

Εισαγωγή στις βασικές έννοιες της Στατιστικής

Γούργουλης Βασίλης, Ph.D.
Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ.

1

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

- μεταβλητή
 - εξαρτημένη μεταβλητή
 - ανεξάρτητη μεταβλητή - παράγοντας
 - διασπορά
- απόκλιση
 - διακύμανση
- τυπική απόκλιση
 - δείγμα
- πληθυσμός
 - σφάλμα δειγματοληψίας
- δειγματική κατανομή
 - κανονική κατανομή
- z - τυπική κανονική κατανομή
 - τυπικό σφάλμα δειγματικού μέσου όρου
- διάστημα εμπιστοσύνης
 - t - κατανομή

2

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Μεταβλητή:

Το οποιοδήποτε χαρακτηριστικό

Αν το μετρήσεις !!! ...

Εξαρτημένη Μεταβλητή:

Αυτό που μετράμε

Από τι επηρεάζεται;

Ανεξάρτητη Μεταβλητή - Παράγοντας:

Επηρεάζει αυτό που μετράμε

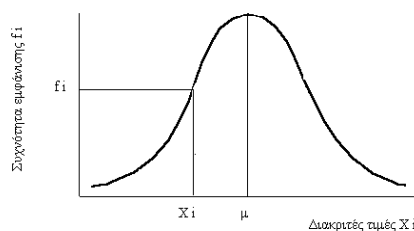
3

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Διασπορά:

πώς κατανέμονται

οι τιμές μιας μεταβλητής γύρω από το μέσο όρο



δηλαδή

πόσο απλωμένες

είναι οι τιμές του δείγματος γύρω από το μέσο όρο

4

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η διασπορά
των τιμών μιας μεταβλητής γύρω από το μέσο όρο
θα μπορούσε να φανεί
από το σύνολο των αποκλίσεων της κάθε τιμής X_i
από τον αριθμητικό μέσο $\bar{X}$

$$\sum (X_i - \bar{X})$$

5

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Παράδειγμα:

X_i	$X_i - \bar{X}$
9	$9 - 6 = 3$
8	$8 - 6 = 2$
8	$8 - 6 = 2$
7	$7 - 6 = 1$
5	$5 - 6 = -1$
5	$5 - 6 = -1$
4	$4 - 6 = -2$
2	$2 - 6 = -4$

Το σύνολο των αποκλίσεων
της κάθε τιμής (X_i)
από το μέσο όρο ($\bar{X}$)
είναι πάντα μηδέν (0)

$N = 8$ $\sum X_i = 48$ $\sum (X_i - \bar{X}) = 0$

$$\bar{X} = \frac{\sum X_i}{N} = 6$$

6

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Για να δηλώσουμε
τη διασπορά μιας μεταβλητής
(πώς κατανέμονται οι τιμές της γύρω από το μέσο όρο)
χρησιμοποιείται
η **διακύμανση** (Variance)

σ^2  (για τον πληθυσμό)

s^2  (για το δείγμα)

(που είναι μια τιμή πάντοτε θετική)

7

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Διακύμανση πληθυσμού

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

άθροισμα των τετραγώνων των αποκλίσεων
της κάθε τιμής (X_i) (του πληθυσμού)
από το μέσο όρο του πληθυσμού (μ)
δια
του συνολικού αριθμού των τιμών (N)

8

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Διακύμανση δείγματος

$$s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}$$

άθροισμα των τετραγώνων των αποκλίσεων

της κάθε τιμής (X_i) (του δείγματος)

από το μέσο όρο του δείγματος ($\bar{X}$)

δια

του συνολικού αριθμού των τιμών του δείγματος μείον 1 ($N - 1$)

9

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Παράδειγμα:

X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
9	$9 - 6 = 3$	$3^2 = 9$
8	$8 - 6 = 2$	$2^2 = 4$
8	$8 - 6 = 2$	$2^2 = 4$
7	$7 - 6 = 1$	$1^2 = 1$
5	$5 - 6 = -1$	$(-1)^2 = 1$
5	$5 - 6 = -1$	$(-1)^2 = 1$
4	$4 - 6 = -2$	$(-2)^2 = 4$
2	$2 - 6 = -4$	$(-4)^2 = 16$

$$N = 8 \quad \sum_{i=1}^{N=8} X_i = 48 \quad \sum_{i=1}^{N=8} (X_i - \bar{X}) = 0 \quad \sum_{i=1}^{N=8} (X_i - \bar{X})^2 = 40$$

$$\bar{X} = \frac{\sum_{i=1}^{N=8} X_i}{N} = 6 \quad s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1} = \frac{40}{8 - 1} \cong 5.71$$

10

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Αν
όλες οι παρατηρήσεις (X_i)
έχουν την ίδια τιμή ($X_i = \bar{X}$)

τότε

η διακύμανση θα είναι μηδέν ($s^2 = 0$)

ενώ

όσο περισσότερο διασκορπίζονται γύρω από το μέσο όρο
τόσο μεγαλύτερη θα είναι η διακύμανση.

