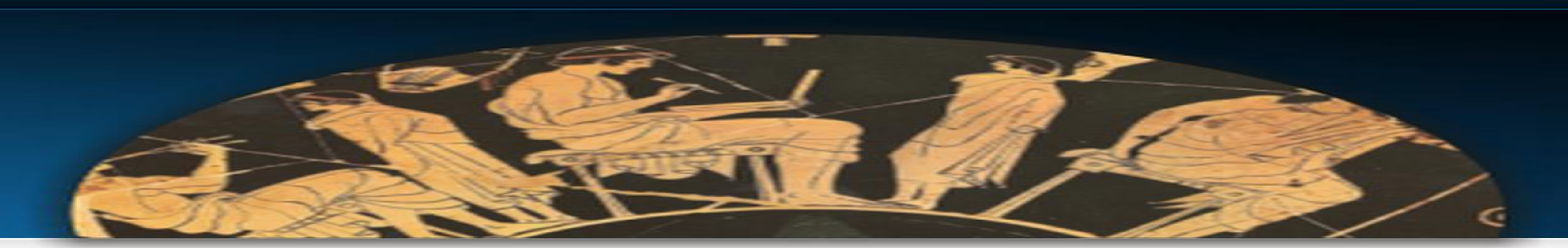
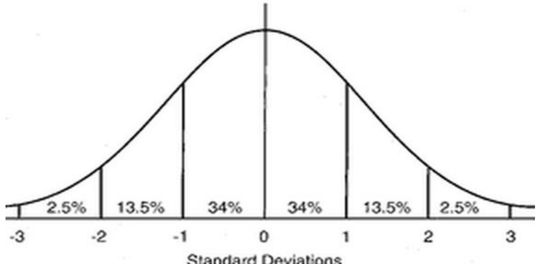


