

SUPPLEMENT AU MÉMOIRE  
SUR LES  
APPROXIMATIONS DES FORMULES

QUI SONT FONCTIONS DE TRÈS GRANDS NOMBRES.

P.S. Laplace

*Mémoires de l'Académie royale des Sciences*, 1st Series, T. X, year 1809; 1810.

I have shown, in article VI of this memoir, that, if we suppose in each observation the positive and negative errors equally facile, the probability that the mean error of a number  $n$  of observations will be contained within the limits  $\pm \frac{rh}{n}$  is equal to

$$\frac{2}{\sqrt{\pi}} \sqrt{\frac{k}{2k'}} \int dr e^{-\frac{k}{2k'} r^2};$$

$h$  is the interval in which the errors of each observation can be extended. If we designate next by  $\phi\left(\frac{x}{h}\right)$  the probability of the error  $\pm x$ ,  $k$  is the integral  $\int dx \phi\left(\frac{x}{h}\right)$  extended from  $x = -\frac{1}{2}h$  to  $x = \frac{1}{2}h$ ;  $k'$  is the integral  $\int \frac{x^2}{h^2} dx \phi\left(\frac{x}{h}\right)$  taken in the same interval;  $\pi$  is the semi-circumference of which the radius is unity, and  $e$  is the number of which the hyperbolic logarithm is unity.

We suppose now that a like element is given by  $n$  observations of a first kind, in which the law of facility of the errors is the same for each observation, and that it is found equal to  $A$  out of a mean among all these observations. We suppose next that it is found equal to  $A + q$  by  $n'$  observations of a second kind, in which the law of facility of the errors is not the same as in the first kind; that it is found equal to  $A + q'$  by  $n''$  observations of a third kind, and thus in sequence. We demand the mean that it is necessary to choose among these diverse results.

If we suppose that  $A + x$  is the true result, the error of the mean result of the  $n$  observations will be  $-x$ , and the probability of this error will be, by that which precedes,

$$\frac{1}{\sqrt{\pi}} \sqrt{\frac{k}{2k'}} \frac{dr}{dx} e^{-\frac{k}{2k'}r^2};$$

we have here

$$x = \frac{rh}{\sqrt{n}},$$

that which transforms the preceding function into this one

$$\frac{1}{\sqrt{\pi}} a \sqrt{n} e^{-na^2x^2},$$

$a$  being equal to  $\frac{1}{h} \sqrt{\frac{k}{2k'}}$ .

The error of the mean result of the  $n'$  observations is  $\pm (q - x)$ , the  $+$  sign having place, if  $q$  surpasses  $x$ , and the  $-$  sign if it is surpassed by it. The probability of this error is

$$\frac{1}{\sqrt{\pi}} a' \sqrt{n'} e^{-n'a'^2(q-x)^2},$$

$a'$  expressing with respect to these observations that which  $a$  expresses relatively to the observations  $n$ .

Likewise the error of the mean result of the observations  $n''$  is  $\pm (q' - x)$ , and the probability of this error is

$$\frac{1}{\sqrt{\pi}} a'' \sqrt{n''} e^{-n''a''^2(q'-x)^2},$$

$a''$  being that which  $a$  becomes relatively to these observations, and thus of the rest.

Now, if we designate generally by  $\Psi(-x)$ ,  $\Psi'(q - x)$ ,  $\Psi''(q' - x)$ , ... these diverse probabilities, the probability that the error of the first result will be  $-x$ , and that the other results will be separated from the first, respectively by  $q$ ,  $q'$ , ..., will be, by the theory of the probabilities, equal to the product  $\Psi(-x)\Psi'(q - x)\Psi''(q' - x)\cdots$ ; therefore, if we construct a curve of which the

ordinate  $y$  is equal to this product, the ordinates of this curve will be proportional to the probabilities of the abscissas, and, for this reason, we will name it *curve of probabilities*.

