DISCRETE PROBABILITY

How do we work with probabilities when the number of possible outcomes is countable?

The French mathematician Laplace once claimed that probability theory is nothing but “common sense reduced to calculation.” This is true many times, but not always ….

LESSON 11: BASIC PROBABILITY

How Likely is an Event?

PART A: EVENTS and PROBABILITIES

An event is something that occurs or doesn’t occur.

• We use capital letters such as A, B, C, etc. to denote events.

• An event A could be “it rains in my neighborhood today,” or “the die comes up a ‘4’ after it is rolled,” etc.

Let \( P(A) \) = the probability of event A occurring.

\[ P(A) \text{ must be a real number between 0 and 1, inclusive:} \]

\[ 0 \leq P(A) \leq 1 \]

Here is a scale for \( P(A) \):
PART B: COMPLEMENTARY EVENTS

If the event \( A \) does not occur, then the event “not \( A \)” occurs.

The event “not \( A \)” is called the complement of \( A \), denoted by \( \bar{A} \) or \( A^C \).

\[
P(\text{not } A) = P(\bar{A}) = P(A^C) = 1 - P(A)
\]

• This is the probability of event \( A \) not occurring.

Example 1 (Complementary Events)

If the probability that it rains in my neighborhood today is 0.3, then the probability that it does not rain in my neighborhood today is:

\[1 - 0.3 = 0.7\]

PART C: Rounding

• Probabilities may be written as fractions or as decimal numbers. (Using percents is more controversial.)

• We often round off probabilities to three significant digits (or significant figures – “sig figs”), such as for 0.333 (33.3%) or 0.00703 (0.703%).

• Remember that “leading zeros” don’t count when counting “sig figs.” Note that 0.00703 has five decimal places (digits after the decimal point).

\[0.00703\]

leading zeros (red)
don’t count for “sig figs”

• If final answers are rounded off to, say, three “sig figs,” then intermediate results should be either exact or rounded off to more “sig figs” (at least twice as many, six, perhaps).

• You may want more than three “sig figs” if a probability is close to 1, such as 0.9999987, or if it is very close to a related probability in a problem.
PART D: THREE APPROACHES TO PROBABILITY

Approach 1): Classical Approach

- A trial is an experiment such as rolling a die, flipping a fair coin, etc.

- Assume that the trial must result in exactly one of \( N \) equally likely outcomes (we’ll say “elos”) that are simple events, which can’t really be broken down further.

- The elos make up the sample space, \( S \).

- Then, for an event \( A \), \( P(A) = \frac{\# \text{ of elos for which } A \text{ occurs}}{N} \).

Example 2 (Roll One Die)

We roll one standard six-sided die.

\( S = \{1, 2, 3, 4, 5, 6\} \). This is the sample space. 
\( N = 6 \) elos.

As a shorthand, we will say that the event “4” means “the die comes up ‘4’ when it is rolled.” It is a simple event.

\[ P(4) = \frac{1}{6}. \]

\[ \begin{array}{cccc}
1 & 2 & 3 & 4 \\
\hline
\hline
5 & 6
\end{array} \]

- The odds against the event happening are 5 to 1.

\[ P(\text{even}) = \frac{3}{6} = \frac{1}{2}. \]

\[ \begin{array}{ccc}
2 & 4 & 6 \\
\hline
\hline
\end{array} \]

- “Even” is a compound event, because it combines two or more simple events. \( \S \)
Example 3 (Roll Two Dice – 1 red, 1 green)

We roll two standard six-sided dice.

- It turns out to be far more convenient if we distinguish between the dice. For example, we can color one die red and the other die green.

Think About It: Are dice totals equally likely?

