DATA 602 - Conditional Probability and Unconditional Probability

© Jim Stallard 2019

Conditional Probability

The Canadian Cancer Society reports that the chance of a male contracting prostate cancer is approximately 1 in 7, or 16.7%. If a male contracts prostate cancer, is treated, and then diagnosed with remission, the chance of a reaccurrance is approximately 25%. This figure, 25% or 0.25, is referred to as a *conditional probability*, and is to be comprehended in the following manner: If a male contracts prostate cancer, is treated and then diagnosed to be cancer-free, the probability this male will have a reaccurrence of prostate cancer is 0.25, which is *higher* than the initial probability of contracting cancer of 0.167.

Now, let's consider a situation where you pick a card at random from a standard deck of $n(S) = 52$ cards. If the card is a ♠, then it is one of the following $Spade = \{A♠, 2♠, 3♠, \cdots, Q♠, K♠\}$.

Question 1: What is the chance/probability of this ♠ card is an Ace?

To consider the likelihood/chance that the chosen cards, which turns out to be a ♠, is *also* an Ace, consider all the outcomes that satisfy the event ♠ as the *sample space of relevance*. Why? At this point you have some information about the card - it is a ♠ - so you do not have to be concerned with all cards that are not ♠s. As a result, there are only 13 cards that are now relevant. How many of these 13 ♠s will also satisfy the event tha that the card is an Ace? Well, just one card, right? That card being the Ace♠.

From this condensed sample space, the probability of the chosen ♠ being an Ace is $\frac{1}{13} = 0.076923 \approx 0.0769$, or 7.69%.

Question 2: Now, consider the same ♠ card. Compute the chance/probability that this card is not an Ace, or Ace^c. Think about this for a moment before you read on, how would you compute the probabilty of this chosen ♠ card is Ace^c?

If you computed this probability to $\frac{12}{13} = 0.9231$, or 92.31%, then your though process is correct. Also, notice how the probability that the chosen ♠ card not being an Ace is the probabilty that the chosen ♠ card *is* an Ace, subtracted from 1?

Conditional Probability Notation

In the two questions posed above, a conditional probability has been computed. In **Question 1**, the $P(Ace \ GIVEN \ ♠)$ was computed; in **Question 2** the $P(Ace^c \ GIVEN \ ♠)$ was computed. Probability notation replaces the $GIVEN$ with a vertical bar and

$$P(Ace \ GIVEN \ ♠) = P(Ace|♠) = \frac{1}{13} = 0.0769 \qquad \text{and} \qquad P(Ace^c \ GIVEN \ ♠) = P(Ace^c|♠) = \frac{12}{13} = 0.9231$$

In general terms, the conditional probability of an event A when event B *has occurred** is represented by

$$P(A| \underbrace{B}_{\text{given event}}) = P(A|B)$$

The **conditional probability** of an event A, given the occurrence of another event B, is represented by the notation $P(A|B)$. In order to compute this probability, we revisit the Venn Diagram
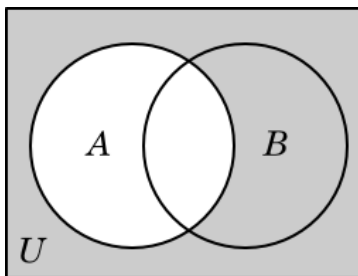
From this, we can derive a probability expression that will allow us to compute the conditional probability $P(A|B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0$$

This result can be applied to *confirm* the conditional probabilities computed in Questions 1 and 2 above:

$$P(Ace|♠) = \frac{P(Ace \ \cap ♠)}{P(♠)} = \frac{\overbrace{\frac{1}{52}}^{\text{chance of an Ace and Spade}}}{\frac{13}{52}} = \frac{1}{13} = 0.0769$$

and

$$P(Ace^c|\spadesuit) = \frac{P(Ace^c \cap \spadesuit)}{P(\spadesuit)} = \frac{\overbrace{\frac{12}{52}}^{\text{chance of not an Ace and Spade}}}{\frac{13}{52}} = \frac{12}{13} = 0.9231$$

Recall **Example 6** from the last module, where you were given the following breakdown of the students in a certain undergraduate course that had 60 students: - 28 are 2nd-year students - 34 are Statistics majors - 7 are 2nd-year students that are not Statistics majors

From this, a probability table was produced:

Event	*P(Stat)*	*P(Stat^c)*	*Row Totals*
P(2nd)	$\frac{21}{60}$	$\frac{7}{60}$	$\frac{28}{60}$
P(2nd^c)	$\frac{13}{60}$	$\frac{19}{60}$	$\frac{32}{60}$
Colum Totals	$\frac{34}{60}$	$\frac{26}{60}$	1

Suppose the randomly chosen student was a Statistics major student, meaning the event *Stat* has occurred. Compute the probability this student is

a. a 2nd year student.

