STATISTICA, anno LXXV, n. 2, 2015

CONSIDERATIONS ON A POSTERIORI PROBABILITY

Corrado Gini (1911)

This paper is the "First part" of an assay published in the "Studi Economico-Giuridici della Universit di Cagliari", year III, 1911. It was reprinted in "Metron" vol. XV, 1-4,1949.

1. To determine the probability that a phenomenon A takes place x times in s future observations, is a problem that frequently occurs and the importance of which is such that it would be unnecessary to emphasize it.

Most of the time we are not in a position to solve it on the basis of *a priori* notions of the probability of A taking place. Notions of this type are found in the games that human beings have conventionally set up, provided that the instruments used do not show detectable irregularity in their makeup. But usually, to assess the probability of phenomenon A taking place in s future observations, we must base our reasoning on its frequency in the previous n observations, by resorting to the hypothesis that the probability of its taking place is constant during the $n + s$ observations. In such a case, we proceed to the *a posteriori* determination of the probability.

The problem of *a posteriori* determining the probability of future events, in its more general form, can be formulated as follows: *Determine the probability that the phenomenon A occurs x times in s future observations, knowing that in A previous observations it did occur in times and supposing that, during the $n+s$ observations, its probability to occur remains unchanged.*

2. In this study, we shall generally define $P_{b,c}$, the probability[1] that the phenomenon A takes place b times in c observations; specifically $P_{1,1}$, more briefly P, will denote the probability of the occurrence of A, with ${}_{d,e}P_{b,c}$ let us denote the probability that A takes place b times in c observations, when it has taken place d times in e observations. Hence, the probability that A occurs x times in s future observations, when it has occurred m times in A past observations, will be ${}_{m,n}P_{x,s}$; in particular ${}_{m,n}P_{1,1}$, that is more briefly ${}_{m.n}P$, will denote the probability of the occurrence of A when it has occurred m times in n previous observations. ${}_{m,n}P$

[1] During the course of this work, when we speak of "probability" of a phenomenon, we mean the "mathematical probability", given by the ratio of the number of favourable cases to the number of equally possible cases of the phenomenon; to define its "empirical probability" or "statistical probability", the word "frequency" has been used.

is usually called the *a posteriori* probability of A.

3. To say that the probability of occurrence of A remains unchanged during $n+s$ observations, is equivalent to saying that what remains unchanged is the ratio of favorable events of A occurring to the unfavorable events. We will than say that phenomenon A is dependent on A same system of causes, during the course of $n+s$ observations[2].

Now we can present two situations: either phenomenon A allows only one system of causes; or it allows several. For example, sunrise, from the beginning till the end of the world, is the consequence of A unique system of causes; the proportion of wins of A competitor (cyclist, racing driver, fencer or target shooter) over another, in one-day competitive games, permits instead as many systems of causes, as are the possible combinations of strength, with which, in one day, the two competitors may find themselves faced with. But, the first can be looked at as the limiting case of the second, which occurs when, among the various systems of causes on which A may depend, one has a probability equal to 1 and the others a probability equal to 0. One can therefore consider the second case only as the general one.

4. Let ν be the systems of causes compatible with phenomenon A. Let us denote with ${}_{m,n}r_y$ the probability that, when A has occurred m times in n observations, the system of causes y^{th} is in action and, with ${}^{y}P_{x,s}$ the probability that, if it is acting, phenomenon A takes place x times in s observations.

It will be

$$ {}_{m,n}P_{x,s} = \sum_{y=1}^{\nu} {}_{m,n}r_y \cdot {}^{y}P_{x,s} \tag{1} $$

If p_y is the probability that generally the system of causes y^{th} takes place and ${}^{y}P_{m,n}$ is the probability that, if it is acting, A takes place m times in n observations, it will be (for the well known Bayes theorem):

$$ {}_{m,n}r_y = \frac{p_y \; {}^{y}P_{m,n}}{\sum_{y=1}^{\nu} p_y \; {}^{y}P_{m,n}} \tag{2} $$

[2] Usually, in the treatises on probability calculus, the phenomenon A is said to depend, in such a case, on the same "cause" during the $n+s$ observations. The word cause has a different meaning than that currently used: "cause" is not the complex of circumstances which necessarily determines the occurrence of A, but the complex of circumstances which gives A a certain probability of occurrence. There are as many "causes" of A as there are probabilities of A occurring during $n+s$ observations. We preferred to speak of "system of causes", so as not to deviate from the common meaning of the words.

hence

$$ {}_{m,n}P_{x,s} = \frac{\sum_{y=1}^{\nu} p_y \; {}^yP_{m,n} \; {}^yP_{x,s}}{\sum_{y=1}^{\nu} p_y \; {}^yP_{m,n}} \tag{3} $$

The practical importance of this result is at any rate almost nil. The value of ${}^yP_{m,n}$ and ${}^yP_{x,s}$ could easily be determined if the probability ${}^yP$ was known, that A occurs when the system of causes y^{th} is in action. But (except for certain conventionally arranged games) we lack in this case the knowledge of ${}^yP$, as well as that of p_y, becoming thus impossible to determine the value of ${}_{m,n}P_{x,s}$, on the base of (3).

