Probability

Introduction

The probability of a specified event is the chance or likelihood that it will occur. There are several ways of viewing probability. One would be **experimental** in nature, where we repeatedly conduct an experiment. Suppose we flipped a coin over and over and over again and it came up heads about half of the time; we would expect that in the future whenever we flipped the coin it would turn up heads about half of the time. When a weather reporter says "there is a 10% chance of rain tomorrow," she is basing that on prior evidence; that out of all days with similar weather patterns, it has rained on 1 out of 10 of those days.

Another view would be **subjective** in nature, in other words an educated guess. If someone asked you the probability that the Seattle Mariners would win their next baseball game, it would be impossible to conduct an experiment where the same two teams played each other repeatedly, each time with the same starting lineup and starting pitchers, each starting at the same time of day on the same field under the precisely the same conditions. Since there are so many variables to take into account, someone familiar with baseball and with the two teams involved might make an educated guess that there is a 75% chance they will win the game; that is, *if* the same two teams were to play each other repeatedly under identical conditions, the Mariners would win about three out of every four games. But this is just a guess, with no way to verify its accuracy, and depending upon how educated the educated guesser is, a subjective probability may not be worth very much.

We will return to the experimental and subjective probabilities from time to time, but in this course we will mostly be concerned with **theoretical** probability, which is defined as follows: Suppose there is a situation with n equally likely possible outcomes and that m of those n outcomes correspond to a particular event; then the **probability** of that event is defined as $\frac{m}{n}$.

Basic Concepts

If you roll a die, pick a card from deck of playing cards, or randomly select a person and observe their hair color, we are executing an experiment or procedure. In probability, we look at the likelihood of different outcomes. We begin with some terminology.

Events and Outcomes

The result of an experiment is called an **outcome**.

An **event** is any particular outcome or group of outcomes.

A **simple event** is an event that cannot be broken down further

The **sample space** is the set of all possible simple events.

Example 1

If we roll a standard 6-sided die, describe the sample space and some simple events.

The sample space is the set of all possible simple events: {1,2,3,4,5,6}

Some examples of simple events:
We roll a 1
We roll a 5

Some compound events:
We roll a number bigger than 4
We roll an even number

Two dice

One die

Basic Probability
Given that all outcomes are equally likely, we can compute the probability of an event *E* using this formula:

$$P(E) = \frac{\text{Number of outcomes corresponding to the event E}}{\text{Total number of equally - likely outcomes}}$$

Example 2

If we roll a 6-sided die, calculate
a) P(rolling a 1)
b) P(rolling a number bigger than 4)

Recall that the sample space is {1,2,3,4,5,6}

a) There is one outcome corresponding to "rolling a 1", so the probability is $\frac{1}{6}$

b) There are two outcomes bigger than a 4, so the probability is $\frac{2}{6} = \frac{1}{3}$

Probabilities are essentially fractions, and can be reduced to lower terms like fractions.

Example 3

Let's say you have a bag with 20 cherries, 14 sweet and 6 sour. If you pick a cherry at random, what is the probability that it will be sweet?

There are 20 possible cherries that could be picked, so the number of possible outcomes is 20. Of these 20 possible outcomes, 14 are favorable (sweet), so the probability that the cherry will be sweet is $\frac{14}{20} = \frac{7}{10}$.

There is one potential complication to this example, however. It must be assumed that the probability of picking any of the cherries is the same as the probability of picking any other. This wouldn't be true if (let us imagine) the sweet cherries are smaller than the sour ones. (The sour cherries would come to hand more readily when you sampled from the bag.) Let us keep in mind, therefore, that when we assess probabilities in terms of the ratio of favorable to all potential cases, we rely heavily on the assumption of equal probability for all outcomes.

Try it Now 1

At some random moment, you look at your clock and note the minutes reading.

a. What is probability the minutes reading is 15?

b. What is the probability the minutes reading is 15 or less?

Cards

A standard deck of 52 playing cards consists of four **suits** (hearts, spades, diamonds and clubs). Spades and clubs are black while hearts and diamonds are red. Each suit contains 13 cards, each of a different **rank**: an Ace (which in many games functions as both a low card and a high card), cards numbered 2 through 10, a Jack, a Queen and a King.

Example 4

Compute the probability of randomly drawing one card from a deck and getting an Ace.

There are 52 cards in the deck and 4 Aces so $P(Ace) = \frac{4}{52} = \frac{1}{13} \approx 0.0769$

We can also think of probabilities as percents: There is a 7.69% chance that a randomly selected card will be an Ace.

Notice that the smallest possible probability is 0 – if there are no outcomes that correspond with the event. The largest possible probability is 1 – if all possible outcomes correspond with the event.

Certain and Impossible events

An impossible event has a probability of 0.

A certain event has a probability of 1.

The probability of any event must be $0 \le P(E) \le 1$

In the course of this chapter, *if you compute a probability and get an answer that is negative or greater than 1, you have made a mistake and should check your work.*

Working with Events

Complementary Events

Now let us examine the probability that an event does **not** happen. As in the previous section, consider the situation of rolling a six-sided die and first compute the probability of rolling a six: the answer is P(six) =1/6. Now consider the probability that we do *not* roll a six: there are 5 outcomes that are not a six, so the answer is $P(\text{not a six}) = \frac{5}{6}$. Notice that

$$P(\text{six}) + P(\text{not a six}) = \frac{1}{6} + \frac{5}{6} = \frac{6}{6} = 1$$

This is not a coincidence. Consider a generic situation with n possible outcomes and an event E that corresponds to m of these outcomes. Then the remaining n - m outcomes correspond to E not happening, thus

$$P(\text{not } E) = \frac{n-m}{n} = \frac{n}{n} - \frac{m}{n} = 1 - \frac{m}{n} = 1 - P(E)$$

Complement of an Event

The **complement** of an event is the event "E doesn't happen"

The notation $\bar{E}$ is used for the complement of event E.

We can compute the probability of the complement using $P\left(\bar{E}\right) = 1 - P(E)$

Notice also that $P(E) = 1 - P\left(\bar{E}\right)$

Example 5

If you pull a random card from a deck of playing cards, what is the probability it is not a heart?

There are 13 hearts in the deck, so $P(\text{heart}) = \frac{13}{52} = \frac{1}{4}$.

The probability of *not* drawing a heart is the complement:

$$P(\text{not heart}) = 1 - P(\text{heart}) = 1 - \frac{1}{4} = \frac{3}{4}$$

Probability of two independent events

Example 6

Suppose we flipped a coin and rolled a die, and wanted to know the probability of getting a head on the coin and a 6 on the die.

We could list all possible outcomes: $\{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$.

Notice there are $2 \cdot 6 = 12$ total outcomes. Out of these, only 1 is the desired outcome, so the probability is $\frac{1}{12}$.

The prior example was looking at two independent events.

Independent Events
Events A and B are **independent events** if the probability of Event B occurring is the same whether or not Event A occurs.

Example 7

Are these events independent?
a) A fair coin is tossed two times. The two events are (1) first toss is a head and (2) second toss is a head.

b) The two events (1) "It will rain tomorrow in Houston" and (2) "It will rain tomorrow in Galveston" (a city near Houston).

c) You draw a card from a deck, then draw a second card without replacing the first.

a) The probability that a head comes up on the second toss is 1/2 regardless of whether or not a head came up on the first toss, so these events are independent.

b) These events are not independent because it is more likely that it will rain in Galveston on days it rains in Houston than on days it does not.

c) The probability of the second card being red depends on whether the first card is red or not, so these events are not independent.

When two events are independent, the probability of both occurring is the product of the probabilities of the individual events.

***P*(*A* and *B*) for independent events**
If events A and B are independent, then the probability of both A and B occurring is

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

where $P(A \text{ and } B)$ is the probability of events A and B both occurring, $P(A)$ is the probability of event A occurring, and $P(B)$ is the probability of event B occurring

If you look back at the coin and die example from earlier, you can see how the number of outcomes of the first event multiplied by the number of outcomes in the second event multiplied to equal the total number of possible outcomes in the combined event.

Example 8

In your drawer you have 10 pairs of socks, 6 of which are white, and 7 tee shirts, 3 of which are white. If you randomly reach in and pull out a pair of socks and a tee shirt, what is the probability both are white?

The probability of choosing a white pair of socks is $\frac{6}{10}$.

The probability of choosing a white tee shirt is $\frac{3}{7}$.

The probability of both being white is $\frac{6}{10} \cdot \frac{3}{7} = \frac{18}{70} = \frac{9}{35}$

Try it Now 2

A card is pulled a deck of cards and noted. The card is then replaced, the deck is shuffled, and a second card is removed and noted. What is the probability that both cards are Aces?

The previous examples looked at the probability of *both* events occurring. Now we will look at the probability of *either* event occurring.

Example 9

Suppose we flipped a coin and rolled a die, and wanted to know the probability of getting a head on the coin *or* a 6 on the die.

Here, there are still 12 possible outcomes: {H1,H2,H3,H4,H5,H6,T1,T2,T3,T4,T5,T6}

By simply counting, we can see that 7 of the outcomes have a head on the coin *or* a 6 on the die *or* both – we use *or* inclusively here (these 7 outcomes are H1, H2, H3, H4, H5, H6, T6), so the probability is $\frac{7}{12}$. How could we have found this from the individual probabilities?

