ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΟΥ ΜΕΓΕΘΟΥΣ ΤΟΥ ΔΕΙΓΜΑΤΟΣ ΤΟΥ ΑΠΑΙΤΟΥΜΕΝΟΥ ΔΙΑ ΤΗΝ ΕΚΤΙΜΗΣΙΝ ΕΝΟΣ ΑΡΙΘΜΗΤΙΚΟΥ ΜΕΣΟΥ ΟΡΟΥ ΜΕΤΑ ΠΡΟΚΑΘΩΡΙΣΜΕΝΟΥ ΒΑΘΜΟΥ ΑΚΡΙΒΕΙΑΣ ΕΙΣ ΟΡΟΥΣ ΠΙΘΑΝΟΤΗΤΟΣ

Ὑπὸ Ε. Β. ΓΕΩΡΓΟΥΛΗ

Δασολόγου παρὰ τῷ Ἰνστιτούτῳ Δασικῶν Ἐρευνῶν τοῦ Ὑπ. Γεωργίας

Ἓν κοινότατον πρόβλημα τὸ ὁποῖον συχνάκις ἀντιμετωπίζει ὁ δασολόγος, εἶναι καὶ ἐκεῖνο κατὰ τὸ ὁποῖον οὗτος διερωτᾶται, πόσας παρατηρήσεις πρέπει νὰ κάμῃ διὰ νὰ ἐκτιμήσῃ τὸν ἀριθμητικὸν μέσον ὅρον ἑνὸς στατιστικοῦ πλήθους ἢ μεγέθους. Τὸ πρόβλημα τοῦτο κέκτηται ἰδιαιτέραν σημασίαν, διότι ἐὰν μὲν γίνουν περισσότεραι τοῦ δέοντος παρατηρήσεις, τότε ὑπάρχει σπατάλη χρήματος, ἐὰν δὲ ὀλιγώτεραι, τότε τὸ σχετικὸν συμπέρασμα δὲν εἶναι ἐπαρκῶς ἀξιόπιστον.

Παραδείγματα τοιούτου προβλήματος θὰ ἠδύναντο νὰ εἶναι καὶ τὰ ἀκόλουθα:

1. Πόσα δένδρα, στηθιαίας διαμέτρου ἀπὸ 0,08 μ. καὶ ἄνω, πρέπει νὰ παχυμετρήσῃ εἷς δασολόγος, διαχειριστὴς εἰς μίαν κηπευτὴν συστάδα ἐλάτης, διὰ νὰ ἐκτιμήσῃ τὸν ἀριθμητικὸν μέσον ὅρον τῶν στηθιαίων διαμέτρων ὅλων τῶν ἀπὸ τῆς τοιαύτης διαμέτρου 0,08 μ. καὶ ἀνωτέρω δένδρων ὁλοκλήρου τῆς συστάδος;

2. Πόσας παρατηρήσεις πρέπει νὰ κάμῃ ὁ δασολόγος ἑνὸς πριστηρίου, διὰ νὰ ἐκτιμήσῃ, κατὰ μέσον ὅρον, τὸ μέγεθος τῆς συστολῆς τὴν ὁποίαν ὑφίστανται τὰ τεμάχια τῆς πριστῆς ξυλείας μιᾶς ὡρισμένης κατηγορίας (μετὰ τὴν πρίσιν των) συνεπείᾳ ἀπωλείας ὕδατος, διὰ νὰ δώσῃ εἰς αὐτὰ προηγουμένως μίαν ἀνάλογον ἀνοχὴν (ὑπὲρ-διάστασιν);

3. Πόσας δοκιμαστικὰς ἐπιφανείας, ἐκτάσεως ἑκάστην ἑνὸς τετραγωνικοῦ μέτρου (1,00 μ. × 1,00 μ. = 1,0000 $μ^2$), πρέπει νὰ λάβῃ εἷς δασολόγος εἰς μίαν δασικὴν συστάδα διὰ νὰ ἐκτιμήσῃ τὸν ἀνὰ μίαν τοιαύτην μονάδα ἐπιφανείας ἀριθμητικὸν μέσον ὅρον τῶν φυομένων ἐκεῖ ἀρτιφύτρων καὶ φυταρίων;

Εἰς τὴν ἰδίαν κατηγορίαν προβλημάτων θὰ ἠδύνατο νὰ ὑπαχθῇ, ὑπὸ τὴν γενικήν του μορφήν, καὶ ἐκεῖνο καθ' ὃ διερωτώμεθα πόσας δοκιμαστικὰς ἐπιφανείας, ἐκτάσεως ἑκάστην ἑνὸς στρέμματος, πρέπει νὰ λάβωμεν εἰς μίαν δασικὴν συστάδα διὰ νὰ ἐκτιμήσωμεν τὸν κατὰ μέσον ὅρον καὶ στρέμμα ὑπάρχοντα ἐκεῖ ξυλώδη ὄγκον;

Δίδομεν κατωτέρω μίαν περιωρισμένην λύσιν τοῦ προβλήματος, ἀποφεύγοντες τὴν εὐρεῖαν θεωρητικὴν ἀνάπτυξιν αὐτοῦ, διότι αὕτη δὲν εἶναι δυνατὴ ἀπὸ τῶν στηλῶν τούτων.

Ἐὰν ἀρχίσωμεν μίαν ἀπεριόριστον σειρὰν τυχαίων ἰσοπληθῶν δειγματοληψιῶν ἀπό τινος στατιστικοῦ πληθυσμοῦ, θὰ παρατηρήσωμεν κατὰ τὴν διάρκειαν ταύτης, ὅτι οἱ ἀπὸ τούτων λαμβανόμενοι ἀντίστοιχοι ἀριθμητικοὶ μέσοι ὅροι διαφέρουν κατὰ τὸ μᾶλλον ἢ ἧττον ἀλλήλων, χωρὶς ἐκ τούτου ν' ἀποκλείεται, μολονότι σπανιώτατα, ὅπως δύο ἢ καὶ περισσότεροι ἐξ αὐτῶν συμπίπτουν εἰς τὴν αὐτὴν τιμήν. Ἐφ' ὅσον προχωρεῖ ἡ δειγματοληψία θὰ διαπιστώσωμεν ἀκόμη ὅτι οἱ οἰκεῖοι ἀριθμητικοὶ μέσοι ὅροι δὲν διασπείρονται εἰκῇ καὶ ὡς ἔτυχεν ἀπὸ ἀπόψεως μεγέθους, ἀλλὰ κατὰ τὴν διασποράν των ταύτην ἀκολουθοῦν ὡρισμένον νόμον, τὸν γνωστὸν ὑπὸ τὸ ὄνομα κανονικὸν νόμον, ἤτοι κατανέμονται συμμέτρως ἑκατέρωθεν μιᾶς μέσης-κεντρικῆς τιμῆς, ἡ ὁποία πάντως δὲν προσδιορίζεται ἐπακριβῶς διὰ τῶν δειγματοληψιῶν. Ἀποδεικνύεται ὅτι ἡ στατιστικὴ κανονικὴ κατανομὴ συχνοτήτων τῶν μέσων τούτων ἔχει ἀριθμητικὸν μέσον ὅρον*, τὸν ἀριθμητικὸν μέσον ὅρον μ τοῦ ὡς ἄνω στατιστικοῦ πληθυσμοῦ καὶ μέσην ἀπόκλισιν τετραγώνου* τὴν ποσότητα $\frac{\sigma}{\sqrt{N}}$, ὅπου σ ἡ μέση ἀπόκλισις τετραγώνου τῶν μονάδων τοῦ ἰδίου πληθυσμοῦ ἀπὸ τοῦ ἀριθμητικοῦ μέσου ὅρου των καὶ Ν ὁ ἀριθμὸς τῶν μονάδων ἑκάστου ἐκ τῶν ἰσοπληθῶν, ἀπεριορίστων εἰς ἀριθμόν, δειγμάτων. Βασικὴ προϋπόθεσις τῆς λειτουργίας τοῦ νόμου τούτου εἶναι ὅτι ὁ στατιστικὸς πληθυσμὸς εἶναι ἄπειρος, ἢ τοὐλάχιστον εἰς τὰς περιπτώσεις τῆς πράξεως, ἐπαρκῶς μέγας, ὅπως εἶναι, ἐπὶ παραδείγματι, ὁ ἀριθμὸς τῶν δένδρων μιᾶς ἱκανῆς εἰς ἔκτασιν καὶ πυκνότητα δασικῆς συστάδος καὶ ὅτι τὸ μέγεθος τοῦ δείγματος εἶναι μικρὸν ποσοστὸν τοῦ μεγέθους τοῦ πληθυσμοῦ. Ὁ νόμος οὗτος ἰσχύει ἀκόμη καὶ ὅταν αἱ μονάδες, εἴτε τοῦ πληθυσμοῦ, εἴτε τοῦ δείγματος, εἴτε καὶ ἀμφοτέρων δὲν κατανέμονται κανονικῶς, ἀρκεῖ τὰ δείγματα νὰ εἶναι ἀπεριόριστα εἰς ἀριθμόν, τυχαῖα καὶ πολυάριθμα. Κατ' ἀρχὴν λέγομεν ὅτι ἓν δεῖγμα εἶναι τυχαῖον, ὅταν αἱ μονάδες τοῦ οἰκείου πληθυσμοῦ ἔχουν ἴσην καὶ ἀνεξάρτητον πιθανότητα νὰ συμπεριληφθοῦν εἰς τὸ δεῖγμα, εἶναι δὲ πολυάριθμον ὅταν περιλαμβάνῃ ἄνω τῶν 30 μονάδων. Εἰς τὰς βιολογικὰς ἐπιστήμας καί, ἑπομένως, εἰς τὴν δασολογίαν, ὅταν ἰδίᾳ πρόκειται περὶ δένδρων, χρησιμοποιοῦνται κατὰ κανόνα πολυάριθμα δείγματα.

