ΔΥΟ ΖΗΤΗΜΑΤΑ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΕΩΣ

ΥΠΟ ΤΟΥ ΚΑΘΗΓΗΤΟΥ κ. Κ. Α. ΑΘΑΝΑΣΙΑΔΟΥ

Α' ΕΞΟΜΑΛΥΝΣΙΣ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΔΙΑ ΤΩΝ ΟΡΘΟΓΩΝΙΩΝ ΠΟΛΥΩΝΥΜΩΝ

1. Ἓν τῶν σημαντικωτέρων ἅμα δὲ καὶ δυσκολωτέρων προβλημάτων ἅτινα ἀντιμετωπίζει ἡ Στατιστικὴ εἶναι ἡ ἐξομάλυνσις τῶν καλουμένων χρονολογικῶν σειρῶν (Smoothing of Time Series), τουτέστιν ἡ ἀναλυτικὴ των ἀπεικόνισις, συντομογραφικῶς διδομένων ὑπὸ τῆς σχέσεως : $y = y_t$, ὅπου t ἡ χρονικὴ στιγμὴ πρὸς ἣν ἀναφέρεται ἡ ποσοτικὴ ἐκδήλωσις τῶν ὑπ' ὄψει φαινομένων.

Ἐὰν αἱ παρατηρήσεις (ποσοτικαὶ ἐκδηλώσεις) εἶναι κατ' ἀριθμὸν ἐπαρκεῖς, τότε εἴθισται νὰ χρησιμοποιῆται ὁ συναρτησιακὸς συμβολισμός : $y = y(t)$ ἀντὶ τοῦ $y = y_t$, τῆς μεταβλητῆς t (χρόνος) ὑποτιθεμένης συνεχοῦς ἐν τῷ διαστήματι $t_0 \leqslant t \leqslant t_1$.

2. Ὡς γνωστόν, εἰς τὰς χρονολογικὰς σειρὰς κατὰ κανόνα διαπιστοῦται μία ***τάσις αἰωνόβιος*** ἢ τουλάχιστον ***μακροχρόνιος***, δηλαδὴ μία ἔμμονος τάσις πρὸς αὔξησιν ἢ μείωσιν, ἡ κληθεῖσα ὑπὸ τοῦ Carl Snyder ***ἀδράνεια*** τῶν χρονολογικῶν σειρῶν καὶ ἡ ὁποία, ἀναλυτικῶς, δὲν δίδεται πάντοτε ὑπὸ εὐθείας γραμμῆς (εὐθύγραμμος τάσις) οὔτε καὶ πάντοτε εἶναι θετική. Διὰ τὴν ἑρμηνείαν τῶν χρονολογικῶν σειρῶν ἐκ τῶν ὧν οὐκ ἄνευ θεωρεῖται ἡ ἀπαλειφὴ τῆς ἐπιδράσεως τῆς τάσεως, ἥτις προϋποθέτει τὴν ἀντικατάστασιν αὐτῶν ὑπὸ ἀναλυτικῆς ἐκφράσεως, ἀποτελούσης τὸν ἐμπειρικὸν ποσοτικὸν νόμον τῆς ἐκδηλώσεως τοῦ ἀντιστοίχου πρὸς ὃ αἱ παρατηρήσεις φαινομένου μακροχρονίως.

3. Ἐξ ὑποθέσεως δεχόμεθα ὅτι αἱ χρονολογικαὶ σειραί, καὶ δὴ αἱ οἰκονομικαὶ τοιαῦται, δύνανται νὰ ἀναπαρασταθοῦν ὑπὸ ἀκεραίων πολυωνύμων ἢ παραβολῶν τοῦ νυοστοῦ βαθμοῦ δηλ. ὑπὸ ἀναλυτικῶν ἐκφράσεων τῆς μορφῆς :

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_n t^n = \Sigma \beta_i t^i \qquad (i = 0, 1, 2, \ldots, n)$$

Εἰς ταύτας, αἱ παράμετροι β_i ἐπιδιώκεται νὰ ἐκφράζωνται συναρτήσει τῶν

Οὕτω π.χ. ἔχομεν

$$y_1 = \alpha_1 x_1^{\alpha_2} + \alpha_3, \qquad y_2 = \alpha_1 x_2^{\alpha_2} + \alpha_3, \qquad y_3 = \alpha_1 x_3^{\alpha_2} + \alpha_3.$$

Δι' ἀπαλοιφῆς τοῦ α_1 εὑρίσκομεν

$$(y_1 - \alpha_3) : (y_2 - \alpha_3) = \left(\frac{x_1}{x_2}\right)^{\alpha_2}$$

$$(y_2 - \alpha_3) : (y_3 - \alpha_3) = \left(\frac{x_2}{x_3}\right)^{\alpha_2}.$$

Ὑποτίθεται, συνήθως $\quad x_2^2 = x_1 x_3$

ὅτε εἶνε $\quad (y_1 - \alpha_3) : (y_2 - \alpha_3) = \left(\frac{x_1}{x_2}\right)^{\alpha_2} = \left(\frac{x_2}{x_3}\right)^{\alpha_2} = (y_2 - \alpha_3) : (y_3 - \alpha_3)$

καὶ ἐκ τούτων εὑρίσκομεν : $\quad \alpha_3 = (y_1 y_3 - y_2^2) : (y_1 + y_3 - 2y_2)$, δεδομένων ($y_t$, t), πρὸς καθορισμὸν αὐτῶν ἐφαρμοζομένης τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων, τοῦ t ἐκφράζοντος τὰς διαφόρους χρονικάς στιγμὰς ὡς ἀποκλίσεις (διαφορὰς) ἀπὸ τῆς κεντρικῆς χρονικῆς στιγμῆς ἥτις ἐλήφθη, αὐθαιρέτως, ὡς ἀρχὴ καὶ ἐπὶ τῇ συνθήκῃ, πρὸς ἁπλούστευσιν τῶν ὑπολογισμῶν, ὅπως τὰ ἀθροίσματα $\Sigma t = \Sigma t^3 = \Sigma t^k = \ldots = 0$.

4. Ἡ μέθοδος ὅμως τῶν ἐλαχίστων τετραγώνων, ὅταν $i > 3$, καθίσταται κοπιώδης, καθ' ὃ ἀπαιτοῦσα σχοινοτενεῖς ὑπολογισμοὺς καὶ συνεπῶς ἀπώλειαν χρόνου. Ἐκτὸς τούτου, ἂν αὐξηθῇ ὁ βαθμὸς τοῦ χρησιμοποιουμένου πολυωνύμου κατὰ μονάδα, τότε ἀπαιτεῖται ἐξ ὑπαρχῆς νὰ ἐκτιμηθοῦν, συναρτήσει τῶν δεδομένων, αἱ ἀντίστοιχοι τούτου παράμετροι β_i καί, συνεπῶς, ἡ μέχρι τότε γενομένη ὑπολογιστικὴ ἐργασία καθίσταται ἄχρηστος. Ἀλλὰ καὶ παρὰ τὰ μειονεκτήματα ταῦτα, ἅτινα θὰ ἠδύναντο νὰ παροραθοῦν, δὲν ὑφίσταται ἀντικειμενικόν τι κριτήριον, ἐπιτρέπον εἰς τὸν ἐρευνητὴν νὰ προσδιορίσῃ, ἅμα τῇ χρησιμοποιήσει, διὰ τὴν ἐξομάλυνσιν τῶν δεδομένων, πολυωνύμου δεδομένου βαθμοῦ, ἂν τοῦτο ἐνδείκνυται ἢ ὄχι, τουτέστιν ἂν θὰ πρέπῃ νὰ προχωρήσῃ χρησιμοποιῶν πολυώνυμον ἀνωτέρου ἔτι βαθμοῦ, καὶ τοῦτο, διότι τὸ σφάλμα ἐκτιμήσεως δὲν περιεστάλη ἱκανῶς διὰ τοῦ χρησιμοποιηθέντος πολυωνύμου.

5. Ἐφ' ὅσον ὅθεν πρὸς ἐξομάλυνσιν χρησιμοποιεῖται ἡ μέθοδος τῶν ἐλαχίστων τετραγώνων, δὲν διαθέτομεν πρακτικῶς μέθοδον ἐπιτρέπουσαν νὰ ἐλέγχωμεν, ἀνὰ πᾶν βῆμα, ἂν τὸ χρησιμοποιούμενον πολυώνυμον εἶναι τὸ ἐνδεικνυόμενον διὰ τὴν ἐξομάλυνσιν ἢ ὄχι, εἰμὴ μόνον καθορίζοντες ἐκ νέου τὰς παραμέτρους τοῦ πολυωνύμου καὶ συναρτήσει τούτων καθορίζοντες τὴν διακύμανσιν τῆς ἐκτιμήσεως, δηλαδὴ τὸν μέσον τοῦ ἀθροίσματος τῶν τετραγώνων τῶν ἀποκλίσεων μεταξὺ τιμῶν παρατηρήσεως καὶ τιμῶν ὑπολογισμοῦ διὰ τοῦ πολυωνύμου ἐξομαλύνσεως.

6. Ἕνεκεν τῶν ἀνωτέρω λόγων ἐπιδιώκεται ἡ ὑποκατάστασις τῆς ἐκφράσεως: $y = \Sigma\beta_i t^i$ $(i=0, 1, 2, \ldots, n)$ δι' ἄλλης, ἰσοδυνάμου, τῆς μορφῆς: $y = \Sigma\alpha_i\,\xi_i$ $(i = 0, 1, 2, \ldots, n)$ καὶ τοιαύτης ὥστε νὰ πληροῦνται αἱ τρεῖς συνθῆκαι: α) $\xi_0 = 1$, β) $\Sigma\xi_i = 0$, $\xi_i\,\xi_\kappa = 0$, ἂν $i \neq \kappa$ τῶν ξ_i ὄντων πολυωνύμων τῶν t^i καὶ καλουμένων ***ὀρθογωνίων*** τοιούτων.

ἐπειδὴ δὲ εἶνε $y - \alpha_3 = \alpha_1 x^{\alpha_2}$ ἐργαζόμεθα ἐπὶ ταύτης ὡς ἀνωτέρω διὰ τὴν εὕρεσιν τῶν α_1, α_2 *.

