

Επανάληψη - 1

Άσκηση Α

Μια μελέτη με ερωτηματολόγια, διερεύνησε την επαγγελματική ικανοποίηση σε απλό τυχαίο δείγμα $n=300$ κοινωνικών λειτουργών που εργάζονται στην Αθήνα.

Η επαγγελματική ικανοποίηση καθορίστηκε από το σκορ των ατόμων (**δυνατό εύρος από 0 μέχρι 50**). – **ΠΟΣΟΤΙΚΟ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ**

Οι δειγματικές στατιστικές για το σκορ επαγγελματικής ικανοποίησης υπολογίστηκαν ως εξής

μέση τιμή=16 -- τυπική απόκλιση=4 ---- Διάμεσος = 14

Ενδοτεταρτομοριακό εύρος: από 12,5 έως 21

Απαντήστε σύντομα στις ερωτήσεις:

μέση τιμή=16 -- τυπική απόκλιση=4 ---- Διάμεσος = 14

Ενδοτεταρτομοριακό εύρος: από 12,5 έως 21

α) Τι πληροφορία δίνει η τιμή 12,5 (και τι πληροφορία η τιμή 21);

- Το 25% των κοινωνικών λειτουργών είχε σκορ επαγγελματικής ικανοποίησης μικρότερο από 12,5 – είναι το Q1
- Το 75% των κοινωνικών λειτουργών είχε σκορ επαγγελματικής ικανοποίησης μικρότερο από 21 – είναι το Q3

(ισοδύναμα μπορούμε να πούμε «Το 25% των κοινωνικών λειτουργών είχε σκορ επαγγελματικής ικανοποίησης μεγαλύτερο από 21»)

μέση τιμή=16 -- τυπική απόκλιση=4 ---- Διάμεσος = 14

Ενδοτεταρτομοριακό εύρος: από 12,5 έως 21

β) Είναι η κατανομή του σκορ επαγγελματικής ικανοποίησης συμμετρική; Δικαιολογήστε την απάντησή σας. Μπορείτε να ζωγραφίσετε ένα θηκόγραμμα;

Η μέση τιμή 16 είναι πιο μεγάλη από τη διάμεσο, συνεπώς έχουμε θετική ασυμμετρία

Όχι!! Λείπει η μικρότερη και η μεγαλύτερη τιμή. Αν και γνωρίζουμε ότι η κλίμακα είναι από το 0 μέχρι το 50, δεν ξέρουμε ποιες τιμές (μικρότερη-μεγαλύτερη) βρήκαμε στην πραγματικότητα στο δείγμα μας.

μέση τιμή=16 -- τυπική απόκλιση=4 ---- Διάμεσος = 14

Ενδοτεταρτομοριακό εύρος: από 12,5 έως 21

γ) Να βρείτε την διακύμανση

Διακύμανση = το τετράγωνο της τυπικής απόκλισης = $4^2=16$

μέση τιμή =16 -- τυπική απόκλιση=4 ---- Διάμεσος = 14

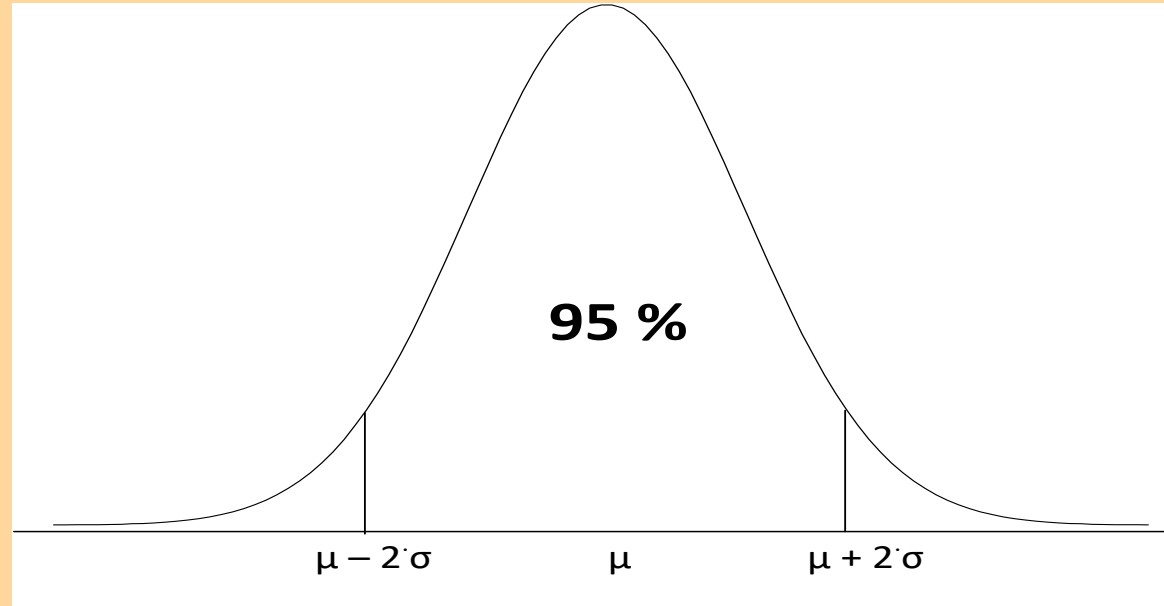
Ενδοτεταρτομοριακό εύρος: από 12,5 έως 21