Ἐλέχθη ἀνωτέρω ὅτι ἡ κατανομὴ τῶν μέσων τῶν δειγμάτων εἶναι κανονική. Ἀποδεικνύεται ὅμως ὅτι εἰς μίαν κανονικὴν κατανομήν, τὰ 68 % τῶν παρατηρήσεων, κεῖνται ἀπὸ ἀπόψεως μεγέθους, ἑκατέρωθεν τοῦ ἀριθμητικοῦ μέσου ὅρου καὶ εἰς ἀπόστασιν μιᾶς μονάδος μέσης ἀποκλίσεως τετραγώνου, δηλ. εἰς τὸ διάστημα τὸ ὁριζόμενον ὑπὸ τῶν ἀριθμῶν $\bar{x} - 1{,}00\sigma$ καὶ $\bar{x} + 1{,}00\sigma$, ὅπου $\bar{x}$ ὁ μέσος τοῦ δείγματος ἢ τοῦ πληθυσμοῦ. Τὰ 95 % τῶν περιπτώσεων κεῖνται εἰς τὸ διάστημα $\bar{x} - 1{,}96\sigma$ καὶ $\bar{x} + 1{,}96$. Τέλος τὰ 99 % τῶν περιπτώσεων κεῖνται εἰς τὸ διάστημα μεταξὺ $\bar{x} - 2{,}58\sigma$ καὶ $\bar{x} + 2{,}58\sigma$.

Παράδειγμα:

Εἰς μίαν κηπευτὴν συστάδα ἐλάτης, τὰ ὕψη τῶν ἄνω τῶν 0,08 μ. στηθιαίας

* Περὶ τῆς ἐννοίας καὶ τοῦ τρόπου τοῦ ὑπολογισμοῦ τῶν παραμέτρων τούτων, ἰδὲ βοηθητικὸν σημείωμα εἰς τὸ τέλος τοῦ παρόντος.

διαμέτρου δένδρων, κατανέμονται κανονικῶς μὲ μέσον ὅρον $\bar{x} = 20{,}50$ μ. καὶ μέσην ἀπόκλισιν τετραγώνου $\sigma = 3{,}20$ μ. Τοῦτο σημαίνει ὅτι, ἐκ τοῦ συνολικοῦ ἀριθμοῦ τῶν ὡς ἄνω δένδρων τῆς συστάδος τὰ 68 % ἔχουν ὕψος περιλαμβανόμενον μεταξὺ τῶν μεγεθῶν :

20,50 μ. — 1,00 × 3,20 μ. = 17,30 μ. καὶ 20,50 μ. + 1,00 × 3,20 μ. = 23,70 μ.

Τὰ 95 % τῶν δένδρων θὰ ἔχουν ὕψος κυμαινόμενον μεταξὺ τῶν ἀριθμῶν :

20,50 μ. — 1,96 × 3,20 = 14,20 μ. καὶ 20,50 μ. + 1,96 × 3,20 = 26,80 μ. κ.ο.κ.

Κάτι ἀνάλογον πρέπει νὰ σκεφθῇ ὁ ἀναγνώστης καὶ προκειμένου περὶ τῶν μέσων τῶν δειγμάτων μιᾶς σειρᾶς δειγματοληψιῶν, ὡς προεξετέθη. Τὰ 68 % ἐκ τῶν μέσων τούτων δὲν ἀπέχουν ἐκ τοῦ ἀληθοῦς μέσου τοῦ πληθυσμοῦ, τὸν ὁποῖον, πάντως δὲν γνωρίζομεν, περισσότερον ἢ ὀλιγώτερον τῆς ποσότητος 1,00σ, ἔνθα σ ἡ μέση ἀπόκλισις τετραγώνου τοῦ πληθυσμοῦ. Τὰ 95 % ἐκ τῶν ἰδίων μέσων δὲν θ' ἀπέχουν ἐκ τοῦ ἰδίου μ περισσότερον ἢ ὀλιγώτερον τῆς ποσότητος 1,96 σ κ.ο.κ.

Ἐπανερχόμενοι ἐπὶ τῶν ἀνωτέρω λεχθέντων, περὶ κανονικῆς κατανομῆς τῶν μέσων τῶν δειγμάτων μιᾶς ἀπεριορίστου σειρᾶς τυχαίων ἰσοπληθῶν δειγματοληψιῶν, μὲ μέσον τὸν μ καὶ μέσην ἀπόκλισιν τετραγώνου τὴν ποσότητα $\frac{\sigma}{\sqrt{N}}$, ὡς τὰ μεγέθη ταῦτα προσδιωρίσθησαν ἀνωτέρω, δυνάμεθα προφανῶς νὰ συναγάγωμεν, ὅτι ἡ μέση ἀπόκλισις τετραγώνου $\frac{\sigma}{\sqrt{N}}$ τῆς στατιστικῆς κατανομῆς τῶν μέσων τῶν δειγμάτων εἶναι συγχρόνως καὶ μέσον σφάλμα τετραγώνου ἐκτιμήσεως τοῦ ἀριθμητικοῦ μέσου ὅρου ἑνὸς καὶ μόνον ἑνὸς τυχόντος δείγματος ἐκ τῶν ἀνωτέρω. Ἔχομεν ὅθεν:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \qquad (1)$$

Ἐκ τούτου λαμβάνομεν:

$$N = \left(\frac{\sigma}{\sigma_{\bar{x}}}\right)^2 \qquad (1\alpha)$$

εἰς ὅρους πιθανότητος 0,68.

Ἐκ τοῦ τύπου τούτου συνάγεται ὅτι δύο παράγοντες ἐπηρεάζουν τὸ μέγεθος τοῦ δείγματος, ἡ μέση ἀπόκλισις τετραγώνου τῶν μονάδων τοῦ διερευνωμένου στατιστικοῦ πληθυσμοῦ, ἤτοι τὸ σ καὶ τὸ μέγεθος τοῦ σφάλματος, ἤτοι τὸ $\sigma_{\bar{x}}$. Ὅσον μεγαλυτέρα εἶναι ἡ διασπορὰ τῶν τιμῶν τῶν μονάδων τοῦ πληθυσμοῦ περὶ τὸν ἀριθμητικὸν μέσον ὅρον των, ἤτοι ὅσον μεγαλυτέρα εἶναι ἡ ποικιλομορφία αὐτῶν, τόσον μεγαλύτερον πρέπει νὰ εἶναι τὸ δεῖγμα.

Ἀντιθέτως, ὅσον τὸ ἀνεκτὸν ὅριον τοῦ σφάλματος εἶναι μεγαλύτερον, ἤτοι ἡ ἐπιδιωκομένη προσέγγισις μικροτέρα, τόσον τὸ μέγεθος τοῦ δείγματος δέον νὰ εἶναι μικρότερον. Καὶ τὸ μὲν σ τοῦ πληθυσμοῦ δὲν δυνάμεθα βεβαίως νὰ μεταβάλωμεν, τὸ μέγεθος ὅμως τοῦ σφάλματος δυνάμεθα ἡμεῖς νὰ ἐκλέξωμεν a priori, προαιρετικῶς ἀναλόγως τοῦ ἐπιδιωκομένου σκοποῦ, τῆς ἐπιθυμουμένης ἀκριβείας καὶ τῶν διατεθειμένων οἰκονομικῶν μέσων.

