Orchid Academia Siraha
Vol. 3, Issue 1, Dec. 2024

ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

An Axiomatic Approach to Prove the Converse of Bayes' Theorem in Probability

Aanand Kumar Yadav

Department of Math Education, Tribhuvan University, Suryanarayan Satyanarayan Morbaita Yadav Multiple Campus, Siraha, Nepal
Email: aanandkumaryadav365@gmail.com

https://doi.org/10.3126/oas.v3i1.78106

Abstract

The converse of Bayes' theorem, have been proved using axiomatic approach to probability. This approach to probability utilizes the relations and theorems of set theory. Simply presentation of the converse of Bayes' theorem has been possible due to the correspondence theorem in set theory. This theorem is seen to be more applicable in the proof of Bayes' theorem and its converse. So at first, the correspondence theorem of set theory with its proof has been presented here and then has been applied to prove the Bayes' theorem and its converse. Some important applications of correspondence theorem for set theory have also been presented in this article. The use of correspondence theorem in proving Bayes' theorem and its converse makes the proof easily understandable and reduces the steps proving them. After this work, the statement of Baye's theorem can be stated with conditions necessary and sufficient both.

***Keywords*:** Correspondence theorem, family of sets, prior and posterior probability, Bayes' theorem, axiomatic approach

Introduction

This article majorly concerns with presenting two things: the correspondence theorem of set theorem and the converse of Bayes' theorem. Bayes' theorem in probability is based on the concept of conditional probability that uses dependent events. Conditional probability of any event, say A, defined on the given event B (known as the Hypothesis of the conditional probability), denoted by $P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$. This probability is also called the probability of event A conditional on event B. Therefore, $P(A \cap B) or\ P(AB) = P(B)P\left(\frac{A}{B}\right)$ is called the theorem of compound probabilities. A conditional probability is also known as the probability for prediction and inference.

There are so many books and literatures that describes only about Bayes' theorem but there can be hardly find any of these books mentioned about the converse of this theorem. Therefore an effort has been done towards the converse of Bayes' theorem by this work.

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

Review of Literatures

Correspondence theorem has been presented as providing a foundation in the proof of Bayes' theorem and its converse. So the review of some literatures related to both these concepts has been presented here.

The presentation of the theory of probability by defining a system of axioms, which was developed by a Russian mathematician A.N. Kolmogorov in 1933, is known as Axiomatic approach to probability (https://www.ck12.org/book/cbse_maths_book_class_11/section/17.4/). It has presented a consistent system of five axioms to define probability of an event. A system of sets together with a definite assignment of numbers (probability of an event, i.e. P (A), satisfying all these five axioms, was called a field of probability (Kolmogorov, 1950, pp. 1-8). Bayes' theorem, which was published posthumously in 1763, is the foundation of Bayesian inference. Bayes' theorem presents the relation between two probabilities that are the reverse of each other. Bayes' theorem, which is also referred to as Bayes' law or Bayes' rule, was named after reverend Thomas Bayes' (1702-1761).

Ash, Robert B. (2008, pp. 4-7, 33-38) has mentioned that the algebra of events is similar to the algebra of real numbers, with union corresponding to addition and intersection to multiplication. Further, the theorem of total probability has been mentioned by defining a finite or countable infinite family of mutually exclusive and exhaustive events.

Bayes' theorem expresses the conditional probability, or 'posterior probability' of an event B after A is observed in terms of the 'prior probability' of B, prior probability of A, and the conditional probability of A given B, denoted byA/B. Bayes's theorem is valid in all common interpretations of probability.

Halmos (1974, pp. 34-35) has presented an example addressing an interesting algebraic identities that looks closely similar to correspondence theorem. In one of these examples, he has presented, only as an identity, the distribution of union and intersection operation with a set over intersection of family of sets and union of family of sets respectively. In another example, he has presented, only as an identity, the distribution of the union of one family of sets and the intersection of one family of sets operation over the intersection of another family of sets and the union of another family of sets respectively. But he has nothing mentioned about their proof s and has expressed nothing about their applications.

Maskey (1997, pp. 56-66) has mentioned detailed explanation on set theory and probability theory. He has only mentioned distributive law for sets limited only to the distribution of the union and intersection operations with a set over the intersection of two sets and union of two sets respectively. But he has nothing mentioned about the extension of distributive law over family of sets.

Maskey has mentioned in more detailed explanation on probability theory including Bayes' theorem only but he has explained nothing about its converse. Freiwald (2014, pp. 7-12) has presented the correspondence theorem of set as generalized form of distributive law. But he has given its proof in the form of mathematical analysis which seems more abstract. So there was a need to present it in simple form and which may be easiest in teaching and learning and simply use the concept in others field of mathematics like probability. Bolstad and Curran (2017, pp. 66-75) have expressed Bayes' theorem in probability including the terms prior probability, likelihood of events and posterior probability. They have summarized Bayes' theorem the posterior probability as the prior probability times likelihood divided by the sum of the prior times likelihood. He has also presented Bayes' theorem in its proportional form as $posterior \propto prior \times likelihood$. But he has not mentioned any more about its converse theorem.