   Answer You seek to compute $P(2nd|Stat)$. Using the foundational conditional probabilty formula,

$$P(2nd|Stat) = \frac{P(2nd \cap Stat)}{P(Stat)}$$

Since the event *Stat* has occurred, you only need to be interested in the first-column, the *Stat* column. This column, the sample space of relevance, is taken out below. The chance of this event, or landing in this column, is $P(Stat) = \frac{34}{60}$.

Event	*P(Stat)*
P(2nd)	$\frac{21}{60}$
P(2nd^c)	$\frac{13}{60}$
Colum Totals	$\frac{34}{60}$

and

$$P(2nd|Stat) = \frac{P(2nd \cap Stat)}{\frac{34}{60}}$$

Returning to the column of probabilities, observe which one of these represents the intersection of the *Stat*-event and the *2nd*-event. This is $P(2nd \cap Stat)$ which was found to be $\frac{21}{60}$. This probability goes into the numerator of the expression above and the conditional probabilty is then computed.

$$\begin{aligned} P(2nd|Stat) &= \frac{P(2nd \cap Stat)}{\frac{34}{60}} \\ &= \frac{\frac{21}{60}}{\frac{34}{60}} \\ &= \frac{21}{34} \\ &= 0.617647 \\ &\approx 0.6177 \end{aligned}$$

In summary, if the student chosen from this class *is* a Statistics major, the probability they are also a 2nd-year student is 0.6177. This is greater than the *initial probability* that the student chosen at random is a 2nd-year student, which is $P(2nd) = \frac{28}{60} = 0.4667$. What this information indicates is that a student chosen from this class whom is a *Stat* major is more likely to be a *2nd*-year student.

   Answer with R: From the information re-summarized above, we have the following data: $P(Stat)$ in **pstat**, $P(2nd)$ in **psecond**, $P(Stat \cap 2nd)$ in **psecondstat**. You wish to find $P(2nd|Stat)$, which is equal to

Hide

```
psecondstat = 21/60
pstat = 34/60
psecondstat/pstat   #P(second AND Stat)/P(Stat)
```

```
[1] 0.6176471
```

What this information indicates is that a student chosen from this class whom is a *Stat* major is more likely to be a $2nd$-year student when compared to being initially chosen as a $2nd$-year student.

    b. not a 2nd year student.

      Answer 1: In this instance, you are required to find $P(2nd^C|Stat)$:

$$P(2nd^c|Stat) = \frac{P(2nd^c \cap Stat)}{\frac{34}{60}}$$
$$= \frac{\frac{13}{60}}{\frac{34}{60}}$$
$$= \frac{13}{34}$$
$$= 0.382353$$
$$\approx 0.3823$$

The chance/probability that a student randomly chosen from this class, a student whom is a Statistics major, is not a $2nd$-year student is 0.3823, or about 38.23%. It more more likely for a Statistics major chosen from this class to be a $2nd$-year student (0.6177) than not a $2nd$-year student (0.3823).

If you did not know the make up of this class, what would these to conditional probabilities imply? There are likely more 2nd-year Statistics students in this class than not, which perhaps means that this is a class that 2nd-year Statistics students need to take?

Answer with R:

<div align="right">Hide</div>

```
(pstat - psecondstat)/(pstat)   #P(not 2nd AND Stat)/P(Stat)
```

```
[1] 0.3823529
```

Notice in (b) the relation between $P(2nd|Stat)$ and $P(2nd^c|Stat)$, in that $P(2nd^c|Stat) = 1 - P(2nd|Stat)$? This is not be fluke, and holds true in the general sense. A proof of this result is provided below:

$$P(A^c|B) = \frac{P(A^c \cap B)}{P(B)}$$
$$= \frac{P(B) - P(A \cap B)}{P(B)} \quad (P(B) = P(A \cap B) + P(A^c \cap B))$$
$$= \frac{P(B)}{P(B)} - \frac{P(A \cap B)}{P(B)}$$
$$P(A^c|B) = 1 - P(A|B)$$

Question: Suppose two events A and B are mutually exclusive. Compute the probability of A given B, $P(A|B)$.

Answer

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{P(B)} = 0$$

Therefore, if two events A and B are mutually exclusive events, then $P(A|B) = 0$. Likewise, $P(B|A) = 0$.

Example 3a:. Two presumed to be fair die are tossed and the top-sides are observed. Suppose the top-sides sum to 7. Compute the probablity that one of the top sides was a 6?

Answer: To start, define the event sum of 7 as the sum of the dice is 7, and event one 6 appears as one of the two top sides is a 6. You can compute that $P(\text{sum of } 7) = \frac{6}{36}$ and $P(\text{one 6 appears}) = \frac{6}{36}$. The $P(\text{sum of 7} \cap \text{one 6 appears}) = \frac{2}{36}$ as the intersection of these two events can occur $\{(6,1),(1,6)\}$. From these data,

$$P(\text{one 6 appears}|\text{sum of 7}) = \frac{P(\text{one 6 appears} \cap \text{sum of 7})}{P(\text{sum of 7})}$$
$$= \frac{\frac{2}{36}}{\frac{6}{36}}$$
$$= \frac{2}{6}$$
$$= 0.3333$$

As you might expect, this probability is computed to be 0.3333, or 33.33%. If the sum of the two die is 7, you can verify that of the six different outcomes that will result in a sum of 7, two will satisfy the one 6 occurs. Hence, $\frac{2}{6}$.

