ΚΑΤΑΝΟΜΑΙ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ

(ΠΟΛΥΜΕΤΑΒΛΗΤΟΣ ΚΑΝΟΝΙΚΟΣ ΝΟΜΟΣ)

Τοῦ κ. ΒΑΣΙΛΕΙΟΥ Κ. ΜΠΕΝΟΥ

Ἐπιστημ. Βοηθοῦ τῆς ἕδρας τῆς Στατιστικῆς εἰς τὴν Ἀνωτάτην Βιομηχανικὴν Σχολὴν Πειραιῶς

I. Εἰσαγωγὴ

Σκοπὸς τοῦ παρόντος εἶναι ἡ ἐντὸς τῶν πλαισίων τῆς Τεχνικῆς τῆς Δειγματοληψίας σπουδὴ πληθυσμοῦ τινος, οὐχὶ πλέον ὡς συνήθως γίνεται ὑπὸ τὴν ἔποψιν τῆς ἐπὶ τῇ βάσει κριτηρίου τινός, διακρίσεως αὐτοῦ εἰς δύο τάξεις, ἀλλὰ ὑπὸ τὴν ἔποψιν περισσοτέρων τοιούτων.

Διότι, συνήθως κατὰ τὴν παρουσίασιν ἀποτελεσμάτων, αἱ μονάδες τοῦ ὑπὸ ἐξέτασιν πληθυσμοῦ, ταξινομοῦνται εἰς περισσοτέρας τῶν δύο τάξεις ὑπὸ τὴν ἔποψιν πάντοτε μιᾶς χαρακτηριστικῆς ἰδιότητος. Οὕτως, ἐπὶ παραδείγματι, δεῖγμα ἐξ ἀνθρωπίνου πληθυσμοῦ δύναται νὰ ταξινομηθῇ εἰς 15 ὁμάδας τῶν πέντε (5) ἐτῶν.

II. Πολυωνυμικὸς νόμος

Ἔστω ὅτι ὁ (ὑποθετικὸς) πληθυσμὸς μεγέθους N, ὑπὸ τὴν ἔποψιν κριτηρίου τινὸς (συγκεκριμένης χαρακτηριστικῆς ἰδιότητος) διακρίνεται εἰς τὰς τάξεις.

$$A_1, A_2, A_3, \ldots, A_i, \ldots, A_k$$

εἰς ἃς ἀντιστοιχοῦν

$$N_1, N_2, N_3, \ldots, N_i, \ldots, N_k \text{ μονάδες,}$$

$$\text{ὅπου} \quad \sum_{i=1}^{k} N_i = N, \quad P_i = \frac{N_i}{N} \quad \text{καὶ} \quad \sum_{i=1}^{k} P_i = 1.$$

Ἐὰν ἐκ τοῦ πληθυσμοῦ τούτου λάβωμεν δεῖγμα μεγέθους n (διὰ τῆς μεθόδου τῆς ἐπαναθέσεως), καὶ καλέσωμεν n_i τὰς μονάδας τοῦ δείγματος, αἵτινες ἀναφέρονται εἰς τὴν τάξιν A_i, μὲ ἀντικειμενικὸν σκοπὸν νὰ συμπεράνωμεν τὴν κατανομὴν τῆς μεταβλητῆς X (ἥτις ἀντιστοιχεῖ εἰς τὴν συγκεκριμένην χαρακτηριστικὴν ἰδιότητα τῶν μονάδων τοῦ πληθυσμοῦ καὶ λαμβάνει τὰς τιμὰς $x_1, x_2, \ldots x_k$), θὰ ἔχωμεν n τιμὰς τῆς μεταβλητῆς X :

$$x_1, x_1, x_2, x_1, \ldots, x_2, x_4, \ldots x_k, x_k$$

ἐκ τῶν ὁποίων : n_1 θὰ ἔχουν τὴν τιμὴν x_1
n_2 » » » » x_2
n_3 » » » » x_3
. . . .
n_k θὰ ἔχουν τὴν τιμὴν x_k

$$\text{καὶ} \quad \sum_{i=1}^{k} n_i = n.$$

Ἐὰν δὲ ἡ πιθανότης ἵνα ἡ μεταβλητὴ X λάβῃ γενικῶς τὴν τιμὴν x_i ($i = 1,2,3, \ldots, k$) εἶναι P_i ($i = 1,2,3, \ldots, k$), τότε ἡ πιθανότης ἵνα ἡ μεταβλητὴ X λάβῃ τὴν συγκεκριμένην τιμὴν x_1, n_1 φοράς, τὴν τιμὴν x_2, n_2 φορὰς κ.ο.κ. ἀποδεικνύεται ὅτι εἶναι :

$$Pn_i\,(x_i) = \frac{n!}{n_1!\, n_2! \ldots n_k!} P_1^{n_1} \cdot P_2^{n_2} \cdots P_k^{n_k} =$$

$$= \frac{n!}{\prod_{i=1}^{k} n_i!} \cdot \prod_{i=1}^{k} P_i^{n_i} \qquad (1)$$

{ ἔνθα Π τὸ σύμβολον τοῦ πολλαπλασιασμοῦ }.

Παρατηροῦμεν ὅτι ὁ τύπος (1) συνίσταται ἐκ δύο μερῶν :

α) Τὸν πολυωνυμικὸν συντελεστήν :

$$\frac{n!}{\prod_{i=1}^{k} n_i!} = \binom{n}{n_1, n_2, \ldots n_k}$$, ὅστις εἶναι ὁ ἀριθμὸς τῶν μεταθέσεων μετ' ἐπαναλήψεως τῶν n ἀνὰ $n_1, n_2, \ldots n_k$, ἔνθα $n_1 + n_2 + \ldots + n_k = n$ καὶ δίδει τὸν ἀριθμὸν τῶν διαφορετικῶν τρόπων κατὰ τοὺς ὁποίους ἔχομεν ἀκριβῶς n_1 φορὰς τὴν τιμὴν x_1, n_2 φορὰς τὴν τιμὴν x_2 κ.ο.κ.

