



ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

επίπεδο εμπιστοσύνης, μέγεθος του
δείγματος, κάτω/άνω όριο εμπιστοσύνης

Ορολογία και σημαντικές έννοιες

- **Ποσοτικές** λέγονται οι **μεταβλητές** οι οποίες παίρνουν μόνο αριθμητικές τιμές και μπορούν να ταξινομηθούν σε διακριτές (π.χ. αριθμός παιδιών ανά οικογένεια) ή συνεχείς (π.χ. βάρος).
- Το **πλήθος** των δεδομένων ενός δείγματος που εξετάζονται για την μελέτη ενός φαινομένου ονομάζεται **μέγεθος του δείγματος** και συμβολίζεται με n (n).
- **Τυπική απόκλιση** (σ) είναι ένα μέτρο που χρησιμοποιείται για να υπολογιστεί το ποσό της μεταβολής ή της διασποράς ενός συνόλου τιμών δεδομένων.
- Το α (**άλφα**) είναι το **επίπεδο σημαντικότητας** που χρησιμοποιείται για τον υπολογισμό του **επιπέδου εμπιστοσύνης**. Συνήθως $\alpha = 0,05$

Ορολογία και σημαντικές έννοιες

- Το **επίπεδο εμπιστοσύνης** ισούται με $100 \cdot (1 - \alpha)\%$, οπότε το επίπεδο εμπιστοσύνης είναι συνήθως 95%
- Το **διάστημα εμπιστοσύνης** είναι ένα εύρος τιμών που περιέχει τον πραγματικό Μ.Ο. του πληθυσμού με δηλωμένη πιθανότητα (επίπεδο εμπιστοσύνης, π.χ. 95%).
- **Κάτω όριο εμπιστοσύνης** είναι η ελάχιστη τιμή (βάσει του επιπέδου εμπιστοσύνης)
- **Άνω όρια εμπιστοσύνης** είναι η μέγιστη τιμή (βάσει του επιπέδου εμπιστοσύνης)
- Το **εύρος** του διαστήματος εμπιστοσύνης εξαρτάται από το επίπεδο εμπιστοσύνης και από το μέγεθος του δείγματος
 - Μεγάλο επίπεδο εμπιστοσύνης -> μεγάλο εύρος του διαστήματος εμπιστοσύνης
 - Μικρό μέγεθος του δείγματος -> μεγάλο εύρος του διαστήματος εμπιστοσύνης
 - Γιατί;

Παράδειγμα 1 (έναν πληθυσμός)

- Ερώτημα: ποιος είναι ο πραγματικός M.O. της παραμέτρου του πληθυσμού;
- Απάντηση:
 - εάν δεν μετράμε το κάθε άτομο δεν μπορούμε να ξέρουμε
 - αλλά εάν κάνουμε μία δειγματοληψία μπορούμε να μάθουμε τουλάχιστον το εύρος τιμών στο οποίο βρίσκεται ο πραγματικός M.O. (με δεδομένη πιθανότητα)!

Παράδειγμα 1 (έναν πληθυσμός)

- Μετά από μια τυχαία δειγματοληψία (επιλέγουμε τυχαία κάποια άτομα από το σύνολο του πληθυσμού):
 - υπολογίζουμε το Μ.Ο. του δείγματος (\bar{x})
 - υπολογίζουμε την τιμή εμπιστοσύνης ($z \cdot \frac{\sigma}{\sqrt{n}}$)
 - υπολογίζουμε το διάστημα εμπιστοσύνης (c.i.)

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

$$c.i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

Παράδειγμα 1 (έναν πληθυσμός)

- Ένας αγρότης θέλει να μάθει το βάρος των καρπουζιών στο χωράφι του.
- Θέλει να είναι 95% σίγουρος για το αποτέλεσμα
- Ποια βήματα πρέπει να ακολουθήσει;
- Δειγματοληψία:

Καρπούζια	Βάρος καρπουζιών σε κιλά
1ο άτομο	4.6
2ο άτομο	2.5
3ο άτομο	5.9
4ο άτομο	7.3
5ο άτομο	6.1
6ο άτομο	6.8
7ο άτομο	5.7
8ο άτομο	6.3
9ο άτομο	5.9
10ο άτομο	4.9

Παράδειγμα 1 (έναν πληθυσμός)

