

WILEY

On the Nature and Uses of Averages

Author(s): John Venn

Source: *Journal of the Royal Statistical Society*, Vol. 54, No. 3 (Sep., 1891), pp. 429-456

Published by: Wiley for the Royal Statistical Society

Stable URL: <http://www.jstor.org/stable/2979569>

Accessed: 27-06-2016 08:16 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Royal Statistical Society, Wiley are collaborating with JSTOR to digitize, preserve and extend access to *Journal of the Royal Statistical Society*

On the NATURE and USES of AVERAGES.

By DR. JOHN VENN, D.Sc., F.R.S.

[Read before the Royal Statistical Society, 26th May, 1891.
The President, DR. F. J. MOUAT, M.D., F.R.C.S., in the Chair.]

OUR President has kindly asked me to make some remarks this evening upon the "Nature and Uses of Averages." I feel that I must choose the ground warily, and speak with some caution upon this subject and in this place. On the one hand I am addressing a gathering most of whom are far better acquainted than myself with the sort of materials to which averages are commonly applied, so that the cautions which might be appropriate to a lay assembly would become an impertinence here. On the other hand, what may be called the higher theory of the subject—the full justification, for example, of the doctrine and method of least squares—is a subject which the profoundest mathematicians of the last hundred years have made so completely their own, that it would be an impertinence for any one who is not a first rate mathematician to meddle with the subject.

There is, however, an intermediate ground which, though not extensive, is apt to be rather neglected. I refer here to what may be called the *logical* account of the nature and uses of averages; the answers, that is, to such questions as these: Why do we resort to averages at all? What do we gain and lose respectively by doing so? What different kinds of average are there, and how and why does one such kind become more appropriate than another? Such questions as these, which might be multiplied further, are apt to be overlooked by the practical man, owing partly to his long and early familiarity with one particular kind of average; and though they are by no means left unregarded by the mathematician, the treatment they receive at his hands needs a good deal of simplification and selection in order to make it readily intelligible.¹

What, then, is an average, in the most general sense of the term? What it seems to presuppose is always this: that we have a plurality of values or magnitudes of some kind set before us which are, regarded in themselves, actual concrete matters of fact. We put these aside and substitute for them an intermediate value

¹ The mathematical justification of almost every kind of average which can be proposed, will be found discussed in Mr. F. Y. Edgeworth's paper in the "Cambridge Philosophical Transactions" (vol. xiv, part ii).

of our own construction, which is, by contrast with the former, fictitious and artificial. This is the most general account we can give of the matter, and will presumably cover every kind of average which can be taken into account. But it at once raises the question, How can a single introduction of our own, and that a fictitious one, possibly take the place of the many values which were actually given to us? And the answer surely is, that it can *not* possibly do so; the one thing cannot take the place of the other for purposes in general, but only for this or that specific purpose. And, what is more, the object proposed in such substitution must be clearly defined, and, in consequence, the suitable kind of average must be selected for the purpose in view. This seems so obvious when stated in words, that nothing but extreme familiarity with the subject-matter, and the satisfaction which comes of long and successful practice within definite limits, could have brought us to overlook it. But it will be useful to work out the contrast in detail by considering successively the principal purposes which we may have in view in resorting to one of these averages. These are prominently three:—

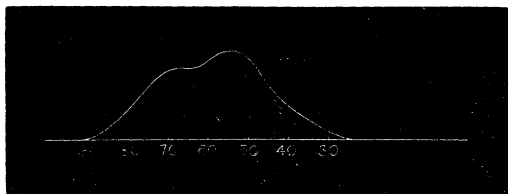
I. The first of these purposes, I apprehend, is merely that of general non-quantitative comparison. We have a number of groups of things before us, and we want to compare them one with another; to put the groups in their order, for instance, at any given time, or to say, when comparing them at different times, which of them are increasing and which are diminishing.

For this, which may be called the merely *comparative* purpose, I apprehend that almost any kind of average will answer our end equally well. To take two that are most familiar to the statistician—the *mean*, or arithmetical average, and the *median*, or middlemost value—the occasional employment of the latter is sometimes defended on the ground of its general close agreement with the former. But such agreement does not seem necessary. For purposes of general comparison any kind of intermediate value would answer the purpose, provided this be deduced according to the same rule in each case. Begin with a simple physical example or analogy. Suppose a lot of pictures of different sizes hanging on a wall, and we were asked to decide as to their relative height, we should require some common point of comparison. We might take the centre of each as a simple test, but we might equally take a point, say, nine-tenths of the way above or below the centre, if we took the same relative point with all. Any such question as what was the order, in respect of their height, of all the pictures, or whether any particular picture had been raised or lowered within a given time, would receive exactly the same answer, whichever such test we adopted.

Reverting then from this example to a more serious application. When we are comparing the stature of different nations, or of different classes in the same nation, or are inquiring whether that stature is increasing or diminishing as time advances, it does not seem to me to be at all necessary that the median and the mean should coincide, for us to feel at ease in resorting to the former. For these merely comparative purposes the two might differ considerably, and the answer to our questions would not be in the least affected.

Consider, for instance, the results of the measurements of the Cambridge students obtained at the Anthropometrical Laboratory started by Mr. Galton ("Journal of Anthropological Institute," November, 1888). A number of conclusions were drawn, of which the two following may be taken as samples: The stature of the class to which university students belong ceases to increase at a comparatively early age, viz., about 20; and when the students are divided into three classes, in respect of their mental attainments, no physical difference in respect of height can be detected between these classes. When these conclusions were being drawn I resorted to the *median*, in great part, in order to save time; but this was only done after specific examination had shown, what a wide induction supported elsewhere, that the curve of error being normal here, the median and the mean were practically coincident. What I now want to point out is that for conclusions of this kind such coincidence seems quite unnecessary.

In respect of *stature*, as just remarked, the median and mean do coincide, but in respect of *temperature* curves (and probably most other meteorological phenomena would present the same characteristic), they do not coincide. The difference undoubtedly is not great, but it is enough to illustrate the point in view. For example, this curve represents the relative frequency of the different degrees of temperature (maximum) at Cambridge during the twelve years from 1873 to 1885, including, therefore, about 4,400 observations. (The curve has been somewhat "smoothed.")



The curve is obviously unsymmetrical or lop-sided, and in consequence the median and the mean differ appreciably, by as much, in fact, as about 0.7 of a degree Fahrenheit. But almost every

question which the meteorologist can want to answer could be determined just as well by one of these as by the other, and determined moreover to a degree of precision far exceeding the admitted difference between the two kinds of average in any individual case. A graduated list, for instance, of the various English or foreign health resorts, in respect of their annual temperature, would be almost exactly the same whichever scheme we resorted to; and so would the answer to any such question as whether the climate now, as compared with former years, was improving or deteriorating so far as temperature was considered.

II. In the previous case we are only concerned to a very slight extent with quantitative considerations; that is, we are doing nothing more than putting a set of groups in their order of precedence, or determining whether any particular group is rising or falling. But the moment we take a step further and reach the ground of definite quantitative measurement, by asking such questions as, How much is one above another? By what amount has this one risen and that one fallen? a fresh set of considerations comes in, and the selection of the kind of mean is no longer so freely at our choice.

But we must first inquire precisely what it is we want to do. The mean is a single fictitious value substituted for a plurality of actual values. It stands to reason therefore that the former cannot take the place of the latter for general purposes, any more than the centre of gravity can for purposes in general take the place of the system of material points to which it corresponds.

