Théoreme sur la probabilité des résultats moyens des observations*

J. Bienaymé

L'Institut, Journal General de Sociétés et travaux scientifiques Sciences Mathématiques, Physiques et Naturelle 1839, pp. 24–26
Unedited extract from the procès-verbaux,

Addition to the session of 4 May 1839

M. J. Bienaymé communicated to the Society a theorem on the probability of the mean results of observations, and in general on the probability of some events.

The author has frequently applied to natural events, to statistical observations, for example, the famous theorem of Jacques Bernoulli, or rather the theorem of Bayes which is the reciprocal of it. He has nearly always found that, despite the magnitude of the number of observations compared, the deviations of many mean results of the same nature deviate from the limits that these two theorems would assign to them with a very great probability. He is likewise assured by the aid of a rather simple formula given by Laplace, that there are strong reasons to think that the hypothesis on which the rules of Bernoulli and Bayes repose, are realized rarely in nature. Thus the annual elements furnished by the judicial statistic, offering a very remarkable stability in their mean values, will satisfy nearly without effort the limits deduced from these rules, and however their deviations are rather great in order that the rule of Laplace give a probability which presumes small variations in the values of the annual possibilities of these elements.

In a crowd of other researches, the deviations of the annual results are so great, that the variation of the possibilities which determine them, would not be doubted. Finally for some the modifications of the possibilities are evident: such are the atmospheric influences on the mean results of a numerous physical phenomena.

It is acceptable thence to seek to represent the effects that were able to produce some causes or some variable possibilities, when the extent of their variations is known, and when one knows that each of the values of the possibility has been able to endure during a certain number of trials or of observations making part of the great number that one has gathered.

The author has been led to these considerations by a formula according to which the extent of the gaps of the mean results of the observations is no longer proportional

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. June 20, 2010 to the radical that Moivre has so ingeniously introduced into this question, by reducing the sum of a certain number of terms of the development of the power of a binomial to a definite integral, and by giving the precise measure of the probability of which Bernoulli had shown as the approximation.

Mr. Bienaymé finds, indeed, that it is still to a radical that the deviations are proportional; but this new radical, instead of containing only a small fraction offering for denominator the total number of observations supposed very great, this new radical contains beyond a second fraction of the same kind, multiplied by the duration of each different possibility, during the course of the observations.

The existence of this multiplier makes imagine without difficulty that if this duration is expressed by a number rather elevated in order to become comparable to the total number of observations, it will increase considerably the extent of the probable deviations. One imagines that it suffices even that each different possibility acts during a rather small number of trials, ten or twelve for example, in order to render the deviations two or three times greater than they would be if each possibility acted many times in sequence.

Finally to clarify that which is just exposed, and that which there remains yet to explain, M. Bienaymé reports that if x is the possibility of an event, the probability that in a total of n trials, this event will be reproduced a number r times such that the mean ratio $\frac{r}{n}$ is contained between the limits:

$$\frac{r}{n} = x \pm c\sqrt{\frac{2x(1-x)}{n}} \tag{1}$$

is expressed by the integral

$$\frac{1}{\sqrt{\pi}}\int_{-c}^{+c} dt\, e^{-t^2} + \frac{e^{-c^2}}{\sqrt{2\pi x(1-x)n}} \tag{2}$$

to the quantities nearly of order of $\frac{1}{n}$.

These formulas suppose the possibility x constant during all the duration of the n trials, and they contain the theorem of Bernoulli. It is well understood that c must be taken in such manner that r is a whole number.

