ΤΑΧΕΙΣ ΣΤΑΤΙΣΤΙΚΟΙ ΥΠΟΛΟΓΙΣΜΟΙ

Ὑπὸ Ι. Α. ΣΑΚΑΛΗ Ἀντισμηνάρχου
Διευθυντοῦ Στατιστικῆς Γεν. Ἐπιτελείου Ἀεροπορίας καὶ Ἐθνικοῦ Στατιστικοῦ Συντονιστοῦ Ἀεροπορίας παρὰ τῷ Ν. Α. Τ. Ο.

Ι. ΕΙΣΑΓΩΓΙΚΑ

Ἡ ἀπόφασις περὶ χρησιμοποιήσεως ἐν δεδομένῃ στιγμῇ τῆς καταλληλοτέρας στατιστικῆς μεθόδου ἐξαρτᾶται κατὰ τὸ πλεῖστον ἐκ τῶν δοθησομένων ἀπαντήσεων εἰς τὰ τρία ἐρωτήματα: α) Εἰς τί χρησιμεύει ἡ μέθοδος; β) Πόσον καλὴ εἶναι ἡ μέθοδος; καὶ γ) Ποῖαι αἱ συνέπειαι οἱωνδήποτε ὑποθέσεων ἐπὶ τῆς ἐφαρμογῆς τῆς μεθόδου; Μία ἐκτίμησις τῆς πιθανῆς ἐκτάσεως τῶν ἀναμενομένων ἀπαντήσεων ἐφ' ἑκάστης τῶν ἐρωτήσεων αὐτῶν εἶναι ἀναγκαία διὰ τὴν ἀποτελεσματικὴν χρῆσιν τῶν διαθεσίμων μεθόδων.

Ἡ συνηθεστέρα ὁμὰς στατιστικῶν μεθόδων εἶναι ἐκείνη, κατὰ τὴν ὁποίαν ὑποτίθεται ὅτι αἱ παρατηρήσεις κατανέμονται κανονικῶς. Γενικῶς, αἱ μέθοδοι αὐταὶ ἐκτιμῶσιν ἢ ἐλέγχουσι παραμέτρους τῆς Κανονικῆς Κατανομῆς, ὡς $\overline{X}$ καὶ σ, διὰ τὴν λῆψιν ὅλων τῶν δυνατῶν πληροφοριῶν ἐκ τῶν παρατηρήσεων μὲ ἀποτελεσματικότητα 100%. Ἀλλὰ καὶ τὸ θέμα τῆς ὑποθέσεως δὲν εἶναι ἁπλοῦν. Ἐν πρώτοις ἡ ὑπόθεσις τῆς κανονικότητος, πιθανὸν νὰ ἀποδειχθῇ μὴ ὀρθή, ὁπότε εἶναι συνήθως ἀληθὲς ὅτι ἡ ἐπὶ τῆς ὑποθέσεως αὐτῆς βασισθεῖσα μέθοδος θέλει δώσει ἀρκετὰ ἀπατηλὰς ἀπαντήσεις. Πολλαὶ κανονικαὶ μέθοδοι εἶναι ρωμαλέαι. Ὅ,τι εἶναι πλέον σημαντικὸν εἶναι ὅτι ἡ μέθοδος πιθανὸν νὰ δώσῃ ἀπάντησιν εἰς τὴν ἐσφαλμένην ἐρώτησιν ἢ ὁμοίως ὅταν ἡ ἐρώτησις εἶναι λογικὴ πιθανὸν ν' ἀπαντήσῃ εἰς ταύτην ἀνεπαρκῶς. Ἐπὶ παραδείγματι ἡ κατάλληλος ἐρώτησις θὰ ἦτο ὄχι ἐὰν οἱ Μέσοι ἢ αἱ Μέσαι Τετραγωνικαὶ Ἀποκλίσεις δύο πληθυσμῶν εἶναι ὅμοιαι, ἀλλ' ἐὰν οἱ πληθυσμοὶ διαφέρωσι καθόλου. Ἐπίσης ἐπὶ τοῦ ἐρωτήματος ἐὰν οἱ Μέσοι τῶν πληθυσμῶν εἶναι οἱ αὐτοὶ τὴν ἀπάντησιν καλύτερον θέλει δώσει ἡ χρησιμοποίησις τῶν Διαμέσων. Αἱ διαπιστώσεις αὐταὶ πολλάκις ἀναγκάζουσιν ἡμᾶς νὰ χρησιμοποιήσωμεν μεθόδους ἀπηλλαγμένας τῆς ὑποθέσεως τῆς Κατανομῆς (Distribution - Free Methods). Δεύτερον ἀκόμη καὶ ἐὰν ἡ ὑπόθεσις τῆς κανονικότητος εὐσταθῇ οἱ ἀπαιτούμενοι ὑπολογισμοὶ διὰ τὴν ἐφαρμογὴν μιᾶς πλήρως ἀποτελεσματικῆς μεθόδου δυνατὸν νὰ εἶναι ὑπερβολικοὶ σχετικῶς πρὸς τὸ ἐπιδιωκόμενον διὰ τῆς ὑπ' ὄψιν μεθόδου κέρδος. Εἶναι πολὺ ἁπλούστερον νὰ χρησιμοποιήσωμεν μίαν ὀλιγώτερον ἀποτελεσματικὴν μέθοδον, ἀπαιτοῦσαν ἐλαχίστους ὑπολογισμοὺς καὶ ἐνδεχομένως νὰ λάβωμεν ὀλίγας τινὰς παρατηρήσεις εἰς ἀντιστάθμισμα τῆς ἀπωλείας τῆς ἀποτελεσματικότητος, τὴν ὁποίαν ἐξασφαλίζουσι σχοινοτενεῖς ὑπολογισμοί. Τέλος καὶ εἰς τὴν περίπτωσιν, καθ' ἣν ἐπιβάλλονται πλήρεις ὑπολογισμοί, ταχυτέρα τις μέθοδος δύναται νὰ χρησιμεύσῃ ὡς ὁδηγὸς περὶ τοῦ ἐὰν ἀξίζῃ τὸν κόπον νὰ καταναλώσωμεν ἀρκετὴν ἐγκεφαλικὴν οὐσίαν ἢ ἐὰν οἱ ἐνεργηθέντες ὑπολογισμοὶ ἦσαν ἀκριβεῖς.

Τὸ μέτρον τῆς ἀποτελεσματικότητος προσδιορίζει τὸν ἀπαιτούμενον σχετικὸν ἀριθμὸν παρατηρήσεων διὰ νὰ ἐπιτύχωμεν ἐξ ἴσου ἀκριβῆ ἀποτελέσματα διὰ τῆς δοθείσης μεθόδου ὅσον καὶ τῆς καλυτέρας δυνατῆς μεθόδου. Οὕτω μέθοδος ἀποτελεσματικὴ κατὰ 50 % ἀπαιτεῖ κατὰ μέσον ὅρον διπλασίας παρατηρήσεις διὰ νὰ ἐπιτύχωμεν ἀκρίβειαν ἔναντι τῆς κατὰ 100 % ἀποτελεσματικῆς μεθόδου. Ὁπωσδήποτε θὰ πρέπει νὰ τονισθῇ ὅτι τὸ μέτρον αὐτὸ ἐμβάλλει πολλὰς φορὰς εἰς ἀμφισβητήσεις καὶ ὑπονοίας καὶ μόνον ἡ ὀξυδέρκεια τοῦ Στατιστικοῦ θὰ σώσῃ τὴν κατάστασιν.

Ὡς ἀποτελεσματικωτέρα μέθοδος ἐκλαμβάνεται ἐκείνη, κατὰ τὴν ὁποίαν ἡ ὑπόθεσις τῆς κανονικότητος εἶναι ἀληθής. Ἔστω ὡς παράδειγμα ἡ Συσχέτισις. Ἐπὶ τῇ ὑποθέσει ὅτι αἱ παρατηρήσεις ἐλήφθησαν ἔκ τινος Διμεταβλητῆς Κανονικῆς Κατανομῆς ἡ ἀποτελεσματικότης τῆς μεθόδου παρίσταται εἰς $4:\pi^2 = 41\,\%$. Ὁσάκις ἡ ἀνεξάρτητος μεταβλητὴ ἔχει κατανεμηθῆ ὀρθογωνίως καὶ ἡ ἐξηρτημένη κανονικῶς ἡ ἀποτελεσματικότης θὰ εἶναι $3:2\pi = 48\,\%$. Ὁμοίως ἐὰν ὑποτεθῇ ὅτι ἡ ἀνεξάρτητος μεταβλητὴ ἔχει κατανεμηθῆ ὁμαλῶς, ὅτι περαιτέρω παρατηρήσεις ἔχουσι γίνει διὰ τῆς ἐπεκτάσεως τῆς σειρᾶς ταύτης καὶ ὅτι ἡ ἐξηρτημένη μεταβλητὴ ἔχει κατανεμηθῆ κανονικῶς ἡ ἀποτελεσματικότης θὰ εἶναι $3:2\pi = 78\,\%$.

II. ΕΙΔΙΚΩΤΕΡΑ ΕΞΕΤΑΣΙΣ ΤΩΝ ΜΕΘΟΔΩΝ ΥΠΟΛΟΓΙΣΜΟΥ

Αἱ παρατιθέμεναι κατωτέρω μέθοδοι προχείρων ὑπολογισμῶν ἐλήφθησαν ἐκ τοῦ προσφάτου ἔργου τοῦ διαπρεποῦς Ἄγγλου Στατιστικοῦ κ. M. H. Quenouille «Rapid Statistical Calculations» (1959), τῆς Ὁμάδος Μεθόδων Ἐρευνῶν τῆς London School of Economics, περιλαμβάνουσι δὲ τρεῖς κατηγορίας: α) Ἐπιπέδου Μέσου ἢ Διασπορᾶς ὁμάδων παρατηρήσεων. β) Συγκρίσεις Συχνοτήτων ἢ Ἀναλογιῶν καὶ γ) Συσχετίσεως, συμπληρουμένας ὑπὸ ἐννέα στατιστικῶν πινάκων καὶ ἐπὶ πλέον χαρακτηριζομένας διὰ συμβόλων ὡς Ταχεῖαι ∗ , Λίαν Ταχεῖται ‡ , Ταχύτεραι τῶν συνήθων † καὶ Ἐλεύθεραι Κατανομῆς §.

Α΄. Μέθοδοι ἐπιπέδου Μέσου ἢ Διασπορᾶς ὁμάδων παρατηρήσεων

1. Ἐκτίμησις Ἀριθμητικοῦ Μέσου

Παράδειγμα :

$$\overline{X} = \overline{Z} + \frac{\Sigma(d)}{N} \qquad \sigma_{\overline{x}} = \frac{S}{\sqrt{N}}$$

2	2	6	9	9	10	15	18	20	20
21	21	23	23	26	27	27	27	28	28
29	29	30	33	34	34	34	34	34	35
35	37	37	37	37	37	38	38	38	39
39	39	40	42	44	44	45	47	51	58

Ἔστω ὡς Ὑποτιθέμενος Μέσος $\overline{Z}$ ὁ ἀριθμὸς 30 καὶ αἱ ἀπὸ τούτου ἀποκλίσεις μιᾶς ἑκάστης τῶν πεντήκοντα παρατηρήσεων –28, –28, –24, –21, –21, 20, –15, –12, –10, –10, –9, –9, –7, –7, –4, –3, –3, –3, –2, –2, –1, –1, 0, 3, 4, 4, 4, 4, 4, 5, 5, 7, 7, 7, 7, 7, 8, 8, 8, 9, 9, 9, 10, 12, 14, 14, 15, 17, 21, 28. Σύνολον ἀρνητικῶν ἀποκλίσεων –240 καὶ θετικῶν + 250, ἄρα $\Sigma(d) = 10$ καὶ κατ' ἐφαρμογὴν τοῦ τύπου θὰ ἔχωμεν: $30 + \frac{10}{50} = 30.20$ μὲ Κανονικὸν Σφάλμα Ἀριθμ. Μέσου $\frac{12.37}{\sqrt{50}} = \pm\ 1.75$. (Διὰ τρόπον ὑπολογισμοῦ τῆς σ ὅρα μέθοδον 2).

Παρατηρήσεις :

α) Ἡ μέθοδος εἶναι ταχεῖα καὶ 100% ἀποτελεσματικὴ διὰ κανονικῶς κατανεμηθείσας παρατηρήσεις.

β) Ἐὰν αἱ παρατηρήσεις ἐμφανίζωνται ὑπὸ τὴν μορφὴν Κατανομῆς Συχνότητος ὁ $\overline{X}$ δύναται νὰ ὑπολογισθῇ ἄνευ διορθώσεως διὰ τὴν ὁμαδικοποίησιν.

γ) Ὁ $\overline{X}$, θὰ τείνῃ νὰ κατανέμηται κανονικῶς (ἐξαιρέσει μικροῦ ἀριθμοῦ παρατηρήσεων ἀσυμμέτρου κατανομῆς) καὶ τὰ ὅρια τούτου θὰ καθορίζωνται διὰ τῆς χρησιμοποιήσεως τῆς Κατανομῆς t Student βάσει τοῦ οἰκείου Πίνακος ἢ ἐπὶ μεγάλου ἀριθμοῦ παρατηρήσεων ὁμοίως διὰ τοῦ ὑπολογισμοῦ τῆς κανονικῆς ἀποκλίσεως τοῦ Πίνακος.

2. Ἐκτίμησις Μέσης Τετραγωνικῆς Ἀποκλίσεως

$$\widehat{S} = \sqrt{\left[\frac{1}{N-1}\left(\Sigma d^2 - \frac{d^2}{N}\right)\right]}$$

Ἐκτίμησις τῆς σ τοῦ πληθυσμοῦ.

$$\boxed{\widehat{S} = \frac{S}{1 + \frac{1}{4(N-1)}}}$$ 'Αμερόληπτος ἐκτίμησις τῆς S

$$\boxed{\sigma_S = \frac{S}{\sqrt{[2(N-0{,}75)]}}}$$ Κανονικὸν Σφάλμα τῆς S

Παράδειγμα:

'Εκ τῶν δεδομένων τοῦ παραδείγματος 1 ἔχομεν $\Sigma d^2 = 7502$ καὶ $d^2 = 10^2$

καὶ συνεπῶς $$\widehat{S} = \sqrt{\frac{1}{49}\left(7502 - \frac{100}{50}\right)} = 12.37$$

$$\sigma_S = 12.37 : \sqrt{[2 \times 49.25]} = \pm 1.25$$

$$\widehat{S} = 12.37 : 1 + 0.005 = 12.31$$

Παρατηρήσεις:

α) Ἡ μέθοδος εἶναι ἀποτελεσματικὴ 100% μόνον ἐπὶ κανονικῶς κατανεμηθεισῶν παρατηρήσεων.

