

**ΠΑΝΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΟΙΚΟΝΟΜΙΚΩΝ
ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΣΠΟΥΔΩΝ**
Τμήμα Οικονομικής και Περιφερειακής Ανάπτυξης

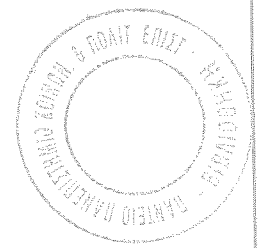
**ΠΛΗΡΟΦΟΡΙΑΚΟ ΣΥΣΤΗΜΑ ΜΕΤΑΔΕΔΟΜΕΝΩΝ ΑΓΟΡΑΣ
ΕΡΓΑΣΙΑΣ ΣΕ ΧΩΡΙΚΟ ΕΠΙΠΕΔΟ-ΣΤΑΤΙΣΤΙΚΕΣ ΕΚΤΙΜΗΣΕΙΣ
ΑΠΟ ΤΗΝ ΕΡΕΥΝΑ ΕΡΓΑΤΙΚΟΥ ΔΥΝΑΜΙΚΟΥ**

Του ΙΩΑΝΝΟΥ ΓΕΡΑΣΙΜΟΥ



Επιβλέπουσα: καθ. Αντ.Λυμπεράκη

ΙΟΥΛΙΟΣ 2005



Στους γονείς μου

ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	Σελίδα
ΠΡΟΛΟΓΟΣ	4
ΕΙΣΑΓΩΓΗ	5
1.ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ ΤΩΝ ΕΚΤΙΜΗΣΕΩΝ ΜΙΚΡΟΠΕΡΙΟΧΩΝ (SMALL AREA ESTIMATION)	7
1.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ – ΟΡΟΛΟΓΙΑ	7
1.2 ΖΗΤΗΣΗ ΓΙΑ ΕΚΤΙΜΗΣΕΙΣ ΜΙΚΡΟΠΕΡΙΟΧΩΝ	8
1.3 ΤΟ ΠΑΡΑΔΕΙΓΜΑ ΤΟΥ EURAREA	10
2.ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΚΑΙ ΕΚΤΙΜΗΣΗ	13
2.1 ΔΕΙΓΜΑΤΟΛΗΨΙΑ	13
2.2 ΕΚΤΙΜΗΣΗ	15
2.2.1 ΚΡΙΤΗΡΙΑ ΕΠΙΛΟΓΗΣ-ΑΞΙΟΛΟΓΗΣΗΣ ΕΚΤΙΜΗΤΩΝ	16
3.ΤΕΧΝΙΚΕΣ ΕΚΤΙΜΗΣΗΣ ΜΙΚΡΟΠΕΡΙΟΧΩΝ	19
3.1 ΑΜΕΣΟΣ ΕΚΤΙΜΗΤΗΣ	19
3.2 ΈΜΜΕΣΟΙ ΕΚΤΙΜΗΤΕΣ	22
3.2.1 ΣΥΝΘΕΤΙΚΟΣ ΕΚΤΙΜΗΤΗΣ	22
3.2.2 ΣΥΝΘΕΤΟΣ ΕΚΤΙΜΗΤΗΣ	27
4.ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ – ΕΦΑΡΜΟΓΕΣ	32
5.ΕΜΠΕΙΡΙΚΗ ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ– ΑΜΕΣΟΙ ΚΑΙ ΈΜΜΕΣΟΙ ΕΚΤΙΜΗΤΕΣ: ΕΚΤΙΜΗΣΕΙΣ ΓΙΑ ΑΠΑΣΧΟΛΗΣΗ ΣΕ ΑΕ & ΕΠΕ	35
5.1 ΣΤΑΤΙΣΤΙΚΗ ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΕΡΕΥΝΗΤΙΚΟΥ ΠΡΟΒΛΗΜΑΤΟΣ	35
5.2 ΔΙΑΔΙΚΑΣΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ ΓΙΑ ΕΚΤΙΜΗΣΗ ΜΙΚΡΟΠΕΡΙΟΧΩΝ	37
5.3 ΕΜΠΕΙΡΙΚΗ ΑΝΑΛΥΣΗ & ΑΠΟΤΕΛΕΣΜΑΤΑ	39
5.4 ΣΥΜΠΕΡΑΣΜΑΤΑ	42
6.ΠΑΡΑΡΤΗΜΑ	43
ΒΙΒΛΙΟΓΡΑΦΙΑ	

ΠΡΟΛΟΓΟΣ

Η μελέτη απευθύνεται σε όσους ενδιαφέρονται για ειδικά θέματα εφαρμοσμένης στατιστικής, καθώς το αντικείμενο εκτιμήσεων μικροπροπεριοχών αποτελεί παρακλάδι της κλασσικής στατιστικής. Η σύγχρονη αντίληψη σε θέματα εκτιμήσεων ξεφεύγει από την παραδοσιακή πρακτική. Η ανάγκη για περισσότερες εκτιμήσεις σε μικρότερο χρονικό διάστημα και με μικρότερο κόστος κάνει επιτακτική την ανάγκη χρησιμοποίησης τεχνικών εκτίμησης, που ξεφεύγουν από τις παραδοσιακές. Άλλωστε δεν είναι τυχαίο ότι οι στατιστικές υπηρεσίες σε πολλά ευρωπαϊκά κράτη χρησιμοποιούν όλο και περισσότερο εναλλακτικές τεχνικές εκτίμησης. Παράλληλα, η σύγχρονη μορφή του Περιφερειακού Προγραμματισμού στην Ελλάδα ισχυροποιεί την ανάγκη για εκτιμήσεις σε μικροπεριοχές ή μικροπεριφέρειες (π.χ νομοί, δήμοι κτλ).

Η μελέτη αυτή είναι ένα από αποτέλεσμα συνεργασίας δύο πολύ σημαντικών ανθρώπων για εμένα, για ξεχωριστούς λόγους ο καθένας. Αρχικά το ερέθισμα για την ενασχόληση με το συγκεκριμένο αντικείμενο δόθηκε από τον Αναστάσιο Καραγάνη Λέκτορα του Παντείου Πανεπιστημίου, η εμπειρία του οποίου βοήθησε ιδιαίτερα κατά την διαδικασία της εμπειρικής εφαρμογής. Το δεύτερο πρόσωπο είναι αυτό της Αντιγόνης Λυμπεράκη καθηγήτριας του Παντείου Πανεπιστημίου, στην οποία οφείλω πολλά στην συνολική παρουσία μου στο ακαδημαϊκό χώρο, καθώς είναι αυτή που έχει πιστέψει περισσότερο από κάθε άλλο στις δυνατότητες μου και δέχθηκε να είναι επιβλέπουσα καθηγήτρια μου.

ΕΙΣΑΓΩΓΗ

Οι εθνικές στατιστικές υπηρεσίες παράγουν εκτιμήσεις τόσο για το σύνολο της χώρας όσο και για μεγάλες χωρικές μονάδες, όπως οι περιφέρειες. Τα τελευταία χρόνια παρουσιάζεται ιδιαίτερο ενδιαφέρον για εκτιμήσεις σε επίπεδο νομών, δήμων και κοινοτήτων, με σκοπό τον καλύτερο σχεδιασμό της περιφερειακής ανάπτυξης.

Το ενδιαφέρον γύρω από εκτιμήσεις μικρών περιοχών δεν περιορίζεται μόνο στον δημόσιο τομέα, καθώς βρίσκει πρακτικό ενδιαφέρον και στον ιδιωτικό τομέα κατά την λήψη επιχειρηματικών αποφάσεων. Η απαίτηση για επιπρόσθετες πληροφορίες και η έλλειψη πόρων για να σχεδιαστούν τα απαραίτητα συστήματα δεδομένων, για εκτιμήσεις κοινωνικοοικονομικών μεγεθών μικροπεριοχών με γεωγραφικά ή δημογραφικά όρια, έχουν οδηγήσει τις εθνικές στατιστικές υπηρεσίες να εξετάσουν εναλλακτικές μεθόδους εκτιμήσεων. Οι έμμεσοι εκτιμητές αποτελούν την εναλλακτική πρόταση απέναντι στους κλασικούς άμεσους εκτιμητές.

Οι έμμεσοι εκτιμητές αντίθετα με τους άμεσους δεν περιορίζονται ως προς την πηγή των δεδομένων που χρησιμοποιούνται για την παραγωγή εκτιμήσεων. Το αυξανόμενο ενδιαφέρον για τους έμμεσους (μη παραδοσιακούς) εκτιμητές οδήγησε στην ευρύτερη χρήση του όρου «μικροπεριοχή εκτίμησης» (small area estimation) που αποτελεί ξεχωριστό αντικείμενο ενδιαφέροντος της στατιστικής.

Σκοπός αυτής της μελέτης είναι να τεκμηριώσει τις πρακτικές και τις μεθόδους εκτίμησης, με τις οποίες παράγονται έμμεσες εκτιμήσεις και πως αυτές βρίσκουν εφαρμογή στα πραγματικά δεδομένα. Στην μελέτη περιλαμβάνονται οι ορισμοί και τα χαρακτηριστικά των άμεσων και έμμεσων εκτιμητών, καθώς και η εφαρμογή τους για εκτιμήσεις απασχόλησης.

Συγκεκριμένα στο πρώτο κεφάλαιο παρουσιάζεται αντικείμενο των εκτιμήσεων μικροπεριοχής και το εννοιολογικό περιεχόμενο των όρων που σχετίζονται με το αντικείμενο αυτό. Παράλληλα στο πρώτο κεφάλαιο καλύπτεται η αιτιολόγηση της ανάγκης για εκτιμήσεις μικροπεριοχής. Η ευελιξία των τεχνικών εκτιμήσεων μικροπεριοχής δικαιολογούν την αυξημένη τους ζήτηση τόσο από τον δημόσιο, όσο και από τον ιδιωτικό τομέα.

Στο δεύτερο κεφάλαιο αναπτύσσονται θεμελιώδεις γνώσεις δειγματοληψίας και εκτιμητικής. Άλλωστε, η προσέγγιση κάθε νέου τύπου στατιστικής ανάλυσης βασίζεται στις θεμελιώδεις γνώσεις στατιστικής, ανεξάρτητα με τον τελικό σκοπό της ανάλυσης. Στο κεφάλαιο αυτό ξεχωριστό ενδιαφέρον έχουν τα κριτήρια αξιολόγησης των εκτιμητών και ειδικά το μέσο τετραγωνικό σφάλμα και οι παραλλαγές του, που αποτελούν το βασικό κριτήριο αξιολόγησης στο αντικείμενο των εκτιμητών μικροπεριοχών.

Το τρίτο κεφάλαιο αποτελεί το σημαντικότερο κομμάτι της μελέτης, καθώς παρουσιάζονται ο άμεσος και ο έμμεσος εκτιμητής. Όσον αφορά μάλιστα τους έμμεσους εκτιμητές παρουσιάζονται οι περιπτώσεις του συνθετικού και του σύνθετου εκτιμητή. Σκοπός του κεφαλαίου είναι η πληρέστερη ανάλυση των άμεσων και έμμεσων εκτιμητών, έτσι ώστε να είναι δυνατή η αξιολόγηση για την χρήση τους κατά περίπτωση. Το μόνο σίγουρο είναι ότι όσο πιο μικρή η περιοχή ενδιαφέροντος, και όσο πιο λίγα τα διαθέσιμα στοιχεία για αυτή, τόσο πιο αναγκαία κρίνεται η χρησιμοποίηση των έμμεσων εκτιμητών.

Στο τέταρτο κεφάλαιο γίνεται μια βιβλιογραφική επισκόπηση στα σημαντικότερα και πιο συναφή συγγραφικά κείμενα με αναφορά στις εκτιμήσεις μικροπεριοχών. Αξίζει να σημειώσουμε ότι το σύνολο των αναφορών είναι ξενόγλωσσες.

Στο πέμπτο και τελευταίο κεφάλαιο αναλύεται η εμπειρική εφαρμογή των τεχνικών εκτίμησης μικροπεριοχών. Η εφαρμογή έχει ως στόχο την εξέταση της συμπεριφοράς των έμμεσων εκτιμητών σε σχέση με τον παραδοσιακό άμεσο εκτιμητή. Η μελέτη περίπτωσης αφορά εκτιμήσεις για την συνολική απασχόληση της μικροπεριοχής νομός βασιζόμενη σε δεδομένα του συνόλου των επιχειρήσεων ΑΕ & ΕΠΕ από την βάση δεδομένων της ICAP. Αυτό έγινε προκειμένου να προσομοιώσουμε πειραματικά την Ερευνα Εργατικού Δυναμικού της ΕΣΥΕ, έτσι ώστε να μπορούμε να ελέγξουμε της ερευνητικές υποθέσεις. Κατά συνέπεια η επαλήθευση των υποθέσεων οδηγεί στην χρήση των μεθόδων στην Ερευνα Εργατικού Δυναμικού.

ΚΕΦΑΛΑΙΟ 1^ο

ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ ΤΩΝ ΕΚΤΙΜΗΣΕΩΝ ΜΙΚΡΟΠΕΡΙΟΧΩΝ (SMALL AREA ESTIMATION)

1.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ - ΟΡΟΛΟΓΙΑ

Η περιγραφή των βασικών όρων που αναφέρονται στη μελέτη βοηθάει στην καλύτερη κατανόηση του τρόπου και του λόγου που αυτοί χρησιμοποιούνται. Όροι όπως «μικροπεριοχή», «άμεσοι» και «έμμεσοι» εκτιμητές, σκιαγραφούν το αντικείμενο των τεχνικών εκτίμησης που περιγράφονται στην παρούσα εργασία.

Σε κάθε έρευνα που διενεργείται υπάρχει ένα αντικείμενο μελέτης. Για την στατιστική ανάλυση του αντικειμένου ενδιαφέροντος του δίνουμε μορφή μεταβλητής που παίρνει συγκεκριμένες τιμές. Η μεταβλητή αυτή καλείται μεταβλητή ενδιαφέροντος. Στις περισσότερες μελέτες οι εκτιμήσεις για την μεταβλητή ενδιαφέροντος αναφέρονται στον μέσο όρο, στο σύνολο ή ακόμα και στην αναλογία (ποσοστό). Για παράδειγμα μας ενδιαφέρει να εκτιμήσουμε τον μέσο όρο, το σύνολο ή ακόμα και την αναλογία των απασχολούμενων σε μια πόλη.

Η επικράτηση του όρου «μικροπεριοχή» εκτίμησης (*small area estimation*) είναι αποτέλεσμα δύο παραγόντων. Ο πρώτος παράγοντας δικαιολογεί την χρήση του όρου «μικρή», και αναφέρεται στο γεγονός της ύπαρξης μικρού αριθμού δειγματικών παρατηρήσεων που είναι διαθέσιμες για την εκτίμηση μιας μεταβλητής στην περιοχή ενδιαφέροντος. Ουσιαστικά με τον όρο μικρή τονίζεται η ανάγκη για την χρησιμοποίηση τεχνικών που να αντιμετωπίζουν το πρόβλημα της μεγάλης διακύμανσης των άμεσων εκτιμητών, που στηρίζονται μόνο στον μικρό αριθμό των παρατηρήσεων του δείγματος που αντιστοιχούν στην περιοχή μελέτης. Ο δεύτερος παράγοντας δικαιολογεί την χρήση όρου «περιοχή», και αναφέρεται στο γεγονός ότι στις περισσότερες εφαρμογές εκτιμήσεων μεταβλητών ενδιαφέροντος το πεδίο ενδιαφέροντος είναι μια σχετικά μικρή γεωγραφική περιοχή. Με άλλα λόγια θα λέγαμε ότι με τις τεχνικές εκτίμησης που αναλύονται στην συγκεκριμένη μελέτη αντιμετωπίζονται κυρίως προβλήματα με χωρική διάσταση.

Βέβαια, θα πρέπει να τονίσουμε ότι ο όρος μικροπεριοχή είναι δυνατόν να αναφέρεται και στην ανάγκη εκτιμήσεων για μια κοινωνική και δημογραφική ομάδα, χωρίς να είναι απαραίτητη η χωρική διάσταση. Μάλιστα, συχνά στην βιβλιογραφία χρησιμοποιείται ο όρος πεδίο ή τομέας (domain) ενδιαφέροντος για την προσέγγιση τέτοιων προβλημάτων εκτίμησης μικροπεριοχών.

Στην σχετική βιβλιογραφία γίνεται βασική διάκριση μεταξύ άμεσων (*direct*) και έμμεσων (*indirect*) εκτιμήσεων μικρών περιοχών. Ένας άμεσος εκτιμητής χρησιμοποιεί παρατηρήσεις της μεταβλητής ενδιαφέροντος που προέρχονται από την χρονική περίοδο ενδιαφέροντος και τις μονάδες της περιοχής ενδιαφέροντος και μόνο. Αντίθετα, ένας έμμεσος εκτιμητής είναι δυνατόν να χρησιμοποιεί παρατηρήσεις της μεταβλητής ενδιαφέροντος που προέρχονται από διαφορετική περιοχή ή/και χρονική περίοδο σε σχέση με την περιοχή και την χρονική περίοδο ενδιαφέροντος.

Ο όρος «άμεσος» συναντάται και ως «αμερόληπτος», «έγκυρος», «τυπικός» κ.α. Γενικά ο άμεσος εκτιμητής αποτελεί τον παραδοσιακό εκτιμητή, που χρησιμοποιεί δεδομένα της περιοχής ενδιαφέροντος και μόνο, που προέρχονται από γενικές έρευνες (π.χ Ερευνα Εργατικού Δυναμικού, Απογραφή πληθυσμού κ.α). Η χρήση του όρου «έμμεσος» ή «μη παραδοσιακός» εκτιμητής εκφράζει την διαφορά που έχει από έναν άμεσο εκτιμητή, καθώς ο έμμεσος εκτιμητής χρησιμοποιεί κάτω από προϋποθέσεις, μία ή περισσότερες βοηθητικές μεταβλητές που είναι γνωστές για την περιοχή ενδιαφέροντος. Άλλωστε αυτός είναι και ο λόγος που συχνά αναφέρονται και ως εκτιμητές δανειζόμενης δύναμης (*borrow strength estimator*).

