

Ατομική εργασία Νο 4 και 5

ΑΣΚΗΣΗ 4

Τα ακόλουθα δεδομένα μας δίνουν τον αριθμό νέων οικοδομικών αδειών (2011) για 19 Δημοτικές Ενότητες. Συνολικά η Ελλάδα αποτελείται από 1036 Δημοτικές Ενότητες.

	Δ.Ε.	Αριθμός οικοδομικών αδειών
1	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΓΡΙΑΣ	22
2	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΙΣΩΝΙΑΣ	17
3	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΡΤΕΜΙΔΑΣ	23
4	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΙΩΛΚΟΥ	11
5	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΜΑΚΡΙΝΙΤΣΗΣ	8
6	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΝΕΑΣ ΑΓΧΙΑΛΟΥ	18
7	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΝΕΑΣ ΙΩΝΙΑΣ	54
8	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΠΟΡΤΑΡΙΑΣ	12
9	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΛΜΥΡΟΥ	57
10	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΠΤΕΛΕΟΥ	8
11	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΖΑΓΟΡΑΣ	36
12	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΜΟΥΡΕΣΙΟΥ	37
13	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΡΓΑΛΑΣΤΗΣ	49
14	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΑΦΕΤΩΝ	15
15	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΜΗΛΕΩΝ	24
16	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΣΗΠΙΑΔΟΣ	28
17	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΤΡΙΚΕΡΙΟΥ	8
18	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΚΑΡΛΑΣ	10
19	ΔΗΜΟΤΙΚΗ ΕΝΟΤΗΤΑ ΚΕΡΑΜΙΔΙΟΥ	3

4.1. /95% και 99% Διάστημα Εμπιστοσύνης (Δ.Ε.) για το μέσο αριθμό οικοδομικών αδειών.

Δεδομένου ότι $n = 19 < 30$, το $(1-\alpha)\%$ Δ.Ε. υπολογίζεται ως εξής:

$$\bar{X} \pm t(n-1; \alpha/2) \times \frac{s}{\sqrt{n}} \cdot \sqrt{d}$$

Όπου $d = 1 - (n/N) = 1 - (19/1034) = 0,982$

$$\bar{X} = \frac{1}{19} \times 440 = 23,16$$

$\sum_{i=1}^{19} (X_i - \bar{X})^2 = 4818,53 \rightarrow s^2 = 4818,53 / 18 = 267,696$ (προσοχή $n < 30$, επομένως η διαίρεση είναι με βάση $n-1$)

$$s = 16,36$$

Για $\alpha = 5\%$, $t(18;0,025) = 2,101$ και για $\alpha = 1\%$, $t(18;0,005) = 2,878$

$$95\% \text{ Δ.Ε. : } 23,16 \pm 2,101 \frac{16,36}{\sqrt{19}} \sqrt{0,982} = 23,16 \pm 7,814$$

$$\Leftrightarrow 15,35 \leq \bar{X} \leq 30,97$$

Το εύρος του 95% Δ.Ε. = 15,63 δηλαδή: εύρος / μέση τιμή = 0,675 (67,5%)

$$99\% \text{ Δ.Ε. : } 23,16 \pm 2,878 \frac{16,36}{\sqrt{19}} \sqrt{0,982} = 23,16 \pm 10,704$$

$$\Leftrightarrow 12,46 \leq \bar{X} \leq 33,86$$

Το εύρος του 99% Δ.Ε. = 21,40 δηλαδή: εύρος / μέση τιμή = 0,924 (92,4%)

Το εύρος του 95% Δ.Ε. όπως και του 99% Δ.Ε. είναι πολύ μεγάλο, κατά συνέπεια το δείγμα δεν μπορεί να αντανakλά με αξιοπιστία τον πραγματικό μέσο αριθμό νέων οικοδομικών αδειών για το σύνολο των 1034 Δημοτικές Ενότητες της Χώρας.

