e-ISSN: 2564-7954 CUSJE 18(1): 061-071 (2021) Short Communication

Çankaya University

Journal of Science and Engineering

CUJSE

https://dergipark.org.tr/cankujse

Innovative Teaching Pedagogy For Teaching and Learning of Bayes' Theorem

Ikram E Khuda

Faculty of Engineering Sciences and Technology, Iqra University, Karachi, Pakistan

Keywords

Probability and Statistics
Bayes' Theorem
Probability Trees
Teacher Education
Pedagogy

Abstract

Probability and Statistics are difficult courses to learn by the statistics students and thus giving a challenging job for the instructors to teach. Among the many topics, most commonly, Bayes' theorem and its associated probabilities is one common area which causes a lot of stir in teaching and learning. In such difficult scenarios, the instructor's teaching pedagogy plays a vital role to inculcate correct knowledge to the students, as required. In this paper we have discussed probability tree diagrams and presented that they are better off to help understand and teach Bayes' theorem with a completeness of process thinking. Tree diagrams methodology is presented on varying examples –simple to difficult as a prove of its generalization and entireness in statistics learning of Bayes' probabilities. This is an instructional article, in which an innovative pedagogy for teaching Bayes' theorem is discussed. We have provided a scholarly argument to justify for considering its wider use. There is no claim that the proposed method is simpler from presently used methods but however this is claimed that this will unleash the details of Bayes' probabilities which are neither taught in the class nor emphasized in the texts.

1. Introduction

Bayes' theorem and the associated probabilities are a demanding topic in the courses of probability, statistics and quantitative analysis subjects. For example, there are numerous issues in modern economics where they are used, Nguyen et al., [1]. Bayes' theorem is key technique in decision making theory. It has diverse applications in engineering, business and also in medical diagnosis, Sahai, [2]. Therefore, it is very important for the instructors to not only have a sound knowledge of the subject but also on its effective teaching pedagogy.

Bayes' theorem is a probabilistic model encompassing learning from data. It helps to improve the uncertainties associated to the occurrences of events using new evidence. The resultant probability therefore does not depend on the number of trials performed in a random experiment, opposite to the frequentists' approach**.** Hence it is also called as probabilistic model of learning from experience. The decisioning strategy of human brain follows the similar patterns as is used by Bayes' theorem. Kolb, [3]. The popular mathematical formula of Bayes' theorem is indicated as, Florens, [4],

$$\boldsymbol{P}(\boldsymbol{x}|\boldsymbol{y}) = \frac{\boldsymbol{P}(\boldsymbol{y}|\boldsymbol{x})\boldsymbol{P}(\boldsymbol{x})}{\boldsymbol{P}(\boldsymbol{y})} \tag{1}$$

Here, $\boldsymbol{P}(\boldsymbol{x}|\boldsymbol{y})$ and $\boldsymbol{P}(\boldsymbol{x})$ are the posterior and prior probabilities or distributions respectively for the event or variable $\boldsymbol{x}$. $\boldsymbol{P}(\boldsymbol{y})$ and $\boldsymbol{P}(\boldsymbol{y}|\boldsymbol{x})$ are the likelihood and marginal likelihood of the data variable $\boldsymbol{y}$ respectively. Prior probability is the *belief* about the variable being measured before some evidence is taken into account.

* Corresponding Author: mailto:ikram@iqra.edu.pk
Received: February 4, 2021, Accepted: March 18, 2021

Upon receiving and using the likelihood of evidence, prior probability is updated into posterior probability. The posterior and marginal likelihood are conditional probabilities, indicating that they both show probabilities of statistically dependent events. Marginal likelihood is a normalizing constant used in Bayes' equation (1). Marginal likelihood therefore does not change the shape of posterior probability but only scales it to create it a valid probability by making its area sum to 1.

Equation (1) is a basic model to use and teach Bayes' theorem in undergraduate probability and statistics courses involving descriptive statistics. This equation is also applicable to make Baysian inferences in inferential statistics.

In the literature study it was found that mathematicians and statisticians have given careful consideration on teaching Bayes' Theorem so that the difficulty level of its learning could significantly reduce. Educationists and researchers in educational science and particularly from statistics education domain have explored different teaching methodologies to make Bayes' theorem in (1) and its probabilities understandable to the students.

