Chapter 4

Probability

Chapter Outline

4.1 The Concept of Probability

4.2 Sample Spaces and Events

4.3 Some Elementary Probability Rules

4.4 Conditional Probability and Independence

4.5 Bayes' Theorem (Optional)

Review and outlook

Chapter 3: how to use sample statistics as point estimates of population parameters.

Starting in Chapter 7, we will focus on using sample statistics to make more sophisticated **statistical inferences** about population parameters. We will see that these statistical inferences are generalizations—based on calculating **probabilities**—about population parameters.

In this chapter and in Chapters 5 and 6, we present the fundamental concepts about **probability that are needed to make such statistical inferences**. We begin our discussions in this chapter by considering rules for calculating probabilities.

4.1 The Concept of Probability

- An **experiment** is any process of observation with an uncertain outcome
- The **sample space** of an experiment is the set of all possible outcomes for the experiment
- The possible outcomes are sometimes called the **experimental outcomes or sample space outcomes**
- **Probability** is a measure of the chance that an experimental outcome will occur when an experiment is carried out

Probability

- If E is a sample space outcome, then P(E) denotes the probability that E will occur and:
- **Conditions:**
  - $0 \leq P(E) \leq 1$ such that:
    - If E can never occur, then P(E) = 0
    - If E is certain to occur, then P(E) = 1
  - The probabilities of all the sample space outcomes must sum to 1

Example

Let A, B, C, D, and E be sample space outcomes forming a sample space. Suppose that $P(A) = .2$, $P(B) = .15$, $P(C) = .3$, and $P(D) = .2$. What is $P(E)$? Explain how you got your answer.

Assigning Probabilities to Sample Space Outcomes

1. **Classical method**
   - For equally likely outcomes
2. **Relative frequency method**
   - Using the long run relative frequency
3. **Subjective method**
   - Assessment based on experience, expertise or intuition

Classical method

- Example1: consider the experiment of **tossing a fair coin**. Here, there are two equally likely sample space outcomes—head (H) and tail (T). Therefore, logic suggests that the probability of observing a head, denoted P(H), is ½=0.5, and that the probability of observing a tail, denoted P(T), is also ½=0.5.

Classical method

- Example 2: consider the experiment of **rolling a fair die**. It would seem reasonable to think that the six sample space outcomes 1, 2, 3, 4, 5, and 6 are equally likely, and thus each outcome is assigned a probability of 1/6. If P(1) denotes the probability that one dot appears on the upward face of the die, then P(1)=1/6. Similarly, P(2)=1/6, P(3)=1/6, P(4)=1/6, P(5)=1/6, and P(6)=1/6.

Relative frequency method

- For example, to estimate the probability that a randomly selected consumer prefers Coca-Cola to all other soft drinks, we perform an experiment in which we ask a randomly selected consumer for his or her preference. There are **two possible experimental outcomes:** “prefers Coca-Cola” and “does not prefer Coca-Cola.”
- We might perform the experiment, say, 1,000 times by surveying 1,000 randomly selected consumers.
- Then, if 140 of those surveyed said that they prefer Coca-Cola, we would **estimate the probability** that a randomly selected consumer prefers Coca-Cola to all other soft drinks to be 140/1,000=0.14.
- This is an example of the relative frequency method of assigning probability.

Subjective method

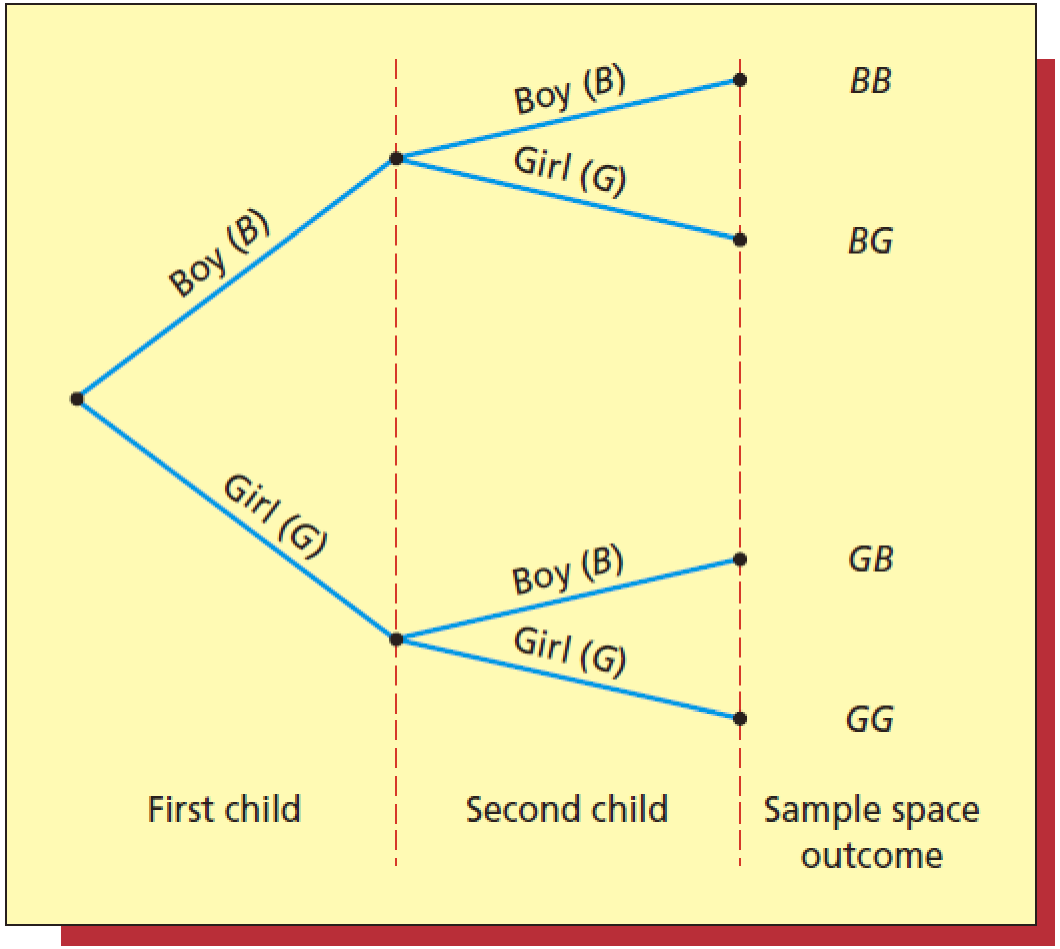
- When we use experience, intuitive judgement, or expertise to assess a probability, we call this **the subjective method** of assigning probability.
- For instance, when the company president estimates that the probability of a successful business venture is 0.7
- This may mean that, if business conditions similar to those that are about to be encountered could be repeated many times, then the business venture would be successful in 70 percent of the repetitions.

