of the bipartite number (22), namely

$$\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}=\frac{1}{n}\kappa_{22}+\frac{2}{n-1}\kappa_{11}^2 \tag{1 b}$$

and

$$\kappa\begin{pmatrix}1&1\\1&1\end{pmatrix}=\frac{1}{n}\kappa_{22}+\frac{1}{n-1}\kappa_{11}^2+\frac{1}{n-1}\kappa_{02}\kappa_{20}, \tag{1 c}$$

representing the product moment of the estimates of variance of the two correlated variates, and the variance of the estimated product moment.

It will be observed that by equating the two variates, which is carried out by summing the columns of the partition, and replacing the two suffixes of each κ by their sum, equations (1a), (1b), and (1c) are reduced to equation (1). As with univariate formulae, the partitions involving * parts of the first degree may be directly derived from formulae of lower degree and therefore need receive no separate consideration.

With more than two variates the bivariate notation may be extended to the use of three or more rows in the representation of a partition of a tripartite number, and three or more suffixes to the parameters κ. The remaining formulae of the fourth degree are therefore

$$\kappa\begin{pmatrix}2&0\\0&1\\0&1\end{pmatrix}=\frac{1}{n}\kappa_{211}+\frac{2}{n-1}\kappa_{110}\kappa_{101}, \tag{1 d}$$

$$\kappa\begin{pmatrix}1&1\\1&0\\0&1\end{pmatrix}=\frac{1}{n}\kappa_{211}+\frac{1}{n-1}\kappa_{200}\kappa_{011}+\frac{1}{n-1}\kappa_{101}\kappa_{110}, \tag{1 e}$$

$$\kappa\begin{pmatrix}1&0\\1&0\\0&1\\0&1\end{pmatrix}=\frac{1}{n}\kappa_{1111}+\frac{1}{n-1}\kappa_{1010}\kappa_{0101}+\frac{1}{n-1}\kappa_{1001}\kappa_{0110}, \tag{1*}$$

representing the partitions of the tripartite number (211) and of the quadrupartite (1111), ignoring such as have unitary parts.

Just as equation (1) may be derived from either of equations (1a), (1b), or (1c) by identifying the variates, so, by equating appropriate variates, (1a) may be derived from (1d), or (1b) from (1d), or (1c) from (1e), and finally all can be derived from the general multivariate formula (1*).

It appears, therefore, that the formulae appropriate for both univariate and multivariate distributions may all be expressed in terms of those representing partitions of the multipartite number (1^k). Thus of the

* For parts of the first degree, read unit parts.

sixth degree, a series of formulae, of which formula (3) is the final condensation, will be given by the partition of the multipartite (1^6) into parts $(1^4 0^2)$ and $(0^4 1^2)$, a series of formulae reducing to (4) by the partition into the parts $(1^3 0^3)$ and $(0^3 1^3)$, and a series of formulae reducing to (5) by the partition into parts $(1^2 0^4)$, $(0^2 1^2 0^2)$ and $(0^4 1^2)$. The presentation of formulae of the type here discussed for the case of many variates might therefore be completed by the tabulation of the general multivariate formulae (2*), (3*), etc.

The disadvantage of such a course is that such general formulae will consist of a large number of terms equal to the sum of the coefficients (of the highest powers of n) of the formulae already tabulated, and that each term will consist of a product of κ's, each having as many suffixes as the degree of the equation. The general formulae are therefore extremely cumbrous, and, as the suffixes will consist merely of repetitions in different orders of the numbers 0 and 1, it will be of more value if general rules can be found by which these particular combinations are to be selected. Such rules will then apply to the univariate and less general multivariate cases, the coefficients being merely the number of ways in which each selection can be made.

Now the suffixes of the product terms are merely other partitions of the same number, whether unipartite or multipartite, of which one particular partition specifies our formula; we are therefore concerned with the difficult question of the relations which can exist between different partitions of the same number. This question may be considered solely with respect to unipartite numbers, for if the rules can be made out which govern the coefficients in such cases, the same identical rules must apply to multipartite numbers by reason of the methods by which one formula may be condensed into another. For example, if we start with the partition (2^2) of the number 4, in conjunction with the rule that only such partitions are to be considered as in each part involve elements from both parts of the old partition, we should obtain equally the coefficient 2 of the term κ_2^2, and by applying the same rule to the partition of the multipartite number (1^4) into parts $(1^2 0^2)$ and $(0^2 1^2)$ should obtain the terms κ_{1010}, κ_{0101}, and κ_{1001}, κ_{0110}, having in both cases the same divisor $n-1$.

9. *Empirical statement of the rules for the direct evaluation of the coefficients.*

Although the rules of the combinatorial procedure were not completed before the development of the method of Section 10, yet so much can be learned by an empirical study of the formulae that it is convenient to make a complete statement of the rules in an empirical form, prior to the demonstration of their validity.

(1) The coefficient of $\kappa_{q_1}^{\chi_1}\kappa_{q_2}^{\chi_2}\ldots$ in the expression for $\kappa(p_1^{\pi_1}p_2^{\pi_2}\ldots)$ depends on the possible partitions of the second order of which the column totals give the partition $(p_1^{\pi_1}p_2^{\pi_2}\ldots)$, and the row totals give the partition $(q_1^{\chi_1}q_2^{\chi_2}\ldots)$.

For example, the coefficient of $\kappa_6\kappa_2^2$ in the expression for $\kappa(4^2 2)$ may be obtained by inspection of the partitions of the second order

$$\begin{array}{ccc|c} 2 & 2 & 2 & 6 \\ 1 & 1 & . & 2 \\ 1 & 1 & . & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} \qquad \begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 1 & 1 & . & 2 \\ 1 & . & 1 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} \qquad \begin{array}{ccc|c} 3 & 3 & . & 6 \\ 1 & . & 1 & 2 \\ . & 1 & 1 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array}$$

in each of which the sums of the rows constitute the partition (62^2), while the sums of the columns constitute the partition $(4^2 2)$.

(2) The numerical factor in the contribution made by any partition of the second order is the number of ways in which the totals in the lower margin may be allocated to form a partition of the type considered. The numerical factors corresponding to the three partitions set out above are 72, 192, and 32 respectively. In the first case, for example, the number may be arrived at from the consideration that the pair of units to be separated in the first four may be chosen in six ways, and that these may be assigned partners from the second four in twelve ways. In the second case we may choose either of the two fours to be parted into (21^2), as in the first column, and, whichever is chosen, we may allocate the units in the three columns in twelve, four, and two ways respectively; while, in the third case, we may choose the units from the two fours in sixteen ways and associate them in two ways with the units of the two.