In order to determine the point of the axis of the abscissas where we must fix the mean among the results of the observations  $n, n', n'', \dots$ , we will observe that this point is the one where the gap from the truth, which we can fear, is a minimum; now, just as, in the theory of probabilities, we evaluate the loss to fear by multiplying each loss that we can experience by its probability, and by making a sum of all these products, just as we will have the value of the gap to fear by multiplying each gap from the truth, or each error, setting aside the sign, by its probability, and by making a sum of all these products. Let therefore  $l$  be the distance from the point which it is necessary to choose to the origin of the curve of the probabilities, and  $z$  the abscissa corresponding to  $y$  and counted from the same origin; the product of each error by its probability, setting aside the sign, will be  $(l - z)y$ , from  $z = 0$  to  $z = l$ , and this product will be  $(z - l)y$  from  $z = l$  to the extremity of the curve. We will have therefore

$$\int (l - z)y dz + \int (z - l)y dz$$

for the sum of all these products, the first integral being taken from  $z$  null to  $z = l$ , and the second taken from  $z = l$  to the last value of  $z$ . By differentiating the preceding sum with respect to  $l$ , it is easy to assure ourselves that we will have

$$dl \int y dz - dl \int y dz$$

for this differential, which must be null in the case of the minimum; we have therefore then

$$\int y dz = \int y dz,$$

that is to say that the area of the curve, contained from  $z$  null to the abscissa which it is necessary to choose, is equal to the area contained from  $z$  equal to this abscissa to the last value of  $z$ ; the ordinate corresponding to the abscissa which it is necessary to choose divides therefore the area of the curve of the

probabilities into two equal parts. [See the *Mémoires de l'Académie des Sciences*, year 1778, page 324.<sup>1</sup>]

Daniel Bernoulli, next Euler and Mr. Gauss have taken for this ordinate the greatest of all. Their result coincides with the preceding when this greater ordinate divides the area of the curve into two equal parts, this which, as we just saw it, holds in the present question; but, in the general case, it seems to me that the manner by which I come to view the thing results from the same theory of probabilities.

In the present case, we have, by making  $x = X + z$ ,

$$y = p p' p'' \dots e^{-p^2 \pi (X+z)^2 - p'^2 \pi (q-X-z)^2 - p'' \pi (q'-X-z)^2 - \dots},$$

$p$  being equal to  $\frac{a\sqrt{n}}{\sqrt{\pi}}$ , and consequently expresses the greatest probability of the result given by the  $n$  observations;  $p'$  expresses likewise the greatest ordinate relative to the  $n'$  observations, and thus of the rest;  $r$  can, without sensible error, be extended from  $-\infty$  to  $+\infty$ , as we have seen in article VII of the memoir cited, we can take  $z$  in the same limits, and then if we choose  $X$  in the manner that the first power of  $z$  disappears from the exponent of  $e$ , the ordinate  $y$  corresponding to  $z$  null will divide the area of the curve into two equal parts, and will be at the same time the greatest ordinate. In effect, we have, in this case,

$$X = \frac{p'^2 q + p''^2 q' + \dots}{p^2 + p'^2 + p''^2 + \dots},$$

and then  $y$  takes this form

$$y = p p' p'' \dots e^{-M - N z^2};$$

whence it follows that the ordinate which corresponds to  $z$  null is the greatest, and divides the entire area of the curve into equal parts. Thus  $A + X$  is the mean result that it is necessary to take among the results  $A$ ,  $A + q$ ,  $A + q'$ ,  $\dots$ . The preceding value of  $X$  is that which renders a minimum the function

$$(pX)^2 + [p'(q - X)]^2 + [p''(q' - X)]^2 + \dots,$$

that is to say the sum of the squares of the errors of each result, multiplied respectively by the greatest ordinate of the curve of the facility of these errors. Thus this property, which is only hypothetical when we consider only the results

---

<sup>1</sup> *Mémoire sur les probabilités*, §XXX.

given by a single observation or by a small number of observations, becomes necessary when the results among which we must take a mean is given each by a very great number of observations, whatever be moreover the law of facility of errors of these observations. This is a reason in order to use it in all cases.

We will have the probability that the error of the result  $A + X$  will be contained within the limits  $\pm Z$ , by taking within these limits the integral  $\int dz e^{-Nz^2}$  and by dividing it by the same integral, taken from  $z = -\infty$  to  $+\infty$ . This last integral is  $\frac{\sqrt{\pi}}{\sqrt{N}}$ ; by making therefore  $z\sqrt{N} = t$  and  $Z\sqrt{N} = T$ , the probability that the error of the chosen result  $A + X$  will be contained within the limits  $\pm \frac{T}{\sqrt{N}}$  will be

$$\frac{2 \int dt e^{-t^2}}{\sqrt{\pi}},$$

the integral being taken from  $t$  null to  $t = T$ . The value of  $N$  is, by that which precedes,

$$\pi(p^2 + p'^2 + p''^2 + \dots).$$