4.2 Sample Spaces and Events

- **Sample Space:** The set of all possible experimental outcomes
- **Sample Space Outcomes:** The experimental outcomes in the sample space
- **Event:** A set of sample space outcomes (a ***subset*** of sample space)
- **Probability:** The probability of an event is the sum of the probabilities of the sample space outcomes that correspond to the event

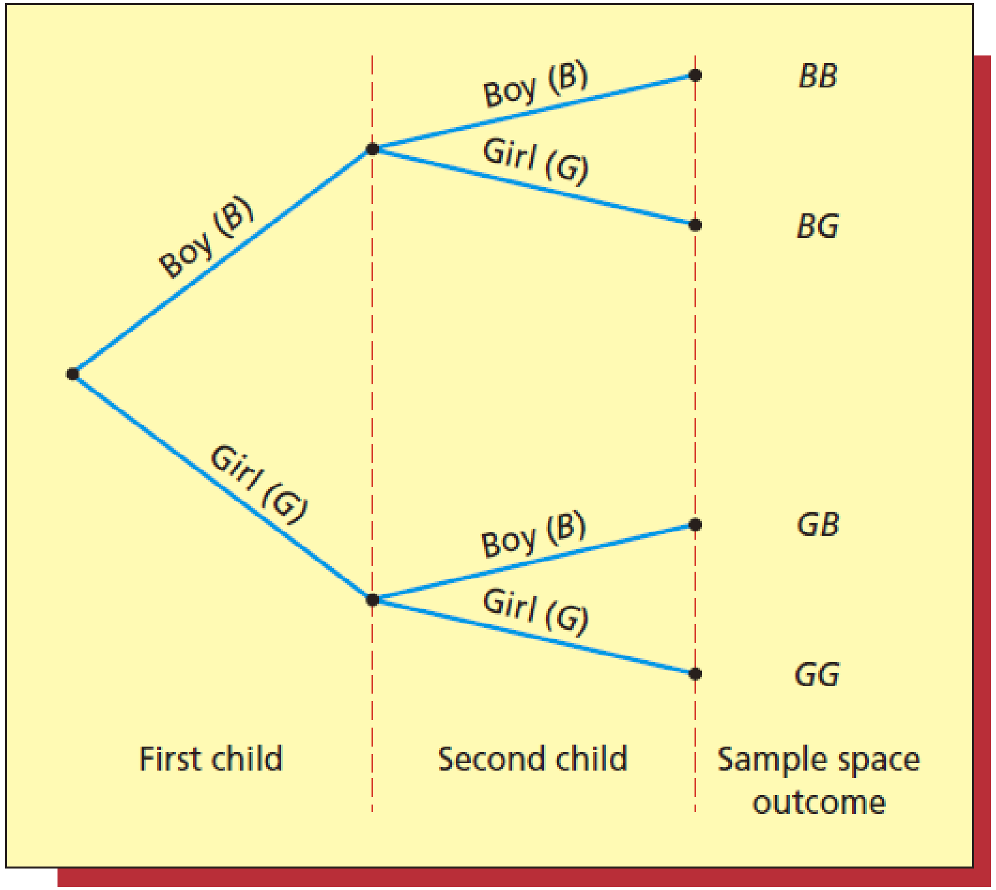
Example 4.1 Boys and girls

- A newly married couple plans to have two children. In order to find the sample space of this experiment, we let B denote that a child is a boy and G denote that a child is a girl. The diagram on the right pictures the experiment as a two-step process—having the first child, which could be either a boy or a girl (B or G), and then having the second child, which could also be either a boy or a girl (B or G). Each branch of the tree leads to a sample space outcome. We see that there are four sample space outcomes. Therefore, the sample space (that is, the set of all the sample space outcomes) is
- BB BG GB GG



Example 4.1Boys and girls

- sample space outcomes: BB BG GB GG
- suppose that boys and girls are equally likely each time a child is born. This implies that
- $P(BB) = \; P(BG) = \; P(GB) = P(GG) = \frac{1}{4}$
- The probability that the couple will have two boys is $P(BB) = \frac{1}{4}$
- The probability having one boy and one girl is$P(BG) + P(GB) = \frac{1}{2}$
- The probability that the couple will have two girls is $P(BB) = \frac{1}{4}$
- The probability that the couple will have at least one girl is$P(BG) + \; P(GB) + P(GG) = \frac{3}{4}$