As we would expect, $\frac{1}{2}$ of these outcomes have a head, and $\frac{1}{6}$ of these outcomes have a 6 on the die. If we add these, $\frac{1}{2} + \frac{1}{6} = \frac{6}{12} + \frac{2}{12} = \frac{8}{12}$, which is not the correct probability. Looking at the outcomes we can see why: the outcome H6 would have been counted twice, since it contains both a head and a 6; the probability of both a head *and* rolling a 6 is $\frac{1}{12}$.

If we subtract out this double count, we have the correct probability: $\frac{8}{12} - \frac{1}{12} = \frac{7}{12}$.

***P*(*A* or *B*)**
The probability of either A or B occurring (or both) is

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Example 10

Suppose we draw one card from a standard deck. What is the probability that we get a Queen or a King?

There are 4 Queens and 4 Kings in the deck, hence 8 outcomes corresponding to a Queen or King out of 52 possible outcomes. Thus the probability of drawing a Queen or a King is:

$$P(\text{King or Queen}) = \frac{8}{52}$$

Note that in this case, there are no cards that are both a Queen and a King, so $P(\text{King and Queen}) = 0$. Using our probability rule, we could have said:

$$P(\text{King or Queen}) = P(\text{King}) + P(\text{Queen}) - P(\text{King and Queen}) = \frac{4}{52} + \frac{4}{52} - 0 = \frac{8}{52}$$

In the last example, the events were **mutually exclusive**, so $P(A \text{ or } B) = P(A) + P(B)$.

Example 11

Suppose we draw one card from a standard deck. What is the probability that we get a red card or a King?

Half the cards are red, so $P(\text{red}) = \frac{26}{52}$

There are four kings, so $P(\text{King}) = \frac{4}{52}$

There are two red kings, so $P(\text{Red and King}) = \frac{2}{52}$

We can then calculate

$$P(\text{Red or King}) = P(\text{Red}) + P(\text{King}) - P(\text{Red and King}) = \frac{26}{52} + \frac{4}{52} - \frac{2}{52} = \frac{28}{52}$$

Try it Now 3
In your drawer you have 10 pairs of socks, 6 of which are white, and 7 tee shirts, 3 of which are white. If you reach in and randomly grab a pair of socks and a tee shirt, what the probability at least one is white?

Example 12

The table below shows the number of survey subjects who have received and not received a speeding ticket in the last year, and the color of their car. Find the probability that a randomly chosen person:
a) Has a red car *and* got a speeding ticket
b) Has a red car *or* got a speeding ticket.

	Speeding ticket	No speeding ticket	Total
Red car	15	135	150
Not red car	45	470	515
Total	60	605	665

We can see that 15 people of the 665 surveyed had both a red car and got a speeding ticket, so the probability is $\frac{15}{665} \approx 0.0226$.

Notice that having a red car and getting a speeding ticket are not independent events, so the probability of both of them occurring is not simply the product of probabilities of each one occurring.

We could answer this question by simply adding up the numbers: 15 people with red cars and speeding tickets + 135 with red cars but no ticket + 45 with a ticket but no red car = 195 people. So the probability is $\frac{195}{665} \approx 0.2932$.

We also could have found this probability by:
P(had a red car) + P(got a speeding ticket) – P(had a red car and got a speeding ticket)
$= \frac{150}{665} + \frac{60}{665} - \frac{15}{665} = \frac{195}{665}$.

Conditional Probability

Often it is required to compute the probability of an event given that another event has occurred.

Example 13

What is the probability that two cards drawn at random from a deck of playing cards will both be aces?

It might seem that you could use the formula for the probability of two independent events and simply multiply $\frac{4}{52} \cdot \frac{4}{52} = \frac{1}{169}$. This would be incorrect, however, because the two events are not independent. If the first card drawn is an ace, then the probability that the second card is also an ace would be lower because there would only be three aces left in the deck.

Once the first card chosen is an ace, the probability that the second card chosen is also an ace is called the **conditional probability** of drawing an ace. In this case the "condition" is that the first card is an ace. Symbolically, we write this as:
P(ace on second draw | an ace on the first draw).

The vertical bar "|" is read as "given," so the above expression is short for "The probability that an ace is drawn on the second draw given that an ace was drawn on the first draw." What is this probability? After an ace is drawn on the first draw, there are 3 aces out of 51 total cards left. This means that the conditional probability of drawing an ace after one ace has already been drawn is $\frac{3}{51} = \frac{1}{17}$.

Thus, the probability of both cards being aces is $\frac{4}{52} \cdot \frac{3}{51} = \frac{12}{2652} = \frac{1}{221}$.

Conditional Probability
The probability the event B occurs, given that event A has happened, is represented as
$P(B \mid A)$
This is read as "the probability of B given A"

Example 14

Find the probability that a die rolled shows a 6, given that a flipped coin shows a head.

These are two independent events, so the probability of the die rolling a 6 is $\frac{1}{6}$, regardless of the result of the coin flip.

Example 15

The table below shows the number of survey subjects who have received and not received a speeding ticket in the last year, and the color of their car. Find the probability that a randomly chosen person:
a) Has a speeding ticket *given* they have a red car
b) Has a red car *given* they have a speeding ticket

	Speeding ticket	No speeding ticket	Total
Red car	15	135	150
Not red car	45	470	515
Total	60	605	665

a) Since we know the person has a red car, we are only considering the 150 people in the first row of the table. Of those, 15 have a speeding ticket, so

$$\text{P(ticket} \mid \text{red car)} = \frac{15}{150} = \frac{1}{10} = 0.1$$

b) Since we know the person has a speeding ticket, we are only considering the 60 people in the first column of the table. Of those, 15 have a red car, so
P(red car | ticket) = $\dfrac{15}{60} = \dfrac{1}{4} = 0.25$.

Notice from the last example that P(B | A) is **not** equal to P(A | B).

These kinds of conditional probabilities are what insurance companies use to determine your insurance rates. They look at the conditional probability of you having accident, given your age, your car, your car color, your driving history, etc., and price your policy based on that likelihood.

Conditional Probability Formula
If Events *A* and *B* are not independent, then
$P(A \text{ and } B) = P(A) \cdot P(B \mid A)$

Example 16

If you pull 2 cards out of a deck, what is the probability that both are spades?

The probability that the first card is a spade is $\dfrac{13}{52}$.

The probability that the second card is a spade, given the first was a spade, is $\dfrac{12}{51}$, since there is one less spade in the deck, and one less total cards.

The probability that both cards are spades is $\dfrac{13}{52} \cdot \dfrac{12}{51} = \dfrac{156}{2652} \approx 0.0588$

Example 17

If you draw two cards from a deck, what is the probability that you will get the Ace of Diamonds and a black card?

You can satisfy this condition by having Case A or Case B, as follows:
Case A) you can get the Ace of Diamonds first and then a black card or
Case B) you can get a black card first and then the Ace of Diamonds.

Let's calculate the probability of Case A. The probability that the first card is the Ace of Diamonds is $\dfrac{1}{52}$. The probability that the second card is black given that the first card is the Ace of Diamonds is $\dfrac{26}{51}$ because 26 of the remaining 51 cards are black. The probability is therefore $\dfrac{1}{52} \cdot \dfrac{26}{51} = \dfrac{1}{102}$.

Now for Case B: the probability that the first card is black is $\frac{26}{52}=\frac{1}{2}$. The probability that the second card is the Ace of Diamonds given that the first card is black is $\frac{1}{51}$. The probability of Case B is therefore $\frac{1}{2}\cdot\frac{1}{51}=\frac{1}{102}$, the same as the probability of Case 1.

Recall that the probability of A or B is $P(A) + P(B) - P(A \text{ and } B)$. In this problem, $P(A \text{ and } B) = 0$ since the first card cannot be the Ace of Diamonds and be a black card. Therefore, the probability of Case A or Case B is $\frac{1}{102}+\frac{1}{102}=\frac{2}{102}=\frac{1}{51}$. The probability that you will get the Ace of Diamonds and a black card when drawing two cards from a deck is $\frac{1}{51}$.

Try it Now 4

In your drawer you have 10 pairs of socks, 6 of which are white. If you reach in and randomly grab two pairs of socks, what is the probability that both are white?

Example 18

A home pregnancy test was given to women, then pregnancy was verified through blood tests. The following table shows the home pregnancy test results. Find

a) $P(\text{not pregnant} \mid \text{positive test result})$

b) $P(\text{positive test result} \mid \text{not pregnant})$

	Positive test	Negative test	Total
Pregnant	70	4	74
Not Pregnant	5	14	19
Total	75	18	93

a) Since we know the test result was positive, we're limited to the 75 women in the first column, of which 5 were not pregnant. $P(\text{not pregnant} \mid \text{positive test result}) = \frac{5}{75} \approx 0.067$.

b) Since we know the woman is not pregnant, we are limited to the 19 women in the second row, of which 5 had a positive test. $P(\text{positive test result} \mid \text{not pregnant}) = \frac{5}{19} \approx 0.263$

The second result is what is usually called a false positive: A positive result when the woman is not actually pregnant.

Bayes Theorem

In this section we concentrate on the more complex conditional probability problems we began looking at in the last section.

Example 19

Suppose a certain disease has an incidence rate of 0.1% (that is, it afflicts 0.1% of the population). A test has been devised to detect this disease. The test does not produce false negatives (that is, anyone who has the disease will test positive for it), but the false positive rate is 5% (that is, about 5% of people who take the test will test positive, even though they do not have the disease). Suppose a randomly selected person takes the test and tests positive. What is the probability that this person actually has the disease?

