MOMENTS AND PRODUCT MOMENTS OF SAMPLING DISTRIBUTIONS

By R. A. FISHER.

[Received 5 December, 1928.—Read 13 December, 1928.]

1. *Introductory.*

If a random sample of n observations be taken from a univariate distribution, and the sample values obtained be designated by $x_1, x_2, \ldots, x_n$, then any symmetric function of these sample values of degree r may be termed a moment function of the sample of the r-th degree. If the coefficients of the symmetric function involve the sample number n in such a way that, as n tends to infinity, the value of the function tends to a finite limit, in the sense that the probability of exceeding or falling short of that limit by a positive quantity ϵ, however small, tends to zero, then the limit to which it tends is a moment function of the population sampled, and the moment function of the sample may be regarded as a statistical estimate of the corresponding moment function of the population.

If we consider the random sampling distribution of such a statistic it is evident that the moment functions of this distribution will be expressible in terms of the moment functions of the original distribution, in so far as these are finite, by means of formulae which will be independent of the nature of this distribution. For example, a moment function of degree s of the sampling distribution of a moment function of degree r will involve only symmetric functions of the observations of degree rs, and will therefore be expressible as a moment function of the population of this degree, irrespective of the moments of higher degree.

Numerous researches have been made into the moments, chiefly of the second order, of moment statistics. The algebraic method was developed by Sheppard [**1**], and used extensively by Pearson [**2**, **3**] and Isserlis [**4**, **5**]; in all these researches, however, owing to the supposition that the mean of the sample coincides with the mean of the population, or for other similar reasons, the results are only first approximations neglecting n^{-1}. In 1913 [**6**] Soper obtained a number of approximations as far as n^{-2}. In 1908 "Student" [**7**] derived an exact formula for the second moment of the variance as estimated, which corresponds in a different notation to equation (1) of this paper for the univariate case. Later, much work, by the exact algebraic method, was carried out by Tchouproff [**8**], who obtained in this way the first eight moments of the mean, in addition to the univariate formulae corresponding to numbers (5) and (14). Tchouproff's version of (14) in the univariate problem was subsequently corrected by Church [**9**]. The application of the combinatorial method developed below to the general moments of the distribution of statistics of the second degree from normal multivariate populations has already appeared in a paper by J. Wishart [**11**].

Apart from the last, these results are subject to two somewhat serious limitations; the great complexity of the results attained detracts largely from the possibility either of a theoretical comprehension of their meaning, or of numerical applications; it has also led to great difficulties in the detection of errors, which have had on more than one occasion to be corrected by subsequent workers. Secondly, partly no doubt in consequence of this complexity, attention has been almost solely confined to the direct moments of single statistics, and the product moments, specifying the simultaneous distribution of two or more statistics, have been largely neglected. The total number of formulae of degree no higher than 12 is large, and it is scarcely possible that the whole body should be made available, either for study or for use, unless an improved notation can be found which will greatly simplify the algebraic expressions. It will be shown that the formulae are much simplified by the use of the cumulative moment functions, or semi-invariants, in place of the crude moments.

The importance of the formulae lies in their generality; they are applicable to all distributions for which the expressions have a meaning. In the present state of our knowledge any information, however incomplete, as to sampling distributions is likely to be of frequent use, irrespective of the fact that moment functions only provide statistical estimates of high efficiency for a special type of distribution [**10**].

2. *The cumulative moment functions.*

If the probability that a single sample value falls in the range dx is

$$\phi(x)\, dx,$$

then the function

$$M = \int e^{tx} \phi(x)\, dx,$$

taken over all possible values of the variate x, may, or may not, have a meaning for real values of t. If it has a meaning we may expand the exponential term, and, writing

$$\mu_r = \int x^r \phi(x)\, dx,$$

we have

$$M = 1 + \mu_1 t + \mu_2 \frac{t^2}{2\,!} + \mu_3 \frac{t^3}{3\,!} + \ldots.$$

If we expand the logarithm of M in powers of t we may write

$$K = \log M = \kappa_1 t + \kappa_2 \frac{t^2}{2\,!} + \kappa_3 \frac{t^3}{3\,!} + \ldots,$$

where the cumulative moment functions κ are determinate functions of the moments μ, whether the series converges or not; moreover, since κ_r involves only μ_r, and lower orders, it follows that, if $\mu_1, \ldots, \mu_r$ are finite, so will $\kappa_1, \ldots, \kappa_r$ be finite.

The expression of κ_r in terms of μ will involve the term

$$\mu_{p_1}^{\pi_1} \mu_{p_2}^{\pi_2} \ldots \mu_{p_h}^{\pi_h}$$

corresponding to any partition

$$(p_1^{\pi_1} p_2^{\pi_2} \ldots p_h^{\pi_h})$$

of the integer r, with coefficient

$$\frac{(-)^{\rho-1}(\rho-1)\,!}{\pi_1\,!\, \pi_2\,! \ldots \pi_h\,!}\, \frac{r\,!}{(p_1\,!)^{\pi_1} (p_2\,!)^{\pi_2} \ldots (p_h\,!)^{\pi_h}},$$

where $\rho = \Sigma(\pi)$ is the number of parts.

Similarly, the expression μ_r in terms of κ will involve the term

$$\kappa_{p_1}^{\pi_1} \kappa_{p_2}^{\pi_2} \ldots \kappa_{p_h}^{\pi_h}$$

with coefficient

$$\frac{1}{\pi_1!\,\pi_2!\,\dots\,\pi_h!}\;\frac{r!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\dots(p_h!)^{\pi_h}}.$$

The simplification of moment formulae obtained by referring the moments to the mean of the distribution is due to the fact that, when $\mu_1 = 0$, no subsequent μ involves κ_1, and the number of partitions required is much reduced; thus

$$\mu_2 = \kappa_2, \qquad \mu_3 = \kappa_3, \qquad \mu_4 = \kappa_4 + 3\kappa_2^2,$$

and so on. The advantage of this simplification may be carried to higher orders by consistently using the cumulative moment functions κ in place of the moments μ.

The cumulative moment functions supply an immediate solution of the problem of the distribution of the mean, for, using the well known cumulative property, that, if x and y are independent variates,

$$K(x+y) = K(x)+K(y),$$

where $K(x)$ stands for the K function specifying the distribution of x, we find that, if $s_1 = S(x)$ is the sum of n independent values constituting a sample from a given distribution, then

$$\begin{aligned} K(s_1) &= nK(x) \\ &= n\kappa_1 t + n\kappa_2 \frac{t^2}{2!} + nK_3 \frac{t^3}{3!} + \dots; \end{aligned}$$

but the mean is $\bar{x} = (1/n)s_1$; consequently the K function of the mean is found by substituting t/n for t in the series for $K(s_1)$, giving

$$K(\bar{x}) = \kappa_1 t + \frac{\kappa_2}{n}\frac{t^2}{2!} + \frac{\kappa_3}{n^2}\frac{t^3}{3!} + \dots.$$

The value of κ_r in the distribution of the mean is thus found from that of the sampled distribution by dividing by n^{r-1}.