(3) Before considering the general rule for determining the function of n by which the numerical factor is to be multiplied, it is convenient to note that certain partitions of the second order make no contribution whatever to the coefficient, and so may be neglected at once. The most useful class consists of those in which any row has only one entry other than zero; for example, such partitions as

$$\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ . & 1 & 1 & 2 \\ 2 & . & . & 2 \\ \hline 4 & 4 & 2 & 10 \end{array}$$

are to be ignored. It is obvious for statistical reasons, as has been mentioned above, that κ_1 cannot appear in any of these formulae, and as it will be seen that the function of n involved depends only upon the configuration of the zeros of the partition of the second order, the necessity for this rule will become apparent. More generally, we may exclude any partition in which any set of rows is connected to its complementary set by a single column only.

(4) The usefulness of rule (3) for excluding superfluous partitions is extended by employing it in conjunction with the rule which holds when any column has only one entry other than zero; for in these cases we may introduce the factor n^{-1} and ignore the column concerned. For example, the partition pattern

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & . \\ \times & \times & . \end{array}$$

irrespective of its numerical coefficient, is associated with a function of n which is one n-th of that associated with

$$\begin{array}{cc} \times & \times \\ \times & \times \\ \times & \times \end{array}$$

Moreover, such a partition as

$$\begin{array}{ccc|c} 4 & 2 & . & 6 \\ . & 1 & 1 & 2 \\ . & 1 & 1 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array}$$

is to be ignored (although every row has two entries) by reason of its connection with

$$\begin{array}{cc} \times & . \\ \times & \times \\ \times & \times \end{array}$$

in which this condition is not fulfilled.

With these criteria of rejection one may easily assure oneself that the three partitions set out above are the only ones which need be considered in that case.

(5) To find, in general, the function of n with which any pattern is associated, we consider all the possible ways in which the rows can be

separated into 1, 2, 3, ... separate groups, or separates. Thus with three rows we have one separation into one separate, with which is associated the factor n; three separations into two separates, with which is associated the factor $n(n-1)$; and one separation into three separates, with which is associated the factor $n(n-1)(n-2)$. In each of these five separations we count in how many separates each column is represented by entries other than zero. If in one separate, that column contributes a factor n^{-1}; if in 2, 3, 4, ... separates, the factors are

$$\frac{-1}{n(n-1)}, \quad \frac{2!}{n(n-1)(n-2)}, \quad \frac{-3!}{n(n-1)(n-2)(n-3)}.$$

In applying this rule all patterns which are resolvable into two parts, each confined to separable sets of rows and columns, must be ignored.

As an example, consider the five possible separations of the pattern

$$\begin{matrix} \times & \times \\ \times & \times \\ \times & \times \end{matrix}\;;$$

the first supplies the term

$$\frac{n}{n^2} = \frac{1}{n},$$

the separations into two separates supply

$$\frac{3n(n-1)}{n^2(n-1)^2} = \frac{3}{n(n-1)},$$

while the separation into three separates gives

$$\frac{4n(n-1)(n-2)}{n^2(n-1)^2(n-2)^2} = \frac{4}{n(n-1)(n-2)},$$

the total being $n/\{(n-1)(n-2)\}$, the function appropriate to this pattern.

It is equally easy to verify that the functions appropriate to the patterns

$$\begin{matrix} \times & \times & \times \\ \times & \times & . \\ \times & . & \times \end{matrix} \qquad \begin{matrix} \times & \times & . \\ \times & . & \times \\ . & \times & \times \end{matrix}$$

both reduce to $1/(n-1)^2$. The required coefficient is therefore

$$\frac{72}{(n-1)(n-2)}+\frac{224}{(n-1)^2}=\frac{8(37n-65)}{(n-1)^2(n-2)},$$

as appears in formula 28.

It will be obvious from the preceding section that the same rules must be applicable to multivariate problems, the only difference being that the column totals are then regarded as consisting of objects of two or more kinds. For example, to find the coefficient of $\kappa_{33}\kappa_{11}^2$ in the expression for $\kappa\begin{pmatrix}2&2&1\\2&2&1\end{pmatrix}$, it is merely necessary to note that the second order partitions of the bipartite (55) corresponding to the three partitions of 10, used above, can be allocated in 20, 48, and 8 ways respectively, yielding a coefficient

$$\frac{4(19n-33)}{(n-1)^2(n-2)}.$$

Alternatively the contributions to the coefficient of the univariate formula may be each split up among the six coefficients by which it is replaced in the bivariate formula, giving in this case

$$4(19n-33)\kappa_{33}\kappa_{11}^2+8(11n-20)\kappa_{33}\kappa_{20}\kappa_{02}+8(7n-12)\kappa_{42}\kappa_{11}\kappa_{02}$$
$$+8(7n-12)\kappa_{24}\kappa_{11}\kappa_{20}+2(5n-9)\kappa_{81}\kappa_{20}^2+2(5n-9)\kappa_{15}\kappa_{20}^2$$

in place of $\qquad 8(37n-65)\kappa_6\kappa_2^2.$

In the same way the appropriate subdivision of the other bivariate and multivariate formulae may be obtained from an examination of the same set of two-way partitions, and it will evidently be sufficient for practical purposes to tabulate all the univariate formulae up to a given degree in order that all the corresponding multivariate formulae should be rapidly obtainable.

The algebraic equivalents of a number of the more commonly occurring patterns are given on pages 223–226.

Some useful patterns.

Two rows.

$$\begin{matrix}\times\times\\\times\times\end{matrix}\ \frac{1}{n-1}\qquad \begin{matrix}\times\times\times\\\times\times\times\end{matrix}\ \frac{n-2}{n(n-1)^2}\qquad \begin{matrix}\times\times\times\times\\\times\times\times\times\end{matrix}\ \frac{n^2-3n+3}{n^2(n-1)^3}$$

In general, if $a=-\{1/(n-1)\}$, we have $\dfrac{1-a^{p-1}}{n^{p-1}}$.