Επαγωγική Στατιστική

Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης

Δημήτρης Μαυρίδης



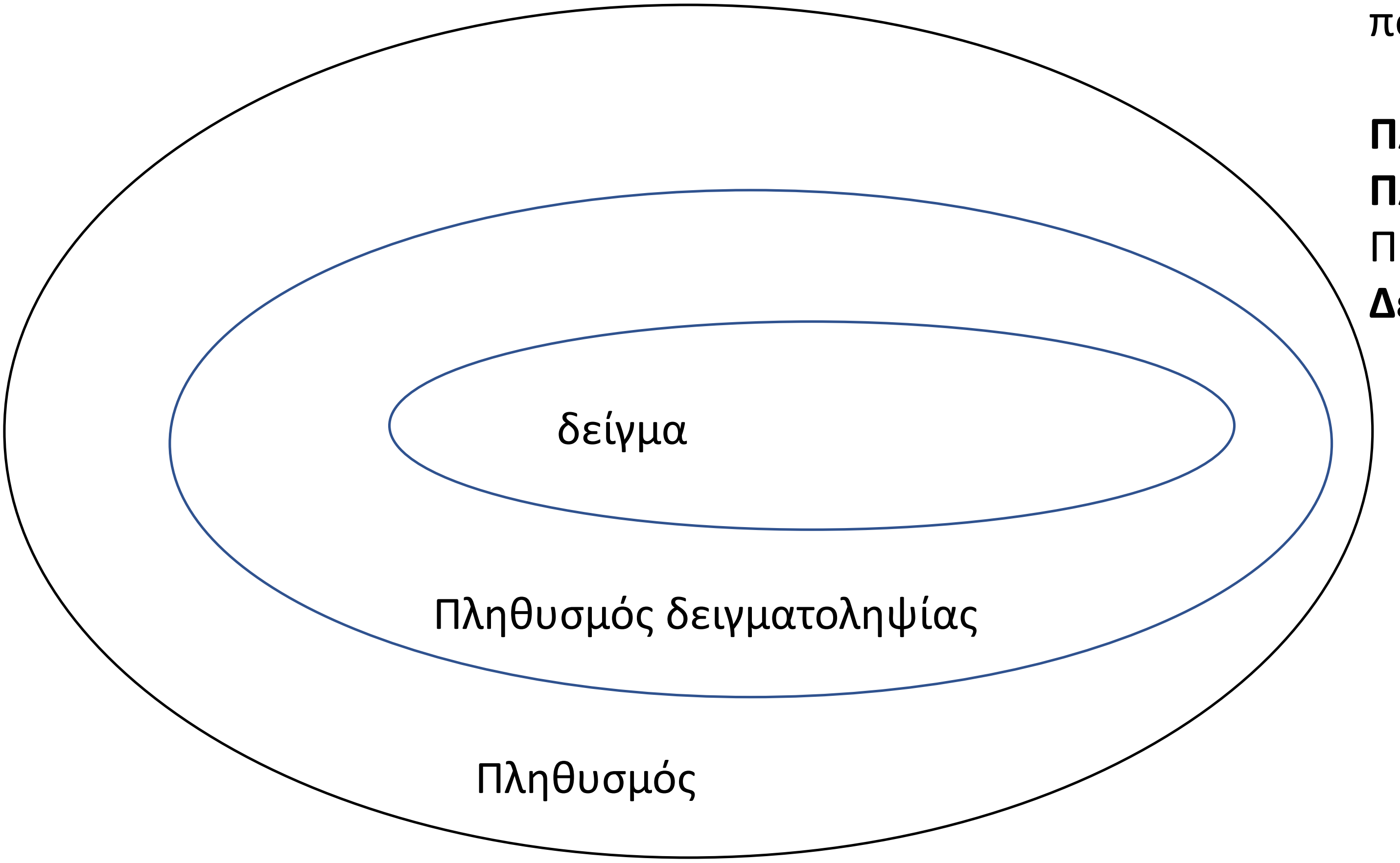
Δείγμα και πληθυσμός

- **Πληθυσμός** (target/study population) είναι το σύνολο των στοιχείων για τα οποία ενδιαφερόμαστε να εξάγουμε συμπεράσματα (όλοι οι Έλληνες/άνδρες/γυναίκες/μαθητές Δημοτικού κλπ όλα τα ελαιόδενδρα, παραγόμενα προϊόντα κλπ όλα τα κοτόπουλα/βοειδή κλπ)
- **Δείγμα** (sample) είναι ένα υποσύνολο του πληθυσμού που μας ενδιαφέρει

Συνήθως, λόγω χρηματικών και χρονικών περιορισμών, δεν μπορούμε να αναλύσουμε όλο τον πληθυσμό και καταφεύγουμε στην ανάλυση ενός αντιπροσωπευτικού δείγματος αυτού.

Ενδέχεται, ο ερευνητής, να μην έχει πρόσβαση σε όλα τα υποκείμενα του πληθυσμούς παρά μόνο σε ένα υποσύνολό του (sampled population)

Δείγμα και πληθυσμός



Παράδειγμα: με ενδιαφέρουν οι απόψεις των φοιτητών ΠΤΔΕ για ένα θέμα για την πρωτοβάθμια εκπαίδευση/ ποιότητα των σπουδών τους

Πληθυσμός: όλοι οι φοιτητές ΠΤΔΕ της Ελλάδας
Πληθυσμός δειγματοληψίας: οι φοιτητές ΠΤΔΕ του Πανεπιστημίου Ιωαννίνων
Δείγμα: αυτοί που παρακολουθούν

Επαγωγική Στατιστική
Inferential Statistics

- Η εξαγωγή συμπερασμάτων από το δείγμα στον πληθυσμό που μας ενδιαφέρει
- Μας ενδιαφέρει να εξάγουμε συμπεράσματα για συγκεκριμένους παραμέτρους του πληθυσμού π.χ. για το μέσο όρο ή το ποσοστό
- Από τα δείγματα, θα πάρουμε μια σημειοεκτίμηση (point estimation) π.χ. ένα μέσο όρο, ένα ποσοστό…
- Κάθε δείγμα, θα μας δίνει και μια διαφορετική εκτίμηση
- Δεν έχουμε τη δυνατότητα να πάρουμε πολλά διαφορετικά δείγματα
- Ιδανικά θα θέλαμε ένα διάστημα τιμών που θα συμπεριλαμβάνει με μεγάλη πιθανότητα (συνήθως την ορίζουμε ως 95%) τις τιμές που θα προέκυπταν από τα πολλαπλά δείγματα ή θα περιλαμβάνει τις πιθανές τιμές της παραμέτρου που με ενδιαφέρει
- Η θεωρία πιθανοτήτων μας προσφέρει αυτό το εργαλείο με τα διαστήματα εμπιστοσύνης

