

ΑΡΙΣΤΟΠΟΙΗΣΗ ΤΟΥ ΣΧΕΔΙΟΥ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ ΓΙΑ ΤΗΝ ΔΕΙΓΜΑΤΟΛΗΠΤΙΚΗ ΕΚΤΙΜΗΣΗ ΤΗΣ ΑΞΙΑΣ ΤΩΝ ΟΦΕΙΛΟΜΕΝΩΝ ΕΙΣΦΟΡΩΝ ΕΝΟΣ ΟΡΓΑΝΙΣΜΟΥ ΚΟΙΝΩΝΙΚΗΣ ΑΣΦΑΛΙΣΕΩΣ

Υπό

*Βασιλείου Μπένου**

Πανεπιστήμιο Πειραιώς

Abstract

During the study of the financial position of a Social Security Organization, the Method of Census is basically applied. Nevertheless, there are values which demand a large amount of expenses as well as of time in order to be verified with the census. This need has led us to the investigation of the possibility to apply the Sampling Method for the estimation, at a certain confidence level, of the amount of these values needed for the Actuarial Balance. (JEL H55)

Περίληψη

Κατά την κατάρτιση αναλογιστικής μελέτης για τον έλεγχο της Οικονομικής κατάστασης ενός Οργανισμού Κοινωνικής Ασφάλισης, βασικά εφαρμόζεται η μέθοδος της Απογραφής. Υπάρχουν όμως αξίες οι οποίες για να διαπιστωθούν απογραφικά απαιτείται πολύ μεγάλο κόστος αλλά και χρόνος. Η ανάγκη αυτή μας οδήγησε στην διερεύνηση της δυνατότητας εφαρμογής της δειγματοληπτικής μεθόδου για την εκτίμηση με συγκεκριμένο βαθμό εμπιστοσύνης του ύψους αυτών των αξιών, οι οποίες είναι απαραίτητες για την κατάρτιση του Αναλογιστικού Ισοζυγίου. (JEL H55))

1. Εισαγωγή - Παραδοχές

Από την ανάλυση που θα ακολουθήσει, πιστεύουμε ότι επιτυγχάνεται η ανωτέρω επιδίωξη και μάλιστα με την μέθοδο της στρωματοποιημένης δειγμα-

* Ευχαριστώ τους ανώνυμους κριτές για τις χρήσιμες παρατηρήσεις και διορθώσεις.

τοληψίας, η οποία οδηγεί σε αριστοποίηση του μεγέθους του δείγματος αφ' ενός αλλά και του βαθμού αξιοπιστίας των εκτιμouμένων αξιών.

Πράγματι το μέγεθος του δείγματος, που απαιτείται από τον αναλογιστή μπορεί να μειωθεί σημαντικά από μία έξυπνη χρησιμοποίηση της στρωματοποίησης. Ανεπιφύλακτα δε θα μπορούσαμε να υποστηρίξουμε ότι η χρησιμοποίηση των μεθόδων της δειγματοληψίας δεν θα παρουσίαζε ενδιαφέρον, εάν δεν χρησιμοποιείτο η μέθοδος της στρωματοποίησης.

Είναι γνωστό ότι η τυπική απόκλιση είναι ένας πολύ σπουδαίος παράγων υπολογισμού του μεγέθους του δείγματος. Με δεδομένα το μέγεθος του δειγματοληπτούμενου πληθυσμού και τον βαθμό εμπιστοσύνης, το μέγεθος το δείγματος, καθορίζεται από τον λόγο: (όριο του σφάλματος ακρίβειας) / (τυπική απόκλιση)

Το όριο του σφάλματος ακρίβειας καθορίζει την απαιτούμενη ακρίβεια μιας δειγματικής εκτιμήσεως και η τιμή του εξαρτάται από τον εκτιμouμένο μέσο του πληθυσμού. Επίσης, ο λόγος: (τυπική απόκλιση / Μέσος), είναι πολύ σημαντικός στην δειγματοληψία, και ονομάζεται «συντελεστής μεταβλητικότητας».

