Bayes' Theorem
by Mario F. Triola

The concept of *conditional probability* is introduced in *Elementary Statistics*. We noted that the conditional probability of an event is a probability obtained with the additional information that some other event has already occurred. We used $P(B|A)$ to denoted the conditional probability of event B occurring, given that event A has already occurred. The following formula was provided for finding $P(B|A)$:

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

In addition to the above formal rule, the textbook also included this "intuitive approach for finding a conditional probability":

> The conditional probability of B given A can be found by assuming that event A has occurred and, working under that assumption, calculating the probability that event B will occur.

In this section we extend the discussion of conditional probability to include applications of *Bayes' theorem* (or *Bayes' rule*), which we use for revising a probability value based on additional information that is later obtained. One key to understanding the essence of Bayes' theorem is to recognize that we are dealing with *sequential* events, whereby new additional information is obtained for a subsequent event, and that new information is used to revise the probability of the initial event. In this context, the terms *prior probability* and *posterior probability* are commonly used.

Definitions
A **prior probability** is an initial probability value originally obtained before any additional information is obtained.
A **posterior probability** is a probability value that has been revised by using additional information that is later obtained.

Example 1
The Gallup organization randomly selects an adult American for a survey about credit card usage. Use subjective probabilities to estimate the following.

a.   What is the probability that the selected subject is a male?

b.   After selecting a subject, it is later learned that this person was smoking a cigar during the interview. What is the probability that the selected subject is a male?

c.   Which of the preceding two results is a prior probability? Which is a posterior probability?

Solution

a.   Roughly half of all Americans are males, so we estimate the probability of selecting a male subject to be 0.5. Denoting a male by M, we can express this probability as follows: $P(M) = 0.5$.

b.   Although some women smoke cigars, the vast majority of cigar smokers are males. A reasonable guess is that 85% of cigar smokers are males. Based on this additional subsequent information that the survey respondent was smoking a cigar, we estimate the probability of this person being a male as 0.85. Denoting a male by M and denoting a cigar smoker by C, we can express this result as follows: $P(M \mid C) = 0.85$.

c.   In part (a), the value of 0.5 is the initial probability, so we refer to it as the prior probability. Because the probability of 0.85 in part (b) is a revised probability based on the additional information that the survey subject was smoking a cigar, this value of 0.85 is referred to a posterior probability.

The Reverend Thomas Bayes [1701 (approximately) – 1761] was an English minister and mathematician. Although none of his work was published during his lifetime, later (posterior?) publications included the following theorem (or rule) that he developed for determining probabilities of events by incorporating information about subsequent events.

Bayes' Theorem

The probability of event *A,* given that event *B* has subsequently occurred, is

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{[P(A) \cdot P(B|A)] + [P(\overline{A}) \cdot P(B|\overline{A})]}$$

That's a formidable expression, but we will simplify its calculation. See the following example, which illustrates use of the above expression, but also see the alternative method based on a more intuitive application of Bayes' theorem.

Example 2

In Orange County, 51% of the adults are males. (It doesn't take too much advanced mathematics to deduce that the other 49% are females.) One adult is randomly selected for a survey involving credit card usage.

a.   Find the prior probability that the selected person is a male.

b.   It is later learned that the selected survey subject was smoking a cigar. Also, 9.5% of males smoke cigars, whereas 1.7% of females smoke cigars (based on data from the Substance Abuse and Mental Health Services Administration). Use this additional information to find the probability that the selected subject is a male.

Solution

Let's use the following notation:

      M = male $\overline{M}$ = female (or not male)

      C = cigar smoker $\overline{C}$ = not a cigar smoker.

a.   Before using the information given in part b, we know only that 51% of the adults in Orange County are males, so the probability of randomly selecting an adult and getting a male is given by $P(M) = 0.51$.

b.   Based on the additional given information, we have the following:

         $P(M) = 0.51$ because 51% of the adults are males

         $P(\overline{M}) = 0.49$ because 49% of the adults are females (not males)

         $P(C|M) = 0.095$ because 9.5% of the males smoke cigars (That is, the probability of getting someone who smokes cigars, given that the person is a male, is 0.095.)