A survey of the textbooks used for Freshman Business Mathematics (FBM) courses showed that they include topics for example counting principles, basic probability and Bayes Theorem along with basic algebra and pre-calculus etc, Green [5].

Bolstad, [6] explicitely made a one semester introductory course on Baysian Statistics with the objective to produce better statisticians for future needs. Lecoutre [7], supported the similar experience that students find it hard to learn Baysian probabilities and that the Baysian approach, in context to experimental data analysis, is the requirement for future data science.

Holt [8], emphasized on the importance of Bayes' rule for handling economic data in the study of uncertain markets. Using ball/ urn example as teaching aid, the practical drawing of colored marbles as the teaching method of counting heuristic was used to help students understand and learn conditional probabilities and Baye's theorem. The experiment comprised of a single die and 6 marbles; three of one color and other three of another color. To encapsulate the findings; for a prior probability of half that each urn is used and if the urns hold equivalent numbers of colored marbles, then the posterior probabilities can be computed as percentages of numbers of marbles of the color drawn.

Sedlmeir, [9] and Kurzenhäuser, [10], also supported the similar idea that problems involving Baysian calculations are easier when data is provided in natural frequencies as compared to probability and ratio representations. Accordingly, natural frequencies are the natural way humans have encountered problems historically.

Gelman, [11], Jackman, [12], Kruschke, [13], Faulkenberry, [14], and Rouder, [15], have advocated the idea of teaching Bayes' theorem in probability and ratio forms. Accordingly, this representation helps to better interpret the Baysian probabilities as improvement in prior beliefs to posterior beliefs. Furthermore, using probability idea, the concept of level of significance in statistical inference is easier to understand.

Faulkenberry,[14], and Rouder,[15] have emphasized that marginal probability of data variable i.e. $\boldsymbol{p}(\boldsymbol{y})$ in (1) is a unique Baysian concept and should not be missed when performing significant testing in statistical inference. They considered that $\boldsymbol{p}(\boldsymbol{x}|\boldsymbol{y})$ in (1) is the probability of null hypothesis $\boldsymbol{H_o}$ given the data $\boldsymbol{D}$ (the alternate hypothesis, $\boldsymbol{H_a}$). So $\boldsymbol{p}(\boldsymbol{x}|\boldsymbol{y})$ can be written as $\boldsymbol{p}(\boldsymbol{H_o}|\boldsymbol{D})$ which is the $\boldsymbol{p}$ values quoted in statistics texts and used to test for the rejection or not rejection of null hupothesis for hypothesis/ significant testing in statistical inference. Hence in this way Baysian inferencing can be readily used to test the relative probability of correctness for the two contending hypotheses; null hypothersis ($\boldsymbol{H_o}$) and alternate hypothesis ($\boldsymbol{H_a}$). Thus, relatively approximating the hypothesis testing that an effect is existing when it just changes from zero value, it can be concluded that the likelihood of actuality exterior to an explicit range signifying for "essentially/practically no effect". In Baysian inferencing this interval is called the region of practical equivalence (ROPE). The ROPE is the region corresponding to null hypothesis and is used to perform equivalence testing to test if the parameter is significant. This is the take of Baysian inferencing in contrast to statistical inference.

Frequency and ratios approach to find conditional and Bayes' probabilities can be connected using cross tabulation or probability contingency tables. The frequency information which is used to find Bayes' probabilities can be extended in the form of a cross tabulation or probability contingency table and then using simple calculations the probabilities or ratios can be easily determined. Hence these probability contingency tables serve as bridge between the two approaches to solve for Bayes' theorem, Rässler [16].

There is yet another way of finding and learning Bayes' theorem. This is the use of frequency diagram, also called as probability tree diagrams. Researchers quoted above have discussed about this approach as well but seldom this method is used in statistics text books. In this statistics education research we have explored further into the probability tree diagrams and highlighted some interesting facts giving further insight and conceptual meaning to Bayes' theorem. Probability tree diagrams open up to illustrate the chances of existence of all possible events from a random experiment which are not visible otherwise. We have also proposed a teaching method to teach these tree diagrams effectively in the context of learning Bayes' theorem. We support that probability tree diagrams should be the method to use in text books because they give the completeness to Bayes' theorem as not done by any other method.

