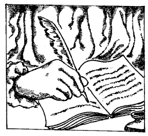


Reproduced below is a substantial part of the famous essay of Thomas Bayes. The results and conclusions are fully given, but some of the long and elaborate derivations based on geometric considerations have been left out due to page constraints.

An Essay Towards Solving a Problem in the Doctrine of Chances[1]

Problem

Given the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named.

SECTION 1

Definition

1. Several events are *inconsistent*, when if one of them happens, none of the rest can.
2. Two events are *contrary* when one, or other of them must; and both together cannot happen.
3. An event is said to *fail*, when it cannot happen; or, which comes to the same thing, when its contrary has happened.
4. An event is said to be determined when it has either happened or failed.
5. The *probability of any event* is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the value of the thing expected upon it's [2] happening.

[1] Reprinted with permission from publishers of *Biometrika*, Parts 3 and 4, Vol.45, pp.298-315, December 1958. The article originally appeared in *The Philosophical Transactions*, Vol. 53, pp.370-418, 1763 and was reprinted in *Biometrika*.

[2] The spelling "it's" was correct and appropriate form in Bayes' time even though today we would use "its."

6. By *chance* I mean the same as probability.
7. Events are independent when the happening of any one of them does neither increase nor abate the probability of the rest.

Proposition 1

When several events are inconsistent the probability of the happening of one or other of them is the sum of the probabilities of each of them.

Suppose there be three such events, and whichever of them happens I am to receive N, and that the probability of the 1st, 2nd and 3rd are respectively a/N, b/N, c/N. Then (by the definition of probability) the value of my expectation from the 1st will be a, from the 2nd b, and from the 3rd c. Wherefore the value of my expectations from all three will be $a+b+c$. But the sum of my expectations from all three is in this case an expectation of receiving N upon the happening of one or other of them. Wherefore (by definition 5) the probability of one or other of them is $(a+b+c)/N$ or $a/N+b/N+c/N$, the sum of the probabilities of each of them.

Corollary

If it be certain that one or other of the three events must happen, then $a+b+c=N$. For in this case all the expectations together amounting to a certain expectation of receiving N, their values together must be equal to N. And from hence it is plain that the probability of an event added to the probability of its failure (or of its contrary) is the ratio of equality. For these are two inconsistent events, one of which necessarily happens. Wherefore if the probability of an event is P/N that of it's failure will be $(N-P)/N$.

Proposition 2

If a person has an expectation depending on the happening of an event, the probability of the event is to the probability of its failure as his loss if it fails to his gain if it happens.

Suppose a person has an expectation of receiving N, depending on an event the probability of which is P/N. Then (by definition 5) the value of his expectation is P, and therefore if the event fails, he loses that which in value is P; and if it happens he receives N, but his expectation ceases. His gain therefore is $N-P$. Likewise since the probability of the event is P/N, that of its failure (by corollary prop. 1) is $(N-P)/N$. But P/N is to $(N-P)/N$ as P is to $N-P$, i.e. the probability of the event is to the probability of it's failure, as his loss if it fails to his gain if it happens.

Proposition 3

The probability that two subsequent events will both happen is a ratio compounded of the probability of the 1st, and the probability of the 2nd on supposition the 1st happens.

Suppose that, if both events happen, I am to receive N, that the probability both will happen is P/N, that the 1st will is a/N (and consequently that the 1st will not is $(N-a)/N$) and that the 2nd will happen upon supposition the 1st does is b/N. Then (by definition 5) P will be the value of my expectation, which will become b if the 1st happens. Consequently if the 1st happens, my gain by it is $b-P$, and if it fails my loss is P. Wherefore, by the foregoing proposition, a/N is to $(N-a)/N$, i.e. a is to $N-a$ as P is to $b-P$. Wherefore (*componendo inverse*) a is to N as P is to b. But the ratio of P to N is compounded of the ratio of P to b, and that of b to N. Wherefore the same ratio of P to N is compounded of the ratio of a to N and that of b to N, i.e. the probability that the two subsequent events will both happen is compounded of the probability of the 1st and the probability of the 2nd on supposition the 1st happens.

Corollary

Hence if of two subsequent events the probability of the 1st be a/N, and the probability of both together be P/N, then the probability of the 2nd on supposition the 1st happens is P/a.