δ) να υπολογίσετε τον συντελεστή μεταβλητότητας (CV)

$$CV = 4/16 * 100 =$$

$$0,25 * 100 = 25\%$$

ΘΥΜΙΖΟΥΜΕ: Η σημασία της Τυπικής Απόκλισης (σ) – περίπτωση 2



Σε απόσταση **$\Delta\mathbf{ΥΟ}$** τυπικών αποκλίσεων από τη μέση τιμή, υπάρχει περίπου το 95% των περιπτώσεων **μέση τιμή = μ ή $\bar{X}$ - τυπική απόκλιση (σ ή S)**

μέση τιμή ($\bar{X}$)=16 -- τυπική απόκλιση (S) =4

ε) Ανάμεσα σε ποιες τιμές σκορ, βρίσκεται το 95% των κοινωνικών λειτουργών που μετείχαν στην έρευνα;

Είναι γνωστό ότι μεταξύ $\bar{X} \pm 2 * S$ βρίσκεται το 95%, δηλαδή
 $16 \pm 2 * 4 = 16 \pm 8$ δηλαδή

ανάμεσα στο 8 και στο 24

95% Διάστημα Εμπιστοσύνης για τη μέση τιμή (μ)
--- προσεγγιστικό, για μεγάλα δείγματα ---

$$\bar{X} \pm 2 \cdot SE(\bar{X})$$

$$\left(\bar{X} - 2 \cdot SE(\bar{X}), \bar{X} + 2 \cdot SE(\bar{X}) \right)$$

Με λόγια (αν δεν σας αρέσουν τα ...μαθηματικά)

Στη μέση τιμή προσθέτουμε 2 φορές το τυπικό σφάλμα και βγάζουμε το πάνω άκρο,
Στη μέση τιμή αφαιρούμε 2 φορές το τυπικό σφάλμα και βρίσκουμε το κάτω άκρο

μέση τιμή ($\bar{X}$)=16 -- τυπική απόκλιση (S) =4 ---- n=300

στ) να βρείτε το 95% Διάστημα εμπιστοσύνης για τη μέση τιμή του σκόρ ικανοποίησης των κοινωνικών λειτουργών.

Το Δ.Ε. είναι $\bar{X} \pm 2 * SE$, οπότε πρέπει να βρούμε το SE (τυπικό σφάλμα)

$$SE = \frac{S}{\sqrt{n}} \text{ οπότε είναι } SE = \frac{4}{\sqrt{300}} = \frac{4}{17,32} = \mathbf{0,23}$$

Άρα το 95% Δ.Ε. για τη μέση τιμή είναι

$$16 \pm 2 * 0,23 = 16 \pm 0,46$$

Δηλαδή από 15,54 μέχρι 16,46

Πότε χρησιμοποιείται η τυπική απόκλιση (S) και πότε το τυπικό σφάλμα (SE);;

Ανάμεσα σε ποιες τιμές σκορ, βρίσκεται το 95% των κοινωνικών λειτουργών που μετείχαν στην έρευνα;