Εἰς τὸ σημεῖον τοῦτο ὁ ἀναγνώστης πρέπει νὰ γνωρίζῃ καλῶς, ὅτι ὅταν ἐργαζώμεθα ἐπὶ τῇ βάσει δειγμάτων, δὲν εἶναι δυνατὸν νὰ διατυποῦμεν τὰ ἐκ τούτων συμπεράσματα εἰς ὅρους βεβαιότητος. Τοιαῦτα συμπεράσματα διατυποῦμεν πάντοτε εἰς ὅρους πιθανότητος. Ὁ ἄνθρωπος πρέπει νὰ ἀρκεῖται ἐνίοτε εἰς πιθανότητας, τουλάχιστον κατὰ τὸ παρὸν στάδιον τῶν γνώσεών του, ὅπως ἄλλως τε τοῦτο κάμνει εἰς τὸν καθημερινόν του βίον, ὄχι σπανίως καὶ χωρὶς ἴσως νὰ τὸ ὑποπτεύηται. Ἐπὶ παραδείγματι, ὅταν λέγωμεν: «Θὰ συναντηθῶμεν εἰς Θεσσαλονίκην τὴν προσεχῆ Δευτέραν», δὲν ἐκφραζόμεθα εἰμὴ εἰς ὅρους πιθανότητος, ἀφοῦ ἡ συνάντησις αὕτη δὲν ἐξαρτᾶται ἀπὸ ἡμᾶς, ἀλλ' ἀπὸ πλῆθος παραγόντων, ὡς εἶναι ἡ ὑγεία ἡμῶν, ἡ ὑγεία τῶν οἰκείων μας, ἡ μηχανικὴ κατάστασις τῶν μεταφορικῶν μέσων, οἱ μετεωρολογικοὶ παράγοντες κλπ., πράγματα τὰ ὁποῖα διαφεύγουν τὸν ἔλεγχόν μας.

Ἤδη, διὰ νὰ κατορθώσωμεν νὰ προσεγγίσωμεν περισσότερον τὴν λύσιν τοῦ προβλήματος τὸ ὁποῖον μᾶς ἀπασχολεῖ, εἶναι ἀνάγκη νὰ κατανοήσωμεν καὶ τὴν ἔννοιαν τῆς στατιστικῆς ὑποθέσεως, ἐν συνδυασμῷ μὲ τὴν ἔννοιαν τῆς πιθανότητος.

Λέγομεν ὅτι μία στατιστικὴ ὑπόθεσις εἶναι ἀληθής, ὅταν εἰς μίαν ἀπεριόριστον σειρὰν τυχαίων δειγματοληψιῶν αὕτη ἐπαληθεύεται τουλάχιστον 95 φορὰς ἐπὶ τῶν 100. Οἱ ἄνθρωποι, ἐν τούτοις, διαφέρουν κατὰ τὸν βαθμὸν τῆς πίστεως τὸν ὁποῖον τοποθετοῦν ἐπὶ ἑνὸς γεγονότος, ὡς τυχαίου ἢ μή. Κάτι ἀνάλογον συμβαίνει καὶ εἰς τὸν καθημερινὸν βίον: Ἄλλοι εἶναι δύσπιστοι, ἄλλοι εὔπιστοι κλπ. Οἱ στατιστικολόγοι φαίνεται ὅτι κατ' ἀρχὴν συμφωνοῦν ἐπὶ τοῦ ὅτι μία στατιστικὴ ὑπόθεσις εἶναι ἀληθής, ὡς ἀνωτέρω ἐλέχθη, ὅταν ἐκ τύχης καὶ μόνον ἐκ ταύτης, ἐπαληθεύεται 95 φορὰς εἰς τὰς 100, καὶ δὲν ἐπαληθεύεται μόνον 5 φορὰς εἰς τὰς 100.

Παράδειγμα:

Ἐλέγχεται τὸ ἀληθὲς ἢ μὴ τῆς στατιστικῆς ὑποθέσεως, ὅτι εἰς μίαν δεδομένην κηπευτὴν συστάδα ἐλάτης, ὁ ἀριθμητικὸς μέσος ὅρος τῶν ἀνὰ στρέμμα δένδρων ταύτης, ἀπὸ στηθιαίας διαμέτρου 0,08 μ. καὶ ἀνωτέρω εἶναι 50 μὲ μέσην ἀπόκλισιν τετραγώνου 10 δένδρα. Ἐὰν εἰς μίαν ἀπεριόριστον σειρὰν τυχαίων πραγματικῶν δειγματοληψιῶν ἡ ὑπόθεσις αὕτη εὑρεθῇ ἀληθής, τουλάχιστον 95 φορὰς εἰς τὰς 100, τότε λέγομεν ὅτι πράγματι ἡ ὡς ἀνωτέρω ὑπόθεσις εἶναι βάσιμος, πραγματικὴ καὶ ἀληθής. Ἐὰν τοῦτο δὲν συμβῇ, τότε ἡ ὑπόθεσις ἀπορρίπτεται.

Ἐξ ὅσων ἐλέχθησαν ἀνωτέρω, περὶ κανονικῆς κατανομῆς τῶν μέσων τῶν δειγμάτων μιᾶς ἀπεριορίστου σειρᾶς τυχαίων δειγματοληψιῶν δυνάμεθα, προφανῶς, νὰ συναγάγωμεν, ὅτι εἶναι δυνατὴ ἡ ἐκτίμησις ἑνὸς ἀριθμητικοῦ μέσου ὅρου ἑνὸς στατιστικοῦ πληθυσμοῦ μετὰ προκαθωρισμένου βαθμοῦ ἀκριβείας εἰς ὅρους πιθανότητος, διὰ μιᾶς καὶ μόνον μιᾶς τυχαίας δειγματοληψίας, χωρὶς νὰ ἔχωμεν ἀνάγκην μιᾶς ἀπεριορίστου σειρᾶς τοιούτων. Τοῦτο, ἄλλως τε, ἐγένετο ἤδη καὶ ἐδόθη ὁ τύπος (1), ἰσχύων εἰς ὅρους πιθανότητος 68 %. Ἐπειδὴ ὅμως, συμφώνως πρὸς ὅσα ἐλέχθησαν ἀνωτέρω, ὁ βαθμὸς οὗτος δὲν κρίνεται ἐπαρκής, διὰ τοῦτο λαμβάνομεν τὸν διὰ τὸν βαθμὸν πιθανότητος 95 % ἰσχύοντα τύπον, ὁ ὁποῖος εἶναι:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \cdot 1{,}96 \qquad (2)$$

ἐξ οὗ πάλιν ἔχομεν:

$$N = \left(\frac{\sigma}{\sigma_{\bar{x}}} \cdot 1{,}96 \right)^2 \qquad (2\alpha)$$

Ἀποδεικνύεται ἀκόμη ὅτι διὰ νὰ ὑποβιβάσωμεν τὸ μέσον σφάλμα τετραγώνου τοῦ μέσου εἰς τὸ νυοστὸν ἀρχικῶς δοθέντος μεγέθους, πρέπει νὰ αὐξήσωμεν τὸν ἀρχικὸν ἀριθμὸν τῶν παρατηρήσεων κατὰ ν^2 φοράς.

Ἐκ τῶν τύπων (1α) καὶ (2α) προκύπτει ὅτι, διὰ νὰ λύσωμεν τὸ πρόβλημα τὸ ὁποῖον μᾶς ἀπασχολεῖ, ἔχομεν ἀνάγκην νὰ γνωρίζωμεν τὴν ποσότητα σ, τουτέστιν τὴν μέσην ἀπόκλισιν τετραγώνου τοῦ πληθυσμοῦ τὴν ὁποίαν, συνήθως, δὲν γνωρίζομεν. Καὶ πάλιν εἰς τὸ σημεῖον τοῦτο πρέπει νὰ προσέξῃ ὁ ἀναγνώστης: Ἡ Στατιστικὴ Μεθοδολογία δὲν εἶναι πανάκεια οὔτε ἔχει μαντικὰς ἱκανότητας. Δὲν εἶναι δυνατὸν νὰ δώσῃ στοιχεῖα ἐπὶ ἀγνώστων πληθυσμῶν. Ὅπως αἱ ἄλλαι ἐπιστῆμαι, οὕτω καὶ αὕτη ἔχει ἀνάγκην τῆς παρατηρήσεως διὰ νὰ λύσῃ τὰ προβλήματά της. Ἐν πάσῃ περιπτώσει τὴν ὡς ἄνω παράμετρον δυνάμεθα νὰ πορισθῶμεν εἴτε ἐξ ἱστορικῶν στοιχείων, εἴτε ἐκ προκαταρκτικοῦ δείγματος, λαμβανομένου εἰδικῶς πρὸς τοῦτο ἐκ τοῦ οἰκείου πληθυσμοῦ. Λέγομεν ὅτι κατέχομεν μίαν στατιστικὴν πληροφορίαν ἐξ ἱστορικῶν στοιχείων, ὅταν ἡ πληροφορία αὕτη εἶναι γνωστὴ ἐκ προγενεστέρων ἐργασιῶν εἰς τὸν ἴδιον στατιστικὸν πληθυσμόν, εἰς τὴν ἰδίαν συστάδα, ἐπὶ παραδείγματι, ἢ εἰς ἄλλην παρακειμένην ὑπὸ ὁμοειδεῖς συνθήκας τελοῦσαν. Ἐξ ἄλλου, ὀνομάζομεν ἐν προκειμένῳ ἓν δεῖγμα προκαταρκτικόν, ὅταν τοῦτο ἔχῃ ἁπλῶς βοηθητικὸν ἢ ἐξερευνητικὸν χαρακτῆρα καὶ δὲν εἶναι τὸ ζητούμενον τοιοῦτον, τὸ ὁποῖον θὰ ἠδυνάμεθα, πρὸς διάκρισιν, νὰ καλέσωμεν ὁριστικόν.

