαθητής αν και έχασε την υποτροφία στην πραγματικότητα ίσως να πήγε καλύτερα από τον πρώτο γιατί τα θέματα στα οποία διαγωνίστηκε ήταν δυσκολότερα από αυτά του πρώτου
- Η τυποποίηση εξασφαλίζει το ότι οι διάφορες τιμές έχουν την ίδια κατανομή και επομένως είναι συγκρίσιμες

# Τυπικη κανονική κατανομή

# Αν $x \sim N(\mu, \sigma^2)$ τότε $Z = \frac{x-\mu}{\sigma} \sim \mathrm{N}(0,1)$



# Τυπική κανονική κατανομή
# Εμπειρικός κανόνας

$$P(-1 < z < 1) = 0.68$$

$$P(-2 < z < 2) = 0.955$$

Βασικά είναι

$$P(-1.96 < z < 1.96) = 0.95$$

$$P(-3 < z < 3) = 0.997$$

# Για να υπολογίσουμε πιθανότητες όταν πρέπει $x \sim N(\mu, \sigma^2)$ να τυποποιήσουμε

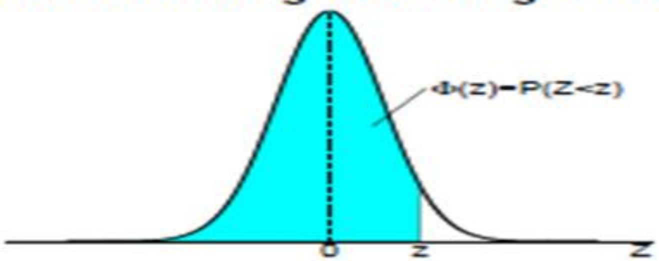
- Αν $x \sim N(\mu, \sigma^2)$ τότε $z = \frac{x-\mu}{\sigma} \sim \mathrm{N}(0,1)$

- Αν $x \sim N(\mu, \sigma^2)$ τότε $z = \frac{x-15}{2} \sim \mathrm{N}(0,1)$

$$P(x < 17) = P\left(\frac{x-\mu}{\sigma} < \frac{17-\mu}{\sigma}\right) = P\left(z < \frac{17-15}{2}\right)$$
$$= P(z < 1) = 0.84$$

$$P(x < 17.5) = P\left(\frac{x-\mu}{\sigma} < \frac{17.5-\mu}{\sigma}\right) = P\left(z < \frac{17.5-15}{2}\right) =$$
$$P(z < 1.25) =;$$

## Στατιστικός Πίνακας Τυπικής Κανονική Κατανομής




Παράδειγμα:

$z = 1.28 \Longleftrightarrow \Phi(z) = 0.90$
$z = 1.65 \Longleftrightarrow \Phi(z) = 0.95$
$z = 2.33 \Longleftrightarrow \Phi(z) = 0.99$
$z = 3.08 \Longleftrightarrow \Phi(z) = 0.999$

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

$$P(Z < -3.5) = P(Z > 3.5) = 0$$

$$P(Z > -3.5) = P(Z < 3.5) = 1$$

$$P(a < Z < \beta) = P(Z < \beta) - \mathrm{P}(\mathrm{z} < \alpha)$$

$$P(Z > \alpha) = P(Z < -\alpha) = 1 - P(Z < \alpha)$$

Η είσοδος σε ένα Πανεπιστήμιο εξαρτάται από την επίδοση σε ένα διαγώνισμα. Η κατανομή των βαθμών σε αυτό το διαγώνισμα είναι κανονική με μέσο όρο 500 και τυπική απόκλιση 100. Η βάση ορίζεται με τέτοιο τρόπο ώστε γίνονται δεκτοί το 30% των φοιτητών με τους καλύτερους βαθμούς. Ένας μαθητής βαθμολογήθηκε με 585. Θα γίνει δεκτός;

Για να γίνει δεκτός, θα πρέπει η πιθανότητα να πάρει κάποιος μικρότερο βαθμό να είναι τουλάχιστον ίση με 0.7

Έστω x η βαθμολογία στο συγκεκριμένο διαγώνισμα

$$x \sim N(500, 100^2)$$

$$P(x < 585) = P\left(\frac{x-\mu}{\sigma} < \frac{585-\mu}{\sigma}\right) = P\left(Z < \frac{585-500}{100}\right) = P(Z < 0.85)$$

$$P(Z < 0.85) = 0.8023$$

Άρα θα γίνει δεκτός

Σε τι βαθμό πρέπει να οριστεί η βάση έτσι ώστε να περνάει στο Πανεπιστήμιο το 30% των υποψηφίων

$$x \sim N(500, 100^2)$$

Ζητάμε το σημείο α για το οποίο

$$P(x < a) = 0.7$$

$$P\left(\frac{x-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right) = 0.7$$

$$P\left(Z < \frac{a-500}{100}\right) = 0.7$$

$$\frac{a-500}{100} = 0.525$$

$$a = 552.5$$

# Ποιά είναι η πιθανότητα ότι στους 5 υποψηφίους θα περάσουν οι 2

- Η πιθανότητα να γίνει κάποιος δεκτός είναι p=0.3
- Κάθε εξέταση ενός μαθητή είναι ένα πείραμα Bernoulli με δυο πιθανά αποτελέσματα
- Έχουμε 5 μαθητές (n=5) και θέλουμε επιτυχίες x=2
- Άρα

$$P(X=2)=\binom{5}{2}\times 0.3^2 \times 0.7^3$$

Έστω ότι η βάση ορίζεται στο 600. Ένας εξεταστής διορθώνει τα γραπτά διαδοχικά. Ποιά είναι η πιθανότητα το πρώτο επιτυχές γραπτό να είναι το τρίτο

Η πιθανότητα να περάσει κάποιος είναι

$$P(x > 600) = 1 - P(x < 600) = 1 - P\left(z < \frac{600 - 500}{100}\right)$$
$$= 1 - \mathrm{P}(\mathrm{Z} < 1) = 0.1587$$

Έστω  ο αριθμός των γραπτών που διόρθωσε ο εξεταστής μέχρι να βρει το πρώτο σωστό

Θέλω 2 αποτυχίες και 1 επιτυχία (γεωμετρική κατανομή)

$$P(y = 3) = (1 - 0.1587)^2 \times 0.1587$$