Expressed by half-invariants in this manner the explicit form of equation (6) is

$$\vartheta(x) = \frac{s_0}{\sqrt{2\pi\lambda_2}} e^{-\frac{1}{2}\frac{(x-\lambda_1)^2}{\lambda_2}} \left\{ \begin{array}{l} 1 + \frac{\lambda_3}{6\lambda_2^3}((x-\lambda_1)^3 - 3\lambda_2(x-\lambda_1)) + \\ + \frac{\lambda_4}{24\lambda_2^4}((x-\lambda_1)^4 - 6\lambda_2(x-\lambda_1)^2 + 3\lambda_2^2) + \\ + \frac{\lambda_5}{120\lambda_2^5}((x-\lambda_1)^5 - 10\lambda_2(x-\lambda_1)^3 + 15\lambda_2^2(x-\lambda_1)) + \ldots \end{array} \right\} \quad (31)$$

VIII. LAWS OF ERRORS OF FUNCTIONS OF OBSERVATIONS.

§ 26. There is nothing inconsistent with our definitions in speaking of laws of errors relating to any group of quantities which, though not obtained by repeated observations, have the like property, namely, that repeated estimations of a single thing give rise, owing to errors of one kind or other, to multiple and slightly differing results which are *prima facie* equally valid. The various forms of laws of actual errors are indeed only summary expressions for such multiplicity; and the transition to the law of presumptive errors requires, besides this, only that the multiplicity is caused by fixed but unknown circumstances, and that the values must be mutually independent in that sense that none of the circumstances have connected some repetitions to others in a manner which cannot be common to all. Compare § 24, Example 6.

It is, consequently, not difficult to define the law of errors for a function of *one* single observation. Provided only that the function is univocal, we can from each of the observed values $o_1, o_2 \ldots o_n$ determine the corresponding value of the function, and

$$f(o_1),\ f(o_2),\ \ldots\ f(o_n)$$

will then be the series of repetitions in the law of errors of the function, and can be treated quite like observations.

With respect, however, to those forms of laws of errors which make use of the idea of frequency (probability) we must make one little reservation. Even though o_i and o_k are different, we can have $f(o_i) = f(o_k)$, and in this case the frequencies must evidently be added together. Here, however, we need only just mention this, and remark that the laws of errors when expressed by half-invariants or other symmetrical functions are not influenced by it.

Otherwise the frequency is the same for $f(o_i)$ as for o_i, and therefore also the probability. The ordinates of the curves of errors are not changed by observations with discontinuous values; but the abscissa o_i is replaced by $f(o_i)$, and likewise the argument in the functional law of errors. In continuous functions, on the other hand, it is the areas between corresponding ordinates which must remain unchanged.

In the form of symmetrical functions the law of errors of functions of observations may be computed, and not only when we know all the several observed values, and can therefore compute, for each of them, the corresponding value of the function, and at last the symmetrical functions of the latter. In many and important cases it is sufficient if we know the symmetrical functions of the observations, as we can compute the symmetrical functions of the functions directly from these. For instance, if $f(o) = o^2$; for then the sums of the powers s'_n of the squares are also sums of the powers s_m of the observations, if only constantly $m = 2n$; $s'_0 = s_0$, $s'_1 = s_2$, $s'_2 = s_4$, etc.

§ 27. The principal thing is here a proposition as to laws of errors of the *linear functions* by half-invariants.

It is almost self-evident that if $o' = ao + b$

$$\left.\begin{aligned} \mu'_1 &= a\mu_1 + b \\ \mu'_2 &= a^2\mu_2 \\ \mu'_3 &= a^3\mu_3 \\ &\text{etc.} \\ \mu'_r &= a^r\mu_r \ (r > 1) \end{aligned}\right\} \tag{32}$$

For the linear functions can always be considered as produced by the change of both zero and unity of the observations (Compare (24)).

However special the linear function $ao + b$ may be, we always in practice manage to get on with the formula (32). That we can succeed in this is owing to a happy circumstance, the very same as, in numerical solutions of the problems of exact mathematics, brings it about that we are but rarely, in the neighbourhood of equal roots, compelled to employ the formulæ for the solution of other equations than those of the first degree. Here we are favoured by the fact that we may suppose the errors in *good* observations to be small, so small — to speak more exactly — that we may generally in repetitions for each series of observations $o_1, o_2, \ldots o_n$ assign a number c, so near them all that the squares and products and higher powers of the differences

$$o_1 - c,\ o_2 - c,\ \ldots\ o_n - c$$

without any perceptible error may be left out of consideration in computing the function: i. e., these differences are treated like differentials. The differential calculus gives a definite method, in such circumstances, for transforming any function $f(o)$ into a linear one

$$f(o) = f(c) + f'(c)\cdot(o - c).$$

The law of errors then becomes

$$\left.\begin{aligned} \mu_1(f(o)) &= f(c) + f'(c)(\mu_1(o) - c) = f(\mu_1(o)) \\ \mu_r(f(o)) &= (f'(c))^r \mu_r(o) \end{aligned}\right\} \tag{33}$$

But also by quite elementary means and easy artifices we may often transform functions into others of linear form. If for instance $f(o) = \frac{1}{o}$, then we write

$$\frac{1}{o} = \frac{1}{c + (o - c)} = \frac{c - (o - c)}{c^2 - (o - c)^2} = \frac{1}{c} - \frac{1}{c^2}(o - c),$$

and the law of errors is then

$$\mu_1\left(\frac{1}{o}\right) = \frac{1}{c} - \frac{1}{c^2}(\mu_1(o) - c)$$

$$\mu_2\left(\frac{1}{o}\right) = \frac{1}{c^4}\mu_2(o)$$

$$\mu_r\left(\frac{1}{o}\right) = \frac{(-1)^r}{c^{2r}}\mu_r(o).$$

§ 28. With respect to *functions of two or more observed quantities* we may also, in case of repetitions, speak of laws of errors, only we must define more closely what we are to understand by repetitions. For then another consideration comes in, which was out of the question in the simpler case. It is still necessary for the idea of the law of errors of $f(o, o')$ that we should have, for each of the observed quantities o and o', a series of statements which severally may be looked upon as repetitions:

$$o_1, \; o_2, \; \ldots\ldots \; o_m$$
$$o'_1, \; o'_2, \; \ldots\ldots \; o'_n.$$

But here this is not sufficient. Now it makes a difference if, among the special circumstances by o and o', there are or are not such as are common to observations of the different series. We want a technical expression for this. Here it is not appropriate only to speak of observations which are, respectively, dependent on one another or independent; we are led to mistake the partial dependence of observations for the functional dependence of exact quantities. I shall propose to designate these particular interdependences of repetitions of different observations by the word "bond", which presumably cannot cause any misunderstanding.