11

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Επειδή
η **διακύμανση** (s^2) δίνεται σε μονάδες υψωμένες στο τετράγωνο

για καθαρά υπολογιστικούς λόγους
χρησιμοποιείται

η τετραγωνική ρίζα της διακύμανσης

που καλείται

τυπική απόκλιση (standard deviation)

(έχει τις ίδιες μονάδες μέτρησης όπως και οι παρατηρήσεις)

12

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Τυπική απόκλιση πληθυσμού

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$$

Τυπική απόκλιση δείγματος

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}}$$

13

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Σε μια έρευνα
προσπαθούμε

- από τις τιμές ενός δείγματος (στατιστικά)

επαγωγικά

να βγάλουμε συμπεράσματα για

- την αντίστοιχη τιμή του πληθυσμού (παράμετρος)
(από τον οποίο προέρχεται το δείγμα)

14

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Υπάρχει κίνδυνος
λανθασμένου συμπεράσματος
για την παράμετρο του πληθυσμού.

Παράδειγμα: Μηδενική υπόθεση $H_0 : \mu = \bar{X}$

Μπορεί να συμβούν δύο ειδών σφάλματα:

1. **Σφάλμα τύπου I:**
να απορρίψουμε τη μηδενική υπόθεση ενώ είναι αληθής
2. **Σφάλμα τύπου II:**
να αποδεχτούμε τη μηδενική υπόθεση ενώ είναι ψευδής

15

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η πιθανότητα να συμβεί **σφάλμα τύπου I**
συμβολίζεται με α
και καλείται **επίπεδο σημαντικότητας**

Παράδειγμα: Καθορίζοντας επίπεδο σημαντικότητας $\alpha = 0.05$
καθορίζουμε την πιθανότητα να διαπράξουμε
σφάλμα τύπου I
δηλ. να απορρίψουμε τη μηδενική υπόθεση
ενώ είναι αληθής, σε ποσοστό 5%.

16

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Παράδειγμα:

$H_0 : \mu = \bar{X}$ δηλαδή ο μέσος όρος του δείγματος
είναι ίσος με το μέσο όρο του πληθυσμού
από τον οποίο προέρχεται το δείγμα

Αν απορρίψουμε τη μηδενική υπόθεση $H_0 : \mu = \bar{X}$

σε επίπεδο σημαντικότητας $\alpha = 0.05$

αποδεχόμαστε την εναλλακτική υπόθεση $H_A : \mu \neq \bar{X}$

$H_A : \mu \neq \bar{X}$ δηλαδή ο μέσος όρος του δείγματος
είναι διαφορετικός από το μέσο όρο του πληθυσμού
από τον οποίο προέρχεται το δείγμα,
ενώ υπάρχει πιθανότητα
από τις 100 περιπτώσεις
να έχουμε κάνει λάθος στις 5 περιπτώσεις
(όπου εκεί οι μέσοι όροι θα είναι ίσοι)

17

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Διαφορετικά δείγματα

οδηγούν

σε διαφορετικές εκτιμήσεις
των παραμέτρων του πληθυσμού

Συνεπώς

υπάρχει πάντα ένα σφάλμα δειγματοληψίας

και είναι πάρα πολύ σπάνιο

η μέση τιμή του δείγματος ($\bar{X}$)

να συμπίπτει

με τη μέση τιμή του πληθυσμού (μ)

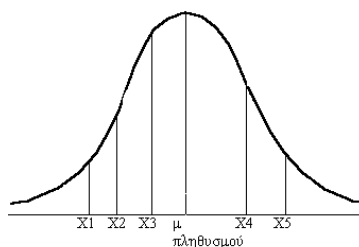
18

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Αν όμως
 πάρουμε από έναν πληθυσμό πολλά δείγματα ίσου μεγέθους N
 και υπολογίσουμε το μέσο όρο του κάθε δείγματος
 τότε η κατανομή των μέσων όρων των δειγμάτων (δειγματική κατανομή)

θα ακολουθεί την κανονική κατανομή

και
 ο μέσος όρος των μέσων όρων των επιμέρους δειγμάτων
 θα είναι ίσος με το μέσο όρο του πληθυσμού
 (σύμφωνα με το **Κεντρικό Οριακό Θεώρημα**)



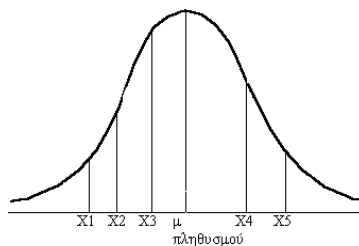
19

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Συνεπώς

αν επιλέξουμε ένα δείγμα από κάποιον πληθυσμό
 και υπολογίσουμε
 τη μέση τιμή κάποιου χαρακτηριστικού (μεταβλητής)

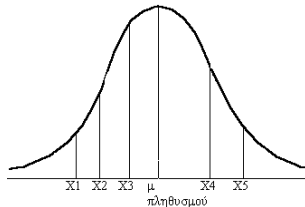
αυτή η μέση τιμή θα είναι
 μια τιμή από τη **θεωρητική (δειγματική) κατανομή** των δειγματικών μέσων τιμών
 η οποία ακολουθεί την κανονική κατανομή.



