BOOK II
CHAPTER VI
DE LA PROBABILITÉ DES CAUSES ET DES ÉVÈNEMENTS FUTURS, TIRÉE DES ÉVÈNEMENTS OBSERVÉS

Pierre Simon Laplace

Théorie Analytique des Probabilités
3rd Edition (1820), §§26–33, pp. 363–401

ON THE PROBABILITY OF CAUSES AND OF FUTURE EVENTS, DEDUCED FROM OBSERVED EVENTS

§26. The probability of the greater part of simple events, is unknown: by considering it *a priori*, it appears to us susceptible of all the values comprehended between zero and unity; but, if we have observed a result composed of many of these events, the manner by which they enter there, renders some of these values more probable than the others. Thus in measure as the observed result is composed by the development of the simple events, their true possibility is made more and more known, and it becomes more and more probable that it falls within some limits which being tightened without ceasing, would end by coinciding, if the number of simple events became infinite. In order to determine the laws according to which this possibility is discovered, we will name it x. The theory exposed in the preceding chapters, will give the probability of the observed result, as a function of x. Let y be this function; if we consider the different values of x as so many causes of this result, the probability of x will be, by the third principal of §1, equal to a fraction of which the numerator is y, and of which the denominator is the sum of all the values of y; by multiplying therefore the numerator and the denominator of this fraction by dx, this probability will be [363]

$$\frac{y\,dx}{\int y\,dx},$$

the integral of the denominator being taken from $x = 0$ to $x = 1$. The probability that the value of x is comprehended within the limits $x = \theta$, and $x = \theta'$ is consequently equal to

$$\frac{\int y\,dx}{\int y\,dx}, \tag{1}$$

the integral of the numerator being taken from $x = \theta$ to $x = \theta'$, and that of the denominator [364]
being taken from $x = 0$ to $x = 1$.

The most probable value of x, is that which renders y a *maximum*. We will designate it by a. If at the limits of x, y is null, then each value of y has a corresponding equal value on the other side of the *maximum*.

When the values of x, considered independently of the observed result, are not equally possible; by naming z the function of x which expresses their probability; it is easy to see, by that which has been said in first chapter of this Book, that by changing in formula (1), y into yz, we will have the probability that the value of x is comprehended within the limits $x = \theta$ and $x = \theta'$. This reverts to supposing all the values of x equally possible *a priori*, and by considering the observed result, as being formed of two independent results, of which the probabilities are y and z. We are able to restore thus all the cases to the one where we suppose *a priori*, before the event, an equal possibility to the different values of x, and, by this reason, we will adopt this hypothesis in that which will follow.

We have given in §22 and the following of the first Book, the formulas necessary in order to determine by some convergent approximations, the integrals of the numerator and of the denominator of formula (1), when the simple events of which the observed event is composed, are repeated a very great number of times; because then y has for factors, functions of x raised to very great powers. We will, by means of these formulas, determine the law of probability of the values of x, in measure as they deviate from the value a, the most probable, or which renders y a *maximum*. For that, let us resume formula (c) of §27 of the first Book,

$$\begin{aligned}\int y\,dx =& Y\left\{U+\frac{1}{2}\cdot\frac{d^2U^3}{1.2\,dx^2}+\frac{1.3}{2^2}\cdot\frac{d^4U^5}{1.2.3.4\,dx^4}+\text{etc.}\right\}\int dt\,c^{-t^2}\\ &+\frac{Y}{2}c^{-T^2}\left\{\frac{dU^2}{dx}-T\cdot\frac{d^2U^3}{1.2\,dx^2}+(T^2+1)\cdot\frac{d^3U^4}{1.2.3\,dx^3}-\text{etc.}\right\}\\ &-\frac{Y}{2}c^{-T'^2}\left\{\frac{dU^2}{dx}+T'\cdot\frac{d^2U^3}{1.2\,dx^2}+(T'^2+1)\cdot\frac{d^3U^4}{1.2.3\,dx^3}+\text{etc.}\right\};\end{aligned} \tag{2}$$

ν is equal to $\frac{x-a}{\sqrt{\log Y-\log y}}$, and U, $\frac{dU^2}{dx}$, $\frac{d^2U^3}{dx^2}$, etc. are that which ν, $\frac{d\nu^2}{dx}$, $\frac{d^2\nu^3}{dx^2}$, etc. become, [365]
when we change after the differentiations, x into a, a being the value of x which renders y a *maximum*: T is equal to that which the function $\sqrt{\log Y-\log y}$ becomes, when we change x into $a-\theta$ in y, and T' is that which the same function becomes, when we change x into $a+\theta'$. The preceding expression of $\int y\,dx$ gives the value of this integral, within the limits $x=a-\theta$ and $x=a+\theta'$, the integral $\int dt\,c^{-t^2}$ being taken from $t=-T$ to $t=T'$.

Most often, at the limits of the integral $\int y\,dx$, extended from $x=0$ to $x=1$, y is null; now when y is not null, it becomes so small at these limits, that we are able to suppose it null. Then, we can make at these limits T and T' infinite, that which gives for the integral $\int y\,dx$, extended from $x=0$ to $x=1$,

$$\int y\,dx = Y\left\{U+\frac{1}{2}\cdot\frac{d^2U^3}{1.2\,dx^2}+\frac{1.3}{2^2}\cdot\frac{d^4U^5}{1.2.3.4\,dx^4}+\text{etc.}\right\}\sqrt{\pi};$$

thus the probability that the value of x is comprehended within the limits $x=a-\theta$ and $x=a+\theta'$ is equal to

$$\frac{\int dt\,c^{-t^2}}{\sqrt{\pi}}+\frac{\left\{\begin{array}{l}\frac{1}{2}c^{-T^2}\left\{\frac{dU^2}{dx}-T\cdot\frac{d^2U^3}{1.2\,dx^2}+(T^2+1)\cdot\frac{d^3U^4}{1.2.3\,dx^3}-\text{etc.}\right\}\\ -\frac{1}{2}c^{-T'^2}\left\{\frac{dU^2}{dx}+T'\cdot\frac{d^2U^3}{1.2\,dx^2}+(T'^2+1)\cdot\frac{d^3U^4}{1.2.3\,dx^3}+\text{etc.}\right\}\end{array}\right\}}{\left\{U+\frac{1}{2}\cdot\frac{d^2U^3}{1.2\,dx^2}+\frac{1.3}{2^2}\cdot\frac{d^4U^5}{1.2.3.4\,dx^4}+\text{etc.}\right\}\sqrt{\pi}}; \tag{3}$$

We see, by §23 of the first Book, that in the case where y has for factors, some functions of x raised to great powers of order $\frac{1}{\alpha}$, α being an extremely small fraction, then U is most often of order $\sqrt{\alpha}$, so that its successive differences; U, $\frac{dU^2}{dx}$, $\frac{d^2U^3}{dx^2}$, etc. are respectively of the orders $\sqrt{\alpha}$, α, $\alpha^{\frac{3}{2}}$, etc.; whence it follows that the convergence of the series of formula (3), requires that T and T' are not of an order superior to $\frac{1}{\sqrt{\alpha}}$.

If we suppose $\theta=\theta'$, then we have very nearly $T=T'$, and formula (3) is reduced, by [366]
neglecting the terms of order α, to the integral $\frac{\int dt\,c^{-t^2}}{\sqrt{\pi}}$, taken from $t=-T$ to $t=T'$; that which reverts in neglecting the square of the difference T'^2-T^2, to doubling the preceding

integral, and to taking it from t null to

$$t = \sqrt{\frac{T^2 + T'^2}{2}}:$$

now we have

$$T^2 = \log Y - \log y,$$

and we can suppose

$$\log y = \frac{1}{\alpha}\log\phi,$$

ϕ being a function of x or of $a-\theta$, which no longer contains factors raised to great powers; by naming therefore Φ, $\frac{d\Phi}{dx}$, $\frac{d^2\Phi}{dx^2}$, etc., that which ϕ, $\frac{d\phi}{dx}$, $\frac{d^2\phi}{dx^2}$, etc. become, when θ is null; by observing next that the condition of Y or Φ, a *maximum*, gives $\frac{d\Phi}{dx} = 0$, we will have

$$\alpha T^2 = -\theta^2\frac{dd\Phi}{2\Phi\, dx^2} + \theta^3\frac{d^3\Phi}{6\Phi\, dx^3} - \frac{\theta^4}{8}\left[\frac{d^4\Phi}{3\Phi\, dx^4} - \left(\frac{dd\Phi}{\Phi\, dx^2}\right)^4\right] + \text{etc.}$$

By changing θ into $-\theta$, we will have the value of $\alpha T'^2$; we will have therefore, by neglecting the terms of order α^2,

$$\frac{\alpha(T^2 + T'^2)}{2} = -\theta^2\frac{dd\Phi}{2\Phi\, dx^2};$$

hence,

$$\sqrt{\frac{T^2 + T'^2}{2}} = \frac{\theta}{\sqrt{\alpha}}\sqrt{-\frac{dd\Phi}{2\Phi\, dx^2}}.$$

Let us make

$$k = \sqrt{-\frac{dd\Phi}{2\Phi\, dx^2}} = \sqrt{-\frac{\alpha ddY}{2Y dx^2}},$$

$$\theta = \frac{t\sqrt{\alpha}}{k};$$

the probability that the value of x is comprehended within the limits $a \pm \frac{t\sqrt{\alpha}}{k}$ will be [367]

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

the integral being taken from $t = 0$, and being able to be obtained in a very close manner, from the formulas of §27 from the first Book.

