MÉMOIRE SUR LA PROPORTION DES NAISSANCES DES FILLES ET DES GARÇONS*

M. Poisson

Mém. Acad. Sci. IX (1830) pp. 239-308
Read to the Academy, 8 February 1829

In considering the births of the two sexes during six consecutive years, either in France entire, or in the part more southern of the realm, I have remarked, it is already four years ago:

1° That the ratio of the births of boys and girls is $\frac{16}{15}$, instead of $\frac{22}{21}$, as one believed it before.

2° That this ratio is very nearly the same for the middle of France and for France entire, so that it appeared independent of the variation of climate, at least in the extent of our country.

3° That its value, among the infants born outside of marriage, is sensibly less than for legitimate infants, and nearly equal to $\frac{21}{20}$.

These results have been inserted for the first time in the *Annuaire* of the bureau of longitudes of the year 1825; and from this period, one has verified them on some numbers more and more considerable. Here is the ratio of the births of boys and of girls, for each of the six years running from 1817 to 1826 and for France entire, by having regard to all the births, legitimate or outside of marriage:

for	1817	....	1.0720,
	1818	....	1.0644,
	1819	....	1.0642,
	1820	....	1.0642,
	1821	....	1.0685,
	1822	....	1.0623,
	1823	....	1.0621,
	1824	....	1.0659,
	1825	....	1.0703,
	1826	....	1.0614.
		Mean	1.0656.

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. January 18, 2010

Its mean value results from around ten million births. It differs by very nearly two hundredths, on both sides, from the extreme values which correspond to 1817 and 1826. The mean value of this same ratio for the thirty departments most south of the realm, is 1.0655, which deviates only by one thousandth of that which holds for France entire. But if one considers in isolation the annual births of each of the 86 departments, one finds that the ratio of which there is concern varies in the same year, from one department to another, and for one same department, from one year to another, in such a way that there happens sometimes that the births of the girls have surpassed those of the boys.

Relative to the natural infants, their number is raised to nearly seven hundred thousand for France entire during the ten years that we consider; and in this number, the ratio of the male births to the female births has been 1.0484. The fraction 0.0172 by which this quantity deviates from the general ratio 1.0656, in not small enough, and the numbers employed are too great, in order that one is able to attribute this difference to chance; and howsoever singular that it appears, one is grounded to believe that there exists, in regard to natural infants, any cause whatsoever which diminishes the preponderance of the births of boys over those of girls. This influence is impressed even on the annual births, so that one is able to be assured of it by calculating for each of the ten years comprehended from 1817 to 1826, the ratio of the births of the two sexes, a defect made of the natural infants. One finds then:

for	1817	....	1.0743,
	1818	....	1.0644,
	1819	....	1.0650,
	1820	....	1.0656,
	1821	....	1.0699,
	1822	....	1.0628,
	1823	....	1.0629,
	1824	....	1.0680,
	1825	....	1.0727,
	1826	....	1.0659.
		Mean	1.0671.

and if one compares the preceding ratios to these new values, one sees that excepting the one which corresponds to 1818 and which has not changed, all the others have a slight increase.

The proportion of births of the two sexes is no longer the same at Paris and in the departments, either among the legitimate infants, or among those which are born outside marriage. During the thirteen years elapsed from 1815 to 1827, there is born at Paris around 215,000 legitimate infants, and the ratio of the male births to the female births has been 1.0408, or nearly $\frac{26}{25}$, instead of $\frac{16}{15}$ which corresponds to France entire. There is born in that city and in the same interval of time, very nearly 122,000 natural infants, among which the ratio of the number of boys to the one of girls, has been 1.0345, or around $\frac{30}{29}$, instead of $\frac{21}{20}$ which holds for that class of infants in the rest of France. It is therefore presumable that there exists also in a great capitol as Paris, a particular cause which diminishes the preponderance of male births, and which drives at the same time on the legitimate infants and on the natural infants.

Our mind is naturally carried to admit the results of experience with so much more confidence as they are deduced from a greater number of observations; but if we wish to increase the probability and to understand that of our future reproduction, we are obliged to recur to the formulas which mathematical analysis furnishes for this object: the perfection of these methods in general, and their application to the facts that I just cited, are the object of the Memoir that I present today to the Academy. If I have added something to the numerous works of geometers who have occupied themselves on the calculation of chances, since Pascal has given the first examples of it, I owe it to the analysis that I have employed, and of which I have pushed the principle in the *Théorie analytique des probabilités*; a work so eminently remarkable by the variety of questions which are treated, as by the generality of the methods that Laplace has imagined in order to resolve them.

§ I.
The probability of the repetition of an event of which the chance is given.

(1) Let p be the probability of an event A, and q that of the contrary event B, so that one has $p+q=1$. We designate by P the probability that on a number n of trials, during which p and q are invariable, A will arrive a number x times and consequently B a number $n-x$. This probability is equal to p^xq^{n-x} for each of the different combinations of which the n trials are susceptible taken x by x, or $n-x$ by $n-x$; one will have therefore the value of P by multiplying p^xq^{n-x} by the number of these combinations; this which gives

$$\mathrm{P} = \frac{1.2.3\ldots n}{12.3\ldots x.1.2.3\ldots n-x} p^x q^{n-x}.$$

But when x and $n-x$, in the same way their sum n, are very great numbers, the calculation of this formula becomes impractical, and one is obliged to recur to the methods of approximation in order to obtain the value.

According to a known formula, one has then

$$\begin{aligned}
1.2.3\ldots n &= n^n e^{-n}\sqrt{2\pi n}\left(1+\tfrac{1}{12n}+\tfrac{1}{288n^2}+\text{etc.}\right),\\
1.2.3\ldots x &= x^x e^{-x}\sqrt{2\pi x}\left(1+\tfrac{1}{12x}+\tfrac{1}{288x^2}+\text{etc.}\right),\\
1.2.3\ldots n-x &= (n-x)^{n-x} e^{-n+x}\sqrt{2\pi(n-x)}\left(1+\tfrac{1}{12(n-x)}\right.\\
&\qquad\left.+\tfrac{1}{288(n-x)^2}+\text{etc.}\right);
\end{aligned}$$

e designating the base of the Naperian logarithms, and π the ratio of the circumference to the diameter, this which will hold in all this Memoir. The series contained between the parentheses are so much more convergent as the numbers n, x, $n-x$, are greater; by conserving only the first term of each series, one will conclude from it

$$\mathrm{P} = \left(\frac{pn}{x}\right)^x \left(\frac{qn}{n-x}\right)^{n-x} \sqrt{\frac{n}{2\pi x(n-x)}}, \tag{1}$$

for the approximate value of P which will be useful to us thereafter.

(2) We designate now by X the probability that A will not arrive more than x times out of the number n trials, and we call C this composite event. It will take place in the $x+1$ following ways:

1° If the $n-x$ first trials bring forth B; because then there will remain no more than x trials which would not be able to bring forth A more than x times. The probability of this last case will be q^m, by making $n-x=m$.

2° If the first $m+1$ trials bring forth B m times and A once, without that A occupy the last place, a necessary condition in order that this second case not return into the first. It is evident that then the $x-1$ following trials may bring forth A only $x-1$ times more, this event will not arrive more than x times in the n trials. The probability of m events B and of an event A which would occupy a determined rank, is $q^m p$, and this rank may be able to be the first m, the probability of the second case favorable to C will be $mq^m p$.

3° If the first $m+2$ trials bring forth B m times and A twice, without that A occupy the rank $m+2$, this which is necessary and sufficient in order that that this third case neither returns to the first, nor to the second. The probability of B m times and A two times in some determined ranks, is $q^m p^2$; by taking two by two the first $m+1$ ranks in order to place A, one has $\frac{1}{2}m(m+1)$ different combinations; the probability of the third case favorable to C will be therefore $\frac{1}{2}m(m+1)q^m p^2$.

By continuing thus, one will arrive finally to an $x+1^{\text{st}}$ case, in which the $m+x$, or n proofs, will bring forth B m times and A x times, without that A occupy the n^{th} rank, so that this case not return into any of the x preceding; and its probability will be

$$\frac{m(m+1)(m+2)\cdots(m+x-1)}{1.2.3\ldots n}q^m p^x.$$

These $x+1$ cases being distinct from one another, and presenting all the different ways by which the event C is able to happen, its probability X will be the sum of their respective probabilities. By putting back $n-x$ in the place of m, one will have consequently

$$\begin{aligned} \mathrm{X} = q^{n-x}\Bigg[&1+(n-x)p+\frac{(n-x)(n-x+1)}{1.2}p^2+\cdots \\ &\cdots+\frac{(n-x)(n-x+1)\cdots(n-2)(n-1)}{1.2.3\ldots x}p^x\Bigg]. \end{aligned} \tag{2}$$

This probability is expressed also, as one knows, by the sum of the $x+1$ first terms of the development of $(q+p)^n$, that is to say that one has equally

$$\begin{aligned} \mathrm{X} = q^n + nq^{n-1}p + \frac{n(n-1)}{2}q^{n-2}p^2+\cdots \\ \cdots+\frac{n(n-1)(n-2)\cdots(n-x+1)}{1.2.3\ldots x}q^{n-x}p^x. \end{aligned}$$

The numerical calculation of both of these equivalent expressions, is able to be regarded as impossible when x and $n-x$ are very great numbers. It is then preferable to employ formula (2), because it is transformed immediately into a definite integral which is reduced next to a very convergent series.

(3) By integrating $x+1$ times in succession by parts, and designating by c an arbitrary constant, there comes

$$n\int \frac{y^x dy}{(1+y)^{n+1}} = c - \frac{y^x}{(1+y)^n} - \frac{x}{n-1}\frac{y^{x-1}}{(1+y)^{n-1}} - \frac{x.x-1}{n-1.n-2}\frac{y^{x-2}}{(1+y)^{n-2}} \cdots - \frac{x.x-1\ldots 2.1}{n-1.n-2\ldots n-x}\frac{1}{(1+y)^{n-x}}$$

As one has $n > x$, all the terms of this formula, c excepted, vanish for $y=\infty$; if therefore one designates by α any positive quantity whatsoever, or zero, one will have

$$n\int_\alpha^\infty \frac{y^x dy}{(1+y)^{n+1}} = c - \frac{\alpha^x}{(1+\alpha)^n} - \frac{x}{n-1}\frac{\alpha^{x-1}}{(1+\alpha)^{n-1}} - \frac{x.x-1}{n-1.n-2}\frac{\alpha^{x-2}}{(1+\alpha)^{n-2}} \cdots - \frac{x.x-1\ldots 2.1}{n-1.n-2\ldots n-x}\frac{1}{(1+\alpha)^{n-x}}$$

In the case $\alpha = 0$, this equation is reduced to

$$n\int_\alpha^\infty \frac{y^x dy}{(1+y)^{n+1}} = \frac{x.x-1\ldots 1}{n-1.n-2\ldots n-x}$$

by dividing the preceding equation by that here, and making, for brevity,

$$\frac{y^x}{(1+y)^{n+1}} = \mathrm{Y},$$

one concludes from it

$$\frac{\int_\alpha^\infty \mathrm{Y}dy}{\int_0^\infty \mathrm{Y}dy} = \frac{1}{(1+\alpha)^{n-x}}\left[1+(n-x)\frac{\alpha}{(1+\alpha)} + \frac{(n-x)(n-x+1)}{(1+\alpha)^2} + \cdots + \frac{(n-x)(x-x+1)\cdots(n-1)}{1.2.3\ldots x}\frac{\alpha^x}{(1+\alpha)^x}\right]$$

now, if one takes $\alpha = \frac{p}{q}$, and if one observes that p+q= 1, the second member of this last equation coincides with formula (2): for this value of α, we will have therefore

$$\mathrm{X} = \frac{\int_\alpha^\infty \mathrm{Y}dy}{\int_0^\infty \mathrm{Y}dy}. \tag{3}$$

(4) In order to reduce into series, the integrals contained in this new expression of X, I call h the value of y which renders Y a *maximum*: by equating dY to zero, one has

$$x(1+h) - (n+1)h = 0;$$

and if one calls H the corresponding value of Y, one will have

$$h = \frac{x}{n+1-x}, \quad \mathrm{H} = \frac{x^x(n+1-x)^{n+1-x}}{(n+1)^{n+1}}.$$

The function Y is infinite for no positive value of y, it is null for $y = 0$ and for $y = \infty$, and has only a single *maximum* between these two limits; one will have therefore to put

$$\mathrm{Y} = \mathrm{H}e^{-t^2}; \tag{4}$$

t being a new variable that one will make increase from $-\infty$ to $+\infty$, and of which the particular values $t = -\infty$, $t = 0$, $t = \infty$, correspond respectively to $y = 0$, $y = h$, $y = \infty$. One will have next

$$\log \mathrm{Y} = \log \mathrm{H} - t^2.$$

By making $y = h + y'$, and developing according to the powers of y', there will result from it

$$t^2 + \frac{1}{2}\frac{d^2 \log \mathrm{Y}}{dy^2} y'^2 + \frac{1}{6}\frac{d^3 \log \mathrm{Y}}{dy^3} y'^3 + \text{etc.} = 0,$$

where one will make $y = h$ after the differentiations, this which will render null $\frac{d\mathrm{Y}}{dy}$. The value of y' that one will draw from this equation will be able to be represented by a series of the form:

$$y' = h't + h''t^2 + h'''t^3 + \text{etc.};$$

$h', h'', h'''.$ etc., being some coefficients independent of t, that one will determine the ones by means of the others by substituting this value into the preceding equation, and equating next to zero the sum of the coefficients of each power of t in its first member. One will have in this manner

$$\begin{aligned}
&1 + \frac{1}{2}\frac{d^2 \log \mathrm{Y}}{dy^2} h'^2 = 0,\\
&\frac{d^2 \log \mathrm{Y}}{dy^2} h'h'' + \frac{1}{6}\frac{d^3 \log \mathrm{Y}}{dy^3} h'^3 = 0,\\
&\text{etc.};
\end{aligned}$$

and by having regard to the value of h, one deduces from it

$$\begin{aligned}
h' &= \sqrt{\frac{2(n+1)x}{(n+1-x)^3}},\\
h'' &= \frac{2(n+1+x)}{3(n+1-x)^2},\\
&\text{etc.}
\end{aligned}$$

If the numbers x, $n-x$, n, are very great, and of the same order of magnitude, it is easy to see that the values of h', h'', h''', etc., will form a very rapidly decreasing sequence of which the first term will be of the same order of smallness as the fraction $\frac{1}{\sqrt{n}}$, the second of the order $\frac{1}{n}$, the third of order $\frac{1}{n\sqrt{n}}$, and thus in sequence; this which will be able to excuse from forming those values above from the first two that we just gave.