Downey (2012, pp. 2-6), has explained Bayes' theorem as *diachronic interpretation* in which "diachronic" means something is happening over time and in this case probability of the hypothesis changes over time as we see new data [10]. While rewriting Bayes' theorem with A and B as $P\left(\frac{A}{B}\right) = \frac{P(A)P\left(\frac{B}{A}\right)}{P(B)}$, then the probability $P(H)$ of the hypothesis before we the data, is called the prior probability or just prior and the probability $P\left(\frac{A}{B}\right)$ that we want to compute, the probability of the hypothesis after we see the data, is called Posterior.

Bertsekas and Tsitsiklis (2000, pp. 25-31) have mentioned that Bayes' rule is often used for inference and there are a number of "causes" that may result in certain "effect". The events $A_1, A_2, A_3, \dots \dots, A_n$ are associated with the causes and the event B represents the event. Thus the probability$P\left(\frac{B}{A_i}\right)$ that the effect will be observed when the cause A_i is present amounts to a probabilistic model of the cause-effect relation. We wish to evaluate the probability $P\left(\frac{A_i}{B}\right)$ that the cause A_i is present. He has mentioned no more about the converse of Bayes' theorem. Spiegel et al. (2013, pp. 7-8), have mentioned that the theorem of Bayes' enables us to find the probabilities of the various events $A_1, A_2, A_3, \dots \dots, A_n$ that can cause A to occur and due to this reason Bayes' theorem is often referred to as a theorem on the probability of causes. But they have not specified anything about its converse part. Pinter (2014) has also mentioned the correspondence theorem of set theory as the generalized distributive law and has given analytical proof different from that is intended to present here in this article.

After reviewing the literatures, it is found that most of the literatures have explained in detailed about the theorems except the correspondence theorem and only a few of them have mentioned it in different form. Only a statement of the correspondence theorem can be found in some of these few literatures. The proof,

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

If any in these literatures contains, has been presented in more complex form. So there is a need to present a proof of this theorem in simple form and easily understandable and applicable to similar context. Most of the literatures on probability theory contain Bayes' theorem but almost there has not been discussed about the converse of Bayes' theorem. Stating of the converse of Bayes' theorem was felt necessary and its proof was found to be possible through axiomatic approach of probability using the correspondence theorem.

Methodology

This study follows the mixed method approach of inductive and deductive research methodologies. This study follows the pattern of inductive method while establishing the proof of correspondence theorem of set theory. But further in the same study, the pattern of deductive method has been used while establishing the proof of the converse of Baye's theorem. Thus inductive and deductive both methods have been applied according as the nature and need of the content to be derived as conclusion.

Results and Discussions

Theory

Before proving the theorems mentioned above, some basic rules from algebra of sets, which are applicable in proving, can be put here. These are as follows-

If A and B are two subsets of a universal set U then concerning to the definition of union and intersection of two sets we have the following properties:

(i) $x \in A \cup B \Leftrightarrow x \in A \; or \; x \in B$
(ii) $x \notin A \cup B \Leftrightarrow x \notin A$ and $x \notin B$
(iii) $x \in A \cap B \Leftrightarrow x \in A$ and $x \in B$
(iv) $x \notin A \cap B \Leftrightarrow x \notin A$ or $x \notin B$

Let us define the operation of union and intersection between two families of sets each containing the equal number of two sets. This can be shown by the following example:

Example 1: *If* $\mathcal{A} = \{A_1, A_2\}$ *and* $\mathcal{B} = \{B_1, B_2\}$ *are two families of sets which are subsets of a universal set* U, *then*

(I) $(A_1 \cap A_2) \cup (B_1 \cap B_2) = (A_1 \cup B_1) \cap (A_2 \cup B_2) and$
(II) $(A_1 \cup A_2) \cap (B_1 \cup B_2) = (A_1 \cap B_1) \cup (A_2 \cap B_2)$

Solution:

The union over intersections of two families of sets and each family containing only two sets (as shown in figure 1) can be verified as follows:

