

LETTERS TO THE EDITOR

| | PAGE | | PAGE |
|--|------|---|------|
| Improved Formulæ for the Coefficient of Contingency. BY A. A. KRISHNASWAMI AYYANGAR | 51 | On Some Squilla Larvæ from the Madras Plankton. BY K. H. ALIKUNHI AND R. GOPALA AIYAR | 56 |
| Absorption Spectra of Potassium Permanganate in Different Media. BY M. PRABHANJANA MURTHY AND JAGDEO SINGH | 52 | Chromosome Number in Carica. BY L. S. S. KUMAR AND A. ABRAHAM .. | 58 |
| A Sample of Shark Liver Oil Unusually Rich in Vitamin A. BY K. RAJAGOPAL | 52 | A Probable Case of Translocation during Mitosis Involving the Satellite Thread. BY K. V. SRINATH | 59 |
| Studies in Insect Nutrition—Assay of 'Quality' in Crops. BY S. RAMASWAMI, P. S. SARMA AND M. SREENIVASAYA .. | 53 | Two New Species of Aspiromitus St. from Bor-Ghat (Lonavala & Khandala). BY V. V. APTE AND P. V. SANE | 59 |
| Celestite in Cis-Indus Salt Range. BY B. S. LAMBA | 54 | Some Observations on an "Improved Method for the Determination of Prothrombin Time". BY D. V. S. REDDY AND C. VENKATARAMIAH | 60 |
| Road Tests of Automobiles with Petrol-Alcohol Mixtures. BY G. NARASIMHA IYENGAR | 55 | — — B. MUKERJI | 61 |

IMPROVED FORMULÆ FOR THE COEFFICIENT OF CONTINGENCY

IN Statistics, the number which summarises the mutual dependence of two attributes, say, A, B, of the elements of a universe is called a coefficient of contingency. If the classification of the A's is s-fold and that of the B's t-fold, there will be st classes of the type $A_m B_n$, $m \leq s$, $n \leq t$, whose frequencies a_{mn} may be set out in the form of a contingency table of s rows and t columns. If we write $\sum_{n=1}^t a_{mn} = x_m \neq 0$, $\sum_{m=1}^s a_{mn} = y_n \neq 0$, $S = \sum_{m=1}^s \sum_{n=1}^t a_{mn}^2 / x_m y_n$, the coefficient of mean square contingency, C, proposed by Karl Pearson is, in effect, equivalent to $\sqrt{1 - \frac{1}{S}}$. It is known that the coefficients calculated from the formula by varying s and t are not strictly comparable, as they are related to varying maxima. Tschuprow's modification

$$T = C / (1 - C^2)^{\frac{1}{2}} (s - 1)^{\frac{1}{2}} (t - 1)^{\frac{1}{2}}$$

remedies the defect only when $s = t$.

It is not perhaps generally known that the maximum value of S can never exceed the lesser of the two numbers s and t. Since

$$a_{mn} / x_m \leq 1, S \text{ does not exceed } \sum_{m=1}^s \sum_{n=1}^t a_{mn} / y_n = t;$$

and $S = t$ only when $a_{mn} = x_m$, i.e., in each row, all the frequencies vanish except one. Similarly $S \leq s$. It will now be clear that the maximum values of C and T are $\sqrt{(s-1)/s}$, and $\sqrt[3]{(s-1)/(t-1)}$ respectively, when $s < t$. To illustrate the point numerically, let $s = 7$, $t = 20$; then C will be less than $\sqrt{6/7} < .93$, while T is less than $\sqrt[3]{6/19} < .75$; Tschuprow's coefficient fares worse than Pearson's and underestimates the association.

It is therefore desirable that C and T should be replaced by a coefficient whose maximum is exactly unity. We suggest two formulæ, one based on C and the other on T:

$$1. K_1 = \sqrt{\frac{(S-1)s}{S(s-1)}} = C / \sqrt{\left(1 - \frac{1}{s}\right)};$$

$$2. K_2 = \sqrt{\frac{S-1}{s-1}} = T / \sqrt[3]{(s-1)/(t-1)};$$

where s is the number of cells in the shorter array of the contingency-table.

A. A. KRISHNASWAMI AYYANGAR.

Maharaja's College,
Mysore,
November 30, 1941.