Κατὰ τὸν προσδιορισμὸν τοῦ N, καλὸν εἶναι νὰ λαμβάνηται εἰς ἐλαφρῶς μεγαλύτερος ἀριθμὸς μονάδων δείγματος, ἐκείνου ὁ ὁποῖος ὑπολογίζεται ἐκ τοῦ ἀνωτέρω τύπου, ἀποκλειστικῶς καὶ μόνον πρὸς ἐξουδετέρωσιν τοῦ ἐνδεχομένου ὑπάρξεως παρὰ τῷ ἐρευνωμένῳ πληθυσμῷ μιᾶς μέσης ἀποκλίσεως τετραγώνου μεγαλυτέρας τῆς ἀρχικῶς ἐν τῷ τύπῳ τούτῳ τεθείσης, εἴτε ἐξ ἱστορικῶν στοιχείων εἴτε ἐκ προκαταρκτικοῦ δείγματος. Τέλος, δὲν πρέπει νὰ διαφεύγῃ ἡμᾶς ὁ χαρακτὴρ τοῦ ἐκτιμωμένου σφάλματος. Τοῦτο, διὰ τὴν περίπτωσίν μας, μετρᾷ μόνον τὸ σφάλμα δειγματοληψίας καὶ εἶναι ἄσχετον καὶ ἀνεξάρτητον ἀπὸ τὰ σφάλματα μετρήσεως, ἀπὸ τὰ συστηματικὰ σφάλματα καί, τέλος, πολὺ περισσότερον, ἀπὸ τὰ λάθη. Ὁ ἀναγνώστης πρέπει νὰ προφυλάσσηται ἰδιαιτέρως ἀπὸ τὴν (ὄχι ὀρθὴν) ἀντίληψιν, ὅτι διὰ τῆς αὐξήσεως τοῦ μεγέθους τοῦ δείγματος δύναται νὰ μετριάσῃ οἱονδήποτε εἶδος σφάλματος. Μόνον τὸ δειγματοληπτικὸν σφάλμα δύναται νὰ μετριάσῃ καὶ ὄχι ὅλα τὰ σφάλματα ἢ λάθη τὰ ὁποῖα μάλιστα, κατὰ τὸ πλεῖστον αὐξάνονται μὲ τὴν αὔξησιν τοῦ δείγματος.

Δίδομεν κατωτέρω μερικὰ παραδείγματα ἐκ τῆς δασικῆς πράξεως ἐπὶ τοῦ ὑπὸ διαπραγμάτευσιν θέματος.

ΠΑΡΑΔΕΙΓΜΑ 1ον

Ζητεῖται νὰ ἐκτιμηθῇ ὁ ἀριθμητικὸς μέσος ὅρος τῶν στηθιαίων διαμέτρων τῶν δένδρων μιᾶς κηπευτῆς συστάδος ἐλάτης, ἀπὸ 0,08 μ. καὶ ἀνωτέρω, ὅταν εἶναι γνωστὸν ἐξ ἱστορικῶν στοιχείων, ὅτι ἡ μέση ἀπόκλισις τετραγώνου τῶν στηθιαίων διαμέτρων τῶν μονάδων τοῦ στατιστικοῦ τούτου πληθυσμοῦ ἀπὸ τοῦ ἀριθμητικοῦ μέσου ὅρου των εἶναι 0,07 μ. Ἡ συστὰς εἶναι γνωστῆς ἐκτάσεως ἐξ 180 στρεμμάτων καὶ περιλαμβάνει, εἰς ἁδρὰν ἐκτίμησιν, ἄνω τῶν 8500 δένδρων, ὡς ταῦτα προσδιωρίσθησαν ἀνωτέρω. Ἐπειδὴ ἡ συστὰς πρόκειται νὰ πωληθῇ, λόγῳ προσαρτήσεως εἰς τὸ γειτονικὸν δάσος, διὰ τὴν ἀποστρογγύλωσιν τῶν ὁρίων τοῦ τελευταίου τούτου, ἐπιθυμοῦμεν νὰ ἐκτιμήσωμεν διὰ δειγματοληψίας τὸν ὡς ἀνωτέρω ἀριθμητικὸν μέσον ὅρον, γνωρίζοντες ἀκόμη, ἐκ συγκεκριμένων τινῶν ἐνδείξεων, ὅτι ἡ κατανομὴ τῶν ἐν λόγῳ διαμέτρων εἰς τὴν συστάδα εἶναι κανονική. Ἐρωτᾶται: Πόσα δένδρα θὰ παχυμετρήσωμεν διὰ τὸν ἀνωτέρω σκοπόν;

Ἐπειδὴ ὁ σκοπὸς τῆς ζητουμένης ἐκτιμήσεως εἶναι σοβαρός, ἀποφασίζομεν ν' ἀνεχθῶμεν ἓν μέγιστον ὅριον σφάλματος ἐξ 0,01 μ. εἰς ὅρους πιθανότητος 95 %. Τοῦτο σημαίνει ὅτι εἴμεθα διατεθειμένοι νὰ ἀνεχθῶμεν ὅπως ὁ ἀριθμητικὸς μέσος ὅρος τῶν στηθιαίων διαμέτρων τῶν δένδρων τοῦ οἱουδήποτε τυχαίου καὶ πολυαρίθμου δείγματος τὸ ὁποῖον θὰ ἠδυνάμεθα νὰ λάβωμεν ἐκ τοῦ ἀνωτέρω πληθυσμοῦ, ἀπέχῃ τὸ πολὺ 0,01 μ. πλέον ἢ ἔλαττον τοῦ ἀληθοῦς καὶ πραγματικοῦ μέσου ὅρου τῶν στηθιαίων διαμέτρων ὁλοκλήρου τοῦ πληθυσμοῦ τούτου, τὸν ὁποῖον, βεβαίως, δὲν γνωρίζομεν οὐδ' εἶναι δυνατὸν νὰ μάθωμεν, ὡς προεξετέθη, ἐκτὸς ἐὰν καταμετρήσωμεν ὅλα τὰ δένδρα τῆς συστάδος, πρᾶγμα τὸ ὁποῖον οὐσιαστικῶς ἀπεκλείσαμεν διὰ τῆς ἤδη ἀποφασισθείσης δειγματοληψίας. Ἐφαρμόζοντες τὸν τύπον (2α) λαμβάνομεν:

$$N = \left(\frac{0{,}07\ \mu.}{0{,}01\ \mu.} \cdot 1{,}96\right)^2 = 188 \text{ δένδρα}$$

Ἐὰν ἤδη λάβωμεν καὶ παχυμετρήσωμεν τὰ δένδρα ταῦτα καὶ ἀνεύρωμεν ὡς ἀριθμητικὸν μέσον ὅρον τῶν στηθιαίων διαμέτρων των τὸ μέγεθος 0,30 μ. τοῦτο οὐδὲν ἄλλο θὰ ἐσήμαινεν εἰ μὴ ὅτι ὁ ἀληθὴς καὶ πραγματικὸς μέσος ὅρος τῶν στηθ. δ. ὅλων τῶν δένδρων τῆς συστάδος θὰ ἔκειτο εἰς τὸ διάστημα τὸ ὁριζόμενον ὑπὸ τῶν ἀριθμῶν 0,30 μ. – 0,01 μ. καὶ 0,30 μ. + 0,01 μ. ἤτοι εἰς τὸ διάστημα μεταξὺ 0,29 μ. καὶ 0,31 μ. δηλ. θὰ ἦτο εἰς ἐκ τῶν τριῶν ἀριθμῶν 0,29 μ., 0,30 μ. καὶ 0,31 μ. (Αἱ παραστάσεις καὶ οἱ ἀριθμοὶ οὗτοι οὐδεμίαν, κατ' ἀρχήν, ἔχουν ἄμεσον σχέσιν μὲ τ' ἀνωτέρω τεθέντα σύμβολα καὶ ἀριθμούς: $\bar{x} \pm 1{,}00\sigma$ κλπ.).