(I) Let $x \in \{(A_1 \cap A_2) \cup (B_1 \cap B_2)\}$

$\Leftrightarrow x \in (A_1 \cap A_2) or, x \in (B_1 \cap B_2)$

$\Leftrightarrow x \in A_1$ and $x \in A_2 or, x \in B_1$ and $x \in B_2$

$\Leftrightarrow x \in A_1 or, x \in B_1$ and $x \in A_2 or, x \in B_2$

$\Leftrightarrow x \in (A_1 \cup B_1)$ and $x \in (A_2 \cup B_2)$

$\Leftrightarrow x \in \{(A_1 \cup B_1) \cap (A_2 \cup B_2)\}$

Thus, $x \in \{(A_1 \cap A_2) \cup (B_1 \cap B_2)\}$

$\Leftrightarrow x \in \{(A_1 \cup B_1) \cap (A_2 \cup B_2)\}$

$\therefore (A_1 \cap A_2) \cup (B_1 \cap B_2) = (A_1 \cup B_1) \cap (A_2 \cup B_2)$

Figure 1 *Union over Intersections*

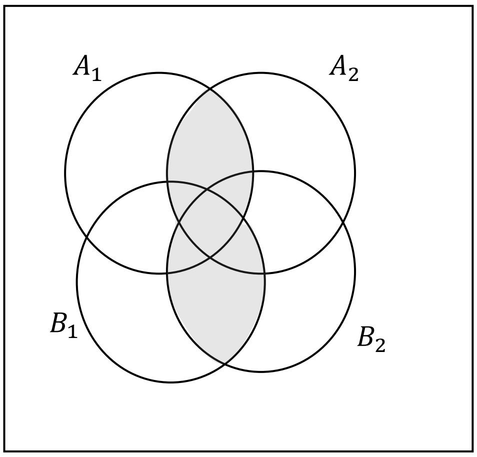
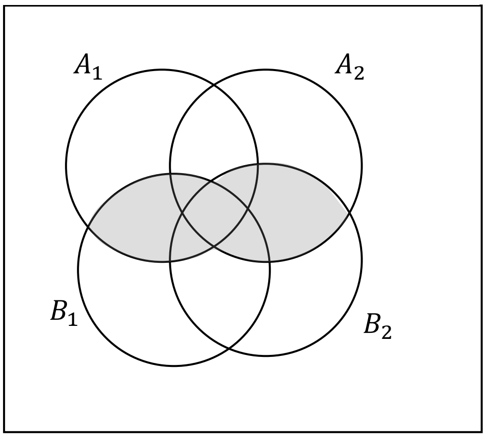


Figure 1 presents the results after the union over the intersection of two family of sets containing only two sets on both families in Venn diagram.

(II)

The intersection over unions of two families of sets and each family containing only two sets (as shown in figure2) can be verified as follows-

Figure 2 *Intersection over Unions*



Let $x \in \{(A_1 \cup A_2) \cap (B_1 \cup B_2)\}$

$\Leftrightarrow x \in (A_1 \cup A_2)$ and $x \in (B_1 \cup B_2)$

$\Leftrightarrow x \in A_1 or, x \in A_2$ and $x \in B_1 or, x \in B_2$

$\Leftrightarrow x \in A_1$ and $x \in B_1 or, x \in A_2$ and $x \in B_2$

$\Leftrightarrow x \in (A_1 \cap B_1) or, x \in (A_2 \cap B_2)$

$\Leftrightarrow x \in \{(A_1 \cap B_1) \cup (A_2 \cap B_2)\}$

Thus, $x \in \{(A_1 \cup A_2) \cap (B_1 \cup B_2)\}$

$\Leftrightarrow x \in \{(A_1 \cap B_1) \cup (A_2 \cap B_2)\}$

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

$$\therefore (A_1 \cup A_2) \cap (B_1 \cup B_2) = (A_1 \cap B_1) \cup (A_2 \cap B_2)$$

Figure 2 presents the result after the intersection over the union of two family of sets containing only two sets on both families in Venn diagram.

Thus we see that the operation of union and intersection can easily be defined between two family of sets each containing only two sets. Now, we shall define the same operations between two families of sets containing infinite number of sets. We have called, this relation between two families of sets, as correspondence theorem of sets. For this purpose we shall consider first, two equivalent families of sets and the theorem proceeds as follows:

Correspondence theorem for two families of sets

Theorem 1: *If* $A = \{A_i : i \in I\}$ *and* $B = \{B_i : i \in I\}$, *where I is an index set, are any two indexed families of equivalent sets then*

(i)$(\bigcup_{i\in I} A_i) \cap (\bigcup_{i\in I} A_i) = \bigcup_{i\in I}(A_i \cap B_i)$*and*

(ii)$(\bigcap_{i\in I} A_i) \cup (\bigcap_{i\in I} B_i) = \bigcap_{i\in I}(A_i \cup B_i)$

Proof:

We shall prove this theorem by the method of induction.

If $i = 2$, then the relation is true as shown in example 1.