There results from this expression, that the most probable value of x is a, or that which renders the observed event, the most probable; and that by multiplying to infinity the simple events of which the observed event is composed, we are able at the same time to narrow the limits $a \pm \frac{t\sqrt{\alpha}}{k}$, and to increase the probability that the value of x will fall between these limits; so that at infinity, this interval becomes null, and the probability is confounded with certitude.

If the observed event depends on simple events of two different kinds, by naming x and x' the possibilities of these two kinds of events, we will see by the preceding reasonings, that y being then the probability of the composite event, the fraction

$$\frac{y\,dx\,dx'}{\iint y\,dx\,dx'}, \tag{4}$$

will be the probability of the simultaneous values of x and of x', the integrals of the denominator being taken from $x = 0$ to $x = 1$, and from $x' = 0$ to $x' = 1$. By naming a and a' the values of x and x' which render y a *maximum*, and making $x = a + \theta$, $x' = a' + \theta'$, we will find, by the analysis of §27 from the first Book, that if we suppose

$$\frac{\theta}{\sqrt{2Y}}\sqrt{-\left(\frac{ddY}{dx^2}\right)} - \theta'\frac{\left(\frac{ddY}{dx\,dx'}\right)}{2Y}\sqrt{\frac{2Y}{-\left(\frac{ddY}{dx^2}\right)}} = t,$$
$$\frac{\theta'}{\sqrt{-2Y\left(\frac{ddY}{dx^2}\right)}}\sqrt{\left(\frac{ddY}{dx^2}\right)\left(\frac{ddY}{dx'^2}\right) - \left(\frac{ddY}{dxdx'}\right)^2} = t',$$

the fraction (4) will take form [368]

$$\frac{dt\,dt'c^{-t^2-t'^2}}{\iint dt\,dt'c^{-t^2-t'^2}}.$$

The integrals of the denominator must be taken from $t = -\infty$ to $t = \infty$, and from $t' = -\infty$ to $t' = \infty$; because the integrals relative to x and x' of the fraction (4) being taken from $x = 0$ and $x' = 0$ to x and x' equal to unity, and at these limits, the values of θ and of θ' being $-a$ and $1 - a$, $-a'$ and $1 - a'$, the limits of t and of t' are equal to these last limits multiplied by some quantities of order $\frac{1}{\sqrt{\alpha}}$: thus the exponential $c^{-t^2-t'^2}$ is excessively small at these limits, and we can without sensible error, extend the integrals of the denominator of the preceding fraction, to the positive and negative infinite values of the variables t and t'. This denominator becomes thus equal to π; and the probability that the values of θ' and of θ are comprehended within the limits

$$\theta' = 0, \qquad \theta' = \frac{t'\sqrt{-2Y\left(\frac{ddY}{dx^2}\right)}}{\sqrt{\left(\frac{ddY}{dx^2}\right)\left(\frac{ddY}{dx'^2}\right) - \left(\frac{ddY}{dxdx'}\right)^2}},$$
$$\theta = 0, \qquad \theta = \frac{t\sqrt{2Y}}{\sqrt{-\left(\frac{ddY}{dx^2}\right)}} + \frac{t'\left(\frac{ddY}{dxdx'}\right)}{\sqrt{\left(\frac{ddY}{dx^2}\right)\left(\frac{ddY}{dx'^2}\right) - \left(\frac{ddY}{dxdx'}\right)^2}}\sqrt{\frac{2Y}{-\left(\frac{ddY}{dx^2}\right)}},$$

is equal to

$$\frac{1}{\pi}\iint dt\,dt'c^{-t^2-t'^2},$$

the integrals being taken from t and t' nulls.

We see by this formula, that in the case of two different kinds of simple events, the probability that their respective possibilities are those which render the composite event, most probable, becomes more and more great, and ends by being confounded with certitude; that which holds generally for any number whatsoever of different kinds of simple events, which enter into the observed event.

If we imagine an urn containing an infinity of balls of many different colors, and if after having drawn a great number n from it, p out of this number, had been of the first color, q of the second, r of the third, etc.; by designating by x, x', x'', etc. the respective probabilities to bring forth in a single drawing, one of these colors, the probability of the observed event will be the term which has for factor $x^p x'^q x''^r$.etc., in the development of the polynomial [369]

$$(x + x' + x'' + \cdots)^n,$$

where we have

$$x + x' + x'' + \text{etc.} = 1,$$
$$p + q + r + \text{etc.} = n;$$

we will be able therefore to suppose here $y = x^p x'^q x''^r$.etc.; and then we have for the values of x, x', x'', etc. which render the observed event the most probable

$$x = \frac{p}{n}, \quad x' = \frac{q}{n}, \quad x'' = \frac{r}{n}, \quad \text{etc.}$$

Thus the most probable values are proportional to the numbers of the arrivals of the colors; and when the number n is a great number, the respective probabilities of the colors, are very nearly equal to the numbers of times that they have arrived, divided by the number of drawings.

§27. In order to give an application of the preceding formula, let us consider the case where two players A and B play together with this condition, that the one who out of three coups will have won two of them, wins the game; and let us suppose that out of a very great number n of games, A has won a number i of them. By naming x the probability of A to win a coup, and consequently $1 - x$, the corresponding probability of B; the probability of A to win a game, will be the sum of the first two terms of the binomial $(x + 1 - x)^3$, and the corresponding probability of B, will be the sum of the last two terms. These probabilities are therefore $x^2(3 - 2x)$ and $(1 - x)^2(1 + 2x)$; thus the probability that out of n games, A will win i of them, and B, $n - i$, will be proportional to $x^{2i}(3-2x)^i(1-x)^{2n-2i}(1+2x)^{n-i}$. By naming therefore y this function, and a the value of x which renders it a *maximum*, the [370]
probability that the value of x is comprehended within the limits $a - \theta$ and $a + \theta$ will be

$$\frac{\int y\,dx}{\int y\,dx},$$

the integral of the numerator being taken from $x = a - \theta$ to $x = a + \theta$, and that of the denominator being taken from $x = 0$ to $x = 1$. If we make

$$\frac{1}{n} = \alpha, \quad \frac{i}{n} = i',$$

we will have by the preceding section,

$$\phi = x^{2i'}(3-2x)^{i'}(1-x)^{2-2i'}(1+2x)^{1-i'}.$$

The condition of the *maximum* of y or of ϕ, gives $d\phi = 0$; consequently a being the value of x corresponding to this *maximum*, we will have

$$0 = \frac{2i'}{a} - \frac{2i'}{3-2a} - \frac{2(1-i')}{1-a} + \frac{2(1-i')}{1+2a};$$

whence we deduce

$$i' = a^2(3-2a), \quad 1-i' = (1-a)^2(1+2a);$$

next we have

$$\frac{-dd\Phi}{2\Phi dx^2} = \frac{18}{(3-2a)(1+2a)} = k^2.$$

The probability that the value of x is comprehended within the limits $a \pm \frac{r}{\sqrt{n}}$, will be therefore, by the preceding section, equal to

$$\frac{6\sqrt{2}}{\sqrt{\pi(3-2a)(1+2a)}}\int dr\, c^{\frac{-18r^2}{(3-2a)(1+2a)}}.$$

We will see easily that this result agrees with the one that we have found in §16, by an analysis less direct than this one.

The game ends in two coups, if A or B wins the first two coups, the third coup not being played, because it becomes useless. Thus the numbers of games won by one and the [371] other of the players, does not indicate the number of games played; but they indicate that this last number is contained within some given limits, with a probability that increases without ceasing, in measure as the games are multiplied. The investigation of this number and of this probability being very proper to clarify the preceding analysis; we will occupy ourselves with it.

