ΑΝΩΤΑΤΟ ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΑΤΡΩΝ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΔΕΙΓΜΑΤΟΛΗΨΙΑ : ΤΕΧΝΙΚΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΤΗΣ

ΣΠΟΥΔΑΣΤΡΙΕΣ
ΜΗΤΣΑΚΑΚΗ ΣΤΕΛΛΑ
ΠΑΘΙΑΚΗ ΜΑΡΙΑ

ΕΠΟΠΤΕΥΩΝ ΚΑΘΗΓΗΤΗΣ
ΤΣΑΓΚΑΝΟΣ ΑΘΑΝΑΣΙΟΣ

ΠΑΤΡΑ 2008

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ
ΕΙΣΑΓΩΓΗ

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ
ΑΠΛΗ ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ
ΣΥΣΤΗΜΑΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ
ΣΤΡΩΜΑΤΟΠΟΙΗΜΕΝΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ
ΚΑΤΑ ΟΜΑΔΕΣ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

ΚΕΦΑΛΑΙΟ ΈΚΤΟ
ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΜΕΓΕΘΟΥΣ ΔΕΙΓΑΜΤΟΣ

ΚΕΦΑΛΑΙΟ ΈΒΔΟΜΟ
ΈΡΕΥΝΑ- ΜΕΛΕΤΗ ΠΛΗΘΥΣΜΟΥ

ΠΑΡΑΡΤΗΜΑ

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΕΙΣΑΓΩΓΗ

1.1. Οι Κυριότερες Μέθοδοι Συλλογής Δεδομένων

Συχνά βλέπουμε όλοι μας στον ημερήσιο και περιοδικό τύπο τα αποτελέσματα ερευνών που αναφέρονται σε οικονομικά, κοινωνικά και πολιτικά φαινόμενα, όπως για παράδειγμα το μέσο μηνιαίο εισόδημα των εργαζομένων σε διάφορες χώρες, την επίδραση της ανεργίας στο κοινωνικό σύνολο, τις προτιμήσεις των ψηφοφόρων για τους υποψήφιους και άλλα.

Για τη διεξαγωγή μιας έρευνας απαιτείται η συλλογή δεδομένων τα οποία αναφέρονται σε ορισμένες χαρακτηριστικές ιδιότητες του προς μελέτη πληθυσμού. Για τη συλλογή αυτών των στοιχείων εφαρμόζονται διάφορες τεχνικές συλλογής δεδομένων, οι κυριότερες από τις οποίες είναι η *απογραφή* και η *δειγματοληψία*. (βλ. Κιόχος Α.Πέτρος, 1993)

Η απογραφή συνίσταται στη συγκέντρωση στοιχείων τα οποία αναφέρονται σε ορισμένα χαρακτηριστικά γνωρίσματα από όλα τα άτομα του προς μελέτη πληθυσμού. Γενικότερα οι απογραφές πραγματοποιούνται σε πανελλήνιο επίπεδο μία φορά κάθε δέκα χρόνια. Σημειώνεται ότι η γενική απογραφή που διενήργησε η Γ.Γ. ΕΣΥΕ (Γενική Γραμματεία Εθνικής Στατιστικής Υπηρεσίας Ελλάδος) το 2001 προετοιμαζόταν από τα μέλη της για πολλά χρόνια και στοίχισε περίπου 16 δις δρχ, πράγμα που σημαίνει οτι η απογραφή απαιτεί αρκετά χρήματα, προετοιμασία και ανθρώπινο δυναμικό. Επομένως , η απογραφή είναι πρακτικά μια ασύμφορη ,άν όχι αδύνατη τεχνική. Για αυτό το λόγο συχνά καταφεύγουμε σε εναλλακτικές τεχνικές.

Στις περιπτώσεις αυτές εφαρμάζουμε δειγματοληπτικές έρευνες, οι οποίες καλύπτουν μέρος του ερευνώμενου πληθυσμού, δηλαδή του συνόλου των ατόμων ή αντικειμένων που μας ενδιαφέρει να μελετήσουμε. Μέσω της δειγματοληψίας επιλέγεται ενα υποσύνολο του πληθυσμού, το οποίο καλείται *δείγμα*. Το δείγμα επιλέγεται με τέτοιο τρόπο, ώστε να περιέχει όσο το δυνατόν περισσότερα χαρακτηριστικά του πληθυσμού στον οποίο ανήκει. Αυτό σημαίνει ότι οι πληροφορίες, οι εκτημίσεις και τα συμπεράσματα που θα βγάλουμε από το δείγμα μπορούν να γενικευτούν στον πληθυσμό.

Τα βασικά πλεονεκτήματα της δειγματοληψίας είναι:

- Εξοικονόμηση χρόνου στη συλλογή και επεξεργασία των στοιχείων.
- Έχει μεγαλύτερη ταχύτητα διεξαγωγής με συνέπεια τα αποτελέσματα της να αποδίδουν την τρέχουσα κατάσταση.
- Σε ορισμένες περιπτώσεις είναι η μοναδική μέθοδος έρευνας (π.χ. σε ποιοτικό έλεγχο προϊόντων , όπου απαιτείται καταστροφή των υλικών).
- Χαμηλό κόστος.

  Είναι λογικό να απαιτούνται λιγότερες χρηματικές μονάδες όταν εξετάζεται ένα υποσύνολο του πληθυσμού παρά όταν εξετάζουμε ολόκληρο τον πληθυσμό.

1.2. Στάδια της Δειγματοληψίας

Η διαδικασία δειγματοληψίας μιας έρευνας περιλαμβάνει τα ακόλουθα στάδια, τα οποία αποτελούν βασική προϋπόθεση για την επιτυχία της:

- Προσδιορισμός του υπό μελέτη πληθυσμού
- Εξασφάλιση δειγματοληπτικού πλαισίου
- Επιλογή μεγέθους δείγματος
- Επιλογή κατάλληλης μεθόδου δειγματοληψίας

1.2.1. Προσδιορισμός του υπό μελέτη Πληθυσμού

Το πρώτο πράγμα που πρέπει να γίνει σε μια σχεδιαζόμενη δειγματοληψία είναι να ορισθούν ακριβώς όλες οι μονάδες που περιέχονται στον πληθυσμό.

Εάν η τιμή του χαρακτηριστικού που μετράμε παίρνει πολύ μικρές και πολύ μεγάλες τιμές, θα πρέπει οι ακραίες περιπτώσεις ν'απομονωθούν και να μελετηθούν χωριστά γιατί αλλιώς τα συμπεράσματα μας θα είναι εντελώς λανθασμένα.(βλ. Δαμιανός Χ.,2006)

1.2.2. Εξασφάλιση Δειγματοληπτικού Πλαισίου

Το πλαίσιο δειγματοληψίας είναι ένας κατάλογος που περιέχει όλες τις μονάδες που έχουν τη χαρακτηριστική ιδιότητα που θέλουμε να μελετήσουμε, από τον οποίο με την εφαρμογή διαφόρων μεθόδων θα επιλέξουμε το επιθυμητό δείγμα.

Ως πλαίσια δειγματοληψίας χρησιμοποιούνται, συνήθως, αρχεία που διατηρούνται για διοικητικούς σκοπούς (μητρώα αρρένων, πελατολόγιο μιας εταιρείας κ.λπ) ή μητρώα που προκύπτουν από τις γενικές απογραφές πληθυσμού.

Δεν είναι πάντα εφικτό να υπάρχει ένας τέτοιος κατάλογος όπως για παράδειγμα στην περίπτωση κατά την οποία έχουμε ποιοτικό έλεγχο. Σε αυτή την περίπτωση ο στόχος μας είναι να τον κατασκευάσουμε. Έστω οτι μελετάμε τις αφίξεις ασθενών σε νοσοκομείο που έχουν ενα συγκεκριμένο πρόβλημα υγείας, το δειγματοληπτικό μας πλαίσιο είναι το βιβλίο εισερχομένων ασθενών όπου αναγράφονται αναλυτικά τα στοιχεία του ασθενούς καθώς και τα συμπτώματα που παρουσιάζει.

Η εξασφάλιση ενός κατάλληλου καταλόγου-πλαισίου, είναι το πρώτο βήμα στη διαδικασία επιλογής δείγματος. Γι'αυτό και ο προσδιορισμός του πλαισίου παίζει σημαντικό ρόλο στο σχεδιασμό μιας δειγματοληπτικής έρευνας. Επηρεάζει τη συλλογή της πληροφορίας και επιδρά στην αποτελεσματικότητα με την οποία λαμβάνεται ένα δείγμα. Ένας κατάλογος-πλαίσιο θεωρείται ιδανικός όταν είναι πλήρης, δεν υπάρχουν παραληψείς ή διπλές καταχωρήσεις και έχουν αφαιρεθεί οι μονάδες που δεν ανήκουν πλέον στον πληθυσμό της έρευνας. Διαφορετικά, τα συμπεράσματα που θα γενικευθούν επαγωγικά στο σύνολο του ερευνώμενου πληθυσμού δεν θα είναι αξιόπιστα.

1.2.3. Προσδιορισμός Μεγέθους Δείγματος

Κατά το σχεδιασμό μιας δειγματοληπτικής έρευνας, ο προσδιορισμός του μεγέθους του δείγματος αποτελεί μια από τις σπουδαιότερες αποφάσεις. Πολύ μεγάλο δείγμα συνεπάγεται απώλεια χρόνου και χρήματος, ενώ πολύ μικρό δείγμα συνεπάγεται αναξιόπιστα αποτελέσματα.

Γενικά μπορούμε να πούμε οτι ο προσδιορισμός του μεγέθους του δείγματος καθορίζεται απο τους εξής παράγοντες:

Α) τον επιθυμητό βαθμό ακρίβειας των αποτελεσμάτων μιας έρευνας. Η επάρκεια δείγματος εξαρτάται κυρίως από τον επιθυμητό βαθμό ακρίβειας, με άλλα λόγια τι περιθώρια σφάλματος πρόκειται να δεχθούμε σε μια δειγματοληπτική εκτίμηση. Ο βαθμός ακρίβειας της εκτίμησης είναι το τυπικό σφάλμα το οποίο ισούται με $\frac{s}{\sqrt{n}}$ όπου s η τυπική απόκλιση του χαρακτηριστικού στον πληθυσμό και n το μέγεθος του δείγματος, παραδείγματος χάριν, στην περίπτωση εκτίμησης του μέσου αριθμητικού ενός χαρακτηριστικού, το τυπικό σφάλμα εκφράζει τη διασπορά του πλήθους όλων των δυνατών εκτιμήσεων $\bar{x}$ από το πραγματικό μέσο του πληθυσμού m. Όσο μικρότερο είναι το τυπικό σφάλμα, τόσο μικρότερο θα είναι το δειγματοληπτικό σφάλμα $|\bar{x}-m|$ και άρα πιο ακριβής η εκτίμηση του μέσου αριθμητικού.

Επίσης το τυπικό σφάλμα εξαρτάται από το μέγεθος του δείγματος και από τη διακύμανση του χαρακτηριστικού στον ερευνώμενο πληθυσμό. Όσο μικρότερη είναι η διακύμανση ενός πληθυσμού, τόσο μικρότερο δείγμα απαιτείται για κάποια προκαθορισμένη ακρίβεια.

Η διακύμανση του ερευνώμενου πληθυσμού είναι συνήθως άγνωστη. Στην περίπτωση αυτή καταφεύγουμε σε αποτελέσματα προηγούμενης έρευνας ή στη διεξαγωγή μίας πιλοτικής έρευνας.

Β) τον αριθμό των μεταβλητών που πρόκειται να ερευνηθούν. Μια άλλη δυσκολία που αντιμετωπίζεται κατά τον προσδιορισμό του μεγέθους του δείγματος, είναι ότι στην πράξη συγκεντρώνονται πληροφορίες για περισσότερες από μια μεταβλητές (χαρακτηριστικά) ταυτόχρονα. Έτσι ένα δείγμα το οποίο μπορεί να είναι ικανοποιητικό για μια μεταβλητή, ενδέχεται να μην είναι για τις υπόλοιπες. Ένας τρόπος προσδιορισμού του μεγέθους του δείγματος είναι να καθοριστούν περιθώρια σφάλματος για τις μεταβλητές που θεωρούνται ως πιο σπουδαίες για την έρευνα. Στην συνέχεια προσδιορίζεται το μέγεθος του δείγματος που χρειάζεται, ξεχωριστά για κάθε μία από τις σπουδαιότερες μεταβλητές. Αν τα μεγέθη (n) δεν διαφέρουν σημαντικά μεταξύ τους και το μεγαλύτερο από αυτά περιέχεται εντός του καθορισμένου περιθωρίου σφάλματος, επιλέγεται ως μέγεθος δείγματος. Συνήθως τα μεγέθη (n) των μεταβλητών διαφέρουν μεταξύ τους οπότε επιλέγεται το μεγαλύτερο (n) γιατί δίνει τη μεγαλύτερη ακρίβεια.

Γ) τα διαθέσιμα οικονομικά και τα χρονικά περιθώρια. Μερικές φορές τα μεγέθη (n) που απαιτούνται για διάφορες μεταβλητές είναι τόσο ασύμφωνα μεταξύ τους, ώστε μερικά από αυτά πρέπει να παραλειφθούν. Στην πράξη το επιθυμητό μέγεθος είναι απραγματοποίητο λόγω περιορισμών στα οικονομικά μέσα και στον απιτούμενο χρόνο. Σε τέτοιες περιπτώσεις είναι καλύτερα να πάρει κανείς το μεγαλύτερο δυνατό δείγμα, σε σχέση με τα διαθέσιμα οικονομικά και να απορρίψει ερωτήματα για τα οποία χρειάζεται μεγαλύτερο δείγμα, ώστε να επιτευχθούν χρήσιμα αποτελέσματα από τη δειγματοληπτική έρευνα. (βλ. Δαμιανός Χ.,2006)

1.2.4. Επιλογή Μεθόδου Δειγματοληψίας

Οι δειγματοληπτικές έρευνες διακρίνονται σε δύο κατηγορίες:

1. Την τυχαία δειγματοληψία ή δειγματοληψία με πιθανότητα, όπου κάθε μονάδα του πληθυσμού έχει κάποια πιθανότητα να επιλεγεί στο δείγμα.
2. Τη μη τυχαία δειγματοληψία ή δειγματοληψία χωρίς πιθανότητα, όπου η επιλογή των δειγματοληπτικών μονάδων γίνεται κατά τρόπο μη τυχαίο.

Λόγω του ότι στη μη τυχαία δειγματοληψία η επιλογή του δείγματος γίνεται με ντετερμινιστικό τρόπο, τα αποτελέσματά της δεν είναι αντιπροσωπευτικά και άρα είναι αναξιόπιστα. Γι’αυτό έχει καθιερωθεί στην πράξη η εφαρμογή της τυχαίας δειγματοληψίας ή δειγματοληψίας με πιθανότητα, η οποία δίνει περισσότερο αντικειμενικά αποτελέσματα καθώς και τη δυνατότητα μέτρησης του δειγματοληπτικού σφάλματος.

Υπάρχουν πάρα πολλές τεχνικές τυχαίας δειγματοληψίας. Οι πιο συνηθισμένες απο αυτές είναι οι ακόλουθες:

- Απλή τυχαία δειγματοληψία
- Συστηματική δειγματοληψία
- Στρωματοποιημένη δειγματοληψία
- Κατά Ομάδες δειγματοληψία

Προτού αναλύσουμε διαξοδικά τις παραπάνω τεχνικές, είναι απαραίτητο να αναφερθούμε στα σφάλματα δειγματοληψίας, τα οποία επηρεάζουν την ακρίβεια των αποτελεσμάτων μιας δειγματοληπτικής έρευνας, τα οποία ταξινομούνται σε δύο βασικές κατηγορίες:

1.3. Δειγματοληπτικά σφάλματα

Μια καλή δειγματοληπτική έρευνα πρέπει να περιλαμβάνει των υπολογισμό των δειγματοληπτικών σφαλμάτων. Το δειγματοληπτικό σφάλμα, προκύπτει στην περίπτωση που δεν ερευνάται ολόκληρος ο πληθυσμός, αλλά ένα δείγμα αυτού. Οφείλεται στις τυχαίες κυμάνσεις της δειγματοληψίας, δηλαδή στο γεγονός ότι το δείγμα, δεν είναι δυνατόν να είναι απολύτως αντιπροσωπευτικό του πληθυσμού από τον οποίο προέρχεται και επομένως οι μετρήσεις που προκύπτουν από το δείγμα θα αποκλίνουν από τις αντίστοιχες πραγματικές του πληθυσμού.

Το δειγματοληπτικό σφάλμα γίνεται μικρότερο όταν αυξάνει το μέγεθος του δείγματος. Γίνεται επομένως σαφές ότι η ύπαρξη του δειγματοληπτικού σφάλματος είναι εγγενής της διαδικασίας της δειγματοληψίας και ενώ είναι δυνατό να μειωθεί, δεν δύναται να μηδενιστεί παρά μόνο αν η έρευνα είναι απογραφική.(βλ. Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής, 2000)

Θα πρέπει να γίνεται σωστή επιλογή της μεθόδου δειγματοληψίας και κατάλληλου δείγματος, ώστε τα δειγματοληπτικά σφάλματα να είναι μειωμένα.

1.4. Μη δειγματοληπτικά σφάλματα

Ο όρος «μη δειγματοληπτικά σφάλματα» προέρχεται από το γεγονός ότι αυτός ο τύπος σφάλματος μπορεί να συμβεί σε οποιαδήποτε έρευνα, είτε αυτή γίνεται με τη μέθοδο της απογραφής, είτε με τη μέθοδο της δειγματοληψίας.

Τα μη δειγματοληπτικά σφάλματα οφείλονται συνήθως σε:

- Ατέλειες του σχεδιασμού και της οργάνωσης της έρευνας για λόγους ταχύτητας και περιορισμού του κόστους
- Λάθη των ερευνητών, κωδικογράφων κ.λπ.
- ‘Αλλες αντικειμενικές δυσκολίες που ανακύπτουν κατά την εκτέλεση μιας στατιστικής έρευνας

Τα μη δειγματοληπτικά σφάλματα περιορίζουν την ποιότητα των συλλεγόμενων στοιχείων και επομένως την αξία των αποτελεσμάτων που προκύπτουν από αυτά.

1.5. Συμβολισμοί

N : αριθμός μονάδων πληθυσμού ή ανάλογα ο αριθμός των ομάδων στον πληθυσμό

n : αριθμός μονάδων δείγματος ή ανάλογα ο αριθμός των ομάδων που έχουν επιλεγεί στο δείγμα

x_i : η τιμή του υπό μελέτη χαρακτηρικτηριστικού της μονάδας i

$X = \sum_{i=1}^{N} X_i$: Άθροισμα των τιμών του χαρακτηριστικού X στον πληθυσμό

$x = \sum_{i=1}^{n} x_i$: Άθροισμα των τιμών του χαρακτηριστικού x στο δείγμα

m : η μέση τιμή του πληθυσμού

$\bar{x}$: η μέση τιμή του δείγματος

s^2 : η διακύμανση του πληθυσμού

s^2 : η διακύμανση του δείγματος

$V(\bar{x})$: η διακύμανση της δειγματικής μέσης τιμής

$f = \frac{n}{N}$: κλάσμα δειγματοληψίας

s : τυπικό σφάλμα ή τυπική απόκλιση του πληθυσμού

S : τυπικό σφάλμα ή τυπική απόκλιση του δείγματος

SE : η εκτίμηση του τυπικού σφάλματος

$W_i = \frac{N_i}{N}$: στάθμιση στρώματος i

A : ο αριθμός των μονάδων του πληθυσμού με μια συγκεκριμένη ιδιότητα

P : ποσοστό ενός χαρακτηριστικού στον πληθυσμό

a : ο αριθμός των μονάδων του δείγματος με την ιδιότητα αυτή

p : ποσοστό ενός χαρακτηριστικού στο δείγμα

Q : ποσοστό στον πληθυσμό που δεν παρουσιάζει την ιδιότητα ($Q = 1 - P$)

q : ποσοστό στο δείγμα που δεν παρουσιάζει την ιδιότητα ($q = 1 - p$)

l : βήμα ή διάστημα δειγματοληψίας(το αντίστροφο του κλάσματος δειγματοληψίας)

s_w^2 : διακύμανση μεταξύ των μονάδων που βρίσκονται στο ίδιο συστηματικό δείγμα

k : συνολικός αριθμός των στρωμάτων

N_i : μέγεθος του στρώματος i

$\overline{X}_i$: ο μέσος όρος των N_i μονάδων στο στρώμα i (μέσος στρώματος)

n_i : μέγεθος του δείγματος για το στρώμα i

x_{ij} : η τιμή της μεταβλητής X της j μονάδας ($j = 1,2,\ldots,N_i$) του συστηματικού δείγματος ή στρώματος i

M : ο αριθμός των στοιχείων στον πληθυσμό στην κατά Ομάδες Δειγματοληψία

m_i : ο αριθμός των στοιχείων στην ομάδα i

$\overline{M} = \frac{N}{M}$ το μέσο μέγεθος ομάδων στον πληθυσμό

a_i: ο αριθμός των στοιχείων στην ομάδα i που έχουν το συγκεκριμένο χαρακτηριστικό που μας ενδιαφέρει.

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

ΑΠΛΗ ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

2.1. Εισαγωγή

Απλή τυχαία δειγματοληψία ονομάζεται η μέθοδος επιλογής n μονάδων από έναν πληθυσμό N μονάδων, με τέτοιο τρόπο ώστε κάθε δυνατό δείγμα μεγέθους n έχει την ίδια πιθανότητα να επιλεγεί, δηλαδή όλες οι μονάδες του πληθυσμού να έχουν την ίδια πιθανότητα να συμπεριληφθούν στο δείγμα. (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997). Η απλή τυχαία δειγματοληψία είναι μια αποτελεσματική διαδικασία όταν ο πληθυσμός δεν είναι μεγάλος και όταν είναι σχετικά εύκολο και όχι δαπανηρό να καταγράψουμε και να απαριθμήσουμε τις δειγματοληπτικές μονάδες.

Προκειμένου να επιτευχθεί το παραπάνω, πρέπει να επιλέξουμε τυχαία το δείγμα από τον ερευνώμενο πληθυσμό. Η τυχαία αυτή επιλογή εξασφαλίζεται με δύο τρόπους:

i. Μέθοδος της Κλήρωσης

Η επιλογή των ατόμων του δείγματος γίνεται με κλήρωση. Για παράδειγμα, γράφονται τα ονόματα των ατόμων που αποτελούν τον ερευνώμενο πληθυσμό σε χαρτάκια ή κάρτες, τοποθετούνται σε ένα δοχείο και στη συνέχεια εξάγεται το επιθυμητό μέγεθος δείγματος. Όταν οι κάρτες που εξάγονται από το δοχείο κλήρωσης δεν επανατοποθετούνται, έχουμε δειγματοληψία χωρίς επανατοποθέτηση, συνήθως γίνεται χρήση της όταν έχουμε δειγματοληψία ενός σταδίου, διαλέγουμε δηλάδή τα στοιχεία απευθείας μέσα απο τον πληθυσμό. Εάν στο δείγμα χρειάζεται να περιληφθεί κάθε στατιστική μονάδα του πληθυσμού μέχρι n φορές, τότε επανατοποθετούμε κάθε φορά την κάρτα που εξάγεται πριν προχωρήσουμε στην επόμενη εξαγωγή κάρτας. Στην περίπτωση αυτή έχουμε δειγματοληψία με επανατοποθέτηση. Η κλήρωση με επανατοποθέτηση χρησιμοποιείται συνήθως στις πολυσταδιακές δειγματοληψίες όπου ο πληθυσμός είναι αρκετά μεγάλος ή άπειρος. (βλ. Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής , 2000)

ii. Μέθοδος Τυχαίων Αριθμών

Οι πίνακες τυχαίων αριθμών μας δίνουν τη δυνατότητα να επιλέξουμε μια σειρά τυχαίων αριθμών, πράγμα που σημαίνει ότι όλοι οι πολυψήφιοι (μονοψήφιοι, διψήφιοι κλπ.) αριθμοί έχουν την ίδια πιθανότητα να επιλεγούν. Η επιλογή γίνεται ως εξής:

Έστω οτι επιθυμούμε να επιλέξουμε ένα δείγμα n απο έναν πληθυσμό με N μονάδες. Αρχικά αριθμούμε τον πληθυσμό απο 1 μεχρι N, στη συνέχεια θα επιλέξουμε n αριθμούς οι οποίοι θα είναι μικρότεροι ή ίσοι απο το N. Η επιλογή αρχίζει απο κάποιον αριθμό στήλης ή γραμμής και προχωράμε είτε οριζόντια είτε κατακόρυφα μέχρι να επιλέξουμε n αριθμούς. Αριθμοί μεγαλύτεροι απο το N ή αριθμοί που έχουν επιλεγεί μια φορά παραλείπονται (βλ. Παράρτημα).

