Analytic Theory of Probabilities

P.S. Laplace

Book II Chapter II, pp. 194–203

4. A lottery being composed of n numbered tickets of which r come forth at each drawing, one requires the probability that after i drawings all the tickets will come forth.

We name $z_{n,q}$ the number of cases in which, after i drawings, the totality of the tickets $1,\ 2,\ 3,\ \ldots q$ will come forth. It is clear that this number is equal to the number $z_{n,q-1}$ of cases in which the tickets $1,\ 2,\ 3,\ \ldots, q-1$ come forth, less the number of cases in which, these tickets being brought out, the ticket q is not drawn; now this last number is evidently the same as the one of the cases in which the tickets $1,\ 2,\ 3,\ \ldots, q-1$ would be extracted, if one removed the ticket q from the n tickets of the lottery, and this number is $z_{n-1,q-1}$; one has therefore

$$z_{n,q} = z_{n,q-1} - z_{n-1,q-1}. \tag{i}$$

Now the number of all possible cases in a single drawing being $\frac{n(n-1)(n-2)\cdots(n-r+1)}{1.2.3\ldots r}$, the one of all possible cases in i drawings is

$$\left[\frac{n(n-1)(n-2)\cdots(n-r+1)}{1.2.3\ldots r}\right]^i.$$

The number of all the cases in which the ticket 1 will not come forth in these i drawings is the number of all possible cases, when one subtracts this ticket from the n tickets in the lottery, and this number is

$$\left[\frac{(n-1)(n-2)\cdots(n-r)}{1.2.3\ldots r}\right]^i;$$

the number of cases in which the ticket 1 will come forth in i drawings is therefore

$$\left[\frac{n(n-1)(n-2)\cdots(n-r+1)}{1.2.3\ldots r}\right]^i - \left[\frac{(n-1)(n-2)\cdots(n-r)}{1.2.3\ldots r}\right]^i,$$

or

$$\Delta\left[\frac{(n-1)(n-2)\cdots(n-r)}{1.2.3\ldots r}\right]^i;$$

this is the value of $z_{n,1}$. This put, equation (i) will give, by making successively $q=2$, $q=3,\ \ldots,$

$$z_{n,2} = \Delta^2 \left[\frac{(n-2)(n-3)\cdots(n-r-1)}{1.2.3\ldots r} \right]^i,$$
$$z_{n,3} = \Delta^3 \left[\frac{(n-3)(n-4)\cdots(n-r-2)}{1.2.3\ldots r} \right]^i,$$
$$\cdots$$

and generally

$$z_{n,q} = \Delta^q \left[\frac{(n-q)(n-q-1)\cdots(n-r-q+1)}{1.2.3\ldots r} \right]^i.$$

Thus the probability that the tickets 1, 2, 3, $\ldots q$ will come forth in i drawings being equal to $z_{n,q}$ divided by the number of all possible cases, it will be

$$\frac{\Delta^q[(n-q)(n-q-1)\cdots(n-r-q+1)]^i}{[n(n-1)(n-2)\cdots(n-r+1)]^i}$$

If one makes in this expression $q = n$, one will have, s being here the variable which must be supposed null in the result,

$$\frac{\Delta^n[s(s-1)\cdots(s-r+1)]^i}{[n(n-1)\cdots(n-r+1)]^i}$$

for the expression of the probability that all the tickets of the lottery will come forth in i drawings.

If n and i are very great numbers, one will have, by the formulas of No. 40[1] of Book I, the value of this probability by means of a highly convergent series. We suppose, for example, that he brings forth only one ticket at each drawing; the preceding probability becomes

$$\frac{\Delta^n s^i}{n^i}.$$

Let us propose to determine the number i of drawings in which this probability is $\frac{1}{k}$, n and i being very great numbers. By following the analysis of the section cited, one will determine first a by the equation

$$0 = \frac{i+1}{a} - s - \frac{nc^a}{c^a - 1},$$

this which gives

$$a = \frac{i+1}{n+s} \left\{ \frac{1-c^{-a}}{1 - \frac{sc^{-a}}{n+s}} \right\}.$$

[1] Pages 160-165.