1.2 ΖΗΤΗΣΗ ΓΙΑ ΕΚΤΙΜΗΣΕΙΣ ΜΙΚΡΟΠΕΡΙΟΧΩΝ

Ο παράγων χώρος αποτελείται από ένα σύνολο μικροπεριοχών. Η διαίρεση του χώρου σε μικροπεριφέρειες είναι αποτέλεσμα συνήθως οργανωτικών αναγκαιοτήτων της διοίκησης των επιχειρήσεων ή του δημόσιου τομέα.

Η γνώση δεδομένων μικροπεριοχών (π.χ. νομοί, πόλεις) κάνει ευκολότερη την εμπειρική μελέτη και κατανόηση των κοινωνικοοικονομικών και δημογραφικών διαφορών μεταξύ τους. Ο πρωταρχικός σκοπός τέτοιων μελετών είναι η διεύρυνση

της γνώσης για τους φορείς λήψεων αποφάσεων και χάραξης πολιτικής. Κοινωνιολόγοι, δημογράφοι, οικονομολόγοι, γεωγράφοι, ιατροί, ερευνητές αγοράς είναι μερικές μόνο από τις κατηγορίες επιστημόνων που αναζητούν την γνώση εκτιμήσεων μικροπεριοχών. Γενικότερα η ζήτηση για εκτιμήσεις μικροπεριοχών προέρχεται τόσο από τον ιδιωτικό όσο και από τον δημόσιο τομέα.

Με την γνώση αυτή ενισχύεται η λήψη αποφάσεων κρατικών και τοπικών φορέων. Η γνώση είναι απαραίτητη για το σχεδιασμό της περιφερειακής ανάπτυξης, και ειδικότερα για την κατάρτιση του προϋπολογισμού όσον αφορά στη διάθεση των δημοσίων δαπανών και την επιλεξιμότητα των δημοσίων έργων. Ειδικά για την περιφερειακή ανάπτυξη, ο Steinberg (1979) αναφέρει ότι ένα καλό τοπικό πρόγραμμα ανάπτυξης απαιτεί καλές τοπικές εκτιμήσεις. Η ανάλυση των επιπτώσεων της πολιτικής που ασκείται από τις διοικητικές μονάδες πρέπει να γίνεται σε όσο το δυνατόν χαμηλότερο γεωγραφικό επίπεδο αναφοράς και σίγουρα μικρότερο από τις διοικητικές περιφέρειες.

Η γνώση δεδομένων για μικροπεριοχές ενισχύουν μια σειρά επιχειρηματικών αποφάσεων, που ως γνωστόν συνυπολογίζουν σε μεγάλο βαθμό τα κοινωνικοοικονομικά χαρακτηριστικά μιας περιοχής. Η αξία της διαθεσιμότητας τέτοιων δεδομένων ενισχύεται αν αναλογιστεί κανείς πόσο σημαντικά είναι για την επιλογή του τόπου εγκατάστασης, την εκτίμηση των πωλήσεων, την διερεύνηση του καταναλωτικού προτύπου κ.α. Ειδικά η επιλογή του τόπου εγκατάστασης αποτελεί κομβική απόφαση τόσο για την επιχείρηση όσο και για μια κυβέρνηση καθώς δημιουργεί πολλαπλασιαστικές επιδράσεις σε μια περιοχή. Η απόφαση αυτή επηρεάζεται σε μεγάλο βαθμό από την γνώση της σύνθεσης και της ποιότητας του εργατικού δυναμικού σε τοπικό επίπεδο.

Ξεχωριστό ενδιαφέρον δίνει η ζήτηση εκτιμήσεων μικροπεριοχών από χώρες τις Κεντρικής και Ανατολικής Ευρώπης (συμπεριλαμβανομένης και της πρώην Σοβιετικής Ένωσης), που βρίσκονται στο στάδιο της μετάβασης. Κατά την δεκαετία του 90 οι χώρες αυτές πέρασαν από ένα συγκεντρωτικό σε ένα αποκεντρωτικό σύστημα λήψης αποφάσεων. Μια σειρά από άρθρα (*Statistics in transition*¹) για την

¹ Kordos, J. (Editor), (2000-2004), *Statistics in transition*, Journal of the Polish Statistical Association

εφαρμογή μεθόδων εκτίμησης μικρών περιοχών σε βασικά κοινωνικοοικονομικά μεγέθη (π.χ απασχόληση, ανεργία κτλ), καθώς και η διεξαγωγή διεθνών συνεδρίων στην Βαρσοβία (1992) και στη Ρίγα (1999) με το ίδιο θέμα αποτελεί την καλύτερη απόδειξη για το ενδιαφέρον των χωρών αυτών για το θέμα.

Συνοψίζοντας, οι εκτιμήσεις μικροπεριοχών ενισχύουν τη λήψη αποφάσεων στο δημόσιο και στον ιδιωτικό τομέα. Οι εκτιμήσεις αυτές αποτελούν τις πρώτες ύλες για την κατασκευή ποικίλων κοινωνικοοικονομικών δεικτών για μεμονωμένες γεωγραφικές περιοχές, οι οποίοι είναι χρήσιμοι για διάφορες κατηγορίες επιστημόνων.

1.3 ΤΟ ΠΑΡΑΔΕΙΓΜΑ ΤΟΥ EURAREA²

Το EURAREA είναι ένα τριετές ερευνητικό πρόγραμμα (2001-2004), το οποίο χρηματοδοτείται από την Ευρωπαϊκή Ένωση και στο οποίο συμμετέχουν έξι ευρωπαϊκά εθνικά στατιστικά ινστιτούτα. Σκοπός του προγράμματος είναι η αξιολόγηση της αποτελεσματικότητας των «τυπικών»³ και καινοτόμων προσεγγίσεων για εκτιμήσεις μικροπεριοχών και πώς αυτές είναι δυνατόν να εφαρμοστούν ως εναλλακτικές μέθοδοι εκτιμήσεων για την αντικατάσταση παραδοσιακών στατιστικών ερευνών (π.χ απογραφές). Η εφαρμογή των εκτιμητών βασίζεται σε πραγματικά δεδομένα των συμμετέχουσων χωρών στο πρόγραμμα και η συνολική αξιολόγηση θα αφορά την εμπειρία από την εφαρμογή των εκτιμητών στις χώρες αυτές.

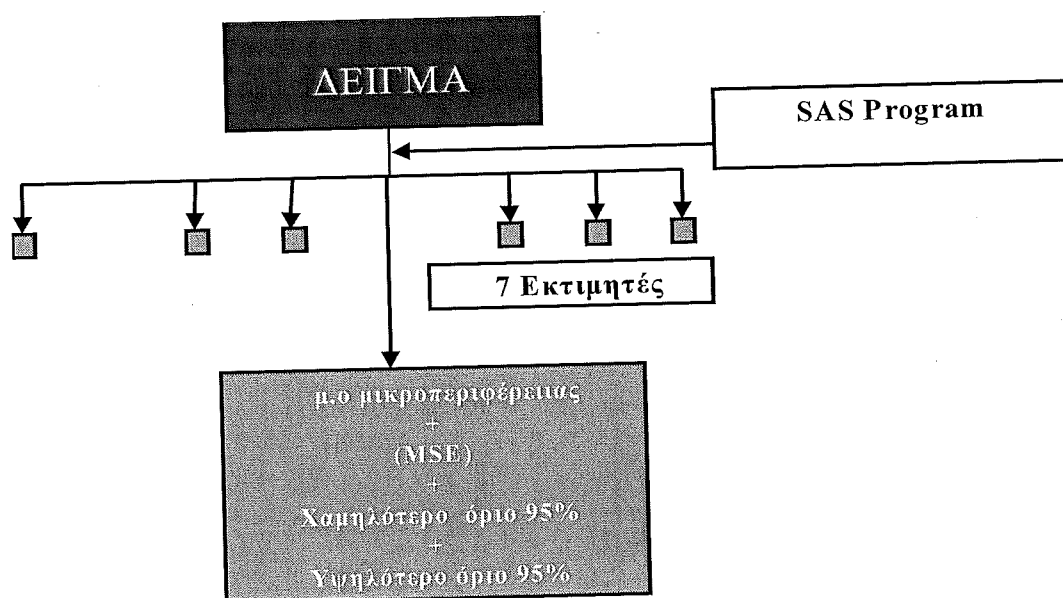
Ένας επιπλέον στόχος του προγράμματος είναι η επιβεβαίωση της λειτουργικότητας και αποτελεσματικότητας των «τυπικών» εκτιμητών και σε άλλες χώρες που δεν λαμβάνουν μέρος στο πρόγραμμα. Για το λόγο αυτό χρησιμοποιείται κώδικας προγραμματισμού στο στατιστικό πακέτο SAS, ο οποίος θα μπορεί να εφαρμοστεί εύκολα από άλλα εθνικά στατιστικά ινστιτούτα.

² Αντίληψη του προγράμματος και περισσότερες πληροφορίες στο Office for National Statistics

³ Τυπικών με την έννοια ότι αφορούν αποκλειστικά εκτιμητές για εκτιμήσεις μικροπεριοχών (συνθετικός, σύνθετος GRE εκτιμητές κτλ).

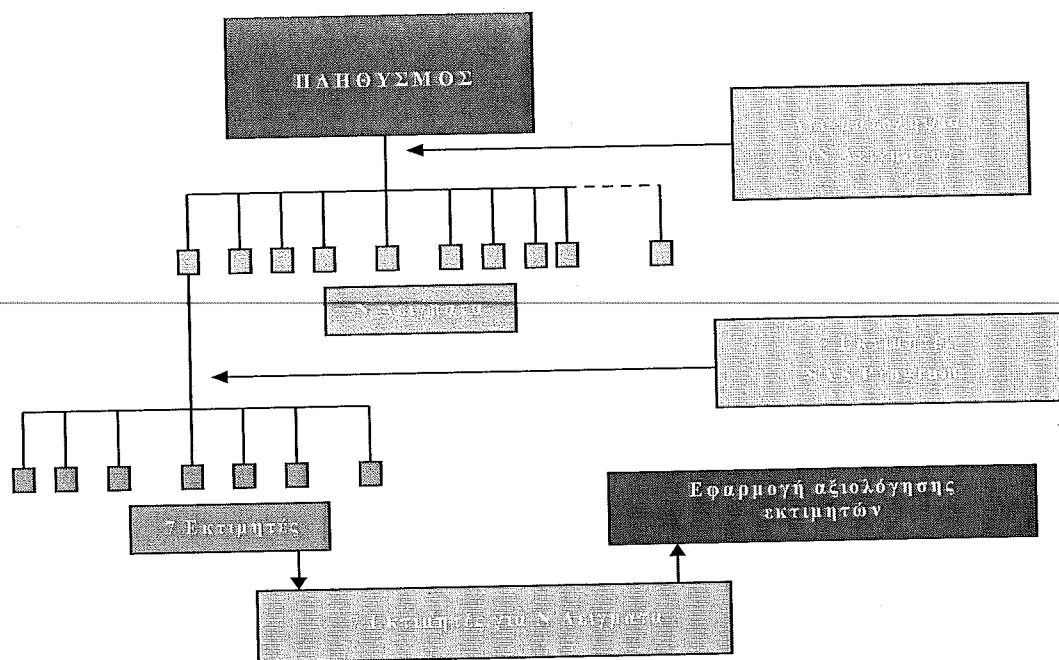
Οι μεταβλητές ενδιαφέροντος είναι η σύνθεση των νοικοκυριών, η ανεργία και το εισόδημα ανά νοικοκυριό. Οι δύο πρώτες αναφέρονται σε δυαδική μεταβλητή, ενώ η τρίτη σε συνεχή. Η κάθε μεταβλητή ενδιαφέροντος εκτιμάται σε κάθε χώρα σε δύο χωρικά επίπεδα μικροπεριοχών (π.χ NUTS 3, NUTS 5 για την Αγγλία).

Στο πρώτο στάδιο του EURAREA εφαρμόζονται 7 τυπικοί εκτιμητές μέσω προγραμματισμού στο SAS, οι οποίοι και αξιολογούνται. Οι εκτιμητές που υπολογίζονται είναι οι εξής: άμεσος, GREG, τρία μοντέλα συνθετικών, και δύο μοντέλα EBLUP εκτιμητών. Έτσι αξιοποιώντας την πληροφορία από το δείγμα παράγονται εκτιμήσεις για τον μέσο όρο της μεταβλητής ενδιαφέροντος, το μέσο τετραγωνικό σφάλμα αλλά και το διάστημα εμπιστοσύνης.



Διαδικασία εκτίμησης των 7 τυπικών εκτιμητών EURAREA

Η αξιολόγηση της μεθοδολογίας του EURAREA προσεγγίζεται με μια διαδικασία προσομοίωσης κατά την οποία λαμβάνονται N δείγματα και παράγονται οι 7 εκτιμητές. Η διαδικασία ολοκληρώνεται με την αξιολόγηση των εκτιμητών στη βάση στατιστικών κριτηρίων.



Διαδικασία Προσομοίωσης των 7 τυπικών εκτιμητών EURAREA

Τα κριτήρια αξιολόγησης που χρησιμοποιούνται αναφέρονται στο μέσο τετραγωνικό σφάλμα, στο συντελεστή Gini καθώς και σε εναλλακτικά κριτήρια που έχουν ως βάση τους τα προηγούμενα δύο, όπως η ρίζα του μέσου τετραγωνικού σφάλματος ή σχετική μεροληψία του συντελεστή Gini.

ΚΕΦΑΛΑΙΟ 2^ο

ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΚΑΙ ΕΚΤΙΜΗΣΗ

Η στατιστική συμπερασματολογία περιλαμβάνει μεθόδους με τις οποίες είναι δυνατή η προσέγγιση των χαρακτηριστικών ενός πληθυσμού μέσα από την μελέτη ενός δείγματος. Η επαγωγή ορθών συμπερασμάτων είναι συνάρτηση των επιλογών μας για τις τεχνικές εκτίμησης και δειγματοληψίας που θα χρησιμοποιήσουμε. Έτσι, μια καλή εκτίμηση της μέσης τιμής μιας μεταβλητής ενδιαφέροντος προϋποθέτει μια καλή επιλογή της τεχνικής δειγματοληψίας.

2.1 ΔΕΙΓΜΑΤΟΛΗΨΙΑ

Μια δειγματοληπτική έρευνα βασίζεται σε ένα ή περισσότερα δείγματα, τα οποία αντλούνται από έναν πληθυσμό. Η διαδικασία άντλησης δεν είναι όσο απλή φαίνεται, καθώς θα πρέπει να πληρούνται κάποιες βασικές προϋποθέσεις-στάδια.

Τα βασικότερα στάδια μιας διαδικασίας δειγματοληψίας είναι:

Πληθυσμός στόχος: Ο πληθυσμός από τον οποίο επιλέγουμε το δείγμα αποτελεί τον υπό δειγματοληψία πληθυσμό και πρέπει να συμπίπτει με τον πληθυσμό στόχο, στον οποίο αναφέρονται τα στατιστικά αποτελέσματα. Στην έρευνα για την συνολική απασχόληση σε επιχειρήσεις του ιδιωτικού τομέα, το πλήθος των επιχειρήσεων ιδιωτικού τομέα της χώρας αποτελεί και τον πληθυσμό στόχο.

Δειγματοληπτικό πλαίσιο και Βοηθητικές πληροφορίες: Ένα δείγμα μπορεί να επιλεγεί από έναν κατάλογο που απαριθμεί τις μονάδες (άτομα, επιχειρήσεις, κ.λπ.), που αποτελούν τον πληθυσμό. Αυτός ο κατάλογος καλείται δειγματοληπτικό πλαίσιο ή πλαίσιο δειγμάτων για την έρευνα. Η ύπαρξη ενός τέτοιου πλαισίου είναι απαραίτητη για τη χρήση πολλών δειγματοληπτικών μεθόδων.

Επιπλέον, η εφαρμογή αυτών των μεθόδων συχνά απαιτεί ότι ένα πλαίσιο περιέχει περισσότερα από ένα προσδιοριστικά στοιχεία (παραδείγματος χάριν

διευθύνσεις, ονόματα) για τις μονάδες που αποτελούν έναν πληθυσμό. Παραδείγματος χάριν, η στρωματοποιημένη δειγματοληψία απαιτεί από το πλαίσιο να περιέχει αρκετές πληροφορίες για κάθε μονάδα πληθυσμού. Γενικά, αναφερόμαστε σε αυτές τις πληροφορίες ως *βοηθητικές πληροφορίες*. Χαρακτηριστικά, αυτές οι βοηθητικές πληροφορίες αποτελούν χαρακτηριστικά του πληθυσμού, που συσχετίζονται με τις μεταβλητές ενδιαφέροντος που μετριοούνται στην έρευνα. Ο βαθμός κάλυψης του πληθυσμού από το δειγματοληπτικό πλαίσιο είναι ένα βασικό μέτρο της ποιότητας δειγματοληπτικών ερευνών.

Μέθοδος Πιθανοθεωρητικής Δειγματοληψίας⁴ - Μέγεθος Δείγματος: Με την δειγματοληψία με πιθανότητα η επιλογή των μονάδων από το δειγματοληπτικό πλαίσιο που θα αποτελέσουν το δείγμα βασίζεται στην πιθανότητα που έχει κάθε μονάδα να εμφανίζεται στο δείγμα. Με αυτόν τον τύπο μεθόδου επιλογής, δεν είναι δυνατό να διευκρινιστεί εκ των προτέρων ποιες μονάδες του δειγματοληπτικού πλαισίου θα αποτελούν το δείγμα.

Στην πραγματικότητα είναι δυνατόν να έχουμε δειγματοληψία με ίσες και δειγματοληψία με άνισες πιθανότητες. Στην πρώτη περίπτωση κάθε μονάδα έχει ίση πιθανότητα να είναι στο δείγμα, ενώ στην δεύτερη κάθε μονάδα έχει πιθανότητα ανάλογη ενός μετρήσιμου μεγέθους που σχετίζεται με αυτήν. Σε ότι ακολουθεί κάνουμε τη βασική υπόθεση ότι η δειγματοληπτική μέθοδος πιθανότητας είναι τέτοια που κάθε μονάδα στο πλαίσιο έχει μια γνωστή μη μηδενική και διαφορετική πιθανότητα επιλογής στο δείγμα.