2.2. Σχέση Παραμέτρου – Εκτίμησης

Αυτό που μας ενδιαφέρει να επιτύχουμε από μια δειγματοληπτική έρευνα, δηλαδή από τη μελέτη ενός δείγματος, είναι η εξαγωγή συμπερασμάτων για το σύνολο του πληθυσμού. Συνήθως, τα χαρακτηριστικά του πληθυσμού που μελετάμε τα εκφράζουμε με τη βοήθεια ποσοστών, μέσων όρων, αθροισμάτων κλπ. Τις πραγματικές τιμές που έχουν τα μεγέθη αυτά μπορούμε να τις βρούμε μόνο αν κάνουμε απογραφή του πληθυσμού. Τις πραγματικές τιμές του πληθυσμού που μελετάμε τις ονομάζουμε παραμέτρους του πληθυσμού.

Τις τιμές των χαρακτηριστικών που υπολογίζουμε από το δείγμα τις ονομάζουμε εκτιμήσεις. Όλο το πρόβλημα της δειγματοληψίας έγκειται στο πώς θα πάρουμε εκτιμητές για τις παραμέτρους του πληθυσμού που μας ενδιαφέρουν με μια προκαθορισμένη ακρίβεια (βλ. Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής, 2000).

2.3. Εκτίμηση Μέσου Πληθυσμού

Η εκτίμηση της δειγματικής μέσης τιμής είναι:

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Η διακύμανση της δειγματικής μέσης τιμής ισούται με:

$$V(\bar{x}) = E(\bar{x} - m)^2 = \frac{s^2}{n} \frac{(N-n)}{N}$$

Η εκτίμηση της τυπικής απόκλισης ή τυπικού σφάλματος είναι:

$$SE(\bar{x}) = \sqrt{V(\bar{x})} = \sqrt{\frac{s^2}{n} \frac{(N-n)}{N}} = \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N}}$$

Στην πράξη όμως η διακύμανση του πληθυσμού s^2 είναι άγνωστη, οπότε για την εκτίμηση των τυπικών σφαλμάτων χρησιμοποιείται η δειγματική διακύμανση $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$, η οποία προκύπτει από τα στοιχεία του δείγματος και είναι αμερόληπτη εκτίμηση της s^2.

Επομένως στην περίπτωση της εκτίμησης του τυπικού σφάλματος του μέσου του δείγματος θα έχουμε:

εκτίμηση διακύμανσης

$$V(\bar{x}) = \frac{s^2}{n}\left(\frac{N-n}{N}\right)$$

και εκτίμηση τυπικού σφάλματος

$$SE(\bar{x}) = \sqrt{V(\bar{x})} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}}.$$

Ο συντελεστής $\frac{N-n}{N}$ ονομάζεται διόρθωση πεπερασμένου πληθυσμού. Ο συντελεστής αυτός μπορεί να παραληφθεί όταν το κλάσμα δειγματοληψίας $\frac{n}{N}$ είναι πολύ μικρό, οπότε ο συντελεστής $\frac{N-n}{N}=1-\frac{n}{N}$ πλησιάζει τη μονάδα (βλ.Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997). Αυτό συμβαίνει όταν ο πληθυσμός είναι αρκετά μεγάλος ή όταν πρόκειται για άπειρους πληθυσμούς (δηλαδή σε περίπτωση δειγματοληψίας με επανατοποθέτηση).

Στην πράξη η διόρθωση πεπερασμένου πληθυσμού παραλείπεται όταν το κλάσμα δειγματοληψίας δεν ξεπερνά το 5%.

Κατά την περιγραφή της μεταβλητικότητας ενός δείγματος, ένα μέτρο που συχνά χρησιμοποιείται είναι ο συντελεστής μεταβλητικότητας. Η τυπική απόκλιση εκφράζεται σαν ένα κλάσμα, ή μερικές φορές σαν ένα ποσοστό του μέσου. Η χρησιμότητα αυτού του μέτρου βρίσκεται εν μέρει στο γεγονός ότι σε πολλές περιπτώσεις μεταβλητών ο μέσος και η τυπική απόκλιση τείνουν να μεταβάλλονται μαζί (βλ. Βασίλειος Κιμ. Μπένος, 1991) Συμβολίζοντας το συντελεστή μεταβλητικότητας με CV, για την εκτίμηση του $\bar{x}$ έχουμε:

$$CV(\bar{x})=\frac{s}{\bar{x}}$$

Ο συντελεστής μεταβλητικότητας συνήθως εκφράζεται σε ποσοστό, δηλαδή:

$$CV(\bar{x})=\frac{s}{\bar{x}}100\%$$

2.4. Εκτίμηση Ποσοστού ή Αναλογίας Πληθυσμού

Σε μία δειγματοληπτική έρευνα πολλές φορές μας ενδιαφέρει να εκτιμήσουμε το ποσοστό του πληθυσμού με μία συγκεκριμένη ιδιότητα. Για παράδειγμα, ένας καθηγητής επιθυμεί να εκτιμήσει το ποσοστό των φοιτητών που συνεχίζουν για μεταπτυχιακές σπουδές ή ο Οργανισμός Απασχόλησης Εργατικού Δυναμικού ενδιαφέρεται να μάθει το το ποσοστό ανέργων μιας πόλης κ.ο.κ.

Στη γενική περίπτωση συμβολίζουμε με x_i την τιμή του χαρακτηριστικού στην i δειγματοληπτική μονάδα . Επομένως για κατηγορικά (διχοτομικά) χαρακτηριστικά έχουμε:

$x_i=1$, αν η μονάδα i έχει τη συγκεκριμένη ιδιότητα

$x_i = 0$, διαφορετικά.

Έχουμε τον αντίστοιχο συμβολισμό:

$A \equiv X = \sum_{i=1}^{N} x_i =$ αριθμός του πληθυσμού με τη συγκεκριμένη ιδιότητα

$a \equiv x = \sum_{i=1}^{n} x_i =$ αριθμός των μονάδων του δείγματος με την ιδιότητα αυτή

$m = \frac{A}{N} = P =$ ποσοστό στον πληθυσμό με την ιδιότητα αυτή

$\bar{x} = \frac{a}{n} = p$ ποσοστό στο δείγμα με την ιδιότητα αυτή.

Επειδή κάθε x_i λαμβάνει τιμές 1 ή 0, το x_i^2 λαμβάνει επίσης τις τιμές 1 ή 0, επομένως

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i^2 = A = NP$$

και αντίστοιχα για το δείγμα

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i^2 = a = np.$$

Η διακύμανση s^2 του πληθυσμού παίρνει τη μορφή

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - m)^2}{N} = \frac{\sum_{i=1}^{N} x_i^2 - 2m\sum_{i=1}^{N} x_i + Nm^2}{N} = \frac{NP - 2p \cdot NP + NP^2}{N} =$$

$$= \frac{NP - NP^2}{N} = P - P^2 = P(1-P) = PQ$$

όπου $Q = 1 - P$

Όμοια η δειγματική διακύμανση γίνεται:

$$s^2 = \frac{n}{n-1} pq, \quad q = 1 - p$$

Με τη μέθοδο της απλής τυχαίας δειγματοληψίας χρησιμοποιείται σαν εκτιμητής του πληθυσμιακού ποσοστού P το δειγματικό ποσοστό p για το οποίο ισχύουν τα παρακάτω:

- Το δειγματικό ποσοστό $p=\frac{a}{n}$ είναι αμερόληπτος εκτιμητής του πληθυσμιακού ποσοστού P, δηλαδή

$$E(p)=P.$$

- Η διακύμανση του P είναι

$$s^{2}=V(p)=\frac{s^{2}}{n}(1-f)=\frac{PQ}{n}\left(\frac{N-n}{N-1}\right)$$

- Μία αμερόληπτη εκτίμηση της δισποράς s^{2} είναι

$$s^{2}(p)=\frac{pq}{n-1}(1-f)$$

Η διακύμανση της εκτίμησης του συνολικού αριθμού $(\hat{A}=Np)$ των μονάδων που ανήκουν στην τάξη με τη συγκεκριμένη χαρακτηριστική ιδιότητα που μας ενδιαφέρει είναι :

$$V(\hat{A})=V(Np)=\frac{N^{2}s^{2}}{n}\left(\frac{N-n}{N}\right)=\frac{N^{2}PQ}{n}\left(\frac{N-n}{N}\right)$$

ή

$$V(\hat{A})=V(Np)=\frac{N^{2}s^{2}}{n}\left(\frac{N-n}{N}\right)=\frac{N^{2}PQ}{n}\left(\frac{N-n}{N-1}\right)$$ για N πολύ μεγάλο

Αντίστοιχα η εκτίμηση του τυπικού σφάλματος είναι:

$$SE(Np)=\sqrt{V(Np)}$$

Επειδή στην πράξη η διακύμανση του πληθυσμού είναι συνήθως άγνωστη χρησιμοποιούμε την αμερόληπτη εκτίμηση της διακύμανσης του p που προκύπτει από το δείγμα. Έτσι η εκτίμηση της διακύμανσης του $\hat{A}$ θα είναι:

$$V(\hat{A})=V(Np)=\frac{N^{2}s^{2}}{n}\cdot\frac{N-n}{N}=\frac{pq}{n-1}N(N-n)$$

(βλ. Δαμιανός Χ.,2006)

2.5. Εκτίμηση Μέσων Μεγεθών Υποπληθυσμών

Σε πολλές έρευνες οι εκτιμήσεις γίνονται κατά κατηγορία ενός πληθυσμού (υποπληθυσμοί). Για παράδειγμα σε μια έρευνα νοικοκυριών μπορεί να εκτιμηθούν τα νοικοκυριά κατά μέγεθος, κατά εισόδημα, κατά επάγγελμα του υπευθύνου του νοικοκυριού κ.λπ. Επίσης πολλές φορές στα διαθέσιμα δειγματοληπτικά πλαίσια περιέχονται μονάδες οι οποίες δεν ανήκουν στις υπό μελέτη μονάδες και δεν είναι εύκολο να γίνει διαχωρισμός, παρά μόνο μετά την επιλογή του δείγματος. Για παράδειγμα, έστω οτι έχουμε στη διάθεσή μας ένα δειγματοληπτικό πλαίσιο (κατάλογο) των νοικοκυριών μιας πόλης και θέλουμε να συγκεντρώσουμε πληροφορίες από νοικοκυριά στα οποία υπάρχουν άτομα ηλικίας άνω των 65 ετών. Στην περίπτωση αυτή δεν μπορούμε εκ των προτέρων να διακρίνουμε τα νοικοκυριά που έχουν άτομα ηλικίας πάνω των 65 ετών. Για το λόγο αυτό παίρνουμε ένα τυχαίο δείγμα από το σύνολο του πληθυσμού και εξετάζουμε τον αριθμό των νοικοκυριών τα οποία έχουν το χαρακτηριστικό που μας ενδιαφέρει.

Έστω οτι ο αριθμός των μονάδων του πληθυσμού είναι N και N_j είναι ο αριθμός των μονάδων της κατηγορίας (υποπληθυσμού) που μας ενδιαφέρει. Αν ληφθεί δείγμα μεγέθους n από τον πληθυσμό N και ο αριθμός των μονάδων του δείγματος που εμπίπτει στον υποπληθυσμό N_j είναι n_j, τότε συμβολίζοντας με x_{jk} τις μετρήσεις των μονάδων αυτών (όπου $k = 1,2,\dots,n_j$), ο μέσος του υποπληθυσμού για την κατηγορία j εκτιμάται από την σχέση :

$$\bar{x}_j = \frac{\sum_{k=1}^{n_j} x_{jk}}{n_j}$$

Η διακύμανση της εκτίμησης $\bar{x}_j$ δίνεται απο τη σχέση:

$$V(\bar{x}_j) = \left(\frac{N_j - n_j}{N_j} \right) \frac{\sum_{k=1}^{n_j} (x_{jk} - \bar{x}_j)^2}{n_j(n_j - 1)}$$

Και η εκτίμηση του τυπικού σφάλματος είναι:

$$SE(\bar{x}_j) = \sqrt{V(\bar{x}_j)}$$

Σε περίπτωση βέβαια που η τιμή του N_j είναι άγνωστη μπορεί να χρησιμοποιηθεί το $\frac{n}{N}$ αντί του $\frac{n_j}{N_j}$ για τον υπολογισμό της διόρθωσης πεπερασμένου πληθυσμού.(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

2.6. Παράδειγμα Εκτίμησης Μέσου Υποπληθυσμού

Ο στόχος της έρευνας είναι να μελετήσουμε το μέσο ετήσιο εισόδημα των νοικοκυριών που έχουν έναν ιδιωτικό υπάλληλο. Επιλέγουμε μια τυχαία πόλη όπου διαμένουν 1000 νοικοκυριά. Στην περίπτωση αυτή δεν μπορούμε εκ των προτέρων να διακρίνουμε τα νοικοκυριά που έχουν έναν ιδιωτικό υπάλληλο και για αυτό εφαρμόζουμε την μέθοδο της απλής τυχαίας δειγματοληψίας υποπληθυσμού. Για το λόγο αυτό παίρνουμε από διαθέσιμο κατάλογο των νοικοκυριών της πόλης ένα τυχαίο δείγμα n=100 και εξετάζουμε τον αριθμό των νοικοκυριών τα οποία έχουν το χαρακτηριστικό που μας ενδιαφέρει. Από αυτόν τον κατάλογο βρίσκουμε ότι τα νοικοκυριά που έχουν έναν ιδιωτικό υπάλληλο ανέρχονται σε $n_j = 60$. Από την έρευνα που έγινε στα 60 νοικοκυριά διαπιστώθηκε οτι το ετήσιο εισόδημα κατά νοικοκυριό (σε χιλιάδες ευρώ) είναι:

10	9	11	15	11	14	10	13	12	12	14	12
15	12	13	14	12	13	12	14	11	13	15	14
12	14	15	12	9	15	14	11	10	15	13	13
13	15	12	14	14	9	15	9	14	14	14	10
11	10	9	10	13	10	11	10	15	12	12	9

Η εκτίμηση του μέσου ετήσιου εισοδήματος των νοικοκυριών της πόλης που έχουν έναν ιδιωτικό υπάλληλο είναι:

$$\bar{x}_j = \frac{\sum_{k=1}^{n_j} x_{jk}}{n_j} = \frac{10+9+11+...+12+9}{60} = \frac{739}{60} = 12{,}3$$

δηλαδή **12.300€** το χρόνο.

Εάν ο αριθμός των νοικοκυριών της πόλης που έχουν έναν ιδιωτικό υπάλληλο ανέρχεται σε $N_j = 650$, τότε η διακύμανση του $\bar{x}_j$ είναι:

$$V(\bar{x}_j) = \left(\frac{N_j - n_j}{N_j} \right) \frac{\sum_{k=1}^{n_j} (x_{jk} - \bar{x}_j)^2}{n_j(n_j - 1)} = \frac{650 - 60}{650} \cdot \frac{223}{60(60-1)} = 0{,}057$$

Το τυπικό σφάλμα είναι:

$$SE(\bar{x}_j) = \sqrt{V(\bar{x}_j)} = \sqrt{0{,}057} = 0{,}238$$

Ο συντελεστής μεταβλητικότητας είναι:

$$CV(\bar{x}_i) = \frac{SE(\bar{x}_i)}{\bar{x}_i} 100\% = \frac{0.238}{12.3} 100\% = 1.93\%$$

Το διάστημα εμπιστοσύνης στο οποίο βρίσκεται το μέσο ετήσιο εισόδημα των νοικοκυριών της πόλης με πιθανότητα 95% είναι:

$$\bar{x}_j \pm Z_{a/2} \cdot SE(\bar{x}_j) \Leftrightarrow$$

$$12{,}3 \pm 1{,}96 \cdot 0{,}238 \Leftrightarrow$$

$$12{,}3 \pm 0{,}47$$

Δηλαδή, βρίσκεται στο δίαστημα από **11.830€** έως **12.770€** με πιθανότητα 95%.

2.7. Πλεονεκτήματα – Μειονεκτήματα Απλής Τυχαίας Δειγματοληψίας

Η απλή τυχαία δειγματοληψία αποτελεί τη βασικότερη και απλούστερη μέθοδο τυχαίας δειγματοληψίας. Με αυτήν συγκρίνονται όλες οι άλλες τεχνικές τυχαίας δειγματοληψίας ως προς τη διακύμανση των εκτιμήσεών τους, δηλαδή ως προς την αποτελεσματικότητά τους.

Η απλή τυχαία δειγματοληψία συνήθως εφαρμόζεται όταν ο πληθυσμός της έρευνας είναι μικρός και όταν είναι σχετικά εύκολο και όχι δαπανηρό να καταγράψουμε και να απαριθμήσουμε τις μονάδες που θα συμπεριληφθούν στο δείγμα.

Επίσης, σε περίπτωση που το δείγμα διασκορπίζεται σε μία ευρεία περιοχή, η απλή τυχαία δειγματοληψία μπορεί να γίνει αρκετά επίπονη και οικονομικά ασύμφορη διαδικασία.

Μια περίπτωση όπου η απλή τύχαία δειγματοληψία δεν είναι η κατάλληλη μέθοδος είναι όταν οι δειγματοληπτικές μονάδες διαφέρουν σημαντικά σε μέγεθος και υπάρχει έντονη συσχέτιση μεταξύ του μεγέθους της μονάδας και των υπό εξέταση χαρακτηριστικών. Για παράδειγμα, η μεταβλητή ¨αριθμός υπαλλήλων που απουσιάζουν απο τη θέση τους λόγω ασθένειας¨ επηρεάζεται έντονα από το μέγεθος της επιχείρησης (δηλαδή απο τον αριθμό υπαλλήλων της επιχείρησης). Στην περίπτωση αυτή, η εφαρμογή της απλής τυχαίας δειγματοληψίας με πιθανότητα επιλογής ίση για όλες τις μονάδες, μπορεί να οδηγήσει σε εσφαλμένες εκτιμήσεις, αφού δεν λαμβάνεται υπόψη η επίδραση των μεγαλύτερων δειγματοληπτικών μονάδων στα υπό μελέτη χαρακτηριστικά του ερευνώμενου πληθυσμού.

Το πρόβλημα αυτό μπορεί να βελτιωθεί σε κάποιο βαθμό με την εφαρμογή της σρτωματοποιημένης δειγματοληψίας. Ένας άλλος τρόπος βελτίωσης του προβλήματος είναι η επιλογή των δειγματοληπτικών μονάδων με πιθανότητα ανάλογη προς το μέγεθος των μονάδων και με επανατοποθέτηση (δηλαδή μια δειγματοληπτική μονάδα μπορεί να επιλεγεί περισσότερες απο μια φορά). Με τον τρόπο αυτό δίνεται μεγαλύτερη ευκαιρία στις μεγαλύτερες μονάδες να επιλεγούν στο δείγμα (βλ. . Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Τέλος, προϋπόθεση για την εφαρμογή της απλής τυχαίας δειγματοληψίας αποτελεί η ύπαρξη ενός πλήρους , χωρίς επαναλήψεις δειγματοληπτικού πλαισίου. Στην πράξη όμως, τα δειγματοληπτικά πλαίσια που έχουμε στη διάθεσή μας σπάνια ικανοποιούν τις παραπάνω απαιτήσεις, με άμεσο αποτέλεσμα να δυσχεραίνουν την εφαρμογή της μεθόδου.

2.8. Παράδειγμα Απλής Τυχαίας Δειγματοληψίας

Έστω ότι θέλουμε να εκτιμήσουμε το μέσο μέγεθος των νοικοκυριών και το διάστημα εμπιστοσύνης του μέσου σε μια πόλη όπου διαμένουν συνολικά 800 νοικοκυριά. Επειδή ο πληθυσμός είναι μικρός ειναι εύκολο να απαριθμήσουμε τις μονάδες που θα συμπεριληφθούν στο δείγμα. Σε αυτή την περίπτωση, αν εφαρμόσουμε τη μέθοδο της απλής τυχαίας δειγματοληψίας για την επιλογή του δείγματος οι εκτιμήσεις που θα προκύψουν θα είναι πιο ακριβείς παρά αν χρησιμοποιούσαμε οποιαδήποτε άλλη μέθοδο δειγματοληψίας. Από κατάλογο όπου είναι καταγεγραμμένα όλα τα νοικοκυριά της πόλης (δειγματοληπτικό πλαίσιο), επιλέγουμε με τη μέθοδο της κλήρωσης 40 νοικοκυριά, τα οποία αποτελούν το δείγμα μας.

Από τη δειγματοληπτική αυτή έρευνα προέκυψε ο αριθμός μελών κατά νοικοκυριο ως εξής:

4	3	2	5	3	4	3	2	1	6
5	4	6	2	2	4	5	6	3	2
7	6	5	3	4	3	5	7	4	3
3	4	4	5	6	3	2	1	7	6

Η εκτίμηση του μέσου μεγέθους των νοικοκυριών είναι:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{4+3+2+...+7+6}{40} = \frac{160}{40} = 4$$

δηλαδή το μέσο μέγεθος των νοικοκυριών είναι 4 άτομα.

Το τυπικό σφάλμα στην απλή τυχαία δειγματοληψία είναι:

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}}$$

Επειδή η διακύμανση του πληθυσμού s^2 είναι άγνωστη χρησιμοποίουμε την αμερόληπτη εκτίμηση της s^2 , η οποία είναι:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(4-4)^2 + (3-4)^2 + ... + (7-4)^2 + (6-4)^2}{39} = \frac{108}{39} = 2,77$$

Επομένως το τυπικό σφάλμα είναι:

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}} = \frac{1,66}{6,32}\sqrt{0,95} = 0,256$$

Το σχετικό τυπικό σφάλμα είναι:

$$CV(\bar{x}) = \frac{SE(\bar{x})}{\bar{x}}100\% = \frac{0.256}{4}100\% = 6.4\%$$

Το διάστημα εμπιστοσύνης στο οποίο βρίσκεται το μέσο μέγεθος των νοικοκυριών m με πιθανότητα 95% είναι:

$$\bar{x} \pm Z_{a/2} \cdot SE(\bar{x}) \Leftrightarrow$$

$$4 \pm 1,96 \cdot 0,256 \Leftrightarrow$$

$$4 \pm 0,5$$

δηλαδή το μέσο μέγεθος των νοικοκυριών βρίσκεται στο διάστημα 3,5 εως 4,5, με πιθανότητα 95%.

<u>ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ</u>

ΣΥΣΤΗΜΑΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

3.1. Εισαγωγή

Η συστηματική δειγματοληψία χρησιμοποιείται συχνά αντί της απλής τυχαίας δειγματοληψίας, ιδίως σε μεγάλους πληθυσμούς, επειδή η διαδικασία επιλογής του δείγματος είναι πολύ απλή. Τα N μέλη του πληθυσμού πρέπει να είναι αριθμημένα και καταγεγραμμένα στο πλαίσιο δειγματοληψίας.

Για τη συστηματική δειγματοληψία χρειαζόμαστε δύο πράγματα:

i. το διάστημα ή βήμα δειγματοληψίας, το οποίο θα παριστούμε με l και
ii. ένα τυχαίο ξεκίνημα i.

Εάν N είναι το μέγεθος του πληθυσμού, n το απαιτούμενο μέγεθος δείγματος και υποθέτοντας ότι $N = n \cdot l$, τότε το βήμα δειγματοληψίας είναι ακέραιος αριθμός και δίνεται από τη σχέση :

$$l = \frac{N}{n} = \frac{1}{f}.$$

Εκλέγουμε τώρα τυχαία (π.χ. με τη βοήθεια των πινάκων τυχαίων αριθμών) ένα αριθμό μεταξύ των πρώτων l αριθμών της λίστας. Έστω i $(1 \le i \le l)$ ο επιλεγμένος αριθμός, δηλαδή το τυχαίο ξεκίνημα. Τότε οι μονάδες του πληθυσμού που έχουν αύξοντες αριθμούς

$$i, i + l, i + 2l, \ldots, i + (n-1)l$$

απαρτίζουν το επιλεγόμενο συστηματικό δείγμα μεγέθους n.

