ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΟΜΑΔΩΝ ΣΤΑΤΙΣΤΙΚΟΥ ΠΛΗΘΥΣΜΟΥ

Ὑπὸ ΣΑΡΑΝΤΟΥ Κ. ΔΗΜΑ

Ἡ περὶ δειγματοληπτικῶν στατιστικῶν ἐρευνῶν θεωρία ἀναφέρεται, κατὰ τὸ μεγαλύτερον αὐτῆς μέρος εἰς τὸ πρόβλημα τῆς ἐκτιμήσεως παραμέτρων τινῶν ἑνὸς μόνον εἰδικοῦ χαρακτηριστικοῦ τοῦ ἐρευνωμένου πληθυσμοῦ : Αὕτη εἶναι ἡ περίπτωσις τῆς λεγομένης «Μονοπαραμετρικῆς Δειγματοληψίας». Εἰς τὴν πρᾶξιν, ὅμως, αἱ δειγματοληπτικαὶ ἔρευναι σπανίως σχεδιάζονται καὶ ἐκτελοῦνται πρὸς τὸν σκοπὸν καὶ μόνον ἐκτιμήσεως ἑνὸς χαρακτηριστικοῦ. Τοῦτο ἄλλωστε θὰ ἀνεβίβαζε τὸ κόστος τῆς ἐρεύνης ἀνὰ ἐκτιμώμενον χαρακτηριστικόν. Ἡ περίπτωσις τῆς συγχρόνου ἐκτιμήσεως πλειόνων χαρακτηριστικῶν εἶναι συνήθης κατὰ τὴν διενέργειαν δειγματοληπτικῆς τινος ἐρεύνης, ὅτε ἔχομεν τὴν περίπτωσιν τῆς «Πολυπαραμετρικῆς Δειγματοληψίας». Ἀλλὰ τὸ γεγονὸς τοῦτο δημιουργεῖ σοβαρὰς δυσχερείας εἰς τὴν ἀντιμετώπισιν τῶν σχετικῶν προβλημάτων. Ἡ χρησιμοποίησις, λ.χ., μεθόδου τινός, διὰ τῆς ὁποίας ἐπιτυγχάνεται ἐπαρκὴς ἀκρίβεια εἰς τὴν ἐκτίμησιν ἑνὸς χαρακτηριστικοῦ, πιθανὸν νὰ μὴ θεωρῆται κατάλληλος διὰ τὴν ἐκτίμησιν ἑτέρου χαρακτηριστικοῦ, καὶ ἀντιστρόφως.

Ἐκτὸς τῶν ἀνωτέρω, αἱ ἀπαιτήσεις μας κατὰ τὴν διενέργειαν δειγματοληπτικῆς τινος ἐρεύνης δὲν περιορίζονται μόνον εἰς τὴν ἐκτίμησιν τῶν παραμέτρων ἑνὸς ἢ πλειόνων χαρακτηριστικῶν ἀναφερομένων εἰς ὁλόκληρον τὸν ὑπὸ ἔρευναν πληθυσμόν. Συχνὰ τὸ ἐνδιαφέρον τοῦ ἐρευνητοῦ προεκτείνεται καὶ εἰς ἐκτιμήσεις διὰ διαφόρους ὑποδιαιρέσεις τοῦ πληθυσμοῦ ἢ τμήματα αὐτοῦ, τὰς καλουμένας εἰς τὸ ἑξῆς «ὁμάδας» (Domains of Study) [1]. Ἐπὶ παραδείγματι, εἰς μίαν ἔρευναν «νοικοκυριῶν», ἡ ὁποία ἐσχεδιάσθη διὰ τὴν ἐκτίμησιν τοῦ μέσου ἢ καὶ τοῦ συνολικοῦ εἰσοδήματος ἢ τῶν δαπανῶν αὐτῶν, ἐνδιαφερόμεθα νὰ ἐκτιμήσωμεν τὰ χαρακτηριστικὰ ταῦτα, τόσον διὰ τὸ σύνολον τῶν «νοικοκυριῶν», ὅσον καὶ διὰ ἐπὶ μέρους κατηγορίας αὐτῶν. Λ.χ., διὰ τὰς ὁμάδας τῶν «νοικοκυριῶν», αἱ ὁποῖαι ἔχουν 0, 1, 2, 3 κ.λ.π. τέκνα ἢ διὰ τὰς ὁμάδας τῶν «νοικοκυριῶν», τὰ ὁποῖα διαμένουν εἰς ἰδιοκτήτους οἰκίας κ.λ.π. Ἐπίσης, εἰς μίαν ἔρευναν τοῦ συνολικοῦ πληθυσμοῦ τῆς Χώρας, εἶναι πολὺ πιθανὸν ν' ἀναζητῆται ἡ ἐκτίμησις παραμέτρων τινῶν διὰ τμήματα αὐτοῦ, ἀποτελούμενα ἀπὸ ἄτομα συγκεκριμένων ὁμάδων ἡλικιῶν ἢ ὡρισμένων κοινωνικῶν τάξεων.

---

1) Ὁ ὅρος Domains of Study κατὰ πρῶτον εἰσήχθη ὑπὸ τοῦ Στατιστικοῦ Γραφείου τῶν Ἡνωμένων Ἐθνῶν τὸ 1949 καὶ σημαίνει ὑποδιαιρέσεις τοῦ πληθυσμοῦ, διὰ τὰς ὁποίας ἡ ἔρευνα σχεδιάζεται καὶ ἐκτελεῖται πρὸς ἀπόκτησιν ἀριθμητικῶν πληροφοριῶν.

Κατὰ τὸν σχεδιασμὸν καὶ τὴν ἐκτέλεσιν μιᾶς τοιαύτης δειγματοληπτικῆς ἐρεύνης θὰ πρέπει νὰ καταβληθῇ προσπάθεια, ὥστε νὰ εἶναι δυνατὴ ἡ ἐκτίμησις, μὲ ἐπαρκῆ ἀκρίβειαν, τῶν χαρακτηριστικῶν, τόσον τοῦ συνόλου τοῦ πληθυσμοῦ, ὅσον καὶ τῶν ἐπὶ μέρους ὑποδιαιρέσεων, διὰ τὰς ὁποίας ὑπάρχει σχετικὸν ἐνδιαφέρον.

