

## 5ο Φροντιστήριο – Στατιστική Ι

**1. 2η Εργασία**

**2. Επαναληπτικό Προόδου:**

**Ερωτήσεις πολλαπλής Επιλογής-Άσκηση ανάπτυξης**

## 2η Εργασία

Παράδοση μέχρι το Σάββατο 29/3/25

Ποια είναι η σημασία των περιγραφικών στατιστικών στην ανάλυση δεδομένων και ποια στοιχεία από τα περιγραφικά στατιστικά μπορούν να μας βοηθήσουν να κατανοήσουμε την κατανομή των δεδομένων σε μια έρευνα; Συζητήστε με παραδείγματα.

## 2η Εργασία

Παράδοση μέχρι το Σάββατο 29/3/25

Για να απαντήσουμε σκεφτόμαστε:

- Από τι αποτελείται η περιγραφική στατιστική;
    - Κατανομή συχνοτήτων
    - Μέτρα κεντρικής τάσης (μέσος όρος, διάμεσος, επικρατούσα τιμή)
    - Μέτρα διασποράς (εύρος, ενδοτεταρτ.εύρος, τυπική απόκλιση)
- Αναπαράσταση (ολπτικοποίηση), κατανόηση, περιγραφή...

# Ερωτήσεις πολλαπλής επιλογής

**1. Ποια είναι η βασική διαφορά μεταξύ της περιγραφικής και της επαγωγικής στατιστικής;**

- α) Η περιγραφική στατιστική χρησιμοποιεί υποθέσεις, ενώ η επαγωγική τα δεδομένα.
- β) Η περιγραφική στατιστική αναλύει τα δεδομένα του δείγματος, ενώ η επαγωγική κάνει συμπεράσματα για τον πληθυσμό.
- γ) Η περιγραφική στατιστική αφορά μόνο τις ποιοτικές μεταβλητές.
- δ) Η επαγωγική στατιστική δεν επιτρέπει τη γενίκευση των αποτελεσμάτων.

# Ερωτήσεις πολλαπλής επιλογής

**1. Ποια είναι η βασική διαφορά μεταξύ της περιγραφικής και της επαγωγικής στατιστικής;**

- α) Η περιγραφική στατιστική χρησιμοποιεί υποθέσεις, ενώ η επαγωγική τα δεδομένα.
- β) Η περιγραφική στατιστική αναλύει τα δεδομένα του δείγματος, ενώ η επαγωγική κάνει συμπεράσματα για τον πληθυσμό.
- γ) Η περιγραφική στατιστική αφορά μόνο τις ποιοτικές μεταβλητές.
- δ) Η επαγωγική στατιστική δεν επιτρέπει τη γενίκευση των αποτελεσμάτων.

**Απάντηση: β)**

**2. Ποια από τις παρακάτω είναι ποιοτική μεταβλητή.**

α) Ύψος

β) Χρόνος

γ) Κατηγορία επαγγέλματος

δ) Βάρος

**2. Ποια από τις παρακάτω είναι ποιοτική μεταβλητή.**

α) Ύψος

β) Χρόνος

γ) Κατηγορία επαγγέλματος

δ) Βάρος

**Απάντηση:** γ) Κατηγορία επαγγέλματος

**3. Τί μεταβλητή είναι ο αριθμός των παιδιών σε μια οικογένεια;**

α) Συνεχής

β) Ασυνεχής

γ) Ιεραρχική

δ) Κατηγορική

**3. Τί μεταβλητή είναι ο αριθμός των παιδιών σε μια οικογένεια;**

α) Συνεχής

β) Ασυνεχής

γ) Ιεραρχική

δ) Κατηγορική

**Απάντηση:** β) Ασυνεχής

#### **4. Ποια είναι η διαφορά μεταξύ συνεχών και ασυνεχών μεταβλητών;**

- α) Οι ασυνεχείς παίρνουν μόνο ακέραιες τιμές, ενώ οι συνεχείς οποιαδήποτε τιμή.
- β) Οι συνεχείς μεταβλητές παίρνουν οποιαδήποτε τιμή, ενώ οι ασυνεχείς μόνο συγκεκριμένες.
- γ) Οι ασυνεχείς μεταβλητές είναι πιο εύκολες να μετρηθούν.
- δ) Οι συνεχείς μεταβλητές είναι κατηγορικές.