Εμπειρικά έχει διαπιστωθεί, όταν ο συντελεστής μεταβλητικότητας είναι μικρότερος του 0,50, είναι ενδεχόμενο τα δείγματα που προκύπτουν με τον κλασικό τρόπο υπολογισμού (απλή τυχαία δειγματοληψία) να θεωρούνται επαρκή για τον έλεγχο και εκτίμηση των πληθυσμιακών αξιών, αλλά όταν ο συντελεστής μεταβλητικότητας υπερβαίνει το 0,5 και προσεγγίζει το 1,0 είναι ενδεχόμενο κάποια μορφή στρωματοποίησης να καθίσταται απαραίτητη. Τέλος, όταν ο συντελεστής μεταβλητικότητας υπερβαίνει το 1,0 η στρωματοποίηση είναι επιβεβλημένη¹.

2. Μία Θεωρητική Αντιμετώπιση του Άριστου Μεγέθους του Δείγματος

Κατά τον καθορισμό του μεγέθους του δείγματος επικρατεί βασικά η αρχή ότι «όσο μεγαλύτερο είναι το μέγεθος του δείγματος τόσο μεγαλύτερη ακρίβεια των επαγωγικών συμπερασμάτων εκ του δείγματος είναι πιθανόν να επιτευχθεί: «πάντως, στην περίπτωση της δειγματοληψίας, δεν πρέπει να μας διαφεύγει ότι ισχύει ο νόμος της μειώσεως των αποδόσεων»². Δηλαδή, μετά από ένα δεδομένο σημείο η αξία της οριακής πληροφορίας που παράγεται από την επόμενη δειγματοληπτούμενη μονάδα είναι μικρότερη από το οριακό κό-

στος δειγματοληψίας αυτής της μονάδας. Στο σημείο αυτό θα πρέπει να σταματήσει η δειγματοληψία, διότι έχει επιτευχθεί το άριστο μέγεθος του δείγματος. Αυτό απεικονίζεται στο σχήμα 1. Στο μέγεθος δείγματος n^* , όπου η μείωση της οριακής αξίας της n -μονάδας δειγματοληψίας ισούται με το σταθερό οριακό κόστος λήψεως της μονάδας, προσδιορίζουμε το άριστο μέγεθος του δείγματος.

Ενώ είναι εύκολο να προσδιορίσουμε ένα θεωρητικά άριστο μέγεθος δείγματος είναι μάλλον πολύ δύσκολο να διαπιστωθεί το αντίστοιχο και στην πράξη. Το οριακό κόστος λήψεως μιας επιπρόσθετης μονάδας του δείγματος είναι σχετικά εύκολο να υπολογισθεί, αλλά το πρόβλημα είναι πώς θα υπολογίσουμε την μείωση της οριακής αξίας. Αυτό πράγματι είναι δύσκολο να υπολογισθεί με ακρίβεια.

3. Το Άριστο Σχέδιο Στρωματοποίησης του Υπό Έρευνα Πληθυσμού

Επειδή το μέγεθος του δείγματος εξαρτάται από την τιμή της τυπικής απόκλισης των στρωμάτων, το άριστο σχέδιο στρωματοποίησης είναι αυτό που ελαχιστοποιεί την τυπική απόκλιση του στρώματος. Δηλαδή, μειώνει την μεταβλητικότητα των τιμών μέσα στο στρώμα στο ελάχιστο. Όσο περισσότερα στρώματα δημιουργούνται τόσο μεγαλύτερη ομοιογένεια εξασφαλίζεται σε κάθε στρώμα και επομένως θα είναι χαμηλότερη η μέση τυπική απόκλιση όλων των στρωμάτων.

Πάντως, το κόστος της δειγματοληψίας αυξάνει καθώς αυξάνει ο αριθμός των στρωμάτων επειδή ο σχεδιάζων την δειγματοληψία (εν προκειμένω ο αναλογιστής) δεν μπορεί να πάρει ένα απλό «άνευ ορίου στο μέγεθος» δείγμα αλλά πρέπει να πάρει $< n >$ μονάδες από κάθε στρώμα, και κάθε στρώμα στον πληθυσμό δεν μπορεί να προσδιοριστεί εκ των προτέρων από τον αναλογιστή.

Υπάρχει έτσι μία διακριτική ευχέρεια, επιλογής μικρού κόστους μεγάλων δειγμάτων από σχετικά μη-διαφοροποιημένους πληθυσμούς ή επιλογής μεγάλου κόστους μικρών δειγμάτων από υψηλά διαφοροποιημένους πληθυσμούς.

