1) the formula (1) or

$$y = e^{\alpha+\beta x+\gamma x^2+\cdots-\kappa x^{2r}},$$

2) the products of integral algebraic functions by a typical function or (6)

$$y = k_0\varphi - \frac{k_1}{|\underline{1}}D\varphi + \frac{k_2}{|\underline{2}}D^2\varphi - \frac{k_3}{|\underline{3}}D^3\varphi + \cdots, \qquad \varphi = e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2},$$

3) a sum of several typical functions

$$y = \Sigma_1^r k_i e^{-\frac{1}{2}\left(\frac{x-m_i}{n_i}\right)^2}. \tag{14}$$

This account of the more prominent among the functional forms, which we have at our disposal for the representation of laws of errors, may prove that we certainly possess good instruments, by means of which we can even in more than one form find general series adapted for the representation of laws of errors. We do not want forms for the series, required in theoretical speculations upon laws of errors; nor is the exact representation of the actual frequencies more than reasonably difficult. If anything, we have too many forms and too few means of estimating their value correctly.

As to the important transition from laws of actual errors to those of presumptive errors, the functional form of the law leaves us quite uncertain. The convergency of the series is too irregular, and cannot in the least be foreseen.

We ask in vain for a fixed rule, by which we can select the most important and trustworthy forms with limited numbers of constants, to be used in predictions. And even if we should have decided to use only the typical form by the laws of presumptive errors, we still lack a method by which we can compute its constants. The answer, that the "adjustment" of the law of errors must be made by the "method of least squares", may not be given till we have attained a satisfactory proof of that method; and the attempts that have been made to deduce it by speculations on the functional laws of errors must, I think, all be regarded as failures.

## VI. LAWS OF ERRORS

### EXPRESSED BY SYMMETRICAL FUNCTIONS.

§ 21. All constants in a functional law of errors, every general property of a curve of errors or, generally, of a law of numerical errors, must be symmetrical functions of the several results of the repetitions, i. e. functions which are not altered by interchanging two or more of the results. For, as all the values found by the repetitions correspond to the same essential circumstances, no interchanging whatever can have any influence on the law of errors. Conversely, any symmetrical function of the values of the observations will represent some property or other of the law of errors. And we must be able to express the whole law of errors itself by every such collection of symmetrical functions, by which every property of the law of errors can be expressed as unambiguously as by the very values found by the repetitions.

We have such a collection in the coefficients of that equation of the $n^{th}$ degree, whose roots are the $n$ observed values. For if we know these coefficients, and solve the equation, we get an unambiguous determination of all the values resulting from the repetitions, i. e. the law of errors. But other collections also fulfil the same requirements; the essential thing is that the $n$ symmetrical functions are rational and integral, and that one of them has each of the degrees $1, 2 \dots n$, and that none of them can be deduced from the others.

The collection of this sort that is easiest to compute, is *the sums of the powers.* With the observed values

$$o_1, o_2, o_3, \dots o_n$$

we have

$$\left.\begin{array}{l} s_0 = o_1^0 + o_2^0 + \dots + o_n^0 = n \\ s_1 = o_1 + o_2 + \dots + o_n \\ s_2 = o_1^2 + o_2^2 + \dots + o_n^2 \\ \dots\dots\dots\dots\dots\dots \\ s_r = o_1^r + o_2^r + \dots + o_n^r \end{array}\right\} \quad (15)$$

and the fractions $\frac{s_r}{n}$ may also be employed as an expression for the law of errors; it is only important to reduce the observations to a suitable zero which must be an average value of $o_1 \dots o_n$; for if the differences between the observations are small, as compared with their differences from the average, then

$$\frac{s_1}{s_0}, \sqrt{\frac{s_2}{s_0}}, \dots \sqrt[r]{\frac{s_r}{s_0}}$$

may become practically identical, and therefore unable to express more than one property of the law of errors.

