From this reason, and owing to the proportionally slight difficulties attached to it, the graphical adjustment becomes particularly suitable where we are to lay down new empirical laws. In such cases we have to work through, to check, and to reject series of hypotheses as to the functional interdependency of observations and their essential circumstances. We save much labour, and illustrate our results, if we work by graphical adjustment.

Of course, we are not obliged to subject observations to adjustment. In the preliminary stages, or as long as it is doubtful whether a greater number of essential circumstances ought not to be taken into consideration, it may even be the best thing to give the observations just as they are.

But if we use the graphical form in order to illustrate such statements by the drawing of a line which connects the several observed points, then we ought to give this line the form of a continuous curve and not, according to a fashion which unfortunately is widely spread, the form of a rectilinear polygon which is broken in every observed point. Discontinuity in the curve is such a marked geometrical peculiarity that it ought, even more than cusps, double-points, and asymptotes, to be reserved for those cases in which the author expressly wants to give his opinion on its occurrence in reality.

XIV. THE THEORY OF PROBABILITY.

§ 65. We have already, in § 9, defined "*probability*" as the limit to which — the law of the large numbers taken for granted — the relative frequency of an event approaches, when the number of repetitions is increasing indefinitely; or in other words, as the limit of the ratio of the number of favourable events to the total number of trials.

The theory of probabilities treats especially of such observations whose events cannot be naturally or immediately expressed in numbers. But there is no compulsion in this limitation. When an observation can result in different numerical values, then for each of these events we may very well speak of its probability, imagining as the opposite event all the other possible ones. In this way the theory of probabilities has served as the constant foundation of the theory of observation as a whole.

But, on the other hand, it is important to notice that the determination of the law of errors by symmetrical functions may also be employed in the non-numerical cases without the intervention of the notion of probability. For as we can always indicate the mutually complementary opposite events as the "fortunate" or "unfortunate" one, or as "Yes" and "No", we may also use the numbers 0 and 1 as such a formal indication. If then we identify 1 with the favourable "Yes"-event, 0 with the unfavourable "No", the sums of the numbers got in a series of repetitions will give the frequency of affirmative events. This relation, which has been used already in some of the foregoing examples, we must here consider more explicitly.

If repetitions of the same observation, which admits of only two alternatives, give the result "Yes" = 1 m times, against n times "No" = 0, then the relative frequency for the favourable event is $\frac{m}{m+n}$. But if we employ the form of the symmetrical functions for the same law of actual errors, then the sums of the powers are

$$s_0 = m+n, \quad s_1 = s_2 \ldots\ldots = s_r = m. \tag{121}$$

In order to determine the half-invariants by means of this, we solve the equations

$$\begin{aligned}
m &= (m+n)\mu_1 \\
m &= \quad m\cdot\mu_1 + (m+n)\mu_2 \\
m &= \quad m\cdot\mu_1 + \quad 2m\cdot\mu_2 + (m+n)\mu_3 \\
m &= \quad m\cdot\mu_1 + \quad 3m\cdot\mu_2 + \quad 3m\cdot\mu_3 + (m+n)\mu_4,
\end{aligned}$$

and find then

$$\left.\begin{aligned}
\mu_1 &= \frac{m}{m+n} \\
\mu_2 &= \frac{mn}{(m+n)^2} \\
\mu_3 &= \frac{mn(n-m)}{(m+n)^3} \\
\mu_4 &= \frac{mn(n^2-4mn+m^2)}{(m+n)^4}.
\end{aligned}\right\} \tag{122}$$

Compare § 23, example 2, and § 24, example 3.

All the half-invariants are integral functions of the relative frequency, which is itself equal to μ_1. The relative frequency of the opposite result is $\frac{n}{m+n} = 1-\mu_1$; by interchanging m and n, none of the half-invariants of even degree are changed, and those of odd degree (from μ_3 upwards) only change their signs.

In order to represent the connection between the laws of presumptive errors, we need only assume, in (122), that m and n increase indefinitely, while the probability of the event becomes $p = \frac{m}{m+n}$, and the probability of the opposite event is represented by $\frac{n}{m+n} = 1-p = q$. The half invariants are then:

$$\left.\begin{aligned}
\lambda_1 &= p \\
\lambda_2 &= pq \\
\lambda_3 &= pq(q-p) \\
\lambda_4 &= pq(q^2-4pq+p^2).
\end{aligned}\right\} \tag{123}$$

Our mean values are therefore, respectively, the relative frequency and the probability itself.