There are two ways to approach the solution to this problem. One involves an important result in probability theory called Bayes' theorem. We will discuss this theorem a bit later, but for now we will use an alternative and, we hope, much more intuitive approach.

Let's break down the information in the problem piece by piece.

Suppose a certain disease has an incidence rate of 0.1% (that is, it afflicts 0.1% of the population). The percentage 0.1% can be converted to a decimal number by moving the decimal place two places to the left, to get 0.001. In turn, 0.001 can be rewritten as a fraction: 1/1000. This tells us that about 1 in every 1000 people has the disease. (If we wanted we could write P(disease)=0.001.)

A test has been devised to detect this disease. The test does not produce false negatives (that is, anyone who has the disease will test positive for it). This part is fairly straightforward: everyone who has the disease will test positive, or alternatively everyone who tests negative does not have the disease. (We could also say P(positive | disease)=1.)

The false positive rate is 5% (that is, about 5% of people who take the test will test positive, even though they do not have the disease). This is even more straightforward. Another way of looking at it is that of every 100 people who are tested and do not have the disease, 5 will test positive even though they do not have the disease. (We could also say that P(positive | no disease)=0.05.)

Suppose a randomly selected person takes the test and tests positive. What is the probability that this person actually has the disease? Here we want to compute P(disease|positive). We already know that P(positive|disease)=1, but remember that conditional probabilities are not equal if the conditions are switched.
Rather than thinking in terms of all these probabilities we have developed, let's create a hypothetical situation and apply the facts as set out above. First, suppose we randomly select 1000 people and administer the test. How many do we expect to have the disease? Since about 1/1000 of all people are afflicted with the disease, 1/1000 of 1000 people is 1. (Now you know why we chose 1000.) Only 1 of 1000 test subjects actually has the disease; the other 999 do not.

We also know that 5% of all people who do not have the disease will test positive. There are 999 disease-free people, so we would expect (0.05)(999)=49.95 (so, about 50) people to test positive who do not have the disease.

Now back to the original question, computing *P*(disease|positive). There are 51 people who test positive in our example (the one unfortunate person who actually has the disease, plus the 50 people who tested positive but don't). Only one of these people has the disease, so

$$\text{P(disease} \mid \text{positive)} \approx \frac{1}{51} \approx 0.0196$$

or less than 2%. Does this surprise you? This means that of all people who test positive, over 98% *do not have the disease*.

The answer we got was slightly approximate, since we rounded 49.95 to 50. We could redo the problem with 100,000 test subjects, 100 of whom would have the disease and (0.05)(99,900)=4995 test positive but do not have the disease, so the exact probability of having the disease if you test positive is

$$\text{P(disease} \mid \text{positive)} \approx \frac{100}{5095} \approx 0.0196$$

which is pretty much the same answer.

But back to the surprising result. *Of all people who test positive, over 98% do not have the disease.* If your guess for the probability a person who tests positive has the disease was wildly different from the right answer (2%), don't feel bad. The exact same problem was posed to doctors and medical students at the Harvard Medical School 25 years ago and the results revealed in a 1978 *New England Journal of Medicine* article. Only about 18% of the participants got the right answer. Most of the rest thought the answer was closer to 95% (perhaps they were misled by the false positive rate of 5%).

So at least you should feel a little better that a bunch of doctors didn't get the right answer either (assuming you thought the answer was much higher). But the significance of this finding and similar results from other studies in the intervening years lies not in making math students feel better but in the possibly catastrophic consequences it might have for patient care. If a doctor thinks the chances that a positive test result nearly guarantees that a patient has a disease, they might begin an unnecessary and possibly harmful treatment regimen on a healthy patient. Or worse, as in the early days of the AIDS crisis when being HIV-positive was often equated with a death sentence, the patient might take a drastic action and commit suicide.

As we have seen in this hypothetical example, the most responsible course of action for treating a patient who tests positive would be to counsel the patient that they most likely do *not* have the disease and to order further, more reliable, tests to verify the diagnosis.

One of the reasons that the doctors and medical students in the study did so poorly is that such problems, when presented in the types of statistics courses that medical students often take, are solved by use of Bayes' theorem, which is stated as follows:

Bayes' Theorem

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

In our earlier example, this translates to

$$P(\text{disease} \mid \text{positive})=\frac{P(\text{disease})P(\text{positive} \mid \text{disease})}{P(\text{disease})P(\text{positive} \mid \text{disease})+P(\text{no disease})P(\text{positive} \mid \text{no disease})}$$

Plugging in the numbers gives

$$P(\text{disease} \mid \text{positive})=\frac{(0.001)(1)}{(0.001)(1)+(0.999)(0.05)} \approx 0.0196$$

which is exactly the same answer as our original solution.

The problem is that you (or the typical medical student, or even the typical math professor) are much more likely to be able to remember the original solution than to remember Bayes' theorem. Psychologists, such as Gerd Gigerenzer, author of *Calculated Risks: How to Know When Numbers Deceive You*, have advocated that the method involved in the original solution (which Gigerenzer calls the method of "natural frequencies") be employed in place of Bayes' Theorem. Gigerenzer performed a study and found that those educated in the natural frequency method were able to recall it far longer than those who were taught Bayes' theorem. When one considers the possible life-and-death consequences associated with such calculations it seems wise to heed his advice.

Example 20

A certain disease has an incidence rate of 2%. If the false negative rate is 10% and the false positive rate is 1%, compute the probability that a person who tests positive actually has the disease.

Imagine 10,000 people who are tested. Of these 10,000, 200 will have the disease; 10% of them, or 20, will test negative and the remaining 180 will test positive. Of the 9800 who do not have the disease, 98 will test positive. So of the 278 total people who test positive, 180 will have the disease. Thus

$$P(\text{disease} \mid \text{positive})=\frac{180}{278} \approx 0.647$$

so about 65% of the people who test positive will have the disease.

Using Bayes theorem directly would give the same result:

$$P(\text{disease} \mid \text{positive})=\frac{(0.02)(0.90)}{(0.02)(0.90)+(0.98)(0.01)}=\frac{0.018}{0.0278} \approx 0.647$$

Try it Now 5
A certain disease has an incidence rate of 0.5%. If there are no false negatives and if the false positive rate is 3%, compute the probability that a person who tests positive actually has the disease.

Counting

Counting? You already know how to count or you wouldn't be taking a college-level math class, right? Well yes, but what we'll really be investigating here are ways of counting *efficiently*. When we get to the probability situations a bit later in this chapter we will need to count some *very* large numbers, like the number of possible winning lottery tickets. One way to do this would be to write down every possible set of numbers that might show up on a lottery ticket, but believe me: you don't want to do this.

Basic Counting

We will start, however, with some more reasonable sorts of counting problems in order to develop the ideas that we will soon need.

Example 21

Suppose at a particular restaurant you have three choices for an appetizer (soup, salad or breadsticks) and five choices for a main course (hamburger, sandwich, quiche, fajita or pizza). If you are allowed to choose exactly one item from each category for your meal, how many different meal options do you have?

Solution 1: One way to solve this problem would be to systematically list each possible meal:

soup + hamburger	soup + sandwich	soup + quiche
soup + fajita	soup + pizza	salad + hamburger
salad + sandwich	salad + quiche	salad + fajita
salad + pizza	breadsticks + hamburger	breadsticks + sandwich
breadsticks + quiche	breadsticks + fajita	breadsticks + pizza

Assuming that we did this systematically and that we neither missed any possibilities nor listed any possibility more than once, the answer would be 15. Thus you could go to the restaurant 15 nights in a row and have a different meal each night.

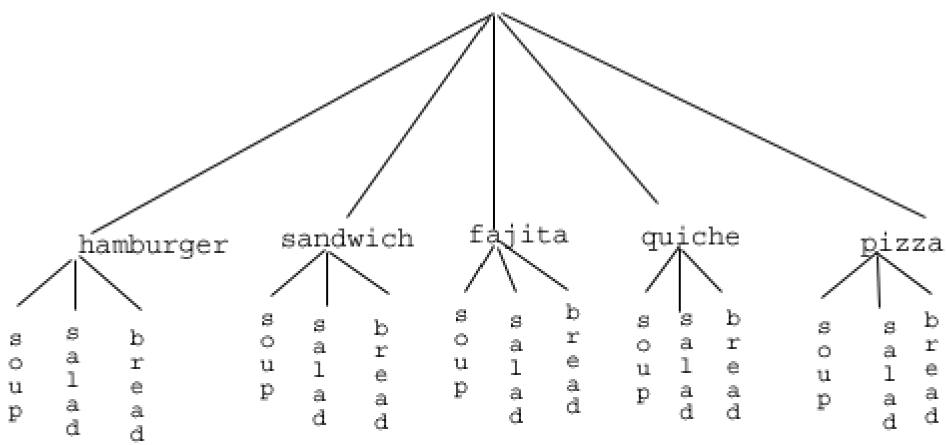
Solution 2: Another way to solve this problem would be to list all the possibilities in a table:

	hamburger	**sandwich**	**quiche**	**fajita**	**pizza**
soup	soup+burger				
salad	salad+burger				
bread	*etc.*				

In each of the cells in the table we could list the corresponding meal: soup + hamburger in the upper left corner, salad + hamburger below it, etc. But if we didn't really care *what* the possible meals are, only *how many* possible meals there are, we could just count the number of cells and arrive at an answer of 15, which matches our answer from the first solution. (It's always good when you solve a problem two different ways and get the same answer!)

