ON A PROPERTY OF THE SEMI-INVARIANTS OF THIELE

By

CECIL C. CRAIG
National Research Fellow

Given a general linear form

$$(1) \qquad a_1 x_1 + a_2 x_2 + \cdots\cdots + a_n x_n$$

of a set of statistical variables, $x_1, x_2, \cdots\cdots, x_n$,[1] it is well-known that in case the variables, $x_1, x_2, \cdots\cdots, x_n$ are independent, in the sense of the theory of probability, that the r'th semi-invariant of this form is simply

$$(2) \qquad a_1^r \lambda_r^{(1)} + a_2^r \lambda_r^{(2)} + \cdots\cdots + a_n^r \lambda_r^{(n)},^2$$

in which $\lambda_r^{(i)}$ is the r'th semi-invariant of x_i. This is perhaps the most important and useful property of semi-invariants.

Each semi-invariant is defined as a certain isobaric function of the moments of weight equal to the order of the semi-invariant. The question to which this note is devoted is whether among such isobaric functions, the property given above belongs uniquely to the semi-invariant. This problem is equivalent to another which

---

[1]There is no loss in generality in supposing the origin so chosen for each x_i that the constant in the form is zero.

[2]Thiele, T. N., Theory of Observations (C. & E. Layton, London, 1903) p. 39.

seems more difficult to state verbally. The r'th semi-invariant L_r of the form (1) is itself found in terms of the semi-invariants, $\lambda_{rst}\cdots$, of the n-way probability function $F(x_1, x_2, \cdots, x_n)$ by means of a symbolic multinomial expansion. Now in order that the above property may hold generally it is necessary and sufficient that the cross-semi-invariants of $F(x_1, x_2, \cdots x_n)$ should vanish if x_1, $x_2, \cdots, x_n$ are independent; that is, that each $\lambda_{rst}\cdots$ in which at least two of the quantities $r, s, t, \cdots$ are different from zero, should vanish identically. Now are semi-invariants the only such functions of moments, whose "cross" members behave in this way?

The semi-invariants L_r of the given linear form are defined by

$$e^{L_1 t + \frac{1}{2} L_2 t^2 + \frac{1}{3!} L_3 t^3 + \cdots}$$

$$(3) \qquad = \int_{-\infty, \cdots, -\infty}^{\infty, \cdots, \infty} dF(x_1, x_2, \cdots, x_n) e^{t(\sum_1^n a_i x_i)} \quad [1]$$

which is to be regarded as a formal identity in t. And the semi-invariants of $x_1, x_2, \cdots, x_n$ are given by

$$e^{(\sum_1^n \lambda_i t_i) + \frac{1}{2}(\sum_1^n \lambda_i t_i)^{(2)} + \frac{1}{3!}(\sum_1^n \lambda_i t_i)^{(3)} + \cdots}$$

$$(4) \qquad = \int_{-\infty, \cdots, -\infty}^{\infty, \cdots, \infty} dF(x_1, x_2, \cdots, x_n) e^{(\sum_1^n x_i t_i)}$$

$$= 1 + (\sum_1^n \nu_i t_i) + \frac{1}{2}(\sum_1^n \nu_i t_i)^{(2)} + \frac{1}{3!}(\sum_1^n \nu_i t_i)^{(3)} + \cdots$$

[1]We shall observe the distinction between probability functions and frequency functions suggested by H. Cramér in his important memoir: "On the Composition of Elementary Errors," Skandinavisk Aktuarietidskrift, 1928, p. 13. By a probability function we mean what has been called the cumulative frequency function and thus in the above we are using an n-way Stieltjes integral.

which is also a formal identity in $t_1, t_2, \cdots, t_n$.

