74

MOMENTS AND PRODUCT MOMENTS OF SAMPLING DISTRIBUTIONS

Author's Note (CMS 20.198a) *

To study frequency distributions with generality a number of writers, such as Thiele, had been led to give attention to the symmetric functions of samples from such distributions. The most obvious of these are the moments. It was gradually realised that the symmetric functions of the distributions of such functions were necessarily expressible in terms of corresponding functions of the parent distribution. That is that a symmetric function of degree s of a symmetric function of degree r of a sample must be expressible as a symmetric function of degree rs. The pioneers in the development of the relevant formulae were Sheppard and "Student," but by about 1919 an immense amount of algebraic material of this sort had been published by Tchrouproff, using what appeared to the author a very clumsy approach. C. C. Craig had indeed called attention to the need, if the algebraic formulation was to be made manageable, of the use of functions other than crude moments.

In this paper are defined the functions which provide the necessary simplification, namely the symmetric functions k, the mean values of which are unconditionally equal to the cumulants of the parent distribution. On the basis of the simpler forms so obtained for some of the expressions already known, and others which are made easily accessible, the paper develops an approach in which the mechanical simplification of overwhelming algebraic formulae is replaced by a consideration of the properties of certain bipartitional functions, which, apart from the sample number n, are purely arithmetical.

 Proceedings of the London Mathematical Society, Series 2, 30: 199-238, (1929).

This form of approach has the advantage that it is immediately applicable to the complex extension offered by bivariate and multivariate distributions, for we have merely to consider under the same rules of procedure the bipartitions of multipartite numbers to obtain for them equivalent formulae. Section 11 on measures of departure from normality may be ignored, as it has been superseded by a more exact treatment (Paper 83). Complete univariate formulae are given up to the 10th degree; these afford a commodious check for developing any multivariate formulae that may be required.

MOMENTS AND PRODUCT MOMENTS OF SAMPLING DISTRIBUTIONS

By R. A. FISHER.

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

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

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

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

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 $\mu_1, \dots, \mu_r$ are finite, so will $\kappa_1, \dots, \kappa_r$ be finite.

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

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

corresponding to any partition

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

of the integer r, with coefficient

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

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

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

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

with coefficient

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

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

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

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

$$\begin{aligned} K(s_1) &= nK(x) \\ &= n\kappa_1 t+n\kappa_2\frac{t^2}{2!}+n\kappa_3\frac{t^3}{3!}+\ldots; \end{aligned}$$

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

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

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,

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

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

$$m_3 = n^{-1} s_3 - 3n^{-2} s_1 s_2 + 2n^{-3} s_1^3,$$

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 κ_1, κ_2, κ_3, ... They will be represented by k_1, k_2, k_3, ...

The first few statistics which fulfil this condition are

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

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

$$k_3 = \frac{n^2}{(n-1)(n-2)} m_3 = \frac{n}{(n-1)(n-2)} (s_3 - 3n^{-1} s_1 s_2 + 2n^{-2} s_1^3),$$

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

$$= \frac{n}{(n-1)(n-2)(n-3)} \{(n+1) s_4 - 4n^{-1}(n+1) s_1 s_3 - 3n^{-1}(n-1) s_2^2 + 12n^{-1} s_1^2 s_2 - 6n^{-2} s_1^4\},$$

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

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

$$\times \left\{(n+5)\,s_5 - 5\,\frac{n+5}{n}\,s_1 s_4 - 10\,\frac{n-1}{n}\,s_2 s_3 + 20\,\frac{n+2}{n^2}\,s_1^2 s_3 + 30\,\frac{n-1}{n^2}\,s_1 s_2^2 - \frac{60}{n^2}\,s_1^3 s_2 + \frac{24}{n^3}\,s_1^5\right\},$$

$$k_6 = \frac{n^2}{(n-1)\ldots(n-5)} \{(n+1)(n^2+15n-4)\,m_6 - 15(n-1)^2(n+4)\,m_2 m_4 - 10(n-1)(n^2-n+4)\,m_3^2 + 30n(n-1)(n-2)\,m_2^3\}$$

$$= \frac{n}{(n-1)\ldots(n-5)} \left\{(n+1)(n^2+15n-4)\,s_6 - 6\,\frac{n+1}{n}(n^2+15n-4)\,s_1 s_5 - 15\,\frac{(n-1)^2}{n}(n+4)\,s_2 s_4 - 10\,\frac{n-1}{n}(n^2-n+4)\,s_3^2 + 30\,\frac{n^2+9n+2}{n}\,s_1^2 s_4 + 120\,\frac{n^2-1}{n}\,s_1 s_2 s_3 + 30\,\frac{(n-1)(n-2)}{n}\,s_2^3 - 120\,\frac{n+3}{n}\,s_1^3 s_3 - 270\,\frac{n-1}{n}\,s_1^2 s_2^2 + \frac{360}{n}\,s_1^4 s_2 - \frac{120}{n^2}\,s_1^6\right\}.$$