Solution 3: We already have two perfectly good solutions. Why do we need a third? The first method was not very systematic, and we might easily have made an omission. The second method was better, but suppose that in addition to the appetizer and the main course we further complicated the problem by adding desserts to the menu: we've used the rows of the table for the appetizers and the columns for the main courses—where will the desserts go? We would need a third dimension, and since drawing 3-D tables on a 2-D page or computer screen isn't terribly easy, we need a better way in case we have three categories to choose form instead of just two.

So, back to the problem in the example. What else can we do? Let's draw a **tree diagram**:



This is called a "tree" diagram because at each stage we branch out, like the branches on a tree. In this case, we first drew five branches (one for each main course) and then for each of those branches we drew three more branches (one for each appetizer). We count the number of branches at the final level and get (surprise, surprise!) 15.

If we wanted, we could instead draw three branches at the first stage for the three appetizers and then five branches (one for each main course) branching out of each of those three branches.

OK, so now we know how to count possibilities using tables and tree diagrams. These methods will continue to be useful in certain cases, but imagine a game where you have two decks of cards (with 52 cards in each deck) and you select one card from each deck. Would you really want to draw a table or tree diagram to determine the number of outcomes of this game?

Let's go back to the previous example that involved selecting a meal from three appetizers and five main courses, and look at the second solution that used a table. Notice that one way to count the number of possible meals is simply to number each of the appropriate cells in the table, as we have done above. But another way to count the number of cells in the table would be multiply the number of rows (3) by the number of columns (5) to get 15. Notice that we could have arrived at the same result without making a table at all by simply multiplying the number of choices for the appetizer (3) by the number of choices for the main course (5). We generalize this technique as the *basic counting rule*:

Basic Counting Rule

If we are asked to choose one item from each of two separate categories where there are m items in the first category and n items in the second category, then the total number of available choices is $m \cdot n$.

This is sometimes called the multiplication rule for probabilities.

Example 22

There are 21 novels and 18 volumes of poetry on a reading list for a college English course. How many different ways can a student select one novel and one volume of poetry to read during the quarter?

There are 21 choices from the first category and 18 for the second, so there are $21 \cdot 18 = 378$ possibilities.

The Basic Counting Rule can be extended when there are more than two categories by applying it repeatedly, as we see in the next example.

Example 23

Suppose at a particular restaurant you have three choices for an appetizer (soup, salad or breadsticks), five choices for a main course (hamburger, sandwich, quiche, fajita or pasta) and two choices for dessert (pie or ice cream). If you are allowed to choose exactly one item from each category for your meal, how many different meal options do you have?

There are 3 choices for an appetizer, 5 for the main course and 2 for dessert, so there are $3 \cdot 5 \cdot 2 = 30$ possibilities.

Example 24

A quiz consists of 3 true-or-false questions. In how many ways can a student answer the quiz?

There are 3 questions. Each question has 2 possible answers (true or false), so the quiz may be answered in $2 \cdot 2 \cdot 2 = 8$ different ways. Recall that another way to write $2 \cdot 2 \cdot 2$ is 2^3, which is much more compact.

Try it Now 6

Suppose at a particular restaurant you have eight choices for an appetizer, eleven choices for a main course and five choices for dessert. If you are allowed to choose exactly one item from each category for your meal, how many different meal options do you have?

Permutations

In this section we will develop an even faster way to solve some of the problems we have already learned to solve by other means. Let's start with a couple examples.

Example 25

How many different ways can the letters of the word MATH be rearranged to form a four-letter code word?

This problem is a bit different. Instead of choosing one item from each of several different categories, we are repeatedly choosing items from the *same* category (the category is: the letters of the word MATH) and each time we choose an item we *do not replace* it, so there is one fewer choice at the next stage: we have 4 choices for the first letter (say we choose A), then 3 choices for the second (M, T and H; say we choose H), then 2 choices for the next letter (M and T; say we choose M) and only one choice at the last stage (T). Thus there are $4 \cdot 3 \cdot 2 \cdot 1 = 24$ ways to spell a code worth with the letters MATH.

In this example, we needed to calculate $n \cdot (n - 1) \cdot (n - 2) \cdots 3 \cdot 2 \cdot 1$. This calculation shows up often in mathematics, and is called the **factorial**, and is notated $n!$

Factorial

$n! = n \cdot (n - 1) \cdot (n - 2) \cdots 3 \cdot 2 \cdot 1$

Example 26

How many ways can five different door prizes be distributed among five people?

There are 5 choices of prize for the first person, 4 choices for the second, and so on. The number of ways the prizes can be distributed will be $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ ways.

Now we will consider some slightly different examples.

Example 27

A charity benefit is attended by 25 people and three gift certificates are given away as door prizes: one gift certificate is in the amount of \$100, the second is worth \$25 and the third is worth \$10. Assuming that no person receives more than one prize, how many different ways can the three gift certificates be awarded?

Using the Basic Counting Rule, there are 25 choices for the person who receives the \$100 certificate, 24 remaining choices for the \$25 certificate and 23 choices for the \$10 certificate, so there are $25 \cdot 24 \cdot 23 = 13{,}800$ ways in which the prizes can be awarded.

Example 28

Eight sprinters have made it to the Olympic finals in the 100-meter race. In how many different ways can the gold, silver and bronze medals be awarded?

Using the Basic Counting Rule, there are 8 choices for the gold medal winner, 7 remaining choices for the silver, and 6 for the bronze, so there are $8 \cdot 7 \cdot 6 = 336$ ways the three medals can be awarded to the 8 runners.

Note that in these preceding examples, the gift certificates and the Olympic medals were awarded *without replacement*; that is, once we have chosen a winner of the first door prize or the gold medal, they are not eligible for the other prizes. Thus, at each succeeding stage of the solution there is one fewer choice (25, then 24, then 23 in the first example; 8, then 7, then 6 in the second). Contrast this with the situation of a multiple choice test, where there might be five possible answers — A, B, C, D or E — for each question on the test.

Note also that *the order of selection was important* in each example: for the three door prizes, being chosen first means that you receive substantially more money; in the Olympics example, coming in first means that you get the gold medal instead of the silver or bronze. In each case, if we had chosen the same three people in a different order there might have been a different person who received the \$100 prize, or a different goldmedalist. (Contrast this with the situation where we might draw three names out of a hat to each receive a \$10 gift certificate; in this case the order of selection is *not* important since each of the three people receive the same prize. Situations where the order is *not* important will be discussed in the next section.)

We can generalize the situation in the two examples above to any problem *without replacement* where the *order of selection is important*. If we are arranging in order r items out of n possibilities (instead of 3 out of 25 or 3 out of 8 as in the previous examples), the number of possible arrangements will be given by

$$n \cdot (n-1) \cdot (n-2) \cdots (n-r+1)$$

If you don't see why $(n - r + 1)$ is the right number to use for the last factor, just think back to the first example in this section, where we calculated $25 \cdot 24 \cdot 23$ to get 13,800. In this case $n = 25$ and $r = 3$, so $n - r + 1 = 25 - 3 + 1 = 23$, which is exactly the right number for the final factor.

Now, why would we want to use this complicated formula when it's actually easier to use the Basic Counting Rule, as we did in the first two examples? Well, we won't actually use this formula all that often, we only developed it so that we could attach a special notation and a special definition to this situation where we are choosing r items out of n possibilities *without replacement* and where the *order of selection is important*. In this situation we write:

Permutations

$_nP_r = n \cdot (n-1) \cdot (n-2) \cdots (n-r+1)$

We say that there are $_nP_r$ **permutations** of size r that may be selected from among n choices *without replacement* when *order matters*.

It turns out that we can express this result more simply using factorials.

$$_nP_r = \frac{n!}{(n-r)!}$$

In practicality, we usually use technology rather than factorials or repeated multiplication to compute permutations.

Example 29

I have nine paintings and have room to display only four of them at a time on my wall. How many different ways could I do this?

Since we are choosing 4 paintings out of 9 *without replacement* where the *order of selection is important* there are $_9P_4 = 9 \cdot 8 \cdot 7 \cdot 6 = 3{,}024$ permutations.

Example 30

How many ways can a four-person executive committee (president, vice-president, secretary, treasurer) be selected from a 16-member board of directors of a non-profit organization?

We want to choose 4 people out of 16 without replacement and where the order of selection is important. So the answer is $_{16}P_4 = 16 \cdot 15 \cdot 14 \cdot 13 = 43{,}680$.

Try it Now 7

How many 5 character passwords can be made using the letters A through Z

a. if repeats are allowed

b. if no repeats are allowed

Combinations

In the previous section we considered the situation where we chose r items out of n possibilities *without replacement* and where the *order of selection was important*. We now consider a similar situation in which the order of selection is *not* important.

Example 31

A charity benefit is attended by 25 people at which three \$50 gift certificates are given away as door prizes. Assuming no person receives more than one prize, how many different ways can the gift certificates be awarded?