Proposition 4

If there be two subsequent events to be determined every day, and each day the probability of the 2nd is b/N and the probability of both P/N, and I am to receive N if both the events happen the first day on which the 2nd does; I say, according to these conditions, the probability of my obtaining N is P/b. For if not, let the probability of my obtaining N be x/N and let y be to x as $N-b$ to N. Then since x/N is the probability of my obtaining N (by definition 1) x is the value of my expectation. And again, because according to the foregoing conditions the first day I have an expectation of obtaining N depending on the happening of both the events together, the probability of which is P/N, the value of this expectation is P. Likewise, if this coincident should not happen I have an expectation of being reinstated in my former circumstances, i.e. of receiving that which in value is x depending on the failure of the 2nd event the probability of which (by corollary to prop.1) is $(N-b)/N$ or y/x, because y is to x as $N-b$ to N. Wherefore since x is the thing expected and y/x the probability of obtaining it, the value of this expectation is y. But these two last expectations together are evidently the same with my original expectation, the value of which is x, and therefore $P+y=x$. But y is to x as $N–b$ is to N. Wherefore x is to P as N is to b, and x/N (the probability of my obtaining N) is P/b.

Corollary

Suppose after the expectation given me in the foregoing proposition, and before it is at all known whether the 1st event has happened or not, I should find that the 2nd event has happened; from hence I can only infer that the event is determined on which my expectation depended, and have no reason to esteem the value of my expectation either greater or less than it was before. For if I have reason to think it less, it would be reasonable for me to give something to be reinstated in my former circumstances, and this over and over again as often as I should be informed that the 2nd event had happened, which is evidently absurd. And the like absurdity plainly follows if you say I ought to set a greater value on my expectation than before, for then it would be reasonable for me to refuse something if offered me upon condition I would relinquish it, and be reinstated in my former circumstances; and this likewise over and over again as often as (nothing being known concerning the 1st event) it should appear that the 2nd had happened. Notwithstanding therefore this discovery that the 2nd event has happened, my expectation ought to be esteemed the same in value as before, i.e. x, and consequently the probability of my obtaining N is (by definition 5) still x/N or P/b. [3] But after this discovery the probability of my obtaining N is the probability that the 1st of two subsequent events has happened upon the supposition that the 2nd has, whose probabilities were as before specified. But the probability that an event has happened is the same as the probability I have to guess right if I guess it has happened. Wherefore the following proposition is evident.

Proposition 5

If there be two subsequent events, the probability of the 2nd b/N and the probability both together P/N, and it being first 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 P/b.[4]

[3] What is here said may perhaps be a little illustrated by considering that all that can be lost by the happening of the 2nd event is the chance I should have had of being reinstated in my former circumstances, if the event on which my expectation depended had been determined in the manner expressed in the proposition. But this chance is always as much *against* me as it is *for me*. If the 1st event happens, it is *against* me, and equal to the chance for the 2nd event's failing. If the 1st event does not happen, it is *for me*, and equal also to the chance for the 2nd event's failing. The loss of it, therefore, can be no disadvantage.

Proposition 6

The probability that several independent events shall all happen is a ratio compounded of the probabilities of each.

Corollary 1

If there be several independent events, the probability that the 1st happens, the 2nd fails, the 3rd fails and the 4th happens, etc. is a ratio compounded of the probability of the 1st, and the probability of the failure of the 2nd, and the probability of the failure of the 3rd, and the probability of the 4th, etc. For the failure of an event may always be considered as the happening of its contrary.

Corollary 2

If there be several independent events, and the probability of each one be a, and that of its failing be b, the probability that the 1st happens and the 2nd fails, and the 3rd fails and the 4th happens, etc. will be $abba$, etc. For, according to the algebraic way of notation, if a denotes any ratio and b another, $abba$ denotes the ratio compounded of the ratios a,b,b,a. This corollary therefore is only a particular case of the foregoing.

Definition

If in consequence of certain data there arises a probability that a certain event should happen, its happening or failing, in consequence of these data, I call it's happening or failing in the 1st trial. And if the same data be again repeated, the happening or failing of the event in consequence of them I call its happening or failing in the 2nd trial; and so on as often as the same data are repeated. And hence it is manifest that the happening or failing of the same event in so many different trials, is in reality the happening or failing of so many distinct independent events exactly similar to each other.

[4] What is proved by Mr. Bayes in this and the preceding proposition is the same with the answer to the following question. What is the probability that a certain event, when it happens, will be accompanied with another to be determined at the same time? In this case, as one of the events is given, nothing can be due for the expectation of it; and, consequently, the value of an expectation depending on the happening of both events must be the same with the value of an expectation depending on the happening of one of them. In other words: the probability that, when one of two events happens, the other will, is the same with the probability of this other. Call x then the probability of this other, and if b/N be the probability of the given event, and p/N the probability of both, because $p/N=(b/N) \times x$, $x = p/b =$ the probability mentioned in these propositions.