Επαγωγική Στατιστική
Inferential Statistics

- Συνήθως με ενδιαφέρει να ελέγξω μια συγκεκριμένη υπόθεση

π.χ. η πολιτική παράταξη Α θα πάρει πάνω από 35%, η μέση σχολική επίδοση αυτών που χρησιμοποίησαν επιπροσθέτως μια συγκεκριμένη εκπαιδευτική παρέμβαση είναι μεγαλύτερη από αυτούς που δεν την χρησιμοποίησαν

- Υπάρχει κλάδος της Στατιστικής που εστιάζει στον έλεγχο υποθέσεων (hypothesis testing)
- Η στατιστική μας δίνει τη δυνατότητα να υπολογίσουμε την πιθανότητα σφάλματος (να απορρίψουμε μια υπόθεση που είναι σωστή ή να μην απορρίψουμε μια υπόθεση που δεν είναι σωστή)

Εργασία

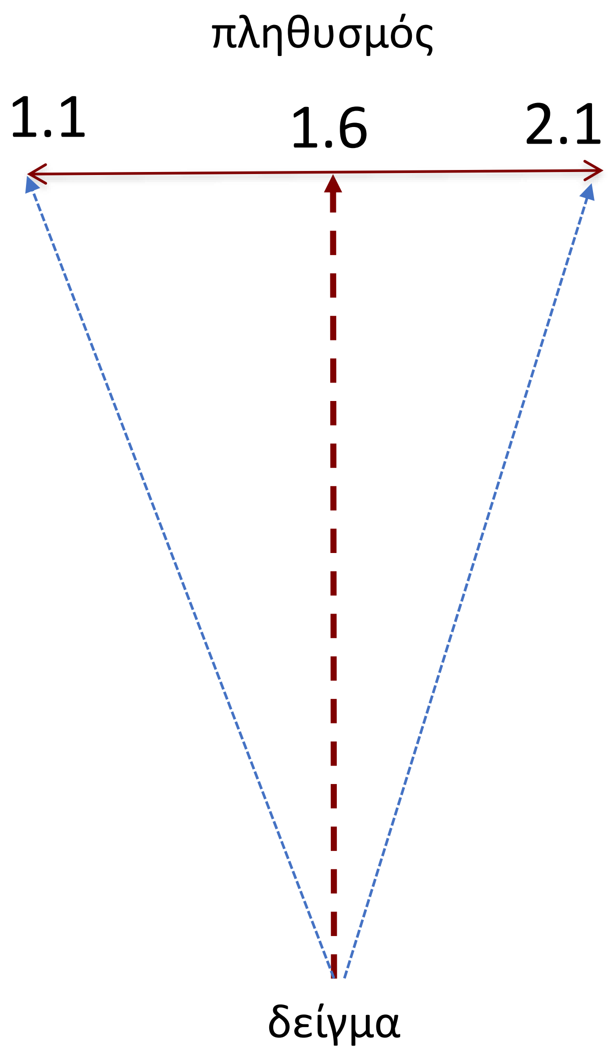
- Δίνει η **μέθοδος με τη χρήση ΤΠΕ** καλύτερα μαθησιακά αποτελέσματα σε σχέση με την **παραδοσιακή διδασκαλία** όσον αφορά στην κατανόηση των μαθηματικών;
- Δοκιμάζουμε για 3 μήνες τη διδασκαλία με τη χρήση ΤΠΕ σε μια τάξη και την παραδοσιακή διδασκαλία σε μια άλλη και μετράμε τη βελτίωση χρησιμοποιώντας ένα τυποποιημένο τεστ
- Χρειαζόμαστε την ομάδα ελέγχου (παραδοσιακή διδασκαλία) γιατί επιθυμούμε να δούμε αν οι μαθητές βελτιώθηκαν σε σχέση με αυτό που τυπικά διδάσκεται
- Υπολογίζουμε το μέσο όρο στις 2 ομάδες (μ_T **και** μ_C **αντίστοιχα) και τη διαφορά** $\Delta = \mu_T - \mu_C$