The quantities $(\sum_1^n \lambda_i t_i)^{(r)}$ and $(\sum_1^n \nu_i t_i)^{(r)}$ refer to symbolic multinomial expansions, perhaps most easily explained by means of examples. Thus

$$(\sum_1^3 \lambda_i t_i)^{(2)} = \lambda_{200} t_1^2 + \lambda_{020} t_2^2 + \lambda_{002} t_3^2$$
$$+ 2\lambda_{110} t_1 t_2 + 2\lambda_{101} t_1 t_3 + 2\lambda_{011} t_2 t_3 ,$$

and

$$(\sum_1^2 \lambda_i t_i)^{(3)} = \lambda_{30} t_1^3 + \lambda_{03} t_2^3 + 3\lambda_{21} t_1^2 t_2 + 3\lambda_{12} t_1 t_2^2$$

in which $\lambda_{k00\cdots\cdots0} = \lambda_k^{(1)}, \lambda_{0k0\cdots\cdots0} = \lambda_k^{(2)}, \cdots\cdots,$ in our first used notation, and $\lambda_{110\cdots0}, \lambda_{210\cdots\cdots0}$; etc. are cross-semi-invariants of x_1 and x_2.

Then by inspection of (3) and (4) it is evident that

$$L_k = (\sum_1^n a_i \lambda_i)^{(k)}, \quad k = 1, 2, 3, \cdots\cdots \tag{5}$$

In case the variables $x_1, x_2, \cdots\cdots, x_n$ are all independent of each other $F(x_1, x_2, \cdots\cdots x_n)$ splits up into the product $F_1(x_1)\ F_2(x_2) \cdots F_n(x_n)$ of the probability functions of the separate variables, L_k becomes equal to the expression (2), and all the cross-semi-invariants in the expansion of the right member of (4) become identically zero. That the vanishing of these cross-semi-invariants is not only a sufficient but is also a necessary condition that L_k assume the value (2) is evident from the absence of any restrictions on $F(x_1, x_2, \cdots, x_n)$ (except that it be an n-way probability function) or on the set $a_1, a_2, \cdots\cdots, a_n$.

Now each cross-semi-invariant is expressed as a certain isobaric function of moments, some of them cross-moments. But in the case of independent variables,

$$\nu_{rst\ldots\ldots} = \nu_r \, \nu_s \, \nu_t \cdots\cdots,$$

and when this is true, the value of each cross-semi-invariant becomes identically zero. To illustrate this and for use in the demonstration that the semi-invariants are the only such functions, let us write out the fourth order semi-invariants of $F(x_1, x_2, x_3, \cdots, x_n)$ in terms of moments. These are obtained by equating coefficients of like terms in[1]

$$(6)\qquad \begin{aligned}(\sum_1^n \lambda_i t_i)^{(4)} = (\sum_1^n \nu_i t_i)^{(4)} - 4(\sum_1^n \nu_i t_i)^{(3)}(\sum_1^n \nu_i t_i)\\ -3[(\sum_1^n \nu_i t_i)^{(2)}]^2 + 12(\sum_1^n \nu_i t_i)^{(2)}(\sum_1^n \nu_i t_i)^2\\ -6(\sum_1^n \nu_i t_i)^4.\end{aligned}$$

Leaving off superfluous zeros in the subscripts, this gives for example

$$\lambda_{40} = \nu_{40} - 4\nu_{30}\nu_{10} - 3\nu_{20}^2 + 12\nu_{20}\nu_{10}^2 - 6\nu_{10}^4$$

$$\begin{aligned}\lambda_{22} = \nu_{22} - (2\nu_{21}\nu_{01} + 2\nu_{12}\nu_{10}) - (\nu_{20}\nu_{02} + 2\nu_{11}^2)\\ + (2\nu_{20}\nu_{01}^2 + 2\nu_{02}\nu_{10}^2 + 8\nu_{11}\nu_{10}\nu_{01}) - 6\nu_{10}^2\nu_{01}^2.\end{aligned}$$

If in the value of λ_{22} we set $\nu_{22} = \nu_{20}\nu_{02}$, $\nu_{21} = \nu_{20}\nu_{01}$, etc., then $\lambda_{22} \equiv 0$ as it was already known must happen.

