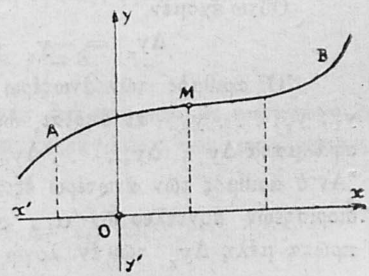
Η ΜΕΘΟΔΟΣ ΤΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ ΕΙΣ ΤΗΝ ΣΤΑΤΙΣΤΙΚΗΝ

ΥΠΟ ΤΟΥ ΚΑΘΗΓΗΤΟΥ Κ. ΝΕΙΛΟΥ ΣΑΚΕΛΛΑΡΙΟΥ

Δύο ποσότητες ἢ μεγέθη, ἔστωσαν x, y, σχετίζονται μεταξύ των ἐνίοτε κατὰ τοιοῦτον τρόπον, ὥστε εἰς μεταβολὴν τοῦ ἑνὸς ἐξ αὐτῶν νὰ ἀντιστοιχῇ μεταβολὴ καὶ τοῦ ἄλλου. Ἐὰν εἶνε καθωρισμένος ὁ τρόπος τῆς ἐξαρτήσεως αὐτῶν καὶ δοθῇ τιμή τις εἰς τὸ ἓν ἐκ τούτων π.χ. εἰς τὸ x, ἠμποροῦμεν νὰ εὕρωμεν τὴν ἀντίστοιχον της τιμὴν τοῦ y. Ἐνίοτε, ὅμως, ὁ τρόπος τῆς ἐξαρτήσεως δὲν εἶνε πλήρως καθωρισμένος καὶ ἰδίᾳ ὅταν πρόκειται περὶ ζητήματος στατιστικῆς. Ἐὰν π.χ. πρόκειται περὶ τῶν ἐσόδων μιᾶς κοινότητος, ἢ ἑνὸς κράτους, ἢ μιᾶς ἐπιχειρήσεως κατὰ διάφορα χρονικὰ διαστήματα, π.χ. κατὰ τὰ διάφορα ἔτη, δὲν γνωρίζομεν τὸν τρόπον τῆς ἐξαρτήσεως τῶν δύο ποσῶν, χρόνου (εἰς ἔτη π.χ.) καὶ εἰσοδήματος (εἰς χρῆμα π.χ.), ἐκ παρατηρήσεων δ᾽ ἔχομεν ἀντιστοίχους τιμὰς αὐτῶν. Π.χ. ἔστω ὅτι κατὰ τὰ ἔτη 1948, 1949, 1950, 1951, 1952 ἔχομεν ἀντίστοιχα εἰσοδήματα παριστάμενα μὲ E_{48}, E_{49}, E_{50}, E_{51}, E_{52}, δισεκατομύρια δραχμῶν. Καθ᾽ ὅμοιον τρόπον, ἀπὸ μίαν σειρὰν μετρήσεων π.χ. προκύπτουν δύο σειραὶ ἀριθμῶν, ἐκ τῶν ὁποίων ἡ μία παριστάνει τὰς τιμὰς τοῦ x καὶ ἡ ἄλλη τὰς ἀντιστοίχους τοῦ y. Ἐπιδιώκομεν ἐπὶ τῇ βάσει τῶν δύο ἐν λόγῳ σειρῶν τιμῶν τοῦ x καὶ y νὰ εὕρωμεν τὸν τρόπον τῆς ἐξαρτήσεως αὐτῶν, ἤτοι μίαν ἐξίσωσιν μεταξὺ τῶν x καὶ y, π.χ. τὴν $y=\phi(x)$, τοιαύτην ὥστε εἰς διδομένην τιμὴν τοῦ x νὰ ὑπολογίζεται ἡ ἀντίστοιχος τιμὴ τοῦ y, ἂν δὲ θεωρηθοῦν τὰ x, y ὡς συντεταγμέναι σημείου τοῦ ἐπιπέδου xoy ὡς πρὸς ἄξονας (ὀρθογωνίους) x′Ox, y′Oy, νὰ κατασκευάσωμεν κατὰ τὸ δυνατόν, τὴν ὑπὸ τῆς $y=\phi(x)$ παριστανομένην καμπύλην, ἔστω ΑΜΒ τοῦ ἔναντι σχήματος.