-0.5 1.6 **2** 2.2 2.5

πείραμα 4 πείραμα 1 πείραμα 3 πείραμα 2

ΤΠΕ vs Παραδοσιακής διδασκαλίας

Εκτίμηση $\pm 1.96 \times$ τυπικό σφάλμα



95% διαστήματα εμπιστοσύνης :

εκφράζουν την αβεβαιότητά μας για την πραγματική τιμή

Αντιπροσωπευτικότητα δείγματος

- Το δείγμα μας πρέπει να είναι αντιπροσωπευτικό του πληθυσμού για να είναι έγκυρη η γενίκευση των αποτελεσμάτων του
- Δείγμα υπολογιζόμενο με πιθανότητες (probability sampling): κάθε μέλος του πληθυσμού (ο οποίος θεωρείται γνωστός εκ των προτέρων) έχει την ίδια πιθανότητα συμμετοχής. Σε μεγάλα δείγματα, η θεωρία πιθανοτήτων εξασφαλίζει ότι το δείγμα μας έχει τα ίδια χαρακτηριστικά με τον πληθυσμό
- Δείγμα υπολογιζόμενο χωρίς πιθανότητες (non-probability samples): Δεν έχουν όλα τα υποκείμενα την ίδια πιθανότητα συμμετοχής στο δείγμα.

Δείγμα ευκολίας (convenient sampling)

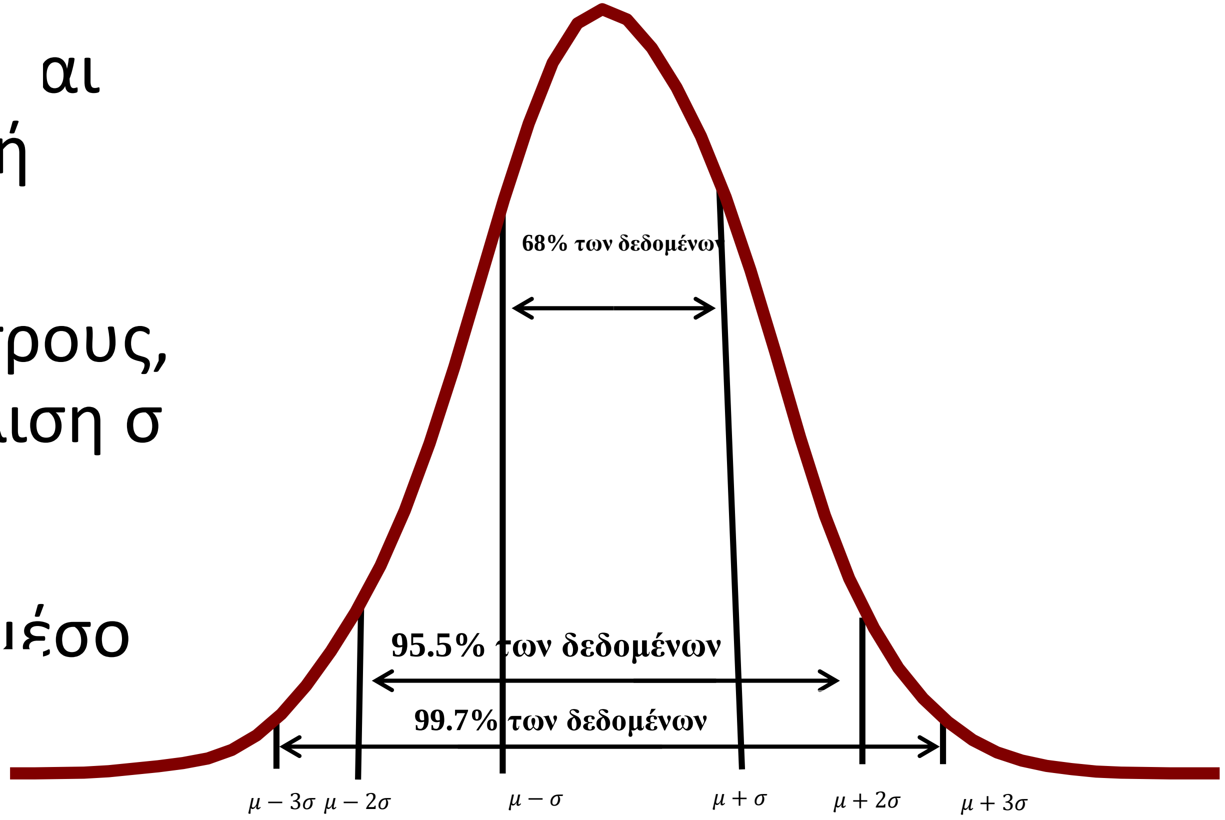
- Πολλές φορές καταφεύγουμε σε δείγματα ευκολίας, σε υποσύνολα του πληθυσμού στα οποία έχουμε εύκολη πρόσβαση.
- Στην εκπαιδευτική έρευνα, ο ερευνητής καταφεύγει σε μια τάξη στην οποία έχει εύκολη πρόσβαση.
- Η απόκτηση δείγματος υπολογιζόμενο με πιθανότητες μπορεί να είναι χρονοβόρα, να έχει υψηλό κόστος ή να είναι ανέφικτη
- Έστω ότι ένας φοιτητής συγκεντρώνει απαντήσεις από ένα ερωτηματολόγιο που έδωσε την ώρα κάποιου μαθήματος. Τα αποτελέσματα ενδεχομένως να μη γενικεύονται στους φοιτητές που δεν παρακολουθούν το μάθημα