For the sake of simplicity let us suppose, at first, that the component variables in (1) are all "equal," that is, that $F(x_1, x_2, \cdots\cdots, x_n) \equiv F(x, x, \cdots\cdots, x)$. In the case of

[1]The general formula giving semi-invariants in terms of moments is to be found in several places. See e. g., C. Jordan, Statistique Mathématique (Gauthier-Villars, Paris, 1927), p. 41. For an elementary derivation and also for an extended example of the use of semi-invariants of a correlation function of several variables see the author's "An Application of Thiele's Semi-invariants to the Sampling Problem," Metron, Vol. VII, No. 4 (1928), pp. 3-74.

independence among $x_1, x_2, \cdots, x_n$ we can write also $F_1(x_1) = F_2(x_2) = \cdots\cdots = F_n(x_n) = F(x)$. An equivalent assumption is that all moments and hence all semi-invariants of the same *type* of $F(x_1, x_2, \cdots, x_n)$ are equal. (Moments of the same type are all those with the same combination of digits in their subscripts.) Then the expressions for all the semi-invariants of the fourth order of $F(x_1, x_2, \cdots \cdots, x_n)$ are equivalent to the following:

$$
(7)\quad
\begin{aligned}
\lambda_{40} &= \nu_{40} - 4\nu_{30}\nu_{10} - 3\nu_{20}^2 + 12\nu_{20}\nu_{10}^2 - 6\nu_{10}^4 \\
\lambda_{31} &= \nu_{31} - (\nu_{30}\nu_{10} + 3\nu_{21}\nu_{10}) - 3\nu_{20}\nu_{11} + (6\nu_{20}\nu_{10}^2 + 6\nu_{11}\nu_{10}^2) - 6\nu_{10}^4 \\
\lambda_{22} &= \nu_{22} - 4\nu_{21}\nu_{10} - (\nu_{20}^2 + 2\nu_{11}^2) + (4\nu_{20}\nu_{10}^2 + 8\nu_{11}\nu_{10}^2) - 6\nu_{10}^4 \\
\lambda_{211} &= \nu_{211} - (2\nu_{21}\nu_{10} + 2\nu_{111}\nu_{10}) - (\nu_{20}\nu_{11} + 2\nu_{11}^2) + (2\nu_{20}\nu_{10}^2 + 10\nu_{11}\nu_{10}^2) - 6\nu_{10}^4 \\
\lambda_{1111} &= \nu_{1111} - 4\nu_{111}\nu_{10} - 3\nu_{11}^2 + 12\nu_{11}\nu_{10}^2 - 6\nu_{10}^4
\end{aligned}
$$

Now, our general isobaric function of the moments of weight four can be written

$$
(8)\quad
\begin{aligned}
f(z_1, z_2, \cdots, z_5) &= z_1\Big(\sum_1^n \nu_i t_i\Big)^{(4)} - 4z_2\Big(\sum_1^n \nu_i t_i\Big)^{(3)}\Big(\sum_1^n \nu_i t_i\Big) \\
&\quad - 3z_3\Big[\Big(\sum_1^n \nu_i t_i\Big)^{(2)}\Big]^2 + 12z_4\Big(\sum_1^n \nu_i t_i\Big)^{(2)}\Big(\sum_1^n \nu_i t_i\Big)^2 - 6z_5\Big(\sum_1^n \nu_i t_i\Big)^4
\end{aligned}
$$

And in our special case of equal component variables $x_1, x_2, \cdots, x_n$ our problem is to determine for what sets of values of $z_1, z_2, \cdots, z_5$ the coefficients of $t_1^3 t_2$, $t_1^2 t_2^2$, $t_1^2 t_2 t_3$ and $t_1 t_2 t_3 t_4$ in the right member of (8) vanish identically if $x_1, x_2, \cdots, x_n$ are independent.