From a well known theorem of the theory of symmetrical functions, the equations

$$\begin{aligned} 1 + a_1\omega + a_2\omega^2 + \dots &= (1 - o_1\omega)(1 - o_2\omega)\dots(1 - o_n\omega) \\ &= e^{\Sigma \log(1 - o_r\omega)} \\ &= e^{-(s_1\omega + \frac{1}{2}s_2\omega^2 + \frac{1}{3}s_3\omega^3 + \dots)}, \end{aligned}$$

which are identical with regard to every value of $\omega$, we learn that the sum of the powers $s_r$ can be computed without ambiguity, if we know the coefficients $a_r$ of the equation, whose roots are the $n$ observations; and vice versâ, by differentiating the last equation with regard to $\omega$, and equating the coefficients we get

$$\left.\begin{array}{l} 0 = a_1 + s_1 \\ 0 = 2a_2 + a_1 s_1 + s_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ 0 = na_n + a_{n-1} s_1 + \dots + a_1 s_{n-1} + s_n \end{array}\right\} \qquad (16)$$

from which the coefficients $a_n$ are unambiguously and very easily computed, when the $s_n$ are directly calculated.

§ 22. But from the sums of powers we can easily compute also another serviceable collection of symmetrical functions, which for brevity we shall call *the half-invariants*.

Starting from the sums of powers $s_r$, these can be defined as $\mu_1, \mu_2, \mu_3$, by the equation

$$s_0 e^{\frac{\mu_1}{\underline{|1}}\tau + \frac{\mu_2}{\underline{|2}}\tau^2 + \frac{\mu_3}{\underline{|3}}\tau^3 \dots} = s_0 + \frac{s_1}{\underline{|1}}\tau + \frac{s_2}{\underline{|2}}\tau^2 + \frac{s_3}{\underline{|3}}\tau^3 + \dots, \qquad (17)$$

which we suppose identical with regard to $\tau$.

As $s_r = \Sigma\rho^r$, this can be written

$$s_0 e^{\frac{\mu_1}{\underline{|1}}\tau + \frac{\mu_2}{\underline{|2}}\tau^2 + \frac{\mu_3}{\underline{|3}}\tau^3 + \dots} = e^{\rho_1\tau} + e^{\rho_2\tau} + \dots e^{\rho_n\tau}. \qquad (18)$$

By developing the first term of (17) as $\Sigma k_r\tau_r$ and equating the coefficients of each power of $\tau$, we get each $\frac{s_r}{s_0}$ expressed as a function of $\mu_1 \dots \mu_r$:

$$\left.\begin{array}{l} s_1 = s_0\mu_1 \\ s_2 = s_0(\mu_2 + \mu_1^2) \\ s_3 = s_0(\mu_3 + 3\mu_2\mu_1 + \mu_1^3) \\ s_4 = s_0(\mu_4 + 4\mu_3\mu_1 + 3\mu_2^2 + 6\mu_2\mu_1^2 + \mu_1^4) \\ \dots\dots\dots \end{array}\right\} \qquad (19)$$

Taking the logarithms of (17) we get

$$\frac{\mu_1}{\underline{|1}}\tau + \frac{\mu_2}{\underline{|2}}\tau^2 + \frac{\mu_3}{\underline{|3}}\tau^3 + \dots = \log\left(1 + \frac{s_1}{s_0}\frac{\tau}{\underline{|1}} + \frac{s_2}{s_0}\frac{\tau^2}{\underline{|2}} + \frac{s_3}{s_0}\frac{\tau^3}{\underline{|3}} + \dots\right) \qquad (20)$$

and hence

$$\left.\begin{array}{l} \mu_1 = s_1 : s_0 \\ \mu_2 = (s_2 s_0 - s_1^2) : s_0^2 \\ \mu_3 = (s_3 s_0^2 - 3s_2 s_1 s_0 + 2s_1^3) : s_0^3 \\ \mu_4 = (s_4 s_0^3 - 4s_3 s_1 s_0^2 - 3s_2^2 s_0^2 + 12s_2 s_1^2 s_0 - 6s_1^4) : s_0^4 \\ \dots\dots\dots \end{array}\right\} \qquad (21)$$