Proposition 7

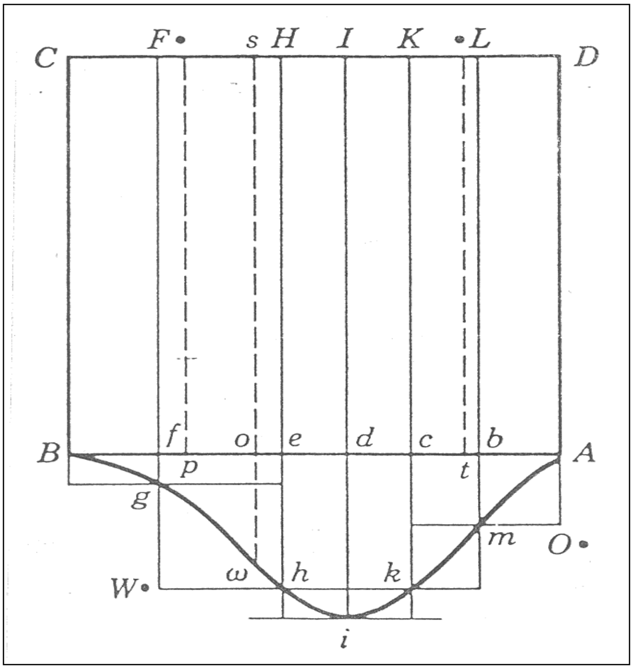
If the probability of an event be a, and that of its failure be b in each single trial, the probability of its happening p times, and failing q times in $p+q$ trials is $Ea^p b^q$ if E be the coefficient of the term in which occurs $a^p b^q$ when the binomial $(a+b)^{p+q}$ is expanded.

SECTION II

Postulate

1. I suppose the square table or plane *ABCD* to be so made and levelled, that if either of the balls *o* or *W* be thrown upon it, there shall be the same probability that it rests upon any one equal part of the plane as another, and that it must necessarily rest somewhere upon it.

2. I suppose that the ball *W* shall be first thrown, and through the point where it rests a line *os* shall be drawn parallel to *AD*, and meeting *CD* and *AB* in *s* and *o*; and that afterwards the ball *O* shall be thrown $p+q$ or n times, and that its resting between *AD* and *os* after a single throw be called the happening of the event *M* in a single trial. These things supposed:



Lemma 1

The probability that the point *o* will fall between any two points in the line *AB* is the ratio of the distance between the two points to the whole line *AB*.

Let any two points be named, as *f* and *b* in the line *AB*, and through them parallel to *AD* draw *fF*, *bL* meeting *CD* in *F* and *L*. Then if the rectangles *Cf*, *Fb*, *LA* are commensurable to each other, they may each be divided into the same equal parts, which being done, and the ball *W* thrown, the probability it will rest somewhere upon any number of these equal parts will be the sum of the probabilities it has to rest upon each one of them, because its resting upon any different parts of the plane AC are so many inconsistent events; and this sum, because the probability it should rest upon any one equal part as another is the same, is the probability it should rest upon any one equal part multiplied by the number of parts. Consequently, the probability there is that the ball *W* should rest somewhere upon *Fb* is the probability it has to rest upon one equal part multiplied by the number of equal parts in *Fb*; and the probability it rests somewhere upon *Cf* or LA, i.e. that it does not rest upon *FB* (because it must rest somewhere upon AC) is the probability it rests upon one equal part multiplied by the number of equal parts in *Cf*, *LA* taken together. Wherefore, the probability it rests upon *Fb* is to the probability it does not as the number of equal parts in *Fb* is to the number of equal parts in *Cf*, *LA* together, or as *Fb* to *Cf*, *LA* together, or as *fb* to *Bf*, *Ab* together. Wherefore the probability it rests upon *Fb* is to the probability it does not as *fb* to *Bf*, *Ab* together. And (*componendo inverse*) the probability it rests upon *Fb* is to the probability it rests upon *Fb* added to the probability it does not, as *fb* to AB, or as the ratio of *fb* to AB to the ratio of AB to AB. But the probability of any event added to the probability of its failure is the ratio of equality; wherefore, the probability it rests upon *FB*bis to the ratio of equality as the ratio of *fb* to AB to the ratio of AB to AB, or the ratio of equality; and therefore the probability it rests upon *Fb* is the ratio of *fb* to AB. But *ex hypothesi* according as the ball *W* falls upon *Fb* or not the point *o* will lie between *f* and *b* or not, and therefore the probability the point *o* will lie between *f* and *b* is the ratio of *fb* to AB.