Κατ' ἄλλην διατύπωσιν: Εἰς μίαν ἀπεριόριστον σειρὰν τυχαίων δειγματοληψιῶν, ἑκάστη τῶν ὁποίων θὰ περιελάμβανεν ἓν πολυάριθμον δεῖγμα ἐκ N μονάδων ἕκαστον, οἱ ἀπὸ τῶν δειγμάτων τούτων λαμβανόμενοι μέσοι θὰ ἔκειντο εἰς τὸ διάστημα 0,29 μ. ἕως 0,31 μ., 95 φορὰς ἐπὶ τῶν 100, ὄχι μεμονωμένως ἐξεταζόμενοι ἀλλὰ κατὰ μέσον ὅρον καὶ δι᾽ ὁλόκληρον τὴν σειρὰν ταύτην τῶν δειγματοληψιῶν. Εἰς ἐκείνους τυχὸν ἐκ τῶν ἀναγνωστῶν, οἱ ὁποῖοι καὶ πάλιν θὰ εἶχον ἀντιρρήσεις ἐπὶ τοῦ τρόπου τούτου τῆς ἐκτιμήσεως τοῦ μέσου ὅρου εἰς ὅρους πιθανότητος, ἃς μᾶς ἐπιτραπῇ ἀκόμη νὰ παρατηρήσωμεν, ὅτι ἐὰν εἷς δασολόγος κατεμέτρα τὰς στηθιαίας διαμέτρους ὅ λ ω ν καὶ τῶν 8500 δένδρων τῆς συστάδος δ ύ ο φ ο ρ ά ς, πάλιν δὲν θὰ εὕρισκε τὸν αὐτὸν μέσον ὅρον καὶ τὰς δύο φοράς, ὅπως τοῦτο τουλάχιστον συνάγεται ἀπὸ τοὺς σημερινοὺς ὅρους τῆς σχετικῆς ἀνθρωπίνης ἐργασίας. Πολὺ περισσότερον δὲν θὰ εὕρισκον τὸν αὐτὸν μέσον ὅρον δ ύ ο συνάδελφοι ἐργαζόμενοι κεχωρισμένως ἐπὶ τῆς ἰδίας ὁλοκληρωτικῆς ἀπογραφῆς.

ΠΑΡΑΔΕΙΓΜΑ 2ον

Ὁ δασολόγος ἑνὸς πριστηρίου ἐπιθυμεῖ νὰ γνωρίζῃ πόση εἶναι, κατὰ μέσον ὅρον, ἡ ἀπομείωσις τοῦ πλάτους τῶν σανίδων ἐλάτης τῆς κατηγορίας 4,00 μ. × 0,25 μ. × 0,024 μ. συνεπείᾳ ἀπωλείας ὕδατος, 3 μῆνας μετὰ τὴν ἔξοδόν των ἐκ τοῦ πριστηρίου καὶ τὴν ἰσόχρονον ἔκθεσίν των εἰς τὸ ἐλεύθερον περιβάλλον κατὰ τὸ θέρος. Ὁ ζητούμενος μέσος ὅρος χρειάζεται διὰ νὰ δοθῇ ἀνάλογος ἀνοχὴ εἰς τὰς οἰκείας σανίδας κατὰ τὴν πρίσιν των. Ἐρωτᾶται: πόσων σανίδων πρέπει νὰ καταμετρήσωμεν τὴν ἀπομείωσιν τοῦ πλάτους, ὡς προεξετέθη, διὰ νὰ λάβωμεν τὸν ἐπιζητούμενον μέσον ὅρον;

Οὐδεμίαν πληροφορίαν ἔχομεν περὶ τοῦ εὔρους τῆς ἀπομειώσεως τοῦ πλάτους τῶν σανίδων συνεπείᾳ ἀπωλείας ὕδατος. Διὰ νὰ διερευνήσωμεν τὸν πληθυσμὸν ὡς πρὸς τὸ σημεῖον τοῦτο ἀποφασίζομεν νὰ λάβωμεν ἓν π ρ ο κ α τ α ρ κ τ ι κ ὸ ν δεῖγμα. Ὄντως, τὸ δεῖγμα τοῦτο ληφθέν, ἔδωσεν εἰς ἡμᾶς τ᾽ ἀκόλουθα ἀποτελέσματα:

Διαστ. τάξ. χ, μ.	Συχν. f	ξ	ξ. f	(ξ).²f	Κεντρικαὶ Τιμαὶ x	f. x
0,002—0,004	4	—2	—8	16	0,003	0,012
4— 6	6	—1	—6	6	5	30
6— 8	12	0	0	0	7	84
8— 10	7	+1	+7	7	9	63
0,010—0,012	3	+2	+6	12	0,011	0,033
Σύνολον	32		—1	41	—	0,222

Ὁ πίναξ οὗτος δέον νὰ ἑρμηνευθῇ ὡς ἑξῆς: Τὸ δεῖγμα ἀπετελέσθη ἐκ 32 σανίδων. Αὗται ἐκτεθεῖσαι δοκιμαστικῶς εἰς τὸ ὕπαιθρον κατὰ τὸ θέρος ἐπὶ 3μηνον, παρουσίασαν τὴν κάτωθι ἀπομείωσιν: 4 ἐξ αὐτῶν παρουσίασαν ἀπομείωσιν κυμανθεῖσαν μεταξὺ 0,002 μ. καὶ 0,004 μ., 6 ἐξ αὐτῶν μεταξὺ 0,004 μ. καὶ 0,006, 12 μεταξὺ 0,006 μ. καὶ 0,008 μ. κ.ο.κ. Ἡ μέση ἀπόκλισις τετραγώνου τῆς ὡς ἀνωτέρω κατανομῆς ὑπολογισθεῖσα εὑρέθη ὡς ἴση πρὸς

$$\sigma = 0{,}002 \,.\, \sqrt{\frac{(32.41)-(-1)^2}{32^2}} = 0{,}00227 \text{ μ.}$$

Ἐπειδὴ τὸ ἐμπόριον εἶναι αὐστηρὸν εἰς τὴν κατηγορίαν ταύτην τῆς ξυλείας, ἀποφασίζομεν νὰ δεχθῶμεν ἕνα τόσον πολυάριθμον δεῖγμα, ὥστε ὁ ἀπὸ τούτου ὑπολογιζόμενος ἀριθμητικὸς μέσος ὅρος τοῦ μεγέθους τῆς ἀπομειώσεως, νὰ μὴ ἀπέχῃ τοῦ ἀληθοῦς καὶ πραγματικοῦ τοιούτου πλέον ἢ ἔλαττον τοῦ 0,001 μ. εἰς τὸ σύνολον σχεδὸν τῶν παρατηρήσεων μὲ μίαν ἀνοχὴν 5 % περίπου, δηλ. εἰς ὅρους πιθανότητος 95 %.

Ὁ τύπος (2α) μᾶς δίδει τὸ ἀκόλουθον μέγεθος δείγματος:

$$N = \left(\frac{0,00227}{0,00100} \cdot 1,96\right)^2 = 20.$$

Ἐρωτᾶται τώρα: Ποία εἶναι ἡ ἔννοια τοῦ εὑρεθέντος ἀριθμοῦ 20 ; Ἡ ἔννοια εἶναι ὅτι ἐκ τῆς παραγωγῆς τοῦ πριστηρίου θὰ πρέπει νὰ λάβωμεν τυχαίως 20 σανίδας τῆς κατηγορίας 4,00 μ. × 0,25 μ. × 0,024 μ., νὰ ἐκθέσωμεν ταύτας εἰς τὸ ἐλεύθερον περιβάλλον κατὰ τὸ θέρος ἐπὶ 3μηνον καὶ νὰ ὑπολογίσωμεν τὸ μέγεθος τῆς ἀπομειώσεως τὴν ὁποίαν θὰ ὑποστοῦν αὗται, λαμβάνοντες τὸν οἰκεῖον μέσον ὅρον. Ἂς ὑποθέσωμεν ὅτι ἐπράξαμεν τοῦτο καὶ εὕρομεν ἀριθμητικὸν μέσον ὅρον ἀπομειώσεως ἴσον πρὸς τὴν ποσότητα $\bar{x} = 0,00594$ μ. Τοῦτο θὰ ἐσήμαινεν ὅτι θὰ ἔπρεπε νὰ κανονίσωμεν τὰ οἰκεῖα ἐργαλεῖα πρίσεως κατὰ τοιοῦτον τρόπον, ὥστε αἱ σανίδες νὰ ἐξήρχοντο τοῦ πριστηρίου μὲ διάστασιν πλάτους 0,25 μ. + 0,00594 μ. = 0,256 μ. περίπου. Τούτου γενομένου, θὰ ἀνεμένετο ὅτι αἱ σανίδες τῆς κατηγορίας ταύτης, δὲν θ' ἀπέβαινον πλέον ἐλλειποδιάστατοι κατὰ τὸ θέρος, εἰμὴ κατὰ μίαν μικρὰν ἀναλογίαν 5 % περίπου καὶ κατὰ μέσον ὅρον δι' ὁλόκληρον τὴν ἀπεριόριστον σειρὰν παραγωγῆς τοῦ πριστηρίου.