Among the repetitions of a single observation, no other bonds must be found than such as equally bind all the repetitions together, and consequently belong to the pecularities of the method. But while, for instance, several pieces cast in the same mould may be fair repetitions of one another, and likewise one dimension measured once on each piece, two or more dimensions measured on the same piece must generally be supposed to be bound together. And thus there may easily exist bonds which, by community in a circumstance, as here the particularities in the several castings, bind some or all the repetitions of a series each to its repetition of another observation; and if observations thus connected are to enter into the same calculation, we must generally take these bonds into account. This, as a rule, can only be done by proposing a theory or hypothesis as to the mathematical dependence between the observed objects and their common circumstance, and whether the number which expresses this is known from observation or quite unknown, the right treatment falls under those methods of adjustment which will be mentioned later on.

It is then in a few special cases only that we can determine laws of errors for functions of two or more observed quantities, in ways analogous to what holds good of a single observation and its functions.

If the observations o, o', o'' ..., which are to enter into the calculation of $f(o, o', o'', \ldots)$, are repeated in such a way that, in general, o_i, o'_i, o''_i, ... of the i'th repetition are connected by a common circumstance, the same for each i, but otherwise without any other bonds, we can for each i compute a value of the function $y_i = f(o_i, o'_i, o''_i, \ldots)$, and laws of errors can be determined for this, in just the same way as for o separately. To do so we need no knowledge at all of the special nature of the bonds.

§ 29. If, on the contrary, there is no bond at all between the repetitions of the observations o, o', o'', ... — and this is the principal case to which we must try to reduce the others — then we must, in order to represent all the equally valid values of $y = f(o, o', o'', \ldots)$, herein combine every observed value for o with every one for o', for o'', etc., and all such values of y must be treated analogously to the simple repetitions of one single observed quantity. But while it may here easily become too great a task to compute y for each of the numerous combinations, we shall in this case be able to compute y's law of errors by means of the laws of errors for o, o', o'' ...

Concerning this a number of propositions might be laid down; but one of them is of special importance and will be almost sufficient for us in what follows, viz., that which teaches us to determine the law of errors for the sum O of the observed quantities o and o'.

If the law of errors is given in the form of relative frequencies or probabilities, $\varphi(o)$ for o and $\psi(o')$ for o', then it is obvious that the product $\varphi(o)\psi(o')$ must be the frequency of the special sum $o + o'$.

In the calculus of probabilities we shall consider this form more closely, and there some cases of bound observations will find their solution; here we shall confine ourselves to the treatment of the said case with half-invariants.

If o occurs with the observed values

$$o_1,\ o_2,\ \ldots\ldots\ o_m$$

and o' with

$$o'_1,\ o'_2,\ \ldots\ldots\ o'_n,$$

then by the mn repetitions of the operation $O = o + o'$ we get:

$$o_1 + o'_1,\ o_1 + o'_2,\ \dots\dots\ o_1 + o'_n,$$
$$o_2 + o'_1,\ o_2 + o'_2,\ \dots\dots\ o_2 + o'_n,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$o_m + o'_1,\ o_m + o'_2,\ \dots\dots\ o_m + o'_n.$$

Indicating by M_r the half-invariants of the sum $O = o + o'$, we get by (18)

$$m \cdot n \cdot e^{\frac{M_1}{1!}\tau + \frac{M_2}{2!}\tau^2 + \frac{M_3}{3!}\tau^3 + \dots} = \Sigma e^{(o+o')\tau} = (e^{o_1\tau} + \dots e^{o_m\tau})(e^{o'_1\tau} + \dots e^{o'_n\tau})$$

where m and n are the numbers of repetitions of o and o'. Consequently, if μ_r represent the half-invariants of o, and μ'_r of o', we get

$$e^{\frac{M_1}{1!}\tau + \frac{M_2}{2!}\tau^2 + \dots} = e^{\frac{\mu_1}{1!}\tau + \frac{\mu_2}{2!}\tau^2 + \dots} \, e^{\frac{\mu'_1}{1!}\tau + \frac{\mu'_2}{2!}\tau^2 + \dots}$$

and finally

$$\left.\begin{array}{l} M_1 = \mu_1 + \mu'_1 \\ \dots\dots\dots\dots \\ M_r = \mu_r + \mu'_r \end{array}\right\} \qquad (34)$$

Employing the equation (17) instead of (18) we can also obtain fairly simple expressions for the sums of powers of $(o + o')$ analogous to the binomial formula. But the extreme simplicity of (34) renders the half-invariants unrivalled as the most suitable symmetrical functions and the most powerful instrument of the theory of observations.

More generally, for every linear function of observations not connected by any bond,

$$O = a + bo + co' + \dots do''',$$

we obtain in the same manner and by (32)

$$\left.\begin{array}{l} M_1(o) = a + b\mu_1 + c\mu'_1 + \dots + d\mu'''_1 \\ M_2(o) = \qquad b^2\mu_2 + c^2\mu'_2 + \dots + d^2\mu'''_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ M_r(o) = \qquad b^r\mu_r + c^r\mu_r + \dots + d^r\mu'''_r \\ \qquad\qquad r > 1. \end{array}\right\} \qquad (35)$$

When the errors of observation are sufficiently small, we shall also here generally be able to give the most different functions a linear form. In consequence of this, the propositions (34) and (35) acquire an almost universal importance, and afford nearly the whole necessary foundation for the theory of the laws of errors of functions.

