

# Sur la probabilité des erreurs qui affectent des résultats moyens d'observations de même nature

Augustin Cauchy\*

*Comptes Rendus Hebd. Séances Acad. Sci.* 37, 264-272.  
OC I, 12 (527), 104-114. Read 16 August 1853.

§ I. *On the probability of the errors which affect quantities determined by some observations of like nature.*

Let, as in the preceding Memoir:

$k_1, k_2, \dots, k_n$  the quantities furnished by observations of like nature;

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  the errors that they involve;

$l$  any one of the whole numbers  $1, 2, 3, \dots, n$ ;

We suppose, besides, the positive or negative errors equally probable, and let, under this hypothesis:

$-\kappa, \kappa$  be the limits between which the error  $\varepsilon_l$  is certainly contained;

$f(\varepsilon)d\varepsilon$  be the probability of the coincidence of this error with a quantity contained between the infinitely close limits  $\varepsilon, \varepsilon + d\varepsilon$ .

The function  $f(\varepsilon)$ , which we will name the *index of probability of the error  $\varepsilon$* , will be able to be transformed into a definite integral by aid of the formula

$$(1) \quad f(\varepsilon) = \frac{1}{\pi} \int_0^{\infty} \phi(\theta) \cos \theta \varepsilon d\theta,$$

in which one will have

$$(2) \quad \phi(\theta) = 2 \int_0^{\kappa} f(\varepsilon) \cos \theta \varepsilon d\varepsilon.$$

The function  $\phi(\theta)$ , which formula (2) determines, is therefore linked to the probability  $f(\varepsilon)$ , in such a way that, one of these functions being given, the other is deduced from it. Besides, if, in formula (2), one puts  $\theta = 0$ , one will have

$$(3) \quad \phi(0) = 1,$$

---

\*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. July 7, 2010

or, that which reverts to the same,

$$(4) \quad 2 \int_0^{\kappa} f(\varepsilon) \cos \theta \varepsilon d\varepsilon = 1.$$

The function  $\phi(\theta)$  being supposed known, one is able without difficulty to deduce, not only the value of  $f(\varepsilon)$ , that is to say the index of probability of the error  $\varepsilon$ , but also the probability  $p$  of the coincidence of the error  $\varepsilon_i$  with a quantity contained between the limits  $-\varepsilon, \varepsilon$ . In fact, this last [265] probability will be evidently represented by the integral

$$\int_{-\varepsilon}^{\varepsilon} f(\theta) d\theta = 2 \int_0^{\varepsilon} f(\theta) d\theta,$$

or, that which reverts to the same, by the double of the integral

$$\int f(\varepsilon) d\varepsilon,$$

taken starting from  $\varepsilon = 0$ . Besides, by virtue of formula (1), this last integral will be equivalent to

$$\frac{1}{\pi} \int_0^{\infty} \phi(\theta) \frac{\sin \theta \varepsilon}{\theta} d\theta.$$

One will have therefore

$$(5) \quad p = \frac{2}{\pi} \int_0^{\infty} \phi(\theta) \frac{\sin \theta \varepsilon}{\theta} d\theta.$$

Let now

$$(6) \quad \omega = \lambda_1 \varepsilon_1 + \lambda_2 \varepsilon_2 + \dots + \lambda_n \varepsilon_n$$

be a linear function of the errors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ . The probability  $P$  of the coincidence of the error  $\omega$  with a quantity contained between the limits  $-v, v$  will be furnished by an equation analogous to formula (5), namely, by that which one deduces from it by replacing the limit  $\varepsilon$  by the limit  $v$ , and the function  $\phi(\theta)$  by the function

$$(7) \quad \Phi(\theta) = \phi(\lambda_1 \theta) \phi(\lambda_2 \theta) \dots \phi(\lambda_n \theta).$$

One will have, in consequence,

$$(8) \quad P = \frac{2}{\pi} \int_0^{\infty} \Phi(\theta) \frac{\sin \theta v}{\theta} d\theta.$$

In other terms, one will have

$$(9) \quad P = \int_0^v F(\tau) d\tau,$$

the form of the function that the letter  $F$  indicates being determined by the equation

$$(10) \quad F(v) = \frac{2}{\pi} \int_0^{\infty} \Phi(\theta) \cos \theta v d\theta.$$

This put, the product  $F(v)dv$  will represent the probability of the coincidence of the error  $\omega$  with a quantity contained between the infinitely near limits  $v, v + dv$ , and the first factor of this product or the function  $F(v)$ , will be [266] that which we name the *index of probability of error*  $v$ , considered as a particular value of  $\omega$ . The index of probability of a null value of  $\omega$  will be therefore

$$(11) \quad F(0) = \frac{2}{\pi} \int_0^{\infty} \Phi(\theta) d\theta.$$

In the particular case where the function  $\omega$  is reduced to the arithmetic mean between the errors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  and where one has, hence

$$(12) \quad \omega = \frac{\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n}{n},$$

formula (7) gives simply

$$(13) \quad \Phi(\theta) = \left[ \phi \left( \frac{\theta}{n} \right) \right]^n.$$

Formulas (7), (8), (9) and (10) make the quantities  $F(v)$  and  $P$  to depend on the function  $\phi(\theta)$  determined itself by formula (2). Besides, from this last formula one is able to deduce many others which are able to be substituted for it more or less usefully, according as one attributes to the positive variable  $\theta$  some values more or less great.