Due to this ignorance, some believe they can resort to the hypotheses that ν is infinitely large, that p_y is equal for all the ν systems of causes and that ${}^yP$ assumes, as y varies, all the infinite values between 0 and 1.

In such A case it is demonstrated[3] that it is

$$ {}_{m,n}P_{x,s} = \frac{s!(m+x)!(n-m+s-x)!(n+1)!}{x!(s-x)!(s+n+1)!m!(n-m)!} \tag{4} $$

and in particular

$$ {}_{m,n}P = \frac{m+1}{n+2} \tag{5} $$

But, it is evident that these hypotheses are totally arbitrary. It is true that some authors have stated that, if n is very large, one might assume to use formula (5), even when these hypotheses are not satisfied[4], but there is no valid proof of such a statement and we will later note (§ 9) how, generally, this is not correct.

5. The determination of *a posteriori* probability, according to the outline considered in the previous paragraph, presumes some *a priori* information, regarding the number and the probability of the different systems of causes and regarding the probability that when each of them takes place, phenomenon A occurs. We may say that this is an *a priori outline for the a posteriori determination of probability.* We shall call it: method for the *a posteriori* probability determination of A phenomenon on the base of the probability of causes.

I believe that to this, one can conveniently counterpoise an *a posteriori way for the a posteriori determination of probability.*

Instead of starting from the consideration of the various systems of causes, which might affect the occurrence of A in $n+s$ observations, when it has occurred m times in n observations, we start from the consideration of the various times A can occur in $n+s$ observations, when it has occurred m times in n observations.

[3] See, for the proof, E. Czuber, *Warhrscheinlichkeitsrechnung.* 1. Hälfte Leipzig, Teubner, 1902, p. 165.

[4] See, for example, E. Czuber, op cit. p. 166.

Each one of these numbers represents a *direct result.*

When A has taken place m times in n observations, in the $n+s$ observations it can occur $m, m+1, m+2, \ldots, m+s$ times. Therefore there are $s+1$ possible *direct results.*

We shall call i^{th} direct result the one according to which the phenomenon A occurs $m+i$ times in $n+s$ observations.

We shall call this scheme: *method for the a posteriori determination of the probability of phenomenon based on the probability of direct results.*

${}_{m+1,n+s}P_{i,s}$ denotes the probability that, when one obtains the i^{th} direct result, that is when A has occurred $m+1$ times in $n+s$ observations, A has occurred m times in the first n observations, i.e. i times in the last s observations. This holds, in the hypothesis that the probability of A occurring keeps constant during the $n+s$ observations.

With $f_i = P_{m+1,n+s}$ we denote the probability that the i^{th} result occurs, with $P_{m,n}$, the probability that phenomenon A occurs m times in n observations.

The probability that A occurs m times in n observations and i times in the following s observations will be, by the theorem of conditional probability, $P_{m,n}\ {}_{m,n}P_{i,s}$.

The probability that A occurs to $m+i$ times in $n+s$ observations and i times in the last s observations will be, by the same theorem, $f_i\ {}_{m+i,n+s}P_{i,s}$.

On the other hand, to say that phenomenon A occurs m times in n observations and i times in the following s observations, amounts to say that A occurs $m+i$ times in the $n+s$ observations and i times in the last s observations.

Hence we can write:

$$f_i\ {}_{m+i,n+s}P_{i,s} = P_{m,n}\ {}_{m,n}P_{i,s}$$

from which

$${}_{m,n}P_{i,s} = \frac{f_i\ {}_{m+i,n+s}P_{i,s}}{P_{m,n}}$$

It follows that

$$\frac{{}_{m,n}P_{x,s}}{\sum_{i=0}^{s} {}_{m,n}P_{i,s}} = \frac{f_x\ {}_{m+x,n+s}P_{x,s}}{\sum_{i=0}^{s} f_i\ {}_{m+i,n+s}P_{i,s}}$$

and, as it is $\sum_{i=0}^{s}\ {}_{m,n}P_{i,s} = 1$

$${}_{m,n}P_{x,s} = \frac{f_x\ {}_{m+x,n+s}P_{x,s}}{\sum_{i=0}^{s} f_i\ {}_{m+i,n+s}P_{i,s}} \tag{6}$$

When the probability P of the phenomenon A occurring remains constant in the $n+s$ observations, the value of ${}_{m+i,n+s}P_{i,s}$, is given by the formula[5]