Take the simplest of possible cases. A man disposes of 1,000 packets of tea at half-a-crown apiece, and says that on an average they weigh one pound each, with individual variations. Here, as before, we are substituting a single intermediate value for the many that were actually assigned. And what can the former do for us? Just this: It assures us that if the seller had put precisely one pound into every packet, instead of the variable amount that he actually did put, he would have been in exactly the same position; and in this way subsequent calculation may be much simplified. But it is only for this particular purpose, and even for this purpose only so far as the seller and not the individual buyers are concerned, that such a substitution of an average, or fictitious construction of our own, in place of the concrete actual magnitudes, can be safely trusted. When a quantity of rations of food are issued out, the supplying party—the government or contractor—may with perfect accuracy replace the many actual variable quantities by a single fixed fictitious quantity in the shape of a mean; but to the receiving parties the substitution may let slip facts of vital importance. But, knowing precisely

what we thus intend, we see that one kind of average, and one only, viz., the arithmetical, can be employed without error.

Now take another equally simple case: A population is found to have just doubled itself in the course of a century. There was an actual ratio of increase of course for each year which might have been determined, and if we had had all these ratios before us we should have had, as in all cases where we are dealing with an average, a plurality of actual magnitudes, possibly all differing slightly from each other. But we want for simplicity to have a *single* such value, and this we call the "average rate of increase." Will it do to resort here also to the arithmetical average, and to say that the average ratio of increase which will in one hundred years raise 1 million to 2 millions, is $\frac{1}{100}$ th? clearly not. An increase of $\frac{1}{100}$ th carried on annually for a century would raise the population from 10 millions to about 27 millions. The factor we want is, of course, the $\sqrt[100]{2}$ or about 1.007; this is familiar enough to every statistician.* The only purpose for which I bring it forward is to remind ourselves that an average in such cases as these can only be trusted to subserve one definite purpose, and that so soon as this purpose is assigned, we are necessarily bound down to one particular kind of average.

As a further illustration take the following case. The curve of temperature, as we saw, is markedly non-symmetrical. Suppose such a sample as that afforded by the curve above. What sort of average ought we to appeal to? The word "ought" I apprehend implies simply this—having defined precisely the end we want to secure, what process will serve to attain this end? As always, in adhering to a mean we are letting slip an enormous amount of information that might under other circumstances be necessary, in order to retain a more economical and convenient hold of the one characteristic which we then and there happen to want. Now what is this characteristic in the case of the temperature? I can conceive several which might correspond to our wants:—

(1.) Suppose, for instance, that it were ascertained that the flowering of plants or ripening of fruit was determined by what may be loosely called the total temperature they have enjoyed in

* In this particular form the error is hardly likely to be committed by any statistician. But substantially the same misconception, in a closely analogous form, is by no means unknown. For instance, I have seen it assumed that if (say) a population increases, during five successive equal periods, by 7, 6, 5, 4, and 3 *per cent.* respectively, we may talk of an "average" increase of 5 *per cent.* during each of those periods. This is not so, of course. The real average is, in strictness, not $\frac{1}{5}$ (107+106+105+104+103), but $\sqrt[5]{107 \times 106 \times 105 \times 104 \times 103}$. That is, not 5 *per cent.*, but 4.99. The difference here is very small owing to the fact that we are dealing with such a small number (five) of intervals; the distinctive characteristic of the geometrical average, being cumulative in its effect, is far less prominent here than if we were dealing, say, with one hundred such intervals.

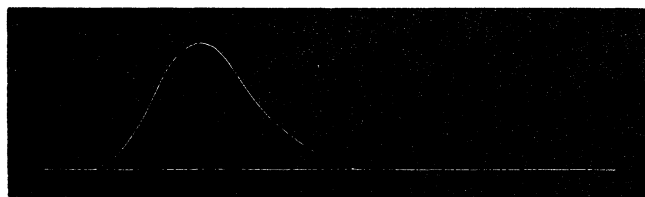
the course of that year. That is, we might conceive that two successive days on which the temperature stood throughout at 32° and 52° respectively were precisely equivalent in their effect on vegetation to two days each of which marked 42° . If this were so, and this were the characteristic of importance to us, then clearly the arithmetic average is the only one fitted for our use.

(2.) Or suppose again, that all temperatures below a certain point counted for just nothing. I presume that for the ripening of corn or fruit this is somewhat nearer the mark. If so, we should simply reject all occasions on which the temperature fell below this point, and take the arithmetical average of the rest. We should be seeking the mean point not of the whole curve of error, but of that very lop-sided part of it which fell beyond a certain mark.

(3.) But if the horticulturist ascertained that the lower temperatures counted something, but not much, towards the result, what we should then take would be not the common average but a "*weighted*" one, this being weighted according to the law, whatever this might be, by which lower temperatures were effective.

(4.) Again, we might suppose that the point of importance was, what temperature is *most frequent*? Speaking here for mere purposes of illustration, we might put the case that invalids were only allowed to go out, or that certain plants would only flourish, when the temperature lay between certain limits, say 60° and 70° . If so, the point of the "curve of error" to be attended to is what is called the "maximum ordinate" or most frequent value, rather than that indicated by the arithmetical average. No doubt these two points generally are very near each other, and often almost precisely coincide; but they are theoretically distinct, and for such purposes as we have now in hand they must be distinguished.

(5.) Once more, take a case in which what we are concerned with is the mere presence or absence of a quality above a certain degree, the *extent* of excess being a matter of indifference. For instance, if men are grouped according to their muscular strength as shown by a dynamometer, we get a curve somewhat like this:—



The difference here between the median, and the ordinary mean, indicates this abnormality, there being a separation of about one

pound between the two. The former is 83.6 lb., the latter 82.6 lb. The difference is not great, but I prefer to select an actual example from nature rather than a fanciful one which should display the difference in an extreme form.

Owing to the presence of a comparatively small number of instances of very great strength—ordinates straggling, so to say, far up to the right, and unbalanced by corresponding defects of equal magnitude in the opposite direction—we get a curve of decided unsymmetry. "*Ought*" we to take the common average here? If the population were being estimated by their capacity to furnish blacksmiths or draymen I suppose we ought, because the practical value of absolute strength here bears some proportion to its amount. Without going so far as to assert that one man who can lift 800 lbs. is equivalent to two who can each lift 400 lbs., it will be admitted that the facts, as regards efficient work, tend in that direction. But for army purposes, say, is the Samson worth a bit more than the man of good muscular power? If not, then the arithmetical mean shifts the centre of importance too far upwards. It would be quite possible for a group of men, or a nation, to show a decidedly higher average in this respect, owing to the exceptional predominance of such giants, without their being any good grounds for rating them higher for purposes of practical utility. The *median*, in accordance with which these extreme values are merely counted, but not weighed, would seem to be the truer point of reference. That is—to sum up in one sentence—for all purposes of mere comparison, one kind of average will do almost as well as any other. If, therefore, one of these kinds is situated at a more critical or important point of the group of magnitudes in question than some other, then there seems sufficient ground here for the preference of the former. Our *special* purpose may have a peculiar or predominant reference to that point. And such an illustration as I have just offered seems to indicate that a preference of this kind may reasonably exist.