The general law of the relation between the $\mu$ and $s$ is more easily understood through the equations

$$\left.\begin{aligned}
s_1 &= \mu_1 s_0 \\
s_2 &= \mu_1 s_1 + \mu_2 s_0 \\
s_3 &= \mu_1 s_2 + 2\mu_2 s_1 + \mu_3 s_0 \\
s_4 &= \mu_1 s_3 + 3\mu_2 s_2 + 3\mu_3 s_1 + \mu_4 s_0 \\
&\cdots\cdots\cdots
\end{aligned}\right\} \quad (22)$$

where the numerical coefficients are those of the binomial theorem. These equations can be demonstrated by differentiation of (17) with regard to $\tau$, the resulting equation

$$s_1 + \frac{s_2}{1!}\tau + \frac{s_3}{2!}\tau^2 + \frac{s_4}{3!}\tau^3 + \ldots = \left(\mu_1 + \frac{\mu_2}{1!}\tau + \frac{\mu_3}{2!}\tau^2 + \ldots\right)\left(s_0 + \frac{s_1}{1!}\tau + \frac{s_2}{2!}\tau^2 + \ldots\right) \quad (23)$$

being satisfied fòr all values of $\tau$ by (22).

These half-invariants possess several remarkable properties. From (18) we get

$$s_0 e^{\frac{\mu_2}{2!}\tau^2 + \frac{\mu_3}{3!}\tau^3 + \ldots} = e^{(o_1 - \mu_1)\tau} + \ldots + e^{(o_{s_0} - \mu_1)\tau} \quad (24)$$

consequently any transformation $o' = o + c$, any change of the zero of all observations $o_1 \ldots o_n$, affects only $\mu_1$ in the same manner, but leaves $\mu_2, \mu_3, \mu_4, \ldots$ unaltered; any change of the unit of all observations can be compensated by the reciprocal change of the unit of $\tau$, and becomes therefore indifferent to $\mu_2\tau^2, \mu_3\tau^3, \ldots$

Not only the ratios

$$\frac{s_1}{s_0}, \frac{s_2}{s_0}, \ldots \frac{s_n}{s_0}$$

but also the half-invariants have the property which is so important in a law of errors, of remaining unchanged when the whole series of repetitions is repeated unchanged.

We have seen that the typical character of a law of errors reveals itself in the elegant functional form

$$\varphi(x) = e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2}.$$

Now we shall see that it is fully as easy to recognize the typical laws of errors by means of their half-invariants. Here the criterion is that $\mu_r = 0$ if $r \geq 3$, while $\mu_1 = m$ and $\mu_2 = n^2$. This remarkable proposition has originally led me to prefer the half-invariants to every other system of symmetrical functions; it is easily demonstrated by means of (5), if we take $m$ for the zero of the observations.

We begin by forming the sums of powers $s_r$ of that law of errors where the frequency of an observed $x$ is proportional to $\varphi(x) = e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2}$; as this law is continuous we get

$$s_r = \int_{-\infty}^{+\infty} x^r \varphi(x)\, dx.$$

For every differential coefficient $D^m\varphi(x)$ we have

$$\int_{-\infty}^{+\infty} D^m\varphi(x)\cdot dx = D^{m-1}\varphi(\infty) - D^{m-1}\varphi(-\infty) = 0,$$

consequently we learn from (5) that $s_{2r+1} = 0$, but

$$\begin{aligned} s_2 &= 1\cdot n^2 s_0 \\ s_4 &= 1\cdot 3\cdot n^4 s_0 \\ s_6 &= 1\cdot 3\cdot 5\cdot n^6 s_0 \\ &\cdots\cdots\cdots \end{aligned}$$