[5] If the probability P of phenomenon A is constant in the $n+s$ observations, then the probability that it takes place i time in the last s observations is independent of the

$$ {}_{m+i,n+s}P_{i,s} = \frac{s!(m+i)!(n+s-m-i)!n!}{i!(s-i)!(n+s)!m!(n-m)!} \tag{7} $$

Then (6) becomes

$$ {}_{m,n}P_{x,s} = \frac{f_x \dfrac{(m+x)!(n+s-m-x)!}{x!(s-x)!}}{\displaystyle\sum_{i=0}^{s} f_i \dfrac{(m+i)!(n+s-m-i)!}{i!(s-i)!}} \tag{8} $$

In the particular case $s = 1$, $x = 1$, it is

$$ {}_{m,n+1}P_{0,1} = \frac{n-m+1}{n+1} \qquad {}_{m+1,n+1}P_{1,1} = \frac{m+1}{n+1} $$

hence

$$ {}_{m,n}P = \frac{R(m+1)}{R(m+1)+n-m+1} \tag{9} $$

where $R = \frac{f_1}{f_0}$

Often statistical surveys allow to determine, from A large enough number of observations, the frequencies with which the various direct results occur. Such frequencies can be thought of, as approximate values of the related probabilities f_i.

Then it becomes easy to calculate the value of ${}_{m,n}P_{x,s}$.

Let us give some examples.

6. Example 1. We are dealing with A trigeminous birth. The woman concerned has already given birth to 2 males. The third baby is expected. What is the probability that it will also be a male?

There are 2 possible direct results: 3 males; 2 males and 1 female.

---

frequency with which it has taken place in the previous n observations. It is then

$$ P_{m,n}\ {}_{m,n}P_{i,s} = P_{m,n}P_{i,s} $$

And remembering that it is $f_i = P_{m+1,n+s}$

$$ P_{m+i,n+s}\ {}_{m+i,n+s}P_{i,s} = P_{m,n}P_{i,s} $$

$$ {}_{m+i,n+s}P_{i,s} = \frac{P_{m,n}P_{i,s}}{P_{m+i,n+s}} $$

By substituting, in this equation, for $P_{m,n}, P_{i,s}, P_{m+i,n+s}$ the values given by the known formula

$$ P_{b,c} = \frac{c!}{b!(c-b)!}P^b(1-P)^{c-b} $$

(7) is obtained.

Statistics points out that in 100 trigeminous births, in 24 cases all 3 were males and in 27 cases there were 2 males and 1 female (Italy: 1884-95).

When the direct result obtained is: 3 males, the probability that the first two born were males is equal to 1.

When the direct result obtained is: 2 males and 1 female, the probability that the first two born are males is 1/3, on the hypothesis that the probability of a male birth does not vary with the generation order.

The theoretical probability that the third born in a trigeminous birth is a male, when the first two were males, will be

$$\frac{24 \times 1}{24 \times 1 + 27 \times 1/3} = \frac{24}{33}$$

that is, approximately, 3/4.

From here onwards, I use the expression "theoretical probability" to define the probability ${}_{m,n}P_{x,s}$ calculated on the hypothesis that probability P remains constant for the $n + s$ observations.

EXAMPLE 2. A two-race match is run between A and B. Competitor A has already won one race. What is the probability that he will win the second as well?

The race bulletin allows us to establish that, in the course of the numerous encounters between A and B, A won both races 35 out of the 100 times, 50 out of the 100 he won one race only, and 15 out of the 100 he lost both races.

In our case, only two direct results are possible: either A wins both of the races, or he wins only one. If A wins both the races, there is a probability equal to 1 that he has won the first race. If A wins one race out of two, and the probability of winning is independent of the order of the races, there is a probability equal to 1/2 that he has won the first race.

The theoretical probability that he wins the second race, when he has won the first, will therefore be

$$\frac{35}{35 \times 50/2} = \frac{7}{12}.$$

EXAMPLE 3. A couple has up to now given birth to 3 males and 1 female: what is the probability that they will have, in the next two deliveries: 2 males; 1 male 1 female; 2 females?

There are 3 possible direct results: 5 males and 1 female; 4 males and 2 females; 3 males and 3 females.

Statistics shows that the frequencies of these three direct results are related to each other almost at the ratio 10:23:30 (Saxony 1876-85).

If the probability of having a male or a female is constant during the entire fertile period, the probability of having had, in the first 4 births, 3 males and 1 female is easily found by (7), to be 2/3 for those who, in 6 births, had 5 males and 1 female, 8/15 for those who, in 6 births, had 4 males and 2 females, 1/5 for those who, in 6 births, had 3 males and 3 females.

The couple that has so far given birth to 3 males and 1 female then has the theoretical probability

$$\frac{23 \times 8/15}{10 \times 2/3 + 23 \times 8/15 + 30 \times 1/5} = \frac{92}{187}$$

of giving birth in the next two deliveries to one male and one female. The couple has the theoretical probability 50/187 of giving birth to 2 males, and the theoretical probability 45/187 of giving birth to 2 females.