Ἡ σημασία τῆς εὑρέσεως τῆς $y=\phi(x)$ καὶ τῆς κατασκευῆς τῆς παραστατικῆς καμπύλης ΑΜΒ τῆς ἐξισώσεως αὐτῆς ἔγκειται εἰς τὸ ὅτι, θὰ δυνάμεθα νὰ ἔχωμεν τὴν τιμὴν τοῦ y δι᾽ οἱανδήποτε τιμὴν (πραγματικὴν) τοῦ x, ἐπὶ πλέον δὲ θὰ ἔχωμεν τὴν εἰκόνα τῆς πορείας τῆς τιμῆς τοῦ y. Οὕτω π. χ. ἂν ἔχωμεν τὴν ἐν λόγῳ ἐξίσωσιν $y=\phi(x)$ καὶ τὴν παριστάνουσαν αὐτὴν καμπύλην διὰ τὴν περίπτωσιν τοῦ ἀνωτέρω παραδείγματος τοῦ χρόνου καὶ τῶν ἐσόδων κοινότητος ἢ κράτους π.χ., θὰ δυνάμεθα νὰ εὕρωμεν τὸ εἰσόδημα εἰς οἱανδήποτε ἐποχήν, π.χ. κατὰ τὰ ἔτη 1953, 1954,..., 1960, καθὼς καὶ κατὰ παρελθόντα ἔτη 1947, 1946, 1945 κλπ., διότι αἱ ἀντίστοιχοι τιμαὶ τοῦ y θὰ δίδωνται ὑπὸ τοῦ $\phi(1953)$, $\phi(1954)$, . . . , $\phi(1960)$ καὶ τῶν $\phi(1947)$, $\phi(1946)$, $\phi(1945)$ κλπ. καθὼς ἐπίσης καὶ ἀπὸ τὰς τεταγμένας τῶν σημείων τῆς καμπύλης ΑΜΒ, τὰ ὁποῖα ἔχουν τετμημένας ἀντιστοίχους εἰς τὰ ἔτη 1953, 1964, . . . , 1960 κλπ.

Ἐνίοτε, ὅταν ζητοῦμεν νὰ εὕρωμεν τὸν τρόπον τῆς ἀμοιβαίας ἐξαρτήσεως δύο μεταβλητῶν π.χ. τῶν x, y, δεχόμεθα ὅτι μεταξὺ αὐτῶν ὑπάρχει σχέσις τῆς μορφῆς π.χ.

$$(1) \qquad y=\alpha_1\varphi_1(x)+\alpha_2\varphi_2(x)+\ldots+\alpha_\nu\varphi_\nu(x)=\sum_{i=1}^{\nu}\alpha_i\varphi_i(x),$$

ὅπου τὰ $\varphi_1(x), \varphi_2(x), \ldots, \varphi_\nu(x)$ εἶνε γνωσταὶ συναρτήσεις τῆς x, τὰ δὲ $\alpha_1, \alpha_2, \ldots \alpha_\nu$ συντελεσταὶ σταθεροὶ μὲν ἀλλ' ***ἄγνωστοι*** καὶ προσδιοριστέοι, ἀνεξάρτητοι δ' ἐν γένει μεταξύ των, καὶ ὁ προσδιορισμὸς αὐτῶν ἐξαρτᾶται συνήθως ἀπὸ τὰ ἀποτελέσματα παρατηρήσεων, πειραμάτων, μετρήσεων κλπ.

Διὰ τὸν προσδιορισμὸν τῶν α_κ $(\kappa=1, 2, \ldots, \nu)$ ἐφαρμόζομεν τὴν καλουμένην ***μέθοδον τῶν ἐλαχίστων τετραγώνων***, ἡ ὁποία ἔχει ὡς ἐκτίθεται κατωτέρω ἐν συντομίᾳ.