(compare problem 3, § 18). Now the half-invariants can be found by (22) or by (17). If we use (22) we remark that $s_{2r} = n^2(2r-1)s_{2r-2}$; then writing for (22)

$$\begin{array}{llr} s_1 & = \mu_1 s_0 & = 0 \\ s_2 - \mu_2 s_0 & = \mu_1 s_1 & = 0 \\ s_3 - 2\mu_2 s_1 & = \mu_1 s_2 + \mu_3 s_0 & = 0 \\ s_4 - 3\mu_2 s_2 & = \mu_1 s_3 + 3\mu_3 s_1 + \mu_4 s_0 & = 0 \\ s_5 - 4\mu_2 s_3 & = \mu_1 s_4 + 6\mu_3 s_2 + 4\mu_4 s_1 + \mu_5 s_0 & = 0 \\ s_6 - 5\mu_2 s_4 & = \mu_1 s_5 + 10\mu_3 s_3 + 10\mu_4 s_2 + 5\mu_5 s_1 + \mu_6 s_0 & = 0 \end{array}$$

we see that the solution is $\mu_2 = n^2$ and $\mu_1 = \mu_3 = \mu_4 = \ldots = 0$.

By (17) we get

$$e^{\frac{\mu_1}{|1}\tau + \frac{\mu_2}{|2}\tau^2 + \frac{\mu_3}{|3}\tau^3 + \cdots} = 1 + \frac{(n\tau)^2}{2|1} + \frac{(n\tau)^4}{4|2} + \cdots$$

$$= e^{\frac{n^2\tau^2}{2}}.$$

Equating the coefficients of $\tau^r$ we get here also $\mu_1 = 0 = m$, $\mu_2 = n^2$, $\mu_r = 0$ if $r \overline{>} 3$.

If we wish to demonstrate this important proposition without change of the zero, and without the use of the equations (3) whose general demonstration is somewhat difficult, we can commence by the lemma that, for each integral and positive value of $r$, and also for $r = 0$, we have for the typical law of errors

$$s_{r+1} = m s_r + r n^2 s_{r-1}.$$

The function $\psi(x) = n^2 x^r e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2}$ is equal to zero both for $x = \infty$ and for $x = -\infty$; if we now between these limits integrate its differential equation

$$\frac{d\psi(x)}{dx} = \left(r n^2 x^{r-1} - (x-m)x^r\right)e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2},$$

we get

$$0 = -s_{r+1} + m s_r + r n^2 s_{r-1},$$

where

$$s_r = \int_{-\infty}^{+\infty} x^r e^{-\frac{1}{2}\left(\frac{x-m}{n}\right)^2} dx.$$

If we now from (22) subtract, term by term, the equations

$$s_1 = ms_0$$
$$s_2 = ms_1 + n^2 s_0$$
$$s_3 = ms_2 + 2n^2 s_1$$
$$s_4 = ms_3 + 3n^2 s_2$$
$$\cdots\cdots\cdots$$

it is obvious that $\mu_1 - m = 0$, $\mu_2 = n^2$, $\mu_3 = \mu_4 = \ldots = 0$.

By computation of $\mu_1$ and $\mu_2$ we find consequently, in the simplest way, the constants of a typical law of errors.

If the law of errors deviates only a little from the typical form, $\mu_3$, $\mu_4$, etc., will also, all of them, be relatively small numbers; and each of them may be either positive or negative.

On the whole, a law of errors can be determined without ambiguity by the values $\mu_1, \mu_2, \ldots \mu_r$, $r$ being the number of repetitions. From any such $\mu$'s we can compute the sums of the powers $s$ unambiguously, and from these again the coefficients of the equation whose roots are the observed values.

But for real laws of errors it is a necessary condition that no imaginary root can be admitted. If an infinite number of repetitions is considered, the equation ceases to be algebraic, and then the convergency of the series necessary for its solution is a further condition.

§ 23. *The mean value* $\mu_1 = \frac{s_1}{s_0} = \frac{o_1 + o_2 + \ldots + o_n}{n}$ is always greater than the least, less than the greatest of the observed values $o_1, o_2, \ldots o_n$; under typical circumstances we shall find almost the same number of greater and less values of the observations. The majority of them lie rather near to $\mu_1$; only few very distant from it. The mean value is the *simplest* representative of what is common in a series of values found by repetition; its application as such is most likely exceedingly old, and marks in the history of science the first trace of a theory of observations.