Δείγμα και πληθυσμός

παράμετρος	Δείγμα	Πληθυσμός
μέσος	$\bar{x}$	μ
διασπορά	s^2	σ^2
Πιθανότητα/ποσοστό	p	π

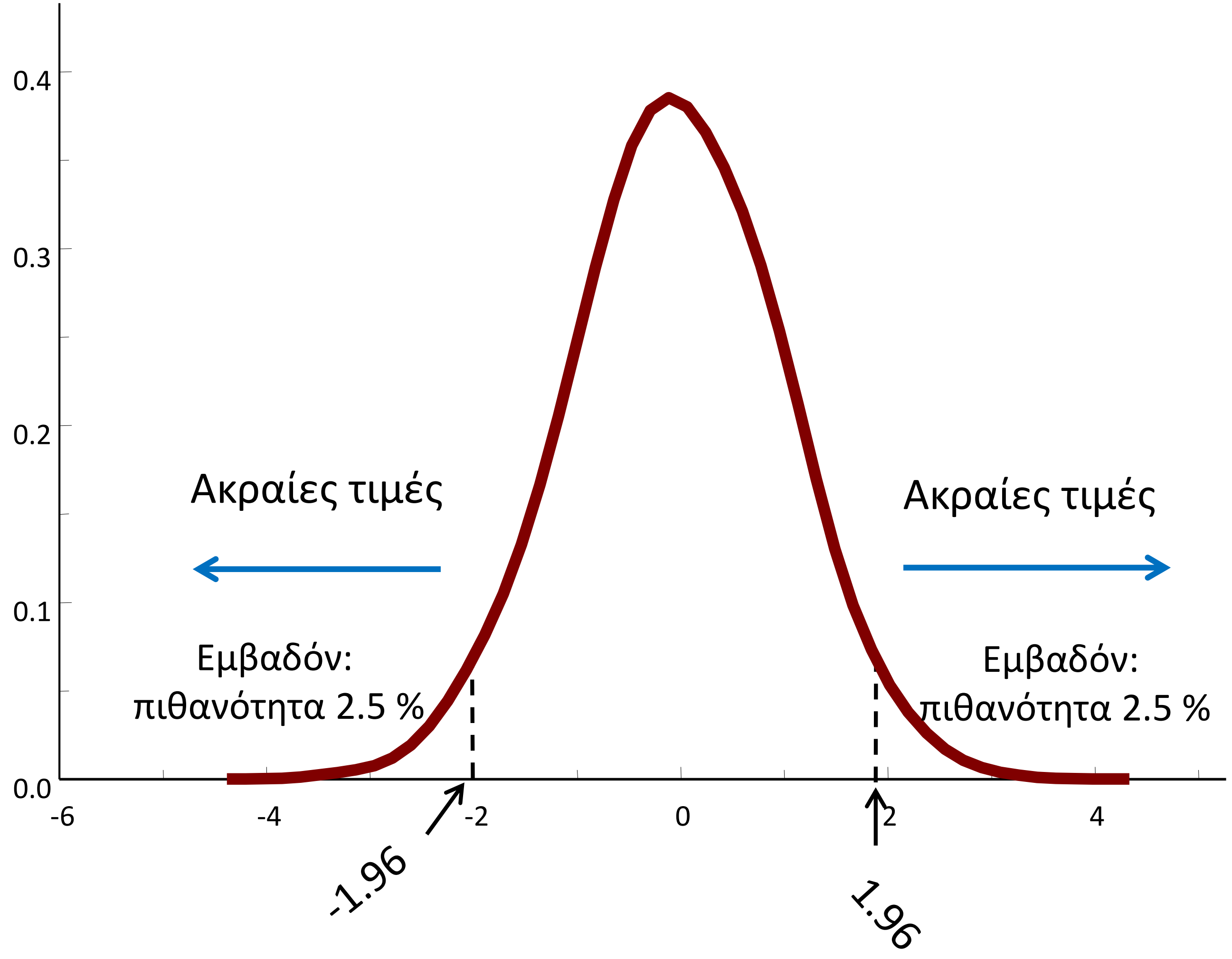
Κανονική κατανομή (Normal distribution)