#### **4. Ποια είναι η διαφορά μεταξύ συνεχών και ασυνεχών μεταβλητών;**

- α) Οι ασυνεχείς παίρνουν μόνο ακέραιες τιμές, ενώ οι συνεχείς οποιαδήποτε τιμή.
- β) Οι συνεχείς μεταβλητές παίρνουν οποιαδήποτε τιμή, ενώ οι ασυνεχείς μόνο συγκεκριμένες.
- γ) Οι ασυνεχείς μεταβλητές είναι πιο εύκολες να μετρηθούν.
- δ) Οι συνεχείς μεταβλητές είναι κατηγορικές.

**Απάντηση:** α)

**5. Ποια από τις παρακάτω είναι ποιοτική μεταβλητή.**

α) Βάρος

β) Χρόνος ύπνου

γ) Αριθμός παιδιών σε οικογένεια

δ) Κατηγορία επαγγέλματος

**5. Ποια από τις παρακάτω είναι ποιοτική μεταβλητή.**

α) Βάρος

β) Χρόνος ύπνου

γ) Αριθμός παιδιών σε οικογένεια

δ) Κατηγορία επαγγέλματος

**Απάντηση:** δ) Κατηγορία επαγγέλματος

**6. Ποιο από τα εξής είναι παράδειγμα περιγραφικής στατιστικής;**

α) Υπολογισμός του μέσου όρου της ηλικίας του δείγματος.

β) Έλεγχος υποθέσεων για το μέσο όρο του πληθυσμού.

γ) Εκτίμηση των παραμέτρων του πληθυσμού.

δ) Δημιουργία υποθέσεων για τη σχέση μεταβλητών.

**6. Ποιο από τα εξής είναι παράδειγμα περιγραφικής στατιστικής;**

- α) Υπολογισμός του μέσου όρου της ηλικίας του δείγματος.
- β) Έλεγχος υποθέσεων για το μέσο όρο του πληθυσμού.
- γ) Εκτίμηση των παραμέτρων του πληθυσμού.
- δ) Δημιουργία υποθέσεων για τη σχέση μεταβλητών.

**Απάντηση: α)**

**7. Ποιο από τα εξής είναι παράδειγμα περιγραφικής στατιστικής;**

- α) Δημιουργία υποθέσεων για τη σχέση μεταβλητών.
- β) Έλεγχος υποθέσεων για το μέσο όρο του πληθυσμού.
- γ) Εκτίμηση των παραμέτρων του πληθυσμού.
- δ) Υπολογισμός του εύρους των τιμών.

**7. Ποιο από τα εξής είναι παράδειγμα περιγραφικής στατιστικής;**

- α) Δημιουργία υποθέσεων για τη σχέση μεταβλητών.
- β) Έλεγχος υποθέσεων για το μέσο όρο του πληθυσμού.
- γ) Εκτίμηση των παραμέτρων του πληθυσμού.
- δ) Υπολογισμός του εύρους των τιμών.