β) Τὸν παράγοντα $\prod_{i=1}^{k} P_i^{n_i}$, ὁ ὁποῖος δίδει τὴν πιθανότητα δι' ἕκαστον ἐξ αὐτῶν τῶν διαφορετικῶν τρόπων.

Ἡ σχέσις (1) προκύπτει ἐκ τῆς γενικεύσεως τοῦ διωνυμικοῦ νόμου.

Ἐπειδὴ δὲ εἶναι τὸ ἀνάπτυγμα τοῦ πολυωνύμου $(P_1 + P_2 + \cdots + P_k)^n$ καλεῖται πολυωνυμικὸς νόμος.

Ὡς εἶναι προφανές, ἂν ὁρισθῇ μία τάξις ὡς ἐπιτυχία, π.χ. ἡ τάξις A_i* καὶ αἱ ὑπόλοιποι ὑπαχθοῦν εἰς τὴν σύνθετον τάξιν καλουμένην «Τάξις ὄχι A_i», τότε, ἐὰν καλέσωμεν P_i τὸ ποσοστὸν τῶν μονάδων τῆς τάξεως A_i, εἰς τὸν πληθυσμὸν N, ἡ τάξις A_i θὰ περιλαμβάνῃ NP_i μονάδας. Τὸ ποσοστὸν τοῦτο P_i χαρακτηριστικὸν τῆς συνθέσεως τοῦ πληθυσμοῦ συμβαίνει πολλάκις νὰ εἶναι ἄγνωστον. Πρὸς τοῦτο ἐπιδιώκεται ἡ εὕρεσις μιᾶς ἐκτιμήσεως αὐτοῦ $\hat{P}_i$ ἐκ τῶν μονάδων ἑνὸς δείγματος. Ἔστω ὅτι εἰς δεῖγμα μεγέθους n εὑρέθησαν n_i μονάδες ἀνήκουσαι εἰς τὴν τάξιν A_i· τότε ἡ ἀναλογία $n_i/n = f_i$ συνιστᾷ μεταβλητήν, ἥτις κατανέμεται, ὡς γνωστόν, καθ' ὅμοιον τρόπον ὡς καὶ ἡ μεταβλητὴ $x/n = f$ (τοῦ διωνυμικοῦ νόμου) καὶ συνεπῶς ὁ λόγος $f_i = n_i/n$ ἀποτελεῖ ἀμερόληπτον ἐκτιμητὴν τοῦ $P_i = NP_i/N$ ἤτοι, $f_i = \hat{P}_i$**.

$$\text{Συνεπῶς,}\quad P_i = \frac{NP_i}{N} = \frac{\Sigma x_i}{N} = \overline{X}$$

Ομοίως διὰ τὸ δεῖγμα θὰ ἔχωμεν:

$$1 + 0 + 1 + 1 + 1 + 0 \cdots + 0 + 1 = nf_i = n_i$$

$$\text{καὶ}\quad f_i = \frac{n_i}{n} = \frac{\Sigma x_i}{n} = \overline{x}$$

$$\text{ἀλλὰ}\quad E(\overline{x}) = \overline{X}$$

$$\text{ὁπότε,}\quad E(f_i) = E\left(\frac{ni}{n}\right) = \Sigma\left\{\frac{n_i}{n} \cdot Pn_i(x_i)\right\} = P_i$$

---

*) Θεωροῦμεν δηλαδὴ ἐν προκειμένῳ ὡς συγκεκριμένην χαρακτηριστικὴν ἰδιότητα διακρίσεως τοῦ πληθυσμοῦ εἰς δύο τάξεις τὴν «νὰ ἀνήκῃ μία μονὰς εἰς τὴν τάξιν A_i ἢ ὄχι».

**) Ὑποθέσωμεν ὅτι εἰς ἑκάστην μονάδα τοῦ πληθυσμοῦ ἀντιστοιχεῖ καὶ μία τιμὴ τῆς μεταβλητῆς X ἥτις λαμβάνει τὰς τιμὰς 0 καὶ 1 (ὅπου $x_i = 1$ ἐὰν ἡ μονὰς τοῦ πληθυσμοῦ ἀνήκει εἰς τὴν τάξιν A_i) μὲ ἀντιστοίχους πιθανότητας

$$P_r(x_i = 1) = P_i \quad \text{καὶ} \quad P_r(x_i = 0) = 1 - P_i = q_i,$$

τότε τὸ σύνολον τῶν τιμῶν x_i τοῦ πληθυσμοῦ θὰ εἶναι:

$$1 + 0 + 1 + \cdots + 1 + 0 + 0 = NP_i = \Sigma x_i$$

διότι ὑπάρχουν τόσαι τιμαὶ $x_i = 1$, ὅσον εἶναι τὸ πλῆθος τῶν μονάδων τῆς τάξεως A_i.

$$\mathrm{var}\,(f_i) = \mathrm{var}\left(\frac{n_i}{n}\right) = E\left\{\left(\frac{n_i}{n} - P_i\right)^2\right\} =$$

$$= \Sigma\left\{\left(\frac{n_i}{n} - P_i\right)^2 \cdot Pn_i\,(x_i)\right\} =$$

$$= \frac{P_i\,(1 - P_i)}{n} = \frac{P_i q_i}{n}$$

ἤτοι f_i : ἀμερόληπτος ἐκτιμητὴς τοῦ P_i

καὶ $\hat{\sigma}^2_{p_i} = \dfrac{f_i\,(1 - f_i)}{n - 1}$: ἀμερόληπτος ἐκτιμητὴς τῆς $\sigma^2_{p_i}$ *