- **Δειγματοληψία:**

Καρπούζια	Βάρος καρπουζιών σε κιλά
1ο άτομο	4.6
2ο άτομο	2.5
3ο άτομο	5.9
4ο άτομο	7.3
5ο άτομο	6.1
6ο άτομο	6.8
7ο άτομο	5.7
8ο άτομο	6.3
9ο άτομο	5.9
10ο άτομο	4.9

Μ.Ο.	5.60
Μέγεθος δείγματος	10.00
άλφα	0.05
Τυπική απόκλιση (σ)	1.35
Τιμή εμπιστοσύνης	0.84
Κάτω όριο εμπιστοσύνης	4.76
Άνω όριο εμπιστοσύνης	6.44

Ο αγρότης μπορεί να είναι 95% σίγουρος ότι ο Μ.Ο. όλων των καρπουζιών του χωραφιού του είναι μεταξύ 4.76 και 6.44 κιλά.

Άσκηση 1 (έναν πληθυσμός)

- Ένα ψητοπωλείο θέλει να εκτιμήσει τον χρόνο παράδοσης ενός διανομέα.
- Θέλουν να είναι 90% σίγουροι για το αποτέλεσμα
- Ποια βήματα πρέπει να ακολουθήσουν;
- Δειγματοληψία:

Δείγματα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Άσκηση 1 (έναν πληθυσμός)

- Δειγματοληψία:

$$c.i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Δείγματα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Μ.Ο.	
Μέγεθος δείγματος	
άλφα	
Τυπική απόκλιση (σ)	
Τιμή εμπιστοσύνης	
Κάτω όριο εμπιστοσύνης	
Άνω όριο εμπιστοσύνης	

$$\bar{x} = \frac{\sum x_i}{n}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Άσκηση 1 (έναν πληθυσμός)

- **Δειγματοληψία:**

Δείγματα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Μ.Ο.	54.88
Μέγεθος δείγματος	8
άλφα	0.1
Τυπική απόκλιση (σ)	28.23
Τιμή εμπιστοσύνης	16.42
Κάτω όριο εμπιστοσύνης	38.46
Άνω όριο εμπιστοσύνης	71.29

Μπορούν να είναι 90% σίγουροι ότι ο Μ.Ο. όλων των παραδόσεων του διανομέα να είναι μεταξύ 38.46 και 71.29 λεπτών.

Παράδειγμα 2 (δύο πληθυσμοί)

- Ερώτημα 1: ποιοι είναι οι πραγματικοί Μ.Ο. της παραμέτρου των πληθυσμών;
- Ερώτημα 2: διαφέρουν στατιστικά σημαντικά οι πραγματικοί Μ.Ο. των πληθυσμών;
- Απάντηση:
 - εάν δεν μετράμε το κάθε άτομο δεν μπορούμε να ξέρουμε
 - αλλά εάν κάνουμε μια δειγματοληψία στο κάθε πληθυσμό μπορούμε να μάθουμε τουλάχιστον τα εύρη τιμών στα οποία βρίσκονται οι πραγματικοί Μ.Ο. (με δεδομένη πιθανότητα)!
 - εάν τα διαστήματα εμπιστοσύνης των πληθυσμών δεν επικαλύπτονται τότε διαφέρουν στατιστικά σημαντικά

Παράδειγμα 2 (δύο πληθυσμοί)

- Μετά από δυο τυχαίες δειγματοληψίες (επιλέγουμε τυχαία κάποια άτομα από το σύνολο του πληθυσμού):
 - υπολογίζουμε το Μ.Ο. του κάθε δείγματος (\bar{x})
 - υπολογίζουμε τις τιμές εμπιστοσύνης ($z \cdot \frac{\sigma}{\sqrt{n}}$)
 - υπολογίζουμε τα διαστήματα εμπιστοσύνης (c.i.)
 - Ελέγχουμε εάν επικαλύπτονται τα διαστήματα εμπιστοσύνης

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

$$c. i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

Παράδειγμα 2 (δύο πληθυσμοί)

- Ένας αγρότης θέλει να μάθει τα βάρη δύο ποικιλιών καρπουζιών στο χωράφι του.
- Θέλει να μάθει ποια ποικιλία έχει την καλύτερη απόδοση.
- Θέλει να είναι 95% σίγουρος για το αποτέλεσμα
- Ποια βήματα πρέπει να ακολουθήσει;
- Δειγματοληψίες:

Παράδειγμα 2 (δύο πληθυσμοί)

◦ Δειγματοληψίες:

$$c.i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} = \frac{\sum x_i}{n}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