Οι πιο συνηθισμένες μέθοδοι δειγματοληψίας για την επιλογή τυχαίων δειγμάτων είναι:

- Απλή τυχαία δειγματοληψία
- Στρωματοποιημένη δειγματοληψία
- Συστηματική δειγματοληψία
- Πολυσταδιακή δειγματοληψία συστάδων

⁴ Πέραν της πιθανοθεωρητικής δειγματοληψίας υπάρχει και η δειγματοληψία χωρίς πιθανότητα με την οποία παράγονται μη τυχαία δείγματα ή όπως αλλιώς χαρακτηρίζονται δείγματα ευκολίας. Σε τέτοιες περιπτώσεις τα δειγματοληπτικά σφάλματα δεν έχουν νόημα και δεν μπορούν να υπολογιστούν.

Η τυχαία δειγματοληψία αναφέρεται στην λήψη του δείγματος από έναν πληθυσμό με τυχαίο τρόπο. Αντίθετα στην στρωματοποιημένη δειγματοληψία διαμερίζονται τα στοιχεία του αρχικού πληθυσμού σε υποσύνολα (στρώματα), ενώ το δείγμα είναι το σύνολο των επιμέρους δειγμάτων ανά στρώμα που λαμβάνεται με απλή τυχαία δειγματοληψία. Μια ουσιαστική διαφοροποίηση, τόσο της απλής όσο και της στρωματοποιημένης δειγματοληψίας, αποτελεί η συστηματική δειγματοληψία. Η μέθοδος βασίζεται σε ένα δείγμα που αντλείται με συστηματικό τρόπο (βήμα), που προσδιορίζεται από το μέγεθος του δείγματος σε σχέση με τον πληθυσμό. Σε πληθυσμούς μεγάλου μεγέθους αποδεικνύεται χρήσιμη η εφαρμογή των αρχών της τυχαίας δειγματοληψίας σε δύο ή περισσότερα στάδια (πολυσταδιακή δειγματοληψία).

Τα παραπάνω αναφέρονται στον τρόπο που θα αντληθεί το δείγμα. Όμως θα πρέπει να λαμβάνεται υπόψη και το μέγεθος του δείγματος που θα αντληθεί. Στην επιλογή του κατάλληλου μεγέθους δείγματος τα δύο κεντρικά ζητήματα είναι το διαθέσιμο ποσό (κόστος) για την δειγματοληψία και η απαιτούμενη ακρίβεια των εκτιμήσεων. Ειδικότερα το κατάλληλο μέγεθος είναι συνάρτηση του μεγέθους και της διασποράς του πληθυσμού.

2.2 ΕΚΤΙΜΗΣΗ

Η εκτίμηση αποτελεί ένα από τα δύο κομμάτια του περιεχομένου της στατιστικής επαγωγής (το άλλο είναι ο στατιστικός έλεγχος), δηλαδή της διαδικασίας συναγωγής συμπερασμάτων για τον πληθυσμό με βάση τις πληροφορίες από το δείγμα. Οι πληροφορίες του δείγματος χρησιμοποιούνται με αντικειμενικό σκοπό την εκτίμηση της άγνωστης τιμής μιας ή περισσότερων παραμέτρων του πληθυσμού (σημειακή εκτίμηση) ή την εκτίμηση ενός διαστήματος (εκτίμηση διαστήματος), μέσα στο οποίο θα περιλαμβάνεται η αληθινή αλλά άγνωστη τιμή της παραμέτρου.

Η παράμετρος θ μπορεί να είναι ο μέσος, η διακύμανση κλπ και σχετίζεται με μια τυχαία μεταβλητή (y) την οποία μελετάμε. Για την εκτίμηση της άγνωστης παραμέτρου, συνδυάζουμε τις εκ των προτέρων πληροφορίες που είναι δυνατόν να έχουμε, με τις πληροφορίες που μας δίνει ένα τυχαίο δείγμα ($y_1, \dots, y_n$). Η επιλογή της συνάρτησης, που συνδέει της παρατηρήσεις των διαφόρων δειγμάτων με την

άγνωστη τιμή της παραμέτρου, αποτελεί σημαντικό παράγοντα για την παραγωγή αποδεκτών εκτιμήσεων της παραμέτρου του πληθυσμού. Η συνάρτηση αυτή καλείται εκτιμητής, ενώ η τιμή που παίρνει ο εκτιμητής θ για ένα δεδομένο δείγμα ονομάζεται εκτίμηση της παραμέτρου θ και συμβολίζεται με $\hat{\theta}$.

Οι τιμές που παίρνει ένας εκτιμητής θα διαφέρουν από δείγμα σε δείγμα και συνεπώς δεν μπορεί να παίρνει πάντα την πραγματική τιμή της παραμέτρου του πληθυσμού. Η διαφορά μεταξύ της τιμής του εκτιμητή για ένα δεδομένο δείγμα και της αληθινής τιμής της παραμέτρου στον πληθυσμό ονομάζεται σφάλμα δειγματοληψίας και είναι διαφορετικό για κάθε δείγμα. Η διαφορά μεταξύ της προσδοκώμενης τιμής του εκτιμητή και της αληθινής τιμής της παραμέτρου ονομάζεται σφάλμα μεροληψίας και είναι σταθερό για κάθε δείγμα. Σκοπός μιας διαδικασίας εκτίμησης είναι η ελαχιστοποίηση του συνολικού σφάλματος του εκτιμητή της μεταβλητής ενδιαφέροντος.

2.2.1 ΚΡΙΤΗΡΙΑ ΕΠΙΛΟΓΗΣ-ΑΞΙΟΛΟΓΗΣΗΣ ΕΚΤΙΜΗΤΩΝ

Για μια συγκεκριμένη παράμετρο μπορεί να προτείνονται περισσότεροι από ένας εκτιμητές για αυτό και χρειαζόμαστε κάποια κριτήρια για να μπορούμε να αξιολογούμε και να επιλέξουμε τον καλύτερο. Τα κριτήρια που ουσιαστικά είναι οι ιδιότητες που πρέπει να χαρακτηρίζουν έναν «καλό» εκτιμητή αναφέρονται στην: α) αμεροληψία, β) αποτελεσματικότητα, γ) ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος και δ) συνέπεια.

Αμεροληψία (unbiasedness)

Ένας εκτιμητής είναι αμερόληπτος όταν η δειγματική κατανομή του έχει μέση τιμή ίση με την τιμή της παραμέτρου θ . Με άλλα λόγια όταν η μέση τιμή όλων των δυνατών δειγμάτων είναι ίση με την πραγματική τιμή στον πληθυσμό. Από έναν αμερόληπτο εκτιμητή παίρνουμε αμερόληπτες εκτιμήσεις. Η αμεροληψία ως κριτήριο είναι επαρκή αλλά όχι ικανό να μας οδηγήσει στην επιλογή του καλύτερου εκτιμητή, καθώς είναι πολύ πιθανό να έχουμε παραπάνω από έναν αμερόληπτο εκτιμητή.

Αποτελεσματικότητα (*efficiency*)

Το κριτήριο της αποτελεσματικότητας αφορά στη διακύμανση της δειγματικής κατανομής ενός εκτιμητή. Αποτελεσματικότερος είναι ο εκτιμητής με την μικρότερη διακύμανση. Παρά το γεγονός ότι περιμένουμε ένας αμερόληπτος εκτιμητής να έχει μικρότερη διακύμανση από έναν μεροληπτικό, κάτι τέτοιο δεν συμβαίνει πάντα. Ένας μεροληπτικός εκτιμητής είναι δυνατόν να έχει μικρότερη διακύμανση από έναν αμερόληπτο εκτιμητή, αν η διακύμανση του είναι αρκετά μικρή ώστε να αντισταθμίζει το σφάλμα μεροληψίας..

Ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (*Mean square error*)

Λόγω της φαινομενικής αδυναμίας της διακύμανσης να αποτελέσει σαφή κριτήριο επιλογής μεταξύ δύο ή περισσότερων εκτιμητών χρησιμοποιείται το κριτήριο της ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος. Το κριτήριο αποτελεί μια γενικότερη περίπτωση του κριτηρίου της αποτελεσματικότητας. Το ενδιαφέρον εντοπίζεται στην επιλογή του εκτιμητή που κατά μέσο όρο δίνει εκτιμήσεις που απέχουν λιγότερο από την πραγματική τιμή της υπό εκτίμηση παραμέτρου. Δηλαδή υποδεικνύεται ο εκτιμητής που «σφάλει» λιγότερο. Με τον όρο σφάλμα εννοούμε την διαφορά μιας εκτίμησης με την τιμή της παραμέτρου πληθυσμού είτε σε απόλυτη τιμή $|g - \theta|$, είτε στο τετράγωνο $(\theta - \theta)^2$, που αποτελεί το τετραγωνικό σφάλμα.

Το μέσο τετραγωνικό σφάλμα ενός εκτιμητής ισούται με την διακύμανση του εκτιμητή σύν το τετράγωνο της μεροληψίας του⁵.

$$\text{Μέσο τετραγωνικό σφάλμα} = \text{Διακύμανση} + (\text{Μεροληψία})^2$$

$$E (\theta - \theta)^2 = \text{Var } \theta + \text{Bias } \theta^2$$

⁵ Αυτό είναι λογικό αν αναλογιστεί κανείς τον ρόλο που παίζει το μέσο τετραγωνικό σφάλμα όσο αφορά τον συνυπολογισμό της μεροληψίας και της διακύμανσης που από μόνες τους δεν μπορούν πάντα να δώσουν τον καλύτερο εκτιμητή.

Αν γνωρίζω ότι η συγκρινόμενοι εκτιμητές είναι αμερόληπτοι τότε το ποσό της μεροληψίας θα είναι μηδέν όποτε η επιλογή μπορεί να γίνει στην βάση της αποτελεσματικότητας.

Η σύγκριση δύο εκτιμητών θ_1, θ_2 , στην βάση του κριτηρίου του μέσου τετραγωνικού σφάλματος μπορεί να γίνει και με την χρήση της σχετικής αποτελεσματικότητας που αποτελεί το πηλίκο του μέσου τετραγωνικού σφάλματος του θ_1 και του μέσου τετραγωνικού σφάλματος του θ_2 :

$$RE = \frac{MSE \theta_1}{MSE \theta_2}$$

Το κριτήριο της ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος χρησιμοποιείται πολύ συχνά για την σύγκριση μεταξύ των άμεσων και των έμμεσων εκτιμητών. Ειδικά στις μελέτες που αναπτύσσουν έμμεσους εκτιμητές μικρών περιοχών (SAE) το μέσο τετραγωνικό αποτελεί βασικό κριτήριο αξιολόγησης-απόδοσης του εκτιμητή της μεταβλητής ενδιαφέροντος.

Συνέπεια (consistency)

Η συνέπεια αναφέρεται στην συμπεριφορά που έχει ένας εκτιμητής καθώς το μέγεθος του δείγματος αυξάνεται. Ένας συνεπής εκτιμητής αυξάνει την ακρίβεια των εκτιμήσεων όσο μεγαλώνει το δείγμα. Δηλαδή, οι τιμές που δίνει ο εκτιμητής για την παράμετρο του πληθυσμού θα πλησιάζουν όλο και περισσότερο την πραγματική τιμή της παραμέτρου. Παράδειγμα συνεπής εκτιμητή της μέσης τιμής του πληθυσμού αποτελεί ο αριθμητικός μέσος και η διάμεσος όταν ο πληθυσμός ακολουθεί την κανονική κατανομή.

ΚΕΦΑΛΑΙΟ 3^ο

ΤΕΧΝΙΚΕΣ ΕΚΤΙΜΗΣΗΣ ΜΙΚΡΟΠΕΡΙΟΧΩΝ

3.1 ΑΜΕΣΟΣ ΕΚΤΙΜΗΤΗΣ

Παρουσίαση

Ένας άμεσος εκτιμητής χρησιμοποιεί παρατηρήσεις της μεταβλητής ενδιαφέροντος, που προέρχονται από τις δειγματοληπτικές μονάδες της περιοχής ενδιαφέροντος, και μόνο. Ο άμεσος εκτιμητής προκύπτει στη βάση της δειγματοληψίας με πιθανότητα - πιθανοθεωρητικής δειγματοληψίας (Design-Based Estimators) και ικανοποιεί τις κλασικές ιδιότητες της αμεροληψίας, αποτελεσματικότητας και της συνέπειας.

Οι άμεσοι εκτιμητές χρησιμοποιούνται γενικά όταν το μέγεθος δείγματος για κάθε περιοχή είναι αρκετά μεγάλο για να δώσει ακριβείς εκτιμήσεις. Εντούτοις, δεδομένου ότι οι πηγές στοιχείων είναι συνήθως δειγματοληπτικές έρευνες, που έχουν σκοπό να δώσουν εθνικές και περιφερειακές στατιστικές εκτιμήσεις, τα μεγέθη δειγμάτων για τοπικές περιοχές είναι συνήθως πολύ μικρά. Αυτό έχει σαν αποτέλεσμα μεγάλη πιθανότητα μεγάλων διακυμάνσεων των εκτιμητών.

Στην βάση της δειγματοληπτικής θεωρίας και των δειγματοληπτικών σχεδίων προκύπτει το δείγμα για την εκτίμηση της μεταβλητής ενδιαφέροντος. Κάθε στοιχείο του δείγματος έχει μια συγκεκριμένη πιθανότητα να είναι στο δείγμα. Οι μικροπεριοχές υποεκπροσωπούνται στα δειγματοληπτικά σχέδια που αφορούν μεγάλες περιοχές (π.χ περιφέρειες) με αποτέλεσμα να εμφανίζονται μεγάλες διακυμάνσεις στην εκτίμηση της μεταβλητής ενδιαφέροντος y για την μικροπεριοχή. Για μικροπεριοχές, που τους αντιστοιχεί μικρό δείγμα, δημιουργούμε στρώματα με βάση ένα κοινό χαρακτηριστικό τους για την καλύτερη εκπροσώπηση των περιοχών ενδιαφέροντος στο δείγμα.

Άμεσοι εκτιμητές μέσης τιμής και αθροίσματος

Έστω i μια μικροπεριοχή του αρχικού πληθυσμού, N_i το μέγεθος του και Y_{ik} , $k=1,2,\dots,N_i$, οι τιμές του. Έστω n_i ο αριθμός των μονάδων του δείγματος που ανήκουν στην i μικροπεριοχή και y_{ik} , $k=1,2,\dots,n_i$, οι τιμές της υπό μελέτη μεταβλητής στο δείγμα. Στην πραγματικότητα μπορεί να γνωρίζουμε ή να μην γνωρίζουμε το πραγματικό μέγεθος N_i .

Ο αριθμητικός μέσος του δείγματος δίνεται από τον τύπο:

$$\bar{y}_i = \frac{\sum_{k=1}^{n_i} y_{ik}}{n_i}$$

και αποτελεί έναν αμερόληπτο εκτιμητή της μέσης τιμής του πληθυσμού, δηλαδή ισχύει ότι, $E(\bar{y}_i) = \bar{Y}_i$.

Ένας αμερόληπτος εκτιμητής για την διακύμανση του εκτιμητή της μέσης τιμής δίνεται από τον τύπο:

$$\text{var}(\bar{y}_i) = \frac{s_i^2}{n_i} \frac{N_i - n_i}{N_i} = \frac{s_i^2}{n_i} (1 - f_i)^6$$

Ο αμερόληπτος εκτιμητής για το ολικό ενός υποπληθυσμού δίνεται από τον τύπο:

$$Y_i = N_i \bar{y}_i = N_i \frac{\sum_{k=1}^{n_i} y_{ik}}{n_i}$$

Ένας αμερόληπτος εκτιμητής για την διακύμανση του εκτιμητή του αθροίσματος δίνεται από τον τύπο:

$$\text{var } Y_i = \frac{N_i^2 s_i^2}{n_i} (1 - f_i)$$

Για την εκτίμηση της μέσης τιμής και του αθροίσματος έχουμε ότι: s_i^2 αμερόληπτος εκτιμητής της διακύμανσης του πληθυσμού, f το δειγματοληπτικό

⁶ Όταν το δειγματοληπτικό σφάλμα είναι πολύ μικρό η διόρθωση του πεπερασμένου του πληθυσμού είναι κοντά στην μονάδα και παραλείπεται. Έτσι το μέγεθος του πληθυσμού δεν επηρεάζει την διακύμανση και κατά συνέπεια και το τυπικό σφάλμα του εκτιμητή.

κλάσμα, $f_i = \frac{n_i}{N_i}$, ενώ το $(1 - f_i) = \frac{N_i - n_i}{N_i}$ καλείται διόρθωση πεπερασμένου πληθυσμού.

Επιπρόσθετα, η τετραγωνική ρίζα του εκτιμητή δίνει εκτιμήσεις για το δειγματοληπτικό σφάλμα. Το δειγματοληπτικό αυτό σφάλμα είναι γνωστό με τον όρο τυπικό σφάλμα της μέσης τιμής και αποτελεί μέτρο του πόσο «καλή» είναι η εκτίμηση μας.

Τα παραπάνω ισχύουν όταν το N_i είναι γνωστό από κάποιο γνωστό κατάλογο για κάθε μικροπεριοχή. Όταν το μέγεθος των μικροπεριοχών που μελετάμε είναι άγνωστο χρησιμοποιούμε την ποσότητα n/N στην θέση της ποσότητας n_i/N_i .

N_i άγνωστο	
Εκτιμητής απλής μεγέθυνσης Simple expansion estimator	$Y_i = N \frac{\sum_{k=1}^{n_i} y_{ik}}{n}$
N_i γνωστό	
Μεταστρωματοποιημένος εκτιμητής Post-stratified estimator	$Y_i = N_i \bar{y}_i = N_i \frac{\sum_{k=1}^{n_i} y_{ik}}{n_i}$

Χαρακτηριστικά

Οι άμεσοι εκτιμητές χρησιμοποιούνται κυρίως για εκτιμήσεις με μεγάλα δείγματα όπως οι εκτιμήσεις για εθνικές στατιστικές. Εντούτοις σε πολλές μελέτες μικροπεριοχών ενδιαφέροντος χρησιμοποιείται σαν συγκριτικός δείκτης σε σχέση με τους έμμεσους εκτιμητές. Η διακύμανση του εκτιμητή για εκτιμήσεις μικροπεριοχών είναι μεγαλύτερη όσο μικρότερο το δείγμα που αντιστοιχεί στην μικροπεριοχή. Ο άμεσος εκτιμητής δεν μπορεί να εκτιμηθεί εάν στο δείγμα δεν υπάρχει καμία τιμή για την μεταβλητή ενδιαφέροντος της μικροπεριοχής.