Η πιθανότητα επιλογής κάθε μονάδας του πληθυσμού είναι ίση με την πιθανότητα επιλογής του τυχαίου αριθμού i, δηλαδή ίση με $\frac{1}{l} = \frac{1}{N/n} = \frac{n}{N} = f$ όπως δηλαδή και στην περίπτωση της απλής τυχαίας δειγματοληψίας.

Στην περίπτωση που το μέγεθος N του πληθυσμού είναι άγνωστο ή πολύ μεγάλο, το μέγεθος n του δείγματος δεν μπορεί να προσδιορισθεί. Εκείνο που προσδιορίζεται είναι το κλάσμα δειγματοληψίας f. Έστω για παράδειγμα ότι θέλουμε να εξετάσουμε τη γνώμη των πελατών ενός υποκαταστήματος μιας τράπεζας. Προκαθορίζουμε ότι $f = 0.1$, οπότε το βήμα της συστηματικής δειγματοληψίας είναι $l = \frac{1}{f} = 10$ και ερευνάμε κάθε 10 πελάτες που βγαίνουν από το υποκατάστημα, μετά από μια αρχή i $(1 \leq i \leq l)$.

3.1.1 Παράδειγμα

Ας πάρουμε για παράδειγμα τη Νομική Σχολή Αθηνών. Σ' αυτή φοιτούν 9.200 φοιτητές, και θέλουμε δείγμα από 400. Τότε το διάστημα δειγματοληψίας θα είναι:

$$\frac{N}{n} = \frac{9.200}{400} = 23,$$

δηλαδή το *λ = 23.*

Μπορούμε τώρα να ενοποιήσουμε τα μητρώα των ετών της σχολής και, αφού βρούμε έναν αριθμό μεταξύ 1 και 23, π.χ. το 12, να πάρουμε τους φοιτητές με τη σειρά αρχίζοντας από τον 12:

12, 35, 58, 81, ..., 9.189.

Αυτοί οι αριθμοί ανταποκρίνονται στη συνεχή αρίθμηση των φοιτητών που κάναμε μετά την ενοποίηση των καταλόγων.

Στην πραγματικότητα κατά τη συστηματική δειγματοληψία, ο πληθυσμός χωρίζεται σε l ισοπληθείς ομάδες των *n* μονάδων η κάθε μία. Δηλαδή έχουμε $N = nl$. Τα δυνατά συστηματικά δείγματα δίνονται στον παρακάτω πίνακα.

Συστηματικά Δείγματα	Επιλεγόμενες Ομάδες
1	X_1 $X_{1+\lambda}$ $X_{1+2\lambda}$ ... $X_{1+(n-1)\lambda}$
2	X_2 $X_{2+\lambda}$ $X_{2+2\lambda}$ ... $X_{2+(n-1)\lambda}$
...	... ...
i	X_i $X_{i+\lambda}$ $X_{i+2\lambda}$ ... $X_{i+(n-1)\lambda}$
...	... ...
λ	X_λ $X_{2\lambda}$ $X_{3\lambda}$ ... $X_{n\lambda}$

Η πιθανότητα να επιλεγεί ένα από τα λ συστηματικά δείγματα είναι $\frac{1}{l}$.

Επομένως, κάθε μονάδα του πλυθησμού έχει την ίδια πιθανότητα να περιληφθεί στο δείγμα κατά τη μέθοδο της συστηματικής δειγματοληψίας.

3.2. Διακύμανση της Εκτίμησης του Μέσου

Για την αποφυγή των ''πολύπλοκων'' δεικτών στις τιμές των l συστηματικών δειγμάτων συμβολίζουμε με x_{ij} την j μονάδα του i συστηματικού δείγματος, όπου $j\ 1,2,...\ n$ και $i=1,2,...l$. Για την εκτίμηση $\bar{x}$ στη συστηματική δειγματοληψία, ισχύει για $N=n\cdot l$ το εξής:

$$E(\bar{x})=m$$

$$\mathrm{E}(\bar{x}_i)=\frac{1}{nl}\sum_{i=1}^{l}\sum_{j=1}^{n}x_{ij}=\frac{\mathrm{X}}{nl}=\frac{\mathrm{X}}{\mathrm{N}}=m$$

Η εκτίμηση του μέσου του πληθυσμού m, γίνεται απο τον μέσο $\bar{x}$ του συστηματικού δείγματος ,δηλαδή έχουμε:

$$\bar{x}=\frac{\sum_{j=1}^{n}x_{ij}}{n}$$

Οι δυνατοί δειγματικοί μέσοι είναι $\bar{x}_1,\bar{x}_2,...,\bar{x}_i,...,\bar{x}_l$,οπότε η διακύμανση του μέσου $\bar{x}$ είναι εξ ορισμού:

$$V(\bar{x})=\frac{1}{l}\sum_{i=1}^{l}(\bar{x}_i-m)^2$$

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Από τα στοιχεία ενός μόνο συστηματικού δείγματος δεν μπορεί να βρεθεί αμερόληπτη εκτίμηση της διακύμανσης του δειγματικού μέσου $\bar{x}$. Επομένως η διακύμανση της εκτίμησης του μέσου $\bar{x}$, δίνεται απο τη σχέση :

$$V(\bar{x})=\frac{\mathrm{N}-1}{\mathrm{N}}s^2-\frac{l(n-1)}{\mathrm{N}}s_w^2$$

Και επειδη : $l=\frac{N}{n}\Leftrightarrow nl=N\Leftrightarrow n=\frac{N}{l}\Leftrightarrow\frac{1}{n}=\frac{1}{N/l}\Leftrightarrow\frac{1}{n}=\frac{l}{N}$

Προκύπτει : $V(\bar{x})=\frac{N-1}{N}s^2-\frac{n-1}{n}s_w^2$

και

$$SE(\bar{x}) = \sqrt{V(\bar{x})}$$

όπου

$$s^{2} = \frac{\sum_{i=1}^{l}\sum_{j=1}^{n}(x_{ij} - m)^{2}}{\mathrm{N} - 1}$$

είναι η διακύμανση του συνολικού πληθυσμού και

$$s_{w}^{2} = \frac{1}{l\,(n-1)}\sum_{i-1}^{l}\sum_{j=1}^{n}(x_{ij} - \bar{x})^{2}$$

είναι η διακύμανση μεταξύ των μονάδων που βρίσκονται στο ίδιο συστηματικό δείγμα, δηλαδή ο μέσος όρος των λ διακύμανσεων, αφού η διακύμανση εντός του i συστηματικού δείγματος είναι:

$$s^{2} = \frac{\sum_{j=1}^{n}(x_{ij} - \bar{x})^{2}}{n-1}$$

Ο μέσος ενός συστηματικού δείγματος είναι πιο ακριβής από τον μέσο ενός απλού τυχαίου δείγματος, του ίδιου μεγέθους εάν και μόνο εάν $s^{2} < s_{w}^{2}$.

Αυτό δείχνει οτι η συστηματική δειγματοληψία είναι περισσότερο ακριβής από την απλή τυχαία δειγματοληψία, εάν η διακύμανση μέσα στα συστηματικά δείγματα είναι μεγαλύτερη απο τη διακύμανση του πληθυσμού. Δηλαδή η συστηματική δειγματοληψία είναι περισσότερο ακριβής όταν οι μονάδες των συστηματικών δειγμάτων είναι ανομοιογενείς.(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Σε περίπτωση που εντός των συστηματικών δειγμάτων η διακύμανση είναι μικρή ή μηδέν τότε πρέπει να αποφεύγεται η συστηματική δειγματοληψία, διότι οι εκτιμήσεις δεν θα είναι αξιόπιστες.

Μια περίπτωση που η διακύμανση των συστηματικών δειγμάτων ισούται με μηδέν, είναι όταν οι μονάδες του πληθυσμού παρουσιάζουν περιοδικότητα, ως προς το μέγεθος του ερευνώμενου χαρακτηριστικού, με συνέπεια κάθε παρατήρηση να είναι ίδια εντός του συστηματικού δείγματος. Έτσι το σύνολο των παρατηρήσεων δίνει την ίδια πληροφορία που θα έδινε και μία μόνο παρατήρηση.

3.3. Εκτίμηση του Μέσου Πληθυσμού

Όπως αναφέρθηκε και προηγούμενα, η εκτίμηση του μέσου πληθυσμού m από ένα συστηματικό δείγμα δίνεται απο τον τύπο

$$\overline{x} = \frac{\sum_{j=1}^{n} x_{ij}}{n}$$

Η διακύμανση του $\overline{x}$ δεν μπορεί να εκτιμηθεί από τα στοιχεία ενός μόνο συστηματικού δείγματος, εκτός αν ο πληθυσμός είναι τυχαίος. Ένας πληθυσμός είναι τυχαίος όταν οι μονάδες του είναι διατεταγμένες τυχαία. Οι μονάδες ενός συστηματικού δείγματος που επιλέγονται από έναν τυχαίο πληθυσμό αναμένονται να είναι ετερογενείς. Όταν το N είναι μεγάλο η διακύμανση του $\overline{x}$ είναι περίπου ίση με τη διακύμανση του $\overline{x}$ που βασίζεται στην απλή τυχαία δειγματοληψία, οπότε μπορούμε να χρησιμοποίησουμε τον τύπο : $V(\overline{x}) = \frac{s^2}{n}\left(\frac{N-n}{N}\right)$

Η διακύμανση s^2 είναι μία αμερόληπτη εκτίμηση του s^2, δηλαδή:

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})}{n-1}$$

3.4. Εκτίμηση Ποσοστού ενός Πληθυσμού

Η εκτίμηση ποσοστού ενός πληθυσμού από ένα συστηματικό δείγμα, όταν ο πληθυσμός είναι τυχαίος, γίνεται όπως και στην περίπτωση της απλής τυχαίας δειγματοληψίας, δηλαδή έχουμε: $p = \frac{\sum_{i=1}^{n} x_i}{n}$

Η διακύμανση του ποσοστού p όταν κάνουμε επιλογή μονάδων, χωρίς επανατοποθέτηση είναι:

$$V(p)=E(p-P)^2=\frac{PQ}{n}\left(\frac{N-n}{N-1}\right)$$

Όταν η διακύμανση του πληθυσμού $s^2=PQ$ είναι άγνωστη, μια αμερόληπτη εκτίμηση της διακύμανσης του ποσοστού p, προκύπτει από το δείγμα, δηλαδή:

$$V(p)=\frac{s^2}{n}\left(\frac{N-n}{N}\right)=\frac{pq}{n-1}\left(\frac{N-n}{N}\right)$$

3.5. Πλεονεκτήματα – Μειονεκτήματα Συστηματικής Δειγματοληψίας

Η συστηματική δειγματοληψία χρησιμοποιείται συχνά αντί της απλής τυχαίας δειγματοληψίας, ιδίως σε μεγάλους πληθυσμούς, επειδή η διαδικασία επιλογής του δείγματος είναι πολύ απλή.

Η ευκολία στην εφαρμογή της συνεπάγεται *μείωση σφαλμάτων συλλογής* καθώς επίσης και *περιορισμό της μεροληπτικής επιλογής των δειγματοληπτικών μονάδων*, με αποτέλεσμα τη μεγαλύτερη δυνατότητα για αντιπροσωπευτικό δείγμα.

Επίσης η συστηματική δειγματοληψία μπορεί να δώσει περισσότερες πληροφορίες, σε δεδομένο κατά μονάδα κόστος, διότι ένα συστηματικό δείγμα κατά κανόνα είναι διεσπαρμένο με μεγαλύτερη ομοιογένεια στον ερευνώμενο πληθυσμό.

Επιπλέον, η συστηματική δειγματοληψία είναι πιο αποτελεσματική από την απλή τυχαία δειγματοληψία, εάν η διακύμανση μέσα στα συστηματικά δείγματα είναι μεγαλύτερη απο τη διακύμανση του πληθυσμού. Δηλαδή η συστηματική δειγματοληψία είναι περισσότερο ακριβής όταν οι μονάδες των συστηματικών δειγμάτων είναι ανομοιογενείς. Όταν η διακύμανση των συστηματικών δειγμάτων είναι πολύ μικρή ή μηδεν, η εφαρμογή της συστηματικής δειγματοληψίας για την επιλογή του δείγματος πρέπει να αποφεύγεται. Μία τέτοια περίπτωση είναι όταν οι μονάδες του ερευνώμενου πληθυσμού παρουσιάζουν περιοδικότητα.(βλ. Ξεκαλάκη Ε.,1995).

Για παράδειγμα αν οι μονάδες που έχουν επιλεγεί σε μια συστηματική δειγματοληψία

$$x_2\,,x_{12}\,,x_{22}\,,x_{32}\ldots$$

παρουσιάζουν περιοδικότητα, τότε θα παρέχουν την ίδια πληροφορία και θα έχουμε:

$$x_2 = x_{12} = x_{22} = \ldots = \overline{x}$$

δηλαδή η εκτίμηση του μέσου m του πληθυσμού από δείγμα μεγέθους n συνεπάγεται την ίδια ακρίβεια με μία εκτίμηση που θα βασίζεται σε δείγμα μεγέθους $n=1$. Επομένως θα έχουμε υπερεκτίμηση ή υποεκτίμηση των υπό μελέτη χαρακτηριστικών (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

3.6. Παράδειγμα Συστηματικής Δειγματοληψίας

Ο στόχος της συγκεκριμένης έρευνας είναι να εκτιμήσουμε κατά μέσο όρο τον αριθμό των ατόμων που εισέρχονται ημερησίως σε ένα κατάστημα κατά τη διάρκεια 120 ημερών. Επειδή η επιλογή τυχαίων αριθμών απαιτεί πολύ χρόνο θα εφαρμόσουμε την συστηματική δειγματοληψία η οποία είναι πιο εύκολη όσον αφορά την επιλογή του δείγματος. Λόγω της ευκολίας στην εφαρμογή της, αποφεύγονται τα σφάλματα και η μεροληψία, με αποτέλεσμα να έχουμε πιο αντιπροσωπευτικό δείγμα σε σχέση με την απλή τυχαία δειγματοληψία.

Έτσι λαμβάνοντας υπόψιν και το κόστος αποφασίσαμε να καταμετρήσουμε τα άτομα που εισέρχονται στο κατάστημα κάθε 10 ημέρες, δηλαδή το διάστημα δειγματοληψίας είναι: $l = \frac{N}{n} = \frac{120}{12} = 10$.

Επομένως, ο αριθμός των ατόμων που εισέρχονται στο κατάστημα κάθε δέκατη μέρα, με τυχαίο ξεκίνημα τη δεύτερη μέρα δίνεται στον παρακάτω πίνακα:

Ημέρα	Αριθμός ατόμων που εισέρχονται στο κατάστημα	Ημέρα	Αριθμός ατόμων που εισέρχονται στο κατάστημα
2	20	62	20
12	18	72	22
22	24	82	25
32	23	92	28
42	26	102	19
52	17	112	22

Η εκτίμηση του μέσου όρου των ατόμων που εισέρχονται ημερησίως στο κατάστημα είναι:

$$\bar{x} = \frac{\sum_{j=1}^{n} x_{ij}}{n} = \frac{20+18+\ldots+19+22}{12} = \frac{264}{12} = 22$$

Η διακύμανση του $\bar{x}$ είναι:

$$V(\bar{x}) = \frac{s^2}{n}\frac{N-n}{N}$$

Όπου:

$$s^2 = \frac{\sum_{j=1}^{n}(x_{ij}-\bar{x})^2}{n-1} = \frac{(20-22)^2+(18-22)^2+\ldots+(22-22)^2}{12-1} = 11{,}3$$

$$\text{και} \quad V(\bar{x}) = \frac{s^2}{n}\left(\frac{N-n}{N}\right) = \frac{11{,}3}{12}\left(\frac{120-12}{120}\right) = 0{,}85$$

Επομένως το τυπικό σφάλμα της εκτίμησης είναι:

$$SE(\bar{x}) = \sqrt{V(\bar{x})} = \sqrt{0{,}85} = 0{,}92$$

Το διάστημα εμπιστοσύνης στο οποίο βρίσκεται ο μέσος όρος των ατόμων m που εισέρχονται στο κατάστημα με πιθανότητα 95% είναι:

$$\bar{x} \pm z_{a/2} \cdot SE(\bar{x})$$

$$22 \pm 1{,}96 \cdot 0{,}92$$

$$22 \pm 1{,}8$$

Δηλαδή ο μέσος αριθμός ατόμων που εισέρχονται στο κατάστημα ημερησίως βρίσκεται στο διάστημα 20 – 24, με πιθανότητα 95%.

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

ΣΤΡΩΜΑΤΟΠΟΙΗΜΕΝΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

4.1. Εισαγωγή

Ο σκοπός της στρωματοποιημένης δειγματοληψίας είναι να χωρίσει ένα ανομοιογενή πληθυσμό σε ομοιογενείς υποπληθυσμούς. Έτσι με αυτόν τον τρόπο επιτυγχάνεται μεγαλύτερη ακρίβεια στις εκτιμήσεις των χαρακτηριστικών του ερευνώμενου πληθυσμού, επειδή οι μετρήσεις που προκύπτουν απο ομοιογενή στρώματα διαφέρουν ελάχιστα μεταξύ τους. Επίσης εξασφαλίζεται η αντιπροσωπευτικότητα του ερευνώμενου πληθυσμού επειδή στο συνολικό δείγμα συμπεριλαμβάνονται μονάδες από όλα τα στρώματα. Ενώ αντίθετα, στην απλή τυχαία δειγματοληψία υπάρχει περίπτωση να παραληφθούν μονάδες που αποτελούν ένα ομοιογενές στρώμα, με αποτέλεσμα να μην έχουμε καθόλου πληροφόρηση για την κατηγορία αυτή (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997).

Αυξάνοντας την ομοιογένεια κάθε στρώματος κερδίζουμε σε ακρίβεια, άρα μας συμφέρει να διαιρέσουμε τον πληθυσμό σε όσο το δυνατόν περισσότερα στρώματα, δεδομένου οτι η ομοιογένεια των στρωμάτων θα αυξάνει γενικά με τη μείωση του μεγέθους τους.

Επομένως στη στρωματοποιημένη δειγματοληψία ο ερευνώμενος πληθυσμός χωρίζεται σε κ υποπληθυσμούς οι οποίοι ονομάζονται στρώματα.

Δηλαδή έχουμε $\mathrm{N}_1 + \mathrm{N}_2 + \mathrm{N}_3 + ... + \mathrm{N}_\mathrm{K} = \mathrm{N}$. Στη συνέχεια επιλέγεται μέσα από κάθε στρώμα με τυχαίο τρόπο ένα δείγμα, το άθροισμα των δειγμάτων αυτών ($n_1 + n_2 + ... + n_K$) δίνει το συνολικό δείγμα *n*.

4.2. Κατανομή Δείγματος κατά Στρώματα

Μετά τον διαχωρισμό του ερευνώμενου πληθυσμού σε ομοιγενή στρώματα, γίνεται κατανομή του δείγματος μεταξύ των στρωμάτων. Στη συνέχεια επιλέγεται το δείγμα με τυχαίο τρόπο μέσα σε κάθε στρώμα, είτε με τη μέθοδο της απλής τυχαίας ή της συστηματικής δειγματοληψίας.

Η επιλογή του κατάλληλου μεγέθους μέσα σε κάθε στρώμα, αποσκοπεί στο να πετύχουμε εκτιμήτριες με μικρές διακυμάνσεις στο ελάχιστο δυνατό κόστος.

Με βάση το σκοπό αυτό για την κατανομή δείγματος μεταξύ των στρωμάτων, πρέπει να λαμβάνονται υπόψη οι παρακάτω παράγοντες:

i. Ο συνολικός αριθμός μονάδων σε κάθε στρώμα.
ii. Η μεταβλητικότητα των μονάδων σε κάθε στρώμα
iii. Το κόστος κατά δειγματοληπτική μονάδα σε κάθε στρώμα

Ο αριθμός των δειγματοληπτικών μονάδων που επιλέγονται σε κάθε στρώμα επιδρά στην αξιοπιστία των αποτελεσμάτων και στην ακρίβεια των εκτιμήσεων. Για παράδειγμα, ένα δείγμα 30 μονάδων από πληθυσμό 300 μονάδων δίνει περισσότερες πληροφορίες, από ότι ένα δείγμα 30 μονάδων που λαμβάνεται από πληθυσμό 30.000 μονάδων. Επομένως, πρέπει να λαμβάνεται μεγαλύτερο δείγμα από στρώματα με μεγαλύτερους πληθυσμούς.

Η μεταβλητικότητα επιδρά στη αξιοπιστία των εκτιμήσεων και επομένως πρέπει να λαμβάνεται μεγαλύτερο δείγμα από στρώματα που παρουσιάζουν μεγαλύτερη ανομοιογένεια.

Εάν το κατά μονάδα κόστος μιας δειγματοληπτικής έρευνας διαφέρει από στρώμα σε στρώμα, πρέπει να πάρουμε μικρότερο δείγμα από το στρώμα με υψηλότερο κόστος, για να κρατήσουμε το κόστος της δειγματοληπτικής έρευνας σε χαμηλά επίπεδα, υπό την προϋπόθεση ότι δεν επηρεάζεται η αξιοπιστία των αποτελεσμάτων (βλ. Πανταζίδης Ν.,1976).

4.3. Μέθοδοι Κατανομής Δείγματος

Η κατανομή του συνολικού δείγματος στα στρώματα πρέπει να γίνεται με μία από τις παρακάτω μεθόδους, προκειμένου να πετύχουμε εκτιμήσεις με μικρές διακυμάνσεις στο ελάχιστο δυνατό κόστος:

- Αναλογική κατανομή δείγματος και
- Άριστη κατανομή δείγματος.

Πολλές φορές η κατανομή του δείγματος γίνεται αυθαίρετα στα διάφορα στρώματα, δηλαδή το δείγμα που επιλέγεται σε ένα στρώμα είναι ανεξάρτητο από τα χαρακτηριστικά του δείγματος των άλλων στρωμάτων. Ο τρόπος αυτός κατανομής του δείγματος είναι αποδεκτός, με την προϋπόθεση ότι εξασφαλίζεται η ελαχιστοποίηση της διακύμανσης των εκτιμήσεων, ώστε να προκύπτουν αξιόπιστα αποτελέσματα (βλ. Δαμιανός Χ.,2006).

4.3.1. Αναλογική Κατανομή Δείγματος

Η αναλογική κατανομή δείγματος εφαρμόζεται όταν το κόστος κατά μονάδα δείγματος είναι ίδιο και η διακύμανση είναι περίπου ίδια σε όλα τα στρώματα. Με την αναλογική κατανομή, το δείγμα λαμβάνεται αναλογικά μέσα σε κάθε στρώμα προς το μέγεθος του αντίστοιχου στρωμάτος του ερευνώμενου πληθυσμού. Δηλαδή με σταθερό κλάσμα δειγματοληψίας, το οποίο δίνει αυτοσταθμιζόμενο δείγμα, σε κάθε στρώμα ισχύει:

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = ... = \frac{n_k}{N_k} = \frac{n}{N}$$

Οπότε ο τύπος προσδιορισμού του δείγματος στο στρώμα i είναι:

$$\frac{n_i}{N_i} = \frac{n}{N} \quad \text{ή} \quad n_i = \frac{n}{N} N_i$$

4.3.2. Άριστη Κατανομή Δείγματος

Με τον άριστο καταμερισμό του δείγματος n στα διάφορα στρώματα αποβλέπουμε στο να ελαττώσουμε κατά το δυνατό τη διακύμανση του συνολικού μεγέθους της μέσης τιμης χωρίς να αυξήσουμε το συνολικό μέγεθος του δείγματος n, δηλαδή το κόστος της έρευνας.