20

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η δειγματική κατανομή,
δηλαδή η κατανομή των μέσων όρων πολλών δειγμάτων,
ίσου μεγέθους N , που προέρχονται από τον ίδιο πληθυσμό,
ακολουθεί
την κανονική κατανομή:



Η κανονική κατανομή

- έχει τη **μορφή καμπάνας**
- οι περισσότεροι μέσοι όροι των δειγμάτων βρίσκονται **κοντά** στον πραγματικό μέσο όρο του πληθυσμού
- λίγοι μέσοι όροι θα βρίσκονται σε μια μεγαλύτερη απόσταση και
- ακόμη λιγότεροι θα είναι πολύ απομακρυσμένοι από το μέσο όρο του πληθυσμού

21

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Κάθε κανονική κατανομή
μπορεί να μετασχηματιστεί
σε τυπική κανονική κατανομή

όπου:

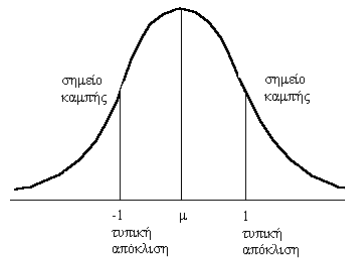
ΔΕΝ μελετούμε πλέον τις **αρχικές τιμές** X_i
ΑΛΛΑ τις μετασχηματίζουμε σε **z - τυπικές τιμές**
και μελετούμε την απόστασή τους από το μέσο όρο μ
σε μια νέα κλίμακα
που έχει ως μονάδα μέτρησης την τυπική απόκλιση (σ).

$$z = \frac{X - \mu}{\sigma}$$

22

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Στην καμπύλη της z - τυπικής κατανομής
υπάρχουν δύο **σημεία καμπής**,
όπου η καμπύλη αλλάζει και
από κυρτή μετατρέπεται σε κοίλη.



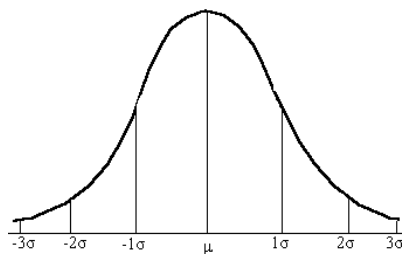
Η απόσταση του κάθε σημείου καμπής από το μέσο όρο
ισούται με μία (1) τυπική απόκλιση

23

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η z - **τυπική κατανομή** διαφέρει από την κανονική κατανομή

- οι μονάδες στον άξονα των X είναι **μονάδες τυπικής απόκλισης** και όχι μονάδες της μεταβλητής X
- ο **μέσος όρος** της τυπικής κατανομής είναι **μηδέν (0)**
- η **τυπική απόκλιση** της τυπικής κατανομής είναι **= 1**



24

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Μετατρέποντας
την κανονική κατανομή
σε z - τυπική κατανομή

μπορούμε

να περιγράψουμε τη σχετική θέση της κάθε τιμής X_i
μέσα στην κατανομή

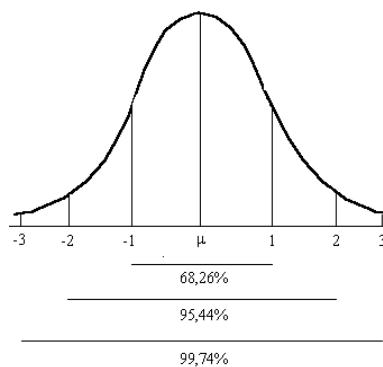
λέγοντας

πόσες τυπικές αποκλίσεις μακριά από το μέσο όρο
βρίσκεται η συγκεκριμένη τιμή X_i

25

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η καμπύλη της z - τυπικής κατανομής έχει συγκεκριμένες μαθηματικές ιδιότητες



- συνολικό εμβαδόν κάτω από την καμπύλη της τυπικής κατανομής = 100%
- το τμήμα μεταξύ -1 και +1 τυπικές αποκλίσεις περιέχει το 68,26%
- το τμήμα μεταξύ -2 και +2 τυπικές αποκλίσεις περιέχει το 95,44%
- το τμήμα μεταξύ -3 και +3 τυπικές αποκλίσεις περιέχει το 99,74%

26

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η τυπική απόκλιση της θεωρητικής δειγματικής κατανομής
είναι ένα μέτρο
του σφάλματος δειγματοληψίας
καλείται
τυπικό σφάλμα του δειγματικού μέσου όρου
και δίνεται από τη σχέση