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, ... 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

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

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

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

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

$$\kappa(r1^s) = \frac{\kappa_{r+s}}{n^s} \qquad \text{(III)}$$

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	−6	−6	−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
	$\frac{1}{n^2}\kappa_8$	$\frac{\kappa_6\kappa_2}{n(n-1)}$	$\frac{\kappa_5\kappa_3}{n(n-1)^2}$	$\frac{\kappa_4^2}{n(n-1)^2}$	$\frac{\kappa_4\kappa_2^2}{(n-1)^2(n-2)}$
$\kappa(42^2)$	1	20	$8(5n-7)$	$4(7n-10)$	$80(n-2)$
$\kappa(3^2 2)$	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	$32(n-2)$	$8(4n^2-9n+6)$	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)
$\frac{\kappa_3^2\kappa_2}{n(n-1)^3}$	$\frac{\kappa_2^4}{(n-1)^3}$		
$96(n-2)$	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)
$\frac{\kappa_4\kappa_3\kappa_2}{(n-1)^3(n-2)}$	$\frac{\kappa_3^3}{(n-1)^2(n-2)^2}$	$\frac{n\kappa_3\kappa_2^3}{(n-1)^2(n-2)^2}$	
$480(n-2)$	$120(n-2)^2$	—	(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)
$\frac{\kappa_4\kappa_3\kappa_2}{n(n-1)^3}$	$\frac{\kappa_3^3}{n(n-1)^3}$	$\frac{\kappa_3\kappa_2^3}{(n-1)^3}$	
$360(2n-3)$	$24(5n-12)$	480	(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	1500
	$\frac{\kappa_{10}}{n^2}$	$\frac{\kappa_8\kappa_2}{n(n-1)}$	$\frac{\kappa_7\kappa_3}{n(n-1)^2}$	$\frac{\kappa_6\kappa_4}{n(n-1)^2}$	$\frac{\kappa_5^2}{n(n-1)^2}$	$\frac{\kappa_6\kappa_2^2}{(n-1)^2(n-2)}$	$\frac{\kappa_5\kappa_3\kappa_2}{(n-1)^2(n-2)}$
$\kappa(62^2)$	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	$4(23n-37)$	$4(47n^2-120n+81)$	$12(9n^2-24n+17)$	$360(n-2)$	$288(5n-7)(n-2)$
$\kappa(3^22^2)$	1	37	$6(17n-27)$	$3(61n^2-166n+117)$	$2(59n^2-154n+113)$	$6(67n-131)$	$24(71n^2-246n+202)$
	$\frac{\kappa_{10}}{n^4}$	$\frac{\kappa_8\kappa_2}{n^3(n-1)}$	$\frac{\kappa_7\kappa_3}{n^3(n-1)^2}$	$\frac{\kappa_6\kappa_4}{n^3(n-1)^3}$	$\frac{\kappa_5^2}{n^3(n-1)^4}$	$\frac{\kappa_6\kappa_2^2}{n^2(n-1)^2}$	$\frac{\kappa_5\kappa_3\kappa_2}{n^2(n-1)^3}$
$\kappa(2^5)$	1	40	$80(n-2)$	$40(5n^2-12n+9)$	$16(n-2)(6n^2-12n+7)$	480	$1280(n-2)$

and the sixth moment of k_2. These are :—

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

$$+\frac{n}{(n-1)(n-2)}(450\kappa_8\kappa_2^2 + 3600\kappa_7\kappa_3\kappa_2 + 7200\kappa_6\kappa_4\kappa_2 + 6300\kappa_6\kappa_3^2 + 4500\kappa_5^2\kappa_2 + 21600\kappa_5\kappa_4\kappa_3 + 4950\kappa_4^3)$$

$$+\frac{n(n+1)}{(n-1)(n-2)(n-3)}(2400\kappa_6\kappa_2^3 + 21600\kappa_5\kappa_3\kappa_2^2 + 15300\kappa_4^2\kappa_2^2 + 54000\kappa_4\kappa_3^2\kappa_2 + 8100\kappa_3^4)$$

