APPENDIX I ADDITIONAL TOPICS

PART I

Bayes's Theorem

The Reverend Thomas Bayes (1702–1761) was an English mathematician who discovered an important relation for conditional probabilities. This relation is referred to as *Bayes's rule* or *Bayes's theorem*. It uses conditional probabilities to adjust calculations so that we can accommodate new relevant information. We will restrict our attention to a special case of Bayes's theorem in which an event B is partitioned into only *two* mutually exclusive events (see Figure AI-1). The general formula is a bit complicated but is a straightforward extension of the basic ideas we will present here. Most advanced texts contain such an extension.

Note: We use the following compact notation in the statement of Bayes's theorem:

Notation	Meaning
A^c	complement of *A; not A*
$P(B\|A)$	probability of event *B, given* event *A; P(B, given A)*
$P(B\|A^c)$	probability of event *B, given* the complement of *A; P(B, given not A)*

We will use Figure AI-1 to motivate Bayes's theorem. Let A and B be events in a sample space that have probabilities not equal to 0 or 1. Let A^c be the complement of A.

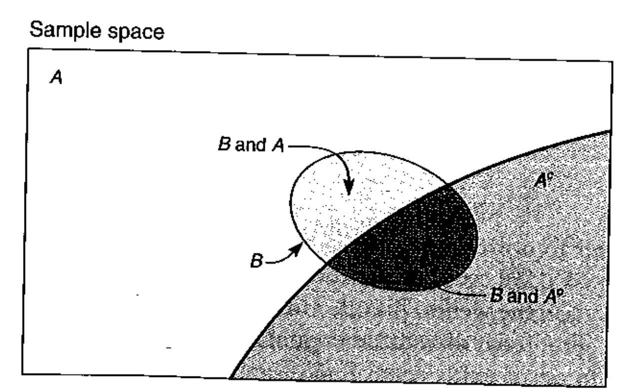
Here is Bayes's theorem: $$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \quad (1)$$

Overview of Bayes's Theorem

Suppose we have an event A and we calculate $P(A)$, the unconditional probability of A standing by itself. Now suppose we have a "new" event B and we know the probability of B given that A occurs $P(B|A)$, as well as the probability of B given that A does not occur $P(B|A^c)$. Where does such an event B come from? The event B can be constructed in many possible ways. For example, B can be constructed as

FIGURE AI-1

A Typical Setup for Bayes's Theorem



the result of a consulting service, a testing procedure, or a sorting activity. In the examples and problems, you will find more ways to construct such an event B.

How can we use this "new" information concerning the event B to adjust our calculation of the probability of event A, given B? That is, how can we make our calculation of the probability of A more realistic by including information about the event B? The answer is that we will use Equation (1) of Bayes's theorem.

Let's look at some examples that use Equation (1) of Bayes's theorem. We are grateful to personal friends in the oil and natural gas business in Colorado who provided the basic information in the following example.

EXAMPLE 1 BAYES'S THEOREM

A geologist has examined seismic data and other geologic formations in the vicinity of a proposed site for an oil well. Based on this information, the geologist reports a 65% chance of finding oil. The oil company decides to go ahead and start drilling. As the drilling progresses, sample cores are taken from the well and studied by the geologist. These sample cores have a history of predicting oil *when there is oil* about 85% of the time. However, about 6% of the time the sample cores will predict oil *when there is no oil*. (Note that these probabilities need not add up to 1.) Our geologist is delighted because the sample cores predict oil for this well.

Use the "new" information from the sample cores to revise the geologist's original probability that the well will hit oil. What is the new probability?

SOLUTION: To use Bayes's theorem, we need to identify the events A and B. Then we need to find $P(A)$, $P(A^c)$, $P(B|A)$, and $P(B|A^c)$. From the description of the problem, we have

A is the event that the well strikes oil.

A^c is the event that the well is dry (no oil).

B is the event that the core samples indicate oil.

Again, from the description, we have

$$P(A) = 0.65, \quad \text{so} \quad P(A^c) = 1 - 0.65 = 0.35$$

These are our *prior* (before new information) probabilities. New information comes from the sample cores. Probabilities associated with the new information are

$$P(B|A) = 0.85$$

This is the probability that core samples indicate oil when there actually is oil.

$$P(B|A^c) = 0.06$$

This is the probability that core samples indicate oil when there is no oil (dry well).