The probability that A will win a game in two coups, is x^2, x expressing, as above, his probability to win at each coup. The probability that he will win the game in three coups, is $2x^2(1-x)$. The sum $x^2(3-2x)$ of these two probabilities, is the probability that A will win the game. Thus in order to have the probability that out of i games won by player A, s will be of two coups, it is necessary to raise to the power i, the binomial

$$\frac{x^2}{x^2(3-2x)} + \frac{2x^2(1-x)}{x^2(3-2x)}$$

or

$$\frac{1}{3-2x} + \frac{2(1-x)}{3-2x},$$

and the term $i - s + 1$ of the development of this power, will be that probability which is thus equal to

$$\frac{1.2.3\ldots i.2^{i-s}(1-x)^{i-s}}{1.2.3\ldots s.1.2.3\ldots(i-s)(3-2x)^i}.$$

The greatest term of this development is, by §16, the one in which the exponents s and $i - s$ of the first and of the second term of the binomial are very nearly in the ratio of these terms, that which gives

$$s = \frac{i}{3-2x}.$$

We will name s' this quantity, and we will make

$$s = s' + l.$$

We will have, by §16,

$$\sqrt{\frac{i}{2s'\pi(i-s')}}dl\, c^{\frac{-il^2}{2s'(i-s')}}$$

for the probability of s, corresponding to the skill x of player A. [372]

We will find similarly that, if we name z the number of the games of two coups, won by player B, out of the number $n - i$ of games that he has won; the most probable value of z will be $\frac{n-i}{1+2x}$; and that by designating by z' this quantity, and making

$$z = z' + l',$$

the probability of z corresponding to x will be

$$\sqrt{\frac{n-i}{2z'(n-i-z')\pi}}dl'\, c^{\frac{-(n-i)l'^2}{2z'(n-i-z')}}.$$

The product of these two probabilities is therefore the probability corresponding to x, that the number of games of two coups, won by player A, will be $s' + l$, while the number of games of two coups, won by player B, will be $z' + l'$. Let

$$q = \frac{i}{2s'(i-s')}, \quad q' = \frac{n-i}{2z'(n-i-z')};$$

we will have, for this composite probability,

$$\frac{\sqrt{qq'}}{\pi}dl\, dl' c^{-ql^2-q'l'^2}.$$

It is necessary to multiply this probability by that of x, which, as we have seen in the preceding section, is $\frac{y\,dx}{\int y\,dx}$; the product is

$$\frac{\sqrt{qq'}}{\pi}\,\frac{y\,dx}{\int y\,dx}dl\, dl'\, c^{-ql^2-q'l'^2}; \qquad (\epsilon)$$

the integral of the denominator must be taken from $x = 0$ to $x = 1$; and by §27 of the first Book, this integral is very nearly,

$$Y\sqrt{\pi}\sqrt{-\frac{2Y\,dx^2}{ddY}}.$$

If we name X the function

$$\sqrt{qq'}\,c^{-ql^2-q'l'^2},$$

and if we designate by a' the value of x, which renders Xy a *maximum*, and by X' and Y', that which X and y become, when we change x into a' there; we will have, by the preceding section, by making $x = a' + \theta$, [373]

$$ydx\sqrt{qq'}\,c^{-ql^2-q'l'^2} = Y'X'd\theta\,c^{\frac{\theta^2 d^2(X'Y')}{2X'Y'dx^2}}.$$

It is easy to see that a' differs from the value a of x, which renders y a *maximum*, only by a quantity of order α, which we will designate by $f\alpha$; by substituting into Y, $a + f\alpha$ instead of a', in order to form Y', and developing with respect to the powers of α, we will see that $\frac{dY}{da}$ being null, because Y is the *maximum* of y, Y' differs from Y, only by quantities of order α; thus we have, to the quantities near of an order inferior to the one that we conserve, and by observing that $\frac{dX'}{X'dx}$ and $\frac{d^2X'}{X'dx^2}$ can be neglected with respect to $\frac{dY'}{Y'dx}$,

$$\frac{d^2X'Y'}{2X'Y'dx^2} = \frac{d^2Y}{2Y\,dx^2};$$

the function (ϵ) becomes thence

$$\frac{\sqrt{qq'}}{\pi\sqrt{\pi}}\sqrt{-\frac{ddY}{2Y\,dx^2}}\,dl\,dl'd\theta\,c^{-ql^2-q'l'^2+\frac{\theta^2 ddY}{2Y\,dx^2}}. \qquad (\epsilon')$$

We must in this function, suppose $x = a$, that which gives, by substituting for i, its value $na^2(3-2a)$,

$$q = \frac{3-2a}{4na^2(1-a)}, \quad q' = \frac{1+2a}{4na(1-a)^2}.$$

Next, x being equal to $a' + \theta$, it is equal to $a + f\alpha + \theta$; by neglecting therefore the quantities of order α, we will have

$$x = a + \theta.$$

Now the number of games of two coups, being

$$\frac{i}{3-2x} + \frac{n-i}{1+2x} + l + l',$$

this number will be [374]

$$\frac{i}{3-2a} + \frac{n-i}{1+2a} + \left[\frac{2i}{(3-2a)^2} - \frac{2(n-i)}{(1+2a)^2}\right]\theta + l + l'.$$

Let us make

$$t = \left[\frac{2i}{(3-2a)^2} - \frac{2(n-i)}{(1+2a)^2}\right]\theta + l + l';$$

and let us designate by q'' the quantity

$$-\frac{ddY}{2Y\,dx^2\left[\frac{2i}{(3-2a)^2} - \frac{2(n-i)}{(1+2a)^2}\right]^2},$$

which, after all the reductions, is reduced to

$$\frac{9(3-2a)(1+2a)}{2n(1-2a)^2(3-2a+2a^2)^2};$$

the function (ϵ') will become

$$\frac{\sqrt{qq'q''}}{\pi\sqrt{\pi}}\,dt\,dl\,dl'\,c^{-ql^2-q'l'^2-q''(t-l-l')^2}. \qquad (\epsilon'')$$

By integrating it from $l=-\infty$ to $l=\infty$, and from $l'=-\infty$ to $l'=\infty$, we will have the probability that the number of games of two coups, will be equal to

$$\frac{i}{3-2a} + \frac{n-i}{1+2a} + t;$$

now we have

$$\int dl\,c^{-ql^2-q'l'^2-q''(t-l-l')^2}$$
$$= \int dl\,c^{-\frac{qq''}{q+q''}(t-l')^2-q'l'^2-(q+q'')\left[l-\frac{q''}{q+q''}(t-l')\right]^2}.$$

This last integral, taken from $l=-\infty$ to $l=\infty$, is, by that which precedes,

$$\frac{\sqrt{\pi}}{\sqrt{q+q''}}c^{-\frac{qq''}{q+q''}(t-l')^2-q'l'^2}.$$

By multiplying it by dl', and by putting it under this form, [375]

$$\frac{\sqrt{\pi}\,dl'}{\sqrt{q+q''}}c^{-\frac{qq'q''t^2}{qq'+qq''+q'q''}-\frac{qq'+qq''+q'q''}{q+q''}\left(l'-\frac{qq''t}{qq'+qq''+q'q''}\right)^2},$$

and integrating from $l'=-\infty$ to $l'=\infty$; we will have

$$\frac{\pi}{\sqrt{qq'+qq''+q'q''}}\,c^{-\frac{qq'q''t^2}{qq'+qq''+q'q''}}$$

The function (ϵ'') integrated with respect to l and l', within the positive and negative infinite limits of these variables, becomes thus

$$\frac{1}{\sqrt{\pi}}\sqrt{\frac{qq'q''}{qq'+qq''+q'q''}}\, c^{-\frac{qq'q''t^2}{qq'+qq''+q'q''}}.$$

Thus the probability that the number of games of two coups, will be comprehended within the limits

$$\frac{i}{3-2a}+\frac{n-i}{1+2a}\pm t = n(a^2+\overline{1-a}^2)\pm t,$$

is equal to the double of the integral of the preceding differential, taken from t null. We must observe that q, q', q'' are of order $\frac{1}{n}$, so that the quantity $\frac{qq'q''}{qq'+qq''+q'q''}$ is of the same order. Let us represent it by $\frac{k'^2}{n}$, and let us make $t = r\sqrt{n}$; we will have

$$\frac{2}{\sqrt{\pi}}\int k'dr\, c^{-k'^2r^2}, \qquad (\epsilon''')$$

for the expression of the probability that the number of games of two coups, will be comprehended within the limits

$$n(a^2+\overline{1-a}^2)\pm r\sqrt{n},$$

the integral being taken from r null. The interval of these two limits is $2r\sqrt{n}$, and the ratio of this interval to the number n of games, is $\frac{2r}{\sqrt{n}}$; this ratio diminishes without ceasing, [376]
in measure as n increases, and r can increase at the same time indefinitely; so that the preceding integral approaches indefinitely unity.

The total number of coups, is the triple of the number of games of three coups, plus the double of the number of games of two coups, or the triple of the total number n of games, less the number of games of two coups; it is therefore

$$2n(1+a-a^2)\mp r\sqrt{n},$$

The integral (ϵ''') is therefore the expression of the probability that the number of coups will be comprehended within these limits.

If instead of knowing the number i of games won by player A, and the total number n of games, we knew the number i and the total number of coups; the same analysis will be able to serve to determine the unknown number n of games. For this, let us designate by h, the total number of coups; we will have, by that which precedes, the two equations

$$3n-\frac{i}{3-2a}-\frac{n-i}{1+2a} = h\pm r\sqrt{n},$$
$$\frac{i}{a}-\frac{i}{3-2a} = \frac{n-i}{1-a}-\frac{n-i}{1+2a}.$$

These equations give a and n as functions of $h\pm r\sqrt{n}$. Let us suppose

$$n = i\psi\left(\frac{h\pm r\sqrt{n}}{i}\right),\quad a = \Gamma\left(\frac{h\pm r\sqrt{n}}{i}\right);$$

we will have, by reducing into series,

$$n = i\psi\left(\frac{h}{i}\right) \pm ir\sqrt{n}\frac{d\psi\left(\frac{h}{i}\right)}{dh} + \text{etc.};$$

we will substitute into k', instead of n and of a, $i\psi\left(\frac{h}{i}\right)$ and $\Gamma\left(\frac{h}{i}\right)$: the integral (ϵ''') is then the probability that the number n of games, is comprehended within the limits

$$i\psi\left(\frac{h}{i}\right) \pm ir\sqrt{i\psi\left(\frac{h}{i}\right)\frac{d\psi\left(\frac{h}{i}\right)}{dh}}.$$

§28. It is principally in the births, that the preceding analysis is applicable, and we are able to deduce from it, not only for the human race, but for all the kinds of organized beings, some interesting results. Until here the observations of this kind have been made in great number, only on the human race: we will submit the principals to the calculus. [377]

Let us consider first the births observed at Paris, at London and in the realm of Naples. In the space of 40 years elapsed from the commencement of 1745, an epoch where we have begun to distinguish at Paris, out of the registers, the births of two sexes, to the end of 1784, we have baptized in this capital, 393386 boys, and 377555 girls, the found infants being comprehended in this number: this gives nearly $\frac{25}{24}$ for the ratio of the baptisms of the boys to those of the girls.