Three rows.

$$\begin{array}{cc} \times & \times \\ \times & \times \\ \times & \times \end{array} \quad \frac{n}{(n-1)(n-2)} \qquad \begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{array} \quad \frac{n^2-6n+10}{(n-1)^2(n-2)^2} \qquad \begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \cdot \end{array} \quad \frac{n-3}{(n-1)^2(n-2)}$$

$$\begin{array}{ccc} \times & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \end{array} \equiv \begin{array}{ccc} \cdot & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \end{array} \quad \frac{1}{(n-1)^2} \qquad \begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{array} \quad \frac{n^4-9n^3+33n^2-60n+48}{n(n-1)^3(n-2)^3}$$

$$\begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \cdot \end{array} \quad \frac{(n-3)(n^2-4n+6)}{n(n-1)^3(n-2)^2} \qquad \begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \cdot & \cdot \end{array} \quad \frac{n^2-4n+5}{n(n-1)^3(n-2)}$$

$$\begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \cdot & \times \\ \times & \times & \times & \cdot \end{array} \quad \frac{n^2-5n+7}{n(n-1)^3(n-2)} \qquad \begin{array}{cccc} \times & & \times & \times \\ \times & \times & \cdot & \times \\ \times & \times & \times & \cdot \end{array} \quad \frac{n-3}{n(n-1)^3}$$

$$\begin{array}{cccc} \times & \times & \times & \times \\ \times & \cdot & \times & \times \\ \times & \times & \cdot & \cdot \end{array} \equiv \begin{array}{cccc} \cdot & \times & \times & \times \\ \times & \cdot & \times & \times \\ \times & \times & \cdot & \cdot \end{array} \quad \frac{n-2}{n(n-1)^3} \qquad \begin{array}{cccc} \times & \times & \times & \times \\ \cdot & \cdot & \times & \times \\ \times & \times & \cdot & \cdot \end{array} \quad \frac{1}{n(n-1)^2}$$

Four rows.

$$\begin{array}{cc} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{array} \quad \frac{n(n+1)}{(n-1)(n-2)(n-3)} \qquad \begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{array} \quad \frac{n^4-12n^3+51n^2-74n-18}{(n-1)^2(n-2)^2(n-3)^2}$$

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \cdot \end{array} \quad \frac{n^3-8n^2+17n+2}{(n-1)^2(n-2)^2(n-3)} \qquad \begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \cdot \\ \times & \times & \cdot \end{array} \quad \frac{n^2-4n-1}{(n-1)^2(n-2)(n-3)}$$

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \end{array} \quad \frac{n(n-4)}{(n-1)^2(n-2)^2} \qquad \begin{array}{ccc} \times & \times & \times \\ \cdot & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \end{array} \quad \frac{n(n-3)}{(n-1)^2(n-2)^2}$$

$$\begin{array}{ccc} \times & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \\ \times & \times & \cdot \end{array} \equiv \begin{array}{ccc} \cdot & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \\ \times & \times & \cdot \end{array} \quad \frac{n}{(n-1)^2(n-2)}$$

$$\begin{array}{cccc} \times & \times & \cdot & \times \\ \times & \times & \cdot & \times \\ \times & \times & \times & \cdot \\ \times & \times & \times & \cdot \end{array} \equiv \begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \cdot & \times \\ \times & \times & \cdot & \cdot \\ \times & \times & \times & \cdot \end{array} \quad \frac{n^3-7n^2+13n+1}{n(n-1)^3(n-2)(n-3)}$$

$$\begin{array}{cccc}\times&\times&\times&\times\\\times&\times&\times&\times\\\times&\times&\cdot&\cdot\\\times&\times&\cdot&\cdot\end{array}\quad\frac{n^3-5n^2+7n+1}{n(n-1)^3(n-2)(n-3)}\qquad\qquad\begin{array}{cccc}\times&\times&\times&\times\\\cdot&\times&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\cdot\end{array}\quad\frac{n^3-8n^2+23n-24}{(n-1)^3(n-2)^3}$$

$$\begin{array}{cccc}\cdot&\times&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\times&\cdot\end{array}\quad\frac{n^3-9n^2+29n-32}{(n-1)^3(n-2)^3}\qquad\qquad\begin{array}{cccc}\times&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\times&\cdot\\\times&\times&\times&\cdot\end{array}\quad\frac{n^2-7n+14}{(n-1)^3(n-2)^2}$$

$$\begin{array}{cccc}\times&\times&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\times&\times&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\times&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\times&\cdot\\\times&\times&\times&\cdot\end{array}\quad\frac{n^2-6n+10}{(n-1)^3(n-2)^2}$$

$$\begin{array}{cccc}\times&\times&\times&\times\\\times&\times&\times&\times\\\cdot&\cdot&\times&\times\\\times&\times&\cdot&\cdot\end{array}\equiv\begin{array}{cccc}\times&\times&\times&\times\\\times&\times&\times&\times\\\times&\cdot&\times&\cdot\\\times&\times&\cdot&\cdot\end{array}\quad\frac{n^2-5n+8}{(n-1)^3(n-2)^2}\qquad\qquad\begin{array}{cccc}\times&\times&\times&\times\\\cdot&\times&\times&\times\\\times&\cdot&\times&\cdot\\\times&\times&\cdot&\cdot\end{array}\quad\frac{n^2-5n+7}{(n-1)^3(n-2)^2}$$

$$\begin{array}{cccc}\cdot&\times&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}\quad\frac{(n-3)^2}{(n-1)^3(n-2)^2}\qquad\qquad\begin{array}{cccc}\times&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\quad\frac{n-4}{(n-1)^3(n-2)}$$

$$\begin{array}{cccc}\times&\cdot&\cdot&\times\\\times&\times&\cdot&\times\\\times&\times&\times&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\times&\times&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\times&\times&\times&\times\\\cdot&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\times&\times&\times\\\cdot&\times&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}$$

$$\equiv\begin{array}{cccc}\cdot&\times&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\cdot&\times&\times\\\times&\times&\cdot&\times\\\times&\times&\times&\cdot\\\times&\times&\cdot&\cdot\end{array}\quad\frac{n-3}{(n-1)^3(n-2)}$$

$$\begin{array}{cccc}\times&\times&\times&\times\\\times&\times&\cdot&\cdot\\\times&\times&\cdot&\cdot\\\cdot&\cdot&\times&\times\end{array}\quad\frac{1}{(n-1)^2(n-2)}$$

$$\begin{array}{cccc}\times&\times&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\\\times&\cdot&\times&\cdot\end{array}\equiv\begin{array}{cccc}\times&\times&\times&\times\\\cdot&\cdot&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\times&\times&\times\\\cdot&\cdot&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\times&\times&\cdot\\\cdot&\cdot&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\end{array}\equiv\begin{array}{cccc}\cdot&\times&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\cdot\\\times&\times&\cdot&\cdot\end{array}$$

$$\equiv\begin{array}{cccc}\cdot&\times&\times&\times\\\times&\times&\cdot&\cdot\\\times&\cdot&\times&\cdot\\\times&\cdot&\cdot&\times\end{array}\equiv\begin{array}{cccc}\cdot&\cdot&\times&\times\\\times&\cdot&\times&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\equiv\begin{array}{cccc}\times&\cdot&\times&\times\\\times&\cdot&\cdot&\times\\\times&\times&\cdot&\cdot\\\times&\times&\times&\cdot\end{array}\quad\frac{1}{(n-1)^3}$$

Five and six row patterns.