Ἂν αἱ τιμαὶ τῆς x, αἱ $x_1, x_2, \ldots, x_\mu$, $(\mu \geqslant \nu)$ εἶνε ***ἀκριβεῖς***, αἱ δ' ἀντίστοιχοι τούτων τιμαὶ τῆς y, αἱ $y_1, y_2, \ldots, y_\mu$, εἶναι ἐν γένει ***μὴ ἀκριβεῖς***, παριστάνομεν τὰ σφάλματα διὰ τὰς τιμὰς αὐτὰς τῆς y, ἢ τὰς διαφορὰς μεταξὺ τῶν ἀληθῶν ἢ ἀκριβῶν τιμῶν τοῦ y καὶ τῶν ἀντιστοίχων τῶν $y_1, y_2, \ldots, y_\mu$ μὲ $\Delta y_1, \Delta y_2, \ldots, \Delta y_\mu$. Θέτομεν δηλαδὴ

$$y_1+\Delta y_1=\alpha_1\varphi_1(x_1)+\alpha_2\varphi_2(x_1)+\ldots+\alpha_\nu\varphi_\nu(x_1)$$

$$y_2+\Delta y_2=\alpha_1\varphi_1(x_2)+\alpha_2\varphi_2(x_2)+\ldots+\alpha_\nu\varphi_\nu(x_2)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$y_\mu+\Delta y_\mu=\alpha_1\varphi_1(x_\mu)+\alpha_2\varphi_2(x_\mu)+\ldots+\alpha_\nu\varphi_\nu(x_\mu)$$

καὶ γενικῶς

$$y_\kappa+\Delta y_\kappa=\alpha_1\varphi_1(x_\kappa)+\alpha_2\varphi_2(x_\kappa)+\ldots+\alpha_\nu\varphi_\nu(x_\kappa)$$

Οὕτω ἔχομεν

$$\Delta y_\kappa=-y_\kappa+\alpha_1\varphi_1(x_\kappa)+\alpha_2\varphi_2(y_\kappa)+\ldots+\alpha_\nu\varphi_\nu(x_\kappa)$$

Ὁ ἀριθμὸς τῶν ἀνωτέρω ἐξισώσεων εἶνε ἴσος μὲ τὸν ἀριθμὸν τῶν τιμῶν $y_1, y_2, \ldots y_\mu$, αἱ ὁποῖαι, ὡς ἀνωτέρω ἀνεγράφη, δὲν θεωροῦνται ἀκριβεῖς. Τὰ σφάλματα $\Delta y_1, \Delta y_2, \ldots, \Delta y_\mu$ εἶνε ἄγνωστα, ἀλλ' ἐν γένει μικραὶ ποσότητες. Ἂν ὁ ἀριθμὸς τῶν ἀνωτέρω ἐξισώσεων εἶναι ν καὶ ἴσος μὲ τὸν ἀριθμὸν τῶν προσδιοριστέων συντελεστῶν $\alpha_1, \alpha_2, \ldots, \alpha_\nu$, ἠμποροῦμεν νὰ παραλείψωμεν τὰ πρῶτα μέλη Δy_κ τῶν ἐν λόγῳ ἐξισώσεων, λύοντες δ' αὐτὰς ἀκολούθως ὡς πρὸς $\alpha_1, \alpha_2 \ldots, \alpha_\nu$, εὑρίσκομεν τιμὰς τούτων, αἱ ὁποῖαι δὲν θὰ εἶνε ἀκριβεῖς ἀλλὰ θὰ προσεγγίζουν πρὸς αὐτάς.

Εἶνε ὅμως φανερόν, ὅτι θὰ προσεγγίσωμεν περισσότερον τὰς ἀκριβεῖς τιμὰς τῶν $\alpha_1, \alpha_2, \ldots, \alpha_\nu$, ἂν ἔχωμεν τιμὰς τῆς y περισσοτέρας τοῦ ν, ἀλλὰ διὰ τὴν περίπτωσιν αὐτὴν τίθεται τὸ ἐρώτημα, τοῦ πῶς πρέπει νὰ συνδυασθοῦν αἱ οὕτω προκύπτουσαι ὡς ἄνω ἐξισώσεις, διὰ νὰ εὑρεθοῦν τιμαὶ τῶν ἀγνώστων συντελεστῶν κατὰ τὸ δυνατὸν ἀκριβεῖς. Τὴν λύσιν τοῦ προβλήματος αὐτοῦ καθορίζει ὁ Λογισμὸς τῶν πιθανοτήτων διὰ τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων, στηρίζεται δὲ ἐπὶ τῆς ἑξῆς προτάσεως :

