

# 7ο Φροντιστήριο – Στατιστική Ι

## Κεφ. 8ο

**Έλεγχος Υποθέσεων-Διαστήματα εμπιστοσύνης**

**Ασκήσεις Κατανόησης**

## Άσκηση 1

Πού οφείλεται το γεγονός ότι ο μέσος όρος κάθε δείγματος διαφέρει από το μέσο όρο του πληθυσμού; Πώς μπορούμε να αντιμετωπίσουμε αυτό το πρόβλημα;

Απάντηση: Κάθε δείγμα που επιλέγουμε έχει τα δικά του χαρακτηριστικά αφού αποτελείται από άτομα. Για να αντιμετωπίσουμε αυτό το πρόβλημα, θα πρέπει αρχικά η επιλογή του δείγματος να γίνεται με αυστηρούς κανόνες δειγματοληψίας (δειγματοληπτική στρατηγική, π.χ. απλή τυχαία δειγμ/ψία), ώστε να αυξήσουμε την πιθανότητα το δείγμα μας να έχει τα ίδια χαρακτηριστικά με τον πληθυσμό. Στη συνέχεια, θα βασιστούμε στο Κεντρικό Οριακό Θεώρημα και θα ακολουθήσουμε τις στατιστικές διαδικασίες που αναφέρονται σε μια δειγματοληπτική κατανομή.

## **Κεντρικό Οριακό Θεώρημα**

Αν επιλέξουμε με τυχαία δειγματοληψία πολλά μεγάλα σε πλήθος και ίσα μεταξύ τους δείγματα από έναν πληθυσμό, η κατανομή των μέσων όρων των δειγμάτων αυτών θα έχει τη μορφή της κανονικής κατανομής.

## Άσκηση 2

Από επιστημονική άποψη, ποιο επίπεδο σημαντικότητας είναι καλύτερο; Γιατί;

Απάντηση: Το καλύτερο επίπεδο σημαντικότητας είναι αυτό που δίνει τις λιγότερες πιθανότητες σφάλματος (π.χ. 0,001). Ένα τέτοιο επίπεδο σημαντικότητας μας δίνει πληροφορίες ότι η πιθανότητα το αποτέλεσμα που πήραμε να είναι στατιστικό λάθος, δηλαδή να προέρχεται από τυχαίους παράγοντες, είναι μόλις 1‰.

## Άσκηση 3

Γιατί η δειγματοληπτική κατανομή είναι πολύ σημαντική έννοια στον έλεγχο υποθέσεων;

Απάντηση: Είναι σημαντική γιατί μας δίνει πληροφορίες για το τι τιμές θα πρέπει (ή δεν θα πρέπει) να περιμένουμε ότι θα βρεθούν από την εφαρμογή ενός συγκεκριμένου στατιστικού κριτηρίου μέσα σε ένα προκαθορισμένο πλαίσιο κανόνων (π.χ. ποιος θα είναι ο μέσος όρος της επίδοσης ενός δείγματος παιδιών ως προς την αριθμητική ικανότητα, όταν ο μέσος όρος του γενικού πληθυσμού είναι 45).

## Άσκηση 4

Εντοπίστε ποιες από τις παρακάτω υποθέσεις αντανακλούν έναν αμφίπλευρο και ποιες ένα μονόπλευρο έλεγχο:

- i. Οι διαβητικοί ακολουθούν πιο υγιεινή διατροφή από τους υπόλοιπους ανθρώπους.
- ii. Οι εξωστρεφείς και οι εσωστρεφείς άνθρωποι θα διαφέρουν μεταξύ τους ως προς την ικανότητά τους να μαθαίνουν τα ονόματα των άλλων.

iii. Η επαγγελματική ικανοποίηση θα έχει αρνητική συσχέτιση με την απουσία στην εργασία.

iv. Οι άντρες έχουν θετικότερη στάση από τις γυναίκες ως προς την τεχνολογία.

