Probabilités. Principe nouveau du calcul des probabilités avec ses applications aux sciences d'observation.*

Jules Bienaymé

L'Institut 333, Vol. 8, (1840), pp. 167–169
Soc. Philomat. Paris Extraits, Ser. 5, 18–22.

Session of 25 April 1840

Mr. Jules Bienaymé exposes a principle of probabilities which he believes entirely new, and which appeared to him susceptible to receive some continual applications in the sciences of observations. Here is in what this principle consists.

When one makes a great number of experiences, or when one has gathered a mass of statistical information in order to deduce a certain mean result, one is able to divide them into many groups, either according to the order in which the experiences have been effected, or according to each other consideration particular to these experiences. If one determines the mean results of each of these natural groups, one imagined that they will differ more or less between them, and that they will deviate more or less from the general result. Ordinarily there will be found some deviations so much greater as the groups will be more multiplied. It is easy to see that the extent of the deviations must depend on the observed result; but it seems at first glance that it ought depend equally on the possibility that give to the phenomena in question the cause or the system of causes which regulate them. However there is nothing of it, when this system of causes remains constant during all the duration of the experiences. One demonstrates without difficulty that, in this case, the relations of probability which must exist between the general result and the partial results are absolutely independent of the possibility of the phenomena; there enters into the expressions which characterize them only the results of the observations made, even then as the law of possibility of the phenomena is known in advance.

Thus, for example, if one takes note of the results of 120,000 trials of an ordinary die, one will know in advance that the possibility of bringing forth ace is $\frac{1}{6}$, so that this point should present itself very nearly $\frac{1}{6}$ of 120,000, or 20,000 times. However it will able to arrive that the ace, instead of appearing very nearly one time out of 6, shows itself in 120,000 trials only one time out of 20, that is around 6,000 trials. Ah!

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. June 20, 2010 well, despite this difference between the ratio observed and the real possibility, when one will divide the 120,000 casts of a die in many series, and when one will examine the number of the aces in each of these series, as the cause will be remained constant, one will find that the partial series give nearly some number of aces proportional to the number of the dice contained in the total series; and that so much more exact as the partial series will contain each more trials of the die. If the total ratio of the ace itself is found $\frac{1}{6}$ nearly, out of the 120,000 trials, it will be yet this fraction $\frac{1}{6}$ which will regulate the partial series, and each of them will offer, with certain deviations nearly, around $\frac{1}{6}$ of ace. If, on the contrary, the total ratio itself is found $\frac{1}{20}$, in spite of the constant probability $\frac{1}{6}$ of the face of the die which carries the ace, this will be the fraction $\frac{1}{20}$ which will rule the partial series, and it will be encountered in each of those $\frac{1}{20}$ of aces, always with certain deviations nearly, of which the probability is assignable.

However, as it comes to be said, each partial series will have been formed under the influence of the possibility $\frac{1}{6}$; and it would seem *a priori* that this possibility had to give a great number of series where $\frac{1}{6}$ of aces is found, so that it had furnished the extraordinary ratio $\frac{1}{20}$ out of the set of 120,000 trials that through some quite rare series which would deviate excessively from the ratio of possibility $\frac{1}{6}$. But this presumption would be inexact; and on the contrary, when the set of trials of the dice furnish the ratio $\frac{1}{20}$, so removed from the real ratio of possibility $\frac{1}{6}$, it becomes extremely probable, morally certain (in order to employ the expression of Bernoulli) that the partial series, even very multiplied, will deviate little from the ratio $\frac{1}{20}$.

The demonstration of the principle which assures this result is very simple. We suppose that one has executed a great number c of drawings in an urn containing some white balls and some black balls in a known ratio, such that p is the possibility to bring forth a white ball. We suppose next that there may be exited a white balls and b blacks, and that one divides the total number c of drawings into two series, the first of m and the second of n drawings. One knows that the probability to bring forth in the first series of trials r white balls, of which the possibility is p, is expressed by the term of the development of the power m of the binomial $p + (1 - p)$, in which p has the exponent r, is:

$$\frac{m.m-1.m-2\dots m-r+1}{1.2.3\dots r}p^r(1-p)^{m-r}.$$

Similarly the probability to bring forth q white balls in the second series of trials, will be

$$\frac{n.n-1.n-2\dots n-q+1}{1.2.3\dots q}p^q(1-p)^{n-q}.$$

Hence the concurrence of the two events (r white balls in the first series, and q white balls in the second) will have for probability the product of the two preceding, it is:

$$\frac{m.m-1\dots m-r+1}{1.2.\dots r}\times\frac{n.n-1\dots n-q+1}{1.2\dots q}p^{r+q}(1-p)^{n+m-r-q}$$

Now it agrees with observing that the trials are made, and that there is exited a white balls out of the total $c = m + n$ trials: so that the two numbers of whites r and q, in the two partial series, are subject to the condition $r + q = a$. Each of these numbers are able to vary therefore only from 0 to a: this which renders impossible a great number of cases which were able to arrive in two series of trials. There is since then place to consider that the values of the probability given above by the condition $r+q=a$, and since these values become only possibles, it is necessary to make the sum and to divide the preceding expression by this sum. Now, one recognized without difficulty that the sum of which there is concern is

$$\frac{m+n.m+n-1\ldots m+n-a+1}{1.2\ldots a}p^a(1-p)^b$$

or

$$\frac{c.c-1\ldots c-a+1}{1.2\ldots a}p^a(1-p)^b.$$

The quotient of the probability above with this sum is

$$\frac{\frac{m.m-1\ldots m-r+1}{1.2\ldots r}\times\frac{n.n-1\ldots n-q+1}{1.2\ldots q}p^{r+q}(1-p)^{m+n-r-q}}{\frac{c.c-1\ldots c-a+1}{1.2\ldots a}p^a(1-p)^{c-1}}$$

and one sees that because of $c=m+n$ and $a=r+q$, the possibility p disappeared completely in the quotient.