$$\begin{matrix}\times&\times&.\\\times&\times&.\\\times&.&\times\\\times&.&\times\\.&\times&\times\end{matrix}\equiv\begin{matrix}\times&\times&.\\\times&\times&.\\\times&.&\times\\\times&.&\times\\\times&\times&\times\end{matrix}\quad\frac{n^2}{(n-1)^2(n-2)^2}\qquad\begin{matrix}\times&\times&.\\\times&\times&.\\\times&\times&.\\\times&.&\times\\.&\times&\times\end{matrix}\equiv\begin{matrix}\times&\times&.\\\times&\times&.\\\times&\times&.\\\times&.&\times\\\times&\times&\times\end{matrix}\quad\frac{n(n+1)}{(n-1)^2(n-2)(n-3)}$$

$$\begin{matrix}\times&\times\\\times&\times\\\times&\times\\\times&\times\\\times&\times\end{matrix}\quad\frac{n^2(n+5)}{(n-1)(n-2)(n-3)(n-4)}\qquad\begin{matrix}\times&\times&\times\\\times&\times&.\\\times&\times&.\\\times&.&\times\\.&\times&\times\end{matrix}\quad\frac{n(n^2-4n-1)}{(n-1)^2(n-2)^2(n-3)}$$

$$\begin{matrix}\times&\times&\times\\\times&\times&\times\\\times&\times&.\\\times&\times&.\\\times&.&\times\end{matrix}\quad\frac{n(n^2-5n-2)}{(n-1)^2(n-2)^2(n-3)}\qquad\begin{matrix}\times&\times&\times\\\times&\times&\times\\\times&\times&.\\\times&\times&.\\\times&\times&.\end{matrix}\quad\frac{n(n^2-4n-9)}{(n-1)^2(n-2)(n-3)(n-4)}$$

$$\begin{matrix}\times&\times&\times\\\times&\times&\times\\\times&\times&.\\\times&.&\times\\.&\times&\times\end{matrix}\quad\frac{n(n^3-9n^2+19n+5)}{(n-1)^2(n-2)^2(n-3)^2}\qquad\begin{matrix}\times&\times&.\\\times&\times&.\\\times&.&\times\\\times&.&\times\\.&\times&\times\\.&\times&\times\end{matrix}\quad\frac{n(n+1)(n^2-5n+2)}{(n-1)^2(n-2)^2(n-3)^2}$$

$$\begin{matrix}\times&\times&.\\\times&\times&.\\\times&\times&.\\\times&.&\times\\\times&.&\times\\.&\times&\times\end{matrix}\quad\frac{n^2(n+1)}{(n-1)^2(n-2)^2(n-3)}\qquad\begin{matrix}\times&\times\\\times&\times\\\times&\times\\\times&\times\\\times&\times\\\times&\times\end{matrix}\quad\frac{n(n+1)(n^2+15n-4)}{(n-1)(n-2)(n-3)(n-4)(n-5)}$$

The general formula for the two-column pattern with r rows is easily found, by enumerating the separations into 1, 2, 3, ... separates, to be

$$\sum_{p=1}^{r}\frac{(p-1)!}{p}\,\frac{\Delta^p(0^r)}{n(n-1)\ldots(n-p+1)},$$

where $\Delta^p(0^r)$ stands for the leading p-th advancing difference of the series $0^r, 1^r, 2^r, \ldots$.

10. *Demonstration of the combinatorial method.*

To demonstrate the validity of the rules which have been stated, it is useful to consider in what manner the generating function M will be modified by a functional transformation of the variates. In the case of a single variate x we have the function

$$M = 1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \dots,$$

the coefficients of which give the mean values of all powers of x in the population. By what operation should the function M be transformed so as to give the corresponding function appropriate to a new variate ξ, which is a known function of x? Suppose that

$$\xi = f(x) = c_0 + c_1 x + c_2 x^2 + \dots,$$

then the mean value of ξ is

$$\mu_1' = c_0 + c_1 \mu_1 + c_2 \mu_2 + \dots,$$

which may be written

$$c_0 + c_1 \frac{d}{dt} M + c_2 \frac{d^2}{dt^2} M + \dots,$$

or

$$f\left(\frac{d}{dt}\right) M,$$

where t is made to vanish after operation.

Moreover, the mean value of the r-th power of ξ will be given, at least formally, by the equation

$$\mu_r' = \left\{ f\left(\frac{d}{dt}\right) \right\}^r M,$$

and the new generating function,

$$M' = 1 + \mu_1' \tau + \mu_2' \frac{\tau^2}{2!} + \dots,$$

may be written

$$e^{\tau f} M,$$

in which the operator is supposed to be expanded in powers of d/dt before attacking the operand.

The corresponding relationship for simultaneous variation is easily found. In such cases M will be a function of two or more variables $t_1, t_2, \dots$ corresponding to the variates $x, y, \dots$; the new variates will be given functions of the old

$$\xi_1 = f_1(x, y, \dots),$$

$$\xi_2 = f_2(x, y, \dots),$$

$$\dots \quad \dots \quad \dots$$

and the operative expression for the transformation of M is

$$M' = e^{\tau_1 f_1 + \tau_2 f_2 \dots} M.$$

To apply this result to univariate sampling problems, consider the n observations of the sample as our n original variates, and the symmetric functions k_1, k_2, ... as the new variates the generating function of which is required. Then, considering first the operand, for the first observation x, the μ generator is $e^{K(t_1)}$, where K is the κ generator of the population sampled, *i.e.*

$$K(t) = \kappa_1 t + \kappa_2 \frac{t^2}{2!} + \dots.$$

Moreover, since the n observations are independent, their simultaneous κ generator will be merely the sum of the individual generators, so that our operand is

$$\exp\left\{\kappa_1 s_1 + \kappa_2 \frac{s_2}{2!} + \dots\right\},$$

in which

$$s_r = \overset{n}{\underset{\nu=1}{S}} (t_\nu^r).$$

We may note at once that the coefficient of $\kappa_{q_1}^{\chi_1} \kappa_{q_2}^{\chi_2} \dots$ in the operand is

$$\frac{s_{q_1}^{\chi_1}}{(q_1!)^{\chi_1} \chi_1!} \frac{s_{q_2}^{\chi_2}}{(q_2!)^{\chi_2} \chi_2!} \dots.$$