- *ανάμεσα στο 8 και στο 24* $(\bar{X} \pm 2 * S)$ – τυπική απόκλιση

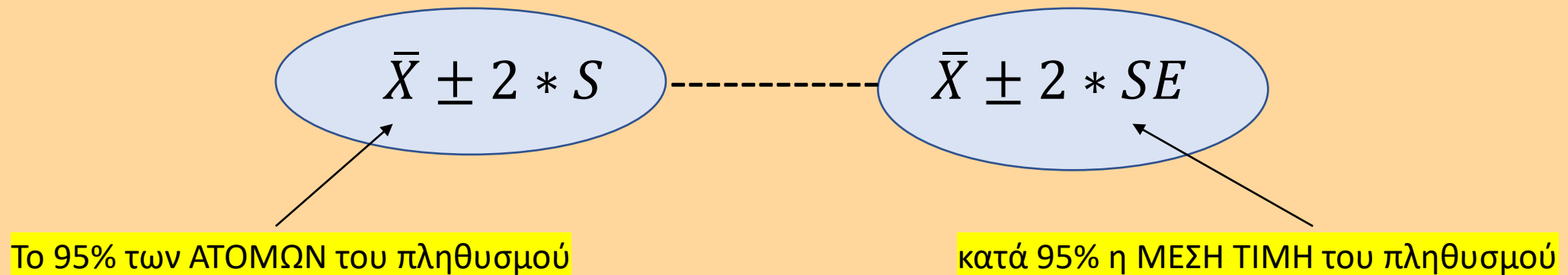
Το 95% Διάστημα εμπιστοσύνης για τη μέση τιμή του σκόρ ικανοποίησης των κοινωνικών λειτουργών

- *ανάμεσα στο 15,54 και στο 16,46* $(\bar{X} \pm 2 * SE)$ – τυπικό σφάλμα

Έχουμε 95% εμπιστοσύνη ότι είναι η μέση τιμή της ικανοποίησης θα βρίσκεται ανάμεσα στις τιμές αυτές.

Άρα:

- Όταν θέλουμε να βρούμε που κυμαίνεται το **95% των ατόμων μιας έρευνας**, χρησιμοποιούμε τον τύπο με την τυπική απόκλιση (S).
- Όταν θέλουμε να δούμε με βεβαιότητα **95% που βρίσκεται η μέση τιμή του πληθυσμού μας**, χρησιμοποιούμε τον τύπο με το τυπικό σφάλμα (SE)



Άσκηση Β

- Μελετώντας 225 άτομα σε ένα ερωτηματολόγιο για την κατανάλωση αλκοόλ, χρησιμοποιήθηκε μια ποσοτική κλίμακα με τιμές από το 0 (καθόλου κατανάλωση), μέχρι το 45 (υπερβολική κατανάλωση), και ο ερευνητής ανέφερε:

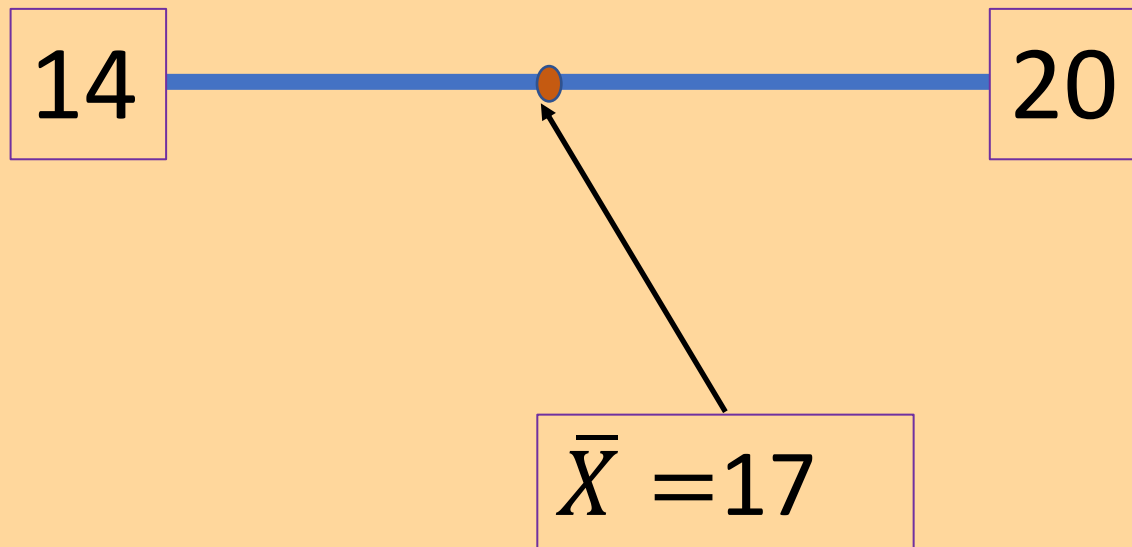
«Είμαστε 95% βέβαιοι ότι η μέση τιμή στον πληθυσμό που μελετήσαμε θα είναι ανάμεσα στην τιμή 14 και στην τιμή 20.»