Mathematical Probability and Probability Computed Through Simulation. Using R, a simulation of 10000 tosses of two fair die was conducted to demonstrate how the *frequentist/empirical* $P(observe\ one\ 6|sum\ of\ 7)$ was computed and appears below.

```
[1] 0.3366
```


Conditional Probability
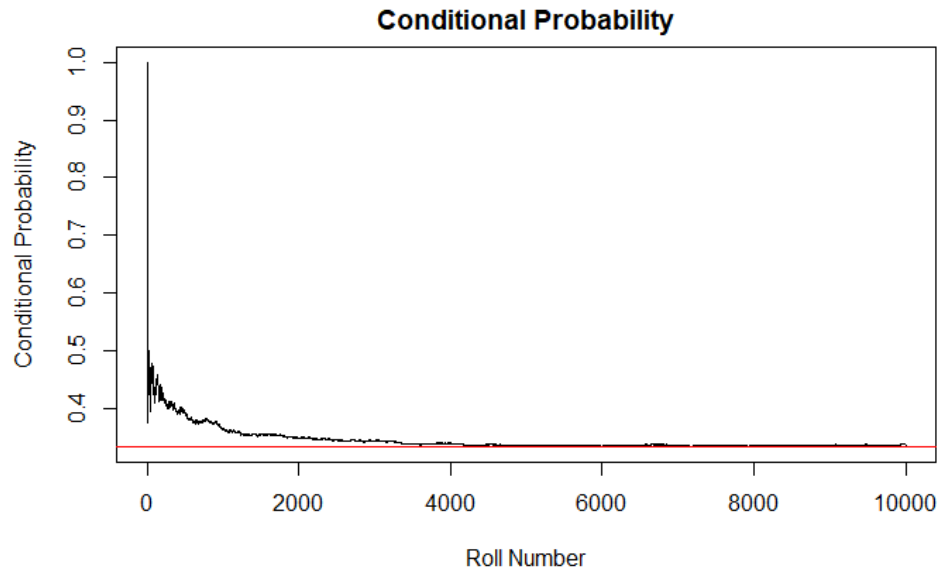
Observe how in the first few throws of the two die, the conditional probability of observing one six, *given* the two uppper-most faces sum to 7, is very erratic. However, as the number of throws increases this conditional probabilit starts to settle down, eventually converging to a probability of $\frac{1}{3}$.

Example 3b: If the sum of the top-sides is 7, compute the probability that no 6 was observed.

Answer: There are two different, and equally correct ways, to arrive at the correct answer. One method uses the basic form of conditional probability. The other employs the `Law of Complement', when applied to conditional probability. In this vain,

$$P((\text{one 6 appears})^c|\text{sum of 7}) = 1 - \underbrace{P(\text{one 6 appears}|\text{sum of 7})}_{\text{answer from Example 3a}}$$

$$= 1 - \left(\frac{1}{3}\right)$$
$$= \frac{2}{3}$$
$$= 0.6666666$$
$$\approx 0.6667$$

If the two top-sides of the die sum to 7, the probability that the are no sixes appearing is 0.6667, or 66.67%.

Independent Events

What does it mean for events to be independent. If we go back to Examples 3a and 3b and the events defined in this example, we can find that

$$P(\text{sum to 7}) = \frac{6}{36} = \frac{1}{6} = 0.16667 \quad \text{and} \quad P(\text{one 6 appears}) = \frac{11}{36} = 0.305556$$

Does the occurrance of a sum of 7 depend, or not depend, on the occurrence getting a six on one of the tosses?

Before this inquiry can be addressed, the notion of *independent events* needs to be established.

Two events A and B are said to be **independent events** if the occurrence of one event (say A) does not affect the occurrence of the other event (B, in this case). As a result,

$$P(A \cap B) = P(A) * P(B) \quad \text{for} \quad P(A) > 0, P(B) > 0$$

Example 1: Suppose you are to play the casino game roulette. There are 38 pockets on a roulette table, each distinctly numbered 1 through 36, with a pocket numbered 0 and another pocket numbered 00. The even numbered pockets have a red background, the odd numbered pockets are on a black background, the two zeros are on a green background. You are to play roulette twice, both times the white ball landed on green. Compute the probability of such an outcome.

Answer: Start by defining A to be the outcome on Game #1 is green, and B as the outcome on Game #2 as green. To find $P(Game\ 1\ is\ green\ \cap\ Game\ 2\ is\ green)$:

$$
\begin{aligned}
P(Game\ 1\ is\ green\ \cap\ Game\ 2\ is\ green) &= P(A \cap B) \\
&= P(A) * P(B) \quad \text{(why?)} \\
&= \left(\frac{2}{38}\right) * \left(\frac{2}{38}\right) \\
&= \left(\frac{2}{38}\right)^2 \\
&= 0.00277
\end{aligned}
$$

This probability computation is demonstrated in R Studio below.