EXAMPLE 4. Let us consider a similar case. Until now a couple had 2 males and 4 females. They lost the 2 males and would like to have another one: what is the probability of their having another child of male sex?

There are two possible direct results: 3 males and 4 females, 2 males and 5 females.

Let us assume that statistics tells us only that the frequency of the first direct result is to that of the second as 191 is to 107 (Saxony: 1876-85).

If the probability of having a male is constant during the entire fertile period, the probability of having 2 males and 4 females in the first 6 births is 3/7 when, in 7 births, we have 3 males and 4 females and 5/7 when in 7 births, we have 2 males and 5 females.

The theoretical probability that the said couple will have a male child in the 7^{th} birth is then

$$\frac{191 \times 3/7}{191 \times 3/7 + 107 \times 5/7 \times 1/5} = 0.517.$$

7. In the previous examples we have always talked about theoretical probabilities. One would expect that these probabilities would not differ systematically from the actual frequencies, if the hypothesis from which we started was correct: the probability of the considered phenomenon occurring remains constant during the $n+s$ observations. Sometimes the statistical data enable us to directly calculate the actual frequencies and the comparison between the latter and the theoretical probabilities allows to verify if the probability of A occurring keeps truly constant during the $n+s$ observations or if it instead varies, and how.

For example, statistics enable us to establish that in Saxony (1876-85), 15,700 couples, who had 2 males and 4 females, had a 7^{th} male in 53% of the cases. The theoretical probability was instead less than 52%. (§ 6, example 4). This difference introduces the doubt that the probability of having a male does not remain constant during the entire fertile period. The subject is far too interesting not to bother to go deeper into it. The second part of the study deals with it.

8. Let us consider the particular case in which all the "direct results" have the same probability. In such a case it will be

$$_{m,n}P_{x,s} = \frac{_{m+x,n+s}P_{x,s}}{\sum_{i=0}^{s} {}_{m+i,n+s}P_{i,s}} \tag{10}$$

Substituting in the numerator the value given by (7) and, in the denominator, the value given by the following formula[6]:

$$\sum_{i=0}^{s} {}_{m+i,n+s}P_{i,s} = \frac{n+s+1}{n+1}$$

we obtain

$$_{m,n}P_{x,s} = \frac{s!(m+x)!(n-m+s-x)!(n+1)!}{x!(s-x)!(s+n+1)!m!(n-m)!} \tag{11}$$

This formula is identical to (4).

The probability that the phenomenon takes place x times in s observations, when it has occurred m times in n previous observations is therefore the same: a) when the $s+1$ direct results are equally probable; b) when the phenomenon allows with the same probability an infinite number of systems of causes, according to which the phenomenon has a probability of occurrence which ranges from 0 to 1.

If $s = 1$, $x = 1$, (11) becomes[7]

$$_{m,n}P = \frac{m+1}{n+2} \tag{12}$$

that gives the *a posteriori* probability of phenomenon A, when it has occurred m times in n previous observations. (12) is identical to (5).

The *a posteriori* probability of phenomenon A, that has occurred m times in n previous observations, is therefore the same: α) when, in $n+1$ observations, A has the same probability of occurring m or $m+1$ times; β) when the phenomenon allows, with the same probability, an infinite number of systems of causes, according to which the phenomenon has a probability of occurring which ranges from 0 to 1.

It is obvious that, having satisfied condition α), condition β) may not be satisfied. It may in fact be that phenomenon A has the same probability of occurring m or $m+1$ times $n+1$ observations and that it does not allow certain systems of causes. For example, a phenomenon may have the same probability of occurring 24 and 26 times in 50 observations and allow only 2 systems of causes, according to which it has a probability of occurring of 24/50 and 26/50 respectively.

It can instead be demonstrated that, if condition β) is fulfilled, condition α) is as well.

If ${}^{y}P$ is the probability of phenomenon A occurring when the system of causes y^{th} is in action, the probability that A takes place in $m+1$ times on $n+1$ observations, will be

$${}^{y}P_{m+1,n+1} = \frac{(n+1)!}{(m+1)!(n-m)!}\,{}^{y}P^{m+1}(1-{}^{y}P)^{n-m}$$

[6] The proof of this equation is attributed to Prof. Luigi Galvani, Assistant Professor of Differential calculus in the R. University of Cagliari. It is shown in the Appendix to [the original] paper.

[7] This formula can be directly deduced from (9), noting that in such a case it is $f_1 = f_0$, $R = 1$.

If there are infinite systems of causes, all equally probable, according to which ${}^yP$ assumes the infinite values from 0 to 1, the probability p_y that system y_{th} becomes active will be constant and it will be possible to set it equal to the difference $d\,{}^yP$ between 2 successive values of ${}^yP$.