In the space of 95 years elapsed from the commencement of 1664 to the end of 1758, there was born at London, 737629 boys, and 698958 girls; that which gives $\frac{19}{18}$ nearly, for the ratio of the births of boys to those of girls.

Finally, in the space of nine years elapsed, from the commencement of 1774 to the end of 1782, there was born in the realm of Naples, Sicily not included, 782352 boys, and 746821 girls; that which gives $\frac{22}{21}$ for the ratio of the births of the boys to those of the girls.

The smallest of these numbers of births, are relative to Paris; besides, it is in this city that the births of the boys and of the girls, more approach equality. For these two reasons, the probability that the possibility of the birth of a boy surpasses $\frac{1}{2}$, must be less than at London and in the realm of Naples. Let us determine numerically this probability.

Let us name p the number of masculine births observed at Paris, q the one of the feminine births, and x the possibility of a masculine birth, that is the probability that an infant who must be born, will be a boy; $1 - x$ will be the possibility of a feminine birth, and we will have the probability that out of $p+q$ births, p will be masculine, and q will be feminine, [378]
equal to

$$\frac{1.2.3\ldots(p+q)}{1.2.3\ldots p.1.2.3\ldots q}x^p(1-x)^q;$$

by making therefore

$$y = x^p(1-x)^q,$$

the probability that the value of x is comprehended within some given limits, will be by §26, equal to

$$\frac{\int y\,dx}{\int y\,dx},$$

the integral of the denominator being taken from $x = 0$ to $x = 1$, and that of the numerator being taken within the given limits. If we take zero and $\frac{1}{2}$ for these limits, we will have the probability that the value of x not surpass $\frac{1}{2}$. The value which corresponds to the *maximum* of y is $\frac{p}{p+q}$; and seeing the magnitude of the numbers p and q, the excess of $\frac{p}{p+q}$ over $\frac{1}{2}$, is too considerable in order to employ here formula (c) from §27 of the first Book, in the approximation of the integral $\int y\,dx$, taken from $x = 0$ to $x = \frac{1}{2}$; it is necessary therefore, in this case, to make use of formula (A) from §22 of the same Book. Here we have

$$\nu = -\frac{y\,dx}{dy} = -\frac{x(1-x)}{p-(p-q)x};$$

the formula cited (A) gives thus for the integral $\int y\,dx$, taken from $x = 0$ to $x = \frac{1}{2}$,

$$\frac{1}{2^{p+q+1}(p-q)}\left[1 - \frac{p+q}{(p-q)^2} + \text{etc.}\right].$$

As for the integral $\int y\,dx$, taken from $x = 0$ to $x = 1$, we have, by §26,

$$\int y\,dx = Y\left[U + \frac{1}{2}\cdot\frac{d^2U^3}{1.2\,dx^2} + \text{etc.}\right]\sqrt{\pi},$$

Y being that which y becomes at its *maximum*, or when we substitute $\frac{p}{p+q}$ for x. ν is here equal to $\frac{x-\frac{p}{p+q}}{\sqrt{\log Y - \log y}}$; and U, $\frac{d^2U^3}{dx^2}$, etc. are that which ν, $\frac{d^2\nu^3}{dx^2}$, etc. become, when we make, [379]
after the differentiations, $x = \frac{p}{p+q}$. We find thus for the integral $\int y\,dx$ taken from x null to $x = 1$,

$$\int y\;dx = \frac{p^{p+\frac{1}{2}}q^{q+\frac{1}{2}}\sqrt{2\pi}}{(p+q)^{p+q+\frac{3}{2}}}\left[1 + \frac{(p+q)^2 - 13pq}{12pq(p+q)} + \text{etc.}\right];$$

the probability that the value of x does not surpass $\frac{1}{2}$, is therefore equal to

$$\frac{(p+q)^{p+q+\frac{3}{2}}}{(p-q)\sqrt{\pi}\,2^{p+q+\frac{3}{2}}p^{p+\frac{1}{2}}q^{q+\frac{1}{2}}} \times\left[1 - \frac{p+q}{(p-q)^2} - \frac{\overline{p+q}^2 - 13pq}{12pq(p+q)} - \text{etc.}\right]. \qquad (o)$$

In order to apply large numbers to this formula, it would be necessary to have the logarithms of p, q and $p-q$, with twelve decimals at least: we are able to supply it in this manner. We have

$$\log\left[\frac{\left(\frac{p+q}{2}\right)^{p+q}}{p^pq^q}\right] = -p\;\log\left(1 + \frac{p-q}{p+q}\right) - q\;\log\left(1 - \frac{p-q}{p+q}\right).$$

When the logarithms are hyperbolic, the second member of this equation, reduced to series, becomes

$$-(p+q)\left[\frac{\left(\frac{p-q}{p+q}\right)^2}{1.2}+\frac{\left(\frac{p-q}{p+q}\right)^4}{3.4}+\frac{\left(\frac{p-q}{p+q}\right)^6}{5.6}+\frac{\left(\frac{p-q}{p+q}\right)^8}{7.8}+\text{etc.}\right];$$

we will have therefore by this very convergent series, the hyperbolic logarithm of $\frac{(p+q)^{p+q}}{2^{p+q}p^p q^q}$. In multiplying it by $0,43429448$, we will convert it into tabular logarithm, and by adding to it the tabular logarithm of $\frac{(p+q)^{\frac{3}{2}}}{2(p-q)\sqrt{2pq\pi}}$, we will have the tabular logarithm of the factor which multiplies series (o). If we name $\frac{1}{\mu}$ this factor, and if we make

$$p = 393386, \quad q = 377555;$$

we find by tabular logarithm [380]

$$\log\mu = 72,2511780,$$

the series (o) becomes

$$\frac{1}{\mu}(1-0,0030761+\text{etc.}).$$

This quantity of an excessive smallness, subtracted from unity, will give the probability that at Paris, the possibility of the births of the boys, surpasses that of girls; whence we see that we must regard this probability as being equal, at least, to that of the most authenticated historical facts.

If we apply formula (o) to the births observed in the principal cities of Europe, we find that the superiority of the births of boys over the births of girls, observed everywhere from Naples to Petersburg, indicates a greater possibility of the births of boys, with a probability extremely near to certitude; this result appears therefore to be a general law, at least in Europe; and if, in some small cities, where we have observed only a not very considerable number of births, nature seems to deviate from it; there is everywhere to believe that this deviation was only apparent, and that at length, the observed births in these cities would offer, in being multiplied, a result similar to the one of the great cities. Many philosophers, deceived by these anomalies, have sought the cause of phenomena which are only the effect of chance; that which proves the necessity to make precede parallel investigations, by that of the probability with which the observations indicate the phenomena of which we wish to determine the cause. I take for example, the small city of Vitteaux, in which, out of 415 births observed during five years, there were born 203 boys and 212 girls. p being here less than q, the natural order appears reversed. Let us see what is according to these observations, the probability that the facilities of the births of boys surpasses in this city, those of the births of girls. This probability is $\frac{\int y\,dx}{\int y\,dx}$, the integral of the numerator being taken from $x=\frac{1}{2}$ to $x=1$, and that of the denominator being taken from $x=0$ to $x=1$. Formula (o) which, subtracted from unity, gives this fraction, becomes here [381] divergent; we will employ then formula (3) from §26, which is reduced very nearly to its first term $\frac{\int dt\, c^{-t^2}}{\sqrt{\pi}}$, the integral being taken from the value of t which corresponds to $x = \frac{1}{2}$ to the value of t which corresponds to $x = 1$. Now we have, by the section cited,

$$t^2 = \log Y - \log y,$$

y being $x^p(1-x)^q$, and Y being the value of y corresponding to the *maximum* of y, which holds when $x = \frac{p}{p+q}$; the value of t^2 which corresponds to $x = \frac{1}{2}$ is $-\log\left[\frac{\left(\frac{p+q}{2}\right)^{p+q}}{p^p q^q}\right]$, this logarithm being hyperbolic, and being given, by that which precedes, by a very convergent series. The value of t^2 which corresponds to $x = 1$, is $t^2 = \infty$; thus we have therefore the two limits of the integral $\int dt\, c^{-t^2}$, an integral which it will be easy to obtain by the formulas which we have given for this object. We find thus the probability that at Vitteaux, the facilities of the births of boys surpasses over those of girls, equal to $0,33$; the superiority of the facility of the births of girls, is therefore indicated by these observations, with a probability equal to $0,67$, a probability much too weak to balance the analogy which carried us to think that at Vitteaux, as in all the cities where we have observed a considerable number of births, the possibility of the births of boys surpasses that of the births of girls.