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$$

$\sigma_{\bar{X}}$ = τυπικό σφάλμα (τυπική απόκλιση) του δειγματικού μέσου όρου

σ = τυπική απόκλιση του πληθυσμού

N = μέγεθος του δείγματος

27

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Συνεπώς

Αν

γνωρίζουμε

την τυπική απόκλιση της θεωρητικής (δειγματικής) κατανομής
των δειγματικών μέσων τιμών

τότε

μπορούμε να προβλέψουμε,
με μεταβλητό βαθμό εμπιστοσύνης (α),
«**πόσο μακριά**» από τη μέση τιμή του πληθυσμού (μ)
βρίσκεται η μέση τιμή του δείγματος ($\bar{X}$)

28

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Επειδή συνήθως δεν γνωρίζουμε
 τις παραμέτρους του πληθυσμού,
 δηλαδή το μέσο όρο του πληθυσμού (μ)
 και την τυπική απόκλιση του πληθυσμού (σ)
εκτιμούμε
 το τυπικό σφάλμα του δειγματικού μέσου όρου
 (δηλαδή πόσο απέχει ο μέσος όρος του δείγματος ($\bar{X}$)
 από τον πραγματικό μέσο όρο του πληθυσμού (μ))
 από την τυπική απόκλιση του δείγματος (s)
 μέσω του τύπου

$$S_{\bar{X}} = \frac{s}{\sqrt{N}}$$

$S_{\bar{X}}$ = τυπικό σφάλμα (τυπική απόκλιση) του δειγματικού μέσου όρου

s = τυπική απόκλιση του δείγματος

N = μέγεθος του δείγματος

29

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η μετατροπή

κανονική κατανομή



z - τυπική κατανομή

χρησιμοποιείται όταν
 το μέγεθος του δείγματος είναι **μεγάλο** $N > 30$
 και ακολουθεί την κανονική κατανομή.

Όταν
 το μέγεθος του δείγματος είναι **μικρό** ($N < 30$)
 χρησιμοποιείται η **t - κατανομή**

30

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Όταν
το μέγεθος του δείγματος είναι **μεγάλο** $N > 30$
και ακολουθεί την κανονική κατανομή
μετατρέποντας την
κανονική κατανομή
σε
 z - τυπική κατανομή
μπορούμε να ξέρουμε **μεταξύ ποιών z - τιμών**,
δηλαδή **πόσες τυπικές αποκλίσεις μακριά από το μέσο όρο**,
βρίσκεται ένα συγκεκριμένο ποσοστό τιμών.
Έτσι μπορούμε να καθορίσουμε
ένα **διάστημα εμπιστοσύνης**,
δηλαδή τις τιμές μεταξύ των οποίων θα βρίσκεται
ο μέσος όρος του πληθυσμού,
με πιθανότητα που ορίζεται από το επίπεδο σημαντικότητας

31

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Παράδειγμα

Αν

ένα δείγμα μεγέθους $N = 400$ άτομα
έχει μέσο όρο $\bar{X} = 61,5$ και
τυπική απόκλιση $s = 20$

ποιο είναι το διάστημα εμπιστοσύνης

μεταξύ των τιμών του οποίου
θα βρίσκεται ο μέσος όρος του πληθυσμού
από τον οποίο προέρχεται το δείγμα;

32

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

1. Καθορίζουμε το επίπεδο σημαντικότητας α

Π.χ. $\alpha = 0.05$

Δηλαδή υπάρχει πιθανότητα 5%
να διαπράξουμε **σφάλμα τύπου I**
και να απορρίψουμε τη μηδενική υπόθεση

$$H_0 : \mu = \bar{X}$$

ενώ είναι αληθής.

Οι τιμές μεταξύ των οποίων θα βρίσκεται το 95% των μέσων όρων,
όλων των τυχαίων δειγμάτων που θα μπορούσαν να είχαν επιλεγεί,
άρα και ο μέσος όρος του πληθυσμού,
αποτελούν το διάστημα εμπιστοσύνης

33

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

2. Υπολογίζουμε το **τυπικό σφάλμα** του μέσου όρου του δείγματος

δηλαδή το **σφάλμα δειγματοληψίας**

$$s_{\bar{X}} = \frac{s}{\sqrt{N}}$$

Τυπική απόκλιση δείγματος $s = 20$

Μέγεθος δείγματος $N = 400$

Τυπικό σφάλμα μέσου όρου του δείγματος

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{20}{\sqrt{400}} = \frac{20}{20} = 1$$

34

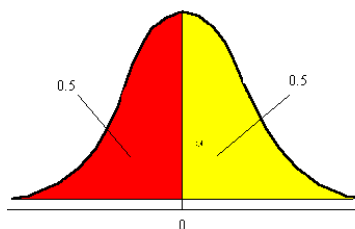
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

3. Καθορίζουμε το **διάστημα εμπιστοσύνης**

Από τον πίνακα της z - τυπικής κατανομής εντοπίζουμε τις **κρίσιμες z - τιμές** (για το συγκεκριμένο επίπεδο σημαντικότητας $\alpha = 0.05$) πέρα από τις οποίες υπάρχει πιθανότητα να κάνουμε λάθος.