We remark first that one draws from equation (2), by having regard to the formula  $\cos x = 1 - 2 \sin^2 \frac{x}{2}$ ,

$$(14) \quad \phi(\theta) = 1 - \int_0^{\kappa} \left( 2 \sin^2 \frac{\theta \varepsilon}{2} \right)^2 f(\varepsilon) d\varepsilon,$$

and by integrating by parts,

$$(15) \quad \phi(\theta) = 2 \frac{f(\kappa) \sin \theta \kappa - \int_0^{\kappa} f'(\varepsilon) \sin \theta \varepsilon d\varepsilon}{\theta}$$

On the other hand, if one names  $x$  a positive variable and  $\chi(x)$  a function which, vanishing for  $x = 0$ , remains continuous, with its derivative  $\chi'(x)$ , for the values of  $x$  inferior to a certain limit, one will have, as one knows, for some such values of  $x$ ,

$$\chi(x) = x \chi'(\eta x),$$

$\eta$  designating a number inferior to unity. There results from it, for example, that, for the values of  $x$  very small,  $\sin x$  is the product of  $x$  by a factor contained between the limits 1,  $\cos x$ ; that similarly,  $\ln(1 - x)$  is the product [267] of  $-x$  by a factor contained between the limits 1,  $\frac{1}{1-x}$ , and that, hence, by naming  $\rho$  one such factor, one has

$$1 - x = e^{-\rho x}.$$

This put, if one makes, for brevity,

$$(16) \quad c = \int_0^{\kappa} \varepsilon^2 f(\varepsilon) d\varepsilon,$$

and if, besides, one attributes to the positive variable  $\theta$  a value small enough in order that the product  $\theta\kappa$  be itself very small, one will see the value of  $\phi(\theta)$  given by formula (15) is reduced sensibly to the exponential  $e^{-c\theta^2}$ , and one will conclude from this formula that one has in total rigor

$$(17) \quad \phi(\theta) = e^{-\zeta\theta^2},$$

$\zeta$  being the product of the constant  $c$  by a factor contained between the limits

$$(18) \quad \cos^2 \frac{\theta\kappa}{2}, \quad \frac{1}{1 - \int_0^\kappa (2 \sin \frac{\theta\varepsilon}{2})^2 f(\varepsilon) d\varepsilon},$$

and with a stronger reason between the limits

$$(19) \quad 1 - \frac{1}{2} \left( \frac{\theta\kappa}{2} \right)^2, \quad \frac{1}{1 - c\theta^2\kappa^2}.$$

Formula (17) permits obtaining easily a value very near the function  $\phi(\theta)$ , in the case where  $\theta$  and  $\phi\kappa$  are very small.

If, on the contrary, one attributes to  $\theta$  a value which is not very small, then, the value of  $\zeta$  being no longer very near to the constant  $c$ , formula (17) must be abandoned. But then, especially if  $\theta$  becomes very great, one will be able usefully to recur to formula (15). We consider, in order to fix the ideas, the case where the function  $f(\varepsilon)$  decreases constantly, while the variable  $\varepsilon$ , supposed positive, increases starting from zero. In this case,  $f'(\varepsilon)$  being negative, formula (15) will furnish immediately a limit superior to  $\phi(\theta)$  and will give

$$(20) \quad \phi(\theta) < \frac{2f(0)}{\theta}.$$

In order to show an application of the formulas that we just obtained, we apply them to the determination of the quantity  $F(v)$ , or, that which reverts [268] to the same, to the determination of the integral

$$(21) \quad \int_0^\infty \Phi(\theta) \cos \theta v d\theta,$$

in the case where,  $n$  being a very great number, the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$  are some very small quantities of order  $\frac{1}{n}$ . Let besides  $\Theta$  be a number which is itself very great, but such, that that the products  $\lambda_1\Theta, \lambda_2\Theta, \dots, \lambda_n\Theta$  remain very small. The integral (21) will be the sum of the two integrals

$$(22) \quad \int_0^\Theta \Phi(\theta) \cos \theta v d\theta,$$

$$(23) \quad \int_\Theta^\infty \Phi(\theta) \cos \theta v d\theta.$$

Besides, by virtue of formulas (7) and (17), one will have, for the values of  $\theta$  inferior to  $\Theta$ ,

$$(24) \quad \Phi(\theta) = e^{-s\theta^2},$$

the value of  $s$  being given by an equation of the form

$$s = \zeta_1 \lambda_1^2 + \zeta_2 \lambda_2^2 + \cdots + \zeta_n \lambda_n^2,$$

and the factors  $\zeta_1, \zeta_2, \dots, \zeta_n$  being very near the constant  $c$ . There is more: one will have also

$$(25) \quad s = \zeta \Lambda, \quad \Lambda = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2,$$