4.2. / 95% Δ.Ε. για τη διαφορά των μέσων τιμών

Τουλάχιστον ένα από τα 2 δείγματα είναι μικρό, επομένως το 95% Δ.Ε. δίνεται από:

$$(\bar{X}_1 - \bar{X}_2) \pm t(v, \alpha/2) \cdot s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

και $v = n_1 + n_2 = 19 + 50 = 69$

	1 ^ο Δείγμα	2 ^ο Δείγμα
Μέγεθος δείγματος	19	50
Μέση τιμή	23,16	20,5
Διασπορά s^2	$(16,36)^2 = 267,65$	$(19,85)^2 = 394,02$

Για 95% , $t(v;0,025) = t(69;0,025) = 1,96$

$$s^2 = \frac{18 \times 267,65 + 49 \times 394,02}{67} = 360,07 \rightarrow s = 18,98$$

$$\sqrt{\frac{1}{19} + \frac{1}{50}} = 0,270$$

95% Δ.Ε. : $(23,16 - 20,5) \pm 1,96 \times 18,98 \times 0,270 = 2,66 \pm 10,04$

$$\rightarrow -7,38 \leq \bar{X}_1 - \bar{X}_2 \leq +12,70$$

Το 95% Δ.Ε. περιλαμβάνει το 0 (αρνητικές και θετικές τιμές) \rightarrow οι 2 μέσες τιμές δεν διαφέρουν μεταξύ τους. Δεν υπάρχει σημαντική στατιστική διαφορά μεταξύ των 2 μέσων τιμών.

ΑΣΚΗΣΗ 5:

Μια δειγματοληπτική έρευνα πραγματοποιήθηκε σε τρεις διαφορετικές ομάδες καταναλωτών της Αττικής (σύνολο: 1000 άτομα). Ζητήσαμε στους καταναλωτές να δηλώνουν τις προτιμήσεις τους για ένα συγκεκριμένο προϊόν. Σύμφωνα με τα αποτελέσματα της έρευνας, έχουμε τα ακόλουθα αποτελέσματα:

	Ομάδες καταναλωτών			Σύνολο
	Ομάδα Α	Ομάδα Β	Ομάδα Γ	
Σύνολο ατόμων ανά ομάδα	500	400	100	1000
Αριθμός καταναλωτών που προτιμούν το προϊόν:	250	240	10	500
Ποσοστό καταναλωτών που προτιμούν το προϊόν	0,5	0,6	0,1	0,5

5.1. Δ.Ε. για το ποσοστό των καταναλωτών που προτιμούν το συγκεκριμένο προϊόν, ανεξάρτητα από την ομάδα στην οποία ανήκουν.

Συνολικά $n = 1000$, N (αριθμός καταναλωτών της Αττικής) άγνωστος και πολύ μεγάλος $\rightarrow n/N \rightarrow 0 \rightarrow d = 1$

Από τους 1000 καταναλωτές, οι 500 δηλώσαν ότι προτιμούν το προϊόν, επομένως $p = 0,5$ ($q = 0,5$).

$$95\% \text{ Δ.Ε. : } 0,5 \pm 1,96 \times \sqrt{\frac{0,5 \times 0,5}{1000}} = 0,5 \pm 0,031$$

$$\mathbf{0,469 \leq p \leq 0,531}$$

Εύρος του Δ.Ε. = 0,062 \rightarrow εύρος / $p = 0,124$ (**12,4%**) πολύ περιορισμένο. Πραγματικά, με 95% βεβαιότητας, μπορούμε να στηρίξουμε ότι, περίπου το 50% των καταναλωτών της Αττικής προτιμούν το συγκεκριμένο προϊόν.

$$99\% \text{ Δ.Ε. : } 0,5 \pm 2,576 \times \sqrt{\frac{0,5 \times 0,5}{1000}} = 0,5 \pm 0,041$$

$$\mathbf{0,459 \leq p \leq 0,541}$$

Εύρος του Δ.Ε. = 0,082 \rightarrow εύρος / $p = 0,163$ (**16,3%**) παραμένει περιορισμένο. Με 99% βεβαιότητας, περίπου το 50% των καταναλωτών της Αττικής προτιμούν το συγκεκριμένο προϊόν.

5.2. Δ.Ε. για το ποσοστό των καταναλωτών της ομάδας Γ που προτιμούν τον προϊόν.

$n = 100$, N άγνωστο $\rightarrow n/N \rightarrow 0 \rightarrow d = 1$

Από τους 100 καταναλωτές, οι 10 προτιμούν το προϊόν, επομένως $p = 0,1$ ($q = 0,9$).

$$95\% \text{ Δ.Ε. : } 0,1 \pm 1,96 \times \sqrt{\frac{0,1 \times 0,9}{100}} = 0,1 \pm 0,059$$

$$\mathbf{0,041 \leq p \leq 0,159}$$

Εύρος του Δ.Ε. = 0,118 \rightarrow εύρος / $p = 1,176$ (**118%**) πάρα πολύ μεγάλο!!!

$$99\% \text{ Δ.Ε. : } 0,1 \pm 2,576 \times \sqrt{\frac{0,1 \times 0,9}{100}} = 0,1 \pm 0,077$$

$$\mathbf{0,023 \leq p \leq 0,177}$$

Εύρος του Δ.Ε. = 0,154 \rightarrow εύρος / $p = 1,546$ (**155%**) πάρα πολύ μεγάλο!!!