3. *The appropriate moment statistics.*

In order to take the full advantage of the properties of the cumulative moment functions, it is necessary to introduce a modification also into the form of the moment statistics; it is usual to employ statistics which may be written $$m_r = \frac{1}{n} S(x-\bar{x})^r,$$

which are called the moments of the sample about its mean, together with the mean itself, $\bar{x}$. These moments may be expressed in terms of the symmetric functions s_r, defined by

$$s_r = S(x^r),$$

by direct expansion; for example,

$$\bar{x} = n^{-1}s_1,$$

$$m_2 = n^{-1}s_2 - n^{-2}s_1^2,$$

$$m_3 = n^{-1}s_3 - 3n^{-2}s_1s_2 + 2n^{-3}s_1^3,$$

and so on. While the coefficients n^{-1}, n^{-2}, etc., are kept simple, we here encounter the complication that the mean value of m_s is not in finite samples equal to μ_s; in order that this should be so we should multiply m_2 by $n/(n-1)$, and m_3 by $n^2/[(n-1)(n-2)]$; further, for functions of the fourth and higher degrees, κ_r is not a linear function of the moments μ, and, in consequence, a moment statistic of which the mean is κ_r will not be exactly the same function of moment statistics, of which the means are μ_r, as κ_r is of μ_r. As a preliminary step, therefore, to the simplification of the formulae to be obtained, it will be desirable to obtain, in terms of the direct summation values s_r, the moment statistics of each degree of which the sampling means shall be $\kappa_1, \kappa_2, \kappa_3, \ldots$. They will be represented by $k_1, k_2, k_3, \ldots$.

The first few statistics which fulfil this condition are

$$k_1 = m_1 = n^{-1}s_1,$$

$$k_2 = \frac{n}{n-1} m_2 = \frac{1}{n-1} (s_2 - n^{-1}s_1^2),$$

$$k_3 = \frac{n^2}{(n-1)(n-2)} m_3 = \frac{n}{(n-1)(n-2)} (s_3 - 3n^{-1}s_1s_2 + 2n^{-2}s_1^3),$$

$$k_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \{(n+1)m_4 - 3(n-1)m_2^2\}$$

$$= \frac{n}{(n-1)(n-2)(n-3)} \{(n+1)s_4 - 4n^{-1}(n+1)s_1s_3 - 3n^{-1}(n-1)s_2^2 + 12n^{-1}s_1^2s_2 - 6n^{-2}s_1^4\}$$

$$k_5 = \frac{n^3}{(n-1)(n-2)(n-3)(n-4)} \{(n+5)\,m_5 - 10(n-1)\,m_2 m_3\}$$

$$= \frac{n^2}{(n-1)(n-2)(n-3)(n-4)}$$

$$\times \left\{(n+5)\,s_5 - 5\,\frac{n+5}{n}\,s_1 s_4 - 10\,\frac{n-1}{n}\,s_2 s_3 \right.$$

$$\left. +20\,\frac{n+2}{n^2}\,s_1^2 s_3 + 30\,\frac{n-1}{n^2}\,s_1 s_2^2 - \frac{60}{n^2}\,s_1^3 s_2 + \frac{24}{n^3}\,s_1^5 \right\},$$

$$k_6 = \frac{n^2}{(n-1)\dots(n-5)} \{(n+1)(n^2+15n-4)\,m_6 - 15(n-1)^2(n+4)\,m_2 m_4$$

$$-10(n-1)(n^2-n+4)\,m_3^2 + 30n(n-1)(n-2)\,m_2^3\}$$

$$= \frac{n}{(n-1)\dots(n-5)} \left\{(n+1)(n^2+15n-4)\,s_6 - 6\,\frac{n+1}{n}\,(n^2+15n-4)\,s_1 s_5 \right.$$

$$-15\,\frac{(n-1)^2}{n}\,(n+4)\,s_2 s_4 - 10\,\frac{n-1}{n}\,(n^2-n+4)\,s_3^2$$

$$+30\,\frac{n^2+9n+2}{n}\,s_1 s_4 + 120\,\frac{n^2-1}{n}\,s_1 s_2 s_3$$

$$+30\,\frac{(n-1)(n-2)}{n}\,s_2^3 - 120\,\frac{n+3}{n}\,s_1^3 s_3$$

$$\left. -270\,\frac{n-1}{n}\,s_1^2 s_2^2 + \frac{360}{n}\,s_1^4 s_2 - \frac{120}{n^2}\,s_1^6 \right\}.$$

If these be employed we have not only the result that the r-th cumulative moment function of the mean is $n^{-(r-1)}\kappa_r$, but also that the mean of k_r is κ_r, thus reducing a second group of the required formulae to its simplest form. It is, however, the effect of their use upon the more complex formulae which is of the greater importance. The general structure of k for any degree will be elucidated in § 10.

4. *The aggregate of moment sampling formulae.*

If we consider in its full generality the simultaneous distribution in random samples of the statistics k_1, k_2, k_3, ..., it is clear that we can represent it by means of cumulative moment functions analogous to those developed for a single variate. To any partition

$$(p_1^{\pi_1} p_2^{\pi_2} \dots p_h^{\pi_h})$$

of the number r, there will correspond a moment

$$\mu(p_1^{\pi_1} p_2^{\pi_2} \dots p_h^{\pi_h}) = \text{mean value of } k_{p_1}^{\pi_1} k_{p_2}^{\pi_2} \dots k_{p_h}^{\pi_h},$$

and, if we write

$$M = \Sigma \mu(p_1^{\pi_1} p_2^{\pi_2} \dots p_h^{\pi_h}) \frac{t_{p_1}^{\pi_1}}{\pi_1!} \frac{t_{p_2}^{\pi_2}}{\pi_2!} \dots \frac{t_{p_h}^{\pi_h}}{\pi_h!},$$

the expansion in terms of $t_1, t_2, \dots$ of $K = \log M$ assumes the form

$$K = \Sigma \kappa(p_1^{\pi_1} p_2^{\pi_2} \dots p_h^{\pi_h}) \frac{t_{p_1}^{\pi_1}}{\pi_1!} \frac{t_{p_2}^{\pi_2}}{\pi_2!} \dots \frac{t_{p_h}^{\pi_h}}{\pi_h!}.$$

There will thus be a separate formula of degree r for every partition of the number r, and for the complete specification of the distribution each must be expanded in terms of the cumulative moment functions of the sampled population. For example, the semi-invariants of the distribution of the second moment statistic k_2 will be given by the terms corresponding to the partitions (2), (2^2), (2^3), (2^4), ..., which we designate by

$$\kappa(2), \ \kappa(2^2), \ \kappa(2^3), \ \kappa(2^4), \text{ and so on.}$$

The well known solution of the distribution of the mean, given above, may now be written

$$\kappa(1^r) = \frac{\kappa_r}{n^{r-1}}, \tag{I}$$

while from the manner in which the statistics k have been constructed we have also

$$\kappa(r) = \kappa_r. \tag{II}$$

In general, the expression for the κ corresponding to any given partition of r will include a term in κ_r together with terms of the form

$$A(q_1^{\chi_1} q_2^{\chi_2} \dots q_h^{\chi_h}) \kappa_{q_1}^{\chi_1} \kappa_{q_2}^{\chi_2} \dots \kappa_{q_h}^{\chi_h},$$

where $q_1^{\chi_1} q_2^{\chi_2} \dots q_h^{\chi_h}$ is any partition of r in which no part is unity. This restriction, which greatly diminishes the number of terms to be evaluated, flows from the consideration that κ_1, unlike all other cumulative moment functions, is altered by a change of origin, and by such a change can be given any desired value, while of the moment statistics also k_1 is the only one affected by such a change, and that by addition of a quantity which is invariable from sample to sample; consequently, κ_1 can only appear in the single formula

$$\kappa(1) = \kappa_1,$$

expressing that the mean of the sample of n will be the mean of the population.

