ΣΤΑΤΙΣΤΙΚΑΙ ΜΕΛΕΤΑΙ

ΕΝΑ ΑΠΛΟΥΣΤΑΤΟ ΤΕΣΤ ΤΥΧΑΙΟΤΗΤΟΣ ΔΕΙΓΜΑΤΟΣ

Ὑπὸ Δρος ΑΛ. Ι. ΠΑΠΠΑ

Τακτικοῦ Καθηγητοῦ τοῦ Ε. Μ. Πολυτεχνείου καὶ τοῦ Κέντρου Στατιστικῆς Ἐκπαιδεύσεως

Ἔχουν προταθῆ ἀρκετὰ στατιστικὰ τέστ διὰ τὸν ἔλεγχον τυχαιότητος δείγματος.

Κατωτέρω προτείνω ἕνα νέον, τὸ ὁποῖον ἐνῶ εἶναι ἁπλούστερον εἶναι καὶ ἀκριβέστερον τῶν, εἰς ἐμὲ τουλάχιστον, γνωστῶν.

Ἔστω δεῖγμα ἐκ N τιμῶν, τὸ ὁποῖον θέλομε νὰ ἐλέγξωμε ἂν εἶναι τυχαῖο, καὶ M ἡ κεντρικὴ ἢ μεσαία τιμὴ τῶν τιμῶν αὐτῶν.

Ἐὰν οἱ N τιμὲς γραφοῦν κατὰ τὴν σειρὰ ποὺ προέκυψαν καὶ καλέσωμε A κάθε τιμὴ ποὺ εἶναι μικρότερη ἀπὸ τὴν M καὶ B κάθε τιμὴ μεγαλύτερη ἀπὸ τὴν M, παραλείποντας τὴν τιμὴ ποὺ ἐνδεχομένως θὰ εἶναι ἴση μὲ τὴν M, θὰ προκύψη μία ἀκολουθία ἀπὸ N ἢ N—1 γράμματα A ἢ B.

Ἐὰν τὸ δεῖγμα εἶναι τυχαῖο ἡ πιθανότης μετὰ μία τιμὴ A (μικρότερη ἀπὸ τὴν μεσαία τιμὴ M) νὰ ἐμφανισθῆ μία τιμὴ B (μεγαλύτερη ἀπὸ τὴν M) εἶναι $1/2$ καὶ ἡ πιθανότης μετὰ μία A νὰ ἐμφανισθῆ πάλι μία τιμὴ A εἶναι ἐπίσης $1/2$, ἀφοῦ ὁ ἀριθμὸς τιμῶν ποὺ εἶναι μικρότερες ἀπὸ τὴν M ἰσοῦται μὲ τὸν ἀριθμὸ τιμῶν ποὺ εἶναι μεγαλύτερες ἀπὸ τὴν M.

Ὁμοίως ἡ πιθανότης τῆς περιπτώσεως BA εἶναι $1/2$ καὶ τῆς BB πάλι $1/2$.

Γενικώτερα, ἐφ' ὅσον τὸ δεῖγμα εἶναι τυχαῖο, ἡ πιθανότης τῆς περιπτώσεως εἴτε AB εἴτε BA, ποὺ θὰ καλέσωμε ἑτερόσημη διαδοχή, εἶναι $1/2$ καὶ ἴση μὲ τὴν πιθανότητα ὁμόσημης διαδοχῆς ὅπως θὰ καλέσωμε τὴν περίπτωση εἴτε AA εἴτε BB.

Ἐὰν καλέσωμε n τὸν ὅλο ἀριθμὸ τῶν διαδοχῶν σὲ μία ἀκολουθία N ἢ N—1 γραμμάτων A ἢ B θὰ εἶναι :

$$\text{εἴτε } n = N-1 \qquad \text{εἴτε } n = N-2$$

Ἐὰν δὲ ὀνομάσωμε κ τὸν ἀριθμὸ τῶν ὁμοσήμων καὶ λ τὸν ἀριθμὸ τῶν ἑτεροσήμων διαδοχῶν, ἡ πιθανότης νὰ ἔχωμε κ ὁμόσημες διαδοχὲς καὶ λ ἑτερόσημες, σύμφωνα μὲ τὴν διωνυμικὴ κατανομή, εἶναι ἴση μέ :

$$(1) \qquad C_\kappa^n \left(\frac{1}{2}\right)^\kappa \left(\frac{1}{2}\right)^\lambda = C_\kappa^n \left(\frac{1}{2}\right)^n \quad \text{ἀφοῦ } \kappa + \lambda = n.$$

Ἡ διασπορὰ ὅμως τῶν τιμῶν κ μπορεῖ νὰ θεωρηθῆ κατὰ προσέγγισιν

κανονικὴ μὲ μέση τιμή : $m_\kappa = np = \frac{1}{2} n$ καὶ

τυπικὴ ἀπόκλισι : $\sigma_\kappa = \sqrt{npq} = \frac{1}{2}\sqrt{n}$

᾿Αφοῦ δὲ $\lambda = n - \kappa$ ἄρα $\kappa - \lambda = \kappa - n + \kappa = 2\kappa - n$ ἕπεται ὅτι καὶ ἡ διασπορὰ τῆς διαφορᾶς $\kappa - \lambda$ μπορεῖ ἐπίσης νὰ θεωρηθῆ κανονικὴ μέ:

μέση τιμή: $$m_{\kappa-\lambda} = 2m_\kappa - n = 2\frac{1}{2}n - n = 0$$

τυπικὴ ἀπόκλισι: $$\sigma_{\kappa-\lambda} = 2\sigma_\kappa = \sqrt{n}$$

῎Αρα μὲ πιθανότητα $1 - \alpha = 95\%$ ἡ διαφορὰ $(\kappa - \lambda)$, ἐὰν τὸ δεῖγμα εἶναι τυχαῖο, πρέπει νὰ περιλαμβάνεται στὰ ὅρια

$$0 \pm 1{,}96\sqrt{n} \quad \text{ἢ στρογγυλὰ} \quad 0 \pm 2\sqrt{n}$$

῾Επομένως ἐάν: (2) $$|\kappa - \lambda| \leqslant 2\sqrt{n}$$

θὰ συμπεράνωμε ὅτι, στὴν στάθμη σημαντικότητος $\alpha = 5\%$, δὲν διαψεύδεται ἡ ὑπόθεσις ὅτι τὸ δεῖγμα εἶναι τυχαῖο εἰδεμὴ τὸ ἐναντίον.