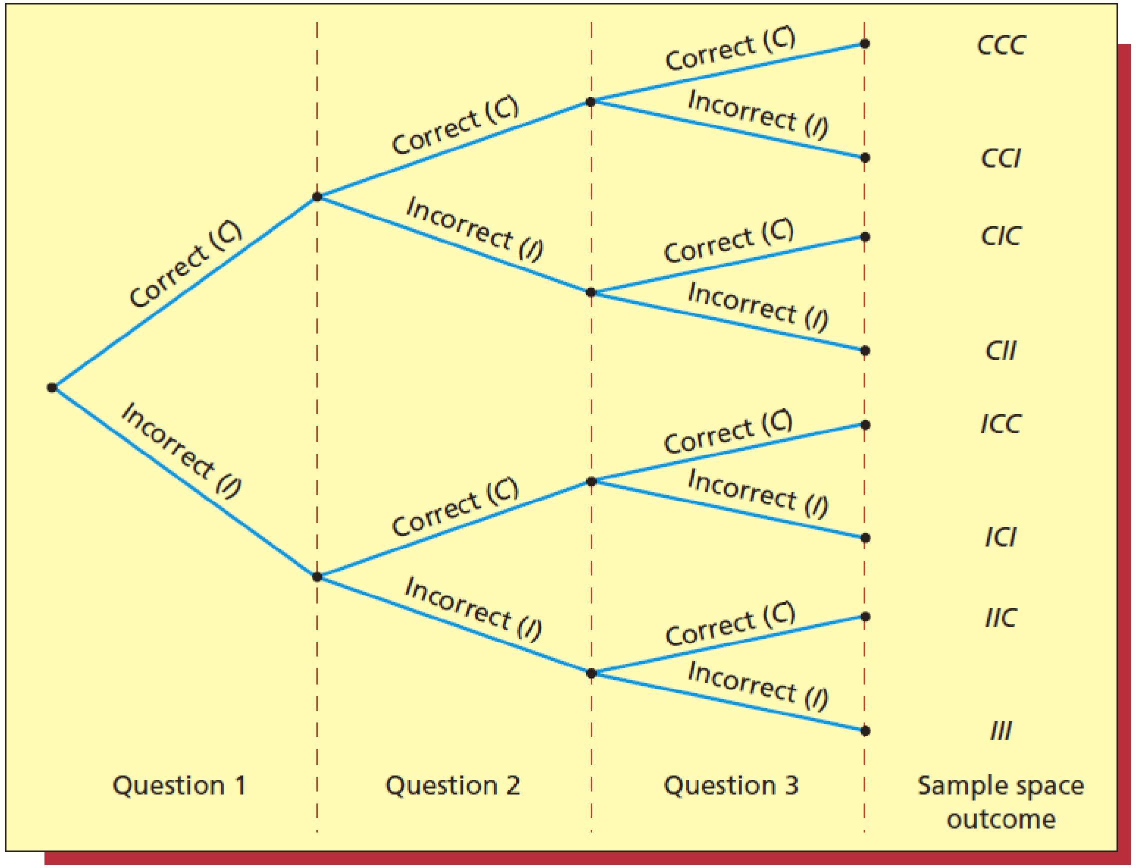


Example 4.2 Pop quizzes

- A student takes a pop quiz that consists of three true–false questions. Each question can be answered correctly (denoted as *C*) or incorrectly (*I*)
- The sample space is

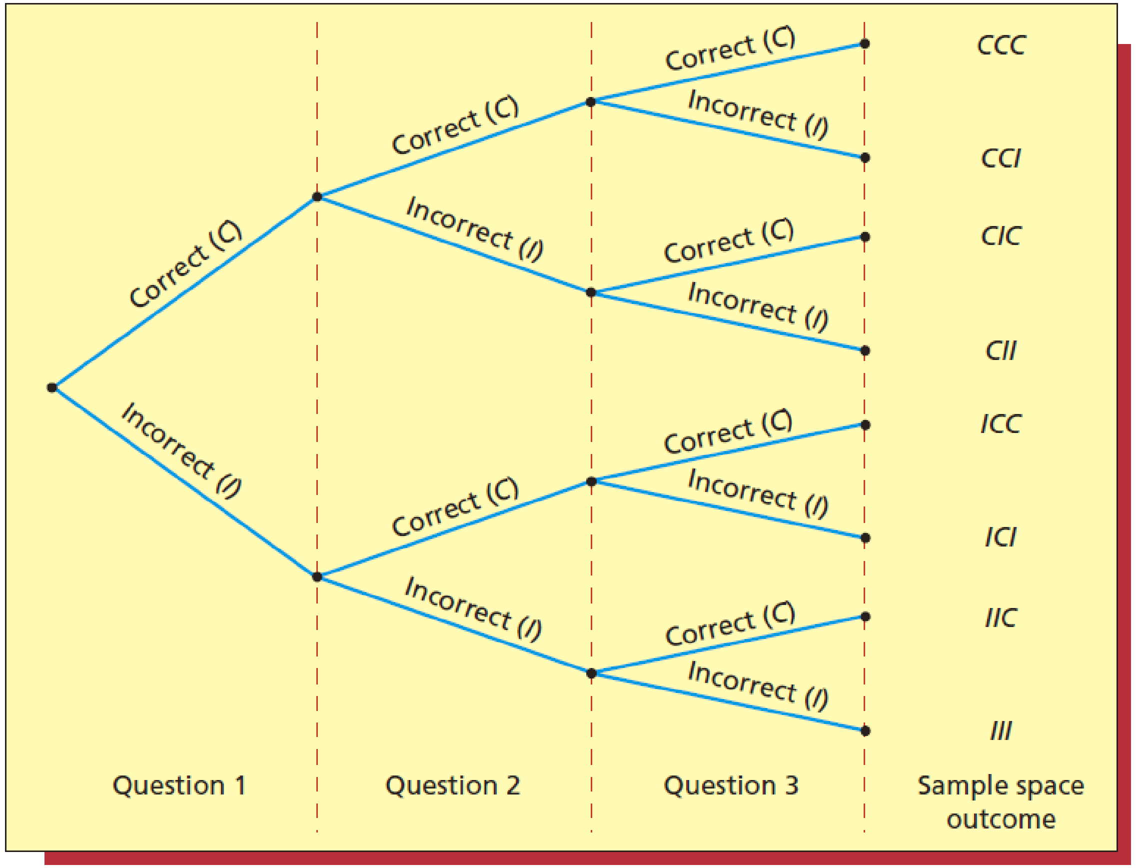
CCC	*CCI*	*CIC*	*CII*
ICC	*ICI*	*IIC*	*III*

- Assume the student had a 50–50 chance (or .5 probability) of correctly answering each question
- $P(CCC) = \ P(CCI) = \ \ldots = P(III) = \frac{1}{8}$



Example 4.2 Pop quizzes

- Assume the student had a 50–50 chance (or .5 probability) of correctly answering each question
- $P(CCC) = P(CCI) = \ldots = P(III) = \frac{1}{8}$
- The probability that three questions are correct is $P(CCC) = \frac{1}{8}$
- The probability that exactly two questions correct is $P(CCI) + P(CIC) + P(ICC) = \frac{3}{8}$
- The probability that one question correct is $P(CII) + P(ICI) + P(IIC) = \frac{3}{8}$
- The probability that all three questions incorrect is $P(III) = \frac{1}{8}$
- The probability that at least two questions correct is $P(CCC) + P(CCI) + P(CIC) + P(ICC) = \frac{4}{8} = \frac{1}{2}$



Finding Simple Probabilities

- Sample space is finite
- All sample space outcomes equally likely
- Probability of an event can be computed using the following formula:

$$\frac{\text{the number of sample space outcomes that correspond to the event}}{\text{the total number of sample space outcomes}}$$