By designating by i any whole and positive number, one will have

$$\int_{-\infty}^{\infty} e^{-t^2} t^{2i+1} dt = 0,$$
$$\int_{-\infty}^{\infty} e^{-t^2} t^{2i} dt = \frac{1.3.5\ldots 2i-1}{2^i} \int_{-\infty}^{\infty} e^{-t^2} dt.$$

Therefore, because of

$$\frac{dy'}{dt} = h' + 2h''t + 3h'''t^2 + \text{etc.},$$

and by observing that

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}, \quad \int_{-\infty}^{\infty} \mathrm{Y} dy = \mathrm{H} \int_{-\infty}^{\infty} e^{-t^2} \frac{dy'}{dt} dt,$$

we will have

$$\int_{-\infty}^{\infty} \mathrm{Y} dy = \mathrm{H}\sqrt{\pi} \left(h' + \frac{1.3}{2} h''' + \frac{1.3.5}{4} h^{v} + \text{etc.} \right). \tag{5}$$

Under the hypothesis which renders the quantities h', h''', h^v, etc., very rapidly decreasing, the series contained between the parentheses will be very convergent, at least in the first terms, this which will suffice in order to calculate by means of this last formula, the approximate value of $\int_{-\infty}^{\infty} \mathrm{Y} dy$. It is to Laplace that the analysis is indebted of this method in order to reduce the integrals into convergent series, when the quantities submitted to integration are affected of very great exponents.

(5) The expression of the other integral $\int_{\alpha}^{\infty} \mathrm{Y} dy$ will be different according as the limit α or $\frac{p}{q}$ will surpass or will be less than the value h of y which corresponds to the *maximum* of Y. If one makes $y = \alpha = \frac{p}{q}$ in equation (4), and if one puts for Y and H their values, one will conclude from it

$$e^{-t^2} = \left(\frac{p(n+1)}{x} \right)^x \left(\frac{q(n+1)}{n+1-x} \right)^{n+1-x};$$

whence one draws $t \pm k$, by making, for brevity,

$$k^2 = x \log \frac{x}{p(n+1)} + (n+1-x) \log \frac{n+1-x}{q(n+1)}. \tag{6}$$

According as the value of α of y will be $>$ or $<$, it will be necessary, by hypothesis, that that of t which corresponds to it is positive or negative. By regarding therefore k as a positive quantity, one will take $t = k$, in the case of $\alpha > h$, and $t = -k$ when one will have $\alpha < h$. In the first case, one will have

$$\int_{\alpha}^{\infty} \mathrm{Y} dy = \mathrm{H} \int_{k}^{\infty} e^{-t^2} \frac{dy'}{dt} dt,$$

and in the second

$$\int_{\alpha}^{\infty} \mathrm{Y} dy = \int_{-\infty}^{\infty} \mathrm{Y} dy - \mathrm{H} \int_{-\infty}^{-k} e^{-t^2} \frac{dy'}{dt} dt,$$

One has besides

$$\int_{-\infty}^{-k} e^{-t^2} t^{2i+1} dt = -\int_{k}^{\infty} e^{-t^2} t^{2i+1} dt,$$

$$\int_{-\infty}^{-k} e^{-t^2} t^{2i} dt = \int_{k}^{\infty} e^{-t^2} t^{2i} dt;$$

i being a whole and positive number, or zero. If therefore one makes generally

$$\int_{k}^{\infty} e^{-t^2} t^{2i} dt = \mathrm{K}_i, \quad \int_{k}^{\infty} e^{-t^2} t^{2i+1} dt = \mathrm{K}'_i,$$

there will result from it

$$\begin{aligned}\int_{\alpha}^{\infty} \mathrm{Y}dy =& \mathrm{H}(h'\mathrm{K}_0 + 3h'''\mathrm{K}_1 + 5h^{v}\mathrm{K}_2 + \text{etc.}) \\ &+ \mathrm{H}(2h''\mathrm{K}'_0 + 4h^{iv}\mathrm{K}'_1 + 6h^{vi}\mathrm{K}'_2 + \text{etc.}),\end{aligned} \tag{7}$$

for the case of α or $\frac{p}{q} > h$, and

$$\begin{aligned}\int_{\alpha}^{\infty} \mathrm{Y}dy = \int_{-\infty}^{\infty} \mathrm{Y}dy &- \mathrm{H}(h'\mathrm{K}_0 + 3h'''\mathrm{K}_1 + 5h^{v}\mathrm{K}_2 + \text{etc.}) \\ &+ \mathrm{H}(2h''\mathrm{K}'_0 + 4h^{iv}\mathrm{K}'_1 + 6h^{vi}\mathrm{K}'_2 + \text{etc.}),\end{aligned} \tag{8}$$

for the case of $\frac{p}{q} > h$.

Each of the series contained in these formulas, will have, in general, the same degree of convergence as series (5). The values of the integrals designated by K'_i will be able to be obtained only by approximation, when k will be different from zero. Those which are represented by K'_i will be obtained under finite form, and one will have

$$\begin{aligned}\mathrm{K}'_2 = \frac{1}{2} e^{-k^2} \big(& k^{2i} + ik^{2i-1} + i.i-1.k^{2i-2} + \cdots \\ &\cdots + i.i-1\ldots 1.k^2 + i.i-1\ldots 2.1\big).\end{aligned}$$

When one will have exactly $\frac{p}{q} = h$, one will have at the same time $k = 0$, and consequently

$$\mathrm{K}_i = 1.3.5\ldots 2i-1.\frac{\sqrt{\pi}}{2}, \quad \mathrm{K}'_i = 1.2.3\ldots i.\frac{1}{2};$$

and according to the value of $\int_{-\infty}^{\infty} \mathrm{Y}dy$, formulas (7) and (8) will coincide and will be reduced to

$$\begin{aligned}\int_{\alpha}^{\infty} \mathrm{Y}dy = \frac{\mathrm{H}\sqrt{\pi}}{2}&(h' + \frac{1.3}{2}h''' + \frac{1.3.5}{4}h^{v} + \text{etc.}) \\ &+ \mathrm{H}(h'' + 1.2h^{iv} + 1.2.3h^{vi} + \text{etc.})\end{aligned} \tag{9}$$

(6) We will suppose actually the numbers n, x, $n-x$, great enough in order that one is able to neglect in these different formulas, the quantities h''', h^{iv}, etc. By putting n in place of $n+1$ into the values of h' and h'', one will have

$$\frac{h''}{h'} = \frac{(n+x)\sqrt{2}}{3\sqrt{nx(n-x)}};$$

equation (3) and formulas (5), (7) and (8) will give next

$$\left.\begin{aligned} \mathrm{X} &= \frac{1}{\sqrt{\pi}}\int_k^\infty e^{-t^2}dt + \frac{(n+x)\sqrt{2}}{3\sqrt{\pi n x(n-x)}}e^{-k^2}, \\ \mathrm{X} &= 1-\frac{1}{\sqrt{\pi}}\int_k^\infty e^{-t^2}dt + \frac{(n+x)\sqrt{2}}{3\sqrt{\pi n x(n-x)}}e^{-k^2}; \end{aligned}\right\} \tag{10}$$

the first or the second of these two values of X taking place according as one has $\frac{p}{q} >$ or $< h$, and k being a positive quantity, given by equation (6). These formulas will make known with a sufficient exactitude the probability X which there is a question to determine.

If n is an even number, let one make $x = \frac{n}{2}$, and let one suppose $p > q$, one will have

$$h = \frac{n}{n+2}, \frac{p}{q} > h;$$

this will be therefore the first equation (10) which it will be necessary to employ: this formula and equation (6) will become

$$\left.\begin{aligned} \mathrm{X} &= \frac{1}{\sqrt{\pi}}\int_k^\infty e^{-t^2}dt + \sqrt{\frac{2}{\pi n}}e^{-k^2}, \\ k^2 &= \frac{n}{2}\log\frac{n}{2p(n+1)} + \frac{n+2}{2}\log\frac{n+2}{2q(n+1)}; \end{aligned}\right\} \tag{11}$$

and X will be the probability that out of a very great number n of trials, the most probable event will not arrive however more often than the contrary event. By calling P the probability that they both will arrive the same number of times, this which is possible, since n is an even number, X−P will be the probability that the first event will arrive less often than the second. In the case of $p = q = \frac{1}{2}$, it is evident that the double of this last probability, added to P, will be certitude; one will have therefore 2X−P= 1, or

$$\mathrm{P} = \frac{2}{\sqrt{\pi}}\int_k^\infty e^{-t^2}dt - 1 + \frac{2\sqrt{2}}{\sqrt{\pi n}}e^{-k^2};$$

and it is, in fact, this which one is able to verify easily.

By reducing into series, one has

$$\begin{aligned} n\log\frac{n}{n+1} &= -\log\left(1+\frac{1}{n}\right) = -1 + \frac{1}{2n} - \text{etc.}, \\ (n+2)\log\frac{n+2}{n+1} &= -(n+2)\log\left(1-\frac{1}{n+2}\right) = 1 + \frac{1}{2(n+2)} + \text{etc.}, \end{aligned}$$

and consequently

$$k^2 = \frac{1}{4n} + \frac{1}{4(n+2)} + \text{etc.};$$

therefore by conserving only the terms of order $\frac{1}{\sqrt{n}}$, we will have

$$k = \frac{1}{\sqrt{2n}}, \quad e^{-k^2} = 1;$$

we will have, at the same time,

$$\int_k^\infty e^{-t^2}dt = \int_0^\infty e^{-t^2}dt - \int_0^k e^{-t^2}dt = \frac{1}{2}\sqrt{\pi} - \frac{1}{\sqrt{2\pi}};$$

whence there will result

$$\mathrm{P} = \sqrt{\frac{2}{\pi n}},$$

this which coincides effectively with formula (1), when one makes there $p = \frac{1}{2}$, $q = \frac{1}{2}$, $x = \frac{n}{2}$.

If n is an odd number, let one make $x = \frac{1}{2}(n-1)$, and let one suppose always $p > q$; one will have further $\frac{p}{q} > h$: the first formula (10) and equation (6) will become

$$\left.\begin{aligned} \mathrm{X} &= \frac{1}{\sqrt{\pi}}\int_k^\infty e^{-t^2}dt + \frac{\sqrt{2}}{\sqrt{\pi n}}e^{-k^2}, \\ k^2 &= \frac{n-1}{2}\log\frac{n-1}{2p(n+1)} + \frac{n+3}{2}\log\frac{n+3}{2q(n+1)}; \end{aligned}\right\} \tag{12}$$

and X will be the probability that out of a very great number of trials, the most probable event will be presented less often than the contrary event; because, n being odd, the case of equality will be impossible. In the case of $p = q = \frac{1}{2}$, this probability X must be equal to $\frac{1}{2}$; and it is also this which we are going to verify.

We will have

$$\begin{aligned} (n-1)\log\frac{n-1}{n+1} &= -(n-1)\log\left(1+\frac{2}{n-1}\right) = -2 + \frac{2}{n-1} - \text{etc.}, \\ (n+3)\log\frac{n+3}{n+1} &= -(n+3)\log\left(1+\frac{2}{n+3}\right) = -2 + \frac{2}{n+3} + \text{etc.}, \end{aligned}$$

and consequently

$$k^2 = \frac{1}{n-1} + \frac{1}{n+3} + \text{etc.}$$

By neglecting the terms of order of $\frac{1}{n}$, there will result from it

$$k = \sqrt{\frac{2}{n}}, \quad e^{-k^2} = 1, \quad \int_k^\infty e^{-t^2}dt = \frac{1}{2}\sqrt{\pi} - \sqrt{\frac{2}{n}};$$

this which reduces to $\frac{1}{2}$, the preceding value of X.