Εἰς τὴν πρᾶξιν ἔχει διαπιστωθῆ, ὅτι ἓν μέγεθος συνολικοῦ δείγματος, ἐκ τοῦ ὁποίου ἐπιτυγχάνονται ἱκανοποιητικαὶ ἐκτιμήσεις διὰ παραμέτρους ὑποδιαιρέσεων τοῦ πληθυσμοῦ, εἶναι πλέον ἢ ἐπαρκὲς διὰ τὴν ἐκτίμησιν τῶν αὐτῶν παραμέτρων, ὅταν ἀναφερώμεθα εἰς τὸ σύνολον τοῦ πληθυσμοῦ. Τὸ ἀντίστροφον, ἐν τούτοις, δὲν ἰσχύει. Ἤτοι δεῖγμα, διὰ τοῦ ὁποίου γίνεται ἐκτίμησις χαρακτηριστικῶν τοῦ συνολικοῦ πληθυσμοῦ μεθ' ὡρισμένης ἀκριβείας, δίδει κατὰ κανόνα ὀλιγώτερον ἀκριβεῖς ἐκτιμήσεις δι' ἀντίστοιχα χαρακτηριστικὰ μιᾶς ὑποδιαιρέσεως αὐτοῦ.

Τὸ πρόβλημα τοῦτο κατὰ τὸ στάδιον τοῦ σχεδιασμοῦ τῆς ἐρεύνης ἀντιμετωπίζεται διὰ διαφόρων μεθόδων, ὡς αἱ ἀκόλουθοι τρεῖς : α) Διὰ τῆς αὐξήσεως τοῦ μεγέθους τοῦ συνολικοῦ δείγματος. β) Διὰ τῆς χρησιμοποιήσεως εἰδικῶν μεθόδων ἐπιλογῆς, ὡς, ἑ.π., διὰ τῆς ἐπιλογῆς δείγματος ἐξ ἑκάστης ὑπὸ σπουδὴν ὁμάδος κεχωρισμένως (κατὰ στρώματα δειγματοληψία). γ) Διὰ τῆς χρησιμοποιήσεως εἰδικῶν μεθόδων ἐκτιμήσεως, ὡς εἶναι ἡ ἐκτίμησις λόγου (ratio estimate), ἡ ἐκτίμησις παλινδρομήσεως (regression estimate) κλπ.

Τὸ πρόβλημα, πάντως, ἐξακολουθεῖ νὰ παραμένῃ εἰς τὰς περιπτώσεις ἐκείνας, εἰς τὰς ὁποίας ζητοῦνται ἐκτιμήσεις διὰ διαφόρους ὑποδιαιρέσεις τοῦ πληθυσμοῦ, ἐνῷ ἡ κατανομὴ τῶν μονάδων τοῦ πληθυσμοῦ εἰς τὰς ὁμάδας ταύτας εἶναι ἄγνωστος. Προφανῶς, πᾶσα ἐκ τῶν προτέρων στρωματοποίησις τῶν μονάδων τοῦ πληθυσμοῦ εἶναι ἀνέφικτος.

Ἐνταῦθα παρέχεται ἡ θεωρητικὴ ἀντιμετώπισις τοῦ ἀκολούθου θέματος :

«Ἐάν, ἀγνοοῦντες τὴν σύνθεσιν τοῦ συνολικοῦ πληθυσμοῦ ὡς πρὸς τὰς ἐπὶ μέρους ὑπὸ σπουδὴν ὑποδιαιρέσεις (ὁμάδας), ἐπιλέξωμεν ἐξ αὐτοῦ ἁπλοῦν τυχαῖον δεῖγμα, τότε ἐρωτᾶται, ποῖαι αἱ μέθοδοι ἐκτιμήσεως διὰ τὰ χαρακτηριστικὰ τῶν ὁμάδων τούτων καὶ ποία ἡ ἀξιοπιστία τῶν ἐκτιμήσεων τούτων ;».

Ἔστω πληθυσμὸς N μονάδων κατανεμημένων εἰς m ὁμάδας, τάς :

$$G_1, G_2, \ldots, G_m$$

Τὸ πλῆθος τῶν μονάδων ἑκάστης ὁμάδος ἔστω ἀντιστοίχως :

$$N_1, N_2, \ldots, N_m$$

Προφανῶς :

$$\sum_{i=1}^{m} N_i = N$$

Ἐκ τοῦ συνολικοῦ πλήθους τῶν N μονάδων ἐπιλέγομεν, κατὰ τὴν μέθοδον τῆς ἁπλῆς τυχαίας δειγματοληψίας, δεῖγμα n μονάδων.

Τὸ δεῖγμα θὰ ἀποτελῆται ἀπὸ μονάδας ἀνηκούσας εἰς τὰς ἐπὶ μέρους ὁμάδας.

Ἔστω n_i τὸ πλῆθος τῶν μονάδων, αἵτινες εὑρέθησαν ὅτι προέρχονται ἐκ τῆς G_i ὁμάδος. Συνεπῶς :

$$\sum_{i=1}^{m} n_i = n$$

Ἂς ὑποθέσωμεν ἤδη, ὅτι ἐπαναλαμβάνομεν ἀπεριορίστως τὴν διαδικασίαν ἐπιλογῆς εἰς τρόπον, ὥστε τὰ κληρούμενα δείγματα νὰ ἔχουν πάντοτε σταθερὸν μέγεθος n μονάδων. Εἰς τὰ διάφορα διαδοχικὰ δείγματα οἱ ἀριθμοί :

$$n_1,\ n_2,\ \ldots,\ n_m$$

θὰ ἔχουν μὲν σταθερὸν ἄθροισμα ἴσον πρὸς n, ἀλλὰ θὰ εἶναι ἐν γένει μεταβλητοί.

Συγκεκριμένως τὸ πλῆθος n_j τῶν μονάδων δείγματος, αἱ ὁποῖαι ἀνήκουν εἰς τὴν ὁμάδα G_j , ἀκολουθεῖ τὴν ὑπερ-γεωμετρικὴν κατανομήν :

$$\left\{ n^j,\quad P(n_j) = \frac{\binom{N_j}{n_j}\binom{N-N_j}{n-n_j}}{\binom{N}{n}} \right\}$$

$$n_j = 0,\ 1,\ 2,\ \ldots,\ n$$

Οὕτως, εἰς μίαν ἀτέρμονα σειρὰν δειγμάτων — ἑκάστου μεγέθους n — θὰ ὑπάρξουν περιπτώσεις δειγμάτων μὴ περιεχόντων μονάδας ἐκ τῆς G_j , μὲ μίαν μονάδα ἐκ τῆς G_j , μὲ δύο κ.ο.κ., ὡς ἐπίσης περίπτωσις, καθ᾿ ἣν ἅπασαι αἱ μονάδες ἀνήκουν εἰς τὴν ὁμάδα G_j ($n = n_j$). Εἰς μίαν οἱανδήποτε τῶν ὡς ἄνω περιπτώσεων, καθ᾿ ἣν λαμβάνει χώραν ἡ συνθήκη :

$$n_j = c \qquad (c = \text{σταθερά, ὡς π.χ. } n_j = 5),$$

ὅλοι οἱ συνδυασμοὶ τῶν N_j μονάδων (τῆς G_j ὁμάδος) ἀνὰ c ἔχουν τὴν αὐτὴν πιθανότητα νὰ περιέχωνται εἰς τὸ συνολικὸν δεῖγμα τῶν n μονάδων.