2. Learning Bayes' Theorem from Probability Tree Diagrams Manuscript Length

Consider two random experiments X and Y which are statistically dependent on each others and both are mutually exclusive in nature. Random experiment X can either produce an outcome X or not X, i.e. it's complement $\overline{X}$. Random experiment Y can likewise either produce an outcome Y or not Y, i.e. it's complement $\overline{Y}$. This random experiment is performed after the first one and is taken to be statistically dependent upon it. The chances of occurrances of all possible events can be shown in the probability tree diagram as in Figure 1. We have quoted this scenario as random expermeiment Y performed after X.

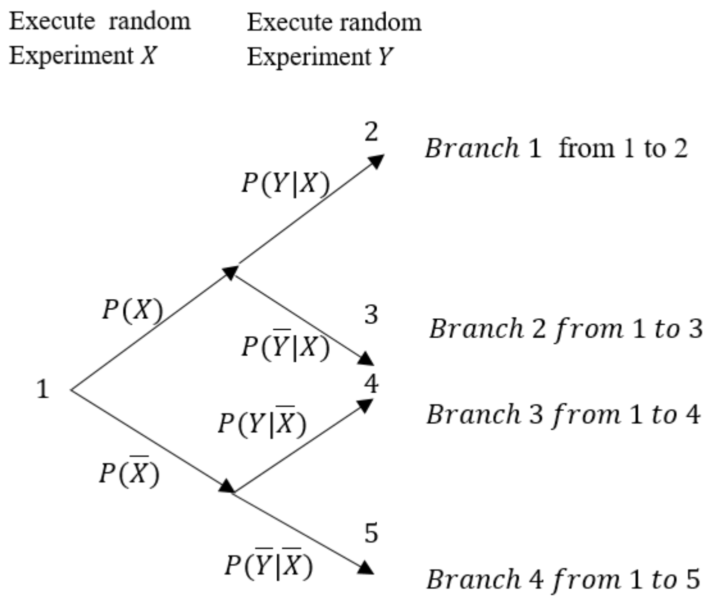

Figure 1. Probability tree diagram of random experiment Y performed after X.

In a similar scenario, if random experiment Y is performed first and afterwards X is performed such that it is also statistically dependent on Y. This second scenario is depicted as in Figure 2.

Now using basic set theory, probability knowledge, rule of sums and rule of products for statistically dependent events, we can relate Figure 1 and Figure 2 and write as the following different cases.

Case 1:

$$Branch\ 1\ (Figure\ 1)\ =\ Branch\ 1\ (Figure\ 2)$$
$$P(X \cap Y) = P(Y \cap X)$$
$$P(X)p(Y|X) = P(Y)P(X|Y)$$

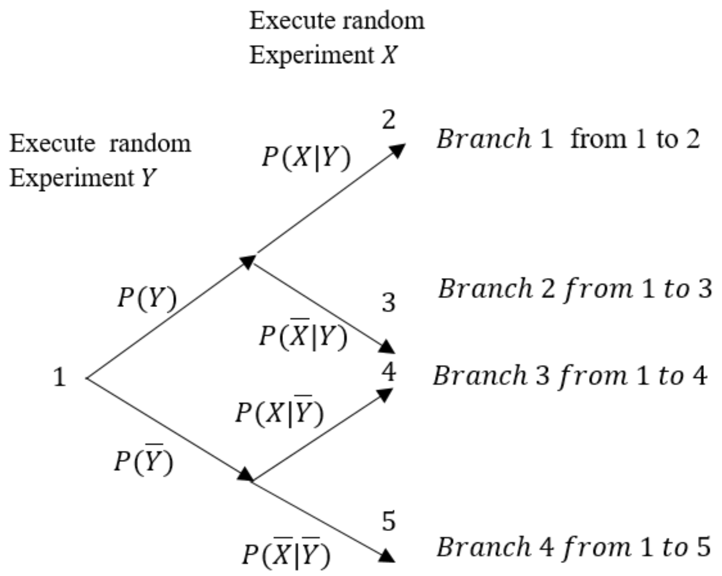

Figure 2. Probability tree diagram of random experiment X performed after Y.

$$P(Y|X) = \frac{P(Y)P(X|Y)}{P(X)} \tag{2}$$

$$P(X) = P(Y)P(X|Y) + P(\overline{Y})P(X|\overline{Y}) \tag{3}$$

Here, $P(Y)$ is the prior probability, $P(X|Y)$ is the likelihood and $P(Y|X)$ is the change or improvement in $P(Y)$ after the occurrence of X. Similarly, we can also have,

$$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)} \tag{4}$$

$$P(Y) = P(X)P(Y|X) + P(\overline{X})P(Y|\overline{X}) \tag{5}$$

Here, $P(X)$ is the prior probability, $P(Y|X)$ is the likelihood and $P(X|Y)$ is the change or improvement in $P(X)$ after the occurrence of Y.