Our simple events are ordered pairs of the form \((r, g)\), where \(r\) is the result on the red die and \(g\) is the result on the green die.

\[
S = \{(1,1), (1,2), \ldots, (6,6)\}.
\]

\(N = 36\) elos.

\[
P((4,2)) = \frac{1}{36}
\]

\[
P(2\text{ on red}) = \frac{6}{36} = \frac{1}{6}
\]

Think About It: Why does this answer make sense? How could you have answered very quickly?
We get **doubles** if both dice come up the **same number**.

\[ P(\text{doubles}) = \frac{6}{36} = \frac{1}{6} \]

**Think About It:** Assume that the red die is rolled first. Regardless of the result, what is \( P(g \text{ matches } r) \)?

\[ P(\text{total of 5}) = \frac{4}{36} = \frac{1}{9} \approx 11.1\% \]
**Example 4 (Roulette)**

18 red slots \[ P(\text{red}) = \frac{18}{38} = \frac{9}{19} \approx 47.4\% \]

18 black slots \[ P(\text{black}) = \frac{18}{38} = \frac{9}{19} \approx 47.4\% \]

2 green slots \[ P(\text{green}) = \frac{2}{38} = \frac{1}{19} \approx 5.26\% \]

38 total

- The casino pays “even money” for red / black bets; in other words, every dollar bet is matched by the casino if the player wins. This is slightly unfair to the player! A fair payoff would be about $1.11 for every dollar bet. But then, the casino wouldn’t be making a profit!

- According to the Law of Large Numbers (LLN), the probabilities built into the game will “crystallize” in the long run. In other words, the observed relative frequencies of red, black, and green results will approach the corresponding theoretical probabilities \( \left(\frac{9}{19}, \frac{9}{19}, \text{ and } \frac{1}{19}\right) \) after many bets. In the long run, on average, the casino will make a profit of about 5.26 cents for every dollar bet by players in red / black bets. We say that the house advantage is 5.26%.

**Think About It**: How do you maximize your chances of winning a huge amount of money?
Approach 2): Frequentist / Empirical Approach

Let $N$ = the total number of trials observed.

If $N$ is “large,” then:

$$P(A) = \frac{\text{# of trials in which } A \text{ occurred}}{N}$$

Example 5 (A Magician's Coin)

A magician’s coin comes up heads (H) 255 times and tails (T) 245 times. Estimate $P(H)$ for the coin, where

$$P(H) = P(\text{coin comes up "heads" when flipped once}).$$

§ Solution

First, $N = 255 + 245 = 500$.

$$P(H) \approx \frac{255}{500} = 0.51$$

**WARNING 1:** We find $N$ and divide by $N$ here. Divide 255 by 500, not by 245. Why can’t $\frac{255}{245}$ be a probability, anyway?

• If the coin is, in fact, fair, then our estimate for $P(H)$ will approach $\frac{1}{2}$, or 0.5, if we flip the coin more and more times; that is, as $N$ gets larger and approaches infinity. This is a consequence of the Law of Large Numbers (LLN).
Approach 3): Subjective Approach

Probabilities can be (hopefully educated) guesstimates.

• For example, what is your estimate for:
  \[ P(\text{your favorite candidate will win the next presidential election}) \]?

An adjustable color wheel may help you visualize probabilities.
LESSON 12: “OR” PROBLEMS IN PROBABILITY

When Do We Add (or Subtract) Probabilities?

PART A: PROBABILITIES FOR MUTUALLY EXCLUSIVE EVENTS (“MEES”)

Example 1 (Mutually Exclusive Events - “Mees”)

Roll one die.

\[
P\left( \text{even or 5} \right) = \frac{\text{# of "elos" for which } A \text{ or } B \text{ occurs}}{N} = \frac{4}{6} = \frac{2}{3}
\]

Here, events \( A \) and \( B \) are disjoint, or mutually exclusive events (we’ll say “mees”) in that they can’t both occur on the same trial.

- If one event occurs, it excludes the possibility that the other event can also occur with it.
- Here, a die can’t come up “even” and “5” on the same roll (that is, for the same trial).

Addition Rule for “Mees”

If events \( A \) and \( B \) are “mees,” then \( P(A \text{ or } B) = P(A) + P(B) \).

Here,

\[
P(\text{even or 5}) = P(\text{even}) + P(5) = \frac{3}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}
\]
PART B: PROBABILITIES FOR NON-“MEES”

Example 2 (Non-“Mees”)

Roll one die.

\[
P\left(\text{even or higher than } 2\right) = \frac{\# \text{ of "elos" for which } A \text{ or } B \text{ occurs}}{N}
\]

\[= \frac{5}{6}\]

Here, events \(A\) and \(B\) are not “mees.” Both events can happen on the same trial, namely if the die comes up a “4” or a “6” on the roll.