β) Δοθέντος ὅτι ἡ S παρέχει μεροληπτικὴν ἐκτίμησιν τῆς Μέσης Τετραγωνικῆς 'Αποκλίσεως ἡ μεροληψία αὐτὴ διορθοῦται προκειμένου περὶ Κανονικῆς Κατανομῆς ἐὰν διαιρέσωμεν τὴν S διὰ $1 + \frac{1}{4(N-1)}$. Πολλοὶ ἔλεγχοι, ἐν οἷς καὶ ὁ τοῦ t Student, ἔχουσιν ὑπ' ὄψιν τὴν μεροληψίαν ταύτην.

γ) Ἡ ἐκτίμησις τῆς S τείνει νὰ κατανέμηται κανονικῶς ἐπὶ μεγάλων δειγμάτων, ὁ δὲ $\log e^S$ πλησιάζει περισσότερον τὴν Κανονικὴν Κατανομήν.

δ) 'Επὶ Κατανομῆς Συχνότητος ἡ 'Αμερόληπτος 'Εκτίμησις τῆς S ἐπιτυγχάνεται ἐὰν ἐκ τῆς συνήθους ἐκτιμήσεως ἀφαιρέσωμεν τὸ 1/12 τοῦ τετραγώνου τοῦ διαστήματος τάξεως.

3. *Ἕτερα Μέτρα 'Επιπέδου Μέσου*

I. Διάμεσος: Ἤτοι ἡ τιμὴ $y\left(\frac{1}{2}\right)$, ἥτις ὑπερβάλλεται ἐκ τοῦ ἡμίσεος τῶν παρατηρήσεων. Τὸ Κανονικὸν Σφάλμα τοῦ Διαμέσου ἐπὶ κανονικῶς κατανεμομένων παρατηρήσεων ὑπολογίζεται περίπου βάσει τοῦ τύπου:

$$\boxed{\sigma_{me} = \frac{5S}{4\sqrt{N}}}$$

II. Τεταρτημοριακὸς Μέσος (Mid—Quartile):

$$\boxed{MQ = \frac{Q_1 + Q_3}{2}}$$

Τούτέστιν ὁ μέσος ὅρος τῶν τιμῶν $y\left(\frac{1}{4}\right)$ καὶ $y\left(\frac{3}{4}\right)$, ὑπερβαλλομένων ἐκ τοῦ ἑνὸς τετάρτου καὶ τριῶν τετάρτων τῶν παρατηρήσεων μὲ Κανονικὸν Σφάλμα εἰς κανονικῶς κατανεμηθείσας παρατηρήσεις ὡς ἑξῆς:

$$\boxed{\sigma_{MQ} = \frac{1 \cdot 1\ S}{\sqrt{N}}}$$

Παράδειγμα: Τὸ αὐτὸ τῆς μεθόδου I.

$M_e = N + 1 : 2 = 25.5$ ἤτοι ὁ ἀριθμὸς 34

$$\sigma_{me} = \frac{5 \times 12.37}{4\sqrt{50}} = \pm 2.19$$

$Q_1 = 50 + 1 : 4 = 12.75$ ἢ 22.5

$Q_3 = 3 \times 51 : 4 = 38.75$ ἢ 38

Ἐνδοτεταρτημοριακὸς Μέσος $= 22.5 + 38 : 2 = 30.25$

Κανονικὸν Σφάλμα Τεταρτημοριακοῦ Μέσου $\frac{1.1 \times 12.37}{\sqrt{50}} = \pm 1.92$

Παρατηρήσεις

Αἱ ἀνωτέρω μέθοδοι ὑπολογισμοῦ εἶναι ταχεῖαι καὶ χρησιμοποιοῦνται διὰ τὴν ἐκτίμησιν τῶν Μέσων τῶν συμμετρικῶν κατανομῶν. Ἡ ἀποτελεσματικότης των διὰ τὴν ἐκτίμησιν τοῦ Ἀριθμ. Μέσου κανονικῶς κατανεμομένων παρατηρήσεων, εἶναι 64 % ἐπὶ τοῦ Διαμέσου καὶ 81 % ἐπὶ Τεταρτημοριακοῦ Μέσου. Ἐπειδὴ παρὰ ταῦτα εἶναι ἀρκούντως εὐπαθεῖς εἰς ἐνδεχομένας ἀσυμμετρίας τῆς κατανομῆς προτιμότερον εἶναι νὰ χρησιμοποιηθῶσιν αἱ μέθοδοι 6 καὶ 7.

4. *Ἕτερα Μέτρα Διασπορᾶς*

Α. Εὖρος. Β. Ἀπόκλισις Μέσου (ἀγνοουμένων τῶν σημείων αἱ ἀποκλίσεις ὑπολογίζονται ἐκ τοῦ Διαμέσου), *Γ. Διαφορὰ μεταξὺ τῶν Μέσων τοῦ μεγαλυτέρου 5 % καὶ μικροτέρου 5 % τῶν παρατηρήσεων.*

Παράδειγμα :

(ι)	2	6	10	23	28	34	38	39	45	47
(ιι)	4	17	20	30	36	39	39	42	44	46
(ιιι)	13	25	31	31	33	33	34	38	41	42
(ιν)	26	27	33	40	41	43	43	44	45	57
(ν)	17	28	31	34	41	42	43	47	52	66

Αἱ πέντε αὐταὶ σειραὶ ἐκ 10 παρατηρήσεων ἑκάστη ἔχουσιν ὡς Εὔρη ἀντιστοίχως 45, 42, 29, 31, 49, ἅτινα διαιρούμενα διὰ 3.1, δίδουσιν ὡς ἐκτιμήσεις τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως 14.5, 13.5, 9.4, 10.0, 15.8, μὲ μέσην τιμὴν 12.64. Διὰ τὰς ἰδίας σειρὰς λαμβανομένας ὁμοῦ ὁ Διάμεσος εἶναι ὁ 37 καὶ τὸ ἄθροισμα τῶν ἀποκλίσεων ἀπὸ τούτου 498. Κατὰ συνέπειαν ἡ ἐκτίμησις τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως τῶν πεντήκοντα παρατηρήσεων θὰ εἶναι :

$$\widehat{S} = \frac{5 \times 498}{4 \times 50} = 12.45$$

Ἑτέρα ἐκτίμησις ἐπιτυγχάνεται κατὰ τὴν Γ' μέθοδον διὰ τοῦ ὑπολογισμοῦ τοῦ μεγαλυτέρου καὶ μικροτέρου 2 1/2 τῶν παρατηρήσεων, αὗται δὲ εἶναι 2/5 $(2 + 4 + 1/2 \times 6) = 3.6$ καὶ 2/5 $(66 + 57 + 1/2 \times 52) = 59.6$ καὶ ἡ ἐκτίμησις τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως θὰ εἶναι :

$$\widehat{S} = \frac{59.6 - 3.6}{4} = 14$$

ἢ καθ' ἡμᾶς ἁπλουστεύοντες τὰς πράξεις ὑπολογίζομεν τὰς 21/2 μικροτέρας καὶ 2 1/2 μεγαλυτέρας παρατηρήσεις, λαμβάνομεν, τὴν διαφορὰν αὐτῶν καὶ διαιροῦντες διὰ 10 εὑρίσκομεν τὴν ἐκτίμησιν. Εἰς τὸ παράδειγμά μας $(2 + 4 + 6/2) = 9$ καὶ $(66 + 57 + 52/2) = 149$, ὁπότε

$$\widehat{S} = \frac{149 - 9}{10} = 14$$

Παρατηρήσεις :

α) Αἱ περιγραφεῖσαι μέθοδοι εἶναι Ταχεῖαι, χρησιμοποιούμεναι διὰ τὴν εὕρεσιν τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως κανονικῶς κατανεμομένων παρατηρήσεων, μὲ ἀποτελεσματικότητα 90%, 88% καὶ 70%, εἶναι ὅμως εὐπαθεῖς εἰς ἀποκλίσεις ἀπὸ τῆς κανονικότητος.

β) Ἡ μέθοδος, τοῦ Εὔρους εἶναι πλέον ἀποτελεσματικὴ διὰ μικρὰ δείγματα, ἐκτιμᾷ δὲ πολλαπλάσια τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως ὡς κατωτέρω :

N	2	3	4	5	6	7	9	10	15
Πολλαπλάσια	1.1	1.7	2.1	2.3	2.5	2.7	3.0	3.1	3.5
Ἀποτελεσματικότης %	100	99	98	96	93	91	89	88	81

γ) Εἰς μεγάλα δείγματα αἱ δύο τελευταῖαι μέθοδοι ἐκτιμῶσι 4/5 καὶ 4 φορὰς τὴν Μέσην Τετραγωνικὴν Ἀπόκλισιν μὲ ἀποτελεσματικότητα ἀντιστοίχως 88% καὶ 70%.

5. *Ἐκτίμησις τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως μιᾶς τυχαίως διαταχθείσης σειρᾶς κανονικῶς κατανεμηθεισῶν παρατηρήσεων*

Α. Κατὰ τὴν μέθοδον ταύτην διαιροῦμεν τὰς N παρατηρήσεις εἰς ὁμάδας ἐξ ἐννέα παρατηρήσεων, ἐπικαλύπτοντες κατὰ τὰς περιστάσεις, ἐκτιμῶμεν τὸ Εὖρος ἑκάστης ὁμάδος παρατηρήσεων, ἐκτιμῶμεν τὴν Μέσην Τετραγωνικὴν Ἀπόκλισιν, διὰ τοῦ ὑπολογισμοῦ τοῦ ἑνὸς τρίτου τοῦ Μέσου τῶν Εὐρῶν καὶ τέλος ἐκτιμῶμεν τὸ Κανονικὸν Σφάλμα τῆς S διὰ τοῦ τύπου :

$$\boxed{\sigma_s = \frac{0.8\ S}{\sqrt{N}}}$$

Β. Καθ' ἡμᾶς ἀποφεύγονται ἐνδεχόμενα σφάλματα ἐὰν λάβωμεν τὰ Εὔρη τῶν ὁμάδων, διαιρέσωμεν ταῦτα διὰ τοῦ ἀριθμοῦ τῶν ὁμάδων πρὸς εὕρεσιν τοῦ Μέσου τῶν Εὐρῶν καὶ περαιτέρω διαιρέσωμεν τὸν εὑρεθέντα Μέσον διὰ τοῦ 3 πρὸς ὑπολογισμὸν τῆς ἐκτιμήσεως τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως :

Παράδειγμα :

(ι)	34	23	38	2	10	6	47	39	45	28
(ιι)	42	15	34	37	44	2	37	28	18	40
(ιιι)	38	37	29	21	27	29	9	34	30	27
(ιv)	21	35	37	39	20	37	51	27	38	34
(v)	34	33	44	26	58	23	20	9	39	35

Κατὰ τὴν Α' Μέθοδον:

Παρατηρήσεις	Μικροτέρα τιμὴ	Μεγαλυτέρα τιμὴ	Εὖρος
1 – 9	2	47	45
9 – 17	2	45	43
17 – 25	18	40	22
25 – 33	9	37	28
33 – 41	20	51	31
41 – 50	9	58	49

Μέσον εὖρος: $45 + 43 + 22 + 28 + 31 + 49 : 6 = 36.3$

$$\widehat{S} = 36.3 : 3 = 12.1 \quad \text{καὶ} \quad \sigma_s = \frac{0.8 \times 12.1}{\sqrt{50}} = \pm 1.4$$

Κατὰ τὴν Β' Μέθοδον:

$$R_1 = 45 \quad R_2 = 42 \quad R_3 = 29 \quad R_4 = 31 \quad R_5 = 49$$

$$45 + 42 + 29 + 31 + 49 : 5 = 39.2$$

$$\widehat{S} = 39.2 : 3 = 13.0 \quad \text{καὶ} \quad \sigma_s = \frac{0.8 \times 13}{\sqrt{50}} = \pm 1.4$$

Παρατηρήσεις:

α) Ἡ μέθοδος εἶναι ταχυτέρα τῆς συνήθους, εἰς Κανονικὴν δὲ Κατανομὴν παρέχει ἐκτίμησιν τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως μὲ ἀποτελεσματικότητα 80–89%.

β) Ἡ μέθοδος αὐτὴ δὲν ἐπηρεάζεται σχεδὸν ἐξ ἀποκλίσεων ἐκ τῆς κανονικότητος.

γ) Εἰς κανονικῶς κατανεμηθείσας παρατηρήσεις ἡ ἐκτίμησις S θὰ κατανέμηται ἀρκετὰ πλησίον τοῦ κανονικοῦ.

δ) Τὸ Εὖρος εἰς ὁμάδας τῶν τεσσάρων ἐκτιμᾷ δύο φορὰς τὴν Μέσην Τετραγωνικὴν Ἀπόκλισιν μὲ ἀποτελεσματικότητα 74–98%.

6. *Ἐκτίμησις τοῦ ἀριθμητικοῦ μέσου καὶ μέσης τετραγωνικῆς ἀποκλίσεως εἰς κανονικῶς κατανεμηθείσας παρατηρήσεις πίνακος συχνοτήτων*

— Ἐκτίμησις τῶν τιμῶν, αἵτινες ὑπερβάλλονται κατὰ 1/16, 1/2, 15/16 τῶν N παρατηρήσεων, δηλαδὴ $y\left(\frac{1}{16}\right)$, $y\left(\frac{1}{2}\right)$, $y\left(\frac{15}{16}\right)$

— Ἐκτίμησις τοῦ Μέσου: $\widehat{m} = 0.2\,y\,(1/16) + 0.6\,y\,(1/2) + 0.2\,y\,(15/16)$

— Ἐκτίμησις τῆς σ : $\widehat{S} = \frac{1}{3}\left[y\left(\frac{1}{16}\right) - y\left(\frac{15}{16}\right)\right.$

— Κανονικὸν Σφάλμα Ἀριθμητικοῦ Μέσου : $\sigma_m = \frac{1.1\,\sigma}{\sqrt{N}}$

— Δείκτης Ἀσυμμετρίας : $S_k = y\left(\frac{1}{16}\right) - 2y\left(\frac{1}{2}\right) + y\left(\frac{15}{16}\right)$

— Κανονικὸν Σφάλμα S_k καὶ S : $\sigma_{s_k} = 3\sigma : \sqrt{N}$, $\sigma_s = \frac{0.9\,\sigma}{\sqrt{N}}$

Παράδειγμα :

Μέσος Δείκτης :	69,	71,	73,	75,	77,	79,	81,	83,	85,	87,	89	N
Συχνότης :	1,	2,	34,	107,	183,	205,	140,	50,	21,	11,	2	756

$y\left(\frac{1}{16}\right) = 75.2, \quad y\left(\frac{1}{2}\right) = 79.5, \quad y\left(\frac{15}{16}\right) = 84.5$

$\widehat{m} = 0.2 \times 75.2 + 0.6 \times 79.5 + 0.2 \times 84.5 = 79.64$

$\widehat{S} = 84.5 - 75.2 : 3 = 3.1$

$\sigma_m = 1.1 \times 3.1 : \sqrt{756} = 0.12$

$\sigma_s = 0.9 \times 3.1 : \sqrt{756} = 0.10$

$S_k = 75.2 - 2 \times 79.5 + 84.5 = 0.7$

$\sigma_{s_k} = 3 \times 3.1 : \sqrt{756} = 0.33$

Παρατηρήσεις :

α) Ἡ μέθοδος εἶναι Ταχεῖα καὶ ἐκτιμᾷ τὸν Ἀριθμητικὸν Μέσον καὶ Μέσην Τετραγωνικὴν Ἀπόκλισιν εἰς Κανονικὴν Κατανομὴν μὲ ἀποτελεσματικότητα 83% καὶ 62% ἀντιστοίχως.