**Απάντηση: δ)**

**8. Ποιο από τα παρακάτω είναι ένα παράδειγμα ακολουθίας στο R;**

α) `c(1, 2, 3, 4)`

β) `matrix(1, 2, 3)`

γ) `data.frame(1, 2, 3)`

δ) `list(1, 2, 3)`

**8. Ποιο από τα παρακάτω είναι ένα παράδειγμα ακολουθίας στο R;**

α) `c(1, 2, 3, 4)`

β) `matrix(1, 2, 3)`

γ) `data.frame(1, 2, 3)`

δ) `list(1, 2, 3)`

**Απάντηση: α)**

**9. Ποιο από τα παρακάτω χαρακτηρίζει αντικείμενο δεδομένων στο R;**

α) list

β) data.frame

γ) matrix

δ) Όλα τα παραπάνω

**9. Ποιο από τα παρακάτω χαρακτηρίζει αντικείμενο δεδομένων στο R;**

α) list

β) data.frame

γ) matrix

δ) Όλα τα παραπάνω

**Απάντηση:** δ) Όλα τα παραπάνω

## 10. Τι είναι το `str()` στο R;

- α) Εμφανίζει την πρώτη γραμμή του `data.frame`
- β) Εμφανίζει τη δομή ενός αντικειμένου
- γ) Εμφανίζει τα στοιχεία ενός διανύσματος
- δ) Υπολογίζει το άθροισμα των στοιχείων

**10. Τι είναι το `str()` στο R;**

- α) Εμφανίζει την πρώτη γραμμή του `data.frame`
- β) Εμφανίζει τη δομή ενός αντικειμένου
- γ) Εμφανίζει τα στοιχεία ενός διανύσματος
- δ) Υπολογίζει το άθροισμα των στοιχείων

**Απάντηση: β)**

**11. Ποια από τις παρακάτω εντολές δημιουργεί ένα πλαίσιο δεδομένων;**

α) `data.frame()`

β) `matrix()`

γ) `vector()`

δ) `list()`

**11. Ποια από τις παρακάτω εντολές δημιουργεί ένα πλαίσιο δεδομένων;**

α) `data.frame()`

β) `matrix()`

γ) `vector()`

δ) `list()`

**Απάντηση:** α) `data.frame()`

**12. Ποια εντολή δημιουργεί λίστα στο R;**

α) `list()`

β) `vector()`

γ) `matrix()`

δ) `data.frame()`

**12. Ποια εντολή δημιουργεί λίστα στο R;**

α) list()

β) vector()

γ) matrix()

δ) data.frame()

**Απάντηση: α)**

**13. Ποιος είναι ο στόχος του R ως γλώσσα προγραμματισμού;**

α) Η δημιουργία ιστοσελίδων

β) Η ανάλυση δεδομένων και στατιστικής

γ) Η σχεδίαση παιχνιδιών

δ) Η δημιουργία γραφικών παραστάσεων

**13. Ποιος είναι ο στόχος του R ως γλώσσα προγραμματισμού;**

α) Η δημιουργία ιστοσελίδων

β) Η ανάλυση δεδομένων και στατιστικής

γ) Η σχεδίαση παιχνιδιών

δ) Η δημιουργία γραφικών παραστάσεων

**Απάντηση: β)**

**14. Ποια είναι η διαφορά μεταξύ απόλυτης συχνότητας και σχετικής συχνότητας;**

α) Η απόλυτη συχνότητα εκφράζεται σε ποσοστά, ενώ η σχετική συχνότητα σε αριθμούς

β) Η απόλυτη συχνότητα είναι ο αριθμός των εμφανίσεων μιας κατηγορίας, ενώ η σχετική συχνότητα είναι το ποσοστό των εμφανίσεων

γ) Η απόλυτη συχνότητα περιλαμβάνει όλα τα δεδομένα, ενώ η σχετική συχνότητα όχι

δ) Δεν υπάρχει διαφορά

**14. Ποια είναι η διαφορά μεταξύ απόλυτης συχνότητας και σχετικής συχνότητας;**

α) Η απόλυτη συχνότητα εκφράζεται σε ποσοστά, ενώ η σχετική συχνότητα σε αριθμούς

β) Η απόλυτη συχνότητα είναι ο αριθμός των εμφανίσεων μιας κατηγορίας, ενώ η σχετική συχνότητα είναι το ποσοστό των εμφανίσεων

γ) Η απόλυτη συχνότητα περιλαμβάνει όλα τα δεδομένα, ενώ η σχετική συχνότητα όχι

δ) Δεν υπάρχει διαφορά

**Απάντηση: β)**

## **15. Η διάμεσος είναι:**

- α) Η τιμή που εμφανίζεται πιο συχνά σε ένα σύνολο δεδομένων
- β) Η τιμή που χωρίζει τα δεδομένα σε δύο ίσα μέρη
- γ) Η μέση τιμή όλων των δεδομένων
- δ) Η απόσταση μεταξύ του μεγαλύτερου και του μικρότερου δεδομένου

