the sun itself to volatilize carbon—why, even if the small comets said, in the Philosophical Transactions, to be throwing off the incandescent vapour of carbon every night they were under observation, even in a dark and cold sky, had been taken thence and placed on the very surface of the sun itself, and had experienced there not only the heat which that other comet had experienced of earth's × 47,000, but earth's × 300,000, they could not have shown a pure carbon-spectrum.

As our sun, according to Father Secchi, ranks only among the yellow stars, and *they* are supposed not to be so hot as the white stars, perhaps the vapour of carbon may exist glowing and incandescent in Sirius, which is so noted a member of the latter class of stars. We may, too, perhaps be privileged to see the actual and real spectral lines of carbon there, in any good telespectroscope—but with the drawback that, however plainly the lines may appear in themselves, we cannot recognize their chemical origin and assign them their true name, because neither has man ever yet volatilized pure carbon, nor has any angel (in default of theory) ever told us the wave-lengths of carbon-lines when the carbon has been volatilized by a higher power.

Hydrocarbon compound it *is* given to man to volatilize and spectroscope; and he should be thankful for its many admirable uses; but as to the spectrum of the pure carbon element being seen in the base of the flame of every little candle made and set alight by human hands, it would be well if certain modern men, and the secret committee of the Royal Society in particular, were to come forward openly and confess with deep contrition in the words of ancient Job,

"I have uttered that I understood not; things too wonderful for me, which I knew not."

"Wherefore I abhor myself, and repent in dust and ashes."

IV. *Statistics by Intercomparison, with Remarks on the Law of Frequency of Error*. *By* Francis Galton, *F.R.S.**

MY object is to describe a method for obtaining simple statistical results which has the merit of being applicable to a multitude of objects lying outside the present limits of statistical inquiry, and which, I believe, may prove of service in various branches of anthropological research. It has already been proposed (Lecture, Royal Institution, Friday evening, February 27, 1874), and in some degree acted upon ('Hereditary Genius,' p. 26), by myself. What I have now to offer is a more complete explanation and a considerable development of previous views.

* Communicated by the Author.

The process of obtaining mean values &c. now consists in measuring each individual with a standard that bears a scale of equal divisions, and afterwards in performing certain arithmetical operations upon the mass of figures derived from these numerous measurements. I wish to point out that, in order to procure a specimen having, in one sense, the mean value of the quality we are investigating, we do not require any one of the appliances just mentioned: that is, we do not require (1) independent measurements, nor (2) arithmetical operations; we are (3) able to dispense with standards of reference, in the common acceptation of the phrase, being able to create and afterwards indirectly to define them; and (4) it will be explained how a rough division of our standard into a scale of degrees may not unfrequently be effected. Therefore it is theoretically possible, in a great degree, to replace the ordinary process of obtaining statistics by another, much simpler in conception, more convenient in certain cases, and of incomparably wider applicability.

Nothing more is required for the due performance of this process than to be able to say which of two objects, placed side by side, or known by description, has the larger share of the quality we are dealing with. Whenever we possess this power of discrimination, it is clear that we can marshal a group of objects in the order in which they severally possess that quality. For example, if we are inquiring into the statistics of height, we can marshal a number of men in the order of their several heights. This I suppose to be effected wholly by *intercomparison,* without the aid of any external standard. The object then found to occupy the middle position of the series must possess the quality in such a degree that the number of objects in the series that have more of it is equal to that of those that have less of it. In other words, it represents the *mean* value of the series in at least one of the many senses in which that term may be used. Recurring to the previous illustration, in order to learn the mean height of the men, we have only to select the middlemost one and measure him; or if no standard of feet and inches is obtainable, we must describe his height with reference to numerous familiar objects, so as to preserve for ourselves and to convey to strangers as just an idea of it as we can. Similarly the mean speed of a number of horses would be that of the horse which was middlemost in the running.