- Πολλές τυχαίες μεταβλητές χαρακτηρίζονται ή προσεγγίζ αι 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



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

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

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

Κατανομή του δειγματικού μέσου
Κεντρικό Οριακό Θεώρημα

- Αν έχω $x_1, x_2, \dots, x_n$ παρατηρήσεις με μέσο μ και διακύμανση σ^2 τότε

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Έχουμε $\bar{x} = \frac{\sum x_i}{n}$ και εκτιμώ το σ^2 από το δείγμα s^2

$$\mathrm{z} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} \sim N(0,1\ )$$

95% διάστημα εμπιστοσύνης

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

$$P(L<\bar{x}<U) = 0.95$$

$$P\left(\frac{L-\mu}{\sigma/\sqrt{n}} < \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{U-\mu}{\sigma/\sqrt{n}}\right) = 0.95$$

$$P\left(\frac{L-\mu}{\sigma/\sqrt{n}} < z < \frac{U-\mu}{\sigma/\sqrt{n}}\right) = 0.95$$

Επομένως, $\frac{L-\mu}{\sigma/\sqrt{n}} = -1.96 \leftrightarrow L = \mu - 1.96 \times \sigma/\sqrt{n}$
και $\frac{\mathrm{U}-\mu}{\sigma/\sqrt{n}} = 1.96 \leftrightarrow U = \mu + 1.96 \times \sigma/\sqrt{n}$ και το 95% διάστημα εμπιστοσύνης είναι (χρησιμοποιούμε τη δειγματική τυπική απόκλιση s)

$$\bar{x} \pm 1.96 \times \frac{s}{\sqrt{n}}$$

Η είσοδος σε ένα Πανεπιστήμιο εξαρτάται από την επίδοση σε ένα διαγώνισμα. Η κατανομή των βαθμών 400 μαθητών σε αυτό το διαγώνισμα είναι κανονική με μέσο όρο 500 και τυπική απόκλιση 100. Μπορούμε να ισχυριστούμε ότι η μέση βαθμολογία είναι ίση με 515;

- Να φτιαχτεί ένα 95% διάστημα εμπιστοσύνης

$$L = \bar{x} - 1.96 \times \frac{s}{\sqrt{n}} = 500 - 1.96 \times \frac{100}{\sqrt{400}} = 490.2$$

$$U = \bar{x} + 1.96 \times \frac{s}{\sqrt{n}} = 500 + 1.96 \times \frac{100}{\sqrt{400}} = 509.8$$

Επομένως το διάστημα εμπιστοσύνης είναι

$$(490.2, 509.8)$$

Διαστήματα εμπιστοσύνης

90% διάστημα εμπιστοσύνης

εκτίμηση $\pm 1.68 \times$ τυπικό σφάλμα

95% διάστημα εμπιστοσύνης

εκτίμηση $\pm \mathbf{1.96} \times$ τυπικό σφάλμα

99% διάστημα εμπιστοσύνης

εκτίμηση $\pm 2.57 \times$ τυπικό σφάλμα

Τυπικό σφάλμα μέσου όρου και τυπική απόκλιση παρατηρήσεων

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Επηρεάζεται **ελάχιστα ή καθόλου** απο το μέγεθος του δείγματος