*The mean deviation*, whose square is $= \mu_2$, measures the magnitude of the deviations, the uncertainty of the repeated actual observations. The square of the mean deviation is the mean of the squares of the deviations of the several observations from their mean value. By addition of

$$(o_1 - \mu_1)^2 = o_1^2 - 2o_1\mu_1 + \mu_1^2$$
$$(o_2 - \mu_1)^2 = o_2^2 - 2o_2\mu_1 + \mu_1^2$$
$$\cdots\cdots\cdots\cdots$$
$$(o_n - \mu_1)^2 = o_n^2 - 2o_n\mu_1 + \mu_1^2$$

we get

$$\Sigma(o - \mu_1)^2 = s_2 - 2s_1\mu_1 + s_0\mu_1^2,$$

and as $\mu_1 = \frac{s_1}{s_0}$

$$\frac{\Sigma(o-\mu_1)^2}{s_0} = \frac{s_0 s_2 - s_1^2}{s_0^2} = \mu_2. \tag{25}$$

The computation of $\mu_2$ by this formula will often be easier than by the equation (21), because $s_2$ in the latter must frequently be computed with more figures. There is however a middle course, which is often to be preferred to either of these methods of computation. As a *change in the zero of the observations* involves the same increase of every $o$ and of $\mu_1$, it will, according to (24), have no influence at all on $\mu_2$. We select therefore as zero a convenient, round number, $c$, very near $\mu_1$, and by reference to this zero the observed values are transformed to

$$o_1' = o_1 - c,\ o_2' = o_2 - c,\ \ldots\ o_n' = o_n - c.$$

When $s_1'$ and $s_2'$ indicate the sums of the transformed observations, and $\mu_1' = \mu_1 - c$, then we have $\mu_1 = c + \frac{\Sigma(o-c)}{n}$ and

$$\left.\begin{aligned} \mu_2 &= \frac{s_2'}{n} - \left(\frac{s_1'}{n}\right)^2 \\ &= \frac{\Sigma(o-c)^2}{n} - (\mu_1 - c)^2. \end{aligned}\right\} \tag{26}$$

We have still to mention a theorem concerning the mean deviation, which, though not useful for computation, is useful for the comprehension and further development of the idea: The square of the mean deviation $\mu_2$ is equal to the sum of squares of the difference between each observed value and each of the others, divided by twice the square of the number. The said squares are:

$$(o_1 - o_1)^2,\ (o_2 - o_1)^2,\ \ldots\ (o_n - o_1)^2,$$
$$(o_1 - o_2)^2,\ (o_2 - o_2)^2,\ \ldots\ (o_n - o_2)^2,$$
$$\cdots\cdots\cdots\cdots$$
$$(o_1 - o_n)^2,\ (o_2 - o_n)^2,\ \ldots\ (o_n - o_n)^2;$$

developing each of these by the formula $(o_m - o_n)^2 = o_m^2 - 2o_m o_n + o_n^2$, and first adding each column separately, we find the sums

$$s_0 o_1^2 - 2 s_1 o_1 + s_2$$
$$s_0 o_2^2 - 2 s_1 o_2 + s_2$$
$$\cdots\cdots$$
$$s_0 o_n^2 - 2 s_1 o_n + s_2$$

and the sum of these

$$s_0 s_2 - 2 s_1 s_1 + s_2 s_0 = 2 (s_0 s_2 - s_1^2),$$

consequently,

$$\Sigma\Sigma(o_r - o_s)^2 = 2 s_0^2 \mu_2. \qquad (27)$$

The mean deviation is greater than the least, less than the greatest of the deviations of the values of repetitions from the mean number, and less than $\sqrt{\frac{1}{2}}$ of the greatest deviation between two observed values.

As to the higher half-invariants it may here be enough to state that they indicate various sorts of deviations from the typical form. Skew curves of errors are indicated by the $\mu_{2r+1}$ being different from zero, peaked or flattened (divided) forms respectively by positive or negative values of $\mu_{4r}$, and inversely by $\mu_{4r+2}$.