If $i = 3$, then we can easily show that for $A = \{A_i\}$ and $B = \{B_i\}$, for $i = 1, 2, 3$ the following relation is true:

(i) $(A_1 \cap A_2 \cap A_3) \cup (B_1 \cap B_2 \cap B_3)$

Let$C = A_1 \cap A_2$and $D = B_1 \cap B_2$ then we get

$$(A_1 \cap A_2 \cap A_3) \cup (B_1 \cap B_2 \cap B_3) = (C \cap A_3) \cup (D \cap B_3)$$

Then by using example 1, we get

$$= (C \cup D) \cap (A_3 \cup B_3)$$

$$= (A_1 \cap A_2) \cup (B_1 \cap B_2) \cap (A_3 \cup B_3)$$

Again, by using example 1, we get

$$= (A_1 \cup B_1) \cap (A_2 \cup B_2) \cap (A_3 \cup B_3)$$

$$i.e., \left(\bigcap_{i=1}^{3} A_i\right) \cup \left(\bigcap_{i=1}^{3} B_i\right) = \bigcap_{i=1}^{3}(A_i \cup B_i)$$

(ii) $(A_1 \cup A_2 \cup A_3) \cap (B_1 \cup B_2 \cup B_3)$

Let$C = A_1 \cup A_2$and $D = B_1 \cup B_2$ then we get

$$(A_1 \cup A_2 \cup A_3) \cap (B_1 \cup B_2 \cup B_3) = (C \cup A_3) \cap (D \cup B_3)$$

Then by using example 1, we get

$$= (C \cap D) \cup (A_3 \cap B_3)$$

$$= (A_1 \cup A_2) \cap (B_1 \cup B_2) \cup (A_3 \cap B_3)$$

Again, by using example 1, we get

$$= (A_1 \cap B_1) \cup (A_2 \cap B_2) \cup (A_3 \cap B_3)$$

$$i.e., \left(\bigcup_{i=1}^{3} A_i\right) \cap \left(\bigcup_{i=1}^{3} B_i\right) = \bigcup_{i=1}^{3} (A_i \cup B_i)$$

Let the relation be true for $i = k$, then we get

$\left(\bigcap_{i=1}^{k} A_i\right) \cup \left(\bigcap_{i=1}^{k} B_i\right) = \bigcap_{i=1}^{k} (A_i \cup B_i) - - - - - (a)$ and

$$\left(\bigcup_{i=1}^{k} A_i\right) \cap \left(\bigcup_{i=1}^{k} B_i\right) = \bigcup_{i=1}^{k} (A_i \cap B_i) - - - - - (b)$$

Now,

$$\left(\bigcap_{i=1}^{k} A_i \cap A_{k+1}\right) \cup \left(\bigcap_{i=1}^{k} B_i \cap B_{k+1}\right)$$

Let$C = \bigcap_{i=1}^{k} A_i$and $D = \bigcap_{i=1}^{k} B_i$ then we get

$$\left(\bigcap_{i=1}^{k} A_i \cap A_{k+1}\right) \cup \left(\bigcap_{i=1}^{k} B_i \cap B_{k+1}\right) = (C \cap A_{k+1}) \cup (D \cap B_{k+1})$$

Then by using example 1, we get

$$= (C \cup D) \cap (A_{k+1} \cup B_{k+1})$$

$$= (\bigcap_{i=1}^{k} A_i) \cup (\bigcap_{i=1}^{k} B_i) \cap (A_3 \cup B_3)$$

Again, by using (a), we get

$$= \bigcap_{i=1}^{k} (A_i \cup B_i) \cap (A_{k+1} \cup B_{k+1})$$

$$= \bigcap_{i=1}^{k+1} (A_i \cup B_i)$$

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

And similarly, we have

$$i.e., \left(\bigcup_{i=1}^{k} A_i \cup A_{k+1}\right) \cap \left(\bigcup_{i=1}^{k} B_i \cup B_{k+1}\right) = \bigcup_{i=1}^{k+1} (A_i \cap B_i)$$

This implies that the relation is true for $i = k + 1$. Then by Mathematical induction the relation is true for all $i \in I$, where I is an index set. Hence the theorem is proved.

This theorem is applicable in many set theory relations. Some of them are given in the following examples:

Example 2: If M and N are any two subsets of a universal set U, then $\mathrm{M\Delta N} = (\mathrm{M} \cup \mathrm{N}) - (\mathrm{M} \cap \mathrm{N})$.