β) Ἡ μέθοδος δὲν ἐπηρεάζεται ἐξ ἀποκλίσεων ἀπὸ τῆς κανονικότητος ὡς ἐπίσης ἐξ ἀποκλίσεων κατὰ προσέγγισιν συμμετρίας τῆς Κατανομῆς. Πάντως δι' ὑψηλῶς ἀσυμμέτρους κατανομὰς συνιστᾶται ἡ ἑπομένη μέθοδος 7.

7. *Ἐκτίμησις ἀριθμητικοῦ μέσου καὶ μέσης τετραγωνικῆς ἀποκλίσεως διὰ παρατηρήσεις κατὰ μέγεθος ἢ κατανομὴν συχνότητος*

— Ἐκτίμησις τῶν τιμῶν, αἵτινες ὑπερβάλλονται κατὰ 1/16, 1/4, 1/2, 3/4, 15/16 τῶν N παρατηρήσεων, ἤτοι

$$y\left(\frac{1}{16}\right),\ y\left(\frac{1}{4}\right),\ y\left(\frac{1}{2}\right),\ y\left(\frac{3}{4}\right),\ y\left(\frac{15}{16}\right)$$

— Ἐκτίμησις τοῦ Μέσου :

$$\widehat{m} = \left[y\left(\frac{1}{16}\right) + y\left(\frac{1}{4}\right) + 2y\left(\frac{1}{2}\right) + y\left(\frac{3}{4}\right) + y\left(\frac{15}{16}\right)\right] : 6$$

— Ἐκτίμησις τῆς σ :

$$\widehat{S} = \left[y\left(\frac{1}{16}\right) + \frac{3}{4}y\left(\frac{1}{4}\right) - \frac{3}{4}y\left(\frac{3}{4}\right) - y\left(\frac{15}{16}\right)\right] : 4$$

— Κανονικὸν Σφάλμα Ἀριθμητικοῦ Μέσου: $\sigma_m = 1.04\sigma : \sqrt{N}$

— Κανονικὸν Σφάλμα Μέσης Τετραγωνικῆς Ἀποκλίσεως: $\sigma_s = 5\sigma : 6\sqrt{N}$.

Παράδειγμα :

2	2	6	9	9	10	15	18	20	20
21	21	23	23	26	27	27	27	28	28
29	29	30	33	34	34	34	34	34	35
35	37	37	37	37	37	38	38	38	39
39	39	40	42	44	44	45	47	51	58

Εἰς αὐτὴν τὴν σειρὰν τῶν 50 παρατηρήσεων χρειαζόμεθα τιμάς, ὑπερβαλλομένας κατὰ 3, 12.5, 25, 37.5, 47 παρατηρήσεις καὶ αὗται εἶναι περίπου αἱ $y\left(\frac{1}{16}\right) = 46$, $y\left(\frac{1}{4}\right) = 38$, $y\left(\frac{1}{2}\right) = 34$, $y\left(\frac{3}{4}\right) = 22.5$ καὶ $y\left(\frac{15}{16}\right) = 7.5$.

$$\widehat{m} = 46 + 38 + 2 \times 34 + 22.5 + 7.5 : 6 = 30.3$$

$$\widehat{S} = 46 + 38 \times 3/4 - 22.5 \times 3/4 - 7.5 : 4 = 12.5$$

$$\sigma_m = 1.04 \times 12.5 : \sqrt{50} = \pm 1.83$$

$$\sigma_s = 5 \times 12.5 : 6\sqrt{50} \pm 1.47.$$

Παρατηρήσεις:

α) Ἡ μέθοδος αὐτὴ εἶναι Ταχυτέρα τῆς συνήθους, ἀποτελοῦσα ἐπεξεργασίαν τῆς μεθόδου 6.

β) Ἐπὶ Κανονικῆς Κατανομῆς ἡ μέθοδος ἐκτιμᾷ τὸν Ἀριθμητικὸν Μέσον καὶ Μέσην Τετραγωνικὴν Ἀπόκλισιν μὲ ἀποτελεσματικότητα 93% καὶ 73% ἀντιστοίχως.

γ) Ἑτέρα παραλλαγὴ τῆν μεθόδου ταύτης μὲ ἀποτελεσματικότητα 81% ἀλλὰ μεγαλυτέραν εὐπάθειαν εἰς ἀποκλίσεις ἀπὸ τῆς κανονικότητος ἔχει ὡς ἑξῆς:

$$\widehat{S}\left[y\left(\frac{1}{33}\right) + y\left(\frac{1}{7}\right) - y\left(\frac{1}{7}\right) - y\left(\frac{1}{33}\right) : 6 \quad \text{καὶ} \quad \sigma_s = 0.8\,\sigma : \sqrt{N}.\right.$$

8. *Ὅρια ἐκτιμήσεως ἀριθμητικοῦ μέσου ἢ διαφορᾶς μέσων*

Ἀποτελεσματικαὶ Μέθοδοι: Α. Διὰ τὸν προσδιορισμὸν ὁρίων τοῦ Ἀριθμητικοῦ Μέσου χρησιμοποιοῦμεν τὴν Κατανομὴν Student $t = (m - \mu) : \sigma_m$ ἔνθα μ εἶναι ὁ ἀληθὴς Μέσος, καὶ $\nu = N - 1$ βαθμοὶ ἐλευθερίας. Β. Ὁσάκις οἱ δύο Μέσοι προέρχωνται ἐξ ὁμάδων παρατηρήσεων μὲ ἴσας Μέσας, Τετραγωνικὰς Ἀποκλίσεις (πρᾶγμα ὅπερ ἐλέγχεται διὰ τῆς μεθόδου 9) τὰ ὅρια διὰ τὴν διαφορὰν τῶν δύο Μέσων προσδιορίζονται διὰ τοῦ ὑπολογισμοῦ τοῦ t Student ὡς ἀκολούθως:

$$\boxed{t = \frac{[(m_1 - m_2) - (\mu_1 - \mu_2)]}{\sigma_{m_1 - m_2}}} \qquad \text{καὶ μὲ } \nu = N_1 + N_2 - 2$$

ἔνθα $$(N_1 + N_2 - 2)\sigma^2_{m_1 - m_2} = [(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2]\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$$

Γ. Εἰς Μεγάλα Δείγματα ἔχομεν $$d = \frac{[(m_1 - m_2) - (\mu_1 - \mu_2)]}{\sigma_{m_1 - m_2}}$$

Σημ. $\sigma_{m_1 - m_2} = \sigma_{m_1}{}^2 + \sigma_{m_2}{}^2$ εἶναι κανονικὴ ἀπόκλισις, ἄσχετος τῆς ἰδιότητος τῶν Μέσων Τετραγωνικῶν Ἀποκλίσεων.

Παράδειγμα:

(1)

2.70	4.01	4.14	4.16	4.22	4.59	4.66	4.82	5.66
4.98	5.01	5.03	5.11	5.34	5.35	5.62	5.64	
5.76	5.78	5.83	6.16	6.37	7.12	7.82	4.97	

(ιι)

1.22	1.73	2.50	2.74	2.83	3.18	3.18	3.36	4.24
3.66	3.68	3.70	3.77	3.85	3.86	3.87	3.89	
4.29	4.41	4.54	4.81	4.87	4.89	5.48	3.58	

Δι' αὐτὰς τὰς δύο ὁμάδας τῶν 25 παρατηρήσεων θὰ ἔχωμεν :

$m_1 = 5.234 \qquad S_1 = 1.047 \qquad \sigma_{m_1} = 1.047 : \sqrt{25} = 0.209$

$m_2 = 3.685 \qquad S_2 = 0.978 \qquad \sigma_{m_2} = 0.978 : \sqrt{25} = 0.196$

Δοθέντος ὅτι ἡ τιμὴ 5 % διὰ t μὲ $\nu = 24$ εἶναι 2.06 τὰ ὅρια 95 % εἶναι :

$$5.234 \pm 2.06 \times 0.209 = 4.803 - 5.665$$

$$3.685 \pm 2.06 \times 0.196 = 3.281 - 4.089$$

Ἀλλ' ἐπειδὴ καὶ αἱ Μέσαι Τετραγωνικαὶ Ἀποκλίσεις δὲν διαφέρουσι σημαντικῶς, τὰ ὅρια 95% διὰ τὴν διαφορὰν μεταξὺ τῶν Μέσων ἐπιβάλλουσι τὴν χρῆσιν τῆς τιμῆς t μὲ $\nu = 48$ βαθμοὺς ἐλευθερίας, ἥτις εἶναι 2,01 καὶ τὸ Κανονικὸν Σφάλμα Διαφορᾶς :

$$\sigma_{m_1 - m_2} = \sqrt{(0.209)^2 + (0.196)^2} = 0.287$$

$(m_1 - m_2) = 5.234 - 3.685 = 1.549$ καὶ $1.549 \pm 2.01 \times 0.287 = 0.972 = 2.126$. Ἐὰν ἡ διαφορὰ τῶν Μέσων Τετραγωνικῶν Ἀποκλίσεων ἦτο μεγάλη μόνον ἡ Γ' μέθοδος θὰ συνιστᾶτο.

Παρατηρήσεις :

α) Αἱ μέθοδοι Α καὶ Β, ἀποτελεσματικαὶ 100%, ὑποθέτουσι κανονικότητα καὶ εἶναι εὐπαθεῖς εἰς ἀσυμμετρίαν ἐπὶ μικρῶν δειγμάτων. Τοῦτο δὲν συμβαίνει διὰ δύο ἰσομεγέθεις καὶ ἐξ ἴσου κατανεμηθείσας ὁμάδας παρατηρήσεων.

β) Ἡ μέθοδος Γ' δὲν ἔχει ἀνάγκην ἀποτελεσματικῶν ἐκτιμήσεων, τῶν Μέσων καὶ Μέσων Τετραγωνικῶν Ἀποκλίσεων.

9. Ὅρια ἐκτιμήσεως Μέσων Τετραγωνικῶν Ἀποκλίσεων (σφαλμάτων) ἢ ὁ λόγος F τῶν Μέσων Τετραγωνικῶν Ἀποκλίσεων

Ἀποτελεσματικαὶ Μέθοδοι 100%: Α. Ἡ $(N-1)S^2 : \sigma^2$ μὲ $\nu = N - 1$ βαθμοὺς ἐλευθερίας ἀκολουθεῖ τὴν Κατανομὴν χ^2 διὰ τὸν προσδιορισμὸν τῶν ἀκριβῶν ὁρίων τῆς σ. Παράδειγμα: Εἰς τὰς κατὰ τὰς μεθόδους 1 καὶ 2 ἀναλυθείσας 50 παρατηρήσεις ἔχομεν $S = 12.37$ καὶ $(N - 1)S^2 = 7500$. Χρησιμοποιοὔντες τὰς τιμὰς χ^2 μὲ 49 βαθμοὺς ἐλευθερίας διὰ $P = 0.975$ καὶ $P = 0.025$ θὰ ἔχωμεν:

$$31.6 \leq \frac{7500}{\sigma} \leq 70.2 \quad \text{ἤτοι} \quad 10.34 < \sigma < 15.41 \quad \text{μὲ πιθανότητα } 0.95.$$

Β. Κατὰ προσέγγισιν μέθοδος ὁρίων τῆς σ χρησιμοποιεῖ τὸ Κανονικὸν Σφάλμα τῆς S εἰς ὅρους τῆς σ καὶ οὕτω προκύπτει:

$$\boxed{\frac{S}{1+\alpha d} \leq \sigma \leq \frac{S}{1-\alpha d}}$$

ἔνθα d εἶναι ἡ κανονικὴ ἀπόκλισις προσδιωρισμένης πιθανότητος:

καὶ

$$\boxed{\alpha = \frac{1}{\sqrt{2 \times (N - 0.75)}}}$$

Παράδειγμα: Ὡς καὶ εἰς τὴν προηγουμένην μέθοδον

$$\alpha = \frac{1}{\sqrt{2 \times 49.25}} = 0.101 \qquad d = 1.96$$

καὶ

$$\frac{12.37}{1 + 1.96 \times 0.101} \leq \sigma \leq \frac{12.37}{1 - 1.96 \times 0.101}$$

$$10.33 \leq \sigma \leq 15.42 \quad \text{μὲ πιθανότητα } 0.95$$

Γ. Ὁ λόγος $S_1^2\, \sigma_2^2 : S_2^2\, \sigma_1^2$ ἀκολουθεῖ τὴν Κατανομὴν F μὲ $\nu = N_1 - 1$ καὶ $N_2 - 1$ βαθμοὺς ἐλευθερίας καὶ τὸ ἀποτέλεσμα τοῦτο δίδει ἀκριβῆ ὅρια τοῦ λόγου $\sigma_1 : \sigma_2$.

Παράδειγμα: Ἐὰν εἰς τὰς 50 παρατηρήσεις τῆς Μεθόδου 1 προσθέσωμεν καὶ δευτέραν σειρὰν 20 παρατηρήσεων μὲ ἐκτιμηθεῖσαν Μέσην Τετραγωνικὴν Ἀπόκλισιν 17.45 τότε χρησιμοποιοῦντες τιμὰς τῆς F μὲ 49 καὶ 19 βαθμοὺς ἐλευθερίας διὰ $P = 0.975$ καὶ $P = 0.025$ θὰ ἔχωμεν:

$$0.438 \leq \left(\frac{17.45\, \sigma_1}{12.37\, \sigma_2} \right)^2 \leq 2.06$$

$$0.469 \leq \frac{\sigma_1}{\sigma_2} \leq 1.017 \quad \text{μὲ πιθανότητα } 0.95$$

Δ. Κατὰ προσέγγισιν ὅρια τοῦ λόγου $\sigma_1 : \sigma_2$ παρέχονται διὰ τῶν Κανονικῶν Σφαλμάτων $\alpha_1 \sigma_1$ καὶ $\alpha_2 \sigma_2$ τῶν S_1 καὶ S_2.