(7) From equation $\frac{p}{q} = h$, one draws

$$x = (n+1)p, \quad n+1-x = (n+1)q,$$

because $p+q=1$. We designate by z a positive quantity, such that this value of x diminished from z is a whole number; we will be able to take

$$x = (n+1)p - z, \quad n+1-x = (n+1)q + z;$$

and we will have $\frac{p}{q} > h$. By developing the second member of equation (6) according to the powers of z, one finds

$$k^2 = \frac{z^2}{2(n+1)pq}\left(1 - \frac{(p-q)z}{3pq(n+1)} + \text{etc.}\right);$$

and if one makes

$$z = r\sqrt{2(n+1)pq},$$

one deduces from it

$$k = r\left(1 - \frac{(p-q)z}{3pq(n+1)} + \text{etc.}\right).$$

The series contained between the parentheses proceeds according to the powers of $\frac{r}{\sqrt{n+1}}$; it will be very convergent if r is not a very great number, and if any of the two fractions p and q are not very small; one will have then to conserve only its first two terms, or take simply $k = r - \delta$, by making,

$$\frac{(p-q)r^2}{3\sqrt{2(n+1)pq}} = \delta.$$

One will have, at the same time,

$$x = (n+1)p - r\sqrt{2(n+1)pq};$$

but in the second term of the first formula (10), it will suffice to make $k = r$ and $x = np$; and this formula will become

$$\mathrm{X} = \frac{1}{\sqrt{\pi}}\int_{r-\delta}^{\infty} e^{-t^2}dt + \frac{(1+p)\sqrt{2}}{2\sqrt{\pi npq}}e^{-r^2}.$$

We designate by r' another positive quantity, which is not a very great number either. If one supposes that one has

$$x = (n+1)p - r'\sqrt{2(n+1)pq},$$

the corresponding value of k, drawn from equation (6), will be $k = r' + \delta'$, by making, for brevity,

$$\frac{(p-q)r'^2}{3\sqrt{2(n+1)pq}} = \delta'.$$

One will have, in this case, $\frac{p}{q} < h$; it will be necessary therefore to employ the second formula (10), which will become

$$\mathrm{X} = 1 - \frac{1}{\sqrt{\pi}}\int_{r'+\delta'}^{\infty} e^{-t^2}dt + \frac{(1+p)\sqrt{2}}{2\sqrt{\pi npq}}e^{-r'^2}.$$

If one subtracts from this here, the preceding value of X, and if one calls U the difference, there comes

$$\mathrm{U} = 1 - \frac{1}{\sqrt{\pi}}\int_{r-\delta}^{\infty} e^{-t^2}dt - \frac{1}{\sqrt{\pi}}\int_{r'+\delta'}^{\infty} e^{-t^2}dt$$
$$+\frac{(1+p)\sqrt{2}}{2\sqrt{\pi npq}}\left(e^{-r'^2} - e^{-r^2}\right),$$

for the probability that the event A will arrive a number of times which will not exceed the second value of x, and will surpass the first at least by one unit.

(8) In order to simplify this result, let N be the greatest number contained in np, and f a fraction such that one has $np = N + f$; we designate by u a quantity such that $u\sqrt{2(n+1)pq}$ is a whole number, very small with respect to N; we make next

$$\begin{aligned} p + f - r\sqrt{2(n+1)pq} &= -u\sqrt{2(n+1)pq} - 1, \\ p + f + r'\sqrt{2(n+1)pq} &= u\sqrt{2(n+1)pq}; \end{aligned}$$

U will express the probability that the number of times of which the question will be contained between the limits

$$N \pm u\sqrt{2(n+1)pq},$$

equidistant from N, or that it will be equal to one of them. The values of $r - \delta$ and $r' + \delta'$ will be of the form:

$$\begin{aligned} r - q &= u + \varepsilon + \frac{1}{\sqrt{2(n+1)pq}}, \\ r' + \delta' &= u - \varepsilon; \end{aligned}$$

ε being a quantity of order $\frac{1}{\sqrt{n}}$. Now, by designating by v a quantity of this order, of which one neglects the square, one has

$$\int_{u+v}^{\infty} e^{-t^2} dt = \int_{u}^{\infty} e^{-t^2} dt - e^{-u^2} v;$$

if therefore one applies this transformation to the two integrals which contain U, and if one makes $r' = r$ in the terms contained outside of the sign $\int$, which are already divided by $\sqrt{n}$, one will have

$$\mathrm{U} = 1 - \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} e^{-t^2} dt - \frac{e^{-u^2}}{\sqrt{2\pi npq}}. \tag{13}$$

If one had wished that the values of x not comprehend their inferior limit, it would have been necessary to make $r - \delta = n + \varepsilon$, and the value of X would not contain its last term. Likewise in order that the superior limit of these values of x, were excluded from it, one would need to diminish $r' + \delta'$ from $\frac{1}{\sqrt{2(n+1)pq}}$, this which would have again made vanish the last term of U. It follows therefore that this last term must be the probability that one has precisely

$$x = N + u\sqrt{2(n+1)pq};$$

u being a positive or negative quantity, such that the second term of x is very small with respect to the first. It is also this which results from formula (1).

In fact, by neglecting the quantities of order $\frac{1}{n}$, one will have

$$\frac{x}{n} = p + u\sqrt{\frac{2pq}{n}}, \quad \frac{n-x}{n} = q - u\sqrt{\frac{2pq}{n}};$$

whence one concludes

$$\log\left(\frac{np}{x}\right)^x\left(\frac{nq}{nn-x}\right)^{n-x} = -x\log\left(1+\frac{u}{p}\sqrt{\frac{2pq}{n}}\right)$$
$$-(n-x)\log\left(1+\frac{u}{p}\sqrt{\frac{2pq}{n}}\right);$$

by developing these logarithms and reducing, one finds, to the degree of approximation where we have stopped ourselves, $-u^2$ for the value of the second member of this equation; one will have consequently

$$\left(\frac{np}{x}\right)^x\left(\frac{nq}{nn-x}\right)^{n-x} = e^{-u^2};$$

and as one has at the same time

$$\frac{x(n-x)}{n^2} = pq,$$

formula (1) will become

$$\mathrm{P} = \frac{e^{-u^2}}{\sqrt{2\pi npq}};$$

this which it was the concern to verify.

(9) By calling x' the number of times that A will arrive out of the number n of trials, we will be able to say that U is the probability that the difference $\frac{x'}{n} - p$ will be contained between the two limits:

$$u \pm \sqrt{\frac{2pq}{n}},$$

which will be equally those of the difference $\frac{n-x'}{n} - q$, by changing their signs.

One will be able always to take u great enough in order that this probability U differs as little as one will wish from certitude. It will not even be necessary to give to u a great value, in order to render very small the difference $1-\mathrm{U}$: it will suffice, for example, to take u equal to five or six, in order that the exponential e^{-u^2}, the integral $\int_u^\infty e^{-t^2}dt$, and hence the value of $1-\mathrm{U}$, be nearly insensible. This remark is important; because the preceding analysis requires, effectively, that u is not a considerable number, since u differs very little from r, and that by calculating the value of k in n$^\mathrm{o}$ 7, we have neglected the quantities of order $\frac{r^2}{n}$; this which supposes that r or u is not comparable to $\sqrt{n}$.

The value of u having been conveniently chosen, and remaining constant, the limits of $\frac{x'}{n} - p$ will be tightened more and more in measure as the number n will increase; the ratio $\frac{x'}{n}$ of the number of times that the event A will arrive to the total number of trials, will differ therefore less and less from the probability p of this event; and one will be able always to take n great enough in order that it has the probability U that the difference $\frac{x'}{n} - p$ will be as small as one will wish; this which is, as one knows, the theorem of Jacques Bernoulli on the repetition, in a number of trials, of an event of which the chance is given *a priori*.

(10) We have supposed, in order to arrive to this theorem, that x and $n-x$ are very great numbers, as also that their sum; according to the values of x and $n-x$ of n^{o} 7, it will be necessary therefore that the products pn and qn are very great: but if the probability p is very small, of such sort that pu is a fraction, or a number of little consequence, it will be very probable that A will arrive only a very small number of times out of a very great number n of trials; and in this case, formula (2) will make known without difficulty, the probability X that this number of times will not exceed x. In fact, let $pn = \omega$; by neglecting the ratio $\frac{x}{n}$, the quantity composed between the parentheses in formula (2), will become

$$1+\omega+\frac{\omega^2}{2}+\frac{\omega^3}{2.3}+\cdots+\frac{\omega^x}{1.2.3\ldots x};$$

one will have, at the same time,

$$q = 1-\frac{\omega}{n}, \quad q^{-x} = 1, \quad q^n = e^{-\omega};$$

there will result from it therefore

$$\mathrm{X} = \left(1+\omega+\frac{\omega^2}{2}+\frac{\omega^3}{2.3}+\cdots+\frac{\omega^x}{1.2.3\ldots x}\right)e^{-\omega},$$

or, this which is the same thing,

$$\mathrm{X} = 1-\frac{\omega^{x+1}e^{-\omega}}{1.2.3\ldots x+1}\left(1+\frac{\omega}{x+2}+\frac{w^2}{x+2.x+3}+\text{etc.}\right).$$

Now, one sees that if x is not a small number, this value of X will differ very little from unity. If one has, for example, $x = 10$ and $\omega = 1$, the difference $1-$X will be nearly one hundred millionth, that is to say that it is near certain that an event of which the chance is $\frac{1}{n}$, will not arrive more than 10 times, out of a very great number n of trials.

In the case $x = 0$, one has X$= e^{-\omega}$, for the probability that out of a very great number n of trials, an event of which the change is $\frac{\omega}{n}$ will not arrive a single time.

(11) The integral $\int_u^\infty e^{-t^2}dt$ which contains formula (13), will be calculated, in general, by the method of quadratures. One finds, at the end of the *Analyse des réfractions* of Kramp, a table of its values which extends from $u = 0$ to $u = 3$, and according to which, one has

$$\int_u^\infty e^{-t^2}dt = 0.000019577,$$

for $u = 3$. By means of integration by parts, one finds

$$\int_u^\infty e^{-t^2}dt = \frac{e^{-u^2}}{2u}\left(1-\frac{1}{2u^2}+\frac{1.3}{2^2u^4}-\frac{1.2.3}{2^3u^6}+\text{etc.}\right); \qquad (14)$$

for $u > 3$, the series comprehended between the parentheses will be sufficiently convergent, and this formula will serve to calculate the values of the integral. One has also

$$\int_u^\infty e^{-t^2} dt = \frac{1}{2}\sqrt{\pi} - \int_0^u e^{-t^2} dt;$$

and by developing the exponential e^{-t^2} according to the powers of t^2, one will have

$$\int_0^u e^{-t^2} dt = u - \frac{u^3}{1.3} + \frac{u^5}{1.2.5} - \frac{u^7}{1.2.3.7} + \text{etc.};$$

a series which will be very convergent for the values of u less than unity.

If one wishes to calculate the value of u for which one has U$= \frac{1}{2}$, one will make use of this last series, and according to equation (13), one will have

$$u - \frac{u^3}{1.3} + \frac{u^5}{1.2.5} - \frac{u^7}{1.2.3.7} + \text{etc.} = \frac{1}{4}\sqrt{\pi} + \frac{e^{-u^2}}{2\sqrt{2npq}}.$$

By designating by a the value of u which satisfies this equation, setting aside the second term of its second member, we will have next

$$u = a + \frac{1}{2\sqrt{2npq}},$$

to the quantities nearly of order $\frac{1}{n}$. After some tests, one finds $a = 0.4765$ for the approximate value of a; whence there results that it will be equally probable that the difference $\frac{x'}{n} - p$ will fall outside of or within some limits:

$$\pm\left(0.4765\sqrt{\frac{2pq}{n}} + \frac{1}{2n}\right).$$

For any value of u whatsoever, the difference of the two quantities $\frac{x'}{n} - p$ and $\frac{n-x'}{n} - q$, will have for limits the double of $\pm u\sqrt{\frac{2pq}{n}}$; if therefore one has $p = q = \frac{1}{2}$, there will be a probability equal to $\frac{1}{2}$, that the quantity $\frac{2x'-n}{n}$, will be contained between the limits

$$\pm\left(\frac{0.6739}{\sqrt{n}} + \frac{1}{n}\right);$$

consequently it will be equally probable that the difference $x' - (n - x')$ between the numbers of events A and B, of which the chances are equals, will surpass or will be less than $0.6739\sqrt{n} + 1$, setting aside the sign. According to formula (1), one will have

$$\text{P} = \sqrt{\frac{2}{\pi n}} = \frac{0.7979}{\sqrt{n}},$$

for the probability that this difference will be precisely null.