Following observations can be made from case 1.

a. In case 1 it can be seen that (1) and (2) are Bayes rules.
b. Every term on the right side of (2) can be seen in Figure 2. Although $P(X)$ is present in Figure 1 and in the fist glimpse seems to be absent from Figure 2 but by using (3) , $P(X)$ can also be found from Figure 2 only.
c. Similarly, every term on the right side of (4) can be seen in Figure 1. Although $P(Y)$ is present in Figure 2 and in the fist glimpse seems to be absent from Figure 1 but by using (5) , $P(Y)$ can also be found from Figure 1 only.

Case 2:

$$Branch\ 2\ (Figure\ 1) = Branch\ 3\ (Figure\ 2)$$

$$P(X \cap \overline{Y}) = P(\overline{Y} \cap X)$$

$$P(X)P(\overline{Y}|X) = P(\overline{Y})P(X|\overline{Y})$$

$$P(\overline{Y}|X) = \frac{P(\overline{Y})P(X|\overline{Y})}{P(X)} \tag{6}$$

$$P(X) = P(Y)P(X|Y) + P(\overline{Y})P(X|\overline{Y}) \tag{7}$$

Here $P(\overline{Y})$ is the prior probability $P(X|\overline{Y})$ is the likelihood and $P(\overline{Y}|X)$ is the change or improvement in $P(\overline{Y})$ after the occurrence of X. Similarly, we can also have,

$$P(X|\overline{Y}) = \frac{P(X)P(\overline{Y}|X)}{P(\overline{Y})} \tag{8}$$

$$P(\overline{Y}) = P(X)P(\overline{Y}|X) + P(\overline{X})P(\overline{Y}|\overline{X}) \tag{9}$$

Here $P(X)$ is the prior probability, $P((\overline{Y}|X))$ is the likelihood and $P(X|\overline{Y})$ is the change or improvement in $P(X)$ after the occurrence of $\overline{Y}$.

Following observations can be made from case 2.

a. In case 2 it can be seen that (6) and (8) are Bayes rules.
b. Every term on the right side of (6) can be seen in Figure 2. Although $P(X)$ is present in Figure 1 and in the fist glimpse seems to be absent from Figure 2 but by using (7) , $p(X)$ can also be found from Figure 2 only.
c. Similarly, every term on the right side of (8) can be seen in Figure 1. Although $P(\overline{Y})$ is present in Figure 2 and in the fist glimpse seems to be absent from Figure 1 but by using (9) , $P(\overline{Y})$ can also be found from Figure 1 only.

Case 3:

$$Branch\ 3\ (Figure\ 1) = \ Branch\ 2\ (Figure\ 2)$$

$$P(\overline{X} \cap Y) = P(Y \cap \overline{X})$$

$$P(\overline{X})P(Y|\overline{X}) = P(Y)P(\overline{X}|Y)$$

$$P(Y|\overline{X}) = \frac{P(Y)P(\overline{X}|Y)}{P(\overline{X})} \tag{10}$$

$$P(\overline{X}) = P(Y)P(\overline{X}|Y) + P(\overline{Y})P(\overline{X}|\overline{Y}) \tag{11}$$

Here, $P(Y)$ is the prior probability, $P(\overline{X}|Y)$ is the likelihood and $P(Y|\overline{X})$ is the change or improvement in $P(Y)$ after the occurrence of $\overline{X}$. Similarly, we can also have,

$$P(\overline{X}|Y) = \frac{P(\overline{X})P(Y|\overline{X})}{P(Y)} \tag{12}$$

$$P(Y) = P(X)P(Y|X) + P(\overline{X})P(Y|\overline{X}) \tag{13}$$

Here, $P(\overline{X})$ is the prior probability, $P(Y|\overline{X})$ is the likelihood and $P(\overline{X}|Y)$ is the change or improvement in $P(\overline{X})$ after the occurrence of Y.