One has next, by No. 40 of Book I, when c^{-a} is a very small quantity of the order $\frac{1}{i}$, as that takes place in the present question, one has, I say, to the quantities nearly of order $\frac{1}{i^2}$, s being supposed null in the result of the calculation,

$$\frac{\Delta^n s^i}{n^i} = \frac{\left(\frac{i}{i+1}\right)^{i+\frac{1}{2}} c^{na-i}(1-c^{-a})^{n-i}}{\sqrt{1-\frac{i+1}{n}c^{-a}}}.$$

Now one has, to the quantities nearly of the order $\frac{1}{i^2}$,

$$\left(\frac{i}{i+1}\right)^{i+\frac{1}{2}} = c^{-1};$$

by supposing next $c^{-a} = z$, one has

$$(1-c^{-a})^{n-1} = c^{(i-n)z}\left(1+\frac{i-n}{2}z^2\right);$$

moreover, the equation which determines a gives

$$i+1-na = (i+1)z,$$

whence one deduces

$$c^{na-i-1} = c^{-iz}(1-z);$$

one will have therefore, to the quantities nearly of order $\frac{1}{i^2}$,

$$\frac{\Delta^n s^i}{n^i} = c^{-nz}\left(1+\frac{i-2n+1}{2n}z+\frac{i-n}{2}z^2\right).$$

In order to determine z, we take up again the equation

$$a = \frac{i+1}{n} - \frac{i+1}{n}c^{-a};$$

one will have, by formula (p) of No. 21 of Book II of the *Mécanique céleste*,

$$z = c^{-a} = q + \frac{i+1}{n}q^2 + \frac{3\left(\frac{i+1}{n}\right)^2}{1.2}q^3 + \frac{4^2\left(\frac{i+1}{n}\right)^3}{1.2.3}q^4 + \cdots,$$

q being supposed equal to $c^{-\frac{i+1}{n}}$. This value of z gives

$$c^{-nz} = c^{-nq}[1-(i+1)q^2];$$

consequently,

$$\frac{\Delta^n s^i}{n^i} = c^{-nq}\left(1+\frac{i+1-2n}{2n}q-\frac{n+i+2}{2}q^2\right).$$

By equating this quantity to the fraction $\frac{1}{k}$, one will have

$$q = \frac{\log k}{n}\left(1 + \frac{i+1-2n}{2n^2} - \frac{n+i+2}{2n^2}\log k\right);$$

now one has

$$i + 1 = -n\log q;$$

one will have therefore very nearly, for the expression of the number i of drawings, according to which the probability that all the tickets will come forth is $\frac{1}{k}$,

$$i = (\log n - \log\log k)(n - \frac{1}{2} + \frac{1}{2}\log k) + \frac{1}{2}\log k;$$

one must observe that all these logarithms are hyperbolic.

We suppose the lottery composed of 10000 tickets, or $n = 10000$, and $k = 2$, this formula gives

$$i = 95767,4$$

for the expression of the number of drawings, in which one can wager one against one, that the ten thousand tickets of the lottery will come forth; it is therefore a little less than one against one to wager that they will come forth in 95767 drawings, and a little more than one against one to wager that they will come forth in 95768 drawings.

One will determine by a similar analysis the number of drawings in which one is able to wager one against one that all the tickets of the lottery of France will come forth. This lottery is, as one knows, composed of 90 tickets of which five come forth at each drawing. The probability that all the tickets will come forth in i drawings is then, by that which precedes,

$$\frac{\Delta^n[s'(s'-1)(s'-2)(s'-3)(s'-4)]^i}{[n(n-1)(n-2)(n-3)(n-4)]^i},$$

n being here equal to 90, and s' previously being supposed null in the result of the calculation. If one makes $s = s' - 2$, this function becomes

$$\frac{\Delta^n[s(s^2-1)(s^2-4)]^i}{\{(n-2)[(n-2)^2-1][(n-2)^2-4]\}^i},$$

where, by developing in series,

$$\frac{(\Delta^n s^{5i} - 5i\Delta^n s^{5i-2} + \cdots)}{(n-2)^{5i}}\left[1 + \frac{5i}{(n-2)^2} + \cdots\right],$$

s previously being supposed equal to -2 in the result of the calculation.