§II.

Probabilities of simple events and of future events after the events observed.

(12) Until here we have supposed known *a priori*, the chance p of the event A, and we have concluded from it the probability of a future event, relative to the repetition of A out of a very great number of trials; but in the applications of the calculation of chances to natural phenomena, and particularly in the question indicated by the title of this Memoir, the value of p must, on the contrary, be deduced as much as it is possible, from events observed in great numbers, in order to serve next to calculate the probability of future events. It is this problem which is now going to occupy us.

We suppose first that the unknown probability p of the event A, is susceptible only of m different values which are able to be differently probable. We represent these m values by

$$v_1, v_2, v_3, \dots v_n, \dots v_m,$$

and respective probabilities by

$$\mathrm{R}_1, \mathrm{R}_2, \mathrm{R}_3, \dots \mathrm{R}_n, \dots \mathrm{R}_m.$$

Let also

$$\mathrm{V}_1, \mathrm{V}_2, \mathrm{V}_3, \dots \mathrm{V}_n, \dots \mathrm{V}_m,$$

be the corresponding probabilities of a composite event C, so that V_n designates the probability of C in function of v_n which will hold if it was certain that one had $p = v_n$. By hypothesis, the composite event C has been observed; and one demands the probability R_n that its arrival corresponds to the probability v_n of the simple event A.

In order to determine R_n, I suppose that one reduces all the fractions $\mathrm{V}_1, \mathrm{V}_2$, etc., to the same denominator, and that one replaces them by

$$\frac{\mathrm{N}_1}{\mu}, \frac{\mathrm{N}_2}{\mu}, \frac{\mathrm{N}_3}{\mu}, \dots \frac{\mathrm{N}_n}{\mu}, \dots \frac{\mathrm{N}_m}{\mu},$$

$\mu, \mathrm{N}_1, \mathrm{N}_2$, etc., being whole numbers. The question proposed is evidently the same as if one had a number of urns m, containing each a number μ of balls; of which the first contained the number N_1 of white balls, the second contained a number N_2 of them, the third a number N_3, and thus in sequence; that one had extracted one white ball from these vases, and that one demanded the probability that this ball is exited from the n^{th} urn. The extraction of a white ball is the observed fact, or the event C, and the exit from the n^{th} urn is the case where this fact coincides with the hypothesis $p = v_n$ which gives to C a probability $\frac{\mathrm{N}_n}{\mu}$, or V_n.

This put, we mark the balls of the first urn, with the n$^{\text{o}}$ 1; those of the second urn, with the n$^{\text{o}}$ 2; etc. Since the number of balls is the same and equal to μ for the different numerals, it is evident that one is able to reunite them all in one same vase, with nothing changing to the probability to bring forth a white ball carrying the n$^{\text{o}}$ n, or arising from the n^{th} urn. Now, if one makes

$$\mathrm{N}_1 + \mathrm{N}_2 + \cdots + \mathrm{N}_m = \lambda,$$

λ being the total number of white balls contained in this unique vase, and, consequently, if one supposes that one has extracted from it one white ball, the ratio $\frac{N_n}{\lambda}$ will be the probability that this ball is marked with the n° n, or the value demanded of R_n. Therefore by dividing the two terms of this fraction by μ, one will have

$$R_n = \frac{V_n}{\sum V_n};$$

the sum $\sum$ extending to all the values of the index n, from $n = 1$ to $n = m$.

Let C' be another composite event and dependent on A; we call V'_n the probability of C' as function of v_n, which would take place if one had certainly $p = v_n$; as this value of p has itself only a probability R_n, the coincidence of the event C' and of $p = v_n$, is an event composed of two others, which will have for probability the product $V'_n R_n$, of those of these two events. This being, if one designates by T the probability of C', relative to the m different values of p, one will have

$$T = \sum V'_n R_n,$$

the sum $\sum$ being the same signification as above; and by substituting for R_n its preceding value, there will result from it

$$T = \frac{\sum V'_n V_n}{\sum V_n},$$

(13) We suppose actually that the probability p of A is susceptible of all the possible values from zero to unity; their number m will be infinite, and the probability of each of them will be infinitely small. In representing by v any value of p, and by V, V', R, this which becomes V_n, V'_n, R_n, when one puts v there in the place of v_n; multiplying high and low by dv, the preceding formulas; observing finally that the sums $\sum$ will be changed into definite integrals, taken from $v = 0$ to $v = 1$; we will have

$$R = \frac{V dv}{\int_0^1 V dv}, \quad T = \frac{\int_0^1 V' V dv}{\int_0^1 V dv}.$$

If one designates by Z the probability that the value of p will be contained between some given limits a and b, Z will have for value a finite quantity, namely:

$$Z = \frac{\int_a^b V' V dv}{\int_0^1 V dv}.$$

Let, at the same time Q, be the probability that the event C' will correspond to the one of the values of p comprehended between these limits; one will have also

$$Q = \frac{\int_1^b V' V dv}{\int_0^1 V dv}.$$

According to these expressions of T, Z, Q, one will have

$$T < Q + M(1 - Z) \quad \text{and} \quad > Q + M'(1 - Z),$$

by naming M and M′ the greatest and the smallest value of V′ which corresponds to the values of p comprehended from $p=0$ to $p=a$ and from $p=b$ to $p=1$, or which falls outside of the given limits a and b. Now, M and M′ being some positive quantities which are not able to surpass unity, if the difference $1-\mathrm{Z}$ is small enough provided that one neglects it, the total probability T of the event C′, will coincide with Q, this which will simplify the calculation of it. This case will take place in the diverse applications which we are going to make of the preceding formulas.

(14) If the observed event C consists in this that, out of a number m of trials, A is arrived a number s of times, one will have, according to the first equation of nº 1,

$$\mathrm{V}=\frac{1.2.3\ldots m}{1.2.3\ldots s.1.2.3\ldots m-s}v^s(1-v)^{m-s},$$

and consequently,

$$\mathrm{R}=\frac{v^s(1-v)^{m-s}ds}{\int_0^1 v^s(1-v)^{m-s}ds},\quad \mathrm{Z}=\frac{\int_a^b v^s(1-v)^{m-s}ds}{\int_0^1 v^s(1-v)^{m-s}ds}.$$

We call g the value of v which renders a maximum, the coefficient of dv under the sign $\int$, and G the corresponding value of this coefficient; we will have

$$g=\frac{s}{m},\quad \mathrm{G}=\left(\frac{s}{m}\right)^s\left(\frac{m-s}{m}\right)^{m-s}.$$

We make next

$$v^s(1-v)^{m-1}=\mathrm{G}e^{-t^2};$$

t being a new variable, of which the values $t=-\infty, t=0, t=\infty$, will correspond to $v=0, v=g, v=1$. By taking the logarithms of the two members of this equation, one will deduce from it next for v, a value in series of the form:

$$v=g+g't+g''t^2+g'''t^3+\text{etc.},$$

g', g'', g''', etc., being some coefficients independent of t, of which the values will be determined by the substitution of this series into the logarithmic equation, and by the comparison of the similar terms in the two members. We will suppose that s and $m-s$, are very large numbers, comparable to their sum m; and then these coefficients g', g'', g''', etc., will form a very rapidly decreasing series, of which the terms will be of order to the fractions $\frac{1}{\sqrt{m}}, \frac{1}{m}, \frac{1}{m\sqrt{m}}$, etc. The first two will have for values:

$$g'=\sqrt{\frac{2(m-s)s}{m^3}},\quad g''=\frac{2(m-2s)}{3m^2};$$

this which gives for their ratio:

$$\frac{g''}{g'}=\frac{2(m-2s)}{3\sqrt{2ms(m-s)}}.$$

By means of this transformation and by neglecting the quantities of order of $\frac{1}{m}$, one will have

$$\int_0^1 v^s(1-v)^{m-s}dv = \mathrm{G}\int_{-\infty}^{\infty} e^{-t^2}\frac{dv}{dt}dt = \mathrm{G}g'\sqrt{\pi}.$$

We designate by z a positive quantity which is not very great, and we take

$$a = g - g'z, \quad b = g + g'z,$$

for the limits of the integral which forms the numerator of Z; we make next

$$v = g + g'\theta, \quad dv' = g'd\theta;$$

the corresponding values of θ will be $\pm z$; and as one will have

$$t = \theta - \frac{g''}{g'}\theta^2, \quad e^{-t^2} = e^{-\theta^2}\left(1 + \frac{2g''\theta^3}{g'}\right),$$

by neglecting always the quantities of order $\frac{1}{m}$, there will result from it

$$\int_a^b v^s(1-v)^{m-s}dv = \mathrm{G}g'\int_{-z}^{z} e^{-\theta^2}\left(1 + \frac{2g''\theta^3}{g'}\right)d\theta$$
$$= 2\mathrm{G}g'\int_0^z e^{-\theta^2}d\theta.$$

The value of Z will be therefore simply

$$\mathrm{Z} = \frac{1}{\sqrt{\pi}}\int_0^z e^{-\theta^2}d\theta; \qquad (a)$$

and it will express the probability that the value of p is contained between the limits:

$$\frac{s}{m} \pm g'z. \qquad (b)$$

At the same time the probability R of an intermediate value:

$$p = \frac{s}{m} + g'\theta, \qquad (c)$$

will have for expression

$$\mathrm{R} = \frac{1}{\sqrt{\pi}}e^{-\theta^2}\left(1 + \frac{2g''\theta^3}{g'}\right)d\theta. \qquad (d)$$

We always take the quantity z great enough in order that Z differ very little from unity, and if one is able, consequently, to regard Q as expressing, with a sufficient approximation, the probability of the event C′. By making $z = 3$, for example, one will have

$$\mathrm{Z} = 1 - 0.00002209;$$

and one will be able to neglect the difference $1-\mathrm{Z}$. From this manner the probability of C′ will be

$$\mathrm{Q}=\frac{1}{\sqrt{\pi}}\int_{-z}^{z}\Pi e^{-\theta^2}\left(1+\frac{2g''\theta^3}{g'}\right)d\theta; \tag{e}$$

Π designating the probability of the same event which would have place if the preceding value of p were certain, or that which the function V′ of the preceding number becomes when one replaces the variable v by this value of p.

(15) We take for C′ the event to which the formulas are brought back, that is to say, the case where, out of a number n of trials, the event A will arrive a number of times which will not surpass x, the two parts x and $n-x$ of n being supposed very great numbers. We will give below some examples in which the two ratios $\frac{x}{n+1}$ and $\frac{s}{m}$ will not be very little different from one another; now we are going to suppose that their difference is very small and of order $\frac{1}{\sqrt{m}}$; and we will represent it by $\gamma g'$, so that one has

$$\frac{x}{n+1}=\frac{s}{m}-\gamma g', \tag{f}$$

γ designating a fraction or a number of little consequence.

According to equation (c), we will have

$$x=p(n+1)-(\gamma+\theta)(n+1)g',$$
$$n+1-x=q(n+1)+(\gamma+\theta)(n+1)g';$$

whence one concludes

$$\frac{p}{q}=\frac{x}{n+1-x}+\frac{(\gamma+\theta)(n+1)^2g'}{(n+1-x)^2},$$

by neglecting the square of g'. If therefore we suppose the quantity γ positive and $=z$ or $>z$, the ratio $\frac{p}{q}$ will surpass the quantity h of n° 4 for all the values of θ contained between the limits $\pm z$ of the integral which contains formula (e); consequently this will be the first equation (10) of which it will be necessary to make use in order to form the value of X as function of θ, that we will have to substitute in the place of Π in the expression of Q.

