BAYES THEOREM AND ITS RECENT APPLICATIONS

Nicholas Stylianides

Eleni Kontou

March 2020

Abstract

Individuals who have studied maths to a specific level have come across Bayes' Theorem or Bayes Formula. Bayes' Theorem has many applications in areas such as mathematics, medicine, finance, marketing, engineering and many other. This paper covers Bayes' Theorem at a basic level and explores how the formula was derived. We also, look at some extended forms of the formula and give an explicit example. Lastly, we discuss recent applications of Bayes' Theorem in valuating depression and predicting water quality conditions. The goal of this expository paper is to break down the interesting topic of Bayes' Theorem to a basic level which can be easily understood by a wide audience.

What is Bayes Theorem?

Thomas Bayes was an 18th-century British mathematician who developed a mathematical formula to calculate conditional probability in order to provide a way to re-examine current expectations. This mathematical formula is well known as Bayes Theorem or Bayes' rule or Bayes' Law. It is also the basis of a whole field of statistics known as Bayesian Statistics. For example, in finance, Bayes' Theorem can be utilized to rate the risk of loaning cash to possible borrowers. The formula for Bayes' Theorem is: [1]

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) * P(B|A)}{P(B)}$$

P(A) =The probability of A occurring

P(B) = The probability of B occurring

P(A|B) = The probability of A given B

P(B|A) = The probability of B given A

P(A∩B) = The probability of both A and B occurring

Bayes' Theorem:

$Let\ D = \{A_1, \dots, A_n\}\ be\ a\ complete\ set\ of\ disjoint\ events, and\ B\ be\ an\ event. Then$

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}.$$

,

$\$\%\quad \&((!)\ >\ 0\quad \%,\text{-}\quad .//\quad 1 \le i \le n\ and\ P(B) > 0.$

Proof:

From Bayes' Formula: $P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}$ (1)

From the Total Probability Theorem: $P(B) = \sum_{j=1}^{n} P(B|A_j)P(A_j)$ (2)

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$$

Substituting equation (2) into equation (1) we obtain . □

Extended Form of Bayes' Formula

For some partition $\{(^{"}\}$ of the sample space, the event space is given in terms of $\&((^{"})$ and

$\&(5|(^{"})$. It is then suitable to compute P (B) using the law of total probability $\&(5) = \sum_j^n P(B|A_j)P(A_j), \Rightarrow P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$

.

In the special situation where A is a binary variable (meaning, a unit can take on only two possible states, traditionally labelled as 0 and 1 accordingly): [2]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}$$

Example

Boy or Girl Paradox

One great example using the Bayes' Formula is the Boy or Girl Paradox. The Boy or Girl Paradox also known as the "Two Child Problem" covers a lot of questions in probability theory. The original formation of the question goes back to 1959 when Martin Gardner included it in his October 1959 "Numerical Game Section" in Scientific American. He titled it "The Two Child Problem" and stated the paradox as follows:

- "Mr. Jones has two children. The older child is a girl. What is the probability that both children are girls?"
- "Mr. Smith has two children. At least one of them is a boy. What is the probability that both children are boys?"[3] *Is It Logical?*

The sample space is **Ω**:= {(Boy, Boy), (Boy, Girl), (Girl, Boy), (Girl, Girl)}, where the first component of each pair signifies the sex of the first child, and the second component indicates the sex of the second child. Let us assume that each pair is equally likely, which is sensible assumption. On the off chance that you reveal that one of the children is a boy, then the sample space is decreased to **Ω'**:= {(Boy, Boy), (Boy, Girl), (Girl, Boy)}. Since it is sensible to accept that every one of the three sets are equally likely, the probability that the other child is a girl is given by:

P ({Boy, Girl), (Girl, Boy)}) = P ({(Boy, Girl)}) + P ({(Girl, Boy)}) = $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

Using Bayes' Formula, we can derive the same result:

$$P(Girl|Boy) = \frac{P(Girl \cap Boy)}{P(Boy)} = \frac{P((Girl, Boy)\ or\ (Boy, Girl))}{P((Boy, Boy) or\ (Boy, Girl) or\ (Gril, Boy))} = \frac{2/4}{3/4} = \frac{2}{3}$$