Ἐὰν $\alpha = \sqrt{\alpha_1^2 + \alpha_2^2}$ τότε

$$\boxed{\frac{1-1/2\,\alpha d}{1+1/2\,\alpha d} \leq \frac{S_1\sigma_1}{S_2\sigma_2} \leq \frac{1+\frac{1}{2}\,ad}{1-\frac{1}{2}\,ad}}$$

καὶ κατὰ τὸ ἀνωτέρω παράδειγμα: $\alpha = \left(\frac{1}{2\times 49.25} + \frac{1}{2\times 19.25}\right) = 0.190$

καὶ $$\frac{12.37}{17.45}\;\frac{1-0.98\times 0.190}{1+0.98\times 0.190} \leq \frac{\sigma_1}{\sigma_2} \leq \frac{12.37}{17.45}\;\frac{1+0.98\times 0.190}{1-0.98\times 0.190}$$

ἤτοι $$0.486 \leq \frac{\sigma_1}{\sigma_2} \leq 1.033$$ μὲ πιθανότητα 0.95

Παρατηρήσεις:

α) Ὑποτίθεται ὅτι αἱ παρατηρήσεις κατανέμονται κανονικῶς.

β) Αἱ μέθοδοι εἶναι ἀποτελεσματικαὶ 100% ἀλλ' εὐπαθεῖς εἰς ἀποκλίσεις ἀπὸ τῆς κανονικότητος.

γ) Αἱ μέθοδοι Β καὶ Δ δύνανται νὰ χρησιμοποιηθῶσι χωρὶς ἀποτελεσματικὴν ἐκτίμησιν τῶν Μέσων Τετραγωνικῶν Ἀποκλίσεων.

10. *Ἐκτίμησις καὶ ἔλεγχος διαφορᾶς δύο Μέσων μεγάλων ὁμάδων παρατηρήσεων*

Χρησιμοποιοῦμεν τὰς Μεθόδους 6 καὶ 7 πρὸς ἐκτίμησιν τῶν Μέσων τῶν δύο ὁμάδων m_1 m_2 καὶ τὰ Κανονικὰ Σφάλματα τούτων σ_{m_1} καὶ σ_{m_2} μεθ' ὃ ἐλέγχομεν τὴν διαφορὰν $m_1 - m_2$ ὡς κανονικὴν ἀπόκλισιν μὲ Κανονικὸν Σφάλμα $\sigma_{m_1-m_2} = \sqrt{(\sigma_{m_1}^2 + \sigma_{m_2}^2)}$.

Παράδειγμα: Ἔστωσαν τὰ δεδομένα τοῦ παραδείγματος 8. Χρησιμοποιοῦντες τὴν μέθοδον 6 εἰς τὰς δύο αὐτὰς ὁμάδας τῶν 25 παρατηρήσεων θὰ ἔχωμεν κατὰ προσέγγισιν:

$$y_1\left(\frac{1}{16}\right) = 7.2\,,\quad y_1\left(\frac{1}{2}\right) = 5.1\,,\quad y_1\left(\frac{15}{16}\right) = 3.9\,,\qquad m_1 = 5.28$$

$$S_1 = 1.10,\qquad \sigma_{m_1} = \frac{1.1\times 1.10}{\sqrt{25}} = 0.23$$

$$y_2\left(\frac{1}{16}\right)=5.0,\quad y_2\left(\frac{1}{2}\right)=3.8,\quad y_2\left(\frac{15}{16}\right)=1.6,\quad m_2=3.60,$$

$$S_2=1.13,\quad \sigma_{m_2}=\frac{1.1\times 1.13}{\sqrt{25}}=0.25,\quad m_1-m_2=1.68$$

$$\sigma_{m_1-m_2}=\sqrt{[(0.23)^2+(0.25)^2]}=\pm\,0.34$$

Κανονικὴ Ἀπόκλισις 4.94. Ἄρα ἡ διαφορὰ 1.68 εἶναι σημαντικὴ εἰς τὸ ἐπίπεδον 0.1$^0/_0$.

Παρατηρήσεις:

α) Ἡ μέθοδος εἶναι ταχυτέρα τῆς συνήθους καὶ ἀποτελεσματικὴ 83–93$^0/_0$.

β) Εἶναι ὀλιγώτερον εὐπαθὴς εἰς ἀποκλίσεις, ἀπὸ τῆς κανονικότητος ἰδίᾳ τῆς ἀσυμμετρίας τῶν ἄλλων μεθόδων ἐκτιμήσεως.

γ) Ἡ μέθοδος εἶναι ἐπαρκὴς δι' ὁμάδας ἄνω τῶν 20 παρατηρήσεων, δύναται δὲ νὰ χρησιμοποιηθῇ μὲ μειωμένην ἀκρίβειαν εἰς ὁμάδας μέχρι 9 παρατηρήσεων.

11. *Ὅρια ἐκτιμήσεως τῶν διαμέσων ἢ διαφορῶν διαμέσων μεγάλων ὁμάδων παρατηρήσεων*

Διὰ τὸν προσδιορισμὸν τῶν ὁρίων τοῦ Διαμέσου ἐκτιμῶμεν τὴν ποσότητα $N+1:2\pm d\sqrt{N}$, ἔνθα d εἶναι κανονικὴ ἀπόκλισις, ἀντιστοιχοῦσα εἰς προσδιωρισθεῖσαν πιθανότητα καὶ διὰ τὴν ἐκτίμησιν τῶν ὁρίων τῆς διαφορᾶς μεταξὺ τῶν Διαμέσων τῶν δύο ὁμάδων τῶν N_1 καὶ N_2 παρατηρήσεων ὑπολογίζομεν τὰς διαφορὰς ὡς ἀκολούθως:

Διὰ τὴν Α' ὁμάδα $$\left[\frac{N_1+1}{2}\pm\frac{1}{2}d\sqrt{\left(\frac{N_1 N_2}{N_1+N_2}\right)}\right]$$

Διὰ τὴν Β' ὁμάδα $$\left[N_2+1:2\pm\frac{1}{2}d\sqrt{\left(\frac{N_1 N_2}{N_1+N_2}\right)}\right]$$

Παράδειγμα: Τὸ τῆς μεθόδου 8.

$M_{e_1}=25+1:2=13$η παρατήρησις, ἤτοι 5.11

$M_{e_2}=25+1:2=13$η παρατήρησις, ἤτοι 3.77 καὶ $M_{e_1}-M_{e_2}=1.34$·

Τὰ ὅρια 95$^0/_0$ διὰ τὰς τιμὰς αὐτὰς παρέχονται ὑπὸ $(13\pm\sqrt{25})$, ἤτοι τῆς 8ης καὶ 18ης παρατηρήσεως καὶ εἶναι ταῦτα διὰ τὴν α' ὁμάδα 4.82 καὶ 5.66 καὶ διὰ τὴν β' ὁμάδα 3.36 καὶ 4.24. Ἡ διαφορὰ εἶναι σαφῶς σημαντική.

Διὰ νὰ εὕρωμεν τὰ ὅρια 95% τῆς διαφορᾶς δέον νὰ ἐξετάσωμεν τὰς διαφορὰς μεταξὺ $13 + \sqrt{25:2}$ ἢ τὴν 16.5 παρατήρησιν τῆς α΄ ὁμάδος καὶ $13 - \sqrt{25:2}$ ἢ τὴν 9ην τῆς δευτέρας. Τὰ δύο ὅρια εἶναι περίπου $5.62 + 5.64 : 2 - 3.58 + 3.66 : 2 = 2.01$ καὶ $4.97 + 4.98 : 2 - 3.87 + 3.89 : 2 = 1.09$.

Παρατήρησις:

Ἡ μέθοδος εἶναι Ταχεῖα, Ἐλευθέρα Κατανομῆς καὶ περιωρισμένης ἀποτελεσματικότητος 64% εἰς κανονικῶς κατανεμηθείσας παρατηρήσεις.

12. *Ὅρια ἐκτιμήσεως τῶν Μέσων ἢ διαφορῶν μεταξὺ Μέσων δύο μικρῶν ἰσομεγέθων ὁμάδων παρατηρήσεων*

ι) Ὑπολογίζομεν τοὺς Μέσους m_1, m_2 καὶ τὰ Εὔρη R_1, R_2 τῶν δύο ὁμάδων παρατηρήσεων.

ιι) Προσδιορίζομεν τὰ ὅρια τῶν Μέσων $m_i \pm C\,R_i$.

ιιι) Προσδιορίζομεν τὰ ὅρια τῆς διαφορᾶς μεταξὺ τῶν Μέσων:

$$m_1 - m_2 \pm d\,(R_1 + R_2)$$

Αἱ τιμαὶ τῶν σ καὶ d δίδονται ἐκ τοῦ ἀκολούθου Πίνακος διὰ παρατηρήσεις N εἰς ἑκάστην ὁμάδα.

	N	2	3	4	5	6	7	8	9	10	11	12
C	99 %	31.83	3.009	1.316	0.842	0.628	0.502	0.429	0.374	0.333	0.302	0.276
	95 %	6.353	1.303	0.717	0.506	0.399	0.333	0.288	0.255	0.230	0.210	0.194
d	99 %	3.958	1.047	0.618	0.447	0.352	0.299	0.261	0.232	0.210	0.192	0.177
	95 %	1.715	0.637	0.407	0.306	0.249	0.213	0.187	0.167	0.152	0.140	0.130

Παράδειγμα:

(ι) 4.01, 4.16, 4.82, 5.34, 5.35, 6.37, 7.12, 7.82

(ιι) 1.73, 2.50, 3.18, 3.18, 3.58, 3.68, 4.41, 5.48

$m_1 = 5.62 \qquad m_2 = 3.47 \qquad m_1 - m_2 = 2.15$

$R_1 = 3.81 \qquad R_2 = 3.75 \qquad R_1 + R_2 = 7.56$

Τὰ ὅρια 95 % τῶν Μέσων εἶναι :

$$5.62 \pm 0.288 \times 3.81 = 4.52 - 6.72$$

$$3.47 \pm 0.288 \times 3.75 = 2.39 - 4.55$$

Τὰ ὅρια 95 % τῆς διαφορᾶς τῶν Μέσων εἶναι ἀντιστοίχως :

$$2.15 \pm 0.187 \times 7.56 = 0.74 - 3.56$$

Παρατήρησις :

Αἱ μέθοδοι εἶναι Ταχεῖαι, ὑψηλῆς ἀποτελεσματικότητος 95 % διὰ κανονικῶς κατανεμηθείσας παρατηρήσεις ἀλλ' εἶναι ὑποκείμεναι εἰς εὐπάθειαν μεγάλων ἀποκλίσεων ἀπὸ τῆς κανονικότητος.

13. *Ὅρια ἐκτιμήσεως Ἀριθμητικοῦ Μέσου μικροῦ ἀριθμοῦ παρατηρήσεων*

Ἡ μέθοδος συνίσταται εἰς τὴν ἐκλογὴν ἐκείνων τῶν τιμῶν, διὰ τὰς ὁποίας τὸ σύνολον τῶν βαθμολογουμένων ἀποκλίσεων (θετικῶν) τῶν παρατηρήσεων δὲν διαφέρει σημαντικῶς ἀπὸ τοῦ συνόλου τῶν ἀρνητικῶν ἀποκλίσεων (ἀδιαφόρως σημείων). Πρὸς τοῦτο μετρῶμεν κατὰ προσέγγισιν τὰ ὅρια ἐκτιμήσεως, διὰ τὰς πλησίον ἀλλήλων τιμάς, ὑπολογίζομεν καὶ ταξινομοῦμεν κατὰ σειρὰν μεγέθους τὰς ἀποκλίσεις τῶν N παρατηρήσεων καὶ τέλος ἐφαρμόζομεν τὸν τύπον :

$$\boxed{L = R - \frac{N(N+1)}{4} + d\sqrt{\left(\frac{N(N+1)(2N+1)}{24}\right)}}$$

ἔνθα R εἶναι τὸ σύνολον τῶν βαθμολογημένων θετικῶν ἢ ἀρνητικῶν ἀποκλίσεων καὶ d κανονικὴ ἀπόκλισις προσδιωρισμένης πιθανότητος, ἡ ὁποία ἐξαφανίζεται εἰς τὸ μέγιστον ἢ ἐλάχιστον ὅριον.

Παράδειγμα :

4.01,	4.14,	4.16,	4.82,	4.98,	5.11
5.34,	5.35,	5.64,	6.37,	7.12,	7.82

$$N = 12, \quad m = 5.405, \quad S = 1.2, \quad \sigma_s = \frac{1.2}{\sqrt{12}} = 0.35.$$

Τὰ ὅρια 95 % εἶναι περίπου $5.40 \pm 0.7 = 4.7 - 6.1$.

καὶ
$$L = R - \frac{12 \times 13}{4} + 1.96\sqrt{\frac{12 \times 13 \times 25}{24}} = R - 14$$

Αἱ ἀποκλίσεις ἀπὸ τῆς τιμῆς -6.1 εἶναι -2.09, -1.96, -1.94, -1.28, -1.12, -0.99, -0.76, -0.75, -0.46, 0.27, 1.02, 1.72. Αἱ θετικαὶ ἀποκλίσεις ἔχουσι βαθμοὺς 1.6 καὶ 9, ἄρα $L = 1 + 6 + 9 - 14 = 2$.

Αὐξάνοντες τὴν δοκιμαστικὴν τιμὴν εἰς 6.2 ἔχομεν βαθμοὺς 1, 5, 9 καὶ κατὰ συνέπειαν $L = 1$. Τελικῶς οἱαδήποτε τιμὴ μεταξὺ 6.23 καὶ 6.24 δίδει βαθμοὺς 1,4 καὶ 9 μὲ $L = 0$. Ἡ τιμὴ 6.24 παρέχει τὸ μέγιστον ὅριον διὰ τὸν Ἀριθμητικὸν Μέσον. Ὁμοίως τιμαὶ μεταξὺ 4.625 καὶ 4.635 παρέχουσιν ἓν ἐλάχιστον ὅριον διὰ τὸν Ἀριθμητικὸν Μέσον καὶ τὰ 95 % ὅρια θὰ εἶναι οὕτω 4.625 καὶ 6.24 (4.65 καὶ 6.16 ἐπὶ τῇ ὑποθέσει κανονικότητος).

Παρατήρησις:

Ἡ μέθοδος εἶναι ταχυτέρα τῆς συνήθους, ἐλευθέρα κατανομῆς καὶ μὲ ἀποτελεσματικότητα 95 % διὰ κανονικῶς κατανεμηθείσας παρατηρήσεις.

14. *Ἔλεγχος τῆς διαφορᾶς μέσου ἐπιπέδου μεταξὺ δύο ὁμάδων παρατηρήσεων*

Κατὰ τὴν μέθοδον αὐτὴν δίδομεν σκὸρ εἰς τὰς παρατηρήσεις βάσει τῶν βαθμῶν των εἰς τὰς δύο ὁμάδας, λαμβανομένας ὁμοῦ. Ἴσαι παρατηρήσεις λαμβάνουσι ἑκάστη τὸν μέσον τῶν βαθμῶν, τὸν ὁποῖον ἀπὸ κοινοῦ κατέχουσι. Τὰ στάδια ἐργασίας ἔχουσιν ὡς ἀκολούθως :

(ι) Ὑπολογίζομεν τὰ Σύνολα καὶ Μέσους τῶν σκὸρ διὰ τὰς δύο ὁμάδας κεχωρισμένως.