Τοῦτο σημαίνει, ὅτι ὑπὸ τὴν συνθήκην $n_j = c$ ἡ ἁπλῆ τυχαία δειγματοληψία n μονάδων ἐκ τῶν N συνεπάγεται καὶ ἁπλῆν τυχαίαν δειγματοληψίαν n_j μονάδων ἐκ τῶν N_j .

Ἂς παραστήσωμεν περαιτέρω διά :

$$A_{j1},\ A_{j2},\ \ldots,\ A_{jN_j}$$

τὰς μονάδας τῆς ὁμάδος G_j καὶ διά :

$$Y_{j1}\,,\; Y_{j2}\,,\; \ldots\,,\; Y_{jN_j}$$

τὰς ἀντιστοίχους τιμὰς ποσοτικοῦ τινος χαρακτηριστικοῦ τῶν μονάδων τούτων.

Αἱ παράμετροι τῆς ὁμάδος ταύτης : σύνολον Y_j, μέσος $\overline{Y}_j$ καὶ διακύμανσις S_j^2, θὰ ὁρίζωνται, ὡς γνωστόν, διὰ τῶν σχέσεων :

$$Y_j = \sum_{k=1}^{N_j} Y_{jk}$$

$$\overline{Y}_j = \frac{1}{N_j}\, Y_j$$

$$S_j^2 = \frac{1}{N_j - 1} \sum_{k=1}^{N_j} (Y_{jk} - \overline{Y}_j)^2$$

Ἀναφερόμεθα ἤδη εἰς τοὺς ἐκτιμητὰς τῶν παραμέτρων τούτων, ὅταν ἡ δειγματοληψία διενεργῆται ὡς ἀνωτέρω ἐξετέθη, ἤτοι ἐπιλέγονται n μονάδες ἐκ τοῦ συνολικοῦ πληθυσμοῦ μεταξὺ τῶν ὁποίων n_j εὑρίσκονται ὅτι προέρχονται ἐκ τῆς ὑπὸ σπουδὴν ὁμάδος G_j.

Ἐκτίμησις τῶν παραμέτρων $\overline{Y}_j$ καὶ S_j^2

Ἃς παραστήσωμεν διὰ y_j τὸ σύνολον (ἄθροισμα) τῶν τιμῶν $y_{j\lambda}$ τῶν μονάδων τῆς ὁμάδος G_j, αἵτινες περιελήφθησαν εἰς τὸ δεῖγμα :

$$y_j = \sum_{\lambda=1}^{n_j} y_{j\lambda}$$

Ὁ λόγος :

$$\overline{y}_j = \frac{y_j}{n_j}$$

τῶν δύο τυχαίων μεταβλητῶν y_j καὶ n_j, λέγομεν ὅτι ἀποτελεῖ ἀμερόληπτον ἐκτίμησιν τῆς παραμέτρου $\overline{y}_j$. Δηλαδή :

$$E(\overline{y}_j) = \overline{Y}_j$$

Πράγματι, ἡ μαθηματικὴ ἐλπὶς τῆς μεταβλητῆς $\overline{y}_j$ δύναται νὰ ἐκφρασθῇ ὡς ἀποτέλεσμα τῆς διαδοχικῆς ἐφαρμογῆς τῶν «τελεστῶν» (operators) E_1 καὶ E_2, ἤτοι :

$$E(\overline{y}_j) = E_1\,[\,E_2(\overline{y}_j)\,]$$

ἔνθα : $E_2(\overline{y}_j)$ ἡ ὑπὸ (τὴν συγκεκριμένην) συνθήκην $n_j = c$ μαθηματικὴ ἐλπὶς καὶ E_1 ἡ μαθηματικὴ ἐλπὶς τῶν τιμῶν E_2, τῶν ἀντιστοιχουσῶν εἰς τὰς διαφόρους τιμὰς n_j (c_1, c_2, c_3, . . .).

Ἀλλὰ ὑπὸ τὴν συνθήκην $n_j = c$ ἰσχύουν αἱ ἰδιότητες τῶν ἐκτιμητῶν τῆς περιπτώσεως ἁπλῆς τυχαίας δειγματοληψίας, ὁπότε:

$$E_2(\bar{y}_j) = E[\bar{y}_j / \text{διὰ } n_j = c] = \bar{Y}_j$$

Ἀλλὰ τότε καί:

$$E(\bar{y}_j) = E_1(\bar{Y}_j) = \bar{Y}_j$$

Καθ' ὅμοιον τρόπον δύναται νὰ ἀποδειχθῇ, ὅτι ἡ ἔκφρασις:

$$\widehat{S}_j^2 = \frac{1}{n_j - 1} \sum_{\lambda=1}^{n_j} (y_{j\lambda} - \bar{y}_j)^2$$

ἀποτελεῖ ἀμερόληπτον ἐκτιμητὴν τῆς παραμέτρου S_j^2.

Ἡ διακύμανσις τοῦ ἐκτιμητοῦ $\bar{y}_j$ προκύπτει ὡς:

$$V(\bar{y}_j) = E_1[V_2(\bar{y}_j)]$$

ἔνθα V_2 ἡ διακύμανσις ὑπὸ τὴν συνθήκην n_j = σταθερόν, ἤτοι:

$$V_2(\bar{y}_j) = V(\bar{y}_j / n_j = \text{σταθερὸν}) = \left(1 - \frac{n_j}{N_j}\right) \frac{S_j^2}{n_j}$$

Συνεπῶς:

$$V(\bar{y}_j) = E_1(V_2) = \left[E\left(\frac{1}{n_j}\right) - \frac{1}{N_j} \right] S_j^2$$

Εἰς τὴν πρᾶξιν ὁ ὑπολογισμὸς τῆς διακυμάνσεως τοῦ ἐκτιμητοῦ $\bar{y}_j$ ἐκ τῶν δεδομένων αὐτοῦ τούτου τοῦ δείγματος δίδεται (κατὰ προσέγγισιν) διὰ τῆς σχέσεως:

$$\widehat{V}(\bar{y}_j) = \left(1 - \frac{n}{N}\right) \frac{\widehat{S}_j^2}{n_j}$$

Ὁ τύπος οὗτος ἔχει τὸ πλεονέκτημα ὅτι ἐν αὐτῷ δὲν ἐμφανίζεται τὸ μέγεθος τῆς ὁμάδας N_j, τὸ ὁποῖον συμβαίνει εἰς πλείστας περιπτώσεις νὰ εἶναι ἄγνωστον.

Ἐκτίμησις τοῦ συνόλου Y_j

Εἰς τὴν προκειμένην περίπτωσιν χρησιμοποιοῦμεν τὴν ἀκόλουθον τεχνικήν: Συμβολίζομεν τὰς μονάδας τοῦ συνολικοῦ πληθυσμοῦ ὡς:

$$A_1, A_2, \ldots, A_i, \ldots, A_N$$

[Προφανῶς N_j ἐκ τούτων ταυτίζονται μὲ τὰς μονάδας

$$A_{j1}\,,\ A_{j2}\,,\ \ldots\,,\ A_{jN_j}$$

τῆς ὁμάδος G_j].

