

Έρευνα διακοπών

ΕΙΔΟΣ

Δειγματοληπτική έρευνα

Μέθοδος δειγματοληψίας και δειγματοληπτική μονάδα

Κατά την έρευνα εφαρμόστηκε η διασταδιακή στρωματοποιημένη δειγματοληψία με πρωτογενή μονάδα έρευνας την επιφάνεια (ένα ή περισσότερα συνεχόμενα οικοδομικά τετράγωνα) και τελική μονάδα έρευνας το νοικοκυριό.

Κριτήρια στρωμάτωσης και υπο - στρωμάτωσης του πληθυσμού

Πρώτο κριτήριο στρωμάτωσης του ερευνώμενου πληθυσμού ήταν η γεωγραφική διαίρεση της Χώρας. Ως μείζονα στρώματα χρησιμοποιήθηκαν οι 13 Περιφέρειες (NUTS II), η Περιφέρεια Πρωτευούσης και το Πολεοδομικό Συγκρότημα Θεσσαλονίκης.

Κατά το δεύτερο κριτήριο στρωμάτωσης σε κάθε Περιφέρεια (NUTS II), οι Δήμοι / Κοινότητες στρωματώθηκαν (χωρίστηκαν) κατά βαθμό αστικότητας ως εξής:

- Δήμοι με πληθυσμό ≥ 30.000 κατοίκους
- Δήμοι με πληθυσμό 5.000 - 29.999 κατοίκους
- Δήμοι / Κοινότητες με πληθυσμό 1.000 - 4.999 κατοίκους
- Κοινότητες με πληθυσμό 1 - 999 κατοίκους

Το πλήθος των τελικών στρωμάτων των δεκατριών (13) γεωγραφικών περιφερειών ανέρχεται στα 50. Η Περιφέρεια Πρωτευούσης χωρίστηκε σε 31 στρώματα ίσου περίπου μεγέθους (ίσος αριθμός νοικοκυριών) με βάση τους καταλόγους πλαισίου με τα οικοδομικά τετράγωνα των Δήμων και λαμβάνοντας υπόψη κοινωνικοοικονομικά κριτήρια. Όμοια το Πολεοδομικό Συγκρότημα Θεσσαλονίκης χωρίστηκε σε 9 ίσου μεγέθους στρώματα. Τα Πολεοδομικά Συγκροτήματα των δύο Μεγάλων Πόλεων αποτελούν το 40% του συνολικού πληθυσμού και κατέχουν ακόμα υψηλότερο ποσοστό σε κάποιες κοινωνικοοικονομικές μεταβλητές.

Το πλήθος των τελικών στρωμάτων, τα οποία προέκυψαν από την εφαρμογή των δύο κριτηρίων στρωματοποίησης του πληθυσμού, *ανέρχεται στα 90*.

Μέγεθος και κριτήρια κατανομής του δείγματος

Το αρχικό μέγεθος δείγματος είναι 10.000 νοικοκυριά (το κλάσμα δειγματοληψίας είναι περίπου 2,5 ‰). Το κλάσμα αυτό είναι το ίδιο σε κάθε γεωγραφική περιφέρεια. Οι γεωγραφικές περιφέρειες είναι 13 στην Ελλάδα. Όμως σε αυτή τη μελέτη, η 2^η γεωγραφική περιφέρεια (Κεντρική Μακεδονία) λήφθηκε υπόψη χωρίς το Πολεοδομικό Συγκρότημα Θεσσαλονίκης και η 9^η γεωγραφική περιφέρεια (Αττική) λήφθηκε υπόψη χωρίς την Περιφέρεια Πρωτευούσης, ενώ και τα δύο μεγάλα συγκροτήματα θεωρήθηκαν ως ξεχωριστές γεωγραφικές περιφέρειες.