## Απάντηση:

- i) Μονόπλευρος έλεγχος (πιο υγιεινή)
- ii) Αμφίπλευρος έλεγχος (θα διαφέρουν)
- iii) Μονόπλευρος (αρνητική συσχέτιση ,  $<0$ )
- iv) Μονόπλευρος (θετικότερη στάση)

## Άσκηση 5

Σε μία έρευνα που μελετάμε αν υπάρχει διαφορά ανάμεσα στους άντρες και στις γυναίκες ως προς τη συναισθηματική τους νοημοσύνη, βρήκαμε ότι η πιθανότητα το στατιστικό αποτέλεσμα που υπολογίσαμε να οφείλεται σε τυχαίους παράγοντες είναι  $p=0,5$  ( $p$ -value). Σε αυτή την περίπτωση θα πρέπει να απορρίψουμε ή όχι τη μηδενική υπόθεση που λέει ότι δεν υπάρχουν διαφορές ανάμεσα στους άντρες και τις γυναίκες ως προς τη συναισθηματική τους νοημοσύνη;

## Θεωρία:

- Αν η πιθανότητα να έχει εμφανιστεί μια συσχέτιση/διαφορά λόγω τυχαίων παραγόντων είναι μικρή, τότε απορρίπτουμε τη μηδενική υπόθεση (έχουμε στατιστικώς σημαντικό αποτέλεσμα). Η πιθανότητα αυτή ονομάζεται  $p$ -value.
- Όσο χαμηλότερη είναι η τιμή  $p$  τόσο μικρότερη είναι η πιθανότητα τα αποτελέσματα να προέρχονται από τυχαίους παράγοντες.

## Απάντηση:

Αφού η πιθανότητα σφάλματος είναι μεγάλη ( $p=0,5$ ), τότε θα πρέπει να δεχθούμε τη μηδενική υπόθεση που υποστηρίζει ότι δεν υπάρχει διαφορά ανάμεσα στους άντρες και στις γυναίκες ως προς τη συναισθηματική νοημοσύνη.

Στην περίπτωση που παίρναμε πιθανότητα 0,05 (δηλ.5% - συχνά μπερδεύομαστε!) θα πρέπει να απορριφθεί η μηδενική υπόθεση.

## Άσκηση 6

Διατυπώστε τη μηδενική και την εναλλακτική υπόθεση (αμφίπλευρος έλεγχος) για το παρακάτω ερευνητικό παράδειγμα:

Σε ένα πείραμα με σκοπό να μελετηθεί η επίδραση που έχει στην οδήγηση ένα φάρμακο για τη ζαλάδα από το ταξίδι πήραν μέρος 45 άτομα. Όλα τα άτομα έκαναν το πείραμα δύο φορές.

Την πρώτη φορά πήραν το νέο φάρμακο και στη συνέχεια μπήκαν σε εξομοιωτή οδήγησης και οδήγησαν μία συγκεκριμένη διαδρομή.

Τη δεύτερη φορά πήραν ψευδοφάρμακο και στη συνέχεια μπήκαν πάλι στον εξομοιωτή οδήγησης και οδήγησαν μία αντίστοιχη διαδρομή.

Αξιολογήθηκαν τα λάθη.

## Απάντηση: αμφίπλευρος έλεγχος

**H<sub>0</sub>:** Τα άτομα που πήραν το φάρμακο για τη ζαλάδα δεν θα έχουν διαφορετικό αριθμό λαθών στην οδήγηση από τα άτομα που πήραν ψευδο-φάρμακο.

**H<sub>1</sub>:** Τα άτομα που πήραν το φάρμακο για τη ζαλάδα θα έχουν διαφορετικό αριθμό λαθών στην οδήγηση από τα άτομα που πήραν ψευδο-φάρμακο.

## Άσκηση 7

Διατυπώστε τη μηδενική και την εναλλακτική υπόθεση (μονόπλευρος και αμφίπλευρος έλεγχος) για το παρακάτω ερευνητικό παράδειγμα:

Σε ένα πείραμα με σκοπό να μελετηθεί η επίδραση του άγχους στην αριθμητική ικανότητα χρησιμοποιήθηκαν 60 άτομα. Από αυτά, 30 χαρακτηρίστηκαν ως άτομα με χαμηλό άγχος και 30 ως άτομα με υψηλό άγχος.

Στη συνέχεια, όλα τα άτομα συμπλήρωσαν ένα τεστ αριθμητικής ικανότητας. Ο ερευνητής προσπαθούσε να διαπιστώσει κατά πόσο το άγχος επιδρά στην αριθμητική ικανότητα.

## Απάντηση: Μονόπλευρος Έλεγχος

**H<sub>0</sub>:** Τα άτομα που έχουν χαμηλό άγχος δεν θα έχουν καλύτερες επιδόσεις στο τεστ αριθμητικής ικανότητας από τα άτομα που έχουν υψηλό άγχος.

**H<sub>1</sub>:** Τα άτομα που έχουν χαμηλό άγχος θα έχουν καλύτερες επιδόσεις στο τεστ αριθμητικής ικανότητας από τα άτομα που έχουν υψηλό άγχος.