### **15. Η διάμεσος είναι:**

- α) Η τιμή που εμφανίζεται πιο συχνά σε ένα σύνολο δεδομένων
- β) Η τιμή που χωρίζει τα δεδομένα σε δύο ίσα μέρη
- γ) Η μέση τιμή όλων των δεδομένων
- δ) Η απόσταση μεταξύ του μεγαλύτερου και του μικρότερου δεδομένου

**Απάντηση: β)**

**16. Ποιο μέτρο χρησιμοποιείται για να περιγράψει την πιο συχνά εμφανιζόμενη τιμή σε ένα σύνολο δεδομένων;**

α) Η διάμεσος

β) Η επικρατούσα τιμή

γ) Η τυπική απόκλιση

δ) Το εύρος

**16. Ποιο μέτρο χρησιμοποιείται για να περιγράψει την πιο συχνά εμφανιζόμενη τιμή σε ένα σύνολο δεδομένων;**

α) Η διάμεσος

β) Η επικρατούσα τιμή

γ) Η τυπική απόκλιση

δ) Το εύρος

**Απάντηση: β)**

17. Ποιο μέτρο περιγραφικής στατιστικής μας δείχνει πόσο “αραιά” ή “συμπυκνωμένα” είναι τα δεδομένα γύρω από τον μέσο όρο;

- α) Η διακύμανση
- β) Ο μέσος όρος
- γ) Η επικρατούσα τιμή
- δ) Το εύρος

**17. Ποιο μέτρο περιγραφικής στατιστικής μας δείχνει πόσο “αραιά” ή “συμπυκνωμένα” είναι τα δεδομένα γύρω από τον μέσο όρο;**