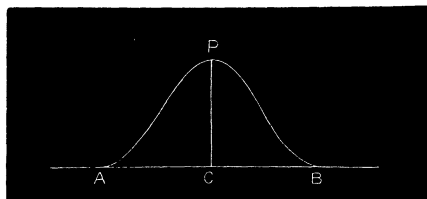
III. In the two foregoing classes of cases we have supposed ourselves simply to have a lot of variable values set before us, and to take them as they so stand, without asking whether there is anything which they are aiming at, any "*true*" value from which they swerve as errors in one direction or the other.

When we get on to this latter ground any but a superior mathematician must be cautious where he intrudes, for as every one knows this province has been profoundly studied by men of consummate power.

But the main *logical* principle, apart from quantitative details, or inquiries as to *how* accurate a result may be reckoned, seems

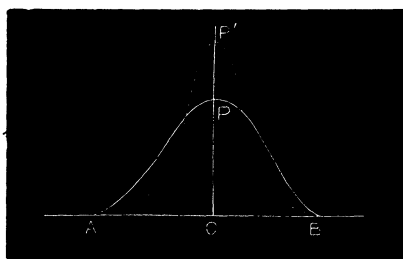
to me to be simple enough, especially when put into graphical language. The way I look at the matter is this :—

Suppose there is some point C on the line AB which is being aimed at. Then we take it for granted—



that, whilst the general circumstances and conditions of the aiming process remain the same, the actual results will group themselves according to some law of frequency about C . The amount of this relative frequency at each assignable point may be represented in the customary manner by the curve APB .

Now the average (of whatever description) of any group of these efforts or aims is simply another such effort or aim. It is generically of precisely the same character as any of the elements from which it is deduced. Represent its curve by the dotted line.



I call attention to this because I cannot but think that there is some misapprehension on this point. The plain man is apt to suppose that the due preparation, according to rule, of an "average" gives him something superior in kind to what he started with. Any assigned average is better, no doubt, than the single values from which it was deduced, but it is generically of the same kind as they were. But it is obvious that it may, on the other hand, be of exactly the same value as *another* set of single values obtained under more favourable circumstances, or even distinctly inferior to such a set obtained under peculiarly advantageous conditions.

At this point the function of the mathematician comes into prominence. He is able to tell us several things which we could not have found out for ourselves. For one thing, he tells us that, given what we may call the primary error curve APB , the derivative or "average" curve $AP'B$ will be more closely squeezed

up about the central vertical CPP', by our selecting the arithmetic mean, than by our selecting any other of the possible means. And more than this: the mathematicians are able to tell us with quantitative accuracy *how much* more squeezed up will be that dotted line according as we take groups of 2, 3, 4, &c., elements as the basis of our average.

It may be taken for granted, and generally *is* taken for granted, that the dotted curve is a better one for us than the plain one. But as we are in the logical mood, we will ask, *How* is it better? *Why* is it better? It will not do, in reply, just to repeat that the one curve is closer to the truth than the other, for two reasons: In the first place, the *curve* represents the limiting result of averaging for ever, whereas what we want to justify is the *single* average. And in the second place, to say that the dotted curve is nearer to the truth than the continuous one, is only another way of saying that little errors are more frequent on the former scheme, and that big errors are less frequent. And we should still have to face the question, *How and why* is this an advantage?

The first of these objections may, however, speedily be set aside. We *cannot* justify a single random selection from the dotted curve as being necessarily better than one from the thick curve. It may be much worse. Take a single random observation, and the average of a hundred similar ones, and the latter may be the worst of the two. All that we can say is, that "on the "average" the one is better than the other.

This brings us to what is, to my thinking, the weak point in the theory of errors as popularly explained and justified. There seem to be two difficulties which may be fairly brought forward here:—

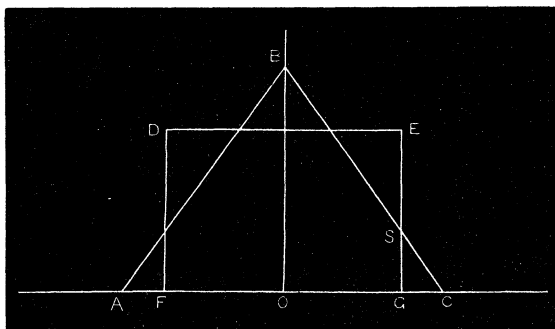
(1.) In the first place, when it is urged that "*the*" average is better—that is, that the average of the errors is smaller—it is fair to ask, *what* average? The answer presumably is, the arithmetic average. But why this one should be selected rather than any other is not, I apprehend, explained with sufficient clearness. And in any case there would seem to be a logical flaw in justifying the result of an arithmetic mean, as compared with any other similar process, by starting with the postulate that the arithmetic average of the departures from the truth to which it leads us is to be the test of its superiority.

As we shall see presently, this difficulty and some others are alleviated by certain tacit conditions or notorious experimental facts commonly prevalent in the class of phenomena dealt with. But as a mere matter of general reasoning there does seem to me something awkward here.

(2.) But, in the second place, is the mere fact that the

arithmetic average (or for that matter any other kind of average) is less on one plan of procedure than on another, a justification of the former, from which there is no appeal?

Take an example or two, as displayed in a diagram. The following represents two of the simplest conceivable laws of error. They are drawn, of course, to the same scale; that is, the area of each is the same. O is the point aimed at; in other words, the "true" value from which, by error, we deviate right and left.



DEFG represents a state of things in which all deviations up to F and G respectively are equally likely, but in which there is no possibility of exceeding these limits. (It corresponds to the familiar case of choosing a single digit at random out of a logarithmic table, or other mass of such like figures, for all the digits, from 0 to 9, will tend to present themselves equally often.) ABC represents a state of things in which all errors are possible up to the wider limits A and C, but in which the probability of an error diminishes uniformly and directly with its magnitude up to those limits. Simple calculation shows that the average of the errors in the case of the parallelogram is slightly greater than that in the case of the triangle. [If $AO = OC = a$, these averages (the "mean errors") are respectively $\frac{a}{2\sqrt{2}}$ and $\frac{a}{3}$].

Is, then, one of these necessarily better than the other? That is, would there be any reason to prefer that the deviations from our aim should occur according to one of these laws rather than according to the other? This, it need hardly be said, is not a question for decision between one kind of average and another, but a question whether any kind of average is, by itself, a sole and sufficient guarantee for selection or rejection. The question seems to me to depend upon the particular circumstances of the case, and upon the precise requirement we have in view.

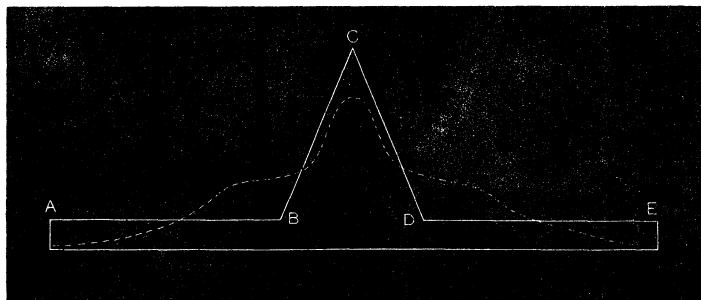
Suppose it was a case of actual aiming at a mark. Let O be

the centre of an embrasure AC, and FG the width of the shoulders of a man posted before the centre of it. If any marksmen from a distance were taking a shot at him down the line BO, he would certainly prefer that the law of error should be that which is represented by DEFG than that which is represented by ABC. But this preference would have nothing to do with an average. It would turn on the fact that one scheme gives a slight chance of escape (the ratio of the triangle SGC to BOC), and the other does not. That is, we have here a case in which the predominant determining factor is the possibility of great errors; the "mark" wants these, whatever may be the wish of the marksman. On the other hand, for many scientific and other purposes the existence, on a certain scheme, of very wide errors may be quite sufficient to decide against a method which opens out such a possibility. And again there would be no necessary reference to an average.