5. *Partitions involving unit parts.*

A relationship exists, of which a proof may be deduced from the general theory to be developed, which enables us to dispense with the separate examination and tabulation of the formulae corresponding to all those partitions which involve unit parts. The effect upon the corresponding formula of adding a new unit part to the partition is (1) to modify every term in the formula by increasing the suffix of one of its κ functions by unity in every possible way, and (2) to divide the whole by n. For example, the formula for the variance of k_2 is

$$\kappa(2^2) = \frac{1}{n}\kappa_4 + \frac{2}{n-1}\kappa_2^2,$$

whence we may deduce, by applying the above rules,

$$\kappa(2^2 1) = \frac{1}{n^2}\kappa_5 + \frac{4}{n(n-1)}\kappa_2\kappa_3,$$

and, by further applications,

$$\kappa(2^2 1^2) = \frac{1}{n^3}\kappa_6 + \frac{4}{n^2(n-1)}\kappa_2\kappa_4 + \frac{4}{n^2(n-1)}\kappa_3^2,$$

$$\kappa(2^2 1^3) = \frac{1}{n^4}\kappa_7 + \frac{4}{n^3(n-1)}\kappa_2\kappa_5 + \frac{12}{n^3(n-1)}\kappa_3\kappa_4,$$

and so on.

An immediate consequence of the same relationship is that

$$\kappa(r1^s) = \frac{\kappa_{r+s}}{n^s} = \frac{1}{n^s}\kappa_{r+s}. \tag{III}$$

The number of formulae remaining of any degree r is the number of partitions of r into parts of 2 or more; these are

r	4	5	6	7	8	9	10	11	12	13	14	15	16	17
partitions	1	1	3	3	6	7	11	13	20	23	33	40	54	65

Up to the 12th degree there are therefore 65 formulae, while 150 more will only reach the 16th degree. It is proposed to put on record, as a basis for discussion, the formulae up to the 10th degree, together with a few others of special interest, with an explanation of the procedure of calculation.

6. *Calculation of formulae.*

In the calculation of the formulae by the algebraic method it is desirable to proceed somewhat formally, although the results for the 4th and 5th degrees may be obtained fairly readily by writing down the algebraical expressions at length. The procedure may be illustrated by the work for the formulae of the eighth degree. There will be six of these, and corresponding to any of these, such as $\kappa(62)$, the k product, $k_6 k_2$, may be written down and expanded in the symmetric functions s. The work proceeds in three steps: (1) the mean value of the k product is expressed in terms of the population moments μ; (2) by substitution, the expression in terms of μ is condensed into its equivalent in terms of κ; (3) from the moment thus obtained, corresponding to the required partition, the corresponding cumulative moment function is found by the use of formulae of lower degree previously prepared.

The first step is carried out by means of easily verified relationships giving the mean value of such a product as $s_p s_q s_r$ in the form

$$n\mu_{p+q+r}+n(n-1)(\mu_p\mu_{q+r}+\mu_q\mu_{r+p}+\mu_r\mu_{p+q})+n(n-1)(n-2)\,\mu_p\mu_q\mu_r.$$

In order to apply these relationships expeditiously a table is prepared for each degree, showing the coefficients with which each μ product, ignoring μ_1, occurs in the expansion of each s product.

To evaluate the mean value of any k product, such as $k_3^2 k_2$, it is first expanded in s products as

$$\frac{n^2}{(n-1)^3(n-2)^2}\left(s_3^2 s_2-\frac{1}{n}\,s_3^2 s_1^2-\frac{6}{n}\,s_3 s_2^2 s_1+\frac{10}{n^2}\,s_3 s_2 s_1^3+\frac{9}{n^2}\,s_2^3 s_1^2-\frac{4}{n^3}\,s_3 s_1^5\right.$$
$$\left.-\frac{21}{n^3}\,s_2^2 s_1^4+\frac{16}{n^4}\,s_2 s_1^6-\frac{4}{n^5}\,s_1^8\right),$$

whence from a table of the separations of 8 the following table may at once be constructed.

TABLE 1.

Calculation of the mean value of $k_3^2 k_2$.

			$n(n-1)$	$n(n-1)$	$n(n-1)$	$n(n-1)(n-2)$	$n(n-1)(n-2)$	$n(n-1)(n-2)(n-3)$
		$n\mu_8$	$\mu_6\mu_2$	$\mu_5\mu_3$	μ_4^2	$\mu_4\mu_2^2$	$\mu_3^2\mu_2$	μ_2^4
$s_3^2 s_2$		1	1	2	—	—	1	—
$s_3^2 s_1^2$	n^{-1}	−1	−1	−2	2	—	−1	—
$s_3 s_2^2 s_1$	n^{-1}	−6	−12	−18	−8	−8	−12	—
$s_3 s_2 s_1^3$	n^{-2}	10	40	50	30	30	40	—
$s_2^3 s_1^2$	n^{-2}	9	36	54	27	54	54	9
$s_3 s_1^5$	n^{-3}	−4	−40	−44	−20	−60	−40	—
$s_2^2 s_1^4$	n^{-3}	−21	−168	−252	−147	−336	−420	−63
$s_2 s_1^6$	n^{-4}	16	256	416	240	960	1120	240
s_1^8	n^{-5}	−4	−112	−224	−140	−840	−1120	−420

Collecting like terms and cancelling the factors $n-1$ and $n-2$ whenever possible, we get

$$\mu(3^2 2) = \frac{\mu_8}{n^2} + \frac{n^2-8n+28}{n^2(n-1)}\mu_6\mu_2 + \frac{2n^3-12n^2+48n-56}{n^2(n-1)^2}\mu_5\mu_3$$

$$+\frac{-8n^2+25n-35}{n^2(n-1)^2}\mu_4^2 + \frac{1}{n^2(n-1)^2(n-2)}$$

$$\times\{(-6n^4+84n^3-396n^2+960n-840)\mu_4\mu_2^2$$

$$+(n^5-13n^4+94n^3-460n^2+1120n+1120)\mu_3^2\mu_2$$

$$+(9n^4-90n^3+429n^2-1140n+1260)\mu_2^4\}.$$

The second step consists in substituting

$$\mu_4 = \kappa_4 + 3\kappa_2^2,$$

$$\mu_5 = \kappa_5 + 10\kappa_2\kappa_3,$$

$$\mu_6 = \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3,$$

$$\mu_8 = \kappa_8 + 28\kappa_6\kappa_2 + 56\kappa_5\kappa_3 + 35\kappa_4^2 + 210\kappa_4\kappa_2^2 + 280\kappa_3^2\kappa_2 + 105\kappa_2^4,$$

which reduces the expression to the simpler form

$$\mu(3^2 2) = \frac{\kappa_8}{n^2} + \frac{n+20}{n(n-1)}\kappa_6\kappa_2 + \frac{2n^2+44n-64}{n(n-1)^2}\kappa_5\kappa_3 + \frac{27n-45}{n(n-1)^2}\kappa_4^2$$

$$+\frac{9n^2+81n-180}{(n-1)^2(n-2)}\kappa_4\kappa_2^2 + \frac{n^3+17n^2+104n-320}{(n-1)^2(n-2)}\kappa_3^2\kappa_2$$

$$+\frac{6n^2+30n}{(n-1)^2(n-2)}\kappa_2^4.$$

The third stage consists in removing from $\mu(3^2 2)$ those terms which do not belong to $\kappa(3^2 2)$; from the general relationship which connects these two groups of functions

$$\mu(3^2 2) = \kappa(3^2 2) + 2\kappa_3 \kappa(32) + \kappa_2 \kappa(3^2) + \kappa_3^2 \kappa_2,$$

and from formulae of lower degree already evaluated we know that

$$\kappa(32) = \frac{\kappa_5}{n} + \frac{6\kappa_2 \kappa_3}{n-1},$$

while
$$\kappa(3^2) = \frac{\kappa_6}{n} + \frac{9\kappa_2 \kappa_4}{n-1} + \frac{9\kappa_3^2}{n-1} + \frac{6n\kappa_2^3}{(n-1)(n-2)}.$$