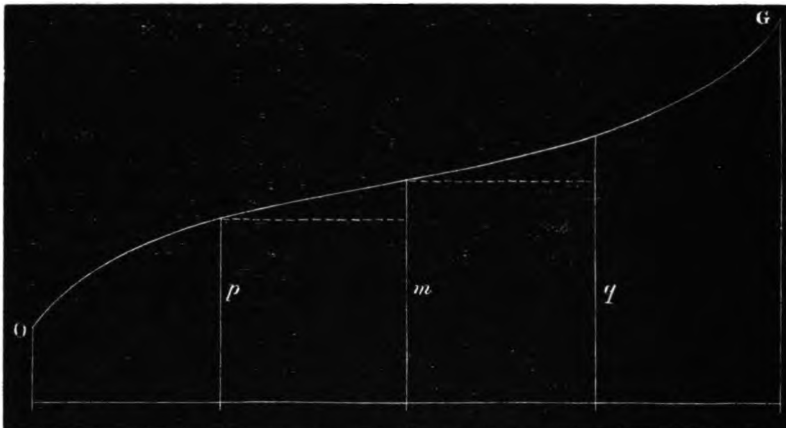
If we proceed a step further and desire to compare the mean height of two populations, we have simply to compare the representative man contributed by each of them. Similarly, if we wish to compare the performances of boys in corresponding classes of different schools, we need only compare together the middle boys in each of those classes.

The next great point to be determined is the divergency of the series—that is, the tendency of individual objects in it to diverge from the mean value of all of them. The most convenient measure of divergency is to take the object that has the mean value, on the one hand, and those objects, on the other, whose divergence in either direction is such that one half of the objects in the series on the same side of the mean diverge more than it does, and the other half less. The difference between the mean and either of these objects is the measure in question, technically and rather absurdly called the "probable error." Statisticians find this by an arithmetical treatment of their numerous measurements; I propose simply to take the objects that occupy respectively the first and third quarter points of the series. I prefer, on principle, to reckon the divergencies in excess separately from those in deficiency. They cannot be the same unless the series is symmetrical, which experience shows me to be very rarely the case. It will be observed that my process fails in giving the difference (probable error) in numerical terms; what it does is to select specimens whose differences are precisely those we seek, and which we must appreciate as we best can.

We have seen how the mean heights &c. of two populations may be compared; in exactly the same way may we compare the divergencies in two populations whose mean height is the same, by collating representative men taken respectively from the first and third quarter points of the series in each case.

We may be confident that if any group be selected with the ordinary precautions well known to statisticians, it will be so far what may be called "generic" that the individual differences of members of that group will be due to various combinations of pretty much the same set of variable influences. Consequently, by the well-known laws of combinations, medium values will occur very much more frequently than extreme ones, the rarity of the latter rapidly increasing as the deviation slowly increases. Therefore, when the objects are marshalled in the order of their magnitude along a level base at equal distances apart, a line drawn freely through the tops of the ordinates which represent their several magnitudes will form a curve of double curvature. It will be nearly horizontal over a long space in the middle, if the objects are very numerous; it will bend down at one end until it is nearly vertical, and it will rise up at the other end until there also it is nearly vertical. Such a curve is called, in the phraseology of architects, an "ogive," and is represented by O G in the diagram (fig. 1), in which the process of statistics by intercomparison is clearly shown. If $n =$ the length of the base of the ogive, whose ordinate y represents the magnitude of the object that stands at a distance x from that end of the base where the ordinates are smallest, then the number of

Fig. 1.



$\frac{1}{4}$ $\frac{1}{2}$ $\frac{3}{4}$

objects less than y : the number of objects greater than $y :: x : n-x$. The ordinate m at $\frac{1}{2}$ represents the mean value of the series, and p, q at $\frac{1}{4}$ and $\frac{3}{4}$, taken in connexion with m, give data for estimating the divergence; thus $q-m$ is the divergence (probable error) of at least that portion of the series that is in excess of the mean, and $m-p$ is that of at least the other portion. When the series is symmetrical, $q-m=p-q$, and either, or the mean of both, may be taken as the divergence of the series generally. No doubt we are liable to deal with cases in which there may be some interruption in the steady sweep of the ogive; but the experience of qualities which we *can* measure, assures us that we need fear no large irregularity of that kind when dealing with those which, as yet, we have no certain means of measuring.