Διαδικασία επιλογής του δείγματος

1^ο στάδιο δειγματοληψίας

Στο 1^ο στάδιο δειγματοληψίας, από κάθε ένα στρώμα (έστω στρώμα h), επιλέχθηκαν n_h πρωτογενείς μονάδες (επιφάνειες). Ο αριθμός n_h των πρωτογενών μονάδων σε κάθε στρώμα είναι τέτοιος, ώστε να είναι ανάλογος του πληθυσμού του στρώματος X_h , δηλαδή κάθε πρωτογενής μονάδα του στρώματος έχει πιθανότητα επιλογής ανάλογη του μεγέθους της. Έτσι, αν X_{hi} είναι ο αριθμός των νοικοκυριών της μονάδας του δείγματος τάξης i ($i = 1, 2, \dots, n_h$) και X_h ο αριθμός των νοικοκυριών του στρώματος h , σύμφωνα με την Απογραφή Πληθυσμού έτους 2001, τότε η πιθανότητα επιλογής είναι

$$: P_{hi} = \frac{X_{hi}}{X_h}$$

Ο συνολικός αριθμός των πρωτογενών μονάδων του δείγματος ανέρχεται στις 1.056 μονάδες. Το δείγμα των πρωτογενών μονάδων (επιφανειών) ισοκατανεμήθηκε μέσα στο έτος, ώστε να επιλεγούν 4 ισοδύναμα, ανεξάρτητα μεταξύ τους δείγματα, που αντιστοιχούν στα 4 τρίμηνα της έρευνας.

2^ο στάδιο δειγματοληψίας

Στο 2^ο στάδιο δειγματοληψίας, σε κάθε επιλεγείσα πρωτογενή δειγματοληπτική μονάδα (μονάδα επιφάνειας) επιλέγεται το δείγμα των τελικών μονάδων (νοικοκυριών). Στην πραγματικότητα, στο δεύτερο στάδιο επιλέγεται δείγμα κατοικιών. Εντούτοις, στις περισσότερες περιπτώσεις, υπάρχει μία προς μία αντιστοίχιση μεταξύ νοικοκυριού και κατοικίας. Εάν η επιλεγείσα κατοικία αποτελείται από περισσότερα του ενός νοικοκυριά, τότε ερευνώνται όλα τα νοικοκυριά.

Έστω M_{hi} ο αριθμός των νοικοκυριών κατά τη διάρκεια της έρευνας στην i επιλεγείσα μονάδα επιφάνειας στο στρώμα h . Από αυτά, ένα συστηματικό δείγμα m_{hi} νοικοκυριών επιλέγεται με ίσες πιθανότητες. Καθένα από τα m_{hi} νοικοκυριά έχει την ίδια πιθανότητα να περιληφθεί στην έρευνα και η οποία είναι ίση με $\frac{m_{hi}}{M_{hi}}$.

Σε κάθε επιλεγείσα πρωτογενή μονάδα, προσδιορίζεται το μέγεθος του δείγματος m_{hi} . Ο συνολικός αριθμός των νοικοκυριών που πρόκειται να ερευνηθούν από τις n_h

επιλεγείσες πρωτογενείς δειγματοληπτικές μονάδες είναι
$$m_h = \sum_{i=1}^{n_h} m_{hi}.$$

Τελικά, εφαρμόζοντας δισταδιακή δειγματοληψία, το ποσοστό των νοικοκυριών που επιλέγονται από το στρώμα ισούται με $\frac{m_h}{M_h}$.

Στην επαναλαμβανόμενη δειγματοληψία, ο αριθμητής αυτού του κλάσματος θα ποικίλει από δείγμα σε δείγμα, δηλαδή συγκεκριμένα το κλάσμα $\frac{m_h}{M_h}$ θα είναι τυχαία μεταβλητή.