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

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

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

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

$\frac{n\kappa_4^2\kappa_2}{(n-1)(n-2)}$	$\frac{n\kappa_4\kappa_3^2}{(n-1)(n-2)}$	$\frac{n(n+1)\kappa_4\kappa_2^3}{(n-1)(n-2)(n-3)}$	$\frac{n(n+1)\kappa_3^2\kappa_2^2}{(n-1)(n-2)(n-3)}$	$\frac{n^2(n+5)\kappa_2^5}{(n-1)(n-2)(n-3)(n-4)}$	
—	—	—	—	—	(22)
420	630	—	—	—	(23)
720	1260	480	1080	—	(24)
850	1200	600	1800	120	(25)
$\frac{\kappa_4^2\kappa_2}{(n-1)^2(n-2)}$	$\frac{\kappa_4\kappa_3^2}{(n-1)^2(n-2)}$	$\frac{\kappa_4\kappa_2^3}{(n-1)^2(n-2)}$	$\frac{\kappa_3^2\kappa_2^2}{(n-1)^2(n-2)}$	$\frac{n\kappa_2^5}{(n-1)^2(n-2)}$	
$560(n-2)$	$840(n-2)$	—	—	—	(26)
$60(15n-31)$	$30(45n-103)$	$720n$	$1620n$	—	(27)
$144(7n-15)$	$72(21n-50)$	$\frac{96(10n^2-27n-1)}{n-3}$	$\frac{144(17n^2-53n-2)}{n-3}$	$\frac{192(n+1)}{n-3}$	(28)
$54(19n-48)$	$\frac{54(33n^2-148n+172)}{n-2}$	$\frac{72n(17n-40)}{n-2}$	$\frac{108n(27n-70)}{n-2}$	$\frac{216n}{n-2}$	(29)
$\frac{\kappa_4^2\kappa_2}{n(n-1)^3(n-2)}$	$\frac{\kappa_4\kappa_3^2}{n(n-1)^3(n-2)}$	$\frac{\kappa_4\kappa_2^3}{(n-1)^3(n-2)}$	$\frac{\kappa_3^2\kappa_2^2}{(n-1)^3(n-2)}$	$\frac{n\kappa_2^5}{(n-1)^3(n-2)}$	
$144(7n-10)(n-2)$	$24(49n-95)(n-2)$	$960(n-2)$	$2160(n-2)$	—	(30)
$36(29n^2-103n+93)$	$36(38n^2-155n+160)$	$72(14n-23)$	$144(19n-44)$	288	(31)
$\frac{\kappa_4^2\kappa_2}{n^2(n-1)^4}$	$\frac{\kappa_4\kappa_3^2}{n^2(n-1)^4}$	$\frac{\kappa_4\kappa_2^3}{n(n-1)^3}$	$\frac{\kappa_3^2\kappa_2^3}{n(n-1)^4}$	$\frac{\kappa_2^5}{(n-1)^4}$	
$320(4n^2-9n+6)$	$480(2n^2-7n+6)$	1920	$1920(n-2)$	384	(32)

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

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

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

$$+\frac{288(19n^3-98n^2+125n+2)}{(n-1)^2(n-2)^2(n-3)}\kappa_6\kappa_2^3+\frac{1728(24n^3-140n^2+200n+4)}{(n-1)^2(n-2)^2(n-3)}\kappa_5\kappa_3\kappa_2^2$$

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

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

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

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

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

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

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

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

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

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

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

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

$$+648n\,\frac{103n^2-510n+640}{(n-1)^3(n-2)^3}\kappa_3^2\kappa_2^3+648n^2\,\frac{5n-12}{(n-1)^3(n-2)^3}\kappa_2^6, \qquad (62)$$

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

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

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

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

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

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

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

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

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:

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

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 last term of (62) 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

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

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

$$\kappa_{r,s} = S\left\{\frac{(-)^{\rho-1}(\rho-1)\,!}{\pi_1!\,\pi_2!\ldots} \frac{r\,!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\ldots} \frac{s\,!}{(p_1'!)^{\pi_1}(p_2'!)^{\pi_2}\ldots} \mu_{p_1p_1'}^{\pi_1} \mu_{p_2p_2'}^{\pi_2} \ldots\right\}$$

and

$$\mu_{r,s} = S\left\{\frac{1}{\pi_1!\,\pi_2!\ldots} \frac{r\,!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\ldots} \frac{s\,!}{(p_1'!)^{\pi_1}(p_2'!)^{\pi_2}\ldots} \kappa_{p_1p_1'}^{\pi_1} \kappa_{p_2p_2'}^{\pi_2} \ldots\right\},$$

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

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

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

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

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

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

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

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