Hence, the probability that A occurs $m+1$ times in $n+1$ observations will be

$$P_{m+1,n+1} = \frac{(n+1)!}{(m+1)!(n-m)!}\int_0^1 {}^yP^{m+1}(1-{}^yP)^{n-m}\,d\,{}^yP \tag{13}$$

By partial integrations, it is found that

$$\int_0^1 x^h(1-x)^k dx = \frac{h!k!}{(h+k+1)!}$$

hence

$$P_{m+1,n+1} = \frac{1}{n+2}$$

Likewise it is demonstrated that

$$P_{m,n+1} = \frac{(n+1)!}{m!(n-m+1)!}\int_0^1 {}^yP^m(1-{}^yP)^{n-m+1}d\,{}^yP = \frac{1}{n+2} \tag{14}$$

hence

$$P_{m,n+1} = P_{m+1,n+1}$$

Condition α) is then weaker than β).

Thus, by means of formula (5), we have found a validity condition which is more general than the one which is commonly proposed, that the phenomenon allows, with equal probability, infinite systems of causes, according to which its probability of occurring may take any of the infinite values from 0 to 1.

It is true that some authors maintain that, as stated earlier, when n is very large, the formula (5) remains valid even when the various systems of causes do not have the same probability of acting; but how this is not generally true will be demonstrated in the following paragraph.

9. Let us consider a more general situation than that examined in the previous paragraph.

Let us again assume that there are infinite systems of causes, according to which ${}^yP$ assumes the infinite values from 0 to 1; the probability that the system of causes y^{th} occurs, is

$$P_y = \frac{(k+h+1)!}{k!h!}\,{}^yP^k(1-{}^yP)^h\,d\,{}^yP \tag{15}$$

where the constant k and h are integer positive numbers[8].

[8] In the particular case when $k=0$, $h=0$ it is $p_y = d\,{}^yP$ and one goes back to the case considered in the previous paragraph. (When k or h are not integers, the function Γ is substituted for the factorial. The distribution (15) is known as the "β distribution"; it constitutes a particular case of K Pearson type I curves).

It is $\sum_{y=1}^{\nu} P_y = 1$.

Instead of (13), it is then

$$P_{m+1,n+1} = \frac{(n+1)!}{(m+1)!(n-m)!} \frac{(k+h+1)!}{k!h!} \int_0^1 {}^yP^{m+k+1}(1-{}^yP)^{n-m+h} d\,{}^yP$$

and instead of (14),

$$P_{m,n+1} = \frac{(n+1)!}{m!(n-m+1)!} \frac{(k+h+1)!}{k!h!} \int_0^1 {}^yP^{m+k}(1-{}^yP)^{n-m+h+1} d\,{}^yP$$

From which

$$R = \frac{P_{m+1,n+1}}{P_{m,n+1}} = \frac{m+k+1}{m+1} \frac{n-m+1}{n-m+h+1}$$

and, on the basis of (9),

$$_{m,n}P = \frac{m+k+1}{n+k+h+2} \tag{16}$$

Because k and h may assume any integer positive value, it is obvious that the value of formula (16), for any n, may differ essentially from the value of formula (5).

10. To determine the probability that a phenomenon A, which has occurred m times in n observations, will occur i times in the following s observations, we cannot always make use of the frequency with which phenomenon A occurs $m+i$ times in $n+s$ observations. One often knows only the frequency with which phenomenon A occurs $m+y$ times in $n+t$ observations, where $t > s$.

When A takes place $m+i$ times in $n+s$ observations, we said that the y^{th} direct result occurs; when A takes place $m+y$ times in $n+1$ observations, we will say that the y^{th} *indirect result* occurs.

Let us try to define an *a posteriori* scheme for the *a posteriori* determination of the probability, a scheme more general than the one we considered in § 5, which is based on the frequency of the results, irrespective of these being direct or indirect. In general, we will call it: *scheme to a posteriori determine the probability of a phenomenon, on the basis of the probability of the results*[9].

[9] In a paper on *Os fundamentos e o alcance do metodo estatistico*, in "Revista Brasileira de Estatistica", IX, N. 33, 1948, a formula is given for such a scheme, on the basis of probability of results, which is formally different, but which substantially coincides with (17), In this paper a case is also considered in which we do not know either the direct results or the indirect ones (which are called "superior indirect results") related to a number of cases greater than that of direct results, but only the indirect results (which are called "inferior indirect results") related to a number of cases smaller than that of direct results are known, it is shown how the scheme of inferior indirect results is applied and which hypotheses such a scheme assumes.

