

## **Έρευνα Τεχνολογιών Πληροφόρησης και Επικοινωνίας**

### **ΕΙΔΟΣ**

Δειγματοληπτική έρευνα

Μεθοδολογία για τις Κοινοτικές στατιστικές σχετικά με την κοινωνία της πληροφορίας: Άτομα, νοικοκυριά και η κοινωνία της πληροφορίας – υπό το προσχέδιο του Κανονισμού της Επιτροπής εφαρμόζοντας τον Κανονισμό 808/2004 όπως συμφωνήθηκε από το Στατιστικό Πρόγραμμα της Επιτροπής.

### **Στάδια δειγματοληψίας**

Το δείγμα των νοικοκυριών για την έρευνα του έτους 2005 θα αποτελείται από το δείγμα της Έρευνας Εισοδήματος και Συνθηκών Διαβίωσης των Νοικοκυριών της Χώρας (EU – SILC) έτους 2004. Η EU – SILC είναι μια ετήσια έρευνα νοικοκυριών με εναλλασσόμενο δείγμα, που καλύπτει τον πληθυσμό στον οποίο στοχεύει η Έρευνα Τεχνολογιών, Πληροφόρησης και Επικοινωνίας (ICT).

Επιπροσθέτως, η EU – SILC είναι μια δισταδιακή στρωματοποιημένη δειγματοληπτική έρευνα με πρωτογενή μονάδα δειγματοληψίας την επιφάνεια (ένα ή περισσότερα συνεχόμενα οικοδομικά τετράγωνα) και τελική μονάδα το νοικοκυριό. Ο δειγματοληπτικός σχεδιασμός της EU – SILC του έτους 2004 βασίστηκε στα στοιχεία της Απογραφής Πληθυσμού του έτους 2001.

Ο λόγος χρησιμοποίησης του δείγματος της EU – SILC για τη συλλογή των στατιστικών στοιχείων της ICT ήταν η λειτουργική διευκόλυνση, το χαμηλό κόστος και η ευχέρεια δημιουργίας αντιπροσωπευτικού δείγματος με το απαιτούμενο μέγεθος.

### **Ερευνητική μέθοδος**

Η συλλογή των δεδομένων πραγματοποιείται με τηλεφωνικές συνεντεύξεις.

Η σύνθεση των μελών των νοικοκυριών είναι ήδη γνωστή από τα στατιστικά στοιχεία που συλλέχθηκαν στην έρευνα EU – SILC έτους 2004. Τα μέλη των νοικοκυριών ηλικίας 16 ως 74 ετών, τα οποία θα αποτελέσουν το δείγμα θα επιλεγούν τυχαία χρησιμοποιώντας τυχαίους αριθμούς.

Ο ερευνητής θα τηλεφωνήσει στο νοικοκυριό και θα ζητήσει να μιλήσει με το προεπιλεγμένο άτομο του δείγματος, έτσι ώστε να συλλέξει πληροφορίες για ολόκληρο το νοικοκυριό αλλά και τον ίδιο τον ερευνώμενο (στο υπόλοιπο ερωτηματολόγιο). Εάν η επικοινωνία με το άτομο του δείγματος είναι εφικτή, τότε ο ερευνητής θα συμπληρώσει το ερωτηματολόγιο. Στην περίπτωση που το μέλος του δείγματος απουσιάζει προσωρινά (στη δουλειά, σε κάποια επίσκεψη κλπ.), τότε ο ερευνητής θα πρέπει να προκαθορίσει κάποια άλλη χρονική στιγμή για τη συμπλήρωση του ερωτηματολογίου σε συνεργασία με το νοικοκυριό.

Αν το επιλεγμένο μέλος έχει αποχωρήσει οριστικά από το νοικοκυριό, τότε μπορεί να αντικατασταθεί από κάποιο άλλο μέλος ηλικίας 16 ως 74 ετών. Το ίδιο ισχύει και στην περίπτωση που το επιλεγμένο μέλος έχει καταταγεί στο στρατό ή λείπει για σπουδές.

### **Ερευνώμενος πληθυσμός**

Σύμφωνα με τη γενική οδηγία του σχεδίου της Eurostat για την έρευνα Τεχνολογιών και Πληροφόρησης στα νοικοκυριά, ο ερευνώμενος πληθυσμός είναι τα άτομα ηλικίας 16 ως 74 ετών που διαμένουν σε ιδιωτικά νοικοκυριά στην Ελλάδα. Ο πληθυσμός αυτός αποτελείται από 8.160.045 άτομα που διαμένουν σε 3.726.806 νοικοκυριά.

### **Δειγματοληπτική μονάδα**

Μονάδα έρευνας είναι το ιδιωτικό νοικοκυριό και ένα τυχαία επιλεγμένο άτομο από τα μέλη του, ηλικίας 16 ως 74 ετών.