Ο Αναλογιστής πιθανώς θα καταλήξει στον διαχωρισμό του πληθυσμού σε δύο ή τρία στρώματα.

Μετά τον καθορισμό του αριθμού των στρωμάτων πρέπει να αποφασίζει το εύρος των στρωμάτων.

Για παράδειγμα εάν ένας πληθυσμός κυμαίνεται από 1ν.μ. έως 100.000 ν.μ., πρέπει να αποφασίσει το εύρος των στρωμάτων, ως εξής: (βλέπε περιπτώσεις πίνακα 1).

Η αποδοχή του συγκεκριμένου σχήματος, εξαρτάται από την κατανομή της αξίας στον πληθυσμό. Εάν είναι διαθέσιμη μία καμπύλη Lorentz (ή ισοδύναμα ο αντίστοιχος πίνακας) για ένα προηγούμενο έτος, βοηθά στον καθορισμό του εύρους των στρωμάτων ώστε να επιτυγχάνεται περίπου ένα ίσο ποσοστό, της συνολικής αξίας, σε κάθε στρώμα. Πάντως, δεν πρέπει να διαφεύγει της προσοχής μας ότι αντικειμενικός σκοπός είναι να ελαχιστοποιήσουμε την μεταβλητικότητα μέσα σε κάθε στρώμα και η στρατηγική της κατανομής ίσης αξίας μπορεί να μην επιτύχει αυτό.

Όταν ο αριθμός των στρωμάτων και το εύρος αυτών αποφασισθεί, ένας πολύ απλός κανόνας μπορεί να χρησιμοποιηθεί για την κατανομή ενός δεδομένου δείγματος μεταξύ των στρωμάτων, ώστε να επιτευχθεί η μέγιστη αποτελεσματικότητα της δειγματοληψίας³.

Ο κανόνας είναι:

(Αριθμός των μονάδων στο στρώμα) x (Τυπική απόκλιση του στρώματος)

Άθροισμα των ανωτέρω για όλα τα στρώματα

Η κατανομή βάσει αυτού του κανόνα, επιτρέπει στον αναλογιστή να αντλήσει την μέγιστη πληροφόρηση από το δεδομένου μεγέθους συνολικό δείγμα.

Η ανωτέρω μέθοδος είναι σχετικά εύκολο να ακολουθηθεί. Αλλά ο αναλογιστής δεν αποφασίζει αυθαίρετα για το μέγεθος του δείγματος εκ των προτέρων. Συνήθως αποφασίζει για τον βαθμό εμπιστοσύνης και το όριο του σφάλματος ακρίβειας που χρειάζεται στα συμπεράσματά του και βάσει αυτών υπολογίζει το απαιτούμενο μέγεθος του δείγματος. Επομένως, είναι απαραίτητο να υπολογίσει τα όρια ακρίβειας στην εκτίμησή του⁴.

Μία μέθοδος υπολογισμού του απαιτούμενου μεγέθους του δείγματος στην στρωματοποιημένη δειγματοληψία είναι η ακόλουθη:

Υποθέτουμε ότι ένας πληθυσμός οφειλετών, στρωματοποιείται όπως φαίνεται στον πίνακα 2.

Εάν υποθέσουμε ότι δεν στρωματοποιούμε τον πληθυσμό, και θέλουμε να

εκτιμήσουμε την αξία των οφειλών του πληθυσμού μέσα στα όρια ± 20.000 , με βαθμό εμπιστοσύνης 90%, ποιο είναι το απαιτούμενο μέγεθος του δείγματος;

Από τους σχετικούς υπολογισμούς προκύπτει ότι το μέγεθος του δείγματος υπερβαίνει κατά πολύ τις 1000 μονάδες. Γεννάται όμως αμέσως το ερώτημα, ποιο θα είναι το μέγεθος του δείγματος εάν χρησιμοποιήσουμε την στρωματοποιημένη δειγματοληψία;