3.2 ΈΜΜΕΣΟΙ ΕΚΤΙΜΗΤΕΣ

Η ανάγκη για εκτιμητές μικροπεριοχών με μικρότερη διακύμανση και άρα μικρότερα τυπικά σφάλματα σε σχέση με αυτά των άμεσων εκτιμητών, καθώς και η ανάγκη για εκτιμήσεις μικροπεριοχών όπου δεν τους αντιστοιχεί καμία δειγματοληπτική μονάδα κάνει την χρήση έμμεσων εκτιμητών απαραίτητη. Με την χρήση των έμμεσων εκτιμητών αυξάνει το «αποτελεσματικό⁷» μέγεθος του δείγματος, με την έννοια ότι το δείγμα γίνεται πιο αντιπροσωπευτικό για την εξαγωγή συμπερασμάτων. Αυτό επιτυγχάνεται είτε με την χρήση βοηθητικών μεταβλητών που έχουν υψηλή συσχέτιση με την μεταβλητή ενδιαφέροντος και οι οποίες είναι γνωστές για την μικροπεριοχή, είτε με την χρήση εκτιμήσεων που προϋπάρχουν και αφορούν έμμεσα την περιοχή ενδιαφέροντος. Τέτοιοι έμμεσοι εκτιμητές είναι ο συνθετικός (synthetic) και ο σύνθετος (composite) εκτιμητής.

3.2.1 ΣΥΝΘΕΤΙΚΟΣ ΕΚΤΙΜΗΤΗΣ

Παρουσίαση

Ο συνθετικός εκτιμητής περιγράφεται με τον καλύτερο τρόπο από τον M.E.Gonzales (1973). Ένας αμερόληπτος εκτιμητής (αμερόληπτος άμεσος εκτιμητής) λαμβάνεται από ένα δείγμα μιας περιφέρειας, που περιλαμβάνει μερικές μικροπεριφέρειες. Αυτός ο εκτιμητής χρησιμοποιείται για την παραγωγή εκτιμήσεων για μια μικροπεριοχή στη βάση της υπόθεσης ότι αυτή έχει τα ίδια χαρακτηριστικά με την περιφέρεια στην οποία ανήκει.

Μια άλλη προσέγγιση για την περιγραφή του συνθετικού εκτιμητή είναι αυτή του Simmons (1977) όπου δικαιολογεί τον τίτλο συνθετικός στη βάση της από κοινού χρησιμοποίησης περιγραφικών δεδομένων σε τοπικό επίπεδο και μέσω των τιμών σε περιφερειακό ή εθνικό επίπεδο.

⁷ Αποτελεσματικό δείγμα ορίζεται το δείγμα που ελαχιστοποιεί το συνολικό σφάλμα εκτίμησης.

Αυτή η μέθοδος χαρακτηρίζεται μέθοδος δανειζόμενης πληροφορίας «borrow strength» από μια σχετική περιοχή έτσι ώστε να αυξηθεί το αποτελεσματικό μέγεθος του δείγματος και η ακρίβεια των εκτιμήσεων (Holt, Smith & Tomberlin, 1979).

Ο όρος «συνθετικός εκτιμητής» χρησιμοποιήθηκε για πρώτη φορά το 1968, από το Αμερικανικό Εθνικό Κέντρο για Στατιστικές Υγείας (U.S National Center for Health Statistics) για τον υπολογισμό των ατόμων με μακροχρόνιες και βραχυχρόνιες αναπηρίες και άλλων χαρακτηριστικών υγείας (Synthetic State Estimates of Disability) βασιζόμενο στην Εθνική Έρευνα για την Υγεία (National Health Interview Survey). Από τότε υπήρξε ιδιαίτερο ενδιαφέρον για εκτιμήσεις μικροπεριοχών πράγμα που οδήγησε σε μεγάλη ποικιλία προσεγγίσεων για την παραγωγή τέτοιων εκτιμήσεων. Συγκεκριμένα, οι συνθετικές εκτιμήσεις έχουν χρησιμοποιηθεί με μεγάλη ποικιλία για την παραγωγή εκτιμήσεων μικροπεριοχών με εφαρμογή στην οικονομική, κοινωνική και πολιτική πραγματικότητα..

Η μέθοδος υποθέτει την διαθεσιμότητα εκτιμήσεων από δειγματοληπτικές ή απογραφικές έρευνες που θα μπορούν να χρησιμοποιηθούν για την άντληση της «βοηθητικής» εκτίμησης για την μετέπειτα εκτίμηση του συνθετικού εκτιμητή μιας μικρής περιοχής ενδιαφέροντος. Στην πραγματικότητα οι Εθνικές Έρευνες αναφέρονται σε εκτιμήσεις μεγάλων ομάδων του πληθυσμού όπως είναι μια μεγάλη γεωγραφική περιοχή (π.χ κράτος, περιφέρεια), μια δημογραφική κατηγορία (π.χ φύλο και ηλικιακή ομάδα) ή ακόμα και ολόκληρος τομέας παραγωγής (π.χ βιομηχανία). Στη βάση αυτών των εκτιμήσεων για τις μεγάλες ομάδες του πληθυσμού ο συνθετικός εκτιμητής δίνει την δυνατότητα για εκτιμήσεις υποομάδων του πληθυσμού, που ανήκουν σε μια μικροπεριοχή και για τις οποίες δεν υπάρχουν εκτιμήσεις. Η εκτίμηση της δειγματοληπτικής ή απογραφικής έρευνας αποτελεί το πρώτο συνθετικό του εκτιμητή. Το δεύτερο συνθετικό του εκτιμητή είναι ένα σταθμιστής που τις περισσότερες φορές βασίζεται στις πληθυσμιακές απογραφές που είναι διαθέσιμες εκτιμήσεις για τις μικρές περιοχές ενδιαφέροντος..

Γενικά ο συνθετικός εκτιμητής δίνεται από τον τύπο :

$$Y_{ih} = \sum_h p_{ih} Y_h$$

όπου p_{ih} : ο σταθμιστής που αναφέρεται στην υποομάδα h για μικροπεριοχή i .

Y_h : ο εκτιμητής της μεταβλητής ενδιαφέροντος y για την υποομάδα h της μεγάλης περιοχής, που θεωρούμε ότι έχει όμοια χαρακτηριστικά με την περιοχή ενδιαφέροντος.

Η μέθοδος συνθετικής εκτίμησης που πρότειναν οι Gonzales και Hoza⁸ (1978)

χρησιμοποιεί ως σταθμιστή p_{ih} εκτιμήσεις που αφορούσαν την μικρή περιοχή (συνήθως εκτιμήσεις από απογραφικά δεδομένα πληθυσμού). Στη βάση της υπόθεσης ότι υπάρχουν αξιόπιστες άμεσες εκτιμήσεις Y_h της μεταβλητής ενδιαφέροντος για κάθε υποομάδα (subgroup) h της μεγάλης περιοχής, παρήγαγαν εκτιμήσεις για την μικρή περιοχή. Η μικρή περιοχή αναφερόταν σε μια μικρή γεωγραφική περιοχή που ήταν μια από τις περιοχές που περιλαμβάνονται στην μεγάλη περιοχή.

Αν ο σταθμιστής p_{ih} είναι ένα ποσοστό συμμετοχής του πληθυσμού της κάθε υποομάδας h στον συνολικό πληθυσμό της μικρής περιοχής i συνεπάγεται ότι:

$$p_{ih} = 1 \quad \text{και} \quad p_{ih} = \frac{N_{ih}}{N_i}$$

Η συνολική εκτίμηση για την μεταβλητή ενδιαφέροντος y για την υποομάδα h στην μεγάλη περιοχή είναι, Y_h , ενώ ο μέσος της μεταβλητής ενδιαφέροντος y για την υποομάδα h στην μεγάλη περιοχή είναι ο λόγος των εκτιμήσεων της μεταβλητής ενδιαφέροντος Y_h για την υποομάδα h προς τον αριθμό του πληθυσμού της υποομάδας N_h .

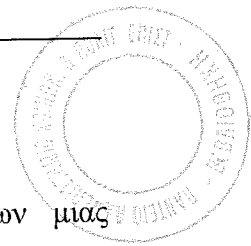
$$\bar{y}_h = \frac{Y_h}{N_h}$$

Οπότε η εκτίμηση του μέσου όρου της μεταβλητής ενδιαφέροντος της συγκεκριμένης υποομάδας για την μικρή περιοχή θα δίνεται από την σχέση:

$$\bar{y}_{ih} = \sum_{h=1} \bar{p}_{ih} \bar{y}_h$$

ενώ ο εκτιμητής του συνόλου της δίνεται από

⁸ Gonzales, M.E. and Hoza, C. (1978). Small-Area Estimation with Application to Unemployment and Housing Estimates. *Journal of the American Statistical Association*, 73, 7-15.



$$Y_i = \sum_h p_{ih} Y_h$$

Στην πραγματικότητα, εφόσον υπάρχει διαθεσιμότητα δεδομένων μιας βοηθητικής μεταβλητής (X), ο σταθμιστής p_{ih} μπορεί να είναι μια αναλογία αυτής της βοηθητικής μεταβλητής που έχει υψηλή συσχέτιση με την μεταβλητή ενδιαφέροντος, δηλαδή

$$p_{ih} = \frac{X_{ih}}{X_i}$$

Χαρακτηριστικό παράδειγμα χρήσης τέτοιας μεταβλητής αποτελεί το παράδειγμα των Gonzales και Hoza που χρησιμοποίησαν τις αναλογίες του εργατικού δυναμικού ανά ηλικιακή ομάδα ως σταθμιστή για την εκτίμηση του ποσοστού ανεργίας ανά ηλικιακή ομάδα σε τοπικό επίπεδο.

Οι Rao και Choudry⁹ (1995) πρότειναν την χρήση του λόγου συνθετικού εκτιμητή (ratio synthetic estimator) που αποτελούσε μια τροποποίηση της μεθόδου των Gonzales και Hoza (1978), καθώς και στις δύο μεθόδους χρησιμοποιήθηκε μια βοηθητική μεταβλητή που είναι εκ των προτέρων γνωστή. Ο λόγο αυτός βασίστηκε στην υπόθεση ότι :

$$R_i = \frac{Y_i}{X_i}$$

όπου Y_i είναι η τιμή της μεταβλητής ενδιαφέροντος στην μικρή περιοχή και X_i η τιμή της συσχετιζόμενης βοηθητικής μεταβλητής για την μικρή περιοχή. Επειδή ο συνθετικός εκτιμητής βασίζεται στην υπόθεση της ομοιογένειας μεταξύ των σχετιζόμενων περιοχών θα ισχύει:

$$R_i = R = \frac{Y}{X} \text{ και } Y_i = R * X_i$$

όπου R, X, Y είναι τιμές που αναφέρονται στον πληθυσμό της μεγάλης περιοχής. Η τιμή του R εκτιμήθηκε από την σχέση:

$$R_i = \frac{\bar{y}}{\bar{x}}$$

⁹ Rao, J.N.K. and Choudry, G. H. (1995). Small Area Estimation: Overview and Empirical Study, in B.G. Cox et al (editors), *Business Survey Methods*, John Wiley & Sons, New York, 527-542.

όπου $\bar{y}$ και $\bar{x}$ είναι μέσοι όροι των μεταβλητών στον πληθυσμό της μεγάλης περιοχής για την οποία υπάρχουν διαθέσιμα στοιχεία. Τέλος οι Rao και Choundry κατέληξαν στον εκτιμητή της μεταβλητής ενδιαφέροντος για γνωστές τιμές της μεταβλητής ενδιαφέροντος στην μικρή περιοχή:

$$Y_i = R_i * X_i$$

Έτσι εκτός από τις αναλογίες και άλλες εκτιμήσεις τις μεταβλητής ενδιαφέροντος που παίρνουμε για τον υπολογισμό του συνθετικού εκτιμητή της μικρής περιοχής ενδιαφέροντος, είναι δυνατόν να χρησιμοποιηθεί και μια βοηθητική μεταβλητή που συσχετίζεται σε μεγάλο βαθμό με την μεταβλητή ενδιαφέροντος. Γενικά, η οποιαδήποτε βοηθητική μεταβλητή που είναι διαθέσιμη από μια αξιόπιστη πηγή μπορεί να αποτελέσει τον σταθμιστή για τον συνθετικό εκτιμητή.

Οι έμμεσοι εκτιμητές μπορούν να ταξινομηθούν σε τρεις τύπους: έμμεσοι περιοχής, έμμεσοι χρόνου, και έμμεσοι περιοχής/χρόνου, δηλαδή ανάλογα με το από πού προέρχεται η δανειζόμενη δύναμη-πληροφορία. Για παράδειγμα αν τα δεδομένα της βοηθητικής μεταβλητής προέρχονται από διαφορετική χρονική περίοδο από την περίοδο ενδιαφέροντος τότε έχουμε έναν έμμεσο εκτιμητή χρόνου.

Χαρακτηριστικά

Ένα μεγάλο πλεονέκτημα του συνθετικού εκτιμητή είναι ο απλός υπολογισμός του. Παράλληλα, έχει χαρακτηριστεί και ως «οικονομικός» λόγω των περιορισμένων πόρων που απαιτεί για την εκτίμηση του. Η διακύμανση του συνθετικού εκτιμητή είναι μικρότερη σε σχέση με αυτή του άμεσου εκτιμητή, καθώς χρησιμοποιεί βοηθητικές μεταβλητές που αυξάνουν το αποτελεσματικό μέγεθος του δείγματος και δίνουν εκτιμήσεις ακόμα και για περιοχές όπου δεν έχουμε παρατηρήσεις από το δείγμα.

Ο συνθετικός εκτιμητής δεν πληροί την βασική ιδιότητα ενός BLUE εκτιμητή (άριστου, γραμμικός, αμερόληπτος) σχετικά με την αμεροληψία του, καθώς είναι μεροληπτικός για τους εξής λόγους. Πρώτον γιατί η υπόθεση για την ομοιογένεια των ποσοστών ή των αναλογιών δύσκολα ικανοποιείται, καθώς οι εκτιμήσεις για την μεγάλη περιοχή (για μια συγκεκριμένη μεταβλητή ενδιαφέροντος) είναι πολύ πιθανό

να διαφέρουν από εκείνες της μικρής περιοχής που ανήκει σε αυτή. Ο βαθμός ομοιογένειας άλλωστε αναγνωρίζεται ως πρωταρχικού ενδιαφέροντος για την παραγωγή συνθετικών εκτιμητών. Με άλλα λόγια η βασική υπόθεση για την χρησιμοποίηση ενός συνθετικού εκτιμητή, που στην βιβλιογραφία εμφανίζεται ως «model assumption», παραβιάζεται. Δεύτερον, η μεροληψία του εκτιμητή οφείλεται στις πιθανές αλλαγές της δομής του πληθυσμού σε σχέση με την περίοδο της χρησιμοποιούμενης απογραφής (συνήθως τις πιο πρόσφατης). Γενικότερα ο συνθετικός εκτιμητής «έλκεται» από τον εκτιμητή της μεγάλης περιοχής δημιουργώντας αυξημένη πιθανότητα μεροληψίας του εκτιμητή για την μεταβλητή ενδιαφέροντος στην μικρή περιοχή.

Ένας μεροληπτικός εκτιμητής έχει υψηλό μέσο τετραγωνικό σφάλμα, καθώς το μέσο τετραγωνικό σφάλμα είναι το άθροισμα της διακύμανσης (variation) του εκτιμητή και της μεροληψίας (bias). Σε κάθε περίπτωση η διακύμανση ενός συνθετικού είναι μικρότερη σε σχέση με έναν άμεσο εκτιμητή, όμως αυτό το πλεονέκτημα αντισταθμίζεται με το μειονέκτημα της ανισορροπίας-αστάθειας του εκτιμητή λόγω της μεροληψίας.

3.2.2 ΣΥΝΘΕΤΟΣ ΕΚΤΙΜΗΤΗΣ

Παρουσίαση

Ανάλογα με το αν το δείγμα της μικρής περιοχής είναι σχετικά μικρό (μεγάλο), προτιμάται ο συνθετικός (άμεσος) εκτιμητής. Αυτό γιατί όσο μικρότερο είναι το δείγμα τόσο εξασθενίζει η εκτιμητική ικανότητα του αμερόληπτου άμεσου εκτιμητή που γίνεται μη αποτελεσματικός (δεν έχει την μικρότερη δυνατή διακύμανση). Αντίθετα, όσο μεγαλύτερο είναι το δείγμα για την μικρή περιοχή τόσο εξασθενίζει η εκτιμητική ικανότητα του μεροληπτικού συνθετικού εκτιμητή σε σχέση με τον αμερόληπτο άμεσο εκτιμητή. Σε αυτό το συμπέρασμα κατέληξαν οι Gonzalez and Waksberg (1973) και Schaible, Brock, Casady and Schnack (1977) συγκρίνοντας τα σφάλματα από τις συνθετικές και άμεσες εκτιμήσεις για τις Μητροπολιτικές περιοχές και τους νομούς των Ηνωμένων Πολιτειών. Στη βάση των αποτελεσμάτων αυτών προτάθηκε ο συνδυασμός των δύο εκτιμητών που οδηγεί σε ακριβέστερες εκτιμήσεις για την μικρή περιοχή. Ο συνδυασμός αυτός είναι γνωστός ως σύνθετος

εκτιμητής (ο σύνθετος εκτιμητής προτάθηκε για πρώτη φορά από τον Royal (1973) αλλά μελετήθηκε από τον Schaible (1978).

Με σύνθετο εκτιμητή εξισορροπείτε η εν δυνάμει αμεροληψία του συνθετικού εκτιμητή με την αστάθεια του άμεσου εκτιμητή. Ο σύνθετος εκτιμητής αποτελεί ένα σταθμισμένο άθροισμα του άμεσου και του συνθετικού εκτιμητή. Γενικά ο σύνθετος εκτιμητής είναι ένας σταθμισμένος μέσος ενός άμεσου και ενός έμμεσου εκτιμητή (Suciu¹⁰,2001), και συγκεκριμένα του άμεσου και του συνθετικού εκτιμητή.