Following observations can be made from case 3.

a. In case 3 it can be seen that (10) and (12) are Bayes rules.
b. Every term on the right side of (12) can be seen in Figure 2. Although $P(\overline{X})$ is present in Figure 1 and in the fist glimpse seems to be absent from Figure 2 but by using (11) , $P(\overline{X})$ can also be found from Figure 2 only.
c. Similarly, every term on the right side of (12) can be seen in Figure 1. Although $P(Y)$ is present in Figure 2 and in the fist glimpse seems to be absent from Figure 1 but by using (13) , $P(Y)$ can also be found from Figure 1 only.

Case 4:

$$Branch\ 4\ (Figure\ 1) = \ Branch\ 4\ (Figure\ 2)$$

$$P(\overline{X} \cap \overline{Y}) = P(\overline{Y} \cap \overline{X})$$

$$P(\overline{X})P(\overline{Y}|\overline{X}) = P(\overline{Y})P(\overline{X}|\overline{Y})$$

$$P\left(\overline{Y}\middle|\overline{X}\right) = \frac{P(\overline{Y})P(\overline{X}|\overline{Y})}{P(\overline{X})} \quad (14)$$

$$P\left(\overline{X}\right) = P(Y)P\left(\overline{X}\middle|Y\right) + P\left(\overline{Y}\right)P\left(\overline{X}\middle|\overline{Y}\right) \quad (15)$$

Here, $P\left(\overline{Y}\right)$ is the prior probability, $P(\overline{X}|\overline{Y})$ is the likelihood and $P\left(\overline{Y}\middle|\overline{X}\right)$ is the change or improvement in $p(\overline{Y})$ after the occurrence of $\overline{X}$. Similarly, we also have,

$$P\left(\overline{X}\middle|\overline{Y}\right) = \frac{P(\overline{X})P\left(\overline{Y}\middle|\overline{X}\right)}{P(\overline{Y})} \quad (16)$$

$$P\left(\overline{Y}\right) = P(X)P\left(\overline{Y}\middle|X\right) + P\left(\overline{X}\right)P\left(\overline{Y}\middle|\overline{X}\right) \quad (17)$$

Here $P\left(\overline{X}\right)$ is the prior probability, $P\left(\overline{Y}\middle|\overline{X}\right)$ is the likelihood and $P(\overline{X}|\overline{Y})$ is the change or improvement in $P(\overline{X})$ after the occurrence of $\overline{Y}$.

Following observations can be made from case 4.

a. In case 4 it can be seen that (14) and (16) are Bayes rules.
b. Every term on the right side of (14) can be seen in Figure 2. Although $P\left(\overline{X}\right)$ is present in Figure 1 and in the fist glimpse seems to be absent from Figure 2 but by using (15) , $P\left(\overline{X}\right)$ can also be found from Figure 2 only.
c. Similarly, every term on the right side of (16) can be seen in Figure 1. Although $P\left(\overline{Y}\right)$ is present in Figure 2 and in the fist glimpse seems to be absent from Figure 1 but by using (17) , $P\left(\overline{Y}\right)$ can also be found from Figure 1 only.

From the above cases it is interesting to note that all the equations including (3), (4), (6), (8), (10), (12), (16) and (18) are the resultant forms of Bayes' theorem. But only the equation (3) from case 1 is stated as the Bayes' theorem mathematical formula in probability, statistics and quantitative analysis text books Walpole, [17], Pishro-Nik, [18], Anderson, [19], Leon Garcia, [20], Render, [21] and Giri, [22]. Also the approaches used to teach Bayes' theorem at present (noted in the previous section) also does not show the presence of these many mathematical formulae of Bayes' theorem. Hence from the above analysis we highlight the very existance of all these different cases of Bayes' theorem, which are only possible to see if probability tree diagram approach is used to teach it. Also, it is the probabilty tree diagram which gives the ultimate complete formulations of Bayes' theorem, as shown above.

In the above analysis a simple case was considered with the existence of only two outcomes in both random experiments. But this can be extended to any level of complexity.

