Journal of Education and Culture Studies
ISSN 2573-0401 (Print) ISSN 2573-041X (Online)
Vol. 7, No. 3, 2023
www.scholink.org/ojs/index.php/jecs

Original Paper

Bayesian Formula and Its Application Analysis

Yuan Zhang[1*], Jia Liu[1] & Saisai Shi[1]

[1] Institute of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, 233030, P. R. China

[*] Corresponding author, Yuan Zhang, E-mail: 120210066@aufe.edu.cn

Received: August 6, 2023 Accepted: August 14, 2023 Online Published: August 16, 2023
doi:10.22158/jecs.v7n3p70 URL: http://dx.doi.org/10.22158/jecs.v7n3p70

Abstract

Bayesian formula is a very important and widely used formula in probability and statistics courses, which can simplify the probability calculation of certain complex events. This article not only introduces the relevant theories of Bayesian formula, but also explores its practical applications in daily life through examples.

Keywords

Complete event group, Bayesian formula

1. Introduction

With the development of society, the applications of probability theory and mathematical statistics in daily life is becoming increasingly widespread. Bayesian formula is an important tool in probability theory, which often used to solve complex conditional probability problems. It can be used to update our probability estimates of certain events based on observed evidence and given a prior probability. This article mainly explores the significance of Bayesian formula and its application in diagnosing diseases, testing product, and optimize decision. At the same time, specific analysis methods are introduced through practical problems, reflecting the basic ideas of Bayesian formula (Xie, 2002; Yang & Hu, 2015).

2. Bayesian Formula

Definition 2.1 (Complete event group)

If event group $B_1, B_2, ..., B_n$ satisfies the following two conditions:

(1) $B_1, B_2, ..., B_n$ are incompatible in pairs;

(2) $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$.

Then call these n events $B_1, B_2, ..., B_n$ is a partition of the sample space Ω, it also known as a complete set of events.

Theorem 2.2 (Total probability formula)

Assume that Ω is a sample space, $B_1, B_2, ..., B_n$ is a partition of Ω and A is an event. If $P(A) > 0, P(B_i) > 0, i = 1,2,..., n$, one has

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i).$$

Proof: Since

$$\begin{aligned} A &= A \bigcap \Omega \\ &= A \bigcap (B_1 \bigcup B_2 \bigcup \cdots \bigcup B_n) \\ &= AB_1 \bigcup AB_2 \bigcup \cdots \bigcup AB_n, \end{aligned}$$

by simultaneously calculating the probability on both sides of the above equation, we can obtain:

$$P(A) = P(AB_1 \bigcup AB_2 \bigcup \cdots \bigcup AB_n).$$

Notice that $AB_1, AB_2, \ldots, AB_n$ are incompatible in pairs, according to the finite additivity of probability:

$$\begin{aligned} P(A) &= P(AB_1 \bigcup AB_2 \bigcup \cdots \bigcup AB_n) \\ &= P(AB_1) + P(AB_2) + \cdots + P(AB_n), \end{aligned}$$

by reusing the multiplication formula, one has:

$$\begin{aligned} P(A) &= P(AB_1) + P(AB_2) + \cdots + P(AB_n) \\ &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_n)P(A|B_n) \\ &= \sum_{i=1}^{n} P(B_i)P(A|B_i). \end{aligned}$$

The total probability formula simplifies the probability of a complex event to calculating the sum of probabilities of several independent events, and the key is to find an appropriate complete set of events.

Theorem 2.3 (Bayesian formula)

Assume that Ω is a sample space, $B_1, B_2, ..., B_n$ is a partition of Ω and A is an event. If $P(A) > 0, P(B_i) > 0, i = 1,2,..., n$, one has

$$P(B_k|A) = \frac{P(AB_k)}{P(A)} = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}, k = 1,2,..., n.$$

In daily life, we often use Bayesian formula to repeatedly estimate the probability of certain events to calculate a new posterior probability, which can improve the accuracy of the obtained probability. The basic idea of Bayesian formula is to find the cause when the results are known. For example, the rise and fall of financial products may be influenced by factors such as policies, funds and the environment. If a certain financial product falls, it is necessary to analyze which factors are most likely to cause similar problems. The application of Bayesian formula is very extensive. Next we will use specific examples to illustrate its application in practical problems (Feng & Han, 2022; Li, 2019; Zhao, 2023).

3. Application Analysis of Bayesian Formula

3.1 Diagnose Disease

Example 1. A blood test is used to identify whether a certain disease is present. It is known that 95% of people with this disease tested positive, while 99% of healthy individuals tested negative. Historical data show that the incidence rate of this disease in a certain area is 0.5%. If a person's test result is positive, what is the probability that the person really suffers from this disease?

Solution: Assume that

A ={Suffering from this disease},

$\overline{A}$ ={Not suffering from this disease},

B ={The test reaction is positive}.