If we make

$$(\gamma+\theta)(n+1)g'=r\sqrt{2pq(n+1)},$$

we will draw from equation (6), as in n° 7,

$$k=r-\frac{(p-q)r^2}{3\sqrt{2pq(n+1)}},$$

by neglecting the quantities of order of $\frac{1}{n}$, and supposing that each of the two quantities p and q none is a very small fraction. By virtue of equation (c) and of the value of g' of which one neglects the square, one will have also

$$\frac{1}{\sqrt{pq}}=\frac{m}{\sqrt{s(m-s)}}-\frac{\theta(m-2s)\sqrt{m}}{s(m-s)\sqrt{2}}.$$

Let further, for brevity,

$$\sqrt{\frac{n+1}{m}} = \alpha;$$

there will result from it

$$r = (\gamma+\theta)\alpha - \frac{(m-2s)(\gamma+\theta)\theta\alpha}{\sqrt{2ms(m-s)}};$$

and if α is not a very great quantity, the second terms of k and r will be of order $\frac{1}{\sqrt{m}}$, that is to say, of order of the quantities that we have conserved to present. One would be able to continue to have regard; but in order to simplify the following calculations, we will neglect actually these quantities, and we will take simply

$$k = r = (\gamma+\theta)\alpha.$$

By putting under the sign $\int$ in the first formula (10), $(\gamma+\theta)\alpha t$ and $(\gamma+\theta)\alpha dt$ in the place of t and dt, and making $x = \frac{ns}{m}$ in its second term which is of order $\frac{1}{\sqrt{n}}$, this formula will become

$$\mathrm{X} = \frac{\alpha}{\sqrt{\pi}} \int_1^\infty e^{-(\gamma+\theta)^2\alpha^2t^2}(\gamma+\theta)dt + \frac{(m+s)\sqrt{2}}{3\sqrt{\pi ns(m-s)}} e^{-(\gamma+\theta)^2\alpha^2};$$

and if one substitutes this value of X in the place of Π in equation (e), and if one suppresses the term multiplied by $\frac{g''}{g'}$, a quantity of order $\frac{1}{\sqrt{m}}$, one will have, by interchanging the order of integrations,

$$\mathrm{Q} = \frac{\alpha}{\pi}\int_\alpha^\infty \left(\int_{-z}^{z} e^{-(\gamma+\theta)^2\alpha^2t^2} e^{-\theta^2}(\gamma+\theta)d\theta\right)dt + \frac{(m+s)\sqrt{2}}{3\sqrt{\pi ns(m-s)}}\int_{-z}^{z} e^{-(\gamma+\theta)^2\alpha^2} e^{-\theta^2} d\theta.$$

By hypothesis, the factor $e^{-\theta^2}$ is nearly null at the limits $\pm z$, this which permits to extend now, without sensible error, the integrals relative to θ from $-\infty$ to $+\infty$. One will have then

$$\int_{-\infty}^{\infty} e^{-(\gamma+\theta)^2\alpha^2t^2-\theta^2}(\gamma+\theta)d\theta = \frac{\gamma\sqrt{\pi}}{(1+\alpha^2t^2)^{\frac{3}{2}}} e^{-\frac{\gamma^2\alpha^2t^2}{1+\alpha^2t^2}},$$

$$\int_{-\infty}^{\infty} e^{-(\gamma+\theta)^2\alpha^2t^2-\theta^2}d\theta = \frac{\sqrt{\pi}}{\sqrt{1+\alpha^2}} e^{\frac{\alpha^2t^2}{1+\alpha^2}},$$

and consequently

$$\mathrm{Q} = \frac{1}{\sqrt{\pi}}\int_1^\infty e^{-\frac{\gamma^2\alpha^2t^2}{1+\alpha^2t^2}} \frac{\gamma\alpha dt}{(1+\alpha^2t^2)^{\frac{3}{2}}} + \Gamma,$$

by making, for brevity,

$$\frac{\gamma\alpha}{\sqrt{1+\alpha^2}} = \beta, \quad \frac{(m+s)\sqrt{2}}{3\sqrt{ns(m-s)(1+\alpha^2)}} e^{-\beta^2} = \Gamma.$$

Let finally

$$\frac{\gamma\alpha t}{\sqrt{1+\alpha^2t^2}} = v, \quad \frac{\gamma\alpha dt}{(1+\alpha^2t^2)^{\frac{3}{2}}} = dv;$$

at the limit $t = 1$, one will have $v = \beta$; at the other limit $t = \infty$, one will have $v = \gamma$, and one will be able to take $v = \infty$, because of $\gamma = z$ or $> z$; there will result from it therefore

$$\text{Q} = \frac{1}{\sqrt{\pi}}\int_\beta^\infty e^{-v^2}dv + \Gamma,$$

for the probability that out of a number n of trials, the event A will arrive a number of times which will not exceed the value of x drawn from equation (f), which value is able to be written thus:

$$x = \frac{(n+1)s}{m} - \frac{\beta}{m}\sqrt{2(n+1)(1+\alpha)^2(m-s)s}.$$

We designate by γ' a second positive quantity, equal or superior to z, and by β' and Γ', that which β and Γ become, when one puts γ' in the place of γ. By substituting $-\gamma'$ in γ in equation (f), and making use of the second formula (10), one will find, by an analysis similar to the preceding

$$\text{Q} = 1 - \frac{1}{\sqrt{\pi}}\int_{\beta'}^\infty e^{-v^2}dv + \Gamma',$$

for the probability that the number of times of which there is question will not exceed the value of x expressed by the formula:

$$x = \frac{(n+1)s}{m} - \frac{\beta'}{m}\sqrt{2(n+1)(1+\alpha)^2(m-s)s}.$$

Therefore, by calling U the excess of this second value of Q over the first, we will have

$$\text{U} = 1 - \frac{1}{\sqrt{\pi}}\int_\beta^\infty e^{-v^2}dv - \frac{1}{\sqrt{\pi}}\int_{\beta'}^\infty e^{-v^2}dv - \Gamma + \Gamma';$$

and U will express the probability that out of a number n of trials, this number of times will not exceed the second value of x and will surpass the first at least by unity.

(16) In order to compare this result to the one of n° 8, we designate by N the greatest whole number contained in $\frac{ns}{m}$, by f the difference $\frac{ns}{m}-\text{N}$, and by u a positive quantity, such that $\frac{u}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s}$ is a whole number, very small with respect to N. Let next

$$\begin{aligned}&\frac{s}{m} + f - \frac{\beta}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s}\\ &= -\frac{u}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s},\\ &\frac{s}{m} + f - \frac{\beta'}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s}\\ &= \frac{u}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s};\end{aligned}$$

U will express the probability that the number of events A out of a number of trials n, will be contained between the limits:

$$\mathrm{N} \pm \frac{u}{m}\sqrt{2(n+1)(1+\alpha^2)(m-s)s}, \tag{g}$$

or equal to one of them. Moreover, one will have very nearly

$$\beta' = \beta = u,$$

and exactly

$$\beta' + \beta = 2u + \frac{m}{\sqrt{2(n+1)(1+\alpha^2)(m-s)s}};$$

by means of which the preceding value of U will become

$$\mathrm{U} = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty e^{-t^2}dt + \frac{me^{-u^2}}{\sqrt{2(n+1)(1+\alpha^2)(m-s)s}},$$

by neglecting the quantities of order $\frac{1}{n}$, this which makes the difference $\Gamma - \Gamma'$ vanish, and employing the letter t instead of v under the sign $\int$. As one has supposed $\gamma = z$ or $> z$, it will be necessary that u not be less than $\frac{\alpha z}{\sqrt{1+\alpha^2}}$; whatever be α or $\sqrt{\frac{n}{m}}$, one will satisfy this condition and one will render the probability U very little different from unity, by taking $u = z$ or $> z$.

If the number n is very small with respect to m, and only of a magnitude comparable to $\sqrt{m}$, the quantity α will be of order $\frac{1}{\sqrt{m}}$, and one must replace the factor $1+\alpha^2$ by unity in the preceding formulas; this which reduces the limits (g) to

$$\mathrm{N} \pm \frac{u}{m}\sqrt{2(n+1)(m-s)s},$$

and their probability to

$$\mathrm{U} = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty e^{-t^2}dt + \frac{me^{-u^2}}{\sqrt{2(n+1)(m-s)s}}.$$

Now, this result coincides with the one of n° 8, when one makes, in the formulas of this numeral,

$$p = \frac{s}{m}, \quad q = \frac{m-s}{m}.$$

When the number n of the future events is very small with regard to the number m of observed events, the limits of the number of times that A will arrive and their probability will be able to be calculated by taking for the probability p of A, the ratio $\frac{s}{m}$ of the number of times that this event is arrived to the total number of observations, as if this value of p were certain and given a priori. But it is not thus when the two n and m are of the same order of magnitude: although the value of p concluded from the experience, is contained, with a probability which differs very little from certitude, between the limits which are separated very little also from $\frac{s}{m}$, however, all things equal besides, the number of times that A will arrive will be contained between some limits less narrow than those which would take place if it were certain that one had $p = \frac{s}{m}$. One will be able then to suppress the last term of the value of U, which will be of order $\frac{1}{\sqrt{m}}$, and take simply

$$\text{U} = 1 - \frac{1}{\sqrt{\pi}} \int_u^\infty e^{-t^2} dt. \qquad (h)$$

It is also to this that formula (13) is reduced when one neglects its last term; but the limits to which it corresponds are more narrowed in the ratio of $\sqrt{1+\alpha^2}$ to unity, or of $\sqrt{m+n}$ to $\sqrt{m}$, that the limits (g) relative to the case of which we occupy ourselves now.

(17) In the applications that one will make, of the preceding results, one must not lose from view the assumption on which they are founded, that the simple event A is always the same, by understanding thence that its unknown probability p remains invariable during all the trials past and future. We suppose, for example, that one has an urn which contains an unknown number and considered as infinite, of white balls and of black balls, and that the event A is the arrival of a white ball. Its probability p will be the ratio of the number of white balls to the total number; it will be unknown, when the proportion of the two kinds of balls will be not in the least given; and moreover, p will be susceptible of all possible values, from zero to unity, because of the number of balls supposed infinite. By the same reason, this probability will not be changed during the finite number of trials, even though at each drawing, one will not return into the urn the ball which will be exited from it. That being, if one has drawn from this urn s white balls and $m-s$ black balls, and if these numbers s and $m-s$ are both very great, it will have the probability Z given by equation (a), if the value of p is contained between the limits (b), and the probability U given by equation (h), if out of a number n also very great, of new trials, the one of the white balls that one will draw from the same urn will be contained between the limits (g).

Instead of a single urn, we suppose that one has a number m of them, and that one draws a ball from each of them. If the proportion of white and black balls is the same in all these vases, the probability p to bring forth a white ball will be invariable during the m drawings; but, in general, it will vary with this proportion in any manner whatsoever; now, one will be able nevertheless to calculate the chance of the composite events, as if the value of p, known or unknown, were constant and equal to the mean of its values for all the urns. In fact, let $p_1, p_2, p_3, \ldots p_m$, be these values; the order of the drawings not being able to have any influence on the result, one is able to suppose that the urns in which they take place, are taken successively at random. The probability, at the first drawing, to bring forth a white ball, or of the event A, will be then $\frac{1}{m}\sum p_i$, the sum $\sum$ extending to all the values of the index i from $i = 1$ to $i = m$. At the second drawing, the probability of A will be

$$\frac{\sum p_i - p_{i'}}{m-1},$$

if the urn in which the first is made, corresponds to $p_{i'}$; but this urn having been taken at random, it will be necessary in this expression, to give to i' all the values from $i' = 1$ to $i' = m$, to take the sum of the results, and to divide by m, in order to have the second probability of A, which will be then

$$\frac{m\sum p_i - \sum p_{i'}}{m(m-1)},$$

a quantity yet equal to $\frac{1}{m}\sum p_{i'}$. A the third drawing, the probability of A will be

$$\frac{\sum p_i - p_{i'} - p_{i''}}{m-2},$$

if the first and the second have taken place in the urns which correspond respectively to $p_{i'}$ and $p_{i''}$; because if these two urns have been taken at random, it will be necessary first, without making i' vary, to give to i'' all the values from unity to m, excepting i', and to divide the sum of the results by the number of these values or by $m-1$, this which gives

$$\frac{(m-1)\sum p_i - (m-1)\sum p_{i'} - \sum p_{i''} + p_{i'}}{(m-1)(m-2)},$$

a quantity which is reduced to

$$\frac{\sum p_i - p_{i'}}{m-1},$$

because $\sum p_{i''} = \sum p_i$. It will be necessary next to give to i' all the values from $i'=1$ to $i'=m$, and to divide by m the sum of the results; whence there will result $\frac{1}{m}\sum p_i$ for the probability of A at the third drawing, as in the first two. By continuing thus, one will see that the value of p will be the same and equal to $\frac{1}{m}\sum p_i$ in all the drawings. But one is able also to be assured, by observing that this value is able to be only a linear function of p_1, p_2, etc., symmetric with regard to these m quantities; one is able therefore to represent by $p = \mu\sum p_i$, μ being a coefficient independent of p_1, p_2, etc., in the case where these m quantities are equal among them, one will have therefore $p = m\mu p_{i'}$, and as then one must have $p = p_i$, it is necessary that the product $m\mu$ be unity; whence there results $p = \frac{1}{m}\sum p_i$, whatever be p_1, p_2, etc.

Thus, when the proportion of the white and black balls will be given for each urn in particular, one will calculate, by the formulas of n^{os} 6 and 8, the probability that the number of arrivals of a white ball will not exceed a given number, or will be contained between some given limits, by putting, in these formulas, in the place of p, the mean of these values relative to all the urns. Reciprocally, if this proportion is unknown, and if the probability p is susceptible to all the possible values for each of the urns, the limits (b) of n^o 14 and their probability will correspond to the mean of the unknown values of p for all the urns, and the formulas of n^o 16 will make known the limits of the arrivals of A and their probability, when the drawings will take place in the same system of urns, or in another system for which the mean of the values of p is supposed the same as for the first.

It is in this case of the different urns that it is necessary to assimilate the questions relative to the births of boys and of girls. The event A will be the birth of a boy of which the probability p is susceptible to all the possible values from zero to unity. When one considers the births of the two sexes during a certain time and in a country of a certain extent, the unknown value of p is able to vary with the epochs and the localities, and without doubt it is not the same for all fathers and mothers. The mean of all these different values is the quantity p of which one determines the limits; and it is by supposing that this mean will not vary, that one calculates the probability of the male births during an interval of time. These observations were not unuseful in order to determine with precision the object of the following calculations.

