Lesson 1.5: Law of Total Probability and Bayes Theorem

Lesson Summary

In this lesson, we'll learn about a classical theorem known as Bayes' Theorem. In short, we'll want to use Bayes' Theorem to find the conditional probability of an event $P(A|B)$, say, when the "reverse" conditional probability $P(B|A)$ is the probability that is known.

Learning Outcomes

At the end of the lesson, students must be able to

1. Determine the probability of an event by using a partition of the sample space S, and
2. Apply Bayes' Theorem to find the conditional probability of an event when the "reverse" conditional probability is known.

Motivation Question

How does prior information affect the probability of occurrence of an event?

Discussion

Law of Total Probability

Consider a partition of the sample space *S*, say B_1, B_2, and B_3 (refer to the diagram below). Then, for any set *A* in *S*, we have $A \cap B_1$ (red), $A \cap B_2$ (green), and $A \cap B_3$ (blue) also forms a partition of A. That is,

$$A = (A \cap B_1) \cup (A \cap B_3) \cup (A \cap B_3).$$

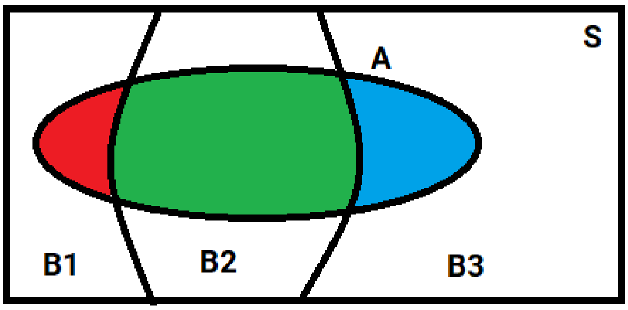

Figure 5.1. Partitions of *S* and $A \subset S$.

Therefore, by Axiom 3, we have

$$P(A)=P(A\cap B_1)+P(A\cap B_2)+P(A\cap B_3).$$

Then using the multiplication rule, we get

$$P(A)=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+P(A|B_3)P(B_3).$$

This is called the **Law of Total Probability**.

In general, if B_1, B_2, B_3, …, B_k form a partition of S, then for any set A in S,

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

Example 5.1:

Suppose there are three bags each containing 100 marbles: Bag 1 has 75 red and 25 blue marbles; Bag 2 has 60 red and 40 blue marbles; and Bag 3 has 45 red and 55 blue marbles. I choose one of the bags at random and then pick a marble from the chosen bag, also at random. What is the probability that the chosen marble is red?

Answer:

There are two activities involved in here. The first one is selecting a bag at random and the second is selecting a red marble from the chosen bag.

Let B_1 be the event that the first bag is selected, B_2 be the event that the second bag is selected, and B_3 be the event that third bag is selected. Further, let R be the event that the selected marble is red. Assuming the three bags are equally likely to be selected, then

$$P(B_i)=\frac{1}{3},\ i=1,2,3.$$

Now, the chance of selecting a red marble depends which bag is selected. For example, if the first bag is selected then, $P(R|B_1)=\frac{75}{100}$. The probabilities of getting a red marble are $P(R|B_2)=\frac{60}{100}$ and $P(R|B_3)=\frac{45}{100}$, respectively for the 2nd and 3rd bags.

Therefore, using the Law of Total Probability, the probability that a red marble is selected is

$$\begin{aligned}P(R)&=P(R|B_1)P(B_1)+P(R|B_2)P(B_2)+P(R|B_3)P(B_3)\\&=\frac{75}{100}\cdot\frac{1}{3}+\frac{60}{100}\cdot\frac{1}{3}+\frac{45}{100}\cdot\frac{1}{3}\\&=\frac{60}{100}\end{aligned}$$

Bayes' Theorem

Now we are ready to state one of the most useful results in conditional probability: *Bayes' rule*. Suppose that we know *P*(*A*|*B*), but we are interested in the probability *P*(*B*|*A*). Using the definition of conditional probability, we have

$$P(A \mid B)P(B) = P(B \mid A)P(A).$$

If divide both sides by $P(A)$, we get

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

which is the famous Bayes' rule. In this formulation, we call $P(B \mid A)$ as the "posterior" probability and $P(B)$ as the "prior" probability.

Often, in order to find *P*(*A*) in Bayes' formula we need to use the law of total probability. We state formally the Bayes' theorem as follows.

Suppose *A* is an event in *S* and suppose that B_1, B_2, B_3, …, B_k forms a partition of *S*. Then

$$P(B_j \mid A) = \frac{P(A \mid B_j)P(B_j)}{\sum_{i=1}^{k} P(A \mid B_i)P(B_i)}.$$

Example 5.2:

Suppose in Example 5.1 we observed that the chosen marble is red. What is the probability that Bag 1 was chosen?