When we marshal a series, we may arrange them roughly, except in the neighbourhood of the critical points; and thus much labour will be saved. But the most practical way of setting to work would probably depend not on the mere discrimination of greater and less, but also on a rough sense of what is much greater or much less. We have called the objects at the $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ distances p, m, and q respectively; let us sort the objects into two equal portions P and Q, of small and great, taking no more pains about the sorting than will ensure that P contains p and all smaller than p, and that Q contains q and all larger than q. Next, beginning, say, with group P, sort away alternately to right and left the larger and the smaller objects, roughly at first, but proceeding with more care as the residuum diminishes and the differences become less obvious. The last remaining object will be *p*. Similarly we find *q*. Then *m* will be found in the same way from the group compounded of those that were sorted to the right from P and to the left from Q.

There are not a few cases where both the ordinary method and that by intercomparison are equally applicable, but in which the latter would prove the more rapid and convenient. I would mention one of some importance to those anthropologists who may hereafter collect data in uncivilized countries. A barbarian chief might often be induced to marshal his men in the order of their heights, or in that of the popular estimate of their skill in any capacity; but it would require some apparatus and a great deal of time to measure each man separately, even supposing it possible to overcome the usually strong repugnance of uncivilized people to any such proceeding.

The practice of sorting objects into classes may be said to be coextensive with commerce, the industries, and the arts. It is adopted in the numerous examinations, whether pass or competitive, some or other of which all youths have now to undergo. It is adopted with every thing that has a money-value; and all acts of morality and of intellectual effort have to submit to a verdict of "good," "indifferent," or "bad."

The specimen values obtained by the process I have described are capable of being reproduced so long as the statistical conditions remain unchanged. They are also capable of being described in various ways, and therefore of forming permanent standards of reference. Their importance then becomes of the same kind as that which the melting-points of well-defined alloys or those of iron and of other metals had to chemists when no reliable thermometer existed for high temperatures. These were excellent for reference, though their relations *inter se* were subject to doubt. But we need never remain wholly in the dark as to the relative value of our specimens, methods appropriate to each case being sure to exist by which we may gain enlightenment. The measurement of work done by any faculty when trained and exerted to its uttermost, would be frequently available as a test of its absolute efficacy.

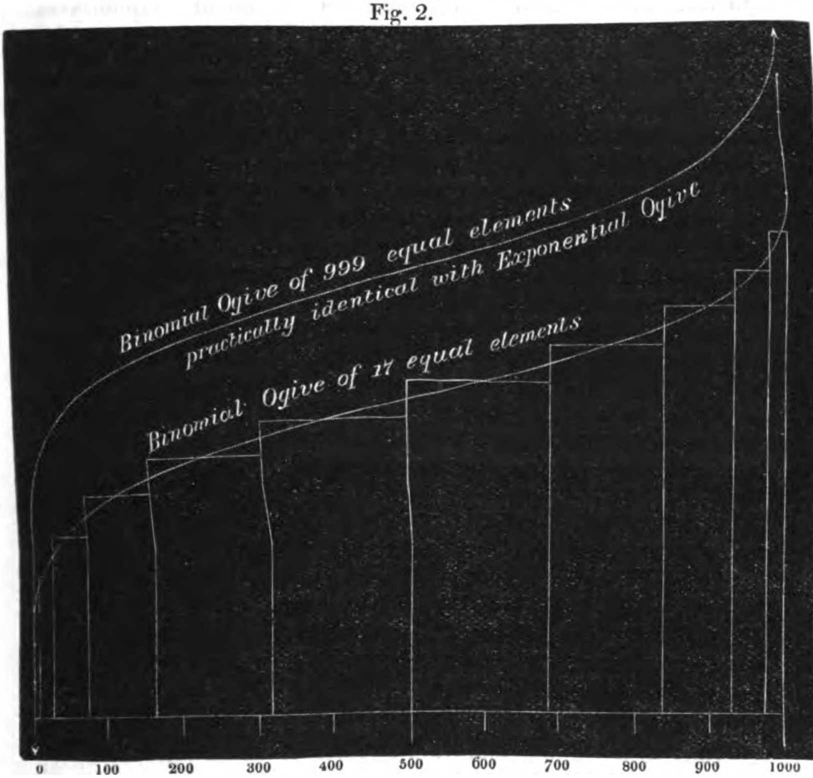
There is another method, which I have already advocated and adopted, for gaining an insight into the absolute efficacies of qualities, on which there remains more to say. Whenever we have grounds for believing the law of frequency of error to apply, we may *work backwards*, and, from the relative frequency of occurrence of various magnitudes, derive a knowledge of the true relative values of those magnitudes, expressed in units of probable error. The law of frequency of error says that "magnitudes differing from the mean value by such and such multiples of the probable error, will occur with such and such degrees of frequency." My proposal is to reverse the process, and to say, "since such and such magnitudes occur with such and such degrees of frequency, therefore the differences between them and the mean value are so and so, as expressed in units of probable error." According to this process, the positions of the first divisions of the scale of divergence, which are those of the mean value *plus* or *minus* one unit of probable error, are of course p and q, lying at the $\frac{1}{4}$ and $\frac{3}{4}$ points of the ogive, or, if the base consist of 1000 units, at the 250th point from the appropriate end. The second divisions being those of mean value *plus* or *minus* two units of probable error, will, according to the usual Tables, be found at the 82nd point from the appropriate end, the third divisions will be at the 17th, and the fourth at the 3rd. If we wished to pursue the scale further, we should require a base long enough to include very many more than 1000 units.