Όταν ο ερευνώμενος πληθυσμός παρουσιάζει σε μερικά στρώματα μεγαλύτερη μεταβλητότητα (μεγαλύτερη διακύμανση), είναι απαραίτητο στα στρώματα αυτά να ληφθεί μεγαλύτερο δείγμα, προκειμένου να αυξηθεί η ακρίβεια των εκτιμήσεων. Επίσης εάν το κόστος συλλογής πληροφοριών είναι μεγαλύτερο σε μερικά στρώματα, μπορεί να ληφθεί μικρότερο δείγμα, προκειμένου να μειωθεί το κόστος της έρευνας. Δηλαδή όταν η διακύμανση της υπό μελέτη μεταβλητής και το κόστος συλλογής πληροφοριών διαφέρουν σημαντικά από στρώμα σε στρώμα, οι μονάδες δείγματος πρέπει να κατανέμονται στα στρώματα κατά τέτοιο τρόπο ώστε η διακύμανση της εκτίμησης να γίνει ελάχιστη για ένα καθορισμένο κόστος ή να γίνει ελαχιστοποίηση του κόστους για μια καθορισμένη τιμή της διακύμανσης της εκτίμησης.

Επομένως για να αντιπροσωπευθούν επαρκώς τα στρώματα αυτά στο δείγμα θα πρέπει ο λόγος n_i / N_i να είναι ανάλογος της τυπικής απόκλισης s_i του στρώματος και αντιστρόφως ανάλογα της τετραγωνικής ρίζας του κόστους κατά μονάδα (C_i). Δηλαδή:

$$\frac{n_1}{N_1 s_1 \big/ \sqrt{C_1}} = \ldots = \frac{n_k}{N_k s_k \big/ \sqrt{C_k}} = \frac{n}{\sum (N_i s_i / C_i)}$$

Οπότε ο τύπος προσδιορισμόυ του δείγματος στο στρώμα i είναι:

$$\frac{n_i}{N_i s_i \big/ \sqrt{C_i}} = \frac{n}{\sum (N_i s_i / C_i)}$$

$$n_i = n \frac{N_i s_i \big/ \sqrt{C_i}}{\sum (N_i s_i \big/ \sqrt{C_i})}$$

Διαιρώντας λοιπόν την παραπάνω σχέση με N και επειδή $W_i = \frac{N_i}{N}$ προκύπτει το εξής:

$$n_i = n \frac{W_i s_i \big/ \sqrt{C_i}}{\sum (W_i s_i \big/ \sqrt{C_i}}$$

Η σχέση αυτή δηλώνει ότι το μέγεθος του απλού τυχαίου δείγματος που επιλέγεται από ένα στρώμα πρέπει να είναι μεγαλύτερο από τα δείγματα άλλων στρωμάτων αν το μέγεθος του στρώματος είναι μεγαλύτερο, η διασπορά του στρώματος είναι μεγαλύτερη και το κόστος ανά μονάδα του στρώματος είναι χαμηλότερο (βλ.Ζαϊρης, 1991).

Εάν υποθέσουμε όμως ότι το δειγματοληπτικό κόστος ανά μονάδα είναι το ίδιο για όλα τα στρώματα αποδεικνύεται ότι η διακύμανση της εκτίμησης γίνεται ελάχιστη αν τα $n_1, n_2, \ldots, n_k$ επιλεγούν έτσι ώστε:

$$\frac{n_1}{N_1 s_1} = \ldots = \frac{n_k}{N_k s_k} = \frac{n}{\sum N_i s_i}$$

Οπότε ο τύπος προσδιορισμού του δείγματος στο στρώμα i είναι:

$$n_i = n \frac{N_i s_i}{\sum N_i s_i} = n \frac{W_i s_i}{\sum W_i s_i}$$

Η ειδική αυτή περίπτωση άριστης κατανομής ονομάζεται *βέλτιστος καταμερισμός του n με σταθερό κόστος ανά δειγματοληπτική μονάδα* ή *καταμερισμός κατά Neyman*.

Αξίζει να σημειωθεί ότι τις περισσότερες φορές υπάρχουν μόνο προσεγγιστικές εκτιμήσεις για τα s_i (τυπική απόκλιση) , C_i (μοναδιαίο κόστος κατά στρώμα), οπότε ο άριστος καταμερισμός δεν αποτελεί παρά θεωρητική έννοια και η εφαρμογή του απλώς θα κάνει πολύ πιο πολύπλοκους τους υπολογισμούς. Μόνο στις περιπτώσεις που υπάρχει πραγματικά μεγάλη διαφορά στα s_i, C_i από στρώμα σε στρώμα αξίζει τον κόπο να προτιμηθεί ο άριστος καταμερισμός απο τον αναλογικό. Επιπλέον, μπορεί αυτός ο άριστος καταμερισμός ως προς ένα χαρακτηριστικό, να ελαττώσει σημαντικά την ακρίβεια της εκτίμησης για ένα άλλο χαρακτηριστικό (βλ. Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής, 2000).

Έτσι όταν σε μια δειγματοληπτική έρευνα υπάρχουν περισσότερα από ένα χαρακτηριστικά, τότε η κατανομή του δείγματος στα διάφορα στρώματα μπορεί να οδηγήσει στην απώλεια ακρίβειας, δεδομένου ότι το δειγματοληπτικό κλάσμα που είναι άριστο για ένα χαρακτηριστικό μπορεί να μην είναι για ένα άλλο. Το πρόβλημα αυτό μπορεί να ξεπεραστεί με κάποια συμβιβαστική λύση. Το πρώτο βήμα είναι να περιορισθεί ο αριθμός των χαρακτηριστικών που θα εξετασθούν σε έναν σχετικά μικρό αριθμό που θεωρούμε ότι είναι τα πιο σπουδαία. Στη συνέχεια αν υπάρχουν διαθέσιμα στοιχεία απο προηγούμενες έρευνες ή από μια πιλοτική έρευνα, μπορουμε να προσδιορίσουμε το δείγμα με την άριστη κατανομή για το κάθε χαρακτηριστικό ξεχωριστά. Οπότε από τα μεγέθη των δειγμάτων που έχουν κατανεμηθεί κατά χαρακτηριστικό, λαμβάνουμε το μεγαλύτερο με το οποίο εξασφαλίζεται η επιθυμητή ακρίβεια σε όλες τις μεταβλητές, το οποίο αποτελεί και μια ικανοποιητική συμβιβαστική λύση κατανομής του συνολικού δείγματος.(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Επίσης μια άλλη δυσκολία που συναντάμε κατά την εφαρμογή της άριστης κατανομής δείγματος είναι ότι συνήθως η διακύμανση του πληθυσμού κατά στρώμα είναι άγνωστη. Ένας τρόπος να ξεπεραστεί η δυσκολία αυτή είναι να ληφθεί ένα δείγμα μεγέθους n_i απο κάθε στρώμα (δηλαδή διεξαγωγή δοκιμαστικής έρευνας), για την εκτίμηση s_i^2 η οποία στη συνέχεια θα χρησιμοποιηθεί για την κατανομή του δείγματος n στα στρώματα. Ακόμα, μπορούν να χρησιμοποιηθούν στοιχεία από προηγούμενες έρευνες ή απογραφές με ίδιους ή παρόμοιους πληθυσμούς, ώστε να πετύχουμε κάποιες εκτιμήσεις της διακύμανσης. Γι΄αυτό συνήθως χρησιμοποιείται ο αναλογικός καταμερισμός με τον αποίο κερδίζουμε σε ακρίβεια, γιατί έτσι παραλείπεται από τον εκτιμητή διακύμανσης ο παράγοντας της διαφοροποίησης που οφείλεται στις διαφορές μεταξύ των μέσων τιμών των διάφορων στρωμάτων. Η διακύμανση είναι ανάλογη προς τις διακύμανσεις στο κάθε στρώμα επομένως όσο μικρότερες είναι οι διακύμανσεις s^2 των στρωμάτων, δηλαδή όσο περισσότερο ομοιογενή είναι τα στρώματα, τόσο μικρότερη είναι η διακύμανση $V(\bar{x})$ της $\bar{x}$, επομένως τόσο μεγαλύτερης ακρίβειας είναι η εκτιμήτρια $\bar{x}$ (βλ.Ξενακης 2002).

Τέλος η αναλογική κατα στρώματα δειγματοληψία χρησιμοποιείται πιο πολύ όχι μόνο λόγω της ακρίβειας αλλά και επειδή η στρωματοποίηση σε πολλές περιπτώσεις έιναι από μόνη της έτοιμη, ο πληθυσμός δηλαδή είναι από μόνος του χωρισμένος σε στρώματα. Επιπλέον στον αναλογικό καταμερισμό το δείγμα είναι αυτοσταθμιζόμενο, έτσι οι υπολογισμοί γίνονται απλούστεροι απ΄ότι σε άλλες κατά στρώματα δειγματοληψίες.

4.4. Εκτίμηση Μέσου Πληθυσμού

Έστω ότι έχουμε k στρώματα και κάθε στρώμα i περιέχει N_i άτομα τότε ο μέσος του πληθυσμού του στρώματος i είναι :

$$\bar{X} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{N_i} X_{ij}}{N} = \frac{\sum_{i=1}^{k} N_i \bar{X}_i}{N} = \sum_{i=1}^{k} W_i \bar{X}_i$$

δηλαδή ο μέσος του πληθυσμού είναι ένας σταθμισμένος μέσος όρος που ισούται με το άθροισμα των μέσων τιμών των στρωμάτων σταθμισμένα με το αντίστοιχο βάρος τους $W_i = \frac{N_i}{N}$ την αναλογία του στρώματος στον πληθυσμό.

Μια αμερόληπτη εκτίμηση του μέσου $\bar{X}$ του πληθυσμού είναι ο σταθμισμένος μέσος των εκτιμήσεων των k στρωμάτων, δηλαδή:

$$\bar{x} = \frac{1}{N}(N_1\bar{x}_1 + N_2\bar{x}_2 + ... + N_k\bar{x}_k) = \frac{\sum_{i=1}^{k} N_i \bar{x}_i}{N} = \sum_{i=1}^{k} W_i \bar{x}_i$$

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997).

Η διακύμανση της εκτίμησης του $\bar{x}$ εξαρτάται από τον τρόπο κατανομής του δείγματος στα στρώματα, επομένως θα έχουμε τις ακόλουθες περιπτώσεις:

- Εάν η **κατανομή του δείγματος γίνεται αυθαίρετα** στα διάφορα στρώματα, τότε μια αμερόληπτη εκτίμηση της διακύμανσης του $\bar{x}$ δίνεται από τη σχέση :

$$V(\bar{x}) = Var\left(\sum_{i=1}^{k} \frac{N_i \bar{x}_i}{N} \right) = \frac{1}{N^2} \sum_{i=1}^{k} N_i^2 V(\bar{x}_i)$$

όπου

$$V(\bar{x}_i) = \frac{s_i^{\ 2}}{n_i} \left(\frac{N_i - n_i}{N_i} \right)$$

Επομένως

$$V(\bar{x}) = \frac{1}{N^2} \sum_{i=1}^{k} N_i^2 \frac{s_i^{\ 2}}{n_i} \left(\frac{N_i - n_i}{N_i} \right) = \sum_{i=1}^{k} W_i^{\ 2} \frac{s_i^{\ 2}}{n_i} \left(\frac{N_i - n_i}{N_i} \right) \qquad (1)$$

Αν η $s_i^{\ 2}$ είναι άγνωστη παίρνουμε την αμερόληπτη εκτίμησή της s_i^2, οπότε μια αμερόληπτη εκτίμηση της διακύμανσης του $\bar{x}$ είναι:

$$V(\bar{x}) = \sum_{i=1}^{k} W_i^{\ 2} \frac{s_i^2}{n_i} \left(\frac{N_i - n_i}{N_i} \right) = \sum_{i=1}^{k} \frac{W_i^{\ 2} s_i^{\ 2}}{n_i} - \sum_{i=1}^{k} \frac{W_i^{\ 2} s_i^{\ 2}}{N}$$

όπου

$$s_i^{\ 2} = \frac{\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}{n_i - 1}$$

η οποία υπολογίζεται από τα δεδομένα του δείγματος.

- Σε περίπτωση **αναλογικής κατανομής του δείγματος**, η διακύμανση της εκτίμησης του $\bar{x}$ προκύπτει από τη σχέση (1) αν αντικαθιστώντας το $n_i = n \frac{N_i}{N}$ οπότε έχουμε :

$$V(\bar{x}) = \sum_{i=1}^{k} W_i^2 \frac{s_i^{\ 2}}{n_i} \left(\frac{N_i - n_i}{N_i} \right) = \sum_{i=1}^{k} W_i^2 \frac{s_i^{\ 2}}{n \cdot N_i / N} \left(\frac{N_i - n \cdot N_i / N}{N_i} \right) =$$

$$= \sum_{i=1}^{k} W_i^2 \frac{s^{\ 2}}{n \cdot W_i} \left(\frac{N_i (N - n)}{N_i \cdot N} \right)$$

Δηλαδή: $V(\bar{x}) = \frac{1}{n} \left(\frac{N - n}{N} \right) \sum_{i=1}^{k} W_i s_i^{\ 2}$ (2)

Στην αναλογική κατανομή δείγματος, εάν οι διακυμάνσεις έχουν την ίδια τιμή s_w^2 σε όλα τα στρώματα, τότε η διακύμανση του $\overline{x}$ υπολογίζεται από τη σχέση:

$$V(\overline{x}) = \frac{s_w^2}{n} = \left(\frac{N-n}{N} \right)$$

Η εκτίμηση της διακύμανσης του $\overline{x}$, όταν η s_i^2 είναι άγνωστη, υπολογίζεται από την αμερόληπτη εκτίμηση αυτής s_i^2, οπότε:

$$V(\overline{x}) = \frac{1}{n}\left(\frac{N-n}{N} \right)\sum_{i=1}^{k} W_i s_i^2$$

- Στην περίπτωση **αρίστης κατανομής του δείγματος,** έχουμε δύο περιπτώσεις όταν το κόστος κατά μονάδα είναι ίδιο σε όλα τα στρώματα (κατανομή Neyman) και όταν το κόστος κατά μονάδα διαφέρει μεταξύ των στρωμάτων.

Η διακύμανση της εκτίμησης του $\overline{x}$ προκύπτει από τη σχέση (1) αντικαθιστώντας:

i. το

$$n_i = n\frac{W_i s_i}{\sum W_i s_i}$$

(Άριστη κατανομή με σταθερό κόστος κατά μονάδα-κατανομή Neyman)

Οπότε θα έχουμε:

$$V(\overline{x}) = \frac{\left(\sum_{i=1}^{k} W_i s_i \right)^2}{n} - \frac{\sum_{i=1}^{k} W_i s_i^2}{N} \qquad (3)$$

ii. το

$$n_i = n \frac{W_i s_i \big/ \sqrt{C_i}}{\sum \left(W_i s_i \big/ \sqrt{C_i} \right)}$$

<u>(Άριστη κατανομή με διαφορετικό κόστος ανά μονάδα)</u>

Οπότε θα έχουμε:

$$V(\bar{x}) = \frac{\left(\sum_{i=1}^{k} W_i s_i \sqrt{C_i} \right) \left(\sum_{i=1}^{k} W_i s_i \big/ \sqrt{C_i} \right)}{n} - \frac{\sum W_i s_i^2}{N} \qquad (4)$$

και τυπικό σφάλμα

$$SE(\bar{x}) = \sqrt{V(\bar{x})}$$

που ισχύει και για όλες τις παραπάνω περιπτώσεις. (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

4.5. Εκτίμηση Ποσοστού ή Αναλογίας Πληθυσμού

Εάν θέλουμε να εκτιμήσουμε κατά στρώμα το ποσοστό ή την αναλογία των μονάδων ενός πληθυσμού, οι οποίες ανήκουν σε μια συγκεκριμένη τάξη C, χωρίζουμε τον πληθυσμό σε στρώματα. Αν υποθέσουμε ότι A_i απο τις N_i μονάδες του στρώματος i , ανήκουν στην τάξη C, τότε το ποσοστό των μονάδων της τάξης C θα είναι $P_i = \frac{A_i}{N_i}$.

Το ποσοστό των μονάδων που δεν ανήκουν στην τάξη C θα είναι: $Q_i = 1 - P_i$.

Βέβαια στην πράξη το ποσοστό P_i συνήθως είναι άγνωστο και παίρνουμε τη δειγματική εκτίμηση αυτού $p_i = \frac{a_i}{n_i}$.

Το γινόμενο $N_i p_i$ εκφράζει το πλήθος των μονάδων της τάξης C εντός του στρώματος i.

Επομένως το άθροισμα : $N_1 p_1 + ... + N_k p_k = \sum_{i=1}^{k} N_i p_i$ είναι μια εκτίμηση του συνολικού αριθμού των μονάδων της τάξης C του πληθυσμού. Αν διαιρέσουμε το άθροισμα αυτό με το σύνολο του πληθυσμού N προκύπτει μια εκτίμηση του ποσοστού P του πληθυσμού:

$$p = \frac{1}{N}\sum_{i=1}^{k} N_i \, p_i$$

Αν η κατανομή του δείγματος *γίνεται* *<u>αυθαίρετα</u>* στα διάφορα στρώματα, η διακύμανση της εκτίμησης p προκύπτει από τη σχέση (1) που είναι $V(\bar{x}) = \sum_{i=1}^{k} W_i^2 \frac{s_i^2}{n_i}\left(\frac{N_i - n_i}{N_i}\right)$, αν αντικατασταθεί η διακύμανση s_i^2. Η διακύμανση του στρώματος i είναι ίση με $s_i^2 = \frac{N_i}{N_i - 1} P_i Q_i$ οπότε έχουμε:

$$V(p) = \sum_{i=1}^{k} W_i^2 \frac{N_i - n_i}{N_i - 1} \cdot \frac{P_i Q_i}{n_i}$$

Για N_i μεγάλο προκύπτει η απλούστερη σχέση:

$$V(p) = \sum_{i=1}^{k} W_i^2 \frac{P_i Q_i}{n_i}\left(\frac{N_i - n_i}{N_i}\right)$$

Και τυπικό σφάλμα: $SE(p) = \sqrt{V(p)}$

Σε περίπτωση που η διακύμανση του p υπολογίζεται απο στοιχεία δείγματος, αντικαθιστούμε το $P_i Q_i / n_i$ με $p_i q_i / n_i - 1$. Αρα έχουμε:

$$V(p_{st}) = \sum_{i=1}^{k} W_i^2 \frac{p_i q_i}{n_i - 1}\left(\frac{N_i - n_i}{N_i}\right)$$

Όπου $q_i = 1 - p_i$

Στην περίπτωση *αναλογικής κατανομής* η διακύμανση της εκτίμησης p προκύπτει απο τη σχέση $V(\bar{x})=\frac{1}{n}\left(\frac{N-n}{N}\right)\sum_{i=1}^{k}W_i s_i^2$, με αντικατάσταση της s_i^2 και για N_i μεγάλο έχουμε:

$$V(p)=\frac{1}{n}\left(\frac{N-n}{n}\right)\sum_{i=1}^{k}W_i P_i Q_i$$

Και τυπικό σφάλμα το ίδιο.

Στην περίπτωση *άριστης κατανομής* η διακύμανση της εκτίμησης p προκύπτει απο τη σχέση $V(\bar{x})=\sum_{i=1}^{k}W_i^2\frac{s_i^2}{n_i}\left(\frac{N_i-n_i}{N_i}\right)$, αντικαθιστώντας για την κατανομή Neyman το $n_i=n\frac{N_i\sqrt{P_iQ_i}}{\sum N_i\sqrt{P_iQ_i}}$ και για την άριστη κατανομή με διαφορετικό κόστος κατά μοναδα το $n_i=n\frac{N_i\sqrt{P_iQ_i/C_i}}{\sum N_i\sqrt{P_iQ_i/C_i}}$.

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997).

4.6. Στρωματοποίηση μετά τη Συλλογή του Δείγματος

Κάποιες φορές δεν είναι δυνατό να γνωρίζουμε εκ των προτέρων σε ποιό στρώμα ανήκει μια δειγματοληπτική μονάδα. Γι'αυτόν ακριβώς το λόγο η στρωματοποίηση του ερευνώμενου πληθυσμού δεν μπορεί να γίνει πριν τη συλλογή του δείγματος .

Σε αυτές τις περιπτώσεις, συλλέγεται το δείγμα με απλή τυχαία δειγματοληψία, διαιρείται σε στρώματα και έπειτα θεωρούμε οτι το δείγμα έχει συλλεχθεί μέσω στρωματοποιημένης δειγματοληψίας.

Ο μέσος του πληθυσμού m θα εκτιμηθεί από τη σχέση:

$$\bar{x}=\sum_{i=1}^{k}W_i\bar{x}_i$$

δεδομένου ότι το κλάσμα $\frac{N_i}{N}$ είναι γνωστό. Αν το $\frac{N_i}{N}$ είναι γνωστό και το δείγμα $n_i \geq 20$ σε κάθε στρώμα, τότε η μέθοδος στρωματοποίησης μετά τη συλλογή του δείγματος έχει περίπου την ακρίβεια της αναλογικής στρωματοποιημένης δειγματοληψίας.

Η διακύμανση της εκτίμησης $\bar{x}$ στην περίπτωση αυτή δίνεται από τη σχέση:

$$V(\bar{x}) = \frac{1}{n}\left(\frac{N-n}{N}\right)\sum_{i=1}^{k} W_i s_i^2 + \frac{1}{n^2}\sum_{i=1}^{k}(1-W_i)s_i^2$$

Ο πρώτος όρος της παραπάνω σχέσης είναι η διακύμανση του $\bar{x}$ στην αναλογική στρωματοποιημένη δειγματοληψία. Ο δεύτερος όρος παρουσιάζει την αύξηση της διακύμανσης που οφείλεται στη στρωματοποίηση μετά τη συλλογή του δείγματος. Επειδή ο δεύτερος όρος είναι μικρός σε σχέση με τον πρώτο, όταν το n είναι μεγάλο, βλέπουμε ότι η στρωματοποίηση μετά τη συλλογή δείγματος δίνει περίπου την ακρίβεια της αναλογικής στρωματοποιημένης δειγματοληψίας.

Μπορεί να αποδειχθεί ότι μια αμερόληπτη εκτιμήτρια της διακύμανσης είναι:

$$V(\bar{x}) = \frac{N-n}{N}\sum_{i=1}^{k} W_i^2 \frac{s_i^2}{n_i}$$

Η μέθοδος της στρωματοποίησης μετά τη συλλογή του δείγματος, μπορεί να εφαρμοσθεί σε ένα δείγμα που ήδη έχει γίνει στρωματοποίηση ως προς κάποια γνωστή μεταβλητή.

4.7. Πλεονεκτήματα – Μειονεκτήματα Στρωματοποιημένης Δειγματοληψίας

Η στρωματοποιημένη δειγματοληψία είναι περισσότερο αποτελεσματική της απλής τυχαίας ή της συστηματικής δειγματοληψίας, όταν χρησιμοποιείται σε ανομοιογενείς ή με μεγαλη ασυμμετρία πληθυσμούς. Σκοπός της είναι να χωρίσει τον ανομοιογενή πληθυσμό σε ομοιγενείς υποπληθυσμούς.

Η στρωματοποιημένη δειγματοληψία παρουσιάζει μία σειρά πλεονεκτημάτων σε σχέση με την απλή τυχαία ή την συστηματική δειγματοληψία τα οποία παρουσιάζονται παρακάτω:

1. Μπορεί να δώσει μικρότερο σφάλμα εκτιμήσεων, δηλαδή μεγαλύτερη ακρίβεια, απ’ ότι η απλή τυχαία δειγματοληψία για ίσου μεγέθους δείγματα, λόγω του ότι οι μετρήσεις από ομοιογενή στρώματα διαφέρουν ελάχιστα.

2. Το κόστος δειγματοληψίας μπορεί να μειωθεί σημαντικά με τη στρωματοποίηση του πληθυσμού σε ομοιογενείς υποπληθυσμούς, διότι είναι δυνατόν να περιορισθεί το μέγεθος του δείγματος στους επιμέρους ομοιογενείς υποπληθυσμούς.