* Ἐφαρμογὰς τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων ὡς καὶ ἀνάπτυξιν αὐτῆς βλ. ἐκτὸς τοῦ βιβλίου τοῦ Σ. Μαργαρίτη, Στατιστικὴ (ὡς ἀνωτέρω) σελ. 199—209, καὶ εἰς

1) Στοιχεῖα Βιομετρίας καὶ Στατιστικῆς Β. Γ. Βαλαώρα, 1942, σελ. 278—292.

2) Κ. Ἀθανασιάδου : Στατιστική, 1931. σελ. 118—142.

3) Emil Borel : Probabilités, Erreurs, 1934. σελ. 171 κ.ἑ.

4) W.Winkler: Gruudriss der Statistik, I, Theoretische Statistik, σελ. 117 κ.ἑ.

5) G. A. Baker : Theory of Budgets based on parabolic Engel Curves : Mathematics Magazine, Vol. 26, No 2, Nov-Dec., 1952, σελ. 67—70 (ὅπου ἐρευνᾶται ἡ ἐξίσωσις $x_i = \alpha - \beta e^{-\mu x}$.

7. Αἱ παράμετροι τῆς σχέσεως $y = \Sigma\alpha_i \xi_i$ καθορίζονται συναρτήσει τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων, καθισταμένου ἐλαχίστου τοῦ ἀθροίσματος τῶν τετραγώνων τῶν ἀποκλίσεων μεταξὺ τιμῶν παρατηρήσεως (y) καὶ τιμῶν ὑπολογισμοῦ $(\Sigma\alpha_i \xi_i)$. Τὸ ἐλάχιστον τοῦτο ὑφίσταται, ὡς γνωστόν, ἂν αἱ μερικαὶ παράγωγοι πρὸς α_i τῆς ἐκφράσεως $\Sigma(y - \Sigma\alpha_i \xi_i)^2$ ἐξισωθοῦν πρὸς τὸ μηδέν, ὅτε προκύπτει τὸ κάτωθι σύστημα ***κανονικῶν ἐξισώσεων***.

$$\Sigma y = N\alpha_0 + \alpha_1\Sigma\xi_1 + \alpha_2\Sigma\xi_2 + \ldots + \alpha_n\Sigma\xi_n = N\alpha_0$$

$$\Sigma\xi_1 y = \alpha_0\Sigma\xi_1 + \alpha_1\Sigma\xi_1^2 + \alpha_2\Sigma\xi_1\xi_2 + \ldots + \alpha_n\Sigma\xi_1\xi_n = \alpha_1\Sigma\xi_1^2$$

$$\Sigma\xi_2 y = \alpha_0\Sigma\xi_2 + \alpha_1\Sigma\xi_1\xi_2 + \alpha_2\Sigma\xi_2^2 + \ldots + \alpha_n\Sigma\xi_2\xi_n = \alpha_2\Sigma\xi_2^2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\Sigma\xi_n y = \alpha_0\Sigma\xi_n + \alpha_1\Sigma\xi_1\xi_n + \alpha_2\Sigma\xi_2\xi_n + \ldots + \alpha_n\Sigma\xi_n^2 = \alpha_n\Sigma\xi_n^2$$

διότι $\xi_0 = 1$ καὶ $\Sigma\xi_0 = N$, $\Sigma\xi_i = 0$ καὶ $\Sigma\xi_i\xi_\kappa = 0$, ἐὰν $i \neq \kappa$.

8. Κατ' ἀκολουθίαν αἱ ζητούμεναι παράμετροι δίδονται ὑπὸ τῶν σχέσεων:

$$\alpha_0 = \frac{\Sigma y}{N}, \quad \alpha_1 = \frac{\Sigma\xi_1 y}{\Sigma\xi_1^2}, \quad \alpha_2 = \frac{\Sigma\xi_2 y}{\Sigma\xi_2^2}, \ldots, \alpha_n = \frac{\Sigma\xi_n y}{\Sigma\xi_n^2}.$$

Ἑπομένως ἡ προσθήκη νέας παραμέτρου α_{n+1} δὲν ἀχρηστεύει τὰς προγενεστέρως ὑπολογισθείσας τοιαύτας, $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n$, διότι $\alpha_{n+1} = \dfrac{\Sigma\xi_{n+1} y}{\Sigma\xi_{n+1}^2}$

Τὸ ἄθροισμα τῶν τετραγώνων τῶν ἀποκλίσεων μεταξὺ τιμῶν παρατηρήσεως καὶ τιμῶν συναρτήσεως δίδεται ὑπὸ τῶν σχέσεων:

$$S = \Sigma(y - \alpha_0\xi_0 - \alpha_1\xi_1 - \alpha_2\xi_2 - \ldots - \alpha_n\xi_n)^2$$

$$= \Sigma y^2 - 2\alpha_0\Sigma\xi_0 y - \ldots - 2\alpha_n\Sigma\xi_n y + \alpha_0^2\Sigma\xi_0^2 + \alpha_1^2\Sigma\xi_2^2 + \ldots + \alpha_n^2\Sigma\xi_n^2$$

ἀλλὰ $\alpha_0^2\Sigma\xi_0^2 = N\alpha_0^2$ καὶ $\Sigma\xi_0 y = \Sigma y = N\alpha_0$, $\xi_0 = 1$, $\Sigma\xi_0 = N$

$\alpha_1\Sigma\xi_1^2 = \Sigma\xi_1 y, \ldots, \alpha_n = \Sigma\xi_n^2 = \Sigma\xi_n y$, ὅθεν ἡ ὡς ἄνω ἔκφρασις γράφεται

$$S = \Sigma y^2 - N\alpha_0^2 - \alpha_1\Sigma\xi_1 y - \alpha_2\Sigma\xi_2 y - \ldots - \alpha_n\Sigma\xi_n y \qquad (\alpha)$$

Εἰς τὴν τελευταίαν ταύτην ἔκφρασιν παρατηροῦμεν ὅτι οἱ δύο πρῶτοι ὅροι τοῦ δευτέρου μέλους δίδουν τὸ ἄθροισμα τῶν τετραγώνων τῶν ἀποκλίσεων τῶν τιμῶν παρατηρήσεως y περὶ τὸν μέσον αὐτῶν $\alpha_0 = \Sigma y/N$. Ὁμοίως ὁ τρίτος ὅρος: $\alpha_1\Sigma\xi_1 y$ δηλοῖ τὴν περιστολὴν τοῦ ἀθροίσματος τούτου, ἥτις προέρχεται ἐκ τῆς ὑφισταμένης γραμμικῆς σχέσεως μεταξὺ y καὶ t, ὁ τέταρτος ὅρος τὴν περαιτέρω περιστολὴν τοῦ αὐτοῦ ἀθροίσματος ἥτις ὀφείλεται εἰς τὴν παραβολικὴν σχέσιν μεταξὺ y καὶ t, καὶ οὕτω καθεξῆς. Κατὰ συνέπειαν, ἡ τελευταία σχέσις (α) παρέχει ἡμῖν, ἀνὰ πᾶν βῆμα, πρακτικὸν μέσον ἐλέγχου τοῦ ὀρθοῦ τῆς ἐξομαλύνσεως τῶν δεδομένων.

9. Ὅσον ἀφορᾷ τὰς ἀναλυτικὰς μορφὰς τῶν ξ_i δυνάμεθα νὰ θέσωμεν

$\xi_0 = 1$, $\xi_1 = \kappa_1 + t$, $\xi_2 = \kappa_2 + \kappa_3 t + t^2$, $\xi_3 = \kappa_3 + \kappa_4 t + \kappa_5 t^2 + t^3$ κλπ.

καὶ νὰ ἐφαρμόσωμεν ἐπὶ τούτου τοὺς ὁρισμοὺς α) $\Sigma\xi_i = 0$ καὶ $\Sigma\xi_i\xi_\kappa = 0$ ἐὰν $i \neq \kappa$ καὶ β) $\xi_0 = 1$, $\Sigma\xi_0 = N$.

Ἐπ' εὐκαιρίᾳ σημειοῦμεν ὅτι, ἐπειδὴ αἱ χρονολογικαὶ σειραὶ ἀναφέρονται εἰς διαφόρους χρονικὰς στιγμὰς T (ἔτη συνήθως), θέτομεν $t = T - \overline{T}$, ὅπου $\overline{T}$ ὁ μέσος τῶν T καὶ ἑπομένως $\Sigma t = 0$, τῶν T βαινόντων κατὰ πρόοδον ἀριθμητικὴν λόγου 1. Τούτου τεθέντος, αἱ τιμαὶ ἃς ἡ t δύναται νὰ λάβῃ θὰ εἶναι $t = \pm 1$ ἕως $t = \pm(n-1)$, ὅπου $n = (N+1)/2$, δηλ. $n = t + 1$. Οὕτω, ἂν ὑφίστανται 11 ἔτη (N=11), τότε $n = (11+1)/2 = 6$ καὶ ἑπομένως t θὰ λάβῃ τὰς τιμὰς 0, ±1, ±2, ±3, ±4, ±5.

Αἱ ῥοπαὶ ἀρτίας τάξεως τῆς t ὑφίστανται πᾶσαι ($\Sigma t^2/N$, $\Sigma t^4/N$, $\Sigma t^6/N$, ...) ἐνῷ αἱ τοιαῦται περιττῆς τάξεως ($\Sigma t/N$, $\Sigma t^3/N$, $\Sigma t^5/N$, ...) εἶναι πᾶσαι μηδέν. Εὐκόλως δὲ ἀποδεικνύεται ὅτι

$$\Sigma t^2 = \frac{n(n-1)(2n-1)}{3} \quad \text{'Αλλὰ} \quad n = \frac{N+1}{2}, \quad \text{ὅθεν}$$

$$\Sigma t^2 = \frac{\frac{N+1}{2}\left(\frac{N-1}{2}\right)N}{3} = \frac{(N^2-1)N}{12} \quad \text{καὶ} \quad \frac{\Sigma t^2}{N} = \frac{N^2-1}{12}$$

ὁμοίως

$$\Sigma t^4 = \Sigma t^2\left(\frac{3n^2-3n-1}{5}\right) = \Sigma t^2 \cdot \frac{3N^2-7}{20} \quad \text{κλπ.}$$

10. Ἑπομένως θὰ ἔχωμεν:

ἐξ ὁρισμοῦ: $\xi_0 = 1$ καὶ $\Sigma\xi_1 = \Sigma(\kappa_1 + t) = N\kappa_1 + \Sigma t = 0$. Ἀλλὰ $\Sigma t = 0$,

ὅθεν $N\kappa_1 = 0$ δηλ. $\kappa_1 = 0$. Ἑπομένως $\xi_1 = t$, $\Sigma\xi_0\xi_1 = \Sigma(\kappa_1 + t) = 0$.