<div style="text-align:right">Hide</div>

```
(2/38)*(2/38) #multiply 2/38 together
```

```
[1] 0.002770083
```

OR

<div style="text-align:right">Hide</div>

```
(2/38)^2 #2/38 to the power of 2
```

```
[1] 0.002770083
```

Example 2: A poll of post-secondary education graduates under the age of 40 in Canada was conducted by BDO in 2017. One of this poll's findings was that two-thirds, or about 66%, graduated with student debt (the average amount of student debt being about $22,000). You randomly pick two Canadians under the age of 40 who are post-secondary education graduates. Compute the probability that both graduated with student debt.

Answer: Defining G_i to represent graduate $i, i = 1, 2$ had student debt, compute $P(G_1 \cap G_2)$.

$$
\begin{aligned}
P(G_1 \cap G_2) &= P(G_1) * P(G_2) \quad \text{(random selection results in Canadian 1 graduating with debt to be free/independent of Canadian 2)} = \quad (0.66) \\
&= 0.4356
\end{aligned}
$$

The chance/probability that two randomly chosen Canadians, both under the age of 40 and both being graduates from a post-secondary program, will graduate with student-debt is 0.4356, or 43.56%.
This can also be solved in R:

<div style="text-align:right">Hide</div>

```
(0.66)^{2}
```

```
[1] 0.4356
```

Result from Independence: If two events A and B are independent, then

1. A and B^c are independent
2. A^c and B are independent
3. A^c and B^c are independent

Demonstration of 1: We want to show that $P(A \cap B^c) = P(A) * P(B^c)$. If this is the case, then A and B^c are independent events. We know there are two ways for A to occur: A with B or $A \cap B$, and A without B ($A \cap B^c$).

$$
\begin{aligned}
P(A \cap B) + P(A \cap B^c) &= P(A) \\
P(A \cap B^c) &= P(A) - \underbrace{P(A \cap B)}_{P(A)*P(B)} \\
P(A \cap B^c) &= P(A) - P(A) * P(B) \\
P(A \cap B^c) &= P(A)(1 - P(B)) \quad (\text{factor out } P(A) - \text{term}) \\
P(A \cap B^c) &= P(A) * P(B^c)
\end{aligned}
$$

This proves that if A and B are independent events, then A and B^c are independent events.

Example 3: Refering to the story in Example 2, compute the probability

a. that the first person chosen graduated with debt and the second person chosen did not graduate with debt, or $P(G_1 \cap G_2^c)$.

Answer: To compute $P(G_1 \cap G_2^c)$

$$P(G_1 \cap G_2^c) = P(G_1) * P(G_2^c) \qquad \text{(Since } G_1 \text{ and } G_2 \text{ are independent, so to are } G_1 \text{ and } G_2^c)$$
$$= (0.66) * (1 - 0.66)$$
$$= 0.2244$$

The chance/probability that the *first* of the two randomly chosen Canadians (each have completed a post-secondary education program and each under the age of 40) graduated with debt and the *second* did not graduate with debt is 0.2244, or 22.44%.
Answer with R: Refer to the R code below

Hide

```
(0.66)*(1-0.66)
```

```
[1] 0.2244
```

b. that at least one of the two graduated with debt.

Answer 1: Using the notation that was defined in Example 3a, you need to compute $P(G_1 \overset{\text{at least one}}{\cup} G_2)$:

$$P(G_1 \cup G_2) = P(G_1) + P(G_2) - P(G_1 \cap G_2) \quad \text{(the Addition Law)}$$
$$= P(G_1) + P(G_2) - \underbrace{P(G_1) * P(G_2)}_{\text{by independence}}$$
$$= 0.66 + 0.66 - \underbrace{(0.66 * 0.66)}_{0.4356}$$
$$= 0.8844$$

Answer 2: Another approach is to observe that the complement of 'at least one of the two graduates with student-debt' is *neither* of the two randomly chosen graduate with student-debt. That is

$$P(G_1 \cup G_2)^c = P(G_1^c \cap G_2^c) \quad \text{(Using DeMorgan's Law)}$$

Therefore, we can express $P(G_1 \cup G_2)$ as

$$P(G_1 \cup G_2) = 1 - P(G_1 \cup G_2)^c \quad \text{(using the Law of Complement))}$$
$$= 1 - P(G_1^c \cap G_2^c) \quad \text{(using DeMorgan's Law))}$$
$$= 1 - \underbrace{(P(G_1^c) * P(G_2^c))}_{\text{by independence}}$$
$$= 1 - (0.34 * 0.34)$$
$$= 1 - 0.1156$$
$$= 0.8844$$

Answer with R: Using the first approach, the answer is obtained with the one line of R code below:

Hide

```
pg1 = 0.66
pg2 = 0.66
pg1 + pg2 - (pg1*pg2) #uses the approach demonstated in Answer 1
```

```
[1] 0.8844
```

Using the second approach, the answer is coded in R below

Hide

```
1 - (0.34^{2}) #uses the approach outlines in Answer 2
```

```
[1] 0.8844
```

How are Conditional Events and Independent/Unconditional Events related Probabilistically?