3. Η δυνατότητα να χρησιμοποιήσουμε μέσα στο κάθε στρώμα διαφορετικό κλάσμα δειγματοληψίας, διαφορετικό δειγματοληπτικό σχέδιο, ακόμα και διαφορετική μέθοδο συλλογής στοιχείων. Επιπλέον μας δίνει αποτελέσματα με την ακρίβεια που επιθυμούμε για συγκεκριμένους υποπληθυσμούς που μας ενδιαφέρουν, δεδομένου οτι κάθε στρώμα μπορούμε να το μελετήσουμε σαν έναν επιμέρους πληθυσμό. Γνωρίζοντας ότι η αναλογία του υποπληθυσμού που μας ενδιαφέρει είναι πολύ μικρή, η απλή τυχαία δειγματοληψία θα έδινε ίσως τόσες λίγες μονάδες που δεν θα ήταν δυνατόν να τις αξιοποιήσουμε στατιστικά. Επομένως, με τη μέθοδο των στρωμάτων έχουμε τη δυνατότητα να αυξήσουμε τον αριθμό των δειγματικών μονάδων του υποπληθυσμού αυτού χωρίς να αλλοιώσουμε τα συνολικά αποτελέσματα.

4. Με την επιλογή δείγματος από καθε στρώμα χωριστά, το συνολικό δείγμα αντιπροσωπεύει καλύτερα τον συνολικό πληθυσμό απ΄ότι αν λαμβάναμε ένα απλό τυχαίο δείγμα (βλ. Ξενάκης,2002).

5. Η στρωματοποίηση δίνει τη δυνατότητα αντιμετώπισης των διαφορών που υπάρχουν μεταξύ πληθυσμιακών ομάδων. Για παράδειγμα οι απαιτήσεις των ατόμων που ζουν σε συλλογικές κατοικίες (π.χ. νοσοκομεία, οικοτροφεία) είναι διαφορετικές απο αυτές των ατόμων που ζουν σε κανονικές κατοικίες.

6. Σε δειγματοληψία γεωγραφικών επιφανειών ο πληθυσμός στρωματοποιείται στις γεωγραφικές περιοχές που εμπίπτει, ακόμα και αν δεν παρατηρείται ομοιογένεια των τιμών κατά γεωγραφική περιοχή. Έτσι εκτός του ότι παρέχονται εκτιμήσεις και κατά γεωγραφική περιοχή, η λήψη των στοιχείων διενεργείται πιο σύντομα και με μικρότερο κόστος διότι διενεργείται ταυτόχρονα σε όλα τα στρώματα από διαφορετικούς ερευνητές, οπότε τα στοιχεία ειναι άμεσα συγκρίσιμα και διότι αποφεύγονται οι δαπάνες μεγάλων μετακινήσεων των ερευνητών από περιοχή σε περιοχή (βλ. Ζαϊρης Εμμ.Ποσειδώνας 1991).

Παρά το ότι η στρωματοποιημένη δειγματοληψία είναι μία μέθοδος που παρουσιάζει αρκετά πλεονεκτήματα, είναι αρκετά δύσκολο ο πληθυσμός να διαχωριστεί σε κατάλληλα στρώματα. Ακόμα και αν γίνει αυτό με επιτυχία, είναι πιο σύνθετη μέθοδος για την οργάνωση και την ανάλυση αποτελεσμάτων.

Επίσης, μερικές φορές υπάρχει μεγαλύτερη μεταβλητότητα σε μερικά στρώματα έναντι κάποιων άλλων. Σε αυτή την περίπτωση, ένα μεγαλύτερο δείγμα πρέπει να προέλθει από εκείνα τα στρώματα με τη μεγαλύτερη μεταβλητότητα.

4.8. Παράδειγμα Στρωματοποιημένης Δειγματοληψίας

Έστω ότι θέλουμε να εκτιμήσουμε τη μέση ετήσια κατανάλωση ενός προϊόντος των νοικοκυριών μιας περιοχής, όπου διαμένουν 2.000 νοικοκυριά. Για το σκοπό αυτό εφαρμόζουμε την τυχαία στρωματοποιημένη δειγματοληψία. Έτσι το σύνολο των νοικοκυριών (πληθυσμός) χωρίζεται σε τρία στρώματα (υποπληθυσμοί): αστικά, ημιαστικά και αγροτικά νοικοκυριά. Είναι λογικό να χρησιμοποιήσουμε την στρωματοποιημένη δειγματοληψία έναντι της απλής τυχαίας δειγματοληψίας ή της συστηματικής διότι η στρωματοποίηση δίνει τη δυνατότητα αντιμετώπισης των διαφορών που υπάρχουν μεταξύ των νοικοκυριών. Δηλαδή οι απαιτήσεις των αστικών νοικοκυριών ειναι διαφορετικές από αυτές των αγροτικών και των ημιαστικών.

Επιπλέον το κόστος δειγματοληψίας μπορεί να μειωθεί σημαντικά με τη στρωματοποίηση του πληθυσμού σε ομοιογενείς υποπληθυσμούς, διότι περιορίζεται το μέγεθος του δείγματος στους επιμέρους ομοιογενείς υποπληθυσμούς. Έστω ότι από προηγούμενη έρευνα γνωρίζουμε τον αριθμό νοικοκυριών και τη διακύμανση της μέσης ετήσιας κατανάλωσης κατά στρώμα.

Στρώμα	Περιοχή	Αριθμός νοικοκυριών	Διακύμανση μέσης ετήσιας κατανάλωσης s_i^2
1ο	Αστική	1.200	25
2ο	Ημιαστική	200	16
3ο	Αγροτική	600	4
Σύνολο		2.000	

Λαμβάνοντας υπόψη μας ότι το κόστος κατά μονάδα δείγματος είναι ίδιο σε όλα τα στρώματα, κατανέμουμε δείγμα 40 μονάδων με την άριστη κατανομή (κατανομή Neyman). Δηλαδή χρησιμοποιούμε τη σχέση $n_i = n\frac{N_i s_i}{\sum N_i s_i}$, οπότε με βάση τα παραπάνω δεδομένα παίρνουμε 30 μονάδες από το πρώτο στρώμα, 4 μονάδες από το δεύτερο και 6 μονάδες από το τρίτο, ώστε να επιτευχθεί ελαχιστοποιήση της διακύμανσης της εκτίμησης.

Από τη δειγματοληπτική έρευνα που έγινε, βρέθηκαν τα παρακάτω μεγέθη μέσης ετήσιας κατανάλωσης του προϊόντος σε κιλά κατά νοικοκυριό, στα τρία στρώματα:

1^{ο} στρώμα	28,36,35,38,27,32,39,29,37,28, 38,26,31,35,37,39,35,36,27,31, 27,36,34,39,40,32,36,28,25,29
2^{ο} στρώμα	24,27,30,21
3^{ο} στρώμα	22,18,20,24,19,23

Από τα αριθμητικά μεγέθη του δείγματος υπολογίζουμε τους αριθμητικούς μέσους και τις διακυμάνσεις κατά στρώματα:

$$\bar{x}_1 = \frac{\sum_{j=1}^{n_i=1} x_{1j}}{n_1} = \frac{28+36+35+....+29}{30} = \frac{990}{30} = 33$$

$$\bar{x}_2 = \frac{\sum_{j=1}^{n_2} x_{2j}}{n_2} = \frac{24+27+30+21}{4} = \frac{102}{4} = 25{,}5$$

$$\bar{x}_3 = \frac{\sum_{j=1}^{n_3} x_{3j}}{n_3} = \frac{22+18+20+24+19+23}{6} = \frac{126}{6} = 21$$

$$s_1^2 = \frac{\sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2}{n_1 - 1} = \frac{(28-33)^2 + (36-33)^2 + ... + (29-33)^2}{30-1} = 19{,}86$$

$$s_2^2 = \frac{\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)}{n_2 - 1} = \frac{(24-25{,}5)^2 + ... + (21-25{,}5)^2}{4-1} = 15$$

$$s_3^2=\frac{\sum_{j=1}^{n_3}(x_{3j}-\overline{x}_3)^2}{n_3-1}=\frac{(22-21)^2+...+(23-21)^2}{6-1}=5,6$$

Η εκτίμηση της μέσης ετήσιας κατανάλωσης του προϊόντος σε κιλά κατά νοικοκυριό στην περιοχή είναι:

$$\overline{x}=\frac{1}{N}\sum_{i=1}^{k}N_i\overline{x}_i=\frac{1}{N}(N_1\overline{x}_1+N_2\overline{x}_2+N_3\overline{x}_3)=$$
$$=\frac{1}{2000}(1200\cdot 33+200\cdot 25,5+600\cdot 21)=28,65$$

Η διακύμανση της εκτίμησης $\overline{x}$ υπολογίζεται από τη σχέση (3) (κατανομή Neyman), δηλαδή:

$$V(\overline{x})=\frac{\left(\sum_{i=1}^{k}W_i s_i\right)^2}{n}-\frac{\sum_{i=1}^{k}W_i s_i^2}{N}=$$
$$=\frac{(W_1 s_1+W_2 s_2+W_3 s_3)^2}{n}-\frac{W_1 s_1^2+W_2 s_2^2+W_3 s_3^2}{N}=$$
$$=\frac{(0,60\cdot 5+0,10\cdot 4+0,30\cdot 2)^2}{40}-\frac{0,60\cdot 25+0,10\cdot 16+0.30\cdot 4}{2000}=0,39$$

Το τυπικό σφάλμα της εκτίμησης $\overline{x}$ είναι:

$$SE(\overline{x})=\sqrt{V(\overline{x})}=\sqrt{0,39}=0,62$$

Ο συντελεστής μεταβλητικότητας είναι:

$$CV(\overline{x})=\frac{SE(\overline{x})}{\overline{x}}=\frac{0,62}{28,65}=0,02=2\%$$

Το διάστημα εμπιστοσύνης στο οποίο βρίσκεται η μέση ετήσια κατανάλωση του προϊόντος των νοικοκυριών με πιθανότητα 95% είναι:

$$\overline{x}\pm Z_{a/2}\cdot SE(\overline{x})$$

ή

$$28{,}65 \pm 1.96 \cdot 0{,}62$$

ή

$$28{,}65 \pm 1{,}22$$

δηλαδή η μέση τιμή βρίσκεται στο διάστημα 27,43 – 29,87 με πιθανότητα 95% και με την προϋπόθεση ότι η εκτίμηση $\overline{x}$ ακολουθεί την κανονική κατανομή.

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

ΚΑΤΑ ΟΜΑΔΕΣ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

5.1. Εισαγωγή

Η δειγματοληψία κατά ομάδες είναι μια απλή τυχαία δειγματοληψία στην οποία κάθε δειγματοληπτική μονάδα είναι μια ομάδα στοιχείων του ερευνώμενου πληθυσμού. Για παράδειγμα αν θέλουμε να εκτιμήσουμε το μέσο εισόδημα των νοικοκυριών μιας πόλης, μπορούμε να ομαδοποιήσουμε όλα τα νοικοκυριά κατά οικοδομικό τετράγωνο. Έτσι το σύνολο των νοικοκυριών ισοδυναμεί με το σύνολο των οικοδομικών τετραγώνων, οπότε παίρνουμε τυχαία έναν αριθμό οικοδομικών τετραγώνων και ερευνούμε το σύνολο των νοικοκυριών που βρίσκονται στα οικοδομικά τετράγωνα του δείγματος (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος,1997).

Με τον τρόπο αυτό συντομεύεται ο χρόνος διεξαγωγής της έρευνας αλλά κυρίως με τη δειγματοληψία κατά ομάδες μειώνεται το κόστος. Συνήθως οι ομάδες αποτελούνται από μονάδες του πληθυσμού που ειναι γεωγραφικά εντοπισμένες με αποτέλεσμα να απαιτούνται λιγότερες μετακινήσεις και ερευνητές για τη διερεύνηση τους.

Με την κατά ομάδες δειγματοληψία λύνουμε πιο εύκολα το πρόβλημα της κατασκευής πλαισίου. Δηλαδή μπορεί να εφαρμοστεί όταν δεν υπάρχουν τα απαιτούμενα δειγματοληπτικά πλαίσια (π.χ. κατάλογοι νοικοκυριών μιας πόλης), ενώ αντίθετα υπάρχουν πλαίσια για ομάδες στοιχείων (π.χ. οικοδομικά τετράγωνα μιας πόλης)

Ομάδες στοιχείων, μπορεί να αποτελέσουν τα οικοδομικά τετράγωνα, όταν ως μονάδα του ερευνώμενου πληθυσμού είναι το νοικοκυριό, τα σχολεία, η κατοικία, το άτομο κ.ο.κ.

Τα στοιχεία μέσα σε μια ομάδα πρέπει ιδανικά να είναι όσο το δυνατόν πιο ετερογενή. Αυτό εχει σαν αποτέλεσμα τη μείωση της διακύμανσης και του τυπικού σφάλματος της μέσης τιμής. Στην πράξη όμως, μεταξύ των στοιχείων των ομάδων παρουσιάζεται ομοιογένεια ως προς τα χαρακτηριστικά που ερευνόνται. Επίσης, οι ομάδες θα πρέπει να είναι κατά το δυνατόν μικρού μεγέθους. Είναι προτιμότερο δηλαδή να έχουμε, για το ίδιο μέγεθος δείγματος, πολλές μικρές ομάδες, αντί για λίγες και μεγάλες. Τέλος, οι ομάδες θα πρέπει να είναι όσο το δυνατόν πιο ισομεγέθεις. Σε αντίθετη περίπτωση, ο εκτιμητής της μέσης τιμής είναι μεροληπτικός.

Είναι σημαντικό να μη γίνεται σύγχηση της διαίρεσης του πληθυσμού σε στρώματα με τη διαίρεσή του σε ομάδες. Ένα στρώμα αποτελεί έναν υποπληθυσμό από τον οποίο θα πάρουμε με τυχαίο τρόπο ορισμένες μονάδες. Επιπλέον πρέπει να χρησιμοποιήσουμε όλα τα στρώματα στα οποία διαιρέσαμε τον πληθυσμό. Στην κατά ομάδες δειγματοληψία θα κληρώσουμε ορισμένες από τις ομάδες και θα διερευνήσουμε όλες τις μονάδες τους. Επίσης το στρώμα προϋποθέτει ομοιογένεια των μονάδων που περιλαμβάνονται σ΄αυτό, ενώ αντίθετα στη δειγματοληψία κατά ομάδες, πρέπει να υπάρχει μέσα σε κάθε ομάδα η ανομοιογένεια που υπάρχει στον ερευνώμενο πληθυσμό, ώστε να προκύπτουν αξιόπιστα αποτελέσματα.

Τέλος, στη στρωματοποιημένη δειγματοληψία προσπαθούμε να πετύχουμε το πιο αξιόπιστο δείγμα σε σχέση με το μέγεθος του, ενώ στην κατά ομάδες δειγματοληψία σκοπός είναι να πετύχουμε τη μεγαλύτερη ακρίβεια σε σχέση με το κόστος ανά στοιχειώδη μονάδα δείγματος (βλ. Ξενάκης, 2002).

5.2. Εκτίμηση του Μέσου Πληθυσμού

Στην ουσία η δειγματοληψία κατά ομάδες είναι μια απλή τυχαία δειγματοληψία στην οποία κάθε δειγματοληπτική μονάδα αποτελείται από μια ομάδα στοιχείων του πληθυσμού, έτσι, η εκτίμηση του μέσου πληθυσμού γίνεται με τον ίδιο τρόπο που γίνεται και στην απλή τυχαία δειγματοληψία. Επομένως μια αμερόληπτη εκτίμηση του μέσου m του πληθυσμού είναι ο μέσος του δείγματος και συμβολίζεται με $\overline{x}$.

Δηλαδή έχουμε:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} m_i}$$

Η διακύμανση της εκτίμησης του $\overline{x}$ είναι:

$$V(\overline{x}) = \frac{1-f}{n\overline{M}^2} s^2 = \frac{1-n/N}{n\overline{M}^2} s^2 = \frac{\frac{N-n}{N}}{n\overline{M}^2} s^2 \Rightarrow$$

$$V(\overline{x}) = \frac{N-n}{Nn\overline{M}^2} s^2$$

Όπου $s^{2} = \dfrac{\sum_{i=1}^{N}(x_i - m \cdot m_i)^2}{N-1}$, σε περίπτωση που το s^{2} είναι άγνωστο τότε

χρησιμοποιείται η εκτιμήτρια που είναι: $s^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x}m_i)^2}{n-1}$

Η εκτίμηση του τυπικού σφάλματος έιναι: $SE(\bar{x}) = \sqrt{V(\bar{x})}$ (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997).

5.3. Εκτίμηση Ποσοστού ενός Πληθυσμού

Η εκτίμηση ποσοστού ενός πληθυσμού με τη μέθοδο της δειγματοληψίας κατά ομάδες, γίνεται όπως και στην απλή τυχαία δειγματοληψία από το ποσοστό των στοιχείων του δείγματος που ανήκουν σε μια συγκεκριμένη κατηγορία, η οποία μας ενδιαφέρει να μελετήσουμε, δηλαδη έχουμε:

$$p = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} m_i}$$

Στη συνέχεια η διακύμανση της εκτίμησης του p είναι : $V(p) = \dfrac{(N-n)}{Nn\bar{M}^2}s^2$

Όπου $s^2 = \dfrac{\sum_{i=1}^{n}(a_i - pm_i)^2}{n-1}$

Η εκτίμηση του τυπικού σφάλματος είναι: $SE(p) = \sqrt{V(p)}$

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

5.4. Πλεονεκτήματα – Μειονεκτήματα Δειγματοληψίας Κατά Ομάδες

Η δειγματοληψία κατά ομάδες είναι μία αποτελεσματική μέθοδος η οποία συνήθως εφαρμόζεται όταν το πλήθος των στοιχείων του πληθυσμού είναι γεωγραφικά εντοπισμένο μεν, αλλά άγνωστο και άρα οι άλλες μέθοδοι που περιγράψαμε παραπάνω δεν μπορούν να χρησιμοποιηθούν. Επίσης δεν είναι απαραίτητο να γνωρίζουμε το πλήθος των στοιχείων των ομάδων, παρά μόνο αυτών που θα επιλεγούν τελικά στο δείγμα. Μπορεί επίσης να εφαρμοστεί όταν δεν υπάρχουν τα απαιτούμενα δειγματοληπτικά πλαίσια για όλα τα στοιχεία του πληθυσμού, ενώ αντίθετα υπάρχουν πλαίσια για ομάδες στοιχείων. Με αυτόν τον τρόπο, ο χρόνος διεξαγωγής της έρευνας περιορίζεται σημαντικά αλλά κυρίως μειώνεται το κόστος σε σχέση με τις άλλες μεθόδους.

Το μειονέκτημα που παρουσιάζει η κατά ομάδες δειγματοληψία είναι ότι η ομαδοποίηση του πληθυσμού έχει σαν αποτέλεσμα οι μονάδες κάθε ομάδας να παρουσιάζουν μεγάλη ομοιογένεια, με αποτέλεσμα να αυξάνεται η διασπορά των εκτιμήσεων και έτσι το δείγμα μας να είναι λιγότερο αντιπροσωπευτικό του πληθυσμού απ'όσο θα ήταν αν εφαρμόζαμε οποιαδήποτε άλλη μέθοδο με το ίδιο μέγεθος δείγματος.

Για ένα δεδομένο όμως προϋπολογισμό κόστους, με την κατά ομάδες δειγματοληψία έχουμε τη δυνατότητα να μελετήσουμε περισσότερες μονάδες του πληθυσμού. Από αυτή την άποψη, αν η μέθοδος χρησιμοποιηθεί σωστά, μπορούμε να έχουμε καλύτερα αποτελέσματα, από ότι αν εφαρμόζαμε, για παράδειγμα απλή τυχαία δειγματοληψία με αυτό το κόστος (βλ. Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής, 2000).

5.5. Παράδειγμα Δειγματοληψίας Κατά Ομάδες

Έστω ότι θέλουμε να εκτιμήσουμε τη μέση μηνιαία κατανάλωση νερού σε κυβικά μέτρα (m^3) των νοικοκυριών μιας πόλης, όπου διαμένουν 3.600 νοικοκυριά. Συνήθως είναι αδύνατο να βρεθεί ένας κατάλογος με καταχωρημένα και αριθμημένα όλα τα νοικοκυριά. Σε αυτή την περίπτωση, εφαρμόζουμε δειγματοληψία κατά ομάδες. Έτσι η συγκεκριμένη πόλη διαιρείται σε οικοδομικά τετράγωνα και κάθ' ένα απ' αυτά αποτελεί μια ομάδα νοικοκυριών. Με αυτόν τον τρόπο, ο χρόνος διεξαγωγής της έρευνας περιορίζεται σημαντικά αλλά κυρίως μειώνεται το κόστος σε σχέση με τις άλλες μεθόδους.

Το σύνολο των οικοδομικών τεραγώνων της πόλης είναι 200. Από το σύνολο των οικοδομικών τετραγώνων επελέγησαν 10 με τυχαίο τρόπο. Από την έρευνα προέκυψαν τα αποτελέσματα που παρουσιάζονται στον παρακάτω πίνακα.

Ομάδα i (οικοδομικό τετράγωνο)	Αριθμός νοικοκυριών n_i	Συνολική κατανάλωση νερού (σε m^3) κατά ομάδα x_i
1	10	1.200
2	15	1.900
3	20	2.500
4	12	1.300
5	25	3.000
6	18	2.400
7	16	1.980
8	21	2.700
9	24	2.900
10	17	2.350
Σύνολο	178	22.230

Η εκτίμηση της μέσης κατανάλωσης νερού των νοικοκυριών της πόλης είναι:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} n_i} = \frac{22.230}{178} = 124{,}88 \quad m^3$$

Η διακύμανση του $\bar{x}$ είναι:

$$V(\bar{x}) = \frac{N-n}{Nn\bar{M}^2} s^2$$

επειδή M είναι γνωστό έχουμε

$$\bar{M} = \frac{N}{M} = \frac{3.600}{200} = 18$$

Καταρτίζουμε τον παρακάτω πίνακα υπολογισμών:

Ομάδα i	x_i	$\bar{x}n_i$	$(x_i - \bar{x}n_i)^2$
1	1.200	1249	2.401
2	1.900	1873	729
3	2.500	2498	4
4	1.300	1498	39.204
5	3.000	3122	14.884
6	2.400	2248	23.104
7	1.980	1998	324
8	2.700	2622	6.084
9	2.900	2997	9.409
10	2.350	2123	51.529
Σύνολο	22.230		147.672

οπότε

$$V(\bar{x}) = \frac{N-n}{Nn\bar{M}^2}\frac{\sum_{i=1}^{n}(x_i - \bar{x}n_i)^2}{n-1} = \frac{200-10}{200\cdot 10\cdot 18^2}\frac{147.672}{10-1} = 4{,}8$$

Το τυπικό σφάλμα είναι:

$$SE(\bar{x}) = \sqrt{V(\bar{x})} = \sqrt{4{,}8} = 2{,}2$$

Ο συντελεστής μεταβλητικότητας είναι:

$$CV(\bar{x}) = \frac{SE(\bar{x})}{\bar{x}} = \frac{2{,}2}{124{,}88} = 0.0188 \text{ ή } 18\%$$

Το διάστημα εμπιστοσύνης στο οποίο βρίσκεται η μέση μηνιαία κατανάλωση νερού των νοικοκυριών της πόλης με πιθανότητα 95% είναι:

$$\bar{x} \pm Z_{a/2}\cdot SE(\bar{x})$$

ή

$$124{,}88 \pm 1{,}96\cdot 2{,}2$$

ή

$$124{,}88 \pm 4{,}3$$

δηλαδή το m βρίσκεται στο διάστημα από 120,58 έως 129,18 με πιθανότητα 95% και με την προϋπόθεση ότι η εκτίμηση $\bar{x}$ ακολουθεί την κανονική κατανομή.

<u>ΚΕΦΑΛΑΙΟ ΈΚΤΟ</u>

ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΜΕΓΕΘΟΥΣ ΔΕΙΓΜΑΤΟΣ

6.1. Προσδιοριμός Μεγέθους Δείγματος για την Εκτίμηση Μέσου Πληθυσμού

Το πρώτο πράγμα που πρέπει να γίνει πριν ξεκινήσει μία δειγματοληπτική έρευνα είναι να οριστεί το μέγεθος δείγματος n. Γι' αυτό πρέπει να δοθεί κάποιο περιθώριο σφάλματος d απόκλισης της εκτιμήτριας από την εκτιμώμενη μέση τιμή του πληθυσμού καθώς και η αξιοπιστία ή εμπιστοσύνη με την οποία θα γίνει η εκτίμηση. Εκλογή μεγαλύτερου δείγματος από το απαιτούμενο οδηγεί ασφαλώς σε σπατάλη χρόνου και χρήματος. Από την άλλη μεριά, εκλέγοντας μικρότερο αριθμό μονάδων στο δείγμα σημαίνει ότι επιλέγουμε ανεπαρκή πληροφορία για την εκτίμηση της παραμέτρου (του πληθυσμού) που ενδιαφερόμαστε. Στην περίπτωση αυτή δεν είμαστε σε θέση να υποστηρίξουμε ότι τα αποτελέσματά μας θα είναι αξιόπιστα, ότι δηλαδή οι παράμετροι έχουν εκτιμηθεί με το προκαθορισμένο περιθώριο σφάλματος και τη δεδομένη αξιοπιστία εκτίμησης.