---

*) Πρὸς τὸν σκοπὸν ἐκτιμήσεως τοῦ γινομένου $P_i\, q_i$ θεωροῦμεν τὴν μεταβλητήν:

$$\left\{\frac{n_i}{n} \cdot \frac{n - n_i}{n}\right\} = f_i\,(1 - f_i) = (P_i - c)\,(q_i - c)$$

$c =$ θετικὸς ἢ ἀρνητικός

$$E\{f_i\,(1 - f_i)\} = E\{(P_i + c)\,(q_i - c)\} = E\,(P_i q_i + c\,q_i - c\,P_i - c^2) =$$

$$= P_i\,q_i + (q_i - P_i)\,E\,(c) - E\,(c^2) \qquad (\alpha)$$

Δοθέντος ὅμως ὅτι :

$$E\,(c) = E\,(f_i - P_i) = 0 \text{ καὶ } E\,(c^2) = E\{(f_i - P_i)^2 = \sigma^2_{p_i} = \frac{P_i\,q_i}{n},$$

δι' ἀντικαταστάσεως εἰς τὴν (α) θὰ ἔχωμεν :

$$E\,\{f_i\,(1 - f_i)\} = P_i\,q_i - \frac{P_i\,q_i}{n} = \frac{nP_i\,q_i}{n} - \frac{P_i\,q_i}{n} = \frac{n - 1}{n}\,P_i\,q_i$$

συνεπῶς ἡ μαθηματικὴ ἐλπὶς τοῦ γινομένου $f_i\,(1 - f_i)$ εἶναι διάφορος τοῦ

$$P_i\,(1 - P_i) = P_i\,q_i$$

ὅπερ σημαίνει ὅτι τὸ γινόμενον $f_i\,(1 - f_i)$ δὲν εἶναι ἀμερόληπτος ἐκτίμησις τοῦ γινομένου $P_i\,q_i$.

διὰ τοῦτο

$$E\left\{\frac{\frac{n}{n-1}\,f_i\,(1 - f_i)}{n}\right\} = \frac{P_i\,q_i}{n} = \sigma^2_{p_i}$$

ἢ $\qquad E\left[\dfrac{f_i\,(1 - f_i)}{n - 1}\right] = \dfrac{P_i\,q_i}{n} = \sigma^2_{p_i} \Longrightarrow \hat{\sigma}^2_{p_i} = \dfrac{f_i\,(1 - f_i)}{n - 1}$

III. Προσεγγιστική μορφή τοῦ πολυωνυμικοῦ νόμου

(Πολυμεταβλητὸς κανονικὸς νόμος)

Ἐπειδὴ ἡ σχέσις (1) ἐμφανίζει τὰς αὐτὰς δυσκολίας μὲ τὴν ἀντίστοιχον τοῦ διωνυμικοῦ νόμου, εἶναι ἀναγκαία ἡ ἀντικατάστασίς της δι' ἑτέρας, ἥτις νὰ δίδῃ ἀφ' ἑνὸς μέν, κατὰ προσέγγισιν, τὰ αὐτὰ ἀποτελέσματα, ἀφ' ἑτέρου δὲ νὰ προσφέρεται εἰς τὴν πρακτικήν. Πρὸς τὸν σκοπὸν τοῦτον καὶ ἔχοντες ὑπ' ὄψιν τὸν τύπον τοῦ Stirling ὅστις δίδει κατὰ προσέγγισιν τὴν τιμὴν τοῦ n!

$$n! \simeq (2\pi)^{1/2} \cdot n^{(n+1/2)} \cdot e^{-n} * \quad \text{θὰ ἔχωμεν}$$

$$P_{n_i}(x_i) \simeq \frac{(2\pi)^{1/2} \cdot n^{n+1/2} \cdot \exp(-n)}{\prod_{i=1}^{k} \{ (2\pi)^{1/2} \cdot n_i^{\,n_i+1/2} \cdot \exp(-n_i) \}} \cdot \prod_{i=1}^{k} P_i^{\,n_i}$$

$$= \frac{n^{n+1/2} \cdot \exp(-n)}{(2\pi)^{(k-1)/2} \cdot \exp\left(-\sum_{i=1}^{k} n_i\right)} \cdot \prod_{i=1}^{k} \{ P_i^{\,n_i} / n_i^{\,n_i+1/2} \}$$

Ἀλλὰ $\sum_{i=1}^{k} n_i = n$ καὶ διὰ τοῦτο,

$$P_{n_i}(x_i) \simeq \frac{n^{\left[\sum_{i=1}^{k}(n_i+1/2)\right]-(k-1)/2}}{(2\pi)^{(k-1)/2} \prod_{i=1}^{k} P_i^{1/2}} \cdot \prod_{i=1}^{k} \{ P_i/n_i \}^{n_i+1/2}$$

$$\text{ἢ} \quad P_{n_i}(x_i) \simeq \frac{1}{(2\pi n)^{(k-1)/2} \cdot \prod_{i=1}^{k} P_i^{1/2}} \cdot \prod_{i=1}^{k} \{ nP_i/n_i \}^{n_i+1/2}$$

Ἡ ποσότης $(2\pi n)^{(k-1)/2} \prod_{i=1}^{k} P_i^{1/2}$ εἶναι ἀνεξάρτητος τῶν n_i καὶ διὰ δεδομένον n καὶ δεδομένον σύνολον πιθανοτήτων P_i εἶναι σταθερά.

Ἐὰν τεθῇ $\frac{1}{C} = (2\pi n)^{(k-1)/2} \cdot \prod_{i=1}^{k} P_i^{1/2}$ ἢ $C = \frac{\sqrt{2\pi n}}{\prod_{i=1}^{k} \sqrt{2\pi n P_i}}$

καὶ $X_i = \frac{n_i - nP_i}{\sqrt{nP_i}}$ ἢ $n_i = nP_i + X_i \sqrt{nP_i}$.