Recent Applications

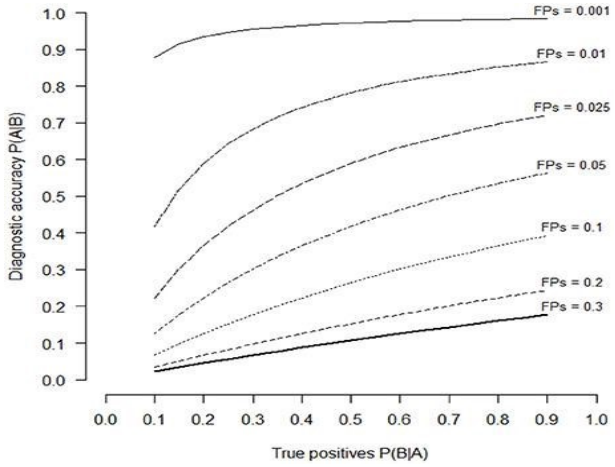
Application of Bayes' Theorem in Valuating Depression Tests Performance:

In this application, Bayes' Theorem is used to assess the subsequent probability that a person has depression or whether they do not have it. It uses, as a basis the prior probability of information about the diffusion of this pathology as well as any information of the sensitivity and specificity values of the scores of psychological tests taken by this person.

Let us assume that a psychological test is used to diagnose the presence of depression in individuals. In this case the probability of a person really being depressed, means her/his score exceeds the cut-off value and is not equivalent to the probability to get scores greater than the cut-of when she/he is really depressed. When carrying out a diagnosis to evaluate the actual risk of failure, one must consider the conditional probability of an individual being depressed when she/he exceeds the cut-off value, and the conditional probability of exceeding the cut-off when she/he is depressed. So if A is the probability of an individual being depressed and B the probability to exceeding the cut-off, then using Bayes' theorem: $P(A|B) = [P(B|A)P(A)] / P(B)$ (Equation 1)

P(B|A), is also known as the conditional probability of surpassing the cut-off when the individual is really depressed whereas P(A) relates to the percentage of people who are actually depressed in the population and P(B) is the probability of getting a test score greater than the cut-off.

That is to say, if the percentage of depressed people in the USA population is 6.7% (source: https://www.nimh.nih.gov/health/statistics/major-depression.shtml) then P(A) = 0.067. To estimate P(B), the conditional probability that an individual surpasses the cut-off, (if he/she is actually depressed) can be estimated from the test scores. Assuming the probability that a person surpasses the cut off, P(B|A)=0.8 and P(B)=0.10 (taking into account that P(A)=0.067), then the conditional probability that an individual is really depressed if he/she overcomes the cut-off is: [P(B|A)P(A)]/P(B)= (0.8*0.067)/0.10=0.54. That is the probability a person is really depressed is only 54%.



The adjacent graph shows the diagnostic accuracy P(A|B) versus the true positives P (B|A) of various selected levels of FPs. This is estimated using the Bayes' Theorem equation 1

Summing up, the example above indicates the use of Bayes' Theorem to diagnose the probability of a person having depression with some accuracy.[4]

Predicting Water Quality Conditions using Bayes' Theorem

Water quality is regularly monitored through the concentration of at least one in situ toxin, (for example, nutrients, microscopic organisms, and natural mixes). Water's suitability for a proposed use, (i. e. drinking water, diversion, or agricultural use) relies upon whether the amounts of the pollutants surpass the water quality standard numeric thresholds. Since these contaminations regularly cannot be estimated straightforwardly, researchers normally measure pointers that fill in as potential substitutes for the pollutant of concern. The correlation between indicator concentration and the amount of the pollutant, it apparently stands for, differs widely depending on the type of the pollutant. For instance, in recreational and shellfish-harvesting waters all over the United States, water properties depends on the amounts of non-pathogenic fecal indicator bacteria (FIB, for example, fecal coliforms and Escherichia coli). These microorganisms form a conservative index of fecal pollution and of the potential existence of hurtful waterborne pathogens, which, while affect human and ecological wellbeing, are additionally considerably more difficult and expensive to calculate. The uncertainty in the relationship between pollutant-indicator, the sampling frequency and other factors collectively can contribute in reliably forecasting environmental conditions. In this case Bayesian approach can be used to evaluate water quality.