Example 1. Determine the square of the mean error for differences of the n'th order of equidistant tabular values, between which there is no bond, the square of the mean error for every value being $= \lambda_2$.

$$\lambda_2(\Delta^1) = \lambda_2(o_1 - o_0) = 2\lambda_2$$
$$\lambda_2(\Delta^2) = \lambda_2(o_2 - 2o_1 + o_0) = 6\lambda_2$$
$$\lambda_2(\Delta^3) = \lambda_2(o_3 - 3o_2 + 3o_1 - o_0) = 20\lambda_2$$
$$\lambda_2(\Delta^4) = \lambda_2(o_4 - 4o_3 + 6o_2 - 4o_1 + o_0) = 70\lambda_2$$

. . . .

$$\lambda_2(\Delta^n) = \frac{2}{1}\cdot\frac{6}{2}\cdot\frac{10}{3}\cdot\frac{14}{4}\cdots\frac{4n-2}{n}\lambda_2.$$

Example 2. By the observation of a meridional transit we observe two quantities, viz. the time, t, when a star is covered behind a thread, and the distance, f, from the meridian at that instant. But as it may be assumed that the time and the distance are not connected by a bond, and as the speed of the star is constant and proportional to the known value sin p (p = polar distance), we always state the observation by the one quantity, the time when the very meridian is passed, which we compute by the formula $o = t + f \operatorname{cosec} p$.

The mean error is

$$\lambda_2(o) = \lambda_2(t) + \operatorname{cosec}^2 p\, \lambda_2(f).$$

Example 3. A scale is constructed by making marks on it at regular intervals, in such a way that the square of the mean error on each interval is $= \lambda_2$.

To measure the distance between two objects, we determine the distance of each object from the nearest mark, the square of the mean error of this observation being $= \lambda_2'$. How great is the mean error in a measurement, by which there are n intervals between the marks we use?

$$\lambda_2(\text{length}) = n\lambda_2 + 2\lambda_2'.$$

Example 4. Two points are supposed to be determined by bond-free and equally good ($\lambda_2 = 1$) measurements of their rectangular co-ordinates. The errors being small in proportion to the distance, how great is the mean error in the distance Δ?

$$\lambda_2(\Delta) = 2.$$

Example 5. Under the same suppositions, what is the mean error in the inclination to the x-axis?

$$\lambda_2(B) = \frac{2}{\Delta^2}.$$

Example 6. Having three points in a plane determined in the same manner by their rectangular co-ordinates (x_1, y_1), (x_2, y_2), (x_3, y_3), find the mean error of the angle at the point (x_1, y_1)

$$\lambda_2(V) = \frac{\Delta_1^2 + \Delta_2^2 + \Delta_3^2}{\Delta_2^2 \Delta_3^2},$$

Δ_1, Δ_2, Δ_3 being the sides of the triangle; Δ_1 opposite to (x_1, y_1).

Examples 7 and 8. Find the mean errors in determinations of the areas of a triangle and a plane quadrangle.

$$\lambda_2 \text{ (triangle)} = \tfrac{1}{4}(\Delta_1^2 + \Delta_2^2 + \Delta_3^2); \quad \lambda_2 \text{ (quadrangle)} = \tfrac{1}{4}\left(\Delta^2_{1,3} + \Delta^2_{2,4}\right).$$

§ 30. Non-linear functions of more than one argument present very great difficulties. Even for integral rational functions no general expression for the law of errors can be found. Nevertheless, even in this case it is possible to indicate a method for computing the half-invariants of the function by means of those of the arguments. To do so it seems indispensable to transform the laws of errors into the form of systems of sums of powers. If $O = f(o, o', \ldots o^{m})$ be integral and rational, both it and its powers O^r can be written as sums of terms of the standard form $\Sigma k o^a \cdot o'^b \ldots o^{(m)d}$, and for every such term the sum resulting from the combination of all repetitions is $ks_a \cdot s'_b \ldots s_d^{(m)}$ (including the cases where a or b or d may be $= 0$), $s_c^{(i)}$ being the sum of all c^{th} powers of the repetitions of $o^{(i)}$. Thus if S_r indicates the sum of the r^{th} powers of the function O, we get

$$S_r = \Sigma k s_a \cdot s'_b \ldots s_d^{(m)}.$$

Of course, this operation is only practicable in the very simplest cases.

Example 1. Determine the mean value and mean deviation of the product $oo' = O$ of two observations without bonds. Here $S_0 = s_0 s'_0$ and generally $S_r = s_r s'_r$, consequently the mean value $M_1 = \mu_1 \mu'_1$ and

$$M_2 = \mu_2 \mu'_2 + \mu_2 \mu'^2_1 + \mu'_2 \mu^2_1.$$

M_3 already takes the cumbersome form

$$M_3 = \mu_3 \mu'_3 + \mu_3 \mu'_1 (3\mu'_2 + \mu'^2_1) + \mu'_3 \mu_1 (3\mu_2 + \mu^2_1) + 6\mu_2 \mu_1 \mu'_2 \mu'_1.$$

Example 2. Express exactly by the half-invariants of the co-ordinates the mean value and the mean deviation of the square of the distance $r^2 = x^2 + y^2$, if x and y are observed without bonds. Here

$$s_0(r^2) = s_0(x) s_0(y)$$
$$s_1(r^2) = s_2(x) s_0(y) + s_0(x) s_2(y)$$
$$s_2(r^2) = s_4(x) s_0(y) + 2 s_2(x) s_2(y) + s_0(x) s_4(y)$$

and

$$\mu_1(r^2) = \mu_2(x) + (\mu_1(x))^2 + \mu_2(y) + (\mu_1(y))^2$$
$$\mu_2(r^2) = \mu_4(x) + 4\mu_3(x)\mu_1(x) + 2(\mu_2(x))^2 + 4\mu_2(x)(\mu_1(x))^2 +$$
$$+ \mu_4(y) + 4\mu_3(y)\mu_1(y) + 2(\mu_2(y))^2 + 4\mu_2(y)(\mu_1(y))^2.$$

§ 31. The most important application of proposition (35) is certainly the determination of the law of errors of the mean value itself. The mean value

$$\mu_1 = \frac{1}{n}(o_1 + o_2 + \ldots o_n)$$

is, we know, a linear function of the observed values, and we may treat the law of errors for μ_1 according to the said proposition, not only where we look upon $o_1, \dots o_n$ as perfectly unconnected, but also where we assume that they result from repetitions made according to the same method. For, just like such repetitions, $o_1, \dots o_n$ must not have any other circumstances in common as connecting bonds than such as bind them all and characterise the method.