3. Learning Bayes' Theorem from Probability Tree Diagrams Manuscript Length

Having understood the idea of Bayes' theorem as shown in the above section, we now propose a teaching methodology of teaching the topic using the approach of probability tree diagrams. This is enumerated as follows:

i. In a given probability problem, the first thing is to identify that whether it is to be solved using Bayes' theorem or not. This can be easily done if the required probability is conditional and the exact opposite of required conditional probability is already provided. This means that if $P(X|Y)$ is required then $P(Y|X)$ would be given in the question or vice versa. Likewise if $P(\bar{X}|Y)$ is required then $P(Y|\bar{X})$ would be given or vice versa. If $P(X|\bar{Y})$ is required then $P(\bar{Y}|X)$ would be given or vice versa. If $P(\bar{X}|\bar{Y})$ is required then $P(\bar{Y}|\bar{X})$ would be given or vice versa.
ii. Now make a probability tree diagram for the given/ provided conditional probability, e.g. $P(Y|X)$. All the other probability values provided in the question would belong to the tree made here, i.e. of $P(Y|X)$ in this example. It will be noted that every probability for the tree made would be known. Therefore, we can also call this tree as "known tree" or simply the input dataset.

iii. Similarly now draw the tree for the required conditional probability, i.e. of $P(X|Y)$ in this case.
iv. Now relate the branches of both trees drawn in (b) and (c) accordingly in the same way as we did in the previous Section 2.

The resultant Bayes' theorem would be any of the equations derived as in (3), (4), (6), (8), (10), (12), (16) and (18). For the considered case it will be equation (3).

Hence following the above stated steps we can not only solve Bayes' theorem problems easily, without memorizing the formula, but also it gives more depth in learning the potential of Bayes' theorem.

4. Examples to Solve for Bayes' Theorem

In the previous section we provided the teaching methodology to use in learning and solving the Bayes' theorem in probability problems. In this section we will use two examples from simple to higher difficulty level, to see how the proposed methodology in learning and solving for Bayes' theorem can used.

Problem 1

Reference: Section 4.5, Anderson, [19].

"Consider a manufacturing firm that receives shipment of parts from two different suppliers. Let A1 denote the event that a part is from supplier 1 and A2 denote an event that a part is from supplier 2. Currently 65% of the parts purchased by the company are from supplier 1 and the remaining 35% are from supplier 2. Let G indicates good parts and B indicates bad parts. Historical data suggests that $\boldsymbol{P(G|A_1) = 98}\%$, $\boldsymbol{P(B|A_1) = 2}\%$, $\boldsymbol{P(G|A_2) = 95}\%$ and $\boldsymbol{P(B|A_2) = 5}\%$.

Given the information that the part is bad, what is the probability that it came from supplier 1 and what is the probability that it came from supplier 2."

Solution:

Data:

$$P(A_1) = 0.65$$
$$P(A_2) = 0.35$$
$$P(G|A_1) = 0.98$$
$$P(B|A_1) = 0.02$$
$$P(G|A_2) = 0.95$$
$$P(B|A_2) = 0.05$$

It is required to find: $P(A_1|B)$

Solution Steps:

a. Since $P(A_1|B)$ is required, which is the exact opposite of what is given ,i.e. $P(B|A_1)$ is given, therefore it is the problem to be solved using Bayes' theorem.
b. The given data shows that A_1 or A_2 event happened first then the event of G or B. This leads us to the following probability tree diagram in Figure 3. This is the known tree.

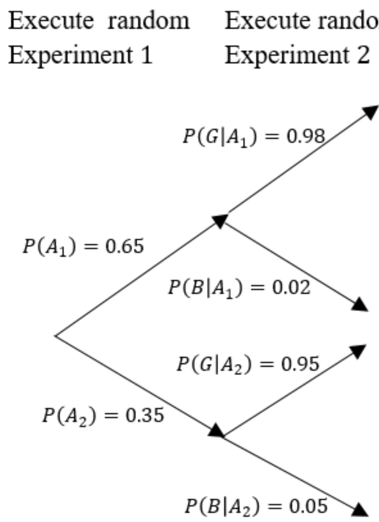

Figure 3. The known tree for problem 1.

c. The probability tree diagrm for the required conditional probabiltiy is as follows in Figure 4.

