A HISTORY

OF THE

MATHEMATICAL THEORY OF PROBABILITY

FROM THE TIME OF PASCAL TO THAT

OF LAPLACE.

BY

I. TODHUNTER, M.A., F.R.S.

Cambridge and London:
MACMILLAN AND CO.
1865.

CHAPTER XIV.

BAYES.

539. THE name of Bayes is associated with one of the most important parts of our subject, namely, the method of estimating the probabilities of the causes by which an observed event may have been produced. As we shall see, Bayes commenced the investigation, and Laplace developed it and enunciated the general principle in the form which it has since retained.

540. We have to notice two memoirs which bear the following titles:

An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, F.R.S. communicated by Mr Price in a Letter to John Canton, A.M. F.R.S. A Demonstration of the Second Rule in the Essay towards the Solution of a Problem in the Doctrine of Chances, published in the Philosophical Transactions, Vol. LIII. *Communicated by the Rev. Mr. Richard Price, in a Letter to Mr. John Canton, M.A. F.R.S.*

The first of these memoirs occupies pages 370—418 of Vol. LIII. of the *Philosophical Transactions*; it is the volume for 1763, and the date of publication is 1764.

The second memoir occupies pages 296—325 of Vol. LIV. of the *Philosophical Transactions*; it is the volume for 1764, and the date of publication is 1765.

541. Bayes proposes to establish the following theorem: If

an event has happened p times and failed q times, the probability that its chance at a single trial lies between a and b is

$$\frac{\int_a^b x^p (1-x)^q \, dx}{\int_0^1 x^p (1-x)^q \, dx}.$$

Bayes does not use this notation; areas of curves, according to the fashion of his time, occur instead of integrals. Moreover we shall see that there is an important condition implied which we have omitted in the above enunciation, for the sake of brevity: we shall return to this point in Art. 552.

Bayes also gives rules for obtaining approximate values of the areas which correspond to our integrals.

542. It will be seen from the title of the first memoir that it was published after the death of Bayes. The Rev. Mr Richard Price is the well known writer, whose name is famous in connexion with politics, science and theology. He begins his letter to Canton thus:

Dear Sir, I now send you an essay which I have found among the papers of our deceased friend Mr Bayes, and which, in my opinion, has great merit, and well deserves to be preserved.

543. The first memoir contains an introductory letter from Price to Canton; the essay by Bayes follows, in which he begins with a brief demonstration of the general laws of the Theory of Probability, and then establishes his theorem. The enunciations are given of two rules which Bayes proposed for finding approximate values of the areas which to him represented our integrals; the demonstrations are not given. Price himself added *An Appendix containing an Application of the foregoing Rules to some particular Cases.*

The second memoir contains Bayes's demonstration of his principal rule for approximation; and some investigations by Price which also relate to the subject of approximation.

544. Bayes begins, as we have said, with a brief demonstration of the general laws of the Theory of Probability; this part of his essay is excessively obscure, and contrasts most unfavourably with the treatment of the same subject by De Moivre.

Bayes gives the principle by which we must calculate the probability of a compound event.

Suppose we denote the probability of the compound event by $\frac{P}{N}$, the probability of the first event by z, and the probability of the second on the supposition of the happening of the first by $\frac{b}{N}$. Then our principle gives us $\frac{P}{N} = z \times \frac{b}{N}$, and therefore $z = \frac{P}{b}$. This result Bayes seems to present as something new and remarkable; he arrives at it by a strange process, and enunciates it as his Proposition 5 in these obscure terms:

If there be two subsequent events, the probability of the 2nd $\frac{b}{N}$ and the probability of both together $\frac{P}{N}$, and it being 1st discovered that the 2nd event has happened, from hence I guess that the 1st event has also happened, the probability I am in the right is $\frac{P}{b}$.

Price himself gives a note which shews a clearer appreciation of the proposition than Bayes had.

545. We pass on now to the remarkable part of the essay. Imagine a rectangular billiard table $ABCD$. Let a ball be rolled on it at random, and when the ball comes to rest let its perpendicular distance from AB be measured; denote this by x. Let a denote the distance between AB and CD. Then the probability that the value of x lies between two assigned values b and c is $\frac{c-b}{a}$. This we should assume as obvious; Bayes, however, demonstrates it very elaborately.

546. Suppose that a ball is rolled in the manner just explained; through the point at which it comes to rest let a line EF be drawn parallel to AB, so that the billiard table is divided into the two portions $AEFB$ and $EDCF$. A second ball is to be rolled on the table; required the probability that it will rest within the

space $AEFB$. If x denote the distance between AB and EF the required probability is $\frac{x}{a}$: this follows from the preceding Article.

547. Bayes now considers the following compound event: The first ball is to be rolled once, and so EF determined; then $p+q$ trials are to be made in succession with the second ball: required the probability, before the first ball is rolled, that the distance of EF from AB will lie between b and c, and that the second ball will rest p times within the space $AEFB$, and q times without that space.

We should proceed thus in the solution: The chance that EF falls at a distance x from AB is $\frac{dx}{a}$; the chance that the second event then happens p times and fails q times is

$$\frac{\lfloor p+q}{\lfloor p \lfloor q} \left(\frac{x}{a}\right)^p \left(1-\frac{x}{a}\right)^q;$$

hence the chance of the occurrence of the two contingencies is

$$\frac{dx}{a} \frac{\lfloor p+q}{\lfloor p \lfloor q} \left(\frac{x}{a}\right)^p \left(1-\frac{x}{a}\right)^q.$$

Therefore the whole probability required is

$$\frac{\lfloor p+q}{a \lfloor p \lfloor q} \int_b^c \left(\frac{x}{a}\right)^p \left(1-\frac{x}{a}\right)^q dx.$$

Bayes's method of solution is of course very different from the above. With him an area takes the place of the integral, and he establishes the result by a rigorous demonstration of the *ex absurdo* kind.