## Απάντηση: Δίπλευρος Έλεγχος

$H_0$ : Οι επιδόσεις στο τεστ αριθμητικής ικανότητας των ατόμων που έχουν χαμηλό άγχος δεν θα διαφέρουν από τις επιδόσεις των ατόμων που έχουν υψηλό άγχος.

$H_1$ : Οι επιδόσεις στο τεστ αριθμητικής ικανότητας των ατόμων που έχουν χαμηλό άγχος θα διαφέρουν από τις επιδόσεις των ατόμων που έχουν υψηλό άγχος.

## Άσκηση 8

Διατυπώστε τη μηδενική και την εναλλακτική υπόθεση (μονόπλευρος έλεγχος) για το παρακάτω ερευνητικό παράδειγμα:

Σε ένα πείραμα με σκοπό να μελετηθεί το επίπεδο εγρήγορσης ύστερα από τη χορήγηση ενός γεύματος πλούσιου σε υδατάνθρακες και ύστερα από ένα γεύμα φτωχό σε υδατάνθρακες χρησιμοποιήθηκαν 75 άτομα. Μετά από κάθε γεύμα, τους ζητήθηκε να συμπληρώσουν ένα ερωτηματολόγιο για το επίπεδο εγρήγορσης στο οποίο βρίσκονταν.

## Απάντηση: Μονόπλευρος Έλεγχος

**H<sub>0</sub>:** Τα άτομα που θα φάνε ένα γεύμα πλούσιο σε υδατάνθρακες δεν θα έχουν υψηλότερο επίπεδο εγρήγορσης από τα άτομα που θα φάνε ένα γεύμα φτωχό σε υδατάνθρακες.

**H<sub>1</sub>:** Τα άτομα που θα φάνε ένα γεύμα πλούσιο σε υδατάνθρακες θα έχουν υψηλότερο επίπεδο εγρήγορσης από τα άτομα που θα φάνε ένα γεύμα φτωχό σε υδατάνθρακες.

## Άσκηση 9

Στην προηγούμενη έρευνα, ποια αποτελέσματα θα σας οδηγήσουν να δεχθείτε τη μηδενική υπόθεση και ποια να μην την δεχθείτε;

Απάντηση: Θα δεχθούμε τη μηδενική υπόθεση αν τα αποτελέσματα δείξουν ότι ο μέσος όρος εγρήγορσης των ατόμων που έφαγαν ένα γεύμα πλούσιο σε υδατάνθρακες δεν είναι στατιστικώς διαφορετικός ( $p > 0,05$ , **αυτή είναι καλή τιμή για να χρησιμοποιήσουμε!**) από το μέσο όρο των ατόμων που έφαγαν ένα γεύμα φτωχό σε υδατάνθρακες. Αντίθετα, δεν θα δεχθούμε τη μηδενική υπόθεση αν τα αποτελέσματα δείξουν ότι ο μέσος όρος εγρήγορσης των ατόμων που έφαγαν ένα γεύμα πλούσιο σε υδατάνθρακες είναι στατιστικώς διαφορετικός ( $p < 0,05$ ) από το μέσο όρο αυτών που έφαγαν ένα γεύμα φτωχό σε υδατάνθρακες.

## Άσκηση 10

Χρησιμοποιώντας καθένα από τα παρακάτω τυχαία δείγματα, υπολογίστε τα 95% και 99% διαστήματα εμπιστοσύνης για τον μέσο όρο του πληθυσμού:

α)  $\bar{X}_A=25, s=6, n=15$

β)  $\bar{X}_B=30,6, s=5,5, n=24$

α)  $\bar{X}_\Gamma=35, s=8, n=210$

## Θεωρία:

Το διάστημα εμπιστοσύνης CI, είναι ένα διάστημα μέσα στο οποίο εκτιμάται ότι βρίσκεται η πραγματική τιμή της παραμέτρου που ενδιαφέρει τον ερευνητή (π.χ. μέσος όρος). Ορίζεται πάντα επί κάπου ποσοστού εμπιστοσύνης (90%, 95% ή 99% ή άλλο, ποτέ 100%!).

$$CI = \left( \bar{X} - z \times \frac{s}{\sqrt{n}}, \bar{X} + z \times \frac{s}{\sqrt{n}} \right)$$

όπου  $\bar{X}$ : είναι ο μέσος όρος του δείγματος,  $s$ : η τυπική απόκλιση και  $n$  το μέγεθος του δείγματος. Το  $z$  εξαρτάται από το επίπεδο εμπιστοσύνης που έχει επιλεγεί. (Αργότερα θα μάθουμε ότι εξαρτάται και από τις πληροφορίες που έχουμε επί του πληθυσμού, και από το πόσο μεγάλο είναι το δείγμα και από το τι μελετάται).