ὁμοίως $\xi_2 = \kappa_2 + \kappa_3 t + t^2$, ἄρα $\Sigma\xi_2 = \Sigma(\kappa_2 + \kappa_3 t + t^2) = N\kappa_2 + \kappa_3\Sigma t + \Sigma t^2 = 0$

ἤτοι $N\kappa_2 + \Sigma t^2 = 0$ ἢ $\kappa_2 = -\dfrac{\Sigma t^2}{N} = -\dfrac{N^2-1}{12}$

Ἀλλὰ $\Sigma\xi_1\xi_2 = \Sigma t(\kappa_2 + \kappa_3 t + t^2) = \kappa_2\Sigma t + K_3\Sigma t^2 + \Sigma t^3 = 0$ ἤτοι $\kappa_3\Sigma t^2 = 0$

καὶ ἐπειδὴ $\Sigma t^2 > 0$, ἕπεται ὅτι $\kappa_3 = 0$, κατ' ἀκολουθίαν

$$\xi_2 = \kappa_2 + \kappa_3 t + t^2 = t^2 - \frac{N^2-1}{12}$$

ὁμοίως $\xi_3 = \kappa_3 + \kappa_4 t + \kappa_5 t^2 + t^3$ καὶ $\Sigma\xi_3 = \Sigma(\kappa_3 + \kappa_4 t + \kappa_5 t^2 + t^3) =$

$= N\kappa_3 + \kappa_4\Sigma t + \kappa_5\Sigma t^2 + \Sigma t^3 = 0$

ἢ

$$\Sigma\xi_3 = N\kappa_3 + \kappa_4\Sigma t + \kappa_5\Sigma t^2 + \Sigma t^3 = 0 \text{ ἢ } N\kappa_3 + \kappa_5\Sigma t^2 = 0 \quad (6)$$

Ἐξ ἄλλου ὅμως $\Sigma\xi_1\xi_3 = \Sigma t(\kappa_3 + \kappa_4 t + \kappa_5 t^2 + t^3) = \kappa_3\Sigma t + \kappa_4\Sigma t^2 + \kappa_5\Sigma t^3 + \Sigma t^4 = 0$ ἢ $\kappa_4 = -\Sigma t^4/\Sigma t^2$. Ἀλλὰ $\Sigma t^4 = \Sigma t^2(3N^2-7)/20$, ὅθεν $\Sigma t^4/\Sigma t^2 = (3N^2-7)/20$

καὶ ἑπομένως $\kappa_4 = -\dfrac{3N^2-7}{20}$.

Εἰς τὴν σχέσιν ὅμως (6) καὶ τὸ N καὶ τὸ Σt^2 εἶναι ἀμφότερα θετικοὶ ἀριθμοί, ὅθεν, ἵνα ἰσχύῃ αὕτη, δέον ὅπως $K_3 = K_5 = 0$. Κατ' ἀκολουθίαν τὸ ξ_3 γράφεται:

$$\xi_3 = t^3 - \frac{3N^2 - t}{20}.\, t$$

Ὁμοίως εὑρίσκομεν ὅτι :

$$\xi_4 = t^4 - \frac{3N^2 - 13}{14}t^2 + \frac{3(N^2-1)(N^2-9)}{560} \quad \text{καὶ}$$

$$\xi_5 = t^5 - \frac{5(N^2-7)}{18}t^3 + \frac{15N^4 - 230N^2 + 407}{1008}.t \quad \text{κ. λ. π.} \quad *.$$

12. Ἐπειδὴ δὲ οἱ ἀριθμοὶ ξ_i δὲν εἶναι πάντοτε ἀκέραιοι, καθιστῶμεν αὐτοὺς τοιούτους διὰ τῆς σχέσεως $\xi'_i = \lambda_i . \xi_i$, ἐκλέγοντες καταλλήλως τὸ λ_i, τῆς τιμῆς τούτου ἐξαρτωμένης τὸ μὲν ἐκ τοῦ i τὸ δὲ ἐκ τοῦ N. Σημειωτέον ἐπίσης ὅτι ὑφίστανται πίνακες τῶν τιμῶν ξ'_i διὰ i=1 ἕως i=5 καὶ N ἀπὸ N=3 ἕως N=52 (Fisher—Yates : Statistical Tables for Biological, Agricultural and Medical Research) καὶ πληρέστεροι τοιοῦτοι ἕως N=104 (Anderson—Houseman : Tables of Orthogonal Polynomial Values extended to N=104). Διὰ τῶν σχέσεων αἵτινες συνδέουν $\xi'_i = \lambda_i \xi_i$ καὶ t δυνάμεθα, ὡς εἰκός, νὰ μεταβῶμεν ἀπὸ τῶν ὀρθογωνίων πολυωνύμων εἰς τὰ ἀκέραια τοιαῦτα εἰς t βαθμοῦ i .

Ἐφαρμογή. Πρὸς κατανόησιν τῶν ἀνωτέρω ἐκτεθέντων παραθέτομεν σχετικὴν ἐφαρμογὴν ἀφορῶσαν τὴν ἐξέλιξιν τοῦ τιμαρίθμου χονδρικῆς πωλήσεως τοῦ Α.Ο.Σ. (βάσις 1913/14=1) διὰ τὰ ἔτη 1928 ἕως 1938.

Ὡς γίνεται σαφὲς ἐκ τοῦ κατωτέρω πίνακος, ἡ ὅλη ἐργασία τῶν ὑπολογισμῶν περιορίζεται εἰς τὴν λῆψιν τῶν τιμῶν τῶν ξ'_i διὰ N=11 ἐκ τῶν πινάκων τῶν

Ἔτος	Τιμάρ. y	t	ξ'_1	ξ'_2	ξ'_3	ξ'_4	$\xi'_1 y$	$\xi'_2 y$	$\xi'_3 y$	$\xi'_4 y$	y^2
1928	17	—5	—5	15	—30	6	—85	255	—510	102	289
1929	18	—4	—4	6	6	—6	—72	108	108	—108	324
1930	16	—3	—3	— 1	22	—6	—48	— 16	352	— 96	256
1931	15	—2	—2	— 6	23	—1	—30	— 90	345	— 15	225
1932	18	—1	—1	— 9	14	4	—18	—162	252	72	324
1933	20	0	0	—10	0	6	0	—200	0	120	400
1934	20	1	1	— 9	—14	4	20	—180	—280	80	400
1935	20	2	2	— 6	—23	—1	40	—120	—460	— 20	400
1936	20	3	3	— 1	—22	—6	60	— 20	—440	—120	400
1937	23	4	4	6	— 6	—6	72	138	—138	—138	529
1938	22	5	5	15	30	6	110	330	660	132	484
Σύνολον...	209	0	0	0	0	0	49	43	—111	9	4031

Fisher—Yates, τὴν μόρφωσιν τῶν $\Sigma\xi'_i y$ καὶ τὸν διὰ διαιρέσεων ὑπολογισμὸν τῶν α_i διδομένων ὑπὸ τῶν πηλίκων $\alpha_i = \dfrac{\Sigma\xi'_i y}{\Sigma(\xi'_i)^2}$ τῶν διαιρετῶν διδομένων ὑπὸ τῶν πινάκων κάτωθι τῶν τιμῶν ἑκάστου ξ'_i .

Ἐπίσης εἶναι φανερὸν ὅτι διὰ τὸν ὑπολογισμὸν τῶν τιμῶν y, πρὸς χάραξιν

---

*) Τὰ πολυώνυμα ξ'_i δύνανται νὰ ὑπολογίζωνται, βαθμηδόν, διὰ τοῦ ἑξῆς ἀναδρομικοῦ τύπου :

$$\xi_{i+1} = \xi_1\xi_i - \frac{i^2(N^2 - i^2)}{4(4i^2-1)}\xi_{i-1}, \quad \text{τιθεμένου} \quad \xi'_0 = 1 \quad \text{καὶ} \quad \xi'_1 = t$$

τῆς σχετικῆς καμπύλης, ἁπλούστερον εἶναι νὰ ἐργαζώμεθα μὲ αὐτὰς ταύτας τὰς τιμὰς ξ'_i παρὰ μὲ τὰς τοιαύτας τοῦ t.