Είναι προφανές ότι το απαιτούμενο μέγεθος δείγματος n, πέρα από το περιθώριο σφάλματος, την αξιοπιστία της εκτίμησης, την ομοιογένεια του δείγματοληπτούμενου πληθυσμού και τη μέθοδο δειγματοληψίας, θα εξαρτάται επίσης και από τη συγκεκριμένη παράμετρο που θέλουμε να εκτιμήσουμε.

Η επιθυμητή ακρίβεια μιας εκτίμησης καθορίζεται από το επιτρεπτό όριο δειγματοληπτικού σφάλματος. Συνήθως ο προσδιορισμός του επιτρεπτού ορίου δειγματοληπτικού σφάλματος γίνεται σε κάποια όρια διαστήματος εμπιστοσύνης και σε ορισμένα επίπεδα σημαντικότητας (π.χ. 95% , 90% κ.ο.κ.). Για παράδειγμα αν θέλουμε να εκτιμήσουμε τον μέσο όρο του πληθυσμού με δειγματοληπτικό σφάλμα $|\bar{x}-m|$ το οποίο δεν πρέπει να υπερβεί ενα περιθώριο d και με συντελεστή εμπιστοσύνης $\{1-a\}$, δηλαδή αν θέλουμε:

$$P\{|\bar{x}-m| \leq d\} = 1-a$$

τότε το απαιτούμενο μέγεθος δείγματος προκύπτει με τον ακόλουθο τρόπο:

Το διάστημα εμπιστοσύνης για το μέσο m ενός πληθυσμού που ακολουθεί την κανονική κατανομή , με συντελεστή εμπιστοσύνης $1-a$ είναι:

$$\bar{x} \pm Z_{a/2} \cdot SE(\bar{x}),$$

όπου $SE(\bar{x})$ είναι το τυπικό σφάλμα και δίνεται απο τη σχέση:

$$SE(\bar{x}) = \sqrt{V(\bar{x})}\,.$$

Από το διάστημα εμπιστοσύνης προκύπτει οτι η ακρίβεια της εκτίμησης του μέσου m του πληθυσμού εξαρτάται απο το τυπικό σφάλμα. Δηλαδή όσο πιο μεγάλο είναι το τυπικό σφάλμα τόσο πιο μεγάλο θα είναι και το διάστημα εμπιστοσύνης και κατά συνέπεια θα έχουμε πιο μικρή ακρίβεια της εκτίμησης του μέσου m. Η ακρίβεια της εκτίμησης αυτής μπορεί να αυξηθεί αν αυξηθεί το μέγεθος του δείγματος (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Για να ισχύει ο περιορισμός ότι το δειγματοληπτικό σφάλμα δεν πρέπει να υπερβεί το περιθώριο d θέτουμε :

$$d = Z_{a/2} \cdot SE(\bar{x})$$

Όπου συγκεκριμένα για την ***απλή τυχαία δειγματοληψία*** το τυπικό σφάλμα δίνεται από τη σχέση:

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}}$$

Επομένως προκύπτει το εξής:

$$d = Z_{a/2}\,\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N}} \qquad \text{σχέση (α)}$$

Λύνοντας την εξίσωση ως προς n και δεδομένου ότι το κλάσμα δειγματοληψίας δεν ξεπερνά το 5% δηλαδή $f = \frac{n}{N} < 0{,}05 \Leftrightarrow n < 0{,}05N$ και άρα παραλείπεται ο συντελεστής διόρθωσης πεπερασμένου πληθυσμού $\frac{N-n}{N}$, το μέγεθος n θα είναι το εξής:

$$n = \frac{Z^2{}_{a/2} \cdot s^2}{d^2} \Leftrightarrow \sqrt{n} = \frac{Z_{a/2} \cdot s}{d}$$

Στις περισσότερες περιπτώσεις το s^2 είναι άγνωστο, τότε μπορούμε να το εκτιμήσουμε από στοιχεία πιλοτικής έρευνας, λαμβάνοντας ενα μικρό δείγμα από τον ερευνώμενο πληθυσμό, ή να χρησιμοποιήσουμε τη διακύμανση από προηγούμενη παρόμοια έρευνα (βλ. Ξενάκης Ανδρέας , 2002).

Στην περίπτωση προσδιορισμού του μεγέθους δείγματος για την εκτίμηση μέσου στη ***<u>συστηματική δειγματοληψία</u>*** ισχύει το εξής:

Επειδή οι μονάδες του πληθυσμού είναι διατεταγμένες σε μια λίστα τυχαία δηλαδή ο πληθυσμός είναι τυχαίος, η διακύμανση του μέσου $\bar{x}$ στην συστηματική δειγματοληψία είναι ισοδύναμη με τη διακύμανση του $\bar{x}$ (στην απλή τυχαία δειγματοληψία), οπότε το μέγεθος του δείγματος μπορεί να προσδιορισθεί από τη σχέση που χρησιμοποιείται και στην περίπτωση της απλής τυχαίας δειγματοληψίας (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997).

Επομένως το απαιτούμενο *μέγεθος δείγματος* n για την εκτίμηση μέσου προσδιορίζεται όμοια από τη σχέση (α) και λύνοντας ως προς n θα έχουμε:

$$n = \frac{\left(\dfrac{Z_{a/2}^2 \cdot s^2}{d^2}\right)}{1 + \dfrac{1}{N}\left(\dfrac{Z_{a/2}^2 \cdot s^2}{d^2}\right)}$$

Στην περίπτωση της ***<u>στρωματοποιημένης δειγματοληψίας</u>*** για να προσδιορίσουμε το μέγεθος του δείγματος που απαιτείται, ώστε να πετύχουμε με καθορισμένο βαθμό ακρίβειας μια εκτίμηση του πληθυσμιακού μέσου, πρέπει να έχουμε μια τιμή της διακύμανσης κατά στρώμα i, εκ των προτέρων κατά προσέγγιση. Μια τιμή της διακύμανσης κατά στρώμα i, μπορεί να προσδιορισθεί απο στοιχεία προηγούμενης έρευνας ή απο διεξαγωγή μιας πιλοτικής έρευνας (βλ. Ξενάκης Ανδρέας ,2002).

Στη συγκεκριμένη μέθοδο ο προσδιορισμός του μεγέθους δείγματος για την εκτίμηση μέσου διαφέρει απο κατηγορία σε κατηγορία δηλαδή:

Για την **Αυθαίρετη Κατανομή** ισχύει:

Εκτίμηση μέσου

$$d = z_{a/2} \cdot SE(\bar{x}) \text{ ή } d = z_{a/2}\sqrt{\sum_{i=1}^{k} W_i^2 \frac{s_i^2}{n_i}\left(\frac{N_i - n_i}{N_i}\right)}$$

Και $$n = \frac{\sum_{i=1}^{k} W_i^2 s_i^2 / w_i}{(d / z_{a/2})^2 + 1/N \sum_{i=1}^{k} W_i s_i^2}$$ όπου $w_i = \dfrac{n_i}{n}$ και $n_i = nw_i$

Για την **Αναλογική Κατανομή** ισχύει:

Εκτίμηση μέσου

$$d = z_{a/2} \cdot SE(\overline{x}) \qquad \text{ή} \qquad d = z_{a/2}\sqrt{\frac{1}{n} \cdot \frac{(N-n)}{N} \sum_{i=1}^{k} W_i \ s_i^2}$$

Και
$$n = \frac{\sum_{i=1}^{k} W_i \ s_i^2}{(d/z_{a/2})^2 + 1/N \sum_{i=1}^{k} W_i s_i^2}$$

Για την **την Άριστη Κατανομή με Σταθερό Κόστος κατά μονάδα (κατανομή Neyman)** ισχύει:

Εκτίμηση μέσου

$$d = z_{a/2} \cdot SE(\overline{x}) \qquad \text{ή} \qquad d = z_{a/2}\sqrt{\frac{\left(\sum_{i=1}^{k} W_i s_i^2\right)^2}{n} - \frac{\left(\sum_{i=1}^{k} W_i s_i^2\right)}{N}}$$

Και
$$n = \frac{\left(\sum_{i=1}^{k} W_i s_i\right)^2}{(d/z_{a/2})^2 + 1/N \sum_{i=1}^{k} W_i s_i^2}$$

Για την **Άριστη Κατανομή με Διαφορετικό Κόστος ανά Μονάδα** ισχύει:

Εκτίμηση μέσου

$$d = z_{a/2} \cdot SE(\overline{x}) \text{ ή}$$

$$d = z_{a/2}\sqrt{\frac{\left(\sum_{i=1}^{k} W_i s_i \sqrt{C_i}\right) \cdot \left(\sum_{i=1}^{k} W_i s_i \big/ \sqrt{C_i}\right)}{n} - \frac{\left(\sum_{i=1}^{k} W_i s_i^2\right)}{N}}$$

Και $$n = \frac{\left(\sum_{i=1}^{k} W_i s_i \sqrt{C_i}\right) \cdot \left(\sum_{i=1}^{k} W_i s_i \Big/ \sqrt{C_i}\right)}{(d / z_{a/2})^2 + 1/N \sum_{i=1}^{k} W_i s_i^2}$$ (βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Τέλος, στην περίπτωση της **Κατά Ομάδες Δειγματοληψίας** το απαιτούμενο μέγεθος δείγματος για την εκτίμηση του μέσου m προσδιορίζεται αντίστοιχα από την εξίσωση :

$$d = z_{a/2} \cdot SE(\bar{x}_c) \Leftrightarrow d = z_{a/2} \cdot \sqrt{\frac{N-n}{Nn\bar{M}^2} s^2}$$

Αν λύσουμε την εξίσωση αυτή ως προς το n θα προκύψει το μέγεθος δείγματος για την εκτίμηση του μέσου m .

$$n = \frac{Ns^2}{(d / z_{a/2})^2 N\bar{M}^2 + s^2}$$

(βλ. Δαμιανός Χ, 2006)

6.2. Προσδιορισμός Μεγέθους Δείγματος για Εκτίμηση Ποσοστών

Για να προσδιορίσουμε το απαιτούμενο μέγεθος του δείγματος για την εκτίμηση του ποσοστου P, καθορίζουμε τον επιθυμητό βαθμό ακρίβειας σε κάποιο διάστημα εμπιστοσύνης.

Έτσι αν δεχθούμε οτι το δειγματοληπτικό σφάλμα $|p-P|$ δεν πρέπει να υπερβεί ενα επίπεδο d με συντελεστή εμπιστοσύνης (1 – α) έχουμε:

$$P\ \{|p-P| \leq d\} = 1 - a$$

Το διάστημα εμπιστοσύνης για το ποσοστό P ενός πληθυσμού με συντελεστή εμπιστοσύνης (1 – α) είναι :

$$p \pm Z_{a/2} \cdot SE(p)$$

όπου $SE(p)$ είναι το τυπικό σφάλμα και για την ***<u>απλή τυχαία δειγματοληψία</u>*** δίνεται απο τη σχέση :

$$SE(p) = \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Γαι να ισχύει ο περιορισμός οτι το δειγματοληπτικό σφάλμα δεν πρέπει να υπερβεί το περιθώριο d θέτουμε :

$$d = Z_{a/2} \cdot SE(p) \quad \text{ή}$$

$$d = Z_{a/2} \cdot \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Αν ο πληθυσμός N είναι αρκετά μεγάλος τότε παραλείπουμε το συντελεστή διόρθωσης πεπεραμένου πληθυσμού και λύνοντας ως προς n έχουμε :

$$n = \frac{Z^2{}_{a/2} pq}{d^2} .$$

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Στην περίπτωση της ***Συστηματικής Δειγματοληψίας***, εάν ο πληθυσμός είναι τυχαίος, το μέγεθος του δείγματος για την εκτίμηση ποσοστού μπορεί να προσδιορισθεί από τη σχέση που χρησιμοποιείται και στην περίπτωση της απλής τυχαίας δειγματοληψίας.

Επομένως ομοίως ισχύει:

$$d = Z_{a/2} \cdot \sqrt{\frac{p \cdot q}{n} \cdot \sqrt{\frac{N-n}{N-1}}}$$

$$n = \frac{Z_{a/2} pq}{d^2}$$

Όσον αναφορά τη ***Στρωματοποιημένη Δειγματοληψία*** το τυπικό σφάλμα $SE(p)$ τίθεται κατά περίπτωση κατανομής του δείγματος (*αναλογική κατανομή, αυθαίρετη κατανομή, άριστη κατανομή*) και λύνουμε εξίσωση ως προς n.

Το μέγεθος του δείγματος στην περίπτωση της *αναλογικής κατανομής* προσδιορίζεται ομοίως από τη σχέση:

$$d = Z_{a/2} \cdot SE(p)$$

Όπου
$$SE(p) = \sqrt{\frac{1}{n}\left(\frac{N-n}{N}\right)\sum_{i=1}^{k} W_i p_i q_i}$$

Δηλαδή:
$$d = Z_{a/2}\sqrt{\frac{1}{n}\left(\frac{N-n}{N}\right)\sum_{i=1}^{k} W_i p_i q_i}$$

Λύνοντας ως προς n έχουμε:

$$n = \frac{\sum_{i=1}^{k} W_i p_i q_i}{(d/Z_{a/2})^2 + \frac{1}{N}\sum_{i=1}^{k} W_i p_i q_i}$$

(βλ. Χαρίσης Ι.Κώστας, Κίοχος Α.Πέτρος ,1997)

Τέλος στην περίπτωση της <u>***Κατά Ομάδες Δειγματοληψίας***</u> το απαιτούμενο μέγεθος δείγματος γαι την εκτίμηση ποσοστού προσδιορίζεται από την εξής σχέση:

$$d = z_{a/2} \cdot \sqrt{V(p)} \qquad \Leftrightarrow \qquad d = z_{a/2} \cdot \sqrt{\frac{N-n}{Nn\overline{M}^2} s^2}$$

Αν λύσουμε την εξίσωση ως προς n θα προκύψει το μέγεθος του δείγματος

$$n = \frac{Ns^2}{(d/z_{a/2})^2 N\overline{M}^2 + s^2}$$

Όπου

$$s^2 = \frac{\sum_{i=1}^{n}(a_i - pm_i)^2}{n-1}$$

ΚΕΦΑΛΑΙΟ ΈΒΔΟΜΟ

ΈΡΕΥΝΑ- ΜΕΛΕΤΗ ΠΛΗΘΥΣΜΟΥ

7.1. Εισαγωγή

Μία ασφαλιστική εταιρεία Χ ανέθεσε σε μία ερευνητική ομάδα τη διεξαγωγή έρευνας αγοράς με θέμα:
«Δειγματοληπτική έρευνα του βαθμού ικανοποίησης από τα υπάρχοντα ασφαλιστικά προγράμματα των ασφαλισμένων στην Ασφαλιστική Εταιρεία Χ και προτάσεις για τη βελτίωση των ασφαλιστικών προγραμμάτων.»

Η ερευνητική ομάδα, αφού έλαβε υπόψη της όλα τα στάδια που απαιτούνται για τη διεξαγωγή μιας δειγματοληπτικής έρευνας, προέβη στη σύνταξη της παρακάτω μελέτης και στην εξαγωγή χρήσιμων συμπερασμάτων.

7.2. Σκοπός και Έκταση της Έρευνας

Σκοπός της έρευνας ήταν η συγκέντρωση στατιστικών στοιχείων από τους πελάτες της εταιρείας Χ, η ανάλυση και ερμηνεία αυτών για την εξαγωγή συμπερασμάτων, τα οποία θα βοηθήσουν στη λήψη αποφάσεων από την ασφαλιστική εταιρεία Χ για τη βελτίωση των ασφαλιστικών της προγραμμάτων.

Τα στοιχεία που συγκεντρώθηκαν αναφέρονται στα χαρακτηριστικά των πελατών της, δηλαδή οικογενειακή κατάσταση, αριθμός παιδιών, αριθμός μελών της οικογένειας που είναι ασφαλισμένα στην εταιρεία Χ, επίπεδο εκπαίδευσης, μηνιαίο εισόδημα κ.λπ.

Ιδιαίτερη έμφαση δόθηκε στη συγκέντρωση στοιχείων για τη διαπίστωση του βαθμού ικανοποίησης των πελατών από τα υπάρχοντα ασφαλιστικά προγράμματα, ώστε να εντοπιστούν οι αδυναμίες και να δοθούν λύσεις για τη βελτίωση των προϊόντων και υπηρεσιών που παρέχονται από την ασφαλιστική εταιρεία.

7.3. Μεθοδολογία της Έρευνας

Η συγκέντρωση των στατιστικών στοιχείων έγινε με τη μέθοδο της ***στρωματοποιημένης τυχαίας δειγματοληψίας***. Η στρωματοποίηση έγινε κατά γεωγραφικές περιοχές στο Νομό Αττικής, όπου έγινε η έρευνα. Ο ερευνώμενος πληθυσμός αποτελείται από το σύνολο των πελατών της εταιρείας ηλικίας 15 ετών και άνω. Δεν συμπεριλήφθηκαν στην έρευνα πελάτες ηλικίας 15 ετών και κάτω, γιατί τα άτομα αυτής της ηλικίας δεν μπορούν να εκφέρουν γνώμη για τα προϊόντα και τις υπηρεσίες της εταιρείας.

Ο προσδιορισμός του μεγέθους του δείγματος έγινε από τη σχέση

$$n = \frac{\sum_{i=1}^{k} W_i p_i q_i}{(d / Z_{a/2})^2 + \frac{1}{N} \sum_{i=1}^{k} W_i p_i q_i}$$

(βλ. Κεφάλαιο Έβδομο**,** Προσδιορισμός Μεγέθους Δείγματος για Εκτίμηση Ποσοστών).

Η κατανομή του δείγματος $n = 660$, έγινε ***αναλογικά*** στα 10 στρώματα του ερευνώμενου πληθυσμού, δηλαδή στις 10 γεωγραφικές περιοχές του Νομού Αττικής.

Η επιλογή των δειγματοληπτικών μονάδων μέσα σε κάθε στρώμα έγινε από ονομαστικό κατάλογο των πελατών της εταιρείας, ο οποίος αποτέλεσαι και το ***δειγματοληπτικό πλαίσιο***, με τη μέθοδο της ***συστηματικής δειγματοληψίας***.

7.4. Ερωτηματολόγιο

Το ερωτηματολόγιο (βλ. Παράρτημα) συντάχθηκε από την ερευνητική ομάδα ύστερα από συνεργασία με τον αρμόδιο Διευθυντή του Μάρκετινγκ.

Λαμβάνοντας υπόψη το σκοπό της συγκεκριμένης έρευνας, η ερευνητική ομάδα φρόντισε ώστε τα ερωτήματα να έχουν μια λογική ακολουθία, να είναι δυνατή η επεξεργασία των στοιχείων, να καλύπτουν όλα τα βασικά θέματα που απασχολούν την εταιρεία. Επίσης, έλαβε υπόψη της την ικανότητα των ατόμων που θα ερευνηθούν, τη θέλησή τους να συνεργαστούν, την αποφυγή ερωτήσεων που οδηγούν σε αποκρύψεις, την αποφυγή κατευθυνόμενων ερωτήσεων, τη σαφήνεια και την απλότητα της γλώσσας. Με αυτόν τον τρόπο, αποφεύγονται κατά το δυνατό τα μη δειγματοληπτικά σφάλματα και κατά συνέπεια περιορίζεται η μεροληψία.

7.5. ΑΠΟΤΕΛΕΣΜΑΤΑ

7.5.1 Δημογραφικά και Κοινωνικά Χαρακτηριστικά των Ασφαλισμένων της Εταιρείας

Από τα 660 άτομα που συμπεριλήφθηκαν στο δείγμα, 377 ήταν άνδρες, δηλαδή ποσοστό 57%, και 283 γυναίκες, δηλαδή ποσοστό 43% (Πίνακας 1).

Πίνακας 1

Κατανομή των ερωτηθέντων κατά φύλο

Φύλο	Αριθμός ατόμων	***Ποσοστό***
Άνδρες	377	57
Γυναίκες	283	43
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Στον Πίνακα 2 απεικονίζεται η σύνθεση των ερωτηθέντων ατόμων κατά οικογενειακή κατάσταση. Από τον πίνακα αυτό παρατηρείται ότι ο αριθμός των έγγαμων ατόμων με δύο παιδιά ανέρχεται σε 227 άτομα, δηλαδή ποσοστό 34% και ακολουθούν οι άγαμοι (ποσοστό 25%) και οι έγγαμοι με ένα παιδί (ποσοστό 18%).

Αξιοπρόσεκτο είναι το γεγονός ότι οι ασφαλισμένοι έγγαμοι με τρία παιδιά και πάνω είναι ελάχιστοι. Αυτό προφανώς οφείλεται σε οικονομικούς λόγους, γιατί αυτές οι οικογένειες αντιμετωπίζουν δυσκολίες στην πληρωμή των ασφαλίστρων.

Πίνακας 2

Αριθμός ερωτηθέντων κατά οικογενειακή κατάσταση

Οικογενειακή κατάσταση	Άτομα	***Ποσοστό%***
Άγαμος	166	25
Έγγαμος ή διαζευγμένος χωρίς παιδιά	89	14
Έγγαμος ή διαζευγμένος με 1 παιδί	120	18
Έγγαμος ή διαζευγμένος με 2 παιδιά	227	34
Έγγαμος ή διαζευγμένος με 3 παιδιά	50	8
Έγγαμος ή διαζευγμένος με 4 παιδιά	6	1
Έγγαμος ή διαζευγμένος με 5 παιδιά	1	0
Έγγαμος ή διαζευγμένος με 6 παιδιά	1	0
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Από τον Πίνακα 3 διαπιστώνεται ότι το μεγαλύτερο ποσοστό που ερωτήθηκαν ήταν ασφαλισμένοι μόνο οι ίδιοι (δηλαδή 51%). Αυτό σημαίνει ότι οι πελάτες της εταιρείας έκαναν ως επί το πλείστον ατομικές ασφάλειες που δεν καλύπτουν άλλα μέλη της οικογένειας. Λαμβάνοντας υπόψη ότι 405 άτομα από αυτά που ερωτήθηκαν είναι έγγαμα ή διαζευγμένα με ένα παιδί και πάνω (βλ. Πίνακα 2), μπορούμε να προτείνουμε στην εταιρεία να προβεί σε ενημέρωση των ασφαλισμένων της κατηγορίας αυτής, ώστε να εξασφαλίσει την επέκταση της ασφάλισης και στα λοιπά μέλη των οικογενειών τους.

Πίνακας 3

Κατανομή ασφαλισμένων μελων της οικογένειας των ερωτηθέντων

Ασφαλισμένος	Αριθμός ατόμων	***Ποσοστό***
Μόνο ο ίδιος	336	51
Μόνο ο/η σύζυγος	15	2
Μόνο τα τέκνα	24	4
Ο ίδος και η σύζυγος	87	13
Ο ίδιος και τα τέκνα	61	9
Μόνο ο/η σύζυγος και τα τέκνα	18	3
Ο ίδιος, η σύζυγος και τα τέκνα	44	7
Περισσότερα από 3 άτομα	75	11
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Από τα στοιχεία του παρακάτω πίνακα (Πίνακας 4) φαίνεται ότι το μεγαλύτερο πόσοστό των ερωτηθέντων (37%) ήταν απόφοιτοι λυκείου και ακολουθούν οι πτυχιούχοι Ανωτέρων και Ανωτάτων σχολών (30%). Από τον ίδιο πίνακα προκύπτει ότι το ποσοστό των αποφοίτων Δημοτικού που είναι ασφαλισμένοι στην εταιρεία είναι σχετικά μικρό.