**WARNING 1:** Do not double-count the “elos” (“4” and “6” here) for which both \(A\) and \(B\) occur. These “elos” make up the intersection of \(A\) and \(B\), or simply \(A \text{ and } B\)”; this is denoted by \(A \cap B\). (“\(A \text{ or } B\)” is denoted by \(A \cup B\).)

### General Addition Rule

For any two events \(A\) and \(B\), \(P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)\).

Here,

\[
P(\text{even or higher than 2}) = P(\text{even}) + P(\text{higher than 2}) - P(\text{both})
\]

\[= \frac{3}{6} + \frac{4}{6} - \frac{2}{6}
\]

\[= \frac{5}{6}\]
Subtracting \( P(A \text{ and } B) \) adjusts for the double-counting from the first two terms. Consider the following Venn Diagram:

![Venn Diagram](image)

**Note:** The General Addition Rule works even if \( A \) and \( B \) are “mees.”

In that special case, \( P(A \text{ and } B) = 0 \), and the rule boils down to:

\[
P(A \text{ or } B) = P(A) + P(B),
\]

as it should. Consider:

![Venn Diagram](image)

**Challenge:** What would be the General Addition Rule for \( P(A \text{ or } B \text{ or } C) \)?
**PART C: USING TABLES TO FIND PROBABILITIES**

**Example 3 (Using a Table for Two Dice)**

Roll two dice.

\[
P(\text{doubles or a "6" on either die}) = \frac{16}{36} = \frac{4}{9} \approx 44.4\%
\]

A formula would be tricky to apply here. We’ll use a diagram, instead.

![Diagram of dice rolls]

\[\]

**Example 4 (Using a Two-Way Frequency (or Contingency) Table)**

All 26 students in a class are passing, but they seek tutoring. One tutor is a Junior who likes helping students who are Juniors or who are receiving a “B” or a “C.” Based on the following two-way frequency (or contingency) table, find the probability that a random student in the class is a Junior or is getting a “B” or a “C.”

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Juniors</td>
<td>5</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>
§ Solution

We will boldface the entries that correspond to Juniors or students receiving “B”s or “C”s. Their sum is 23.

- **Think About It:** What’s an easy way of determining that the sum is 23, aside from adding the boldfaced entries directly?

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Juniors</td>
<td>5</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ P(\text{a random student in the class is a Junior or is getting a "B" or a "C"}) \]

\[ = \frac{23}{26} \]

\[ \approx 0.885 \ (\text{or} \ 88.5\% ) \]

**WARNING 2: Abusing totals.** Although we sometimes want to write down row and column totals, they may be misleading. In the following table, you should be warned against adding 14, 12, and 6. Why would that be wrong?

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Juniors</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>N=26</td>
</tr>
</tbody>
</table>

§
Lesson 13: “And” Problems in Probability

When Do We Multiply Probabilities?

Part A: Probabilities for Independent Events

Example 1 (Independent Events: Card Rank and Suit)

Pick (or “draw”) a card from a standard deck of 52 cards with no Jokers. (Know this setup!)

- **Hearts**
- **Diamonds**
- **Clubs**
- **Spades**

(The 52 cards are equally likely.)

\[
P(3) = \frac{4}{52} = \frac{1}{13} \quad \text{(The 13 ranks are equally likely.)}
\]

\[
P(\text{hearts}) = \frac{13}{52} = \frac{1}{4} \quad \text{(The 4 suits are equally likely.)}
\]

\[
P(3 \text{ and hearts}) = \frac{1}{52} \quad \text{(The 52 cards are equally likely.)}
\]
The events “3” and “Hearts” are independent events, because knowing the rank of a card tells us nothing about its suit, and vice-versa. The occurrence of one event does not change the probability of the other event occurring. The rank and the suit of an unknown card are independent random variables, which we will discuss later.

Dependent events are events that are not independent.