Removing the superfluous terms we are left with

$$\kappa(3^2 2) = \frac{\kappa_8}{n^2} + \frac{21}{n(n-1)} \kappa_6 \kappa_2 + \frac{6(8n-11)}{n(n-1)^2} \kappa_5 \kappa_3 + \frac{9(3n-5)}{n(n-1)^2} \kappa_4^2$$
$$+ \frac{18(6n-11)}{(n-1)^2 (n-2)} \kappa_4 \kappa_2^2 + \frac{18(9n-20)}{(n-1)^2 (n-2)} \kappa_3^2 \kappa_2 + \frac{36n}{(n-1)^2 (n-2)} \kappa_2^4,$$

an expression in which the part played by each of the characteristic coefficients of the original distribution is clearly apparent. In the normal distribution, for example, when every coefficient beyond κ_2 vanishes, only the last term remains to be evaluated.

7. *The univariate formulae.*

In addition to the partitions involving unit parts, which have already been set aside, the numbers 4 and 5 have only one partition each, 6 and 7 have three partitions each, while 8, 9, and 10 bring the total up to 32. These are given in the following Table. Since it is scarcely to be hoped that all of these, especially the heavier formulae, will be entirely free from error, it should be particularly noted that any suspected term may be evaluated separately and independently by means of the combinatorial method elaborated below. I am indebted to Dr. J. Wishart and Prof. Hotelling for checking these formulae.

In addition to these formulae, which are complete up to the tenth degree, four others of the twelfth degree may be put on record, namely those for the variance of k_6, the third moment of k_4, fourth moment of k_3,

TABLE OF FORMULAE.

The 32 univariate formulae up to the 10-th degree.

	$\frac{\kappa_4}{n}$	$\frac{\kappa_2^2}{n-1}$			
$\kappa(2^2)$	1	2			
	$\frac{\kappa_5}{n}$	$\frac{\kappa_3\kappa_2}{n-1}$			
$\kappa(32)$	1	6			
	$\frac{\kappa_6}{n}$	$\frac{\kappa_4\kappa_2}{n-1}$	$\frac{\kappa_3^2}{n-1}$	$\frac{n\kappa_2^3}{(n-1)(n-2)}$	
$\kappa(42)$	1	8	6	—	
$\kappa(3^2)$	1	9	9	6	
	$\frac{\kappa_6}{n^2}$	$\frac{\kappa_4\kappa_2}{n(n-1)}$	$\frac{\kappa_3^2}{n(n-1)^2}$	$\frac{\kappa_2^3}{(n-1)^2}$	
$\kappa(2^3)$	1	12	$4(n-2)$	8	
	$\frac{1}{n}\kappa_7$	$\frac{1}{n-1}\kappa_5\kappa_2$	$\frac{1}{n-1}\kappa_4\kappa_3$	$\frac{n}{(n-1)(n-2)}\kappa_3\kappa_2^2$	
$\kappa(52)$	1	10	20	—	
$\kappa(43)$	1	12	30	36	
	$\frac{\kappa_7}{n^2}$	$\frac{\kappa_5\kappa_2}{n(n-1)}$	$\frac{\kappa_4\kappa_3}{n(n-1)^2}$	$\frac{\kappa_2^2\kappa_3}{(n-1)^2}$	
$\kappa(32^2)$	1	16	$12(2n-3)$	48	
	$\frac{1}{n}\kappa_8$	$\frac{1}{n-1}\kappa_6\kappa_2$	$\frac{1}{n-1}\kappa_5\kappa_3$	$\frac{1}{n-1}\kappa_4^2$	$\frac{n}{(n-1)(n-2)}\kappa_4\kappa_2^2$
$\kappa(62)$	1	12	30	20	—
$\kappa(53)$	1	15	45	30	60
$\kappa(4^2)$	1	16	48	34	72
	$\frac{1}{n}\kappa_8$	$\frac{\kappa_6\kappa_2}{n(n-1)}$	$\frac{\kappa_5\kappa_3}{n(n-1)^2}$	$\frac{\kappa_4^2}{n(n-1)^2}$	$\frac{\kappa_4\kappa_2^2}{(n-1)^2(n-2)}$
$\kappa(42^2)$	1	20	$3(5n-7)$	$4(7n-10)$	$80(n-2)$
$\kappa(3^2 2)$	1	21	$6(8n-11)$	$9(3n-5)$	$18(6n-11)$
	$\frac{\kappa_8}{n^3}$	$\frac{\kappa_6\kappa_2}{n^2(n-1)}$	$\frac{\kappa_5\kappa_3}{n^2(n-1)^2}$	$\frac{\kappa_4^2}{n^2(n-1)3}$	$\frac{\kappa_4\kappa_2^2}{n(n-1)^2}$
$\kappa(2^4)$	1	24	32	$8(4n^2-9n+6)$	144
	$\frac{\kappa_9}{n}$	$\frac{\kappa_7\kappa_2}{n-1}$	$\frac{\kappa_6\kappa_3}{n-1}$	$\frac{\kappa_5\kappa_4}{n-1}$	$\frac{n\kappa_5\kappa_2^2}{(n-1)(n-2)}$
$\kappa(72)$	1	14	42	70	—
$\kappa(63)$	1	18	63	105	90
$\kappa(54)$	1	20	70	120	120
	$\frac{\kappa_9}{n^2}$	$\frac{\kappa_7\kappa_2}{n(n-1)}$	$\frac{\kappa_6\kappa_3}{n(n-1)^2}$	$\frac{\kappa_5\kappa_4}{n(n-1)^2}$	$\frac{\kappa_5\kappa_2^2}{(n-1)^2(n-2)}$
$\kappa(52^2)$	1	24	$20(3n-4)$	$20(5n-7)$	$120(n-2)$
$\kappa(432)$	1	26	$24(3n-4)$	$10(11n-17)$	$36(5n-9)$
$\kappa(3^3)$	1	27	$27(3n-4)$	$27(4n-7)$	$54(4n-7)$
	$\frac{\kappa_9}{n^3}$	$\frac{\kappa_7\kappa_2}{n^2(n-1)}$	$\frac{\kappa_6\kappa_3}{n^2(n-1)^2}$	$\frac{\kappa_5\kappa_4}{n^2(n-1)^3}$	$\frac{\kappa_5\kappa_2^2}{n(n-1)^2}$
$\kappa(32^3)$	1	30	$2(31n-53)$	$12(9n^2-23n+16)$	240