- α) Η διακύμανση
- β) Ο μέσος όρος
- γ) Η επικρατούσα τιμή
- δ) Το εύρος

**Απάντηση:** α) Η διακύμανση

**18. Ποια είναι η βασική διαφορά μεταξύ της διαμέσου και του μέσου όρου;**

α) Ο μέσος όρος επηρεάζεται από τις ακραίες τιμές, ενώ η διάμεσος όχι

β) Η διάμεσος είναι μεγαλύτερη από τον μέσο όρο

γ) Ο μέσος όρος δεν υπολογίζεται εύκολα

δ) Η διάμεσος επηρεάζεται από τις ακραίες τιμές, ενώ ο μέσος όρος όχι

**18. Ποια είναι η βασική διαφορά μεταξύ της διαμέσου και του μέσου όρου;**

α) Ο μέσος όρος επηρεάζεται από τις ακραίες τιμές, ενώ η διάμεσος όχι

β) Η διάμεσος είναι μεγαλύτερη από τον μέσο όρο

γ) Ο μέσος όρος δεν υπολογίζεται εύκολα

δ) Η διάμεσος επηρεάζεται από τις ακραίες τιμές, ενώ ο μέσος όρος όχι

**Απάντηση: α)**

**19. Ποιο μέτρο χρησιμοποιείται για να προσδιορίσει τη διασπορά μεταξύ των τεταρτημορίων των δεδομένων;**

α) Το εύρος

β) Η διακύμανση

γ) Το ενδοτεταρτημοριακό εύρος

δ) Η τυπική απόκλιση

**19. Ποιο μέτρο χρησιμοποιείται για να προσδιορίσει τη διασπορά μεταξύ των τεταρτημορίων των δεδομένων;**

α) Το εύρος

β) Η διακύμανση

γ) Το ενδοτεταρτημοριακό εύρος

δ) Η τυπική απόκλιση

**Απάντηση: γ)**

**20. Ποιο από τα παρακάτω μέτρα δεν σχετίζεται με τη διάσταση των δεδομένων.**

α) Το εύρος

β) Η διακύμανση

γ) Ο μέσος όρος

δ) Η τυπική απόκλιση

**20. Ποιο από τα παρακάτω μέτρα δεν σχετίζεται με τη διάσταση των δεδομένων.**

α) Το εύρος

β) Η διακύμανση

γ) Ο μέσος όρος

δ) Η τυπική απόκλιση

**Απάντηση:** γ) Ο μέσος όρος

## Άσκηση

Δίνεται η ακόλουθη σειρά έξι αριθμών: 5, 9, 8, 1, 1, 2. Να υπολογισθούν τα μέτρα περιγραφικής στατιστικής: **μέσος όρος, διάμεσος, επικρατούσα τιμή, διασπορά (διακύμανση), εύρος, τυπική απόκλιση, ενδοτεταρτημοριακό εύρος.**

Πριν ξεκινήσουμε τη λύση της άσκησης και για να μην το ξεχάσουμε αργότερα, καλύτερα να βάλουμε από την αρχή την ακολουθία των αριθμών σε αύξουσα σειρά: 1, 1, 2, 5, 8, 9.

**Μέσος όρος:**

$$\text{M.O.} = \frac{1 + 1 + 2 + 5 + 8 + 9}{6} = 4,333$$

**Διάμεσος:** Εφόσον έχουμε άρτιο αριθμό τιμών ( $N=6$ ), η διάμεσος θα είναι ανάμεσα στον 3<sup>ο</sup> και στον 4<sup>ο</sup> όρο, δηλαδή το  $\frac{2+5}{2} = 3,5$ .

**Επικρατούσα τιμή:** Είναι η τιμή που εμφανίζεται περισσότερες φορές, δηλαδή το 1.

**Διασπορά:**

$$s^2 = \frac{2(1 - 4,333)^2 + (2 - 4,333)^2 + (5 - 4,333)^2 + (8 - 4,333)^2 + (9 - 4,333)^2}{6}$$

$$\implies s^2 = \frac{22,2178 + 5,4429 + 0,445 + 13,4469 + 21,78}{6} = 10,5554$$

**Τυπική Απόκλιση:**

$$s = \sqrt{s^2} = \sqrt{10,5554} = 3,2489$$

**Εύρος:**  $9 - 1 = 8$

**Ενδοτεταρτημοριακό εύρος:**

Είναι το διάστημα το οποίο βρίσκεται ανάμεσα στο 25% και στο 75% των παρατηρήσεων, δηλ. ανάμεσα στο  $Q_1$  και στο  $Q_3$ .

Για τον υπολογισμό του  $Q_1$  θέλουμε να βρούμε την τιμή που βρίσκεται στο 1ο τέταρτο των παρατηρήσεων, δηλαδή στο  $\frac{N+1}{4} = \frac{6+1}{4} = 1,75$ . Η τιμή αυτή θα βρίσκεται ανάμεσα στον 1<sup>ο</sup> και στον 2<sup>ο</sup> όρο, σε “ απόσταση ” 0,75 ανάμεσά τους. Οπότε θα είναι ο

$$1^{\circ\varsigma} + 0,75(2^{\circ\varsigma} - 1^{\circ\varsigma}) \quad \text{δηλ.} \quad Q_1 = 1 + 0,75(1 - 1) = 1$$

Για τον υπολογισμό του  $Q_3$  θέλουμε να βρούμε την τιμή που βρίσκεται στο 3ο τέταρτο των παρατηρήσεων, δηλαδή στο  $\frac{3(N+1)}{4} = \frac{3(6+1)}{4} = 5,25$ . Η τιμή αυτή θα βρίσκεται ανάμεσα στον 5<sup>ο</sup> και στον 6<sup>ο</sup> όρο, σε 'απόσταση' 0,25 ανάμεσά τους. Οπότε θα είναι ο

$$5^{ος} + 0,25(6^{ος} - 5^{ος}) \quad \text{δηλ.} \quad Q_3 = 8 + 0,25(9 - 8) = 8,25$$

Συνεπώς το ενδοτεταρτημοριακό εύρος είναι  $8,25 - 1 = 7,25$ .



ΤΕΛΟΣ