Πίνακας 4

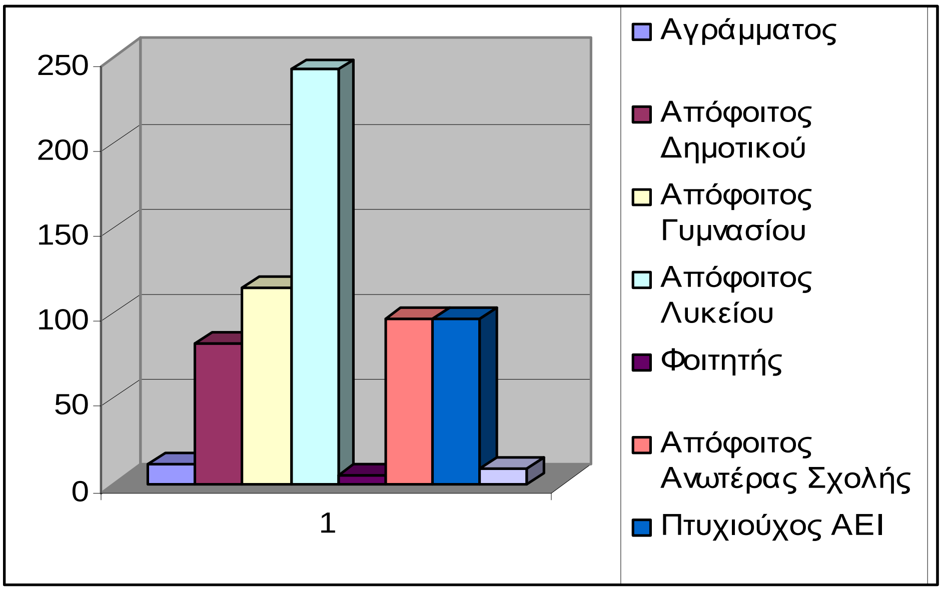
Αριθμός ερωτηθέντων ατόμων κατά επίπεδο εκπαίδευσης

Επίπεδο εκπαίδευσης	Άτομα	***Ποσοστό %***
Αγράμματος	12	2
Απόφοιτος Δημοτικού	82	12
Απόφοιτος Γυμνασίου	115	17
Απόφοιτος Λυκείου	243	37
Φοιτητής	5	1
Απόφοιτος Ανωτέρας Σχολής	97	15
Πτυχιούχος ΑΕΙ	9	15
Κάτοχος Μεταπτυχιακού	9	1
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 1

Επίπεδο εκπαίδευσης



7.5.2. Οικονομικά Χαρακτηριστικά των Ασφαλισμένων

Από την κατανομή των ερωτηθέντων ατόμων κατά μέσο μηνιαίο εισόδημα (Πίνακας 5), προκύπτει ότι το μεγαλύτερο ποσοστό (51%) των ερωτηθέντων ατόμων έχουν μηνιαίο εισόδημα μεταξύ 600-900 ευρώ και 28% αυτών έχουν εισόδημα κάτω των 600 ευρώ. Ενώ, αντίθετα, το ποσοστό των ατόμων με μηνιαίο εισόδημα πάνω από 900 ευρώ ανέρχεται σε μόλις 21%. Η ανάλυση αυτή οδηγεί στο συμπέρασμα ότι οι πελάτες της εταιρίας είναι κατά κανόνα χαμηλού εισοδήματος και δεν μπορούν να πληρώσουν με ευκολία υψηλά ασφάλιστρα. Εάν η εταιρεία ακολουθήσει πολιτική επέκτασης ασφάλισης και στα λοιπά μέλη της οικογένειας, θα πρέπει να διατηρήσει σε χαμηλό επίπεδο τα ασφάλιστρα και να γίνει μια προσπάθεια να ασφαλιστούν άτομα με μεγαλύτερο εισόδημα.

Πίνακας 5
Κατανομή ερωτηθέντων κατά μέσο μηνιαίο καθαρό εισόδημα της οικογένειας τους

Εισόδημα σε ευρώ	Άτομα	***Ποσοστό***
μέχρι 600	186	28
600-900	334	51
900-1500	109	16
1500 και άνω	31	5
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Από την ανάλυση των στοιχείων του Πίνακα 6 φαίνεται ότι οι πελάτες της εταιρείας ασφαλίζονται κατά κανόνα με μικρά χρηματικά ποσά. Συγκεκριμένα, από τους 660 ερωτηθέντες, 144 άτομα (ποσοστό 22%) ασφαλίζονται για χρηματικά ποσά μικρότερα των 1.500 ευρώ, 160 άτομα (ποσοστό 24%) ασφαλίζονται για χρηματικό ποσό μεταξύ 1.500-3.000 ευρώ, ενώ αντίθετα τα άτομα που ασφαλίζονται για χρηματικά ποσά άνω των 7.500 ευρώ αντιπροσωπεύουν μόνο το 14% των ερωτηθέντων. Αυτό εξηγείται από το γεγονός ότι οι ασφαλισμένοι στην εταιρεία είναι κατά κανόνα χαμηλού εισοδήματος.

Πίνακας 6

Κατανομή ερωτηθέντων κατά μέγεθος

Χρηματικού ποσού που ασφαλίζονται

Χρηματικό ποσό σε ευρώ	Άτομα	***Ποσοστό%***
0-1.500	144	22
1.500-3.000	160	24
3.000-4.500	135	20
4.500-6.000	80	13
6.000-7.500	53	8
7.500-10.500	29	4
10.500-12.000	18	3
12.000-13.500	17	3
13.500-15.000	16	2
13.500-15.000	3	1
15.000-18.000	5	1
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

7.5.3. Λοιπά Χαρακτηριστικά των Ασφαλισμένων

Από τα στοιχεία του Πίνακα 7 φαίνεται ότι 336 άτομα από τα 660 που ερωτήθηκαν, δηλαδη ποσοστό 51%, έχουν ως κύρια ασφάλιση την «ασφάλεια ζωής» και «νοσοκομειακή περίθαλψη» και 212 άτομα από τους ερωτηθέντες, δηλαδή ποσοστό 32%, την «συνταξιοδότηση». Τα υπόλοιπα είδη ασφαλίσεων παρατηρείται ότι κυμαίνονται σε πολύ χαμηλά επίπεδα.

Πίνακας 7

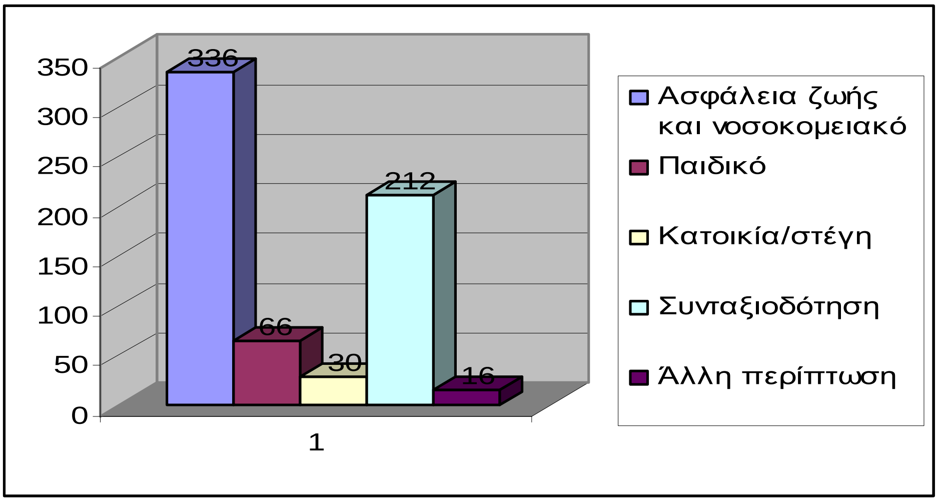
Αριθμό ερωτηθέντων κατά κύρια ασφάλιση

Ασφάλιση	Άτομα	***Ποσοστό%***
Ασφάλεια ζωής και νοσοκομειακό	336	51
Παιδικό	66	10
Κατοικία/στέγη	30	5
Συνταξιοδότηση	212	32
Άλλη περίπτωση	16	2
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 2

Είδος κύριας ασφάλισης



7.5.4. Τρόπος Γνωριμίας Ασφαλισμένων με την Εταιρεία

Από τα 660 που ερωτήθηκαν, τα 480, δηλαδή ποσοστό 73%, δήλωσαν ότι γνώρισαν την εταιρεία και τα προϊόντα της από ασφαλιστή, ενώ 143 άτομα, δηλαδή ποσοστό 22%, από φιλική σύσταση. Αντίθετα, από διαφήμιση γνώρισαν την εταιρεία και τα προϊόντα της μόνο 3 άτομα (βλ. Πίνακα 8). Αυτό σημαίνει ότι η διαφήμιση δεν έχει θετικά αποτελέσματα για απόκτηση νέων πελατών.

Πίνακας 8

Κατανομή ερωτηθέντων σύμφωνα με τον

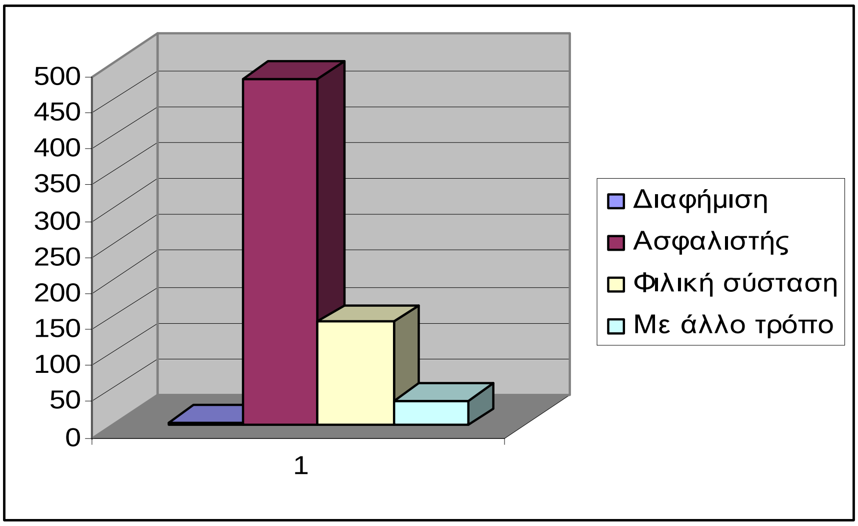
τρόπο γνωριμίας με την εταιρεία

Τρόπος γνωριμίας	Άτομα	***Ποσοστό%***
Από διαφήμιση	3	0
Από ασφαλιστή	480	73
Από φιλική σύσταση	143	22
Με άλλο τρόπο	34	5
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 3

Τρόπος γνωριμίας με την εταιρεία



Από την ανάλυση των στοιχείων αυτών συμπεραίνουμε ότι, σε γενικές γραμμές, οι πελάτες της εταιρείας γνωρίζονται με την εταιρεία μέσω των ασφαλιστών της, οι οποίοι φαίνεται να είναι αρκετά ικανοί στο να φέρνουν νέους πελάτες στην εταιρεία. Αξιοπρόσεκτο όμως είναι το γεγονός ότι το 22% των ερωτηθέντων έγιναν πελάτες της εταιρείας από σύσταση φιλικού προσώπου. Αυτό σημαίνει ότι άτομα τα οποία είναι πάρα πολύ ικανοποιημένα από τα προϊόντα και τις υπηρεσίες της εταιρείας, την έχουν συστήσει σε φίλους τους. Πρέπει λοιπόν να καταβληθεί μεγαλύτερη προσπάθεια όσον αφορά τη δημιουργία φιλικού κλίματος μεταξύ της εταιρείας και των πελατών. Αυτό μπορεί να επιτευχθεί μέσω της καλής συνεργασίας με το τμήμα εξυπηρέτησης της εταιρείας, έτσι ώστε να μπορεί ο πελάτης να συστήσει την εταιρεία σε φίλους του. Θα πρέπει δηλαδή, ο ίδιος ο πελάτης να γίνει διαφημιστής της εταιρείας.

Φυσικά και μια καλή διαφήμιση από τα μέσα ενημέρωσης θα έχει θετικά αποτελέσματα εφόσον τα προϊόντα της εταιρείας θα είναι ανώτερα ή ανταγωνίσιμα σε σχέση με αυτά των άλλων μεγάλων ασφαλιστικών της χώρας.

7.5.5. Ικανοποίηση των Ερωτηθέντων από τα Προϊόντα και τις Υπηρεσίες της Εταιρείας

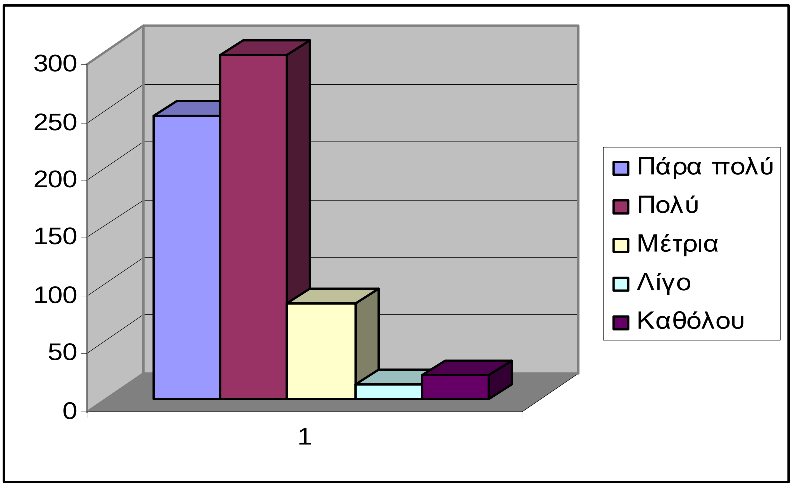
Από τον Πίνακα 9 παρατηρούμε ότι 245 άτομα, δηλαδή ποσοστό 37%, είναι πάρα πολύ ικανοποιημένα απο τα προϊόντα της εταιρείας και 298, ποσοστό 45%, είναι πολύ ικανοποιημένα. Ενώ, αντίθετα, τα άτομα που είναι λίγο ή καθόλου ικανοποιημένα αποτελούν μόλις το 5% του συνόλου των ερωτηθέντων. Από την ανάλυση, λοιπόν, των στοιχείων αυτών μπορούμε να συμπεράνουμε ότι οι πελάτες της εταιρείας είναι ικνοποιημένοι από τα προϊόντα της.

Πίνακας 9
Αριθμός ερωτηθέντων κατά βαθμό ικανοποίησης
από τα προϊόντα της εταιρείας

Βαθμός ικανοποίησης	Άτομα	***Ποσοστό%***
Πάρα πολύ	245	37
Πολύ	298	45
Μέτρια	83	13
Λίγο	13	2
Καθόλου	21	3
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 4
Ποσοστό ικανοποίησης από το προϊόν



Από τα στοιχεία του Πίνακα 10 φαίνεται ότι 193 άτομα από τα 660 τα οποία ερωτήθηκαν, δηλαδή ποσοστό 29%, δήλωσαν ότι είναι πάρα πολύ ικανοποιημένα από τις υπηρεσίες της εταιρείας και 177 δήλωσαν ότι ειναι πολύ ικανοποιημένα. Συγκρίνοντας το αποτέλεσμα ικανοποίησης των πελατών από τα προϊόντα της εταιρείας και τις υπηρεσίες της, παρατηρούμε ότι το επίπεδο ικανοποίησης των πελατών από τις υπηρεσίες είναι χαμηλότερο σε σχέση με αυτό των προϊόντων της. Ειδικότερα, το ποσοστό των ατόμων που δήλωσαν μέτρια ικανοποίηση από τις υπηρεσίες της εταιρείας είναι αρκετά υψηλό, 41%, έναντι 13% που δήλωσαν μέτρια ικανοποίηση από τα προϊόντα. Με την ανάλυση των στοιχείων αυτών επισημαίνεται ότι η εταιρεία πρέπει να καταβάλει προσπάθεια γαι την βελτίωση της εξυπηρέτησης των πελατών της, αφού η πλειονότητα αυτών φαίνεται να είναι μέτρια ικανοποιημένοι.

Πίνακας 10

Αριθμός ερωτηθέντων κατά βαθμό ικανοποίησης

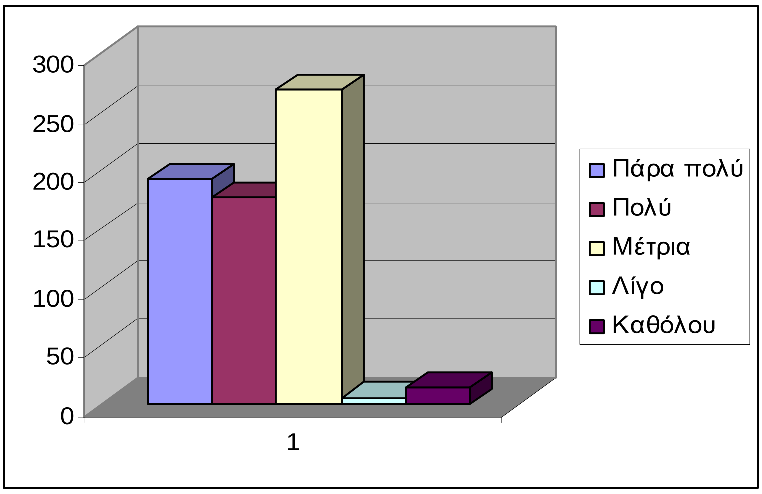
από τις υπηρεσίες της εταιρείας

Βαθμός ικανοποίησης	Άτομα	***Ποσοστό%***
Πάρα πολύ	193	29
Πολύ	177	27
Μέτρια	269	41
Λίγο	6	1
Καθόλου	15	2
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 5

Ικανοποίηση από την εξυπηρέτηση



7.5.6. Ικανοποίηση των Ερωτηθέντων Ατόμων από τη Συνεργασία με τους Ασφαλιστές.

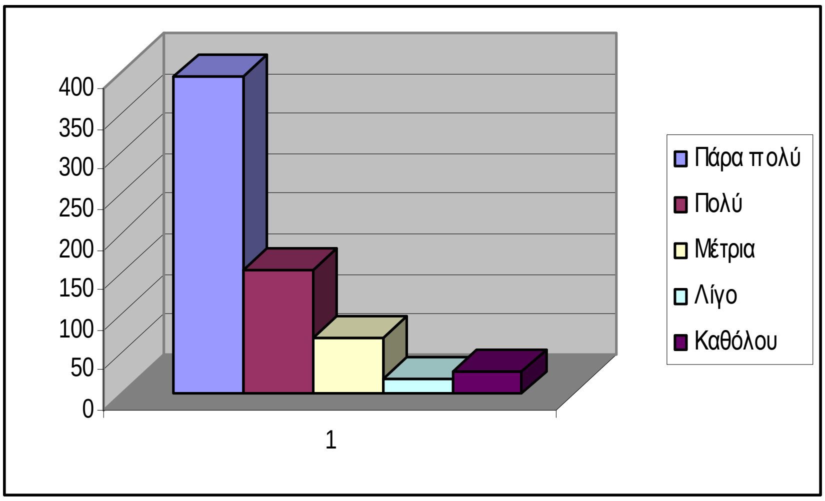
Από την ανάλυση των στοιχείων του Πίνακα 11 προκύπτει ότι 394 άτομα από τα 660 που ερωτήθηκαν (ποσοστό 60%), δήλωσαν ότι είναι πάρα πολύ ικανοποιημένοι από τη συνεργασία με τους ασφαλιστές και 153 άτομα (ποσοστό 23%) δήλωσαν ότι είναι πολύ ικανοποιημένοι. Ενώ, αντίθετα αυτοί που δήλωσαν μέτρια ικανοποίηση ανέρχονται μόνο σε 69 άτομα (ποσοστό 10%) και αυτοί που δήλωσαν ότι είναι λίγο ικανοποιημένοι ή καθόλου αποτελούν το 3% και 4% αντίστοιχα, του συνόλου των ερωτηθέντων.

Πίνακας 11
Αριθμός ερωτηθέντων ατόμων κατά βαθμό ικανοποίησης από τη συνεργασία με τους ασφαλιστές

Βαθμός ικανοποίησης	Άτομα	Ποσοστό %
Πάρα πολύ	394	60
Πολύ	153	23
Μέτρια	69	10
Λίγο	17	3
Καθόλου	27	4
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 6
Ικανοποίηση από τους ασφαλιστές



Για μια ακόμη φορά τα στοιχεία δείχνουν την ικανότητα των ασφαλιστών για την επίτευξη του επιδιωκόμενου στόχου της εταιρείας.

Επίσης από ερώτημα που διατυπώθηκε κατά τη διάρκεια της έρευνας προς τους πελάτες, για το πόσο συχνά επικοινωνεί ο ασφαλιστής μαζί τους, προκύπτει ότι το 21% των πελατών δεν επικοινωνεί καθόλου, το 9% επικοινωνεί λιγότερο απο 3 φορές το χρόνο, το 28% επικοινωνεί 4-9 φορές το χρόνο, το 23% επικοινωνεί 10-19 φορές το χρόνο και το 19% επικοινωνεί πάνω απο 15 φορές το χρόνο.

Σε γενικές γραμμές, μπορούμε να πούμε ότι η επικοινωνία του ασφαλιστή με τους πελάτες δεν ήταν ικανοποιητική και πρέπει να γίνεται πιο συχνά ώστε να επιτυγχάνεται σωστή ενημέρωση των πελατών.

7.5.7. Απόψεις Ερωτηθέντων για την Αναγκαιότητα Απόκτησης Ασφάλειας.

Από τον Πίνακα 12 διαπιστώνεται ότι 246 άτομα απο τα 660 που ερωτήθηκαν (ποσοστό 37%) δήλωσαν ότι θεωρούν πάρα πολύ αναγκαία την απόκτηση μιας ασφάλειας (ατομική, ασφάλεια ζωής, συνταξιοδοτικό κ.λπ.). Επίσης, 135 άτομα (ποσοστό 20%) δήλωσαν ότι θεωρούν πολύ αναγκαία την απόκτηση μιας ασφάλειας. Ενώ, αντίθετα, 215 άτομα (ποσοστό 33%) δήλωσαν ότι δεν θεωρούν καθόλου αναγκαίο. Για μια ακόμα φορά επισημαίνεται η ανάγκη σωστής ενημέρωσης των πελατών για τα πλεονεκτήματα των ασφαλειών. Πρέπει να αντιληφθούν ότι η αναγκαιότητα απόκτησης μιας ασφάλειας με τις σημερινές συνθήκες ζωής είναι επιβεβλημένη και αναγκαία. Πράγματι, ο κάθε άνθρωπος σε όλη τη ζωή του είναι εκτεθειμένος σ‘έναν κόσμο από κινδύνους, οι οποίοι είναι θεμελιώδους σημασίας και καθορίζουν την προσωπική, οικογενειακή, οικονομική και συναισθηματική του ύπαρξη και τη ζωή του καθώς και της οικογένειάς του γενικότερα.

Είναι δύσκολο κανείς να προσδιορίσει τους περισσότερους κινδύνους, *αν ,πότε* και *ποιοί* θα πραγματοποιηθούν και *τι συνέπειες* θα έχουν, όπως θάνατος, ανικανότητα, οικονομικές ζημίες, ασθένειες στο ίδιο το άτομο ή σε τρίτους με τους οποίους έχει άμεση σχέση ή είναι υπεύθυνος για αυτούς. Μέσα στα πλαίσια αυτά ο άνθρωπος, για να εξασφαλίσει τα αγαθά του και να αποκαταστήσει τις βλάβες ή τις ζημιές ή τις ασθένειες που ενδέχεται να υποστεί καθώς και να μειώσει τις συνέπειες, ασφαλίζεται δια μέσου της ιδιωτικής ή της δημόσιας ασφαλιστικής πρωτοβουλίας ή και των δύο μορφών μαζι.