§29. We have seen at London, the observed ratio of the births of boys to those of girls, is equal to $\frac{19}{18}$, while at Paris, the one of the baptisms of boys to those of girls, is only $\frac{25}{24}$. This seems to indicate a constant cause of this difference. Let us determine the probability of this cause.

Let p and q be the numbers of baptisms of boys and girls, made at Paris in the interval from the beginning of 1745 to the end of 1784; by designating by x, the possibility of the baptism of a boy, and making, as in the preceding section, [382]

$$y = x^p(1-x)^q,$$

the most probable value of x, will be that which renders y a *maximum*; it is therefore $\frac{p}{p+q}$; by supposing next

$$x = \frac{p}{p+q} + \theta;$$

the probability of the value of θ will be, by §26, equal to

$$\frac{d\theta}{\sqrt{\pi}}\sqrt{\frac{(p+q)^3}{2pq}}\, c^{-\frac{(p+q)^3}{2pq}\theta^2}.$$

By designating by p', q' and θ' that which p, q and θ become for London, we will have

$$\frac{d\theta'}{\sqrt{\pi}}\sqrt{\frac{(p'+q')^3}{2p'q'}}\, c^{-\frac{(p'+q')^3}{2p'q'}\theta'^2}$$

for the probability of θ'; the product

$$\frac{d\theta\, d\theta'}{\pi}\sqrt{\frac{(p+q)^3(p'+q')^3}{4pqp'q'}}\, c^{-\frac{(p+q)^3}{2pq}\theta^2 - \frac{(p'+q')^3}{2p'q'}\theta'^2}$$

of these two probabilities, will be therefore the probability of the simultaneous existence of θ and θ'. Let us make

$$\frac{p'}{p'+q'}+\theta'=\frac{p}{p+q}+\theta+t;$$

the preceding differential function becomes

$$\frac{d\theta\ dt}{\pi}\sqrt{\frac{(p+q)^3(p'+q')^3}{4pqp'q'}}c^{-\frac{(p+q)^3}{2pq}\theta^2-\frac{(p'+q')^3}{2p'q'}\left(\theta+t-\frac{p'q-pq'}{(p+q)(p'+q')}\right)^2}.$$

By integrating it for all the possible values of θ, and next for all the positive values of t; we will have the probability that the possibility of the baptisms of boys is greater at London than at Paris. The values of θ are able to be extended from θ equal to $-\frac{p}{p+q}$ to θ equal to [383]
$1-\frac{p}{p+q}$; but when p and q are very great numbers, the factor $c^{-\frac{(p+q)^3}{2pq}\theta^2}$ is so small at these two limits, that we are able to regard it as null; we are able therefore to extend the integral relative to θ, from $\theta=-\infty$ to $\theta=\infty$. We see for the same reason, that the integral relative to t, is able to be extended from $t=0$ to $t=\infty$. By following the process from §27 for these multiple integrations, we will find easily that if we make

$$k^2=\frac{(p+q)^3(p'+q')^3}{2p'q'(p+q)^3+2pq(p'+q')^3},$$
$$h=\frac{p'q-pq'}{(p+q)(p'+q')},$$
$$\theta+\frac{2pqk^2}{(p+q)^3}(t-h)=t',$$

that which gives $d\theta=dt'$; the preceding differential integrated first with respect to t' from $t'=-\infty$ to $t'=\infty$, and next from $t=0$ to t infinity, will give

$$\int\frac{k\ dt}{\sqrt{\pi}}c^{-k^2(t-h)^2}$$

for the probability that at London, the possibility of the baptisms of boys is greater than at Paris. If we make

$$k(t-h)=t'',$$

this integral becomes

$$\int\frac{dt''}{\sqrt{\pi}}c^{-t''^2},$$

the integral being taken from $t''=-kh$ to $t''=\infty$; and it is clear that it is equal to

$$1-\int\frac{dt''}{\sqrt{\pi}}c^{-t''^2},$$

the integral being taken from $t'' = kh$ to $t'' = \infty$. Thence it follows, by §27 of the first Book, that if we suppose

$$i^2 = \frac{p'q'(p+q)^3 + pq(p'+q')^3}{(p+q)(p'+q')(p'q-pq')^2};$$

the probability that the possibility of the baptisms of boys is greater at London than at Paris, [384]
has for expression

$$1 - \frac{ic^{-\frac{1}{2i^2}}}{\sqrt{2\pi}} \cfrac{1}{1+\cfrac{i^2}{1+\cfrac{2i^2}{1+\cfrac{3i^2}{1+\cfrac{4i^2}{1+\text{etc.}}}}}} \qquad (\mu)$$

By making in this formula

$$p = 393386, \qquad q = 377555,$$
$$p' = 737629, \qquad q' = 698958,$$

it becomes

$$1 - \frac{1}{328269}.$$

There is therefore odds of 328268 against one, that at London, the possibility of the baptisms of boys was greater than at Paris. This probability approaches so much to certitude, that there is place to investigate the cause of this superiority.

Among the causes which can produce it, it has appeared to me that the baptisms of the found infants, who are part of the annual list of the baptisms at Paris, must have a sensible influence on the ratio of the baptisms of the boys to those of the girls; and that they should diminish this ratio, if, as it is natural to believe, the parents in the surrounding country, finding advantage to retain near to them the male infants, have sent them to the hospice of the Enfants-trouvés[1] of Paris, in a ratio less than the one of the births of the two sexes. This is that which the summary from the registers of this hospice has made me see with a very great probability. From the commencement of 1745 to the end of 1809, we have baptized 163499 boys and 159405 girls, a number of which the ratio is $\frac{39}{38}$, and differs too much from the ratio $\frac{25}{24}$ of the baptisms of the boys and the girls at Paris, in order to be attributed to simple chance.

§30. Let us determine, according to the preceding principles, the probabilities of the [385]
results founded on the tables of mortality or of assurance, constructed on a great number of observations. Let us suppose first that with respect to a number p of individuals of a given age A, we have observed that there exists yet the number q, at the age $A + a$; we demand

[1] *Translator's note*: Foundling Hospital of Paris.

the probability that out of p' individuals of age A, there will exist $q'+z$ of them at the age $A+a$, the ratio of p' and q' being the same as that of p to q.

Let x be the probability of an individual of age A, to survive to age $A+a$; the probability of the observed event is then the term of the binomial $(x+\overline{1-x})^p$ which has x^q for factor; this probability is therefore

$$\frac{1.2.3\ldots p}{1.2.3\ldots\overline{p-q}1.2.3\ldots q}\cdot x^q(1-x)^{p-q};$$

thus the probability of the value of x, taken from the observed event, is

$$\frac{x^q dx\,(1-x)^{p-q}}{\int x^q dx(1-x)^{p-q}},$$

the integral of the denominator being taken from $x=0$ to $x=1$.

The probability that out of the p' individuals of age A, $q'+z$ will live to age $A+a$, is

$$\frac{1.2.3\ldots p'}{1.2.3\ldots(q'+z)1.2.3\ldots(p'-q'-z)}x^{q'+z}(1-x)^{p'-q'-z}.$$

By multiplying this probability by the preceding probability of the value of x; the product integrated from $x=0$ to $x=1$, will be the probability of the existence of $q'+z$ persons at age $A+a$; by naming therefore P this probability, we will have

$$P=\frac{1.2.3\ldots p'\int x^{q+q'+z}dx(1-x)^{p+p'-q-q'-z}}{1.2.3\ldots(q'+z)1.2.3\ldots(p'-q'-z)\int x^q dx(1-x)^{p-q}},$$

the integrals of the numerator and of the denominator being taken from $x=0$ to $x=1$. We have by §28, very nearly,

$$\begin{aligned}&\int x^{q+q'+z}dx(1-x)^{p+p'-q-q'-z}\\ =&\sqrt{2\pi}\left[(q+q')\left(1+\frac{z}{q+q'}\right)\right]^{q+q'+z+\frac{1}{2}}\\ &\times\frac{\left[(p+p'-q-q')\left(1-\frac{z}{p+p'-q-q'}\right)\right]^{p+p'-q-q'-z+\frac{1}{2}}}{(p+p')^{p+p'+\frac{3}{2}}},\\ &\int x^q dx(1-x)^{p-q}=\sqrt{2\pi}\,\frac{q^{q+\frac{1}{2}}(p-q)^{p-q+\frac{1}{2}}}{p^{p+\frac{3}{2}}}.\end{aligned}$$ [386]

Next, by §33 of the first Book, we have

$$\begin{aligned}&1.2.3\ldots\ldots\ldots p'=p'^{p'+\frac{1}{2}}c^{-p'}\sqrt{2\pi},\\ &1.2.3\ldots\ldots(q'+z)=q'^{q'+z+\frac{1}{2}}\left(1+\frac{z}{q'}\right)^{q'+z+\frac{1}{2}}c^{-q'-z}\sqrt{2\pi},\\ &1.2.3\ldots(p'-q'-z)=(p'-q')^{p'-q'-z+\frac{1}{2}}\left(1-\frac{z}{p'-q'}\right)^{p'-q'-z+\frac{1}{2}}c^{-p'+q'+z}\sqrt{2\pi};\end{aligned}$$

finally we have $q' = \frac{qp'}{p}$. This premised, we find after all the reductions,

$$P = \sqrt{\frac{p^3}{qp'(p-q)(p+p')2\pi}} \times \frac{\left(1+\frac{z}{q+q'}\right)^{q+q'+z+\frac{1}{2}}\left(1-\frac{z}{p+p'-q-q'}\right)^{p+p'-q-q'-z+\frac{1}{2}}}{\left(1+\frac{z}{q'}\right)^{q'+z+\frac{1}{2}}\left(1-\frac{z}{p'-q'}\right)^{p'-q'-z+\frac{1}{2}}}.$$