Εἰς τὰς μονάδας $A_i\,(i=1, 2, \ldots, N)$ φανταζόμεθα ὅτι ἀντιστοιχοῦν ἀριθμοὶ Z_j τοιοῦτοι, ὥστε:

$Z_i = 0$ ἐὰν ἡ A_i δὲν ἀνήκει εἰς τὴν ὁμάδα G_j

$Z_i = Y_{jk}$ ἐὰν ἡ A_i ἀνήκῃ εἰς τὴν ὁμάδα G_j δηλαδὴ ταυτίζεται μὲ τὴν μονάδα A_{jk}.

Προφανῶς τὸ σύνολον:

$$Z = \sum_{i=1}^{N} Z_i$$

τῶν τιμῶν τούτων ἰσοῦται μὲ τὸ σύνολον:

$$Y_j = \sum_{k=1}^{N_j} Y_{jk}$$

τῶν τιμῶν τῶν μονάδων τῆς ὁμάδος G_j.

Τὸ αὐτὸ ἰσχύει καὶ διὰ τὸ σύνολον:

$$z = \sum_{\lambda=1}^{n} z_\lambda$$

τῶν τιμῶν z_λ τῶν n μονάδων τοῦ δείγματος καὶ τὸ ὁποῖον ἰσοῦται μὲ τὸ σύνολον:

$$y_j = \sum_{\lambda=1}^{n_j} y_{j\lambda}$$

Κατὰ συνέπειαν ἐκτίμησις τῆς τιμῆς Z σημαίνει κατ᾽ οὐσίαν ἐκτίμησιν τῆς παραμέτρου Y_j.

Ἔχομεν:

$$\widehat{Z} = \frac{N}{n} \sum_{\lambda=1}^{n} z_\lambda$$

καὶ

$$V(\widehat{Z}) = \frac{N(N-n)}{n} S_z^2$$

$$= \frac{N(N-n)}{n(N-1)} \left\{ \sum_{i=1}^{N} Z_i^2 - \frac{Z^2}{N} \right\}$$

Δηλαδή:

$$\widehat{Y}_j = \frac{N}{n} y_j$$

καὶ

$$V(\widehat{Y}_j) = \frac{N(N-n)}{n(N-1)} \left[\sum_{k=1}^{N_j} Y_{jk}^2 - \frac{Y_j^2}{N} \right]$$

$$= \frac{N(N-n)}{n(N-1)} \left[\left(\sum_{k=1}^{N_j} Y_{jk}^2 - N_j \overline{Y}_j^2\right) + \left(N_j \overline{Y}_j^2 - \frac{1}{N} N_j^2 \overline{Y}_j^2 \right) \right]$$

$$= \frac{N(N-n)}{n(N-1)} \left[S_j^2 (N_j - 1) + N_j \overline{Y}_j^2 \left(1 - \frac{N_j}{N} \right) \right]$$

Ἐπειδή:

$$E(\widehat{Z}) = Z$$

ἕπεται ὅτι καὶ ὁ τύπος:

$$E\left(\frac{N}{n} y_j \right) = Y_j .$$

Συνάμα, εὐκόλως ἀποδεικνύεται ὅτι μία ἀμερόληπτος ἐκτίμησις τῆς $V(\widehat{Y}_j)$ ἐκ τῶν δεδομένων τοῦ δείγματος εἶναι ἡ ἔκφρασις:

$$\widehat{V}(\widehat{Y}_j) = \frac{N(N-n)}{n(n-1)} \left[\widehat{S}_j^2 (n_j - 1) + n_j \overline{y}_j^2 \left(1 - \frac{n_j}{n} \right) \right]$$

Τὰ μέχρι τοῦδε ἐκτεθέντα διὰ τὴν ἐκτίμησιν τοῦ συνόλου Y_j καὶ τῆς ἀξιοπιστίας του ἀναφέρονται εἰς τὴν περίπτωσιν ἐκείνην, κατὰ τὴν ὁποίαν τὸ N_j, ἤτοι ἡ κατανομὴ τῶν μονάδων τοῦ πληθυσμοῦ εἰς ἑκάστην ὁμάδα G_j εἶναι ἄγνωστος. Ἐὰν ὅμως εἶναι δυνατὸν ἐκ τῶν προτέρων νὰ γνωρίζωμεν τὴν τιμὴν τοῦ N_j, τότε ὁ ἐκτιμητὴς $\widehat{Y}_j$ δίδεται ὑπὸ τοῦ τύπου:

$$\widehat{Y}_j = N_j \overline{y}_j = \frac{N_j}{n_j} y_j$$

Ἤτοι:

$$E(\widehat{Y}_j) = E_1 E_2 \left(\frac{N_j}{n_j} y_j \,/\, n_j = c \right)$$

$$= E_1 (Y_j)$$

$$= Y_j$$

Ἡ διακύμανσις δὲ τοῦ ἐκτιμητοῦ $\widehat{Y}_j$ εἶναι κατὰ προσέγγισιν:

$$V(\widehat{Y}_j) = \left(\frac{N}{n} - 1 \right) [N_j S_j^2],$$

ἐὰν θεωρηθῇ ὅτι $$E\left(\frac{N_j}{n_j}\right) \simeq \frac{N}{n}.$$

Γεννᾶται, ὅθεν, τὸ ἐρώτημα κατὰ πόσον αὐξάνεται ἡ διακύμανσις τοῦ ἐκτιμητοῦ $\widehat{Y}$, ἐκ τῆς ἀδυναμίας μας νὰ γνωρίζωμεν ἐκ τῶν προτέρων τὸ N_j.