**Multiplication Rule for Independent Events**

If events $A$, $B$, $C$, etc. are independent, then:

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

\[
P(A \text{ and } B \text{ and } C) = P(A) \cdot P(B) \cdot P(C), \text{ etc.}
\]

Because “3” and “Hearts” are independent events,

\[
P(3 \text{ and hearts}) = P(3) \cdot P(\text{hearts})
\]

\[
= \frac{1}{13} \cdot \frac{1}{4}
\]

\[
= \frac{1}{52}
\]

The tree diagram below shows how probabilities can be multiplied along the branches in a path.
Example 2 (Independent Events: Drawing Cards With Replacement)

Draw three cards from a standard deck with replacement, meaning that, after we draw a card, we place it back in the deck before we draw the next card.

Find the probability that we draw an Ace first, a King second, and a King third. Think: AKK sequence.

§ Solution

Because we are drawing cards with replacement, the draws are independent trials. The probabilities involved with one draw are not impacted by the results of the other draws.

\[
\]

\[
= \frac{1}{13} \cdot \frac{1}{13} \cdot \frac{1}{13}
\]

\[
= \frac{1}{2197} \approx 0.000455
\]

Tree Diagram

PART B: CONDITIONAL PROBABILITY

\[
P(B \mid A) = \text{the conditional probability (think: updated probability) that } B \text{ occurs, given that } A \text{ occurs.}
\]

- The idea of “updating” probabilities is a popular idea among Bayesian statisticians, though it is controversial.
PART C: PROBABILITIES FOR DEPENDENT EVENTS

General Multiplication Rule

For events $A$, $B$, $C$, etc.,

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

$$P(A \text{ and } B \text{ and } C) = P(A) \cdot P(B \mid A) \cdot P(C \mid A \text{ and } B), \text{ etc.}$$

- If $A$ and $B$ are independent, then the occurrence of $A$ does not change the updated probability of $B$ occurring, and $P(B \mid A) = P(B)$. In that case, the General Multiplication Rule becomes the Multiplication Rule for Independent Events that we discussed earlier: $P(A \text{ and } B) = P(A) \cdot P(B)$.

Example 3 (Dependent Events: Drawing Cards Without Replacement)

Draw three cards from a standard deck without replacement, meaning that drawn cards are never returned back to the deck.

Find the probability that we draw an Ace first, a King second, and a King third. Think: AKK sequence.

Solution

Because we are drawing cards without replacement, previous draws affect probabilities on later draws, and the draws are dependent trials.

$$P(A-1\text{st and K-2nd and K-3rd})$$

$$= P(A-1\text{st}) \cdot P(K-2\text{nd} \mid A-1\text{st}) \cdot P(K-3\text{rd} \mid A-1\text{st and K-2nd})$$

$$= \frac{4}{52} \cdot \frac{4}{51} \cdot \frac{3}{50} \cdot 51\text{ cards,} \quad 50\text{ cards,}$$

$$4\text{ Ks} \quad 3\text{ Ks}$$

$$= 0.000362$$

Think About It: Why is this probability for the AKK sequence lower than the one we found for Example 2, when we were drawing with replacement? Also, why is $P(K-2\text{nd} \mid A-1\text{st})$ here higher than $P(K-2\text{nd})$ in Example 2?
PART D: SAMPLING RULE FOR TREATING DEPENDENT EVENTS AS INDEPENDENT

When we conduct polls, we sample without replacement, so that the same person is not contacted twice. Technically, the selections are dependent.

Sometimes, to simplify our calculations, we can treat dependent events as independent, and our results will still be reasonably accurate. We can do this when we take relatively small samples from large populations; then, we can practically assume that we are sampling with replacement, and we can ignore the unlikely possibility of the same item (or person) being selected twice.