Ἐνταῦθα $N=11$. Ἄρα, ἐφαρμόζοντες τοὺς σχετικοὺς τύπους, ἔχομεν :

$$\alpha_0 = \frac{\Sigma y}{N} = \frac{209}{11} = 19, \quad \alpha_1 = \frac{39}{110} = 0{,}445, \quad \alpha_2 = \frac{43}{858} = 0{,}0501$$

$$\alpha_3 = -\frac{111}{4290} = -0{,}0258, \qquad \alpha_4 = \frac{9}{286} = 0{,}0314$$

Ἑπομένως ἡ ζητουμένη ἐξίσωσις εἶναι :

$$y = 19 + 0{,}445\xi'_1 + 0{,}0501\xi'_2 - 0{,}0258\xi'_3 + 0{,}0314\xi'_4$$

τὸ ἄθροισμα τῶν τετραγώνων τῶν ἀποκλίσεων περὶ τὸν μέσον τῶν y δίδεται ὑπὸ

$$\Sigma y^2 - \alpha_0 \Sigma y = 4031 - 19.209 = 4031 - 3971 = 60.$$

Ἡ περιστολὴ ἥτις ὀφείλεται εἰς τὴν γραμμικὴν σχέσιν εἶναι

$\alpha_1 \Sigma\xi'_1 y = 0{,}445.49 = 21{,}80.$ Ἄρα $-21{,}80$, ὑπόλοιπον $38{,}20$

ἡ περιστολὴ λόγῳ τῆς παραβολικῆς σχέσεως εἶναι

$\alpha_2 \Sigma\xi'_2 y = 0{,}0501.43 = 2{,}15.$ Ἄρα $-2{,}15$, ὑπόλοιπον $36{,}05$

ἡ περιστολὴ ἐπίσης λόγῳ τῆς κυβικῆς σχέσεως εἶναι

$\alpha_3 \Sigma\xi'_3 y = (-0{,}0258)(-111) = 2{,}86.$ Ἄρα $-2{,}86$, ὑπόλοιπον $33{,}19$

τέλος ἡ περιστολὴ ἐκ τῆς 4βαθμίου σχέσεως εἶναι

$\alpha_4 \Sigma\xi'_4 = 0{,}0314 \,.\, 9 = 0{,}28.$ Ἄρα $0{,}28$, ὑπόλοιπον $32{,}91$

Τὰ ἀνωτέρω δύνανται νὰ συνοψισθοῦν ὡς κάτωθι

Μεταβολὴ	ἄθροισμα τετρ. ἀποκλίσ.	Βαθμοὶ ἐλευθερίας	Διακυμάνσεις
ὁλική	60	10	6
Γραμμικὴ σχέσις	21,80	1	21,8
Ἀποκλίσεις	38,20	9	**4,24**
Παραβολικὴ σχέσις	2,15	1	2,15
Ἀποκλίσεις	36,05	8	4,50
Κυβικὴ σχέσις	2,86	1	2,86
Ἀποκλίσεις	33,19	7	4,74
4βάθμιος σχέσις	0,28	1	0,29
Ἀποκλίσεις	32,91	6	5,48

Ἐκ τῶν διακυμάνσεων μικροτέρα εἶναι ἡ ἀναφερομένη εἰς γραμμικὴν σχέσιν (4,24). Ἑπομένως, ἐπειδὴ $\lambda_1 = 1$, $\xi'_1 = \xi_1 = t$ καὶ ἄρα $y = 19 + 0{,}445\,t$.

Ἐάν, τώρα, θέλωμεν νὰ χρησιμοποιήσωμεν παραβολὴν 2ου βαθμοῦ, ἐπειδὴ

$$\xi'_2 = \lambda_2 \xi_2 = 1\left(t^2 - \frac{N^2-1}{12}\right) = t^2 - 10.$$

καὶ ἄρα: $y=19+0,445\,t+0,0501(t^2-10)=19+0,445+0,0201\,t^2-0,501$ ἢ
$y=18,499+0,445\,t+0,0501\,t^2$.

Ἡ δευτεροβάθμιος παραβολὴ παρατίθεται ἁπλῶς διὰ νὰ φανῇ ὁ τρόπος τῆς μεταβάσεως ἐκ τῶν ξ'_i εἰς πολυώνυμον εἰς t.

Ἐκ τῶν ἀνωτέρω καταφαίνεται πόσον ἁπλῆ, ταχεῖα καὶ εὔκολος εἶναι ἡ μέθοδος τῶν ὀρθογωνίων πολυωνύμων.

Β' ΑΝΑΛΥΣΙΣ ΤΗΣ ΔΙΑΚΥΜΑΝΣΕΩΣ *

1. Ἐὰν δοθοῦν n δείγματα, προτιθέμεθα νὰ προσδιορίσωμεν ἐὰν αἱ παρατηρούμεναι διαφοραὶ εἰς τὴν ὁμάδα τῶν n μέσων δύνανται νὰ ἀποδοθοῦν εἰς τὰ ***τυχαῖα τῆς δειγματοληψίας σφάλματα***, ὅτε ἀπὸ ἀπόψεως ***θεωρητικῆς*** τὸ πρόβλημα τίθεται ὡς ἑξῆς: αἱ παρατηρηθεῖσαι n σειραὶ δύνανται νὰ θεωρηθοῦν ὡς n δείγματα ληφθέντα ***κατὰ τύχην*** ἐκ τοῦ αὐτοῦ πλήθους, τοῦ πλήθους ὑποτιθεμένου κανονικοῦ (τύπου Gauss — Laplace) ἢ προσεγγιζόντως κανονικοῦ. Ἐὰν ὅθεν ὁ ἔλεγχος ἀγάγῃ εἰς ἀπάντησιν ἀρνητικήν, λογικῶς ἕπεται ὅτι αἱ διαφοραὶ φέρονται ἐπὶ τῶν μέσων καὶ ὄχι ἐπὶ τῶν διακυμάνσεων.

2. Ἀπὸ ***πρακτικῆς*** ἀπόψεως ἡ σύγκρισις τῶν μέσων ἐπιτρέπει νὰ ἀπαντήσωμεν εἰς ἓν τῶν ἑπομένων ἐρωτημάτων.

— Τὸ σύνολον τῶν n δειγμάτων ἀποτελεῖ ***ὁμοιογενὲς*** πλῆθος ;

— Ὅταν αἱ παρατηρήσεις κατανεμηθοῦν εἰς n τάξεις, συναρτήσει τῆς δυνατῆς ἐπιδράσεως μιᾶς αἰτίας μεταβολῆς ἐκφραζούσης n τρόπους παρεμβάσεως διακεκριμένους καὶ ἀνεξαρτήτους, τὸ διαπιστούμενον ἀποτέλεσμα, θεωρούμενον ἐν τῷ συνόλῳ, εἶναι σημαντικόν ; Τὸ ἀποτέλεσμα τῆς ἐρευνωμένης ***αἰτίας***—καλουμένης αἰτίας ἐλεγχομένης—προστίθεται εἰς τὰς κυβευτικὰς (aléatoires) μεταβολὰς τοῦ μεταβλητοῦ καὶ ἡ αἰτία αὕτη, ἐξ ὑποθέσεως, ἀσκεῖ τὴν ἐπίδρασίν της ἐπὶ τοῦ μέσου τοῦ ὑπ' ὄψει πλήθους καὶ ὄχι ἐπὶ τῆς διακυμάνσεως.

3. Ἐὰν ὑποτεθῇ ὅτι ποσότης τις x ὑπόκειται εἰς ***τυχαίας μεταβολὰς*** ἀφ' ἑνὸς καὶ ἀφ' ἑτέρου εἰς ἄλλας τοιαύτας προερχομένας ἐκ γνωστῶν παραγόντων Α, Β, Γ, ... τότε ἡ ποσότης αὕτη δύναται νὰ τεθῇ ὑπὸ τὴν μορφήν :

$$x = \overline{x} + \alpha + \beta + \ldots + \xi$$

ὅπου $\overline{x}$ ὁ μέσος τῶν x, α, β, ... αἱ ἐκ τῶν γνωστῶν παραγόντων Α, Β, ... ἀποκλίσεις καὶ ξ ἡ ἀπόκλισις ἥτις ὀφείλεται εἰς ***ἄγνωστον ἢ ἀγνώστους παράγοντας***, ἣν καὶ συμβατικῶς ἀποκαλοῦμεν ***τυχαῖον ἢ σφάλμα τυχαῖον*** ἢ ἁπλῶς σφάλμα.

Ἐκ τῆς ἄνω ἐκφράσεως λαμβάνομεν :

$$\Sigma (x-\overline{x})^2 = \Sigma (\alpha + \beta + \ldots + \xi)^2 \quad \text{ἢ} \quad \sigma_x^2 = \sigma_\alpha^2 + \sigma_\beta^2 + \ldots + \sigma_\xi^2$$

καὶ τοῦτο διότι αἱ ποσότητες τοῦ β' μέλους εἶναι ἀνεξάρτητοι—ἀσυσχέτιστοι—ἀλλήλων.

4. Ἡ ἀνάλυσις τῆς διακυμάνσεως ἀντιμετωπίζει δύο διακεκριμένα προβλήματα :

α) τὴν ἀνάλυσιν τῆς ὁλικῆς διακυμάνσεως εἰς τὰς μερικὰς αὐτῆς συνιστώ-

* Analysis of Variance.

σας—συνιστώσας ἀναφερομένας εἰς γνωστοὺς ἀφ' ἑνὸς παράγοντας καὶ ἀφ' ἑτέρου εἰς ἀγνώστους τοιούτους.

6) τὸν ἔλεγχον τῆς σημαντικότητος τῶν καθ' ἑκάστην συνιστωσῶν διακυμάνσεων.

Τὸ πρῶτον εἶναι καθαρῶς μηχανικὸν καὶ δύναται νὰ ἐφαρμόζηται εἰς κάθε περίπτωσιν, οὐχ ἧττον ἡ ἀξιοπιστία τῆς καθοριζομένης διακυμάνσεως ἐξαρτᾶται ἐν μέρει ἐκ τοῦ τρόπου καθ' ὃν ἔχουν συλλεγῆ αἱ παρατηρήσεις.

Ἐν τῇ ἀναλύσει τῆς διακυμάνσεως θὰ διακρίνωμεν δύο βασικὰς περιπτώσεις, ἀναλόγως ἂν αἱ παρατηρήσεις τάσσωνται πρὸς ἓν ἢ δύο κριτήρια (παράγοντας) κατατάξεως, ἅτινα δέον νὰ ἐλεγχθοῦν ἂν ἀσκοῦν ἢ ὄχι ἐπίδρασιν ἐπὶ τὴν διαμόρφωσιν τῶν τιμῶν τῆς μεταβλητῆς.

5. ***Κατάταξις κατὰ ἓν κριτήριον*** : Θεωροῦμεν ***τυχαῖον δεῖγμα*** ἐκ N τιμῶν ἑνὸς μεταβλητοῦ ἀκολουθοῦντος τὸν Νόμον τοῦ Gauss - Laplace. Τὸ μεταβλητὸν τοῦτο δύναται νὰ ταχθῇ εἰς στήλας, συμμόρφως πρός τ.να παράγοντα ἢ κριτήριον. Ἐὰν π. χ. τὸ μεταβλητὸν εἶναι ἡ τιμὴ ἑνὸς ἀγαθοῦ δεδομένου, τότε δύναται νὰ ταχθῇ εἰς στήλας ἀντιστοιχούσας εἰς τὰς διαφόρους τοῦ ἔτους ἐποχὰς ἢ εἰς τὰς διαφόρους περιοχὰς τῆς χώρας διὰ τὴν αὐτὴν ἐποχήν.