Now we use Bayes's theorem to revise the probability that the well will hit oil based on the "new" information from core samples. The revised probability is the *posterior* probability we compute that uses the new information from the sample cores:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{(0.85)(0.65)}{(0.85)(0.65) + (0.06)(0.35)} = 0.9634$$

We see that the revised (*posterior*) probability indicates about a 96% chance for the well to hit oil. This is why sample cores that are good can attract money in the form of venture capital (for independent drillers) on a big, expensive well.

the result of a consulting service, a testing procedure, or a sorting activity. In the examples and problems, you will find more ways to construct such an event B.

How can we use this "new" information concerning the event B to adjust our calculation of the probability of event A, given B? That is, how can we make our calculation of the probability of A more realistic by including information about the event B? The answer is that we will use Equation (1) of Bayes's theorem.

Let's look at some examples that use Equation (1) of Bayes's theorem. We are grateful to personal friends in the oil and natural gas business in Colorado who provided the basic information in the following example.

EXAMPLE 1 BAYES'S THEOREM

A geologist has examined seismic data and other geologic formations in the vicinity of a proposed site for an oil well. Based on this information, the geologist reports a 65% chance of finding oil. The oil company decides to go ahead and start drilling. As the drilling progresses, sample cores are taken from the well and studied by the geologist. These sample cores have a history of predicting oil *when there is oil* about 85% of the time. However, about 6% of the time the sample cores will predict oil *when there is no oil*. (Note that these probabilities need not add up to 1.) Our geologist is delighted because the sample cores predict oil for this well.

Use the "new" information from the sample cores to revise the geologist's original probability that the well will hit oil. What is the new probability?

SOLUTION: To use Bayes's theorem, we need to identify the events A and B. Then we need to find $P(A)$, $P(A^c)$, $P(B|A)$, and $P(B|A^c)$. From the description of the problem, we have

A is the event that the well strikes oil.

A^c is the event that the well is dry (no oil).

B is the event that the core samples indicate oil.

Again, from the description, we have

$$P(A) = 0.65, \quad \text{so} \quad P(A^c) = 1 - 0.65 = 0.35$$

These are our *prior* (before new information) probabilities. New information comes from the sample cores. Probabilities associated with the new information are

$$P(B|A) = 0.85$$

This is the probability that core samples indicate oil when there actually is oil.

$$P(B|A^c) = 0.06$$

This is the probability that core samples indicate oil when there is no oil (dry well).

Now we use Bayes's theorem to revise the probability that the well will hit oil based on the "new" information from core samples. The revised probability is the *posterior* probability we compute that uses the new information from the sample cores:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{(0.85)(0.65)}{(0.85)(0.65) + (0.06)(0.35)} = 0.9634$$

We see that the revised (*posterior*) probability indicates about a 96% chance for the well to hit oil. This is why sample cores that are good can attract money in the form of venture capital (for independent drillers) on a big, expensive well.

GUIDED EXERCISE 1 **Bayes's theorem**

The Anasazi were prehistoric pueblo people who lived in what is now the southwestern United States. Mesa Verde, Pecos Pueblo, and Chaco Canyon are beautiful national parks and monuments, but long ago they were home to many Anasazi. In prehistoric times, there were several Anasazi migrations, until finally their pueblo homes were completely abandoned. The delightful book *Proceedings of the Anasazi Symposium, 1981,* published by Mesa Verde Museum Association, contains a very interesting discussion about methods anthropologists use to (approximately) date Anasazi objects. There are two popular ways. One is to compare environmental data to other objects of known dates. The other is radioactive carbon dating.

Carbon dating has some variability in its accuracy, depending on how far back in time the age estimate goes and also on the condition of the specimen itself. Suppose experience has shown that the carbon method is correct 75% of the time it is used on an object from a known (given) time period. However, there is a 10% chance that the carbon method will predict that an object is from a certain period when we know the object is not from that period.

Using environmental data, an anthropologist reported the probability to be 40% that a fossilized deer bone bracelet was from a certain Anasazi migration period. Then, as a follow-up study, the carbon method also indicated that the bracelet was from this migration period. How can the anthropologist adjust her estimated probability to include the "new" information from the carbon dating?