Remarks on the Law of Frequency of Error.

Considering the importance of the results which admit of being derived whenever the law of frequency of error can be shown to apply, I will give some reasons why its applicability is more general than might have been expected from the highly artificial hypotheses upon which the law is based. It will be remembered that these are to the effect that individual errors of observation, or individual differences in objects belonging to the same generic group, are entirely due to the aggregate action of variable influences in different combinations, and that these influences must be (1) all independent in their effects, (2) all equal, (3) all admitting of being treated as simple alternatives "above average" or "below average;" and (4) the usual Tables are calculated on the further supposition that the variable influences are infinitely numerous.

As I shall lay much stress on matters connected with the last condition, it will save reiteration if I be permitted the use of a phrase to distinguish between calculations based on the supposition of a moderate number (r) of elements (in which case the frequency of error or the divergence is expressed by the coefficients of the expansion of the binomial $(a+b)^r$) and one based on the supposition of the number being infinite (which is expressed by the exponential $e^{-\frac{x^2}{c^2}}$), by calling the one the binomial and the other the exponential process, the latter being the process to be understood whenever the "law of frequency of error" is spoken of without further qualification. When the results of these two processes have to be protracted, as in figure 2, the unit of vertical measurement in the case of a series of bino-

Fig. 2.



mial grades will be a single grade, or, what comes to the same thing, the difference of the effect produced by the plus and minus phase of any one of the alternative elements, upon the value of the whole. The unit of the exponential curve will be $q-m$ of fig. 1, or the probable error. This latter unit is equally applicable to what we may call the binomial ogive, which is the curve drawn with a free hand through the grades. The justification for such a conception as a binomial ogive will be fully established further on. Suffice it for the present to remark that, by the adoption of a unit of this kind, the middle portion of a binomial ogive of 999 elements is compared in the figure with one of 17.

The first three of the above-mentioned conditions may occur in games of chance, but they assuredly do not occur in vital and social phenomena; nevertheless it has been found in numerous instances, where measurement was possible, that the latter conform very fairly, within the limits of ordinary statistical inquiry, to calculations based on the (exponential) law of frequency of error. It is a curious fact, which I shall endeavour to explain, that in this case a false hypothesis, which is undoubtedly a very convenient one to work upon, yields true results.

In illustration of what occurs in nature, let us consider the causes which determine the size of fruit. Some are important, the chief of which is the Aspect, whose range of influence is too wide to permit us to consider it in one of the simple alternatives "good" or "bad." It is no satisfactory argument to say that variations in aspect are *partly* due to a multitude of petty causes, such as the interposition of leaves and boughs, because, so far as they depend on well-known functions of altitude and azimuth, they cannot be reduced to a multitude of elementary causes. There has been much confusion of ideas on this subject, and also a forgetfulness of another fact—namely, that when we once arrive at a simple alternative, there our subdivision of causes must stop, and we must accept that alternative, however great may be its influence, as one of the primary elements in our calculation.

In addition to important elements, there are others of small, but still of a recognizable value, such as exposure to prevalent winds, the pedigree of the tree, the particular quality of the soil on which it stands, the accident of drains running near to its root, &c. Again, there are a multitude of smaller influences, to the second, third, and fourth orders of minuteness.