(18) Let m be the number of infants born in France from 1817 to 1826 inclusively, and s the number of male births during these ten years. We will have

$$m = 9656135, \quad s = 4981566;$$

by taking $z = 3$, the limits (b) of the probability p of a male birth, such as it comes to be defined, will be

$$0.5159 \pm 0.0007,$$

and according to equation (a), their probability z will be 0.999978, or nearly equal to unity, so that one is able to regard as very nearly certain, that in France and in the actual epoch, the probability of the birth of a boy is contained between 0.5152 and 0.5166.

Let next n be the mean number of annual births, for which one is able to take the 10^{th} of the births from 1817 to 1826; one will have

$$n = \frac{m}{10}, \quad \mathrm{N} = 498156;$$

and if one makes $u = 3$, the limits (g) will be

$$498156(1 \pm 0.004386).$$

They will correspond to the male births in France during one year; and their probability U given by formula (h), will be the same as Z, or nearly certitude. The corresponding limits of the female births will have for expression:

$$467456(1 \mp 0.004679);$$

and those which result from it for the ratio of the annual births of the two sexes, will be

$$1.0656(1 \pm 0.0091),$$

that is to say, 1.0753 and 1.0559. These here comprehend, in fact, the values of this ratio which have taken place during the ten years that we consider and which are cited at the beginning of this Memoir; but that is not forbidden, as one sees further, that it is very probable only in this interval of time, the chance of a male birth has a little variation from one year to another.

(19) According to equation (c) and the preceding values of m and of s, one will have

$$p = 0.5159 + \theta(0.00023),$$
$$q = 0.4841 - \theta(0.00023),$$

and the limits of the variable θ being ± 3, equation (e) will become at the same time

$$\mathrm{Q} = \frac{1}{\sqrt{\pi}} \int_{-3}^{3} \Pi e^{-\theta^2} [1 - \theta^3(0.00002)] d\theta.$$

If one takes for event C′ of which Q is the probability, the case where out of 12,000 births, for example, those of the boys will not exceed the births of the girls, it will be necessary to put in the place of Π in this equation, the value of X determined by equations (11) of n° 6, in which one makes $n = 12000$ and one will substitute the preceding values of p and q.

In developing the second member of the second equation (11) according to the powers of θ, one finds then

$$k^2 = 6.1028 + \theta(0.1761) + \theta^2(0.00127) + \text{etc.};$$

whence one draws

$$\frac{1}{2k} = 0.2024 - \theta(0.0029) + \text{etc.};$$

formula (14) gives next

$$\int_k^\infty e^{-t^2} dt = e^{-k^2}[0.1883 - \theta(0.0023) + \text{etc.}];$$

and the first equation (11) becomes

$$\mathrm{X} = \frac{1}{\sqrt{\pi}} e^{-k^2}[0.2012 - \theta(0.0023) + \text{etc.}].$$

After having put this value of X in the place of Π in the expression of Q, one will be able to extend the integral from $\theta = -\infty$ to $\theta = +\infty$, because of the smallness of the factor $e^{-k^2-\theta^2}$ to the limits $\theta = \pm 3$ that we have supposed. This being, if one makes

$$\theta = \frac{\theta'}{\sqrt{1.00127}} - \frac{0.1761}{2(1.00127)},$$

the limits relative to θ' will be yet $\pm\infty$, and one will have, very nearly

$$k^2 + \theta^2 = \theta'^2 + 6.0951;$$

whence one will conclude

$$\mathrm{Q} = \frac{(0.2014)e^{-6.0951}}{\pi\sqrt{1.00127}} \int_{-\infty}^{\infty} e^{-\theta'^2} d\theta' = 0.000256.$$

The number 12000 which we have taken for n, is nearly the one of the annual births in a department of a population mean; if therefore the unknown probability of a male birth was the same for each department as for France entire, it would be very little probable that in one year and in the extent of a department, the number of births of boys would not exceed the one of the births of the girls. There would be, on the contrary, near odds of 4000 against one that the first number would surpass the second; and as the contrary event is arrived many times during the ten years that we have considered, it would be necessary to conclude that the chance of a male birth depends on the localities, so that it varies, for one same year, from one department to another, and for one same department from one year to another.

(20) As it is at Paris and among the natural infants that the number of female births approach most each year to be equal to the one of the births of the boys, one is able to desire known the probability that the second number will not exceed the first. Now, the numbers m and s relative to these births, during the thirteen years elapsed from 1815 to 1827, are

$$m = 122404, \quad s = 62239;$$

whence there results, according to equation (c),

$$p = 0.50847 + \theta(0.002021),$$
$$q = 0.49153 - \theta(0.002021),$$

and, by virtue of equation (e)

$$\mathrm{Q} = \frac{1}{\sqrt{\pi}} \int_{-3}^{3} \Pi e^{-\theta^2} [1 - \theta^3(0.00002)] d\theta.$$

by taking always ± 3 for the limits of the variable θ. I will take, besides

$$n = 10000,$$

for the mean number of births outside of marriage which take place each year in Paris; by means of equations (11) and from these values of p, q, n, I will form the expression of X that I will substitute next into the preceding formula in the place of Π: the value of Q will be the probability demanded.

If we make

$$0.00847 + \theta(0.002021) = \frac{1}{2}\alpha,$$

we will have

$$p = \frac{1}{2} + \frac{1}{2}\alpha, \quad q = \frac{1}{2} - \frac{1}{2}\alpha;$$

and because one neglects the quantities of order $\frac{1}{n}$, the second equation (11) will become

$$k^2 = -\frac{n+1}{2} \log(1-\alpha^2) + \frac{1}{2} \log \frac{1+\alpha}{1-\alpha}.$$

If one neglects also the cube of α and the product $n\alpha^4$, one will have

$$k^2 = \frac{n+1}{2}\alpha^2 + \alpha = \beta^2 + \frac{2\beta}{\sqrt{2(n+1)}},$$

by making, for brevity,

$$\frac{n+1}{2}\alpha^2 = \beta^2,$$

so that one has

$$\beta = 1.1979 + \theta(0.2858).$$

The development of $\frac{1}{k}$ according to the powers of θ, will not be a series convergent enough so that one is able to employ, as in the preceding n°, formula (14) for the determination of the integral $\int_k^\infty e^{-t^2}dt$; but one will have

$$k=\beta+\frac{1}{\sqrt{2n}};$$

consequently

$$\int_k^\infty e^{-t^2}dt=\int_\beta^\infty e^{-t^2}dt-\frac{1}{\sqrt{2n}}e^{-\beta^2};$$

and if one puts βt and βdt in the place of t and dt under the $\int$ sign, the first equation (11) will become

$$\mathrm{X}=\frac{\beta}{\sqrt{\pi}}\int_1^\infty e^{-\beta^2t^2}dt+\frac{1}{\sqrt{2\pi n}}e^{-\beta^2}.$$

It is therefore this value of X that I substitute into that of Q in the place of Π. I extend next the integrals relative to θ from $-\infty$ to $+\infty$, this which is permitted, because of the magnitude of the exponents $\beta^2t^2+\theta^2$ and $\beta^2+\theta^2$ to the two limits $\theta\pm 3$; I suppress the term depending on θ^3 that its coefficient renders negligible; in this manner, there comes

$$\mathrm{Q}=\frac{1}{\pi}\int_1^\infty\left(\int_{-\infty}^\infty e^{-\beta^2t^2-\theta^2}\beta d\theta\right)dt+\frac{1}{\pi\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\beta^2-\theta^2}d\theta.$$

By the known formulas, one finds

$$\int_{-\infty}^\infty e^{-\beta^2t^2-\theta^2}\beta d\theta=\frac{a\sqrt{\pi}}{(1+b^2t^2)^{\frac{3}{2}}}e^{-\frac{\alpha^2t^2}{1+b^2t^2}}$$
$$\int_{-\infty}^\infty e^{-\beta^2-\theta^2}d\theta=\frac{\sqrt{\pi}}{(1+b^2)^{\frac{1}{2}}},$$

by putting

$$a=1.1979,\quad b=0.2858.$$

Let be next

$$\frac{at}{\sqrt{1+b^2t^2}}=v,\quad \frac{a}{\sqrt{1+b^2}}=c;$$

whence there will result

$$\int_1^\infty\frac{adt}{(1+b^2t^2)^{\frac{3}{2}}}e^{-\frac{\alpha^2t^2}{1+b^2t^2}}=\int_c^{\frac{a}{b}}e^{-v^2}dv.$$

Because of the magnitude of $\frac{a}{b}$ which surpasses four, one will be able to replace this limit by ∞, and then one will have

$$\mathrm{Q}=\frac{1}{\sqrt{\pi}}\int_c^\infty e^{-v^2}dv+\frac{1}{\sqrt{2\pi n(1+b^2)}}.$$

By converting this formula into number, by means of the table of Kramp, one has finally

$$\mathrm{Q} = 0.0638,$$

for the probability that there would be concern to determine. That of the contrary event will be 0.9362, so that there is not completely odds of fifteen against one that in Paris the annual births of the boys will exceed those of the girls among the natural infants. There results from it that there is a little more than two against three odds that in an interval of thirteen years, the number of female births will exceed at least one time the one of the male births; because the probability of this event is $1-(0.9362)^{13}$, a quantity equal to 0.424. From 1815 to 1827, it arrives once, in 1815, that the first number has exceeded the second, and the difference has been of ten units.

(21) The limits (b) applied successively to two distinct events, or to the same event at two different epochs, do not make known if the chance of one surpasses that of the other by a given fraction, and what is the probability of this difference. However, it is interesting to compare the probabilities of two events, that one has deduced from observation; it is the solution of this problem which we are going to occupy ourselves now, and of which we will make next application to the cases which present the births of girls and of boys according to their diverse proportions.

We suppose therefore that the event A is arrived s times out of a number m trials, and another event A$'$, x' times out of m'; we suppose also that the four numbers s, $m-s$, s', $m'-s'$, are very great; and we designate by p and p' the respective probabilities of A and A$'$. By virtue of equations (c) and (d), one of the values of p contained between the limits (b) will be represented by

$$p = \frac{s}{m} + \theta\sqrt{\frac{2s(m-s)}{m^3}},$$

and its infinitely small probability by

$$\mathrm{R}{=}\Theta e^{-\theta^2}\frac{d\theta}{\sqrt{\pi}},$$

by making, for brevity,

$$1+2\lambda\theta^3 = \Theta, \quad \frac{(m-2s)\sqrt{2}}{3\sqrt{ms(m-s)}} = \lambda.$$

The extreme values $\pm z$ of θ must be of little consequence in order that the second term of p be very small with respect to the first; nevertheless we will suppose z great enough in order that the probability Z, given by equation (a), is very close to unity, and that one is able, without sensible error, to regard the unknown probability p as contained, with certitude, between the limits (b).

If p' must surpass p by a quantity given and represented by ω, or by a quantity greater than ω, the unknown value of p' will be able to be expressed by the formula:

$$p' = p+\omega+u(1-p-\omega),$$

u being a variable contained between zero and unity. Besides, the event A being arrived s' times out of a number m' trials, the infinitely small probability of any value whatsoever of p', is, according to n$^{\text{o}}$ 14,

$$\frac{(1-p')^{m'-s'}p'^{s'}dp'}{\int_0^1(1-p')^{m'-s'}p'^{s'}dp'}.$$

But the preceding value of p' having place only if the probability of A has a value p which has itself only one probability R, it is necessary to make the product of the probabilities of p and p' in order to have that of this composite event. If one integrates next this product within the limits of the values of p and p', or of the variables θ and u on which p and p' depend, one will have the probability that the value of p' surpasses that of p, by a fraction equal or superior to ω. By designating it by T, and substituting in the numerator for R and p' their expressions, we will have therefore

$$T=\frac{\int_{-z}^{z}\left(\int_0^1[p+\omega+u(1-p-\omega)]^{s'}(1-u)^{m'-s'}du\right)(1-p-\omega)^{m'-s'+1}e^{-\theta^2}\Theta d\theta}{\sqrt{\pi}\int_0^1(1-p')^{m'-s'}p'^{s'}dp'}$$

This put, we designate by h the value of u which renders the coefficient of du a *maximum*, and by H the corresponding value of this coefficient. We will have

$$\frac{(1-p-\omega)s'}{p+\omega+h(1-p-\omega)}-\frac{m'-s'}{1-h}=0;$$

whence one concludes

$$h=\frac{s'-m'(p+\omega)}{m'(1-p-\omega)},$$
$$\mathrm{H}=\left(\frac{s'}{m'}\right)^{s'}\left(\frac{m'-s'}{m'}\right)^{m'-s'}(1-p-\omega)^{s'-m'}.$$