At this point, you can differentiate between independent/unconditional events and dependent/conditional events: Independent/unconditional events, when they do occur, do not affect the probability of occurrence of each other. However, events that are dependent/conditional *do* affect the probabilty of occurrence of the other event.

If you know that two events, let's call them A and B, *are independent* events, what is the conditional probability of A if event B does occur?

Answer: Start with $P(A|B)$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{(using the basic formula for conditional probability)}$$

$$= \frac{P(A) * P(B)}{P(B)} \quad (P(A \cap B) = P(A) * P(B) \text{ because A and B are independent events)}$$

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

This shows that if two events A and B are independent events, then the occurrence of B *does not affect* the probabity of occurrence of A, and vice versa. So

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

Example 4: Refer to Example 1 and compute the probability that neither of the two games produces a green outcome.

Hide

```
(36/38)*(36/38)
```

```
[1] 0.8975069
```

Example 5: Recall that $P(Stat) = \frac{34}{60}, P(2nd) = \frac{28}{60}, P(Stat \cap 2nd) = \frac{21}{60}, P(2nd|Stat) = \frac{21}{34}$. Are the events a student is a Statistics major and a student is in their 2nd year independent, or dependent, events?

Answer: If the events $Stat$ and $2nd$ are independent events, then

$$P(2nd|Stat) = P(2nd)$$

Is this the case? Let's check:

$$P(2nd|Stat) = \frac{P(2nd \cap Math)}{P(Math)}$$
$$= \frac{21/60}{34/60}$$
$$= \frac{21}{34}$$
$$= 0.6176$$

We found that $P(2nd) = \frac{28}{60} = 0.4667$. Because $P(2nd) = 0.4667 \neq P(2nd|Stat)$, the events a student is a math major $Stat$ and the student is a 2nd-year student $2nd$ *are not independent events* from the class-makeup/probabilities provided and computed.

Example 6: Returning the to the very first example encountered with conditional probability, one where a random experiment consisted of selecting a card at random from a standard deck of 52 cards. Two events were considered, the chosen card is a ♠ and the chosen card is an Ace. Are these events independent, or dependent?

Answer: In Example 1, the $P(Ace|♠)$ was computed to be $\frac{1}{13} = 0.0278$. If these events are independent, then $P(Ace|♠) = P(Ace)$. A check of this wil determine if these events are independent or dependent.

$$P(Ace|♠) \stackrel{?}{=} P(Ace)$$

$$\frac{1}{13} \stackrel{?}{\neq} \frac{4}{52} \quad (\underbrace{\{A\heartsuit, A\diamondsuit, A\spadesuit, A\clubsuit\}}_{4-ways})$$

$$0.0278 = 0.0278$$

Due to this equality, $P(Ace|♠) = P(Ace)$ and the events are *independent* events.

A note about dependent events: Should you discover that events are dependent, by no means does this imply 'causation'. This is a very important distinction to make. Dependent events may be directly related, or indirectly related because they are associated with some other event.

For example, it can be shown that when you toss two die and observe the number of dots appearing on the top-side of each, that

$$P(\text{one 6 appears}|\text{sum of 7}) \neq P(\text{one 6 appears})$$
$$\frac{2}{6} \neq \frac{10}{36} \quad \underbrace{(\{(1,6),(2,6),(3,6),(4,6),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5)\})}_{10-ways}$$
$$0.3333 \neq 0.2778$$

Here, if one observed a sum of seven, this event *will depend* on observing a higher number of dots on one of the two die. A sum of three event depends on both top-sides showing a small number of dots.

# Conditional Probability, Independence and the Law of Total Probability

Consider a random experiment that has two stages in sequence, with stage 1 having one of many possible k-events: B_1 or B_2 or B_3 or $\cdots B_k$. Stage two has two possible events, A or A^c and the probabilty of occurrence of either A or A^c depends on which *prior$ event B_i occurred. Since A can with each of the B_i-events, there are k ways in which A can occur, and the $P(A)$ is computed by

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k) \quad (\text{recall that} P(A|B_i) = \frac{P(A \cap B_i)}{P(B_i)})$$
$$= \underbrace{P(A|B_1)P(B_1)}_{P(A \cap B_1)} + \underbrace{P(A|B_2)P(B_2)}_{P(A \cap B_2)} + \cdots + \underbrace{P(A|B_k)P(B_k)}_{P(A \cap B_k)}$$
$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

This is called the **Law of Total Probability**. A visualization of how this result works is provided below.
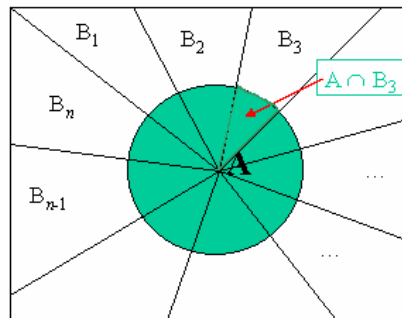
Consider the number of different ways in which A can occur:

- A can occur with B_1, meaning $A \cap B_1$ occurs
- A can occur with B_2, meaning that $A \cap B_2$ occurs
- A can occur with B_3, suggesting that $A \cap B_3$ occurs

However, the occurrence of A will *depend* upon which of the *prior* events $B_1, B_2, B_3, \cdots$, or B_k occur, with each of these prior events occurring with a *prior probability* of $0 < P(B_i) < 1$.