Εἰς τὴν προκειμένην περίπτωσιν τὸ ληφθὲν προκαταρκτικὸν δεῖγμα ἐκ 32 σανίδων, ἐξ ἁπλῆς συμπτώσεως, καὶ μόνον ἐκ ταύτης, ὑπερκαλύπτει τὸ ἐπιθυμούμενον μέγεθος τοῦ δείγματος δι' ἀκρίβειαν 0,001 μ. εἰς ὅρους πιθανότητος 0,95. Ἐὰν ὑπολογίσωμεν τὸν ἀριθμητικὸν μέσον ὅρον τῆς ἀπομειώσεως ἐκ τοῦ δείγματος τούτου τῶν 32 σανίδων, θὰ ἴδωμεν ὅτι οὗτος συμφώνως πρὸς τὸν ἀνωτέρω πίνακα, ἀνέρχεται εἰς $\frac{0,222 \text{ μ.}}{32} = 0,00694$ μ.

Κατόπιν τούτου θὰ ἔπρεπε νὰ ἐκανονίζομεν τὰ οἰκεῖα ἐργαλεῖα πρίσεως εἰς τρόπον ὥστε αἱ σανίδες νὰ ἐπριονίζοντο εἰς ἀρχικὴν διάστασιν πλάτους 0,25 μ. + 0,00694 μ. = 0,257 μ. καὶ ὄχι εἰς διάστασιν πλάτους 0,256 μ., διότι μεταξὺ δύο ἀριθμητικῶν μέσων ὅρων μείζονα ἀξιοπιστίαν παρέχει ἐκεῖνος ὁ ὁποῖος προέρχεται ἀπὸ πολυαριθμωτέρας παρατηρήσεις.

ΠΑΡΑΔΕΙΓΜΑ 3ον

Εἷς δασολόγος-διαχειριστὴς ἐπιθυμεῖ νὰ ἐκτιμήσῃ τὸν ἀριθμητικὸν μέσον ὅρον τῶν ἀνὰ ἓν τετραγωνικὸν μέτρον φυομένων ἀρτιφύτρων καὶ φυταρίων μαύρης πεύκης εἰς μίαν δασικὴν συστάδα τοῦ ἰδίου τούτου δασικοῦ εἴδους, προκειμένου νὰ ρυθμίσῃ τὴν διαχείρισιν ταύτης δι' ὑλοτομιῶν, ἀπαγορεύσεων βοσκῆς κλπ. Πρὸς τοῦτο, ἐφωδιασμένος μὲ μίαν ἁπλῆν ράβδον μήκους 1,14 μ. διατρέχει τὴν ἐρευνωμένην συστάδα καὶ ἀνὰ 30 βήματα ρίπτει τὴν ράβδον κατὰ τύχην πρὸς μίαν κατεύθυνσιν. Ἀκολούθως, θεωρῶν τὴν ὑπὸ τῆς ράβδου καταλαμβανομένην θέσιν καὶ ἀπόστασιν ὡς διάμετρον κύκλου ἐμβαδοῦ 1 μ.² μετρᾷ τὰ ἐπὶ τῆς ἐπιφανείας ταύτης ἀπαντώμενα ἀρτίφυτρα κλπ. καὶ ἀναγράφει τὸ ἀποτέλεσμα εἰς εἰδικὸν σημειωματάριον. Ἐρωτᾶται : Πόσας τοιαύτας ρίψεις καὶ καταμετρήσεις πρέπει νὰ κάμῃ διὰ νὰ ἐκτιμήσῃ τὸν ἐπιζητούμενον ἀριθμητικὸν μέσον ὅρον μὲ προσέγγισιν 2 ἀρτιφύτρων καὶ φυταρίων ἀνὰ τετραγωνικὸν μέτρον εἰς ὅρους πιθανότητος 95 %.

Δὲν γνωρίζομεν τὴν μέσην ἀπόκλισιν τετραγώνου εἰς τὸν ἐρευνώμενον πληθυσμὸν καὶ ἀποφασίζομεν νὰ ἐκτιμήσωμεν ταύτην μέσῳ ἑνὸς προκαταρκτικοῦ δείγματος ἐξ ἀριθμοῦ τινος παρατηρήσεων-ρίψεων τῆς ράβδου. Ὄντως τὸ δεῖγμα τοῦτο ληφθὲν ἀπέδωσε τὰ εἰς τὸν κατωτέρω πίνακα σημειούμενα ἀποτελέσματα.

ΠΙΝΑΞ

Ὑπολογισμοῦ τοῦ ἀριθμητικοῦ μέσου ὅρου καὶ τῆς μ.α.τ. τῆς κατανομῆς τῶν ἀρτιφύτρων καὶ φυταρίων.

Διάστημα τάξ. Ἀριθ. δενδρυλλίων	Κεντρικὴ τιμὴ x	Συχν. f	f. x.	ξ.	ξ. f	$\xi^2 f$
0— 4	2	5	10	—3	—15	45
4— 8	6	7	42	—2	—14	28
8—12	10	10	100	—1	—10	10
12—16	14	12	168	0	0	0
16—20	18	10	180	+1	+10	10
20—24	22	8	176	+2	+16	32
24—28	26	3	78	+3	+ 9	27
Σύνολον		55	754		— 4	152

Ὁ πίναξ οὗτος δέον νὰ ἑρμηνευθῇ ὡς ἑξῆς : Ἐγένοντο ἐν ὅλῳ 55 παρατηρήσεις, δηλ. ρίψεις τῆς ράβδου καὶ καταμετρήσεις τοῦ ἀριθμοῦ τῶν ἀρτιφύτρων ἐπὶ τῆς οἰκείας ἐπιφανείας. Εἰς 5 ἐκ τῶν ἐπιφανειῶν τούτων, ὁ ἀριθμὸς τῶν ἐπὶ τοῦ οἰκείου τετραγωνικοῦ μέτρου καταμετρηθέντων ἀρτιφύτρων ἐκυμαίνετο ἀπὸ 0 ἕως 4. Εἰς 7 ἐξ αὐτῶν ὁ ἀριθμὸς οὗτος ἐκυμαίνετο ἀπὸ 4 ἕως 8 κ.ο.κ. (Ὁ ἀριθμὸς 4 δὲν καταχωρεῖται εἰς τὴν πρώτην τάξιν : 0–4, ἀλλ' εἰς τὴν δευτέραν τοιαύτην : 4–8 κ.ο.κ.). Ἡ μέση ἀπόκλισις τετραγώνου τῆς κατανομῆς ταύτης εἶναι :

$$\sigma = 4\sqrt{\frac{(55\,.\,152) - (-4)^2}{55^2}} = 6{,}6$$

Ὁ τύπος (2α) μᾶς δίδει τὸ ἐπιζητούμενον μέγεθος τοῦ δείγματος :

$$N = \left(\frac{6{,}6}{2}\cdot 1{,}96\right)^2 = 42.$$

Τοῦτο σημαίνει ὅτι διὰ νὰ ἐκτιμήσωμεν τὸν ζητούμενον ἀριθμητικὸν μέσον ὅρον εἰς ὅρους πιθανότητος 95 °/₀ δέον νὰ λάβωμεν διὰ τῆς ράβδου 42 τυχαίας ἐπιφανείας. Κατὰ σύμπτωσιν, ἐν τῇ προκειμένῃ περιπτώσει τὸ ληφθὲν δεῖγμα ἐκ 55 παρατηρήσεων ὑπερκαλύπτει τὸ ἐπιζητούμενον μέγεθος τοῦ δείγματος. Δυνάμεθα ὅθεν νὰ λάβωμεν τὸν μέσον ὅρον τοῦ δείγματος τούτου ἀντὶ τοῦ ἐπιζητουμένου, μὲ ποιάν τινα μάλιστα μείζονα ἐμπιστοσύνην, ἔναντι τοῦ κινδύνου ὑπάρξεως παρὰ τῷ πληθυσμῷ τῶν ἀρτιφύτρων μιᾶς μέσης ἀποκλίσεως τετραγώνου μεγαλυτέρας τῆς ὑπολογισθείσης τοιαύτης ἐξ 6,6 ἀρτιφύτρων. Ὁ ἀριθμητικὸς μέσος ὅρος τῆς ὡς ἀνωτέρω κατανομῆς εἶναι, ὡς ἐμφαίνηται ἐκ τοῦ ἀνωτέρω πίνακος:

$$\bar{x} = \frac{754}{55} = 13{,}7 \simeq 14$$

Τοῦτο σημαίνει ὅτι ὁ ἀνὰ τετραγωνικὸν μέτρον ἀριθμητικὸς μέσος ὅρος τῶν ἀρτιφύτρων κυμαίνεται ἐν τῇ πραγματικότητι μεταξὺ τῶν ἀριθμῶν 12 καὶ 16, εἰς ὅρους πιθανότητος 95 °/₀.