Τούτου τεθέντος ἐὰν x_{ij} παριστᾷ τὴν τιμὴν τῆς παρατηρήσεως j τῆς i στήλης, τότε ὁ πρῶτος δείκτης θὰ ἐμφαίνῃ τὴν στήλην καὶ ὁ δεύτερος τὴν θέσιν εἰς τὴν στήλην ταύτην. Ἐὰν n_i εἶναι ὁ ἀριθμὸς τῶν παρατηρήσεων τῆς i στήλης $\overline{x}_i$ ἡ μέση τιμὴ τῆς στήλης ταύτης καὶ $\overline{x}$ ὁ γενικὸς μέσος τοῦ ὅλου δείγματος τῶν N παρατηρήσεων, τότε :

$$N\overline{x} = \Sigma_i^{\kappa} \Sigma_j^h x_{ij}, \quad \overline{x}_i = \frac{\Sigma_j^h x_{ij}}{n_i} = \frac{T_i}{n_i}, \quad T_i = \Sigma_j^h x_{ij}$$

$$\Sigma_i^{\kappa} \Sigma_j^h (x_{ij} - \overline{x}) = 0, \qquad \Sigma_j^h (x_{ij} - \overline{x}_i) = 0$$

Ἀλλὰ ἡ ἀπὸ τοῦ γενικοῦ μέσου $\overline{x}$ ἀπόκλισις τῆς τυχούσης τιμῆς τοῦ x_{ij} δύναται νὰ γραφῇ καὶ ὡς ἑξῆς :

$$x_{ij} - \overline{x} = (\overline{x}_i - \overline{x}) + (x_{ij} - \overline{x}_i).$$

Ἐὰν ἤδη ὑψώσωμεν τὴν τελευταίαν σχέσιν εἰς τὸ τετράγωνον καὶ ἀθροίσωμεν διὰ πάσας τὰς τιμὰς τοῦ j λαμβάνομεν :

$$\sum_j^h (x_{ij} - \overline{x})^2 = n_i (\overline{x}_i - \overline{x})^2 + 2(\overline{x}_i - \overline{x}) \sum_j^h (x_{ij} - \overline{x}_i) + \sum_j^h (x_{ij} - \overline{x}_i)^2$$

ἀλλὰ $\displaystyle\sum_j^h (x_{ij} - \overline{x}_i) = 0$, ὅθεν :

$$\sum_i^h (x_{ij} - \overline{x})^2 = n_i (\overline{x}_i - \overline{x})^2 + \sum_j^h (x_{ij} - \overline{x}_i)^2$$

ἀθροίζοντες καὶ πάλιν διὰ πάσας τὰς τιμὰς τοῦ i κατὰ μέλη λαμβάνομεν

$$\sum_i^K \sum_j^h (x_{ij} - \overline{x})^2 = \sum_i^K n_i (\overline{x}_i - \overline{x})^2 + \sum_i^K \sum_j^h (x_{ij} - \overline{x}_i)^2 \qquad (1)$$

6. Ἐὰν ὅθεν τὸ πλῆθος ἐξ οὗ τὸ δεῖγμα τῶν N τιμῶν (τυχαῖον δεῖγμα) ἐλήφθη εἶναι ***ὁμοιογενὲς*** ὡς πρὸς τὸν παράγοντα κατατάξεως, τουτέστιν ὁ παράγων δὲν ἀσκεῖ ἐπίδρασιν ἐπὶ τῆς τιμῆς τῆς μεταβλητῆς, τότε, ἂν τὸ πλῆθος ὑποδιαιρεθῇ εἰς στήλας συμμόρφως πρὸς τὸν ὡς εἴρηται παράγοντα, αἱ διάφοροι στῆλαι θὰ ἔχουν τὰς αὐτὰς ***στατιστικὰς ἰδιότητας***. Εἰδικῶς θὰ ἔχουν τὸν αὐτὸν μέσον μ καὶ τὴν αὐτὴν διακύμανσιν σ^2, αἵτινες θὰ εἶναι ὁ μέσος καὶ ἡ διακύμανσις τοῦ ὑπ' ὄψει πλήθους. Συνεπῶς, ἐκ τῶν διαφόρων ἀθροισμάτων τῆς (1) δυνάμεθα νὰ εὕρωμεν τρεῖς ἀμερολήπτους ἐκτιμήσεις τῆς σ^2. Καθ' ὅσον ἡ ***μαθηματικὴ ἐλπὶς*** τοῦ πρώτου μέλους τῆς (1) εἶναι : $(N-1)\,\sigma^2$ ἐν τῇ δειγματοληψίᾳ καὶ ἑπομένως τοῦτο διαιρούμενον διὰ $(N-1)$ βαθμῶν ἐλευθερίας δίδει μίαν ἀμερόληπτον ἐκτίμησιν τοῦ σ^2. Ὁμοίως αἱ n_i τιμαὶ εἰς τὴν i στήλην ἀποτελοῦν ***τυχαῖον*** ἐπίσης, ***δεῖγμα***, οὗτινος ὁ μέσος εἶναι : $\overline{x}_i$.

Ἑπομένως ἡ μαθηματικὴ ἐλπὶς τοῦ $\sum\limits_j^h (x_{ij} - \overline{x}_i) = (n_i - 1)\,\sigma^2$ καὶ

ἄρα ἡ μαθηματικὴ ἐλπὶς τοῦ $\sum\limits_i^\kappa \sum\limits_j^h (x_{ij} - \overline{x})^2 = \sum\limits_i^\kappa (n_i - 1)\sigma^2 = (N-\kappa)\sigma^2$ ὅπου

κ ὁ ἀριθμὸς τῶν στηλῶν, παρέχει ἐπίσης μίαν ἀμερόληπτον ἐκτίμησιν τοῦ σ^2, στηριζομένην ἐπὶ $(N-\kappa)$ βαθμῶν ἐλευθερίας· καὶ ἐπειδὴ αἱ μαθηματικαὶ ἐλπίδες ἀμφοτέρων τῶν μελων τῆς (1) δέον νὰ εἶναι ἴσαι, τὸ πρῶτον ἄθροισμα τοῦ δευτέρου μέλους τῆς (1) θὰ εἶναι :

$$(\kappa - 1)\,\sigma^2 \quad \text{ἤτοι} \quad (N - 1)\,\sigma^2 = (\kappa - 1)\,\sigma^2 + (N - \kappa)\,\sigma^2 .$$

7. Κατ' ἀκολουθίαν τὸ σχῆμα τῆς ἀναλύσεως τῆς διακυμάνσεως λαμβάνει τὴν μορφήν :

πηγὴ μεταβολῆς	ἄθροισμα τετρ. ἀποκλίσεων	βαθμοὶ ἐλευθερίας	μέσον τετράγωνον
μεταξὺ στηλῶν	$\sum\limits_i^\kappa n_i (\overline{x}_i - \overline{x})^2$	$\kappa - 1$	$\sum\limits_i^\kappa n_i (\overline{x}_i - \overline{x}):(\kappa-1) = \sigma_1^2$
ἐντὸς στηλῶν	$\sum\limits_i^\kappa \sum\limits_j^h (x_{ij} - \overline{x})^2$	$N - \kappa$	$\sum\limits_i^\kappa \sum\limits_j^h (x_{ij} - \overline{x})^2 : (N-\kappa) = \sigma_0^2$
σύνολον	$\sum\limits_i^\kappa \sum\limits_j^h (x_{ij} - \overline{x})^2$	$N - 1$	

8. Εἰς τὴν περίπτωσιν ὁμοιογενοῦς πλήθους κανονικοῦ, τὸ πρῶτον μέλος τῆς (1) διαιρεθὲν διὰ σ^2 κατανέμεται ὡς χ^2 τοῦ Helmert—Pearson μὲ $(N-1)$ βαθμοὺς ἐλευθερίας. Ἑπομένως ἡ μέση τιμὴ τούτου εἶναι : $(N-1)\sigma^2$. Ὁμοίως $\sum\limits_j^h (x_{ij} - \overline{x}_i)^2 / \sigma^2$ εἶναι χ^2 μὲ $(n_i - 1)$ βαθμοὺς ἐλευθερίας καὶ συνεπῶς τὸ δεύτερον ἄθροισμα τοῦ β' μέλους τῆς (1) διαιρεθὲν διὰ σ^2 κατανέμεται ὡς χ^2 μὲ $\sum\limits_i^\kappa (n_i - 1) = N-\kappa$ βαθμοὺς ἐλευθερίας καὶ ἄρα ἡ μέση τιμὴ τοῦ ἀθροίσματος

τούτου εἶναι $(N-\kappa)\sigma^2$ · ὁμοίως διὰ τὸ πρῶτον ἄθροισμα τῆς (1) τοῦτο κατανέμεται ὡς χ^2 μὲ $(\kappa-1)$ βαθμοὺς ἐλευθερίας καὶ ἄρα ἡ μέση τιμὴ αὐτοῦ εἶναι $(\kappa-1)\sigma^2$.

9. Διὰ νὰ ἐλέγξωμεν τὴν ὁμοιογένειαν τῶν ἐκτιμήσεων τοῦ σ^2 διὰ τοῦ F τοῦ Snedecor, θὰ ὑποθέσωμεν ὅτι τὸ πλῆθος εἶναι τοῦ τύπου Gauss—Laplace καὶ θὰ συγκρίνωμεν τὴν διακύμανσιν ***μεταξὺ στηλῶν*** πρὸς τὴν τοιαύτην ***ἐντὸς τῶν στηλῶν*** μὲ $(\kappa-1)$ καὶ $(N-\kappa)$ βαθμοὺς ἐλευθερίας καὶ τοῦτο διότι ἡ πρώτη ἐκφράζει τὴν μεταβολὴν ἥτις ἐπέρχεται εἰς τὸν παράγοντα τῆς κατατάξεως καὶ ἡ δευτέρα εἶναι ἡ τῆς καταλοίπου μεταβολῆς μετὰ τὴν ἀφαίρεσιν τῆς προηγουμένης. Ἐὰν σημαντικῶς ἡ $\sigma_1^2 > \sigma_0^2$ τότε δικαιολογούμεθα συμπεραίνοντες ὅτι ὁ παράγων κατατάξεως ἀσκεῖ ἐπίδρασιν ἐπὶ τῆς τιμῆς τῆς μεταβλητῆς καὶ συνεπῶς ἡ ὑπόθεσις τῆς ὁμοιογενείας τοῦ πλήθους ἀπορρίπτεται.

