Chapter 4 Probability
Complements, Conditional Probability, and Bayes' Theorem (4.3)

Objectives

In this lesson, we will learn how to

- use complements to find probabilities involving ``at least one'';
- find and interpret conditional probabilities;
- avoid confusion of the inverse;
- use two-way tables to find conditional probabilities;
- understand the idea behind Bayes' theorem;
- distinguish between prior and posterior probabilities.

1. Key Ideas

This lesson extends the multiplication rule and the complement rule.

Main Ideas

- ``At least one'' means one or more.
- A conditional probability uses extra information that another event has already occurred.
- Bayes' theorem is used to revise a probability after new information is obtained.

2. Complements and "At Least One"

When a problem asks for the probability of at least one, it is often easier to use the complement.

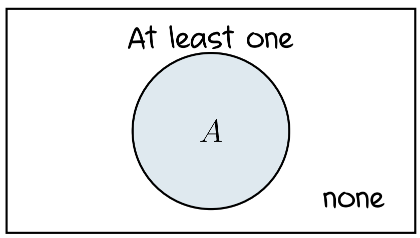
At Least One Rule

Let A be the event of getting at least one occurrence of some event. Then

$$P(A) = 1 - P(\overline{A}).$$

In words,

$$P(\text{at least one}) = 1 - P(\text{none}).$$



$$P(A) = 1 - P(\overline{A})$$

Example 1: Manufacturing: A factory has a defect rate of 15%. If a customer purchases 12 products, what is the probability of getting at least one defective product?

Solution

Let

$$A = \text{at least one of the 12 products is defective}.$$

The complement is

$$\overline{A} = \text{none of the 12 products is defective}.$$

If the defect rate is 0.15, then the probability that one product is good is

$$1 - 0.15 = 0.85.$$

Assuming the products are independent,

$$P(\text{none defective}) = (0.85)^{12}.$$

Therefore,

$$P(\text{at least one defective}) = 1 - (0.85)^{12}.$$

$$= 1 - 0.1422 = 0.8578 \approx 0.858.$$

So the probability of getting at least one defective product is approximately

$$0.858.$$

This is a very high probability, so the company should improve its manufacturing process.

3. Conditional Probability

A conditional probability is a probability found with the additional information that another event has already occurred.

Conditional Probability Notation

$$P(B \mid A)$$

means

the probability of event B occurring, given that event A has already occurred.

The vertical bar $|$ is read as "given."

Formal Conditional Probability Formula

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}, \qquad P(A) \neq 0.$$

Intuitive Approach

To find $P(B \mid A)$:

1. Restrict attention to the outcomes where A has occurred.
2. Within that smaller group, find the probability that B occurs.

4. Example: Pre-Employment Drug Screening

	Positive Test Result	Negative Test Result	Total
Subject Uses Drugs	45	5	50
Subject Does Not Use Drugs	25	480	505
Total	70	485	555

Example 2: Using the table above, find each probability.

(a) $P(\text{positive test result} \mid \text{subject uses drugs})$

(b) $P(\text{subject uses drugs} \mid \text{positive test result})$

Solution

Part (a).
We are given that the subject uses drugs. Therefore, we only look at the row for subjects who use drugs.
There are 50 subjects who use drugs, and 45 of them had positive test results.

$$P(\text{positive} \mid \text{uses drugs}) = \frac{45}{50} = 0.900.$$

So a subject who uses drugs has a 0.900 probability of getting a positive test result.

Part (b).
Now we are given that the subject had a positive test result. Therefore, we only look at the column for positive test results.

There are 70 positive test results, and 45 of those subjects actually use drugs.

$$P(\text{uses drugs} \mid \text{positive}) = \frac{45}{70} = 0.642857 \approx 0.643.$$

So a subject who gets a positive test result has about a 0.643 probability of actually using drugs.

5. Confusion of the Inverse

One common mistake is to switch the order of the events in a conditional probability.

Important Warning

In general,

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

The order matters.

In the drug screening example,

$$P(\text{positive} \mid \text{uses drugs}) = 0.900,$$

while

$$P(\text{uses drugs} \mid \text{positive}) = 0.643.$$

These are not the same probability.

Example 3: Confusion of the Inverse: Let

$$D = \text{it is dark outdoors}, \qquad M = \text{it is midnight}.$$

Compare $P(D \mid M)$ and $P(M \mid D)$.

Solution

If it is midnight, then it is almost certainly dark outdoors, so

$$P(D \mid M) \approx 1.$$

However, if it is dark outdoors, it does not mean that it is exactly midnight. It could be 9:00 p.m., 10:00 p.m., 11:00 p.m., or another time.
So

$$P(M \mid D) \approx 0.$$

Therefore,

$$P(D \mid M) \neq P(M \mid D).$$

6. Bayes' Theorem

Bayes' theorem is used to revise a probability when new information becomes available.

Bayes' Theorem Idea

Bayes' theorem helps us update an original probability after we learn additional information.

Prior and Posterior Probability

- A prior probability is the original probability before additional information is obtained.
- A posterior probability is the revised probability after additional information is obtained.

7. Example: Interpreting Medical Test Results

Suppose a disease has a 1% prevalence rate. Let

$$C = \text{the subject has cancer}.$$

Then

$$P(C) = 0.01.$$

The test has the following characteristics:

- False positive rate: 10%.
- True positive rate: 80%.

Example 4: Find

$$P(C \mid \text{positive test result}).$$

That is, find the probability that a subject actually has cancer, given that the test result is positive.

Solution

A convenient way to understand Bayes' theorem is to use a hypothetical population of 1000 people.
Since the cancer prevalence rate is 1%,

$$1000(0.01) = 10$$

people are expected to have cancer. The other

$$1000 - 10 = 990$$

people do not have cancer.

	Positive Test	Negative Test	Total
Cancer	8	2	10
No Cancer	99	891	990
Total	107	893	1000

Among the 10 people with cancer, 80% test positive:

$$10(0.80) = 8.$$

Among the 990 people without cancer, 10% test positive:

$$990(0.10) = 99.$$

Now focus only on the positive test results. There are

$$8 + 99 = 107$$

positive test results.
Out of these 107 positive test results, only 8 are true cancer cases.
Therefore,

$$P(C \mid \text{positive}) = \frac{8}{107} = 0.0748.$$

So even after a positive test result, the probability of cancer is about

$$7.48\%.$$

8. Interpretation of the Medical Test Example

Important Interpretation

The original probability of cancer was

$$P(C) = 0.01.$$

This is the prior probability.
After receiving a positive test result, the probability becomes

$$P(C \mid \text{positive}) = 0.0748.$$

This is the posterior probability.

The positive test result increases the probability of cancer from 1% to

7.48%, but it does not mean the subject definitely has cancer.

Summary

Key Ideas

- ``At least one'' means one or more.
- The complement of ``at least one'' is ``none.''
- Use $P(\text{at least one}) = 1 - P(\text{none})$.
- $P(B \mid A)$ means the probability of B given that A has occurred.
- In general, $P(B \mid A) \neq P(A \mid B)$.
- Bayes' theorem updates a prior probability after new information is obtained.
- A posterior probability is the revised probability after the new information is used.