Answer:

Using the same notation, we are interested to compute

$$\begin{aligned} P(B_1 \mid R) &= \frac{P(R \mid B_1)P(B_1)}{\sum_{i=1}^{3} P(A \mid B_i)P(B_i)} \\ &= \frac{\frac{75}{100} \cdot \frac{1}{3}}{\frac{75}{100} \cdot \frac{1}{3} + \frac{60}{100} \cdot \frac{1}{3} + \frac{45}{100} \cdot \frac{1}{3}} \\ &= \frac{5}{12} \end{aligned}$$

Example 5.3:

In screening for a certain disease, the probability that a healthy person wrongly gets a positive result is 0.05. The probability that a diseased person wrongly gets a negative result is 0.002. The overall rate of the disease in the population being screened is 1%. If Axel's test gives a positive result, what is the probability he actually has the disease?

Answer:

Define the following events:

D = {have disease}; D^c = {do not have the disease};

T^+ = {positive test}; T^- = {negative test}

We are given these information:

False positive rate is 0.05 ⇒ $P(T^+|D^c) = 0.05$;

False negative rate is 0.002 ⇒ $P(T^-|D) = 0.002$

Disease rate is 1% ⇒ $P(D) = 0.01$.

We wanted to compute the probability that Axel actually has the disease given he gets a positive test result. In other words, we wanted to compute

$$P\left(D \mid T^+\right) = \frac{P\left(T^+ \mid D\right)P(D)}{P\left(T^+\right)}.$$

We use the Law of Total Probability to calculate $P\left(T^+\right)$ as follows:

$$P\left(T^+\right) = P\left(T^+ \mid D\right)P(D) + P\left(T^+ \mid D^c\right)P\left(D^c\right)$$

From the given probabilities, we have $P\left(T^+ \mid D\right) = 1 - P\left(T^- \mid D\right) = 1 - 0.002 = 0.998$, Thus,

$$\begin{aligned} P\left(T^+\right) &= P\left(T^+ \mid D\right)P(D) + P\left(T^+ \mid D^c\right)P\left(D^c\right) \\ &= 0.998 \times 0.01 + 0.05 \times 0.99 \\ &= 0.05948 \end{aligned}$$

Therefore,

$$\begin{aligned} P\left(D \mid T^+\right) &= \frac{P\left(T^+ \mid D\right)P(D)}{P\left(T^+\right)} \\ &= \frac{0.998 \times 0.01}{0.05948} \\ &\cong 0.168 \end{aligned}.$$

Learning Tasks/Activities

Answer the following as indicated.

1. A laboratory test is 95% effective at detecting a disease when itis present. It is 99% effective at declaring a subject negative when the subject is truly negative for the disease. Suppose 8% of the population has the disease.
    a. What is the probability a randomly selected subject will test positively?

   b. What is the probability a subject has the disease if his test is positive?

2. Suppose a statistics class contains 70% male and 30% female students. It is known that in a test, 5% of males and 10% of females got an “A” grade. If one student from this class is randomly selected and observed to have an “A” grade, what is the probability that this is a male student?

3. Of the travelers arriving at a small airport, 60% fly on major airlines, 30% fly on privately owned planes, and the remainder fly on commercially owned planes not belonging to a major airline. Of those traveling on major airlines, 50% are traveling for business reasons, whereas 60% of those arriving on private planes and 90% of those arriving on other commercially owned planes are traveling for business reasons. Suppose that we randomly select one person arriving at this airport. What is the probability that the person
   a. is traveling on business?
   b. arrived on a privately owned plane, given that the person is traveling for business reasons?
   c. is traveling on business, given that the person is flying on a commercially owned plane?

Assessment

Answer the following as indicated.

1. At the Visayas State University 35%, 25% and 40% of the students in STAT 21 are taught by Prof. X, Prof. Y and Prof. Z, respectively. The respective probabilities of failing in STAT 21 for the three professors are 6%, 7% and 5%. A student in STAT 21 in the previous semester was randomly selected and was found out to have failed in the subject. Under which class most likely did he belong?

Instructions on how to submit student output

Write your answers of the *Learning Tasks/Activities* and *Assessment* on a clean sheet of paper using a blue pen and take a picture of your answer sheet with your name on it and send to bertmilla@vsu.edu.ph or to our FB chat group Stat 121.

Submit first your answers of the *Learning Tasks/Activities* and wait for my feedback before you submit your answers of the *Assessment*.