10. Ἐὰν αἱ ἐκτιμήσεις τῶν διακυμάνσεων δὲν εἶναι σημαντικῶς διάφοροι, τὸ κριτήριον δὲν παρέχει τὸ ἐνδόσιμον ἐναντίον τῆς ὑποθέσεως τοῦ ὁμοιογενοῦς πλήθους. Δέον ὅμως νὰ ἔχωμεν ὑπ' ὄψει πάντοτε ὅτι ὁ λόγος τῶν διακυμάνσεων ἐλέγχεται διὰ τοῦ F μόνον ἂν αἱ δύο ἐκτιμήσεις τῶν διακυμάνσεων εἶναι ***στατιστικῶς ἀνεξάρτητοι***, ἐπειδὴ ὁ μέσος ἑνὸς τυχαίου δείγματος ἀπό τινος κανονικοῦ πλήθους κατανέμεται ἀνεξαρτήτως τῶν διακυμάνσεών του, τὰ δύο ἀθροίσματα τοῦ β' μέλους τῆς (1) εἶναι ἀνεξάρτητα ἀλλήλων καὶ συνεπῶς ἡ ἀπαιτουμένη συνθήκη πληροῦται.

11. Πρὸς εὐκολίαν τῶν ὑπολογισμῶν θέτομεν

α) $$\sum_i^\kappa \sum_j^h (x_{ij} - \overline{x})^2 = \sum_i^\kappa \sum_j^h x_{ij}^2 - 2\overline{x}\sum_i^\kappa \sum_j^h + N\overline{x}^2$$

ἀλλὰ

$$\overline{x} = \frac{T}{N} = \frac{\Sigma T_i}{N}, \qquad T_i = \sum_j^h x_{ij}, \qquad T = \Sigma T_i, \qquad \overline{x}_i = \frac{T_i}{n_i}$$

ἄρα $$\sum_i^\kappa \sum_j^h (x_{ij} - \overline{x})^2 = \sum_i^\kappa \sum_j^h x_{ij}^2 - \frac{2T^2}{N} + \frac{T^2}{N} = \sum_i^\kappa \sum_j^h x_{ij}^2 - T^2/N$$

β) $$\sum_i^\kappa n_i(\overline{x}_i - \overline{x})^2 = \sum_i^\kappa n_i \overline{x}_i^2 - 2\overline{x}\sum_i^\kappa n_i \overline{x}_i + N\overline{x}^2$$

ἀλλὰ $$n_i \overline{x}_i^2 = \frac{T_i^2}{n_i}$$ ὅθεν

$$\sum_i^\kappa n_i(\overline{x}_i - \overline{x})^2 = \sum_i^\kappa \frac{T_i^2}{n_i} - \frac{2T}{N}\Sigma T_i + \frac{T^2}{N} = \sum_i^\kappa \frac{T_i^2}{n_i} - \frac{2T^2}{N} + \frac{T^2}{N} =$$

$$= \sum_i^\kappa \frac{T_i^2}{n_i} - \frac{T^2}{N}.$$

γ) $$\sum_i^\kappa \sum_i^h (x_{ij} - \overline{x})^2 = \sum_i^\kappa \sum_j^h (x_{ij} = \overline{x})^2 - \sum_i^\kappa n_i(\overline{x}_i - \overline{x})^2 =$$

$$= \sum_i^\kappa \sum_i^h x_{ij}^2 - \sum_i^\kappa \frac{T_i^2}{n_i},$$

τῆς τελευταίας ταύτης σχέσεως προκυπτούσης ἐκ τῆς (1) εὐκόλως.

12. ***Ἐφαρμογή*** : Ὀκτὼ ποικιλίαι βρώμης ἐσπάρησαν εἰς 4 πειραματικοὺς ἀγροὺς ἴσης ἐκτάσεως καὶ γονιμότητος. Αἱ ἀποδόσεις τούτων δίδονται κατωτέρω.

Π ο ι κ ι λ ί α ι

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
182	196	203	198	171	194	208	183
214	202	212	203	192	218	216	188
216	208	221	207	197	223	218	193
231	224	242	222	204	232	239	198

Γεννᾶται τὸ ἐρώτημα : Διαφέρουν σημαντικῶς ;

Ἐνταῦθα ὑφίστανται $\kappa = 8$ στῆλαι (ποικιλίαι) καὶ $h = 4$ δηλ. ὑπάρχουν $N = \kappa h = 32$ παρατηρήσεις ἐν ὅλῳ. Πρὸς εὐκολίαν τῶν ὑπολογισμῶν ἀφαιροῦμεν ἐξ ἑκάστης παρατηρήσεως τὸ 200 ὁπόταν ὁ μέσος ἐλαττοῦται κατὰ 200, ἀλλὰ ἡ διακύμανσις μένει ἀναλλοίωτος. Κατὰ ταῦτα ἔχομεν

Π ο ι κ ι λ ί α ι

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	
	−18	−4	3	−2	−29	−6	8	−17	
	14	2	12	3	− 8	18	16	−12	
	16	8	21	7	− 3	23	18	− 7	
	31	24	42	22	4	32	39	− 2	
T_i	43	30	78	30	−36	67	81	−38	$T = \Sigma T_i = 255$

θὰ ἔχωμεν $\sum_i^{\kappa} \sum_j^{h} (x_{ij} - \overline{x})^2 = \sum_i^{\kappa} \sum_j^{h} x_{ij}^2 - T^2/N = (-18)^2 + (14)^2 + \ldots$

$$+ (-7)^2 + (-2)^2 - T^2/N = 10795 - \frac{(255)^2}{32} = 10795 - 2032{,}03 = 8762{,}97$$

ὁμοίως ἐνταῦθα $n_i = 4$, ἄρα

$$\sum_i^{\kappa} n_i (\overline{x}_i - \overline{x})^2 = \frac{\sum_i^{\kappa} T_i^2}{n} - \frac{T^2}{N} = \frac{(43)^2 + (30)^2 + (78)^2 + \ldots + (81)^2 + (-38)^2}{4} -$$

$$- 2032{,}03 = \frac{23523}{4} - 2032{,}03 = 3848{,}72$$

$$\sum_i^{\kappa} \sum_j^{h} (x_{ij} - \overline{x}_i)^2 = \sum_i^{\kappa} \sum_j^{h} x_{ij}^2 - \frac{\sum_i^{\kappa} T_i^2}{n} = 10\,795 - \frac{23\,523}{4} = 10\,795 - 5880{,}75 =$$

$$= 4\,914{,}25$$

Κατὰ ταῦτα ἡ ἀνάλυσις τῆς διακυμάνσεως διατυποῦται ὡς κάτωθι :

πηγὴ μεταβολῆς	ἄθροισ. τετραγ. ἀποκλ.	βαθ. ἐλευθερίας	μέσον τετράγωνον
Μεταξὺ ποικιλιῶν	3848,72	7	3848,72 : 7 ≅ 549,17
ἐντὸς ποικιλιῶν	4914,25	24	4914,25 : 24 ≅ 204,76
Σύνολον	8762,97	31	

Σχηματίζομεν τὸ $F_1 = 549{,}17 : 204{,}76 = 2{,}69$ μὲ $\nu_1 = 7$ καὶ $\nu_2 = 24$ βαθμοὺς ἐλευθερίας, ἀλλὰ εἰς τὸ ἐπίπεδον τοῦτο τῶν 5 % σημαντικότητος ἡ τιμὴ τοῦ $F(7,24) = 2{,}42$. Ἑπομένως ***αἱ διαφοραὶ τῶν ἀποδόσεων μεταξὺ τῶν ποικιλιῶν, πιθανώτατα, εἶναι πραγματικαί, μὴ ὀφειλόμεναι εἰς τὰς τυχαίας διακυμάνσεις τῆς δειγματοληψίας.***

13. Ἐπειδὴ ὅθεν ἡ ὑπόθεσις τῆς ὁμοιογενείας ἀπορρίπτεται, ἐπιτρέπεται νῦν νὰ ἀναζητήσωμεν τὴν σύγκρισιν μεταξὺ δύο ὡρισμένων μέσων ἐκ τῶν κ παρατηρηθέντων τοιούτων. Πρὸς τοῦτο ἔστωσαν οἱ $\overline{x}_i$ καὶ $\overline{x}_j$ μέσοι ἀναφερόμενοι εἰς τὸν αὐτὸν ἀριθμὸν παρατηρήσεων ἕκαστος ($n=4$). Ἡ διακύμανσις τῆς μεταβλητῆς ἔχει ὡς ἐκτίμησιν σ_0^2, ἐπομένως ἡ τοιαύτη δείγματος ἐκ n εἶναι σ_0^2/n καὶ συνεπῶς τῆς διαφορᾶς δύο μέσων $\sigma_0^2/n + \sigma_0^2/n = 2\sigma_0^2/n$. Κατ᾽ ἀκολουθίαν ἡ μέση ἀπόκλισις τετραγώνου τῆς διαφορᾶς τῶν δύο μέσων εἶναι:

$$s_0 = \sigma_0 \sqrt{2/n}\,.$$

Ἡ σημαντικότης τῆς διαφορᾶς νῦν ἐλέγχεται διὰ τοῦ t τοῦ Student:

$$t = \frac{\overline{x}_i - \overline{x}_j}{s_0} = \frac{(\overline{x}_i - \overline{x}_j)\sqrt{n}}{\sigma_0\sqrt{2}}$$

μὲ N—κ βαθμοὺς ἐλευθερίας. Ἀλλὰ N—κ=32—8=24 βαθμοὶ ἐλευθερίας· ὅθεν

$$s_0^2 = 2\,(\,204{,}76\,)\,/\,4 = 102{,}38 \quad ἢ \quad s_0 = \sqrt{102{,}30} = \pm\ 10{,}12$$

14. Ἐὰν $\overline{x}_1 = 81/4 = 20{,}25$ $\overline{x}_2 = 30/4 = 7{,}5$, δηλ. 7η καὶ 3η ποικιλία, τότε $t = \dfrac{20{,}25 - 7{,}5}{10{,}12} = \dfrac{12{,}75}{10{,}12} = 1{,}25$ ἥτις δὲν εἶναι σημαντικὴ εἰς $\nu = 24$ βαθμοὺς ἐλευθερίας. Ἐὰν m κληθῇ ἡ διαφορὰ τῶν μέσων ἥτις εἶναι σημαντικὴ εἰς 5 % μὲ $\nu = 24$ βαθμοὺς ἐλευθερίας τότε $t = 2{,}06 = m/10{,}12$ ἢ $m = 2{,}06 \times 10{,}12 = 20{,}85$ ἐπομένως μία διαφορὰ τῶν μέσων ἴση πρὸς $m = 20{,}85$ θὰ ἠδύνατο νὰ ὑπερβληθῇ μόνον 5 %. Οὕτω αἱ μόναι σημαντικαὶ διαφοραὶ εἶναι μεταξὺ τῶν καλυτέρων ἀποδόσεων 7,3 καὶ 6 καὶ τῶν χειροτέρων 5 καὶ 8.