**Sampling Rule for Treating Dependent Events as Independent**

<table>
<thead>
<tr>
<th>Population (Size $N$)</th>
<th>$N \geq 1000$, say</th>
</tr>
</thead>
<tbody>
<tr>
<td>↓</td>
<td>(even if we draw without replacement)</td>
</tr>
<tr>
<td>Sample (Size $n$)</td>
<td>$n \leq 0.05N$</td>
</tr>
</tbody>
</table>

If we are drawing a sample of size $n$ from a population of size $N$ without replacement, then, even though the selections are dependent, we can practically treat them as independent if:

- The sample size is no more than 5% of the population size:
  
  $n \leq \left(5\% \text{ of } N\right)$, or $n \leq 0.05N$,

  and

- The population size, $N$, is large: say, for our exercises, $N \geq 1000$. 
Example 4 (Sampling Rule)

G.W. Bush won 47.8% of the popular vote in 2000.
(Al Gore won 48.4%, and Ralph Nader won 2.7%.)
Over 105 million voters in the U.S. voted for President in 2000.
Of those, three are randomly selected without replacement.
Find the probability that all three selected voters voted for Bush in 2000.

§ Solution

Observe that we are sampling no more than 5% of a huge population.
By the Sampling Rule, we may practically assume that the selections are independent.

Since Bush won 47.8% of the vote, the probability that a randomly selected voter voted for Bush is 47.8%, or 0.478.

**WARNING 1:** Avoid using percents when doing calculations.

\[
P(\text{all three selected voters voted for Bush})
\]
\[
= P(1\text{-st Bush}) \cdot P(2\text{-nd Bush}) \cdot P(3\text{-rd Bush})
\]
\[
= (0.478)(0.478)(0.478), \text{ or } (0.478)^3
\]
\[
\approx 0.109 \text{ (or 10.9%)}
\]
LESSON 14: CALCULATING CONDITIONAL PROBABILITIES

When Do We Divide Probabilities?
How Do We Update Probabilities?

PART A: FORMULAS FOR CONDITIONAL PROBABILITY

For events and $A$ and $B$,

$$P(\text{A and B}) = P(\text{A}) \cdot P(B\mid A)$$  (from Lesson 13)

Assume that $P(A) \neq 0$, meaning that $A$ is a possible event.
We can divide both sides by $P(A)$ and then switch the sides.

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

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

For a frequency table (and for basic diagrams of “elos”), we can use counts:

$$P(B\mid A) = \frac{\#(\text{A and B})}{\#(A)} = \frac{\# \text{ of trials (or "elos") in which A and B occur}}{\# \text{ of trials (or "elos") in which A occurs}}$$

Example 1 (Conditional Probability: Counting “elos”; Revisiting Lesson 13, Example 1)

What is the probability that a randomly drawn card is the “3 of hearts,” given that we know that the card is a “3”?

§ Solution

$$P(3 \text{ of hearts}\mid 3) = \frac{\#(3 \text{ of hearts} \text{ and } 3)}{\#(3)} = \frac{\#(3 \text{ of hearts})}{\#(3)} = \frac{1}{4}$$

• Note that $P(\text{hearts}\mid 3) = P(\text{hearts}) = \frac{1}{4}$. In general, if $A$ and $B$ are independent events, $P(B\mid A) = P(B)$. That is, the occurrence of $A$ does not force us to update the probability of $B$ also occurring.
PART B: USING TABLES TO COMPUTE CONDITIONAL PROBABILITY

Example 2 (Conditional Probability: Using a Two-Way Frequency (or Contingency) Table; Revisiting Lesson 12, Example 4)

Let’s go back to the two-way table of 26 students from Lesson 12, Example 4. One of the students calls the tutor and [honestly] tells the tutor, “I am a Junior in the class.”

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Juniors</td>
<td>5</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

• a) What was the probability that a random student in the class was getting an “A”? (Think: What was the prior probability of an “A” student before the student informed the tutor?)

• b) What is the conditional probability that the student is getting an “A,” given that the student is a Junior? (Think: “Updated probability”; this is a concept in Bayesian statistics.) Compare to a).

• c) Imagine that, instead, the student [honestly] tells the tutor, “I am getting an ‘A’ in the class.” What is the conditional probability that the student is a Junior, given that the student is getting an “A”? Compare to b).