«Ἐξ ὅλων τῶν τιμῶν τῶν ἀγνώστων συντελεστῶν $\alpha_1, \alpha_2, \ldots, \alpha_\nu$, ***θεωροῦνται ἐκεῖναι ὡς πιθανώτατα ἀκριβεῖς, διὰ τὰς ὁποίας τὸ ἄθροισμα τῶν τετραγώνων τῶν σφαλμάτων*** $\sum_{\kappa=1}^{\mu}(\Delta y_\kappa)^2$ ***γίνεται ἐλάχιστον».***

Τὸ ἄθροισμα τῶν τετραγώνων τῶν σφαλμάτων ἢ τῶν διαφορῶν Δy_1, Δy_2, ..., Δy_μ εἶνε $(\Delta y_1)^2 + (\Delta y_2)^2 + \ldots + (\Delta y_\mu)^2 = \sum_{\kappa=1}^{\mu} (\Delta y_K)^2$.

Ἐπιδιώκομεν νὰ προσδιορίσωμεν τὰς τιμὰς τῶν α_1, α_2, ..., α_ν εἰς τρόπον ὥστε τὸ $\sum_{\kappa=1}^{\mu} (\Delta y_\kappa)^2$ νὰ γίνῃ ἐλάχιστον, θεωρουμένων μεταβλητῶν τῶν $\alpha_1 \alpha_2, .. \alpha_\nu$.

Διὰ νὰ γίνῃ τὸ $\sum_{\kappa=1}^{\mu} (\Delta y_\kappa)^2$ ἐλάχιστον, ὅπου μεταβληταὶ θεωροῦνται τὰ α_1, α_2, ..., α_ν, πρέπει αἱ (μερικαὶ) παράγωγοι τοῦ ἐν λόγῳ ἀθροίσματος τῶν τετραγώνων ὡς πρὸς α_1, α_2, ..., α_ν νὰ εἶναι ἴσαι μὲ μηδέν, ἤτοι πρέπει νὰ ἔχωμεν:

$$\frac{d \sum_{\kappa=1}^{\mu} (\Delta y_\kappa)^2}{d \alpha_i} = 0, \qquad (i = 1,\ 2, \ldots, \nu),$$

ἐπὶ πλέον πρέπει νὰ πληροῦνται καὶ ἄλλαι (συμπληρωματικαὶ) συνθῆκαι. Κατὰ ταῦτα πρέπει νὰ ἔχωμεν διὰ τὴν μερικὴν παράγωγον ὡς πρὸς α_1

$$\begin{aligned} & [-y_1 + \alpha_1\varphi_1(x_1) + \alpha_2\varphi_2(x_1) + \ldots + \alpha_\nu \varphi_\nu(x_1)]\varphi_1(x_1)] \\ & + [-y_2 + \alpha_1\varphi_1(x_2) + \ldots + \alpha_\nu \varphi_\nu(x_2)]\ \varphi_1(x_2) \\ & + \ldots \\ & + [-y_\mu + \alpha_1\varphi_1(x_\mu) + \ldots + \alpha_\nu \varphi_\nu(x_\mu)]\varphi_1(x_\mu)] = 0, \end{aligned}$$

$$\text{ἢ } [-y_1 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_1)]\ \varphi_1(x_1) + [-y_2 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_2)]\ \varphi_1(x_2) + \ldots + $$
$$+ [-y_\mu + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_\mu)]\ \varphi_1(x_\mu) = 0.$$

Κατ' ἀναλογίαν ἔχομεν $\nu-1$ ἀκόμη ἐξισώσεις διὰ τὰς μερικὰς παραγώγους ὡς πρὸς α_2, α_3, ..., α_ν τοῦ ἀθροίσματος τῶν τετραγώνων τῶν σφαλμάτων Οὕτω ἔχομεν τὰς κατωτέρω ν ἐξισώσεις.