1. ΠΟΣΗ ΕΙΝΑΙ Η ΜΕΣΗ ΤΙΜΗ ΤΟΥ ΔΕΙΓΜΑΤΟΣ ΠΟΥ ΥΠΟΛΟΓΙΣΕ;
2. ΠΟΣΟ ΗΤΑΝ ΤΟ ΤΥΠΙΚΟ ΣΦΑΛΜΑ ΤΗΣ ΜΕΣΗΣ ΤΙΜΗΣ;
3. ΜΠΟΡΕΙΤΕ ΝΑ ΒΡΕΙΤΕ ΚΑΙ ΤΗΝ ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ;

ΠΟΣΗ ΕΊΝΑΙ Η ΜΕΣΗ ΤΙΜΗ ΤΟΥ ΔΕΙΓΜΑΤΟΣ ΠΟΥ ΥΠΟΛΟΓΙΣΕ;

- Με αυτό που ανέφερε ο ερευνητής, καταλαβαίνουμε ότι υπολόγισε το 95% Διάστημα Εμπιστοσύνης για τη μέση τιμή,
- Γνωρίζουμε από τον τύπο, πως το διάστημα είναι συμμετρικό ως προς τη μέση τιμή, αφού έχει τη μορφή $\bar{X} \pm 2 * SE$
- Άρα η μέση τιμή ($\bar{X}$) που βρήκε είναι στη ΜΕΣΗ
- στο μέσο του διαστήματος από το **14** μέχρι το **20** που βρήκε
- Άρα η μέση τιμή που βρήκε είναι το **17**

$$(14+20)/2 = 34/2 = 17$$



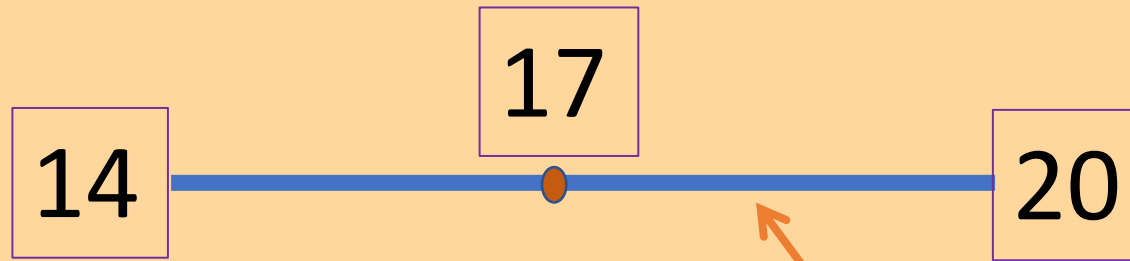
ΠΟΣΟ ΗΤΑΝ ΤΟ ΤΥΠΙΚΟ ΣΦΑΛΜΑ ΤΗΣ ΜΕΣΗΣ ΤΙΜΗΣ;

- Ζητάμε το SE
- Το διάστημα δημιουργείται αφού στη μέση τιμή που βρήκαμε, προσθέσουμε 2 φορές το SE και αφαιρέσουμε 2 φορές το SE.
- Για να φτάσουμε δηλαδή στο πάνω όριο, που είναι το 20, πρέπει να υπολογίσουμε το $\bar{X} + 2 * SE$

Δηλαδή $20 = 17 + 2 * SE$

Άρα $2 * SE = 3$

$SE=1,5$



Αυτό το κομμάτι
είναι 2 φορές το SE

Αφού αυτό το κομμάτι είναι 3 μονάδες,
άρα το SE είναι 1,5 μονάδα

ΜΠΟΡΕΙΤΕ ΝΑ ΒΡΕΙΤΕ ΚΑΙ ΤΗΝ ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ;