			(1)
			(2)
			(3)
			(4)
			(5)
			(6)
			(7)
			(8)
$\frac{n}{(n-1)(n-2)}\kappa_3^2\kappa_2$	$\frac{n(n+1)}{(n-1)(n-2)(n-3)}\kappa_2^4$		
—	—		(9)
90	—		(10)
144	24		(11)
$\frac{\kappa_3^2\kappa_2}{(n-1)^2(n-2)}$	$\frac{\kappa_2^4}{(n-1)^2(n-2)}$		
$120(n-2)$	—		(12)
$18(9n-20)$	$36n$		(13)
$\frac{\kappa_3^2\kappa_2}{2(n-1)^3}$	$\frac{\kappa_2^4}{(n-1)^3}$		
$96(n-2)$	48		(14)
$\frac{n\kappa_4\kappa_3\kappa_2}{(n-1)(n-2)}$	$\frac{n\kappa_3^3}{(n-1)(n-2)}$	$\frac{n(n+1)\kappa_3\kappa_2^3}{(n-1)(n-2)(n-3)}$	
—	—	—	(15)
360	90	—	(16)
600	180	240	(17)
$\frac{\kappa_4\kappa_3\kappa_2}{(n-1)^2(n-2)}$	$\frac{\kappa_3^3}{(n-1)^2(n-2)^2}$	$\frac{n\kappa_3\kappa_2^3}{(n-1)^2(n-2)^2}$	
$480(n-2)$	$120(n-2)^2$	—	(18)
$12(61n-128)$	$36(n-2)(5n-12)$	$360(n-2)$	(19)
$162(5n-12)$	$36(7n^2-30n+34)$	$108(5n-12)$	(20)
$\frac{\kappa_4\kappa_3\kappa_2}{n(n-1)^3}$	$\frac{\kappa_3^3}{n(n-1)^3}$	$\frac{\kappa_3\kappa_2^3}{(n-1)^3}$	
$360(2n-3)$	$24(5n-12)$	336	(21)

TABLE OF FORMULAE—*continued.*

	$\frac{\kappa_{10}}{n}$	$\frac{\kappa_8\kappa_2}{n-1}$	$\frac{\kappa_7\kappa_3}{n-1}$	$\frac{\kappa_6\kappa_4}{n-1}$	$\frac{\kappa_5^2}{n-1}$	$\frac{n\kappa_6\kappa_2^2}{(n-1)(n-2)}$	$\frac{n\kappa_5\kappa_3\kappa_2}{(n-1)(n-2)}$
$\kappa(82)$	1	16	56	112	70	—	—
$\kappa(73)$	1	21	84	168	105	126	630
$\kappa(64)$	1	24	96	194	120	180	1080
$\kappa(5^2)$	1	25	100	200	125	200	1200
	$\frac{\kappa_{10}}{n^2}$	$\frac{\kappa_8\kappa_2}{n-1}$	$\frac{\kappa_7\kappa_3}{n(n-1)^2}$	$\frac{\kappa_6\kappa_4}{n(n-1)^2}$	$\frac{\kappa_5^2}{n(n-1)^2}$	$\frac{\kappa_6\kappa_2^2}{(n-1)^2(n-1)}$	$\frac{\kappa_5\kappa_3\kappa_2}{(n-1)^2(n-2)}$
$\kappa(62^2)$	1	28	$12(7n-9)$	$4(41n-56)$	$20(5n-7)$	$168(n-2)$	$840(n-2)$
$\kappa(532)$	1	31	$101n-131$	$5(37n-55)$	$5(23n-35)$	$30(9n-16)$	$30(45n-92)$
$\kappa(4^22)$	1	32	$8(13n-17)$	$4(49n-73)$	$4(29n-46)$	$8(37n-65)$	$1536(n-2)$
$\kappa(43^2)$	1	33	$6(19n-25)$	$3(65n-107)$	$6(19n-34)$	$18(19n-33)$	$72(23n-52)$
	$\frac{\kappa_{10}}{n^3}$	$\frac{\kappa_8\kappa_2}{n^2(n-1)}$	$\frac{\kappa_7\kappa_3}{n^2(n-1)^2}$	$\frac{\kappa_6\kappa_4}{n^2(n-1)^3}$	$\frac{\kappa_5^2}{n^2(n-1)^3}$	$\frac{\kappa_6\kappa_2^2}{n(n-1)^2(n-2)}$	$\frac{\kappa_5\kappa_3\kappa_2}{n(n-1)^3(n-2)}$
$\kappa(42^3)$	1	36	$4(23n-37)$	$4(47n^2-120n+81)$	$12(9n^2-24n+17)$	$360(n-2)$	$288(5n-7)(n-$
$\kappa(3^22^2)$	1	37	$6(17n-27)$	$3(61n^2-166n+117)$	$2(59n^2-154n+113)$	$6(67n-131)$	$24(71n^2-246n+2$
	$\frac{\kappa_{10}}{n^4}$	$\frac{\kappa_8\kappa_2}{n^3(n-1)}$	$\frac{\kappa_7\kappa_3}{n^3(n-1)^2}$	$\frac{\kappa_6\kappa_4}{n^3(n-1)^3}$	$\frac{\kappa_5^2}{n^3(n-1)^4}$	$\frac{\kappa_6\kappa_2^2}{n^2(n-1)^2}$	$\frac{\kappa_5\kappa_3\kappa_2}{n^2(n-1)^3}$
$\kappa(2^5)$	1	40	$80(n-2)$	$40(5n^2-12n+9)$	$16(n-2)(6n^2-12n+17)$	480	$1280(n-2)$

and the sixth moment of k_2. These are :—

$$\kappa(6^2) = \frac{1}{n}\kappa_{12}+\frac{1}{n-1}(36\kappa_{10}\kappa_2+180\kappa_9\kappa_3+465\kappa_8\kappa_4+780\kappa_7\kappa_5+461\kappa_6^2)$$

$$+\frac{n}{(n-1)(n-2)}(450\kappa_8\kappa_2^2+3600\kappa_7\kappa_3\kappa_2+7200\kappa_6\kappa_4\kappa_2+6300\kappa_6\kappa_3^2$$

$$+4500\kappa_5^2\kappa_2+21600\kappa_5\kappa_4\kappa_3+4950\kappa_4^3)$$

$$+\frac{n(n+1)}{(n-1)(n-2)(n-3)}(2400\kappa_6\kappa_2^3+21600\kappa_5\kappa_3\kappa_2^2$$

$$+15300\kappa_4^2\kappa_2^2+54000\kappa_4\kappa_3^2\kappa_2+8100\kappa_3^4)$$

$$+\frac{n^2(n+5)}{(n-1)(n-2)(n-3)(n-4)}(5400\kappa_4\kappa_2^4+21600\kappa_3^2\kappa_2^3)$$

$$+\frac{n(n+1)(n^2+15n-4)}{(n-1)(n-2)(n-3)(n-4)(n-5)}720\kappa_2^6. \tag{50}$$

$$\kappa(4^3) = \frac{1}{n^2}\kappa_{12}+\frac{48}{n(n-1)}\kappa_{10}\kappa_2\ \frac{16(13n-17)}{n(n-1)^2}\kappa_9\kappa_3+\frac{12(41n-65)}{n(n-1)^2}\kappa_8\kappa_4$$

$$+\frac{48(16n-29)}{n(n-1)^2}\kappa_7\kappa_5+\frac{12(37n-70)}{n(n-1)^2}\kappa_6^2+\frac{72(11n-19)}{(n-1)^2(n-2)}\kappa_8\kappa_2^2+\ldots$$ (p. 213)