Example 4.3 Choosing a CEO

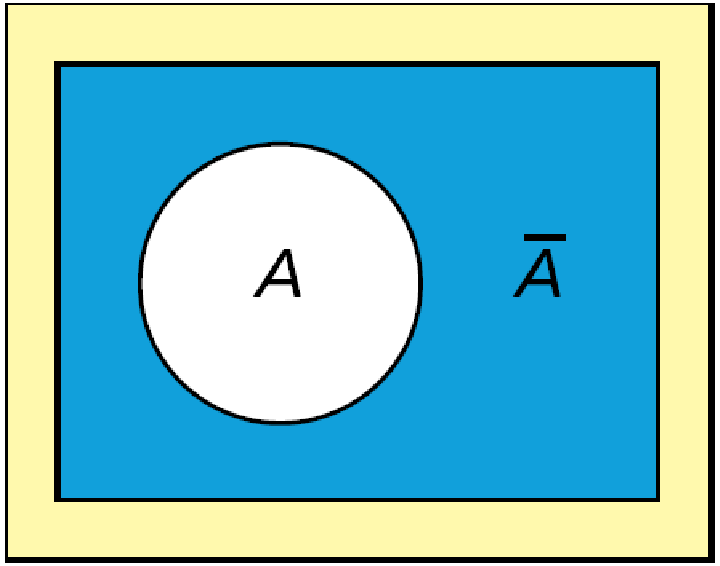
- Acompany is choosing a new chief executive officer (CEO).
- Four finalists (identified by last name only)—Adams (A), Chung (C), Hill (H), and Rankin (R).
- $P(\mathrm{A}) = 0.1$, $P(C) = 0.2$, $P(H) = 0.5$, $P(R) = 0.2$
- Adams and Hill are internal candidates (they already work for the company). Letting INT denote the event that “an internal candidate will be selected for the CEO position,”
- $P(INT) = \frac{the\ number\ of\ internal\ candidates}{the\ total\ number\ of\ candidates} = \frac{2}{4} = 0.5$ (wrong because the sample space outcomes are not equally likely)
- $P(INT) = P(\mathrm{A}) + P(H) = 0.6$ (correct answer)

4.3 Some Elementary Probability Rules

- Complement
- Union
- Intersection
- Addition
- Conditional probability

Complement

- **The complement ($\bar{A}$)** of an event A is the set of all sample space outcomes not in A
- $P(\bar{A}) = 1 – P(A)$



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

Complement

- **The complement ($\bar{A}$)** of an event A is the set of all sample space outcomes not in A

The Rule of Complements

Consider an event A. Then, **the probability that** ***A will not occur*** is

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

Example of the complement

The Rule of Complements

Consider an event A. Then, **the probability that A *will not occur*** is

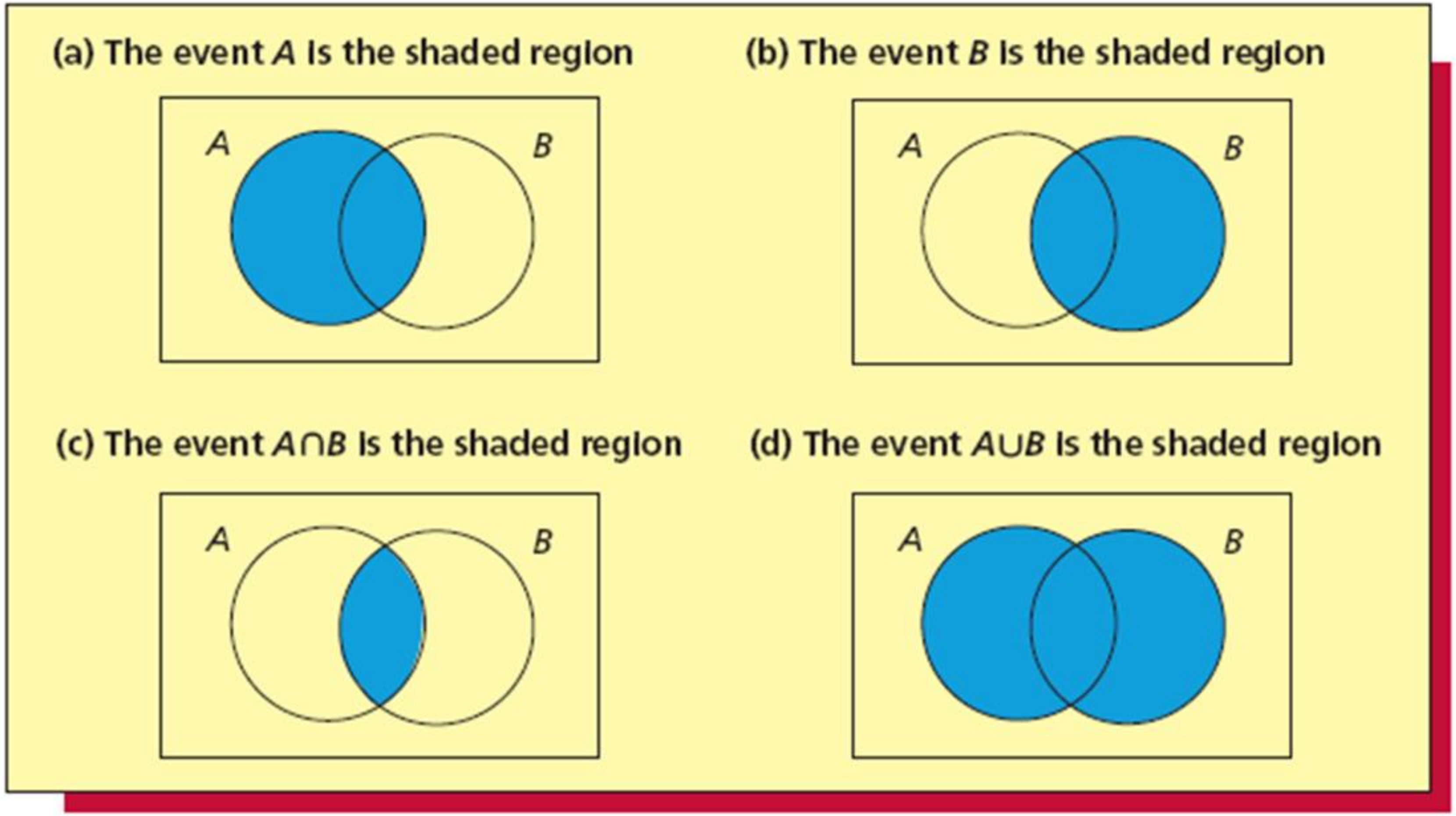
$$P(\overline{A}) = 1 - P(A)$$

- The probability that a randomly selected cable passing has Crystal's cable television service is 0.45. It follows that the probability of the complement of this event (that is, the probability that a randomly selected cable passing does not have Crystal's cable television service) is 1 - 0.45 = 0.55