The full state of things in the balance between the two simple laws of error indicated above is, of course, this: The triangular scheme or curve contains a greater proportion of very small errors, a less proportion of intermediate errors, and, again, a greater proportion (beyond certain limits a monopoly) of extreme errors. The simple assignment of an average blurs over these distinctions; but nothing less than the full statement can meet all the various purposes for which appeal might be made to one rather than the other.

In the former case we have drawn arbitrarily two conceivable "curves of error" for comparison. It may be objected that this is not a fair illustration, because one of them is not derived from the other by the process of averaging. That is, it may be urged that whatever the original curve of error, the derivative curve—*i.e.*, that assigned by averages taken from the former—must be better; better, that is, all through, so that no such contest of opposite qualifications can be possible.

Take, then the following example:—



Conceive a curve of error ABCDE defined by the following very simple conditions: Within certain near limits, B and D, errors are progressively less likely as they increase in magnitude; beyond this, to the remoter limits, A and E, all errors are equally likely. Take an average of *two*, then the dotted line will represent the curve of error of such averages of pairs. Again the same inquiry as before is raised, though here the question comes nearer home to us, for what we are asking is this: Has the process of averaging done us any good? The answer seems to be, as before, that this depends upon what our wants may happen to be. Is the increased prospect of small errors sufficient to outbalance the counter prospect, likewise increased, of errors which are somewhat larger?

As the above example is only proposed by way of illustration, there would be no harm in its being merely fanciful. But I would suggest that similar alternatives for decision might very well arise in purely practical affairs. Put such a case as this. "Conceive that some engineering firm had received a hurried order to export to a distance a single piece of a machine in order to replace a piece that was broken, and that it was absolutely essential that the work should be true to the tenth of an inch for it to serve its purpose. Conceive also that two specifications had been sent, resting on different measurements, in one of which the length of the requisite piece was given as 60 and in the other as 61 inches. On the assumption of the ordinary law of error there can be no doubt that the firm would be making the best of a very bad job by constructing the piece of the length of $60\frac{1}{2}$ inches, *i.e.*, they would have a better chance of getting within the requisite tenth of an inch by so doing than by taking either of the two specifications at random and constructing the work accurately to this. But if the law were of the kind indicated in the diagram above, then they might assume that they would have a *less* chance of success on the former plan. That is, as a mere question of probability—if such estimates were acted upon again and again—there would be fewer failures encountered by simply choosing one of the conflicting measurements at random and accurately working to this than by trusting the result to the average of the two.

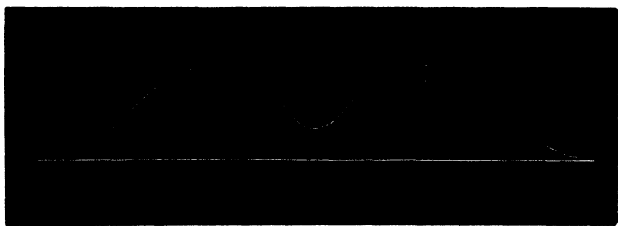
"I may be excused pointing out that even such a law of error as that in question is not outside the bounds of possibility. Suppose, for instance, that one of the two measurements had been made by a careful and skilled mechanic, and the other by some workman who had simply reckoned to the nearest inch—the firm having come to a knowledge of this fact, but being unable at the time to assign the two results to their authors—we

"should really have very much such a law of error as I have "postulated above."³

Now take the class of cases in which the curve of error is hump-backed, *i.e.*, presents *two* maxima with a depression between. Ought we to take the mean, or average of any description, of the two? This turns upon what we want to do with the data, and on the process by which they may have been obtained. Broadly speaking, curves of this particular description may be obtained in two distinct ways, *viz.*, by what may be called artificial or natural processes, and a good deal turns upon this distinction.

(1.) As an artificial process we might take the aiming at any kind of mark, in the widest sense, actual or metaphorical, of the words "aim" and "mark." If in addition to the many trifling causes of disturbance which produce the familiar law of error, we suppose two preponderant causes, of about equal and opposite influence, which tend to act alternatively, we shall get such a result as that of a hump-backed curve of error.

For instance, we know that under ordinary assumptions as to the prevalence of disturbing agencies, the shot marks at a target will cluster about the bull's-eye, and therefore yield what is, diagrammatically, a simple curve of error of the familiar description. But we might conceive that circumstances should give a strong alternative bias to the aim. If, for example, the marksman, firing from a sheltered glen across a broad valley, had no suspicion of the existence of a strong wind blowing up or down the valley, and could not see the spots, where the bullets struck, the effect would simply be to give a decided lateral shift to the whole group of shot marks. And if the wind veered into the opposite direction during the course of the practice, we should get two such groups of shot marks, each clustered about its own centre, but with these centres a considerable distance apart. The diagram for the consequent curve of error would be such as this:—

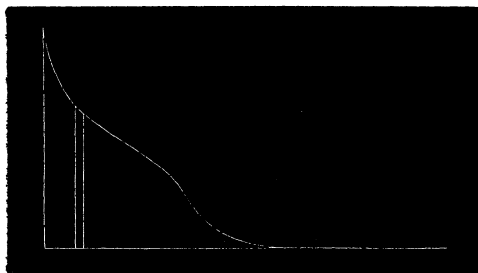


Here, again, the question arises: "Ought" we to appeal to the average? That is, given any two or more shot marks, should we assume that the mean point amongst them is the best available clue to the point really aimed at?

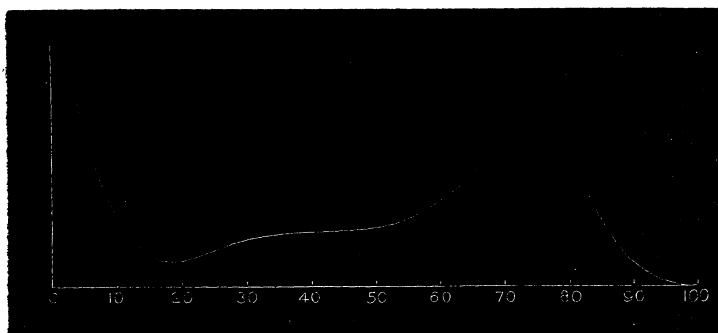
³ "Logic of Chance," p. 495.

In such a case as this, where the phenomenon is an artificial one, that is, where we know that one definite aiming point, and one only, was contemplated throughout, we should of course appeal to the average. And it would in the long run guide us to that central spot, which though almost unattained in execution was constantly present in intention. But then, in such a case as this, since we are dealing with human actions, we know exactly what we are about. We feel tolerably sure that the marksman contemplated one single aim throughout, and we interpret the facts in accordance with this pre-conception.

(2.) Now contrast with this the following case: Every statistician is familiar with the ordinary life table as expressed in a diagram, where the ordinates successively represent the number of some total generation at birth, and the number of survivors in each following year.