One has, by No. 40 of Book I, by neglecting the terms of order $\frac{1}{i^2}$ and supposing c^{-a} very small of order $\frac{1}{i}$,

$$\frac{\Delta^n s^{5i}}{(n-2)^{5i}} = \frac{\left(\frac{5i+1}{a}\right)^{5i}\left(\frac{5i}{5i+1}\right)^{5i} c^{(n-2)a-5i}(1-c^{-a})^n}{(n-2)^{5i}\sqrt{1+\frac{1}{5i}-\frac{na^2c^{-a}}{5i(1-c^{-a})^2}}},$$

a being given by the equation

$$a = \frac{(5i+1)(1-c^{-a})}{(n-2)\left(1+\frac{2c^{-a}}{n-2}\right)}.$$

One has thus, by neglecting the terms of order $\frac{1}{i^2}$,

$$\frac{\Delta^n s^{5i}}{(n-2)^{5i}} = \frac{\left(1+\frac{2c^{-a}}{n-2}\right)^{5i}}{(1-c^{-a})^{5i}}(1-c^{-a})^n c^{1-(5i+1)c^{-a}-\frac{10ic^{-a}}{n-2}}\left(\frac{5i}{5i+1}\right)^{5i}\left(1-\frac{1}{10i}+\frac{na^2c^{-a}}{10i}\right);$$

now one has

$$\left(1+\frac{2c^{-a}}{n-2}\right)^{5i} = c^{\frac{10ic^{-a}}{n-2}},$$

$$(1-c^{-a})^{-5i} = c^{5ic^{-a}}\left(1+\frac{5i}{2}c^{-2a}\right),$$

$$\left(\frac{5i}{5i+1}\right)^{5i} = c^{-1}\left(1+\frac{1}{10i}\right);$$

one will have therefore, to the quantities nearly of order $\frac{1}{i^2}$,

$$\frac{\Delta^n s^{5i}}{(n-2)^{5i}} = (1-c^{-a})^n\left(1-c^{-a}+\frac{5i}{2}c^{-2a}+\frac{na^2c^{-a}}{10i}\right).$$

By substituting for a its value and observing that i is very little different from $n-2$ in the present case, as one will see hereafter, one has, very nearly,

$$\frac{na^2c^{-a}}{10i} = \frac{5i+12}{2(n-2)}c^{-a}.$$

I keep, for greater exactitude, the term $\frac{12c^{-a}}{2(n-2)}$, although of order $\frac{1}{i^2}$, because of the size of its factor 12; one will have therefore

$$\frac{\Delta^n s^{5i}}{(n-2)^{5i}} = (1-c^{-a})^n\left[1+\frac{5i-2n+16}{2(n-2)}c^{-a}+\frac{5i}{2}c^{-2a}\right].$$

If one changes in this equation $5i$ into $5i-2$, one will have that of $\frac{\Delta^n s^{5i-2}}{(n-2)^{5i-2}}$; but the value of a will no longer be the same. Let a' be this new value, one will have

$$a' = \frac{(5i-1)(1-c^{-a'})}{(n-2)\left(1+\frac{2c^{-a'}}{n-2}\right)},$$

this which gives, very nearly,

$$a' = a - \frac{2}{n-2}.$$

Now one has

$$1 - c^{-a'} = 1 - c^{-a} - \frac{2c^{-a}}{n-2},$$

whence one deduces, by neglecting the quantities of order $\frac{1}{i}$,

$$(1 - c^{-a'})^n = (1 - c^{-a})^n;$$

consequently one has, by neglecting the quantities of order $\frac{1}{i}$,

$$\frac{\Delta^n s^{5i-2}}{(n-2)^{5i-2}} = (1 - c^{-a})^n.$$

One will have therefore, to the quantities nearly of order $\frac{1}{i^2}$,

$$\frac{\Delta^n [s(s^2-1)(s^2-4)]^i}{[n(n-1)(n-2)(n-3)(n-4)]^i}$$
$$= (1 - c^{-a})^n \left[1 + \frac{5i - 2n + 16}{2(n-2)} c^{-a} + \frac{5i}{2} c^{-2a}\right].$$

This quantity must, by the condition of the problem, be equal to $\frac{1}{2}$, this which gives

$$1 - c^{-a} = \sqrt[n]{\frac{1}{2}} \left[1 - \frac{5i - 2n + 16}{2n(n-2)} c^{-a} - \frac{5i}{2n} c^{-2a}\right],$$

whence one deduces

$$c^{-a} = \left(1 - \sqrt[n]{\frac{1}{2}}\right) \left[1 + \frac{5i - 2n + 16}{2n(n-2)} + \frac{5i}{2n} c^{-a}\right];$$

consequently one has, by hyperbolic logarithms,

$$a = \log\left(\frac{\sqrt[n]{2}}{\sqrt[n]{2} - 1}\right) - \frac{5i - 2n + 16}{2n(n-2)} - \frac{5i}{2n} c^{-a};$$

now one has, to the quantities nearly of order $\frac{1}{i^2}$,

$$a = \frac{5i + 1}{(n-2)\sqrt[n]{2}};$$

one will have therefore

$$i = \frac{n-2}{5}\sqrt[n]{2}\left[1 - \frac{1}{2n} - \frac{16}{10in} - \frac{1}{2}(\sqrt[n]{2}-1)\right]\log\left(\frac{\sqrt[n]{2}}{\sqrt[n]{2}-1}\right).$$