Union and Intersection

- The **intersection** of A and B are elementary events that belong to both A and B
  - Written as $A \cap B$
- The **union** of A and B are elementary events that belong to either A or B or both
  - Written as $A \cup B$

Union and Intersection Diagram



Union and Intersection

Given two events A and B, the **intersection of A and B** is the event that occurs if both A and B simultaneously occur. The intersection is denoted by $A \cap B$. Furthermore, $P(A \cap B)$ denotes **the probability that *both* A *and* B *will simultaneously occur.***

Given two events A and B, the **union of A and B** is the event that occurs if A or B (or both) occur. The union is denoted $A \cup B$. Furthermore, $P(A \cup B)$ denotes **the probability that *A or B (or both) will occur.***

Example of Union and Intersection

Recall from Example 4.4 that Crystal Cable has 27.4 million cable passings. Consider randomly selecting one of these cable passings, and define the following events:

$A \equiv$ the randomly selected cable passing has Crystal's cable television service.

$\overline{A} \equiv$ the randomly selected cable passing does not have Crystal's cable television service.

$B \equiv$ the randomly selected cable passing has Crystal's cable Internet service.

$\overline{B} \equiv$ the randomly selected cable passing does not have Crystal's cable Internet service.

$A \cap B \equiv$ the randomly selected cable passing has both Crystal's cable television service and Crystal's cable Internet service.

$A \cap \overline{B} \equiv$ the randomly selected cable passing has Crystal's cable television service and does not have Crystal's cable Internet service.

Example of Union and Intersection

A Contingency Table Summarizing Crystal's Cable Television and Internet Penetration (Figures in Millions of Cable Passings)

Events	Has Cable Internet Service, B	Does Not Have Cable Internet Service, $\bar{B}$	Total
Has Cable Television Service, A	6.5	5.9	12.4
Does Not Have Cable Television Service, $\bar{A}$	3.3	11.7	15.0
Total	9.8	17.6	27.4

$$P(A) = \frac{12.4}{27.4} = 0.45$$

$$P(B) = \frac{9.8}{27.4} = 0.36$$

$$P(A \cap B) = \frac{6.5}{27.4} = 0.24$$

$$P(\bar{A} \cap B) = \frac{3.3}{27.4} = 0.12$$

$$P(\bar{A} \cap \bar{B}) = \frac{11.7}{27.4} = 0.43$$

Contingency Table

$P(R_1 \cap C_1)$

$P(R_1)$

	C_1	C_2	Total
R_1	.4	.2	.6
R_2	.1	.3	.4
Total	.5	.5	1.00

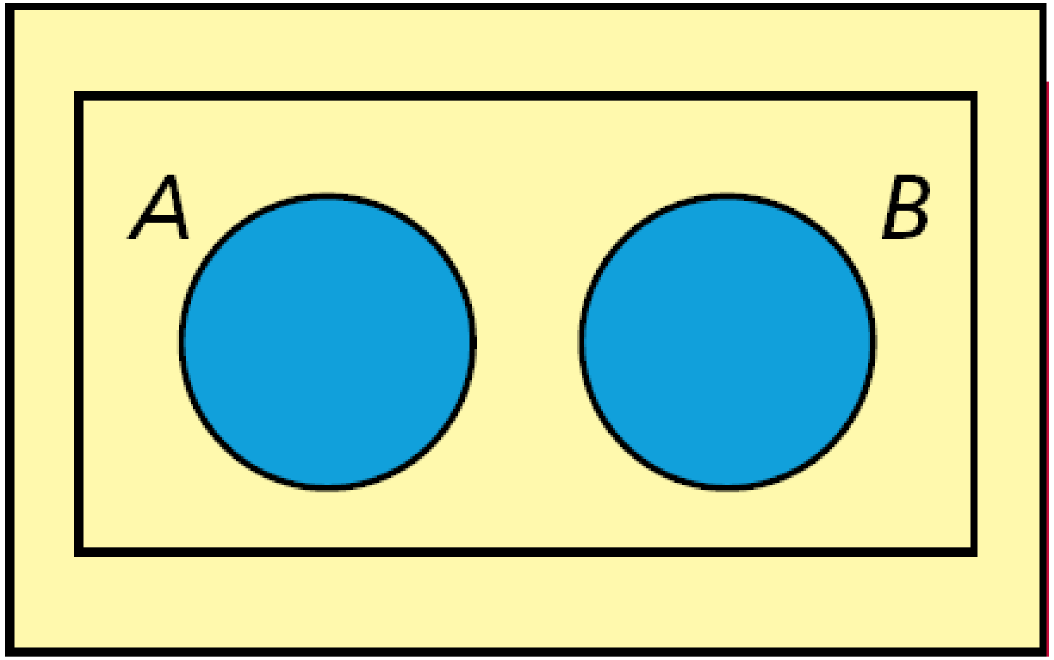
$P(C_2)$

$P(R_2 \cap C_2)$

Mutually Exclusive

- A and B are **mutually exclusive** if they have no sample space outcomes in common
- In other words:

$$P(A \cap B) = 0$$



The Addition Rule

- If A and B are **mutually exclusive** （ $P(A \cap B) = 0$ ）, then the probability that A or B (the union of A and B) will occur is

  $$P(A \cup B) = P(A) + P(B)$$

- If A and B are ***not mutually exclusive***:

  $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

  where $P(A \cap B)$ is the **joint probability** of A and B both occurring together

Example Selecting Playing Cards

- Consider randomly selecting a card from a standard deck of 52 playing cards, and define the events
- J ={the randomly selected card is a jack}.
- Q={the randomly selected card is a queen}.
- R={the randomly selected card is a red card (a diamond or a heart)}.
- Because there are 4 jacks, 4 queens, and 26 red cards, we have P(J)=4/52, P(Q)=4/52, and P(R)=26/52.