As the law of presumptive errors of o_1 is just the same as for $o_2 \dots o_m$, with the known half-invariants $\lambda_1, \lambda_2, \dots \lambda_r \dots$, we get according to (35)

$$\left.\begin{aligned} \lambda_1(\mu_1) &= \frac{1}{m}(\lambda_1 + \dots + \lambda_1) = \lambda_1, \\ \lambda_2(\mu_1) &= \frac{1}{m^2}(\lambda_2 + \dots + \lambda_2) = \frac{1}{m}\lambda_2 \\ \text{and in general} \quad \lambda_r(\mu_1) &= m^{1-r} \cdot \lambda_r. \end{aligned}\right\} \quad (37)$$

While, consequently, the presumptive mean of a mean value for m repetitions is the presumptive mean itself, the mean error on the mean value μ_1 is reduced to $\frac{1}{\sqrt{m}}$ of the mean error on the single observation. When the number m is large, the formation of mean values consequently reduces the uncertainty considerably; the reduction, however, is proportionally greater with small than with large numbers. While already 4 repetitions bring down the uncertainty to half of the original, 100 repetitions are necessary in order to add one significant figure, and a million to add 3 figures to those due to the single observation.

The higher half-invariants of μ_1 are reduced still more. If the λ_3, λ_4, etc., of the single observation are so large that the law of errors cannot be called typical, no very great numbers of m will be necessary to realise the conditions $\lambda_3(\mu_1) = 0 = \lambda_4(\mu_1)$ with an approximation that is sufficient in practice. It ought to be observed that this reduction is not only absolute, but it holds good also in relation to the corresponding power of the mean error $\sqrt{\lambda_2(\mu_1)}^r$; for (37) gives

$$\lambda_r(\mu_1) : (\lambda_2(\mu_1))^{\frac{r}{2}} = m^{1-\frac{r}{2}} \cdot (\lambda_r : \lambda_2^{\frac{r}{2}}),$$

which, for instance when $m = 4$, shows that the deviation of λ_3 from the typical form which appears by means of only 4 repetitions, is halved; that of λ_4 is divided by 4, that of λ_6 is divided by 8, etc. This shows clearly the reason why we *attach great importance to the typical form for the law of errors* and make arrangements to abide by it in practice. For it appears now that *we possess in the formation of mean values a means of making the laws of errors typical,* even where they were not so originally. Therefore the standard rule for all practical observations is this: Take care not to neglect any opportunities of repeating observations and parts of observations, so that you can directly form the mean values which should be substituted for the observed results; and this is to be done especially in the case of observations of a novel character, or with peculiarities which lead us to doubt whether the law of errors will be typical.

This remarkable property is peculiar, however, not to the mean only, but also, though with less certainty, to any linear function of several observations, provided only the coefficient of any single term is not so great relatively to the corresponding deviation from the typical form that it throws all the other terms into the shade. From (35) it is seen that, if the laws of errors of all the observations $o, o', \ldots o^{(m)}$ are typical, the law of errors for any of their linear functions will be typical too. And if the laws of errors are not typical, then that of the linear function will deviate relatively less than any of the observations $o, o', \ldots o_m$.

To avoid unnecessary complication we represent two terms of the linear function simply by o and o'. The deviation from the typical zero, which appears in the r^{th} half-invariants ($r > 2$), measured by the corresponding power of the mean error, will be less for $O = o + o'$ than for the most discrepant of the terms o and o'.

The inequation

$$\frac{\lambda_r^2}{\lambda_2^r} \overline{\gtrless} \frac{\lambda_r'^2}{\lambda_2'^r}$$

says only that, if the laws of errors for o and o' deviate unequally from the typical form, it is the law of errors for o that deviates most. But this involves

$$\left(\frac{\lambda_2'}{\lambda_2}\right)^r \overline{\gtrless} \left(\frac{\lambda_r'}{\lambda_r}\right)^2$$

or more briefly

$$T^r \overline{\gtrless} R^2,$$

where T is positive, $r > 2$.

When we introduce a positive quantity U, so that

$$T^r = U^2 \overline{\gtrless} R^2,$$

it is evident that $(U+1)^2 \overline{\gtrless} (R+1)^2$, and it is easily demonstrated that $(T+1)^r > (U+1)^2$.

Remembering that $x + x^{-1} \overline{\gtrless} 2$, if $x > 0$, we get by the binomial formula

$$\left(U^{\frac{1}{r}} + U^{-\frac{1}{r}}\right)^r \overline{\gtrless} U + U^{-1} + 2^r - 2 > (U^{\frac{1}{2}} + U^{-\frac{1}{2}})^2.$$

Consequently

$$(T+1)^r > (U+1)^2 \overline{\gtrless} (R+1)^2$$

or

$$\left(\frac{\lambda_2 + \lambda_2'}{\lambda_2}\right)^r > \left(\frac{\lambda_r + \lambda_r'}{\lambda_r}\right)^2$$

and

$$\frac{\lambda_r^2}{\lambda_2^r} > \frac{(\lambda_r + \lambda_r')^2}{(\lambda_2 + \lambda_2')^r} - \frac{(\lambda_r(O))^2}{(\lambda_2(O))^r},$$

but this is the proposition we have asserted, for the extension to any number of terms causes no difficulty.

But if it thus becomes a general law that the law of errors of linear functions must more or less approach the typical form, the same must hold good also of all moderately complex observations, such as those whose errors arise from a considerable number of sources. The expression "source of errors" is employed to indicate circumstances which undeniably influence the result, but which we have been obliged to pass over as unessential. If we imagined these circumstances transferred to the class of essential circumstances, and substantiated by subordinate observations, that which is now counted an observation would occur as a function, into which the subordinate observations enter as independent variables; and as we may assume, in the case of good observations, that the influence of each single source of errors is small, this function may be regarded as linear. The approximation to typical form which its law of errors would thus show, if we knew the laws of errors of the sources of error, cannot be lost, simply because we, by passing them over as unessential, must consider the sources of error in the compound observation as unknown. Moreover, we may take it for granted that, in systematically arranged observations, every such source of error as might dominate the rest will be the object of special investigation and, if necessary, will be included among the essential circumstances or removed by corrective calculations. The result then is that great deviations from the typical form of the law of errors are rare in practice.