Let there be now

$$[p+\omega+u(1-p-\omega)]^{s'}(1-u)^{m'-s'}=\mathrm{H}e^{-t^2}. \qquad (i)$$

One will be able to suppose the variable t continually increasing with u, so that $t=-\infty$ corresponds to $u=0$, in the case of $p+\omega=1$, and $t=\infty$ to $u=1$. Whatever be this quantity $p+\omega$, if one designates by k a positive quantity, and by $\pm k$ the value of which corresponds to $u=0$, one will have

$$(p+\omega)^{s'}=\mathrm{H}e^{-t^2}. \qquad (k)$$

Moreover $t=0$ corresponding to $u=h$, one must take $t=-k$ or $t=+k$ for the value of t relative to $u=0$, according as the quantity h will be positive or negative, that is to say, according to the expressions of h and of p, according as

$$\frac{s'}{m'}-\frac{s}{m}-\omega > \quad \text{or} \quad < \theta\sqrt{\frac{2s(m-s)}{m^3}};$$

and one will have at the same time

$$\int_0^1 [p+\omega+u(1-p-\omega)]^{s'}(1-u)^{m'-s'}du = \mathrm{H}\int_{\mp k}^{\infty} e^{-t^2}\frac{du}{dt}dt$$

The value of u deduced from equation (i), will be expressed by a series of the form

$$u = h + h't + h''t^2 + h'''t^3 + \text{etc.},$$

of which the coefficients h', h'', h''', etc., independent of t, will be determined as in n$^{\text{o}}$ 4, and will be very small of order of $\frac{1}{\sqrt{m'}}$, $\frac{1}{m'}$, $\frac{1}{m'\sqrt{m'}}$, etc. One finds, for the value of the first two,

$$h' = \frac{1}{1-p-\omega}\sqrt{\frac{2s'(m'-s')}{m'^3}}, \quad h' = \lambda' h',$$

by making, for brevity,

$$\frac{(m'-2s')\sqrt{2}}{3\sqrt{m's'(m'-s')}} = \lambda'.$$

If therefore one neglects the terms multiplied by the quantities h''', h^{iv}, etc., there will result from it

$$\int_0^1 [p+\omega+u(1-p-\omega)]^{s'}(1-u)^{m'-s'}du = \mathrm{H}h'\left(\int_{\mp k}^{\infty} e^{-t^2}dt + \lambda' e^{-k^2}\right).$$

One will have, in the same manner,

$$\int_0^1 (1-p')^{m'-s'}p'^{s'}dp' = (1-p-\omega)^{m'-s'-1}\mathrm{H}h'\int_{-\infty}^{\infty} e^{-t^2}dt;$$

and the expression of T will become

$$\mathrm{T} = \frac{1}{\pi}\int_{-z}^{z}\left(\int_{\mp k}^{\infty} e^{-t^2}dt + \lambda' e^{-k^2}\right)e^{-\theta^2}\Theta d\theta, \qquad (l)$$

where one must take the superior sign or inferior sign before the limit k of the integral relative to t, according as the quantity h will be positive or negative. As for the value of k, it will be the positive root of the formula:

$$k^2 = s'\log\frac{s'}{m'(p+\omega)} + (m'-s')\log\frac{m'-s'}{m'(q-\omega)}, \qquad (m)$$

which is deduced from equation (k), by taking logarithms of these two members, and observing that $p+q=1$.

(22) In the case where one will have

$$\omega = \frac{s'}{m'} - \frac{s}{m},$$

the value of h will be of the same sign as $-\theta$; consequently equation (l) will become

$$\mathrm{T} = \frac{1}{\pi}\int_0^z \left(\int_k^\infty e^{-t^2}dt\right) e^{-\theta^2}\Theta d\theta + \frac{1}{\pi}\int_{-z}^0 \left(\int_{-k}^\infty e^{-t^2}dt\right) e^{-\theta^2}\Theta d\theta, \\ + \frac{\lambda'}{\pi}\int_{-z}^z e^{-(\theta^2+k^2)}\Theta d\theta;$$

and as one has

$$\int_{-k}^\infty e^{-t^2}dt = \sqrt{\pi} - \int_k^\infty e^{-t^2}dt,$$

this value of T will be the same thing as

$$\mathrm{T} = \frac{1}{\sqrt{\pi}}\int_{-z}^0 e^{-\theta^2}\Theta d\theta + \frac{\lambda'}{\pi}\int_{-z}^z e^{-(\theta^2+k^2)}\Theta d\theta \\ + \frac{1}{\pi}\int_0^z \left(\int_k^\infty e^{-t^2}dt\right) e^{-\theta^2}\Theta d\theta - \frac{1}{\pi}\int_{-z}^0 \left(\int_{-k}^\infty e^{-t^2}dt\right) e^{-\theta^2}\Theta d\theta.$$

In making for a moment

$$\sqrt{\frac{2m(m-s)}{m^3}} = f,$$

we will have

$$p+\omega = \frac{s'}{m'} + \theta f, \quad q-\omega = \frac{m'-s'}{m'} - \theta f,$$

and consequently

$$\log\frac{m'(p+\omega)}{s'} = \frac{m'f}{s'}\theta - \frac{m'^2f^2}{2s'^2}\theta^2 + \frac{m'^3f^3}{3s'^3}\theta^3 - \text{etc.}, \\ \log\frac{m'(q-\omega)}{m'-s'} = \frac{m'f}{m'-s'}\theta - \frac{m'^2f^2}{2(m'-s')^2}\theta^2 + \frac{m'^3f^3}{3(m'-s')^3}\theta^3 - \text{etc.},$$

whence one concludes, by virtue of equation (m),

$$k^2 = \frac{m'^3f^2}{2s'(m'-s')}\theta - \frac{(m'-2s')m'^4f^3}{3s'^2(m'-s')^2}\theta^3 + \text{etc.}$$

The coefficients of this series, departing from the second, are of order $\frac{1}{\sqrt{m'}}$, $\frac{1}{m'}$, etc.; by continuing therefore to neglect the quantities of order $\frac{1}{m}$, making, for brevity,

$$\frac{m'^3s(m-s)}{m^3s'(m'-s')} = \mu^2,$$

and considering μ as a positive quantity, there will result from it

$$k = \mu\theta(1-\mu\lambda'\theta), \quad \text{or} \quad k = -\mu\theta(1-\mu\lambda'\theta),$$

according as the variable θ will be positive or negative, so that the value of k is always positive. It will be necessary therefore to employ the first value of k in the third term of T, and the second in the fourth term; to the degree of approximation where we have stopped ourselves, one has besides

$$\int_{\pm\mu\theta(1-\mu\lambda'\theta)}^{\infty} e^{-t^2}dt = \int_{\pm\mu\theta}^{\infty} e^{-t^2}dt \pm \mu^2\lambda'\theta^2 e^{-\mu^2\theta^3};$$

that being, we will have

$$\begin{aligned}\mathrm{T} =& \frac{1}{\sqrt{\pi}}\int_{-z}^{0} e^{-\theta^2}\Theta d\theta + \frac{\lambda'}{\pi}\int_{-z}^{z} e^{-\theta^2(1+\mu^2)}(1+\mu^2\theta^2)\Theta d\theta \\ &+ \frac{1}{\pi}\int_{0}^{z}\left(\int_{\mu\theta}^{\infty} e^{-t^2}dt\right)e^{-\theta^2}\Theta d\theta - \frac{1}{\pi}\int_{-z}^{0}\left(\int_{-\mu\theta}^{\infty} e^{-t^2}dt\right)e^{-\theta^2}\Theta d\theta;\end{aligned}$$

and, according as one will have put for Θ its value, this expression of T will be able to be written thus

$$\begin{aligned}\mathrm{T} = \frac{1}{\sqrt{\pi}}\int_{0}^{z} e^{-\theta^2}d\theta - \frac{2\lambda}{\sqrt{\pi}}\int_{0}^{z} e^{-\theta^2}\theta^3 d\theta + \frac{2\lambda'}{\pi}\int_{0}^{z} e^{-\theta^2(1+\mu^2)}(1+\mu^2\theta^2)d\theta \\ + \frac{4\lambda}{\pi}\int_{0}^{z}\left(\int_{\mu\theta}^{\infty} e^{-t^2}dt\right)e^{-\theta^2}\theta^3 d\theta.\end{aligned}$$

I put, in the last term, $\mu\theta$ and $\mu\theta dt$ under the sign $\int$ instead of t and dt, this which gives

$$\int_{0}^{z}\left(\int_{\mu\theta}^{\infty} e^{-t^2}dt\right)e^{-\theta^2}\theta^3 d\theta = \mu\int_{1}^{\infty}\left(\int_{0}^{z} e^{-\theta^2(1+\mu^2)}\theta^4 d\theta\right)dt,$$

by interchanging the order of integrations. Because of the magnitude that z must have, one is able, without sensible error, to extend to infinity the integrals relative to θ, this which permits to obtain the values of them, and whence there results

$$\mathrm{T} = \frac{1}{2} - \frac{\lambda}{\sqrt{\pi}} + \frac{\lambda'}{\sqrt{\pi(1+\mu^2)}} + \frac{\mu^2\lambda'}{2\sqrt{\pi}(1+\mu^2)^{\frac{3}{2}}} + \frac{3\mu\lambda}{2\sqrt{\pi}}\int_{1}^{\infty}\frac{dt}{(1+\mu^2t^2)^{\frac{3}{2}}},$$

or else, by effecting the integration indicated,

$$\mathrm{T} = \frac{1}{2} + \frac{\lambda' - \mu\lambda}{\sqrt{\pi(1+\mu^2)}} + \frac{\mu^2\lambda' - \mu\lambda}{2\sqrt{\pi}(1+\mu^2)^{\frac{3}{2}}} \qquad (n)$$

This value of T expresses the probability that the difference $p'-p$ surpasses the difference $\frac{s'}{m'} - \frac{s}{m}$, or that one has

$$p' - p > \frac{s'}{m'} - \frac{s}{m}.$$

The contrary event is

$$p' - p > \frac{s'}{m'} - \frac{s}{m}, \quad \text{or} \quad p - p' > \frac{s}{m} - \frac{s'}{m'}.$$

Its probability will be reduced therefore to the same expression of T, by exchanging m' and m, s' and s, that is to say, by exchanging λ' and λ, and putting $\frac{1}{\mu}$ in the place of μ. Thus, in designating it by T′, one will have

$$\mathrm{T}' = \frac{1}{2} + \frac{\mu\lambda - \lambda'}{\sqrt{\pi(1+\mu^2)}} + \frac{\mu\lambda - \mu^2\lambda'}{2\sqrt{\pi}(1+\mu^2)^{\frac{3}{2}}},$$

and consequently $\mathrm{T}+\mathrm{T}' = 1$; this which is able to serve of verification to our calculations.

(23) If one observes that λ and λ' are respectively of order $\frac{1}{\sqrt{m}}$ and $\frac{1}{\sqrt{m'}}$, one sees that the value of T given by equation (n), will differ little from $\frac{1}{2}$, and that the difference $\mathrm{T}-\frac{1}{2}$ will diminish more and more in measure as the numbers m and m' will increase, so that one will have $\mathrm{T}=\frac{1}{2}$, if m and m' were infinite. It is this that it was easy to anticipate; but the calculation alone would be able to make known the value of $\mathrm{T}-\frac{1}{2}$.

When one will have nearly $s = \frac{1}{2}m$ and $s' = \frac{1}{2}m'$, one will be able to neglect completely λ and λ', this which will render null the difference $\mathrm{T}-\frac{1}{2}$. This case holds when one takes for s and s' some numbers of male births and for m and m' the corresponding numbers of births of the two sexes. If there is a question of the infants born at Paris from 1815 to 1827, one will have

$$m = 122404, \quad s = 62239,$$

for the births outside of marriage, and

$$m' = 215639, \quad s' = 109973,$$

for the legitimate infants; whence there results

$$\frac{s'}{m'} - \frac{s}{m} = 0.0015.$$

Relative to France entire, one has, from 1817 to 1826,

$$m = 673067, \quad s = 344482,$$

for the natural infants, and

$$m' = 8983068, \quad s' = 4637084,$$

for the legitimate births; this which gives

$$\frac{s'}{m'} - \frac{s}{m} = 0.0044.$$

There is therefore odds of one against one that the probability of a male birth is greater among the natural infants than among the legitimate infants, of at least 0.0015 in the city of Paris, and of at least 0.0044 in France entire.