$\dfrac{n\kappa_4^2\kappa_2}{(n-1)(n-2)}$	$\dfrac{n\kappa_4\kappa_3^2}{(n-1)(n-2)}$	$\dfrac{n(n+1)\kappa_4\kappa_2^3}{(n-1)(n-2)(n-3)}$	$\dfrac{n(n+1)\kappa_3^2\kappa_2^3}{(n-1)(n-2)(n-3)}$	$\dfrac{n^2(n+5)\kappa_2^5}{(n-1)(n-2)(n-3)(n-4)}$	
—	—	—	—	—	(22)
420	630	—	—	—	(23)
720	1260	480	1080	—	(24)
850	1200	600	1800	120	(25)
$\dfrac{\kappa_4^2\kappa_2}{(n-1)^2(n-2)}$	$\dfrac{\kappa_4\kappa_3^2}{(n-1)^2(n-2)}$	$\dfrac{\kappa_4\kappa_2}{(n-1)^2(n-2)}$	$\dfrac{\kappa_3^2\kappa_2^2}{(n-1)^2(n-2)}$	$\dfrac{n\kappa_2^5}{(n-1)^2(n-2)}$	
$560(n-2)$	$840(n-2)$	—	—	—	(26)
$60(15n-31)$	$30(45n-103)$	$720n$	$1620n$	—	(27)
$144(7n-15)$	$72(21n-50)$	$\dfrac{96(10n^2-27n-1)}{n-3}$	$\dfrac{144(17n^2-53n-2)}{n-3}$	$\dfrac{192(n+1)}{n-3}$	(28)
$54(19n-48)$	$\dfrac{54(33n^2-148n+172)}{n-2}$	$\dfrac{72n(17n-40)}{n-2}$	$\dfrac{108n(27n-70)}{n-2}$	$\dfrac{216n}{n-2}$	(29)
$\dfrac{\kappa_4^2\kappa_2}{n(n-1)^3(n-2)}$	$\dfrac{\kappa_4\kappa_3^2}{n(n-1)^3(n-2)}$	$\dfrac{\kappa_4\kappa_2^3}{(n-1)^3(n-2)}$	$\dfrac{\kappa_3^2\kappa_2^2}{(n-1)^3(n-2)}$	$\dfrac{n\kappa_2^5}{(n-1)^3(n-2)}$	
$144(7n-10)(n-2)$	$24(49n-95)(n-2)$	$768(n-2)$	$1872(n-2)$	—	(30)
$36(29n^2-103n+93)$	$36(38n^2-155n+160)$	$72(14n-23)$	$144(19n-44)$	288	(31)
$\dfrac{\kappa_4^2\kappa_2}{n^2(n-1)^4}$	$\dfrac{\kappa_4\kappa_3^2}{n^2(n-1)^4}$	$\dfrac{\kappa_4\kappa_2^3}{n(n-1)^3}$	$\dfrac{\kappa_3^2\kappa_2^2}{n(n-1)^4}$	$\dfrac{\kappa_2^5}{(n-1)^4}$	
$320(4n^2-9n+6)$	$480(2n^2-7n+6)$	1920	$1920(n-2)$	384	(32)

$$+\frac{288(19n-41)}{(n-1)^2(n-2)}\kappa_7\kappa_3\kappa_2+\frac{48(203n-523)}{(n-1)^2(n-2)}\kappa_6\kappa_4\kappa_2$$

$$+\frac{144(56n^2-257n+302)}{(n-1)^2(n-2)^2}\kappa_6\kappa_3^2+\frac{1440(4n-11)}{(n-1)^2(n-2)}\kappa_5^2\kappa_2$$

$$+\frac{1152(22n^2-106n+133)}{(n-1)^2(n-2)^2}\kappa_5\kappa_4\kappa_3+\frac{8(709n^2-3430n+4456)}{(n-1)^2(n-2)^2}\kappa_4^3$$

$$+\frac{288(19n^3-98n^2+125n+2)}{(n-1)^2(n-2)^2(n-3)}\kappa_6\kappa_2^3+\frac{1728(21n^3-119n^2+164n+4)}{(n-2)^2(n-2)^2(n-3)}\kappa_5\kappa_3\kappa_2^2$$

$$+\frac{432(49n^3-287n^2+408n+12)}{(n-1)^2(n-2)^2(n-3)}\kappa_4^2\kappa_2^2+\frac{864(103n^3-629n^2+984n+24)}{(n-1)^2(n-2)^2(n-3)}\kappa_4\kappa_3^2\kappa_2$$

$$+\frac{288(41n^4-384n^3+1209n^2-1282n-36)}{(n-1)^2(n-2)^2(n-3)^2}\kappa_3^4+\frac{288(89n^2-323n-88)n}{(n-1)^2(n-2)^2(n-3)}\kappa_4\kappa_2^4$$

$$+\frac{1728(29n^3-196n^2+317n+62)n}{(n-1)^2(n-2)^2(n-3)^2}\kappa_3^2\kappa_2^3+\frac{1728(n^2-5n+2)(n+1)n}{(n-1)^2(n-2)^2(n-3)^2}\kappa_2^6, \quad (57)$$

$$\kappa(3^4) = \frac{1}{n^3}\kappa_{12} + \frac{54}{n^2(n-1)}\kappa_{10}\kappa_2 \; \frac{108(2n-3)}{n^2(n-1)^2}\kappa_9\kappa_3 + 27\,\frac{17n^2-49n+35}{n^2(n-1)^3}\kappa_8\kappa_4$$

$$+108\,\frac{7n^2-20n+16}{n^2(n-1)^3}\kappa_7\kappa_5 + 27\,\frac{17n^2-47n+39}{n^2(n-1)^3}\kappa_6^2 + 27\,\frac{37n-70}{n(n-1)^2(n-2)}\kappa_8\kappa_2^2$$

$$+324\,\frac{19n^2-67n+54}{n(n-1)^3(n-2)}\kappa_7\kappa_3\kappa_2 + 162\,\frac{65n^2-245n+234}{n(n-1)^2(n-2)}\kappa_6\kappa_4\kappa_2$$

$$+108\,\frac{82n^3-481n^2+958-640}{n(n-1)^3(n-2)^2}\kappa_6\kappa_3^2 + 108\,\frac{59n^2-220n+224}{n(n-1)^3(n-2)}\kappa_5^2\kappa_2$$

$$+324\,\frac{75n^3-473n^2+1016n-756}{n(n-1)^3(n-2)^2}\kappa_5\kappa_4\kappa_3$$

$$+27\,\frac{173n^4-1503n^3+4962n^3-7380n+4200}{n(n-1)^3(n-2)^3}\kappa_4^3$$

$$+108\,\frac{71n^2-263n+234}{(n-1)^3(n-2)^2}\kappa_6\kappa_2^3 + 648\,\frac{79n^2-343n+378}{(n-1)^3(n-2)^2}\kappa_5\kappa_3\kappa_2^2$$

$$+486\,\frac{63n^2-290n+352}{(n-1)^3(n-2)^3}\kappa_4^2\kappa_2^2 + 972\,\frac{99n^3-688n^2+1612n-1280}{(n-1)^3(n-2)^3}\kappa_4\kappa_3^2\kappa_2$$

$$+162\,\frac{87n^3-594n^2+1420n-1176}{(n-1)^3(n-3)^3}\kappa_3^4 + 972\,\frac{29n^2-121n+118}{(n-1)^3(n-2)^3}\kappa_4\kappa_2^4$$

$$+648n\,\frac{103n^2-510n+640}{(n-1)^3(n-2)^3}\kappa_2^2\kappa_2^3 + 648n^2\,\frac{5n-12}{(n-1)^3(n-2)^3}\kappa_2^6, \qquad (62)$$

$$\kappa(2^6) = \frac{1}{n^5}\kappa_{12} + \frac{60}{n^4(n-1)}\kappa_{10}\kappa_2 + \frac{160(n-2)}{n^4(n-1)^2}\kappa_9\kappa_3 + 240\,\frac{2n^2-5n+4}{n^4(n-1)^3}\kappa_8\kappa_4$$

$$+96(n-2)\,\frac{7n^2-14n+9}{n^4(n-1)^4}\kappa_7\kappa_5 + 4\,\frac{113n^4-520n^3+950n^2-800n+265}{n^4(n-1)^5}\kappa_6^2$$