Using the Basic Counting Rule, there are 25 choices for the first person, 24 remaining choices for the second person and 23 for the third, so there are $25 \cdot 24 \cdot 23 = 13{,}800$ ways to choose three people. Suppose for a moment that Abe is chosen first, Bea second and Cindy third; this is one of the 13,800 possible outcomes. Another way to award the prizes would be to choose Abe first, Cindy second and Bea third; this is another of the 13,800 possible outcomes. But either way Abe, Bea and Cindy each get \$50, so it doesn't really matter the order in which we select them. In how many different orders can Abe, Bea and Cindy be selected? It turns out there are 6:

ABC ACB BAC BCA CAB CBA

How can we be sure that we have counted them all? We are really just choosing 3 people out of 3, so there are $3 \cdot 2 \cdot 1 = 6$ ways to do this; we didn't really need to list them all, we can just use permutations!

So, out of the 13,800 ways to select 3 people out of 25, six of them involve Abe, Bea and Cindy. The same argument works for any other group of three people (say Abe, Bea and David or Frank, Gloria and Hildy) so each three-person group is counted *six times*. Thus the 13,800 figure is six times too big. The number of distinct three-person groups will be $13{,}800/6 = 2300$.

We can generalize the situation in this example above to any problem of choosing a collection of items *without replacement* where the *order of selection is **not** important*. If we are choosing r items out of n possibilities (instead of 3 out of 25 as in the previous examples), the number of possible choices will be given by $\frac{{}_nP_r}{{}_rP_r}$, and we could use this formula for computation. However this situation arises so frequently that we attach a special notation and a special definition to this situation where we are choosing r items out of n possibilities *without replacement* where the *order of selection is **not** important*.

Combinations

$${}_nC_r = \frac{{}_nP_r}{{}_rP_r}$$

We say that there are ${}_nC_r$ **combinations** of size r that may be selected from among n choices *without replacement* where *order doesn't matter*.

We can also write the combinations formula in terms of factorials:

$${}_nC_r = \frac{n!}{(n-r)!r!}$$

Example 32

A group of four students is to be chosen from a 35-member class to represent the class on the student council. How many ways can this be done?

Since we are choosing 4 people out of 35 *without replacement* where the *order of selection is **not** important* there are ${}_{35}C_4 = \dfrac{35 \cdot 34 \cdot 33 \cdot 32}{4 \cdot 3 \cdot 2 \cdot 1}$ = 52,360 combinations.

Try it Now 8

The United States Senate Appropriations Committee consists of 29 members; the Defense Subcommittee of the Appropriations Committee consists of 19 members. Disregarding party affiliation or any special seats on the Subcommittee, how many different 19-member subcommittees may be chosen from among the 29 Senators on the Appropriations Committee?

In the preceding Try it Now problem we assumed that the 19 members of the Defense Subcommittee were chosen without regard to party affiliation. In reality this would never happen: if Republicans are in the majority they would never let a majority of Democrats sit on (and thus control) any subcommittee. (The same of course would be true if the Democrats were in control.) So let's consider the problem again, in a slightly more complicated form:

Example 33

The United States Senate Appropriations Committee consists of 29 members, 15 Republicans and 14 Democrats. The Defense Subcommittee consists of 19 members, 10 Republicans and 9 Democrats. How many different ways can the members of the Defense Subcommittee be chosen from among the 29 Senators on the Appropriations Committee?

In this case we need to choose 10 of the 15 Republicans and 9 of the 14 Democrats. There are ${}_{15}C_{10}$ = 3003 ways to choose the 10 Republicans and ${}_{14}C_9$ = 2002 ways to choose the 9 Democrats. But now what? How do we finish the problem?

Suppose we listed all of the possible 10-member Republican groups on 3003 slips of red paper and all of the possible 9-member Democratic groups on 2002 slips of blue paper. How many ways can we choose one red slip and one blue slip? This is a job for the Basic Counting Rule! We are simply making one choice from the first category and one choice from the second category, just like in the restaurant menu problems from earlier.

There must be 3003 · 2002 = 6,012,006 possible ways of selecting the members of the Defense Subcommittee.

Probability using Permutations and Combinations

We can use permutations and combinations to help us answer more complex probability questions

Example 34

A 4 digit PIN number is selected. What is the probability that there are no repeated digits?

There are 10 possible values for each digit of the PIN (namely: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9), so there are $10 \cdot 10 \cdot 10 \cdot 10 = 10^4 = 10000$ total possible PIN numbers.

To have no repeated digits, all four digits would have to be different, which is selecting without replacement. We could either compute $10 \cdot 9 \cdot 8 \cdot 7$, or notice that this is the same as the permutation ${}_{10}P_4 = 5040$.

The probability of no repeated digits is the number of 4 digit PIN numbers with no repeated digits divided by the total number of 4 digit PIN numbers. This probability is

$$\frac{{}_{10}P_4}{10^4} = \frac{5040}{10000} = 0.504$$

Example 35

In a certain state's lottery, 48 balls numbered 1 through 48 are placed in a machine and six of them are drawn at random. If the six numbers drawn match the numbers that a player had chosen, the player wins \$1,000,000. In this lottery, the order the numbers are drawn in doesn't matter. Compute the probability that you win the million-dollar prize if you purchase a single lottery ticket.

In order to compute the probability, we need to count the total number of ways six numbers can be drawn, and the number of ways the six numbers on the player's ticket could match the six numbers drawn from the machine. Since there is no stipulation that the numbers be in any particular order, the number of possible outcomes of the lottery drawing is ${}_{48}C_6 =$ 12,271,512. Of these possible outcomes, only one would match all six numbers on the player's ticket, so the probability of winning the grand prize is:

$$\frac{{}_{6}C_6}{{}_{48}C_6} = \frac{1}{12271512} \approx 0.0000000815$$

Example 36

In the state lottery from the previous example, if five of the six numbers drawn match the numbers that a player has chosen, the player wins a second prize of \$1,000. Compute the probability that you win the second prize if you purchase a single lottery ticket.

As above, the number of possible outcomes of the lottery drawing is ${}_{48}C_6 = 12{,}271{,}512$. In order to win the second prize, five of the six numbers on the ticket must match five of the six winning numbers; in other words, we must have chosen five of the six winning numbers and one of the 42 losing numbers. The number of ways to choose 5 out of the 6 winning numbers is given by ${}_6C_5 = 6$ and the number of ways to choose 1 out of the 42 losing numbers is given by ${}_{42}C_1 = 42$. Thus the number of favorable outcomes is then given by the Basic Counting Rule: ${}_6C_5 \cdot {}_{42}C_1 = 6 \cdot 42 = 252$. So the probability of winning the second prize is.

$$\frac{({}_6C_5)({}_{42}C_1)}{{}_{48}C_6} = \frac{252}{12271512} \approx 0.0000205$$

Try it Now 9

A multiple-choice question on an economics quiz contains 10 questions with five possible answers each. Compute the probability of randomly guessing the answers and getting 9 questions correct.

Example 37

Compute the probability of randomly drawing five cards from a deck and getting exactly one Ace.

In many card games (such as poker) the order in which the cards are drawn is not important (since the player may rearrange the cards in his hand any way he chooses); in the problems that follow, we will assume that this is the case unless otherwise stated. Thus we use combinations to compute the possible number of 5-card hands, ${}_{52}C_5$. This number will go in the denominator of our probability formula, since it is the number of possible outcomes.

For the numerator, we need the number of ways to draw one Ace and four other cards (none of them Aces) from the deck. Since there are four Aces and we want exactly one of them, there will be ${}_4C_1$ ways to select one Ace; since there are 48 non-Aces and we want 4 of them, there will be ${}_{48}C_4$ ways to select the four non-Aces. Now we use the Basic Counting Rule to calculate that there will be ${}_4C_1 \cdot {}_{48}C_4$ ways to choose one ace and four non-Aces.

Putting this all together, we have

$$P(\text{one Ace}) = \frac{({}_4C_1)({}_{48}C_4)}{{}_{52}C_5} = \frac{778320}{2598960} \approx 0.299$$

Example 38

Compute the probability of randomly drawing five cards from a deck and getting exactly two Aces.

The solution is similar to the previous example, except now we are choosing 2 Aces out of 4 and 3 non-Aces out of 48; the denominator remains the same:

$$P(\text{two Aces}) = \frac{({}_4C_2)({}_{48}C_3)}{{}_{52}C_5} = \frac{103776}{2598960} \approx 0.0399$$

It is useful to note that these card problems are remarkably similar to the lottery problems discussed earlier.

Try it Now 10
Compute the probability of randomly drawing five cards from a deck of cards and getting three Aces and two Kings.

Birthday Problem

Let's take a pause to consider a famous problem in probability theory:

> Suppose you have a room full of 30 people. What is the probability that there is at least one shared birthday?

Take a guess at the answer to the above problem. Was your guess fairly low, like around 10%? That seems to be the intuitive answer (30/365, perhaps?). Let's see if we should listen to our intuition. Let's start with a simpler problem, however.

Example 39

Suppose three people are in a room. What is the probability that there is at least one shared birthday among these three people?

There are a lot of ways there could be at least one shared birthday. Fortunately there is an easier way. We ask ourselves "What is the alternative to having at least one shared birthday?" In this case, the alternative is that there are **no** shared birthdays. In other words, the alternative to "at least one" is having **none**. In other words, since this is a complementary event,

P(at least one) = 1 – P(none)

We will start, then, by computing the probability that there is no shared birthday. Let's imagine that you are one of these three people. Your birthday can be anything without conflict, so there are 365 choices out of 365 for your birthday. What is the probability that the second person does not share your birthday? There are 365 days in the year (let's ignore leap years) and removing your birthday from contention, there are 364 choices that will guarantee that you do not share a birthday with this person, so the probability that the second person does not share your birthday is 364/365. Now we move to the third person. What is the probability that this third person does not have the same birthday as either you or the second person? There are 363 days that will not duplicate your birthday or the second person's, so the probability that the third person does not share a birthday with the first two is 363/365.