$e^{-n} = \exp(-n)$

Ἐπειδὴ $\sum_{i=1}^{k} n_i = n = \sum_{i=1}^{k} nP_i$, $\sum_{i=1}^{k} X_i \sqrt{nP_i} = 0$, τοῦτο δεικνύει ὅτι μόνον $(k-1)$ ἐκ τῶν X_i εἶναι ἀνεξάρτητοι, ἄρα θὰ ἔχωμεν:

$$Pn_i(x_i) \simeq C \prod_{i=1}^{k} \left[\frac{nP_i}{X_i\sqrt{nP_i} + nP_i} \right]^{nP_i + X_i\sqrt{nP_i} + 1/2}$$

Λογαριθμίζοντες ἀμφότερα τὰ μέλη λαμβάνομεν,

$$\ln Pn_i(x_i) \simeq \ln C + \sum_{i=1}^{k} (nP_i + X_i\sqrt{nP_i} + 1/2) \ln \frac{nP_i}{X_i\sqrt{nP_i} + nP_i} =$$

$$= \ln C - \sum_{i=1}^{k} (nP_i + X_i\sqrt{nP_i} + 1/2) \ln \frac{X_i\sqrt{nP_i} + nP_i}{nP_i} =$$

$$= \ln C - \sum_{i=1}^{k} (nP_i + X_i\sqrt{nP_i} + 1/2) \ln \left(1 + \frac{X_i}{\sqrt{nP_i}} \right).$$

ἀναπτύσσοντες εἰς σειρὰν Taylor τὴν συνάρτησιν $\ln \left(1 + \frac{X_i}{\sqrt{nP_i}} \right)$ καὶ παραλείποντες τὰς δυνάμεις τοῦ $\frac{X_i}{\sqrt{nP_i}}$ τὰς μεγαλυτέρας τῆς δευτέρας, λαμβάνομεν:

$$\ln \left(1 + \frac{X_i}{\sqrt{nP_i}} \right) = \left(\frac{X_i}{\sqrt{nP_i}} - \frac{X_i^2}{2nP_i} \right)$$

Συνεπῶς,

$$\ln Pn_i(x_i) \simeq \ln C - \sum_{i=1}^{k} (nP_i + X_i\sqrt{nP_i} + 1/2) \left(\frac{X_i}{\sqrt{nP_i}} - \frac{X_i^2}{2nP_i} \right) =$$

$$= \ln C - \sum_{i=1}^{k} \left(X_i\sqrt{nP_i} + X_i^2 + \frac{X_i}{2\sqrt{nP_i}} - \frac{X_i^2}{2} - \right.$$

$$\left. - \frac{X_i^3}{2\sqrt{nP_i}} - \frac{X_i^2}{4nP_i} \right)$$

Ἐὰν τὸ μέγεθος τοῦ δείγματος εἶναι μέγα δυνάμεθα νὰ ἀπαλείψωμεν τοὺς ὅρους οἵτινες ἔχουν εἰς τὸν παρονομαστὴν $\sqrt{nP_i}$ καὶ τὰς δυνάμεις τούτου, ὅτε ἔχομεν:

$$Pn_i(x_i) - \ln C \simeq - \sum_{i=1}^{k} \left(\frac{X_i^2}{2} + X_i\sqrt{nP_i} \right) =$$

$$= - \frac{1}{2} \sum_{i=1}^{k} X_i^2 - \sum_{i=1}^{k} X_i \sqrt{nP_i}$$

καὶ ἐπειδὴ $\sum_{i=1}^{k} X_i \sqrt{nP_i} = \sum_{i=1}^{k} \frac{n_i - nP_i}{\sqrt{nP_i}} \cdot \sqrt{nP_i} = 0$

$$\ln Pn_i(x_i) - \ln C \simeq -\frac{1}{2}\sum_{i=1}^{k} X_i^2$$

ἢ $$\ln\left(\frac{Pn_i(x_i)}{C}\right) \simeq -\frac{1}{2}\sum_{i=1}^{k} X_i^2$$

καὶ $$Pn_i(x_i) \simeq C \cdot e^{-1/2\sum_{i=1}^{k} X_i^2}$$

ἢ $$Pn_i(x_i) = \frac{1}{C}\exp\left(-\frac{1}{2}\sum_{i=1}^{k} X_i^2\right) \qquad (2)$$

Οὕτω κατελήξαμεν εἰς τὸν Πολυμεταβλητὸν Κανονικὸν Νόμον, ὅστις εἶναι γενίκευσις τοῦ κανονικοῦ.

Αἱ n_i εἶναι ἀκέραιοι ἀριθμοὶ καὶ ἐπειδὴ

$$n_i = nP_i + X_i (nP_i)^{1/2}$$

εἰς μίαν κατὰ μονάδα μεταβολὴν τῆς n_i, ἡ ἀντίστοιχος μεταβολὴ εἰς τὴν X_i, ΔX_i θὰ εἶναι $\Delta X_i = \frac{1}{\sqrt{nP_i}}$. Ἐνθυμούμενοι ἐπίσης ὅτι μόνον $(k-1)$ μεταβληταὶ $X_i = \frac{n_i - nP_i}{\sqrt{nP_i}}$ εἶναι ἀνεξάρτητοι, θὰ ἔχωμεν :

$$Pn_i(x_i) = \frac{1}{(2\pi n)^{(k-1)/2} \cdot \prod_{i=1}^{k} P_i^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} X_i^2\right)$$

$$= \frac{1}{(2\pi)^{(k-1)/2} \cdot \prod_{i=1}^{k} (nP_i)^{1/2} \cdot P_k^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} X_i^2\right)$$

$$= \frac{1}{(2\pi)^{(k-1)/2} \cdot P_k^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} X_i^2\right) \Delta X_1 \cdot \Delta X_2 \cdots \Delta X_{k-1}.$$