I shall proceed to define what I mean by "small;" then I shall show how this medley of causes may admit of being theoretically sorted into a moderate number of small influences of equal value, giving a first approximation to the truth; then how, by a second approximation, the grades of the binomial expansion thence derived become smoothed into a flowing curve. Lastly, I shall show by quite a different line of argument that the exponential view contains inherent contradictions when nature is appealed to, that the binomial of a moderate power is the truer one, and that we have means of ascertaining a limit which the number of its elements cannot exceed. My conclusion, so far as this source of difficulty is concerned, is that the exponential law applies because it nearly resembles the curve based on a binomial of moderate power, within the limits between which comparisons are usually made.

We observe in fig. 2 how closely the outline of an exponential ogive resembles that of a binomial of a very moderate number of elements, within the narrow limits chiefly used by statisticians. The figure expresses a series of 1000 objects marshalled according to their magnitudes. In the one case the magnitudes are supposed to be wholly due to the various combinations of 17 alternatives, and the elements of the drawing are obtained from the several terms of the expansion of $(1+1)^{17}$, all multiplied into $\frac{1000}{2^{17}}$. These form the following series, reckoning to the nearest integer; and their sum, of course, $=1000$:—0, 0, 1, 5, 18, 47, 95, 148, 186, 186, 148, 95, 47, 18, 5, 1, 0, 0. In the figure these proportions are protracted so far as possible; but the numbers even in the fourth grade are barely capable of being represented on its small scale; after the fourth, the several grades are manifest until we reach the corresponding point at the opposite end of the series. Then, with a free hand, a curve is drawn through them, which gives as their mean value 8·5, as it ought to do. Now, referring to our p and q at the 250th division from either end, I measure the value of $q-m$ (or $m-p$), which is the unit to which I must reduce any other ogive that I may desire to compare with the present one. Also I can find the values for $m+2(q-m)$ and $m+3(q-m)$, which is going as far as a figure on this small scale admits. I now protract the central portion of an exponential ogive to the same scale, horizontally and vertically. Not knowing its base, I start from its middle point, placing it arbitrarily at a convenient position in the prolongation of the m of the binomial; and I lay off, in the prolongation of p and q, points that are respectively 1 unit of probable error less and greater than m. The Tables of the law of error tell me where to lay off the other points; and so the curve is determined. It must be clearly understood that whereas in the figure both the ogive and the base are given for the binomial series of 17 elements, it is only the ogive that is given for the exponential, there being no data to determine the position of its base. The comparison is simply between the middle portions of the ogives. To speak correctly, I have not actually used the exponential Tables to draw the exponential curve, but have used Quetelet's expansion of a binomial of 999 elements, the results of which are identical, as he has shown, with those of the exponential to within extremely minute fractions, utterly insensible in a scale more than a hundred times as great as the present one.

I find the position of the various points in the two ogives, measured from the appropriate end of the base, to be as is expressed in the following Table:—

	In binomial ogive of 17 elements.	In exponential ogive, or in binomial ogive of 999 elements.
The mean	500	500
The mean ±1 unit probable error	250	250
„ ±2 units	71	82
„ ±3 units	16	17

The closeness of the resemblance is striking. It rapidly increases and extends in its range as the number of elements in the binomial increases; there need therefore be no hesitation in recognizing the fact that a binomial of, say, 30 elements or upwards is just as conformable to ordinary statistical observation as is the exponential. If one agrees, the other does, because they agree with one another.

The fewest number of elements that suffice to form a binomial having the above-mentioned conformity is a criterion of the meaning of the word "small," which was lately employed, because each of those elements would be just entitled to rank as small.

I obtain the value of any one of them in an ogive by protracting the series and noticing how many grades are included in the interval $q-m$. It will be found that in a binomial of 17 elements $q-m$ is equal to eight fifths of one grade. Thence I conclude that in any generic series an influence the range of whose mean effects in the two alternatives of above and below average is not greater than, say, one half of the probable error of the series, is entitled to be considered "small."