We want the second person not to share a birthday with you *and* the third person not to share a birthday with the first two people, so we use the multiplication rule:

$$P(\text{no shared birthday}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \approx 0.9918$$

and then subtract from 1 to get

P(shared birthday) = 1 – P(no shared birthday) = 1 – 0.9918 = 0.0082.

This is a pretty small number, so maybe it makes sense that the answer to our original problem will be small. Let's make our group a bit bigger.

Example 40

Suppose five people are in a room. What is the probability that there is at least one shared birthday among these five people?

Continuing the pattern of the previous example, the answer should be

$$P(\text{shared birthday}) = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \frac{361}{365} \approx 0.0271$$

Note that we could rewrite this more compactly as

$$P(\text{shared birthday}) = 1 - \frac{{}_{365}P_5}{365^5} \approx 0.0271$$

which makes it a bit easier to type into a calculator or computer, and which suggests a nice formula as we continue to expand the population of our group.

Example 41

Suppose 30 people are in a room. What is the probability that there is at least one shared birthday among these 30 people?

Here we can calculate

$$P(\text{shared birthday}) = 1 - \frac{{}_{365}P_{30}}{365^{30}} \approx 0.706$$

which gives us the surprising result that when you are in a room with 30 people there is a 70% chance that there will be at least one shared birthday!

If you like to bet, and if you can convince 30 people to reveal their birthdays, you might be able to win some money by betting a friend that there will be at least two people with the same birthday in the room anytime you are in a room of 30 or more people. (Of course, you would need to make sure your friend hasn't studied probability!) You wouldn't be guaranteed to win, but you should win more than half the time.

This is one of many results in probability theory that is counterintuitive; that is, it goes against our gut instincts. If you still don't believe the math, you can carry out a simulation.

Just so you won't have to go around rounding up groups of 30 people, someone has kindly developed a Java applet so that you can conduct a computer simulation. Go to this web page: http://statweb.stanford.edu/~susan/surprise/Birthday.html, and once the applet has loaded, select 30 birthdays and then keep clicking Start and Reset. If you keep track of the number of times that there is a repeated birthday, you should get a repeated birthday about 7 out of every 10 times you run the simulation.

Try it Now 11

Suppose 10 people are in a room. What is the probability that there is at least one shared birthday among these 10 people?

Expected Value

Expected value is perhaps the most useful probability concept we will discuss. It has many applications, from insurance policies to making financial decisions, and it's one thing that the casinos and government agencies that run gambling operations and lotteries hope most people never learn about.

Example 42

[1]In the casino game roulette, a wheel with 38 spaces (18 red, 18 black, and 2 green) is spun. In one possible bet, the player bets \$1 on a single number. If that number is spun on the wheel, then they receive \$36 (their original \$1 + \$35). Otherwise, they lose their \$1. On average, how much money should a player expect to win or lose if they play this game repeatedly?

Suppose you bet \$1 on each of the 38 spaces on the wheel, for a total of \$38 bet. When the winning number is spun, you are paid \$36 on that number. While you won on that one number, overall you've lost \$2. On a per-space basis, you have "won" -\$2/\$38 ≈ -\$0.053. In other words, on average you lose 5.3 cents per space you bet on.

We call this average gain or loss the expected value of playing roulette. Notice that no one ever loses exactly 5.3 cents: most people (in fact, about 37 out of every 38) lose \$1 and a very few people (about 1 person out of every 38) gain \$35 (the \$36 they win minus the \$1 they spent to play the game).

There is another way to compute expected value without imagining what would happen if we play every possible space. There are 38 possible outcomes when the wheel spins, so the probability of winning is $\frac{1}{38}$. The complement, the probability of losing, is $\frac{37}{38}$.

[1] Photo CC-BY-SA http://www.flickr.com/photos/stoneflower/

Summarizing these along with the values, we get this table:

Outcome	Probability of outcome
\$35	$\frac{1}{38}$
-\$1	$\frac{37}{38}$

Notice that if we multiply each outcome by its corresponding probability we get $\$35 \cdot \frac{1}{38} = 0.9211$ and $-\$1 \cdot \frac{37}{38} = -0.9737$, and if we add these numbers we get 0.9211 + (-0.9737) ≈ -0.053, which is the expected value we computed above.

Expected Value

Expected Value is the average gain or loss of an event if the procedure is repeated many times.

We can compute the expected value by multiplying each outcome by the probability of that outcome, then adding up the products.

Try it Now 12

You purchase a raffle ticket to help out a charity. The raffle ticket costs \$5. The charity is selling 2000 tickets. One of them will be drawn and the person holding the ticket will be given a prize worth \$4000. Compute the expected value for this raffle.

Example 43

In a certain state's lottery, 48 balls numbered 1 through 48 are placed in a machine and six of them are drawn at random. If the six numbers drawn match the numbers that a player had chosen, the player wins \$1,000,000. If they match 5 numbers, then win \$1,000. It costs \$1 to buy a ticket. Find the expected value.

Earlier, we calculated the probability of matching all 6 numbers and the probability of matching 5 numbers:

$$\frac{{}_6C_6}{{}_{48}C_6} = \frac{1}{12271512} \approx 0.0000000815 \text{ for all 6 numbers,}$$

$$\frac{\left({}_6C_5\right)\left({}_{42}C_1\right)}{{}_{48}C_6} = \frac{252}{12271512} \approx 0.0000205 \text{ for 5 numbers.}$$

Our probabilities and outcome values are:

Outcome	Probability of outcome
\$999,999	$\frac{1}{12271512}$
\$999	$\frac{252}{12271512}$
-\$1	$1-\frac{253}{12271512}=\frac{12271259}{12271512}$

The expected value, then is:

$$(\$999{,}999)\cdot\frac{1}{12271512}+(\$999)\cdot\frac{252}{12271512}+(-\$1)\cdot\frac{12271259}{12271512}\approx -\$0.898$$

On average, one can expect to lose about 90 cents on a lottery ticket. Of course, most players will lose \$1.

In general, if the expected value of a game is negative, it is not a good idea to play the game, since on average you will lose money. It would be better to play a game with a positive expected value (good luck trying to find one!), although keep in mind that even if the *average* winnings are positive it could be the case that most people lose money and one very fortunate individual wins a great deal of money. If the expected value of a game is 0, we call it a **fair game**, since neither side has an advantage.

Not surprisingly, the expected value for casino games is negative for the player, which is positive for the casino. It must be positive or they would go out of business. Players just need to keep in mind that when they play a game repeatedly, their expected value is negative. That is fine so long as you enjoy playing the game and think it is worth the cost. But it would be wrong to expect to come out ahead.

Try it Now 13

A friend offers to play a game, in which you roll 3 standard 6-sided dice. If all the dice roll different values, you give him \$1. If any two dice match values, you get \$2. What is the expected value of this game? Would you play?

Expected value also has applications outside of gambling. Expected value is very common in making insurance decisions.

Example 44

A 40-year-old man in the U.S. has a 0.242% risk of dying during the next year[2]. An insurance company charges $275 for a life-insurance policy that pays a $100,000 death benefit. What is the expected value for the person buying the insurance?

The probabilities and outcomes are

Outcome	Probability of outcome
$100,000 - $275 = $99,725	0.00242
-$275	1 – 0.00242 = 0.99758

The expected value is ($99,725)(0.00242) + (-$275)(0.99758) = -$33.

Not surprisingly, the expected value is negative; the insurance company can only afford to offer policies if they, on average, make money on each policy. They can afford to pay out the occasional benefit because they offer enough policies that those benefit payouts are balanced by the rest of the insured people.

For people buying the insurance, there is a negative expected value, but there is a security that comes from insurance that is worth that cost.

Try it Now Answers

1. There are 60 possible readings, from 00 to 59. a. $\frac{1}{60}$ b. $\frac{16}{60}$ (counting 00 through 15)

2. Since the second draw is made after replacing the first card, these events are independent. The probability of an ace on each draw is $\frac{4}{52}=\frac{1}{13}$, so the probability of an Ace on both draws is $\frac{1}{13}\cdot\frac{1}{13}=\frac{1}{169}$

3. P(white sock and white tee) = $\frac{6}{10}\cdot\frac{3}{7}=\frac{9}{35}$

   P(white sock or white tee) = $\frac{6}{10}+\frac{3}{7}-\frac{9}{35}=\frac{27}{35}$

4. a. $\frac{6}{10}\cdot\frac{5}{9}=\frac{30}{90}=\frac{1}{3}$

[2] According to the estimator at http://www.numericalexample.com/index.php?view=article&id=91

5. Out of 100,000 people, 500 would have the disease. Of those, all 500 would test positive. Of the 99,500 without the disease, 2,985 would falsely test positive and the other 96,515 would test negative.

$$P(\text{disease} \mid \text{positive}) = \frac{500}{500+2985} = \frac{500}{3485} \approx 14.3\%$$

6. $8 \cdot 11 \cdot 5 = 440$ menu combinations

7. There are 26 characters. a. $26^5 = 11{,}881{,}376$. b. ${}_{26}P_5 = 26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 = 7{,}893{,}600$

8. Order does not matter. ${}_{29}C_{19} = 20{,}030{,}010$ possible subcommittees

9. There are $5^{10} = 9{,}765{,}625$ different ways the exam can be answered. There are 9 possible locations for the one missed question, and in each of those locations there are 4 wrong answers, so there are 36 ways the test could be answered with one wrong answer.

$$P(\text{9 answers correct}) = \frac{36}{5^{10}} \approx 0.0000037 \text{ chance}$$

10. $P(\text{three Aces and two Kings}) = \dfrac{({}_4C_3)({}_4C_2)}{{}_{52}C_5} = \dfrac{24}{2598960} \approx 0.0000092$

11. $P(\text{shared birthday}) = 1 - \dfrac{{}_{365}P_{10}}{365^{10}} \approx 0.117$

12. $(\$3{,}995) \cdot \dfrac{1}{2000} + (-\$5) \cdot \dfrac{1999}{2000} \approx -\$3.00$

13. Suppose you roll the first die. The probability the second will be different is $\frac{5}{6}$. The probability that the third roll is different than the previous two is $\frac{4}{6}$, so the probability that the three dice are different is $\frac{5}{6} \cdot \frac{4}{6} = \frac{20}{36}$. The probability that two dice will match is the complement, $1 - \frac{20}{36} = \frac{16}{36}$.