The μ generator of the simultaneous distribution of the k statistics will be given by the operator

$$e^{\tau_1 k_1 + \tau_2 k_2 + \dots + \tau_n k_n},$$

in which k_ν is interpreted as the same function of d/dt_1, d/dt_2, ... as the corresponding k statistic is of $x_1, x_2, \dots, x_n$. The property by which these statistics were defined, namely that the mean value of k_ν should be κ_ν, is now seen to imply that

$$k_\nu \left(\frac{s_\nu}{\nu!}\right) = 1;$$

but

$$k_\nu \left(\frac{s_{\nu_1}}{\nu_1!} \frac{s_{\nu_2}}{\nu_2!} \dots\right) = 0,$$

where $(\nu_1, \nu_2, \dots)$ is any partition of ν. If, for example, the partition is of two parts,

$$s_{\nu_1} s_{\nu_2} = \overset{n}{\underset{1}{S}} (t^{\nu_1 + \nu_2}) + \overset{n(n+1)}{\underset{1}{S}} (t^{\nu_1} t'^{\nu_2}),$$

in which t and t' are different members of the set $t_1, \ldots, t_n$, it follows that k_ν must contain, in addition to the simple term

$$\frac{1}{n} S\left(\frac{d}{dt}\right)^{\nu},$$

terms for all two-part partitions of the form

$$\frac{-1}{n(n-1)} \frac{\nu!}{\nu_1!\,\nu_2!} S\left\{\left(\frac{d}{dt}\right)^{\nu_1}\left(\frac{d}{dt'}\right)^{\nu_2}\right\}$$

except when $\nu_1 = \nu_2$, when each operator finds two terms on which it can act, and its coefficient is therefore to be halved. Thus, if we write

$$g(p_1^{\pi_1} p_2^{\pi_2} \ldots) = \frac{(-)^{\rho-1}(\rho-1)!}{n(n-1)\ldots(n-\rho+1)}$$

$$\times S\left\{\left(\frac{d}{dt_{\nu_1}}\right)^{p} \cdots \left(\frac{d}{dt_{\nu_{\pi_1}}}\right)^{p_1}\left(\frac{d}{dt_{\nu_1'}}\right)^{p_2} \cdots \left(\frac{d}{dt_{\nu_{\pi_2}'}}\right)^{p_2} \cdots\right\},$$

where $\rho = \pi_1 + \pi_2 + \ldots$ and t_{ν_1}, etc. are any selection of ρ out of the n variables t, the summation being extended over all such selections, then

$$k_p = \Sigma \frac{p!}{(p_1!)^{\pi_1}\pi_1!\,(p_2!)^{\pi_2}\pi_2!\ldots} g(p_1^{\pi_1} p_2^{\pi_2} \ldots),$$

the summation being taken over all partitions of p.

This structure of the k operator makes it possible to think of the p acts of differentiation in each operator as p separate objects, the partitions of which, represented by the g operators, occur each in as many ways as the objects can be arranged in that partition. We may thus use a two-way partition to assign how many of these operations are effective against each of a series of factors $s_{q_1} s_{q_2}$ constituting the operand.

Let now this operand product be expanded in a number of terms of the form

$$z(a, b, c) = S(t^a t'^b t''^c),$$

the summation being taken over all the $n(n-1)(n-2)$ different ways of selecting t, t', and t'' from among the set $t_1, \ldots, t_n$. There will then be a z term for every possible separation of the partition $(q_1^{\chi_1} q_2^{\chi_2} \ldots)$ into one or more separates. For each two-way partition chosen all these separations will contribute to the result, and with the same numerical coefficient, apart from that contained in the g operators, equal to the number of ways of allocating the objects in the two-way partition.

The number of terms in the z corresponding to any separation into a separates is $n(n-1)\ldots(n-a+1)$, and this, combined with the factors in the g operators, gives the functions of n corresponding to any two-way partition according to rule (5). There remains, however, in M' a number of terms corresponding to two-way partitions, in which the columns may be divided into two classes, each confined to different sets of rows. These introduce terms of a higher order in n, which are obliterated when we find $K' = \log M'$, for in these cases the additional term in M' will be of the form AB, where both A and B occur also as other terms in M'.

11. *Measures of departure from normality.*

The statistical inefficiency of moment statistics from distributions differing widely from the normal, except when they are of a special type [**10**], much reduces their practical importance for curve fitting; but since they are fully efficient for the normal distribution, they provide an ideal basis for testing if an observed sample indicates a significant departure from normality in the population sampled. Significant asymmetry should, in the first instance, be shown by an excessive value of k_3; but since the variance of the population is usually unknown, but may be estimated from the value of k_2 observed in the sample, the test of significance will usually involve not the distribution of k_3, the moments of which are given by such formulae as (4) and (20), but the distribution of the ratio $k_3 k_2^{-\frac{3}{2}}$. Since, for the normal distribution, the variance of k_3 is given by

$$\kappa(3^2) = \frac{6n}{(n-1)(n-2)}\kappa_2^3,$$

it will be convenient to show how the moments of such a statistic as

$$x = \sqrt{\left(\frac{(n-1)(n-2)}{6n}\right)}\, k_3 k_2^{-\frac{3}{2}}$$

may be expressed in terms of the general $\kappa(p_1^{\pi_1} p_2^{\pi_2}\ldots)$, for all distributions, and its particular value obtained for the normal distribution. This may be done by expanding the factor $k_2^{-\frac{3}{2}}$ in the form

$$\kappa_2^{-\frac{3}{2}}\left(1+\frac{k_2-\kappa_2}{\kappa_2}\right)^{-\frac{3}{2}},$$

whereupon, in virtue of the expressions connecting the moments μ with the semi-invariants κ, and the mean value of x being zero, we can at once write down the following expansion for its variance :

$$\mu_2(x) = \kappa_2(x) = \frac{(n-1)(n-2)}{6n\kappa_2^3}\left\{\kappa(3^2) - \frac{3}{\kappa_2}\kappa(3^2 2) + \frac{6}{\kappa_2^2}\{\kappa(3^2)\kappa(2^2) + \kappa(3^2 2^2)\}\right.$$
$$\left. - \frac{10}{\kappa_2^3}\{3\kappa(3^2 2)\kappa(2^2) + \kappa(3^2)\kappa(2^3)\} + \frac{15}{\kappa_2^4}\{3\kappa(3^2)\kappa^2(2^2)\}\right\},$$

in which, remembering that a κ of p parts involves $n^{-(p-1)}$, terms beyond n^{-2} have been omitted, as well as the terms of odd degree which vanish for symmetrical distributions.