(ιι) Ὑπολογίζομεν τὴν διαφορὰν τῶν μέσων d καὶ τὸ γενικὸν συνολικὸν σκόρ :

$$\boxed{T = \frac{(M + N)(M + N + 1)}{2}}$$

(ιιι) Ἐλέγχομεν τὸ d ὡς κανονικὴν ἀπόκλισιν μὲ Κανονικὸν Σφάλμα:

$$\boxed{\sigma_d = \sqrt{\left[\frac{T}{6}\left(\frac{1}{M} + \frac{1}{N}\right)\right.}}$$

ἔνθα M καὶ N εἶναι οἱ ἀριθμοὶ τῶν παρατηρήσεων εἰς τὰς δύο ὁμάδας.

Παράδειγμα:

(α) 4.01, 4.14, 4.16, 4.82, 4.98, 5.11, 5.34, 5.35, 5.64, 6.37, 7,12 7.82

(β) 1.73, 2.50, 3.18, 3.18, 3.58, 3.68, 4.41, 5.48

Ἐὰν αἱ δύο αὐταὶ ὁμάδες διαταχθῶσι βαθμολογικῶς κατ' ἀντίστροφον τάξιν, θὰ ἔχωμεν :

		Σύνολον
(α)	7, 8, 9, 11, 12, 13, 14, 15, 17, 18, 19, 20	163
(β)	1, 2, $3\,^1/_2$, $3\,^1/_2$, 5, 6, 10, 16	47

Ἐλέγχοντες ἔχομεν $T = 163 + 47 = 210$ ἢ $20 \times 21 : 2 = 210$

$$d = 163 : 12 - 47 : 8 = 13.6 - 5.9 = 7.7$$

$$\sigma_d = \sqrt{[210 : 6\,(1 : 12 + 1 : 8)]} = \pm 2.7$$

Οὕτως ἡ διαφορὰ 7.7 εἶναι σχεδὸν τριπλασία τοῦ Κανονικοῦ της Σφάλματος καὶ συνεπῶς σημαντικὴ εἰς τὸ ἐπίπεδον 1 %.

Παρατηρήσεις :

— Ἡ μέθοδος εἶναι ἐλευθέρα κατανομῆς, ταχυτέρα τῆς συνήθους καὶ μὲ ἀποτελεσματικότητα 95 % διὰ μεγάλας ὁμάδας κανονικῶς κατανεμηθείσας.

— Ἡ μέθοδος δύναται νὰ χρησιμοποιηθῇ δι' ὁμάδας πλέον τῶν τεσσάρων παρατηρήσεων ἀλλ' εἶναι ὀλιγώτερον ἀποτελεσματικὴ διὰ μικροτέρας τοιαύτας.

15. Ἔλεγχος τῆς διαφορᾶς μεταξὺ δύο μεγάλων ἰσομεγέθων ὁμάδων παρατηρήσεων

Κατὰ τὴν μέθοδον ταύτην θεωροῦμεν τὰς δύο ὁμάδας ὁμοῦ, ὑπολογίζοντες τὸν ἀριθμὸν τῶν παρατηρήσεων ἑκάστης ὁμάδος, ὅστις πίπτει κάτω τῆς μικροτέρας παρατηρήσεως τῆς ἄλλης ὁμάδος καὶ τὸν ἀριθμὸν τῶν παρατηρήσεων ἑκάστης ὁμάδος, ὅστις εἶναι μεγαλύτερος τῆς μεγαλυτέρας παρατηρήσεως τῆς ἄλλης ὁμάδος καὶ τέλος προσθέτοντες τὰς δύο αὐτὰς τιμὰς ὁμοῦ. Ἐννέα ἢ περισσότεραι παρατηρήσεις εἶναι σημαντικαὶ εἰς τὸ ἐπίπεδον 5 % καὶ δώδεκα ἢ πλείονες εἰς τὸ ἐπίπεδον 1 %.

Παράδειγμα : τὸ τῆς μεθόδου 8.

— Ἀριθμὸς παρατηρήσεων Πρώτης Ὁμάδος ὑπεράνω τῆς ὑψηλοτέρας παρατηρήσεως 5.48 τῆς Δευτέρας Ὁμάδος	=	10
— Ἀριθμὸς παρατηρήσεων Δευτέρας Ὁμάδος κάτω τῆς κατωτάτης παρατηρήσεως 2.70 τῆς Πρώτης Ὁμάδος	=	3
Σύνολον		13

Ὁ ἀριθμὸς 13 εἶναι σημαντικὸς εἰς τὸ ἐπίπεδον 1 % :

$$(t \text{ Student} = 5.41 \qquad P < 0.001)$$

Παρατηρήσεις :

– Εἶναι μέθοδος ταχυτάτη, ἐλευθέρα κατανομῆς, ἐπιτρέπουσα τὸν ἔλεγχον τῶν διαφορῶν μέσου ἐπιπέδου ὅσον καὶ μέσης τετραγωνικῆς ἀποκλίσεως.

– Ἰσχύει μὲ ἠλαττωμένην ἀποτελεσματικότητα ἐὰν ἡ μεγαλυτέρα Ὁμὰς εἶναι ἄνω τοῦ ἑνὸς τρίτου ἔναντι τῆς μικροτέρας Ὁμάδος.

16. Ἔλεγχος διαφορῶν μεταξὺ ὁμάδων τινῶν παρατηρήσεων

(ι) Ὑπολογίζομεν τὸ σύνολον T_i τῶν n_i παρατηρήσεων εἰς τὴν ὁμάδα i. Ἔστω ὅτι τὸ σύνολον καὶ συνολικὸν ἄθροισμα τετραγώνων τῶν παρατηρήσεων εἰς ὅλας τὰς m ὁμάδας ὅτι εἶναι T καὶ S ἀντιστοίχως.

(ιι) Ἀναλύομεν τὴν Διακύμανσιν ὡς ἑξῆς :

	Βαθμοὶ Ἐλευθερίας	Ἄθροισμα Τετραγώνων	Μέσον Τετράγωνον
Μεταξὺ Ὁμάδων	$m-1$	$A = \Sigma \frac{T_i^2}{n_i} - \frac{T^2}{N}$	$S_1^2 = A : m-1$
Ἐντὸς Ὁμάδων	$N-m$	$B = S - \Sigma \frac{T_i^2}{n_i}$	$S^2 = B : N-m$
Σύνολον	$N-1$	$S - T^2 : N$	

(ιιι) Ἐλέγχομεν τὰς διαφορὰς διὰ τοῦ Λόγου Διακυμάνσεων $F = S_1^2 : S^2$ μὲ $m-1$ καὶ $N-m$ βαθμοὺς ἐλευθερίας.

(ιv) Τὸ Κανονικὸν Σφάλμα τοῦ Ἀριθμητικοῦ Μέσου οἱασδήποτε ὁμάδος θὰ εἶναι : $\sigma_m = S \sqrt{n_i}$ $(m = T_i : n_i)$

Παράδειγμα :

(αα)	2,	6,	10,	23,	28,	34,	38,	39,	45,	47
(ββ)	4,	17,	20,	30,	36,	39,	39,	42,	44,	46
(γγ)	13,	25,	31,	31,	33,	33,	34,	38,	41,	42
(δδ)	26,	27,	33,	40,	41,	43,	43,	44,	45,	57
(εε)	17,	26,	31,	34,	41,	42,	43,	47,	52,	66

Ἡ ἀνάλυσις τῶν ἄνωθι πέντε ὁμάδων τῶν δέκα παρατηρήσεων γίνεται ὡς ἀκολούθως :

$T_1 = 272$, $T_2 = 317$, $T_3 = 321$, $T_4 = 399$, $T_5 = 401$, $T = 1710$

$S = 66942$, $T^2 = 2924100$, $T^2 : 50 = 58482$, $\Sigma T_i^2 : 10 = 59751.6$

$$F = \frac{317.4}{159,79} = 1.99, \qquad \text{Ἐπίπεδον}: 5\,\% = 2.58,$$

Μέσοι ὁμάδων : 27.2, 31.7, 32.1, 39.9, 40.1

$$\sigma_{m_i} = \sqrt{\frac{159.79}{10}} = \pm 4.$$

	Βαθμοὶ Ἐλευθερίας	Ἄθροισμα Τετραγώνων	Μέσον Τετράγωνον
Μεταξὺ Ὁμάδων	5 – 1	1269.6	317.40
Ἐντὸς Ὁμάδων	50 – 5	7190.4	159.79
Ἄθροισμα	49	8460	

Παρατήρησις :

Ἡ μέθοδος ὑποθέτει ὅτι ὅλαι αἱ παρατηρήσεις ἔχουσι κατανεμηθῆ κανονικῶς μὲ τὴν αὐτὴν Διακύμανσιν, δὲν ἐπηρεάζεται δὲ ἐξ ἐλλείψεως κανονικότητος δοθέντος ὅτι ὅλαι αἱ παρατηρήσεις ἐπηρεάζονται ἐξ ἴσου. Ἀποτελεσματικότης μεθόδου 100 %.

17. *Ἔλεγχος διαφορῶν μέσου ἐπιπέδου μεταξὺ ὁμάδων τινῶν παρατηρήσεων*

Ἐκτιμῶμεν τιμὰς ὑπερβαλλομένας κατὰ 1/4 καὶ 3/4 τῆς συνδεδυασμένης σειρᾶς παρατηρήσεων, ἤτοι y (1/4) καὶ y (3/4), ὑπολογίζομεν τὸν ἀριθμὸν παρατηρήσεων ἑκάστης ὁμάδος ὅστις ὑπερβαίνει y (1/4) ἢ πίπτει κάτω τοῦ y (3/4), καλούμενον n_i (1/4) καὶ n_i (3/4) μεθ' ὃ ἐλέγχομεν τὸ

$$\boxed{\sum_{i=1}^{m} \frac{\left[n_i\left(\frac{1}{4}\right) - n_i\left(\frac{3}{4}\right)\right]^2}{n_i\left(\frac{1}{4}\right) + n_i\left(\frac{3}{4}\right)}}$$, ὡς χ^2 μὲ m–1 βαθμοὺς ἐλευθερίας.

Παράδειγμα :

Τὸ τοιοῦτον τῆς μεθόδου 16. Διὰ νὰ ἐλέγξωμεν τὰς διαφορὰς μεταξὺ τῶν πέντε σειρῶν ἐκ δέκα παρατηρήσεων ὁρίζομεν y (1/4) = 42.5 καὶ y (3/4) = 27.5 καὶ προχωροῦμεν περαιτέρω εἰς τὴν ἀνάλυσιν ὡς κατωτέρω :

i	$n_i\left(\frac{1}{4}\right)$	$n_i\left(\frac{3}{4}\right)$	$\dfrac{\left[n_i\left(\frac{1}{4}\right) - n_i\left(\frac{3}{4}\right)\right]^2}{\left[n_i\left(\frac{1}{4}\right) + n_i\left(\frac{3}{4}\right)\right]}$
1	2	4	0.66
2	2	3	0.20
3	0	2	2.00
4	5	2	1.30
5	4	1	1.80
Σύνολον	13	12	$\chi^2 = 5.96$

Ἀλλὰ διὰ $P = 0.05$ $\chi^2 = 9.5$ καὶ συνεπῶς ἡ τιμὴ 5.96 δὲν εἶναι σημαντική. (Διὰ τὸν Λόγον Διακυμάνσεων $F = 1.99$ $P = 0.11$).

***Παρατηρήσεις*:**

—Ἡ μέθοδος εἶναι Ταχεῖα, Ἐλευθέρα Κατανομῆς εἰς ὁμάδας μεγάλας, κανονικῶς κατανεμηθεισῶν παρατηρήσεων καὶ ἀποτελεσματικότητος 81%.

—Ἀπαιτοῦνται ὁμάδες οὐχὶ ὀλιγώτεραι τῶν πέντε παρατηρήσεων.

18. Ἔλεγχος διαφορῶν μεταξὺ μεγάλων τινῶν ἰσομεγέθων ὁμάδων παρατηρήσεων

Ἐξετάζοντες τὰς ἀκραίας παρατηρήσεις ἑκάστης ὁμάδος ἐκλέγομεν τὰς μεγαλυτέρας τῶν μικροτέρων τιμὰς α καὶ τὰς μικροτέρας τῶν μεγαλυτέρων τιμὰς β μεθ' ὃ ὑπολογίζομεν τοὺς ἀριθμοὺς παρατηρήσεων κάτω τοῦ α καὶ ἄνω τοῦ β, ἀθροίζομεν τοὺς δύο ἀριθμοὺς ὁμοῦ καὶ ἐλέγχομεν τὸ σύνολον βάσει τοῦ ἀκολουθοῦντος Πίνακος:

Ἀριθμὸς Ὁμάδων	2	3	4	5	6	8	10
Μέσος	3	9	14.7	20.8	27.4	41.5	56.6
Ἐπίπεδον 5%	9	17	27	37	47	70	93
Ἐπίπεδον 5%	12	22	33	45	57	83	110

Παράδειγμα:

(ι)	2	4	6	10	17	20	23	28	30	34	36	38	39
	39	39	42	44	45	46	47						
(ιι)	13	25	26	27	31	31	33	33	33	34	38	40	41
	41	42	43	43	44	45	57						
(ιιι)	15	17	23	24	28	31	31	33	34	38	41	42	43
	44	47	48	51	52	59	66						

Εἰς αὐτὰς τὰς τρεῖς ὁμάδας ἐξ εἴκοσι παρατηρήσεων ἡ μεγαλυτέρα ἐκ τῶν χαμηλοτέρων τιμῶν εἶναι ὁ 15 καὶ ἡ μικροτέρα τῶν μεγαλυτέρων τιμῶν ὁ 47. Πέντε παρατηρήσεις πίπτουσι κάτω τοῦ 15 καὶ ἓξ καὶ ἥμισυ ἄνω τοῦ 47, τοῦ συνόλου αὐτῶν ὄντος 11.5.

Τὸ ἐπίπεδον σημαντικότητος 5 % εἶναι ὁ 17 καὶ οὕτω τὸ σύνολον 11.5 δὲν δύναται νὰ θεωρηθῇ σημαντικὸν ($F = 2.58 \quad P = 0.09$). Εἰς τὴν περίπτωσιν αὐτὴν ὁ ἀριθμὸς τῶν παρατηρήσεων ἑκάστης ὁμάδος εἶναι ἀρκετὰ μικρὸς οὕτως ὥστε ἡ σημαντικότης θὰ τείνῃ νὰ ὑποτιμᾶται ἀλλὰ τὸ σύνολον εὑρίσκεται τόσον πλησίον τοῦ Ἀριθμητικοῦ Μέσου 9.0 ὥστε περισσότερον εὐπαθὴς ἔλεγχος φαίνεται ἀπίθανον ὅτι θὰ ἀποφέρῃ σημαντικότητα.