If we take the hyperbolic logarithm of the second member of this equation, if we reduce this logarithm into series ordered with respect to the powers of z, and if we neglect the powers superior to the square, we will have by passing again from the logarithm to the function,

$$P = \sqrt{\frac{p^3}{qp'(p-q)(p+p')2\pi}}\left[1+\frac{(2q-p)p^2z}{2qp'(p-q)(p+p')}\right]c^{\frac{-p^3z^2}{2qp'(p-q)(p+p')}}.$$

p, q, p' being supposed very great numbers of order $\frac{1}{\alpha}$, the coefficient of z is very small of order α; the one of $-z^2$ is very small and of the same order. But if we suppose $\frac{z}{p}$ of the order $\sqrt{\alpha}$, we will be able to neglect in the preceding expression, the term depending on the first power of z, as very small of order $\sqrt{\alpha}$. Moreover, this term is itself destroyed, when we have regard at the same time to the positive and negative values of z. By neglecting it therefore, we will have [387]

$$2\sqrt{\frac{p^3}{qp'(p-q)(p+p')2\pi}}\int dz\, c^{-\frac{p^3z^2}{2qp'(p-q)(p+p')}}$$

for the expression of the probability that out of p' individuals of age A, the number of those who will arrive to age $A+a$ will be comprehended within the limits $q \pm z$, the integral being taken from z null.

Let us suppose now that we have found by observation, that out of p individuals of age A, q lived yet to age $A+a$, and r to age $A+a+a'$; we demand the probability that out of p' individuals of the same age A, $\frac{qp'}{p}+z$ will live to age $A+a$, and $\frac{rp'}{p}+z'$ will live to age $A+a+a'$.

The probability that out of p' individuals of age A, $\frac{qp'}{p}+z$ will live to age $A+a$ is, by that which precedes,

$$\sqrt{\frac{p^3}{2qp'(p-q)(p+p')\pi}}\,c^{-\frac{p^3z^2}{2qp'(p-q)(p+p')}}.$$

We will have the probability that out of $\frac{qp'}{p}+z$ individuals of age $A+a$, $\left(\frac{qp'}{p}+z\right)\frac{r}{q}+u$ will live to age $A+a+a'$, by changing in the preceding function, p' into $\frac{qp'}{p}+z$, p into q, q into r, and z into u; that which gives, by neglecting z with respect to $\frac{qp'}{p}$,

$$\sqrt{\frac{qp^2}{2rp'(q-r)(p+p')\pi}}\; c^{-\frac{qp^2u^2}{2rp'(p-r)(p+p')}}.$$

The product of these two probabilities, is the probability of the simultaneous existence of z and of u; now we have [388]

$$\left(\frac{qp'}{p}+z\right)\frac{r}{q}+u=\frac{rp'}{p}+z';$$

that which gives

$$u=z'-\frac{rz}{q};$$

by making therefore

$$\beta^2=\frac{p^3}{2qp'(p-q)(p+p')},$$
$$\beta'^2=\frac{qp^2}{2rp'(q-r)(p+p')}.$$

The probability P of the simultaneous existence of the values of z and of z' will be

$$P=\int\frac{\beta\,dz}{\sqrt{\pi}}\cdot\frac{\beta'\,dz'}{\sqrt{\pi}}\,c^{-\beta^2z^2-\beta'^2\left(z'-\frac{rz}{q}\right)^2}.$$

By following this analysis, we find generally that, if we make

$$\beta''^2=\frac{rp^2}{2sp'(r-s)(p+p')},$$
$$\beta'''^2=\frac{sp^2}{2tp'(s-t)(p+p')},$$
etc.;

the probability P that out of p' individuals of age A, the numbers of those who will live to ages $A+a$, $A+a+a'$, $A+a+a'+a''$, etc. will be comprehended within the respective limits

$$\frac{qp'}{p},\frac{qp'}{p}+z;\quad \frac{rp'}{p},\frac{rp'}{p}+z';\quad \frac{sp'}{p},\frac{sp'}{p}+z'';\quad \frac{tp'}{p},\frac{tp'}{p}+z''';\quad \text{etc.}$$

is

$$P=\int\frac{\beta\,dz}{\sqrt{\pi}}\cdot\frac{\beta'\,dz'}{\sqrt{\pi}}\cdot\frac{\beta''\,dz''}{\sqrt{\pi}}.\text{etc.}\,c^{-\beta^2z^2-\beta'^2\left(z'-\frac{rz}{q}\right)^2-\beta'''^2\left(z''-\frac{sz'}{r}\right)^2-\text{etc.}}.$$

We can estimate by this formula, the respective probabilities of the numbers of a table of mortality, constructed on a great number of observations. The manner to form these tables, is very simple. We take out of the registers of births and of deaths, a great number of infants [389] who we follow during the course of their life, by determining how many there remain of them at the end of each year of their age; and we write this number vis-à-vis dying each year. But as in the first two or three years of life, mortality is very rapid; it is necessary, for more exactitude, to indicate in this first age, the number of the surviving at the end of each half-year. If the number p of infants were infinite, we would have thus exact tables which would represent the true law of mortality in the place and at the epoch of their formation. But the number of infants that we choose being finite; however great it be, the numbers of the table are susceptible of errors. Let us represent by p', q', r', s', t', etc., these diverse numbers. The true numbers, for a number p' of births, are $\frac{qp'}{p}$, $\frac{rp'}{p}$, $\frac{sp'}{p}$, $\frac{tp'}{p}$, etc. If we make $q' = \frac{qp'}{p} + z$, z will be the error of q'; similarly, if we suppose $r' = \frac{rp'}{p} + z'$, z' will be the error of r', and so forth. The preceding expression of P is therefore the probability that the errors of q', r', s', etc. are comprehended within the limits zero and z, zero and z', zero and z'', etc. The values of β, β', etc. depend on p, q, r, etc. which are unknowns; but the supposition of p infinite gives

$$\beta^2 = \frac{p^2}{2qp'(p-q)}.$$

We are able to substitute without sensible error, $\frac{q'}{p'}$ instead of $\frac{q}{p}$, that which gives

$$\beta^2 = \frac{p'}{2q'(p'-q')}.$$

We will have in the same manner,

$$\beta'^2 = \frac{q'}{2r'(q'-r')},$$

$$\beta''^2 = \frac{r'}{2s'(r'-s')},$$

etc.

If we wish to consider only the error of one of the numbers of the table, such as s', then we will integrate the expression of P, relatively to z''', z^{iv}, etc., from the infinite negative [390]
values of these variables to their infinite positive values; and then we have

$$P = \int \frac{\beta\, dz}{\sqrt{\pi}} \cdot \frac{\beta'\, dz'}{\sqrt{\pi}} \cdot \frac{\beta''\, dz''}{\sqrt{\pi}}\, c^{-\beta^2 z^2 - \beta'^2\left(z' - \frac{r'z}{q'}\right)^2 - \beta'''^2\left(z'' - \frac{s'z'}{r'}\right)^2}.$$

The integrals relative to z and z' must be taken from their negative infinite values, to their positive infinite values; we will find thus, by the process of which we have often made use for this kind of integration, that if we suppose

$$\gamma^2 = \frac{p'}{2s'(p'-s')},$$

we will have

$$P = \int \frac{\gamma\, dz''}{\sqrt{\pi}}\, c^{-\gamma^2 z''^2}.$$

The probability that the error of any number from the table, will be comprehended within the limits zero and any quantity, is therefore independent, either of the intermediate numbers, or of the subsequent numbers.

If we make $\gamma z'' = t$, we will have

$$\frac{z''}{s'} = t\sqrt{\frac{2(p'-s')}{p's'}},$$

and the probability P that the ratio of the error of the number s' from the table, to this number itself, will be comprehended within the limits $\pm t\sqrt{\frac{2(p'-s')}{p's'}}$ is

$$P = 2\int \frac{dt}{\sqrt{\pi}}\, c^{-t^2},$$

the integral being taken from t null. We see thus that the value of t, and consequently the probability P remaining the same, this ratio increases when s' diminishes; thus the numbers from the table are so much less certain, as they are more extended from the first p'. We see further that this ratio diminishes in measure as p' increases, or in measure as we multiply the observations; in a manner that we are able by this multiplication, to diminish at the same time this ratio and to increase t; this ratio becoming null when p' is infinite, and P [391]
becoming then equal to unity.