I now proceed to show how a medley of small and minute causes may, as a first approximation to the truth, be looked upon as an aggregate of a moderate number of "small" and *equal* influences. In doing this, we may accept without hesitation, the usual assumption that all small, and *à fortiori* all minute influences, may be dealt with as simple alternatives of excess or deficiency—the values of this excess and deficiency being the mean of all the values in each of these two phases. The way in which I propose to build up the fictitious groups may be exactly illustrated by a game of odd and even, in which it might be agreed that the *predominance* of "heads" in a throw of three fourpenny pieces, shall count the same as the simple "head" of a shilling. The three fourpenny pieces may fall all heads, 2 heads and 1 tail, 1 head and 2 tails, or all tails—the relative frequency of these events being, as is well known, 1, 3, 3, 1. But by our hypothesis we need not concern ourselves about these minute peculiarities; the question for us is simply the alternative one, are the "heads" in a majority or not? We may therefore treat a ternary system of the third order of smallness exactly as a simple alternative of the first order of smallness. Or, again, suppose a crown were our "small" unit, and we had a medley of 10 crowns, 33 shillings, and 100 fourpenny pieces, with which to make successive throws, throwing the whole number of them at once: we might theoretically sort them into fictitious groups each equivalent to a crown. There would be 29 such groups, viz.:—10 groups, each consisting of 1 crown; 6 groups, each of 5 shillings; 1 group of three shillings and 6 fourpenny pieces; 6 groups each of 15 fourpenny pieces; and a residue of 4 fourpenny pieces, which may be disregarded. Hence, on the already expressed understanding that we do not care to trouble ourselves about smaller sums than a crown, the results of the successive throws of the medley of coins would be approximately the same as those of throwing at a time 29 crowns, and would be expressed by the coefficients of a binomial of the 29th power. Hence I conclude that all miscellaneous influences of a few small and many minute kinds, may be treated for a first approximation exactly as if they consisted of a moderate number of small and equal alternatives.

The second approximation has already been alluded to; it consists in taking some account of the minute influences which we had previously agreed to ignore entirely, the effect of whieh is to turn the binomial grades into a binomial ogive. I effect it by drawing a curve with a free hand through the grades, which affords a better approximation to the truth than any other that can *à priori* be suggested.

I will now show from quite another point of view (1) that the exponential ogive is, on the face of it, fallacious in a vast number of cases, and (2) that we may learn what is the greatest possible number of elements in the binomial whose ogive most nearly represents the generic series we may be considering. The value of $\frac{m}{q-m}$ is directly dependent on the number of elements; hence, by knowing its value, we ought to be able to determine the number of its elements. I have calculated it for binomials of various powers, protracting and interpolating, and obtain the following very rough but sufficient results for their ogives (not *grades*):—

Number of (equal) elements.	Value of $\frac{m}{q-m}$.
17	5
32	10
65	15
107	20
145	25
186	30
999	48

Now, if we apply these results to observed facts, we shall rarely find that the series has been due to any large number of equal elements. Thus, in the stature of man the probable error, $\frac{m}{q-m}$, is about 30, which makes it impossible that it can be looked upon as due to the effect of more than 200 equally small elements. On consideration, however, it will appear that in certain cases the number may be *less*, even considerably less, than the tabular value, though it can never exceed it. As an illustration of the principle upon which this conclusion depends, we may consider what the value of $\frac{m}{q-m}$ would be in the case of a wall built of 17 courses of stone, each stone being 3 inches thick, and subject to a mean error in excess or deficiency of one fifth of an inch. Obviously the mean height m of the wall would be 3×17 inches; and its probable error $q-m$ would be very small, being derived from a binomial ogive of 17 elements, each of the value of only one fifth of an inch. Now we saw from our previous calculation that this would be eight fifths, or 1·6 inch, which would give the value to $\frac{m}{q-m}$ of $\frac{51}{1\cdot6}$, or about 321; consequently we should be greatly misled if, after finding by observation the value of that fraction, and turning to the Table and seeing there that it corresponded to more than 200 equal elements, we should conclude that that was the number of courses of stones. The Table can only be trusted to say that the number of courses certainly does not exceed that number; but it may be less than that.