καὶ ὅταν $n \to \infty$ ἡ πιθανότης ἵνα ἡ μεταβλητὴ X (ἀντιστοιχοῦσα εἰς τὴν συγκεκριμένην χαρακτηριστικὴν ἰδιότητα τοῦ πληθυσμοῦ) λάβῃ τὴν τιμὴν x_1 n_1 φοράς, τὴν τιμὴν x_2 n_2 φορὰς κ.ο.κ. δίδεται κατὰ προσέγγισιν ἀπὸ τὴν στοιχειώδη πιθανότητα (ἢ στοιχεῖον πιθανότητος) εἰς τὸν χῶρον τῶν k – διαστάσεων μιᾶς συνεχοῦς κατανομῆς ὁριζομένης ὑπό :

$$P_{n_i}(x_i) \simeq dP = B\exp\left(-\frac{1}{2}\sum_{i=1}^{k} X_i^2\right) dX_1 \cdot dX_2 \cdot \cdot \cdot dX_{k-1} \cdot \quad (3)$$

$$\sum_{i=1}^{k} X_i^2 = \sum_{i=1}^{k} \frac{(n_i - nP_i)^2}{nP_i} = \frac{(n_1 - nP_1)^2}{nP_1} + \frac{(n_2 - nP_2)^2}{nP_2} + \cdot \cdot \cdot + \frac{(n_k - nP_k)^2}{nP_k} = x^2$$

ἡ ἐξίσωσις αὕτη παριστᾷ μίαν ὑπερσφαῖραν ἀκτῖνος x. Ἐὰν τὸ x αὐξηθῇ κατὰ dx ἤτοι $x + dx$, τότε θὰ ἔχωμεν δύο γειτονικὰς ὑπερσφαίρας ἀκτίνων x καὶ $x + dx$ ἀντιστοίχως. Τὸ χωρίον δὲ τὸ ὁποῖον θὰ περιλαμβάνεται μεταξὺ αὐτῶν θὰ ἔχῃ ὄγκον ἀνάλογον πρὸς $dx_1 \cdot dx_2 \cdot \cdot \cdot dx_{k-1} = d(x^{k-1}) = (k-1)\, x^{k-2}\, dx$. Συνεπῶς, ἡ πιθανότης ἵνα μεταξὺ τῶν τιμῶν τῆς μεταβλητῆς X, n_i ἔχουν τὴν τιμὴν x_i (ὅπου x_i ἡ κεντρικὴ τιμὴ ἑνὸς ταξικοῦ διαστήματος $x_i \pm \frac{1}{2} h_i$) εἶναι προσεγγιστικῶς ἴση πρὸς τὴν πιθανότητα, ὅτι $x = \left(\sum_{i=1}^{k} X_i^2\right)^{1/2}$ εὑρίσκεται μεταξὺ τῶν x καὶ $x + dx$.

$$\text{οὕτω} \quad dP = A\exp\left(-\frac{1}{2}x^2\right) \cdot x^{k-2} \cdot dx$$

καὶ ἐπειδὴ ἡ πιθανότης ἵνα ἡ X λάβῃ τιμὰς μεταξὺ τοῦ 0 καὶ τοῦ ∞ εἶναι μοναδιαία

$$1 = A\int_0^\infty \exp\left(-\frac{1}{2}x^2\right) x^{k-2} \cdot dx$$

$$\frac{1}{A} = \int_0^\infty \exp\left(-\frac{1}{2}x^2\right) \cdot x^{k-2}\, dx$$

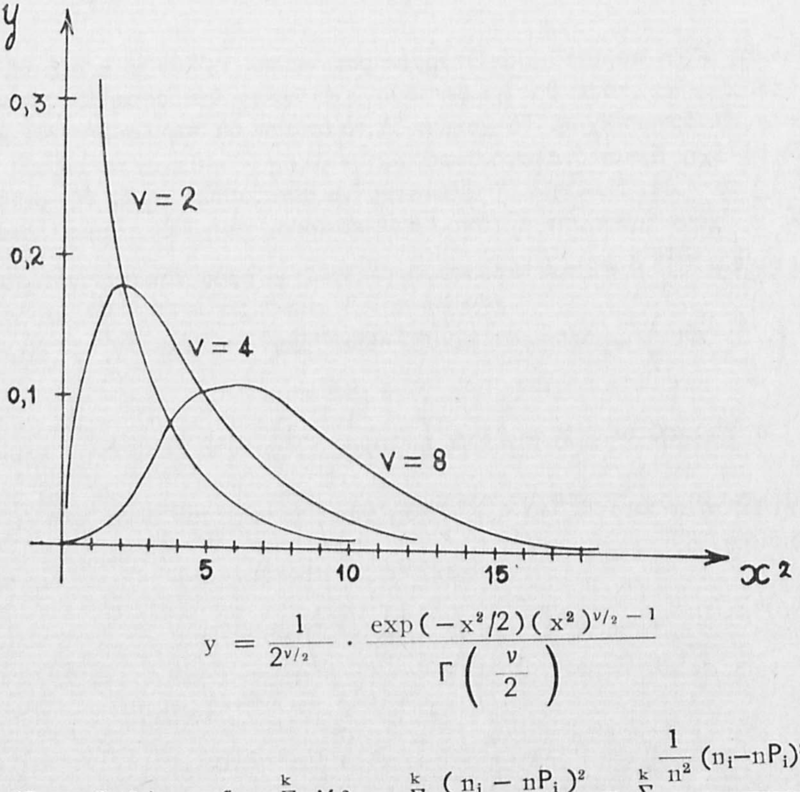
Αὕτη ὁρίζει μίαν συνεχῆ κατανομήν, ἥτις πολλάκις ἀποκαλεῖται x^2–κατανομὴ (ἐσφαλμένως).