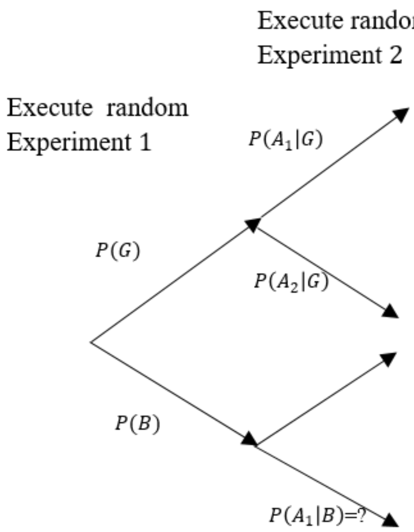

Figure 4. The unknown tree for problem

a. Equating the branches of the two trees using the methods stated in Section 2, we obtain as,

$$P(B \cap A_1) = P(A_1 \cap B)$$
$$P(B)P(A_1|B) = P(A_1)P(B|A_1)$$
$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(B)} \tag{18}$$

Using the method above:

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) \tag{19}$$

Putting the values from Figure 3 in (18) and (19), we get $P(A_1|B) = 0.42623$

Problem 2

Reference: Sedlmeir, [9].

"The probability that a woman who undergoes a mammography will have breast cancer is 1%. If a woman undergoing a mammography has breast cancer, the probability that she will test positive is 80%. If a woman undergoing a mammography does not have cancer, the probability that she will test positive is 10%. What is the probability that a woman who has undergone a mammography actually has breast cancer if she tests positive?"

Solution Steps:

Let there be following events:

C = Presence of breast cancer

$\overline{C}$ = Absence of breast cancer

T^+ = Test positive

T^- = Test negative

a. According to the problem in hand, $P(C\,|T^+)$ is required, which is the exact opposite of what is given ,i.e. $P(T^+|\,C)$ is given, therefore it is the problem to be solved using Bayes' theorem.
b. The given data shows that C or $\overline{C}$ event happened first then the event of T^+ or T^-. This leads us to the following probability tree diagram in Figure 5. This is the known tree.

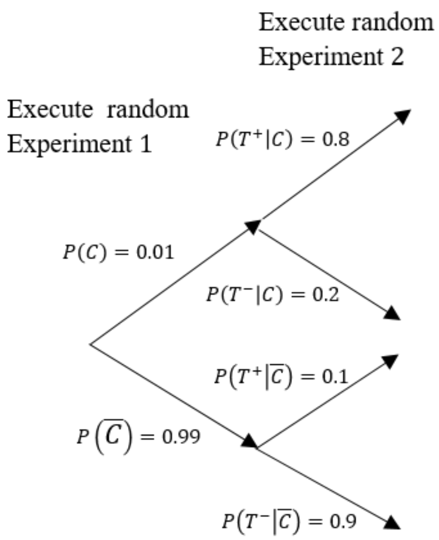

Figure 5. The known tree for problem 2

c. The probability tree diagrm for the required conditional probabiltiy is as follows in Figure 6.
d. Equating the branches of the two trees using the methods stated in Section 2, we obtain as,

$$P(C \cap T^+) = P(T^+ \cap C)$$
$$P(C)P(T^+|C) = P(T^+)P(C|T^+)$$
$$P(C|T^+) = \frac{P(C)\,P(T^+|C)}{P(T^+)} \tag{20}$$

Using the method above:

$$P(T^+) = P(C)\,P(T^+|C) + P(\overline{C})\,P(T^+|\overline{C}) \tag{21}$$

Putting the values from Figure 4in (19) and (20), we get the required conditional probability as, $P(C|T^+) = 0.0747$.