Similarly we have

$$\mu_4(x) = \frac{(n-1)^2(n-2)^2}{36n^2\kappa_2^6}\left\{3\kappa^2(3^2) + \kappa(3^4) - \frac{6}{\kappa_2}\{6\kappa(3^2 2)\kappa(3^2) + \kappa(3^4 2)\}\right.$$
$$+ \frac{21}{\kappa_2^2}\{3\kappa^2(3^2)\kappa(2^2) + 6\kappa^2(3^2 2)$$
$$+ \kappa(3^4)\kappa(2^2) + 6\kappa(3^2 2^2)\kappa(3^2)\}$$
$$- \frac{56}{\kappa_2^3}\{18\kappa(3^2 2)\kappa(3^2)\kappa(2^2) + 3\kappa^2(3^2)\kappa(2^3)\}$$
$$\left. + \frac{126}{\kappa_2^4}\{9\kappa^2(3^2)\kappa^2(2^2)\}\right\}$$

and

$$\mu_6(x) = \frac{(n-1)^3(n-2)^3}{216n^3\kappa_2^9}\left\{15\kappa^3(3^2) + 15\kappa(3^4)\kappa(3^2) + \kappa(3^6)\right.$$
$$- \frac{9}{\kappa_2}\{45\kappa(3^2 2)\kappa^2(3^2) + 15\kappa(3^4 2)\kappa(3^2) + 15\kappa(3^4)\kappa(3^2 2)\}$$
$$+ \frac{45}{\kappa_2^2}\{15\kappa^3(3^2)\kappa(2^2) + 90\kappa^2(3^2 2)\kappa(3^2) + 45\kappa(3^2 2^2)\kappa^2(3^2)$$
$$+ 15\kappa(3^4)\kappa(3^2)\kappa(2^2)\}$$
$$- \frac{165}{\kappa_2^3}\{15\kappa^3(3^2)\kappa(2^3) + 135\kappa(3^2 2)\kappa^2(3^2)\kappa(2^2)$$
$$\left. + \frac{495}{\kappa_2^4}\{45\kappa^3(3^2)\kappa^2(2^2)\}\right\}.$$

From these moments of the distribution of x, the semi-invariants $\kappa_4(x)$ and $\kappa_6(x)$ may be obtained by means of the relations

$$\mu_4(x) = \kappa_4(x) + 3\kappa_2^2(x),$$

$$\mu_6(x) = \kappa_6(x) + 15\kappa_4(x)\kappa_2(x) + 15\kappa_2^3(x),$$

giving

$$\kappa_4(x) = \frac{(n-1)^2(n-2)^2}{36n^2\kappa_2^6}\Big\{\kappa(3^4) - \frac{18}{\kappa_2}\kappa(3^2 2)\kappa(3^2) + \frac{27}{\kappa_2^2}\kappa^2(3^2)\kappa(2^2)$$

$$-\frac{6}{\kappa_2}(3^4 2) + \frac{99}{\kappa_2^2}\kappa^2(3^2 2) + \frac{21}{\kappa_2^2}\kappa(3^4)\kappa(2^2)$$

$$+\frac{90}{\kappa_2^2}\kappa(3^2 2^2)\kappa(3^2) - \frac{720}{\kappa_2^3}\kappa(3^2 2)\kappa(3^2)\kappa(2^2)$$

$$-\frac{108}{\kappa_2^3}\kappa^2(3^2)\kappa(2^3) + \frac{756}{\kappa_2^4}\kappa^2(3^2)\kappa^2(2^2)\Big\}$$

and

$$\kappa_6(x) = \frac{(n-1)^3(n-2)^3}{216n^3\kappa_2^9}\Big\{\kappa(3^6) - \frac{45}{\kappa_2}\kappa(3^4 2)\kappa(3^2) - \frac{90}{\kappa_2}\kappa(3^4)\kappa(3^2 2)$$

$$+\frac{1350}{\kappa_2^2}\kappa^2(3^2 2)\kappa(3^2) + \frac{405}{\kappa_2^2}\kappa(3^2 2^2)\kappa^2(3^2)$$

$$+\frac{270}{\kappa_2^2}\kappa(3^4)\kappa(3^2)\kappa(2^2) - \frac{405}{\kappa_2^3}\kappa^3(3^2)\kappa(2^3)$$

$$-\frac{5670}{\kappa_2^3}\kappa(3^2 2)\kappa^2(3^2)\kappa(2^2) + \frac{4860}{\kappa_2^4}\kappa^3(3^2)\kappa^2(2^2)\Big\},$$

while no higher semi-invariants contain terms involving only n^{-2}.

The formulae tabulated give all the values required for $\kappa_2(x)$; thus for samples from the normal distribution

$$\kappa(2^2) = \frac{2}{n-1}\kappa_2^2, \qquad \kappa(3^2) = \frac{6n}{(n-1)(n-2)}\kappa_2^3, \qquad \kappa(2^3) = \frac{8}{(n-1)^2}\kappa_2^3.$$

$$\kappa(3^2 2) = \frac{6}{n-1}\kappa_2\kappa(3^2), \qquad \kappa(3^2 2^2) = \frac{48}{(n-1)^2}\kappa_2^2\kappa(3^2),$$

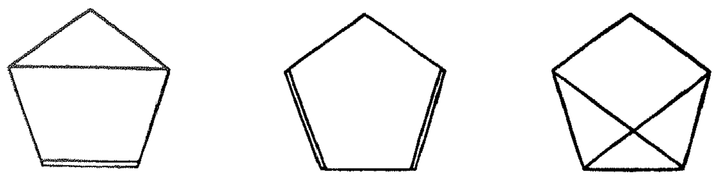
and substituting these values, we find

$$\kappa_2(x) = 1 - \frac{6}{n} + \frac{22}{n^2}.$$

To evaluate $\kappa_4(x)$ we need in addition

$$\kappa(3^4) = \frac{648(5n-12)n^2}{(n-1)^3(n-2)^3}\kappa_2^6,$$

and the leading term in $\kappa(3^4 2)$; this latter only requires the enumeration of the number of ways of building up two-way partitions of $(3^4 2)$ with row totals (2^7), or the number of ways of connecting up the symbolical figures

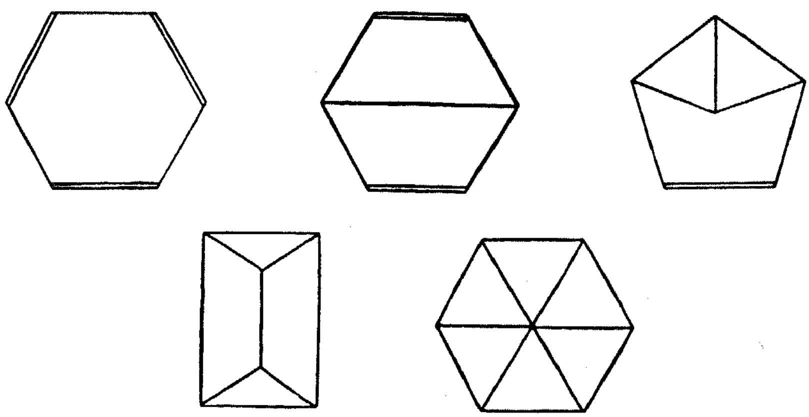


which can be done in 15552, 7776 and 15552 ways respectively, showing that $\kappa(3^4 2)$ from normal samples is approximately $38880n^{-4}$.