We must now first notice here that every half-invariant is its own fixed and simple function of the probability (the frequency). When a result of observation can be stated in the form of one single probability, properly so called, we have thereby given as complete a determination of the law of errors as by the whole series of half-invariants. In such cases it is simpler to employ the theory of probability instead of the symmetrical functions and the method of the least squares.

The theory of probability thereby gets its province determined in a much more natural and suitable way than that employed in the beginning of this paragraph.

But at the same time we see that the form of the half-invariants is not only the general means which must be employed where the conditions for the use of the probability are not fulfilled, but also that, within the theory of probability itself, we shall require, particularly, the notion of the mean error.

Even where the probability can replace all the half-invariants, we shall require all the various sides of the notions which are distinctly expressed in the half-invariants. Now we have particularly to consider the probability as the definite mean value, now the point is to elicit the definite degree of uncertainty which is implied in the probability, and which is particularly emphasised in the mean error. Otherwise, we should constantly be tempted to rely on the predictions of the theory of probability to an extent far beyond what is justly due to them. Finally, we shall see immediately that the laws of error of the probabilities are far from typical, but that they have rather a type of their own, which must sometimes be especially emphasised.

All this we shall be able to do here, where we have the half-invariants in reserve as a means of representing the theory of probability.

§ 66. In particular, we can now, though only in the form of the half-invariants, solve one of the principal problems of the theory of probability, and determine the law of presumptive errors for the frequency m of one of the events of a trial, which can have only two events and which is repeated N times, upon the supposition that the trial follows the law of the large numbers, and that the probability p for a single trial is known.

The equations (123) give us already the corresponding law of error for each trial, and as the total absolute frequency is the sum of the partial ones, we need only use the equations (35) to find:

$$\left.\begin{aligned} \lambda_1(m) &= Np \\ \lambda_2(m) &= Npq = Np(1-p) \\ \lambda_3(m) &= Npq(q-p) = Np(1-p)(1-2p) \\ \lambda_4(m) &= Npq(q^2-4pq+p^2) \\ &= Np(1-p)(1-(3+\sqrt{3})p)(1-(3-\sqrt{3})p). \end{aligned}\right\} \quad (124)$$

The ratio of the mean frequency to the number of trials is therefore the probability itself. When p is small the mean error differs little from the square root $\sqrt{Np}$ of the mean frequency; and if p is nearly $= 1$, the mean error of the opposite event is nearly equal to $\sqrt{Nq}$. When the probability, p, is nearly equal to $\frac{1}{2}$, the mean error will be about $\frac{1}{2}\sqrt{N}$.

The law of error is not strictly typical, although the rational function of the r^{th} degree in $\lambda_r(m)$ vanishes for r different values of p between 0 and 1, the limits included, so that the deviation from the typical form must, on the whole, be small. If, however, we consider the relative magnitude of the higher half-invariants as compared with the powers of the mean error

$$\left.\begin{aligned} &\lambda_3(m)\cdot(\lambda_2(m))^{-\frac{3}{2}} = \frac{q-p}{\sqrt{Npq}} \\ \text{and}\qquad &\lambda_4(m)\cdot(\lambda_2(m))^{-2} = \frac{q^2-4pq+p^2}{Npq} \end{aligned}\right\} \tag{125}$$

the occurence of Npq in the denominators of the abridged fractions shows, not only that great numbers of repetitions, here as always, cause an approximation to the typical form, but also that, in contrast to this, the law of error in the cases of certainty and impossibility, when $q = 0$ and $p = 0$, becomes skew and deviates from the typical in an infinitely high degree, while at the same time the square of the mean errors becomes $= 0$. This remarkable property is still traceable in the cases in which the probability is either very small or very nearly equal to 1. In a hundred trials with the probability $= 99\frac{1}{2}$ per ct. the mean error will be about $= \sqrt{\frac{1}{2}}$. Errors beyond the mean frequency $99\frac{1}{2}$ cannot exceed $\frac{1}{2}$, and are therefore less than the mean error. The great diminishing errors must therefore be more frequent than in typical cases, and frequencies of 97 or 96 will not be rare in the case under consideration, though hey must be fully counter-balanced by numerous cases of 100 per ct. The law of error is consequently skew in a perceptible degree. In applications of adjustment to problems of probability, it is, from this reason, frequently necessary to reject extreme probabilities.

XV. THE FORMAL THEORY OF PROBABILITY.

§ 67. The formal theory of probability teaches us how to determine probabilities that depend upon other probabilities, which are supposed to be given. Of course, there are no mathematical rules specially applicable to computations that deal with probabilities, and there are many computations with probabilities which do not fall under the theory of probability, for instance, adjustments of probabilities. But in view of the direct application