$$\left(-y_1 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_1)\right) \varphi_1(x_1) + \left(-y_2 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_2)\right)\varphi_1(x_2) + \ldots +$$
$$+\left(-y_\mu + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_\mu)\right) \varphi_1(x_\mu) = 0.$$

$$\left(-y_1 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_1)\right) \varphi_2(x_1) + \left(-y_2 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_2)\right)\varphi_2(x_2) + \ldots +$$
$$+\left(-y_\mu + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_\mu)\right) \varphi_2(x_\mu) = 0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left(-y_1 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_1)\right)\varphi_\nu(x_1) + \left(-y_2 + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_2)\right)\varphi_\nu(x_2) + \ldots +$$
$$+\left(-y_\mu + \sum_{i=1}^{\nu} \alpha_i \varphi_i(x_\mu)\right)\varphi_\nu(x_\mu) = 0.$$

Ἐκ τῶν ν ὡς ἀνωτέρω ἐξισώσεων προσδιορίζονται τὰ $\alpha_1, \alpha_2, \ldots, \alpha_\nu$ καὶ μὲ τὰς οὕτω εὑρισκομένας τιμάς των ἀντικαθιστῶμεν αὐτὰ εἰς τὴν

$$y = \alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) + \ldots + \alpha_\nu \varphi_\nu(x),$$

ὑποθέτοντες ὅτι πληροῦνται καὶ ἄλλαι συμπληρωματικαὶ συνθῆκαι, περὶ τῶν ὁποίων ἔγινε λόγος ἀνωτέρω, τὰς ὁποίας δὲν ἀνεφέραμεν ἐπακριβῶς.

Περίπτωσις δύο μεταβλητῶν ἀνεξαρτήτων μεταξύ των (προσδιοριστέων) α_1, α_2. Ἂν $f(\alpha_1, \alpha_2)$ εἶνε συνάρτησις τῶν α_1, α_2 (συνεχὴς καὶ παραγωγίσιμος ὡς πρὸς αὐτάς), ἵνα γίνῃ ἐλαχίστη ἡ τιμὴ αὐτῆς διὰ ζητουμένας τιμὰς τῶν α_1, α_2, παρατηροῦμεν ὅτι ἂν αἱ ζητούμεναι ἀκριβεῖς τιμαὶ τῶν α_1, α_2 εἶνε αἱ $\alpha_1 + \Delta\alpha_1$, $\alpha_2 + \Delta\alpha_2$ τότε ἡ διαφορὰ $f(\alpha_{11} + \Delta\alpha_1, \alpha_{22} + \Delta\alpha_2) - f(\alpha_{11}, \alpha_{22})$ θὰ εἶνε θετικὴ διὰ πάσας τὰς τιμὰς τῶν $\Delta\alpha_1$, $\Delta\alpha_2$ τὰς περιεχομένας μεταξὺ τῶν $-|\varepsilon|$ καὶ $|\varepsilon|$ ὅπου τὸ ε ὑποτίθεται ἀρκούντως (κατὰ τὸ δυνατὸν) μικρὰ ποσότης, ἤτοι ἂν εἶνε: $\quad -|\varepsilon| < \Delta\alpha_1, \ \Delta\alpha_2 < |\varepsilon|$

ἔχομεν ὡς ἀναγκαίας συνθήκας τὰς

$$(1) \qquad f_{\alpha_1}(\alpha_1, \alpha_2) = 0, \qquad f_{\alpha_2}(\alpha_1, \alpha_2) = 0$$

καὶ ἀκόμη τὰς

$$(2) \qquad f_{\alpha_1^2} > 0 \ , \qquad f_{\alpha_1^2} f_{\alpha_2^2} - f^2_{\alpha_1\alpha_2} > 0$$

ὅπου οἱ δεῖκται παριστάνουν μερικὰς παραγώγους πρώτης τάξεως ἢ δευτέρας ὡς πρὸς τὰς ὑπ' αὐτῶν παριστανομένας μεταβλητάς.