Modify this mode of representation slightly, so as to make it exhibit what might be termed a *death* table. That is, let each successive ordinate stand not for the number of survivors, but for the total number who have died in that year of age. This second figure may be obtained at once from the former by the following process. Detach each elemental strip by which one ordinate falls short of the preceding, set all these strips up side by side as a new set of ordinates, and, for convenience, magnify them all in the same ratio, and we get at once such a curve as the following:—



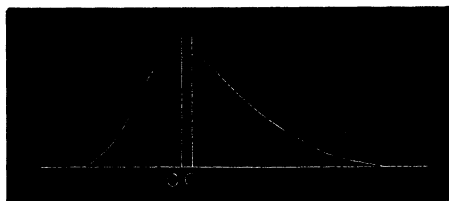
Here, of course, the whole area represents the total generation (corresponding to the first vertical ordinate slip in the preceding figure); and each successive ordinate represents the chance of a person of whom nothing is known but that he has just been born, dying in the year corresponding to that ordinate. The first maximum, to the left, is of course at birth. The second, to the right, is (for the English race) at about the age of 71.

Such a curve of error resembles in essentials that which indicated the comparative prevalence of the shot marks in our last example, in that it displays two maxima with a minimum between them.

Ought we then here also to take the average? In other words can we rationally say that nature was aiming at a life of about 41 years, but was as a rule diverted from her aim, falling short of it in a vast number of cases by some 40 years, and exceeding it in many others by some 30? Such a suggestion is of course merely noticed in order to emphasise the extreme insignificance of the simple average taken by itself. The average duration of life no doubt subserves useful purposes when employed in the way of comparison, though even such use is largely dependent upon the existence of certain tacit assumptions. But taken by itself it really tells us almost nothing. The particular age indicated by the average has none of the significance acquired by the mean point when that point, as commonly, is also the point of maximum frequency.

If we are to use the language of metaphor, and speak about "aims," it would seem more appropriate to recognise *two* such. One at about the age of 71 or 72, which is nature's aim; the length of life for which she builds a man; the dispersion on each side of this point being (as Professor Lexis has shown by the statistics of various European countries) nearly normal. The other is at birth. It is (so to say) the aim of man, as interpreted by his unsanitary proceedings, to kill off the children at once; an aim which he also often fortunately fails to secure. The dispersion at this end of the curve of course cannot be normal, for we reckon only on one side of the point of aim.

Speaking of "nature's aim" in this metaphorical sense, it may be pointed out that a precisely similar question is raised, though in a less sharply alternative form, whenever we have an abnormal, asymmetrical error curve. Take for instance the curve exhibited by the daily readings of the barometer. The result of 5,000 such readings, when plotted on paper, gives a very smooth and steady curve, but one which even to the eye is very distinctly unsymmetrical.



The mean point and that of maximum frequency differ by 0.07 of an inch, a difference which is somewhat considerable in a total extreme range of about 3 inches.⁴ Suppose now that such a curve were set before us, with the accompanying intimation that it was to be regarded as a true "error" curve, but without the slightest intimation of the agency by which it had been produced, how should we read it? Surely we should conclude that the point aimed at was C and not O; and that as regards the marksman, he, or it, or whatever was the agency in question, was subject to an unequal number of opposing diverging influences. As most of those present know well, the ordinary symmetrical error curve is the outcome of a continued aiming where the opposite disturbing causes are *equal* in number, independent, and infinitely numerous. You know also, probably, that if the causes still remaining indefinitely numerous are *unequally* divided, for example, with twice as much tendency to the left as to the right, we still get an error curve of the symmetrical description, the only difference being that the position is shifted two-thirds of the distance from the right limit towards the left limit. Suppose then that such a curve were set before us, as a mere fact of experience, and we were left to conjecture by what actual process of agency this unsymmetrical arrangement of results had been brought about, what process would seem at once the simplest and most natural? I think such a suggestion as this would seem the most natural: Conceive a *small* number of equal and opposite disturbing agencies unequally divided, say in some such ratio as 5:4, we should have an error curve which, *when smoothed out*, would yield much such a shape as we want. The physical counterpart to the graphical "smoothing out" is, I apprehend, the supposition that each of these primary disturbing agencies is itself subject to a large number of minute disturbances which are equally divided in opposite directions.

I need hardly say that this is not intended as a suggestion that the table of barometric heights was brought about by any such set of agencies as this. But what does seem to be suggested is that when, in cases of this unsymmetrical kind, we come to talk about

⁴ Here C marks the maximum ordinate; O the point corresponding to the arithmetic mean.

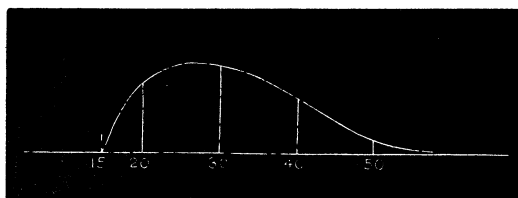
“aims” and “objects,” as we so commonly tend to talk, then the centre of interest and significance must be sought, not at the “average” point, but at the point of “maximum frequency.” And when, as in such a death-table curve as that just displayed, we have *two* such points of maximum frequency, it is these, and not the average point somewhere between them, which can claim this interest and significance.

It may be objected that all these necessities of distinction between one kind of average and another have very little practical importance, for that the familiar arithmetical average works sufficiently well in every case to which we choose to apply it. If what I have urged this evening be true, such an objection has less force than is commonly supposed, but it is well to consider what are the causes to which it owes such force as it has.

I. In the first place the great majority of the error laws with which we practically have to do are approximately of the same normal type, at any rate towards the centre, and it is with the neighbourhood of the centre that we are chiefly concerned.

This, I presume, is the reason why even theoretical requirements seem to be sufficiently secured in the province of mensuration by the assignment of two elements only, viz., “the mean” and the “probable” or “mean error.” The first of these fixes the centre of a Probability Curve, or Error Law (C); and the second assigns the degree to which it is spread out, or crowded up towards the middle. If to these explicit data we add the implicit assumption that all the admissible curves are to belong to the same general class, the two elements in question suffice for complete determination.

When we are dealing with other considerations than those of mensuration, especially with such as concern the social state, we must be careful how we admit this assumption. There is much popular looseness of apprehension in this respect whenever questions of average wages are discussed. To say nothing of the popular agitator who has a purpose to serve, and who, when he describes the ever deteriorating condition of the “proletariate” expressly confines his attention to the stratum at the bottom; even the *average* condition of the labourer is insufficient for social economic purposes. I need not point out to such an assembly as this, that what the social inquirer wants is not merely an average or middle point; but that what may be termed the *shape* of his curve, and any possible alteration in this shape are matters of vital importance. The only satisfactory plan in such a case would be to construct, say, a base line marking amounts of wages, and to set up ordinates marking the relative numbers in receipt of each such amount, in some such fashion as this:—



If a similar curve were plotted out in order to represent the state of things, say, fifty years afterwards, the position of the average point, and whether this had shifted upwards, would be only one subject of inquiry, and possibly not the most important. Characteristics which that average could not tell us might for social purposes be yet more important, such, for instance, as the relative distribution of the various amounts of wage in excess of, and in defect of, the mean.