Here, $M\Delta N = (M - N) \cup (N - M) = (M \cap N') \cup (N \cap M')$

$= (M \cup N) \cap (N' \cup M') = (M \cup N) \cap (M \cap N)'$

$= (M \cup N) - (M \cap N)$ □

Example 3: For Distributive law, if A, B and C are any three subsets of a universal set U, then

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Here,

$$A \cup (B \cap C) = (A \cap A) \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Also, $A \cap (B \cup C) = (A \cup A) \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Example 4: *If A and B are any two subsets of a universal set* U, *then*

$(A \cup B) \cap (A \cup B') = A$ *and* $(A \cap B) \cup (A \cap B') = A$

Here,

$$(A \cup B) \cap (A \cup B') = (A \cap A) \cup (B \cap B') = A \cup \emptyset = A$$

Also, $(A \cap B) \cup (A \cap B')$
$= (A \cup A) \cap (B \cup B') = A \cap U = A$

Correspondence theorem can be used as the extended form of distributive law. In distributive law for sets, the operations of union and intersection are distributive over intersection and union respectively. i.e. union and intersection with only one set are distributive over the intersection and union of only two sets respectively. But the correspondence theorem allows the distribution of union operation of a family of sets over the intersection operation of another family of sets and the intersection operation of a family of sets over the union operation of another family of sets. This is the major characteristic of correspondence theorem for sets.

Now, we shall use correspondence theorem of set theory for the axiomatic proof of Baye's theorem and its converse. For this purpose Bayes' theorem has been presented here in different form.

Theorem 2: *(Bayes' theorem) If* $A_1, A_2, A_3, \dots\dots, A_n$ *are mutually exclusive and exhaustive events of a sample space S of an experiment then for any event B of S with* $p(B) \neq 0$, *we get*

$P\left(\frac{A_i}{B}\right) = \frac{P(A_i)P\left(\frac{B}{A_i}\right)}{P(A_1)P\left(\frac{B}{A_1}\right)+P(A_2)P\left(\frac{B}{A_2}\right)+\cdots\dots\dots+P(A_n)P\left(\frac{B}{A_n}\right)}$, *for any event* A_i, all $i = 1,2,3,\dots,n$

Before the formal proof of this theorem, let us modify this statement with new notation without losing of its meaning and this can be done as follows:

If we put $B_i = B \cap A_i$, for all $i = 1,2,3,\dots,n$ then the above theorem takes the form

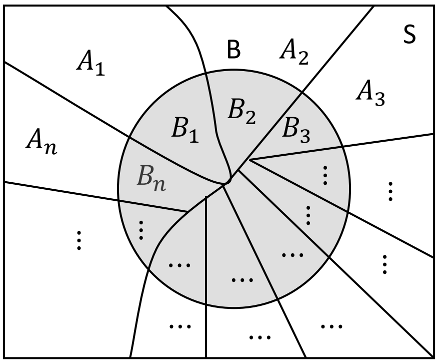
$$P\left(\frac{A_i}{B}\right) = \frac{P(A_i)P\left(\frac{B_i}{A_i}\right)}{P(A_1)P\left(\frac{B_1}{A_1}\right) + P(A_2)P\left(\frac{B_2}{A_2}\right) + \cdots\dots\dots + P(A_n)P\left(\frac{B_n}{A_n}\right)}$$

Proof:

Let $A_1, A_2, A_3, \dots\dots, A_n$ be mutually exclusive and exhaustive events of a sample space S of an experiment.

Figure 3represents the partition of the sample space S of the experiment with the mutually exclusive and exhaustive events $A_1, A_2, A_3, \dots\dots, A_n$ and their intersection with an arbitrary event B of the experiment.

Figure 3 *Bayes' Theorem*



Let B be any event of the sample space S as shown in fig. 3.

Then, $A_1 \cup A_2 \cup \dots\dots \cup A_n = S$, and $A_i \cap A_j = \emptyset$, for all $i \neq j$.

Also, since $B_i = B \cap A_i$, for all $i = 1,2,3,\dots,n$ therefore, $B = B_1 \cup B_2 \cup \dots\dots \cup B_n$

Now, $B = B \cap S$

$$= (B_1 \cup B_2 \cup \dots\dots \cup B_n) \cap (A_1 \cup A_2 \cup \dots\dots \cup A_n)$$

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

Applying correspondence theorem for set, we get

$$B = (B_1 \cap A_1) \cup (B_2 \cap A_2) \cup \dots \dots \cup (B_n \cap A_n)$$

Since $A_i \cap A_j = \emptyset$, for all $i \neq j$and $B_i \subset A_i$, for all $i = 1,2,3, \dots, n$ therefore

$(B_i \cap A_i) \cap \left(B_j \cap A_j\right) = \emptyset$, for all $i \neq j$

Hence, $P(B) = P(B_1 \cap A_1) + P(B_2 \cap A_2) + \cdots \dots \dots + P(B_n \cap A_n)$

Since $P(B_i \cap A_i) = P(A_i)P\left(\frac{B_i}{A_i}\right)$, for all $i = 1,2,3, \dots, n$ therefore, we get

$$P(B) = P(A_1)P\left(\frac{B_1}{A_1}\right) + P(A_2)P\left(\frac{B_2}{A_2}\right) + \cdots \dots + P(A_n)P\left(\frac{B_n}{A_n}\right) \dots \dots \dots \dots \dots (i)$$