By comparison with (7) it is seen that this gives four linear equations with which to determine the five unknowns. But we can add a fifth equation by stating that the coefficient of t_1^4 is in general a parameter which in the case of independence is a function of $F(x)$ and $z_1, z_2, \cdots, z_5$, which we shall designate by ξ_4. Then we have for the determination of z_1 :

$$(9)\; z_1 = \frac{\begin{vmatrix} \xi_4 & -4v_3v_1 & -3v_2^2 & 12v_2v_1^2 & -6v_1^4 \\ 0 & -(v_3v_1+3v_2v_1^2) & -3v_2v_1^2 & 6v_2v_1^2+6v_1^4 & -6v_1^4 \\ 0 & -4v_2v_1^2 & -(v_2^2+2v_1^4) & 4v_2v_1^2+8v_1^4 & -6v_1^4 \\ 0 & -(2v_2v_1^2+2v_1^4) & -(v_2v_1^2+2v_1^4) & 2v_2v_1^2+10v_1^4 & -6v_1^4 \\ 0 & -4v_1^4 & -3v_1^4 & 12v_1^4 & -6v_1^4 \end{vmatrix}}{\begin{vmatrix} v_4 & -v_3v_1 & -3v_2^2 & 12v_2v_1^2 & -6v_1^4 \\ v_3v_1 & -(v_3v_1+3v_2v_1^2) & -3v_2v_1^2 & 6v_2v_1^2+6v_1^4 & -6v_1^4 \\ v_2^2 & -4v_2v_1^2 & -(v_2^2+2v_1^4) & 4v_2v_1^2+8v_1^4 & -6v_1^4 \\ v_2v_1^2 & -(2v_2v_1^2+2v_1^4) & -(v_2v_1^2+2v_1^4) & 2v_2v_1^2+10v_1^4 & -6v_1^4 \\ v_1^4 & -4v_1^4 & -3v_1^4 & 12v_1^4 & -6v_1^4 \end{vmatrix}}$$

By adding each of the four other columns to the first column in the denominator, we have at once in view of (7),

$$z_1 = \xi_4 / \lambda_4$$

unless the identical first minor of numerator and denominator vanishes. But this can happen only if there is linear dependence between the corresponding elements in the four rows of this minor which in turn can happen only if there is a linear relation between the quantities v_3v_1, v_2^2, $v_2v_1^2$, and v_1^4. (Such a linear dependence would exist if the second or third semi-invariant of $F(x)$ is zero.)

Moreover, it is readily seen that we get $z_1 = z_2 = \cdots = z_5 = \dfrac{\xi_4}{\lambda_4}$ (Of course we suppose $\lambda_4 \neq 0$ and moreover $\xi_4 = 0$ could hold only for some $F(x)$'s)

If we no longer suppose the components $x_1, x_2, \cdots, x_n$ "equal" in the sense defined above, the quantities in (7) may be replaced by summations of all terms of the same type or summations of all products of terms which are coefficients of similar terms in t_i 's. Thus in place of $\lambda_{40}, \nu_{40}, \nu_{30}, \nu_{10}$ in the first equation, and λ_{31} and $\nu_{30}\ \nu_{01}$ in the second we now write,

$$\Sigma\, \lambda_{40} = \lambda_{40} + \lambda_{04} + \lambda_{004} + \cdots$$

$$\Sigma\, \nu_{40} = \nu_{40} + \nu_{04} + \nu_{004} + \cdots$$

$$\Sigma\, \nu_{30}\nu_{10} = \nu_{30}\nu_{10} + \nu_{03}\nu_{01} + \nu_{003}\nu_{001} + \cdots$$

$$\Sigma\, \lambda_{31} = \lambda_{31} + \lambda_{13} + \lambda_{031} + \lambda_{013} + \cdots$$

$$\Sigma\, \nu_{30}\nu_{01} = \nu_{30}\nu_{01} + \nu_{03}\nu_{10} + \nu_{030}\nu_{001} + \nu_{003}\nu_{010} + \cdots ,$$

respectively. But otherwise our argument will be the same and lead to the same conclusion.

It is obvious that the argument for weight four is perfectly general and thus that the same kind of conclusions hold for any weight. We conclude that the semi-invariants are the only isobaric functions of the moments of a set of n variables which have the properties described in the first two paragraphs independent of the probability or frequency functions of those variables.