15. ***Κατάταξις κατὰ δύο κριτήρια.*** Κατὰ τὴν ὡς εἴρηται περίπτωσιν αἱ N τιμαὶ x_{ij} τάσσονται κατὰ δύο διάφορα κριτήρια A καὶ B, ἐξ ὧν τὸ A προσδιορίζει h διαφόρους στήλας καὶ τὸ B, κ διαφόρους γραμμάς, ὥστε ὑφίστανται $N = h\kappa$, τιμαὶ τῆς μεταβλητῆς καὶ κατὰ τοιοῦτον τρόπον διατεταγμέναι ὥστε εἰς ἑκάστην τῶν h στηλῶν νὰ ὑφίσταται μία μόνη τιμὴ ἐξ ἑκάστης γραμμῆς καὶ εἰς ἑκάστην γραμμὴν μία μόνη τιμὴ ἐξ ἑκάστης στήλης. Ἡ κατάταξις συνεπῶς λαμβάνει μορφὴν ***ὀρθογωνίου*** μὲ h στήλας καὶ κ γραμμὰς καὶ ὅπου ὁ δείκτης ij τοῦ x_{ij} ἐμφαίνει τιμὴν τῆς παρατηρήσεως εὑρισκομένης εἰς τὴν i στήλην καὶ j γραμμήν. Τούτου τεθέντος, ἡ ἀπόκλισις $x_{ij} - \overline{x}$, ὅπου $\overline{x}$ ὁ γενικὸς μέσος τοῦ συνόλου τῶν παρατηρήσεων, δύναται νὰ γραφῇ:

$$x_{ij} - \overline{x} = (\overline{x}_i - \overline{x}) + (\overline{x}_j - \overline{x}) + (x_{ij} - \overline{x}_i - \overline{x}_j + \overline{x})$$

ὅπου $\overline{x}_i$ ὁ μέσος τῆς i στήλης καὶ $\overline{x}_j$ ὁ μέσος τῆς j γραμμῆς. Ἐὰν νῦν ἀμφότερα τὰ μέλη τῆς ἄνω ἐκφράσεως ὑψώσωμεν εἰς τὸ τετράγωνον καὶ ἀθροίσωμεν διὰ πάσας τὰς τιμὰς τοῦ j θὰ ἔχωμεν :

$$\sum_j^{\kappa} (x_{ij} - \overline{x})^2 = \kappa(\overline{x_i} - \overline{x})^2 + \sum_j^{\kappa} (\overline{x_j} - \overline{x})^2 + \sum_j^{\kappa} (x_{ij} - \overline{x_i} - \overline{x_j} + \overline{x})^2$$

τῶν διπλῶν ἀθροισμάτων εἰς τὸ β΄ μέλος μηδενιζομένων, δι' οὓς λόγους ἐλέχθη προκειμένης τῆς κατατάξεως κατὰ ἓν κριτήριον. Ἀθροίζοντες καὶ αὖθις τὴν ἄνω διὰ πάσας τὰς τιμὰς τοῦ i ἔχομεν :

$$\sum_i^{h}\sum_j^{\kappa}(x_{ij} - \overline{x})^2 = \kappa\sum_i^{h}(\overline{x_i} - \overline{x})^2 + h\sum_j^{\kappa}(\overline{x_j} - \overline{x})^2 + \sum_i^{h}\sum_j^{\kappa}(x_{ij} - \overline{x_i} - \overline{x_j} + \overline{x})^2 \qquad (2)$$

οὕτω ἡ διακύμανσις περὶ τὸν μέσον ἀνελύθη εἰς τρεῖς συνιστώσας, ἤτοι :

α) τὴν διακύμανσιν τῶν μέσων τῶν στηλῶν περὶ τὸν γενικὸν μέσον, β) ὁμοίως τὴν τοιαύτην τῶν μέσων τῶν γραμμῶν περὶ τὸν γενικὸν μέσον καὶ γ) τὴν κατάλοιπον τοιαύτην.

16. Ἡ μαθηματικὴ ἐλπὶς τοῦ α΄ μέλους τῆς (2) εἶναι $(h\kappa-1)\sigma^2 = (N-1)\sigma^2$ ὁμοίως βάσει τῆς ὑποθέσεως ὁμοιογενοῦς πλήθους, αἱ μαθηματικαὶ ἐλπίδες τῶν δύο πρώτων ἀθροισμάτων τοῦ β΄ μέλους τῆς (2) εἶναι ἀντιστοίχως $(h-1)\sigma^2$ καὶ $(\kappa-1)\sigma^2$ καὶ συνεπῶς τοῦ τελευταίου ὅρου

$$[h\kappa - (h-1) - (\kappa-1)]\sigma^2 = (h-1)(\kappa-1)\sigma^2$$

Κατὰ ταῦτα τὰ διάφορα ἀθροίσματα τῆς (2), διαιρούμενα κατὰ σειρὰν διὰ τῶν $(h\kappa-1) = (N-1)$, $(h-1)$, $(\kappa-1)$, $(h-1)(\kappa-1)$ παρέχουν ἀμερολήπτους ἐκτιμήσεις τῆς σ^2, στηριζομένας ἐπὶ βαθμῶν ἐλευθερίας δηλουμένων ὑπὸ τῶν ἀντιστοίχων διαιρετῶν. Συνεπῶς, καὶ τὰ ἀθροίσματα διαιρούμενα διὰ σ^2 κατανέμονται ὡς χ^2 μὲ βαθμοὺς ἐλευθερίας τοὺς $(N-1)$, $(h-1)$, $(\kappa-1)$, $(h-1)(\kappa-1)$ καὶ τοῦτο ὑπὸ τὴν προϋπόθεσιν ὅτι τὸ πλῆθος ἐφ' οὗ τὸ δεῖγμα ἔχει ληφθῆ εἶναι τοῦ τύπου Gauss—Laplace.

17. Τὰ πορίσματα τῆς ἀναλύσεως διατυποῦνται ὡς κάτωθι :

πηγὴ μεταβολῆς	ἄθρ. τετρ. ἀποκλίσεων	βαθμοὶ ἐλευθερίας	μέσον τετράγωνον
μεταξὺ στηλῶν	$\kappa\sum_i^{h}(\overline{x_i} - \overline{x})^2$	$h-1$	$\sigma_2^2 = \kappa\sum_i^{h}(\overline{x_i} - \overline{x})^2 : (h-1)$
μεταξὺ γραμμῶν	$h\sum_j^{\kappa}(\overline{xj} - \overline{x})^2$	$\kappa-1$	$\sigma_1^2 = h\sum_j^{\kappa}(\overline{x_j} - \overline{x})^2 : (\kappa-1)$
σφάλμα ἢ κατάλοιπος μεταβολὴ	$\sum_i^{h}\sum_j^{\kappa}(x_{ij} - \overline{x_i} - \overline{x_j} + \overline{x})^2$	$(h-1)\cdot(\kappa-1)$	$\sigma_0^2 = \sum_i^{h}\sum_j^{\kappa}(x_{ij} - \overline{x_i} - \overline{x_j} + \overline{x})^2 : (h-1)(\kappa-1)$
σύνολον	$\sum_i^{h}\sum_j^{\kappa}(x_{ij} - \overline{x})^2$	$(h\kappa-1)$	

18. Διὰ νὰ ἐλεγχθῆ ἡ ὁμοιογένεια τοῦ πλήθους, συγκρίνομεν τὰ σ_2^2 καὶ

σ_1^2 πρὸς σ_0^2. Ἐὰν ὁ ἀντιστοιχῶν παράγων πρὸς ἑκατέραν τῶν κατατάξεων ἀσκῇ σημαντικὴν ἐπίδρασιν ἐπὶ τῆς τιμῆς τῆς μεταβλητῆς, τοῦτο θὰ ἐμφανισθῇ εἰς τὸ ἀντιστοιχοῦν αὐτῇ μέσον τετράγωνον. Πρὸς ἔλεγχον διὰ τοῦ F τοῦ Snedecor θὰ ὑποτεθῇ τὸ πλῆθος τύπου Gauss — Laplace. Ἐὰν σ_2^2 ἢ σ_1^2 εἶναι σημαντικῶς διάφορα τοῦ σ_0^2, τότε ἡ ὑπόθεσις τῆς ὁμοιογενείας τοῦ πλήθους ἀπορρίπτεται. Ἡ σημαντικότης τῆς διαφορᾶς τῶν μέσων τῶν στηλῶν ἢ τῶν γραμμῶν ἐλέγχεται διὰ τοῦ t τοῦ Student, ὡς ἤδη ἐξετέθη.