Now, the conditional probability of any event A_i given event B is given by

$$p\left(\frac{A_i}{B}\right) = \frac{p(A_i \cap B)}{p(B)} = \frac{p(A_i \cap B_i)}{p(B)}$$

$= \frac{p(A_i)p\left(\frac{B_i}{A_i}\right)}{p(A_1)p\left(\frac{B_1}{A_1}\right) + p(A_2)p\left(\frac{B_2}{A_2}\right) + \cdots \dots + p(A_n)p\left(\frac{B_n}{A_n}\right)}$for all $i = 1,2,3, \dots \dots, n$ $[\therefore using (i)]$

□

Example 5: In a bolt factory, the company has uses three machines A, B and Cthat produces 25%, 35% and 40% of the total output respectively. Out of these outputs, 5%, 4% and 2%, respectively, are defective items. If a bolt is chosen at random from the combined output, what is the probability that it is a defective bolt? If a bolt chosen at random was found to be defective, then find the probability that it was produced by?

Solution:

Here, $P(A) = 25\% = 0.25$, $P(B) = 35\% = 0.35$, $P(C) = 40\% = 0.4$. Let D denotes the event of selecting a defective items. Then

$P\left(\frac{D}{A}\right) = 5\% = 0.05$, $P\left(\frac{D}{B}\right) = 4\% = 0.04$ and $P\left(\frac{D}{C}\right) = 2\% = 0.02$.

$P(D) = ?$, $P\left(\frac{B}{D}\right) = ?$

Now, $P(D) = P(A)P\left(\frac{D}{A}\right) + P(B)P\left(\frac{D}{B}\right) + P(C)P\left(\frac{D}{C}\right)$

$$= 0.25 \times 0.05 + 0.35 \times 0.04 + 0.4 \times 0.02 = 0.0345$$

Also, $P\left(\frac{B}{D}\right) = \frac{P\left(\frac{D}{B}\right)P(B)}{P(D)} = \frac{0.04 \times 0.35}{0.0345} = 0.406$ □

The probability of one of the events$A$, in the joint event (intersection of two events) is called its marginal probability. For example: if $P(A) = P(A \cap B) + P(A \cap B')$, then $P(A)$ is called marginal probability of event A. In

Bayes' theorem, the events A_i for $i = 1,2,3, \dots, n$ are considered unobservable because we don't know that which one of them occurred but the event B is an observable event. Therefore the marginal probabilities $P(A_i)$ for $i = 1,2, \dots \dots, n$ are assumed known before we start and are called *prior probabilities*. The conditional probability $P\left(\frac{B}{A_i}\right)$ of the event B given events A_i, for $i = 1, 2, \dots \dots, n$, is the *likelihood* of the events $A_1, A_2, \dots \dots, A_n$. The conditional probability $P\left(\frac{A_i}{B}\right)$ for $i = 1,2, \dots \dots, n$ is the *posterior probability* of event A_i, given that event B has occurred.

If the prior probabilities $P(A_i)$, $i = 1,2,3, \dots, n$ are equal then it is a constant, say $c.\ i.e.$if $P(A_i) = c$, for $i = 1,2,3, \dots, n$ then from Bayes' theorem, we get $P\left(\frac{A_i}{B}\right) = \frac{P\left(\frac{B}{A_i}\right)}{P(B)}$. Therefore, Bolstad and Curran (2017), p.73, has suggested to write the posterior probability as

$$posterior \propto prior \times likelihood$$

The proof of the converse part of Bayes' theorem has been possible only by using the correspondence theorem of set theory. The axiomatic approach facilitates us to use the theorems of set theory in probability.

Theorem 3: *(Converse of Bayes' theorem) If $A_1, A_2, A_3, \dots \dots, A_n$ are events of a sample space S of a random experiment and any other event B of S satisfies the following condition*

$$P\left(\frac{A_i}{B}\right) = \frac{P(A_i)P\left(\frac{B_i}{A_i}\right)}{P(A_1)P\left(\frac{B_1}{A_1}\right)+P(A_2)P\left(\frac{B_2}{A_2}\right)+\cdots\dots+P(A_n)P\left(\frac{B_n}{A_n}\right)},\text{ where, } B_i \subset A_i, \text{for all } i.$$

Then the events $A_1, A_2, A_3, \dots \dots, A_n$ are mutually exclusive and exhaustive.

Proof:

Here we shall show that $A_1 \cup A_2 \cup \dots \dots \cup A_n = S$ and $A_i \cap A_j = \emptyset$, for all $i \neq j$.