Κατ' αρχήν θα πρέπει να υπολογίσουμε τον συντελεστή στρωματοποίησης (t), από τον τύπο:

$$t = \frac{\Sigma ns}{N^2 (cp)^2 + \Sigma ns^2}$$

όπου,

t = συντελεστής στρωματοποίησης

n = αριθμός των μονάδων σε κάθε στρώμα

s = τυπική απόκλιση σε κάθε στρώμα

p = απαιτούμενο όριο ακριβείας

N = μέγεθος του πληθυσμού

c = 0,606 για συντελεστή εμπιστοσύνης 90%

c = 0,510 για συντελεστή εμπιστοσύνης 95%

c = 0,388 για συντελεστή εμπιστοσύνης 99%

Επειδή το στρώμα III είναι μικρό αλλά με μεγάλη διασπορά, πρέπει να αποφασισθεί 100% δειγματοληψία. Έτσι αποκλείεται σφάλμα δειγματοληψίας από το στρώμα 3 και επικεντρώνεται το ενδιαφέρον μας στα στρώματα 1 και 2 οπότε, προκύπτει ο πίνακας 3. (χρησιμοποιούμε τον δείκτη h , για να δηλώσουμε το στρώμα: δηλαδή $h = I, II, III$).

Στη συνέχεια, με βάση τα δεδομένα του πίνακα 3 και την τιμή του t , προκύπτει ο πίνακας 4.

Με την στρωματοποίηση του πληθυσμού, μειώσαμε το απαιτούμενο μέγεθος του δείγματος από περισσότερες των 1000 μονάδων, σε 78. Πρέπει όμως να τονίσουμε αμέσως ότι η στρωματοποίηση δεν παράγει πάντοτε μία μείωση στο μέγεθος του δείγματος τόσο σημαντική, όπως η ανωτέρω. Απλά το παράδειγμα επιδιώκει να παρουσιάσει τι είναι δυνατόν να επιτευχθεί. Η μεγάλη τυπική απόκλιση 360 ήταν αποτέλεσμα της μεγάλης διακύμανσης του στρώματος III. Με την συμμετοχή ολόκληρου του στρώματος III στο δείγμα εξαλείφουμε όλο το σφάλμα δειγματοληψίας από αυτήν την πηγή και μειώνουμε την τυπική

απόκλιση για τον εναπομείναντα πληθυσμό περίπου στο 25. Αυτό προκάλεσε και την σημαντική μείωση στο μέγεθος του δείγματος.

Συμπεραίνουμε λοιπόν ότι η στρωματοποίηση είναι ένα μέσον μεγάλης σπουδαιότητας, για τον Αναλογιστή που επιθυμεί να χρησιμοποιήσει Στατιστική δειγματοληψία για τον έλεγχο και την εκτίμηση αξιών⁵.

Ένα σημαντικό πρόβλημα που προκύπτει κατά την στρωματοποίηση των οφειλομένων αξιών είναι ο καθορισμός των ορίων των στρωμάτων. Πιστεύουμε ότι για τον σκοπό αυτό σημαντικά βοηθάει η καμπύλη του Lorentz, με την οποία μπορούμε να εκτιμήσουμε την κατανομή της υπό έρευνα αξίας μεταξύ των μονάδων του πληθυσμού.

Η καμπύλη του σχήματος 2, απεικονίζει το αθροιστικό ποσοστό της αξίας που οφείλεται από το αντίστοιχο ποσοστό οφειλετών. Η καμπύλη του Lorentz είναι χρήσιμη για την επιλογή του εύρους των στρωμάτων. Εάν επομένως δεν υπάρχει συμπληρωματική πληροφόρηση, μία επιλογή των ορίων των στρωμάτων θα μπορούσε να προκύψει με αντιστοίχιση του $100/k\%$ της αξίας σε κάθε στρώμα, όπου k ο αριθμός των στρωμάτων. Τα δεδομένα που αφορούν στο σχήμα 2, περιέχονται στον πίνακα 5.

Μία λογική στρωματοποίηση θα καταλήξει σε τρία στρώματα, το πρώτο στρώμα θα περιέχει το πρώτο 70% των οφειλετών, το δεύτερο στρώμα θα περιέχει το επόμενο 20% των οφειλετών και το τρίτο στρώμα θα περιέχει το υπόλοιπο 10% των οφειλετών. Με αυτή την διαδικασία κάθε στρώμα περιέχει περίπου το $1/3$ της συνολικής αξίας των οφειλών.