With this value, that of $\kappa_4(x)$ is evaluated as

$$\kappa_4(x) = \frac{36}{n} - \frac{1296}{n^2}.$$

Finally, for $\kappa_6(x)$ the only new κ required is $\kappa(3^6)$, involving the figures having six points from each of which three lines radiate :—



which supply a contribution of $47520n^{-2}$ to $\kappa_6(x)$, or, with the other terms, lead to the value

$$\kappa_6(x) = \frac{15120}{n^2}.$$

For the practical application of the function x in testing asymmetry we shall now require to construct a function of x which, as far as terms in n^{-2}, is distributed normally. Putting

$$x = \beta\xi + \delta(\xi^3 - 3\xi) + \eta(\xi^5 - 10\xi^3 + 15),$$

where ξ is normally distributed with unit variance, it is easy to obtain

$$\kappa_2(x) = \beta^2 + 6\delta^2,$$
$$\kappa_4(x) = 24\beta^3\delta + 216\beta^2\delta^2,$$
$$\kappa_6(x) = 720\beta^5\eta + 3240\beta^4\delta^2,$$

which are satisfied by

$$\beta = 1-\frac{3}{n}-\frac{1}{4n^2}, \qquad \delta = \frac{3}{2n}\left(1-\frac{81}{2n}\right), \qquad \eta = \frac{87}{8n^2},$$

or, inverting the relation between ξ and x, we have

$$\xi = x\left(1+\frac{3}{n}+\frac{91}{4n^2}\right)-\frac{3}{2n}\left(1-\frac{111}{2n}\right)(x^3-3x)-\frac{33}{8n^2}(x^5-10x^3+15x).$$

This translation formula makes it possible to assess the numerical effects upon tests of significance of the actual distribution; Tables 2 and 3 show the values of various possible formulae for the test deviate in the region, important for tests of significance, $x = 1\cdot 8$ to $2\cdot 2$, and indicate that these effects are very serious.

TABLE 2.

Comparison of deviates in five formulae for testing asymmetry.

$n = 100.$

(a)	(b)	(c)	(d)	(e)
$\sqrt{\frac{n}{6}}\, m_3 m_2^{-\frac{3}{2}}$	$\sqrt{\frac{n}{6}}\, k_3 k_2^{-\frac{3}{2}}$	x	ξ_1	ξ_2
1·7999	1·8274	1·8	1·8475	1·8603
1·9999	2·0305	2·0	2·0300	2·0586
2·1999	2·2335	2·2	2·2053	2·2530

TABLE 3.

Comparison of deviates in five formulae for testing asymmetry.

$n = 50.$

(a)	(b)	(c)	(d)	(e)
$\sqrt{\frac{n}{6}}\, m_3 m_2^{-\frac{3}{2}}$	$\sqrt{\frac{n}{6}}\, k_3 k_2^{-\frac{3}{2}}$	x	ξ_1	ξ_2
1·7996	1·8558	1·8	1·8950	1·9463
1·9996	2·0620	2·0	2·0600	2·1745
2·1995	2·2682	2·2	2·2106	2·4016

In this region an error of $0\cdot 1$ in the deviate produces an error of about 24 per cent. in the probability deduced, and, although high accuracy in the latter is not a necessity, little reliance can be placed upon tests when the deviate may be biased by as much as $0\cdot 2$. Of the formulae tested, the formula (*a*) in terms of crude moments is almost equivalent to the use of x, and these are evidently the most in error. Of the simple formulae (*b*) is least in error, and for samples of 100 this error is only about ·03. The value ξ_1 shows the effect of using terms of the first degree only in the translation formula, while ξ_2 shows the effect of using also terms in n^{-2}. There is evidently little to be gained by using ξ_1 instead of the simple formula $\sqrt{(n/6)}\,k_3 k_2^{-\frac{3}{2}}$, which latter gives apparently the better values for deviations exceeding 2·0. For samples as small as 50 the fully corrected value ξ_2 is evidently required, and in view of the uncertainty of the effect of the omitted terms in n^{-3}, etc., no reliable test of normality for materially smaller samples can be said to be available. As in so many other cases, the adequate treatment even of moderately small samples is not well approached by series in n^{-1}.

12. *The significance of the fourth moment.*

The sampling variance of k_4 from a normal sample is

$$\frac{24n(n+1)}{(n-1)(n-2)(n-3)}\kappa_2^4;$$

in testing the significance of such a value, we should therefore naturally calculate

$$x=\sqrt{\left(\frac{(n-1)(n-2)(n-3)}{24n(n+1)}\right)}\,k_4k_2^{-2}$$

as a variate which, with increasing sample number, tends to be normally distributed with unit variance. With finite samples the distribution is asymmetrical, for $\kappa(4^3)$ is not zero. The true mean value of x is zero, for with a normal distribution $\kappa(42^p)$ is zero for all values of p, whence it follows that the mean of k_4 is zero independently for all values of k_2.

The mean value of x^2 is easily expanded in the form

$$\frac{(n-1)(n-2)(n-3)}{24n(n+1)\kappa_2^4}\left\{\kappa(4^2)-\frac{4}{\kappa_2}\kappa(4^2 2)+\frac{10}{\kappa_2^2}\kappa(4^2)\kappa(2^2)+\ldots\right\}$$

or

$$1-\frac{32}{n-1}+\frac{20}{(n-1)^2}$$

as far as n^{-1}.

The mean value of x^3, as far as $n^{-\frac{3}{2}}$, is

$$\left\{\frac{(n-1)(n-2)(n-3)}{24n(n+1)\kappa_2^4}\right\}^{\frac{3}{2}}\left\{\kappa(4^3)-\frac{6}{\kappa_2}\kappa(4^3 2)+\frac{21}{\kappa_2^2}\kappa(4^3)\kappa(2^2)+\ldots\right\}.$$

Now $\kappa(4^3)$ has been evaluated by the direct combinatorial method, giving (formula 57)

$$\kappa(4^3)=\frac{1728n(n+1)(n^2-5n+2)}{(n-1)^2(n-2)^2\ (n-3)^2}\kappa_2^6,$$

or, as near as needed,

$$\frac{1728}{n^3}(n+8);$$

while the leading term of $\kappa(4^3 2)$ is $1728n^{-3} \times 12$, giving, as the mean value of x^3,

$$\left\{\frac{(n-1)(n-2)(n-3)}{n^3(n+1)}\right\}^{\frac{3}{2}} 6\sqrt{6}\{n+8-72+42\},$$

or

$$\frac{6\sqrt{6}}{\sqrt{n}}\left(1-\frac{65}{2n}\right).$$

Next, the mean value of x^4 is, as far as n^{-1},

$$\left\{\frac{(n-1)(n-2)(n-3)}{24n(n+1)\kappa_2^4}\right\}^2 \left\{3\kappa^2(4^2)+\kappa(4^4) - \frac{8}{\kappa_2}\{6\kappa(4^2)\,\kappa(4^2 2)\} + \frac{36}{\kappa_2^2}\{3\kappa^2(4^2)\,\kappa(2^2)\}\right\};$$

whence, subtracting three times the square of the mean of x^2, there remains

$$\kappa_4(x) = \left\{\frac{(n-1)(n-2)(n-3)}{24n(n+1)\kappa_2^4}\right\}^2 \left\{\kappa(4^4) - \frac{24}{\kappa_2}\kappa(4^2)\,\kappa(4^2 2) + \frac{96}{n-1}\kappa^2(4^2)\right\}.$$