Ἐκ τῆς συγκρίσεως τῶν ἀντιστοίχων διακυμάνσεων ἔχομεν:

$$\frac{V\,(N_j \text{ γνωστὸν})}{V\,(N_j \text{ ἄγνωστον})} = \frac{\left(\frac{N}{n} - 1\right)[N_j S_j^2]}{\frac{N}{N-1}\left(\frac{N}{n} - 1\right)\left[(N_j - 1) S_j^2 + N_j \overline{Y}_j^2 \left(1 - \frac{N_j}{N}\right)\right]}$$

$$= \frac{N_j S_j^2}{N_j S_j^2 + N_j \overline{Y}_j^2 \left(1 - \frac{N_j}{N}\right)}$$

Ἐὰν θεωρηθῇ ὅτι $N \simeq N - 1$ καὶ $N_j \simeq N_j - 1$, τότε:

$$\frac{V\,(N_j \text{ γνωστὸν})}{V\,(N_j \text{ ἄγνωστον})} = \frac{CV^2(Y_{jk})}{CV^2(Y_{jk}) + (1 - w_j)}, \quad \text{ἔνθα} \quad w_j = \frac{N_j}{N}.$$

Συνδιακύμανσις τῶν $\bar{\mathbf{y}}_j$ καὶ $\bar{\mathbf{y}}_i$

Ἐκ τοῦ συνολικοῦ δείγματος τῶν n παρατηρήσεων, αἱ n_j καὶ n_i εὑρέθησαν ἀνήκουσαι εἰς τὰς ὁμάδας G_j καὶ G_i ἀντιστοίχως. Αἱ ἐκτιμήσεις τῶν μέσων $\overline{Y}_j$ καὶ $\overline{Y}_i$ τῶν δύο ὁμάδων ὑπελογίσθησαν κατὰ τὰ ἤδη ἐκτεθέντα:

$$\bar{y}_j = \frac{1}{n_j} y_j$$

$$\bar{y}_i = \frac{1}{n_i} y_i$$

Θὰ δείξωμεν ἤδη, ὅτι ἡ συνδιακύμανσις τῶν δύο τυχαίων μεταβλητῶν $\bar{y}_j$ καὶ $\bar{y}_i$ εἶναι μηδενική:

Ἤτοι: $$\mathrm{Cov}\,(\bar{y}_j, \bar{y}_i) = 0$$

Ἐξ ὁρισμοῦ ἔχομεν:

$$\mathrm{Cov}\,(\bar{y}_j, \bar{y}_i) = E\,(\bar{y}_j\,\bar{y}_i) - E\,(\bar{y}_j)\,E\,(\bar{y}_i)$$

$$= E\,(\bar{y}_j\,\bar{y}_i) - \overline{Y}_j\,\overline{Y}_i$$

Ἤδη τὴν μαθηματικὴν ἐλπίδα τοῦ γινομένου $\bar{y}_j\, \bar{y}_i$ εὑρίσκομεν ἐὰν ἐφαρμόσωμεν διαδοχικῶς τοὺς «τελεστὰς» E_1 καὶ E_2.

Ὡς E_2 θεωροῦμεν ἐνταῦθα τὴν μαθηματικὴν ἐλπίδα τοῦ γινομένου ὑπὸ τὴν ἀκόλουθον συνθήκην: «*Εἰς τὰ διάφορα δείγματα* n *μονάδων ἐκ τῶν* N, *νὰ περικλείηται ἐκ τῆς μονάδος* G_i *ὁ αὐτὸς πάντοτε συνδυασμὸς* n_i *μονάδων*».

Προφανῶς ὑπὸ τὴν συνθήκην ταύτην ἡ τιμὴ $\bar{y}_i$ εἶναι σταθερὰ καὶ συνεπῶς:

$$E_2(\bar{y}_j\, \bar{y}_i) = \bar{y}_i\, E(\bar{y}_j) = \bar{y}_i\, \bar{Y}_j\,.$$

Κατὰ συνέπειαν:

$$\begin{aligned} E(\bar{y}_j\, \bar{y}_i) &= E_1\,[\,E_2(\bar{y}_j\, \bar{y}_i)\,] \\ &= E_1\,[\,\bar{y}_i\ \ \bar{Y}_j\,] \\ &= \bar{Y}_j\, E(\bar{y}_i) \\ &= \bar{Y}_j\, \bar{Y}_i \end{aligned}$$

καί:

$$\mathrm{Cov}\,(\bar{y}_j\,,\, \bar{y}_i) = 0$$

Ἄμεσος συνέπεια τῆς ἰδιότητος ταύτης εἶναι ὅτι:

$$V(\bar{y}_j - \bar{y}_i) = V(\bar{y}_j) + V(\bar{y}_i)$$

Ἐὰν κατ' ἀκολουθίαν ἐπιθυμοῦμεν, ἐπὶ τῇ βάσει τῶν δύο ἐκτιμήσεων, νὰ ἐλέγξωμεν τὴν ὑπόθεσιν:

$$H_0 : \bar{Y}_j - \bar{Y}_i = 0,$$

ἀρκεῖ νὰ ὑπολογίσωμεν τὸν λόγον:

$$t = \frac{|\,\bar{y}_j - \bar{y}_i\,|}{\sqrt{V(\bar{y}_j) + V(\bar{y}_i)}}\,.$$

Συνδιακύμανσις τῶν $\hat{Y}_j$ καὶ $\hat{Y}_i$

Ἔχομεν:

$$\mathrm{Cov}\,(\hat{Y}_j\,,\, \hat{Y}_i) = E(\hat{Y}_j\ \hat{Y}_i) - Y_j\, Y_i$$

καὶ

$$E(\hat{Y}_j\ \hat{Y}_i) = E\left(\frac{N^2}{n^2}\, y_j\, y_i\right) = \frac{N^2}{n^2}\, E(y_j\, y_i)$$

Ἐν προκειμένῳ θὰ εὕρωμεν πάλιν τὴν μαθηματικὴν ἐλπίδα τοῦ γινομένου

$y_j\, y_i$, ἐφαρμόζοντες διαδοχικῶς τοὺς «τελεστὰς» E_1, E_2 καὶ E_3. Ἔνθα :

E_3 ἡ μαθηματικὴ ἐλπὶς ὑπὸ τὴν συνθήκην : *«Εἰς δείγματα μεγέθους n, περικλείονται ἐκ μὲν τῆς ὁμάδος G_i πάντοτε ὁ αὐτὸς συνδυασμὸς n_i μονάδων, ἐκ δὲ τῆς ὁμάδος G_j διάφοροι μὲν συνδυασμοὶ ἀλλὰ τοῦ αὐτοῦ πλήθους $n_j = c$.*

Οὕτω :

$$E_3 (y_j\, y_i) =$$

$$= E \left[y_j\, y_i \middle/ \begin{matrix} y_i = \text{σταθερὸς ἀριθμὸς} \\ n_j = \text{σταθερὸς ἀριθμὸς} \end{matrix} \right]$$

$$= y_i\, E \left[y_j \;/\; n_j = \text{σταθερὸς ἀριθμὸς} \right]$$

Ἀλλ', ὡς γνωστόν, εἰς τὴν περίπτωσιν ἁπλῆς τυχαίας δειγματοληψίας :

$$E (y) = \frac{n}{N}\, Y .$$

Ὁπότε :

$$E_3 (y_j\, y_i) = y_i \frac{n_j}{N_j} Y_j .$$

Ἐπίσης E_2 εἶναι ἡ μαθηματικὴ ἐλπίς, ὑπὸ τὴν συνθήκην «n_j καὶ n_i σταθερά», ὁπότε :

$$E_2 \left[E_3 (y_j\, y_i) \right] =$$

$$= \frac{n_j}{N_j} Y_j\, E \left(y_i / n_i = \text{σταθερὸν} \right)$$

$$= \frac{n_j}{N_j} Y_j \frac{n_i}{N_i} Y_i$$

Τέλος ἡ E_1 θεωρεῖ τὸ n_j καὶ n_i μεταβλητά· καί :

$$E (y_j\, y_i) = E_1 \left[E_2 (E_3) \right]$$

$$= \bar{Y}_j\, \bar{Y}_i\, E (n_j\, n_i) .$$

Ἐναπόκειται ἤδη νὰ ὑπολογίσωμεν τὴν μαθηματικὴν ἐλπίδα τοῦ γινομένου ($n_j\, n_i$).

