Sur le principe de la moyenne arithmétique

J. V. Schiaparelli*

Astronomische Nachrichten Vol. LXXXVII (1875),
Nr. 2068 col. 55-58

I myself take the liberty to communicate to you a demonstration of the principle of the arithmetic mean, which appears to me exempt from the difficulties to which one has objected in many other similar demonstrations. These difficulties have been able to have their origin in the point of view too exclusively analytic, under which the question has been considered. By introducing into the calculations the conditions practically inseparable from the nature of the quantities obtained by observation, all difficulties vanish.

Let a_1, a_2, a_3, $\dots a_n$ be the results of a number n of direct observations of one same quantity, having all the same weight: let $F(a_1, a_2, a_3, \dots, a_n)$ or simply F be the function expressing the mean which it is necessary to adopt. We will determine F by means of the following conditions.

I. The magnitude of the mean must be independent of the unit chosen for the measure; so that if one multiplies $a_1, a_2 \dots$ by any coefficient k, the mean must become also k times times greater. It is necessary therefore that F be a homogeneous function, in one dimension, of the variables a_1, a_2, a_3, $\dots a_n$. By a known theorem, the functions of this specie must all satisfy the equation in the partial derivatives

$$a_1 \frac{dF}{da_1} + a_2 \frac{dF}{da_2} + \cdots + a_n \frac{dF}{da_n} = F \tag{1}$$

which will be our first condition.

II. The position of the mean between the particular values a_1, a_2, $\dots$ must be independent of the origin of the numeration of the measures: in other words, if one adds to all the observed quantities $a_1, a_2 \dots$ one same determined quantity α, the mean must be increased also by the quantity α. One has therefore

$$F(a_1, a_2 \dots) + \alpha = F\{(a_1 + \alpha)(a_2 + \alpha) \dots (a_n + \alpha)\}.$$

This relation must take place whatever be the value of α. By developing the second member according to the powers of α by the series of Taylor (of which the usage is here free of all objection), one will see that it is equivalent to the equation in the partial derivatives

$$\frac{dF}{da_1} + \frac{dF}{da_2} + \frac{dF}{a_3} + \cdots + a_n \frac{dF}{da_n} = 1 \tag{2}$$

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. August 3, 2010 a second condition, which the function F must satisfy.

III. Let a length L (or another measurable quantitiy) be given, divided into two parts A and B. Through n equally exact measures of A one has found the results a_1, $a_2 \ldots a_n$: likewise n equally exact measures of B have given the results b_1, $b_2 \ldots b_n$. The concern is calculating the particular values and the definitive value of L. One will be able to combine the measures of the part A with the measures of the part B, by forming for example the sums $a_1 + b_1$, $a_2 + b_2 \ldots a_n + b_n$. These sums, being composed of parts of equal exactitude, will have the same weight, and will be able to be regarded as so many completed measures of the length L. The definitive value of L will be therefore

$$L = F\{(a_1 + b_1)(a_2 + b_2)\ldots(a_n + b_n)\}.$$

As it is permitted to suppose B as small as one wishes with respect to A, one will be able to develop the second member according to the powers and the products of b_1, $b_2 \ldots b_n$. We will have

$$L = F(a_1, a_2 \ldots a_n) + b_1 \frac{dF}{da_1} + b_2 \frac{dF}{da_2} + b_3 \frac{dF}{da_3} + \cdots + b_n \frac{dF}{da_n} + \text{etc.}; \qquad (3)$$

or $\frac{dF}{da_1}$ $\frac{dF}{da_2}\ldots$ are some functions of a_1, $a_2 \ldots$ only and do not contain b_1, $b_2 \ldots b_n$. In this manner of calculating L nothing indicates which of the values b_1, $b_2 \ldots$ must be combined with a_1 in order to form $a_1 + b_1$; and it is evidently permitted to compose the sums $a_1 + b_1$, $a_2 + b_2 \ldots a_n + b_n$ by combining the different a with the different b in all the possible manners. Whatever be the combination adopted, the value of L must be evidently always the same, as result of the same operations, in which the order alone is changed. This amounts to saying, that it must be permitted to permute the b_1, $b_2 \ldots b_n$ in equation (3) without L changing. This condition will be fulfilled only when the second member of (3) will be a symmetric function of b_1, $b_2 \ldots b_n$. It is necessary and it is sufficient for this object, that one have

$$\frac{dF}{da_1} = \frac{dF}{da_2} = \frac{dF}{da_3} \cdots = \frac{dF}{da_n}. \qquad (4)$$

Equations (1), (2), (4) determine completely the function F. In fact, by substituting (4) into (2) there results

$$\frac{dF}{da_1} = \frac{1}{n}, \frac{dF}{da_2} = \frac{1}{n}, \ldots = \frac{dF}{da_n} = \frac{1}{n} :$$

these values carried into (1) give

$$F = \frac{1}{n}\{a_1 + a_2 + a_3 + \ldots a_n\}$$

this which the concern was to find.

One is able to arrive to equations (4) in a much more simple and more direct manner. In fact, since all the quantities a_1, $a_2 \ldots a_n$ are regarded as being of equal exactitude, if to one of them one attributes the small variation ϵ, the change which results from it in F must be always the same, whatever be the quantity affected with the variation ϵ. For if the effect which results for F by the introduction of ϵ into a_1 was greater than the analogous effect which would result from the introduction of ϵ into a_2, one would conclude from it, that an error of a_1 weighs on the error of the result E in a more considerable manner than an error equal to a_2, or else, that one regards equal of errors a_1 and a_2 as having an unequal importance on the result; or finally, that a_1 and a_2 are not like weights, contrary to the assumption. Now if one adds separately the quantity ϵ to a_1, $a_2 \ldots a_n$, the variations which result from it for F will be respectively

$$\epsilon \frac{dF}{da_1} : \epsilon \frac{dF}{da_2} : \epsilon \frac{dF}{da_3} \ldots \epsilon \frac{dF}{da_n};$$

and since all these variations must be equal, one concludes from it the relations (4).

The manner to consider the question shows, that the arithmetic mean is the only process which permits treating with impartiality a system of observations having all the same weights, without regard to the more or less great accord of each observation with the others. Each other process involves with itself, either a contradiction to one of the equations (1), (2) that is to say an absurdity: or it does not satisfy equations (4), that which reverts in supposing some weights unequal and dependent on the position that each particular result occupies with respect to the others.

It is demonstrated by this, that the arithmetic mean gives the only result plausible and reconcilable with the practical requirements of the question. When the law of probability of the errors is expressed by the exponential function of Gauss, the plausible result is also the most probable result. This is that which happily arrives nearly always. For some other laws of probability the most probable result is different from the arithmetic mean; in this case it is impossible to take account of the equality of the weights of the observed quantities: the method itself assigns to each quantity some weights depending on the position that this quantity occupies in the middle of the analogous others. The condition explicated in number III is no longer fulfilled. The example cited in this same number gives an idea of the singular conclusions and of the inextricable confusion which would derive from the adoption consequent and complete of a law of probability different from that which serves as foundation to the method of least squares.

Observatory of Bréra at Milan, 19 Sept. 1875.