Ἐπειδὴ ἐκ τῶν k μεταβλητῶν X_i μόνον $k-1$ εἶναι ἀνεξάρτητοι, λέγομεν ὅτι ἡ x ἔχει $k-1$ βαθμοὺς ἐλευθερίας. Θέτοντες $\nu = k-1$ λαμβάνομεν:

$$dP = \frac{1}{2^{(\nu-2)/2}\,\Gamma\left(\frac{\nu}{2}\right)} \cdot x^{\nu-1} \cdot \exp\left(-\frac{1}{2}x^2\right) dx$$

Ἡ κατανομὴ x^2 προκύπτει ἐκ τῆς ἀνωτέρω, ἐὰν γραφῇ ὑπὸ τὴν μορφὴν

$$dP = \frac{1}{2^{\nu/2}\,\Gamma\left(\frac{\nu}{2}\right)} \cdot \exp\left(-\frac{1}{2}x^2\right)(x^2)^{\nu/2-1} \cdot dx^2$$



$$y = \frac{1}{2^{\nu/2}} \cdot \frac{\exp(-x^2/2)(x^2)^{\nu/2-1}}{\Gamma\left(\frac{\nu}{2}\right)}$$

Ἐκ τοῦ λόγου ὅτι $\sum\limits_{i=1}^{k} X_i^2 = \sum\limits_{i=1}^{k} \frac{(n_i - nP_i)^2}{nP_i} = \sum\limits_{i=1}^{k} \frac{\frac{1}{n^2}(n_i - nP_i)^2}{nP_i/n^2} =$

$$\sum_{i=1}^{k} \frac{\left(\frac{n_i}{n} - \frac{nP_i}{n}\right)^2}{nP_i/n^2} = \sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i/n} = n \sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}$$

ὁ πολυμεταβλητὸς κανονικὸς νόμος δύναται νὰ γραφῇ καὶ ὡς :

$$Pn_i(x_i) \simeq C \cdot e^{-\frac{1}{2} n \sum\limits_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}} \qquad (4)$$

Συμπεραίνομεν ὅθεν ὅτι ἡ συχνότης τῶν δειγμάτων τὰ ὁποῖα περιέχουν ἐξ ἑκάστης τάξεως

$$A_1,\ A_2,\ A_3,\ \ldots,\ A_i,\ \ldots,\ A_k$$

ἀντιστοίχως :

$$n_1,\ n_2,\ n_3,\ \ldots,\ n_i,\ \ldots,\ n_k \text{ μονάδας,}$$

ἢ ὅπερ τὸ αὐτὸ ἀντίστοιχα ποσοστὰ μονάδων

$$f_1 = \frac{n_1}{n},\ f_2 = \frac{n_2}{n},\ \ldots,\ f_i = \frac{n_i}{n},\ \ldots,\ f_k = \frac{n_k}{n},$$

ἐξαρτᾶται ἀπὸ τὸ ἄθροισμα τῶν τετραγώνων τῶν τιμῶν τῆς μεταβλητῆς $X_i = \dfrac{n_i - nP_i}{\sqrt{nP_i}}$ ἤτοι, ἀπὸ τὴν ποσότητα

$$\sum_{i=1}^{k} X_i^2 = \sum_{i=1}^{k} \frac{(n_i - nP_i)^2}{nP_i} = X_1^2 + X_2^2 + \cdots + X_k^2.$$

Ἐὰν τὸ μέγεθος τοῦ δείγματος n εἶναι μεγάλος ἀριθμὸς τότε αἱ μεταβληταὶ $X_i = (n_i - nP_i) / \sqrt{nP_i}$ εἶναι προσεγγιστικῶς κανονικαὶ μὲ μέσον μηδὲν καὶ διακύμανσιν μοναδιαίαν $N(0,1)$, τὸ δὲ ἄθροισμα τῶν τετραγώνων τῶν ἀνωτέρω μεταβλητῶν ἔχει μίαν κατανομὴν χ^2 μὲ $(k-1)$ βαθμοὺς ἐλευθερίας ἤτοι $\sum_{i=1}^{k} X_i^2 = \chi^2$.

Ἔχοντες ὑπ' ὄψιν τὰ ἀνωτέρω, ἀφ' ἑτέρου δὲ ὅτι ἡ ἀπόστασις δύο διανυσμάτων (σημείων) Α καὶ Β, συμβολικῶς $|A - B|$, ἐκφράζεται διὰ τοῦ μήκους τοῦ διανύσματος $(A - B)$: ἤτοι $|A - B| = |(A-B)' \cdot (A-B)|^{1/2} = \left[\sum_{i=1}^{k} (A_i - B_i)^2 \right]^{1/2}$ * καὶ συνεπῶς, διὰ τὴν περίπτωσίν μας, ὅπου τὸ διάνυσμα Β εἶναι τὸ μηδενικὸν ἤτοι, τὸ ἔχον ὡς ἀρχὴν καὶ πέρας τὴν ἀρχὴν τοῦ συστήματος τῶν συντεταγμένων καὶ τυχοῦσαν κατεύθυνσιν, καὶ τὸ

$$A = (X_1,\ X_2,\ \ldots,\ X_k) \text{ θὰ ἔχωμεν :}$$

$$(|A - B|)^2 = \sum_{i=1}^{k} (A_i - B_i)^2 = \sum_{i=1}^{k} X_i^2.$$

Ἐκ τῶν ἀνωτέρω προκύπτει ὅτι ἕκαστον δεῖγμα ἐκ τῶν δυνατῶν τοιού-

---

*) $(A - B)'$: Εἶναι τὸ ἐνηλλαγμένον διάνυσμα τοῦ $(A - B)$.