$$+\frac{1200}{n^3(n-1)^2}\kappa_8\kappa_2^2 + 4800\,\frac{n-2}{n^3(n-1)^3}\kappa_7\kappa_3\kappa_2 + 2400\,\frac{5n^2-12n+9}{n^3(n-1)^4}\kappa_6\kappa_4\kappa_2$$

$$+160(n-2)\,\frac{31n-53}{n^3(n-1)^4}\kappa_6\kappa_3^2 + 960(n-2)\,\frac{6n^2-12n+7}{n^3(n-1)^5}\kappa_5^2\kappa_2$$

$$+1920(n-2)\,\frac{9n^2-23n+16}{n^3(n-1)^5}\kappa_5\kappa_4\kappa_3 + 480\,\frac{11n^3-41n^2+59n-31}{n^3(n-1)^5}\kappa_4^3$$

$$+\frac{9600}{n^2(n-1)^3}\kappa_6\kappa_2^3 + \frac{38400(n-2)}{n^3(n-1)^4}\kappa_5\kappa_3\kappa_2^2 + 9600\,\frac{4n^2-9n+6}{n^2(n-1)^5}\kappa_4^2\kappa_2^2$$

$$+28800\,\frac{2n^2-7n+6}{n^2(n-1)^5}\kappa_4\kappa_3^2\kappa_2 + 960(n-2)\,\frac{5n-12}{n^2(n-1)^5}\kappa_3^4 + \frac{28800}{n(n-1)^4}\kappa_4\kappa_2^4$$

$$+38400\,\frac{n-2}{n(n-1)^5}\kappa_3^2\kappa_2^3 + \frac{3840}{(n-1)^5}\kappa_2^6. \qquad (65)$$

Some idea of the advantage of using the cumulative moment functions in place of the moments will be obtained by comparing the above formula (14) with the corresponding formula as obtained by Tchouproff, and corrected by Church :

$$
\begin{aligned}
{}_2M_4 = \frac{3}{n^2}(\mu_4-\mu_2^2)^2+\frac{1}{n^3}(\mu_8-4\mu_6\mu_2-24\mu_5\mu_3-15\mu_4^2 \\
+48\mu_4\mu_2^2+96\mu_3^2\mu_2-30\mu_2^4) \\
-\frac{1}{n^4}(4\mu_8-40\mu_6\mu_2-96\mu_5\mu_3-54\mu_4^2+336\mu_4\mu_2^2+528\mu_3^2\mu_2-306\mu_2^4) \\
-\frac{1}{n^5}(6\mu_8-96\mu_6\mu_2-176\mu_5\mu_3-102\mu_4^2+924\mu_4\mu_2^2+1232\mu_3^2\mu_2-1044\mu_2^4) \\
-\frac{1}{n^6}(4\mu_8-88\mu_6\mu_2-160\mu_5\mu_3-95\mu_4^2+1050\mu_4\mu_2^2+1360\mu_3^2\mu_2-1395\mu_2^4) \\
+\frac{1}{n^7}(\mu_8-28\mu_6\mu_2-56\mu_5\mu_3-35\mu_4^2+420\mu_4\mu_2^2+560\mu_3^2\mu_2-630\mu_2^4).
\end{aligned}
$$

The term involving κ_2 only in $\kappa(2^r)$ is already known as the r-th semi-invariant of the distribution of the variance for samples from the normal curve, and is simply $2^{r-1}.(r-1)!/(n-1)^{r-1}$. The corresponding term in $\kappa(3^4)$ is of interest as showing that the distribution of k_3 in samples from a normal distribution, though necessarily symmetrical, yet tends somewhat slowly to normality. Comparing the first term of (17) with that of (4) it is evident that

$$
\frac{\kappa(3^4)}{\{\kappa(3^2)\}^2}=\frac{18(5n-12)}{(n-1)(n-2)},
$$

or is somewhat greater than $90/n$, a fact which indicates that the occurrence of values of k_3 greater than 2 or 3 times its standard error will, except in very large samples, be materially more frequent than one would judge from an assumed normal distribution. The effect upon tests of normality will be examined in §§ 11 and 12.

8. *Bivariate and multivariate distributions.*

The extension to bivariate and multivariate data of the methods of classification and calculation developed above is of both practical and theoretical importance. Apart from the variance the product moment of a bivariate distribution is the most important of all moment statistics.

Moreover, the multivariate formulae, by reason of their greater number, and the confusion caused by the various possible notations in which they may be expressed, are particularly in need of orderly classification. It will be found, in addition, that the examination of the multivariate formulae in their generality throws much light on the expressions already obtained.

It will be seen that just as the univariate formulae correspond to all the possible partitions of unipartite numbers, so the multivariate formulae correspond to all the possible partitions of multipartite numbers, having multiplicities equal to the number of variates.

To make the notation clear let us consider, in the first place, two variates only, and let the frequency with which the two variates x and y fall simultaneously in the ranges dx and dy be

$$df = \phi\, dx\, dy,$$

in which ϕ is the simultaneous frequency function of x and y.

The general moment about any origin is defined as

$$\mu_{pq} = \iint x^p y^q \phi\, dx\, dy$$

over the whole range of possible values of the variates. So far as these moments have a meaning we can build up the expression

$$M = \sum_{p=0} \sum_{q=0} \mu_{pq} \frac{t_1^p}{p\,!} \frac{t_2^q}{q\,!},$$

and equally, with the same limitation, the coefficients of the expression

$$K = \log M = \sum_{p=0} \sum_{q=0} \kappa_{pq} \frac{t_1^p}{p\,!} \frac{t_2^q}{q\,!}$$

will be well defined. The general expressions connecting the cumulative moment functions, κ, with the moments, μ, of the simultaneous distribution are analogous to those given for univariate distribution; if

$$\{(p_1 p_1')^{\pi_1} (p_2 p_2')^{\pi_2} \dots\}$$

is any partition of the bipartite number r, s consisting of ρ parts,

$$\kappa_{r,s} = S\left\{\frac{(-)^{\rho-1}(\rho-1)\,!}{\pi_1!\,\pi_2!\dots} \frac{r\,!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\dots} \frac{s\,!}{(p_1'!)^{\pi_1}(p_2'!)^{\pi_2}\dots} \mu_{p_1 p_1'}^{\pi_1} \mu_{p_2 p_2'}^{\pi_2} \dots\right\}$$

and

$$\mu_{r,s} = S\left\{\frac{1}{\pi_1!\,\pi_2!\dots} \frac{r\,!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\dots} \frac{s\,!}{(p_1'!)^{\pi_1}(p_2')^{\pi_2}\dots} \kappa_{p_1 p_1'}^{\pi_1} \kappa_{p_2 p_2'}^{\pi_2} \dots\right\},$$

the summation being taken over all possible partitions. For any sample, we may define s_{pq} as the sum of the values of $x^p y^q$ for each pair of values in the sample, and obtain, as for single variates, the statistics k_{11}, k_{21}, k_{31}, k_{22}, etc., as expressions in terms of these sums, with such coefficients that the mean value of k_{pq} shall be κ_{pq}. Thus we have

$$k_{11} = \frac{1}{n-1}\left(s_{11} - \frac{1}{n} s_{10} s_{01}\right),$$

$$k_{21} = \frac{n}{(n-1)(n-2)}\left(s_{21} - \frac{2}{n} s_{10} s_{11} - \frac{1}{n} s_{20} s_{01} + \frac{2}{n^2} s_{10}^2 s_{01}\right),$$

$$k_{31} = \frac{n}{(n-1)(n-2)(n-3)}\left\{(n+1) s_{31} - \frac{n+1}{n} s_{30} s_{01} - \frac{3(n-1)}{n} s_{11} s_{20} \right.$$
$$\left. - \frac{3(n+1)}{n} s_{21} s_{10} + \frac{6}{n} s_{11} s_{10}^2 + \frac{6}{n} s_{20} s_{10} s_{01} - \frac{6}{n^2} s_{01} s_{10}^3\right\},$$

$$k_{22} = \frac{n}{(n-1)(n-2)(n-3)}\left\{(n+1) s_{22} - 2\frac{n+1}{n} s_{21} s_{01} - \frac{n+1}{n} s_{12} s_{10} \right.$$
$$- \frac{n-1}{n} s_{20} s_{02} - 2\frac{n-1}{n} s_{11}^2 + \frac{8}{n} s_{11} s_{01} s_{10}$$
$$\left. + \frac{2}{n} s_{02} s_{10}^2 + \frac{2}{n} s_{20} s_{01}^2 - \frac{6}{n^2} s_{10}^2 s_{01}^2\right\}.$$