Αφού

$$SE = \frac{S}{\sqrt{n}}$$

Βρήκαμε $SE=1,5$

Επίσης $\sqrt{n} = \sqrt{225} = 15$

Οπότε $1,5 = S/15$

τελικά $S = 1,5 * 15 = \mathbf{22,5}$

ΑΣΚΗΣΗ Γ

Πήραμε ένα δείγμα 35 ατόμων και καταγράψαμε το αποτέλεσμα σε ένα τεστ γνώσης Πρώτων Βοηθειών (κλίμακα από 0 μέχρι 100)

65	66	57	80	55	64	90	80	82	40
62	98	64	80	42	75	55	60	58	58
49	88	90	91	50	72	45	75	66	70
45	85	77	63	71					

Βρήκαμε μέση τιμή = 67,66

Τυπική απόκλιση = 15,3

Για να δούμε πόσο ποσοστό από τις παραπάνω μετρήσεις βρίσκονται σε απόσταση ΜΙΑ τυπική απόκλιση από τη μέση τιμή. Είναι 68% όπως μας λέει η Θεωρία;

1. Βρίσκουμε $66,67 - 15,3 = 52,36$
2. Βρίσκουμε $66,67 + 15,3 = 82,96$

Μεταξύ του 53 και του 82, πόσες μετρήσεις έχουμε;;

Μεταξύ του 52 και του 83, πόσες μετρήσεις έχουμε;;

65	66	57	80	55	64	90	80	82	40
62	98	64	80	42	75	55	60	58	58
49	88	90	91	50	72	45	75	66	70
45	85	77	63	71					

Είναι 23 μετρήσεις στις 35, δηλαδή σε ποσοστό $23/35 * 100 = 65,7\%$

Δεν είναι ακριβώς 68% που περιμέναμε αλλά είναι αρκετά κοντά.

ΑΣΚΗΣΗ Δ (Παράδειγμα για τυπική κανονική κατανομή Z σκορ – τιμές Z)

Οι τιμές που παίρνουμε σε ένα τεστ, μπορούν να μετασχηματιστούν σε τιμές Z, ως εξής

1. Αφαιρούμε τη μέση τιμή και το αποτέλεσμα το
2. Διαιρούμε με την τυπική απόκλιση

Π.χ. το πρώτο άτομο στο τεστ που πήρε βαθμολογία **65**, έχει τιμή Z ίση με $(65-67,66) / 15,3 = -2,66/15,3 = \mathbf{-0,17 \text{ (αρνητικό)}}$

Ενώ το τελευταίο άτομο που πήρε **71**, έχει τιμή ίση με $(71-67,66)/15,3 = 3,34/15,3 = \mathbf{0,22 \text{ (θετικό)}}$

ΑΣΚΗΣΗ Ε

Σε μια κανονική κατανομή βρέθηκε η μέση τιμή ίση με 40, το $Q1=25$ και το $Q3=45$. Μήπως έγινε κάποιο λάθος; Αιτιολογήστε.

ΝΑΙ

Κανονική κατανομή = συμμετρία = μέση τιμή και διάμεσος ΙΔΙΕΣ

Αρα θα έπρεπε Διάμεσος = 40

Δηλαδή **$Q2 = 40$**

Θα έπρεπε το $Q2$ να είναι ακριβώς στη μεση του $Q1 - Q3$

ΑΣΚΗΣΗ ΣΤ

- Ποιο από τα παρακάτω δύο δείγματα παρουσίασε μεγαλύτερη μεταβλητότητα;

δείγμα A ---- μέση τιμή = 70 --- τυπική απόκλιση = 7

δείγμα B ---- μέση τιμή = 40 --- τυπική απόκλιση = 4

α) το A β) το B γ) παρουσίασαν ίδια μεταβλητότητα

δ) χρειαζόμαστε περισσότερα στοιχεία για να απαντήσουμε

$$CV_A = 7/70 * 100 = 10\%, \quad CV_B = 4/40 * 100 = 10\% \quad -- \text{ ΙΔΙΑ (γ)}$$

ΑΣΚΗΣΗ Ζ

Αν το 95% διάστημα εμπιστοσύνης για το ποσοστό των καπνιστών σε ένα δείγμα ήταν από 30% μέχρι 36%, πόσο ήταν το ποσοστό των καπνιστών στο δείγμα, που βρήκε ο ερευνητής;