The difficulty we have next to consider is that which I first mentioned, but have intentionally postponed. It is due to the presence of influences of extraordinary magnitude, as Aspect in the size of fruit. These influences must be divided into more than two phases, each differing by the same constant amount from the next one, and that difference must not be greater than exists between the opposite phases of the "small" alternatives. If we had to divide an influence into three phases, we should call them "large," "moderate," and "small;" if into four, they would be "very large," "moderately large," "moderately small," and "very small," and so on. Any objects (say, fruit) which are liable to an influence so large as to make it necessary to divide it into three phases, really consist of three series generically different which are entangled together, and ought theoretically to be separated. If there had been two influences of three phases, there would be nine such series, and so on. In short, the fruit, of which we may be considering some hundred or a few thousand specimens, ought to be looked upon as a multitude of different sorts mixed together. The proportions *inter se* of the different sorts may be accepted as constant; there is no difficulty arising from that cause. The question is, why a mixture of series radically different, should in numerous cases give results apparently identical with those of a simple series.

For simplicity's sake, let us begin with considering only one large influence, such as aspect on the size of fruit. Its extreme effect on their growth is shown by the difference in what is grown on the north and south sides of a garden-wall, which in such kinds of fruit as are produced by orchard-trees, is hardly deserving of being divided into more than three phases, "large," "moderate," and "small." Now if it so happens that the "moderate" phase occurs approximately *twice as often* as either of the extreme phases (which is an exceedingly reasonable supposition, taking into account the combined effects of azimuth, altitude, and the minor influences relating to shade from leaves &c.), then the effect of aspect will work in with the rest, just like a binomial of two elements. Generally the coefficients of $(a+b)^n$ are the same as those of $(a+b)^{n-r} \times (a+b)^r$. Now the latter factor may be replaced by any variable function the frequency and number of whose successive phases, into which it is necessary to divide it, happen to correspond with the value of the coefficients of that factor.

It will be understood from what went before, that we are in a position to bring these phases to a common measure with the rest, by the process of fictitious grouping with appropriate doses of minute influences, as already described.

On considering the influences on which such vital phenomena depend as are liable to be treated together statistically, we shall find that their mean values very commonly occur with greater frequency than their extreme ones; and it is to this cause that I ascribe the fact of large influences frequently working in together with a number of small ones without betraying their presence by any sensible disturbance of the series.

The last difficulty I shall consider, arises from the fact that the individuals which compose a statistical group are rarely affected by exactly the same number of variable influences. For this cause they ought to have been sorted into separate series. But when, as is usually the case, the various intruding series are weak in numbers, and when the number of variable influences on which they depend does not differ much from that of the main series, their effect is almost insensible. I have tried how the figures would run in many supposititious cases; here is one taken at haphazard, in which I compare an ordinary series due to 10 alternatives, giving $2^{10}=1024$ events, with a compound series.

The latter also comprises 1024 events; but it is made up of three parts: viz. nine tenths of it are due to a 10-element series; and of the remaining tenth, half are due to a 9 and half to an 11 series. I have reduced all these to the proper ratios, ignoring fractions. It will be observed how close is the correspondence between the compound and the simple series.

Total cases.	Number of elements.	Successive grades in the series.											
		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
52	9	0	1	4	8	13	13	8	4	1	0	0	0
924	10	1	9	41	108	189	227	189	108	41	9	1	0
48	11	0	0	1	4	8	11	11	8	4	1	0	0
1024	Compound series.	1	10	46	120	210	251	208	120	46	10	1	0
1024	10	1	10	45	120	210	252	210	120	45	10	1	0
	Difference	0	0	+1	0	0	−1	−2	0	+1	0	0	0

It appears to me, from the consideration of many series, that the want of symmetry commonly observed in the statistics of vital phenomena is mainly due to the inclusion of small series of the above character, formed by alien elements; also that the disproportionate number of extreme cases, as of giants, is due to this cause.

The general conclusion we are justified in drawing appears to be, that, while each statistical series must be judged according to its peculiarities, a law of frequency of error founded on a binomial ogive is much more likely to be approximately true of it than any other that can be specified *à priori*; also that the exponential law is so closely alike in its results to those derived from the binomial ogive, under the circumstances and within the limits between which statisticians are concerned, that it may safely be used as hitherto, its many well-known properties being very convenient in all cases where it is approximately true. Therefore, if we adopt any uniform system (such as already suggested) of denoting the magnitudes of qualities for the measurement of which no scale of equal parts exists, such system may reasonably be based on an *inverse* application of the law of frequency of error, in the way I have described, to statistical series obtained by the process of *intercomparison.*