$\zeta$  being a quantity contained between the smallest and the largest of the numbers  $\zeta_1, \zeta_2, \dots, \zeta_n$ , consequently, a quantity which will be itself very near to  $c$ . This put, the integral (22) will become

$$(26) \quad \int_0^\Theta e^{-s\theta^2} \cos \theta v d\theta.$$

Now this last will differ very little from the integral

$$(27) \quad \int_0^\infty e^{-s\theta^2} \cos \theta v d\theta,$$

if the product  $s\Theta^2$  is a very great number, this which will happen if  $\Theta^2$  is of an order superior to the order of  $\frac{1}{s}$ , that is to say to the order of  $n$ , or, in other [269] terms, if  $\Theta$  is of an order superior to the order of  $\sqrt{n}$ . Therefore also, under this condition, the integral (22) will be reduced obviously to a product of the form

$$(28) \quad \frac{1}{2} \left( \frac{\pi}{s} \right)^{\frac{1}{2}} e^{-\frac{v^2}{4s}},$$

the value of  $s$  being

$$(29) \quad s = c\Lambda.$$

On the other hand, one will have, by virtue of formula (20),

$$(30) \quad \Phi(\theta) < \frac{[2f(0)]^n}{\lambda_1 \lambda_2 \cdots \lambda_n} \frac{1}{\theta^n},$$

and, hence, the integral (23) will be inferior to the product

$$(31) \quad \frac{[2f(0)]^n}{\lambda_1 \lambda_2 \cdots \lambda_n} \frac{1}{(n+1)\Theta^{n+1}}.$$

Therefore, if this last is able to be neglected vis-a-vis expression (28), that which will happen for example when the quantity  $2f(0)$  will be inferior to each of the products

$\lambda_1\Theta, \lambda_2\Theta, \dots, \lambda_n\Theta$ , the integral (21) will be reduced obviously to the expression (28), and one will have, very nearly,

$$(32) \quad F(v) = \frac{1}{\sqrt{\pi s}} e^{-\frac{v^2}{4s}}.$$

Then also formula (9) will give sensibly

$$(33) \quad P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{v}{2\sqrt{s}}} e^{-\theta^2} d\theta.$$

§ II. — *On the probability of the errors which affect the mean results.*

We suppose that, the  $m$  unknowns  $x, y, \dots, v, w$  being determined by  $n$  approximate linear equations, one deduces from these equations multiplied by certain factors  $\lambda_1, \lambda_2, \dots, \lambda_n$ , next added among them, the final equation which furnishes immediately the value of the unknown  $x$ . This value will be of the form

$$(1) \quad x = \lambda_1 k_1 + \lambda_2 k_2 + \dots + \lambda_n k_n,$$

the first of the equations of condition to which the factors [270]  $\lambda_1, \lambda_2, \dots, \lambda_n$  satisfy being itself of the form

$$(2) \quad a_1 \lambda_1 + a_2 \lambda_2 + \dots + a_n \lambda_n = 1,$$

and if one names  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  the errors which involve the quantities  $k_1, k_2, \dots, k_n$ , the error  $\xi$  of the preceding value of  $x$  will be

$$(3) \quad \xi = \lambda_1 \varepsilon_1 + \lambda_2 \varepsilon_2 + \dots + \lambda_n \varepsilon_n.$$

Finally, if, of the positive and negative errors being equally probable, one supposes the error  $\varepsilon_i$  certainly contained between the limits  $-\kappa, \kappa$ , the probability  $P$  of the coincidence of the error  $\xi$  with a quantity contained between the limits  $-v, v$ , and the index of probability  $F(v)$  of the error  $v$  in the value of the unknown  $x$ , will be determined by formulas (8) and (10) of the first paragraph.