4. Συμπεράσματα

Το συμπέρασμα είναι ότι η μέθοδος της στρωματοποιημένης δειγματοληψίας, μπορεί να χρησιμοποιηθεί με υψηλό βαθμό αξιοπιστίας αλλά και σημαντικό περιορισμό του συνολικού κόστους, για την εκτίμηση του μεγέθους των οφειλωμένων εισφορών ενός Οργανισμού Κοινωνικής Ασφάλισης.

Παράρτημα

ΠΙΝΑΚΑΣ 1
Διάφορες Περιπτώσεις Εύρους των Στρωμάτων

Στρώμα	(α) Αξία	(β) Αξία	(γ) Αξία
I	0 - 99	0 - 999	0 - 9999
II	100 - 499	1000 - 9999	10.000 - 49999
III	500 και άνω	10.000 και άνω	50.000 και άνω

ΠΙΝΑΚΑΣ 2
Πληθυσμός Οφειλετών Στρωματοποιημένος

Στρώμα	Αξία οφειλών	Αριθμός οφειλετών	Ποσοστό %	Εκτίμηση της μέσης αξίας οφειλών του στρώματος	Εκτίμηση της τυπι- κής απόκλισης του στρώματος
I	0 - 200	1450	72,5	100	20
II	201 - 1000	500	25,0	500	70
III	1000 και άνω	50	2,5	2000	1000
Σύνολο		2000	100,0	250	360

ΠΙΝΑΚΑΣ 3
Πίνακας Υπολογισμών Βάσει των Δεδομένων του Πίνακα 2

Στρώμα	Αριθμός μονάδων N_h	Τυπική απόκλιση S_h	$N_h \times S_h$	$N_h \times S_h^2$
I	1450	20	29.000	580.000
II	500	70	35.000	2.450.000
	1950		64.000	3.030.000

όπου,

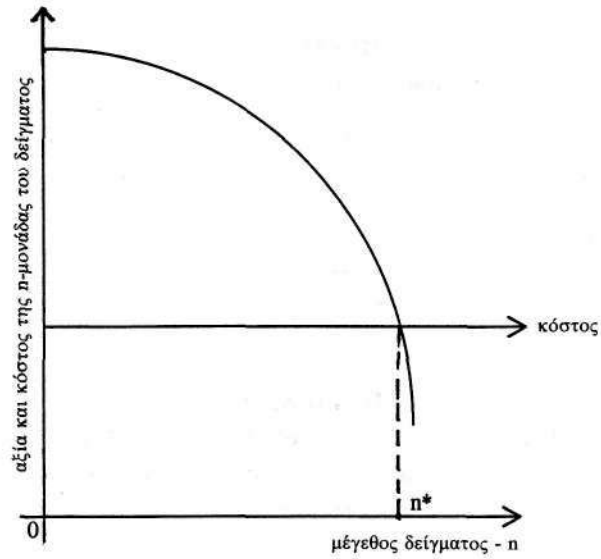
$$t = \frac{64.000}{3.802.500 \times (0.606 \times 10)^2 + 3.030.000} = 0.000447$$

ΠΙΝΑΚΑΣ 4
Υπολογισμός του Συνολικού Δείγματος -n

Στρώμα	N_h	S_h	t^*	n_h
I	1450	20	0,000447	13
II	500	70	»	15
III	100%δειγματοληψία		-	50
Συνολικό μέγεθος δείγματος : n = 78				

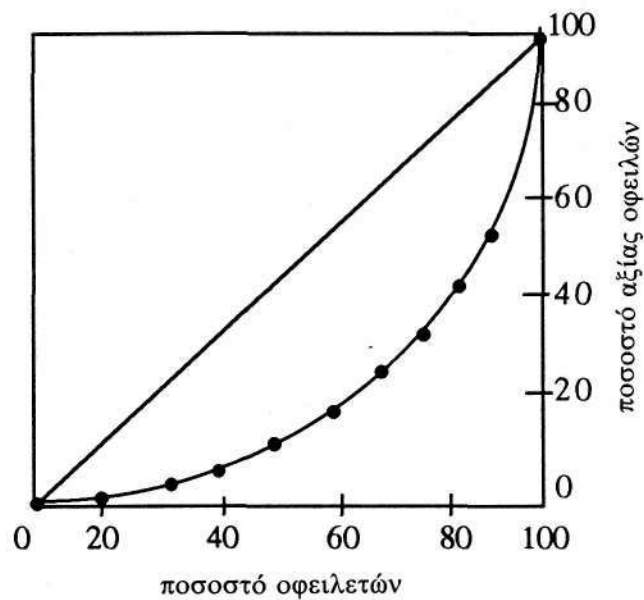
ΠΙΝΑΚΑΣ 5
Αθροιστική Κατανομή των Οφειλετών και Αντίστοιχη Αθροιστική Κατανομή της Αξίας των Οφειλομένων Εισφορών