Example Selecting Playing Cards

- The probability that the randomly selected card is a jack or a queen is

$$P(J \cup Q) = P(J) + P(Q)$$
$$= \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$$

- The probability that the randomly selected card is a jack or a red card is

$$P(J \cup R) = P(J) + P(R) - P(J \cap R)$$
$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52} = \frac{7}{13}$$

The Addition Rule for N mutually exclusive events

The Addition Rule for N Mutually Exclusive Events

The events $A_1, A_2, \ldots, A_N$ are mutually exclusive if no two of the events have any sample space outcomes in common. In this case, no two of the events can occur simultaneously, and

$$P(A_1 \cup A_2 \cup \cdots \cup A_N) = P(A_1) + P(A_2) + \cdots + P(A_N)$$

As an example of using this formula, again consider the playing card situation and the events J and Q. If we define the event

$$K \equiv \text{the randomly selected card is a king}$$

then the events J, Q, and K are mutually exclusive. Therefore,

$$P(J \cup Q \cup K) = P(J) + P(Q) + P(K)$$

4.4 Conditional Probability and Independence

- The probability of an event A, given that the event B has occurred, is **called the conditional probability of A given B**
  - Denoted as $P(A|B)$
- Further, $P(A|B) = P(A \cap B) / P(B)$
  - $P(B) \neq 0$
- Likewise, $P(B|A) = P(A \cap B) / P(A)$

Interpretation

- Restrict sample space to just event B
- The conditional probability $P(A|B)$ is the chance of event A occurring in this new sample space
- In other words, if B occurred, then what is the chance of A occurring

General Multiplication Rule

- Given any two events, A and B

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

- Referred to as **general multiplication rule**

Conditional probability

Conditional Probability

1 The **conditional probability of the event *A* given that the event *B* has occurred** is written ***P*(*A* | *B*)** and is defined to be

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

Here we assume that $P(B)$ is greater than 0.

2 The **conditional probability of the event *B* given that the event *A* has occurred** is written ***P*(*B* | *A*)** and is defined to be

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

Here we assume that $P(A)$ is greater than 0.

Example

A Contingency Table Summarizing Crystal's Cable Television and Internet Penetration (Figures In Millions Of Cable Passings)

Events	Has Cable Internet Service, B	Does Not Have Cable Internet Service, $\bar{B}$	Total
Has Cable Television Service, A	6.5	5.9	12.4
Does Not Have Cable Television Service, $\bar{A}$	3.3	11.7	15.0
Total	9.8	17.6	27.4

$$P(A|B) = \frac{6.5}{9.8} = \frac{P(A \cap B)}{P(B)} = \frac{6.5/27.4}{9.8/27.4} = 0.66$$

$$P(B|A) = \frac{6.5}{12.4} = \frac{P(A \cap B)}{P(A)} = \frac{6.5/27.4}{12.4/27.4} = 0.52$$

Example Gender Issues at a Pharmaceutical Company

- 52 percent of sales representatives(reps) are women $P(W) = 0.52$
- 44 percent of the management sales reps are women $P(W|MGT) = 0.44$
- 25 percent of the sales reps have a management position $P(MGT) = 0.25$
- Let *W* denote the event that the randomly selected sales representative is a woman let *M* denote the event that the randomly selected sales representative is a man. $P(M) = 1 - 0.52 = 0.48$
- Let *MGT* denote the event that the randomly selected sales representative has a management position.

Example Gender Issues at a Pharmaceutical Company

- The percentage of the sales representatives that have a management position and are women.

$$P(MGT \cap W) = P(MGT)P(W|MGT) = 0.25 \times 0.44 = 0.11$$

- The percentage of the female sales representatives that have a management position.

$$P(MGT|W) = \frac{P(MGT \cap W)}{P(W)} = \frac{0.11}{0.52} = 0.2115$$

Example Gender Issues at a Pharmaceutical Company

- The percentage of the sales representatives that have a management position and are men.

$$P(MGT \cap M) = P(MGT)P(M|MGT) = P(MGT)\big(1 - P(W|MGT)\big) = 0.25 \times (1 - 0.44)$$
$$= 0.14$$

- The percentage of the male sales representatives that have a management position.

$$P(MGT|M) = \frac{P(MGT \cap M)}{P(M)} = \frac{0.14}{0.48} = 0.2917$$

Independence

Independence In Example 4.10 the probability of the event MGT is influenced by whether the event W occurs. In such a case, we say that the events MGT and W are **dependent.** If $P(MGT \mid W)$ were equal to $P(MGT)$, then the probability of the event MGT would not be influenced by whether W occurs. In this case we would say that the events MGT and W are **independent.** This leads to the following definition:

Independent Events

Two events A and B are **independent** if and only if

1. $P(A \mid B) = P(A)$ or, equivalently,
2. $P(B \mid A) = P(B)$

Here we assume that $P(A)$ and $P(B)$ are greater than 0.

Example Gender Issues at a Pharmaceutical Company

- 52 percent of sales reps are women
- 44 percent of management are women
- 25 percent have a management position
- If gender and management are independent, would expect 25 percent of both women and men to be management
- This was not the case
  - P(MGT|W) = 0.2115
  - P(MGT|M) = 0.2917
- We conclude that women ***<u>are less likely to have a management position</u>*** at the pharmaceutical company.

Example Gender Issues at a Pharmaceutical Company

- Note that the ratio of P(MGT|M)=0.2917 to P(MGT|W)=0.2115 is 0.2917/0.2115=1.3792.
- This says that the probability that a randomly selected sales representative will have a management position is 37.92 percent higher if the sales representative is a man than it is if the sales representative is a woman.
- In other words, the probability that a randomly selected rep will be management is 37.92 percent higher for a man
- This conclusion describes the actual employment conditions that existed at Novartis Pharmaceutical Company from 2002 to 2007
- Gender discrimination

The Multiplication Rule

- The **joint probability** that A and B (the intersection of A and B) will occur is
  $P(A \cap B) = P(A)\,P(B|A) = P(B)\,P(A|B)$
- If A and B are independent, then the probability that A and B will occur is:

  $P(A \cap B) = P(A)\,P(B) = P(B)\,P(A)$
- For N independent events

  $P(A_1 \cap A_2 \cap \ldots \cap A_N) = P(A_1)\,P(A_2) \ldots P(A_N)$

EXAMPLE An Application of the Independence Rule: Customer Service

This example is based on a real situation encountered by a major producer and marketer of consumer products. The company assessed the service it provides by surveying the attitudes of its customers regarding 10 different aspects of customer service—order filled correctly, billing amount on invoice correct, delivery made on time, and so forth. When the survey results were analyzed, the company was dismayed to learn that only 59 percent of the survey participants indicated that they were satisfied with all 10 aspects of the company's service. Upon investigation, each of the 10 departments responsible for the aspects of service considered in the study insisted that it satisfied its customers 95 percent of the time. That is, each department claimed that its error rate was only 5 percent. Company executives were confused and felt that there was a substantial discrepancy between the survey results and the claims of the departments providing the services. However, a company statistician pointed out that there was no discrepancy. To understand this, consider randomly selecting a customer from among the survey participants, and define 10 events (corresponding to the 10 aspects of service studied):

$A_1 \equiv$ the customer is satisfied that the order is filled correctly (aspect 1).