         $P(C|\overline{M}) = 0.017$.   because 1.7% of the females smoke cigars (That is, the probability of getting someone who smokes cigars, given that the person is a female, is 0.017.)

Let's now apply Bayes' theorem by using the preceding formula with M in place of A, and C in place of B. We get the following result:

$$P(M \mid C) = \frac{P(M) \cdot P(C|M)}{[P(M) \cdot P(C|M)] + [P(\overline{M}) \cdot P(C|\overline{M})]}$$

$$= \frac{0.51 \cdot 0.095}{[0.51 \cdot 0.095] + [0.49 \cdot 0.017]}$$

$$= 0.85329341$$

$$= 0.853 \text{ (rounded)}$$

Before we knew that the survey subject smoked a cigar, there is a 0.51 probability that the survey subject is male (because 51% of the adults in Orange County are males). However, after learning that the subject smoked a cigar, we revised the probability to 0.853. There is a 0.853 probability that the cigar–smoking respondent is a male. This makes sense, because the likelihood of a male increases dramatically with the additional information that the subject smokes cigars (because so many more males smoke cigars than females).

Intuitive Bayes Theorem

The preceding solution illustrates the application of Bayes' theorem with its calculation using the formula. Unfortunately, that calculation is complicated enough to create an abundance of opportunities for errors and/or incorrect substitution of the involved probability values. Fortunately, here is another approach that is much more intuitive and easier:

Assume some convenient value for the total of all items involved, then construct a table of rows and columns with the individual cell frequencies based on the known probabilities.

For the preceding example, simply assume some value for the adult population of Orange County, such as 100,000, then use the given information to construct a table, such as the one shown below.

Finding the number of males who smoke cigars: If 51% of the 100,000 adults are males, then there are 51,000 males. If 9.5% of the males smoke cigars, then the number of cigar–smoking males is 9.5% of 51,000, or $0.095 \times 51,000 = 4845$. See the entry of 4845 in the table. The other males who do *not* smoke cigars must be $51,000 - 4845 = 46,155$. See the value of 46,155 in the table.

Finding the number of females who smoke cigars: Using similar reasoning, 49% of the 100,000 adults are females, so the number of females is 49,000. Given that 1.7% of the females smoke cigars, the number of cigar–smoking females is $0.017 \times 49,000 = 833$. The number of females who do *not* smoke cigars is $49,000 - 833 = 48,167$. See the entries of 833 and 48,167 in the table.

	C (Cigar Smoker)	$\overline{C}$ (Not a Cigar Smoker)	Total
M (male)	4845	46,155	**51,000**
$\overline{M}$ (female)	833	48,167	**49,000**
Total	**5678**	**94,322**	**100,000**

The above table involves relatively simple arithmetic. Simply partition the assumed population into the different cell categories by finding suitable percentages.

Now we can easily address the key question as follows: To find the probability of getting a male subject, given that the subject smokes cigars, simply use the same conditional probability described in the textbook. To find the probability of getting a male given that the subject smokes, restrict the table to the column of cigar smokers, then find the probability of getting a male in that column. Among the 5678 cigar smokers, there are 4845 males, so the probability we seek is $4845/5678 = 0.85329341$. That is, $P(M \mid C) = 4845/5678 = 0.85329341 = 0.853$ (rounded).

Bayes' Theorem Generalized

The preceding formula for Bayes' theorem and the preceding example use exactly two categories for event *A* (male and female), but the formula can be extended to include more than two categories. The following example illustrates this extension and it also illustrates a practical application of Bayes' theorem to quality control in industry. When dealing with more than the two events of A and $\overline{A}$, we must be sure that the multiple events satisfy two important conditions:

1.   The events must be *disjoint* (with no overlapping).

2.   The events must be *exhaustive*, which means that they combine to include all possibilities.

Example 3

An aircraft emergency locator transmitter (ELT) is a device designed to transmit a signal in the case of a crash. The Altigauge Manufacturing Company makes 80% of the ELTs, the Bryant Company makes 15% of them, and the Chartair Company makes the other 5%. The ELTs made by Altigauge have a 4% rate of defects, the Bryant ELTs have a 6% rate of defects, and the Chartair ELTs have a 9% rate of defects (which helps to explain why Chartair has the lowest market share).

a.   If an ELT is randomly selected from the general population of all ELTs, find the probability that it was made by the Altigauge Manufacturing Company.

b.   If a randomly selected ELT is then tested and is found to be defective, find the probability that it was made by the Altigauge Manufacturing Company.