By substituting for n its value 90, one finds

$$i = 85,53,$$

so that it is a little less than one to one to wager that all the tickets will come forth in 85 drawings, and a little more than one to one to wager that they will come forth in 86 drawings.

A quite simple and very approximate way to obtain the value of i is to suppose $\frac{\Delta^n s^i}{n^i}$, or the series

$$1 - n\left(\frac{n-1}{n}\right)^i + \frac{n(n-1)}{2}\left(\frac{n-2}{n}\right)^i - \cdots,$$

equal to the development

$$1 - n\left(\frac{n-1}{n}\right)^i + \frac{n(n-1)}{1.2}\left(\frac{n-1}{n}\right)^{2i} - \cdots$$

of the binomial $\left[1 - \left(\frac{n-1}{n}\right)^i\right]^n$. In reality the two series have the first two terms equal respectively. Their third terms are also, very nearly, equal between themselves; for one has quite nearly $\left(\frac{n-2}{n}\right)^i$ equal to $\left(\frac{n-1}{n}\right)^{2i}$. In reality, their hyperbolic logarithms are, by neglecting the terms of order $\frac{i}{n^2}$, both equal to $-\frac{i}{n}$. One will see in the same way that the fourth terms, the fifth, $\ldots$ are very little different, when n and i are very great numbers; but the difference increases without ceasing in measure as the terms move away from the first, which must in the end produce in them an evident <difference> between the series themselves. In order to estimate it, we determine the value of i concluded from the equality of the two series. By equating to $\frac{1}{k}$ the binomial $\left[1 - \left(\frac{n-1}{n}\right)^i\right]^n$, one will have

$$i = \frac{\log\left(1 - \sqrt[n]{\frac{1}{k}}\right)}{\log\left(\frac{n-1}{n}\right)},$$

these logarithms may be, at will, hyperbolic or tabulated. Let $\sqrt[n]{\frac{1}{k}} = 1 - z$. We will have, by taking the hyperbolic logarithms of each member of this equation

$$\frac{1}{n}\log k = -\log(1-z) = z + \frac{z^2}{2} + \cdots,$$

this which gives, very nearly,

$$z = \frac{\log k}{n}\left(1 - \frac{\log k}{2n}\right);$$

one will have therefore, by hyperbolic logarithms,

$$\log\left(1-\sqrt[n]{\frac{1}{k}}\right)=\log z=\log\,\log\,k-\log\,n-\frac{\log\,k}{2n},$$

One has next

$$\log\frac{n-1}{n}=-\frac{1}{n}-\frac{1}{2n^2}-\cdots$$

The preceding expression for i becomes in this way, very nearly,

$$i=n(\log n-\log\log k)\left(1-\frac{1}{2n}\right)+\frac{1}{2}\log k;$$

the excess of the value found previously for i over this one is

$$\frac{\log k}{2}(\log n-\log\log k);$$

this excess becomes infinite, when n is infinite; but a very great number is necessary in order to render it very evident; and in the case of $n=10000$ and of $k=2$, it is again only of three units.

If one considers likewise the development

$$1-n\left(\frac{n-5}{n}\right)^i+\cdots$$

of the expression $\frac{\Delta^n[s'(s'-1)(s'-2)(s'-3)(s'-4)]^i}{[n(n-1)(n-2)(n-3)(n-4)]^i}$, as the one of the binomial $\left[1-n\left(\frac{n-5}{n}\right)^i\right]^n$, one will have, in order to determine the number i of trials in which one can wager one against one that all the tickets will come forth, the equation

$$\left[1-\left(\frac{n-5}{n}\right)^i\right]^n=\frac{1}{2};$$

this which gives

$$i=\frac{\log\left(\frac{\sqrt[n]{2}}{\sqrt[n]{2}-1}\right)}{\log\left(\frac{n}{n-5}\right)}.$$

These logarithms can be tabulated. By making $n=90$, one finds

$$i=85,204,$$

this which differs very little from the value $i=85,53$ that we have found above.