§ 32. It is of interest, of course, also to acquire knowledge of the laws of errors for the determinations of μ_2 and the higher half-invariants as functions of a given number of repeated observations.

Here the method indicated in § 30 must be applied. But though the symmetry of these functions and the identity of the laws of presumptive errors for $o_1, o_2, \ldots o_m$ afford very essential simplifications, still that method is too difficult. Not even for μ_2 have I discovered the general law of errors. In my *"Almindelig Iagttagelseslære"*, Kobenhavn 1889, I have published tables up to the eighth degree of products of the sums of powers $s_p\, s_q \ldots$, expressed by sums of terms of the form o^i, o^{ij}, o''^k; these are here directly applicable. In W. Fiedler: *"Elemente der neueren Geometrie und der Algebra der binären Formen"*, Leipzig 1862, tables up to the 10th degree will be found. Their use is more difficult, because they require the preliminary transformation of the s_p to the coefficients a_p of the rational equations § 21. There are such tables also in the *Algebra* by Meyer Hirsch, and Cayley has given others in the Philosophical Transactions 1857 (Vol. 147,

p. 480). I have computed the four principal half-invariants of μ_2:

$$
\left.\begin{aligned}
m\lambda_1(\mu_2) &= (m-1)\lambda_2 \\
m^3\lambda_2(\mu_2) &= (m-1)^2\lambda_4 + 2m(m-1)\lambda_2^2 \\
m^5\lambda_3(\mu_2) &= (m-1)^3\lambda_6 + 12m(m-1)^2\lambda_4\lambda_2 + 4m(m-1)(m-2)\lambda_3^2 + \\
&\qquad + 8m^2(m-1)\lambda_2^3 \\
m^7\lambda_4(\mu_2) &= (m-1)^4\lambda_8 + 24m(m-1)^3\lambda_6\lambda_2 + 32m(m-1)^2(m-2)\lambda_5\lambda_3 + \\
&\qquad + 8m(m-1)(4m^2-9m+6)\lambda_4^2 + 144m^2(m-1)^2\lambda_4\lambda_2^2 + \\
&\qquad + 96m^2(m-1)(m-2)\lambda_3^2\lambda_2 + 48m^3(m-1)\lambda_2^4.
\end{aligned}\right\} \quad (38)
$$

Here m is the number of repetitions.

Of μ_3 and μ_4 only the mean values and the mean errors have been found:

$$
\left.\begin{aligned}
m^2\lambda_1(\mu_3) &= (m-1)(m-2)\lambda_3, \\
m^5\lambda_2(\mu_3) &= (m-1)^2(m-2)^2\lambda_6 + 9m(m-1)(m-2)^2(\lambda_4\lambda_2 + \lambda_3^2) + \\
&\qquad + 6m^2(m-1)(m-2)\lambda_2^3;
\end{aligned}\right\} \quad (39)
$$

and

$$
\left.\begin{aligned}
m^3\lambda_1(\mu_4) &= (m-1)(m^2-6m+6)\lambda_4 - 6m(m-1)\lambda_2^2 \\
m^7\lambda_2(\mu_4) &= (m-1)^2(m^2-6m+6)^2\lambda_8 + \\
&+ 8m(m-1)(m^2-6m+6)(2m^2-15m+15)\lambda_6\lambda_2 + \\
&+ 18m(m-1)(m-2)(m-4)(m^2-6m+6)\lambda_5\lambda_3 + \\
&+ 2m(m-1)(17m^4-204m^3+852m^2-1404m+828)\lambda_4^2 + \\
&+ 24m^2(m-1)(3m^3-38m^2+150m-138)\lambda_4\lambda_2^2 + \\
&+ 144m^2(m-1)(m-2)(m-4)(m-5)\lambda_3^2\lambda_2 + \\
&+ 24m^3(m-1)(m^2-6m+24)\lambda_2^4.
\end{aligned}\right\} \quad (40)
$$

Further I know only that

$$
m^4\lambda_1(\mu_5) = (m-1)(m-2)\left\{(m^2-12m+12)\lambda_5 - 60m\,\lambda_3\lambda_2\right\}, \quad (41)
$$

$$
\begin{aligned}
m^5\lambda_1(\mu_6) &= (m-1)(m^4-30m^3+150m^2-240m+120)\lambda_6 - \\
&- 30m(m-1)(7m^2-36m+36)\lambda_4\lambda_2 - \\
&- 60m(m-1)(m-2)(3m-8)\lambda_3^2 - \\
&- 60m^2(m-1)(m-6)\lambda_2^3,
\end{aligned} \quad (42)
$$

$$
\begin{aligned}
m^6\lambda_1(\mu_7) &= (m-1)(m-2)(m^4-60m^3+420m^2-720m+360)\lambda_7 - \\
&- 630m(m-1)(m-2)(m^2-8m+8)\lambda_5\lambda_2 - \\
&- 210m(m-1)(m-2)(7m^2-48m+60)\lambda_4\lambda_3 - \\
&- 1260m^2(m-1)(m-2)(m-10)\lambda_3\lambda_2^2.
\end{aligned} \quad (43)
$$

$$
\begin{aligned}
m^7 \lambda_1(\mu_8) = {} & (m-1)(m^6 - 126m^5 + 1806m^4 - 8400m^3 + 16800m^2 - 15120m + 5040)\lambda_8 - \\
& - 56m(m-1)(31m^4 - 540m^3 + 2340m^2 - 3600m + 1800)\lambda_6\lambda_2 - \\
& - 1680m(m-1)(m-2)(3m^3 - 40m^2 + 120m - 96)\lambda_5\lambda_3 - \\
& - 70m(m-1)(49m^4 - 720m^3 + 3168m^2 - 5400m + 3240)\lambda_4^2 - \\
& - 840m^2(m-1)(7m^3 - 150m^2 + 576m - 540)\lambda_4\lambda_2^2 - \\
& - 10080m^2(m-1)(m-2)(m^2 - 18m + 40)\lambda_3^2\lambda_2 - \\
& - 840m^3(m-1)(m^2 - 30m + 90)\lambda_2^4 .
\end{aligned}
\tag{44}
$$

Some λ_1's of products of the μ_2, μ_3, and μ_4 present in general the same characteristics as the above formulæ. The most prominent of these characteristics are:

1) It is easily explained that λ_1 is only to be found in the equation $\lambda_1(\mu_1) = \lambda_1$; indeed no other half-invariant than the mean value can depend on the zero of the observations. In my computations this characteristic property has afforded a system of multiple checks of the correctness of the above results.