548. As a corollary Bayes gives the following: The probability, before the first ball is rolled, that EF will lie between AB and CD, and that the second event will happen p times and fail q times, is found by putting the limits 0 and a instead of b and c. But it is *certain* that EF will lie between AB and CD. Hence we

have for the probability, before the first ball is thrown, that the second event will happen p times and fail q times

$$\frac{\lfloor p+q}{a \lfloor p \lfloor q} \int_0^a \left(\frac{x}{a}\right)^p \left(1-\frac{x}{a}\right)^q dx.$$

549. We now arrive at the most important point of the essay. Suppose we only know that the second event has happened p times and failed q times, and that we wish to infer from this fact the probable position of the line EF which is to us unknown. The probability that the distance of EF from AB lies between b and c is

$$\frac{\int_b^c x^p (a-x)^q dx}{\int_0^a x^p (a-x)^q dx}.$$

This depends on Bayes's Proposition 5, which we have given in our Art. 544. For let z denote the required probability; then

$z \times$ probability of second event = probability of compound event.

The probability of the compound event is given in Art. 547, and the probability of the second event in Art. 548; hence the value of z follows.

550. Bayes then proceeds to find the area of a certain curve, or as we should say to integrate a certain expression. We have

$$\int x^p (1-x)^q dx = \frac{x^{p+1}}{p+1} - \frac{q}{1}\frac{x^{p+2}}{p+2} + \frac{q(q-1)}{1.2}\frac{x^{p+3}}{p+3} - \dots$$

This series may be put in another form; let u stand for $1-x$, then the series is equivalent to

$$\frac{x^{p+1}u^q}{p+1} + \frac{q}{p+1}\frac{x^{p+2}u^{q-1}}{p+2} + \frac{q(q-1)}{(p+1)(p+2)}\frac{x^{p+3}u^{q-2}}{p+3}$$
$$+ \frac{q(q-1)(q-2)}{(p+1)(p+2)(p+3)}\frac{x^{p+4}u^{q-3}}{p+4} + \dots$$

This may be verified by putting for u its value and rearranging according to powers of x. Or if we differentiate the series with respect to x, we shall find that the terms cancel so as to leave only $x^p u^q$.

551. The general theory of the estimation of the probabilities of causes from observed events was first given by Laplace in the *Mémoires ... par divers Savans*, Vol. VI. 1774. One of Laplace's results is that if an event has happened p times and failed q times, the probability that it will happen at the next trial is

$$\frac{\int_0^1 x^{p+1} (1-x)^q \, dx}{\int_0^1 x^p (1-x)^q \, dx}.$$

Lubbock and Drinkwater think that Bayes, or perhaps rather Price, confounded the probability given by Bayes's theorem with the probability given by the result just taken from Laplace; see *Lubbock and Drinkwater*, page 48. But it appears to me that Price understood correctly what Bayes's theorem really expressed. Price's first example is that in which $p=1$, and $q=0$. Price says that "there would be odds of three to one for somewhat more than an even chance that it would happen on a second trial." His demonstration is then given; it amounts to this:

$$\frac{\int_{\frac{1}{2}}^1 x^p (1-x)^q \, dx}{\int_0^1 x^p (1-x)^q \, dx} = \frac{3}{4},$$

where $p=1$ and $q=0$. Thus there is a probability $\frac{3}{4}$ that the chance of the event lies between $\frac{1}{2}$ and 1, that is a probability $\frac{3}{4}$ that the event is more likely to happen than not.

552. It must be observed with respect to the result in Art. 549, that in Bayes's own problem we *know* that *a priori* any position of EF between AB and CD is equally likely; or at least we know what amount of assumption is involved in this supposition. In the applications which have been made of Bayes's theorem, and of such results as that which we have taken from Laplace in

Art. 551, there has however often been no adequate ground for such knowledge or assumption.

553. We have already stated that Bayes gave two rules for approximating to the value of the area which corresponds to the integral. In the first memoir, Price suppressed the demonstrations to save room; in the second memoir, Bayes's demonstration of the principal rule is given: Price himself also continues the subject. These investigations are very laborious, especially Price's.

The following are among the most definite results which Price gives. Let $n = p + q$, and suppose that neither p nor q is small; let $h = \frac{\sqrt{(pq)}}{n\sqrt{(n-1)}}$. Then if an event has happened p times and failed q times, the odds are about 1 to 1 that its chance at a single trial lies between $\frac{p}{n} + \frac{h}{\sqrt{2}}$ and $\frac{p}{n} - \frac{h}{\sqrt{2}}$; the odds are about 2 to 1 that its chance at a single trial lies between $\frac{p}{n} + h$ and $\frac{p}{n} - h$; the odds are about 5 to 1 that its chance at a single trial lies between $\frac{p}{n} + h\sqrt{2}$ and $\frac{p}{n} - h\sqrt{2}$. These results may be verified by Laplace's method of approximating to the value of the definite integrals on which they depend.

554. We may observe that the curve $y = x^p(1-x)^q$ has two points of inflexion, the ordinates of which are equidistant from the maximum ordinate; the distance is equal to the quantity h of the preceding Article. These points of inflexion are of importance in the methods of Bayes and Price.