But if when the variables are independent the probability function of each one is such that there is an isobaric relation among the moments of order lower than k, the same for each variable, then there are other isobaric functions of order k and higher which enjoy the property of semi-invariants in question. And it will be shown that the only isobaric relations among the moments of order $< k$, mentioned above, which lead to the new isobaric functions of this type of order $\geq k$, are obtained by setting semi-invariants of order $< k$, equal to zero.

Let us return to the case in which the weight is four. Then if $\lambda_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^{\,3} = 0$, the minor D_{11} of our denominator D vanishes, and so, of course, does the corresponding minor in the numerator. Then as a matter of fact there is a double infinity of the sought isobaric functions of weight four.

Some of them are given by the following sets of values of the z's.

z_1	z_2	z_3	z_4	z_5
5	2	5	2	1
6	3	6	3	2
9	3	9	3	1

as may be verified by actual computation.

Now we also have[1]

$$\lambda_4 = \nu_4 - \lambda_1 \nu_3 - 3\lambda_2 \nu_2 - 3\lambda_3 \nu_1$$

from which we can write in place of (8)

$$(10)\quad \begin{aligned} f(y_1, \cdots, y_4) &= y_1 (\sum_1^n \nu_i t_i)^{(4)} - y_2 (\sum_1^n \lambda_i t_i)(\sum_1^n \nu_i t_i)^{(3)} \\ &- 3y_3 (\sum_1^n \lambda_i t_i)^{(2)} (\sum_1^n \nu_i t_i)^{(2)} - 3y_4 (\sum_1^n \lambda_i t_i)^{(3)} (\sum_1^n \nu_i t_i) \end{aligned}$$

in which we can seek to find sets of values of $y_1, \cdots\cdots, y_4$ so that the coefficients of $t_1^3 t_2$, $t_1^2 t_2^2$, $t_2^2 t_3 t_4$ and $t_1 t_2 t_3 t_4$ will vanish when the x's are independent. This will give us four homogeneous linear equations in which the determinant of the coefficients vanishes identically since $y_1 = y_2 = y_3 = y_4 = 1$ is a solution. Addition of the second, third and fourth columns to the first gives a new first column of zeros. But if, say, $\lambda_3 = 0$, in addition to λ_{21}, and λ_{111} which already vanish if the x's are independent, then the elements of the fourth column are all zeros also, and our determinant is of rank not greater than two. But since the solution of the set of equations arising from (10) is equivalent to that arising from (8), the minor D_{11}, of D in (9)

[1]Thiele, T. N., loc. cit., p. 25.

must vanish in case $\lambda_3 = 0$.

But since $z_1 = z_2 = \cdots = z_s = 1$ is a solution of the equations (8), it is easy to see that if in D_{11}, the sum of the last three columns be added to the first column, the resulting first column will be identical, though opposite in sign with the last four elements of the first coulmn of D. Let us indicate the new D_{11} by D_{11}'.

Now there is a linear dependence between the elements of the rows of D_{11}. In fact the elements of the first row minus three times the corresponding elements of the third plus twice the corresponding elements of the fourth ($\lambda_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$) must give zero for each element. For suppose there exists another such linear relationship between rows. This linear relationship must hold between the corresponding elements of the first column of D_{11}', and we have a new isobaric relation between the moments of x. But a probability function $F(x)$can always be found in which

$$(11) \qquad \nu_3\nu_1 - 3\nu_2\nu_1^2 + 2\nu_1^4 = \lambda_3\nu_1 = 0$$

holds and the other relation does not. But for the $F(x)$'s in which (11) holds D_{11}' must vanish, and thus the relation between columns must be that given by (11).

Thus D_{11} contains as factors λ_3, λ_2 and λ_1. That it contains no others can easily be verified directly.

The cases of weights two, three, and four are easily handled directly throughout. If the weight is now k greater than four, our argument readily generalizes. The equations now arising from the relation corresponding to (10) are now greater in number than the unknowns $y_1, y_2, \cdots\cdots, y_k$, but it is obvious that the matrix of the coefficients is of rank not greater than $k-2$. And it follows just as before that $\lambda_{k-1}, \lambda_{k-2}, \cdots, \lambda_1$, are all factors of the new D_{11}.