Επίπεδο Εμπιστοσύνης	Τιμές του $z$
90%	1,64
95%	1,96
99%	2,58

Λύση Άσκησης 10:

**A. Για το 95% επίπεδο εμπιστοσύνης:**

α)  $\bar{X} = 25, s = 6, n = 15$

$$\left( 25 - z_{95\%} \times \frac{6}{\sqrt{15}}, 25 + z_{95\%} \times \frac{6}{\sqrt{15}} \right) = \left( 25 - 1,96 \times \frac{6}{\sqrt{15}}, 25 + 1,96 \times \frac{6}{\sqrt{15}} \right) \\ = (21,96, 28,04)$$

β)  $\bar{X} = 30,6, s = 5,5, n = 24$

$$\left( 30,6 - z_{95\%} \times \frac{5,5}{\sqrt{24}}, 30,6 + z_{95\%} \times \frac{5,5}{\sqrt{24}} \right) = (28,40, 32,80)$$

γ) Άσκηση

Απάντηση: (33,92, 36,08)

**B. Για το 99% επίπεδο εμπιστοσύνης:**

α)  $\bar{X} = 25, s = 6, n = 15$

$$\left(25 - z_{99\%} \times \frac{6}{\sqrt{15}}, 25 + z_{99\%} \times \frac{6}{\sqrt{15}}\right) = \left(25 - 2,58 \times \frac{6}{\sqrt{15}}, 25 + 2,58 \times \frac{6}{\sqrt{15}}\right) \\ = (21,01, 28,99)$$

β) Άσκηση

Απάντηση: (27,71, 33,49)

γ) Άσκηση

Απάντηση: (33,58, 36,42)

## Άσκηση 11

Γιατί είναι σημαντική η εκτίμηση του μεγέθους της επίδρασης πριν από την έναρξη μιας έρευνας;

Απάντηση: Γιατί γνωρίζοντας το μέγεθος της επίδρασης μιας έρευνας μπορούμε να υπολογίσουμε το μέγεθος του δείγματος που απαιτείται για την ασφαλή εκτίμηση των παραμέτρων του πληθυσμού και την αποφυγή σφάλματος τύπου II.

## Θεωρία

### **Μηδενική Υπόθεση, $H_0$ (null hypothesis):**

Συνήθως δεν έχει ιδιαίτερο ενδιαφέρον και χρησιμοποιείται με σκοπό να λειτουργήσει ως κριτήριο σύγκρισης για την εναλλακτική υπόθεση. Στην  $H_0$  δηλώνουμε ότι δεν υπάρχει διαφορά ανάμεσα στις μεταβλητές που μελετάμε. Στην ουσία δηλαδή θέλουμε τελικά να απορριφθεί διότι τότε θα ισχύει το επιχείρημά μας (που διατυπώνεται στην εναλλακτική).

Αν όμως απορρίψουμε τη Μηδενική Υπόθεση ενώ είναι αληθής τότε υποπέφτουμε σε σφάλμα τύπου I.

## Θεωρία

**Εναλλακτική (ή ερευνητική) Υπόθεση,  $H_1$  (alternative/research hypothesis):**

Είναι στην ουσία η ερευνητική υπόθεση του μελετητή, η οποία αναφέρει ότι υπάρχει σχέση ανάμεσα στις μεταβλητές. Για να αποδεχθούμε ότι η υπόθεση είναι σωστή θα πρέπει να απορριφθεί η  $H_0$ .

		Πραγματικότητα	
		Αληθής Μηδενική Υπόθεση	Ψευδής Μηδενική Υπόθεση
Αποσπελέσματα Ελέγχου της Υπόθεσης	<b>Απόρριψη</b> Μηδενικής Υπόθεσης	<b>Σφάλμα τύπου I</b> (ψευδώς θετικά αποτελέσματα)	<b>Ισχύς</b> (αληθώς θετικά αποτελέσματα)
	<b>Μη απόρριψη</b> Μηδενική Υπόθεσης	<b>Σωστό</b> (αληθώς αρνητικά αποτελέσματα)	<b>Σφάλμα τύπου II</b> (ψευδώς αρνητικά αποτελέσματα)

ΤΕΛΟΣ