Ἐκ τῶν ἐξισώσεων (1) προσδιορίζονται αἱ τιμαὶ τῶν α_1, α_2 αἱ ὁποῖαι πρέπει νὰ ἐπαληθεύουν καὶ τὰς συνθήκας (2)

Ἐφαρμογαί. 1) Διὰ τὸν καθορισμὸν σταθερᾶς β π.χ. ὅταν εἶναι $y=\beta$ (ἤτοι εὐθεῖα παράλληλος πρὸς τὸν ἄξονα τῶν x), ἔχομεν:

$$\sum_{\kappa=1}^{\mu} (\Delta y_\kappa)^2 = (-y_1+\beta)^2 + (-y_2+\beta)^2 + \ldots + (-y_\mu+\beta)^2$$

$$\frac{d\sum_{\kappa=1}^{\mu}(\Delta y_\kappa)^2}{d\beta} = \mu\beta - (y_1+y_2+\ldots+y_\mu) = 0,$$

$$\beta = \frac{y_1+y_2+\ldots+y_\mu}{\mu}$$ (ἤτοι ἡ μέση τιμὴ τῶν

$y_1, y_2, \ldots, y_\mu$).

2) Εὕρεσις τῆς συναρτήσεως $y = \alpha_1 x + \alpha_2$.

Ἔχομεν $$\sum_{\kappa=1}^{\mu}(\Delta y_\kappa)^2 = (-y_1+\alpha_1 x_1+\alpha_2)^2 + \ldots + (-y_\mu+\alpha_1 x_\mu+\alpha_2)^2$$

$$\frac{d\sum_{\kappa=1}^{\mu}(\Delta y_\kappa)^2}{d\alpha_1} = (x_1^2+x_2^2+\ldots+x_\mu)\alpha_1 + (x_1+x_2 1-4x_\mu)\alpha_2 - (x_1y_1+\ldots+x_\mu y_\mu) = 0$$

$$\frac{d\sum_{\kappa=1}^{\mu}(\Delta y_\kappa)^2}{d\alpha_2} = (x_1+x_2+\ldots+x_\mu)\alpha_1 + \mu\alpha_2 - (y_1+y_2+\ldots+y_\mu) = 0$$

$$ή \quad (1) \quad \begin{cases} \alpha_1 \sum_{\kappa=1}^{\mu} x_\kappa^2 + \alpha_2 \sum_{\kappa=1}^{\mu} x_\kappa = \sum_{\kappa=1}^{\mu} x_\kappa \; y_\kappa \\ \alpha_1 \sum_{\kappa=1}^{\mu} x_\kappa + \mu\alpha_2 = \sum_{\kappa=1}^{\mu} y_\kappa \, , \end{cases}$$

καὶ

$$\alpha_1 = \left[\left(\sum_{\kappa=1}^{\mu} x_\kappa\right) . \left(\sum_{\kappa=1}^{\mu} y_\kappa\right) - \mu \sum_{\kappa=1}^{\mu} x_\kappa \; y_\kappa\right] : \left[\left(\sum_{\kappa=1}^{\mu} x_\kappa\right)^2 - \mu \sum_{\kappa=1}^{\mu} x_\kappa^2\right]$$

$$\alpha_2 = \left[\left(\sum_{\kappa=1}^{\mu} x_\kappa\right) . \left(\sum_{\kappa=1}^{\mu} x_\kappa \; y_\kappa\right) - \left(\sum_{\kappa=1}^{\mu} x_\kappa^2\right) . \left(\sum_{\kappa=1}^{\mu} y_\kappa\right)\right] : \left[\left(\sum_{\kappa=1}^{\mu} x_\kappa\right) - \sum_{\kappa=1}^{\mu} x_\kappa^2\right]$$

Αἱ (1) καλοῦνται συνήθως (εἰς τὴν στατιστικὴν) ***κανονικαὶ ἐξισώσεις*** *.