Again; if the rectangles *Cf*, *Fb*, *LA* are not commensurable, yet the last mentioned probability can be neither greater nor less than the ratio of *fb* to AB; for, if it be less, let it be the ratio of *fc* to AB, and upon the line *fb* take the points *p* and *t*, so that *pt* shall be greater than *fc*, and the three lines *Bp*, *pt*, *tA* commensurable (which it is evident may be always done by dividing AB into equal parts less than half *cb*, and taking *p* and *t* the nearest points of division of *f* and *c* that lie upon *fb*). Then because *Bp*, *pt*, *tA* are commensurable, so are the rectangles *Cp*, *Dt*, and that upon *pt* completing the square AB. Wherefore, by what has been said, the probability that the point *o* will lie between *p* and *t* is the ratio of *pt* to AB. But if it lies between *p* and *t* it must lie between *f* and *b*. Wherefore, the probability it should lie between *f* and *b* cannot be less than the ratio of *pt* to AB, and therefore must be greater than the ratio of *fc* to AB (since *pt* is greater than *fc*). And after the same manner you may prove that the forementioned probability cannot be greater than the ratio of *fb* to AB, it must therefore be the same.

Lemma 2

The ball W having been thrown, and the line os drawn, the probability of the event M in a single trial is the ratio of Ao to AB.

For, in the same manner as in the foregoing lemma, the probability that the ball o being thrown shall rest somewhere upon Do or between AD and so is the ratio of Ao to AB. But the resting of the ball o between AD and so after a single throw is the happening of the event M in a single trial. Wherefore the lemma is manifest.

Proposition 8

If upon BA you erect the figure $BghikmA$ whose property is this, that (the base BA being divided into any two parts, as Ab, and Bb and at the point of division b a perpendicular being erected and terminated by the figure in m; and y, x, r representing respectively the ratio of bm, Ab, and Bb to AB, and E being the coefficient of the term in which occurs $a^p b^q$ when the binomial $(a+b)^{p+q}$ is expanded) $y=Ex^p r^q$. I say that before the ball W is thrown, the probability the point o should fall between f and b, any two points named in the line AB, and with all that the event M should happen p times and fail q in $p+q$ trials, is the ratio of $fghikmb$, the part of the figure $BghikmA$ intercepted between the perpendiculars fg, bm raised upon the line AB, to CA the square upon AB.

Corollary

Before the ball W is thrown the probability that the point o will lie somewhere between A and B, or somewhere upon the line AB, and withal that the event M will happen p times, and fail q in $p+q$ trials is the ratio of the whole figure AiB to CA. But it is certain that the point o will lie somewhere upon AB. Wherefore, before the ball W is thrown the probaiblity the event M will happen p times and fail q in $p+q$ trials is the ratio of AiB to CA.

Proposition 9

If before anything is discovered concerning the place of the point o, it should appear that the event M had happened p times and failed q in $p+q$ trials, and from hence I guess that the point o lies between any two points in the line AB, as f and b, and consequently that the probability of the event M in a single trial was somewhere between the ratio of Ab to AB and that of Af to AB; the probability I am in the right is the ratio of that part of the figure

AiB described as before which is intercepted between perpendiculars erected upon AB at the points f and b, to the whole figure AiB.

Corollary

The same things supposed, if I guess that the probability of the event M lies somewhere between 0 and the ratio of Ab to AB, my chance to be in the right is the ratio of Abm to AiB.

Scholium

From the preceding proposition it is plain, that in the case of such an event as I there call M, from the number of times it happens and fails in a certain number of trials, without knowing anything more concerning it, one may give a guess whereabouts it's probability is, and, by the usual methods computing the magnitudes of the areas there mentioned, see the chance that the guess is right. And that the same rule is the proper one to be used in the case of an event concerning the probability of which we absolutely know nothing antecedently to any trials made concerning it, seems to appear from the following consideration; viz. that concerning such an event I have no reason to think that, in a certain number of trials, it should rather happen any one possible number of times than another. For, on this account, I may justly reason concerning it as if its probability had been at first unfixed, and then determined in such a manner as to give me no reason to think that, in a certain number of trials, it should rather happen any one possible number of times than another. But this is exactly the case of the event M. For before the ball W is thrown, which determines it's probability in a single trial (by corollary to proposition 8), the probability it has to happen p times and fail q in $p+q$ or n trials is the ratio of AiB to CA, which ratio is the same when $p+q$ or n is given, whatever number p is; as will appear by computing the magnitude of AiB by the method of fluxions. And consequently before the place of the point o is discovered or the number of times the event M has happened in n trials, I can have no reason to think it should rather happen one possible number of times than another.

In what follows therefore I shall take for granted that the rule given concerning the event M in proposition 9 is also the rule to be used in relation to any event concerning the probability of which nothing at all is known antecedently to any trials made or observed concerning it.