The author reports further that Laplace, in his chapter *des bénefices dépendant de la probabilité des évéments futurs*, has modifed the hypothesis of Bernoulli. He has shown that that which was the possibility which presided at each of the trials, the sum of the awaited benefits remained certain despite this variation, provided that the mean possibility of the arrival of the awaited event was superior to the contrary possibility. One does not see why Laplace has not pursued further this application of the variable possibilities. But by calling

$$Sx \text{ and } Sx^2$$

the sum of the possibilities which have taken place in each trial, during the course of n trials relative to each event, and the sum of their squares; the formulas of Laplace give for the deviations of the mean results the limits

$$\frac{r}{n} = \frac{Sx}{n} \pm c\sqrt{2\frac{Sx - Sx^2}{n^2}} \tag{3}$$

with a probability

$$(4) \qquad \frac{1}{\sqrt{\pi}}\int_{-c}^{+c} dt\, e^{-t^2} + \frac{e^{-c^2}}{\sqrt{2\pi Sx - Sx^2)}}$$

On can recognize that if each possibility remained the same during a number m of trials, under multiple of n, such that $mk = n$, the form of the deviations would become:

$$(3^{bis}) \qquad \frac{r}{n} = \frac{Sx}{k} \pm c\sqrt{2\frac{\frac{Sx}{k} - \frac{Sx^2}{k}}{n}}$$

and the probability is changed into

$$(4^{bis}) \qquad \frac{1}{\sqrt{\pi}}\int_{-c}^{+c} dt\, e^{-t^2} + \frac{e^{-c^2}}{\sqrt{2\pi n\left(\frac{Sx}{k} - \frac{Sx^2}{k}\right))}}$$

These formulas suppose determined the possibility which acts during each series of trials, and the means are relative to the only possibilities which have effect.

It is easy to see that the theorem of Bernoulli is only one particular case of it. The formulas (3) and (4) are reduced to formulas (1) and (2) for x constant, or $k = 1$.

Now if one always regards the number $n = km$ of observations as partitioned likewise into many series containing each m trials; and if instead of determining the value of the possibility which comes to regulate each series, one considers it as arising indifferently from a system of a diverse possibilities x_1, x_2, x_3, etc., of which the mean $\frac{Sx}{a}$, the mean of the squares $\frac{Sx^2}{a}$, are some constants; one will find, by a rigorous calculation, that the deviations of the mean result $\frac{r}{n}$ are contained between the limits:

$$(5) \qquad \frac{r}{n} = \frac{Sx}{a} \pm c\sqrt{2\frac{\frac{Sx}{a} - \left(\frac{Sx}{a}\right)^2}{n} + 2(m-1)\frac{\frac{Sx^2}{a} - \left(\frac{Sx}{a}\right)^2}{n}}$$

with the probability

$$(6) \qquad \frac{1}{\sqrt{\pi}}\int_{-c}^{c} dt\, e^{-t^2} + \frac{e^{-t^2}}{\sqrt{2\pi n\left\{\frac{Sx}{a} - \left(\frac{Sx}{a}\right)^2 + (m-1)\left(\frac{Sx^2}{a} - \left(\frac{Sx}{a}\right)^2\right)\right\}}}$$

Such are the formulas which establish the change which the extent of the probable deviations incur, when each cause or possibility emanated from a constant system is able to act during a series m of trials.

In order to know well this change, it is necessary to not neglect to observe that by making $m = 1$, one obtains the case where each cause acts indifferently at each trial; and where then the deviations and their probability are respectively

$$(7) \qquad \frac{r}{n} = \frac{Sx}{z} \pm c\sqrt{2\frac{\frac{Sx}{a}\left(1 - \frac{Sx}{a}\right)}{n}}$$

$$\text{(8)} \qquad \frac{1}{\sqrt{\pi}}\int_{-c}^{c} dt\, e^{-t^2} + \frac{e^{-t^2}}{\sqrt{2\pi n \frac{Sx}{a}\left(1-\frac{Sx}{a}\right)}}$$

that is to say precisely the same as if the events had only their mean possibility in all the trials. Formulas (7) and (8) are effectively only formulas (1) and (2) in which one replaces x by $\frac{Sx}{a}$.

This is besides that which Jacques Bernoulli understood, and that which it was necessary to deduce from the calculation. This great geometer explained very clearly that what he calls the probability of an event depends on very diverse cases which are able to produce it, and it restores these cases, unequally possible, to some cases equally possible, by substituting many cases of equal possibility for each of the composite cases. He cites diverse examples to apply it on that which he wishes to make understood, and among others the example of the mortality resulting from numerous classes of maladies.