(a) To use Bayes's theorem, we must identify the events A and B. From the description of the problem, what are A and B?

→ A is the event that the bracelet is from the given migration period. B is the event that carbon dating indicates that the bracelet is from the given migration period.

(b) Find $P(A)$, $P(A^c)$, $P(B|A)$, and $P(B|A^c)$.

→ From the description,

$$P(A) = 0.40$$
$$P(A^c) = 0.60$$
$$P(B|A) = 0.75$$
$$P(B|A^c) = 0.10$$

(c) Compute $P(A|B)$, and explain the meaning of this number.

→ Using Bayes's theorem and the results of part (b), we have

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$
$$= \frac{(0.75)(0.40)}{(0.75)(0.40) + (0.10)(0.60)} = 0.8333$$

The prior (before carbon dating) probability was only 40%. However, the carbon dating enabled us to revise this probability to 83%. Thus, we are about 83% sure that the bracelet came from the given migration period. Perhaps additional research at the site will uncover more information to which Bayes's theorem could be applied again.

The next example is a classic application of Bayes's theorem. Suppose we are faced with two competing hypotheses. Each hypothesis claims to explain the same phenomenon; however, only one hypothesis can be correct. Which hypothesis should we accept? This situation occurs in the natural sciences, the social sciences, medicine, finance, and many other areas of life. Bayes's theorem will help us compute the probabilities that one or the other hypothesis is correct. Then what do we do? Well, the great mathematician and philosopher René Descartes can guide us. Descartes once said, "When it is not in our power to determine what is true, we ought to follow what is most probable." Just knowing probabilities does not allow us with absolute certainty to choose the correct hypothesis, but it does permit us to identify which hypothesis is *most likely* to be correct.

EXAMPLE 2 COMPETING HYPOTHESES

A large hospital uses two medical labs for blood work, biopsies, throat cultures, and other medical tests. Lab I does 60% of the reports. The other 40% of the reports are done by Lab II. Based on long experience, it is known that about 10% of the reports from Lab I contain errors and about 7% of the reports from Lab II contain errors. The hospital recently received a lab report that, through additional medical work, was revealed to be incorrect. One hypothesis is that the report with the mistake came from Lab I. The competing hypothesis is that the report with the mistake came from Lab II. Which lab do you suspect is the culprit? Why?

SOLUTION: Let's use the following notation.

A = event report is from Lab I

A^c = event report is from Lab II

B = event report contains a mistake

From the information given,

$$P(A) = 0.60 \qquad P(A^c) = 0.40$$
$$P(B|A) = 0.10 \qquad P(B|A^c) = 0.07$$

The probability that the report is from Lab I *given* we have a mistake is $P(A|B)$. Using Bayes's theorem, we get

$$\begin{aligned} P(A|B) &= \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)} \\ &= \frac{(0.10)(0.60)}{(0.10)(0.60) + (0.07)(0.40)} \\ &= \frac{0.06}{0.088} \approx 0.682 \approx 68\% \end{aligned}$$

So, the probability is about 68% that Lab I supplied the report with the error. It follows that the probability is about 100% − 68% = 32% that the erroneous report came from Lab II.

PROBLEM BAYES'S THEOREM APPLIED TO QUALITY CONTROL

A company that makes steel bolts knows from long experience that about 12% of its bolts are defective. If the company simply ships all bolts that it produces, then 12% of the shipment the customer receives will be defective. To decrease the percentage of defective bolts shipped to customers, an electronic scanner is installed. The scanner is positioned over the production line and is supposed to pick out the good bolts. However, the scanner itself is not perfect. To test the scanner, a large number of (pretested) "good" bolts were run under the scanner, and it accepted 90% of the bolts as good.

Continued

us compute the probabilities that one or the other hypothesis is correct. Then what do we do? Well, the great mathematician and philosopher René Descartes can guide us. Descartes once said, "When it is not in our power to determine what is true, we ought to follow what is most probable." Just knowing probabilities does not allow us with absolute certainty to choose the correct hypothesis, but it does permit us to identify which hypothesis is *most likely* to be correct.