των, δύναται νὰ ἀναπαρασταθῇ ὡς σημεῖον εἰς τὸν χῶρον τῶν k — διαστάσεων μὲ συντεταγμένας τὰς τιμὰς τῆς X_i, ἃς ἕκαστον δεῖγμα λαμβάνει καὶ ὅτι ἡ πιθανότης $P_{n_i}(x_i)$ ἐξαρτᾶται ἀπὸ τὴν ἀπόστασιν ἐκ τῆς ἀρχῆς.

Ἐπειδὴ δέ,

$$x^2 = \sum_{i=1}^{k} X_i^2 = \sum_{i=1}^{k} \frac{(n_i - nP_i)^2}{nP_i} = n \sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}$$

ἕπεται ὅτι θὰ εἶναι $x^2 = \sum_{i=1}^{k} X_i^2 = 0$, ἐφ' ὅσον $\sum_{i=1}^{k} (n_i - nP_i)^2 = 0$ ἢ

ὅπερ τὸ αὐτὸ $\sum_{i=1}^{k} (f_i - P_i)^2 = 0$.

Ἀλλὰ διὰ νὰ εἶναι $\sum_{i=1}^{k} (f_i - P_i)^2 = 0$ δέον ὅπως $f_i - P_i = 0$. Καὶ ἐπειδὴ P_i εἶναι τὸ ποσοστὸν τῶν μονάδων τῆς τάξεως A_i εἰς τὸν πληθυσμόν, συμπεραίνομεν ὅτι $(f_i - P_i)^2 = 0$, ἐὰν τὸ ληφθὲν δεῖγμα ἔχει ποσοστὸν μονάδων τῆς τάξεως A_i ἀκριβῶς τὸ ἀντίστοιχον ποσοστὸν τοῦ πληθυσμοῦ, ἤτοι, κατ' ἄλλην διατύπωσιν, ἐὰν τὸ δεῖγμα εἶναι ἀπολύτως ἀντιπροσωπευτικόν.

Τὰ δείγματα ἑπομένως τὰ ὁποῖα ἀναπαρίστανται διὰ τοῦ μηδενικοῦ διανύσματος τοῦ ἔχοντος ἀρχὴν καὶ πέρας τὴν ἀρχὴν τοῦ συστήματος τῶν συντεταγμένων καὶ τυχοῦσαν κατεύθυνσιν, εἶναι ἀπολύτως ἀντιπροσωπευτικὰ τοῦ πληθυσμοῦ, τὰ δὲ ὑπόλοιπα ἀναλόγως ἐὰν ἀναπαρίστανται ὑπὸ διανυσμάτων ἅτινα ἔχουν ἀπόστασιν ἀπὸ τοῦ μηδενικοῦ, μικροτέραν ἢ μεγαλυτέραν εἶναι ἀντιστοίχως περισσότερον ἢ ὀλιγώτερον ἀντιπροσωπευτικά.

Ἐξ ὅσων ἤδη ἐλέχθησαν εἶναι προφανὲς ὅτι δύο δείγματα, ἀνεξαρτήτως μεγέθους, λαμβανόμενα ἐκ τοῦ αὐτοῦ πληθυσμοῦ, θὰ εἶναι ἰσοδύναμα ἀπὸ ἀπόψεως ἀντιπροσωπευτικότητος, ἐὰν καὶ μόνον ἐὰν τὰ διανύσματα ἅτινα ἀναπαριστοῦν ταῦτα εἶναι ἴσα, ἤτοι ἔχουν τὰ ἀντίστοιχα στοιχεῖα αὐτῶν ἀνὰ ἓν ἴσα. Τοῦτο ὅμως θὰ συμβῇ ἐὰν τὰ δύο δείγματα ἔχουν τὸ αὐτὸ ποσοστὸν f_i ($i = 1,2, \cdots k$) μονάδων τῆς τάξεως A_i ($1,2, \cdots k$).

Δύο μὴ ἰσοδύναμα ἀπὸ ἀπόψεως ἀντιπροσωπευτικότητος δείγματα (προερχόμενα ἐκ τοῦ αὐτοῦ πληθυσμοῦ) θὰ ἀναπαρίστανται ὑπὸ διανυσμάτων ἅτινα θὰ διαφέρουν εἰς ἓν ἢ περισσότερα στοιχεῖα, διὰ τὰ ὁποῖα διακρίνομεν τὰς ἑξῆς περιπτώσεις :

α) Ἅπαντα τὰ στοιχεῖα τοῦ ἑνὸς διανύσματος, ἔστω Α, εἶναι μεγαλύτερα ἀπὸ τὰ ἀντίστοιχα στοιχεῖα τοῦ ἑτέρου διανύσματος Β (ἤτοι $A > B$) ὁπότε, ἀπὸ ἀπόψεως ἀντιπροσωπευτικότητος, τὸ δεῖγμα τὸ ὁποῖον ἀναπαρίσταται ὑπὸ τοῦ διανύσματος Β εἶναι καλλίτερον τοῦ δείγματος τοῦ ἀναπαρισταμένου ὑπὸ τοῦ διανύσματος Α.

β) Οὐδὲν στοιχεῖον τοῦ διανύσματος Α εἶναι μικρότερον τοῦ ἀντιστοίχου στοιχείου τοῦ διανύσματος Β καὶ τουλάχιστον ἓν στοιχεῖον τοῦ πρώτου (Α) εἶναι μεγαλύτερον τοῦ ἀντιστοίχου στοιχείου τοῦ δευτέρου διανύσματος (Β), ἤτοι $A \geqslant B$. Τότε λέγομεν ἐπίσης ὅτι, τὸ δεῖγμα τὸ ὁποῖον ἀναπαρίσταται ὑπὸ τοῦ διανύσματος Β εἶναι ἀντιπροσωπευτικώτερον τοῦ δείγματος τοῦ ἀναπαρισταμένου ὑπὸ τοῦ διανύσματος Α.