§31. Let us apply the preceding analysis to the research on the population of a great empire. One of the simplest and most proper ways to determine this population, is the observation of the annual births of which we are obliged to take account in order to determine the civil state of the infants. But this way supposes that we know very nearly the ratio of the population to the annual births, a ratio that we obtain by making at many points of the empire, the exact denumeration of the inhabitants, and by comparing it to the corresponding births observed during some consecutive years: we conclude from it next, by a simple proportion, the population of all the empire. The government has well wished, at my prayer, to give orders to have with precision, these data. In thirty departments distributed over the area of France, in a manner to outweigh the effects of the variety of climates, we have made a choice of the townships of which the mayors, by their zeal and their intelligence, would be able to furnish the most precise information. The exact denumeration of the inhabitants of these townships, for 22 September 1802, is totaled to 2037615 individuals. The summary of the births, of the marriages and of the deaths, from 22 September 1799 to 22 September 1802, has given for these three years,

Births	*Marriages*	*Deaths*
110312 boys,	46037,	103659 males,
105287 girls,		99443 females.

The ratio of the births of boys to those of girls, that this summary presents, is the one of 22 to 21; and the marriages are to the births, as 3 to 14; the ratio of the population to the annual births is 28,352845. In supposing therefore the number of annual births in France, equal to one million, that which deviates little from the truth; we will have, by multiplying by the preceding ratio, this last number, the population of France equal to 28352845 individuals. Let us see the error that we are able to fear in this evaluation.

For this, let us imagine an urn which contains an infinity of white and black balls in an unknown ratio. Let us suppose next that having drawn at random a great number p of these balls, q have been white, and that in a second drawing, out of an unknown number of extracted balls, there are q' of them white. In order to deduce from it this unknown number, we suppose its ratio to q', the same as the one of p to q; that which gives $\frac{pq'}{q}$ for this number. Let us seek the probability that the number of balls extracted in the second drawing, is comprehended within the limits $\frac{pq'}{q} \pm z$. Let us name x the unknown ratio of the number of white balls, to the total number of balls in the urn. The probability of the observed event in the first drawing, will be expressed by the term which has for factor $x^q(1-x)^{p-q}$ in the development of the binomial $(x+\overline{1-x})^p$, whence it is easy to conclude, as in the preceding section, that the probability of x is [392]

$$\frac{x^q dx\,(1-x)^{p-q}}{\int x^q dx\,(1-x)^{p-q}};$$

the integral of the denominator being taken from $x=0$ to $x=1$. Let us imagine now that in the second drawing, the total number of balls extracted is $\frac{pq'}{q}+z$; the probability of the observed number q' of white balls, will be the term of the binomial $(x+\overline{1-x})^{\frac{pq'}{q}+z}$, which has for factor $x^{q'}(1-x)^{\frac{pq'}{q}+z-q'}$; this probability is therefore

$$\frac{1.2.3\ldots\left(\frac{pq'}{q}+z\right)}{1.2.3\ldots q'.1.2.3\ldots\left(\frac{pq'}{q}+z-q'\right)}x^{q'}(1-x)^{\frac{pq'}{q}+z-q'}.$$

By multiplying it by the preceding probability of x, by integrating the product from $x=0$ to $x=1$, and by dividing it by this same product multiplied by dz, and integrated for all the positive and negative values of z, we will have the probability that the total number of balls extracted, is $\frac{pq'}{q}+z$. We will find thus, by the analysis of the preceding section, this probability equal to [393]

$$\sqrt{\frac{q^3}{2pq'(p-q)(q+q')\pi}}\; c^{-\frac{q^3z^2}{2pq'(p-q)(q+q')}};$$

by naming therefore P the probability that the number of balls extracted in the second drawing, is comprehended within the limits $\frac{pq'}{q} \pm z$, we will have

$$P = 1 - 2\int dz\sqrt{\frac{q^3}{2pq'(p-q)(q+q')\pi}}\; c^{-\frac{q^3z^2}{2pq'(p-q)(q+q')}},$$

the integral being taken from $z=z$ to z infinity.

Now, the number p of balls extracted in the first drawing, can represent a denumeration; and the number q of white balls which are comprehended, can express the number of women who, in this denumeration, must become mothers in the year, or the number of annual births, corresponding to the denumeration. Then q' expresses the number of annual births observed in all the empire, and whence we conclude the population $\frac{pq'}{q}$. In this case, the preceding value of P expresses the probability that this population is comprehended within the limits $\frac{pq'}{q} \pm z$.

We will suppose, conformably to the preceding data,

$$p = 2037615, \qquad q = \frac{110313 + 105287}{3};$$

we will suppose next

$$q' = 1500000, \qquad z = 500000;$$

the preceding formula gives then

$$P = 1 - \frac{1}{1162}.$$

There is odds therefore around 1161 against one, that in fixing at 42529267, the population [394]
corresponding to fifteen hundred thousand births, we will not be deceived by a half-million.

The difference between certitude and the probability P diminishes with a very great rapidity, when z increases: it would be insensible, if we suppose $z = 700000$.

§32. Let us consider now the probability of future events, deduced from observed events; and let us suppose that having observed an event composed of any number of simple events, we seek the probability of a future result, composed of similar events.

Let us name x the probability of each simple event, y the corresponding probability of the observed result, and z the one of the future result; the probability of x will be, as we have seen,

$$\frac{y\,dx}{\int y\,dx},$$

the integral being taken from $x = 0$ to $x = 1$; $\frac{yz\,dx}{\int y\,dx}$ is therefore the probability of the future result, taken from the value of x, considered as cause of the simple event; thus, by naming P the entire probability of the future event, we will have

$$P = \frac{\int yz\,dx}{\int y\,dx},$$

the integrals of the numerator and of the denominator being taken from $x = 0$ to $x = 1$.

Let us suppose, for example, that an event having arrived m times consecutively, we demand the probability that it will arrive the following n times. In this case, x being supposed to represent the possibility of the simple event, x^m will be that of the observed event, and x^n that of the future event; that which gives

$$y = x^m, \quad z = x^n;$$

whence we deduce

$$P = \frac{m+1}{m+n+1}.$$

Let us suppose the observed event, composed of a very great number of simple events; let a be the value of x which renders y a *maximum*, and Y that *maximum*; let a' be the value of x which renders yz a *maximum*, and Y' and Z' that which y and z become at this *maximum*. We will have by §27 of the first Book, very nearly [395]

$$\int y\,dx = \frac{Y^{\frac{3}{2}}\sqrt{2\pi}}{\sqrt{-\frac{ddY}{dx^2}}},$$

$$\int yz\,dx = \frac{(Y'Z')^{\frac{3}{2}}\sqrt{2\pi}}{\sqrt{-\frac{dd(Y'Z')}{dx^2}}}.$$

The observed result being composed of a very great number of simple events, let us suppose that the future event is much less composite. The equation which gives the value of a' of x, corresponding to the *maximum* of yz, is

$$0 = \frac{dy}{y\,dx} + \frac{dz}{z\,dx}.$$

$\frac{dy}{y\,dx}$ is a very great quantity, of order $\frac{1}{\alpha}$; and since the future result is much less composite with respect to the observed result, $\frac{dz}{z\,dx}$ will be of a lesser order, which we will designate by $\frac{1}{\alpha^{1-\lambda}}$; thus a being the value of x which satisfies the equation $0 = \frac{dy}{y\,dx}$; the difference between a and a' will be very small of order α^λ, and we will be able to suppose

$$a' = a + \alpha^\lambda \mu.$$

This supposition gives

$$Y' = Y + \alpha^\lambda \mu \cdot \frac{dY}{dx} + \frac{\alpha^{2\lambda}\mu^2}{1.2} \cdot \frac{d^2Y}{dx^2} + \text{etc.}$$

But we have $\frac{dY}{dx} = 0$, and it is easy to conclude from it that $\frac{d^nY}{Y\,dx^n}$ is of an order equal or less than $\frac{1}{\alpha^{\frac{n}{2}}}$; the term $\frac{\alpha^{\lambda n}\mu^n}{1.2.3...n} \cdot \frac{d^nY}{Y\,dx^n}$ will be consequently more than order $\alpha^{n\left(\lambda - \frac{1}{2}\right)}$. Thus the [396] convergence of the expression of Y' in series, requires that λ surpass $\frac{1}{2}$; and in this case, Y' differs from Y, only by quantities of order $\alpha^{2\lambda-1}$.

If we name Z that which z becomes when we make $x = a$; we will be assured in the same manner that Z' can be reduced to Z. Finally, we will prove by a similar reasoning, that $\frac{d^2(Y'Z')}{dx^2}$ is reduced to very nearly $Z\frac{d^2Y}{dx^2}$. By substituting these values into the expression of P, we will have

$$P = Z;$$

that is that we can then determine the probability of the future result, by supposing x equal to the value which renders the observed result most probable. But if it is necessary for that that the future result rather be not very composite, so that the exponents of the factors of z are of an order of magnitude smaller than the square root of the factors of y; otherwise, the preceding supposition would expose some sensible errors.

If the future result is a function of the observed result, z will be a function of y, which we will represent by $\phi(y)$. The value of x, which renders zy a *maximum* is, in this case, the same which renders y a *maximum*; thus we have $a' = a$; and if we designate $\frac{d\phi(y)}{dy}$ by $\phi'(y)$, the expression of P will become, by observing that $\frac{dY}{dx} = 0$,

$$P = \frac{\phi(Y)}{\sqrt{1 + \frac{Y\phi'(y)}{\phi(Y)}}}.$$

If $\phi(Y) = y^n$, so that the future event is n times the repetition of the observed event, we will have

$$P = \frac{Y^n}{\sqrt{n+1}}.$$

The probability P calculated under the supposition that the possibility of the simple events is equal to that which renders the observed result most probable, is Y^n: we see thus that the small errors which result from this supposition, are accumulated at the rate of the simple [397] events which enter into the future result, and become very sensible when these events are in great number.