### **Δειγματοληπτικό πλαίσιο**

Το δειγματοληπτικό πλαίσιο της έρευνας ICT είναι το ίδιο με της έρευνας EU – SILC έτους 2004. Πρωτογενής μονάδα έρευνας είναι η επιφάνεια (ένα ή περισσότερα συνεχόμενα οικοδομικά τετράγωνα), η οποία επιλέγεται χρησιμοποιώντας δεδομένα από την απογραφή του 2001.

Σε κάθε επιλεγμένη πρωτογενή μονάδα έρευνας, το δειγματοληπτικό πλαίσιο των νοικοκυριών ενημερώνεται ένα μήνα πριν από τη συλλογή των στοιχείων.

## Σχέδιο δειγματοληψίας

Στην έρευνα εφαρμόζεται η πολυσταδιακή στρωματοποιημένη δειγματοληψία επιφανειών. Πρωτογενής μονάδα έρευνας είναι η επιφάνεια (ένα ή περισσότερα συνεχόμενα οικοδομικά τετράγωνα) και δευτερογενής μονάδα έρευνας είναι τα νοικοκυριά με μέλη που ανήκουν στον ερευνώμενο πληθυσμό. Η τελική μονάδα έρευνας είναι ένα τυχαία επιλεγμένο μέλος του υπό έρευνα νοικοκυριού, ηλικίας 16 ως 74 ετών.

Η μονάδα επιφάνειας ορίζεται από τμήμα μιας κατοικημένης περιοχής που οριοθετείται από τεχνητά ή φυσικά σύνορα, τα οποία είναι σαφώς σχεδιασμένα και αναγνωρίσιμα όταν χρησιμοποιείται ένας χάρτης της περιοχής. Μια τέτοια μονάδα μπορεί να αποτελείται από ένα ή περισσότερα συνεχόμενα οικοδομικά τετράγωνα ή από μέρος μιας επαρχιακής περιοχής με σαφή σύνορα. Για τη μείωση του κόστους, αλλά και του χρόνου διεξαγωγής των εξωτερικών εργασιών της έρευνας, το μέγεθος των μονάδων επιφάνειας προσδιορίστηκε κατά προσέγγιση σε 50 και 70 νοικοκυριά στις 13 Περιφέρειες (NUTS II) και τις περιφέρειες των δύο μεγάλων πόλεων αντίστοιχα.

## Κριτήρια στρωμάτωσης και υπο – στρωμάτωσης του πληθυσμού

Πρώτο κριτήριο στρωμάτωσης του ερευνώμενου πληθυσμού είναι 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  $\geq 30.000$  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. 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 μέγεθος δείγματος είναι 5.500 νοικοκυριά που ανήκουν σε 1.056 πρωτογενείς μονάδες δειγματοληψίας.

## Διαδικασία επιλογής του δείγματος

Το δείγμα των ιδιωτικών νοικοκυριών και τα μέλη που πληρούν τις προϋποθέσεις της έρευνας επιλέγονται σε τρία στάδια. Συγκεκριμένα, οι πρωτογενείς μονάδες επιλέγονται συστηματικά από το πλαίσιο σε κάθε τελικό στρώμα (= Γεωγραφική Περιφέρεια x Αστικότητα) με πιθανότητα ανάλογη του αριθμού των ιδιωτικών νοικοκυριών στις επιφάνειες (σύμφωνα με τα αποτελέσματα της απογραφής του έτους 2001). Το δείγμα αποτελείται τελικά από 1.056 πρωτογενείς μονάδες οι οποίες περιλαμβάνουν κατά μέσο όρο 60 νοικοκυριά. Στο δεύτερο στάδιο επιλέγεται ένα συστηματικό τυχαίο δείγμα νοικοκυριών από τις πρωτογενείς μονάδες. Στο τρίτο στάδιο επιλέγεται τυχαία ένα μέλος ηλικίας 16 ως 74 ετών για κάθε νοικοκυριό που ανήκει στο δείγμα. Τα στάδια δειγματοληψίας περιγράφονται παρακάτω:

### 1<sup>ο</sup> στάδιο δειγματοληψίας

Η πρωτογενής μονάδα τάξεως  $i$  του στρώματος  $h$  έχει πιθανότητα επιλογής ανάλογη του μεγέθους του ερευνώμενου πληθυσμού, όπως περιγράφεται από την ακόλουθη σχέση:

$$P_{hi} = \frac{N_{hi}}{N_h}, \text{ όπου}$$

$N_{hi}$ : Ο αριθμός των νοικοκυριών (από την έρευνα EU-SILC) στην πρωτογενή μονάδα  $hi$

$N_h$ : Ο αριθμός των νοικοκυριών στο στρώμα  $h$

## 2<sup>ο</sup> στάδιο δειγματοληψίας

Από τα  $N_{hi}$  νοικοκυριά επιλέγεται με ίσες πιθανότητες ένα δείγμα  $n_{hi}$  νοικοκυριών. Καθένα από τα  $n_{hi}$  νοικοκυριά έχει την ίδια πιθανότητα να επιλεγεί και η πιθανότητα αυτή είναι ίση με  $\frac{n_{hi}}{N_{hi}}$ . Για να είναι ο εκτιμητής  $Y_h$  του συνόλου των στρωμάτων (για οποιοδήποτε χαρακτηριστικό) αυτοσταθμιζόμενος, το  $n_{hi}$  ορίζεται από την ακόλουθη σχέση:  $n_{hi} = \frac{n_h}{a_h}$ , όπου  $n_h = \sum_i n_{hi}$  και  $a_h$  είναι ο αριθμός των πρωτογενών μονάδων στο στρώμα  $h$ .