2) All mean $\lambda_1(\mu_r)$ are functions of the 0th degree with regard to m, all squares of mean errors $\lambda_2(\mu_r)$, are of the (-1)st degree, and generally each $\lambda_s(\mu_r)$ is a function of the $(1-s)$th degree, in perfect accordance with the law of large numbers.

3) The factor $m-1$ appears universally as a necessary factor of $\lambda_s(\mu_r)$, if only $r > 1$. If r is an odd number, even the factor $m-2$ appears, and, likewise, if r is an even number, this factor is constantly found in every term that is multiplied by one or more λ's with odd indices. No obliquity of the law of errors can occur unless at least three repetitions are under consideration.

4) Many particulars indicate these functions as compounds of factorials $(m-1)(m-2)\dots(m-r)$ and powers of m.

If, supposing the presumed law of errors to be typical, we put $\lambda_3 = \lambda_4 = \dots = 0$, then some further inductions can be made. In this case the law of errors of μ_2 may be

$$
e^{\frac{\lambda_1(\mu_2)}{\lfloor 1}\tau + \frac{\lambda_2(\mu_2)}{\lfloor 2}\tau^2 + \dots} = \left(1 - \frac{2\lambda_2\tau}{m}\right)^{\frac{1-m}{2}} = \int_{-\infty}^{+\infty} \varphi(o)\, e^{o\tau} do .
\tag{45}
$$

As to the squares of mean errors of μ_r we get under the same supposition:

$$
\left.
\begin{aligned}
\lambda_2(\mu_1) &= \tfrac{1}{m}\lambda_2 \\
\lambda_2(\mu_2) &= \tfrac{2}{m}\lambda_2^2 \\
\lambda_2(\mu_3) &= \tfrac{6}{m}\lambda_2^3 \\
\lambda_2(\mu_4) &= \tfrac{24}{m}\lambda_2^4 ,
\end{aligned}
\right\}
\tag{46}
$$

indicating that generally

$$
\lambda_2(\mu_r) = \frac{\lfloor r}{m}\lambda_2^r .
$$

This proposition is of very great interest. If we have a number m of repetitions at our disposal for the computation of a law of actual errors, then it will be seen that the relative mean errors of $\mu_1, \mu_2, \mu_3 \dots \mu_r$ are by no means uniform, but increase with the index r. If m is large enough to give us μ_1 precisely and μ_2 fairly well, then μ_3 and μ_4 can be only approximately indicated; and the higher half-invariants are only to be guessed, if the repetitions are not counted by thousands or millions.

As all numerical coefficients in $\lambda_2(\mu_r)$ increase with r, almost in the same degree as the coefficients 1, 2, 6, and 24 of λ_2^r, we must presume that the law of increasing uncertainty of the half-invariants has a general character.

We have hitherto been justified in speaking of the principal half-invariants as the complete collection of the μ_r's or λ_r's with the lowest indices, considering a complete series of the first m half-invariants to be necessary to an unambiguous determination of a law of errors for m repetitions.

We now accept that principle as a system of relative rank of the half-invariants with increasing uncertainty and consequently with a decreasing importance of the half-invariants with higher indices.

We need scarcely say that there are some special exceptions to this rule. For instance if $\lambda_4 = -\lambda_2^2$, as in alternative experiments with equal chances for and against (pitch and toss), then $\lambda_2(\mu_2)$ is reduced to $= \frac{2(m-1)}{m^3}\lambda_2^2$, which is only of the $(-2)^{nd}$ order.

§ 33. Now we can undertake to solve the main problem of the theory of observations, the transition from laws of actual errors to those of presumptive errors. Indeed this problem is not a mathematical one, but it is eminently practical. To reason from the actual state of a finite number of observations to the law governing infinitely numerous presumed repetitions is an evident trespass; and it is a mere attempt at prophecy to predict, by means of a law of presumptive errors, the results of future observations.

The struggle for life, however, compels us to consult the oracles. But the modern oracles must be scientific; particularly when they are asked about numbers and quantities, mathematical science does not renounce its right of criticism. We claim that confusion of ideas and every ambiguous use of words must be carefully avoided; and the necessary act of will must be restrained to the acceptation of fixed principles, which must agree with the law of large numbers.

It is hardly possible to propose more satisfactory principles than the following:

The mean value of all available repetitions can be taken directly, without any change, as an approximation to the presumptive mean.

If only one observation without repetition is known, it must itself, consequently, be considered an approximation to the presumptive mean value.

The solitary value of any symmetrical and univocal function of repeated observations must in the same way, as an isolated observation, be considered the presumptive mean of this function, for instance $\mu_r = \lambda_1 (\mu_r)$.

Thus, from the equations 37—41, we get by m repetitions:

$$\left.\begin{aligned}
\lambda_1 &= \mu_1 \\
\lambda_2 &= \frac{m}{m-1}\mu_2 \\
\lambda_3 &= \frac{m^2}{(m-1)(m-2)}\mu_3 \\
\lambda_4 &= \frac{m^3}{(m-1)(m^2-6m+6)}\left(\mu_4 + \frac{6}{m-1}\mu_2^2\right) \\
\lambda_5 &= \frac{m^4}{(m-1)(m-2)(m^2-12m+12)}\left(\mu_5 + \frac{60}{m-1}\mu_2\mu_3\right);
\end{aligned}\right\} \qquad (47)$$

as to λ_6, λ_7, λ_8 it is preferable to use the equations 42—44 themselves, putting only $\lambda_1(\mu_6) = \mu_6$, $\lambda_1(\mu_7) = \mu_7$, and $\lambda_1(\mu_8) = \mu_8$.

Inversely, if the presumptive law of errors is known in this way, or by adoption of any theory or hypothesis, we *predict the future observations, or functions of observations, principally by computing their presumptive mean values.* These predictions however, though univocal, are never to be considered as exact values, but only as the first and most important terms of laws of errors.