$$s.e. = \frac{s}{\sqrt{n}}$$

Επηρεάζεται **πολύ** απο το μέγεθος του δείγματος

Ο χρυσός κανόνας για διαστήματα εμπιστοσύνης

Για μεγέθη που ακολουθούν την κανονική κατανομή, το 95% διάστημα εμπιστοσύνης (95% Confidence Interval) για το **μ** είναι

$$\mu \pm 1.96 \times \frac{s}{\sqrt{n}}$$

τυπικό σφάλμα του **μ**

Διάστημα εμπιστοσύνης

Εκφράζει την ακρίβεια της μέτρησης του μ και είναι συνάρτηση της διασποράς των παρατηρήσεων **και** του μεγέθους δείγματος

Ερμηνεία:
= το πιθανό διάστημα της ‘αλήθειας’ (μ_E)
-Πόσο πιθανό;
-95% πιθανό

Χονδρική ερμηνεία : Το δεδομένο διάστημα έχει πιθανότητα 95% να περιέχει την ‘αληθινή’ τιμή Μ

Παράδειγμα

- Μέτρηση 64 ατόμων με **ΤΠΕ**. Μέση βελτίωση $\mu_{\mathrm{ΤΠΕ}} = 3$ και $s_{\mathrm{ΤΠΕ}} = 1.2$
- Μέτρηση 100 ατόμων με την **παραδοσιακή διδασκαλία**. Μέση βελτίωση $\mu_{\mathrm{ΤΠΕ}} = 2$ και $s_{\mathrm{ΤΠΕ}} = 1.1$

- Σχετίζεται η ΤΠΕ μέθοδος με υψηλότερες βελτιώσεις;

Παράδειγμα: διαστήματα εμπιστοσύνης

- **Παραδοσιακή διδασκαλία:**

SE=1.1/10=0.11
95% CI: (2-1.96·0.11, 2+1.96·0.11)
95% CI: (1.78, 2.22)

$$\mu \pm 1.96 \times \frac{s}{\sqrt{n}}$$

- Χρήση ΤΠΕ:

SE=1.2/9.49=0.13
95% CI: (3-1.96·0.13, 3+1.96·0.13)
95% CI: (2.75, 3.25)

- Υπάρχει διαφορά. Γιατί;

Διάστημα εμπιστοσύνης και μέγεθος δείγματος

Διάστημα εμπιστοσύνης και εμπιστοσύνη					
μέσος	**SD**	**n**	**εμπιστοσύνη**	**CI**	
180	20	100	**99%**	175	185
180	20	100	**95%**	176	184
180	20	100	**90%**	177	183
Διάστημα εμπιστοσύνης και μέγεθος δείγματος					
μέσος	**SD**	**n**	**εμπιστοσύνη**	**CI**	
180	20	**10**	95%	168	192
180	20	**100**	95%	176	184
180	20	**1000**	95%	179	181

Σύγκριση 2 δειγμάτων ως προς ένα μέγεθος (π.χ. μέσο)

Υπολογίζουμε **τα 2 διαστήματα εμπιστοσύνης** για το καθένα δείγμα και βλέπουμε αν αλληλοκαλύπτονται

ή

Υπολογίζουμε **τη διαφορά των δύο δειγμάτων** και **το διάστημα εμπιστοσύνης για την διαφορά** και κοιτάμε αν η τιμή **0** βρίσκεται μεσα στο διάστημα εμπιστοσύνης (ορθότερη προσέγγιση)

Διαφορά δύο μέσων όρων
mean difference (MD)

$$MD = \mu_2 - \mu_1$$

$$s.e.(MD) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

95% διάστημα εμπιστοσύνης:

$$MD \pm 1.96 \times s.e.(MD)$$

Παράδειγμα

Μέτρηση 64 ατόμων απο τη χρήση ΤΠΕ. Μέση βελτίωση =3 και τ.α=1.2

Μέτρηση 100 ατόμων απο την παραδοσιακή διδασκαλία. Μέσο βελτίωση=2 και τ.α=1.1

Το μηδέν δεν περιλαμβάνεται στο CI άρα υπάρχει διαφορά ανάμεσα στις δύο ομάδες.

