

Άσκηση 1: Να περιγράψετε συνοπτικά τις μεθόδους (α) ελέγχου και διόρθωσης (input editing) και (β) αναπλήρωσης των ελλειπουσών τιμών (data imputation, missing values) κατά τη διεξαγωγή μιας δειγματοληπτικής έρευνας.

Άσκηση 2: (Κυκλική συστηματική δειγματοληψία). Στη συστηματική δειγματοληψία όταν το μέγεθος του πληθυσμού N δεν είναι πολλαπλάσιο του μεγέθους δείγματος n , $N \neq nk$, τότε εφαρμόζεται ο λεγόμενος κυκλικός νόμος που συνίσταται στα εξής δύο βήματα:

- Λαμβάνεται ένας τυχαίος αριθμός, έστω ο r , από το σύνολο των N μονάδων του πληθυσμού.
- Συμπεριλαμβάνονται στο δείγμα (μεγέθους n) οι μονάδες y_i του πληθυσμού με δείκτη i , όπου $i = r + km$ εάν $i \leq N$ και $i = r + km - N$ εάν $i > N$, με $m = 0, 1, 2, \dots, n - 1$.

(α) Να δειχθεί ότι η μέση τιμή του δείγματος είναι αμερόληπτη εκτιμήτρια της πληθυσμιακής μέσης τιμής και να βρεθεί η διασπορά της.

(β) Εφαρμογή: Έστω ότι έχουμε $N=33$ και $n=6$ οπότε $k=[N/n]=5$. Έστω επίσης ότι οι τιμές του χαρακτηριστικού y πληθυσμού είναι $y_i = i$, $i = 1, 2, \dots, N$. Να βρεθούν όλα τα δυνατά δείγματα μεγέθους $n=6$.

(β1) με το συνήθη τρόπο, (β2) με τον κυκλικό νόμο. Στη συνέχεια να βρεθούν σε κάθε δείγμα οι μέσες τιμές και οι διασπορές και να γίνει η σχετική σύγκριση (αμεροληψία, διασπορά κ.λ.π).

(γ) Να ληφθεί με επαναλαμβανόμενη συστηματική δειγματοληψία ένα συνολικό δείγμα μεγέθους $n=6$ χρησιμοποιώντας δύο συστηματικά δείγματα μεγέθους τρία το καθένα. Να εκτιμηθεί η μέση τιμή του πληθυσμού και να δοθεί διάστημα εμπιστοσύνης συντελεστή 95%.

Άσκηση 3: Η διάρκεια εξέτασης y (σε λεπτά) $N=12$ ασθενών μιας κλινικής είναι:

15 34 35 36 11 17 49 40 25 46 33 14

(α) (i) Να υπολογισθεί ο πληθυσμιακός μέσος και η πληθυσμιακή διασπορά σ^2 (και S^2).

(ii) Να ληφθούν όλα τα συστηματικά δείγματα μεγέθους $n=6$ και να υπολογισθούν οι μέσες τιμές και οι διασπορές τους.

(iii) Να υπολογισθούν τα $Var(\bar{y}_{sy})$ και ρ_{ω} .

(β) Να επαναληφθούν τα παραπάνω βήματα (α)-(γ) όταν οι τιμές του πληθυσμού είναι:

49 46 40 36 35 34 33 25 17 15 14 11

(γ) Να επαναληφθούν τα παραπάνω βήματα (α)-(γ) όταν οι τιμές του πληθυσμού είναι:

11 17 36 49 14 34 35 46 15 25 33 40

(δ) Να σχολιασθούν τα αποτελέσματα.

Άσκηση 4: Σ' ένα πληθυσμό μεγέθους $N = 8$ που χωρίζεται σε δύο στρώματα μεγέθους $N_1=N_2=4$, οι τιμές των χαρακτηριστικών x και y δίνονται στον παρακάτω πίνακα.

# ατόμου, i	1	2	3	4
Στρώμα 1 x_{1i}	2	5	9	15
y_{1i}	0	3	7	10
Στρώμα 2 x_{2i}	10	18	21	25
y_{2i}	7	15	10	16

Χρησιμοποιώντας τον πίνακα τυχαίων αριθμών να παρθεί απλό τυχαίο δείγμα δύο ατόμων από κάθε στρώμα (μεγέθους $n_1 = n_2 = 2$), να εκτιμηθεί ο πληθυσμιακός μέσος \bar{Y} και το πληθυσμιακό ολικό Y με χωριστή και συνδυασμένη λογοεκτιμήτρια και να δοθούν αντίστοιχα δ.ε. συντελεστού 95% (θεωρήστε ικανοποιητική την κανονική προσέγγιση).

Σημ. Να περιγραφεί η διαδικασία επιλογής του δείγματος από τον πίνακα τυχαίων αριθμών.

Άσκηση 5: Η άσκηση 9 του Κεφ 5 σελ. 278 (στρωματοποιημένη κατά συστάδες δειγματοληψία).