To illustrate this result, let's go through a few different examples.

Example 1 An insurance company provided the following table that summarizes the (i) age distibution and (ii) the probability of an automobile accident in a year for all clients who fall within each age category who have an automobile insurance policy.

Age Group :	< 25	$25 \leq 40$	Over 40
Client % :	0.15	0.35	0.50
P(Car Accident) :	0.20	0.10	0.05

Compute the probability that a randomly chosen client that has an automobile insurance policy will have an automobile accident in the next year.

Solution: Define the event A to represent a client has an auto accident in the next year. From the table above, the provided conditional probabilities are

Conditional probabilities: $P(Accident| < 25) = 0.20, \quad P(Accident|25 \leq 40) = 0.10, \quad P(Accident| > 40) = 0.05$
Probability of a client being in a certain age category: $P(< 25) = 0.15, \quad P(25 \leq 40) = 0.10. \quad P(> 40) = 0.50$

Here the probability of a client being in a car accident in the year, $P(A)$, is necessary to compute.

$$P(\text{Accident}) = P(\text{Accident And} < 25) + P(\text{Accident And } 25 \leq 40) + P(\text{Accident And} > 40)$$
$$= P(\text{Accident}| < 25)P(< 25) + P(\text{Accident}|25 \leq 40)P(25 \leq 40) + P(\text{Accident}| > 40)P(> 40)$$
$$= (0.2 * 0.15) + (0.10 * 0.35) + (0.05 * 0.50)$$
$$= 0.03 + 0.035 + 0.025$$
$$= 0.09$$

Answer with R: Here is some R-code that computes this probability. Try to Copy and paste the R code into an R chunk in an R Studio document to confirm the $P(A)$.

<div align="right">Hide</div>

```
prior = c(0.15, 0.35, 0.50)    #assings three probabilities to a vector called prior for the Bi events
cond = c(0.20, 0.10, 0.05)     #assigns the conditional probabilties to a vector called cond
eachterm = prior*cond          #multiplies each term together (0.15*0.20), (0.35*0.10), (0.50*0.05)
sum(eachterm)                  #adds the three products in the eachterm vector to obain a sum
```

```
[1] 0.09
```

The usage of this formula is quite powerful, but can be cumbersome to use as in some cases it requires an intensive effort to set up notation. A more intuitive method to compute such a probability - a method that uses the Total Law of Probability *without the need to recall the formula* - a method that employs a **Tree Diagram** will now be demonstrated.

In this problem, the probability of the insurance company's client getting into a car accident in a year *depends* on what their age category is. We can see that those who are less than 25 years of age have a probabilty of getting into a car accident over the course of the year of $P(Accident| < 25) = 0.20$. This is four times higher than a person who is over the age of 40, $P(Accident| > 40) = 0.05$. From these data, it appears that the occurrence of an accident *depends on* one's age. Our tree diagram is going to sprout from this notion that there are three different age-branches, one for each age category:
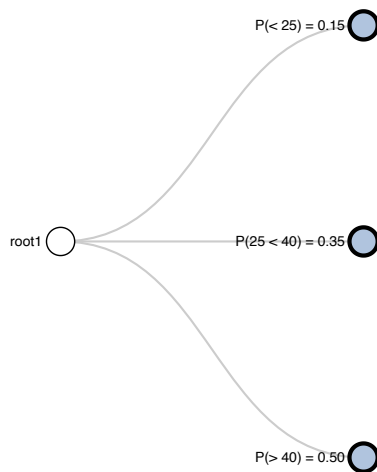
The first stage of this tree diagram shows the branching of the insurance company's clientelle into one of three age groups: < 25 with probability 0.15; $25 < 40$ with probability 0.35, and > 40 with probabilty 0.50. Notice how the sum of the probabilities appearing at the end of the first stage add to 1? This is one property of a tree diagram. This can be shown by selecting on the 'root1' node.

If you select the $P(< 25)$ node, the second stage of events will show. A client who is less than 25 years of age with either have an accident in the next year with probability 0.20, or will not have an accident in the next year - $Accident^c$ - with probability/chance of 0.80. Similar branchings occur when the $P(25 < 40)$ and $P(> 40)$ nodes are selected.

