Volume : 4 | Issue : 4 | April 2015 • ISSN No 2277 - 8179

Research Paper

Some Interesting Case Studies Using Bayes Theorem

Statistics

KEYWORDS : Statistics, Bayes theorem, Case Studies

Sahana Prasad Associate Professor, Department of Statistics, Christ University, Bangalore, India

ABSTRACT

Statistics is basically a mathematical science wherein data is first collected; then is analysed, an explanation/ interpretation is given and finally is converted into a presentation. The data quality is improvised by statisticians and is important to know what information we should pay attention to, and what to do with these and how to use them. Knowledge leads to good decision-making and spurs progress.

More and more researchers are using Bayesian methods as these are best suited to subjective qualities of research. In this approach, the focus is on the probability of the hypothesis, given the data i.e. P (H|D) .The data is considered as fixed and hypotheses as random. This theorem has been used extensively to calculate prior probabilities. In this approach, newly acquired data is combined with prior data to predict an outcome. This paper deals with some interesting case studies which have used Bayes' theorem.

INTRODUCTION:

Bayes' theorem uses the available information and incorporates "conditional probabilities" into conclusions. This was formulated by Reverend Thomas Bayes and he was elected a Fellow of the Royal Society of London which is the most prestigious scientific body of its time in 1742, though he had not published any scientific or mathematical works. It is of great interest that Bayes' Theorem appears in a Bayes manuscript presented to the Royal Society of London in 1764, *three years after Bayes's supposed death in 1761!* Mathematician Pierre-Simon Laplace is credited with popularising this theorem. This theorem can be considered as a fundamental result of probability theory which helps us to calculate "conditional probabilities", which are those probabilities that reveal the influence of one event on the probability of another.

This quote from an economics article is very interesting:

"The essence of the Bayesian approach is to provide a mathematical rule explaining how you should change your existing beliefs in the light of new evidence. In other words, it allows scientists to combine new data with their existing knowledge or expertise. The canonical example is to imagine that a precocious newborn observes his first sunset, and wonders whether the sun will rise again or not. He assigns equal prior probabilities to both possible outcomes, and represents this by placing one white and one black marble into a bag. The following day, when the sun rises, the child places another white marble in the bag. The probability that a marble plucked randomly from the bag will be white (i.e., the child's degree of belief in future sunrises) has thus gone from a half to two-thirds. After sunrise the next day, the child adds another white marble, and the probability (and thus the degree of belief) goes from two-thirds to three-quarters. And so on. Gradually, the initial belief that the sun is just as likely as not to rise each morning is modified to become a near-certainty that the sun will always rise." - http://www.cs.ubc.ca/~murphyk/Bayes/economist.html

If $E_1, E_2, E_3, \ldots, E_n$ are mutually disjoint events with $P(E_i) \neq 0$, $(i = 1, 2, \ldots, n)$, then for any arbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ such that P (A)>0, we have

$$P(E_i \mid A) = \frac{P(E_i)P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i)P(A \mid E_i)} = \frac{P(E_i)P(A \mid E_i)}{P(A)}; \; i = 1, 2, \ldots, n$$

When to Use Bayes theorem:

(i) The sample space is partitioned into a set of mutually exclusive events {A1, A_2, . . . , A_n }.

(ii) Within the sample space, there exists an event B, for which P (B) > 0.

(iii) If there is a prior knowledge of a situation and it is required to predict probabilities of any outcome.

Probability of God?

Richard Dawkins, a professor of University of Oxford, argues in his book *The God Delusion,* against the use of Bayes's theorem for assigning a probability to God's existence. Using Bayes's theorem, if we calculate the probability of God given our experiences in the world (the existence of evil, religious experiences, etc.) and assign numbers to the likelihood of these facts given existence or nonexistence of God, as well as to the prior belief of God's existence i.e. the probability we would assign to the existence of God if we had no data from our experiences. Dawkins's argument is with the lack of data to put into this formula by those employing it to argue for the existence of God. The equation is perfectly accurate, but the numbers inserted are, to quote Dawkins, "not measured quantities but & personal judgments, turned into numbers for the sake of the exercise."

Case Study 1: *Suppose that a woman in her forties goes for a mammogram and receives bad news: a "positive" mammogram. However, since not every positive result is real, what is the probability that she actually has breast cancer? Given that the fraction of women in their forties who have breast cancer is 0.014 and the probability that a woman who has breast cancer will get a positive result on a mammogram is 0.75. The probability that a woman who does not have breast cancer will get a false positive on a mammogram is 0.1* Here, the suitable tool is Bayes theorem as we have prior knowledge of all possible probabilities.

The various factors involved here are:

1. The fraction of women in their forties who have breast cancer is 0.014, which is about one in seventy. The fraction who do not have breast cancer is therefore 1 - 0.014 = 0.986. These fractions are known as the prior probabilities.
2. The probability that a woman who has breast cancer will get a positive result on a mammogram is 0.75. The probability that a woman who does not have breast cancer will get a false positive on a mammogram is 0.1. These are known as the conditional probabilities.
3. Applying Bayes's theorem, we can conclude that, among women who get a positive result, the fraction who actually have breast cancer is (0.014 x 0.75) / ((0.014 x 0.75) + (0.986 x 0.1)) = 0.1, approximately. That is, once we have seen the test result, the chance is about ninety per cent that it is a false positive.