Γενικά ο σύνθετος εκτιμητής δίνεται από τον τύπο:

$$Y_i^C = w_i * Y_i^D + (1 - w_i) * Y_i^S$$

όπου Y_i^C είναι ο σύνθετος εκτιμητής, Y_i^D ο άμεσος εκτιμητής, Y_i^S είναι ο συνθετικός εκτιμητής και w_i είναι ένας σταθμιστής για την μικρή περιοχή i .

$$0 \leq w_i \leq 1$$

Εκτιμητές		w_i	$(1 - w_i)$
Άμεσος	Y_i^D	1	0
Συνθετικός	Y_i^S	0	1
Σύνθετος	$Y_i^C = w_i * Y_i^D + (1 - w_i) * Y_i^S$	w_i	$(1 - w_i)$

Ο Schaible¹¹ (1979), απέδειξε ότι με την επιλογή του κατάλληλου σταθμιστή ο σύνθετος εκτιμητής έχει μικρότερο μέσο τετραγωνικό σφάλμα σε σύγκριση με τους δύο εκτιμητές που τον αποτελούν, τον άμεσο και τον έμμεσο. Ο ιδανικός εκτιμητής προκύπτει από την ελαχιστοποίηση της συνάρτησης του μέσου τετραγωνικού σφάλματος (MSE) του σύνθετου εκτιμητή,

¹⁰ Suciu, G. et al (2001). Uninsured Estimates by County: A Review of Options and Issues, Ohio Department of Health.

¹¹ Schaible, W.L. (1978). Choosing Weights for Composite Estimators for Small Area Statistics, *Proceedings of the Survey Research Methods Section*, 741-746.

$$MSE(Y_i^c) = w_i^2 * MSE(Y_i^s) + (1 - w_i)^2 * MSE(Y_i^s) - 2w_i * (1 - w_i) * Cov(Y_i^s, Y_i^s)$$

ως προς τον σταθμιστή. Ο σταθμιστής που προκύπτει είναι της μορφής:

$$w_i = \frac{MSE(Y_i^s) - Cov(Y_i^s, Y_i^s)}{MSE(Y_i^s) + MSE(Y_i^s) - 2Cov(Y_i^s, Y_i^s)}$$

Στην πραγματικότητα η συνδιακύμανση των δύο εκτιμητών είναι δύσκολα να εκτιμηθεί. Αν η συνδιακύμανση σχετίζεται πολύ λίγο με το μέσο τετραγωνικό σφάλμα του συνθετικού εκτιμητή τότε ο σταθμιστής ο οποίος ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα του συνθετικού εκτιμητή δίνεται από τον τύπο:

$$w_i = \frac{1}{1 + R}, \quad \text{όπου } R = \frac{MSE(Y_i^s)}{MSE(Y_i^s)}$$

Όμοια οι Ghosh and Rao (1994) κατέληξαν στον ίδιο σταθμιστή που προκύπτει από την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE) του σύνθετου εκτιμητή ως προς τον σταθμιστή, κάνοντας την υπόθεση ότι η συνδιακύμανση άμεσου και συνθετικού εκτιμητή είναι μηδέν

$Cov(Y_i^s, Y_i^s) = 0$. Δηλαδή υπέθεσαν ότι και οι δύο εκτιμητές ήταν αμερόληπτοι και ανεξάρτητοι, οπότε και η συνδιακύμανση τους θα είναι ίση με μηδέν

Από την ελαχιστοποίηση της συνάρτησης το ελάχιστο σημείο δίνεται από τον σταθμιστή :

$$w_i = \frac{MSE(Y_i^s)}{MSE(Y_i^s) + V(Y_i^s)}, \quad \text{όπου } MSE(Y_i^s) = V(Y_i^s)$$

δηλαδή από τον λόγο του μέσου τετραγωνικού σφάλματος του συνθετικού εκτιμητή προς το άθροισμα του μέσου τετραγωνικού σφάλματος του συνθετικού εκτιμητή και της διακύμανσης του άμεσου εκτιμητή για την μικρή περιοχή i . Οι σταθμιστές που προκύπτουν στη βάση των παραπάνω υποθέσεων είναι δυνατόν να είναι ιδιαίτερα

ασταθείς λόγω της συμμετοχής του εκτιμητή του μέσου τετραγωνικού σφάλματος του συνθετικού εκτιμητή που είναι πολύ ασταθής.

Ο Schaible (1978) και οι Purcell and Kish (1979) για να προσπεράσουν την αδυναμία αυτή των σταθμιστών πρότειναν την χρήση ενός κοινού σταθμιστή:

$$w = 1 - \frac{\sum_i V_i Y_i}{\sum_i (Y_i^S - Y_i)^2}, \text{ του οποίου η εκτίμηση είναι } w = 1 - \frac{\sum_i v_i Y_i}{\sum_i \frac{s_i^2}{Y_i - Y_i}}$$

Στην περίπτωση μάλιστα που οι διακυμάνσεις των εκτιμήσεων της μεταβλητής ενδιαφέροντος των μικρών περιοχών, $v_i Y_i$, είναι κατά προσέγγιση ίσες (και μόνο τότε), ο εκτιμητής της διακύμανσης μπορεί να αντικατασταθεί με τον μέσο του:

$$\bar{v} = \frac{\sum_i v_i Y_i}{m} \text{ και η εκτίμηση του κοινού σταθμιστή } w = 1 - \frac{\frac{m \bar{v}}{s}}{\sum_i \frac{1}{Y_i - Y_i}}$$

Οι Drew, Singh and Choudry (1982) πρότειναν για τον υπολογισμό του σύνθετου εκτιμητή έναν απλό σταθμιστή που εξαρτάται μόνο από το μέγεθος (πληθυσμός ή βοηθητική μεταβλητή) της μικρής περιοχής (sample size dependent estimator, SDD). Ο εκτιμητής αυτός

$$w_i = \begin{cases} 1, \text{ αν } N_i \geq \delta N_i \\ \frac{N_i}{\delta N_i}, \text{ αν } N_i < \delta N_i \end{cases}$$

όπου N_i είναι ο άμεσος εκτιμητής διόγκωσης (direct expansion estimator) του γνωστού μεγέθους του πληθυσμού της μικρής περιοχής N_i και το δ είναι μια σταθερά που επιλέγεται υποκειμενικά για την συμβολή του συνθετικού εκτιμητή (μια γενική επιλογή για το δ είναι $\delta=1$). Στην περίπτωση της απλής τυχαίας δειγματοληψίας n στοιχείων από ένα πληθυσμό N στοιχείων η εκτίμηση του δίνεται από τον τύπο :

$$N_i = N * \frac{n_i}{n}$$

όπου n_i είναι το μέγεθος του δείγματος για την μικρή περιοχή i . Ο τύπος δείχνει ότι ο άμεσος εκτιμητής αυξάνει όσο αυξάνει το μέγεθος του δείγματος για την μικρή περιοχή.

Χαρακτηριστικά

Η σύνθετη εκτίμηση θα είναι μεγαλύτερη από την άμεση και μικρότερη της συγκεντρωτικής ή της έμμεσης. Αυτό γιατί η σύνθεση γίνεται στην βάση ενός σταθμιστή που όσο πιο μεγάλος τόσο πιο μεγάλη η επίδραση της άμεσης και τόσο μικρότερη η επίδραση της έμμεσης στην σύνθετη εκτίμηση. Η εκτίμηση αυτή θα τείνει να ενισχύει τις περιοχές με μεγάλα δείγματα, καθώς αυτοί δίνουν πληροφορία σχετικά με τον μέσο της κατανομής και ο οποίος επηρεάζει την εκτίμηση της περιοχής με το μικρό δείγμα.

Τα βασικά χαρακτηριστικά του σύνθετου εκτιμητή έχουν να κάνουν με την ακρίβεια αλλά και την δυσκολία υπολογισμού του διαστήματος εμπιστοσύνης ενός τέτοιου εκτιμητή. Όσον αφορά την ακρίβεια, ο σύνθετος εκτιμητής της μεταβλητής ενδιαφέροντος θα έχει μικρότερη διακύμανση, όμως δεν θα είναι αμερόληπτος (όπως ο άμεσος) καθώς δανείζεται πληροφορία από άλλες περιοχές (με την χρήση του συνθετικού εκτιμητή). Παρόλα αυτά το κέρδος από την μείωση της διακύμανσης αντισταθμίζει την απώλεια αμεροληψίας, με αποτέλεσμα το μέσο τετραγωνικό σφάλμα (MSE) του σύνθετου εκτιμητή να είναι μικρότερο του άμεσου εκτιμητή.

Όσον αφορά το διάστημα εμπιστοσύνης, η κλασική προσέγγιση της λήψης του διαστήματος εμπιστοσύνης από μια σημειακή εκτίμηση απαιτεί αυτή η εκτίμηση να είναι αμερόληπτη ή σχεδόν αμερόληπτη. Στην περίπτωση του σύνθετου εκτιμητή κάτι τέτοιο δεν συμβαίνει και μάλιστα το μέγεθος της μεροληψίας είναι άγνωστο και δεν μπορεί να συνυπολογιστεί στον υπολογισμό του διαστήματος εμπιστοσύνης.

ΚΕΦΑΛΑΙΟ 4^ο

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ

Ο όρος «συνθετικός εκτιμητής» χρησιμοποιήθηκε για πρώτη φορά (1968), από το Αμερικανικό Εθνικό Κέντρο για Στατιστικές Υγείας¹² για την εκτίμηση των συνολικών ατόμων με μακροχρόνιες και βραχυχρόνιες αναπηρίες και άλλων χαρακτηριστικών υγείας σε κρατικό επίπεδο, βασιζόμενο στην Εθνική Έρευνα για την Υγεία. Το δείγμα των περισσότερων πόλεων σε αυτή την περίπτωση ήταν πολύ μικρό για να παραχθούν αξιόπιστες άμεσες εκτιμήσεις, οπότε και χρησιμοποιήθηκε ο σύνθετος εκτιμητής, αποτελώντας μια εν δυνάμει λύση εκτιμητή, όπου δεν υπάρχουν πραγματικά τοπικά δεδομένα και δεν μπορούν να παραχθούν εκτιμήσεις.

Μια πρώτη συνολική βιβλιογραφική επισκόπηση στις τεχνικές εκτίμησης μικροπεριοχών είναι αυτή του J.Steinberg (1979) σε συνεργασία με το Εθνικό ίδρυμα εξαρτώμενων ατόμων¹³. Αποτέλεσμα αυτής της συνεργασίας ήταν η δημιουργία μιας σειράς μονογραφιών με αναλυτικό ενδιαφέρον τους έμμεσους εκτιμητές. Το πρώτο μέρος αποτελεί μια αναλυτική παρουσίαση του συνθετικού εκτιμητή από τον P.Levy¹⁴, ενώ στο δεύτερο μέρος γίνεται η παρουσίαση και η τεκμηρίωση του σύνθετου εκτιμητή από τον W.Shaible¹⁵. Ειδικότερα, ο P.Levy αναλύει τις διαφορετικές προσεγγίσεις εκτιμήσεων μικροπεριοχών με έμφαση στα πλεονεκτήματα και τις αδυναμίες του συνθετικού εκτιμητή, ενώ ο W.Shaible αναλύει την συλλογιστική του σύνθετου εκτιμητή και αναπτύσσει μέσα από εμπειρική μελέτη τις μεθόδους εκτιμήσεων των σταθμιστών για τον σύνθετο εκτιμητή.

Μια πιο πρόσφατη συνολική αναφορά στους έμμεσους εκτιμητές αποτέλεσε η έκθεση της ομοσπονδιακής υποεπιτροπής (Federal Subcommittee)¹⁶ (1991), η οποία

¹² National Center for Health Statistics (1968), Synthetic State Estimates of Disability. (PHS Publication No.1759), Washington, D.C.:U.S. Government Printing Office.

¹³ National Institute on Drug Abuse (1979). Synthetic Estimates for Small Areas. (NIDA Research Monograph 24), Washington, D.C.: U.S. Government Printing Office.

¹⁴ Levy, P.S. (1979). Small-area estimation- Synthetic and other procedures, 1968-1978, Proceedings of the Workshop on the Synthetic Estimates for Small Areas, National Institute of Drug Abuse (NIDA), Research Monograph 24, U.S. Government Printing Office, Washington, D.C.

¹⁵ Shaible, W.L. (1979). A composite estimator for small area statistics, in Synthetic Estimates for Small Areas. (NIDA Research Monograph 24), Washington, D.C.: U.S. Government Printing Office.

¹⁶ Statistical Working Paper: Indirect Estimators in the U.S Federal Program, (1993), Office of Management and Budget

δημιουργήθηκε από την Ομοσπονδιακή Επιτροπή Στατιστικής Μεθοδολογίας (Federal Committee on Statistical Methodology) για την εμπειρική μελέτη των έμμεσων εκτιμητών σε πραγματικά δεδομένα. Στην έκθεση περιγράφονται οκτώ προγράμματα¹⁷ που εφάρμοσαν τους έμμεσους εκτιμητές. Αυτά τα προγράμματα ανταποκρίνονται στις ανάγκες κρατικών οργανισμών για δεδομένα μικροπεριοχών.

Σε μερικές περιπτώσεις, ένα πρόγραμμα παράγει εκτιμήσεις για μια μεταβλητή, ενώ σε άλλες οι εκτιμήσεις παράγονται για πολυάριθμες μεταβλητές. Τα κράτη και οι νομοί είναι οι μόνες γεωγραφικές περιοχές για τις οποίες παράγονται οι έμμεσες εκτιμήσεις. Τέσσερα από τα προγράμματα παράγουν εκτιμήσεις για τα κράτη, τρία για τους νομούς, και ένα και για τα κράτη και για τους νομούς. Οι έμμεσοι εκτιμητές που χρησιμοποιούνται για να παραγάγουν τις εκτιμήσεις είναι ο συνθετικός και ο σύνθετος εκτιμητής, που «δανείζονται δύναμη» πέρα από τις περιοχές ενδιαφέροντος. Οι διαδικασίες εκτίμησης για έξι από τα προγράμματα είναι βασισμένες στα στοιχεία δειγματοληπτικές έρευνες, ενώ δεν υπάρχει καμία δειγματοληψία για τις άλλες δύο, με την έννοια ότι χρησιμοποιούνται εναλλακτικές μέθοδοι εύρεσης δεδομένων για την εκτίμηση αυτών των μικροπεριοχών ενδιαφέροντος.

Την τελευταία πενταετία, η ζήτηση εκτιμήσεων μικροπεριοχών από χώρες τις Κεντρικής και Ανατολικής Ευρώπης συμπεριλαμβανομένης και της πρώην Σοβιετικής Ένωσης που βρίσκονται στο στάδιο της μετάβασης οδήγησε σε μια σειρά αρθρογραφίας¹⁸ για εκτιμήσεις μικροπεριοχών, που αφορούν τις χώρες αυτές. Ειδικότερα τα άρθρα της σειράς των G.C Tikkiwal & A.Ghiya¹⁹ αποτελούν μια περίπτωση σύγκρισης άμεσου και έμμεσων εκτιμητών (σύνθετος, συνθετικός). Ο σύνθετος εκτιμητής χρησιμοποιείται για την εκτίμηση της καλλιεργήσιμης έκτασης, ενώ εξετάζεται και η σχετική του αποτελεσματικότητα σε σχέση με τον άμεσο και

¹⁷ Τα προγράμματα αναφέρονταν σε α) χαρακτηριστικά υγείας νηπίων (National Center for Health Statistics (NCHS)), β) προσωπικό εισόδημα, ετήσιο εισόδημα, και ακαθάριστο προϊόν (Bureau of Economic Analysis), γ) πληθυσμό σε επίπεδο νομού (Bureau of the Census (BOC)), δ) απασχόληση και ανεργία σε κρατικό επίπεδο (Bureau of Labor Statistics), ε) καλλιεργήσιμη επιφάνεια σε βαμβάκι, ρύζι, και σόγια (National Agricultural Statistics Service (NASS)); στ) παραγωγή, συγκομιδή ζωικού κεφαλαίου (NASS), ε) χρησιμοποίηση νοσοκομείων, επισκέψεις σε παθολόγους και οδοντιάτρους (NCHS); και ζ) μέσο εισόδημα τετραμελών οικογενειών (BOC).

¹⁸ Kordos, J. (Editor), (2000-2004), *Statistics in transition*, Journal of the Polish Statistical Association
¹⁹ G.C Tikkiwal & A.Ghiya, (2004), "A generalized class of composite estimators with application for small domains", *Statistics in Transition*, Journal of the Polish Statistical Association, Vol.6, No.5, p 697

τον συνθετικό εκτιμητή. Στο επίπεδο μικροπεριοχής που εξετάζεται προτείνεται ο σύνθετος εκτιμητής, καθώς δίνει μικρότερη απόλυτη σχετική μεροληψία (Absolute Relative Bias).

Σε επίπεδο βιβλιογραφίας μια σύγχρονη και συνολική παρουσίαση του αντικειμένου των εκτιμήσεων μικροπεριοχών είναι αυτή του J.N.K Rao²⁰. Στο βιβλίο του εκτός από την κλασική διάκριση μεταξύ των έμμεσων και των άμεσων εκτιμητών, παρουσιάζεται ο μπεϋζιανός εμπειρικός και ιεραρχικός εκτιμητής, που αποτελούν ένα μηχανισμό εκτίμησης που χρησιμοποιεί αυτόματα την δανειζόμενη πληροφορία από σχετικές βοηθητικές μεταβλητές και παράγει με ευκολία ακριβείς εκτιμήσεις και διαστήματα πιθανοτήτων χωρίς επιπλέον κόστος. Αυτό γίνεται με την χρησιμοποίηση των εκ των προτέρων γνώσεων, που έχουμε για την κατανομή της μεταβλητής ενδιαφέροντος, στην παραγωγή της εκ των υστέρων κατανομής της μεταβλητής αυτής. Η εκ των προτέρων γνώση χρησιμοποιείται σε υποδείγματα στην βάση βασικών υποθέσεων από τα οποία λαμβάνεται η μετασχηματισμένη εκ των υστέρων εκτίμηση.

Το ενδιαφέρον για την μελέτη εκτιμήσεων μικροπεριοχών στη βάση εμπειρικών εφαρμογών οδήγησε τους Platek. R., R, J.N.K., Sarndal, C.E., and Singh, M.P. (Editors)²¹ στην έκδοση ενός βιβλίου, το οποίο περιείχε κείμενα εμπειρικών εφαρμογών εκτιμήσεων με τις τεχνικές των έμμεσων εκτιμητών. Τα κείμενα αυτά ήταν το αποτέλεσμα ενός διεθνούς συμποσίου με θέμα εκτιμήσεις μικροπεριοχών.