Let us denote with f_y the probability that the y^{th} result occurs. Given all the values of f_y it is easy to calculate the single values of f_i on the basis of the formula

$$f_i = \sum_{y=0}^{i} f_y \; {}_{m+y,n+1}P_{m+1,n+s}$$

Instead of (6), we must therefore put

$${}_{m,n}P_{x,s} = \frac{\displaystyle\sum_{y=0}^{t} f_y \; {}_{m+y,n+t}P_{m+x,n+s} \; {}_{m+x,n+s}P_{x,s}}{\displaystyle\sum_{i=0}^{s}\sum_{y=0}^{t} f_y \; {}_{m+y,n+t}P_{m+i,n+s} \; {}_{m+1,n+s}P_{i,s}} \tag{17}$$

To calculate the second member of this equality it can be noted that, when the probability of the occurrence A is constant in the $n+t$ observations, it is

$$\begin{aligned} {}_{m+y,n+t}P_{m+i,n+s} = &\frac{(t-s)!}{(y-i)!(t-s-y+1)!} \\ &\frac{(m+y)!(n+t-m-y)!}{(n+t)!} \frac{(n+s)!}{(m+i)!(n+s-m-i)!} \end{aligned} \tag{18}$$

From (18) and from (7), one can easily deduces that the numerator of (17): it is

$$\begin{aligned} = \sum_{y=0}^{t} f_y \frac{s!}{x!(s-x)!} \frac{n!}{m!(n-m)!} \\ \frac{(t-s)!}{(y-x)!(t-s-y+x)!} \frac{(m+y)!(n+t-m-y)!}{(n+t)!} \end{aligned} \tag{19}$$

and the denominator

$$\begin{aligned} = \sum_{y=0}^{t} \Bigg(f_y \frac{n!}{m!(n-m)!} \frac{(m+y)!(n+t-m-y)!}{(n+t)!} \\ \sum_{i=0}^{s} \frac{s!}{i!(s-i)!} \frac{(t-s)!}{(y-i)!(t-s-y+i)!} \Bigg) \end{aligned} \tag{20}$$

Note now that it is

$$\sum_{i=0}^{s} \frac{s!}{i!(s-i)!} \frac{(t-s)!}{(y-i)!(t-s-y+i)!} \frac{y!(t-y)!}{t!} = 1$$

and hence

$$\sum_{i=0}^{s} \frac{s!}{i!(s-i)!} \frac{(t-s)!}{(y-i)!(t-s-y+i)!} = \frac{t!}{y!(t-y)!} \tag{21}$$

Substituting this value in (20), dividing (19) by (20) and by simplification, one obtains

$$
{}_{m,n}P_{x,s} = \frac{s!(t-s)!}{t!} \frac{\displaystyle\sum_{y=0}^{t} f_y \frac{(m+y)!(n+t-m-y)!}{x!(s-x)!(y-x)!(t-s-y+x)!}}{\displaystyle\sum_{y=0}^{t} f_y \frac{(m+y)!(n+t-m-y)!}{y!(t-y)!}} \tag{22}
$$

In the particular case $t = s$, $y = i$, (22) is easily simplified to (8), remembering[10] that it is $1! = 1$, $0! = 1$; $(-K)! = \infty$.

11. Let us see which value (22) assumes, when f_y is constant, that is when the $t+1$ results are equally probable. Because of this, let us observe that it is

$$
\sum_{y=0}^{t} {}_{m+y,n+t}P_{y,t} = \frac{n+t+1}{n+1}
$$

from which it is immediately obtained

$$
\sum_{y=0}^{t} \frac{(m+y)!(n+t-m-y)!}{y!(t-y)!} = \frac{(n+t+1)!m!(n-m)!}{(n+1)t!} \tag{23}
$$

Likewise

$$
\sum_{y=0}^{t} {}_{m+y,n+t}P_{y-x,t-s} = \frac{n+t+1}{n+s+1}
$$

from which it is easily obtained

$$
\sum_{y=0}^{t} \frac{(m+y)!(n+t-m-y)!}{x!(s-x)!(y-x)!(t-s-y+x)!} = \frac{(n+t+1)!(m+x)!(n+s-m-x)!}{x!(s-x)!(n+s+1)!(t-s)!} \tag{24}
$$

If f_y is constant, and the values obtained from (23) and (24) are substituted in (22), this becomes the same as (11).

Then, when the $t+1$ results (direct or indirect) are equally probable, the value of ${}_{m,n}P_{x,s}$ is independent of t; in other words, it is the same, whatever the number of future observations to which the results are referred.

12. Let us apply (22) to an example.

[10] From the combination concept it is immediately derived $\frac{(h-K)!}{h!(-K)!} = 0$ from which $\frac{1}{(-K)!} = 0$, hence $(-K)! = \infty$

The first born in a trigeminous birth was a male: we want to find out the probability that the second born will also be a male.

There are three possible indirect results: 3 males, 2 males and 1 female; 2 females and 1 male; the respective frequencies of which are: $f_2 = 24\%$; $f_1 = 27\%$; $f_0 = 26\%$. (Italy, 1884-85).

On the basis of (22), the probability we are looking for is equal to 89/152, approximately 2/3.

This calculation presumes that the probability of having a male or a female is the same for the first, second or third child born in a trigeminous birth.