19. Διὰ τοὺς πρακτικοὺς ὑπολογισμοὺς θέτομεν :

$$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x})^2 = \sum_i^h \sum_j^\kappa x_{ij}^2 - T^2/N\,, \qquad T = \sum_i^h \sum_j^\kappa x_{ij} = \sum_i^h T_i = \sum_j^\kappa T_j$$

καὶ
$$T_i = \sum_j^\kappa x_{ij}\,, \quad T_j = \sum_i^h x_{ij}\,, \quad \overline{x}_i = \frac{T_i}{\kappa}\,, \quad \overline{x}_j = \frac{T_j}{h}\,, \quad \overline{x} = \frac{T}{N}$$

ὁμοίως

$$\kappa \sum_i^h (\overline{x}_i - \overline{x})^2 = \kappa \sum_i^h \overline{x}_i^{\,2} - 2\,\kappa\, \overline{x} \sum_i^h \overline{x}_i + \kappa h\, \overline{x}^{\,2} =$$

$$\sum_i^h \frac{T_i^2}{\kappa} - \frac{2T^2}{N} + \frac{T^2}{N} = \sum_i^h \frac{T_i^2}{\kappa} - \frac{T^2}{N}$$

$$h \sum_i^\kappa (\overline{x}_j - \overline{x})^2 = \sum_j^\kappa \frac{T_j^2}{h} - \frac{T^2}{N}$$

καὶ τέλος

$$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x}_i - \overline{x}_j - \overline{x})^2 = \sum_i^h \sum_j^\kappa x_{ij}^2 - \left(\frac{\Sigma\, T_i^2}{\kappa} + \frac{\Sigma\, T_j^2}{h} - \frac{T^2}{N} \right)$$

ἥτις προκύπτει εὐκόλως ἐκ τῆς (2) γραφομένης

$$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x})^2 - \kappa \sum_i^h (\overline{x}_i - \overline{x})^2 - h \sum_j^\kappa (\overline{x}_j - \overline{x})^2 =$$

$$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x}_i - \overline{x}_j + \overline{x})^2$$

20. ***Ἐφαρμογή*** : Οἱ μέσοι τριμηνιαῖοι ναῦλοι αὐτοκινήτων εἰς τὰς Η.Π.Α. διὰ τὴν περίοδον 1924—1928 εἶναι :

ἔτη	τρίμηνα 1	2	3	4
1924	890	890	970	970
1925	920	970	1050	1010
1926	940	1010	1100	1060
1927	970	1000	1050	970
1928	900	980	1060	1040

Ἐνταῦθα ὑφίστανται δύο παράγοντες κατατάξεως: α) τὰ τρίμηνα (n = 4) καὶ β) τὰ ἔτη (κ=5) μὲ hκ=20=N ἐν συνόλῳ παρατηρήσεις.

Πρὸς εὐκολίαν τῶν ὑπολογισμῶν ἀφαιροῦμεν ἐξ ἑκάστης τιμῆς τὸ 1 000 καὶ ἔχομεν

τ ρ ί μ η ν α

ἔτη	1	2	3	4	T_j
1924	—110	—110	—30	—30	= —280
1925	— 80	—30	50	10	= — 50
1926	— 60	10	100	60	= 110
1927	— 30	0	50	—30	= — 10
1928	—100	— 20	60	40	= — 20
T_i =	—380	—150	230	50	T = —250

α) $$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x})^2 = \sum_i^h \sum_j^\kappa x_{ij}^2 - T^2/N = (-110)^2 + (-80)^2 + \ldots +$$

$$(-30)^2 + (40)^2 - T^2/N = 73\,100 - 3\,125 = 69\,975$$

β) $$K \sum_i^h (\overline{x_i} - \overline{x})^2 = \sum_i^h T_i^2 / \kappa - T^2/N = \frac{(-380)^2 + (-150)^2 + (230)^2 + (50)^2}{5} -$$

$$- 3125 = 44\,460 - 3\,125 = 41\,335$$

γ) $$h \sum_j^\kappa (\overline{x_j} - \overline{x})^2 = \frac{\sum_i^h T_j^2}{h} - \frac{T^2}{N} = \frac{(-280)^2 + (50)^2 + (110)^2 + (-10)^2 + (-20)^2}{4}$$

$$-3\,125 = 23\,375 - 3\,125 = 20\,250$$

δ) $$\sum_i^h \sum_j^\kappa (x_{ij} - \overline{x_i} - \overline{x_j} + \overline{x})^2 = 69\,975 - (41\,335 + 20\,250) = 8\,390$$

Ἡ ἀνάλυσις τῆς διακυμάνσεως ἔχει ὡς κάτωθι:

πηγὴ μεταβολῆς	ἄθροισμα τετρ. ἀποκλίσεων	βαθμ. ἐλευθερίας	μέσον τετράγωνον
Μεταξὺ τριμήνων	41 335	3	$\sigma_2^2 = 13778{,}3$
Μεταξὺ ἐτῶν	20 250	4	$\sigma_1^2 = 5062{,}5$
Κατάλοιπος μεταβολὴ	8 390	12	$\sigma_0^2 = 699{,}1$
Σύνολον	69 975	19	

Σχηματίζομεν τὸ $F_1 = \dfrac{13778{,}3}{699{,}1} = 19{,}7$ μὲ $\nu_1 = 3$, $\nu_2 = 12$, ἀλλὰ F(3,12)

$=3,49$ εἰς 5 % ἐπίπεδον σημαντικότητος. Ὁμοίως $F_2 = \frac{5062,5}{699,1} = 7,27$ μὲ $\nu_1 = 4$, $\nu_2 = 12$, ἀλλὰ $F(4,12) = 3,11$. δηλ. αἱ δύο διακυμάνσεις εἶναι στατιστικῶς σημαντικαὶ ἑπομένως καὶ αἱ ἐποχαὶ καὶ τὰ ἔτη ἐπηρεάζουν τὴν τιμὴν τῶν ναύλων σημαντικῶς.

21. ***Προσδιορισμὸς συνιστωσῶν διακυμάνσεως.*** Ἐπειδὴ (πρώτη ἐφαρμογὴ) $205,76 \rightarrow \sigma_0^2$ (τὸ σημεῖον $\rightarrow$ ἐμφαίνει ὅτι ἡ πρὸς ἀριστερὰ ποσότης ἀποτελεῖ ἐκτίμησιν τῆς πρὸς τὰ δεξιὰ καὶ ὅτι ἡ πρώτη προσεγγίζει τὴν δευτέραν ὅσον τὸ μέγεθος τοῦ δείγματος αὐξάνει) καὶ $549,17 \rightarrow \sigma_0^2 + 8\sigma_1^2$, ἔπεται ὅτι

$$8\sigma_1^2 = 549,17 - 204,76 = 344,41 \text{ ἢ } \sigma_2^2 = 43,05.$$

Κατ᾽ ἀκολουθίαν

Διακύμανσις μεταξὺ ποικιλιῶν	43,05	ἢ	17,4 %
Διακύμανσις ἐντὸς ποικιλιῶν	204,76	ἢ	82,6 %
σύνολον	247,81		100 %

ἤτοι ἡ διακύμανσις ἐντὸς τῶν ποικιλιῶν εἶναι τὰ 80 % καὶ πλέον τῆς ὅλης τοιαύτης.

22. Προκειμένης ἤδη τῆς ἀναλύσεως τῆς διακυμάνσεως κατατάξεως κατὰ δύο κριτήρια θὰ ἔχωμεν καὶ πάλιν $699,1 \rightarrow \sigma_0^2$, $5062,5 \rightarrow \sigma_0^2 + 4\sigma_1^2$ καὶ $13778,3 \rightarrow \sigma_0^2 + 5\sigma_0^2$ ἤτοι $\sigma_0^2 = 699,1$, $4\sigma_1^2 = 4363,40$ ἢ $\sigma_1^2 = 1090,85$ καὶ

$$5\sigma_2^2 = 13079,2 \text{ ἢ } \sigma_2' = 2615,84$$

Συνεπῶς

Διακύμανσις μεταξὺ τριμήνων	2616,84	ἢ	59,3 %
Διακύμανσις μεταξὺ ἐτῶν	1090,85	ἢ	24,7 %
Κατάλοιπος διακύμανσις	699,10	ἢ	16,0 %
Σύνολον	4406,79		100 %

Δηλαδή, ἡ μείζων μεταβολὴ εἰς τοὺς ναύλους προέρχεται ἐξ ἐποχιακῶν λόγων καὶ κατὰ δεύτερον λόγον ἐκ τῆς παρόδου τῶν ἐτῶν, δηλ. τῆς τάσεως.

Συμπέρασμα. Ἡ ἀνάλυσις τῆς διακυμάνσεως εἶναι πολύτιμον ἐργαλεῖον διότι ἐπιτρέπει νὰ διακριβώνωμεν τὸν βαθμὸν τῆς ἀκριβείας τῶν πορισμάτων.

ΠΑΡΟΡΑΜΑΤΑ

Παρακαλεῖται ὁ ἀναγνώστης νὰ προβῇ εἰς τὰς κάτωθι διορθώσεις λαθῶν τὰ ὁποῖα παρεισέφρυσαν ἐκ τυπογραφικῆς ἀβλεψίας καὶ ἀλλοιώνουν τὴν ἔννοιαν τοῦ κειμένου.

Σελὶς 487 στῖχος 28 ἀντὶ $\xi_i \xi_\kappa = 0$ νὰ γραφῇ $\Sigma\xi_i \xi_\kappa = 0$

» 488 » 21 » $\alpha_n = \Sigma\xi^2_n = \Sigma\xi_n y$ νὰ γραφῇ $\alpha_n \Sigma\xi^2_n = \Sigma\xi_n y$

» 494 » 13 » $\sum_j^h (x_{ij} - \bar{x}_i) =$ νὰ γραφῇ $\sum_j^h (x_{ij} - \bar{x}_i)^2$

» 494 » 22 » $\sum_i^\kappa n_i (\bar{x}_i - \bar{x})$: νὰ γραφῇ $\sum_i^\kappa n_i (\bar{x}_i - \bar{x})^2$:

» 495 » 21 » $\sum_i^\kappa \sum_j^h x_{ij}^3$ νὰ γραφῇ $\sum_i^\kappa \sum_j^h x_{ij}^2$