Solution

We use the following notation:

A = ELT manufactured by Altigauge
B = ELT manufactured by Bryant
C = ELT manufactured by Chartair

D = ELT is defective
$\overline{D}$ = ELT is not defective (or it is good)

a.   If an ELT is randomly selected from the general population of all ELTs, the probability that it was made by Altigauge is 0.8 (because Altigauge manufactures 80% of them).

b.   If we now have the additional information that the ELT was tested and was found to be defective, we want to revise the probability from part (a) so that the new information can be used. We want to find the value of *P*(A|D), which is the probability that the ELT was made by the Altigauge company given that it is defective. Based on the given information, we know these probabilities:

$P(A) = 0.80$ because Altigauge makes 80% of the ELTs
$P(B) = 0.15$ because Bryant makes 15% of the ELTs
$P(C) = 0.05$ because Chartair makes 5% of the ELTs

$P(D|A) = 0.04$ because 4% of the Altigauge ELTs are defective
$P(D|B) = 0.06$ because 6% of the Bryant ELTs are defective
$P(D|C) = 0.09$ because 9% of the Chartair ELTs are defective

Here is Bayes' theorem extended to include three events corresponding to the selection of ELTs from the three manufacturers (A, B, C):

$$P(A|D) = \frac{P(A) \cdot P(D|A)}{[P(A) \cdot P(D|A)] + [P(B) \cdot P(D|B)] + [P(C) \cdot P(D|C)]}$$

$$= \frac{0.80 \cdot 0.04}{[0.80 \cdot 0.04] + [0.15 \cdot 0.06] + [0.05 \cdot 0.09]}$$

$$= 0.703 \text{ (rounded)}$$

Intuitive Baye's Theorem: Now let's find $P(A|D)$ by using a table. Let's arbitrarily assume that 10,000 ELTs were manufactured. (The solution doesn't depend on the number selected, but it's helpful to select a number large enough so that the cells in the table are all whole numbers.) Because 80% of the ELTs are made by Altigauge, we have 8000 ELTs made by Altigauge, and 4% of them (or 320) are defective. Also, if 320 of the Altigauge ELTs are defective, the other 7680 are not defective. See the values of 320 and 7680 in the table below. The other values are found using the same reasoning.

	D (defective)	$\overline{D}$ (not defective)	Total
A (Altigauge)	320	7680	**8,000**
B (Bryant)	90	1410	**1,500**
C (Chartair)	45	455	**500**
Total	**455**	**9545**	**10,000**

We want to find the probability that an ELT was made by Altigauge, given that it is known to be defective. Because we know the condition that the ELT is defective, we can refer to the first column of values where we see that among the 455 total defective ELTs, 320 were made by Altigauge, so that the probability is $320/455 = 0.703$ (rounded). This is the same result obtained with the formula from Bayes' theorem.

The preceding example involve an extension of Bayes' theorem to three events denoted by A, B, C. Based on the format of the formula used in the solution, it is easy to extend Bayes' theorem so that it can be used with four or more events. (See Exercises 11 and 12.)

Exercises

Pregnancy Test Results. *In Exercises 1 and 2, refer to the results summarized in the table below.*

	Positive Test Result (Pregnancy is indicated)	Negative Test Result (Pregnancy is not indicated)
Subject is Pregnant	80	5
Subject is Not Pregnant	3	11

1. **a.** If one of the 99 test subjects is randomly selected, what is the probability of getting a subject who is pregnant?
   b. A test subject is randomly selected and is given a pregnancy test. What is the probability of getting a subject who is pregnant, given that the test result is positive?

2. **a.** One of the 99 test subjects is randomly selected. What is the probability of getting a subject who is not pregnant?
   b. A test subject is randomly selected and is given a pregnancy test. What is the probability of getting a subject who is not pregnant, given that the test result is negative?