Πίνακας 12

Απόψεις ερωτηθέντων για το βαθμό

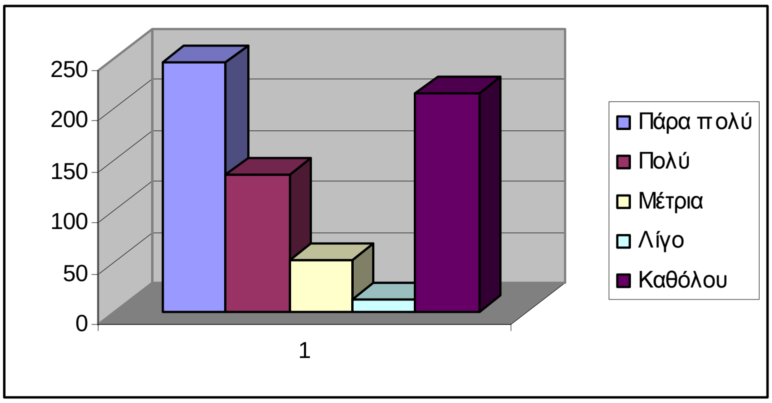
ανάγκης απόκτησης μιας ασφάλειας

Ανάγκη απόκτησης ασφάλειας	Άτομα	Ποσοστό %
Πάρα πολύ	246	37
Πολύ	135	20
Μέτρια	52	8
Λίγο	12	2
Καθόλου	215	33
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Δίαγραμμα 7

Βαθμός ανάγκης απόκτησης ασφάλειας



7.5.8. Απόψεις Ερωτηθέντων για τις Τιμές και τα Πλεονεκτήματα Προϊόντων ή Υπηρεσιών άλλων Εταιρειών.

Από τον Πίνακα 13 φαίνεται ότι η πλειονότητα των ερωτηθέντων ατόμων, 437 άτομα (ποσοστό 66%), δήλωσαν οτι δεν γνωρίζουν αν οι τιμές και τα προϊόντα ή οι υπηρεσίες άλλων εταιρειών είναι καλύτερες από αυτές της εταιρείας. Μόνο το 7% των ερωτηθέντων δήλωσαν ότι οι τιμές και τα προϊόντα ή οι υπηρεσίες άλλων εταιρειών είναι καλύτερες από αυτές της εταιρείας, ενώ, αντίθετα, το 7% δήλωσαν ότι είναι χειρότερες.

Γενικά μπορούμε να πούμε ότι οι δηλώσεις των ερωτηθέντων ατόμων ότι «δεν γνωρίζουν αν τα προϊόντα και οι υπηρεσίες των ανταγωνιστικών εταιρειών είναι καλύτερες ή όχι» είναι ένα γεγονός που υποδηλώνει την άγνοια του Έλληνα ασφαλιζομένου και την κακή ενημέρωσή του.

Πίνακας 13

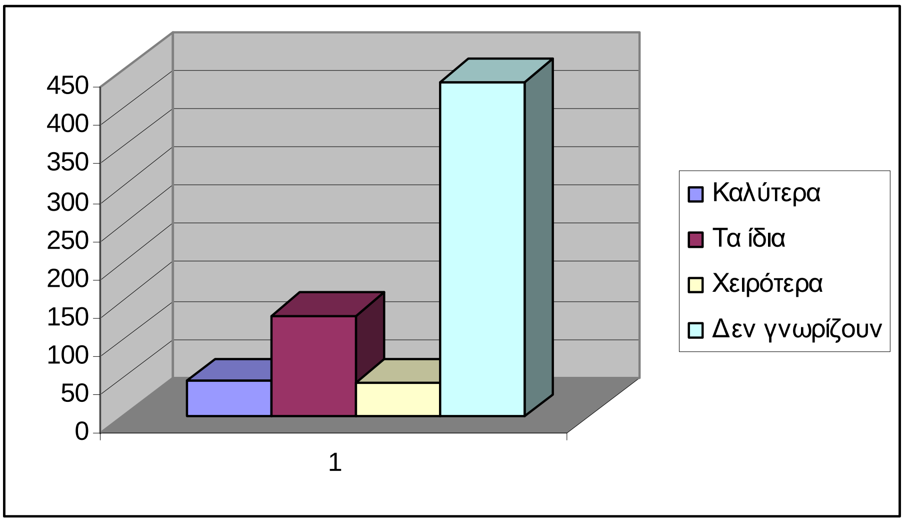
Απόψεις ερωτηθέντων ατόμων για τις τιμές και τα πλεονεκτήματα προϊόντων ή υπηρεσιών άλλων εταιρειών

Προσφορά προϊόντων ή υπηρεσιών άλλων εταιρειών	Άτομα	Ποσοστό %
Καλύτερα	46	7
Τα ίδια	132	20
Χειρότερα	45	7
Δεν γνωρίζουν	437	66
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 8

Γενική σύγκριση ανταγωνιστικών προϊόντων



7.5.9 Συχνότητα Αποζημίωσης των Ασφαλισμένων

Από την ανάλυση του Πίνακα 13 φαίνεται ότι 439 άτομα (ποσοστό 67%), δήλωσαν ότι δεν έχουν αποζημιωθεί από την εταιρεία καμία φορά.

Αυτό εξηγείται από το γεγονός ότι η πλειονότητα των πελατών είναι ασφαλισμένοι για λίγο χρονικό διάστημα και είναι επόμενο να μην έχουν κάνει χρήση του συμβολαίου τους ακόμη. Επίσης, ένας σημαντικός αριθμός ασφαλισμένων, 143 άτομα (ποσοστό 22%), έχουν αποζημιωθεί μόνο μία φορα. Ενώ, αντίθετα, 49 πελάτες (ποσοστό 7%) έχουν αποζημιωθεί δύο φορές και 29 πελάτες (ποσοστό 4%) έχουν αποζημιωθεί πάνω από τρεις φορές.

Πίνακας 13

Κατανομή ερωτηθέντων κατά αριθμό φορών

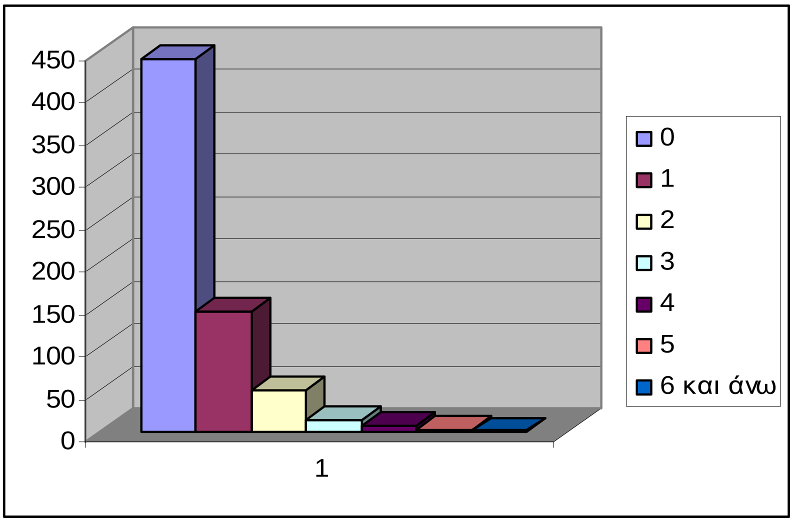
που έχουν αποζημιωθεί από την εταιρεία

Φορές αποζημίωσης	Άτομα	***Ποσοστό%***
0	439	67
1	143	22
2	49	7
3	15	2
4	8	1
5	4	1
6 και άνω	2	0
Σύνολο	660	100

Πηγή: Κιόχος Α. Πέτρος, Αθήνα 1997

Διάγραμμα 9

Αριθμός φορών αποζημίωσης



Γενικά μπορούμε να πούμε ότι οι αποζημιώσεις που πρόκειται να επιβαρύνουν την εταιρεία αναμένεται να μην είναι σημαντικές τουλάχιστο για την επόμενη δεκαετία, γιατί οι αφσαλισμένοι της εταιρείας αυτής, όπως αναφέραμε και παραπάνω, έχουν λίγα έτη ασφάλισης.

7.6. Συμπεράσματα – Προτάσεις

Από τη στατιστική ανάλυση των στοιχείων που συγκεντρώθηκαν με τη δειγματοληπτική έρευνα των ασφαλισμένων σην εταιρεία Χ προκύπτουν τα ακόλουθα:

1. Επισημαίνεται ότι το 51% των πελατών είναι αφαλισμένοι μόνο οι ίδιοι, δηλαδή δεν έχουν ασφαλίσει τα υπόλοιπα μέλη της οικογένειάς τους.

Θα πρέπει να καταβληθεί προσπάθεια από τους ασφαλιστές για επέκταση της ασφάλισης και στα υπόλοιπα μέλη της οικογένειας.

2. Το πελατολόγιο της εταιρείας απαρτίζεται κυρίως από άτομα χαμηλής μόρφωσης και χαμηλού εισοδήματος.

Με τη μεγάλη επέκταση και το σημερινό κύρος της εταιρείας, αυτή η αγορά-στόχος πρέπει να αλλάξει, στρεφόμενη προς άτομα υψηλότερου εισοδήματος (π.χ. μεγαλοεπιχειρηματίες, στελέχη επιχειρήσεων, μεγάλους οργανισμούς κ.λπ.).

3. Οι πελάτες της εταιρείας κατά κανόνα ασφαλίζονται με μικρά χρηματικά ποσά. Το 66% ασφαλίζεται για ποσά μικρότερα των 4.500 ευρώ.

Αυτό οφείλεται στο γεγονός ότι η πλειονότητα των σφαλισμένων είναι χαμηλού εισοδήματος. Η κατάσταση αυτή μπορεί να αλλάξει μόνο αν η εταιρεία στραφεί σε άτομα υψηλότερου εισοδήματος.

4. Το 51% των ασφαλισμένων της εταιρείας έχουν ως κύρια ασφάλιση την «ασφάλεια ζωής και νοσοκομειακή κάλυψη». Αν εξαιρεθεί η συνταξιοδότηση που ανέρχεται σε 32%, τα λοιπά είδη ασφαλειών βρίσκονται σε πολύ χαμηλά επίπεδα. Το γεγονός αυτό πρέπει να ληφθεί υπόψη από τους υπεύθυνους της εταιρείας και να εξεταστεί η περίπτωση βελτίωσης των προγραμμάτων αυτών.
5. Ο τρόπος γνωριμίας των πελατών με την εταιρεία γίνεται σε ποσοστό 73% από τους ασφαλιστές, 22% από φιλική σύσταση και 5% με άλλο τρόπο. Τα ποσοστά γνωριμίας δείχνουν ότι η εταιρεία έχει ικανούς ασφαλιστές , αλλά θα έπρεπε να είναι αντίστροφα. Δηλαδή, το όνομα της εταιρείας να συστήνεται από ευχαριστημένους πελάτες σε φιλικά πρόσωπα, τα οποία στη συνέχεια θα

προστρέχουν σε ασφαλιστές. Η διαφήμιση δεν έχει θετικά αποτελέσματα. Το ποσοστό γνωριμίας με τον τρόπο αυτό ήταν μόνο 0,4%.

6. Η πλειονότητα των πελατών της εταιρείας φαίνεται να είναι ικανοποιημένοι από τα προϊόντα της (ποσοστό 82%). Αντίθετα, οι πελάτες που είναι ικανοποιημένοι από την παροχή υπηρεσιών (γρήγορη-πρόθυμη-σωστή εξυπηρέτηση κ.λπ.) αποτελούν το 56% τους συνόλου.

Είναι απαραίτητο λοιπόν να καταβληθεί προσπάθεια από την εταιρεία για την αναβάθμιση της παροχής υπηρεσιών προς τους πελάτες.

7. Η επικοινωνία των ασφαλιστών με τους πελάτες της εταιρείας δεν είναι αυτή που απαιτείται. Ενώ η εταιρεία διαθέτει ικανούς ασφαλιστές, δεν ενδιαφέρονται για την επαφή με τον πελάτη μετά τη σύναψη του συμβολαίου.
8. Επισημαίνεται ότι το 33% των πελατών δήλωσαν ότι δεν θεωρούν καθόλου αναγκαία την απόκτηση μιας ασφάλειας. Αυτό δείχνει την έλλειψη ενημέρωσης των πελατών από τους ασφαλιστές. Κάθε άτομο πρέπει να γνωρίζει ότι σε όλη τη διάρκεια της ζωής του είναι εκτεθειμένο σε ένα μεγάλο αριθμό κινδύνων.

Είναι λοιπόν επιβεβλημένη ανάγκη για τον άνθρωπο να έχει έναν συμπαραστάτη στη ζωή του ώστε να μπορεί να αντιμετωπίσει έκτακτες ανάγκες και απρόβλεπτα γεγονότα. Ο συμπαραστάτης αυτός μπορεί να είναι η εταιρεία.

9. Η πλειονότητα των πελατών της εταιρείας δήλωσε ότι δε γνωρίζει αν τα προϊόντα και οι υπηρεσίες που προσφέρουν ανταγωνιστικές εταιρείες είναι καλύτερα. Το αποτέλεσμα αυτό υποδηλώνει άγνοια των ασφαλισμένων. Υπήρξε ένα μικρό ποσοστό που θεωρεί ορισμένες υπηρεσίες άλλων εταιρειών καλύτερες.

Μπορεί το ποσοστό αυτό να είναι μικρό, αλλά θα πρέπει να προβληματίσει τους υπεύθυνους και να τους προσανατολίσει στη βελτίωση των παρεχόμενων υπηρεσιών.

10. Το 67% των ασφαλισμένων της εταιρείας δεν έχει αποζημιωθεί από την εταιρεία καμία φορά μέχρι σήμερα. Φυσικά αυτό οφείλεται στο γεγονός ότι η πλειονότητα των πελατών είναι νεοεισερχόμενοι και είναι επόμενο να μην έχουν κάνει ακόμη χρήση του συμβολαίου τους. Αυτό σημαίνει ότι οι αποζημιώσεις που θα επιβαρύνουν την εταιρεία αναμένεται να μην είναι σημαντικές, τουλάχιστον μακροπρόθεσμα.

Μπορεί, λοιπόν, η εταιρεία να εκμεταλλευτεί το γεγονός αυτό και να χρησιμοποιήσει το πλεόνασμα εσόδων για επενδύσεις με απώτερο στόχο την αύξηση των κερδών (βλ. Κιόχος Α. Πέτρος, Αθήνα 1997).

ΠΑΡΑΡΤΗΜΑ

ΠΙΝΑΚΑΣ ΤΥΧΑΙΩΝ ΑΡΙΘΜΩΝ

667	529	156	947	477	875	810	894	469	345	89	716	751
369	249	404	206	719	327	525	715	460	726	285	110	865
455	128	770	444	795	450	667	208	815	50	293	913	128
43	324	603	117	750	226	30	10	822	114	482	16	111
919	166	73	810	896	297	669	604	914	841	372	980	939
6	831	261	557	722	89	92	727	294	210	782	890	385
77	251	504	506	324	997	679	682	969	707	320	340	505
414	351	854	256	738	957	785	618	534	84	149	538	623
49	763	645	94	979	503	472	969	719	364	458	501	95
276	964	68	934	339	550	421	101	628	5	91	912	307
26	61	907	707	532	793	316	735	509	491	471	750	140
64	858	156	341	817	197	91	939	31	131	851	270	910
306	530	135	128	803	284	87	24	699	1	189	21	719
137	651	15	457	852	321	123	200	55	424	125	891	47
99	506	357	19	500	836	752	509	777	530	129	438	489
451	561	879	386	422	15	20	458	237	756	598	258	166
66	779	68	452	149	846	445	555	5	41	745	14	77
364	520	583	442	137	862	874	4	79	54	148	832	896
566	168	957	835	528	974	536	445	69	765	591	500	364
57	359	151	524	567	45	163	96	852	832	198	128	651
63	867	445	472	751	525	741	451	235	15	653	30	93
461	28	546	125	476	542	500	478	359	48	66	199	64
368	761	25	154	553	781	452	553	652	159	37	169	621
579	358	56	63	582	751	254	809	148	478	49	269	783
256	26	752	147	43	445	698	700	235	126	145	687	82
278	898	526	963	463	72	99	530	987	397	9	146	39
554	21	853	496	753	123	674	88	56	452	420	964	231
20	889	539	631	763	556	44	854	530	420	128	201	246
999	523	111	43	800	753	951	852	761	482	935	246	795
606	97	38	117	425	208	381	759	438	128	236	38	98
156	538	293	106	602	228	686	374	538	807	245	525	36
127	858	37	614	145	695	142	462	303	8	473	208	414
85	192	869	285	355	765	352	491	307	325	191	416	852
2	65	994	456	789	123	147	258	369	964	852	485	200

<u>Ερωτηματολόγιο</u>

Δελτίο Έρευνας των ασφαλισμένων στην εταιρεία Χ

Δήμος ή κοινότητα:..

Δ/νση κατοικίας: Οδός..Αριθμός..................

Τηλέφωνο ερωτώμενου...

Α. Γενικά Χαρακτηριστικά

1. Επώνυμο....................................... Όνομα.......................... Πατρώμυνο...................
2. Φύλο: Άρρεν ☐ Θήλυ ☐
3. Έτος γεννήσεως:.................
4. α) οικογενειακή κατάσταση β) αριθμός τέκνων & ηλικία

α) οικογενειακή κατάσταση		β) αριθμός τέκνων & ηλικία
Άγαμος	☐	1. ....
Έγγαμος	☐	2. ....
Διαζευγμένος/η	☐	3. ....
Χήρος/α	☐	4. ....

5. Το ερωτώμενο πρόσωπο είναι υπεύθυνο για το νοικοκυριό;

   ΝΑΙ ☐ ΟΧΙ ☐ Εάν όχι, να γραφεί η σχέση προς το υπεύθυνο πρόσωπο
6. Είναι το υπεύθυνο πρόσωπο ασφαλισμένο στην εταιρεία;

   ΝΑΙ ☐ ΟΧΙ ☐

   Πόσα και ποιά μέλη είναι ασφαλισμένα;

   Σύζυγος ☐ Τέκνον ☐ Άλλα μέλη ☐

*Τα ερωτήματα 7-10 αφορούν το υπεύθυνο πρόσωπο για το νοικοκυριό.

7. Επίπεδο εκπαίδευσης:

Μεταπτυχιακό	☐	Πτυχίο ΑΕΙ/ΤΕΙ	☐
Ανωτέρα σχολή	☐	Φοιτά σε ΑΕΙ/ΤΕΙ	☐
Απολυτήριο Λυκείου	☐		
Απολυτήριο Γυμνασίου	☐		
Απολυτήριο Δημοτικού	☐	Δεν τελείωσε	☐

8. Επάγγελμα (Περιγράψτε το ακριβές επάγγελμα)

....

9. Εργάζεται ο/η σύζυγος; ΝΑΙ ☐ ΟΧΙ ☐

10. Μέσο καθαρό μηνιαίο εισόδημα νοικοκυριού:

0-600 ☐ 600-900 ☐ 900-1.500 ☐ 1.500 και άνω ☐

Β. Απόψεις Ερωτώμενου

11. Πώς γνωρίσατε την εταιρεία και τα προϊόντα της;

Από διαφήμιση ☐ Από ασφαλιστή ☐

Από σύσταση φιλικού προσώπου ☐

Με άλλο τρόπο ☐ Δηλαδή...

12. Είστε ικανοποιημένος/η από τα προϊόντα της εταιρείας;

Πάρα πολύ ☐ Πολύ ☐ Μέτρια ☐ Λίγο ☐ Καθόλου ☐

13. α. Αν δεν είστε ικανοποιημένος/η να γραφούν οι λόγοι δυσαρέσκειας

α)..

β)..

γ)..

δ)..

13. β. Πόσες φορές έχετε αποζημιωθεί από την εταιρεία;............

14. Θεωρείτε ότι η ενημέρωση που σας έγινε ήταν αρκετή;

ΝΑΙ ☐ ΟΧΙ ☐

15. Είστε ικανοποιημένος/η από τις υπηρεσίες της εταιρείας και γιατί;

Πάρα πολύ ☐ Πολύ ☐ Μέτρια ☐ Λίγο ☐ Καθόλου ☐

Γιατί;

α)..

β)..

γ)..

16. Είστε ικανοποιημένος/η από τη συνεργασία με τους ασφαλιστές;

Πάρα πολύ ☐ Πολύ ☐ Μέτρια ☐ Λίγο ☐ Καθόλου ☐

και γιατί;

α)..

β)..

γ)..

17. Πόσο συχνά επικοινωνεί ο ασφαλιστής μαζί σας;

....

18. Ποιό είδος ασφάλισης θεωρείτε πιό αναγκαίο; (ιεραρχικά)

1) Ασφάλεια ζωής και νοσοκομειακή κάλυψη ....

2) Παιδικό πρόγραμμα ....

3) Ασφάλεια κατοικίας / επαγγελματικής στέγης ....

4) Συνταξιοδοτικό ....

5) Άλλο ....

19. Πόσο αναγκαία θεωρείτε την απόκτηση μιας από τις προαναφερθείσες ασφάλειες;

Πάρα πολύ ☐ Πολύ ☐ Μέτρια ☐ Λίγο ☐ Καθόλου ☐

και ποιά είναι αυτή;

………………………………………………………………………………………..

20. Κατά τη γνώμη σας υπάρχουν εταιρείες που προσφέρουν προϊόντα:

Πολύ περισσότερα ☐ Περισσότερα ☐ Τα ίδια ☐

Λιγότερα ☐ Δεν γνωρίζω ☐

21. Εάν προσφέρουν περισσότερα, τα γνωρίζετε;

(και από ποιές εταιρείες;)

α)...

β)...

γ)...

22. Κατά τη γνώμη σας υπάρχουν ασφαλιστικές εταιρείες που προσφέρουν υπηρεσίες:

Πολύ καλύτερες ☐ Καλύτερες ☐ Ίδιες ☐

Χαμηλότερες ☐ Δεν γνωρίζω ☐

23. Ποιά είναι η άποψη σας για τις τιμές και τα πλεονεκτήματα προϊόντων ή υπηρεσιών άλλων ασφαλιστικών εταιρειών;

Καλύτερα ☐ Ίδια ☐ Χειρότερα ☐ Δεν γνωρίζω ☐

Ο υπεύθυνος που έδωσε την συνέντευξη...

Ημερομηνία.. Υπογραφή...

Ονοματεπώνυμο ερευνητή...

ΒΙΒΛΙΟΓΡΑΦΙΑ

<u>Α.ΕΛΛΗΝΙΚΗ</u>

Βιβλιοθήκη Κοινωνικής Επιστήμης και Κοινωνικής Πολιτικής. ‘’Εισαγωγή στη Μεθοδολογία και τις Τεχνικές των Κοινωνικών Ερευνών’’, Αθήνα 2000

Δαμιανός Χ., ‘’Μεθοδολογία Δειγματοληψίας: Τεχνικές και Εφαρμογές 2006’’,

Δαμιανός Χ., ‘’Θεωρία Δειγματοληψίας και Στατιστικού Ελέγχου Ποιότητας’’,

Ζαΐρης Εμμ. Ποσειδώνας, ‘’Τεχνικές Δειγματοληπτικών Ερευνών’’, Αθήνα 1991

Καραδήμας Α. Παναγιώτης, ‘’Εισαγωγή στη Δειγματοληψία’’, Πειραιάς 1979

Κιόχος Α. Πέτρος, ‘’Μεθοδολογία Διεξαγωγής Ερευνών’’, Αθήνα 1997

Κιόχος Α. Πέτρος, ‘’Στατιστική’’, Αθήνα 1993

Μπένου Κιμ. Βασίλειος, ‘’Μέθοδοι και Τεχνικές Δειγματοληψίας’’, Πειραιάς 1989

Μπένου Κιμ. Βασίλειος, ‘’Ανάλυση και Τεχνική της Δειγματοληψίας’’,

Ξεκαλάκη Ε., ‘’Τεχνικές Δειγματοληψίας’’, Αθήνα 1995

Ξενάκης Σ. Ανδρέας,’’ Θεωρία και Τεχνικές Δειγματοληψίας’’, Πειραιάς 2002)

Πανταζίδης Ν., ‘’Εισαγωγή στη Δειγματοληψία’’, Αθήνα 1976

Τζωρτζοπούλου Π., ‘’Δειγματοληψία’’

Φαρμάκης Νίκος, ‘’Εισαγωγή στη Δειγματοληψία’’

Χαρίσης Ι. Κώστας, Κιόχος Α. Πέτρος, ‘’Θεωρία Δειγματοληψίας & Εφαρμογές’’, Αθήνα 1997

<u>Β. ΠΑΓΚΟΣΜΙΟΣ ΙΣΤΟΣ</u>

en.wikipedia.org

psychology.ucdavis.edu

www.ento.vt.edu

www.freewebs.com

www.resample.com

zebu.uoregon.edu