Thus the probability to find r white balls in the first series, and $q=a-r$ whites in the second, when one partitions into two series a total number c of drawings which has given a white balls is simply

$$\frac{\frac{m.m-1\ldots m-r+1}{1.2\ldots r}\times\frac{n.n-1\ldots n-q+1}{1.2\ldots q}}{\frac{c.c-1\ldots c-a+1}{1.2\ldots a}}$$

an expression in which there remains no more than the results of the drawings, that is of the observed facts.

With a little attention, one recognized in this expression the following:

$$\frac{a.a-1.a-2\ldots a-r+1\times b.b-1.b-2\ldots b-m+r+1}{c.c-1.c-2\ldots c-m+1}\times\frac{m.m-1\ldots m-r+1}{1.2\ldots r}$$

which is the possibility to draw r white balls and $(m-r)$ blacks from an urn containing these balls, of which a whites and b blacks.

The relations of probability between the partial series and the total series of trials or of the experiences are therefore not only independent of the real possibility of the events, but moreover they are the same as if the facts of which a partial series is composed had been drawn at random from the total series of observed facts.

The application of this principle (which extends besides to all the cases of constant probability, whatever be the number of kinds of events of which the result is composed) will be very easy to do.

When it will be important to know the cause, or the system of causes, which has regulated a series of experiences has not at all incurred variation during the duration of these experiences, it will suffice to divide them into partial series, and to calculate if the deviations of the mean results of these subdivisions are contained within the limits that the general mean result assign to them, by virtue of the new principle.

It is much worthy to remark, adds Mr. Bienaymé, that one will be able to conclude by this process, that the cause has been constant or variable, without any prejudice on the real possibility that it is able to give to the phenomena. This conclusion will subsist, even when the mean result would deviate completely from the value of this possibility, and that consequently it would give to the observer an idea totally inexact for this value. This is thence an important consequence, for it is evident from it that the statistic, and in general the sciences of observation, are always able to furnish some positive data out of the constancy of natural laws, independently of the value of these laws.

Here is the formula to use when one divides solely into two series of observations, and when the concern is only of two phenomena the one exclusive from the other, as are the exits of a white ball and the exit of a black ball in a sequence of drawings.

By conserving the letters already employed, one will suppose that there has been observed a phenomena of a certain kind out of a great number c of experiences, and that the contrary phenomenon has hence taken place $c - a = b$ times. If one takes a partial series of m of these observations, under the hypothesis of a constant cause, one must find that the number of phenomena of which there is presented a in mass, is, for the partial series, contained between the limits

$$r = N \pm u\sqrt{2m\frac{ab}{c^2} \cdot \frac{c-m}{c}}$$

(N being the greatest whole number contained in $(m+1)\frac{c+1}{+2}$) with a probability, expressed by

$$\frac{1}{\sqrt{\pi}}\int_{-u}^{u} dt\, e^{-t^2} + \frac{e^{-u^2}}{\sqrt{2\pi\frac{ab(c-m)}{c^3}m}}$$

Most often the division into two series will suffice in order to manifest the constancy or the inconstancy of the cause. For one can note that the preceding limits are very narrow.

Fourier, in his *Recherches statistiques sur Paris*, had counseled to separate the observations into groups, in order to recognize by the deviations of the partial series, if one was able to accord some confidence in the general mean result. But he has given no rule on this subject. Uncertainty subsisted therefore. One had even applied to the examination of the partial results a formula of Laplace, which returns to a problem very different from the actual problem: it is this which expresses the probable deviations of a number m of new trials, when already one has made c experiences which have given a times the awaited phenomenon. The limits of the number r of the repetitions of this phenomenon in these m new trials (and not in m of the c trials already made), are:

$$r = N \pm u\sqrt{2\frac{ab}{c^2}\,\frac{c+m}{c}m}$$

with a probability

$$\frac{1}{\sqrt{\pi}}\int_{-u}^{u} dt\, e^{-t^2} + \frac{e^{-u^2}}{\sqrt{w\pi m\frac{ab(c+m)}{c^3}}}.$$

Here N is the greatest whole number contained in $(m+1)\frac{a}{c}$.

One sees that these limits are greater than those which result from the stated principle. They surpass them in the ratio of $\sqrt{c+m}$ to $\sqrt{c-m}$ (for example of $\sqrt{3}$ to 1 if $m = \frac{1}{2}c$). One was next exposed, in employing them, in regarding as results of a constant cause much too considerable deviations, and which indicated positively the existence of a variable cause.

One understands completely that the limits of the deviations of the numbers which have competed to form a mean result, must be much less than are those of numbers which have not contributed, although both are regulated by the same constant possibility. This provision agrees with the formulas which derive from the new principle. They depend no longer but on the terms of the development of the *binomial of the factorials*: one is able to be assured of it. And the greatest term of this binomial, thus as those which adjoin it, are relatively greater than the corresponding terms of the development of the powers. Hence of the lesser deviations for one same probability.