The mean value of any product involving such statistics, as, for example, $k_{20} k_{11}$, may be evaluated in terms of the cumulative moment functions of the bivariate distribution; such mean values may be written

$$\overline{k_{20} k_{11}} \equiv \mu \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix},$$

giving one line to each variate; its value is easily found to be

$$\frac{1}{n} \kappa_{31} + \frac{n+1}{n-1} \kappa_{20} \kappa_{11}.$$

Hence, subtracting the product of the mean values, $\kappa_{20} \kappa_{11}$, we have the formula

$$\kappa \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{31} + \frac{2}{n-1} \kappa_{20} \kappa_{11} \tag{1 a}$$

in which each column represents the particular statistic entering into the product, and the marginal column found by summing each row is the multipartite number (31) representing the degree in which each variate is involved. Similarly, we may deduce the two formulae for partitions 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 (1*a*), (1*b*), and (1*c*) 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 (1*a*), (1*b*), or (1*c*) by identifying the variates, so, by equating appropriate variates, (1*a*) may be derived from (1*d*), or (1*b*) from (1*d*), or (1*c*) from (1*e*), 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^h). Thus of the 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 pair 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

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

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

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

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

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

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

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

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

$$4(19n-33)\,\kappa_{33}\kappa_{11}^2+8(11n-20)\,\kappa_{33}\kappa_{20}\kappa_{02}+8(7n-12)\,\kappa_{42}\kappa_{11}\kappa_{02}$$
$$+8(7n-12)\,\kappa_{24}\kappa_{11}\kappa_{20}+2(5n-9)\,\kappa_{31}\kappa_{20}^2+2(5n-9)\,\kappa_{15}\kappa_{20}^2,$$

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 $1/(n^{p-1})(1-a^{p-1})$.

Three rows.

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

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

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

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

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

Four rows.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Five and six row patterns.

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

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

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

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

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

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

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

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

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

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

the coefficients of which give the mean value 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

$$\xi = f(x) = c_0+c_1x+c_2x^2+\ldots,$$

then the mean value of ξ is

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

which may be written

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

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,

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

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, ... corresponding to the variates x, y, ...; the new variates will be given functions of the old

$$\xi_1 = f_1(x, y, \ldots),$$
$$\xi_2 = f_2(x, y, \ldots),$$
$$\ldots \quad \ldots \quad \ldots$$

and the operative expression for the transformation of M is

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

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.*

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

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

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

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

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

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

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

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

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

where $(\nu_1, \nu_2, \ldots)$ 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

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

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

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 $t_1, \ldots, t_n$. This 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 $n(n-1)\dots(n-a+1)$, 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 $\kappa(p_1^{\pi_1} p_2^{\pi_2}\dots)$, 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 :

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

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

$$\mu_6(x) = \frac{(n-1)^3(n-2)^3}{216n^3\kappa_2^9}\left\{15\kappa^3(3^2)+5\kappa(3^4)\kappa(3^2)+\kappa(3^6)\right.$$
$$-\frac{9}{\kappa_2}\{45\kappa(3^2 2)\kappa^2(3^2)+15\kappa(3^4 2)\kappa(3^2)+15\kappa(3^4)\kappa(3^2 2)\}$$
$$+\frac{45}{\kappa_2^2}\{15\kappa^3(3^2)\kappa(2^2)+90\kappa^2(3^2 2)\kappa(3^2)+45\kappa(3^2 2^2)\kappa^2(3^2)$$
$$+15\kappa(3^4)\kappa(3^2)\kappa(2^2)\}$$
$$-\frac{165}{\kappa_2^3}\{15\kappa^3(3^2)\kappa(2^3)+135\kappa(3^2 2)\kappa^2(3^2)\kappa(2^2)$$
$$\left.+\frac{495}{\kappa_2^4}\{45\kappa^3(3^2)\kappa^2(2^2)\}\right\}.$$

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

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

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

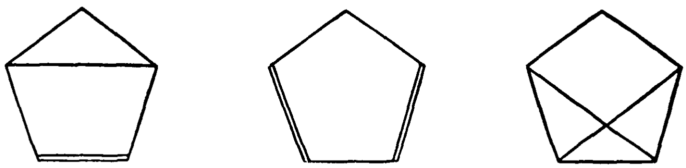
and substituting these values, we find

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

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

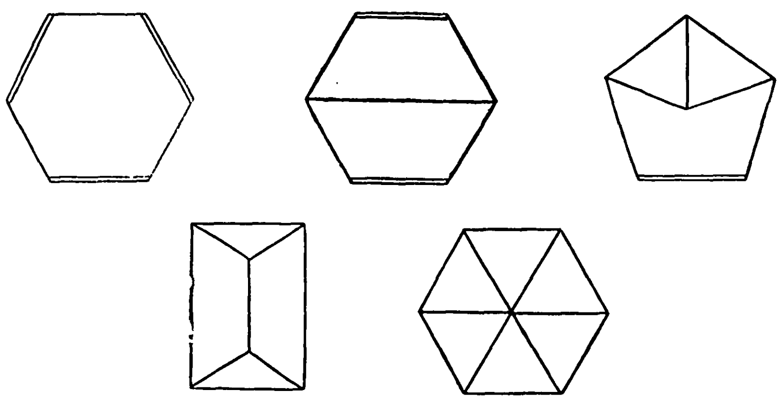


which can be done in 15548, 7774 and 15548 ways respectively, showing that $\kappa(3^4 2)$ from normal samples is approximately $38880n^{-4}$.

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 tc 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

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

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 $\sqrt{(n/6)}\,k_3 k_4^{-\frac{3}{2}}$, which latter gives apparently the better values for deviations exceeding $2\cdot 0$. 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^{-2}, 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

$$x = \sqrt{\left(\frac{(n-1)(n-2)(n-3)}{24n(n+1)}\right)}\,k_4 k_2^{-2}$$

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

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

or

$$1-\frac{32}{n-1}+\frac{20}{n-1}$$

as far as n^{-1}.

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

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

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},

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

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

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

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

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

or

$$\frac{91}{n^{\frac{3}{2}}}\,.\,144\sqrt{6}.$$

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

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

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

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

$$x = \left(1-\frac{12}{n}\right)\xi+\sqrt{\frac{6}{n}}\left(1-\frac{159}{2n}\right)(\xi^2-1)$$
$$+\frac{21}{2n}(\xi^3-3\xi)+\frac{91\sqrt{6}}{5n^{\frac{3}{2}}}(\xi^4-6\xi^2+3),$$

or, inversely,

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

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

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.