$A_2 \equiv$ the customer is satisfied that the billing amount on the invoice is correct (aspect 2).

$\vdots$

$A_{10} \equiv$ the customer is satisfied that the delivery is made on time (aspect 10).

Also, define the event

$S \equiv$ the customer is satisfied with all 10 aspects of customer service.

Because 10 different departments are responsible for the 10 aspects of service being studied, it is reasonable to assume that all 10 aspects of service are independent of each other. For instance, billing amounts would be independent of delivery times. Therefore, $A_1, A_2, \ldots, A_{10}$ are independent events, and

$$\begin{aligned} P(S) &= P(A_1 \cap A_2 \cap \cdots \cap A_{10}) \\ &= P(A_1)P(A_2) \cdots P(A_{10}) \end{aligned}$$

If, as the departments claim, each department satisfies its customers 95 percent of the time, then the probability that the customer is satisfied with all 10 aspects is

$$P(S) = (.95)(.95) \cdots (.95) = (.95)^{10} = .5987$$

This result is almost identical to the 59 percent satisfaction rate reported by the survey participants.

If the company wants to increase the percentage of its customers who are satisfied with all 10 aspects of service, it must improve the quality of service provided by the 10 departments. For example, to satisfy 95 percent of its customers with all 10 aspects of service, the company must require each department to raise the fraction of the time it satisfies its customers to x, where x is such that $(x)^{10} = .95$. It follows that

$$x = (.95)^{\frac{1}{10}} = .9949$$

and that each department must satisfy its customers 99.49 percent of the time (rather than the current 95 percent of the time).

Conclusion

- In this chapter we studied **probability**.
- We began by defining an event to be an **experimental outcome** that may or may not occur and by defining the probability of an event to be a number that measures the likelihood that the event will occur.
- We learned that a probability is often interpreted as a long-run relative frequency, and we saw that probabilities can be found by examining sample spaces and by using probability rules.
- Several important probability rules—**addition rules, multiplication rules, and the rule of complements**. We also studied a special kind of probability called a **conditional probability,** which is the probability that one event will occur given that another event occurs, and we used probabilities to define **independent events.**

4.5 Bayes' Theorem

- Law of Total Probability (optional)
- Bayes' Theorem (optional)

Example-Medical Diagnostic Test

- **Example :** suppose that 0.1% of a human population have one rare cancer and a medical test returns positive or negative for detecting this disease.
- For all the people who have cancer and are tested, 99% of them will get a positive result from the test.
- When the people do not have cancer, only 5% of them will get a positive result of the test.
- **Question:** if a randomly selected patient has the test and it comes back positive, what is the probability that the patient has cancer?

Example-Medical Diagnostic Test

- Denote that A ={the test result is positive}, B ={the patient has cancer}
- 0.1% of a human population have one rare cancer $\Longrightarrow P(B) = 0.001$
- For all the people who have cancer and are tested, 99% of them will get a positive result from the test $\Longrightarrow P(A|B) = 0.99$
- When the people do not have cancer, only 5% of them will get a positive result of the test $\Longrightarrow P(A|\bar{B}) = 0.05$
- **Question:** if a randomly selected patient has the test and it comes back positive, what is the probability that the patient has cancer? $P(B|A)$

$$\boldsymbol{P(B|A) = \frac{P(A \cap B)}{P(A)}} \textbf{ ?}$$

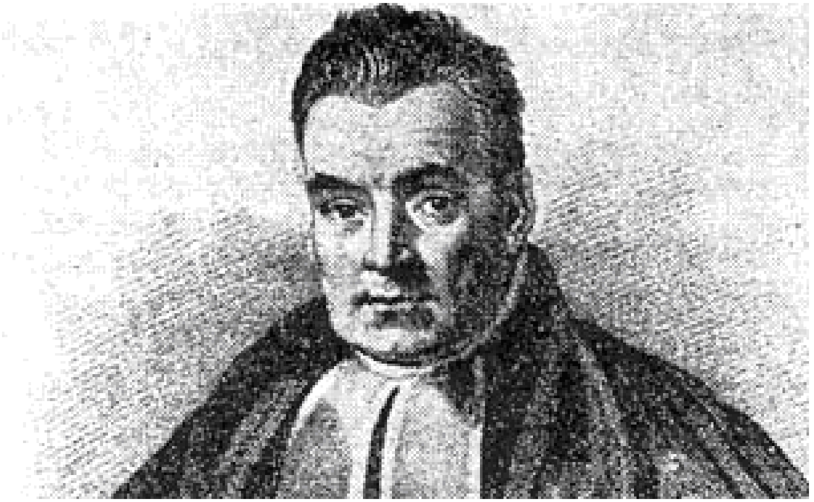
Law of Total Probability

- $P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$, where $\bar{B}$ is the complement of B

- **Law of total probability:** suppose that $\{B_k, k = 1, \cdots, n\}$ be a set of pairwise disjoint events whose union is the entire sample space, then for any event A of the same probability space,

$$P(A) = \sum_{k=1}^{n} P(A \cap B_k) = \sum_{k=1}^{n} P(A|B_k)P(B_k)$$

Bayes’ Theorem



- **Bayes’ Theorem** was firstly proposed by English statistician Thomas Bayes in his paper published in 1763 after his death.