Η ομάδα Γ δεν είναι καθόλου αντιπροσωπευτική.

5.3. Δ.Ε. για το ποσοστό των καταναλωτών της ομάδας Β που δεν προτιμούν τον προϊόν.

Στην ομάδα Β, το 60% προτιμούν το προϊόν επομένως το 40% δεν το προτιμούν. Σε αυτή την περίπτωση θέτουμε $p = 0,4$ και $q = 0,6$.

$$95\% \text{ Δ.Ε. : } 0,4 \pm 1,96 \times \sqrt{\frac{0,4 \times 0,6}{400}} = 0,4 \pm 1,96 \times 0,024 = 0,4 \pm 0,047$$

$$\mathbf{0,353 \leq p \leq 0,447}$$

$$\text{Εύρος του Δ.Ε.} = 0,094 \rightarrow \text{εύρος} / p = 0,235 \text{ (23,5\%)}$$

$$99\% \text{ Δ.Ε. : } 0,4 \pm 2,576 \times \sqrt{\frac{0,4 \times 0,6}{400}} = 0,4 \pm 2,576 \times 0,024 = 0,4 \pm 0,062$$

$$\mathbf{0,338 \leq p \leq 0,462}$$

$$\text{Εύρος του Δ.Ε.} = 0,124 \rightarrow \text{εύρος} / p = 0,31 \text{ (31\%)}$$

Το Δ.Ε. για το ποσοστό των καταναλωτών της ομάδας Β που δεν προτιμούν τον προϊόν είναι σχετικά μεγάλο (ειδικά το 99% Δ.Ε.) και κατά συνέπεια, η ομάδα Β να μην αντιπροσωπεύει καλά τον πληθυσμό.

5.4. 95% Δ.Ε. για τη διαφορά μεταξύ του ποσοστού των καταναλωτών της ομάδας Α που προτιμούν το προϊόν και του αντίστοιχου ποσοστού για την ομάδα Β.

Δεδομένου ότι, τα 2 υπό-δείγματα είναι μεγάλα $n_A = 400$ και $n_B = 500$, το 95% Δ.Ε. για τη διαφορά μεταξύ των 2 ποσοστών δίνεται από:

$$(p_B - p_A) \pm 1,96 \times \sqrt{\frac{p_B q_B}{n_B} + \frac{p_A q_A}{n_A}}$$

Δεδομένου ότι $p_A = 0,5$ και $p_B = 0,6$ επιλέξαμε να υπολογίσουμε τη διαφορά $p_B - p_A$ διότι είναι θετική και διαβάζεται πιο εύκολα.

$$95\% \text{ Δ.Ε.: } (0,6 - 0,5) \pm 1,96 \times \sqrt{\frac{0,6 \times 0,4}{400} + \frac{0,5 \times 0,5}{500}}$$

$$95\% \text{ Δ.Ε.: } 0,1 \pm 1,96 \times 0,033 = 0,1 \pm 0,065$$

$$\mathbf{0,035 \leq (p_B - p_A) \leq 0,165}$$

Το 95% Δ.Ε. για τη διαφορά μεταξύ των 2 ποσοστών είναι συστηματικά θετικό και δεν περιλαμβάνει το 0. Κατά συνέπεια, τα δύο ποσοστά είναι πραγματικά διαφορετικά: οι καταναλωτές της ομάδας Β εκφράζουν μεγαλύτερη προτίμηση για το προϊόν σε σχέση με τους καταναλωτές της ομάδας Α.

5.5. Γενικό συμπέρασμα

Το δείγμα στο σύνολο του είναι αρκετά αντιπροσωπευτικό του συνόλου των καταναλωτών της Αττικής (5.1). Οι υπό-ομάδες δεν είναι αντιπροσωπευτικές του συνόλου των καταναλωτών, ειδικά οι ομάδα Γ (5.2 και 5.3). Το γεγονός αυτό είναι λογικό διότι κάθε υπό-ομάδα έχει διαφορετική συμπεριφορά (% προτίμησης διαφορετικό). Όπως αναφέρθηκε (5.4), η διαφορά μεταξύ ομάδας Β και ομάδας Α - παρότι είναι σχετικά μικρή (10%) - είναι στατιστικά σημαντική.