A second tree diagram of all the possibilies is provided below. Select the first top end-point in Stage 2 of the tree-diagram, $P(\text{Accident}| < 25)$. At this point, you are at one of the final destination points in tree diagram, one of *eight possible* end-points. At this particular end-point, the client is (i) less than 25 years of age and (ii) *given* they are less than 25 years of age, the chance/probability they will be in a car-accident over the course of a year. This results in the event that the randomly chosen client is *both* less than 25 years of age and will be in a car accident in a yaer, $P(\text{Accident}\cap < 25)$. To compute this probability, we use the following relation
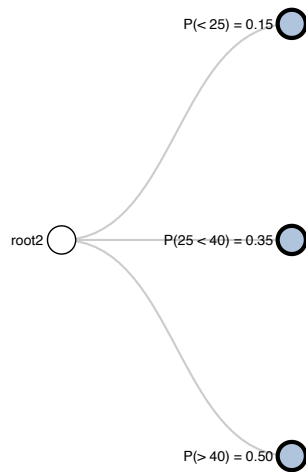
$$P(\text{Accident}| < 25) = \frac{P(\text{Accident} \cap < 25)}{P(< 25)}$$
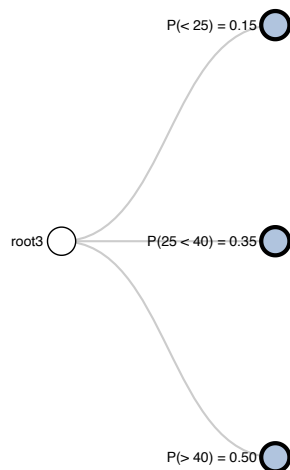$$P(\text{Accident}| < 25) * P(< 25) = P(\text{Accident} \cap < 25)$$
$$(0.20) * (0.15) = P(\text{Accident} \cap < 25)$$
$$0.03 = P(\text{Accident} \cap < 25)$$

This probability is can be computed by simply journeying through the branches of the tree that result in the $P(\text{Accident} \cap < 25)$ outcome/destination. As you travel across the required branches, multiply the probabilities together.

The **completed** probability tree is provided below. Observe that each outcome has a different probability, but all possible outcomes have a total probability of 1.

From all endpoints, identify the outcomes where the evetn of interest - Accident - occurs/will occur. These are at endpoints 1 ($P(\text{Accident} \cap < 25)$), 3 ($P(\text{Accident} \cap 25 < 40)$), and again at endpoint 5 ($P(\text{Accident} \cap > 40)$). There are three *different* outcomes that result in an Accident. The probability of a randomly chosen client having an accident in a year is then

$$P(\text{Accident}) = \underbrace{P(\text{Accident} \cap < 25)}_{Endpoint\ 1} + \underbrace{P(\text{Accident} \cap 25 < 40)}_{Endpoint2} + \underbrace{P(\text{Accident} \cap > 40)}_{Endpoint3}$$

$$= (0.03) + (0.035) + (0.025)$$
$$= 0.09$$

Time to Play 1: Suppose that 5% of all employees of a certain company use illegal drugs. The company has implemented a drug testing policy, where an employee is randomly chosen and tested for illegal drug usage. This particular test returns a positive result (test indicates illegal drug usage) 99% of the time on employees who have used illegal drugs within the past thirty days. However, this test also resturns a positive result on 2% of cases where the employee *had not* used illegal drugs in the past thirty days.

Suppose you are an employee of this company, and you have been chosen for a random drug test. Compute the probability that your test result will be (i) positive (ii) negative.

Please take the time to attempt this problem on your own. It will take you about five minutes. If you need a push, I have provided you with the solution.

Click see the full solution to this Time to Play exercise. (http://people.ucalgary.ca/~jbstall/Answers/SolutionExample2TotalLaw.nb.html)

Conditional Probability and Bayes Theorem

Imagine if you are a human resources manager, and you are faced with a scenario where an employee was randomly chosen and subjected to a drug test, the result of which was positive. But, you are aware of false-positives - a test result indicating the employee has used illegal drugs in the past thirty days, when in fact they have not. How can you decipher if the person who tested positive for illegal drug use *actually did* use illegal drugs in the past thirty days?

Well, there is a theorem for this type of problem and it is called *Bayes Theorem*.

Bayes' Theorem is a result that allows one to update probabilites, based on the latest information. This would also to consider the possibility of an employee whom tested positive , actually did or did not use illegal drugs in the past thirty days.

Here is **Bayes' Theorem**:

Let $B_1, B_2, \cdots, B_k$ represent events that are pairwise mutually exclusive, meaning $B_i \cap B_j = \emptyset$ for all $i \neq j$. Let A represent some other event, which is conditional on some B_i occurring. For example $P(A|B_1) > 0, P(A|B_2) > 0, \cdots, P(A|B_K) > 0$ and $P(A|B_i)$ is not necessarily the same for all $i = 1, 2, \cdots, k$.

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)}$$

It is a very techical formula. In the examples presented below, tree diagrams and the *basics* of conditional probability will be employed.