γ) Τέλος, δυνατὸν ἄλλα μὲν στοιχεῖα τοῦ ἑνὸς διανύσματος νὰ εἶναι μικρότερα, ἄλλα δὲ μεγαλύτερα τῶν ἀντιστοίχων στοιχείων τοῦ ἑτέρου διανύσματος. Ἐν τοιαύτῃ περιπτώσει δὲν ὑφίσταται συγκρισιμότης ἀπὸ ἀπόψεως ἀντιπροσωπευτικότητος.

Ἐὰν ἀμφότερα τὰ μέλη τῆς σχέσεως:

$$x^2 = n \sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}$$

διαιρέσωμεν διὰ n λαμβάνομεν:

$$\frac{x^2}{n} = \sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}$$

ἤτοι, ἡ ἀπόστασις σημείου τινὸς ἀπὸ τῆς ἀρχῆς τῶν ἀξόνων, ἐκφράζεται εἰς μονάδας n.

Ἡ ἀντιπροσωπευτικότης ἑπομένως δειγμάτων, ἢ ὅπερ τὸ αὐτό, ἡ ἀπόστασις τῶν δειγμάτων (σημείων), ἅτινα ἔχουν συντεταγμένας f_i $(i = 1,2, \cdots k)$ ἀπὸ τὴν ἀρχὴν τῶν ἀξόνων δίδεται ἀπὸ τὴν σχέσιν:

$$\frac{x}{\sqrt{n}} = \sqrt{\sum_{i=1}^{k} \frac{(f_i - P_i)^2}{P_i}} \tag{5}$$

Ὅταν τὸ μέγεθος τοῦ δείγματος n τείνῃ εἰς τὸ ἄπειρον $(n \to \infty)$, ἡ ποσότης $\frac{x}{\sqrt{n}}$ τείνει εἰς τὸ μηδέν, γεγονὸς ὅπερ σημαίνει ὅτι αὐξανομένου τοῦ μεγέθους τοῦ δείγματος ἡ ἀντιπροσωπευτικότης τῶν δειγμάτων αὐξάνει. Τοῦτο σημαίνει ὅτι ὁσονδήποτε μικρὰ καὶ ἂν ἐπιθυμοῦμεν νὰ εἶναι ἡ ἀπόστασις (μεγάλη ἡ ἀντιπροσωπευτικότης) ἑνὸς δείγματος ἀπὸ τὴν ἀρχήν, δὲν ἔχομεν εἰμὴ νὰ λάβωμεν δεῖγμα τοσοῦτον μέγα ὥστε τὸ σύνολον τῶν δειγμάτων καὶ ἑπομένως καὶ τὸ ζητούμενον, νὰ ἔχῃ ἀπόστασιν ἀπὸ τὸ μηδὲν οὐχὶ μεγαλυτέραν τοῦ δοθέντος σημείου (ἀντιπροσωπευτικότης ἴση ἢ μεγαλυτέρα τοῦ δοθέντος δείγματος).

Εἰς τὰς πρακτικὰς ἐφαρμογὰς ἀπομονοῦται μία τῶν τάξεων, πασῶν τῶν λοιπῶν θεωρουμένων ὅτι συνθέτουν δευτέραν τάξιν, ἀντιμετωπιζομένων οὕτω τῶν σχετικῶν δειγματοληπτικῶν προβλημάτων ὡς ἐὰν ὁ ὑπὸ ἐξέτασιν πληθυσμὸς νὰ διεκρίνετο ὑπὸ τὴν ἔποψιν κριτηρίου τινὸς (συγκεκριμένης χαρακτηριστικῆς ἰδιότητος) εἰς δύο τάξεις.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ἑλληνική

1. Θεοδωράκη, Γερ.: Εἰσαγωγὴ εἰς τὴν Μαθηματικὴν Στατιστικήν, Ἀθῆναι, 1970.
2. Κανέλου, Σπ.: Εἰσαγωγὴ εἰς τὴν Θεωρίαν τῶν Πιθανοτήτων, Ἀθῆναι. 1959.
3. Κιόχου, Π.: Λογισμὸς Πιθανοτήτων, Ἀθῆναι, 1970.
4. Μαργαρίτη, Εὐστ.: Στατιστική, τόμος ΙΙ.
5. Παναγιωτοπούλου, Α.: Μαθήματα Πιθανοτήτων, Τεῦχος Ι. Πειραιεύς, 1971.
6. Σαραντοπούλου, Σπ.: Λογισμὸς τῶν Πιθανοτήτων καὶ Στατιστικὴ (τόμος Α.', Ἀθῆναι, 1960).

Ξένη

1. Alexander, Howard, W.: Elements of Mathematical Statistics, N. Y., J. Wiley. Inc., 1960.
2. Anderson: An Introduction to Multivariate Statistical Analysis.
3. Apostol, T.: Διαφορικὸς καὶ ὁλοκληρωτικὸς λογισμός, Τόμοι Ι, ΙΙ.
4. Cochran, W.: Sampling techniques, J. Wiley, N.Y., 1958.
5. Cramér, H.: The Elements of Probability Theory, J. Wiley, N.Y., 1958.
6. Cramér, H.: Mathematical Methods of Statistics, Princeton University Press, 1951.
7. Feller, W.: An Introduction to Probability Theory and its Applications vol I, II, J. Wiley, N.Y.
8. Goldberg, S.: Probability, ἑλληνικὴ μετράφρασις ὑπὸ Κ. Παναγόκη, Ἀθῆναι, 1962.
9. Hoel, P. G.: Introduction to Mathematical Statistics, 2d ed., N. Y, J. Wiley, Inc., 1954.
10. Toeplitj, O.: The Calculus, Chicago, 1963.
11. Wilks, S.: Mathematical Statistics, J. Wiley. N. Y. 1962.
12. Wilks, S: Elementary Statistical Analysis, Princeton, University Press, N. J. 1948.
13. Ya-lun Chou Statistical Analysis, N. Y.