In the same way as other different pollutants, FIB concentrations are normally accepted to pursue a lognormal LN (μ, σ) probability distribution with a mean (μ) and a standard deviation (σ). Although this simple probability model recognizes natural spatial and temporal variability in FIB dispersion patterns, it often neglects to clearly recognize other, more not obvious sources of variability from inherent sources coming from FIB concentration estimations and how FIB amounts are determined. These can direct not only to unsureness in FIB concentration expectations, but to unsureness in probability distribution parameters (μ and σ) as well. Using the Bayesian approach, these uncertainties can explicitly be recognized by first putting an earlier probability distribution on the parameters μ and σ (which may represent earlier estimates of their potential values), and building a likelihood function for μ and σ from experiential evidence (in this example, using water quality samples), and, at last, developing a joint posterior probability distribution for both. Figure 1 shows a smoothed contour plot of the joint posterior probability density for the fecal coliform log-concentration mean (μ) and standard deviation (σ) for an example site in eastern North Carolina resulting from this method. [5]

© Nicholas Stylianides, Eleni Kontou

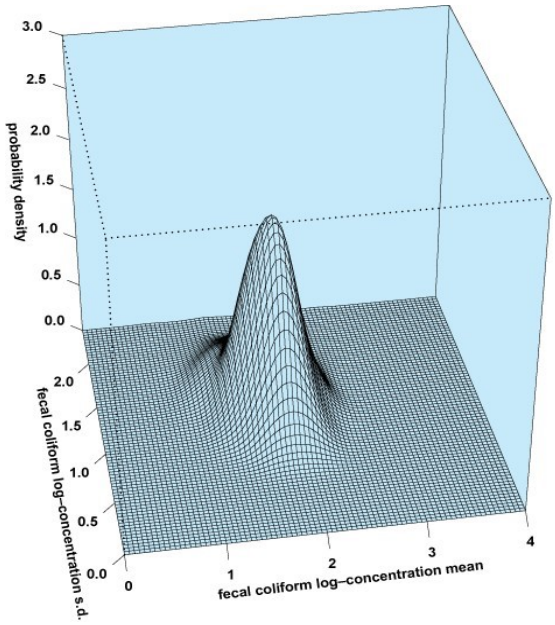
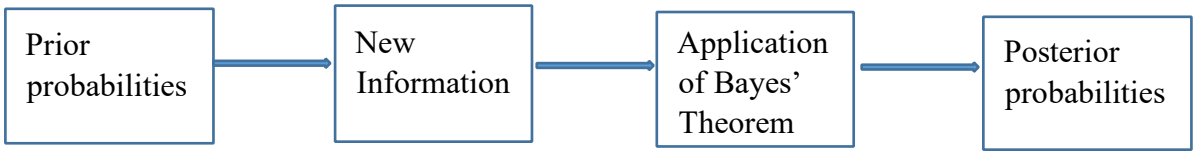


Figure 1 taken from reference 5 shows the use of the Bayes' Theorem in developing posterior probability density functions and graphically shown the result. This is important information in water quality monitoring at a shellfishharvesting area in eastern North Carolina.

Conclusion

Bayes' Theorem provides a method for revising initial or prior probability estimates for specific events of interest taking into account information about the specific events from sources such as a sample, a special report or a product test.



Bayes' theorem does not only apply in mathematics, but it also has many real life applications such as in Internet Marketing to profile visitors to a website, in Decision Analysis and Decision Trees, the "Two Child Problem" explained in the text above. It helped create a test for valuating depression more accurately than before as well as a range of other mental tests. Furthermore, it can be helpful to monitor more accurately water quality properties using the Bayesian theory and then announcing the results to the interested parties.

It has though certain limitations. It is assumed that the event under investigation and its complement are mutually exclusive and their union is the entire sample space.

Even though it was first discovered in the 18th century by Thomas Bayes, we can see that it helps humanity even now in many areas.

© Nicholas Stylianides, Eleni Kontou

References

[1] Hayes, A. (2019). "*Bayes' Theorem Definition".* Available from: https://www.investopedia.com/terms/b/bayes-theorem.asp

[2] En Wikipedia. (2020). "*Bayes' Theorem".* Available from: https://en.wikipedia.org/wiki/Bayes%27_theorem

[3] En Wikipedia. (2020). "*Boy or Girl paradox".* Available from: https://en.wikipedia.org/wiki/Boy_or_Girl_paradox

[4] Tommasi, M., Ferrara, G., Saggino, A. (2018). "*Applications of Bayes' Theorem in Valuating Depression Tests Performance".* Available from: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01240/full

[5] Gronewold, A., Vallero, D. (2010). "*Applications of Bayes' theorem for predicting environmental damage".* Available from: https://www.accessscience.com/content/applications-of-bayes-theorem-for-predictingenvironmental-damage/YB100249

© Nicholas Stylianides, Eleni Kontou