(24) When one will have $\frac{s'}{m'} = \frac{s}{m}$, formula ($n$) will express the probability that p' surpasses p. Under this hypothesis, one has

$$\lambda' = \lambda\sqrt{\frac{m}{m'}}, \quad \mu = \sqrt{\frac{m}{m'}}, \quad \mu\lambda' = \lambda;$$

and formula (n) becomes

$$\mathrm{T} = \frac{1}{2} + \frac{(m-m')\lambda}{\sqrt{m'(m+m')}}.$$

The probability that the difference $p' - p$ is positive, will be therefore $>$ or $< \frac{1}{2}$, according as the product $(m-m')\lambda$, or that $(m-m')(2s-m)$ is positive or negative. In order to render reason from this result, it is necessary to observe that in the case of which there is question, the values of p' and p differ very little from one another, of one same quantity $\frac{s'}{m'}$ or $\frac{s}{m}$. But according to the expression of R of n° 21, the probability that p is above $\frac{s}{m}$, has for value:

$$\frac{1}{\sqrt{\pi}}\int_0^z e^{-\theta^2}\Theta d\theta = \frac{1}{2} + \frac{1}{2}\lambda,$$

by substituting infinity in the limit z; the probability that p' is above $\frac{s'}{m'}$ will be the same $\frac{1}{2} + \frac{1}{2}\lambda'$, or $\frac{1}{2} + \frac{1}{2}\lambda\sqrt{\frac{m}{m'}}$; if therefore one supposes, in order to fix ideas, that the quantity λ is positive, these probabilities will be both a little superior to $\frac{1}{2}$; and moreover the excess over $\frac{1}{2}$ will be greater or lesser relatively to the second as in proportion to the first, according as one will have $m >$ or $< m'$; whence one is able to conclude that the probability T that p' surpasses p, must be greater or lesser as that of the contrary event, according as the difference $m - m'$ will be positive or negative; this which accords with the preceding value of T.

(25) We consider actually the case where the quantity $\frac{s'}{m'} - \frac{s}{m} - \omega$, instead of being null as previously, it a very small fraction of the order $\frac{1}{\sqrt{m}}$, and we make

$$\frac{s'}{m'} - \frac{s}{m} - \omega = \alpha\sqrt{\frac{2s(m-s)}{m^3}},$$

α being a quantity of little consequence, positive or negative. The values of k will be deduced from those of n° 22, by putting there $\theta - \alpha$ in the place of θ. Thus we will have

$$k = \mu(\theta-\alpha)[1-\mu\lambda'(\theta-\alpha)],$$

in the case of $\theta > \alpha$, and

$$k = \mu(\alpha-\theta)[1-\mu\lambda'(\alpha-\theta)],$$

in the case of $\alpha > \theta$, since k must always be a positive quantity; and as the value of h will be of the same sign as $\alpha - \theta$, equation (l) will become

$$\mathrm{T} = \frac{\lambda'}{\pi}\int_{-z}^{z} e^{-[\theta^2+\mu^2(\alpha-\theta)^2]}\Theta d\theta$$
$$+\frac{1}{\pi}\int_\alpha^z\left(\int_{k'}^\infty e^{-t^2}dt\right)e^{-\theta^2}\Theta d\theta + \frac{1}{\pi}\int_{-z}^{\alpha}\left(\int_{-k_1}^\infty e^{-t^2}dt\right)e^{-\theta^2}\Theta d\theta$$

k' designating the first value of k, and k_1 the second. One has besides

$$\int_{-k_1}^{\infty} e^{-t^2} dt = \sqrt{\pi} - \int_{k_1}^{\infty} e^{-t^2} dt,$$
$$\int_{k'}^{\infty} e^{-t^2} dt = \mu(\theta-\alpha)\int_{1}^{\infty} e^{-t^2\mu^2(\alpha-\theta)^2} dt + \lambda'\mu^2(\theta-\alpha)^2 e^{-\mu^2(\alpha-\theta)^2},$$
$$\int_{k_1}^{\infty} e^{-t^2} dt = \mu(\alpha-\theta)\int_{1}^{\infty} e^{-t^2\mu^2(\alpha-\theta)^2} dt - \lambda'\mu^2(\alpha-\theta)^2 e^{-\mu^2(\alpha-\theta)^2},$$

whence one concludes

$$\mathrm{T} = \frac{\lambda'}{\pi}\int_{-z}^{z} e^{-[\theta^2+\mu^2(\alpha-\theta)^2]}[1+\mu^2(\alpha-\theta)^2]d\theta + \frac{1}{\sqrt{\pi}}\int_{-z}^{\alpha} e^{-\theta^2}\Theta d\theta$$
$$+\frac{\mu}{\pi}\int_{1}^{\infty}\left(\int_{-z}^{z} e^{-[\theta^2+\mu^2(\alpha-\theta)^2]}(\theta-\alpha)\Theta d\theta\right)dt.$$

One will be able now to replace z by infinity without altering sensibly the value of T. The integrations relative to θ and which have $\pm z$ for limits, will be executed then under finite form; but in order to simplify the result, we will neglect the quantities of order $\frac{1}{\sqrt{m}}$ or of $\frac{1}{\sqrt{m'}}$ that we will have conserved to the present. In this manner, the term of T which has λ' for factor must be suppressed, Θ will be reduced to unity, and one will have simply

$$\mathrm{T} = \frac{1}{2} + \frac{1}{\sqrt{\pi}}\int_{0}^{\alpha} e^{-\theta^2} d\theta + \frac{\mu}{\pi}\int_{1}^{\infty} \phi dt,$$

by making, for brevity,

$$\phi = \int_{-\infty}^{\infty} e^{-[\theta^2+t^2\mu^2(\alpha-\theta)^2]}(\theta-\alpha)d\theta$$

Let be actually

$$\theta = \frac{\theta'}{\sqrt{1+\mu^2t^2}} + \frac{\mu^2t^2\alpha}{1+\mu^2t^2};$$

the limits relative to θ' will be again $\pm\infty$, and we will have

$$\phi = -\frac{\alpha\sqrt{\pi}}{(1+\mu^2t^2)^{\frac{3}{2}}} e^{-\frac{\mu^2\alpha^2t^2}{1+\mu^2t^2}}.$$

We make next

$$\frac{\mu\alpha}{\sqrt{1+\mu^2}} = \beta, \quad \frac{\mu\alpha t}{\sqrt{1+\mu^2t^2}} = u, \quad \frac{\mu\alpha dt}{(1+\mu^2t^2)^{\frac{3}{2}}} = du;$$

for $t=1$ and $t=\infty$, one will have $u=\beta$ and $u=\alpha$; there will result from it therefore

$$\mathrm{T} = \frac{1}{2} + \frac{1}{\sqrt{\pi}}\int_{0}^{\alpha} e^{-\theta^2} d\theta - \frac{1}{\sqrt{\pi}}\int_{\beta}^{\alpha} e^{-u^2} du,$$

or, this which is the same thing

$$\mathrm{T} = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{\beta} e^{-t^2} dt. \qquad (o)$$

(26) By making

$$\frac{2s(m-s)}{m^3} = f^2, \quad \frac{2s'(m'-s')}{m'^3} = f^2,$$

we will have

$$\mu = \frac{f'}{f}, \quad \beta = \frac{\alpha f}{\sqrt{f^2 + f'^2}},$$

and T will be the probability that one has

$$p' - p > \alpha f + \frac{s'}{m'} - \frac{s}{m}.$$

We put in place of α, another quantity α'; we exchange among them the letters s' and s, m' and m; we make

$$\beta' = \frac{\alpha' f'}{\sqrt{f^2 + f'^2}};$$

and we designate by T' that which T becomes by these changes: there will result from it

$$\mathrm{T}' = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{\beta'} e^{-t^2} dt,$$

for the probability that one has

$$p - p' > \alpha' f' + \frac{s}{m} - \frac{s'}{m'}.$$

If one supposes $\alpha' f' = -\alpha f$, this event will be the contrary of the preceding, and the sum of their probabilities must be unity. It is this which holds effectively; because one has then

$$\beta' = -\beta, \quad \int_0^{\beta'} e^{-t^2} dt = -\int_0^{\beta} e^{-t^2} dt,$$

and consequently $\mathrm{T} + \mathrm{T}' = 1$.

I substitute the value of α into that of β; there comes

$$\beta = \frac{\frac{s'}{m'} - \frac{s}{m} - \omega}{\sqrt{f^2 + f'^2}}. \qquad (p)$$

In the case of $\omega = 0$, T will be the probability that p' surpasses p; now, it is easy to be assured that formula (o) coincides then with that which Laplace has given for the same object.[1] One sees also that this formula depends only on the single quantity β, and that

[1] *Théorie analytique des probabilités*, page 383.

by a very simple modification in the expression of β, this formula (o) extends to the case where p' must surpass p by a fraction equal or superior to ω. In all the cases, the expression of T and the analysis which has led us there suppose that $\frac{s'}{m'} - \frac{s}{m} - \omega$ is a very small fraction; but if that had not held, there would be a very great and useless to calculate probability, that the difference $p' - p$, which differs little from $\frac{s'}{m'} - \frac{s}{m}$, would be greater than ω, setting aside the sign.

(27) We apply now formulas (o) and (p) to the numbers m and m' of births of the two sexes, and to the corresponding numbers s and s' of male births.

In the first example of n° 23, and if one takes $\omega = 0$, one finds

$$\beta = 0.5987, \quad \int_\beta^\infty e^{-t^2} dt = 0.3519;$$

whence T$= 0.8015$ results, or a little more than $\frac{4}{5}$ for the probability that at Paris the chance of a male birth is greater among the legitimate infants than among the natural infants.

In the second example of the same number, and by taking again $\omega = 0$, one has

$$\beta = 4.9186;$$

this which gives for T a value which does not differ sensibly from unity, so that one is able to regard as certain that in France entire, the probability of a male birth is greater among the legitimate infants than among the infants born outside of marriage. In the same example, if one takes $\omega = 0.003$, one will have

$$\beta = 1.5604, \quad \int_\beta^\infty e^{-t^2} dt = 0.02423,$$

and T$= 0.9863$ will result from it; so that it is again very probable that the first chance surpasses the second of a fraction equal or superior to 0.003.

One has found, at the beginning of this century, that there is born in a part of France and in an interval of three years, 110312 boys and 105287 girls. The ratio of the first number to the second is nearly $\frac{22}{21}$ instead of $\frac{16}{15}$ which holds now for France entire; one is able therefore to desire known the probability that the chance of a birth of a boy is greater in the second case than in the first; now, if one makes

$$m = 215599, \qquad s = 110312,$$
$$m' = 9656135, \qquad s' = 4981566,$$

and if one takes $\omega = 0$, there comes

$$\beta = 2.7557;$$

this which gives for $1-$T a quantity of the order of those that we have neglected, so that it is beyond doubt that the chance of a male birth was less in the part of France and in the epoch of which there is a concern, than it is in all this realm and at the actual epoch. If one takes $\omega = 0.003$, one finds

$$\beta = 0.8068, \quad \int_\beta^\infty e^{-t^2} dt = 0.2251;$$

and there will result T= 0.8730 from it, for the probability that the second chance surpasses the first by a fraction equal or superior to 0.003.

For the last example, let there be

$$m = 993191, \qquad s = 511898,$$
$$m' = 944125, \qquad s' = 488457,$$

m and m' will be the numbers of births of the two sexes which have taken place in France in the years 1817 and 1826, and s and s' the corresponding numbers of the male births; and if we take $\omega = 0$, we will have

$$\beta = 0.8258, \quad \int_{\beta}^{\infty} e^{-t^2} dt = 0.2152;$$

whence T= 0.8786 will result, for the probability that the chance of a male birth has been greater in the second year than in the first. In taking $\omega = 0.001$, one finds

$$\beta = 0.4017, \quad \int_{\beta}^{\infty} e^{-t^2} dt = 0.5050;$$

this which gives T= 0.7151, for the probability that the first chance has surpassed the second by a fraction greater than a thousandth. As the difference of the ratios $\frac{s'}{m'}$ and $\frac{s}{m}$ is 0.00196, there results from n° 23 that it is possible that the difference of these two chances has been rather below than above two thousandths. Besides, from 1817 to 1826, these two extreme years have those for which the proportion of the births of the two sexes, is most separated, on both sides, from its mean value. We are able therefore to conclude that at the actual epoch and for France entire, the probability of a male birth sustains only very small variations from one year to another, and take for its value, the mean of the ten years that we have considered, that is to say, 0.5159. In the ignorance where we are of the cause which renders the births of boys dominating, it will be experience alone which will be able to decide if this probability will vary more through the following, or if it will remain nearly constant. Observation has not yet taught us if it changes in one same year with the seasons; we do not know either if it is the same in different nations; we know only that it depends on the state of society, since the number of births outside of marriage influences sensibly on the proportion of the male and female births.

(28) The determination of the ratio which exists between the annual births of the two sexes in a great population is able also to be considered as a problem relative to this part of the calculation of the chances which treats of the mean result of the observations and of its degree of probability. For that, it would be necessary to suppose 1° that there exists one value of this ratio, such that of the equal deviations on both sides, are equally probable; 2° that this unknown value remains constant during all the series of observations. One would take then for this value, the mean result of a long sequence of years; and the calculation would make known, according to the set of observations, the probability that the excess of this result over the exact value, is comprehended between some given limits. The calculation would furnish also some conditions to which the observations must satisfy in order to be compatible with the double hypothesis which

one just enunciated. But in order that the formulas of the calculation of the probability of which there is concern are independent of the law of probability of the deviations which is not given to us, it is necessary that the observations had been made in considerable number; this which does not permit to apply these formulas to the research of the ratio of the annual births of the two sexes, of which we know well only the ten values observed in France from 1817 to 1826. Relative to the probability of the mean results in general, we return to a Memoir inserted in the *Additions à la connaissance des temps pour l'année* 1832.