§33. Since 1745, an epoch where we have commenced to distinguish at Paris upon the registers, the baptisms of boys from those of girls, we have constantly observed that the number of the first has been superior to the one of the second. Let us determine the probability that this superiority will be maintained during a given time, for example, in the space of a century.

Let p be the observed number of baptisms of boys; q the one of girls; $2n$ the number of annual baptisms; x the probability that the infant who will be born and be baptized will be a boy. By raising $x + (1 - x)$ to the power $2n$, and developing this power, we will have

$$x^{2n} + 2nx^{2n-1}(1-x) + \frac{2n(n-1)}{1.2}x^{2n-2}(1-x)^2 + \text{etc.}$$

The sum of the n first terms of this development, will be the probability that each year, the number of baptisms of boys will surpass the one of the baptisms of girls. Let us name z this sum; z^i will be the probability that this superiority will be maintained during the number i of consecutive years; therefore, if we designate by P the entire probability that this will take place; we will have, by the preceding section,

$$P = \frac{\int x^p dx\, z^i (1-x)^q}{\int x^p dx\, (1-x)^q},$$

the integrals of the numerator and of the denominator being taken from $x = 0$ to $x = 1$.

If we name a the value of x which renders $x^p z^i (1-x)^q$ a *maximum*, and if we designate by Z, $\frac{dZ}{dx}$, $\frac{d^2Z}{dx^2}$ that which z, $\frac{dz}{dx}$, $\frac{d^2z}{dx^2}$ become, when we change x into a; we will have, by §26,

$$\int x^p dx\, z^i (1-x)^q = \frac{a^{p+1}(1-a)^{q+1}Z^i\sqrt{2\pi}}{\sqrt{p(1-a)^2 + qa^2 + ia^2(1-a)^2\left(\frac{dZ^2 - ZddZ}{Z^2dx^2}\right)}}.$$

z being the sum of the first n terms of the function [398]

$$x^{2n}\left[1 + 2n\frac{(1-x)}{x} + \frac{2n(2n-1)}{1.2}\cdot\frac{(1-x)^2}{x^2} + \text{etc.}\right],$$

we have by §37 of the first Book,

$$z = \frac{\int \frac{u^{n-1}du}{(1+u)^{2n+1}}}{\int \frac{u^{n-1}du}{(1+u)^{2n+1}}},$$

the integral of the numerator being taken from $u = \frac{1-x}{x}$ to $u = \infty$, and that of the denominator being taken from $u = 0$ to $u = \infty$. Let there be $u = \frac{1-s}{s}$; this value of z will become

$$z = \frac{\int s^n ds\,(1-s)^{n-1}}{\int s^n ds\,(1-s)^{n-1}},$$

the integral of the numerator being taken from $s = 0$ to $s = x$, and that of the denominator taken from $s = 0$ to $s = 1$. Thence we deduce

$$\frac{dz}{z\,dx} = \frac{x^n(1-x)^{n-1}}{\int s^n ds\,(1-s)^{n-1}},$$

the integral of the denominator being taken from $x = 0$ to $s = x$. We will have next

$$\frac{ddz}{z\,dx^2} = \frac{dz}{z\,dx} = \frac{n-(2n-1)x}{x(1-x)}.$$

By changing x into a in these expressions, we will have those of Z, $\frac{dZ}{Z\,dx}$, $\frac{ddZ}{Z\,dx^2}$.

In order to determine a, we will observe that the condition of the *maximum* of $x^p z^i (1-x)^q$ gives

$$0 = \frac{p}{a} - \frac{q}{1-a} + i\frac{dZ}{Z\,dx};$$

whence we deduce, by substituting for $\frac{dZ}{Z\,dx}$, its preceding value,

$$a = \frac{p}{p+q} + \frac{ia^{n+1}(1-a)^n}{(p+q)\int s^n ds\,(1-s)^{n-1}},$$

the integral of the denominator being taken from $s = 0$ to $s = a$. In order to conclude a from this equation, we will observe that the value of s which renders $s^n(1-s)^{n-1}$ a *maximum*, is very nearly $\frac{1}{2}$, and consequently, less than $\frac{p}{p+q}$ which itself is smaller than a. Thus n being supposed a large number, we can, without sensible error, extend the integral of this expression of a, from $s = 0$ to $s = 1$, the term which depends on it being very small. This gives, by §28, [399]

$$\int s^n ds\,(1-s)^{n-1} = \frac{n^{n+\frac{1}{2}}(n-1)^{n-\frac{1}{2}}\sqrt{2\pi}}{(2n-1)^{2n+\frac{1}{2}}} = \frac{\sqrt{\pi}}{2^{2n}\sqrt{n}};$$

the equation which determines a becomes thus quite nearly,

$$a = \frac{p}{p+q} + \frac{ia^{n+1}(1-a)^n 2^{2n}\sqrt{n}}{(p+q)\sqrt{\pi}}.$$

In order to resolve it, we will observe that a differs very little from $\frac{p}{p+q}$: so that if we make

$$a = \frac{p}{p+q} + \mu,$$

μ will be quite small, and we will have in a very close manner,

$$\mu = i\sqrt{n}\frac{p\left[1-\left(\frac{p-q}{p+q}\right)^2\right]^n}{(p+q)^2\sqrt{\pi}}c^{-\frac{n\mu(p+q)(p-q)}{pq}-\frac{(p+q)^2 n\mu^2}{pq}}; \qquad (1)$$

we will have next very nearly,

$$a^p(1-a)^q = \left(\frac{p}{p+q}\right)^p\left(\frac{q}{p+q}\right)^q c^{-\frac{(p+q)^3}{2pq}\mu^2}.$$

By substituting into the radical

$$\sqrt{p(1-a)^2 + qa^2 + ia^2(1-a)^2\left(\frac{dZ^2 - ZddZ}{Z^2dx^2}\right)},$$

for a, its value $\frac{p}{p+q} + \mu$; for $\frac{dZ}{Z\,dx}$, its value $\frac{(p+q)a-p}{ia(1-a)}$ or $\frac{(p+q)\mu}{ia(1-a)}$; and for $\frac{ddZ}{Z\,dx^2}$, its value $\frac{dZ}{Z\,dx}\,\frac{n-(2n-1)a}{a(1-a)}$; this radical becomes very nearly [400]

$$\sqrt{\frac{pq}{p+q}}\sqrt{1+\frac{(p+q)\mu}{pq}[n(p-q)-p]+\frac{(p+q)^2}{pq}\mu^2\left(2n+\frac{p+q}{i}\right)}.$$

Finally, we have by §28,

$$\int x^p dx\,(1-x)^q = \left(\frac{p}{p+q}\right)^p\left(\frac{q}{p+q}\right)^q\sqrt{\frac{pq}{p+q}}\cdot\frac{\sqrt{2\pi}}{p+q}.$$

This premised, the expression of P will become very nearly,

$$P = \frac{Z^i c^{-\frac{(p+q)^3}{2pq}\mu^2}}{\sqrt{1 + \frac{(p+q)\mu}{pq}[n(p-q)-p] + \frac{(p+q)^2\mu^2}{pq}\left(2n + \frac{(p+q)}{i}\right)}}. \tag{2}$$

The concern is therefore no longer but to determine Z. We have

$$Z = \frac{\int s^n ds\,(1-s)^{n-1}}{\int s^n ds\,(1-s)^{n-1}},$$

the integral of the numerator being taken from $s = 0$ to $s = a$, and that of the denominator being taken from $s = 0$ to $s = 1$. It is easy to conclude from it that we have

$$Z = 1 - \frac{\int s^n ds\,(1-s)^{n-1}}{\int s^n ds\,(1-s)^{n-1}},$$

the integral of the numerator being taken from $s = a$ to $s = 1$ and that of the denominator being taken from $s = 0$ to $s = 1$; we will have thus quite nearly, by §29,

$$Z = 1 - \frac{\int dt\, c^{-t^2}}{\sqrt{\pi}}, \tag{3}$$

the integral relative to t being taken from

$$t^2 = \frac{2n-1}{2n(n-1)}\left(\frac{n(p-q)}{p+q} - \frac{p}{p+q} + (2n-1)\mu\right)^2,$$

to $t^2 = \infty$.

In order to apply numbers to these formulas, we will observe that, by that which precedes, in the interval from the commencement of 1745 to the end of 1784, we have by §28, relatively to Paris, [401]

$$p = 393386, \quad q = 377555.$$

By dividing by 40 the sum of these two numbers, we will have 19273,5 for the mean number of annual baptisms; that which gives $n = 9636,75$; we will suppose moreover $i = 100$. By means of these values, we will determine that of μ, by equation (1); we will determine next the value of Z by equation (3); finally equation (2) will give the value of P. We will find thus

$$P = 0,782.$$

There was therefore at the end of 1784, according to these data, odds nearly four against one that in the space of a century, the baptisms of boys at Paris, will surpass, each year, over those of the girls.