Για τον καθορισμό του μεγέθους m_{hi} , αρχικά υπολογίζεται το διάστημα δειγματοληψίας $\delta_{hi} = \frac{M_{hi}}{m_{hi}}$ και ακολούθως $m_{hi} = \frac{M_{hi}}{\delta_{hi}}$, έτσι ώστε να ικανοποιούνται οι δύο ακόλουθες συνθήκες:

α) Το μέγεθος m_{hi} υπολογίζεται σε κάθε πρωτογενή δειγματοληπτική μονάδα hi έτσι, ώστε εφαρμόζοντας τη διαδικασία της δισταδιακής δειγματοληψίας, να επιλέγεται από το στρώμα ποσοστό νοικοκυριών ίσο με: $\frac{m_h}{M_h} = \frac{1}{\lambda}$, όπου $\frac{1}{\lambda} = 2,5\%$ είναι το

γενικό κλάσμα δειγματοληψίας. Δηλαδή, το αναμενόμενο αποτέλεσμα $\frac{m_h}{M_h}$

προκαθορίζεται σε $\frac{1}{\lambda}$ για κάθε περιφέρεια (NUTS II): $E\left(\frac{m_h}{M_h}\right) = \frac{1}{\lambda} = 2,5\%$.

β) Ο εκτιμητής του συνόλου του στρώματος Y_h (για οποιοδήποτε χαρακτηριστικό) θα πρέπει να είναι αυτοσταθμιζόμενος.

Για να ισχύουν οι συνθήκες αυτές, το διάστημα δειγματοληψίας δ_{hi} υπολογίζεται ως εξής:

$$\frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \frac{M_{hi}}{m_{hi}} = \lambda \Rightarrow \frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \delta_{hi} = \lambda \Rightarrow \delta_{hi} = \frac{M_{hi}}{m_{hi}} = \lambda \cdot n_h \cdot P_{hi}$$

Αναφέρουμε ότι σε κάθε τρίμηνο ερευνώνται περίπου 2.500 νοικοκυριά.

Εκτίμηση των χαρακτηριστικών της έρευνας

Για την εκτίμηση των χαρακτηριστικών της έρευνας, τα στοιχεία κάθε νοικοκυριού του δείγματος πολλαπλασιάστηκαν με έναν αναγωγικό συντελεστή. Ο αναγωγικός συντελεστής προκύπτει ως το γινόμενο των ακόλουθων τριών παραγόντων (σταθμίσεων):

α. Η αντίστροφη πιθανότητα επιλογής του νοικοκυριού:

$$\frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \frac{M_{hi}}{m_{hi}} = DW_{hi}, \text{ όπου}$$

M_{hi} = ο αριθμός των νοικοκυριών στον ενημερωμένο κατάλογο–πλαίσιο της πρωτογενούς μονάδας επιφάνειας hi

m_{hi} = ο αριθμός των επιλεγμένων νοικοκυριών στην πρωτογενή μονάδα επιφάνειας hi

n_h = το μέγεθος του δείγματος των πρωτογενών μονάδων στο στρώμα h

P_h = η πιθανότητα επιλογής της πρωτογενούς μονάδας επιφάνειας hi

β. Το αντίστροφο του ποσοστού απόκρισης των νοικοκυριών του δείγματος εντός της πρωτογενούς μονάδας.

γ. Ένας διορθωτικός συντελεστής, ο οποίος καθορίζεται κατά τρόπο ώστε:

- Σε σύνολο χώρας, ο εκτιμημένος πληθυσμός κατά φύλο και ηλικία να συμπίπτει με τον πληθυσμό που υπολογίζεται από την προβολή των στατιστικών στοιχείων φυσικής κίνησης πληθυσμού της περιόδου αναφοράς (απογραφή πληθυσμού, γεννήσεις, θάνατοι, μετανάστευση).
- Σε σύνολο χώρας, τα εκτιμημένα νοικοκυριά κατά τάξη μεγέθους (1, 2, 3, 4, 5+) να συμπίπτουν με αυτά που προέκυψαν με προβολή στο έτος αναφοράς της έρευνας. Η προβολή αυτή πραγματοποιήθηκε λαμβάνοντας υπόψη τη διαχρονική τάση των νοικοκυριών κατά τη δεκαετία 1991 – 2001 και εφαρμόζοντας εκθετικό μοντέλο.