²⁰ Rao J.N.K., (2003), Small area estimation, Wiley-Interscience.

²¹ Platek. R., R, J.N.K., Sarndal, C.E., and Singh, M.P. (Editors) (1987). *Small Area Statistics: An International Symposium*. John Wiley and Sons: New York.

ΚΕΦΑΛΑΙΟ 5^ο

ΕΜΠΕΙΡΙΚΗ ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ

ΆΜΜΕΣΟΙ ΚΑΙ ΈΜΜΕΣΟΙ ΕΚΤΙΜΗΤΕΣ: ΕΚΤΙΜΗΣΕΙΣ ΓΙΑ ΤΗΝ ΑΠΑΣΧΟΛΗΣΗ ΣΕ ΑΕ & ΕΠΕ ΑΝΑ ΝΟΜΟ

5.1 ΣΤΑΤΙΣΤΙΚΗ ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΕΡΕΥΝΗΤΙΚΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Για την μελέτη της έρευνας εργατικού δυναμικού χρειαζόμαστε έναν πληθυσμό που να προσομοιάζει η δομή του με αυτήν. Αυτό σημαίνει ότι θα ακολουθήσουμε το δειγματοληπτικό σχέδιο της ΕΣΥΕ στην Έρευνα Εργατικού δυναμικού και κατόπιν θα προχωρήσουμε στις εκτιμήσεις μικρής περιοχής. Με αυτόν τον τρόπο εξασφαλίζουμε ότι τα συμπεράσματά μας θα ταυτίζονται με αυτά της ΕΣΥΕ. Το στρώμα της ΕΣΥΕ αντιστοιχεί στην περιφέρεια του πειράματός μας. Ο νομός στην μικροπεριοχή ενδιαφέροντος. Το νοικοκυριό στην επιχείρηση, και το σύνολο των εργαζομένων στο σύνολο των ραζομένων στην επιχείρηση.

Σε αυτήν την ενότητα παρουσιάζονται τα ζητήματα που αναφέρονται στη στατιστική διατύπωση του ερευνητικού προβλήματος. Πιο συγκεκριμένα τα ζητήματα αυτά χωρίζονται σε τέσσερις ενότητες.

- Πληθυσμός ενδιαφέροντος
- Συλλογή δεδομένων
- Μεταβλητή ενδιαφέροντος πληθυσμού
- Ακρίβεια των εκτιμήσεων

Πληθυσμός ενδιαφέροντος: Ο πληθυσμός ενδιαφέροντος είναι οι επιχειρήσεις ΑΕ και ΕΠΕ, οι οποίες αποτελούν ένα πεπερασμένο σύνολο αντικειμένων και βρίσκονται στον ελληνικό χώρο την χρονική περίοδο 2003.

Συλλογή Δεδομένων: Η κατάστρωση ενός καταλόγου με όλα τα στοιχεία του πληθυσμού και κατόπιν η συλλογή πληροφοριών, που ενδιαφέρουν την έρευνα,

αποτελεί μια ακριβή διαδικασία ιδιαίτερα στην περίπτωση που εξετάζονται μεγάλοι σε μέγεθος πληθυσμοί που κατανέμονται σε μεγάλες εδαφικές επιφάνειες. Ένας τέτοιος κατάλογος αποτελεί το δειγματοληπτικό πλαίσιο μιας έρευνας από το οποίο αντλείται το δείγμα. Στην περίπτωση μελέτης το κόστος συλλογής στοιχείων μπορεί να παραβλεφθεί, καθώς το δειγματοληπτικό πλαίσιο λαμβάνεται σε μορφή ηλεκτρονικού καταλόγου από την ICAP. Το δειγματοληπτικό πλαίσιο είναι οργανωμένο και σχεδιασμένο γεωγραφικά με συστηματικό τρόπο, ενώ περιέχει και βοηθητικά δεδομένα που μπορούν να χρησιμεύσουν στο σχεδιασμό της έρευνας. Τα στοιχεία του καταλόγου είναι οι επιχειρήσεις που αποτελούν τις δειγματοληπτικές μονάδες. Η παρατήρηση των στοιχείων ενός δείγματος γίνεται με την επιλογή ενός δείγματος και κατόπιν με την παρατήρηση και καταγραφή όλων ή ενός μέρους από τα στοιχεία που απαρτίζουν τις μονάδες που τελικά επελέγησαν.

Μεταβλητή ενδιαφέροντος πληθυσμού: Συνήθως ο σκοπός της δειγματοληπτικής έρευνας είναι περιγραφικός. Αποσκοπεί δηλαδή στην στατιστική περιγραφή ενός πληθυσμού (ή ενός μέρους του), στην κατάσταση που βρίσκεται αυτός ο πληθυσμός μια συγκεκριμένη χρονική στιγμή. Ο περιγραφικός σκοπός της έρευνας σημαίνει ότι πρέπει να εκτιμηθούν ορισμένα αριθμητικά χαρακτηριστικά ή διαφορετικά να εκτιμηθεί μια μεταβλητή ενδιαφέροντος. Τα χαρακτηριστικά αυτά αν αφορούν τον πληθυσμό στο σύνολο του καλούνται συνήθως *χαρακτηριστικά του πληθυσμού*, ή *πληθυσμιακές παράμετροι*. Αν τα χαρακτηριστικά αυτά αφορούν τμήματα μόνο του πληθυσμού, τότε καλούνται συνήθως, *χαρακτηριστικά, παράμετροι, πεδία μελέτης (domain) ή υποπληθυσμοί (subgroup)*.

Τα χαρακτηριστικά που εξετάζονται συνήθως από τους ερευνητές είναι:

- οι μέσοι όροι
- τα ποσοστά
- τα συνολικά ή επίμερους αθροίσματα
- οι αναλογίες των χαρακτηριστικών

Στην μελέτη περίπτωσης η προς εκτίμηση μεταβλητή ενδιαφέροντος είναι η απασχόληση στις ΑΕ και ΕΠΕ της Ελλάδας και συγκεκριμένα η συνολική απασχόληση σε επίπεδο νομού.

Ακρίβεια εκτιμήσεων: Το μέγεθος του δείγματος σχετίζεται με την απαιτούμενη ακρίβεια της εκτίμησης. Με τον όρο «ακρίβεια» εκτίμησης προσδιορίζουμε τον απαιτούμενο βαθμό απόκλισης της εκτίμησης από την πραγματική τιμή της μεταβλητής ενδιαφέροντος στον πληθυσμό. Συνήθως, επιθυμούμε σταθερές τυπικές αποκλίσεις περίπου 2 (1,96) με πιθανότητα σφάλματος της 5% (επίπεδο εμπιστοσύνης 95%).

5.2 ΔΙΑΔΙΚΑΣΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ ΓΙΑ ΕΚΤΙΜΗΣΗ ΜΙΚΡΟΠΕΡΙΟΧΩΝ

Η διαδικασία για την επιλογή του δείγματος περιγράφεται από τα ακόλουθα βήματα:

1. Απόκτηση σχετικού καταλόγου.
2. Καθορισμός μεγέθους του δείγματος (κλάσμα δειγματοληψίας). Για τις ανάγκες της έρευνας το κλάσμα δειγματοληψίας κρίνεται ότι πρέπει να είναι 10%.
3. Καθορισμός των στρωμάτων και του μεγέθους κάθε στρώματος των καταλόγων. Πρόσθεση όλων των στρωμάτων από όλους των καταλόγων για τον καθορισμό του μεγέθους του πληθυσμού.
4. Καθορισμός του ποσοστού του συνολικού πληθυσμού που αντιπροσωπεύεται σε κάθε στρώμα στο δείγμα (κλάσμα δειγματοληψίας κατά στρώμα). Αυτά τα ποσοστά, δηλαδή το κλάσμα δειγματοληψίας κατά στρώμα θα πρέπει να αντιπροσωπεύεται στο δείγμα. Το αποτέλεσμα θα είναι ένα στρωματοποιημένο δείγμα που θα βασίζεται στη μεταβλητή στρωματογράφησης και στο μέγεθος του καταλόγου.
5. Σχεδιασμός και εξαγωγή ενός δείγματος μέσα σε κάθε στρώμα. Το δειγματοληπτικό σχέδιο που θα ακολουθηθεί για την δειγματοληψία σε κάθε στρώμα είναι εκείνο της απλής τυχαίας δειγματοληψίας. Επειδή ο κατάλογος είναι ηλεκτρονικός (βάση δεδομένων) μπορεί εύκολα να εξαχθεί ένα απλό τυχαίο δείγμα.

Ειδικότερα, για την επιλογή του δείγματος χρησιμοποιήθηκε η τεχνική της στρωματοποιημένης τυχαίας δειγματοληψίας (stratified random sampling). Με την στρωματοποιημένη δειγματοληψία επιτυγχάνεται καλύτερη αντιπροσώπευση του

πληθυσμού, ενώ μειώνεται το σφάλμα εκτίμησης. Η στρωματοποίηση έγινε ανάλογα με την γεωγραφική περιοχή στην οποία βρίσκονται οι επιχειρήσεις. Οι καταγραφωμένες επιχειρήσεις συνιστούν τον στατιστικό πληθυσμό από τον οποίο επιλέχτηκε το τελικό δείγμα.

Πίνακας 1: Κατανομή των επιχειρήσεων ΑΕ και ΕΠΕ ανά περιφέρεια

	Επιχειρήσεις	Ποσοστό
ΑΝΑΤΟΛΙΚΗ ΜΑΚΕΔΟΝΙΑ ΘΡΑΚΗ	666	3,1
ΑΤΤΙΚΗ	12008	56,4
ΔΥΤΙΚΗ ΕΛΛΑΔΑ	560	2,6
ΔΥΤΙΚΗ ΜΑΚΕΔΟΝΙΑ	234	1,1
ΗΠΕΙΡΟΣ	275	1,3
ΘΕΣΣΑΛΙΑ	682	3,2
ΙΟΝΙΟΙ ΝΗΣΟΙ	373	1,8
ΚΕΝΤΡΙΚΗ ΜΑΚΕΔΟΝΙΑ	3188	15,0
ΚΡΗΤΗ	1260	5,9
ΝΗΣΙΑ ΒΟΡΕΙΟΥ ΑΙΓΑΙΟΥ	231	1,1
ΝΗΣΙΑ ΝΟΤΙΟΥ ΑΙΓΑΙΟΥ	837	3,9
ΠΕΛΟΠΟΝΝΗΣΟΣ	425	2,0
ΣΤΕΡΕΑ ΕΛΛΑΔΑ	566	2,7
Σύνολο	21305	100,0

Σύμφωνα με την στρωματοποιημένη δειγματοληψία, οι περιφέρειες χρησιμοποιούνται ως στρώματα για τη δειγματοληψία, αποτελούν δηλαδή υποομάδες μέσα στις οποίες πραγματοποιείται απλή τυχαία δειγματοληψία. Με τον τρόπο αυτό επιτυγχάνεται καλύτερη αντιπροσώπευση του πληθυσμού στο δείγμα. Αναλυτικά, αφού οριστούν οι περιοχές-στρώματα, γίνονται επιμέρους δειγματοληπτικές έρευνες, μία σε κάθε στρώμα, με μεγέθη δειγμάτων που περιγράφονται παρακάτω. Ο αριθμός στρωμάτων είναι 13, όσες και οι περιφέρειες της Ελλάδας.

Για να προσδιοριστεί το μέγεθος του δείγματος σε κάθε στρώμα, εφαρμόζεται αναλογική δειγματοληψία (proportionate stratified sampling). Σύμφωνα με αυτή, το σχετικό μέγεθος σε κάθε ένα από τα επιμέρους δείγματα, προσδιορίζεται αναλογικά με το σχετικό μέγεθος των στρωμάτων στον πληθυσμό. Με άλλα λόγια κάθε περιοχή διατηρεί στο δείγμα το ίδιο ποσοστό αντιπροσώπευσης που έχει στον πληθυσμό των καταγραφωμένων επιχειρήσεων. Η αναλογική στρωματοποιημένη δειγματοληψία επιλέχτηκε με το σκεπτικό ότι όλα τα δεδομένα πρέπει να είναι διαθέσιμα για επεξεργασία σε πολλούς χρήστες, να αποφεύγονται διαφορετικές σταθμίσεις και να μπορούν να συμπληρωθούν με νέα δεδομένα. Το μέγεθος του δείγματος σε κάθε στρώμα, δηλαδή η κατανομή του δείγματος ορίστηκε στο 10% ανά μέγεθος στρώματος. Αυτό σημαίνει ότι αν ένα στρώμα έχει 100 επιχειρήσεις, τότε θα επιλεγθούν οι 10. Η επιλογή του 10% των επιχειρήσεων ανά στρώμα αντλήθηκε με

τυχαία δειγματοληψία χωρίς επανατοποθέτηση. Η απλή τυχαία δειγματοληψία υλοποιείται με τη βοήθεια τυχαίων αριθμών που παράγονται από αυτόματα με το στατιστικό πακέτο SPSS.

Με τον όρο μικροπεριοχή περιγράφεται το ενδιαφέρον μας για εκτιμήσεις απασχόλησης σε επίπεδο νομού. Οι εκτιμήσεις απαιτούν πληροφορίες για την απασχόληση επιχειρήσεων ανά νομό. Οι πληροφορίες αυτές αντλούνται από το αρχικό δείγμα των επιχειρήσεων ανά περιφέρεια. Σε αυτή την περίπτωση μερικές μικροπεριφέρειες εκτίμησης υποεκπροσωπούνται ή δεν εμφανίζουν καμία επιχείρηση στο δείγμα τους. Θα πρέπει να τονιστεί ότι για τον πεπερασμένο πληθυσμό γνωρίζουμε τον αριθμό επιχειρήσεων ΑΕ και ΕΠΕ ανά νομό, οπότε και η πληροφορία αυτή μεταφέρεται και για τις δειγματικές παρατηρήσεις. Με την αξιοποίηση βοηθητικών στοιχείων βελτιώνονται οι εκτιμήσεις από το δείγμα.

5.3 ΕΜΠΕΙΡΙΚΗ ΑΝΑΛΥΣΗ & ΑΠΟΤΕΛΕΣΜΑΤΑ

Στην μελέτη ως μικροπεριοχές ορίζονται οι νομοί με την έννοια ότι το στρώμα που βασίζεται η δειγματοληψία είναι η περιφέρεια. Ο αριθμός των παρατηρήσεων του δείγματος που αντιστοιχούν σε κάθε νομό είναι πολύ μικρός ή δεν υπάρχουν καθόλου. Ο αριθμός των μικροπεριοχών στην μελέτη περίπτωσης είναι 51 νομοί της Ελλάδας, ενώ θα πρέπει να διασαφηνιστεί ότι ο αριθμός των μικροπεριοχών θα διαφέρει ανάλογα με την μελέτη περίπτωσης (η εκτίμηση της απασχόλησης ανά νομό και τους τρεις κλάδους θα σήμαινε 51×3 μικροπεριοχές).

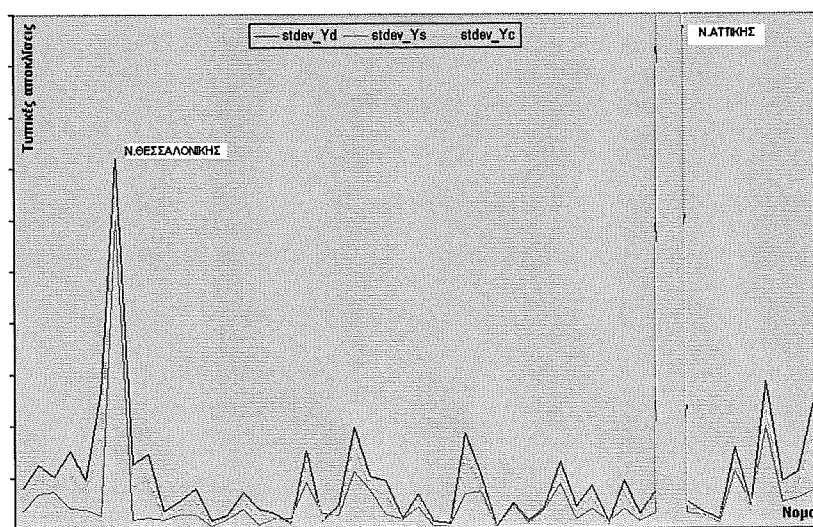
Η εμπειρική ανάλυση των άμεσων και των έμμεσων εκτιμητών έχει σαν στόχο την αξιολόγηση της συμπεριφοράς των εκτιμητών στην μικροπεριοχή νομός. Η αξιολόγηση αυτή βασίστηκε στην ανάλυση των χαρακτηριστικών των εκτιμητών και στην σύγκριση του μέσου τετραγωνικού σφάλματος (MSE) των εκτιμητών. Η ανάλυση βασίστηκε σε 100 δείγματα του πληθυσμού των επιχειρήσεων, και για τον υπολογισμό των εκτιμητών χρησιμοποιήθηκε λογισμικό Gauss.

Ο άμεσος εκτιμητής έδωσε καλύτερες εκτιμήσεις από την άποψη ότι είναι ένας αμερόληπτος εκτιμητής με τον μέσο της δειγματοληπτικής κατανομής να προσεγγίζει καλύτερα τον μέσο του πληθυσμού από τον σχεδιασμό του πειράματος.

Όμως η διακύμανση του άμεσου εκτιμητή της συνολικής απασχόλησης ανά νομό είναι μεγαλύτερη από την διακύμανση των έμμεσων εκτιμητών. Αυτό είναι αναμενόμενο, καθώς οι εκτιμήσεις βασίζονται σε τυχαία δείγματα ανά νομό που στις περισσότερες περιπτώσεις είναι και μικρά ή μηδενικά. Αυτό πρακτικά σημαίνει ότι η επιλογή ενός δείγματος από τον πληθυσμό είναι δυνατόν να μας δώσει εκτιμήσεις με μεγάλη κύμανση για κάθε μικροπεριοχή.