Since $P(B) = P(A_1)P\left(\frac{B_1}{A_1}\right) + P(A_2)P\left(\frac{B_2}{A_2}\right) + \cdots \dots + P(A_n)P\left(\frac{B_n}{A_n}\right)$

$= \sum_{i=1}^{n} P(A_i)P\left(\frac{B_i}{A_i}\right) = \sum_{i=1}^{n} P(A_i \cap B_i) = \sum_{i=1}^{n} P(B_i)$

$[\because B_i \subset A_i$ for all $i]$

$$\Rightarrow P(B_1) + P(B_2) + \cdots \dots + P(B_n) = P(B)$$

$\Rightarrow B_i \cap B_j = \emptyset$, for all $i \neq j$ and $B_1 \cup B_2 \cup \dots \dots \cup B_n = B$

Thus the set $\{B_1, B_2, B_3, \dots \dots, B_n\} - - - (i)$

Form a partition on the set B

 Vol.3, Issue 1, Dec. 2024 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

Also, applying correspondence theorem for set, we get

$$B = \bigcup_{i=1}^{n}(A_i \cap B_i) = \left(\bigcup_{i=1}^{n} A_i\right) \cap \left(\bigcup_{i=1}^{n} B_i\right) = \left(\bigcup_{i=1}^{n} A_i\right) \cap B$$

$$\Rightarrow \bigcup_{i=1}^{n} A_i = S$$

Again, since $B_i \subset A_i$, for all $i = 1,2,3,\ldots,n$ therefore $A_i = B_i \cup (A_i - B_i)$

Then, $P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = \sum_{i=1}^{n} P\{B_i \cup (A_i - B_i)\}$

$= \sum_{i=1}^{n}\{P(B_i) + P(A_i - B_i)\}$

$[\therefore B_i \cap (A_i - B_i) = \emptyset$, for all $i]$

$$= \sum_{i=1}^{n} P(B_i) + \sum_{i=1}^{n} P(A_i - B_i)$$

$$= P(B) + \sum_{i=1}^{n} P(A_i - B_i)$$

$$or, P(S) = P(B) + \sum_{i=1}^{n} P(A_i - B_i)$$

$$or, P(S) - P(B) = \sum_{i=1}^{n} P(A_i - B_i)$$

$$or, P(S - B) = \sum_{i=1}^{n} P(A_i - B_i)\ [\therefore B \subseteq S]$$

$$\Rightarrow \{(A_i - B_i), (A_i - B_i), \ldots\ldots, (A_i - B_i)\} ----(ii)$$

Form a partition on set $S - B$.

Then, from (i) and (ii), we get

$\{A_1, A_2, A_3, \ldots\ldots, A_n\}$Form a partition on set S.

$i.e. A_1 \cup A_2 \cup \ldots\ldots \cup A_n$
$= S$ and $A_i \cap A_j = \emptyset$, for all $i \neq j$.

Note:

(1) By the definition of dependent events we get

$P(A_i \cap B) = P(A_i)P\left(\frac{B}{A_i}\right) ----- (i)$ and also

$$P(A_i \cap B) = P(B)P\left(\frac{A_i}{B}\right) - - - - - - (ii)$$

Then from (i) and (ii), we get

$$P(B)P\left(\frac{A_i}{B}\right) = P(A_i)P\left(\frac{B}{A_i}\right)$$

$$Or, P\left(\frac{A_i}{B}\right) = \frac{P(A_i)P\left(\frac{B}{A_i}\right)}{P(B)}$$

$$Or, P(A_i) = \frac{P\left(\frac{A_i}{B}\right)P(B)}{P\left(\frac{B}{A_i}\right)}$$

(2) Similarly, for the dependent events A_i and B', we get

$$P(B')P\left(\frac{A_i}{B'}\right) = P(A_i)P\left(\frac{B'}{A_i}\right)$$

$$Or, P\left(\frac{A_i}{B'}\right) = \frac{P(A_i)P\left(\frac{B'}{A_i}\right)}{P(B')}$$

$$Or, P(A_i) = \frac{P\left(\frac{A_i}{B'}\right)P(B')}{P\left(\frac{B'}{A_i}\right)}$$

Applications

Example 6: Three people X, Y, Z have been nominated for the post of manager's. If they are selected then the probability of X, Y and Z will introduce Bonus scheme are $0.6, 0.3$ and 0.4 respectively. The probability of Bonus scheme is 0.455. If the Bonus scheme has been introduced and the probability of selection of X, Y and Z as manager were 0.5275, 0.2308 and 0.213 respectively then find the probability of getting elected for the candidates X, Y and Z for the post of manager.

Solution:

Let the B denotes the event of introducing Bonus scheme.

Here, $P\left(\frac{B}{X}\right) = 0.6, P\left(\frac{B}{Y}\right) = 0.3$ and $P\left(\frac{B}{Z}\right) = 0.4$. Also, $P\left(\frac{X}{B}\right) = 0.5275$, $P\left(\frac{Y}{B}\right) = 0.2308$ and $P\left(\frac{Z}{B}\right) = 0.22$. We have given that $P(B) = 0.455$.