II. Another condition which, as it seems to me, is commonly taken for granted, and necessarily so, is continuity of possible values between the prescribed limits. An extremely simple illustration will show how this characteristic is instinctively recognised in certain cases, even by those who have made no study of the nature of an average. If two persons respectively reckoned the distance of Cambridge from London, as being 50 and 60 miles, a third person, who had no reason for suspecting one distance rather than another, would in default of further information naturally guess 55 miles to be the most probable distance. But if two such persons respectively gave it as their opinion that some one known to them was at Cambridge and at Oxford, the "most probable" situation for him would hardly be in one of the small villages between Bedford and Bletchley. In other words, to speak more precisely, it is commonly taken for granted in most kinds of measurement that the thing which we have to evaluate is prior to, or entirely independent of, our scale of measurement; so that it is just as likely to fall at one point of our scale as at another. Such a supposition is generally valid when we are concerned with natural objects, but needs some caution when it is applied to artificial objects. A simple example will serve to show the contrast here intended: if two measurements of the circumference of a *tree* are set before us, we should, in accordance with common usage, select the mean between them as the most likely value. But if two measurements of the circumference of a *drain pipe* showed the same magnitudes and discrepancy, should we be equally justified in selecting their mean? We should be very cautious in doing so, for I should think it extremely probable that drain pipes are made in even numbers of inches in respect of their diameter. If

therefore the two estimates gave respectively $21\frac{1}{4}$ and $22\frac{1}{2}$ inches of diameter, I should select—not $21\frac{7}{8}$ —but 22 inches. And if the mean fell exactly between two clear inches, I should think that it was better to reject this intermediate value, on the ground that we are almost certainly wrong in adhering to it. And in regard to the complete inches between which it lies, it would be as safe a course as any to choose at random between them, as we should then have very nearly an even chance of being exactly right.

Where we know nothing of the scale of measurement according to which the object is likely to have been constructed, this principle cannot be applied. If the drain pipe came from Japan, and I knew nothing of the Japanese units of length, though it still remains likely that certain lengths which may have been exactly yielded by the average are to be set aside, and some other length which was only approximately yielded is to be selected instead, yet in our ignorance as to what these latter are, the average again asserts its probable superiority. At least it does so when we have but a few results to deal with. Where the results become very numerous, I need hardly remind those acquainted with Mr. Flinder Petrie's brilliant investigations into the subject of "Inductive Metrology," that a mere study of the measurements obtained, and observation of what lengths show a preponderant frequency, will by itself yield strong evidence that this or that particular magnitude is the unit according to which the materials were worked. As soon as this knowledge is acquired, the average may again stand in need of occasional correction or even supersession.

The general conclusions involved in the suggestions above offered might be thus summed up. Every sort of average—and there are many such sorts—is a single fictitious substitute of our own for the plurality of actual values existent in the results which are naturally or artificially set before us. It is impossible, therefore, for the former, in any case, effectually to take the place of the latter. But the extent to which it may succeed or fail in doing so will depend upon the nature of the facts presented to us, and still more upon the precise object we have in view. Three cases may be suggested:—

- (1.) If we are only concerned with comparative results, and these presumably form the bulk of statistical inquiry, then almost any kind of average will answer the purpose.
- (2.) If we want accurate quantitative results, then the selection of the kind of average is no longer at our disposal, but must depend upon the precise object we have in view. The common arithmetic mean *may* be the best, and

probably in most cases is so, but circumstances can easily be suggested in which some other kind of mean becomes not merely more convenient but distinctly more accurate.

- (3.) There is the well known class of cases, of which the Science of Mensuration offers the most familiar examples, in which the "mean" is resorted to in order to discover what is regarded, accurately or metaphorically, as the true value. Without in the least wishing to criticise the customary methods and conclusions, I have ventured to suggest that it may depend upon the object we have in view, not merely what kind of average should be selected, but whether there is a necessary gain in employing any kind of average at all.

When, indeed, the "true value" at which the aim is supposed to be directed becomes purely metaphorical, as in the curve of mortality, it is very doubtful whether the resort to a single mean or average is not, for some purposes, not only unsuitable, but actually misleading. Where the Curve of Error is very unsymmetrical, and still more where it is hump-backed or double-backed, the mere assignment of an average, without definitely indicating the law of grouping about this average, may misdirect the attention from some of the most important characteristics of the phenomena.

DISCUSSION *on* DR. VENN'S PAPER.

Mr. FREDERICK HENDRIKS said he was sure all present would concur in thinking they were very much indebted to the learned author of the paper for his clever and ingenious exposition of principles, which were of very great importance to that Society. He had done very great service in recommending the continuance of that caution which they had more or less always exercised as to the particular methods by which averages were taken. It was to some extent gratifying to find the author did not object to the system ordinarily adopted of taking the arithmetical mean, because, in that Society at least, taking the geometrical mean, in such averages as they usually had to investigate, was a process in which they did not often indulge. The paper contained however a great many very startling premisses, and a great many startling corollaries drawn from those premisses. The points which particularly attracted his notice as an actuary, were the observations made respecting the life table, stated in the form of a table showing

the expectation of life both at birth and at every other age. The author evidently thought very small pumpkins of it, but this suggestion was directly in contradistinction to the ideas accepted by the general body of actuaries. As a member of that fraternity he (Mr. Hendriks) thought he might be allowed to say that it generally regarded the method of construction of the life table as the perfection of human reason, and that, given the accuracy of the data, nothing could be devised by scientific skill which could in any way afford a more perfect rendering of what was most important in vital statistics. He must therefore join issue with the author as to the expectation of life indicated by the life table being of extreme insignificance even when taken by itself. It was the most concrete form in which the results for comparison of the ever-varying longevity of the people of the country could be presented. Let them consider how the ordinary table of expectation of life was made. During the last month they had been giving to the census enumerators returns telling the ages of the people of this country, so that the numbers exposed to the risk of mortality at each age on 1st April, 1891, were ascertainable from these returns, and in the course of the current year the deaths at every age would also be recorded. That would give sufficiently exact data for estimating what was the rate of mortality in the year from 1st April, 1891, to 1st April, 1892, at every age they could name. Both practically and scientifically there could be no more perfect means of getting at the mean rate of mortality at every age in this country. The table of expectation of life simply meant that if the ratios of mortality at all ages continued during the whole future possible years of lifetime of those people as they were at present, a certain definite amount of longevity would on the average attend to each person. That was not insignificant, because it was the best means of comparison which was comprehensible to the public mind in relation to the effects of sanitary or other circumstances affecting health and length of life, and it was also the means of comparison with the expectation of life in other countries. If there were any changes hereafter occurring in the observed ratios of mortality from which the expectation of life is deduced, precisely analogous observation went on at the next census, and they were thus able to correct the calculation from period to period. It was indeed somewhat difficult to see how Dr. Venn had arrived at a conclusion so very adverse to that particular branch of statistical inquiry. He (Mr. Hendriks) had only had an opportunity of considering the paper for about an hour, and was not sure that he could quite find the key to the riddle; but it seemed to be this: The author first took the curve showing the ordinary life table. He then proceeded to say that he wished to modify this mode of representation slightly, so as to make it exhibit what might be termed a death table. Now if the curve of the life table were simply turned upside down it was clear the diagram would represent the death table, since life or death was the only possible alternative to be illustrated in the two curves. But instead of doing that, the author had made a diagram which expressed the compound probability of surviving from birth to any given year of age, and then dying in that

particular year of age. Naturally any curve representing a compound probability of that kind would show very irregular results; in fact if they deduced a curve from such compound probabilities they might have them wriggling about in their delineation something in the shape of an eel or a snake; no doubt there would be a great many maxima and minima in it. He was also at issue with a great many of the collateral statements of the author in justifying his views on this very important subject of criticism of the merits of the life table. The author of the paper seemed to attribute the irregularities of his curve to the great death-rate at birth caused by unsanitary proceedings, and to something special surrounding the age of 71 or 72, which age he describes as nature's aim, the length of life for which she builds a man. This normal space of life would, according to the theory of the late M. Flourens of the French Academy, have to be extended to 100 years, or we might go farther back, and with Plato (in his "Timæus") allege that all deaths by disease prior to the period prescribed by nature, are violent deaths. Now of course we know that in the first year after birth the mortality is really greater than at any age until we arrive at the ages of from 80 to 84. But Dr. Venn assumes that it is not nature's law that it should be so. But it was very extraordinary that as far as the records of the human race went it had always been so, and it was most questionable whether it was not rather nature's law that infants should die in proportionately greater numbers soon after birth than at any other period of life except extreme old age. There were many causes which showed that nature, or as he would rather call it, providence or the divine decree, designed that children should die in comparatively large numbers, as otherwise the world would be very much over-peopled. If children did not die in something like the ratio shown by the life table, and if all people lived to the average expected age of 70 or 71, for as he understood the author that was the natural average length of life, the result would be the world could not hold the population that would exist, and they would either starve or else perhaps be eating each other out of the means of life.