Η διακύμανση του συνθετικού εκτιμητή είναι μικρότερη σε σχέση με αυτή του άμεσου εκτιμητή, καθώς χρησιμοποιήσαμε ως βοηθητική μεταβλητή τον αριθμό των επιχειρήσεων ανά νομό και με αναγωγές στα δεδομένα του πληθυσμού πήραμε εκτιμήσεις ακόμα και για περιοχές όπου δεν είχαμε παρατηρήσεις από το δείγμα. Σε κάθε περίπτωση η διακύμανση ενός συνθετικού είναι μικρότερη σε σχέση τον άμεσο εκτιμητή, όμως αυτό το πλεονέκτημα ενδέχεται να αντισταθμίζεται με το μειονέκτημα της ανισορροπίας-αστάθειας του εκτιμητή λόγω της μεροληψίας.

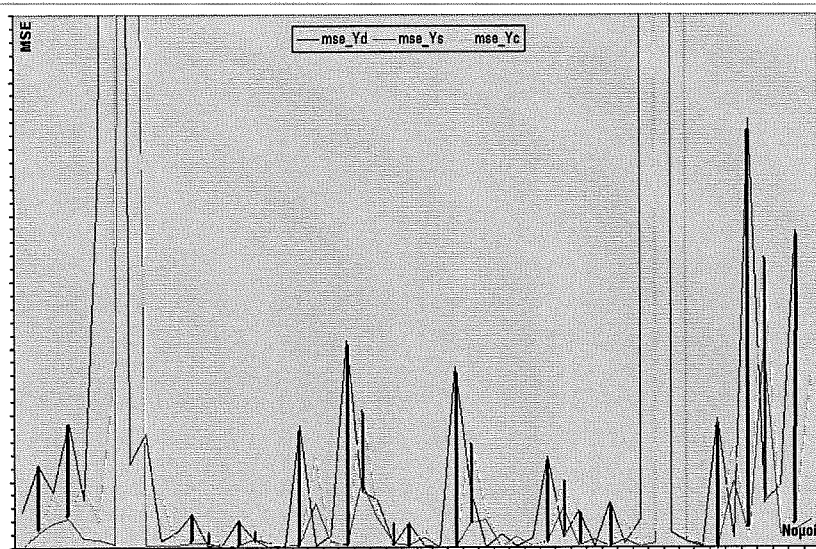
Τυπικές αποκλίσεις άμεσων και έμμεσων εκτιμητών



Το μέσο τετραγωνικό σφάλμα χρησιμοποιήθηκε ως κριτήριο καλής συμπεριφοράς για την αξιολόγηση της καταλληλότητας των εκτιμητών σε περιπτώσεις εκτίμησης μικροπεριοχών. Η χρησιμότητα του είναι πρόδηλη καθώς μετρά την διακύμανση και την μεροληψία του εκτιμητή. Στην περίπτωση του άμεσου εκτιμητή το μέσο τετραγωνικό σφάλμα υπολογίστηκε με την διακύμανση του εκτιμητή, λόγω της αμεροληψίας του εκτιμητή. Αντίθετα για τον υπολογισμό του μέσου τετραγωνικού σφάλματος των έμμεσων εκτιμητών απαιτήθηκε ο

συνυπολογισμός της μεροληψίας λόγω της φύσης των εκτιμητών. Συγκεκριμένα η υπόθεση για την ομοιογένεια στο χώρο, είναι δύσκολο να ικανοποιείται απόλυτα, μεταξύ χαρακτηριστικών των νομών και των περιφερειών στον οποίο ανήκουν, καθώς οι εκτιμήσεις για την συνολική απασχόληση για την περιφέρεια διαφέρουν από εκείνες των νομών που ανήκουν σε αυτή.

Μέσο τετραγωνικό σφάλμα άμεσων και έμμεσων εκτιμητών



Απόσταση όταν $MSE Y_d > MSE Y_s \text{ \& } MSE Y_c$ ———

Απόσταση όταν $MSE Y_d < MSE Y_c$ ———

Το μέσο τετραγωνικό σφάλμα (διακύμανση) για τον άμεσο εκτιμητή υπολογίσθηκε μεγαλύτερο από αυτό των έμμεσων εκτιμητών για τις περισσότερες των περιπτώσεων ανά νομό. Εξαίρεση αποτέλεσαν οι νομοί Θεσσαλονίκης, Αττικής, Κυκλάδων, Πρέβεζας, Αρκαδίας, Ευρυτανίας. Αυτό εξηγεί την καταλληλότητα του άμεσου εκτιμητή όταν έχουμε μικροπεριοχή με μεγάλο δείγμα παρατηρήσεων (Θεσσαλονίκης, Αττικής, Κυκλάδων) και ανομοιογένεια της απασχόλησης στο χώρο (περιφέρεια), όπου βρίσκεται ο νομός (Πρέβεζας, Αρκαδίας, Ευρυτανίας). Από την ανάλυση εκτιμήθηκε ότι καταλληλότεροι εκτιμητές για την εκτίμηση της απασχόλησης ανά νομό ως μικροπεριοχή είναι οι έμμεσοι και συγκεκριμένα ο συνθετικός καθώς δίνει το μικρότερο μέσο τετραγωνικό σφάλμα ανά νομό. Αυτό σημαίνει ότι η διακύμανση ενός συνθετικού είναι μικρότερη σε σχέση με τον άμεσο εκτιμητή, και το πλεονέκτημα αυτό δεν αντισταθμίζεται από το μειονέκτημα της αστάθειας του εκτιμητή λόγω της μεροληψίας.

5.4 ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι εκτιμήσεις σε κατώτερο επίπεδο ή επίπεδο μικροπεριοχής αποτελούν τον πυρήνα του στατιστικού ενδιαφέροντος. Η στατιστική προσέγγιση για την παραγωγή τέτοιων εκτιμήσεων από τους άμεσους εκτιμητές συχνά αποδεικνύεται αναξιόπιστη. Αυτό αποδεικνύεται και από την στατιστική επεξεργασία για την εκτίμηση της απασχόλησης στην μικροπεριοχή νομός. Ο εναλλακτικός τρόπος για την παραγωγής εκτιμήσεων μικροπεριοχής είναι οι έμμεσοι εκτιμητές. Η εφαρμογή της τεχνικής των έμμεσων εκτιμητών προϋποθέτει τον όσο το δυνατό πιο ομοιογενή χώρο από την άποψη της μεταβλητής ενδιαφέροντος. Η προσεχτική επιλογή των βοηθητικών μεταβλητών και της μεθόδου δειγματοληψίας κάνουν πιο εφαρμόσιμες τις τεχνικές αυτές

Σε αυτήν την μελέτη χρησιμοποιήθηκαν οι έμμεσοι εκτιμητές σύνθετος και συνθετικός εκτιμητής, με τους οποίους συγκρίναμε τον άμεσο εκτιμητή και αξιολογήσαμε την καταλληλότητα τους κατά περίπτωση. Η εφαρμογή ανέδειξε τις αδυναμίες και τα ισχυρά σημεία των άμεσων και των έμμεσων εκτιμητών. Οι έμμεσοι εκτιμητές αποδείχθηκαν καταλληλότεροι στις περιπτώσεις ομοιογενών ως προς την απασχόληση περιφερειών, ακόμα και με μικρό δείγμα, ενώ οι άμεσοι εκτιμητές προσαρμόζονται ικανοποιητικά τόσο σε ομοιογενείς, όσο και σε ανομοιογενείς περιφέρειες αρκεί να έχουμε πολλές παρατηρήσεις. Σε αντίθετη περίπτωση όπως και συνέβη στην εφαρμογή θα έχουμε εκτιμήσεις με μεγάλες διακυμάνσεις. Συνολικά στην μελέτη εφαρμογής κρίθηκε ως καταλληλότερος εκτιμητής ο συνθετικός κυρίως λόγω της σταθερότητας των διακυμάνσεων και του χαμηλού μέσου τετραγωνικού σφάλματος.

6. ΠΑΡΑΡΤΗΜΑ

```

closeall;
cls;

fname1 = "c:\\tank\\ioannou\\analysis\\sae02.xls";
range1 = "b2:da21306";
xdata = xlsreadm(fname1, range1, 1, "");          /* Δειγματοληπτικό πλαίσιο (ανά επιχείρηση) και 100 δείγματα (21306 γραμμές x 105 στήλες) */

fname2 = "c:\\tank\\ioannou\\analysis\\sae03.xls";
range21 = "b2:b52";
range22 = "c2:c52";
range23 = "b64:b76";
range24 = "c64:c76";
nomsa = xlsreadsa(fname2, range21, 3, "");
nomnu = xlsreadsa(fname2, range22, 3, "");
persa = xlsreadsa(fname2, range23, 3, "");
pernu = xlsreadsa(fname2, range24, 3, "");

fname3 = "c:\\tank\\ioannou\\analysis\\sae04.xls";
range3 = "a2:e52";
realdata= xlsreadm(fname3, range3, 1, "");          /* Πραγματικά στοιχεία πληθυσμού ανά περιφέρεια και νομό (51 γραμμές x 5 στήλες)
                                                    Στήλη 1: Περιφέρεια και νομός
                                                    Στήλη 2: Περιφέρεια
                                                    Στήλη 3: Νομός
                                                    Στήλη 4: Πλήθος επιχειρήσεων
                                                    Στήλη 5: Σύνολο απασχόλησης */

yd=zeros(rows(nomnu),100); /* Μηδενισμός μήτρας Direct */
ys=zeros(rows(nomnu),100); /* Μηδενισμός μήτρας Synthetic */
yc=zeros(rows(nomnu),100); /* Μηδενισμός μήτρας Composite */

vyd=zeros(rows(nomnu),100); /* Μηδενισμός μήτρας διακύμανσης Direct */
vys=zeros(rows(nomnu),100); /* Μηδενισμός μήτρας διακύμανσης Synthetic */

for j (5, 104, 1);
    sample_init=xdata(:,1:4)-xdata(:,j);
    y=sample(sample_init);
    {yd[:,j-4], vyd[:,j-4]}=direct(y, realdata(:,1:3)); /* Η ρουτίνα επιστρέφει 2 μήτρες, τη μήτρα direct και τη μήτρα των διακυμάνσεων κατά νομό */
    ys[:,j-4]=synthetic(y);
endfor;

for k (1,100,1);
    for j (1,51,1);
        if yd[j,k]>0;
            vys[j,k]=(ys[j,k]-yd[j,k])^2/realdata[j,4]^2;
        elseif ys[j,k]==0;
            vys[j,k]=0;
        endif;
    endfor;
endfor;

```

```

mseys=zeros(51,100);
for k (1,100,1);
    for j (1,51,1);
        mseys[j,k]=vys[j,k]+vyd[j,k];
    endfor;
endfor;

f=zeros(51,100);
for k (1,100,1);
    for j (1,51,1);
        if mseys[j,k]>0;
            f[j,k]=1/(1+(vyd[j,k]/mseys[j,k]));
        else;
            f[j,k]=1;
        endif;
    endfor;
endfor;

fi=zeros(51,100);
for k (1,100,1);
    for j (1,51,1);
        fi[j,k]=1/(1+f[j,k]);
    endfor;
endfor;

yc=zeros(51,100);
mseyc=zeros(51,100);
for k (1,100,1);
    for j (1,51,1);
        yc[j,k]=fi[j,k]*yd[j,k]+(1-fi[j,k])*ys[j,k];
        mseyc[j,k]=fi[j,k]^2*vyd[j,k]+(1-fi[j,k])^2*mseys[j,k];
    endfor;
endfor;

/* Εκτίμηση συνολικής απασχόλησης με τον direct (το πλήθος των επιχειρήσεων ανά νομό είναι
/* Εκτίμηση συνολικής απασχόλησης με τον synthetic (το πλήθος των επιχειρήσεων ανά νομό ε
/* Εκτίμηση συνολικής απασχόλησης με τον composite (το πλήθος των επιχειρήσεων ανά νομό ε

for k (1,100,1);
    for j (1,51,1);
        yd[j,k]=yd[j,k]*realdata[j,4];
        ys[j,k]=ys[j,k]*realdata[j,4];
        yc[j,k]=yc[j,k]*realdata[j,4];
    endfor;
endfor;

print realdata[:,1]~yd;
print realdata[:,1]~ys;
print realdata[:,1]~yc;

end;

```

/* Η ρουτίνα διαλέγει από τη μήτρα δεδομένων xdata τα δειγματικά σημεία.

Ταξινομεί ως προς τους άσους (δειγματικό σημείο) και διαλέγει μέχρις εκείνη τη γραμμή που είναι ίση με το άθροισμα των άσων. Επίσης επαναταξινομεί ως προς περιφέρειες και νομούς.

Τελική γραμμογράφηση y

Ετήλη 1: Περιφέρεια+Νομός, Ετήλη 2: Περιφέρεια, Ετήλη 3:Νομός, Ετήλη 4:Ελάδος, Ετήλη 5:Προσωπικό */

```
proc (1) = sample(y);
    local c;
    y= rev(sortc(y,5));
    c= sumc(y[,5]);
    y= y[1:c,1:4];
    y= (y[,1]*1000+y[,2])~y;
    y= sortc(y,1);
    retp(y);
endp;
```

/* Υπολογισμός του Direct. Υπολογίζω τα δειγματικά σημεία ανά νομό με το μετρητή. Ελέγχω σε διπλό λουπ τον κωδικό νομού και αυξάνω τον μετρητή. Με νέο λουπ και κατάλληλη χρήση του αθροιστή για συγκεκριμένο αριθμό γραμμών υπολογίζω τον direct */

```
proc (2) = direct(y, realdata);
    local nom, first, last, c, yd, vyd;
    nom=realdata[,3];
    c=zeros(51,1); /* αρχικός μηδενισμός μετρητή */
    yd=zeros(51,1); /* μηδενισμός διανύσματος direct */
    vyd=zeros(51,1); /* μηδενισμός διανύσματος διακυμάνσεων direct */
    y=sortc(y,1); /* ταξινόμηση διανύσματος δείγματος ανά νομό */
    for k (1, 51, 1); /* εξωτερικό λουπ για τους κωδικούς νομών */
        for j (1, rows(y), 1); /* εσωτερικό λουπ για τα δειγματικά σημεία ανά νομό */
            if nom[k] == y[j,3];
                c[k]=c[k]+1;
            endif;
        endfor;
        yd[k]=sumc(y[first:last,5])/c[k]; /* Υπολογισμός του Direct */
        first=first+c[k];
    endfor;
    first=1;
    last=0;
    for k (1, 51, 1);
        last=last+c[k];
        if c[k] > 0;
            vyd[k]=((y[first:last,5]-yd[k])^2*(y[first:last,5]-yd[k]))/c[k]^2; /* Υπολογισμός της διακύμανσης του Direct */
        endif;
        first=first+c[k];
    endfor;
```

```

endfor;
retp(yd, vvd);
endp;

proc (1) = synthetic(y);
local per, ys, c, nom1, nom2;
per={5, 7, 4, 4, 4, 4, 5, 3, 5, 1, 3, 2, 4};
per=seqa(1, 1, 13)~per;
ys=zeros(51,1);
y=sortc(y,1);
c=zeros(13,1);
for k (1, 13,1);          /* εξωτερικό loop για τους κωδικούς περιφερειών */
    for j (1, rows(y), 1); /* εσωτερικό loop για τα δειγματικά σημεία ανά περιφέρεια */
        if k == y[j,2];
            c[k]=c[k]+1;
        endif;
    endfor;
endfor;
nom1=1;                  /* Μετρητής που χρησιμοποιείται για να αλλάζει τους νομούς στη μήτρα δεδομένων y */
nom2=0;                  /* Μετρητής που χρησιμοποιείται για να αλλάζει τους νομούς στον synthetic */
for k (1,13,1);
    for j (1,per[k,2],1);
        ys[nom2+j]=sumc(y[nom1:(nom1-1+c[k]),5])/c[k];
    endfor;
    nom1=nom1+c[k];
    nom2=nom2+per[k,2];
endfor;
retp(ys);
endp;