Παρατήρησις:

Ἡ μέθοδος εἶναι Ταχυτάτη, Ἐλευθέρα Κατανομῆς μὲ μειωμένην ἀποτελεσματικότητα καὶ ὡς ἐπέκτασις τῆς μνημονευθείσης μεθόδου 15 ἔχει τὰς αὐτὰς ἰδιότητας.

Β΄. Μέθοδοι Συγκρίσεων Συχνοτήτων ἢ Ἀναλογιῶν

1. Ἔλεγχος διαφόρων σειρῶν ἀναλογιῶν (πίναξ συμπτώσεως $p \times m$*)*

Διατάσσομεν τὰς συχνότητας εἰς ἕνα Πίνακα Συμπτώσεως, τῶν σειρῶν ἀντιστοιχουσῶν εἰς ταξινομήσεις ἑκάστης ὁμάδος καὶ τῶν στηλῶν ἀντιστοιχουσῶν εἰς τὰς ὁμάδας. Ἔστωσαν O_{ij} ἡ συχνότης τῆς σειρᾶς i καὶ στήλης j, R_i C_j τὰ σύνολα σειρᾶς καὶ στήλης καὶ N τὸ γενικὸν σύνολον. Κατόπιν ὑπολογίζομεν τὴν ἀναμενομένην συχνότητα ἑκάστου φατνίου $E_{ij} = R_i C_j : N$ καὶ ἐλέγχομεν τὴν ποσότητα $\Sigma\Sigma(O_{ij} - E_{ij})^2 : E_{ij} = \Sigma\Sigma O_{ij}{}^2 : E_{ij} - N$ ὡς χ^2 μὲ $(p-1)(m-1)$ βαθμοὺς ἐλευθερίας.

Παράδειγμα :

	Ὁμὰς O_{ij} 1	2	3	4	R	Ὁμὰς E_{ij} 1	2	3	4
1	8	9	14	9	40	10.0	8	12	10.0
Τάξις 2	10	4	6	10	30	7.5	6	9	7.5
3	7	7	10	6	30	7.5	6	9	7.5
C_i	25	20	30	25	100	25	20	30	25

Ἐὰν αἱ τιμαὶ O_{ij} παριστῶσι τὰς συχνότητας εἰς 4 ὁμάδας 3 ἀναλογιῶν τότε ἡ E_{ij} ὑπολογίζεται ὡς εἰς τὸν ἀνωτέρω Πίνακα, ὁπότε μὲ βαθμοὺς ἐλευθερίας $(4-1)\,(3-1) = C$ θὰ ἔχωμεν :

$$\frac{(2)^2}{10} + \frac{(1)^2}{8} + \frac{(2)^2}{12} + \ldots\ldots\ldots \frac{(1.5)^2}{7.5} = 4.90$$

ἢ
$$\frac{100}{400}\left[\frac{8^2}{25} + \frac{9^2}{20} + \frac{14^2}{30} + \frac{9^2}{25}\right] + \frac{100}{30}\left[\frac{10^2}{25} + \frac{4^2}{20} + \frac{6^2}{30} + \frac{10^2}{25}\right] +$$

$$+ \frac{100}{30}\left[\frac{7^2}{25} + \frac{7^2}{20} + \frac{10^2}{30} + \frac{6^2}{25}\right] - 100 = 4.90$$

Ἡ τιμὴ $\chi^2 = 4.90$ δὲν εἶναι μεγαλυτέρα ἐκείνης, ἥτις θὰ ἀνεμένετο ἐκ τύχης $(P=0.5)$ οὕτως ὥστε αἱ διαφοραὶ δὲν εἶναι σημαντικαί.

Παρατηρήσεις :

(ι) Εἰς Πίνακα 2×2 μὲ $\nu = N-1$ βαθμοὺς ἐλευθερίας θὰ ἔχωμεν :

$$\frac{N\,(O_{11}\,O_{22} - O_{12}\,O_{21})^2}{R_1\,R_2\,C_1\,C_2}$$

(ιι) Εἰς Πίνακα $2 \times m$ μὲ $\nu = m-1$ βαθμοὺς ἐλευθερίας ἐλέγχεται ἡ ποσότης :

$$\frac{N}{R_1\,R_2}\left[\Sigma\,\frac{O_{1i}^{\,2}}{C_i} - \frac{R_1^{\,2}}{N}\right]$$

(ιιι) Ἡ μέθοδος εἶναι ἐλευθέρα κατανομῆς καὶ ἀπαιτεῖ ὅπως ἡ E_{ij} εἶναι μεγαλυτέρα τοῦ 2.

(ιν) Ἀρχικὸς ὑπολογισμὸς μεγαλυτέρων τιμῶν τοῦ $(O-E)^2$: E προδίδει σημαντικότητα τοῦ συνολικοῦ χ^2, ὁπότε παρέλκει ὁ πλήρης ὑπολογισμός του.

2. Ἔλεγχος διαφορᾶς δύο ἀναλογιῶν (Πίναξ συμπτώσεως 2×2)

α) Διατάσσομεν τὰς παρατηρήσεις εἰς τὸν κατωτέρω Πίνακα Συμπτώσεως 2×2:

O_{11}	O_{12}	R_1
O_{21}	O_{22}	R_2
C_1	C_2	N

καὶ ἐλέγχομεν τὴν $\dfrac{N(O_{11}O_{22}-O_{12}O_{21})^2}{R_1 R_2 C_1 C_2}$ μὲ $\nu = N-1$ βαθμοὺς ἐλευθερίας.

β) Διὰ τὰ κατὰ προσέγγισιν ἴσα ὁριακὰ σύνολα ἐλέγχομεν τὸ $O_{11}+O_{22}-O_{12}-O_{21}$, ὅπερ κατανέμεται κανονικῶς μὲ Μέσον $\dfrac{(C_1-C_2)(R_1-R_2)}{N}$

καὶ Κανονικὸν Σφάλμα $4\sqrt{\left(\dfrac{R_1 R_2 C_1 C_2}{N_3}\right)}$ ἢ περίπου $\sqrt{N}$.

γ) Δι' ἓν ζεῦγος ὁριακῶν συνόλων διαφερόντων σημαντικῶς π.χ. $R_1 < R_2$ ἐλέγχομεν $O_{11}-O_{12}$, ὅπερ κατανέμεται μὲ Μέσον $\dfrac{(C_1-C_2)R_1}{N}$ καὶ Κανονικὸν Σφάλμα $2\sqrt{\left(\dfrac{R_1 C_1 C_2}{N^2}\right)}$ ἢ περίπου $\sqrt{R_1}$.

Παραδείγματα:

I

	1	2	Σύνολον
1	25	42	67
2	56	33	89
Σύνολον	81	75	156

$$\chi^2 = \frac{156\,(33 \times 25 - 42 \times 56)^2}{67 \times 89 \times 81 \times 75} = 10.04 \text{ (Μετὰ τὴν διόρθωσιν YATES 9.04)}$$

Ἡ τιμὴ 10.04 ὑπερβαίνει τὸ 1 % ἐπίπεδον σημαντικότητος.

Κατὰ τὴν τροποποίησιν (β) ἔχομεν $25 + 33 - 42 - 56 = -40$ μὲ Μέσον $\frac{(81-75)(67-89)}{156} = -1$ καὶ Κανονικὸν Σφάλμα περίπου $\sqrt{156} = 12.5$. Ἡ διαφορὰ $-40-(-1) = -39$ εἶναι μεγαλυτέρα τοῦ τριπλασίου τοῦ Κανονικοῦ Σφάλματός της καὶ κατὰ συνέπειαν σημαντικὴ εἰς τὸ ἐπίπεδον 1 %.

II

.	1	2	Σύνολον
1	13	26	39
2	134	107	241
Σύνολον	147	133	280

$$\chi^2 = \frac{280\,(13 \times 107 - 26 \times 134)^2}{39 \times 241 \times 147 \times 133} = 6.68. \text{ (Μετὰ τὴν διόρθωσιν YATES 5.81)}$$

Ἡ τιμὴ 6.68 ὑπερβαίνει τὸ 1 % ἐπίπεδον σημαντικότητος.

Κατὰ τὴν τροποποίησιν (γ) θὰ ἔχωμεν: $13 - 26 - 13$ μὲ Μέσον $\frac{(147-133)\,39}{280} = 2$ καὶ Κανονικὸν Σφάλμα περίπου $\sqrt{39} = 6.25$. Ἡ διαφορὰ $-13-2 = -15$ εἶναι σαφῶς σημαντικὴ ($P < 0.02$).

Παρατηρήσεις :

(ι) Ἡ μέθοδος (α) εἶναι ἐλευθέρα κατανομῆς καὶ πλήρως ἀποτελεσματικὴ ἔναντι τῶν τροποποιήσεων (β) καὶ (γ), αἵτινες ἐνδείκνυνται εἰς τὰς ἀναφερομένας περιπτώσεις.

(ιι) Τὰ ὁριακὰ σύνολα δέον νὰ εἶναι τουλάχιστον πέντε.

3. *Ἔλεγχος καταλληλότητος προσαρμογῆς*

Διατάσσομεν τὰς παρατηρήσεις εἰς Πίνακα Συχνοτήτων μὲ n ὁμάδας, τῆς O_i παριστώσης τὴν συχνότητα τῆς i ὁμάδος, ὑπολογίζομεν τὰς ἀναμενομένας συχνότητας E_i ἑκάστης ὁμάδος μὲ τὴν ἐλεγχομένην ὑπόθεσιν καὶ ὑπὸ τὸν ὅρον ὅτι αὗται ἀπαιτοῦσι τὴν ἐκτίμησιν τῶν μεταβλητῶν $m-1$, ἐκ τῶν παρατηρήσεων καὶ ὅτι τὰ σύνολά των εἶναι ἴσα $\Sigma E_i = \Sigma O_i$ τέλος δὲ ἐλέγχομεν τὴν $\Sigma(O_i - E_i)^2 : E_i$ μὲ $\nu - m$ βαθμοὺς ἐλευθερίας.

Παράδειγμα :

Ὁμάδες i	Παρατηρηθεῖσαι Συχνότητες O_i	Ἀναμενόμεναι Συχνότητες E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	40	50	− 10	100	2
1	171	200	− 29	841	4.20
2	305	300	5	25	0.08
3	222	200	22	484	2.42
4	62	50	12	144	2.88
Σύνολον	800	800	0		$\chi^2 = 11.58$

Αἱ τιμαὶ E_i τοῦ ἀνωτέρω Πίνακος παριστῶσι μίαν θεωρητικὴν σειρὰν παρατηρήσεων τῆς διωνυμικῆς μεταβλητῆς, λαμβανούσης τὰς τιμάς, Ο ἕως 4. Ἐπὶ τῇ ὑποθέσει ὅτι ἡ πιθανότης ἐπιτυχίας εἶναι ἥμισυ αἱ ἀναμενόμεναι συχνότητες θὰ εἶναι οἱ λόγοι 1 : 4 : 6 : 4 : 1.

Δοθέντος ὅτι οὐδεμία μεταβλητὴ ἐξετιμήθη ἐκ τῶν δεδομένων θὰ ἔχωμεν τέσσαρας βαθμοὺς ἐλευθερίας καὶ ἡ τιμὴ 11.58 εἶναι σημαντικὴ εἰς τὸ ἐπίπεδον 2.5 %.

Παρατηρήσεις :

(ι) Ἡ μέθοδος ἀπαιτεῖ ὅπως τὸ E_i εἶναι τουλάχιστον 2.

(ιι) Ἡ μέθοδος εἶναι ἐλευθέρα κατανομῆς, ἐλέγχουσα ἀποκλίσεις ὅλων τῶν εἰδῶν καὶ κατὰ συνέπειαν εἶναι ἀνεπηρέαστος ἐξ ἀποκλίσεων εἰδικοῦ τινος τύπου.

4. *Ἔλεγχος καταλληλότητος προσαρμογῆς μιᾶς συναρτήσεως κατανομῆς ἢ διαφορᾶς μεταξὺ συναρτήσεων κατανομῆς*

Ὑπολογίζομεν τὰς διαφορὰς τῶν πραγματικῶν ἀπὸ τῶν θεωρητικῶν παρατηρήσεων καὶ προσδιορίζομεν τὴν τιμήν, ἥτις μεγιστοποιεῖ τὴν διαφορὰν ταύτην ἀσχέτως σημείου. Τὰ σημεῖα 5 % καὶ 1 % διὰ τὴν μεγίστην διαφορὰν d εἶναι ἀντιστοίχως $1.36\sqrt{N}$ καὶ $1.63\sqrt{N}$ ἔνθα N ὁ ἀριθμὸς τῶν παρατηρήσεων.

Παράδειγμα :

i	O_i	E_i	$O_i - E_i$	Ἀθροιστικαὶ Διαφοραὶ $\Sigma(O_i - E_i)$
0	40	50	− 10	−10
1	171	200	− 29	−39
2	305	300	5	− 34
3	222	200	22	− 12
4	62	50	12	0
Σύνολον	800	800	0	0

Ἡ μεγίστη τιμὴ τῶν ἀθροιστικῶν συχνοτήτων (διαφορῶν) εἶναι 39 καὶ τὸ 5 % ἐπίπεδον σημαντικότητος $1.36\sqrt{800} = 38.5$ πλησιάζει ἀκριβῶς τὴν σημαντικότητα.

Παρατηρήσεις :

(ι) Ἡ μέθοδος εἶναι ταχυτέρα τῆς συνήθους καὶ ἐλευθέρα κατανομῆς.

(ιι) Προϋπόθεσις ἐφαρμογῆς τῆς μεθόδου εἶναι ἡ ὕπαρξις τουλάχιστον τεσσαράκοντα παρατηρήσεων καὶ ἀσυνεχὴς κατανομή.

(ιιι) Ἡ μέθοδος εἶναι περισσότερον εὐπαθὴς εἰς μίαν μεταβολὴν μέσου ἐπιπέδου καὶ δύναται νὰ παραμείνῃ ἀνεπηρέαστος εἰς πολλὰς μορφὰς διαφορῶν.

(ιv) Παρόμοιος ἔλεγχος τῆς διαφορᾶς μεταξὺ δύο παρατηρηθεισῶν συναρτήσεων κατανομῆς, τοὐτέστιν σωρευτικὰ ποσοστά, εἶναι ὁ ἑξῆς :

$$\boxed{1.36\sqrt{\left(\frac{N_1 + N_2}{N_1 N_2}\right)}} \quad \text{καὶ} \quad \boxed{1.63\sqrt{\left(\frac{N_1 + N_2}{N_1 N_2}\right)}}$$

Γ΄. Μέθοδοι Συσχετίσεως

1. *Ἐκτίμησις Συνδιακυμάνσεως, Συσχετίσεως καὶ Γραμμικῆς Παλινδρομήσεως*

(α) Ὑπολογίζομεν τὰ σύνολα τῶν μεταβλητῶν ΣΧ καὶ ΣΨ, τὰ ἀθροίσματα τῶν τετραγώνων τῶν αὐτῶν μεταβλητῶν ΣX^2, $\Sigma \Psi^2$ καὶ τὸ ἄθροισμα τῶν γινομένων τῶν Ν ζευγῶν παρατηρήσεων ΣΧΨ.