If necessary, we complete our predictions with the mean errors and higher half-invariants, computed for the predicted functions of observations by the presumed law of errors, which itself belongs to the single observations. These supplements may often be useful, nay necessary, for the correct interpretation of the prediction. The ancient oracles did not release the questioner from thinking and from responsibility, nor do the modern ones; yet there is a difference in the manner. If the crossing of a desert is calculated to last 20 days, with a mean error of one day, then you would be very unwise, to be sure, if you provided for exactly 20 days; by so doing you incur as great a probability of dying as of living. Even with provisions for 21 days the journey is evidently dangerous. But if you can carry with you provisions for 23—25 days, the undertaking may be reasonable. Your life must be at stake to make you set out with provisions for only 17 days or less.

In addition to the uncertainty provided against by the presumptive law of error, the prediction may be vitiated by the uncertainty of the data of the presumptive law itself. When this law has resulted from purely theoretical speculation, it is always impossible to calculate its uncertainty. It may be quite exact, or partially or absolutely false, we are left to choose between its admission and its rejection, as long as no trial of the prediction by repeated observations has given us a corresponding law of actual errors, by which it can be improved on.

If the law of presumptive errors has been computed by means of a law of actual errors, we can, according to (37), employ the values λ_2, λ_3, ... and the number m of actual observations for the determination of $\lambda_r(\mu_1)$. In this case the complete half-invariants of a predicted single observation are given analogously to the law of errors of the sum of two bondless observations by

$$\begin{array}{l} \lambda_1 \\ \lambda_2 + \lambda_2(\mu_1) \\ \cdots\cdots\cdots \\ \lambda_r + \lambda_r(\mu_1). \end{array}$$

Though we can in the same way compute the uncertainties of λ_2, λ_3, and λ_4, it is far more difficult, or rather impossible, to make use of these results for the improvement of general predictions.

Of the higher half-invariants we can very seldom, if ever, get so much as a rough estimate by the method of laws of actual errors. The same reasons that cause this difficulty, render it a matter of less importance to obtain any precise determination. Therefore the general rule of the formation of good laws of presumptive errors must be:

1. In determining λ_1 and λ_2, rely almost entirely upon the actual observed values.

2. As to the half-invariants with high indices, say from λ_6 upwards, rely as exclusively upon theoretical considerations.

3. Employ the indications obtainable by actual observed values for the intermediate half-invariants as far as possible when you have the choice between the theories in (2).

From what is said above of the properties of *the typical law of errors*, it is evident that no other theory can fairly rival it in the multiplicity and importance of applications. It is not only constantly applied when λ_3, λ_4, and λ_5 are proved to be very small, but it is used almost universally as long as the deviations are not very conspicuous. In these cases also great efforts will be made to reduce the observations to the typical form by modifying the methods or by substituting means of many observed values instead of the non-typical single observations. The preference for the typical observations is intensified by the difficulty of establishing an entirely correct method of adjustment (see the following chapters) of observations which are not typical.

In those particular cases where λ_3 or λ_4 or λ_5 cannot be regarded as small, the theoretical considerations (proposition 2 above) as to λ_6 and the higher half-invariants ought not to result in putting the latter $= 0$. As shown in "*Videnskabernes Selskabs Oversigter*", 1899, p. 140, such laws of errors correspond to divergent series or imply the existence of imaginary observations. The coefficients k_r of the functional law of errors (equation (6)) have this merit in preference to the half-invariants, that no term implies the existence of any other.

This series

$$\vartheta(x) = k_0 \varphi(x) - \frac{k_3}{\underline{|3}} D^3 \varphi(x) + \frac{k_4}{\underline{|4}} D^4 \varphi(x) - \frac{k_5}{\underline{|5}} D^5 \varphi(x) + \frac{k_6}{\underline{|6}} D^6 \varphi(x) \dots,$$

where $\varphi(x) = e^{-\frac{(x-\lambda_1)^2}{2\lambda_2}}$ (the direct expression (31) is found p. 35), is therefore recommended as perhaps the best general expression for non-typical laws of errors. The functional form of the law of errors has here, and in every prediction of future results, the advantage of showing directly the probabilities of the different possible values.

The skew and other non-typical laws of errors seem to have some very interesting applications to biological observations, especially to the variations of species. The scientific treatment of such variations seems altogether to require a methodical use of the notion of laws of errors. Mr. K. Pearson has given a series of skillful computations of biological and other similar laws of errors (*Contributions to the Math. Theory of Evolution, Phil. Trans.* V. 186, p. 343). Here he makes very interesting efforts to develop the refractory binomial functions into a basis for the treatment of skew laws of errors. But there are evidently no natural links between these functions and the biological problems, and the above formulæ (31) will prove to be easier and more powerful instruments. In cases of very abnormal or discontinuous laws of errors, more refined methods of adjustment are required.

Example 1. From the 500 experiments given in § 14 are to be calculated the presumptive half-invariants up to λ_5 and by (31) the frequencies of the special events out of a number of $s_0 = 500$ new repetitions. You will find $\lambda_1 = 11{\cdot}86$, $\lambda_2 = 4{\cdot}1647$, $\lambda_3 = 4{\cdot}708$, $\lambda_4 = 3{\cdot}895$, and $\lambda_5 = -26{\cdot}946$. A comparison of the computed frequencies with the observed ones gives:

Events	Frequency computed	Frequency observed	$o-c$
4	0·0	0	− 0·0
5	− 0·1	0	+ 0·1
6	− 0·3	0	+ 0·3
7	1·6	3	+ 1·4
8	12·3	7	− 5·3
9	39·6	35	− 4·6
10	78·2	101	+ 22·8
11	104·1	89	− 15·1
12	97·7	94	− 3·7
13	69·4	70	+ 0·6
14	42·8	46	+ 3·2

Events	Frequency computed	observed	o—c
15	26·7	30	+ 3·3
16	16·0	15	— 1·0
17	8·0	4	— 4·0
18	3·0	5	+ 2·0
19	0·8	1	+ 0·2
20	0·2	0	— 0·2
21	0·0	0	0·0

Example 2. Determine the law of errors by experiments with alternative results, either "yes" observed m times and every time indicated by 1, or "no" observed n times and indicated by 0. What is the square of the mean error for the single experiment?

$$\lambda_2 = \frac{mn}{(m+n)(m+n-1)};$$

for the probability determined by the whole series?