The expected value is: $(\$2) \cdot \frac{16}{36} + (-\$1) \cdot \frac{20}{36} = \frac{12}{36} \approx \$0.33$. Yes, it is in your advantage to play. On average, you'd win \$0.33 per play.

Exercises

1. A ball is drawn randomly from a jar that contains 6 red balls, 2 white balls, and 5 yellow balls. Find the probability of the given event.
    a. A red ball is drawn
    b. A white ball is drawn

2. Suppose you write each letter of the alphabet on a different slip of paper and put the slips into a hat. What is the probability of drawing one slip of paper from the hat at random and getting:
    a. A consonant
    b. A vowel

3. A group of people were asked if they had run a red light in the last year. 150 responded "yes", and 185 responded "no". Find the probability that if a person is chosen at random, they have run a red light in the last year.

4. In a survey, 205 people indicated they prefer cats, 160 indicated they prefer dots, and 40 indicated they don't enjoy either pet. Find the probability that if a person is chosen at random, they prefer cats.

5. Compute the probability of tossing a six-sided die (with sides numbered 1 through 6) and getting a 5.

6. Compute the probability of tossing a six-sided die and getting a 7.

7. Giving a test to a group of students, the grades and gender are summarized below. If one student was chosen at random, find the probability that the student was female.

	A	B	C	Total
Male	8	18	13	39
Female	10	4	12	26
Total	18	22	25	65

8. The table below shows the number of credit cards owned by a group of individuals. If one person was chosen at random, find the probability that the person had no credit cards.

	Zero	One	Two or more	Total
Male	9	5	19	33
Female	18	10	20	48
Total	27	15	39	81

9. Compute the probability of tossing a six-sided die and getting an even number.

10. Compute the probability of tossing a six-sided die and getting a number less than 3.

11. If you pick one card at random from a standard deck of cards, what is the probability it will be a King?

12. If you pick one card at random from a standard deck of cards, what is the probability it will be a Diamond?

13. Compute the probability of rolling a 12-sided die and getting a number other than 8.

14. If you pick one card at random from a standard deck of cards, what is the probability it is not the Ace of Spades?

15. Referring to the grade table from question #7, what is the probability that a student chosen at random did NOT earn a C?

16. Referring to the credit card table from question #8, what is the probability that a person chosen at random has at least one credit card?

17. A six-sided die is rolled twice. What is the probability of showing a 6 on both rolls?

18. A fair coin is flipped twice. What is the probability of showing heads on both flips?

19. A die is rolled twice. What is the probability of showing a 5 on the first roll and an even number on the second roll?

20. Suppose that 21% of people own dogs. If you pick two people at random, what is the probability that they both own a dog?

21. Suppose a jar contains 17 red marbles and 32 blue marbles. If you reach in the jar and pull out 2 marbles at random, find the probability that both are red.

22. Suppose you write each letter of the alphabet on a different slip of paper and put the slips into a hat. If you pull out two slips at random, find the probability that both are vowels.

23. Bert and Ernie each have a well-shuffled standard deck of 52 cards. They each draw one card from their own deck. Compute the probability that:
    a. Bert and Ernie both draw an Ace.
    b. Bert draws an Ace but Ernie does not.
    c. neither Bert nor Ernie draws an Ace.
    d. Bert and Ernie both draw a heart.
    e. Bert gets a card that is not a Jack and Ernie draws a card that is not a heart.

24. Bert has a well-shuffled standard deck of 52 cards, from which he draws one card; Ernie has a 12-sided die, which he rolls at the same time Bert draws a card. Compute the probability that:
    a. Bert gets a Jack and Ernie rolls a five.
    b. Bert gets a heart and Ernie rolls a number less than six.
    c. Bert gets a face card (Jack, Queen or King) and Ernie rolls an even number.
    d. Bert gets a red card and Ernie rolls a fifteen.
    e. Bert gets a card that is not a Jack and Ernie rolls a number that is not twelve.

25. Compute the probability of drawing a King from a deck of cards and then drawing a Queen.

26. Compute the probability of drawing two spades from a deck of cards.

27. A math class consists of 25 students, 14 female and 11 male. Two students are selected at random to participate in a probability experiment. Compute the probability that
    a. a male is selected, then a female.
    b. a female is selected, then a male.
    c. two males are selected.
    d. two females are selected.
    e. no males are selected.

28. A math class consists of 25 students, 14 female and 11 male. Three students are selected at random to participate in a probability experiment. Compute the probability that
    a. a male is selected, then two females.
    b. a female is selected, then two males.
    c. two females are selected, then one male.
    d. three males are selected.
    e. three females are selected.

29. Giving a test to a group of students, the grades and gender are summarized below. If one student was chosen at random, find the probability that the student was female and earned an A.

	A	B	C	Total
Male	8	18	13	39
Female	10	4	12	26
Total	18	22	25	65

30. The table below shows the number of credit cards owned by a group of individuals. If one person was chosen at random, find the probability that the person was male and had two or more credit cards.

	Zero	One	Two or more	Total
Male	9	5	19	33
Female	18	10	20	48
Total	27	15	39	81

31. A jar contains 6 red marbles numbered 1 to 6 and 8 blue marbles numbered 1 to 8. A marble is drawn at random from the jar. Find the probability the marble is red or odd-numbered.

32. A jar contains 4 red marbles numbered 1 to 4 and 10 blue marbles numbered 1 to 10. A marble is drawn at random from the jar. Find the probability the marble is blue or even-numbered.

33. Referring to the table from #29, find the probability that a student chosen at random is female or earned a B.

34. Referring to the table from #30, find the probability that a person chosen at random is male or has no credit cards.

35. Compute the probability of drawing the King of hearts or a Queen from a deck of cards.

36. Compute the probability of drawing a King or a heart from a deck of cards.

37. A jar contains 5 red marbles numbered 1 to 5 and 8 blue marbles numbered 1 to 8. A marble is drawn at random from the jar. Find the probability the marble is
    a. Even-numbered given that the marble is red.
    b. Red given that the marble is even-numbered.

38. A jar contains 4 red marbles numbered 1 to 4 and 8 blue marbles numbered 1 to 8. A marble is drawn at random from the jar. Find the probability the marble is
    a. Odd-numbered given that the marble is blue.
    b. Blue given that the marble is odd-numbered.

39. Compute the probability of flipping a coin and getting heads, given that the previous flip was tails.

40. Find the probability of rolling a "1" on a fair die, given that the last 3 rolls were all ones.

41. Suppose a math class contains 25 students, 14 females (three of whom speak French) and 11 males (two of whom speak French). Compute the probability that a randomly selected student speaks French, given that the student is female.

42. Suppose a math class contains 25 students, 14 females (three of whom speak French) and 11 males (two of whom speak French). Compute the probability that a randomly selected student is male, given that the student speaks French.

43. A certain virus infects one in every 400 people. A test used to detect the virus in a person is positive 90% of the time if the person has the virus and 10% of the time if the person does not have the virus. Let A be the event "the person is infected" and B be the event "the person tests positive".
    a. Find the probability that a person has the virus given that they have tested positive, i.e. find P(A | B).
    b. Find the probability that a person does not have the virus given that they test negative, i.e. find P(not A | not B).

44. A certain virus infects one in every 2000 people. A test used to detect the virus in a person is positive 96% of the time if the person has the virus and 4% of the time if the person does not have the virus. Let A be the event "the person is infected" and B be the event "the person tests positive".
    a. Find the probability that a person has the virus given that they have tested positive, i.e. find P(A | B).
    b. Find the probability that a person does not have the virus given that they test negative, i.e. find P(not A | not B).

45. A certain disease has an incidence rate of 0.3%. If the false negative rate is 6% and the false positive rate is 4%, compute the probability that a person who tests positive actually has the disease.

46. A certain disease has an incidence rate of 0.1%. If the false negative rate is 8% and the false positive rate is 3%, compute the probability that a person who tests positive actually has the disease.

47. A certain group of symptom-free women between the ages of 40 and 50 are randomly selected to participate in mammography screening. The incidence rate of breast cancer among such women is 0.8%. The false negative rate for the mammogram is 10%. The false positive rate is 7%. If a the mammogram results for a particular woman are positive (indicating that she has breast cancer), what is the probability that she actually has breast cancer?