From the question to know

$P(A)=0.005$, $P(\overline{A})=0.995$, $P(\overline{B} \mid \overline{A})=0.99$,

$P(B \mid \overline{A})=1-P(\overline{B} \mid \overline{A})=1-0.99=0.01$.

From Bayesian formula, we can obtain the following result:

$$P(A|B)=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}=0.323.$$

3.2 Testing Product

Example 2. Assume that a factory has three workshops *A*, *B* and *C* producing the same product, with production accounting for 45%, 35% and 20% of the entire factory respectively. And the defect rates of each workshop are 4%, 2% and 5% respectively. Now that one defective product has been detected from a batch of products, it is most likely to be produced in which workshop?

Solution: Assume that

A_1 ={The product comes from workshop *A*},

A_2 ={The product comes from workshop *B*},

A_3 ={The product comes from workshop *C*},

E ={The product is defective}.

It can be inferred from the question, A_1 , A_2 , A_3 is a partition of the simple space and

$$P(A_1)=0.45,\quad P(A_2)=0.35,\quad P(A_3)=0.2,$$

$$P(E|A_1)=0.04,\quad P(E|A_2)=0.02,\quad P(E|A_3)=0.05.$$

From Bayesian formula, we conclude that:

$$P(E|A_1)=\frac{P(A_1)P(E|A_1)}{P(A_1)P(E|A_1)+P(A_2)P(E|A_2)+P(A_3)P(E|A_3)}=0.514,$$

$$P(E|A_2)=\frac{P(A_2)P(E|A_2)}{P(A_1)P(E|A_1)+P(A_2)P(E|A_2)+P(A_3)P(E|A_3)}=0.200,$$

$$P(E|A_3)=\frac{P(A_3)P(E|A_3)}{P(A_1)P(E|A_1)+P(A_2)P(E|A_2)+P(A_3)P(E|A_3)}=0.286.$$

It can be seen that the possibility of this defective product being produced by workshop *A* is the highest.

3.3 Optimize Decision

Example 3. A student's bank card was lost, and the location of the lost bank card was locked in the dormitory, cafeteria and classroom. Assuming that the probability of a bank card falling into the dormitory, cafeteria and classroom is 10%, 30% and 60% respectively, and the probability of being found in these three places is 20%, 30% and 40% respectively. Where should the student start searching first?

Solution: Assume that

B_1 ={Bank card dropped in the dormitory},

B_2 ={Bank card dropped in the cafeteria},

B_3 ={Bank card dropped in the classroom},

C ={Bank card found}.

It can be inferred from the question, B_1, B_2, B_3 is a partition of the simple space and

$$P(B_1)=0.1,\quad P(B_2)=0.3,\quad P(B_3)=0.6,$$

$$P(C|B_1)=0.2,\quad P(C|B_2)=0.3,\quad P(C|B_3)=0.4.$$

From Bayesian formula, one has

$$P(C|B_1)=\frac{P(B_1)P(C|B_1)}{P(B_1)P(C|B_1)+P(B_2)P(C|B_2)+P(B_3)P(C|B_3)}=0.058,$$

$$P(C|B_2)=\frac{P(B_2)P(C|B_2)}{P(B_1)P(C|B_1)+P(B_2)P(C|B_2)+P(B_3)P(C|B_3)}=0.257,$$

$$P(C|B_3)=\frac{P(B_3)P(C|B_3)}{P(B_1)P(C|B_1)+P(B_2)P(C|B_2)+P(B_3)P(C|B_3)}=0.685.$$

Therefore, the student can first go to the classroom to search.

4. Conclusion

Bayesian formula can not only determine the most likely cause of an event, but also be applied to many fields, such as machine learning, artificial intelligence, and information retrieval. This article mainly uses specific examples to understand the application of Bayesian formula in practical life, stimulate learning interest and achieve the goal of deeply mastering the essence of Bayesian formula.

References

Feng, G. Q., & Han, C. Y. (2022). Exploration of the Application of Total Probability Formula and Bayesian Formula. *Journal of Jiaozuo Normal College*, *38*(03), 73-76. (in Chinese)

Li, Y. Q. (2019). Application of Bayesian Formula. *Science and Education Wenhui*, *2019*(31), 75-76. (in Chinese) https://doi.org/10.15358/0935-0381-2019-5-75

Xie, G. R. (2002). *Probability Theory and Mathematical Statistics*. Beijing: Higher Education Press. (in Chinese)

Yang, J. Q., & Hu, X. P. (2015). *Probability Theory and Mathematical Statistics*. Shanghai: Shanghai Jiao Tong University Press. (in Chinese)

Zhao, M. R. (2023). About Bayesian formula and its practical application. *Scientific and technological innovation*, *2023*(18), 67-70. (in Chinese)