$$MD = \mu_2 - \mu_1 = 3 - 2 = 1$$

$$s.e.(MD) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{1.1^2}{100} + \frac{1.2^2}{64}} = 0.19$$

$$95\%\ CI\ 1 \pm 1.96 \times 0.19 = (0.63,1.37)$$

Παράδειγμα

Από μια κλινική δοκιμή που τυχαιοποίησε τους ασθενείς σε 2 ομάδες των 100 ατόμων προέκυψαν τα παρακάτω αποτελέσματα

Δίαιτα 1, μέσο βάρος 83 κιλά, SD_1= 31

Δίαιτα 2, μέσο βάρος 80 κιλά, SD_2=30

Ποιά δίαιτα είναι στατιστικά καλύτερη;

MD=-3

95% CI (-11.46, 5.46)

Τυπικό σφάλμα ποσοστού

- r οι θετικές απαντήσεις (όσοι έχουν το γεγονός)
- n το μέγεθος δείγματος
- Το ποσοστό (ή πιθανότητα) **p**=r/n

$$s.e.(p) = \sqrt{\frac{p \times (1-p)}{n}}$$

$$p \pm 1.96 \times s.e.(p)$$

$$p \pm 1.96 \times \sqrt{\frac{p \times (1-p)}{n}}$$

Παράδειγμα

Βελτίωση	**Ναι**	**Όχι**	**Σύνολο**
Χρήση ΤΠΕ	70	30	*100*

$$p = \frac{70}{100} = 0.7$$

$$s.e. = \sqrt{\frac{0.7 \times 0.3}{100}} = 0.046$$

$$p \pm 1.96 \times \sqrt{\frac{p \times (1-p)}{n}}$$

$$0.7 \pm 1.96 \times \sqrt{\frac{0.3 \times 0.7}{100}}$$

$$(0.61, 0.79)$$

Έστω ότι σε μια δημοσκόπηση πήραμε ένα δείγμα 400 ατομών και οι 180 είπαν ότι ψηφίζουν το κόμμα Α

Ποσοστό $p = \frac{180}{400} = 0.45$

Για να έχει αυτοδυναμία το κόμμα Α θα πρέπει να πάρει πάνω από το 50% των ψήφων

Έχει αυτοδυναμία;

90% διάστημα εμπιστοσύνης
(0.4091,   0.4909)

95% διάστημα εμπιστοσύνης
(0.4012,   0.4988)

99% διάστημα εμπιστοσύνης
(0.3859 ,   0.5141)

Έστω ότι ρωτήσαμε 800 άτομα και πήραμε το ίδιο ποσοστό (οι 360 είπαν ότι ψηφίζουν το κόμμα Α)

Το ποσοστό είναι 45%

90% διάστημα εμπιστοσύνης
(0.4211 ,   0.4789)

95% διάστημα εμπιστοσύνης
(0.4155 ,   0.4845)

99% διάστημα εμπιστοσύνης
(0.4047 ,   0.4953)

Τι θα γίνει αν πάρω ένα 100% διάστημα εμπιστοσύνης
(0,1)

Μετρήθηκε η κατάθλιψη χρησιμοποιώντας μια επικυρωμένη κλίμακα σε μια ομάδα Α που ακολουθούσε ένα πρόγραμμα γυμναστικής και σε μια άλλη ομάδα Β που δεν χρησιμοποιούσε το εν λόγω πρόγραμμα (ομάδα ελέγχου). Η ομάδα Α μείωσε τα επίπεδα κατάθλιψης κατά μια μονάδα περισσότερο σε σχέση με την ομάδα Β ενώ το 95% διάστημα εμπιστοσύνης για τη μείωση της κατάθλιψης κυμαίνεται από 0.2 έως 1.8 μονάδες

- Γιατί χρειάζεται η ομάδα ελέγχου;
- Το πρόγραμμα γυμναστικής βοήθησε τους συμμετέχοντες;
- Θα άλλαζαν τα αποτελέσματα με ένα 99% διάστημα εμπιστοσύνης; Με ένα 90%;
- Τι θα συμπεραίνετε αν το 95% διάστημα εμπιστόσυνης κυμαινόταν από

-0.5 έως 2.5;

- Τι πρέπει να εξασφαλίσει ο ερευνητής για να κάνει τις 2 ομάδες συγκρίσιμες;