Ἐὰν ἠθέλαμεν μεγαλυτέραν ἀκρίβειαν, π.χ. 1 ἀρτίφυτρον ἀνὰ τετραγωνικὸν μέτρον, τότε τὸ προκαταρκτικὸν δεῖγμα θ' ἀπεδεικνύετο ἀνεπαρκὲς νὰ καλύψῃ τὸ ἐπιθυμούμενον μέγεθος τοῦ δείγματος, τὸ ὁποῖον, τότε, θὰ ἔπρεπε νὰ εἶναι τὸ ἀκόλουθον:

$$N = \left(\frac{6{,}6}{1{,}0} \cdot 1{,}96\right)^2 = 169$$

Δηλ. θὰ ἔπρεπε τότε νὰ λάβωμεν τὸν ἀριθμητικὸν μέσον ὅρον 169 παρατηρήσεων, ὁ ὁποῖος θὰ ἦτο καὶ ὁ ζητούμενος, ἐπίσης εἰς ὅρους πιθανότητος 95°/₀, ἀλλὰ μὲ διάστημα ἐμπιστοσύνης 1 ἀρτίφυτρον, ἄνω ἢ κάτω τοῦ ἀληθοῦς ἀριθμητικοῦ μέσου ὅρου.

Ἐκ τῆς συγκρίσεως τῶν ἀριθμῶν 42 καὶ 169 ἐλέγχεται ὡς ἀληθὴς ἡ προμνησθεῖσα σχέσις σφάλματος καὶ μεγέθους δείγματος.

Πρὶν κλείσωμεν τὴν παροῦσαν μας δίδομεν κατωτέρω ἓν βοηθητικὸν σημείωμα περὶ τῆς ἐννοίας καὶ τοῦ τρόπου τοῦ ὑπολογισμοῦ τῶν παραμέτρων $\bar{x}$ καὶ σ.

ΣΗΜΕΙΩΜΑ

βοηθητικόν, περὶ τῆς ἐννοίας καὶ τοῦ τρόπου τοῦ ὑπολογισμοῦ τοῦ ἀριθμητικοῦ μέσου ὅρου $\bar{x}$ καὶ τῆς μέσης ἀποκλίσεως τετραγώνου σ τῶν μονάδων ἑνὸς στατιστικοῦ δείγματος.

Ἡ πλήρης ἀνάπτυξις τοῦ κεφαλαίου τούτου δὲν εἶναι δυνατὴ ἀπὸ τῶν προκειμένων στηλῶν. Διὰ τοῦτο θὰ περιορισθῶμεν εἰς μίαν βραχεῖαν περίληψιν ἀποβλέπουσαν ἰδίᾳ νὰ δείξῃ εἰς τὸν δασολόγον τὸν τρόπον τοῦ ὑπολογισμοῦ τῶν παραμέτρων τούτων. Διὰ πλείονας πληροφορίας παραπέμπομεν τὸν ἀναγνώστην εἰς τὴν σχετικὴν βιβλιογραφίαν:

(Λ.χ.: Κ. Ἀθανασιάδη, Στατιστική, Μέρος Α΄, σελὶς 112, 142, Ἀθῆναι 1957).

Ἡ ἔννοια τοῦ ἀριθμητικοῦ μέσου ὅρου εἶναι ἁπλῆ καὶ γνωστή. Ἡ μέση ἀπόκλισις τετραγώνου εἶναι ἓν χαρακτηριστικὸν μέτρον τὸ ὁποῖον μετρᾷ τὸν βαθμὸν συγκεντρώσεως τῶν μονάδων ἑνὸς στατιστικοῦ δείγματος πέριξ τοῦ ἀριθμητικοῦ μέσου ὅρου των. Κατ' ἄλλην διατύπωσιν θὰ ἠδυνάμεθα νὰ γράψωμεν ὅτι ἡ παράμετρος αὕτη μετρᾷ τὴν μεταβλητότητα τῶν μονάδων ἑνὸς δείγματος.

Παράδειγμα:

Οἱ ἀριθμοί: 5, 6, 7, 8 καὶ 9 ἔχουν μέσον τὸν 7.

Οἱ ἀριθμοί: 2, 4, 7, 9 καὶ 13 ἔχουν ἐπίσης μέσον τὸ 7. Εἶναι ὅμως προφανὲς ὅτι οἱ ἀριθμοὶ τῆς πρώτης ὁμάδος συγκεντροῦνται πυκνότερον καὶ συμμετρικώτερον περὶ τὸν ἀριθμητικὸν μέσον παρὰ ἐκεῖνοι τῆς δευτέρας ὁμάδος. Τοῦτο δύναται νὰ ἐκφρασθῇ καὶ ἀριθμητικῶς ὡς κατωτέρω προκύπτει. Ἡ μέση ἀπόκλισις τετραγώνου ὁρίζεται ὡς ἡ τετραγωνικὴ ρίζα τοῦ ἀριθμητικοῦ μέσου ὅρου τῶν τετραγώνων τῶν ἀποκλίσεων τῶν μονάδων τοῦ πληθυσμοῦ ἀπὸ τῆς μέσης τιμῆς αὐτῶν. Ἡ παράμετρος αὕτη ὑπολογίζεται ὡς ἀκολούθως:

1. Διὰ τὴν α'. ὁμάδα:

$$\sigma = \pm\sqrt{\frac{(5-7)^2+(6-7)^2+(7-7)^2+(8-7)^2+(9-7)^2}{5}} = \pm 1,41$$

2. Διὰ τὴν β'. ὁμάδα:

$$\sigma = \pm\sqrt{\frac{(2-7)^2+(4-7)^2+(7-7)^2+(9-7)^2+(13-7)^2}{5}} = \pm 3,85$$

(5, ὁ ἀριθμὸς τῶν παραγόντων).

Ὥστε βλέπομεν ὅτι ἡ μεταβλητότης ἐντὸς τῶν μονάδων τοῦ ὑπὸ στοιχεῖον β'. πληθυσμοῦ εἶναι $2\,{}^1/_2$ καὶ πλέον φορὰς μεγαλυτέρα ἐκείνης τῶν ὑπὸ στοιχεῖον α'. τοιούτου, μολονότι ἀμφότεροι οἱ πληθυσμοὶ ἔχουν τὸν αὐτὸν ἀριθμητικὸν μέσον ὅρον.

Δίδομεν κατωτέρω ἓν παράδειγμα, περὶ τοῦ τρόπου τοῦ ὑπολογισμοῦ τοῦ ἀριθμητικοῦ μέσου ὅρου καὶ τῆς μέσης ἀποκλίσεως τετραγώνου, ὅταν τὰ δεδομένα τῶν παρατηρήσεων κατατάσσονται εἰς μίαν στατιστικὴν κατανομὴν συχνοτήτων. Ὁ τρόπος τῆς κατατάξεως παραλείπεται ὡς ἁπλοῦς.

Εἷς δασολόγος ἐπιθυμεῖ νὰ ἐξακριβώσῃ τὸν ἀριθμητικὸν μέσον ὅρον καὶ τὴν μέσην ἀπόκλισιν τετραγώνου τῶν στηθιαίων διαμέτρων ὅλων τῶν ἀπὸ 0,30 μ. καὶ ἀνωτέρω δένδρων μιᾶς γηραιᾶς λόχμης ἐλάτης, μιᾶς ὡρισμένης δασικῆς συστάδος. Προβαίνει πρὸς τοῦτο εἰς γενικὴν παχυμέτρησιν ὅλων καὶ τῶν 38 δένδρων τῆς λόχμης καὶ καταχωρίζει τὰ ἀποτελέσματα εἰς τὸν ἀκόλουθον πίνακα:

ΠΙΝΑΞ

ἐμφαίνων τὸν τρόπον τοῦ ὑπολογισμοῦ τοῦ ἀριθμητικοῦ μέσου ὅρου $\overline{x}$ καὶ τῆς μέσης ἀποκλίσεως τετραγώνου σ, τῶν στηθιαίων διαμέτρων τῶν δένδρων μιᾶς δασικῆς λόχμης, τεταγμένων εἰς μίαν στατιστικὴν κατανομὴν συχνοτήτων.