Η **καμπύλη της z - τυπικής κατανομής**

- είναι μια **συμμετρική** καμπύλη
- το συνολικό εμβαδόν ισούται με 1
- το εμβαδόν της μισής καμπύλης ισούται με 0.5.

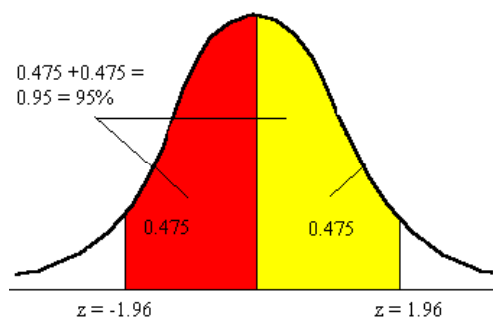


Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

35

Εφόσον το **επίπεδο σημαντικότητας** έχει οριστεί ως $\alpha = 0.05$ αυτό σημαίνει ότι υπάρχει **πιθανότητα 5% να διαπράξουμε σφάλμα** ενώ υπάρχει **πιθανότητα 95% να πάρουμε τη σωστή απόφαση**.

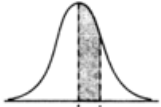
Άρα ο μέσος όρος του πληθυσμού θα πρέπει να εντοπίζεται μεταξύ των κρίσιμων z - τιμών που οριοθετούν το 95% των τιμών κάτω από την καμπύλη της z - τυπικής κατανομής



Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

36

Η κανονική κατανομή

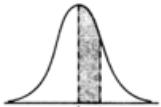


z	0	1	2	3	4	5	6	7	8	9
0.0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0.1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0.7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2996	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3254	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1.4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1.6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4676	4683	4690	4696	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2.1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2.2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2.3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2.4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2.9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986
3.0	4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3.1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3.2	4993	4993	4994	4994	4994	4994	4994	4995	4995	4995
3.3	4995	4995	4995	4996	4996	4996	4996	4996	4996	4997
3.4	4997	4997	4997	4997	4997	4997	4997	4997	4997	4998
3.5	4998	4998	4998	4998	4998	4998	4998	4998	4998	4998
3.6	4998	4998	4999	4999	4999	4999	4999	4999	4999	4999
3.7	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3.8	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3.9	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Γούργουλης Βασίλειος, Ph.D.

37

Η κανονική κατανομή

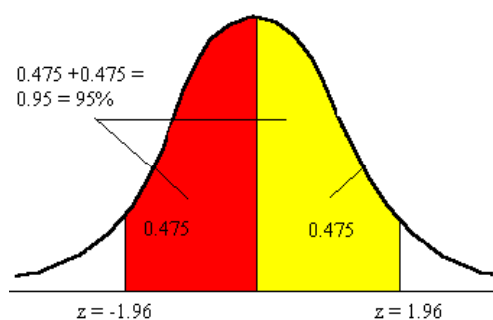


z	0	1	2	3	4	5	6	7	8	9
0.0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0.1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0.7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2996	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3254	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1.4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1.6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4676	4683	4690	4696	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2.1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2.2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2.3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2.4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2.9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986
3.0	4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3.1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3.2	4993	4993	4994	4994	4994	4994	4994	4995	4995	4995
3.3	4995	4995	4995	4996	4996	4996	4996	4996	4996	4997
3.4	4997	4997	4997	4997	4997	4997	4997	4997	4997	4998
3.5	4998	4998	4998	4998	4998	4998	4998	4998	4998	4998
3.6	4998	4998	4999	4999	4999	4999	4999	4999	4999	4999
3.7	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3.8	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3.9	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Γούργουλης Βασίλειος, Ph.D.

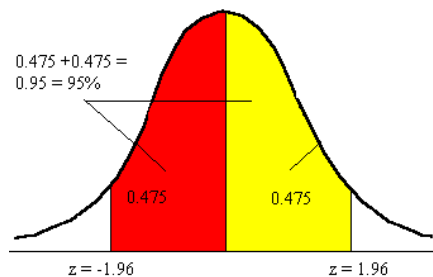
38

z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
1.0	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857



39

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ



Το 95% όλων των πιθανών δειγμάτων
θα βρίσκεται 1.96 τυπικές αποκλίσεις
μακριά από το μέσο όρο του πληθυσμού

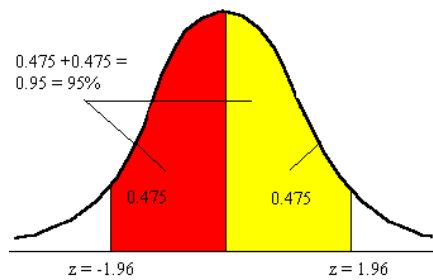
Το 95% όλων των τυχαίων δειγμάτων
θα έχει μέσο όρο που θα κυμαίνεται μεταξύ των τιμών

$$\bar{X}_i + z_{0.05} \cdot s_{\bar{X}} > \bar{X}_o > \bar{X}_i - z_{0.05} \cdot s_{\bar{X}}$$