As regarded the method of taking averages of barometric observations, Dr. Venn recommended that there the points to be looked to should be varied according to the ideas of people. He (Mr. Hendriks) could not conceive any better method of registering barometric observations than the usual one of showing the average height during the day, month, or year, for if the other plan of recording maximum frequency of a particular reading as more important were pursued, agriculturists on the one hand desiring wet weather, and railway companies on the other hand wishing fine weather, would frequently be in conflict of judgment according to their own particular view of what the "error curve" really consisted in. As to the author's recommendations with regard to the averages applicable to a branch of social economics, although the suggestion in the diagram was a very good one, it only referred to the amount of wages and the relative numbers in receipt of those wages. Such a diagram would be very defective

if taken at intervals if it did not refer to the percentages of the working classes receiving wages of given magnitude at each time. If it was confined merely to the amount, it was clear it would not be of much value. He (Mr. Hendriks) could not quite understand the process of reasoning connected with the author's example of average measurements of the "drain pipes." The author spoke of two estimates, presumably those of two surveyors, who gave figures respectively the one of $21\frac{1}{4}$ and the other of $22\frac{1}{2}$ inches. He then said he should not select the arithmetic mean of $21\frac{7}{8}$, but 22 inches; but the reasons he assigned for that view appeared to be unscientific, and such as they ought not to adopt. It was only two measurements or two estimates that could come into play in determining the probable average number of inches. But any hypothesis such as Dr. Venn's, of its being extremely probable that this was a mistake, because he thought that drain pipes are made in even number of inches in respect of their diameter, and that therefore we should be almost certainly wrong in adhering to the arithmetical or intermediate value of the measurement, was not an admissible element. Indeed it was opposed to what was scientifically settled at the earliest time when the bearing of arithmetical means on the science of probabilities was first investigated. This was a question of some importance, and was by no means a new one. As long ago as the year 1700, Leibnitz, in his "Nouveaux Essais sur l'entendement humain," stated that in the valuation of lands, when any heritage had to be sold, the peasantry of Lower Saxony were accustomed to form three companies of valuers, and each company made a separate valuation of the property. The arithmetical mean between the three equally admissible suppositions, anything extraneous to the actual figures not being admitted, was then taken by the axiom *æqualibus æqualia*, for equal suppositions we must have equal considerations. This, as Leibnitz remarked, was the foundation built upon by the grand pensionary John de Witt in the first mathematical application of the doctrine of probabilities to the construction of a life table, in his short treatise on life annuities printed in 1672. This treatise Leibnitz was unable to find, but on its recovery and translation by himself (Mr. Hendriks) in 1851, it was seen that Leibnitz was right, and that the proper application of the arithmetic mean was the true basis of the first application of the theory of probabilities to the life table and to the scientific calculation of life annuity tables.

Mr. FRANCIS GALTON said the author's extremely lucid paper had the great merit of treating his audience not merely as recipients, but of making them think, and thereby widening their views. There could be no doubt that there was a kind of idolatry of the word "mean" amongst statisticians generally, who too frequently used it for purposes to which it was not suited. Their late President, Dr. Graham Balfour, commented strongly upon a common form of a misuse of means in his latest annual address. No outsider could read current statistics without being painfully impressed with the frequent want in them of logical clearness.

Men of science were often reproached as a class with being deficient in the logical faculty, and he feared the accusation was true. No science stood more in need of the criticisms of logicians than statistics. They all knew that Dr. Venn possessed a logical mind in an eminent degree. In his admirable work "The Logic of Chance" he had tapped, as it were, at a multitude of current assumptions to find how far they were sound, and had pointed out numerous flaws or ambiguities; this he had also done in the present paper. The view that the author took of the arithmetic mean did not precisely coincide with that of his (Mr. Galton's) own, which had, he thought, the merit of being a somewhat more general conception, and yet less liable to misapplication. It was as follows: They all agreed that groups of events, in order to be of use statistically, must possess statistical uniformity; in other words, that all large and equally numerous samples of them must be statistically alike. It was not to be expected that they should be mathematically identical, because all statistical results were only approximate. Still, so long as the constitution of all large samples was so nearly alike that whatever differences there might be between them did not lead to appreciable error, then they might be described as statistically uniform. His own view of the arithmetic mean was to look at it merely as a multiple obtained by dividing the sum of the values by their number. In other words, the arithmetic mean was a coefficient which, multiplied into any given number of values taken at random, gave their sum. From this point of view the mean was not regarded as a fictitious interpolated value, but a coefficient which gave the sum of any given number of values by multiplying it into their number. The mean so regarded was simply a method by which a summed value could be obtained. As no mean can adequately represent a plurality of values, it appeared advisable to have schemes of distribution before the mind, as far as was possible, when devising the most appropriate form of statistical treatment in intricate problems. He himself always adopted this simple method of considering a series as a whole, until he felt that he knew exactly what he was able to get out of it. Among the many different possible means it must be remembered that the advantage of the arithmetic mean was that it was the only one suitable to the case of absolute ignorance of the natures of the disturbing conditions. If that absolute ignorance existed, then it was as nearly an axiomatic truth as could be that the arithmetic mean was the best of all; but statisticians were very seldom in absolute ignorance of the conditions by which the events they dealt with had been moulded. It was in the neglect of partial knowledge that the chief pitfalls in the path of statisticians appeared to consist. Statisticians ought to take every scrap of their knowledge into account, and to utilise it to the extent it fairly admitted, not giving it too much weight, but never ignoring it. He had tried to make a selection of marked forms of lop-sided curves, and endeavoured to find out the more prominent theoretical conditions under which they might have been formed. It was a suggestive process, for whenever a new lop-sided curve had to be considered, a reference to such a list would afford a

certain number of suggestions, by no means exhaustive, but which might be helpful. He could only refer once more to the extreme lucidity of Dr. Venn's paper.

