

Άσκηση 1: Να περιγράψετε αναλυτικά με τι ασχολείται η δειγματοληψία. Αναφέρετε πλεονεκτήματα και μειονεκτήματα αυτής (της δειγματοληψίας) έναντι της ολικής απογραφής.

Άσκηση 2: (Α) Με **α.τ.δ.** πάρθηκε δείγμα 40 φοιτητών και ρωτήθηκαν εάν επιθυμούν ενδιάμεση εξέταση στα μαθήματα. Εάν 25 φοιτητές απάντησαν καταφατικά, να εκτιμηθεί **(α)** το ποσοστό και **(β)** το σύνολο των φοιτητών που είναι υπέρ της ενδιάμεσης εξέτασης και να δοθούν αντίστοιχα 95% διαστήματα εμπιστοσύνης.

Υποθέστε ότι το σύνολο των φοιτητών είναι **(ι)** 2000, **(ii)** 100000.

(Β) Να απαντηθούν τα παραπάνω ερωτήματα εάν το δείγμα πάρθηκε με **α.τ.δ.ε.**

Άσκηση 3: Σε μία στρωματοποιημένη δειγματοληψία το κόστος είναι της μορφής $C = c_0 + \sum_{h=1}^L c_h \ln(n_h)$.

(α) Να βρεθεί το μέγεθος δείγματος που απαιτείται και η κατανομή αυτού στα L στρώματα για την ελαχιστοποίηση του κόστους με δεδομένη διασπορά της εκτιμήτριας του πληθυσμιακού μέσου $V(\bar{y}_{st}) = V_0$.

(β) Να υπολογιστεί και το ελάχιστο κόστος.

(γ) Εφαρμογή: Δίνονται τα δεδομένα του παρακάτω πίνακα και επί πλέον $c_0 = 0$, $V_0 = 0,1$ και $N = 200$.

Στρώμα	h	1	2	3
Βάρος	W_h	0,5	0,3	0,2
Διασπορά	S_h^2	2,25	3,24	3,24
Κόστος	c_h	9	25	36

Άσκηση 4: Σε μία εταιρία το 31% των υπαλλήλων είναι ΥΕ, το 62% είναι ΜΕ και το 7% είναι ΑΕ. Θέλοντας να εκτιμήσουμε το ποσοστό των υπαλλήλων που χρησιμοποιούν συγκεκριμένες εγκαταστάσεις αναψυχής λαμβάνουμε με στρωματοποιημένη δειγματοληψία ένα συνολικό δείγμα 400 υπαλλήλων. Αρχικές υποθέσεις είναι ότι οι εγκαταστάσεις χρησιμοποιούνται από το 20%-30% των υπαλλήλων ΥΕ, από το 40%-50% των υπαλλήλων ΜΕ και από το 5%-10% των υπαλλήλων ΑΕ.

(α) Πως θα κατανεμηθεί το δείγμα στις τρεις κατηγορίες υπαλλήλων;

(β) Αν τα πραγματικά ποσοστά ήταν 21%, 48% και 4%, αντίστοιχα, ποιο θα ήταν το τυπικό σφάλμα της εκτίμησης;

(γ) Ποιο θα ήταν το τυπικό σφάλμα του συνολικού ποσοστού των υπαλλήλων (όλων μαζί) εάν είχαμε απλή τυχαία δειγματοληψία $n=400$ υπαλλήλων;

Δίνεται: δ.π.π. παραλείπεται.

Άσκηση 5: Για την εκτίμηση του πληθυσμιακού μέσου \bar{Y} χρησιμοποιούμε 1-κάθε- k συστηματική δειγματοληψία με βάση δείγμα μεγέθους n από πληθυσμό μεγέθους N (υποθέτουμε ότι $kn = N$). Έστω ότι $y_i = N - i + 1$, $i = 1, \dots, N$.

Έστω \bar{y}_{sy} η εκτιμήτρια του πληθυσμιακού μέσου.

(α) Να βρεθεί ο \bar{Y}_{r_1} (ο μέσος του r συστηματικού δείγματος).

(β) Να δειχθεί ότι $Var(\bar{y}_{sy}) = \frac{1}{k} \sum_{r=1}^k (\bar{y}_{r_1} - \bar{Y})^2 = \frac{k^2 - 1}{12}$.

(γ) Ο πληθυσμός χωρίζεται σε στρώματα. Το πρώτο στρώμα αποτελείται από τις πρώτες $2k$ παρατηρήσεις, το δεύτερο από τις επόμενες $2k$ παρατηρήσεις, κ.λπ. (υποθέτουμε ότι $N/2k$ είναι ακέραιος). Χρησιμοποιούμε στρωματοποιημένη δειγματοληψία παίρνοντας 2 παρατηρήσεις από κάθε στρώμα. Έστω \bar{y}_{st} η εκτιμήτρια του μέσου.

Να δειχθεί ότι $Var(\bar{y}_{st}) < Var(\bar{x}_{sy})$ αν $n > \frac{4k+2}{k+1}$.