Τάξεις στ. διαμ., x (μέτρα)	Κεντρικαὶ Τιμαὶ x (μέτρα)	Συχνότης f	f. x	$\xi = \frac{x-0,55}{0,10}$	f. ξ	f. $(\xi)^2$
0,30 – 0,40	0,35	5	1,75	– 2	– 10	20
0,40 – 0,50	0,45	8	3,60	– 1	– 8	8
0,50 – 0,60	0,55	15	8,25	0	0	0
0,60 – 0,70	0,65	6	3,90	+ 1	+ 6	6
0,70 – 0,80	0,75	4	3,00	+ 2	+ 8	16
Σύνολον	–	38	20,50		– 4	50

Οἱ ἀριθμοὶ τοῦ πίνακος τούτου δέον νὰ ἑρμηνευθοῦν ὡς ἑξῆς:

Αἱ στηθιαῖοι διάμετροι 5 δένδρων ἔκειντο εἰς τὸ διάστημα ἀπὸ 0,30 μ. ἕως 0,40 μ., ἐκεῖναι 8 δένδρων εἰς τὸ διάστημα ἀπὸ 0,40 μ. ἕως 0,50 μ., εἰς 15 ἄλλα δένδρα, αἱ στηθ. διάμετροι ἔκειντο εἰς τὸ διάστημα ἀπὸ 0,50 μ. ἕως 0,60 μ. κ.ο.κ. μέχρι τέλους τοῦ πίνακος. Σημειωτέον ὅτι τὰ δένδρα τὰ ὁποῖα ἔχουν στηθ. διάμετρον π.χ. 0,40 μ. δὲν κατατάσσονται εἰς τὴν πρώτην τάξιν 0,30 μ. ἕως 0,40 μ. ἀλλ' εἰς τὴν δευτέραν τοιαύτην 0,40 μ. ἕως 0,50 μ. κ.ο.κ. δι' ὅλας τὰς τάξεις.

Ὁ τύπος ὁ ὁποῖος μᾶς δίδει τὸν ἀριθμητικὸν μέσον ὅρον ὅλων καὶ τῶν 38 στηθιαίων διαμέτρων, εἶναι ὁ ἑξῆς:

$$\overline{x} = \frac{\Sigma(f\,.\,x)}{\Sigma(f)} = \frac{20,50}{38} = 0,539 \text{ μ.}$$

(Εἰς τὸν τύπον τοῦτον, ὡς καὶ εἰς τὸν κατωτέρω, τὸ γράμμα Σ σημαίνει: ἀλγεβρικὸν ἄθροισμα).

Ὁ τύπος ὁ ὁποῖος μᾶς δίδει τὴν μέσην ἀπόκλισιν τετραγώνου τῶν καθ' ἕκαστα στηθιαίων διαμέτρων ὅλων καὶ τῶν 38 δένδρων τῆς λόχμης, ἀπὸ τοῦ ἀριθμητικοῦ μέσου ὅρου των εἶναι:

$$\sigma = \pm\, \delta\,.\sqrt{\frac{N\,.\,\Sigma f(\xi)^2 - (\Sigma f \xi)^2}{N^2}}$$

Εἰς τὸν τύπον τοῦτον εἶναι: δ, τὸ διάστημα τάξεως, δηλ. ἡ ἀπόστασις μεταξὺ τοῦ μείζονος καὶ τοῦ ἐλάσσονος ὁρίου ἑκάστης τάξεως, ἐν προκειμένῳ 0,10 μ., N, ὁ ἀριθμὸς τῶν παρατηρήσεων, ἐν προκειμένῳ ὁ ἀριθμὸς 38.

Κατὰ τὰ ἄλλα οἱ ἀριθμοὶ εἶναι ἀφ' ἑαυτῶν ἀντιληπτοί, ἐὰν ἀνατρέξῃ τις εἰς τὸν οἰκεῖον πίνακα. Ἀντικαθιστῶντες λαμβάνομεν:

$$\sigma = \pm\, 0,10\,.\sqrt{\frac{(38\,.\,50) - (-4)^2}{38^2}} = 0,114 \text{ μ.}$$

Τὸ σημείωμα τοῦτο ὡς καὶ ὁ τρόπος τοῦ ὑπολογισμοῦ τοῦ ἀριθμητικοῦ μέσου ὅρου καὶ τῆς μέσης ἀποκλίσεως τετραγώνου εἰς τὰ δύο τελευταῖα παραδείγματα καθιστοῦν τελείως προσιτὰς τὰς ἐννοίας ταύτας καὶ τὸν τρόπον τοῦ ὑπολογισμοῦ τῶν ἀντιστοίχων παραμέτρων.

ΠΕΡΙΛΗΨΙΣ

Ὁ ἀριθμὸς τῶν ἀπαιτουμένων παρατηρήσεων διὰ τὴν ἐκτίμησιν τοῦ ἀριθμητικοῦ μέσου ὅρου ἑνὸς στατιστικοῦ πλήθους ἢ μεγέθους δίδεται ἀπὸ τὸν τύπον : $N = \left(\frac{\sigma}{\sigma\bar{x}}\right)^2$ εἰς ὅρους πιθανότητος 0,68 καὶ ἀπὸ τὸν τύπον :

$$N = \left(\frac{\sigma}{\sigma\bar{x}} \cdot 1{,}96\right)^2$$ εἰς ὅρους πιθανότητος 0,95.

Εἰς τὸν τύπον τοῦτον σ εἶναι ἡ μέση ἀπόκλισις τετραγώνου τοῦ στατιστικοῦ πλήθους ἢ μεγέθους καὶ $\sigma\bar{x}$ εἶναι τὸ μέσον σφάλμα τετραγώνου τοῦ μέσου ὅρου, τὸ ὁποῖον λαμβάνεται ὡς δεδομένον ἐκ τῶν προτέρων, a priori. Τὸ μέγεθος τοῦτο ἐξαρτᾶται ἀπὸ τὸν ἐπιδιωκόμενον σκοπὸν καὶ εἶναι ὁ κυριώτερος παράγων ὁ ὁποῖος προσδιορίζει τὴν ἀκρίβειαν τῆς σχετικῆς ἐργασίας.

Τὸ μέγεθος σ ποριζόμεθα εἴτε ἐξ ἱστορικῶν στοιχείων εἴτε ἐκ προκαταρκτικοῦ δείγματος.

Τὸ μέσον σφάλμα τετραγώνου τοῦ μέσου ἀντιπροσωπεύει τὸ σφάλμα δειγματοληψίας καὶ ὄχι τὰ λοιπὰ σφάλματα μετρήσεως κλπ. ἢ τὰ λάθη. Ἡ αὔξησις τοῦ ἀριθμοῦ τῶν παρατηρήσεων N σμικρύνει τὸ σφάλμα δειγματοληψίας οὐχὶ ὅμως καὶ τὰ λοιπὰ σφάλματα ἢ λάθη τὰ ὁποῖα οὐχὶ σπανίως αὐξάνονται μετὰ τοῦ μεγέθους τοῦ δείγματος.

SUMMARY

The size of the sample necessary to estimate a mean of a population with a stated degree of accuracy is given by the formula :

$$N = \left(\frac{\sigma}{\sigma\bar{x}}\right)^2$$, in terms of probabilities 0,68 or

$$N = \left(\frac{\sigma}{\sigma\bar{x}} \cdot 1{,}96\right)^2$$ in terms of probabilities 0,95

(σ is the standard deviation of the population and $\sigma\bar{x}$ is the standard error of the mean taken arbitrary-a priori acording to the purpose sought).

We procure σ either from historical data or from a preliminary sample.

The standard error of the mean measures only the sampling error; by increasing the size of the sample we lessen the standard error but we do not lessen mistakes in the selection of the sample or measurment errors that may increase with the size of the sample. Three examples with hypothetical data serve to illustrate the subject.

ΒΙΒΛΙΟΓΡΑΦΙΑ

D. Bruce and F. Schumacher, Forest Mensuration, N. York, 1950.

J. Freund and F. Williams, Modern Business Statistics, Englewood Cliffs, N. J. 1958.

K. Ἀθανασιάδου, Στατιστική, Μέρος Α', Ἀθῆναι 1957.