$$\lambda_2(\mu_1) = \frac{mn}{(m+n)^2(m+n-1)};$$

and for the frequency of "yes" in the $m+n$ experiments?

$$\lambda_2(s_1) = \frac{mn}{m+n-1}.$$

§ 34. If observations are made and repeated, although their presumptive mean value is previously known, exactly or very accurately, the law of presumptive errors of the half-invariants $\mu_2, \mu_3 \ldots$ must be computed by reducing the zero of the observation to the known λ_1. Putting thus $s_1 = 0$ and $\mu_1 = 0$ in the equations (19) and (21) we obtain in analogy to (38)—(41) the following modified equations, the number of repetitions being $= m$:

$$\left.\begin{aligned} \lambda_1(\mu_2) &= \mu_2 = \lambda_2 \\ \lambda_1(\mu_3) &= \mu_3 = \lambda_3 \\ \lambda_1(\mu_4) &= \mu_4 = \frac{m-3}{m}\lambda_4 - \frac{6}{m}\lambda_2^2 \\ \lambda_1(\mu_5) &= \mu_5 = \frac{m-10}{m}\lambda_5 - \frac{90}{m}\lambda_3\lambda_2. \end{aligned}\right\} \quad (48)$$

From the first of these equations we deduce the very important principle, that every mean of the squares of differences between repeated bond-free observations and their presumptive mean value is approximately equal to the square of the mean error

$$\frac{\Sigma(o-\lambda_1)^2}{m} = \lambda_2. \quad (49)$$

Consequently, for any isolated observed value we must expect that

$$(o - \lambda_1(o))^2 = \lambda_2(o). \tag{50}$$

§ 35. In the following chapters, and in almost all practical applications, we shall work only with the typical law of errors as our constant supposition. This gives simplicity and clearness, and thus $a \pm b$ may be recommended as a short statement of the law of errors, $a = \lambda_1$ indicating a result of an observation found directly or indirectly by computation with observations, and $b = \sqrt{\lambda_2}$ expressing the mean error of the same result.

By the *"weights"* of observations we understand numbers inversely proportional to the squares of the mean errors, consequently $v = \frac{k}{\lambda_2}$. The idea presents itself when we speak of the means of various numbers of observed values which have been obtained by the same method, as the latter numbers here, according to (37), represent the weights When v_r is the weight of the partial mean value m_r, the total mean value m must be computed according to the formula

$$m = \frac{m_1 v_1 + m_2 v_2 + \ldots + m_r v_r}{v_1 + v_2 + \ldots + v_r}, \tag{51}$$

which is analogous to the formula for the abscissa of the centre of gravity, if m_r is the abscissa of any single body, v_r its *weight*. We speak also of the weights of single observations, according to the above definition, and particularly in cases where we can only estimate the relative goodness of several observations in comparison to the trustworthiness of the means of various numbers of equally good observations.

The phrase *probable error*, which we still find frequently employed by authors and observers, is for several reasons objectionable. It can be used only with typical or at any rate symmetrical laws of errors, and indicates then the magnitude of errors for which the probabilities of smaller and larger errors are both equal to $\frac{1}{2}$. The simultaneous use of the ideas "mean error" and "probable error" causes confusion, and it is evidently the latter that must be abandoned, as it is less commonly applicable, and as it can only be computed in the cases of the typical law of errors by the previously computed mean error as $0{\cdot}6745\sqrt{\lambda_2}$, while on the other hand the computation of the mean error is quite independent of that of the probable error. As errors which are larger than the probable ones, still frequently occur, this idea is not so well adapted as the mean error to serve as a limit between the frequent "small" errors and the rarer "large" ones. The use of the probable error tempts us constantly to overvalue the degree of accuracy we have attained.

More dangerous still is another confusion which now and then occurs, when the very expression mean error is used in the sense of the average error of the observed values according to their numerical values without regard to the signs. This gives no sense, except when we are certain of a law of typical errors, and with such a one this „mean error" is $\sqrt{\frac{2}{\pi}}\,\lambda_2$. The only reason which may be advanced in defence of the use of this idea is that we are spared some little computations, viz. some squarings and the extraction of a square root, which, however, we rarely need work out with more than three significant figures.

IX. FREE FUNCTIONS.

§ 36. The foregoing propositions concerning the laws of errors of functions — especially of linear functions — form the basis of the theory of computation with observed values, a theory which in several important things differs from exact mathematics. The result, particularly, is not an exact quantity, but always a law of errors which can be represented by its mean value and its mean error, just like the single observation. Moreover, the computation must be founded on a correct apprehension of what observations we may consider mutually unbound, another thing which is quite foreign to exact mathematics. For it is only upon the supposition that the result $R = r_1 o_1 + \ldots r_n o_n = [ro]$ — observe the abbreviated notation — is a linear function of unbound observations only, $o_1 \ldots o_n$, that we have demonstrated the rules of computation (35)

$$\lambda_1(R) = r_1 \lambda_1(o_1) + \ldots + r_n \lambda_1(o_n) = [r\lambda_1(o)] \tag{52}$$

$$\lambda_2(R) = r_1^2 \lambda_2(o_1) + \ldots + r_n^2 \lambda_2(o_n) = [r^2 \lambda_2(o)]. \tag{53}$$

While the results of computations with observed quantities, taken singly, have laws of errors in the same way as the observations, they also resemble the observations in the circumstances that there can be bonds between them, and, unfortunately, there can be bonds between "results", even though they are derived from unbound observations. If only some observations have been employed in the computation of both $R' = [r'o]$ and $R'' = [r''o]$, these results will generally be bound to each other. This, however, does not prevent us from computing a law of errors, for instance for $aR' + bR''$. We can, at any rate, represent the function of the results directly as a function of the unbound observations, $o_1, \ldots o_n$,

$$aR' + bR'' = [(ar' + br'')\,o]. \tag{54}$$

This possibility is of some importance for the treatment of those cases in which the single observations are bound. They must be treated then just like results, and we must try to represent them as functions of the circumstances which they have in common, and which must be given instead of them as original observations. This may be difficult to do, but as a principle it must be possible, and functions of bound observations must therefore always have laws of errors as well as others; only, in general, it is not possible to compute these laws of errors correctly simply by means of the laws of errors of the