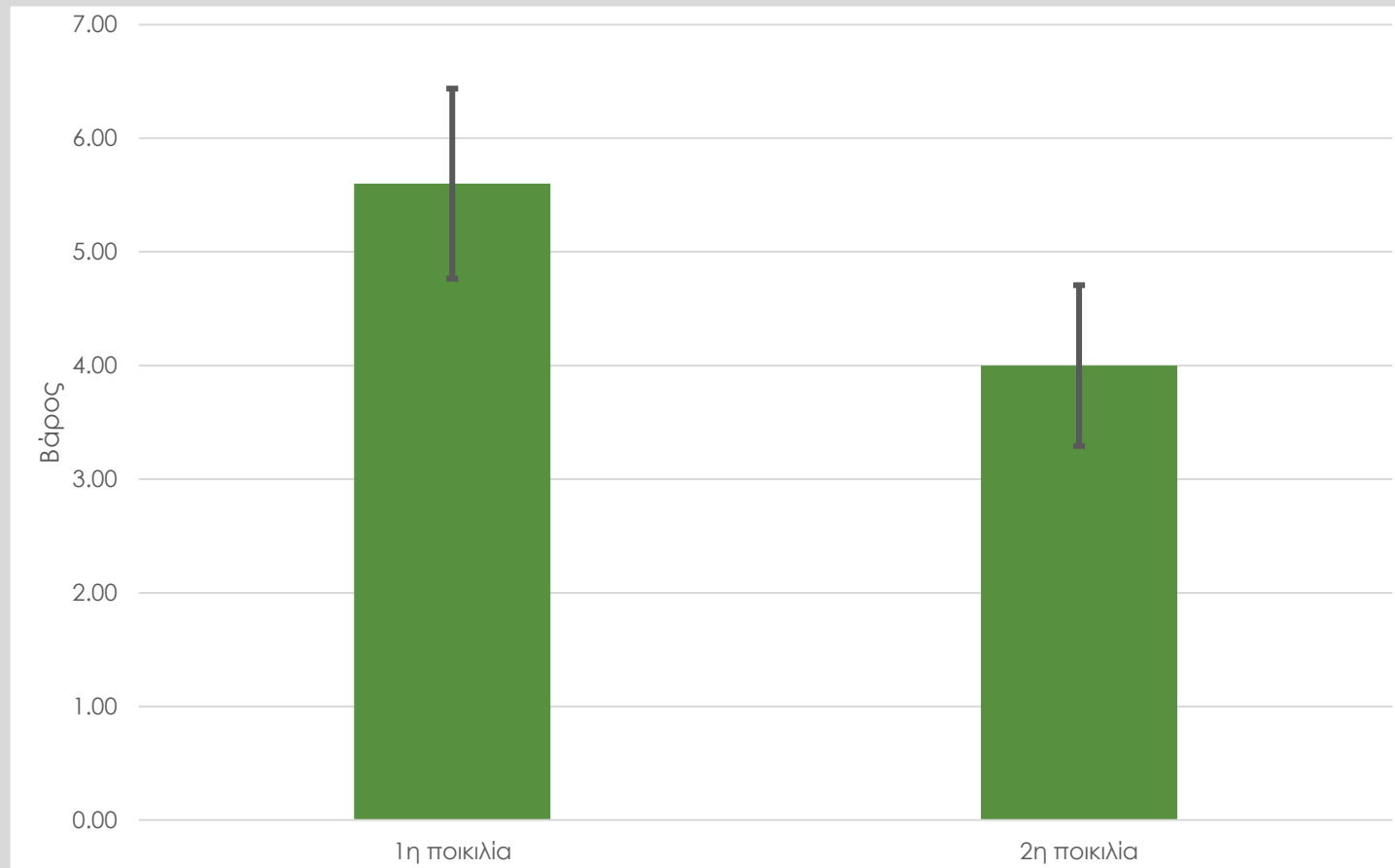
Καρπούζια 1ης ποικιλίας	Βάρος καρπουζιών σε κιλά
1ο άτομο	4.6
2ο άτομο	2.5
3ο άτομο	5.9
4ο άτομο	7.3
5ο άτομο	6.1
6ο άτομο	6.8
7ο άτομο	5.7
8ο άτομο	6.3
9ο άτομο	5.9
10ο άτομο	4.9