Case Study 2 : One famous case of a failure to apply Bayes' Theorem involves a British woman, Sally Clark. After two of her children died of sudden infant death syndrome (SIDS), she was arrested and charged with murdering her children. Paediatrician Roy Meadow testified that the chances that both children died of SIDS were 1 in 73 million. He got this number by squaring the odds of one child dying of SIDS in similar circumstances (1 in 8500).

Volume : 4 | Issue : 4 | April 2015 • ISSN No 2277 - 8179

Because of this testimony, Sally Clark was convicted. The Royal Statistics Society issued a public statement decrying this "misuse of statistics in court," but Sally's first appeal was rejected. She was released after nearly 4 years in a woman's prison where everyone else thought she had murdered her own children. She never recovered from her experience, developed an alcohol dependency, and died of alcohol poisoning in 2007.

The statistical error made by Roy Meadow was, among other things, to fail to consider the *prior probability* that Sally Clark had murdered her children. While two sudden infant deaths may be rare, a mother murdering her two children is even rarer.

Here, the suitable tool is Bayes theorem as we need to find the *conditional probabilities* of the various possible causes of death, *given* the fact that the children died.

If H is some hypothesis, for example, that both of Sally Clark's children died of cot death - and D is some data, that both children are dead we want to find the probability of the hypothesis given the data, which is written as P(H/ D).

Let A be for the alternate hypothesis - that the children did not die of cot death.

Discounting all other possibilities, for example that someone else murdered both children, or that Sally Clark murdered only one of them, or that they died of natural causes other than cot death.

P (D/H) = 1 as it is the probability that two of the children are dead, given that that two of the children have died of natural causes.

P (H) = 1/100,000. (1/73 million can be approximated to this figure)

P (D/A) is the probability that the children died given that they did not die of natural causes. In other words, it is the probability that a randomly chosen pair of siblings will both be murdered. This is the most difficult figure to estimate. Statistics on such double murders are difficult to get, because child murders are so rare (far, far more rare than cot deaths) and because in most cases, someone known to have murdered once is not free to murder again. So we take the Home Office statistic that fewer than 30 children are known to be murdered by their mother each year in England and Wales. Since 650,000 are born each year, and murders of pairs of siblings are clearly rarer than single murders, we should use a figure much smaller than 30/650,000=0.000046.

So, we can take a number ten times as small here i.e. 0.000046*10= 0.0000046. Thus,

P(H/D) > 2/3

This is the probability that Sally Clark is innocent. Thus, Bayes theorem gives a correct picture.

Case Study 3: We want to know how a change in interest rates would affect the value of a stock market index. All major stock market indexes have a plethora of historical data available so you should have no problem finding the outcomes for these events with a little bit of research. For our example we will use the data below to find out how a stock market index will react to a rise in interest rates.

(Adapted from**:** Article from http://www.investopedia.com)

		Interest Rates		
		Decline	Increase	Unit Frequency
Stock Price	Decline	200	950	1150
	Increase	800	50	850
		1000	1000	2000

Here, the suitable tool is Bayes theorem as we have prior probabilities and we need to update it with new information.

P (SI) = the probability of the stock index increasing=
P (SD) = the probability of the stock index decreasing
P (ID) = the probability of interest rates decreasing
P (II) = the probability of interest rates increasing

P (SD) = 1150/2000 (prior probability)

What we find after applying Bayes theorem is posterior probability, P (SD/II) =0.9499~95%

Thus, we can use the outcomes of historical data to base our beliefs on from which we can derive new updated probabilities.

Results and Discussions:

As with any tool, Bayes' theorem also has to be used with caution. If there is authentic prior information, it can be used but one needs to exercise caution if that is not the case. Professor Bradley Efron of Stanford University, who also works as an editor of a statistics journal, states that he found around 25 percent of papers used Bayes' theorem and most were based on uninformative priors. There have been many examples of Bayes theorem being used in clinical arena to demonstrate how, even the most accurate tests can less than acceptable positive predictive values if the prevalence of the disease being tested is low.

REFERENCE

1) B Y Brase, C. H. (n.d.). Understanding basic statistics . | 2) Chawla, D. N. (n.d.). Research Methodology- concepts and cases. | 3) Healey, J. (n.d.). Statistics- A tool for Social Research . Newport University. | 4) http://sxxz.blogspot.in/2005/03/bayes-theorem-and-why-it-matters.html | 5) + magazine, URL: http://plus.maths.org/content/os/ issue21/features/clark/index | 6) The New Yorker, Jan 25, 2013 | 7) http://www.investopedia.com | 8) Bayes' Theorem in the 21st Century , Bradley Efron, Science 7 June 2013, Vol. 340 no. 6137 pp. 1177-1178 , DOI: 10.1126/science.1236536