που καλείται διάστημα εμπιστοσύνης

40

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ



$$\bar{X}_i + z_{0.05} \cdot s_{\bar{X}} > \mu > \bar{X}_i - z_{0.05} \cdot s_{\bar{X}}$$

$$61.5 + 1.96 \cdot 1 > \mu > 61.5 - 1.96 \cdot 1$$

$$63.46 > \mu > 59.54$$

41

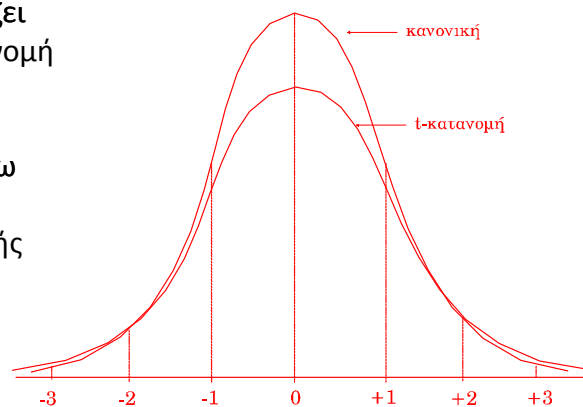
Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. - Σ.Ε.Φ.Α.Α. - Δ.Π.Θ

Όταν
το μέγεθος του δείγματος είναι **μικρό ($N < 30$)**
για τον καθορισμό των διαστημάτων εμπιστοσύνης
μετατρέπουμε την
κανονική κατανομή
σε
***t* - κατανομή**

42

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. - Σ.Ε.Φ.Α.Α. - Δ.Π.Θ

Η **t - κατανομή** μοιάζει
με την κανονική κατανομή
αλλά
όλα τα σημεία της
βρίσκονται πιο κάτω
από τα σημεία της
κανονικής κατανομής



Αν θέλουμε να βρούμε
μεταξύ ποιών τιμών βρίσκεται το 95% διάστημα εμπιστοσύνης,
χρησιμοποιώντας την t - κατανομή, θα πρέπει να «μετακινηθούμε»
σε απόσταση ακόμη μεγαλύτερη από τα σημεία που ορίζονται από το
 $z_{0.05} \cdot s_{\bar{X}} (= 1.96 \cdot s_{\bar{X}})$ (που είναι το όριο της κανονικής κατανομής)

43

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Για **διαφορετικό μέγεθος δείγματος (N)**
υπάρχει και μια **διαφορετική t - κατανομή**.

Στην t - κατανομή
οι αποστάσεις των τιμών μιας μεταβλητής από το μέσο όρο,
μετρημένες σε μονάδες τυπικής απόκλισης
(σφάλματος δειγματοληψίας)
εκφράζονται από τις t - τιμές

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N}}}$$

44

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Παράδειγμα

Αν

ένα δείγμα μεγέθους $N = 16$ άτομα
έχει μέσο όρο $\bar{X} = 61,5$ και
τυπική απόκλιση $s = 20$

ποιο είναι το διάστημα εμπιστοσύνης

μεταξύ των τιμών του οποίου
θα βρίσκεται ο μέσος όρος του πληθυσμού
από τον οποίο προέρχεται το δείγμα;

45

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

1. Καθορίζουμε το επίπεδο σημαντικότητας α

Π.χ. $\alpha = 0.05$

Δηλαδή υπάρχει πιθανότητα 5%
να διαπράξουμε **σφάλμα τύπου I**
και να απορρίψουμε τη μηδενική υπόθεση

$$H_0 : \mu = \bar{X}$$

ενώ είναι αληθής.

Οι τιμές μεταξύ των οποίων θα βρίσκεται το 95% των μέσων όρων,
όλων των τυχαίων δειγμάτων που θα μπορούσαν να είχαν επιλεγεί,
άρα και ο μέσος όρος του πληθυσμού,
αποτελούν το διάστημα εμπιστοσύνης

46

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

2. Υπολογίζουμε το **τυπικό σφάλμα** του μέσου όρου του δείγματος

Δηλαδή το **σφάλμα δειγματοληψίας**

$$s_{\bar{x}} = \frac{s}{\sqrt{N}}$$

Τυπική απόκλιση δείγματος $s = 20$

Μέγεθος δείγματος $N = 16$

Τυπικό σφάλμα μέσου όρου του δείγματος

$$s_{\bar{x}} = \frac{s}{\sqrt{N}} = \frac{20}{\sqrt{16}} = \frac{20}{4} = 5$$

47

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

3. Καθορίζουμε το **διάστημα εμπιστοσύνης**

Από τον πίνακα της t - τυπικής κατανομής
εντοπίζουμε τις **κρίσιμες t - τιμές**

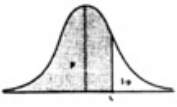
- για το συγκεκριμένο επίπεδο σημαντικότητας $\alpha = 0.05$
- και **N-1** ($16 - 1 = 15$) **βαθμούς ελευθερίας**
πέρα από τις οποίες υπάρχει πιθανότητα να κάνουμε λάθος.