Καρπούζια 2ης ποικιλίας	Βάρος καρπουζιών σε κιλά
1ο άτομο	3.6
2ο άτομο	1.1
3ο άτομο	4.9
4ο άτομο	4.4
5ο άτομο	3.8
6ο άτομο	4.1
7ο άτομο	4.5
8ο άτομο	5.3
9ο άτομο	4.4
10ο άτομο	3.9

Μ.Ο.	5.6
Μέγεθος δείγματος	10
άλφα	0.05
Τυπική απόκλιση (σ)	1.35
Τιμή εμπιστοσύνης	0.84
Κάτω όριο εμπιστοσύνης	4.76
Άνω όριο εμπιστοσύνης	6.44

Μ.Ο.	4
Μέγεθος δείγματος	10
άλφα	0.05
Τυπική απόκλιση (σ)	1.14
Τιμή εμπιστοσύνης	0.71
Κάτω όριο εμπιστοσύνης	3.29
Άνω όριο εμπιστοσύνης	4.71

Ο αγρότης μπορεί να είναι 95% σίγουρος ότι η πρώτη ποικιλία είναι καλύτερη από τη δεύτερη.



Άσκηση 2 (δύο πληθυσμοί)

- Ένα ψητοπωλείο θέλει να συγκρίνει τους χρόνους παράδοσης δύο διανομέων.
- Θέλουν να είναι 90% σίγουροι για το αποτέλεσμα
- Ποια βήματα πρέπει να ακολουθήσουν;
- Δειγματοληψίες:

Άσκηση 2 (δύο πληθυσμοί)

- Δειγματοληψίες:

Δείγματα πρώτου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Μ.Ο.	
Μέγεθος δείγματος	
άλφα	
Τυπική απόκλιση (σ)	
Τιμή εμπιστοσύνης	
Κάτω όριο εμπιστοσύνης	
Άνω όριο εμπιστοσύνης	

Δείγματα δεύτερου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	40
2ο	43
3ο	19
4ο	32
5ο	36
6ο	22
7ο	17
8ο	42

Μ.Ο.	
Μέγεθος δείγματος	
άλφα	
Τυπική απόκλιση (σ)	
Τιμή εμπιστοσύνης	
Κάτω όριο εμπιστοσύνης	
Άνω όριο εμπιστοσύνης	

$$\bar{x} = \frac{\sum x_i}{n} \quad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} \quad c.i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

ο Δειγματοληψίες:

Δείγματα πρώτου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Μ.Ο.	
Μέγεθος δείγματος	
άλφα	
Τυπική απόκλιση (σ)	
Τιμή εμπιστοσύνης	
Κάτω όριο εμπιστοσύνης	
Άνω όριο εμπιστοσύνης	

Δείγματα δεύτερου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	40
2ο	43
3ο	19
4ο	32
5ο	36
6ο	22
7ο	17
8ο	42

Μ.Ο.	
Μέγεθος δείγματος	
άλφα	
Τυπική απόκλιση (σ)	
Τιμή εμπιστοσύνης	
Κάτω όριο εμπιστοσύνης	
Άνω όριο εμπιστοσύνης	

$$\bar{x} = \frac{\sum x_i}{n} \quad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} \quad c.i. = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Percentage Confidence	z*-Value
80	1.28
90	1.645
95	1.96
98	2.33
99	2.58

Μπορούν να είναι 90% σίγουροι ότι οι Μ.Ο. χρόνου παράδοσης των διανομέων διαφέρουν.

ο **Δειγματοληψίες:**

Δείγματα πρώτου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	55
2ο	43
3ο	69
4ο	37
5ο	36
6ο	38
7ο	119
8ο	42

Μ.Ο.	54.88
Μέγεθος δείγματος	8
άλφα	0.10
Τυπική απόκλιση (σ)	28.23
Τιμή εμπιστοσύνης	16.42
Κάτω όριο εμπιστοσύνης	38.46
Άνω όριο εμπιστοσύνης	71.29

Δείγματα δεύτερου διανομέα	Χρόνος παράδοσης σε λεπτά
1ο	40
2ο	43
3ο	19
4ο	32
5ο	36
6ο	22
7ο	17
8ο	42

Μ.Ο.	31.38
Μέγεθος δείγματος	8
άλφα	0.10
Τυπική απόκλιση (σ)	10.64
Τιμή εμπιστοσύνης	6.19
Κάτω όριο εμπιστοσύνης	25.19
Άνω όριο εμπιστοσύνης	37.56