Professor EDGEWORTH remarked that the author had done very useful work in inducing members of the Society to meditate about that statistical practice in which they were constantly occupied. He forgot what proportion of life Matthew Arnold said should be employed in metaphysical reflection about conduct, but he was sure that what might be called meta-statistical contemplation was occasionally very necessary. He wished to connect this remark with some observations on the two stages of statistical theory which the author appeared to have constructed. He meant first, that more familiar region of the common average or mean value which was, so to speak, accessible to all; and secondly, the higher storey only accessible by means of the scaling ladder of mathematics, which consisted of deviations from averages or mean values. Under the first head he might take as an instance the determination of mean variation in prices. For example, there was a special mean correlated with the question: how much the price of what were considered necessities, the price of the board and kit of the general population, might have varied. But apart from that there was a general mean, the mean value of the given price variations, regarded as so many figures, and without reference to any special purpose. The author was, no doubt, quite right in saying that the general average need not be always the arithmetic mean; in fact, for some purposes other means were better. He took as an example a recent Parliamentary paper in which the average number of hours worked per week, in several trades at different periods, was shown. They had there some of the characteristic dangers of averages. In the first place the average week did not state the amount of overtime. He (Professor Edgeworth) was at present concerned to speak of the convenience and accuracy of one of the means of which the author spoke—the median as contrasted with the arithmetic mean in a case like the following. In the case of the stone-masons the standard hours worked in some 132 different places were given for the years 1890 and 1870. It would be very laborious work to take the arithmetic means of these two sets of figures, and moreover they would be rather apt to be inaccurate. There might be some little pottering place which had experienced a large variation in the number of hours worked, and by making it count as much as the others in the formation of the arithmetic mean they might very much distort the average; that little place might unduly impress its individuality upon the whole average. The better mean was the median. The median for all the places proved to be 56 (hours per week), for 1870; if they split up the series into two parts the median for each part would still be 56. For the series of 132 returns for 1890 the median was 54; the same result being obtained both for the parts and the whole. That was as good a measure, under the circumstances, as they could get of the change of the length of hours per week. He would invite some practical statisticians to state objections to this median; apart, of course,

from the special purposes for which an arithmetic mean was suited. In such a case as he had put, another mean mentioned by the author was also appropriate, namely, the greatest ordinate, the centre of greatest condensation. But there was a practical difficulty about the use of this mean, the difficulty of determining it definitely, especially when the number of returns is small. Upon the second and higher stages of averages he might say how very fortunate they were in having Dr. Venn present to lead them to reflect upon first principles. Dr. Venn occupied rather a peculiar position with respect to this higher and more abstract theory of statistics—a position much like that of the historical economists with respect to the earlier and severer deductive reasoners—criticising the received axioms and rounding off their too sharp corners. They had an instance of this in much of what Dr. Venn had brought before them with respect to the character of the deviations from the true mean: saying that they must look to the purpose and end in view in measuring the disadvantage of inaccuracy, in fact introducing an element of utilitarian philosophy. That element seemed to be lost sight of by the practitioners to whom Dr. Venn alluded as “starting with the postulate that the arithmetic average of the departures from the truth to which it (a proposed process of taking an average) leads us is to be the test of its superiority.” It is true that the phraseology of some of the mathematicians was open to that objection. Yet it would be found on careful examination of Laplace’s reasoning, that he had not neglected the utilitarian aspect of the problem of errors. “Interpreted generously,” Laplace, like Ricardo, generally came out right.

In further illustration of the criticism to which the older theorists were open, Professor Edgeworth then referred to the calculations of life tables. First, he denied that Dr. Venn’s remarks were meant to disparage that triumph of ingenuity. It would be generally admitted that, as the author intimated, an increase in the average duration of life was a most imperfect test of improvement, unless they knew at what ages the gain took place. In his idea of a death table, the author had in view the abstract theory of type. The author was of course alluding to Lexis’ theory, that the distribution of these decrements fell into a certain typical form; and there was a kind of reality about that. There was, in fact, as Lexis said, some kind of presumption that nature was aiming at the length of life which the Psalmist assigned. With respect to these actuarial calculations, it was a most profound remark of De Morgan’s, that there was a very great deal of waste of labour in pursuing different decimals to the seventh place, without having any test of the “error” to which those figures were liable. De Morgan himself, walking in the way of the older classical probabilists, calculated a rule for the same error (in his *Calculus of Probabilities*), and he calculated it upon a principle which Dr. Venn and his German colleague, Mr. Lexis, had shown to be generally inapplicable, namely, the analogy between the ideal cases of balls and bags and these cases of concrete statistics. De Morgan’s calculation as to the probable error of these actuarial calculations might remind them of the saying of the witty French-

man: that after estimating the error of the calculation, one ought to estimate the error to which that estimate was liable. For it would seem just for want of this meta-physical or meta-statistical reflection to which the author had introduced them, that the error assigned by De Morgan was very much below the mark; the actuarial calculations were really even more uncertain than De Morgan admitted.

In conclusion, Professor Edgeworth admired the felicity with which Dr. Venn extracted profound truths from homely illustrations. In this respect he might be compared to Socrates, who, we read, was always talking of "smiths and cobblers," and common objects. Professor Edgeworth adverted particularly to Dr. Venn's instance of drain pipes. He had there conveyed with remarkable simplicity the important principle that the methods prescribed by the mathematicians for obtaining the best mean, are only valid on the assumption (generally, but not necessarily always, justified) that the *à priori* probability of the true mean having any particular value was the same or not materially different for all possible values.

Dr. VENN (in reply) said he wished to clear up a misconception. He must have expressed himself badly in the paper if he had seemed to say a word derogatory to the familiar life table of which he had given a diagram. He had the greatest possible respect for it, and if he passed it over in five or six words, it was merely because he took it for granted in such an assembly that its merits would be so universally recognised that no more was necessary with reference to it. He did not quite understand in what sense it was said that the next diagram was obtained by turning the previous diagram topsey-turvey.

Mr. HENDRIKS said his remark was that the author could have done so if he had wanted to show a death table. Instead of that he had constructed a diagram showing, *not* a simple death table, but the compound probabilities of a life at birth attaining each possible age of human existence, and then having survived that age, dying in the year following.

Dr. VENN said the point shown by the table was one which was simple but also very important. Each respective ordinate in the bottom table allowed the relative chance of a person (just born) dying in a given year, and there were certain characteristics thus brought out which, although they could be deduced from the previous table, were by no means so obvious. The bottom diagram in fact corresponded in the height of the ordinates to the inclination of the tangent of the curve in the previous diagram. He could not accept the doctrine as to the natural age of man being 100. A mere glance at such a table as this showed that the natural or normal age of the English race was rather about 71 or 72.

Mr. BAILEY: The extreme age?

Dr. VENN: The age about which the survivors from youth centred. Given the mere knowledge that a child was born, if they were asked to name successively the ages in their order and the probability at which it would die, they would begin 1, 2, 3, 4, and would then go on to 71 or 72. After centreing for a time about that latter point they would revert again to 5, 6, and so on. One reason why he called attention to this figure was that there were certain characteristics which, although they might be obtained by other means, were only prominently exhibited to the eye by this curve; for instance, what had been marked out by Professor Lexis in considerable detail, that they might have different races of men where the average of life was identically the same, though the second maximum or "normal" age was decidedly different. He wished to clear himself of any imputation of having spoken in disrespect of the common life table, which for certain special purposes, and for the purposes of an actuary, he would fully admit was all important. The point about the drain pipe he confessed he must adhere to. The illustration quoted from Leibnitz, that when the question of the value of a property was to be determined it was done by taking the average of the estimates of two or three valuers, belonged to a class of cases which were well understood. He had taken for granted a good deal of what was familiar and common, and had laid stress upon the comparatively few exceptional cases in which their *à priori* knowledge of the particular circumstances in hand enabled them to set aside and to correct the familiar arithmetic average.

The PRESIDENT said there remained but one other duty to be performed, namely, to give their most cordial thanks to Dr. Venn for the admirable paper which he had brought before them, and also to those who had taken part in the very profitable discussion that had followed.