Αἱ περιπτώσεις, καθ' ἃς εἰς ἓν συνολικὸν δεῖγμα n θὰ εὑρεθοῦν ἐκ τῆς G_j ὁμάδος n_j μονάδες καὶ ἐκ τῆς G_i ὁμάδος n_i μονάδες (ὁπότε ἐκ τῶν ὑπολοίπων ὁμάδων τὸ δεῖγμα θὰ περιλάβῃ $n - n_j - n_i$ μονάδας) εἶναι :

$$\binom{N_j}{n_j} \binom{N_i}{n_i} \binom{N - N_j - N_i}{n - n_j - n_i}$$

Ἀφ' ἑτέρου τὸ πλῆθος ὅλων τῶν δυνατῶν δειγμάτων εἶναι : $\binom{N}{n}$ κατὰ συνέπειαν :

$$P(n_j, n_i) = \frac{\binom{N_j}{n_j}\binom{N_i}{n_i}\binom{N - N_j - N_i}{n - n_j - n_i}}{\binom{N}{n}}$$

ἔνθα $(n_j, n_i = 0, 1, \ldots, n)$

Ἄρα:

$$E(n_j n_i) = \sum_{n_j, n_i = 1}^{n} n_j n_i P(n_j, n_i)$$

$$= \sum_{n_j, n_i = 1}^{n} n_j n_i \frac{\frac{N_j}{n_j}\binom{N_j - 1}{n_j - 1}\frac{N_i}{n_i}\binom{N_i - 1}{n_i - 1}\binom{N - N_j - N_i}{n - n_j - n_i}}{\frac{N(N-1)}{n(n-1)}\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)}{N(N-1)} N_j N_i \sum \frac{\binom{N_j - 1}{n_j - 1}\binom{N_i - 1}{n_i - 1}\binom{N - N_j - N_i}{n - n_j - n_i}}{\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)}{N(N-1)} N_j N_i .$$

Ἐπανερχόμενοι εἰς τὴν συνδιακύμανσιν τῶν $(\widehat{Y}_j, \widehat{Y}_i)$ ἔχομεν:

$$\mathrm{Cov}(\widehat{Y}_j, \widehat{Y}_i) = \frac{N^2}{n^2} \overline{Y}_j \overline{Y}_i E(n_j n_i) - Y_j Y_i$$

$$= \frac{N(n-1)}{n(N-1)} \overline{Y}_j \overline{Y}_i N_j N_i - Y_j Y_i$$

$$= Y_j Y_i \left[\frac{N(n-1)}{n(N-1)} - 1 \right]$$

$$= - \frac{(N-n)}{n(N-1)} Y_j Y_i$$

Μία ἀμερόληπτος δὲ ἐκτίμησις τῆς συνδιακυμάνσεως ταύτης δίδεται ὑπὸ τῆς ἐκφράσεως:

$$\widehat{\mathrm{Cov}}(\widehat{Y}_j, \widehat{Y}_i) = - \frac{(N-n)}{N(n-1)} \widehat{Y}_j \widehat{Y}_i .$$

Ἄμεσος συνέπεια τῶν ἀνωτέρω εἶναι ὅτι:

$$V(Y_j - Y_i) = V(Y_j) + V(Y_i) - 2\,\mathrm{Cov}(Y_j, Y_i)$$

καὶ ἑπομένως ἐὰν ἐπιθυμοῦμεν, ἐπὶ τῇ βάσει δύο ἐκτιμήσεων, νὰ ἐλέγξωμεν τὴν ὑπόθεσιν :

$$H_0 : Y_j - Y_i = 0$$

τότε πρέπει νὰ ὑπολογίσωμεν τὴν τιμὴν τοῦ λόγου :

$$t = \frac{|Y_j - Y_i|}{\sqrt{V(Y_j) + V(Y_i) - 2\,\mathrm{Cov}(Y_j, Y_i)}}$$

Μέγεθος δείγματος ἀπαιτούμενον διὰ ἐκτιμήσεις χαρακτηριστικοῦ εἰς ὁμάδας πληθυσμοῦ

Ἓν τελευταῖον πρόβλημα, τὸ ὁποῖον θὰ πρέπει νὰ ἐξετασθῇ διὰ τὴν ὁλοκλήρωσιν τοῦ θέματος, εἶναι ὁ καθορισμὸς τοῦ μεγέθους n τοῦ δείγματος, τὸ ὁποῖον θὰ ἐπιλεγῇ ἐκ τοῦ συνολικοῦ πληθυσμοῦ εἰς τρόπον, ὥστε αἱ ἐκτιμήσεις τῶν χαρακτηριστικῶν ὁμάδος τινὸς (λ.χ. τῆς ὁμάδος G_j) νὰ ἔχουν τὴν ἐπιθυμητὴν ἀκρίβειαν. Καθὼς ἤδη ἐλέχθη καὶ προηγουμένως, τὸ μέγεθος τοῦ δείγματος, χρησιμοποιούμενον διὰ τὴν παροχὴν ἐκτιμήσεων εἰς ὁλόκληρον τὸν πληθυσμόν, δὲν εἶναι πάντοτε ἐπαρκὲς διὰ νὰ δώσῃ ἀκριβεῖς ἐκτιμήσεις καὶ διὰ ὑποδιαιρέσεις τοῦ πληθυσμοῦ.

Ὑποθέσωμεν ὅτι ἐπιθυμοῦμεν ὁ μέσος $\overline{Y}_j$ τῆς G_j ὁμάδος νὰ ἐκτιμηθῇ μὲ ἀκρίβειαν $\frac{1}{V}$ ἤ, ὅπερ τὸ αὐτό, μὲ διακύμανσιν $V(\bar{y}_j) = V$ (V ἡ ἐπιθυμητὴ τιμὴ τῆς διακυμάνσεως).