```

estimator	nomos	realdB	s00	s01	s02	s03	s04	s05	mean	stdev	mse	cv
yd	1009	3204	4745	5604	2791	3588	1728	2254	3306	1639	2686519	0,49573105
yd	1011	4126	2708	5441	10167	2213	9856	2288	4208	2490	6202557	0,59182792
yd	1021	5364	3268	3808	8684	10095	9922	7136	5411	2043	4172235	0,37749819
yd	1037	4830	3620	5421	3139	2383	2541	6912	4569	3053	9318770	0,66811581
yd	1042	3375	7360	1818	4737	4235	4444	2283	3558	1897	3597985	0,53310042
yd	2016	5387	1528	1595	2090	19316	4595	2471	5265	5334	28447317	1,01309497
yd	2019	85377	74299	84048	89546	81676	105472	74434	88987	14378	206736115	0,16157716
yd	2026	4244	5236	2187	1831	2578	1958	2168	4246	2510	6299983	0,59117361
yd	2038	4084	6432	4488	10447	7828	6136	4176	4295	2932	8596839	0,68273104
yd	2039	2100	3526	2016	1691	2077	2378	2797	2023	698	486630	0,34488495
yd	2044	2851	2976	2221	1927	1298	2201	3933	2823	1099	1208637	0,38943602
yd	2049	2907	2288	1930	2390	1864	2193	1882	2752	1580	2495172	0,57402429
yd	3008	472	198	0	123	403	143	198	385	339	115090	0,88165245
yd	3023	1049	824	1506	1271	1116	971	1312	962	525	275429	0,54538129
yd	3027	3104	3010	5242	3732	5250	4048	3221	3251	1433	2053472	0,44077803
yd	3047	595	200	455	292	355	283	1710	543	787	619415	1,44827028
yd	4004	817	789	408	638	442	2550	517	741	590	348383	0,79697845
yd	4018	343	396	231	223	231	253	242	346	208	43466	0,60173366
yd	4020	4633	3444	4768	2403	4185	2850	4683	4667	3025	9151480	0,64817022
yd	4040	853	1259	939	808	1020	1078	451	795	342	116901	0,43029476
yd	5022	1557	579	1662	1225	1568	1475	770	1498	950	902185	0,63412277
yd	5031	10122	10600	8436	10642	7679	9856	11130	10167	3945	15559893	0,3879677
yd	5035	4461	12430	4127	5981	3692	6646	4802	5063	2044	4176876	0,40367264
yd	5045	2579	923	1287	2707	5042	1099	891	2472	1877	3523722	0,75924113
yd	6014	1323	1368	2172	1800	1248	1591	1812	1367	399	159507	0,29218742
yd	6024	4325	2413	3745	4360	6400	4970	7402	4432	1346	1810836	0,30365038
yd	6025	586	406	532	378	292	600	613	543	301	90891	0,55532505
yd	6034	272	432	547	459	234	169	0	251	185	34263	0,73693174
yd	7007	7593	5278	9064	4428	5840	3810	5210	7358	3684	13572528	0,50067302
yd	7012	5435	6813	5808	6205	10266	4015	1779	5466	2091	4372816	0,38256641
yd	7013	123	11	248	154	44	44	44	70	77	5870	1,0875486
yd	7046	2651	2521	4218	2932	3325	4193	2018	2835	1008	1016323	0,35557255
yd	7048	446	330	883	121	158	237	176	418	304	92488	0,72722939
yd	8001	2343	4126	1441	1486	2279	1193	3002	2356	826	681986	0,35055856
yd	8006	8160	8770	6404	15676	8666	12766	8371	8692	2604	6782397	0,29960726
yd	8015	1392	1913	1547	1258	1228	1193	1199	1520	856	732078	0,56282996
yd	9002	3419	2338	2640	1743	2135	3325	3782	3193	1667	2779183	0,5221864
yd	9003	503	410	266	641	609	638	651	498	248	61347	0,49778548
yd	9028	4156	5732	2091	4200	6216	4637	3721	4370	1862	3468669	0,42622676
yd	9030	637	565	0	719	425	587	374	661	584	341278	0,88433854
yd	9036	1893	1880	3137	1657	1577	1453	1214	1959	1450	2101286	0,74001649
yd	10005	640849	598597	508522	634858	591765	708644	514617	659005	112082	12562383063	0,17007764
yd	11033	1964	820	1236	1685	1921	1321	1473	1617	1063	1129387	0,65702888
yd	11043	1193	1127	2590	1057	653	1210	1068	1218	638	407524	0,52426365
yd	11051	531	595	532	481	456	374	372	532	350	122777	0,65883699
yd	12010	11388	12238	13086	7906	9555	7198	11090	10904	3108	9661040	0,28505074
yd	12029	2896	2982	4998	3567	2033	2843	1823	2937	891	794306	0,30344354
yd	13017	16099	12397	10658	31286	12982	21628	12853	16236	5673	32186386	0,34941746
yd	13032	2994	2482	5063	2751	1493	1943	2484	3055	1828	3339778	0,59813133
yd	13041	4045	3666	2992	2646	2428	2404	6932	4277	2171	4714679	0,5077219
yd	13050	6060	4284	3010	4744	2124	6309	5421	5802	4889	23901136	0,84264033
ys	1009	3204	2890	2987	4042	3337	3934	2943	2943	713	1079148	0,24219011
ys	1011	4126	5529	5714	7733	6384	7525	5629	5631	1364	1859560	0,24219011
ys	1021	5364	5969	6168	8348	6892	8124	6077	6078	1472	2167165	0,24219011
ys	1037	4830	3362	3474	4701	3881	4575	3422	3423	829	687309	0,24219011
ys	1042	3375	3173	3279	4438	3664	4319	3230	3231	783	612389	0,24219011
ys	2016	5387	3317	3424	3838	4017	4311	3167	3813	512	262101	0,13427696
ys	2019	85377	77639	80162	89851	94031	100925	74138	89252	11984	143627640	0,13427696
ys	2026	4244	2472	2553	2861	2994	3214	2361	2842	382	145650	0,13427696
ys	2038	4084	2895	2989	3350	3506	3763	2764	3327	447	199630	0,13427696

estimator	nomos	realdB	s00	s01	s02	s03	s04	s05	mean	stdev	mse	cv
ys	2039	2100	2472	2553	2861	2994	3214	2361	2842	382	145650	0,13427696
ys	2044	2851	3739	3860	4327	4528	4860	3570	4298	577	333063	0,13427696
ys	2049	2907	3588	3705	4152	4346	4664	3426	4125	554	306745	0,13427697
ys	3008	472	413	732	453	608	462	715	487	135	18287	0,27759344
ys	3023	1049	1165	2064	1278	1714	1302	2016	1373	381	145238	0,27759345
ys	3027	3104	2348	4161	2576	3455	2625	4065	2768	768	590358	0,27759344
ys	3047	595	470	832	515	691	525	813	554	154	23614	0,27759345
ys	4004	817	723	840	491	796	921	774	818	402	161981	0,49214304
ys	4018	343	702	815	477	773	894	752	794	391	152593	0,49214303
ys	4020	4633	3295	3828	2240	3628	4198	3530	3728	1835	3366441	0,49214303
ys	4040	853	1127	1309	766	1241	1436	1207	1275	627	393604	0,49214303
ys	5022	1557	2604	1637	2270	2051	2089	1774	2119	533	284144	0,25154744
ys	5031	10122	10866	6830	9472	8558	8719	7402	8844	2225	4948848	0,25154744
ys	5035	4461	7394	4648	6446	5824	5934	5037	6018	1514	2291789	0,25154744
ys	5045	2579	2812	1767	2451	2215	2256	1915	2289	576	331425	0,25154744
ys	6014	1323	1159	1811	1812	2181	1881	2815	1717	382	146058	0,22264027
ys	6024	4325	2595	4055	4058	4884	4212	6304	3844	856	732588	0,22264027
ys	6025	586	423	660	661	795	686	1026	626	139	19414	0,22264027
ys	6034	272	326	509	510	613	529	792	483	107	11553	0,22264027
ys	7007	7593	5220	6725	4693	6143	3833	3219	5387	1352	1828986	0,25105296
ys	7012	5435	5577	7185	5014	6564	4095	3439	5755	1445	2087830	0,25105295
ys	7013	123	302	389	272	356	222	186	312	78	6130	0,25105295
ys	7046	2651	3544	4566	3186	4171	2602	2185	3657	918	843107	0,25105295
ys	7048	446	907	1168	815	1067	666	559	936	235	55174	0,25105295
ys	8001	2343	3808	2417	4777	3186	3923	3307	3235	736	541857	0,2275535
ys	8006	8160	8956	5683	11234	7492	9226	7776	7608	1731	2996796	0,2275535
ys	8015	1392	1943	1233	2438	1626	2002	1688	1651	376	141128	0,2275535
ys	9002	3419	2576	2664	2482	3413	2752	2648	3061	771	594100	0,25182533
ys	9003	503	887	917	854	1175	947	912	1054	265	70411	0,25182533
ys	9028	4156	2597	2686	2502	3441	2774	2669	3086	777	603879	0,25182533
ys	9030	637	929	961	895	1231	992	955	1104	278	77276	0,25182533
ys	9036	1893	1985	2053	1912	2629	2120	2040	2358	594	352692	0,25182533
ys	10005	640849	598597	508522	634858	591765	708644	514617	659005	112082	12562383063	0,17007764
ys	11033	1964	1082	2142	1461	1547	1295	1259	1556	622	387099	0,39981221
ys	11043	1193	926	1833	1250	1324	1108	1077	1332	532	283489	0,39981221
ys	11051	531	395	783	534	565	473	460	569	227	51680	0,39981221
ys	12010	11388	10529	12836	8070	8339	7120	9067	9813	2217	4916730	0,22596476
ys	12029	2896	4383	5343	3359	3471	2964	3774	4085	923	851867	0,22596476
ys	13017	16099	11724	11484	21323	9930	16539	14487	15193	3908	15274917	0,25723624
ys	13032	2994	3016	2954	5486	2555	4255	3727	3909	1006	1011046	0,25723624
ys	13041	4045	3519	3447	6400	2980	4964	4348	4560	1173	1376146	0,25723624
ys	13050	6060	4363	4273	7935	3695	6155	5391	5654	1454	2115266	0,25723625
yc	1009	3204	4126	4731	3208	3504	2464	2484	3187	1120	1253499	0,35126795
yc	1011	4126	3649	5532	9355	3604	9079	3402	4682	1972	3889509	0,42118453
yc	1021	5364	4168	4595	8572	9027	9323	6783	5633	1696	2874919	0,30098103
yc	1037	4830	3534	4772	3660	2883	3219	5749	4187	2178	4744097	0,52018622
yc	1042	3375	5965	2305	4637	4044	4402	2599	3449	1405	1973069	0,40724529
yc	2016	5387	2125	2205	2673	14216	4500	2703	4781	3590	12889833	0,75095652
yc	2019	85377	75412	82753	89648	85794	103957	74335	89076	13291	176662922	0,14921563
yc	2026	4244	4315	2309	2175	2717	2377	2233	3778	1684	2836752	0,44582103
yc	2038	4084	5253	3988	8081	6387	5345	3705	3972	1976	3903234	0,49734875
yc	2039	2100	3175	2195	2081	2383	2657	2652	2296	499	249284	0,21743738
yc	2044	2851	3230	2767	2727	2377	3088	3812	3315	796	633439	0,24009182
yc	2049	2907	2722	2522	2977	2692	3017	2397	3210	1082	1170490	0,33703927
yc	3008	472	270	366	244	472	256	457	439	216	46489	0,49155888
yc	3023	1049	937	1692	1273	1315	1082	1547	1103	394	155477	0,35750386
yc	3027	3104	2789	4882	3347	4652	3574	3502	3090	1109	1229072	0,3587811
yc	3047	595	335	581	367	467	364	1411	548	479	229137	0,87330235
yc	4004	817	767	552	589	560	2007	603	776	447	199946	0,57650406
yc	4018	343	498	427	308	502	474	417	512	196	38528	0,38367131

estimator	nomos	realdB	s00	s01	s02	s03	s04	s05	mean	stdev	mse	cv
yc	4020	4633	3395	4455	2348	3999	3299	4298	4354	2610	6814237	0,59951717
yc	4040	853	1215	1062	794	1094	1197	703	964	376	141130	0,38976891
yc	5022	1557	1261	1653	1573	1729	1680	1106	1706	661	436765	0,3872738
yc	5031	10122	10689	7901	10252	7972	9477	9887	9726	3271	10698578	0,33629733
yc	5035	4461	10751	4301	6136	4403	6408	4880	5381	1625	2640905	0,30198301
yc	5045	2579	1554	1447	2622	4100	1486	1233	2411	1339	1792897	0,55528386
yc	6014	1323	1298	2052	1804	1559	1687	2146	1484	322	103851	0,21721603
yc	6024	4325	2474	3848	4259	5894	4717	7036	4236	1167	1360884	0,27540096
yc	6025	586	412	575	473	460	629	751	574	207	42713	0,36017185
yc	6034	272	397	534	476	361	292	396	350	110	12133	0,31515624
yc	7007	7593	5259	8284	4517	5941	3818	4547	6701	2821	7956056	0,42092001
yc	7012	5435	6401	6267	5808	9032	4042	2332	5563	1657	2745701	0,2978875
yc	7013	123	157	296	194	200	133	115	185	47	2215	0,25507543
yc	7046	2651	2862	4334	3017	3607	3662	2074	3109	805	648109	0,25890752
yc	7048	446	618	978	366	483	382	305	610	220	48514	0,36124065
yc	8001	2343	4020	1767	2585	2581	2104	3103	2649	659	433826	0,24865089
yc	8006	8160	8832	6164	14195	8275	11586	8173	8331	2279	5193510	0,27355516
yc	8015	1392	1923	1443	1651	1361	1463	1362	1562	570	325277	0,36501703
yc	9002	3419	2418	2648	1989	2561	3134	3404	3149	1208	1459106	0,38363313
yc	9003	503	569	488	712	799	741	738	698	176	30860	0,25166404
yc	9028	4156	4687	2289	3634	5291	4016	3370	3942	1406	1976163	0,35664098
yc	9030	637	686	480	778	694	722	568	814	376	141080	0,46170383
yc	9036	1893	1915	2776	1742	1928	1675	1490	2092	1089	1187010	0,52068684
yc	10005	640849	598597	508522	634858	591765	708644	514617	659005	112082	12562383063	0,17007764
yc	11033	1964	908	1538	1610	1796	1313	1402	1597	894	798950	0,55967891
yc	11043	1193	1060	2338	1121	876	1176	1071	1256	502	252382	0,40002246
yc	11051	531	529	657	499	492	407	402	547	268	71862	0,48969898
yc	12010	11388	11669	13003	7960	9149	7172	10416	10540	2789	7781215	0,26464763
yc	12029	2896	3449	5113	3498	2512	2883	2473	3320	697	486171	0,2100396
yc	13017	16099	12172	10933	27965	11965	19932	13398	15889	4752	22584618	0,29909911
yc	13032	2994	2660	4360	3663	1847	2714	2898	3340	1241	1541111	0,37163723
yc	13041	4045	3617	3143	3898	2612	3257	6071	4371	1578	2490988	0,36105541
yc	13050	6060	4310	3431	5808	2647	6257	5411	5753	3606	12999847	0,6267649

Σημειώσεις
 yd= άμεσος εκτιμητής, ys=συνθετικός εκτιμητής, yc=σύνθετος εκτιμητής, realdB=αληθής μέτρηση στον πληθυσμό
 sxx=xx-οστό δείγμα, όπου xx=00, 01, ..., 99
 mean=μέσος stdev=τυπική απόκλιση mse=μέσο τετραγωνικό σφάλμα cv=συντελεστής μεταβλητότητας

nomos	νομός
1009	ΔΡΑΜΑ
1011	ΕΒΡΟΣ
1021	ΚΑΒΑΛΑ
1037	ΞΑΝΘΗ
1042	ΡΟΔΟΠΗ
2016	ΗΜΑΘΕΙΑ
2019	ΘΕΣΣΑΛΟΝΙΚΗ
2026	ΚΙΛΚΙΣ
2038	ΠΕΛΛΑ
2039	ΠΙΕΡΙΑ
2044	ΣΕΡΡΕΣ
2049	ΧΑΛΚΙΔΙΚΗ
3008	ΓΡΕΒΕΝΑ
3023	ΚΑΣΤΟΡΙΑ
3027	ΚΟΖΑΝΗ
3047	ΦΛΩΡΙΝΑ
4004	ΑΡΤΑ
4018	ΘΕΣΠΡΩΤΙΑ
4020	ΙΩΑΝΝΙΝΑ
4040	ΠΡΕΒΕΖΑ
5022	ΚΑΡΔΙΤΣΑ
5031	ΛΑΡΙΣΑ
5035	ΜΑΓΝΗΣΙΑ
5045	ΤΡΙΚΑΛΑ
6014	ΖΑΚΥΝΘΟΣ
6024	ΚΕΡΚΥΡΑ
6025	ΚΕΦΑΛΛΗΝΙΑ
6034	ΛΕΥΚΑΣ
7007	ΒΟΙΩΤΙΑ
7012	ΕΥΒΟΙΑ
7013	ΕΥΡΥΤΑΝΙΑ
7046	ΦΘΙΩΤΗΣ
7048	ΦΩΚΙΣ
8001	ΑΙΤΩΛΟΑΚΑΡΝΑΝΙΑ
8006	ΑΧΑΪΑ
8015	ΗΛΕΙΑ
9002	ΑΡΓΟΛΙΣ
9003	ΑΡΚΑΔΙΑ
9028	ΚΟΡΙΝΘΙΑ
9030	ΛΑΚΩΝΙΑ
9036	ΜΕΣΣΗΝΙΑ
10005	ΑΤΤΙΚΗ
11033	ΛΕΣΒΟΣ
11043	ΣΑΜΟΣ
11051	ΧΙΟΣ
12010	ΔΩΔΕΚΑΝΗΣΣΑ
12029	ΚΥΚΛΑΔΕΣ
13017	ΗΡΑΚΛΕΙΟ
13032	ΛΑΣΙΘΙ
13041	ΡΕΘΥΜΝΟ
13050	ΧΑΝΙΑ

ΒΙΒΛΙΟΓΡΑΦΙΑ

Kordos, J. (Editor), (2000-2004), *Statistics in transition, Journal of the Polish Statistical Association*

Cochran, W.G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.

G.C Tikkiwal & A.Ghiya, (2004), " A generalized class of composite estimators with application for small domains", *Statistics in Transition, Journal of the Polish Statistical Association*, Vol.6, No.5,p 697

Gonzales, M.E. and Hoza, C. (1978). *Small-Area Estimation with Application to Unemployment and Housing Estimates. Journal of the American Statistical Association*, 73, 7-15.

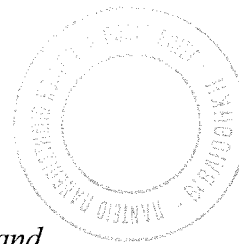
Levy, P.S. (1979). *Small-area estimation- Synthetic and other procedures, 1968-1978, Proceedings of the Workshop on the Synthetic Estimates for Small Areas, National Institute of Drug Abuse (NIDA), Research Monography 24, U.S. Government Printing Office, Washington, D.C.*

National Center for Health Statistics (1968), *Synthetic State Estimates of Disability. (PHS Publication No.1759), Washington, D.C.:U.S. Government Printing Office.*

National Institute on Drug Abuse (1979). *Synthetic Estimates for Small Areas. (NIDA Research Monograph 24), Washington, D.C.: U.S. Government Printing Office.*

Office for National Statistics (ONS)

Platek, R., R, J.N.K., Sarndal, C.E., and Singh, M.P. (Editors) (1987). *Small Area Statistics: An International Symposium. John Wiley and Sons: New York.*



Rao J.N.K., (2003), *Small area estimation*, Wiley-Interscience.

Rao, J.N.K. and Choudry, G. H. (1995). *Small Area Estimation: Overview and Empirical Study*, in B.G. Cox et al (editors), *Business Survey Methods*, John Wiley & Sons, New York, 527-542.

Shaible, W.L. (1979). *A composite estimator for small area statistics*, in *Synthetic Estimates for Small Areas*. (NIDA Research Monograph 24), Washington, D.C.: U.S. Government Printing Office.

Statistical Working Paper: Indirect Estimators in the U.S Federal Program, (1993), Office of Management and Budget

Suciu, G. (2001). *Uninsured Estimates by County: A Review of Options and Issues*, Ohio Department of Health.

ΕΣΥΕ, Έρευνα Εργατικού Δυναμικού, σειρά ετών

Καλαματιανού Γ. Αγλαΐα, (2003, *Κοινωνική στατιστική: μέθοδοι μονοδιάστατης ανάλυσης*, Εκδ. Παπαζήσης

Χρήστου Κ. Γεώργιος, (2002), *Εισαγωγή στην Οικονομετρία*, Εκδ. Gutenberg