3. ***Survey Results*** In Orange County, 51% of the adults are males. One adult is randomly selected for a survey involving credit card usage. (See Example 2 in this section.)
   a. Find the prior probability that the selected person is a female.
   b. It is later learned that the selected survey subject was smoking a cigar. Also, 9.5% of males smoke cigars, whereas 1.7% of females smoke cigars (based on data from the Substance Abuse and Mental Health Services Administration). Use this additional information to find the probability that the selected subject is a female.

4. ***Emergency Locator Transmitters*** An aircraft emergency locator transmitter (ELT) is a device designed to transmit a signal in the case of a crash. The Altigauge Manufacturing Company makes 80% of the ELTs, the Bryant Company makes 15% of them, and the Chartair Company makes the other 5%. The ELTs made by Altigauge have a 4% rate of defects, the Bryant ELTs have a 6% rate of defects, and the Chartair ELTs have a 9% rate of defects. (These are the same results from Example 3 in this section.)
   a. Find the probability of randomly selecting an ELT and getting one manufactured by the Bryant Company.
   b. If an ELT is randomly selected and tested, find the probability that it was manufactured by the Bryant Company if the test indicates that the ELT is defective.

5. **_Emergency Locator Transmitters_** Use the same ELT data from Exercise 4.
   a. Find the probability of randomly selecting an ELT and getting one manufactured by the Chartair Company.
   b. An ELT is randomly selected and tested. If the test indicates that the ELT is defective, find the probability that it was manufactured by the Chartair Company.

6. **_Emergency Locator Transmitters_** Use the same ELT data from Exercise 4. An ELT is randomly selected and tested. If the test indicates that the ELT is _not_ defective, find the probability that it is from the Altigauge Company.

7. **_Pleas and Sentences_** In a study of pleas and prison sentences, it is found that 45% of the subjects studied were sent to prison. Among those sent to prison, 40% chose to plead guilty. Among those not sent to prison, 55% chose to plead guilty.
   a. If one of the study subjects is randomly selected, find the probability of getting someone who was not sent to prison.
   b. If a study subject is randomly selected and it is then found that the subject entered a guilty plea, find the probability that this person was not sent to prison.

8. **_Pleas and Sentences_** Use the same data given in Exercise 7.
   a. If one of the study subjects is randomly selected, find the probability of getting someone who was sent to prison.
   b. If a study subject is randomly selected and it is then found that the subject entered a guilty plea, find the probability that this person was sent to prison.

9. **_HIV_** The New York State Health Department reports a 10% rate of the HIV virus for the "at-risk" population. Under certain conditions, a preliminary screening test for the HIV virus is correct 95% of the time. (Subjects are not told that they are HIV infected until additional tests verify the results.) If someone is randomly selected from the at-risk population, what is the probability that they have the HIV virus if it is known that they have tested positive in the initial screening?

10. **_HIV_** Use the same data from Exercise 9. If someone is randomly selected from the at-risk population, what is the probability that they have the HIV virus if it is known that they have tested negative in the initial screening?

11. **_Extending Bayes' Theorem_** Example 3 in this section included an extension of Bayes' theorem to include three events, denoted by A, B, C. Write an expression that extends Bayes' theorem so that it can be used to find $P(A|Z)$, assuming that the initial event can occur in one of four ways: A, B, C, D.

12. **_Extensions of Bayes' Theorem_** In Example 2, we used only the initial events of A and $\overline{A}$. In Example 3, we used initial events of A, B, and C. If events B and C in Example 3 are combined and denoted as $\overline{A}$, we can find $P(A|D)$ using the simpler format of Bayes' theorem given in Example 2. How would the resulting value of $P(A|D)$ in Example 3 be affected by using this simplified approach?

Answers to Odd–Numbered Exercises

1. a. 85/99 or 0.859

    b. 80/83 or 0.964

3. a. 0.49

    b. 0.147

5. a. 0.05

    b. 0.0989

7. a. 0.55

    b. 0.627

9. 0.679

11. $P(A|Z) = \dfrac{P(A) \cdot P(Z|A)}{[P(A) \cdot P(Z|A)] + [P(B) \cdot P(Z|B)] + [P(C) \cdot P(Z|C)] + [P(D) \cdot P(Z|D)]}$