But from the note 1, mentioned above, we know that

 ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

$P(A_i) = \frac{P\left(\frac{A_i}{B}\right)P(B)}{P\left(\frac{B}{A_i}\right)}$, therefore we get

$$P(X) = \frac{P\left(\frac{X}{B}\right)P(B)}{P\left(\frac{B}{X}\right)} = \frac{0.5275 \times 0.455}{0.6} = 0.4$$

$$P(Y) = \frac{P\left(\frac{Y}{B}\right)P(B)}{P\left(\frac{B}{Y}\right)} = \frac{0.2308 \times 0.455}{0.3} = 0.35$$

$P(Z) = \frac{P\left(\frac{Z}{B}\right)P(B)}{P\left(\frac{B}{Z}\right)} = \frac{0.22 \times 0.455}{0.4} = 0.25$□

Example 7: A motorcycle company uses tyres manufactured by three different companies, A, B and C . The tyres supplied by these companies A, B and C respectively 50%, 30% and 20% of the total demand of tyres. Some tyres supplied by these companies to the motorcycle company are found to be defective with their probability $0.017$and on those defective tyres supplied by A, B and C have probabilities 0.29, 0.35 and 0.35 respectively. If a tyre was selected at random from A then find the probability that it is a defective one.

Solution:

Here, if$D$ denotes the event of selection of a defective tyres, then

$P(D) = 0.017,\ P\left(\frac{A}{D}\right) = 0.29, P\left(\frac{B}{D}\right) = 0.35$ and $P\left(\frac{C}{D}\right) = 0.35$

Then $P\left(\frac{D}{A}\right) = ?$

But from the note 1, mentioned above, we know that

$P\left(\frac{B}{A_i}\right) = \frac{P\left(\frac{A_i}{B}\right)P(B)}{P(A_i)}$, therefore we get

$P\left(\frac{D}{A}\right) = \frac{0.29 \times 0.017}{0.5} = 0.01$ □

This article has been presented the appropriate use of correspondence theorem of set theory. Correspondence theorem obviates the repeated use of distributive law between the sets of two families of sets. An inductive proof has been provided to the correspondence theorem in this study. Bayes' theorem has been mentioned almost every Statistics books containing probability theorem in its content. But it has been hardly seen any books mentioned about the converse of Bayes theorem and its proof. So this article has completely stated the converse of Bayes theorem and has provided a deductive proof of the theorem.

Conclusion

An axiomatic approach to probability proceeds on the basic concepts of subsets as events in probability and the sample space of a random experiment as their universal set. Then the axioms and theorems for subsets and sets are also applicable with the events in probability theory. Correspondence theorem of set theory has one of the most important applications in probability. It was seen that correspondence theorem is not only applicable for the solution to set theory's example and in proving theorems but it is also vital in proving the theorems regarding probability theory based on the axiomatic approach to probability. Here, Bayes's theorem and converse of Bayes's theorem can also be proved using correspondence theorem that also represents the application of the correspondence theorem in probability. This article presents that the Bayes's theorem is not only true necessarily but it is also true conversely. The operation of union and intersection extended from two distinct sets to two families of sets, has make it possible to prove this important relation of converse theorem of Baye's.

ISSN: 2976-1379 (Print)
ISSN: 2976-1387 (Online)

References

Kolmogorov, A. N. (1950). *Foundations of the theory of Probability*. Chelsea Publishing Company.

Bayes, R. T. (2004), Statistical Science. *19*(1), 3–43. https://doi.org/10.1214/088342304000000189

Bolstad, W.M., & Curran, J. M. (2017). *Introduction to Bayesian Statistics* (3rd ed.). John Wiley & Sons.

Ash, R. B. (2008). *Basic probability theory*. Dover publication.

Halmos, P. R. (1974). *Naive Set Theory*. Springer-Verlag.

Maskey, S. M. (1997). *Introduction to Modern Mathematics*, Vol. 1 (2nd ed.). Ratna Pustak Bhandar.

Freiwald, R. C. (2014). *An Introduction to Set Theory and Topology*. Washington University in St. Louis. Retrieved from: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1020&context=books

Downey, A. B. (2012). *Think Bayes Bayesian Statistics made simple*, Green Tea press Needham Massachusetts.

Bertsekas, D. P., & Tsitsiklis, J. N. (2000). *Introduction to Probability*, *Lecture Notes*. Massachusetts Institute of Technology, Cambridge Massachusetts. Retrieved from https://www.vfu.bg/en/e-Learning/Math--Bertsekas_Tsitsiklis_Introduction_to_probability.pdf

Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). S*chaum's Outlines of Probability and Statistics*, (4th ed.). The McGraw-Hill Companies.

Pinter, C. C. (2014). *A Book of Set Theory*. Dover Publications.