3) Ἂν τιμαὶ $x_1, x_2, \ldots, x_\mu$ ἀποτελοῦν ἀριθμητικὴν πρόοδον. Ἔχομεν κατὰ τὴν περίπτωσιν αὐτήν :

$$x_2 - x_1 = x_3 - x_2 = \ldots = x_\mu - x_{\mu-1} = \lambda \quad \pi.\chi., \text{ ὅτε}$$

$$x_\sigma = x_1 + (\sigma - 1) . \lambda \, , \qquad (\sigma = 1, 2 \ldots, \mu)$$

$$y_\sigma = (x_1 - \lambda + \sigma . \lambda)\alpha_1 + \alpha_2 = \lambda\sigma\alpha_1 + (\alpha_1 x_1 - \lambda\alpha_1 + \alpha_2).$$

Ἐπὶ τῶν ἐξισώσεων τούτων ἐργαζόμεθα ὡς ἀνωτέρω πρὸς εὕρεσιν τῶν τιμῶν τῶν α_1, α_2.

4) Κατ᾿ ἀναλογίαν πρὸς τὴν ἀνωτέρω περίπτωσιν 2 διακρίνομεν τὴν περίπτωσιν $y = \alpha_0 x^2 + \alpha_1 x + \alpha_2$ (περίπτωσις παραβολῆς β' βαθμοῦ) κλπ.

5) Ἀξιοσημείωτος εἶνε ἡ περίπτωσις κατὰ τὴν ὁποίαν ἔχομεν

$$y = \alpha_1 x^{\alpha_2} \, , \qquad \alpha_2 \lesseqgtr 0 \, ,$$

εἰς τὴν ὁποίαν διὰ τὸν προσδιορισμὸν τῶν α_1, α_2 ἐφαρμόζομεν τὴν μέθοδον τῶν ἐλαχίστων τετραγώνων εἰς τὴν προκύπτουσαν ἐκ τῆς ἀνωτέρω ἐξισώσεως ἀφοῦ λάβωμεν τοὺς λογαρίθμους (κοινοὺς ἢ νεπερείους) τῶν ἴσων μελῶν, ὅτε ἔχομεν :

$$\lambda o\gamma \; y = \lambda o\gamma \; \alpha_1 + \alpha_2 \; \lambda o\gamma \; x.$$

6) Διὰ τὴν περίπτωσιν τῆς εὑρέσεως τοῦ τρόπου κατανομῆς εἰσοδήματος ἔχομεν τὴν ἐξίσωσιν **

$$y = \alpha_1 \, x^{-\alpha_2} \, , \qquad (\alpha_2 > 0),$$

ὅπου x παριστάνει τὸ εἰσόδημα καὶ y τὸν ἀριθμὸν τῶν ἀτόμων τῶν ἐχόντων αὐτό. Διὰ τὸν προσδιορισμὸν τῶν α_1, α_2, ἐφαρμόζομεν τὴν μέθοδον τῶν ἐλαχίστων τετραγώνων, ἀκολούθως δὲ γίνεται σύγκρισις τῆς κατανομῆς τοῦ εἰσοδήματος α') κατὰ διαφόρους ἐποχὰς εἰς τὸν αὐτὸν τόπον (χώραν)· β') εἰς διαφόρους τόπους κατὰ τὴν αὐτὴν χρονικὴν ἐποχὴν ἢ περίοδον· γ') εἰς διαφόρους τόπους καὶ κατὰ διαφόρους ἐποχάς.

7) Διὰ τὴν περίπτωσιν $\quad y = \alpha_1 x^{\alpha_2} + \alpha_3$

ἡ ἐφαρμογὴ τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων παρουσιάζει αἰσθητὴν δυσκολίαν καὶ διὰ τοῦτο καταφεύγουν συνήθως εἰς τὴν χρησιμοποίησιν τεχνασμάτων.

---

* (Βλ. ἀριθμητικὰς ἐφαρμογὰς τῆς μεθόδου τῶν ἐλαχίστων τετραγώνων εἰς τὴν Στατιστικήν, Σ. Μαργαρίτη, 1952, σελ. 207 κ.ἑ).

** (Κατὰ V. Pareto : La courbe de Revenues : Cours d' Economie Politique, 2 Lausane, 1897).

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 χρημιμοποιούμενον 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 $x_i = \alpha - \beta e^{-\mu x}$.