It is important to observe that one draws from formulas (1) and (3), joined to condition (2),

$$(4) \quad x = \frac{\lambda_1 k_1 + \lambda_2 k_2 + \dots + \lambda_n k_n}{a_1 \lambda_1 + a_2 \lambda_2 + \dots + a_n \lambda_n},$$

$$(5) \quad \xi = \frac{\lambda_1 \varepsilon_1 + \lambda_2 \varepsilon_2 + \dots + \lambda_n \varepsilon_n}{a_1 \lambda_1 + a_2 \lambda_2 + \dots + a_n \lambda_n}.$$

These last values of  $x$  and of  $\xi$  depend uniquely on the ratios between the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$ , and are also those which one would obtain if one ceased to subject these factors to condition (2). We admit this last hypothesis, and we imagine that, the signs of the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$  remaining arbitrary, one assigns to these same factors

the determined numerical values. Let, besides,  $\lambda$  be the arithmetic mean among the numerical values, and we name  $A$  the greatest numerical value that the sum  $a_1\lambda_1 + a_2\lambda_2 + \dots + a_n\lambda_n$  is able to acquire. The greatest of the numerical values that the error  $\xi$  will be able to take will be the smallest possible, and precisely equal to the ratio

$$(6) \quad \frac{n\lambda\kappa}{A},$$

when the signs of the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$  will be chosen in a manner that one has

$$(7) \quad a_1\lambda_1 + a_2\lambda_2 + \dots + a_n\lambda_n = A.$$

Besides, being given the coefficients  $a_1, a_2, \dots, a_n$  and the numerical values of the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$ , one will know the numerical value [271] of each of the products

$$(8) \quad a_1\lambda_1, \quad a_2\lambda_2, \dots, \quad a_n\lambda_n;$$

and the quantity  $A$ , determined by equation (7), will be the greatest possible when all these products will be positive, that is to say, in other terms, when the signs of the factors

$$\lambda_1, \quad \lambda_2, \dots, \quad \lambda_n$$

will be those of the quantities

$$a_1, \quad a_2, \dots, \quad a_n,$$

which represent the coefficients of the unknown  $x$  in the given linear equations. Therefore a system of factors  $\lambda_1, \lambda_2, \dots, \lambda_n$  which will satisfy this condition, being compared to the systems that one is able to deduce by changing the signs of one or of many of these factors, will be precisely the system for which the greatest error to fear in the value of the unknown  $x$ , will become the smallest possible. It will arrive also, in the research of the final equation which will determine the unknown  $x$ , to attribute to the factor  $\lambda_l$ , by which one will multiply a linear equation, the sign which will affect, in this equation, the coefficient  $a_l$  of  $x$ .

We imagine now that, the products (8) being entirely positive, one may make the numerical values of the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$  vary, by supposing, as one is able to do it, these factors subject to the condition (2). The values of  $x$  and  $\xi$  given by formulas (1) and (3) will vary, and the most probable value  $x$  of the unknown  $x$  will be that for which the probability  $P$  will become the greatest possible. Besides, in order to determine the value  $x$  of the unknown  $x$  with the corresponding values of the factors  $\lambda_1, \lambda_2, \dots, \lambda_n$ , it will suffice, in general, to recur to the condition

$$(9) \quad \delta P = 0.$$

The quantity  $x$  thus determined, that is to say the most probable value of the unknown  $x$ , will be independent of the limit  $\nu$  below which one wishes to lower the error  $\xi$  of this unknown, if the function  $f(\varepsilon)$  is of the form

$$(10) \quad f(\varepsilon) = \frac{1}{\pi} \int_0^\infty e^{-c\theta^N} \cos \theta \varepsilon d\theta,$$

the letters  $c, N$  designating two positive constants, and if, on the other hand, one [272] assigns to the errors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  no limit; in a way that one is able to put  $x = \infty$ . Then one will have

$$(11) \quad \phi(\theta) = e^{-c\theta^N},$$

$$(12) \quad \Phi(\theta) = e^{-s\theta^N},$$

the value of  $s$  being determined by the formulas

$$(13) \quad s = c\Lambda, \quad \Lambda = \lambda_1^N + \lambda_2^N + \dots + \lambda_n^N,$$

Thus also the index of probability  $F(0)$  of a null error in the value of  $\xi$  will be given by the equation

$$(14) \quad F(0) = \frac{2}{\pi} \Gamma\left(1 + \frac{1}{N}\right) s^{-\frac{1}{N}}.$$

The most probable value  $x$  of the unknown  $x$  is, by virtue of formula (12), and when one supposes  $N = 2$ , that which the method of least squares gives. But the same formula leads to other values of  $x$  when  $N$  is superior to 2. Therefore the most probable value  $x$  of the unknown  $x$  is able to differ sensibly from that which the method of least squares furnishes.

This method has at least the property to furnish the most probable value of  $x$ , in the case where, the limit  $\kappa$  being a finite quantity, the number of observations become very considerable. In order to clarify this question, it is proper to examine especially the case of which there is question. This is that which I will do in the next article.

**Mr. Cauchy** presents also to the Academy a *Mémoire sur la probabilité des erreurs qui affectent les résultats moyens d'un grand nombre d'observations.*