EXAMPLE 2 COMPETING HYPOTHESES

A large hospital uses two medical labs for blood work, biopsies, throat cultures, and other medical tests. Lab I does 60% of the reports. The other 40% of the reports are done by Lab II. Based on long experience, it is known that about 10% of the reports from Lab I contain errors and about 7% of the reports from Lab II contain errors. The hospital recently received a lab report that, through additional medical work, was revealed to be incorrect. One hypothesis is that the report with the mistake came from Lab I. The competing hypothesis is that the report with the mistake came from Lab II. Which lab do you suspect is the culprit? Why?

SOLUTION: Let's use the following notation.

A = event report is from Lab I

A^c = event report is from Lab II

B = event report contains a mistake

From the information given,

$$P(A) = 0.60 \qquad P(A^c) = 0.40$$

$$P(B|A) = 0.10 \qquad P(B|A^c) = 0.07$$

The probability that the report is from Lab I *given* we have a mistake is $P(A|B)$. Using Bayes's theorem, we get

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}$$

$$= \frac{(0.10)(0.60)}{(0.10)(0.60) + (0.07)(0.40)}$$

$$= \frac{0.06}{0.088} \approx 0.682 \approx 68\%$$

So, the probability is about 68% that Lab I supplied the report with the error. It follows that the probability is about 100% − 68% = 32% that the erroneous report came from Lab II.

PROBLEM BAYES'S THEOREM APPLIED TO QUALITY CONTROL

A company that makes steel bolts knows from long experience that about 12% of its bolts are defective. If the company simply ships all bolts that it produces, then 12% of the shipment the customer receives will be defective. To decrease the percentage of defective bolts shipped to customers, an electronic scanner is installed. The scanner is positioned over the production line and is supposed to pick out the good bolts. However, the scanner itself is not perfect. To test the scanner, a large number of (pretested) "good" bolts were run under the scanner, and it accepted 90% of the bolts as good.

Continued

Then a large number of (pretested) defective bolts were run under the scanner, and it accepted 3% of these as good bolts.

(a) If the company does not use the scanner, what percentage of a shipment is expected to be good? What percentage is expected to be defective?

(b) The scanner itself makes mistakes, and the company is questioning the value of using it. Suppose the company does use the scanner and ships only what the scanner passes as "good" bolts. In this case, what percentage of the shipment is expected to be good? What percentage is expected to be defective?

Partial Answer

To solve this problem, we use Bayes's theorem. The result of using the scanner is a dramatic improvement in the quality of the shipped product. If the scanner is not used, only 88% of the shipped bolts will be good. However, if the scanner is used and only the bolts it passes as good are shipped, then 99.6% of the shipment is expected to be good. Even though the scanner itself makes a considerable number of mistakes, it is definitely worth using. Not only does it increase the quality of a shipment, the bolts it rejects can be recycled into new bolts.

PART II The Hypergeometric Probability Distribution

In Chapter 5, we examined the binomial distribution. The binomial probability distribution assumes *independent trials.* If the trials are constructed by drawing samples from a population, then we have two possibilities: We sample either *with replacement* or *without replacement.* If we draw random samples with replacement, the trials can be taken to be independent. If we draw random samples without replacement and the population is very large, then it is reasonable to say that the trials are approximately independent. In this case, we go ahead and use the binomial distribution. However, if the population is relatively small and we draw samples without replacement, the assumption of independent trials is not valid, and we should not use the binomial distribution.

The *hypergeometric distribution* is a probability distribution of a random variable that has two outcomes when sampling is done *without replacement.*

Consider the following notational setup (see Figure AI-2). Suppose we have a population with only *two* distinct types of objects. Such a population might be

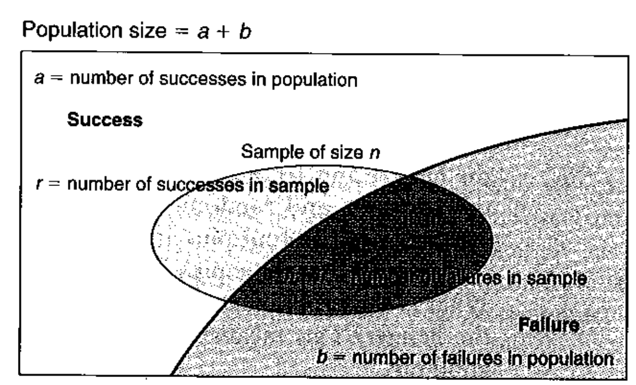


FIGURE AI-2

Notational Setup for Hypergeometric Distribution