13. The *a posteriori* determination of the probability of A, on the basis of the probability of the "direct results", assumes that the probability of A remains constant in the n observations already taken and in the s observations yet to be taken. The *a posteriori* determination of the probability of A on the basis of the probability of the indirect results assumes that the probability of the occurrence of A remains constant in the n observation already taken and in a number t of future observations, which is greater than the number s of observations yet to be taken.

The hypothesis that the probability of A remains constant for a given number of observations will evidently be more plausible, the smaller the number of observations considered.

This does not necessarily mean that the *a posteriori* determination of the probability of A occurring i times in s future observations is less precise when it is obtained on the basis of the probability of "indirect results", than when it is obtained on the basis of the probability of the "direct results". Nor that, when it is obtained on the basis of "indirect results", it is much less precise, the greater is the number t of future observations, to which the results are related. This conclusion would be true when the probability of A does not vary in the first future observations and does vary in the subsequent observations, and likewise when it would vary continuously throughout all future observations, always increasing or always decreasing. But if, during future observations, it changes first in one direction and then in the opposite one, it might be that the *a posteriori* determination of the probability is more precise when obtained on the basis of "indirect results", than when it is based on "direct results". It might also be that, when based on "indirect results", it is more precise when the number of observations, to which the "indirect results" are related, is greater.

Let us verify this for the data on the sex of the newly born in Saxony (1876-85).

The frequency with which a male is born in the second birth, when in the first A male was born, is ${}_{1,1}P" = 0,51931$.

The theoretical probability ${}_{1,1}P'$, calculated on the basis of the frequency of the result related to t future observations, varies with t as follows [see Table 1].

As t increases, the difference between ${}_{1,1}P'$ and ${}_{1,1}P"$ at first increases, then decreases. This makes us think that, in the families where a male was born first, the tendency to have males in the following births increases at first and decreases afterwards. This conclusion is consistent with the results obtained in the second part of this study: in fact it turns out that in the families, which had all males

Table 1 *Table 2*

t	${}_{1,1}P'$	t	${}_{3,4}P$"
1	0.51908	1	0.50955
2	0.51747	2	0.51205
3	0.51533	3	0.51633
4	0.51452	4	0.51709
5	0.51552	5	0.51967
6	0.51659	6	0.52365

in the first birth or in the first two births, the actual frequency of having a male in the following delivery is greater than the theoretical one; instead, the families which had all males in more than two births, in the next one show a frequency of having males which is less than the theoretical probability.

Instead, when in the past births the males are in excess, but there are some females as well, the actual frequency of having males in the following births is always less than the theoretical probability. Hence, it appears that the frequency of a male birth, after having had m males and $n - m < m$ females, differs from the theoretical probability determined on the basis of the frequency of the results in $n + t$ observations, and it differs more when t is greater. As an example, let us consider a family which, until now, had 3 males and 1 female. The frequency of having a male in the 5th birth is ${}_{3,4}P$" $= 0.50350$. The theoretical probability is as follows [see Table 2].

The difference between ${}_{3,4}P$" and ${}_{3,4}P'$ is very small for $t = 1$, increases with t and becomes quite considerable for $t = 6$.

14. Formula (17) can be written in a different form. It must be noted that when the probability of A remains constant in the $n + t$ observations, it is

$$ {}_{m+y,n+t}P_{m,n} = \frac{n!}{m!(n-m-)!} \frac{(m+y)!(n+t-m+y)!}{(n+t)!} \frac{t!}{t!(t-y)!} $$

$$ {}_{y,t}p_{i,s} = \frac{s!}{i!(s-i)!} \frac{y!(t-y)!}{t!} \frac{(t-s)!}{(y-i)!(t-s-y+i)!} $$

From these formulae and from (19), (20), (21) it is

$$ \sum_{y=0}^{t} f_y \; {}_{m+y,n+t}P_{m,n} \; {}_{y,t}P_{x,s} = \sum_{y=0}^{t} f_y \; {}_{m+y,n+t}P_{m+x,n+s} \; {}_{m+x,n+s}P_{x,s} $$

$$ \sum_{y=0}^{t} f_y \; {}_{m+y,n+t}P_{m,n} \; {}_{y,t}P_{x,s} = \sum_{i=0}^{s} \sum_{y=0}^{t} f_y \; {}_{m+y,n+s}P_{m+i,n+s} \; {}_{m+i,n+s}P_{i,s} $$

Hence, (17) can also be written as follows:

$$ {}_{m,n}P_{x,s} = \frac{\sum_{y=0}^{t} f_y \ {}_{m+y,n+t}P_{m,n} \ {}_{y,t}P_{x,s}}{\sum_{y=0}^{t} f_y \ {}_{m+y,n+t}P_{m,n}} \tag{25} $$

15. Formula (25) is useful to show the relationship existing between the scheme for the *a posteriori* determination of the probability of a phenomenon A, on the basis of the probability of causes, and the scheme for the *a posteriori* determination of the probability A, on the basis of the probability of results.