Ὡς ἐξετέθη :

$$V(\bar{y}_j) = \left[E\left(\frac{1}{n_j}\right) - \frac{1}{N_j} \right] S_j^2$$

$$\simeq S_j^2\, E\left(\frac{1}{n_j}\right) \simeq \frac{S_j^2}{n W_j} \qquad \left(W_j = \frac{N_j}{N}\right)$$

Οὕτω :

$$V = \frac{S_j^2}{n\, W_j}$$

ὁπότε τὸ ὁλικὸν μέγεθος n τοῦ δείγματος θὰ προκύπτῃ ὡς :

$$n = \frac{S_j^2}{W_j\, V}$$

Ὁ τύπος οὗτος προϋποθέτει γνωστὰς ἐκ τῶν προτέρων (ἔστω καὶ κατὰ προσέγγισιν) τὰς τιμὰς S_j^2 καὶ W_j .

Ἐὰν δεχθῶμεν ὅτι (ἢ ὑπάρχουν πληροφορίαι ὅτι) αἱ διάφοροι ὁμάδες G_j ($j = 1, 2, 3, \ldots, m$) εἶναι ἰσοπληθεῖς ἢ περίπου ἰσοπληθεῖς, τότε:

$$W_j \simeq \frac{1}{m}$$

καὶ ἑπομένως:

$$n = m \frac{S_j^2}{V}$$

Ἐκ τοῦ τύπου τούτου συνάγεται, ὅτι ἐὰν θέλωμεν νὰ ἐκτιμήσωμεν χαρακτηριστικόν τι εἰς μίαν ἐκ τῶν m ὑποδιαιρέσεων τοῦ πληθυσμοῦ, μεθ' ὡρισμένης ἐπιθυμητῆς διακυμάνσεως, τότε τὸ ἀπαιτούμενον μέγεθος δείγματος θὰ πρέπει νὰ εἶναι m περίπου φορὰς μεγαλύτερον ἐκείνου ποὺ θὰ ἐχρειάζετο διὰ τὴν ἐκτίμησιν τοῦ χαρακτηριστικοῦ τούτου εἰς ὁλόκληρον τὸν πληθυσμὸν μετὰ τῆς αὐτῆς ὡς ἄνω ἀκριβείας.

Παράδειγμα

Διὰ τὴν ἐφαρμογὴν τῶν ἀνωτέρω ἐκτεθέντων παραθέτομεν τὸ ἀκόλουθον παράδειγμα, τὸ ὁποῖον λαμβάνεται ἐκ τῆς κατ' ἔτος διενεργουμένης δειγματοληπτικῆς ἐρεύνης ἐπὶ τῆς φορολογίας εἰσοδήματος φυσικῶν προσώπων. Σκοπὸς τῆς ἐρεύνης ταύτης εἶναι ἡ παροχὴ στατιστικῶν πληροφοριῶν σχετικῶν μὲ τὸ μέσον καὶ συνολικὸν εἰσόδημα τῶν φορολογουμένων, τὸ φορολογητέον τοιοῦτον καὶ τὸν ἀναλογοῦντα φόρον. Συμπληρωματικῶς, πρὸς τὰ λαμβανόμενα ἐκ τῆς ἐρεύνης ταύτης στοιχεῖα, τὰ ἀναφερόμενα εἰς τὸν συνολικὸν ἀριθμὸν τῶν φορολογουμένων, ὑπάρχει πάντοτε τὸ ἐνδιαφέρον διὰ τὴν ἀπόκτησιν παρομοίων στατιστικῶν πληροφοριῶν δι' ὡρισμένα τμήματα τοῦ πληθυσμοῦ τῶν φορολογουμένων, ὡς π.χ. συγκεκριμένων εἰσοδηματικῶν ὁμάδων ἢ ὁμάδων ἐπαγγελμάτων κλπ. Αἱ ὑποδιαιρέσεις αὗται τῶν φορολογουμένων ἀποτελοῦν ὁμάδας στατιστικοῦ πληθυσμοῦ καὶ διὰ τὴν ἐκτίμησιν τῶν παραμέτρων αὐτῶν θὰ πρέπει νὰ ἐφαρμόζεται ἡ ἀνωτέρω ἐκτεθεῖσα θεωρία.

Ἡ ἔρευνα αὕτη διεξάγεται ὡς ἀκολούθως: Τὸ σύνολον τῶν ὑποβαλλομένων ὑπὸ τῶν φορολογουμένων φυσικῶν προσώπων δηλώσεων κατατάσσεται εἰς πέντε στρώματα μὲ βάσιν τὸ ἀναγεγραμμένον συνολικὸν οἰκογενειακὸν εἰσόδημα. Ἐξ ἑκάστου στρώματος ἐπιλέγεται τυχαῖον δεῖγμα μὲ διάφορον κατὰ στρῶμα κλάσμα δειγματοληψίας, συμφώνως πρὸς τὴν ἀρχὴν τοῦ ἀρίστου καταμερισμοῦ τοῦ δείγματος (optimum allocation).

Δεδομένου δὲ ὅτι εἰς ἕκαστον στρῶμα ἡ ἐπιλογὴ τοῦ δείγματος εἶναι τυχαία καὶ ἀνεξάρτητος ἑτέρου στρώματος, ὁ πληθυσμὸς ἑνὸς μόνον στρώματος, καὶ συγκεκριμένως τοῦ στρώματος ἀπὸ 50 - 100 χιλ. εἰσόδημα, θὰ ἦτο δυνατὸν νὰ ληφθῇ ὡς αὐτοτελὴς πληθυσμός, διὰ τὸν ὁποῖον πραγματοποιοῦμεν ἁπλῆν τυχαίαν δειγματοληψίαν.

Ὡς ὑποδιαιρέσεις τοῦ πληθυσμοῦ (ὁμάδες) τῶν φορολογουμένων τοῦ στρώματος τούτου λαμβάνονται αἱ πέντε κύριαι ἐπαγγελματικαὶ ὁμάδες τῶν φορολογουμένων ἤτοι : Α΄. εἰσοδηματιῶν, Β΄. ἐμπόρων, βιομηχάνων κλπ., Γ΄. μισθωτῶν ἐν γένει, Δ΄. ἐλευθέρων ἐπαγγελματιῶν καὶ Ε΄. συνταξιούχων. Διὰ τοὺς φορολογουμένους ἑκάστης τῶν ἀνωτέρω ὁμάδων ζητεῖται ἐκτίμησις τῶν μέσων καὶ συνολικῶν οἰκογενειακῶν εἰσοδημάτων, καθὼς καὶ ἡ ἀξιοπιστία τῶν ἐν λόγῳ ἐκτιμήσεων. Ἐπίσης ἐλέγχεται ἡ σημαντικότης τῆς διαφορᾶς τῶν μέσων εἰσοδημάτων, τῶν διαφόρων ὁμάδων λαμβανομένων ἀνὰ ζεύγη. Τὰ βασικὰ στοιχεῖα παρέχονται κατωτέρω εἰς τοὺς πίνακας Α΄ — Δ΄, ἐλήφθησαν δὲ ἐκ τῆς διενεργηθείσης ἐρεύνης διὰ τὸ δηλωθὲν εἰσόδημα τοῦ οἰκονομικοῦ ἔτους 1964.