Για το συγκεκριμένο επίπεδο σημαντικότητας $\alpha = 0.05$
και N-1 ($16 - 1 = 15$) βαθμούς ελευθερίας
 $t_{\text{κρίσιμη - τιμή}} = 2.13$

48

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Η t-κατανομή (Student)



v	t.55	t.60	t.70	t.75	t.80	t.90	t.95	t.975	t.99	t.995
1	158	325	727	1.000	1.376	3.08	6.31	12.71	31.82	63.66
2	142	289	617	.816	1.061	1.89	2.92	4.30	6.96	9.92
3	137	277	584	.765	.978	1.64	2.35	3.18	4.54	5.84
4	134	271	569	.741	.941	1.53	2.13	2.78	3.75	4.60
5	132	267	559	.727	.920	1.48	2.02	2.57	3.36	4.03
6	131	265	553	.718	.906	1.44	1.94	2.45	3.14	3.71
7	130	263	549	.711	.896	1.42	1.90	2.36	3.00	3.50
8	130	262	546	.706	.889	1.40	1.86	2.31	2.90	3.36
9	129	261	543	.703	.883	1.38	1.83	2.26	2.82	3.25
10	129	260	542	.700	.879	1.37	1.81	2.23	2.76	3.17
11	129	260	540	.697	.876	1.36	1.80	2.20	2.72	3.11
12	128	259	539	.695	.873	1.36	1.78	2.18	2.68	3.06
13	128	259	538	.694	.870	1.35	1.77	2.16	2.65	3.01
14	128	258	537	.692	.868	1.34	1.76	2.14	2.62	2.98
15	128	258	536	.691	.866	1.34	1.75	2.13	2.60	2.95
16	128	258	535	.690	.865	1.34	1.75	2.12	2.58	2.92
17	128	257	534	.689	.863	1.33	1.74	2.11	2.57	2.90
18	127	257	534	.688	.862	1.33	1.73	2.10	2.55	2.88
19	127	257	533	.688	.861	1.33	1.73	2.09	2.54	2.86
20	127	257	533	.687	.860	1.32	1.72	2.09	2.53	2.84
21	127	257	532	.686	.859	1.32	1.72	2.08	2.52	2.83
22	127	256	532	.686	.858	1.32	1.72	2.07	2.51	2.82
23	127	256	532	.685	.858	1.32	1.71	2.07	2.50	2.81
24	127	256	531	.685	.857	1.32	1.71	2.06	2.49	2.80
25	127	256	531	.684	.856	1.32	1.71	2.06	2.48	2.79
26	127	256	531	.684	.856	1.32	1.71	2.06	2.48	2.78
27	127	256	531	.684	.855	1.31	1.70	2.05	2.47	2.77
28	127	256	530	.683	.855	1.31	1.70	2.05	2.47	2.76
29	127	256	530	.683	.854	1.31	1.70	2.04	2.46	2.76
30	127	256	530	.683	.854	1.31	1.70	2.04	2.46	2.75
40	126	255	529	.681	.851	1.30	1.68	2.02	2.42	2.70
60	126	254	527	.679	.848	1.30	1.67	2.00	2.39	2.66
120	126	254	526	.677	.845	1.29	1.66	1.98	2.36	2.62
∞	126	253	524	.674	.842	1.28	1.64	1.96	2.33	2.58

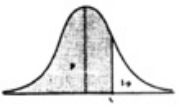
Βαθμοί ελευθερίας
 $N - 1 =$
 $16 - 1 =$
 15

Γούργουλης Βασίλειος, Ph.D.

Δίπλευρος έλεγχος
 $1 - (\alpha/2) =$
 $1 - (0.05/2) =$
 $1 - 0.025 =$
 0.975

49

Η t-κατανομή (Student)