Let us consider a specific case in which t is infinitely large, x and m are negligible compared to y, and s and n are negligible compared to t.

As the probability is the limit to which the frequency of a phenomenon tends, as the number of observations increases, we could say that, when t is infinitely large, to different results correspond different probabilities of phenomenon A occurring, i.e. different systems of causes. The first system of causes will correspond to result 0; the second system of causes to result 1; generally, the system of causes $(y+1)^{th}$ will correspond to result y^{th}. We could that put

$$ p_{y+1} = f_y $$

and (see note (5))

$$ {}_{m+y,n+t}P_{m,n} = \frac{{}^{y+1}P_{m,n} \ {}^{y+1}P_{y,t}}{{}^{y+1}P_{m+y,n+t}} $$

$$ {}_{y,t}P_{x,s} = \frac{{}^{y+1}P_{x,s} \ {}^{x+1}P_{y-x,t-s}}{{}^{y+1}P_{y,t}} $$

But, as m and n are negligible compared to y and to t respectively, we could put

$$ \frac{{}^{y+1}P_{y,t}}{{}^{y+1}P_{m+y,n+t}} = 1 $$

hence

$$ {}_{m+y,n+t}P_{m,n} = {}^{y+1} P_{m,n} $$

Similarly, as x and s are negligible compared to y and t respectively, we could put

$$ \frac{{}^{y+1}P_{y-x,t-s}}{{}^{y+1}P_{y,t}} $$

hence

$$ {}_{y,t}P_{x,s} = {}^{y+1}P_{x,s} $$

If in (25) f^y, ${}_{m+y,n+t}P_{m,n}$ and ${}_{y,t}P_{x,s}$ are substituted by, p_{y+1}, ${}^{y+1}P_{m,n}$ and ${}^{y+1}P_{x,s}$ respectively, then (25) becomes the same as (3).

The scheme for the *a posteriori* determination of the probability of A, on the basis of the probability of causes may therefore be looked at as a particular case of the scheme of the *a posteriori* determination of the probability of A, on the basis of the probability of the results; a situation which occurs when the probability determination is based on the probability of the "indirect results" from a very large number of observations.

16. By introducing a slight modification in (3), one obtains a formula for the *a posteriori* determination of probabilities which is valid either for the method of the probability of results, or for the method of the probability of causes.

The occurrence of phenomenon A in the s future observations may depend on several systems of causes, or it may correspond to different "direct" or "indirect" results. Every system of causes, and similarly every result, may therefore be looked upon as an eventuality[11].

We will indicate by ν the number of the eventuality, by p_y the probability that the y^{th} eventuality will occur, by ${}^yP$ the probability of the occurring of A, according to the y^{th} eventuality, by ${}^yP_{x,s}$ the probability that, in the y^{th} eventuality, A will occur x times in s future observations, with ${}^{\,y}_{m,n}P_{x,s}$ the probability that, in the y^{th} eventuality, A will occur x times in s future observations, when it has occurred m times in n past observations.

Let us first suppose that the eventualities of which we know the single probabilities p_y consist in results. Adopting the new symbols, (25) becomes:

$$ {}_{m,n}P_{x,s} = \frac{\sum_{y=1}^{\nu} p_y \; {}^yP_{m,n} \; {}^{\,y}_{m,n}P_{x,s}}{\sum_{y=1}^{\nu} p_y \; {}_yP_{m,n}} \tag{26} $$

Let us now suppose that the evantualities, of which we know the single probabilities p_y consist in *system of causes*. As the probability ${}^yP$ will be maintained costant in $n+s$ observations, whatever the acting system of causes, it will be:

$$ {}^{\,y}_{m,n}P_{x,s} =^y P_{x,s} $$

and so (3) may be written as (26).

Formula (26) can therefore he seen as the general formula to determine probability *a posterioni.* Confronted with it, (3) and (25) represent particular formulae, of very different usage, valid when the eventualities, on the basis of which probability is *a posteriori* determined, consist in systems of causes or results respectively.

[11] Instead of the word "eventuality" it would be helpful to use the word "hypothesis". I preferred to speak of "eventuality", because the word "hypothesis" is generally used in the theory of probability, in a more limited way, as synonymous of "cause", that is, according to our terminology, of "systems of causes".

Summary

In this first paper of 1911 relating to the sex ratio at birth, Gini repurposed a Laplaces succession rule according to a Bayesian version. The Gini's intuition consisted in assuming for prior probability a Beta type distribution and introducing the "method of results (direct and indirect)" for the determination of prior probabilities according to the statistical frequency obtained from statistical data.

Keywords: Prior and posterior probabilities; Bayes-Laplace solution; Beta distribution; Sex ratio at birth.