§ Solution

• a) Examine the table.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Juniors</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>(N=26)</td>
</tr>
</tbody>
</table>

The prior probability

\[ P(\text{“A”}) = \frac{\#(\text{“A”})}{N} = \frac{8}{26} \approx 0.308 \]
b) Since the student is a Junior, we restrict our attention to the 14 students who are Juniors. Of those 14 Juniors, 5 are getting “A”s.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Juniors</strong></td>
<td><strong>5</strong></td>
<td><strong>8</strong></td>
<td><strong>1</strong></td>
<td><strong>14</strong></td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>N=26</td>
</tr>
</tbody>
</table>

The conditional probability

\[ P("A"|\text{Junior}) = \frac{\#(\text{Junior and } "A")}{\#(\text{Junior})} = \frac{5}{14} \approx 0.357 \]

The condition that the student is a Junior **slightly increases** the probability that the student is getting an “A”; the Juniors did better than the other levels of students.

c) Since the student is getting an “A,” we restrict our attention to the 8 students who are getting “A”s. Of those 8 “A” students, 5 are Juniors.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Juniors</strong></td>
<td><strong>5</strong></td>
<td><strong>8</strong></td>
<td><strong>1</strong></td>
<td><strong>14</strong></td>
</tr>
<tr>
<td>Seniors</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>N=26</td>
</tr>
</tbody>
</table>

The conditional probability

\[ P(\text{Junior} | "A") = \frac{\#("A" \text{ and Junior})}{\#("A")} = \frac{5}{8} = 0.625 \]

This is **very different** from \( P("A" | \text{Junior}) \).
**WARNING 1: Confusion of the inverse.** People often incorrectly assume that $P(B|A)$ must be close to $P(A|B)$. Our results in b) and c) show that $P(B|A)$ could be very different from $P(A|B)$; the numerators of the fractions were the same (5), but the denominators were very different (8 vs. 14). A random “A” student is much more likely to be both an “A” student and a Junior than a random Junior is; there are 14 Juniors but only 8 “A” students.

§

**Example 3 (Confusion of the Inverse: Medical Diagnosis)**

Assume 100,000 adults live in Fredonia. A test exists for the disease “cooties.” The following probabilities apply to a randomly selected adult Fredonian.

- $P(\text{has cooties}) = 0.01$; this is called the prior probability.
- $P(\text{tests positive} | \text{has cooties}) = 0.80$
- $P(\text{tests negative} | \text{does not have cooties}) = 0.90$

Let’s say a random adult Fredonian tests positive for cooties. What is the conditional probability that the adult has cooties, given the positive test result? That is, find $P(\text{has cooties} | \text{tests positive})$. (What would you guess?)

§ **Solution**

The given probabilities suggest the following two-way table for the 100,000 adult Fredonians.

<table>
<thead>
<tr>
<th>Has cooties?</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test result</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>800</td>
<td>9900</td>
<td>10,700</td>
</tr>
<tr>
<td>Negative</td>
<td>200</td>
<td>89,100</td>
<td>89,300</td>
</tr>
<tr>
<td>Total</td>
<td>1000</td>
<td>99,000</td>
<td>100,000</td>
</tr>
</tbody>
</table>
\[ P(\text{has cooties} \mid \text{tests positive}) = \frac{\#(\text{tests positive and has cooties})}{\#(\text{tests positive})} = \frac{800}{10,700} \approx 0.0748 \text{ (or about 7.48%)} \]

- Note that \( P(\text{has cooties} \mid \text{tests positive}) \) is very different from \( P(\text{tests positive} \mid \text{has cooties}) \).

- Note the 9900 false positives, the cases where the test incorrectly comes up positive even though the adult does not have cooties. An adult Fredonian with a positive test result is much more likely to be a false positive than to actually have cooties.

- In one survey about a similar problem, 95% of American doctors said that the answer was about 75%, ten times the actual answer!

**WARNING 2: Base rate fallacy, or ignoring the prior probability.**
People often ignore prior probabilities. Although the cooties test seems pretty good, it does not justify inflating the prior probability (0.01) so dramatically (to 0.75 or thereabouts) when updating the prior to account for the positive test result.

See also: *Calculated Risks* by Gerd Gigerenzer.