v	t.55	t.60	t.70	t.75	t.80	t.90	t.95	t.975	t.99	t.995
1	158	325	727	1.000	1.376	3.08	6.31	12.71	31.82	63.66
2	142	289	617	.816	1.061	1.89	2.92	4.30	6.96	9.92
3	137	277	584	.765	.978	1.64	2.35	3.18	4.54	5.84
4	134	271	569	.741	.941	1.53	2.13	2.78	3.75	4.60
5	132	267	559	.727	.920	1.48	2.02	2.57	3.36	4.03
6	131	265	553	.718	.906	1.44	1.94	2.45	3.14	3.71
7	130	263	549	.711	.896	1.42	1.90	2.36	3.00	3.50
8	130	262	546	.706	.889	1.40	1.86	2.31	2.90	3.36
9	129	261	543	.703	.883	1.38	1.83	2.26	2.82	3.25
10	129	260	542	.700	.879	1.37	1.81	2.23	2.76	3.17
11	129	260	540	.697	.876	1.36	1.80	2.20	2.72	3.11
12	128	259	539	.695	.873	1.36	1.78	2.18	2.68	3.06
13	128	259	538	.694	.870	1.35	1.77	2.16	2.65	3.01
14	128	258	537	.692	.868	1.34	1.76	2.14	2.62	2.98
15	128	258	536	.691	.866	1.34	1.75	2.13	2.60	2.95
16	128	258	535	.690	.865	1.34	1.75	2.12	2.58	2.92
17	128	257	534	.689	.863	1.33	1.74	2.11	2.57	2.90
18	127	257	534	.688	.862	1.33	1.73	2.10	2.55	2.88
19	127	257	533	.688	.861	1.33	1.73	2.09	2.54	2.86
20	127	257	533	.687	.860	1.32	1.72	2.09	2.53	2.84
21	127	257	532	.686	.859	1.32	1.72	2.08	2.52	2.83
22	127	256	532	.686	.858	1.32	1.72	2.07	2.51	2.82
23	127	256	532	.685	.858	1.32	1.71	2.07	2.50	2.81
24	127	256	531	.685	.857	1.32	1.71	2.06	2.49	2.80
25	127	256	531	.684	.856	1.32	1.71	2.06	2.48	2.79
26	127	256	531	.684	.856	1.32	1.71	2.06	2.48	2.78
27	127	256	531	.684	.855	1.31	1.70	2.05	2.47	2.77
28	127	256	530	.683	.855	1.31	1.70	2.05	2.47	2.76
29	127	256	530	.683	.854	1.31	1.70	2.04	2.46	2.76
30	127	256	530	.683	.854	1.31	1.70	2.04	2.46	2.75
40	126	255	529	.681	.851	1.30	1.68	2.02	2.42	2.70
60	126	254	527	.679	.848	1.30	1.67	2.00	2.39	2.66
120	126	254	526	.677	.845	1.29	1.66	1.98	2.36	2.62
∞	126	253	524	.674	.842	1.28	1.64	1.96	2.33	2.58

Βαθμοί ελευθερίας
 $N - 1 =$
 $16 - 1 =$
 15

Γούργουλης Βασίλειος, Ph.D.

Δίπλευρος έλεγχος
 $1 - (\alpha/2) =$
 $1 - (0.05/2) =$
 $1 - 0.025 =$
 0.975

50

Βαθμοί ελευθερίας
 $N - 1 =$
 $16 - 1 =$
 15

Δίπλευρος έλεγχος
 $1 - (\alpha/2) =$
 $1 - (0.05/2) =$
 $1 - 0.025 =$
 0.975

v	t.55	t.60	t.70	t.75	t.80	t.90	t.95	t.975	t.99	t.995
15	.128	.258	.536	.691	.866	1.34	1.75	2.13	2.60	2.95

51

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Το 95% όλων των τυχαίων δειγμάτων
 θα έχει μέσο όρο που θα κυμαίνεται μεταξύ των τιμών

$$\bar{X}_i + t_{0.05} \cdot s_{\bar{X}} > \bar{X}_o > \bar{X}_i - t_{0.05} \cdot s_{\bar{X}}$$

που καλείται διάστημα εμπιστοσύνης

$$\bar{X}_i + t_{0.05} \cdot s_{\bar{X}} > \mu > \bar{X}_i - t_{0.05} \cdot s_{\bar{X}}$$

$$61.5 + 2.13 \cdot 5 > \mu > 61.5 - 2.13 \cdot 5$$

$$72.15 > \mu > 50.85$$

52

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Συνεπώς
έχοντας ένα δείγμα
μπορούμε (λόγω του Κ.Ο.Θ.) να ελέγξουμε
αν η μέση τιμή του δείγματος
βρίσκεται στο διάστημα εμπιστοσύνης της μέσης τιμής του πληθυσμού
και
να απορρίψουμε ή να αποδεχθούμε
τη μηδενική υπόθεση που έχει οριστεί

53

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Βιβλιογραφία

- **Γούργουλης Β., Μαυρομάτης Γ. (2002).**
Βασικές έννοιες εφαρμοσμένης Στατιστικής στη Φυσική Αγωγή. Εκδόσεις ΣΑΛΤΟ. Θεσσαλονίκη.
Κεφ. 5: Κανονική κατανομή (Σελ. 61 - 69)
Κεφ. 6: Εκτίμηση παραμέτρων και διαστημάτων (Σελ. 71 - 89)
Κεφ. 7: Έλεγχος υποθέσεων (Σελ. 91 - 108)
Πίνακας Z - κανονικής κατανομής (Σελ. 169)
Πίνακας t - κατανομής (Σελ. 170)

54

Γούργουλης Βασίλειος, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. – Δ.Π.Θ

Ερωτήσεις ;



55

Γούργουλης Βασίλης, Ph.D., Καθηγητής Τ.Ε.Φ.Α.Α. – Σ.Ε.Φ.Α.Α. - Α.Π.Θ.