The argument above which shows for the weight four, that

λ_3 is a factor of $D_{\prime\prime}$ does not show that there cannot be other linear relations between the elements of the first column which are also factors of $D_{\prime\prime}$. It only shows that if there is such a factor, the corresponding linear dependence holds for certain rows of $D_{\prime\prime}$

Let us consider the case of weight five. The elements of the first column of D are now $\nu_5, \nu_4\nu_1, \nu_3\nu_2, \nu_3\nu_1^2, \nu_2^2\nu_1, \nu_2\nu_1^3$ and ν_1^5 and the elements of the first column of $D'_{\prime\prime}$ are the last six of these with opposite sign, and they thus correspond to the partitions of 5. We know that one of the two sets of three rows of $D_{\prime\prime}$, the second, fourth, and fifth or the third, fifth, and sixth, are connected by the linear relation corresponding to $\nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 = \lambda_3 = 0$ so that λ_3 is at least once a factor of $D_{\prime\prime}$. If we suppose that the first set of three rows are so related, does it follow that this same relation holds for the second set? Now it is easy to see that if in the second row ν_1^2 be everywhere substituted for ν_2 the resulting row will be identical with the third and that the same is true of the fourth and fifth rows and of the fifth and sixth. Then if a certain linear relation holds for the first set of three rows, by the substitution of ν_1^2 for ν_2 everywhere in it, it follows that the same relation holds for the second set of three rows also. Thus λ_3 is twice a factor in $D_{\prime\prime}$ for weight five. We note also that the partitions of 3 (counting 3 as a partition of 3) are twice found with common factors among the partitions of 5, that is, 32, 221, 2111; and 311, 2111, 11111.

The argument is readily generalized[1] and in case of $D_{\prime\prime}$ of weight k, each semi-invariant of weight $r < k$ is a factor of $D_{\prime\prime}$

---

[1]The general argument is based on the principle that the second row of D is obtained from the process which gives the first by replacing one factor t_1 by t_2, the third from the first by replacing t_1^2 by t_2^2, the fourth from the first by replacing t_1^2 by $t_2 t_3$, and so on (see (6) and (7)). Thus in the case of weight six, to compare the three rows beginning with $\nu_3^2, \nu_3\nu_2\nu_1, \nu_3\nu_1^3$ with the three beginning with $\nu_3\nu_1^3, \nu_2\nu_1^4, \nu_1^6$ we replace the ν_3 in the first set which arises as a coefficient of t_1^3 by ν_1^3 and the two sets of rows become identical.

as often as the partitions of r are found with common factors among the partitions of k. (We count r as a partition of r.) Thus for weight four, $D_{11} = \lambda_3 \lambda_2^3 \lambda_1^7$ which gives D_{11} the correct weight sixteen. In case of weight five, $D_{11} = \lambda_4 \lambda_3^2 \lambda_2^4 \lambda_1^{12}$ which again gives D_{11} the correct weight thirty. And it is easy to show by induction that in case of weight k this method gives D_{11} its proper weight. Among the partitions of k are found all the partitions of $k-1$ with a part 1 added to each. Thus each of these adds k to the total weight. For the partition $k-2, 2$, it is seen that the remaining partitions of $k-2$ with the common additional part 2 will be found among the remaining partitions of k and that the remaining partitions of 2 with the common additional part $k-2$ will also be found. Thus this partition contributes the weight k to the total. And similarly it can be seen that every partition of k contributes k to the total weight of D_{11}, which was to be proved.

Finally, then, we have the additional result that the necessary and sufficient condition that more than one isobaric function of weight k of the moments of the probability variables $x_1, x_2, \cdots, x_n$ exists which has the semi-invariant properties in question, is that the probability functions of $x_1, x_2, \cdots, x_n$ in case of independence are such that for some $r < k$, λ_r vanishes for each of them.

Stanford University.