(β) Ὑπολογίζομεν τά:

$$\boxed{A = \Sigma X^2 - \frac{(\Sigma X)^2}{N}}$$

$$\boxed{B = \Sigma \Psi^2 - \frac{(\Sigma \Psi)^2}{N}}$$

$$\boxed{C = \Sigma X \Psi - \frac{\Sigma X \cdot \Sigma \Psi}{N}}$$

καὶ τὴν Συνδιακύμανσιν τῆς $\boxed{X \text{ καὶ } \Psi = \frac{C}{N-1}}$

(γ) Ὑπολογισμὸς τοῦ Συντελεστοῦ Συσχετίσεως μεταξὺ X καὶ Ψ καὶ ἀναζήτησις ἐκ τοῦ οἰκείου Πίνακος τῶν διδομένων ἐπιπέδων σημαντικότητος:

$$r = C : \sqrt{A \cdot B}$$

(δ) Ὑπολογισμὸς τῆς Παλινδρομήσεως ὡς κάτωθι:

– Τῆς Ψ ἐπὶ τῆς X $\boxed{y - \frac{\Sigma \Psi}{N} = \beta_y \left(x - \frac{\Sigma X}{N} \right)}$ ἔνθα $\boxed{\beta_y = C : A}$

– Τῆς X ἐπὶ τῆς Ψ $\boxed{x - \frac{\Sigma X}{N} = \beta_x \left(y - \frac{\Sigma \Psi}{N} \right)}$ ἔνθα $\boxed{\beta_x = C : B}$

(ε) Ὑπολογισμὸς τῆς Μέσης Τετραγωνικῆς Ἀποκλίσεως τῶν παρατηρήσεων περὶ τὴν Γραμμὴν Παλινδρομήσεως τῶν Ψ ἐπὶ X (Συντελεστὴς Παλινδρομήσεως:

$$\boxed{S = \sqrt{\left[\frac{1}{N-2} \left(B - \frac{C^2}{A} \right) \right]}}$$

ὡς καὶ τοῦ Κανονικοῦ Σφάλματος τοῦ Συντελεστοῦ Παλινδρομήσεως:

$$\boxed{\sigma_{\beta_y} = S : \sqrt{A}}$$

Παράδειγμα :

X	Ψ	X^2	$Ψ^2$	XΨ
31	42	961	1764	1302
26	53	676	2809	1378
1	51	1	2601	51
42	11	1761	121	462
33	43	1089	1849	1419
51	4	2601	16	204
2	48	4	2304	96
19	52	361	2704	988
30	33	900	1089	990
41	32	1681	1024	1312
17	33	289	1089	561
33	47	1089	2209	1551
20	42	400	1764	840
36	19	1296	361	684
48	19	2304	361	912
47	21	2209	441	987
26	41	676	1681	1066
11	50	361	2500	950
46	21	2116	441	966
22	31	484	961	682
35	30	1225	900	1050
40	2	1600	4	80
10	50	100	2500	500
44	30	1936	900	1320
8	42	64	1764	336
45	1	2025	1	45
5	52	25	2704	260
13	57	169	3249	741
ΣX = 790	ΣΨ = 957	$ΣX^2$ = 28406	$ΣΨ^2$ = 40111	ΣXΨ = 21733

Διὰ τὰ ἀνωτέρω 28 ζεύγη παρατηρήσεων θὰ ἔχωμεν : $A = 28406 - 790^2 : 28 = 6117$, $B = 40111 - 957^2 : 28 = 7402$, $C = 21733 - 790 \times 957 : 28 = -5268$, Συνδιακύμανσις τῶν X καὶ Ψ $= -5268 : 27 = 195$, Συντελεστὴς Συσχετίσεως $r = -5268 : \sqrt{6117 \times 7402} = -0.783$, ὅστις εἶναι σημαντικὸς εἰς τὸ ἐπίπεδον $1^0/_0$, Παλινδρόμησις τῆς Ψ ἐπὶ X $y - 957 : 28 = -5268 : 6117 (X - 790 : 28)$ καὶ $y - 34.18 = -0.861 (X - 28.21)$ ἢ $y = 58.47 - 0.861 X$, Παλινδρόμησις τῆς X ἐπὶ Ψ $X - 790 : 28 = -5268 : 7402 (y - 957 : 28)$ καὶ $X - 28.21 = -0.711 (y - 34.17)$ ἤτοι $\chi = 52.504 - 0.711 y$.

Συντελεστὴς Παλινδρομήσεως S

$$S = \sqrt{\left[\frac{1}{26} (7402 - 5268^2 : 6117) \right]} = \pm 10.50$$

$$\sigma_{\beta y} = 10.50 : \sqrt{6117)} \pm 0.134$$

Τὰ ὡς ἀνωτέρω ἀποτελέσματα προσεγγίζουσι τὰ τοιαῦτα διὰ τῆς Μεθόδου product moment.

Παρατήρησις:

Ἡ προκειμένη μέθοδος εἶναι ἀποτελεσματικὴ 100% μόνον διὰ παρατηρήσεις κανονικῶς κατανεμομένας καὶ σταθερᾶς διακυμάνσεως περὶ τὴν εὐθεῖαν γραμμὴν παλινδρομήσεως.

2. *Μέθοδος μονοτονικῆς σχέσεως*

Διατάσσομεν τὰ Ν ζεύγη τῶν παρατηρήσεων συμφώνως πρὸς τὰς τιμὰς τῆς μιᾶς μεταβλητῆς, π.χ. τῆς Χ, καὶ βαθμολογοῦμεν τὰς τιμὰς τῆς ἑτέρας μεταβλητῆς Ψ. Ἴσαι παρατηρήσεις λαμβάνουσι τὸν μέσον ὅρον τῶν οἰκείων βαθμῶν. Μεθ' ὃ ὑπολογίζομεν τὰ σύνολα τῶν πρώτων καὶ τελευταίων p βαθμῶν, δηλ. τὰ R_1 καὶ R_2, ὅπου τὸ p εἶναι περίπου τὰ Ν/3 καὶ ἐλέγχομεν τὴν διαφορὰν $R_1 - R_2$ ὡς κανονικὴν ἀπόκλισιν μὲ Κανονικὸν Σφάλμα (Ν + 1/2) $\sqrt{p : 6}$.

Παράδειγμα:

Χ	Ψ	R_1	Χ	ψ	R	Χ	Ψ	R_2
1	51	6	20	42	13	40	19	24.5
2	48	9	22	31	19	41	32	18
5	52	4.5	26	41	15	42	11	26
8	42	13	26	53	2.5	44	30	20.5
10	50	7.5	30	53	2.5	45	1	28
13	57	1	31	33	16.5	46	21	22.5
17	33	16.5	33	42	13	47	21	22.5
19	50	7.5	33	43	11	48	19	24.5
19	52	4.5	35	47	10	51	4	27
			36	30	20.5			
		69.5						213.5

Ἡ σειρὰ αὐτὴ τῶν 28 ζευγῶν παρατηρήσεων ἔχει διαταχθῆ συμφώνως πρὸς τὰς τιμὰς τῶν Χ. Οἱ βαθμοὶ R οἵτινες ἀντιστοιχοῦσιν εἰς τὰς τιμὰς τῶν

Ψ ἔχουσι καταγραφῆ εἰς τὴν γ στήλην καὶ ἔχουσιν ὑπολογισθῆ τὰ σύνολα τῶν πρώτων καὶ τελευταίων 9 βαθμῶν, ὁπότε ἔχομεν διαφορὰν $R_1 - R_2 = 69.5 - 213.5 = -144$ καὶ Κανονικὸν Σφάλμα $28.5 \sqrt{9:6} = 28.5 \times 1.22 = \pm 34.8$. Ἡ διαφορὰ -144 εἶναι κατὰ συνέπειαν ὑψηλῶς σημαντική. Ὁ Συντελεστὴς Σχέσεως θὰ εἶναι $-144 : 9 \times 19 = -0.842$ ($r = -0.783$).

Παρατηρήσεις:

(α) Ἡ μέθοδος εἶναι ταχυτέρα τῆς συνήθους καὶ ἐλευθέρα κατανομῆς.

(β) Προϋποτίθεται ὅτι τὸ Ν δέον νὰ εἶναι τουλάχιστον 9.

(γ) Ἐπὶ Μεγάλων Δειγμάτων ἔκ τινος διμεταβλητῆς κανονικῆς κατανομῆς ἡ ἀποτελεσματικότης ἀνέρχεται περίπου εἰς 81 %.

(δ) Ὁ Συντελεστὴς Σχέσεως $R_1 - R_2 : p(N - p)$ ὑπερεκτιμᾷ ἐλαφρῶς τὸν ἀντίστοιχον Συντελεστὴν Συσχετίσεως ἐπὶ γραμμικῶν συσχετίσεων.

3. *Ἔλεγχος γενικῆς σχέσεως*

Κατὰ τὴν μέθοδον ταύτην διαιροῦμεν τὰ ζεύγη τῶν παρατηρήσεων εἰς ἰσομεγέθεις περίπου ὁμάδας συμφώνως πρὸς τὰς τιμὰς τῆς μιᾶς τῶν μεταβλητῶν μεθ' ὃ ἐλέγχομεν τὰς διαφορὰς μεταξὺ τῶν Μέσων Τιμῶν τῆς ἑτέρας μεταβλητῆς κατὰ τὰς προεκτεθείσας μεθόδους 16, 17 ἢ 18.

Παράδειγμα:

x	1	2	5	8	10	13	17	19	19	20	22	26	26	30
y	51	48	52	42	50	57	33	50	52	42	31	41	53	33
x	31	33	33	35	36	40	41	42	44	45	46	47	48	51
y	42	43	47	30	19	2	32	11	30	1	21	21	19	4

Ἡ σειρὰ αὐτὴ τῶν 28 ζευγῶν, παρατηρήσεων ἔχει διαιρεθῆ εἰς δύο ὁμάδας βάσει τῶν τιμῶν τῆς Χ, ἐφαρμοζομένης ἐν συνεχείᾳ τῆς μεθόδου 18. Ἡ μικροτέρα τιμὴ τῆς y εἰς τὴν πρώτην ὁμάδα εἶναι ὁ 31 καὶ δέκα τιμαὶ τῆς y εἶναι μικρότεραι αὐτοῦ εἰς τὴν δευτέραν ὁμάδα. Ἡ μεγαλυτέρα τιμὴ τῶν y τῆς δευτέρας ὁμάδος εἶναι ὁ 47 καὶ ὀκτὼ τιμαὶ τῶν y εἶναι μεγαλύτεραι αὐτοῦ εἰς τὴν πρώτην ὁμάδα. Ἄρα τὸ σύνολον $10 + 8 = 18$ εἶναι σαφῶς σημαντικὸν εἰς τὸ ἐπίπεδον 1 %.

Παρατηρήσεις:

(α) Ἡ μέθοδος στηρίζεται ἐπὶ τοῦ γεγονότος ὅτι γειτνιάζουσαι παρατηρήσεις θὰ εἶναι περισσότερον ὅμοιαι ἢ αἱ ἀπομεμακρυσμέναι τοιαῦται. Τὸ καλλίτερον μέγεθος ὁμάδος, θὰ ἐξαρτηθῇ ἐκ τοῦ κατὰ πόσον ἡ κυρτότης, οἱασδήποτε σχέσεως ἐπηρεάζει τὴν ὑπόθεσιν αὐτήν. Ὁπωσδήποτε ἐὰν τὸ μέγεθος τῆς ὁμάδος ἔχει ἐκλεγῆ διὰ θεωρήσεως τῶν παρατηρήσεων ἡ σημαντικότης θὰ τείνῃ νὰ ὑπερεκτιμηθῇ.

(β) Ἡ ἀποτελεσματικότης τῆς μεθόδου θὰ κυμαίνηται ἀναλόγως τοῦ ὑφισταμένου τύπου τῆς σχέσεως καὶ τοῦ χρησιμοποιηθέντος ἐλέγχου.

3. 116

4. *Ἔλεγχος μονοτονικῆς σχέσεως*

Διατάσσομεν τὰ N ζεύγη παρατηρήσεων συμφώνως πρὸς τὰς τιμὰς τῆς μιᾶς μεταβλητῆς π.χ. τῆς X, συγκρίνομεν τὰς πρώτας καὶ τελευταίας p τιμὰς τῆς ἑτέρας μεταβλητῆς y, ὅπου τὸ p ἰσοῦται κατὰ προσέγγισιν πρὸς N/3 καὶ σημειοῦμεν $+1$ ἐὰν ἡ τιμὴ $(N-p+i)$ ὑπερβαίνῃ τὴν τιμὴν i καὶ -1 ἀντιθέτως, τέλος ἐλέγχομεν τὸ σύνολον T τῶν p βαθμῶν ὡς κανονικὴν ἀπόκλισιν μὲ Μέσην Τετραγωνικὴν Ἀπόκλισιν ἴσην πρὸς $\sqrt{p}$.

Ἡ μέθοδος ἀπαιτεῖ τὴν κανονικὴν προσέγγισιν τῆς Διωνυμικῆς Κατανομῆς.

Ἐπὶ μικροῦ p χρησιμοποιοῦνται τὰ ἐπίπεδα σημαντικότητος διὰ T βάσει τοῦ ἀκολούθου Πίνακος :

p	6	7	8	9	10	11	12	13	14	15	16
5 %	± 6	7	8	7	8	9	8	9	10	9	10
1 %	–	–	± 8	9	1	11	10	11	12	11	12

p	17	18	19	20	21	22	23	24	25	26
5 %	± 9	10	11	10	11	12	11	12	11	12
1 %	± 13	12	13	14	13	14	15	14	15	14

Παράδειγμα :

x	y	x	y	x	y
1	51	20	42	40	19
2	48	22	31	41	32
5	52	26	41	42	11
8	42	26	53	44	30
10	50	30	53	45	1
13	57	31	33	46	21
17	33	33	42	47	21
19	50	33	43	48	19
19	52	35	47	51	4
		36	30		

Ἡ σειρὰ αὐτὴ τῶν 28 ζευγῶν παρατηρήσεων διετάχθη συμφώνως πρὸς τὰς τιμὰς τῆς X καὶ κατόπιν διηρέθη εἰς τρεῖς ὁμάδας ἐξ 9, 10 καὶ 9 ζευγῶν παρατηρήσεων. Μία θεώρησις ἁπλῆ τοῦ Πίνακος δεικνύει ὅτι ἑκάστη τιμὴ τῶν Ψ εἰς τὸ πρῶτον ζεῦγος στηλῶν ὑπερβαίνει τὴν ἀντίστοιχον τιμὴν τῶν Ψ εἰς τὸ τελευταῖον ζεῦγος στηλῶν. Τὸ συνολικὸν σκὸρ T εἶναι κατὰ συνέπειαν -9 καὶ τοῦτο διὰ $p=9$ εἶναι σημαντικὸν εἰς τὸ ἐπίπεδον 1 %.