For these higher half-invariants we shall propose no special names. But we have already introduced double names "relative frequency" and "probability" in order to accentuate the distinction between the laws of actual errors and those of presumptive errors, and the same we ought to do for the half-invariants. In what follows we shall indicate the half-invariants in laws of presumptive errors by the signs $\lambda_r$ instead of $\mu_r$ which will be reserved for laws of actual errors, particularly when we shall treat of the transition from laws of actual errors to those of presumptive ones. For special reasons, to be explained later on, the name mean value can be used without confusion both for $\mu_1$ and $\lambda_1$, for actual as well as for presumptive means; but instead of "mean deviation" we say "mean error", when we speak of laws of presumptive errors. Thus, if $n = s_0$,

$$\lambda_2 = \text{Lim}_{n=\infty}(\mu_2)$$

is called the square of the mean error.

In speculations upon ideal laws of errors, when the laws are supposed to be continuous or to relate to infinite numbers of observations, this distinction is of course insignificant.

Examples:

1. Professor Jul. Thomsen found for the constant of a calorimeter, in experiments with pure water, in seven repetitions, the values

2649, 2647, 2645, 2653, 2653, 2646, 2649.

If we take here 2650 as zero, we read the observations as

$$-1, -3, -5, +3, +3, -4, -1$$

so that

$$s'_0 = 7,\ s'_1 = -8, \text{ and } s'_2 = 70;$$

consequently

$$\mu_1 = 2650 - \tfrac{8}{7} = 2649.$$

$$\mu_2 = \tfrac{70}{7} - \left(-\tfrac{8}{7}\right)^2 = 9.$$

The mean deviation is consequently $\pm 3$.

2. In an alternative experiment the result is either "yes", which counts 1, or "no", which counts 0. Out of $m+n$ repetitions the $m$ have given "yes", the $n$ "no". What then is the expression for the law of errors in half-invariants?

Answer: $\mu_1 = \frac{m}{m+n}$, $\mu_2 = \frac{mn}{(m+n)^2}$, $\mu_3 = \frac{mn(n-m)}{(m+n)^3}$, $\mu_4 = \frac{mn(m^2-4mn+n^2)}{(m+n)^4}$.

3. Determine the law of errors, in half-invariants, of a voting in which $a$ voters have voted for a motion $(+1)$, $c$ against $(-1)$, while $b$ have not voted $(0)$, and examine what values for $a$, $b$, and $c$ give the nearest approximation to the typical form.

$$\mu_1 = \frac{a-c}{a+b+c},\quad \mu_2 = \frac{ab+4ca+bc}{(a+b+c)^2},\quad \mu_3 = \frac{(c-a)(ab+8ca+bc-b^2)}{(a+b+c)^3},$$

$$\mu_4 = -\frac{((a+c)(a+b+c)-4(a-c)^2)(a+b+c)(2a-b+2c)+6(a-c)^4}{(a+b+c)^4}.$$

Disregarding the case when the vote is unanimous, the double condition $\mu_3 = \mu_4 = 0$ is only satisfied when one sixth of the votes is for, another sixth against, while two thirds do not give their votes. If $\mu_3$ is to be $= 0$, without $a$ being $= c$, $b^2 - b(a+c) - 8ac$ must be $= 0$. But then $\mu_4 = -2\mu_2\left(\frac{a-c}{a+b+c}\right)^2$, which does not disappear unless two of the numbers $a$, $b$, and $c$, and consequently $\mu_2$, are $= 0$.

4. Six repetitions give the quite symmetrical and almost typical law of errors, $\mu_1 = 0$, $\mu_2 = \frac{1}{3}$, $\mu_3 = \mu_4 = \mu_5 = 0$, but $\mu_6 = -\frac{2}{9}$. What are the observed values?

Answer: $-1, 0, 0, 0, 0, +1$.

## VII. RELATIONS BETWEEN FUNCTIONAL LAWS OF ERRORS AND HALF-INVARIANTS.

§ 24. The multiplicity of forms of the laws of errors makes it impossible to write a Theory of Observations in a short manner. For though these forms are of very different value, none of them can be considered as absolutely superior to the others. The functional form which has been universally employed hitherto, and by the most prominent writers, has in my opinion proved insufficient. I shall here endeavour to replace it by the half-invariants.