48. About 0.01% of men with no known risk behavior are infected with HIV. The false negative rate for the standard HIV test 0.01% and the false positive rate is also 0.01%. If a randomly selected man with no known risk behavior tests positive for HIV, what is the probability that he is actually infected with HIV?

49. A boy owns 2 pairs of pants, 3 shirts, 8 ties, and 2 jackets. How many different outfits can he wear to school if he must wear one of each item?

50. At a restaurant you can choose from 3 appetizers, 8 entrees, and 2 desserts. How many different three-course meals can you have?

51. How many three-letter "words" can be made from 4 letters "FGHI" if
    a. repetition of letters is allowed
    b. repetition of letters is not allowed

52. How many four-letter "words" can be made from 6 letters "AEBWDP" if
    a. repetition of letters is allowed
    b. repetition of letters is not allowed

53. All of the license plates in a particular state feature three letters followed by three digits (e.g. ABC 123). How many different license plate numbers are available to the state's Department of Motor Vehicles?

54. A computer password must be eight characters long. How many passwords are possible if only the 26 letters of the alphabet are allowed?

55. A pianist plans to play 4 pieces at a recital. In how many ways can she arrange these pieces in the program?

56. In how many ways can first, second, and third prizes be awarded in a contest with 210 contestants?

57. Seven Olympic sprinters are eligible to compete in the 4 x 100 m relay race for the USA Olympic team. How many four-person relay teams can be selected from among the seven athletes?

58. A computer user has downloaded 25 songs using an online file-sharing program and wants to create a CD-R with ten songs to use in his portable CD player. If the order that the songs are placed on the CD-R is important to him, how many different CD-Rs could he make from the 25 songs available to him?

59. In western music, an octave is divided into 12 pitches. For the film *Close Encounters of the Third Kind*, director Steven Spielberg asked composer John Williams to write a five-note theme, which aliens would use to communicate with people on Earth. Disregarding rhythm and octave changes, how many five-note themes are possible if no note is repeated?

60. In the early twentieth century, proponents of the Second Viennese School of musical composition (including Arnold Schönberg, Anton Webern and Alban Berg) devised the twelve-tone technique, which utilized a tone row consisting of all 12 pitches from the chromatic scale in any order, but with not pitches repeated in the row. Disregarding rhythm and octave changes, how many tone rows are possible?

61. In how many ways can 4 pizza toppings be chosen from 12 available toppings?

62. At a baby shower 17 guests are in attendance and 5 of them are randomly selected to receive a door prize. If all 5 prizes are identical, in how many ways can the prizes be awarded?

63. In the 6/50 lottery game, a player picks six numbers from 1 to 50. How many different choices does the player have if order doesn't matter?

64. In a lottery daily game, a player picks three numbers from 0 to 9. How many different choices does the player have if order doesn't matter?

65. A jury pool consists of 27 people. How many different ways can 11 people be chosen to serve on a jury and one additional person be chosen to serve as the jury foreman?

66. The United States Senate Committee on Commerce, Science, and Transportation consists of 23 members, 12 Republicans and 11 Democrats. The Surface Transportation and Merchant Marine Subcommittee consists of 8 Republicans and 7 Democrats. How many ways can members of the Subcommittee be chosen from the Committee?

67. You own 16 CDs. You want to randomly arrange 5 of them in a CD rack. What is the probability that the rack ends up in alphabetical order?

68. A jury pool consists of 27 people, 14 men and 13 women. Compute the probability that a randomly selected jury of 12 people is all male.

69. In a lottery game, a player picks six numbers from 1 to 48. If 5 of the 6 numbers match those drawn, they player wins second prize. What is the probability of winning this prize?

70. In a lottery game, a player picks six numbers from 1 to 48. If 4 of the 6 numbers match those drawn, they player wins third prize. What is the probability of winning this prize?

71. Compute the probability that a 5-card poker hand is dealt to you that contains all hearts.

72. Compute the probability that a 5-card poker hand is dealt to you that contains four Aces.

73. A bag contains 3 gold marbles, 6 silver marbles, and 28 black marbles. Someone offers to play this game: You randomly select on marble from the bag. If it is gold, you win \$3. If it is silver, you win \$2. If it is black, you lose \$1. What is your expected value if you play this game?

74. A friend devises a game that is played by rolling a single six-sided die once. If you roll a 6, he pays you \$3; if you roll a 5, he pays you nothing; if you roll a number less than 5, you pay him \$1. Compute the expected value for this game. Should you play this game?

75. In a lottery game, a player picks six numbers from 1 to 23. If the player matches all six numbers, they win 30,000 dollars. Otherwise, they lose \$1. Find the expected value of this game.

76. A game is played by picking two cards from a deck. If they are the same value, then you win \$5, otherwise you lose \$1. What is the expected value of this game?

77. A company estimates that 0.7% of their products will fail after the original warranty period but within 2 years of the purchase, with a replacement cost of \$350. If they offer a 2 year extended warranty for \$48, what is the company's expected value of each warranty sold?

78. An insurance company estimates the probability of an earthquake in the next year to be 0.0013. The average damage done by an earthquake it estimates to be \$60,000. If the company offers earthquake insurance for \$100, what is their expected value of the policy?

Exploration

Some of these questions were adapted from puzzles at mindyourdecisions.com.

79. A small college has been accused of gender bias in its admissions to graduate programs.
    a. Out of 500 men who applied, 255 were accepted. Out of 700 women who applied, 240 were accepted. Find the acceptance rate for each gender. Does this suggest bias?
    b. The college then looked at each of the two departments with graduate programs, and found the data below. Compute the acceptance rate within each department by gender. Does this suggest bias?

Department	Men		Women	
	Applied	Admitted	Applied	Admitted
Dept A	400	240	100	90
Dept B	100	15	600	150

    c. Looking at our results from Parts *a* and *b*, what can you conclude? Is there gender bias in this college's admissions? If so, in which direction?

80. A bet on "black" in Roulette has a probability of 18/38 of winning. If you win, you double your money. You can bet anywhere from \$1 to \$100 on each spin.
    a. Suppose you have \$10, and are going to play until you go broke or have \$20. What is your best strategy for playing?
    b. Suppose you have \$10, and are going to play until you go broke or have \$30. What is your best strategy for playing?

81. Your friend proposes a game: You flip a coin. If it's heads, you win \$1. If it's tails, you lose \$1. However, you are worried the coin might not be fair coin. How could you change the game to make the game fair, without replacing the coin?

82. Fifty people are in a line. The first person in the line to have a birthday matching someone in front of them will win a prize. Of course, this means the first person in the line has no chance of winning. Which person has the highest likelihood of winning?

83. Three people put their names in a hat, then each draw a name, as part of a randomized gift exchange. What is the probability that no one draws their own name? What about with four people?

84. How many different "words" can be formed by using all the letters of each of the following words exactly once?
    a. "ALICE"
    b. "APPLE"

85. How many different "words" can be formed by using all the letters of each of the following words exactly once?
    a. "TRUMPS"
    b. "TEETER"

86. The *Monty Hall problem* is named for the host of the game show *Let's make a Deal*. In this game, there would be three doors, behind one of which there was a prize. The contestant was asked to choose one of the doors. Monty Hall would then open one of the other doors to show there was no prize there. The contestant was then asked if they wanted to stay with their original door, or switch to the other unopened door. Is it better to stay or switch, or does it matter?

87. Suppose you have two coins, where one is a fair coin, and the other coin comes up heads 70% of the time. What is the probability you have the fair coin given each of the following outcomes from a series of flips?
    a. 5 Heads and 0 Tails
    b. 8 Heads and 3 Tails
    c. 10 Heads and 10 Tails
    d. 3 Heads and 8 Tails

88. Suppose you have six coins, where five are fair coins, and one coin comes up heads 80% of the time. What is the probability you have a fair coin given each of the following outcomes from a series of flips?
    a. 5 Heads and 0 Tails
    b. 8 Heads and 3 Tails
    c. 10 Heads and 10 Tails
    d. 3 Heads and 8 Tails

89. In this problem, we will explore probabilities from a series of events.
    a. If you flip 20 coins, how many would you *expect* to come up "heads", on average? Would you expect *every* flip of 20 coins to come up with exactly that many heads?
    b. If you were to flip 20 coins, what would you consider a "usual" result? An "unusual" result?
    c. Flip 20 coins (or one coin 20 times) and record how many come up "heads". Repeat this experiment 9 more times. Collect the data from the entire class.
    d. When flipping 20 coins, what is the theoretic probability of flipping 20 heads?
    e. Based on the class's experimental data, what appears to be the probability of flipping 10 heads out of 20 coins?
    f. The formula ${}_nC_x p^x (1-p)^{n-x}$ will compute the probability of an event with probability p occurring x times out of n, such as flipping x heads out of n coins where the probability of heads is $p = ½$. Use this to compute the theoretic probability of flipping 10 heads out of 20 coins.
    g. If you were to flip 20 coins, based on the class's experimental data, what range of values would you consider a "usual" result? What is the combined probability of these results? What would you consider an "unusual" result? What is the combined probability of these results?
    h. We'll now consider a simplification of a case from the 1960s. In the area, about 26% of the jury eligible population was black. In the court case, there were 100 men on the juror panel, of which 8 were black. Does this provide evidence of racial bias in jury selection?