Πίναξ Α΄

Τιμαὶ τῶν n_j, $\sum_{k=1}^{n_j} y_{jk}$, $\sum_{k=1}^{n_j} y_{jk}^2$

ἔνθα y_{jk} δηλωθὲν εἰσόδημα φορολογουμένου τινὸς τῆς G_j ὁμάδος

Ὁμάδες ἐπαγγελμάτων G_j		Ἀριθμὸς φορολογουμένων n_j	Δηλωθὲν εἰσόδημα $\sum_k y_{jk} = y_j$	Ἄθροισμα τετραγώνων $\sum_k y_{jk}^2$
Εἰσοδηματίαι	Α	75	5.082	363.214
Ἔμποροι - Βιομήχανοι κλπ.	Β	512	37 087	2.495.481
Μισθωτοὶ	Γ	577	40.745	2.989.110
Ἐλεύθεροι ἐπαγγελματίαι	Δ	116	8.356	628.026
Συνταξιοῦχοι	Ε	227	15.157	1.059.198
		$n=1.507$		

Πίναξ Β΄

Τιμαὶ τῶν $W_j = \frac{n_j}{n}$, $\widehat{\widehat{Y}}_j$, $\widehat{Y}_j$, καὶ $\widehat{S}_j^2$

	$W_j = \frac{n_j}{n}$	$\widehat{\widehat{Y}}_j = \bar{y}_j$ χιλ. δρχ.	$\widehat{Y}_j$ χιλ. δρχ.	$\widehat{S}_j^2$
Α	0,05	67,9	229.838	222
Β	0,33	68,5	2.188.133	181
Γ	0,38	70,6	2.403.955	196
Δ	0,08	72,0	493.004	230
Ε	0,15	66,7	894.263	218

Πίναξ Γ΄

Διακυμάνσεις καὶ συντελεσταὶ μεταβλητικότητος μέσου καὶ συνόλου

	$\widehat{V}(\bar{y}_j)$	$\widehat{V}(\widehat{Y}_j)$	$\widehat{CV}(\bar{y}_j)$ %	$\widehat{CV}(\widehat{Y}_j)$ %
Α	2,840	1.187.223.272	2,40	11,5
Β	0,349	5.844.140.280	0,86	3,5
Γ	0,333	6.508.522 201	0,82	3,3
Δ	1,941	1.991.438.232	1,94	8,8
Ε	0,939	3.117.189.463	1,35	6,2

Πίναξ Δ΄

Συνδιακυμάνσεις τῶν $\widehat{Y}_j$, $\widehat{Y}_i$

$$\mathrm{Cov}(\widehat{Y}_A\ \widehat{Y}_B) = -\ 426.455.524$$

$$\mathrm{Cov}(\widehat{Y}_A\ \widehat{Y}_\Gamma) = -\ 468.518.088$$

$$\mathrm{Cov}(\widehat{Y}_A\ \widehat{Y}_\Delta) = -\ 96.083.866$$

$$\mathrm{Cov}(\widehat{Y}_A\ \widehat{Y}_E) = -\ 174.287.118$$

$$\mathrm{Cov}(\widehat{Y}_B\ \widehat{Y}_\Gamma) = -3\ 419.107.935$$

$$\mathrm{Cov}(\widehat{Y}_B\ \widehat{Y}_\Delta) = -\ 701.191.947$$

$$\mathrm{Cov}(\widehat{Y}_B\ \widehat{Y}_E) = -1.272.038.632$$

$$\mathrm{Cov}(\widehat{Y}_\Gamma\ \widehat{Y}_\Delta) = -\ 770.353.629$$

$$\mathrm{Cov}(\widehat{Y}_\Gamma\ \widehat{Y}_E) = -\ 1.397.349.206$$

$$\mathrm{Cov}(\widehat{Y}_\Delta\ \widehat{Y}_E) = -\ 286.568.903$$

Προκειμένου ἤδη νὰ ἐλέγξωμεν τὴν ὑπόθεσιν ὅτι τὰ μέσα εἰσοδήματα τῶν φορολογουμένων τῶν ἀνηκόντων εἰς δύο διαφόρους ἐπαγγελματικὰς ὁμάδας εἶναι τὰ αὐτά, ὑπολογίζομεν τὴν τιμήν :

$$t = \frac{|\bar{y}_j - \bar{y}_i|}{\sqrt{V(\bar{y}_j) + V(\bar{y}_i)}}$$

ἡ ὁποία δι' ἕκαστον ζεῦγος ἐπαγγελματικῶν ὁμάδων λαμβάνει τὰς τιμάς :

(1)	$t_{AB} = 0{,}33$	(6)	$t_{B\Delta} = 2{,}27$
(2)	$t_{A\Gamma} = 1{,}78$	(7)	$t_{BE} = 1{,}59$
(3)	$t_{A\Delta} = 1{,}88$	(8)	$t_{\Gamma\Delta} = 1{,}06$
(4)	$t_{AE} = 0{,}63$	(9)	$t_{\Gamma E} = 3{,}45$
(5)	$t_{B\Gamma} = 2{,}12$	(10)	$t_{\Delta E} = 3{,}13$

Εἰς ἐπίπεδον δὲ σημαντικότητος 5 %, ἐκ τῶν ἀνωτέρω τιμῶν t, εἶναι στατιστικῶς σημαντικαὶ αἱ τῶν περιπτώσεων (5), (6), (9) καὶ (10). Τοῦτο σημαίνει ὅτι δὲν δυνάμεθα νὰ δεχθῶμεν τὴν ὑπόθεσιν, ὅτι τὸ μέσον εἰσόδημα τῶν φορολογουμένων τοῦ στρώματος 50 - 100 χιλ. δραχ. τῆς ὁμάδος «Ἔμποροι κλπ.» εἶναι ἴσον μὲ τὸ μέσον εἰσόδημα τῶν φορολογουμένων τῆς ὁμάδος «Μισθωτοὶ» ἢ τὸ μέσον εἰσόδημα τῶν ἀνηκόντων εἰς τὴν ὁμάδα «Ἐλεύθεροι ἐπαγγελματίαι» εἶναι ἴσον μὲ τὸ μέσον εἰσόδημα τῆς ὁμάδος «Συνταξιοῦχοι».

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. *Πανταζίδη Ν.* : Εἰσαγωγὴ εἰς τὴν Θεωρίαν τῆς Δειγματοληψίας.
2. *Sukhatme, P. V.* : Sampling Theory of Surveys with Applications.
3. *Yates, F.* : Sampling Methods for Censuses and Surveys.
4. *Cochram, W. G.* : Sampling Techniques.
5. *Hartley, H. O.* : Analytic Studies of Survey Dada, Instituto di Statistica, Rome, Volume in onora di Corrado Gini.
6. *U. N. Statistical Office* : The preparation et Sample Survey Report. Stat. Papers, Series C, no 1.