The hypothesis of the diverse and even very numerous causes, but forming a constant set, reproduce therefore always the rules of Bernoulli, when one admits that at each trial, at each event or observed phenomenon, the cause or the possibility, very variable without doubt, has been able to be taken indifferently in the general constant system.

But since one comes to suppose that the possibility that presents indifferently this system, regulates many successive trials, the quantity contained within the radical of Bernoulli increases by a second term

$$2(m-1)\frac{\frac{Sx^2}{a}-\left(\frac{Sx}{a}\right)^2}{n}$$

which extends more or less the limits of the probable deviations, according as the number m or the duration of the cause is more or less great. Now, the events of which we know better the circumstances, offer us precisely certain examples of this duration of the causes.

One can write the limits (5) under the form

$$\frac{r}{n} = \frac{Sx}{a} \pm c\sqrt{2\frac{\frac{Sx}{a}-\left(\frac{Sx}{a}\right)^2}{n} + 2m\frac{\frac{Sx^2}{a}-\left(\frac{Sx}{a}\right)^2}{n}}$$

and as $\frac{m}{n} = \frac{1}{k}$, one sees that the deviations are no longer of the order of $\frac{1}{\sqrt{n}}$, but also of order $\frac{1}{\sqrt{k}}$.

It is to remark that the terms neglected in the expression of the probability of these limits, are equally of the order of $\frac{1}{k}$ in general, instead of the order of $\frac{1}{n}$ that the magnitude of the number n permits to omit. So that, in order to employ these formulas, it will be necessary that k or $\frac{n}{m}$ is yet a great number, at least relative to the questions to resolve.

One concludes immediately from it that the deviations will be very great, and the probability to see the mean results contain themselves there, very small, when the number m will exceed the first numbers of the natural series.

In this case, in order to arrive to some nearly constant mean results, it will be necessary to reunite the numbers of observations quite superior to those which would be sufficient, if the possibilities or the causes (as one will wish) which compose the mean action by virtue of which the events are produced, had not been subject to act many times in sequence: for it will be necessary to render the divisor k a great number, and it is only the quotient of the duration of the observations by the duration of each cause.

On the contrary, if m is a number not comparable to n, the quotient k will remain very great, and the mean results will deviate little from the constant value around which they oscillate. However their deviations will exceed still much those which the law of Bernoulli would assign. It is this which will depend especially then on the difference between the mean of the squares and the square of the mean; a difference which influences besides equally on the results when m is comparable to n.

One perceives thence how the set of causes which rules a class of events, are able to remain exactly in the same conditions, and hence offer the same appearances to the observer, while the mean results of these causes will take some values extremely different from one another, and of the constant mean value of their possibilities.

It will suffice so that this singular effect occurs that the causes, each of them indifferently following one another, are able to have a duration prolonged more or less during some trials. The differences will be able to be very marked, even when the number of causes acting will be very small, provided that the excess of the mean of the squares of the possibilities over the square of the mean is considerable, and that the number m is rather great, without however exceeding much the first numbers.

These considerations must be always presented in the mind of the statistician and of each observer, for one is only too carried to attribute to some new and unforeseen causes, that which is without doubt only the consequence of the possible combinations of the ordinary causes more or less constant.

In the given formulas the possibilities have been supposed to be presented indifferently, that is all gifted with one same number of chances. Nothing is more simple that to attribute the diverse chances to them. One is able also to suppose them in infinite number, and to replace the finite sums by some integrals.

The number m has been regarded as the same for each series influenced by the same possibility. One is able easily to render it variable, either from one partial series to another, independently of the cause which is presented; or in linking it to that cause and by supposing to it a certain probability for each chance of the different values which it could take successively. The formulas become then a little more complicated.

It has seemed, adds Mr. Bienaymé, more convenient to avoid here this complication, finally to facilitate the exposition of the order of ideas in which it is necessary to enter in order to follow the consequences of the hypothesis of the duration of the causes, a length added to the hypothesis of Bernoulli. The author is limited therefore to give the formulas of a very simple case chosen among those that diverse problems of statistic have carried to resolve around six years ago: he will develop later the consequences of these formulas.