Αθροιστικό ποσοστό οφειλετών	Αθροιστικό ποσοστό αξίας οφειλών
10	1
20	4
30	9
40	14
50	20
60	28
70	38
80	49
90	65
100	100



Σχήμα 1

Μία Θεωρητική Προσέγγιση του Άριστου Μεγέθους του Δείγματος. Στο Σημείο όπου η Μειούμενη Οριακή Αξία της n-μονάδας Δειγματοληψίας Ισούται με το Σταθερό Οριακό Κόστος Λήψεως Αυτής της Μονάδας, έχουμε το Άριστο Μέγεθος Δείγματος n^* .



Σχήμα 2

Καμπύλη του Lorenz, Κατανομή της Συνολικής Αξίας Οφειλών Μεταξύ των Οφειλετών.

Υποσημειώσεις

1. Slomin, M.J. Guide to Sampling, Pan Piper, London (1968). (An elementary introduction to the basic principles of Sampling). — Stuart, A. Basic ideas of Scientific Sampling, Griffin, London (1962). (Never was so such information about the basic principles of sampling packed into so short a space. Non-mathematical but quite tough reading). — Yates, F. Sampling Methods for Censuses and Surveys, Griffin, London (1960). (A most comprehensive introduction to the practical problems of designing sample surveys).
2. Trueblood, R. M. and Cyert, R. M. Sampling Techniques in Accounting, Prentice Hall, Englewood Cliffs, New Jersey (1957). (An early book, but clearly written with some good examples). — Cyert, R. M., and Davidson, H. Justin. Statistical Sampling for Accounting Information, Prentice Hall, Englewood Cliffs, New Jersey (1962). (Strong emphasis on theory of Sampling).
3. Μπένου Β., Μέθοδοι και Τεχνικές Δειγματοληψίας, Εκδόσεις Α. Σταμούλης, Πειραιάς **1991**.
4. Barnett, V., Sample Survey Principles and Methods, Edward Arnold, London (1991).
5. The new business statistics. H. V. Roberts, Journal of Business, January 1960, 29. — Statistical sampling for auditors. A new look. W. H. Kraft. Journal of Accountancy, 126. August 1968, 49-56.

Βιβλιογραφία

- Anderson, P. H.* (1947) "Distribution in stratified Sampling". Ann. Math. Statist. 13, 42-52.
- Amitage, P.*, (1947). "A comparison of stratified with understricted random sampling form a finite population". Biometrika, 34, 273-280.
- Barnett, V.* (1982). "Comparative Statistical Inference", John Wiley and Sons Inc., New York.
- Cochram, W. G.* (1977). "Sampling Techniques". John Wiley and Sons Inc., New York.
- Deming, W. E.* (1950). "Some Theory of Sampling". John Wiley and Sons Inc., New York.
- Deming, W. E.*, (1960). "Sample design in business research". John Wiley and Sons Inc., New York.
- Frank, Yates,* (1981). "Sampling Methods for Censuses and Surveys". Charles Griffin and Company LTD, London.
- Groves, R. M.* (1989). "Survey Errors and Survey Costs". John Wiley and Sons Inc., New York.
- Kish, Leslie,* (1965). "Survey Sampling". John Wiley and Sons Inc., New York.
- Konijn, H. S.*, (1973). "Statistical Theory of Sample Survey Design and Analysis". American Elsevier Publishing Company Inc., New York.
- Williams, Bill,* (1978). "A Sampler on Sampling". John Wiley and Sons Inc., New York.