᾿Εὰν ὡς στάθμη σημαντικότητος ληφθῆ τὸ $\alpha = 1\%$ τότε ἡ ὑπόθεσις ὅτι τὸ δεῖγμα εἶναι τυχαῖο διαψεύδεται ἐάν:

$$|\kappa - \lambda| > 2{,}5\sqrt{n}$$

ἢ ἀντίστοιχα διὰ ὁποιαδήποτε ἄλλη στάθμη σημαντικότητος.

ΠΑΡΑΔΕΙΓΜΑ 1ον

᾿Απὸ μιὰ σειρὰ δοκιμῶν προέκυψαν, κατὰ τὴν ἀκόλουθη τάξι, οἱ τιμές:

(3) 20, 24, 17, 18, 22, 14, 16, 21, 22, 27, 15, 15

῾Η κεντρικὴ ἢ μεσαία τιμὴ τῶν $N = 12$ αὐτῶν τιμῶν εἶναι ἡ $M = 19$, ποὺ εἶναι ὁ μέσος ὅρος τῶν δύο μεσαίων, ἂν τὶς γράψομε κατὰ τάξι μεγέθους:

(4) 14, 15, 15, 16, 17, 18, 20, 21, 22, 22, 24, 27

᾿Απὸ τὴν ἀκολουθία (3), σύμφωνα μὲ ὅσα εἴπαμε, προκύπτει ἡ ἑξῆς ἀκολουθία γραμμάτων Α ἢ Β:

(5) Β Β Α Α Β Α Α Β Β Β Α Α

| | | | | | = 6 ὁμόσημες διαδοχὲς

| | | | | = 5 ἑτερόσημες διαδοχὲς

ἄρα $|\kappa - \lambda| = |6 - 5| = 1$ καὶ $2\sqrt{n} = 2\sqrt{11} = 7{,}93$

ἀφοῦ δὲ ἐδῶ $|\kappa - \lambda| < 2\sqrt{n}$

ἕπεται ὅτι δὲν διαψεύδεται ἡ ὑπόθεσις ὅτι τὸ δεῖγμα, ποὺ ἀποτελοῦν οἱ τιμὲς (3), εἶναι τυχαῖο.

ΠΑΡΑΔΕΙΓΜΑ 2ον

᾿Εὰν ἀπὸ ἕνα δεῖγμα προέκυψαν κατὰ σειρὰν οἱ τιμές:

(6)
44, 45, 51, 43, 37, 41, 41, 39, 42 42, 50
45, 49, 47, 45, 38, 42, 40, 41, 48, 46

καὶ τὶς γράψομε κατὰ τάξι μεγέθους, θὰ ἔχωμε τὸ ἀκόλουθο διατεταγμένο δεῖγμα :

$$(7) \qquad \begin{array}{l} 37,\ 38,\ 39,\ 40,\ 41,\ 41,\ 41,\ 42,\ 42,\ 42,\ 43 \\ 44,\ 45,\ 45,\ 45,\ 46,\ 47,\ 48,\ 49,\ 50,\ 51 \end{array}$$

τοῦ ὁποίου ἡ μεσαία τιμὴ εἶναι ἡ ἑνδεκάτη $M = 43$.

Ἄρα ἡ ἀκολουθία (6), παραλειπομένης τῆς τιμῆς 43, γράφεται:

```
     B B B A A A A A A B B B B B A A A A B B
(8)   | |   | | | | |   | | | |   | | |   |    ἤτοι κ = 15
         |         |         |       |         »  λ =  4
```

ἄρα ἀφοῦ ἔχομε $n = \kappa + \lambda = 19$ ἡ δὲ διαφορὰ $|\kappa - \lambda| = 15 - 5 = 11$ εἶναι μεγαλύτερη τοῦ

$$2\sqrt{n} = 2\sqrt{19} = 8{,}7$$

ἕπεται ὅτι τὸ δεῖγμα δὲν μπορεῖ νὰ θεωρηθῆ τυχαῖο μὲ τὴν στάθμη σημαντικότητος $\alpha = 5\,\%$ ποὺ δεχθήκαμε.

Π α ρ α τ ή ρ η σ ι ς. Μὲ τὸ τὲστ αὐτό, ὅπως καὶ μὲ ὅλα τὰ τὲστ τυχαιότητος, δὲν ἀποδεικνύεται βέβαια ἂν ἕνα δεῖγμα εἶναι ἢ ὄχι τυχαῖο. Ὅμως ἡ μὴ ἐκπλήρωσις τῆς συνθήκης ποὺ τίθεται ἀποτελεῖ σαφῆ ἔνδειξι μὴ τυχαιότητος. Πρέπει δὲ ἀκόμη νὰ τονισθῆ ὅτι τὸ συμπέρασμα ἐξαρτᾶται ἀπὸ τὴν στάθμη σημαντικότητος ποὺ θὰ γίνη παραδεκτή, ὅπως συμβαίνει μὲ ὅλα τὰ στατιστικὰ τέστ.