## 3<sup>ο</sup> στάδιο δειγματοληψίας

Από κάθε νοικοκυριό του δείγματος επιλέγεται ένα άτομο από τον ερευνητέο πληθυσμό με ίσες πιθανότητες, συγκεκριμένα:

$p_{hij}$ : Η πιθανότητα επιλογής του ατόμου  $hij$ , που ανήκει στο νοικοκυριό  $hi$ .

Δεδομένου ότι το άτομο επιλέχθηκε με ίσες πιθανότητες από τα  $m_{hi}$  μέλη που ανήκουν στον ερευνητέο πληθυσμό, η πιθανότητα  $p_{hij}$  ορίζεται από την ακόλουθη σχέση:

$$p_{hij} = \frac{1}{m_{hi}}$$

## Εκτίμηση των χαρακτηριστικών της έρευνας

Έστω  $h$  ένα από τα τελικά στρώματα νοικοκυριών (Τελικό στρώμα = Γεωγραφική Περιφέρεια Χ Αστικότητα), το οποίο λαμβάνει τις ακόλουθες τιμές  $h = 1, 2, \dots, H$  (όπου  $H = 90$ ). Σε κάθε τελικό στρώμα (έστω  $h$ ), αν η στατιστική πληροφορία επιλέχθηκε από ένα δείγμα  $n'_h$  νοικοκυριών, ο αναγωγικός συντελεστής του νοικοκυριού τάξεως  $i$  ορίζεται από την ακόλουθη σχέση:

$$w_{hi} = \frac{N_h}{n_h} \cdot \frac{1}{r_h} t_{hi} \quad \text{όπου:}$$

$N_h$ : Ο αριθμός των νοικοκυριών στο στρώμα  $h$

$n_h$ : Το αρχικό μέγεθος δείγματος νοικοκυριών στο στρώμα  $h$

$\frac{N_h}{n_h}$ : Η αντίστροφη αρχική πιθανότητα επιλογής του δείγματος νοικοκυριών στο στρώμα  $h$

$r_h = \frac{n'_h}{n_h}$ : Το ποσοστό απόκρισης των νοικοκυριών του δείγματος στο στρώμα  $h$

$t_{hi}$ : Διορθωτικός συντελεστής που καθορίζεται κατά τρόπο ώστε η εκτίμηση των νοικοκυριών κατά τάξη μεγέθους (1, 2, 3, 4 ή 5+ μέλη) να συμπίπτει με αυτή του έτους

αναφοράς των στοιχείων, που υπολογίστηκε από τα στοιχεία της έρευνας Εργατικού Δυναμικού του ίδιου έτους αναφοράς.

Σε κάθε τελικό στρώμα νοικοκυριών (έστω  $h$ ), αν η στατιστική πληροφορία επιλέχθηκε από ένα δείγμα  $m_h$  ατόμων, ο αναγωγικός συντελεστής του ατόμου τάξεως  $j$  που ανήκει στο νοικοκυριό  $hi$  ορίζεται από την ακόλουθη σχέση:

$$w_{hij} = w_{hi} \cdot \frac{1}{p_{hij}} \cdot g_{hij}, \text{ όπου :}$$

$w_{hi}$  : Ο αναγωγικός συντελεστής του νοικοκυριού  $hi$ , όπου ανήκει το άτομο  $hij$

$g_{hij}$  : Διορθωτικός συντελεστής, ο οποίος προσαρμόζει τις σταθμίσεις του δείγματος των ατόμων, έτσι ώστε η εκτίμηση της κατανομής των ατόμων να συμφωνεί με την αντίστοιχη κατανομή του πληθυσμού στο σύνολο των τάξεων. Οι τάξεις είναι 24 και ορίζονται από το φύλο και την ηλικιακή ομάδα (2 φύλα X 12 ηλικιακές ομάδες). Οι ηλικιακές ομάδες ορίζονται από τα χρονικά διαστήματα των ετών: 16-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69 και 70-74. Η κατανομή του πληθυσμού των ατόμων κατά φύλο και ηλικιακή ομάδα εκτιμήθηκε από στοιχεία της έρευνας Εργατικού Δυναμικού με την ίδια περίοδο αναφοράς των στοιχείων.