The leading term in $\kappa(4^4) \div 576\kappa_2^8$ comes to $636n^{-3}$, to which the other terms add -192 and $+96$ respectively, leaving

$$\kappa_4(x) = \frac{540}{n}.$$

For the mean value of x^5 we shall need

$$\left(\frac{n}{24\kappa_2^4}\right)^{\frac{5}{2}} \left\{10\kappa(4^2)\,\kappa(4^3)+\kappa(4^5) - \frac{10}{\kappa_2}\left\{10\kappa(4^2 2)\,\kappa(4^3) + 10\kappa(4^2)\,\kappa(4^3 2)\right\} + \frac{550}{\kappa_2^2}\kappa(4^3)\,\kappa(4^2)\,\kappa(2^2)\right\},$$

whence, deducting $10\kappa_2(x)\,.\,\kappa_3(x)$, there remains

$$\kappa_5(x) = \left(\frac{n}{24\kappa_2^4}\right)^{\frac{5}{2}} \left\{\kappa(4^5) - \frac{60}{\kappa_2}\kappa(4^3)\,\kappa(4^2 2) - \frac{40}{\kappa_2}\kappa(4^3 2)\,\kappa(4^2) + \frac{240}{\kappa_2^2}\kappa(4^3)\,\kappa(4^2)\,\kappa(2^2)\right\},$$

or

$$\frac{71}{n^{\frac{3}{2}}}\,.\,144\sqrt{6}.$$

Now if ξ is normally distributed with unit variance, and x can be expressed approximately in the form

$$x = \beta\xi + \gamma(\xi^2-1) + \delta(\xi^3-3\xi) + \epsilon(\xi^4-6\xi^2+3) + \ldots,$$

we have

$$\kappa_1(x) = \mu_1(x) = 0,$$

$$\kappa_2(x) = \mu_2(x) = \beta^2+2\gamma^2+6\delta^2+24\epsilon^2+\ldots,$$

$$\kappa_3(x) = \mu_3(x) = 6\beta^2\gamma+8\gamma^3+36\beta\gamma\delta+\ldots,$$

$$\mu_4(x) = 3\beta^4+24\beta^3\delta+48\beta^2\gamma^2,$$

whence $$\kappa_4(x) = 24\beta^3\delta+48\beta^2\gamma^2,$$

and $$\mu_5(x) = 60\beta^4\gamma+120\beta^4\epsilon+1080\beta^3\gamma\delta+680\beta^2\gamma^3,$$

whence $$\kappa_5(x) = 120\beta^4\epsilon+720\beta^3\gamma\delta+560\beta^2\gamma^3;$$

and equating these to the actual values, neglecting n^{-2}, we have the translation formula

$$x = \left(1-\frac{12}{n}\right)\xi+\sqrt{\frac{6}{n}}\left(1-\frac{47}{n}\right)(\xi^2-1) +\frac{21}{2n}(\xi^3-3\xi)-\frac{29\sqrt{6}}{5n^{3/2}}(\xi^4-6\xi^2+3),$$

* or, inversely,

$$\xi = -\frac{21\sqrt{6}}{n^{3/2}}+x\left(1+\frac{36}{n}\right)-\sqrt{\frac{6}{n}}\left(1-\frac{201}{2n}\right)(x^2-1) +\frac{3}{2n}(x^3-3x)+\frac{43}{10n}\sqrt{\frac{6}{n}}(x^4-6x^2+3).$$

Summary.

The equations which connect the moment functions of the sampling distribution of moment statistics with the moment functions of the population from which the samples are drawn correspond in univariate problems to all the partitions of all the natural numbers, and in multivariate problems to all the partitions of all multipartite numbers. Very few of this system of equations have hitherto been obtained owing to the algebraical complexity of their direct evaluation. The formulae are very much simplified (i) by using the semi-invariants instead of the moments of the population, and (ii) by using the system of moment statistics, the mean sampling value of each of which is the corresponding semi-invariant. The relations which necessarily exist between the different multivariate formulae demonstrate that all of these, as well as the univariate formulae, must be derivable from a system of rules associating

* Not quite correct but the method is now superseded - R.A.F.

different two-way partitions of multipartite and unipartite numbers with corresponding functions of the sample number n.

Rules are given and illustrated which enable any term of any of these formulae to be obtained directly from an examination of the appropriate partition. Their general validity is demonstrated by a theorem which connects the moment generating function of any distribution with the corresponding function of any functionally related set of variates. Complete univariate formulae are given up to the tenth degree, and some new results are applied to the theory of samples from a normal population.

References.

1. W. F. Sheppard, "On the application of the theory of error to cases of normal distribution and normal correlation", *Phil. Trans.* (A), 192 (1897), 101–167.
2. K. Pearson, "On the probable errors of frequency constants", *Biometrika*, 2 (1902), 273–281.
3. K. Pearson, "On the probable errors of frequency constants", *Biometrika*, 9 (1913), 1–10.
4. L. Isserlis, "On certain probable errors and correlation coefficients of multiple frequency distribution with skew regression", *Biometrika*, 11 (1916), 185–190.
5. L. Isserlis, "Formulae for determining the mean value of products of deviations of mixed moment coefficients in two to eight variables in samples from a limited population", *Biometrika*, 12 (1919), 183–184.
6. H. E. Soper, "On the probable error of the correlation coefficient to the second approximation", *Biometrika*, 9 (1913), 91–115.
7. "Student", "The probable error of a mean", *Biometrika*, 6 (1908), 1–25.
8. A. A. Tchouproff, "On the mathematical expectation of the moments of frequency distributions", *Biometrika*, 12 (1919), 140–169 and 185–210.
9. A. E. R. Church, "On the moments of the distribution of squared standard deviations for samples of n drawn from an indefinitely large population", *Biometrika*, 17 (1925), 79–83.
10. R. A. Fisher, "On the mathematical foundations of theoretical statistics", *Phil. Trans.* (A), 222 (1921), 309–368.
11. J. Wishart, "A problem in combinatorial analysis giving the distribution of certain moment statistics", *Proc. London Math. Soc.* (2), 29 (1929), 309–321.