- In fact, it is another way to ***calculate the conditional probability.***

Bayes' Theorem

- By the definition of conditional probability, for two events A and B and $P(A) \neq 0$,

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

- By the multiplication rule **(Bayes' Theorem)**

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

- By the law of total probability,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

Bayes' Theorem

- **Bayes' Theorem:** one principled way of calculating a conditional probability **without** the joint probability.

$$P(B|A) = \frac{P(A|B)}{P(A)} P(B) = \frac{P(A|B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} P(B)$$

- $P(B)$: **prior probability**
- $P(B|A)$: **posterior probability**
- **Bayes' Theorem:** one way to obtain the posterior probability from the prior probability

Example-Medical Diagnostic Test

- **Example:** Denote that A ={the test result is positive}, B ={the patient has cancer}
- Suppose that **0.1%** of a human population have one rare cancer. $\boldsymbol{P(B)}$
- For all the people who have cancer and are tested, **99%** of them will get a positive result from the test. $\boldsymbol{P(A|B)}$ **(True positve rate)**
- When the people do not have cancer, only **5%** of them will get a positive result of the test. $\boldsymbol{P(A|\bar{B})}$ **(False positive rate)**
- **Question:** if a randomly selected patient has the test and it comes back positive, what is the probability that the patient has cancer? $\boldsymbol{P(B|A)}$
- Intuition: this probability $\boldsymbol{P(B|A)}$ should be large (*Correct or not Correct?*)

Example-Medical Diagnostic Test

- **Prior probability**: $P(B) = 0.001 \Longrightarrow P(\bar{B}) = 0.999$
- $P(A|B) = 0.99$ **(True positve rate)**
- $P(A|\bar{B}) = 0.05$ **(False positive rate)**
- **Posterior probability**: $P(B|A)$

$$P(B|A) = P(B) \times \frac{P(A|B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

$$= 0.001 \times \frac{0.99}{0.99 \times 0.001 + 0.05 * 0.999} \approx 0.019$$

- **Surprising!** Even if the test result is positive, the probability of having cancer is still small (0.019).
- A positive result is not enough to indicate that the patient is sick
- Why? This is caused by **the false positives.**

Example-Medical Diagnostic Test

- If the false positive rate is reduced to **1%,** what is $P(B|A)$?
- **Posterior probability**: $P(B|A)$

$$P(B|A) = P(B) \times \frac{P(A|B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

$$= 0.001 \times \frac{0.99}{0.99 \times 0.001 + 0.01 * 0.999} \approx 0.091$$

- $P(B|A) = 0.019 \nearrow P(B|A) = 0.091$
- An excellent and widely used example of the benefit of Bayes' Theorem is in the analysis of a **medical diagnostic test.**

EXAMPLE 4.14 The Oil Drilling Case: Site Selection

An oil company is attempting to decide whether to drill for oil on a particular site. There are three possible states of nature:

1. No oil (state of nature S_1, which we will denote as *none*)
2. Some oil (state of nature S_2, which we will denote as *some*)
3. Much oil (state of nature S_3, which we will denote as *much*)

Based on experience and knowledge concerning the site's geological characteristics, the oil company feels that the prior probabilities of these states of nature are as follows:

$$P(S_1 \equiv \text{none}) = .7 \qquad P(S_2 \equiv \text{some}) = .2 \qquad P(S_3 \equiv \text{much}) = .1$$

In order to obtain more information about the potential drilling site, the oil company can perform a seismic experiment, which has three readings—low, medium, and high. Moreover, information exists concerning the accuracy of the seismic experiment. The company's historical records tell us that

1. Of 100 past sites that were drilled and produced no oil, 4 sites gave a high reading. Therefore,

$$P(\text{high} \mid \text{none}) = \frac{4}{100} = .04$$

2. Of 400 past sites that were drilled and produced some oil, 8 sites gave a high reading. Therefore,

$$P(\text{high} \mid \text{some}) = \frac{8}{400} = .02$$

3. Of 300 past sites that were drilled and produced much oil, 288 sites gave a high reading. Therefore,

Example 4.18

- Oil drilling on a particular site
  - $P(S_1 = \text{none}) = .7$
  - $P(S_2 = \text{some}) = .2$
  - $P(S_3 = \text{much}) = .1$
- Can perform a seismic experiment
  - $P(\text{high}|\text{none}) = .04$
  - $P(\text{high}|\text{some}) = .02$
  - $P(\text{high}|\text{much}) = .96$

Example 4.18 Continued

$$\begin{aligned} P(high) &= P(none \cap high) + P(some \cap high) + P(much \cap high) \\ &= P(none)P(high \mid none) + P(some)P(high \mid some) + P(much)P(high \mid much) \\ &= (.7)(.04) + (.2)(.02) + (.1)(.96) = .128 \end{aligned}$$

$$P(none \mid high) = \frac{P(none \cap high)}{P(high)} = \frac{P(none)P(high \mid none)}{P(high)} = \frac{.7(.04)}{.128} = .21875$$

$$P(some \mid high) = \frac{P(some \cap high)}{P(high)} = \frac{P(some)P(high \mid some)}{P(high)} = \frac{.2(.02)}{.128} = .03125$$

$$P(much \mid high) = \frac{P(much \cap high)}{P(high)} = \frac{P(much)P(high \mid much)}{P(high)} = \frac{.1(.96)}{.128} = .75$$

Bayes' Theorem

Let $S_1, S_2, \ldots, S_k$ be k mutually exclusive states of nature, one of which must be true, and suppose that $P(S_1)$, $P(S_2), \ldots, P(S_k)$ are the prior probabilities of these states of nature. Also, let E be a particular outcome of an experiment designed to help determine which state of nature is really true. Then, the **posterior probability** of a particular state of nature, say S_i, given the experimental outcome E, is

$$P(S_i \mid E) = \frac{P(S_i \cap E)}{P(E)} = \frac{P(S_i)P(E \mid S_i)}{P(E)}$$

where

$$\begin{aligned} P(E) &= P(S_1 \cap E) + P(S_2 \cap E) + \cdots + P(S_k \cap E) \\ &= P(S_1)P(E \mid S_1) + P(S_2)P(E \mid S_2) + \cdots + P(S_k)P(E \mid S_k) \end{aligned}$$

Specifically, if there are two mutually exclusive states of nature, S_1 and S_2, one of which must be true, then

$$P(S_i \mid E) = \frac{P(S_i)P(E \mid S_i)}{P(S_1)P(E \mid S_1) + P(S_2)P(E \mid S_2)}$$

Thank you!