Παρατήρησις :

Ἐπὶ Μεγάλων Δειγμάτων ἔκ τινος διμεταβλητῆς κανονικῆς κατανομῆς ἡ μέθοδος εἶναι ἀποτελεσματικὴ περίπου κατὰ 50 %.

5. *Ἐκτίμησις συντελεστῶν γραμμικῆς παλινδρομήσεως*

Διευθετοῦμεν τὰ Ν ζεύγη παρατηρήσεων συμφώνως πρὸς τὰς τιμὰς τῆς Ἀνεξαρτήτου Μεταβλητῆς Χ καὶ διαιροῦμεν ταῦτα εἰς τρεῖς ἰσομεγέθεις κατὰ τὸ δυνατὸν ὁμάδας ὑπὸ τὸν ὅρον ὅμως ὅτι αἱ πρώτη καὶ τελευταία θὰ εἶναι ἀπολύτως ἴσαι. Ὑπολογίζομεν τὰ Σύνολα τῶν Ἀνεξαρτήτων καὶ Ἐξηρτημένων Μεταβλητῶν ἑκάστης ὁμάδος καὶ τέλος ἐκτιμῶμεν τὸν Συντελεστὴν Παλινδρομήσεως καὶ Κανονικὸν Σφάλμα τούτου ὡς κατωτέρω.

Α. Ἐπὶ 3 ὁμάδων :

$$\boxed{\beta = \frac{\Psi_3 - \Psi_1}{X_3 - X_1}} \quad \text{καὶ} \quad \boxed{\sigma_\beta = \frac{0\cdot 8S\sqrt{N}}{X_1 - X_3}}$$

Β. Ἐπὶ 4 ὁμάδων : (Αὗται περιέχουσι τὰ 1/6, 1/3, 1/3 καὶ 1/6 τῶν παρατηρήσεων) :

$$\boxed{\beta = \frac{3\Psi_1 + \Psi_2 - \Psi_3 - 3\Psi_4}{3X_1 + X_2 - X_3 - 3X_4}}$$

$$\boxed{\sigma_\beta = \frac{1.9\, S\sqrt{N}}{3X_1 + X_2 - X_3 - 3X_4}}$$

Ἡ Μέση Τετραγωνικὴ Ἀπόκλισις ἐπ' ἀμφοτέρων τῶν περιπτώσεων θὰ ἰσοῦται πρὸς τὰ 8/9 τῆς μέσης διαφορᾶς μεταξὺ τῶν γειτνιαζουσῶν τιμῶν τῆς Ἐξηρτημένης Μεταβλητῆς Ψ.

Παράδειγμα Τριῶν Ὁμάδων :

x_1	y_1		x_2	y_2		x_3	y_3	
1	51		20	42	10	40	2	17
2	48	3	22	31		41	32	
5	52		26	41	10	42	11	21
8	42	10	26	53		44	30	
10	50		30	33	20	45	1	29
13	57	7	31	42		46	21	
17	33		33	43	1	47	21	0
19	50	17	33	47		48	19	
19	52		35	30	17	51	4	15
			36	19				
Σύνολα 94	435		292	381		404	141	

Ἔστω ἡ ἀνωτέρω σειρὰ 28 ζευγῶν παρατηρήσεων, κατατμηθεῖσα εἰς 3 ὁμάδας ἐξ ἐννέα, δέκα καὶ ἐννέα παρατηρήσεων:

$$\beta = \frac{141 - 435}{404 - 94} = \frac{-294}{310} = -0.948$$

$$S\frac{8}{9}\left[3+10+7+17+10+10+20+1+17+17+21+29+15\right] : 13$$

$$= 12.10 \text{ καὶ } \sigma_\beta = \frac{0.8 \times 12.10 \sqrt{28}}{310} = 0.164$$

Παράδειγμα Τεσσάρων Ὁμάδων:

x_1	y_1		x_2	y_2		x_3	y_3		x_4	y_4	
1	51		13	57	7	31	42		45	1	29
2	48	3	17	33		33	43	1	46	21	
5	52		19	50	17	33	47		47	21	0
8	42	10	19	52		35	30	17	48	19	
10	50		20	42	10	36	19		51	4	15
			22	31		40	2	17			
			26	41	10	41	32				
			26	53		42	11	21			
			30	33	20	44	30				
26	243		192	392		335	256		237	66	

$$\beta = \frac{729 + 392 - 256 - 198}{78 + 192 - 335 - 711} = \frac{-667}{-776} = -0.859$$

$$\sigma_\beta = \frac{1.90 \times 12.10 \sqrt{28}}{78 + 192 - 335 - 711} = \frac{22.99 \times 5.28}{-776} = 0.156$$

Παρατηρήσεις:

(α) Ἡ μέθοδος τῶν 3 ὁμάδων εἶναι ἀποτελεσματικὴ 79 % διὰ παρατηρήσεις μιᾶς διμεταβλητῆς κανονικῆς κατανομῆς.

(β) Ἡ μέθοδος τῶν 4 ὁμάδων ἔχει ἀποτελεσματικότητα 88 %.

6. *Γραφικαὶ Μέθοδοι Συσχετίσεως*

Αἱ μέθοδοι αὐταὶ λόγῳ τῆς πενιχρᾶς ἀποτελεσματικότητός των δὲν δύνανται νὰ ληφθῶσιν σοβαρῶς ὑπ' ὄψιν εἰς μίαν στατιστικὴν ἔρευναν ἐφ' ὅσον ὁ γνωστὸς πρόχειρος ἔλεγχος τοῦ Διαγράμματος Διασπορᾶς παρέχει ἡμῖν ἔνδειξιν τινὰ ἐπὶ τῆς ὑπάρξεως ἢ μὴ συσχετίσεως ὡς καὶ τῆς φορᾶς ταύτης.

III. ΣΥΣΤΗΜΑΤΙΚΟΣ ΚΑΙ ΤΑΧΥΣ ΤΡΟΠΟΣ ΥΠΟΛΟΓΙΣΜΟΥ ΚΕΝΤΡΙΚΗΣ ΤΑΣΕΩΣ, ΔΙΑΣΠΟΡΑΣ, ΑΣΥΜΜΕΤΡΙΑΣ, ΚΥΡΤΩΣΕΩΣ

Ἐκ τῆς ὀκταετοῦς πείρας τοῦ ὑπογράφοντος κατὰ τὴν διδασκαλίαν τῆς Στατιστικῆς ἀλλὰ καὶ τῆς παραλλήλου ἐντατικῆς περὶ τὴν θεωρίαν ἐνασχολήσεως διεπιστώθη ἡ ἀνάγκη ὅπως Σπουδασταί τε καὶ Ἐρευνηταὶ ἐπιταχύνωσι τοὺς ὑπολογισμούς των κατὰ τρόπον τόσον ἀκριβῆ ὅσον καὶ μεθοδικόν. Ὁ κατωτέρω συνοπτικὸς Πίναξ Κατανομῆς ἱκανοποιεῖ τὴν μνημονευθεῖσαν ἀνάγκην ἀπολύτως. Ἐπισημαίνεται ὅτι οἱ $\overline{X}$, M_e καὶ M_o ὑπολογίζονται εὐκόλως ἐὰν θεωρήσωμεν τὴν τάξιν ἐκείνην ἥτις ἐμφανίζει τὴν μεγίστην συχνότητα, καὶ ἡ ὁποία καθ' ὑπόθεσιν περιλαμβάνει τὰς αἰτουμένας παραμέτρους ἐνῷ ἐκ παραλλήλου ὡς πρὸς τὰ Τεταρτημόρια θὰ θεωρήσωμεν τὴν στήλην (8) πρὸς ἐντοπισμὸν τοῦ αἰτουμένου Τεταρτημορίου. Ἡ εἰς τὸν ἀκολουθοῦντα πίνακα Κατανομὴ παριστᾷ τὰς ὑποθετικὰς ὥρας πτήσεως 220 Χειριστῶν Ἀεριωθουμένων Ἀεροσκαφῶν.

Φάσις Α' : Κατάστρωσις τοῦ Πίνακος Κατανομῆς ὡς ἕπεται :

Τάξεις ὡρῶν πτήσεως (1)	Συχνότης Χειρισταὶ f (2)	Ἀποκλίσεις διαστημάτων τάξεως d' (3)	Ροπαὶ περὶ τὴν αὐθαίρετον ἀρχὴν f(d') (4)	$f(d'^2)$ (5)	$f(d'^3)$ (6)	$f(d'^4)$ (7)	Ἀθροιστικαὶ Συχνότητες Σi (8)	Μέσα σημεῖα τάξεως Μ.Ρ. (9)
100 – 199	5	— 5	— 25	125	−625	3125	5	150
200 – 299	10	— 4	— 40	160	−640	2560	15	250
300 – 399	20	— 3	— 60	180	−540	1620	35	350
400 – 499	25	— 2	— 50	100	−200	400	60	450
500 – 599	32	— 1	— 32	32	— 32	32	92	550
600 – 699	38	0	0	0	0	0	130	650
700 – 799	31	+ 1	31	31	31	31	161	750
800 – 899	25	+ 2	50	100	200	400	186	850
900 – 999	18	+ 3	54	162	486	1458	204	950
1000 – 1099	12	+ 4	48	192	768	3072	216	1050
1100 – 1199	4	+ 5	20	100	500	2500	220	1150
	N=220		$\Sigma f(d) = -4$	$\Sigma f(d'^2) = 1182$	$\Sigma f(d'^3) = -52$	$\Sigma f(d'^4) = 15198$		

Φάσις Β' : Καταγραφή τῶν Δεδομένων :

$C = 100$ $\qquad \frac{N}{4} = 55$ $\qquad \nu_1^3 = -0{,}000005832$

$\overline{Z} = 650$ $\qquad L_{Q_1} = 400$ $\qquad \nu_1^4 = 0{,}000000105$

$\left.\begin{matrix} L_{me} \\ L_{mo} \end{matrix}\right\} = 600$ $\qquad \frac{3N}{4} = 165$ $\qquad \nu_2 = \frac{1182}{220} = 5.37$

$\frac{N}{2} = 110$ $\qquad L_{Q_3} = 800$ $\qquad \nu_1\nu_2 = -0{,}0967$

$d_1 = 6$ $\qquad \nu_1 = \frac{-4}{220} = -0{,}018$ $\qquad \nu_3 = \frac{-52}{220} = -0{,}236$

$d_2 = 7$ $\qquad \nu_1^2 = 0{,}00324$ $\qquad \nu_1\nu_3 = 0{,}00425$

$\frac{1}{12} = 0{,}083,$ $\qquad \frac{7}{240} = 0{,}029,$ $\qquad \nu_1^2\,\nu_2 = 0{,}01740,$ $\qquad \nu_4 = \frac{15198}{220} = 69.08$

$3(\nu_1\,\nu_2) = -0{,}2901,$ $\qquad 2(\nu_1^3) = -0{,}0000117,$ $\qquad 4(\nu_1\,\nu_3) = 0{,}017$

$6(\nu_1^2\nu_2) = 0{,}10440,$ $\qquad 3(\nu_1^4) = 0{,}000000315,$ $\qquad \sqrt{N} = 14{,}83$

Φάσις Γ' : Ὑπολογισμὸς τῶν Παραμέτρων ὡς ἑξῆς :

— Ἀριθμητικὸς Μέσος $\qquad \overline{X} = \overline{Z} + (\nu_1 \cdot c) = 648.20$

— Διάμεσος $\qquad M_e = L_{me} + \left(\frac{\frac{N}{2} - \Sigma_i}{f} \cdot C \right) = 647.37$

— Ἐπικρατοῦσα Τιμὴ $\qquad M_0 = L_{m0} + \left(\frac{d_1}{d_1 + d_2} \cdot C \right) = 646.15$

— Πρῶτον Τεταρτημόριον $\qquad Q_1 = L_{Q_1} + \left(\frac{\frac{N}{4} - \Sigma_i}{f} \cdot C \right) = 480$

— Τρίτον Τεταρτημόριον $\qquad Q_3 = L_{Q_3} + \left(\frac{\frac{3N}{4} - \Sigma_i}{f} \cdot C \right) = 816$

– Μέση Τετραγ. Ἀπόκλισις $\sigma = C\sqrt{\nu_2 - \nu_1^2} = 232$

– Συντελεστὴς Μεταβλητικότητος $V = \dfrac{\sigma \cdot 100}{\bar{X}} = 35.79\,\%$

– Κανονικὸν Σφάλμα $\bar{X}$ $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{N}} = 15.64$

– Ὅρια Διακυμάνσεως Ἀληθοῦς Μέσης Τιμῆς εἰς ἐπίπεδον 95 %

$$\bar{X} - 1{,}96 \cdot \sigma_{\bar{X}} < \mu < \bar{X} + 1{,}96 \cdot \sigma_{\bar{X}}$$

$$617.75 < \mu < 678.85$$

– Δευτέρα Ροπὴ περὶ τὸν $\bar{X}$ διορθωμένη κατὰ Sheppard

$$\bar{\mu}_2 = \nu_2 - \nu_1^2 - \frac{1}{12} = 5.28$$

– Τρίτη Ροπὴ περὶ τὸν $\bar{X}$

$$\mu_3 = \nu_3 - 3\,(\nu_1\,\nu_2) + 2\,(\nu_1^3) = 0{,}0541$$

– Τετάρτη Ροπὴ περὶ τὸν $\bar{X}$ διορθωμένη κατὰ τὸν Sheppard

$$\bar{\mu} = \nu_4 - 4(\nu_1\nu_3) + 6(\nu_1^2\nu_2) - 3(\nu_1^4) - \frac{\bar{\mu}_2}{2} + \frac{7}{240} = 66.55$$

– Συντελεστὴς Ἀσυμμετρίας $\beta_1 = \dfrac{\mu_3^2}{\bar{\mu}_2^3} = \dfrac{0{,}0029}{147.20} = 0{,}002$

– Συντελεστὴς Κυρτώσεως $\beta_2 = \dfrac{\bar{\mu}_4}{\bar{\mu}_2^2} + 3 = \dfrac{66.55}{27.88} + 3 = 5.38$

Φάσις Δ' : Γραφικὴ Ἀπεικόνισις τῆς Κατανομῆς μὲ τὰ Μέσα Σημεῖα Τάξεων ἐπὶ τῆς Τετμημένης.

Πόρισμα : Ἡ ὑπ' ὄψιν Κατανομὴ εἶναι Κανονική, Συμμετρικὴ καὶ Λεπτόκυρτος.

Ι. Α. ΣΑΚΑΛΗΣ