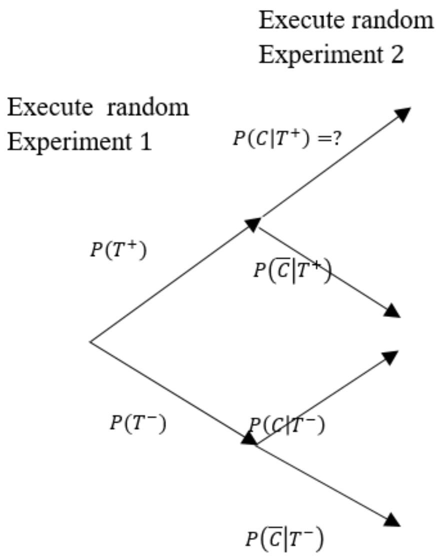

Figure 6. The unknown tree for problem 2

5. Conclusions

Through this article we have tried to improve the teaching pedagogy and thus the learning of Bayes' theorem and inferencing using the probability tree diagram approach (known and unknown trees), thus to remove any ambiguities because of non-understanding of Mathematics symbols and its typical language. This approach, in addition to develop learning has also highlighted insight to Bayes' theorem which is otherwise not prominent. We have also enumerated the steps to identify Bayes' theorem requirement for a probabilistic scenario. Because of the mentioned advantages we propose to include the stated method of teaching Bayes' theorem and its associated probabilities as a regular text book method. This will further help students understand the decision theories which are prevalently used in computer science, business statistics and econometrics involving conditional events having conditional or posterior probabilities for example in finding expected monetary values (EMV) under risk conditions, using the utility assessment theory, in predicting future states probabilities in Markov analysis and simulation and modeling in Monte Carlo simulations.

Declaration of Competing Interest

No conflict of interest was declared by the authors.

Authorship Contribution Statement

Ikram E Khuda: Writing, Reviewing, Methodology, Supervision, Data Preparation, Writing, Reviewing and Editing

References

[1] Nguyen, H. T., & Thach, N. N., "A Panorama of Applied Mathematical Problems in Economics", *Thai Journal of Mathematics*, 2018.

[2] Sahai, H., "Teaching Bayes' Theorem using examples in medical diagnosis", *Teaching Mathematics and its Applications: An International Journal of the IMA*, vol. 11(4), pp. 175-179, 1992

[3] Kolb, B. G., "Brain plasticity and behavior", *Current Directions in Psychological Science*, vol. 12(1), pp. 1-5, 2003

[4] Florens, J. P., "Elements of Bayesian Statistics", *New York: Taylor & Francis Inc*, 2019

[5] Green, K. a., "Reorganizing freshman business mathematics, I: background and philosophy", *Teaching Mathematics and its Applications. An International Journal of the IMA*, vol. 27(1), pp. 11-23, 2008.

[6] Bolstad, W. M., “Teaching Bayesian statistics to undergraduates: Who, what, where, when, why, and how”, *Proceedings of the Sixth International Conference on Teaching of Statistics*, 2002.

[7] Lecoutre, B., “Training students and researchers in Bayesian methods”, *Journal of Data Science*, vol. 4(2), pp. 207-232, 2006.

[8] Holt, C. A., “Classroom games: Understanding bayes' rule”, *Journal of Economic Perspectives*, vol. 10(2), pp. 179-187, 1996.

[9] Sedlmeier, P. &., “Teaching Bayesian reasoning in less than two hours”, *Journal of experimental psychology: general*, vol. 130(3), pp. 380, 2001.

[10] Kurzenhäuser, S. &., “Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students”, *Medical Teacher*, vol. 24(5), pp. 516-521, 2002.

[11] Gelman, A. C., “Bayesian data analysis (2nd ed.)”, *London: Chapman and Hall*, 2004.

[12] Jackman, S., “Bayesian analysis for the social sciences”, *Chichester, United Kingdom: John Wiley & Sons*, 2009

[13] Kruschke, J. K., “Bayesian estimation supersedes the t test”, *Journal of Experimental Psychology: General*, vol. 142(2), pp. 573, 2013.

[14] Faulkenberry, T. J., “A Simple Method for Teaching Bayesian Hypothesis Testing in the Brain and Behavioral Sciences”, *Journal of Undergraduate Neuroscience Education*, vol. 16(2), pp. 126-130, 2018.

[15] Rouder, J. N., “Teaching Bayes’ theorem: Strength of evidence as predictive accuracy”, *The American Statistician*, vol. 73(2), pp. 186-190, 2019.

[16] Rässler, S., “Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches”, *Springer Science & Business Media*, vol. 168, 2012.

[17] Walpole, R. E., “Essentials of Probability and Statistics”, *Pearson*, 2013.

[18] Pishro-Nik, H., “Introduction to probability, statistics, and random processes”, *Kappa Research LLC*, 2014.

[19] Anderson, D. R., “Statistics for business & economics”, *Cengage*, 2015.

[20] Leon-Garcia, A., “Probability, statistics, and random processes for electrical engineering”, *Pearson*, 2017.

[21] Render, B. S., “Quantitative analysis for management (13 ed.)”, *Pearson*, 2018.

[22] Giri, N. C., “Introduction to Probability and Statistics”, *Taylor and Francis*, 2019.

[23] Walpole, R. E., “Introduction to Statistics”, *New York: Macmillan*, 1982.