Example 1(a): In an attempt to collect the opinions of Albertans regarding the management of the Alberta Heritage Fund, a polling company is to survey Alberta households. The survey company will distribute, via mail, questionnaires to 300 randomly selected households in Alberta. In an attempt to ensure the sample is representative, 200 surveys were sent to urban households, the remainder send to rural households. Each questionnaire is classified as a response or a non-response - the latter occurring when the questionnaire is not returned. It is known that the non-response rate of mail-out questionnaires for urban households is 40%. The non-response rate for rural households is 20%.

A survey is returned. Compute the probability this survey was sent to an urban household.

Answer A tree diagam of this random process is provided below
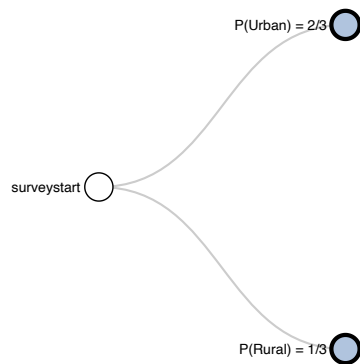
The given event was that a survey was returned, or *Returned* occurred. The probability that this returned survey was *sent* to an urban household-requires computation, or compute $P(Urban|Returned)$:

$$P(Urban|Returned) = \frac{P(Urban \cap Returned)}{P(Returned)} =?$$

First consider the denominator $P(Returned)$.

$$P(Returned) = \underbrace{P(Urban \cap Returned)}_{Endpoint\ 1} + \underbrace{P(Rural \cap Returned)}_{Endpoint\ 3}$$
$$= P(Returned|Urban)P(Urban) + P(Returned|Rural)P(Rural)$$
$$= \left(0.60 * \frac{2}{3}\right) + \left(0.80 * \frac{1}{3}\right)$$
$$= 0.40 + 0.26667$$
$$= 0.66667$$

This can be computed in an R chunk with the following:

<div style="text-align:right">Hide</div>

```
pwhere = c(2/3,1/3) #assign probabilites, P(Urban) then P(Rural)
pcond = c(0.60, 0.80) #assign conditional P(Ret|Urban) then P(Ret|Rural)
ex1den = sum(pwhere*pcond) #multiplies, then sums the terms
ex1den #returns the answer
```

```
[1] 0.6666667
```

Now, returning to our initial probability:

$$P(Urban|Returned) = \frac{P(Urban \cap Returned)}{P(Returned)} = \frac{\left(0.60 * \frac{2}{3}\right)}{0.66667} = 0.60$$

The R chunk to compute this is

If a questionnaire *was* returned, the probablity that the questionnaire was *sent* to an urban household is 0.60. This is an adjustment from the *prior* probability of $P(Urban) = \frac{2}{3} = 0.66687$.

Example 1(b) Suppose one of the 300 questionnaires was tracked, and it was not returned. Compute the probability that this unreturned questionnaire was send to a rural household.

Answer: Here the $P(Urban|Not\ Returned)$ needs to be computed. Returning to the tree diagram appearing in Example 1(a), it can be observed that there are two outcomes that satisfy the event that a questionnaire is not returned. These occur at endpoints 2 and 4. This will help in computing

$$P(Rural|Not\ Returned) = \frac{P(Rural \cap Not\ Returned)}{P(Not\ Returned)}$$

$$= \frac{P(Rural \cap Not\ Returned)}{\underbrace{P(Urban \cap Not\ Returned)}_{endpoint\ 2} + \underbrace{P(Rural \cap Not\ Returned)}_{endoint\ 4}}$$

(one of the basement-terms matches with the numerator!)

$$= \frac{\left(\frac{1}{3} * 0.20\right)}{\left(\frac{2}{3} * (0.40)\right) + \left(\frac{1}{3} * 0.20\right)}$$

$$= \frac{0.0667}{0.2667 + 0.0667}$$

$$= \frac{0.0667}{0.3334}$$

$$= 0.20006$$

$$\approx 0.20$$

If a questionnaire is not returned, then the chance it was sent to a rural household is 0.20, or 20%. This is modified from the initial, or prior probability of a questionnaire being sent to a rural household which was $P(Rural) = \frac{1}{3} = 0.3333$.

Answer wtih R Code

Hide

```
pwhere = c(2/3, 1/3) #assign probabilites, P(Urban) then P(Rural)
pcond = c(0.40, 0.20) #assign conditional P(Ret|Urban) then P(Ret|Rural)
ex1den = sum(pwhere*pcond) #multiplies, then sums the terms
ex1den #returns the answer of the denominator
```

```
[1] 0.3333333
```

Hide

```
ex1num = (1/3*0.20) #computes the numerator
ex1banswer = ex1num/ex1den #computes the answer stored to ex1banswer
ex1banswer #returns the answer
```

```
[1] 0.2
```

Time to Play 2: A certain type of cancer is to be found in 0.5% of a population of individuals. Once a person is suspected of having this type of cancer, they are sent to have a test done. The test comes back either positive - affirming that the person suspected of having this cancer *does* in fact have the cancer; or negative - indicating that the person suspected of having the cancer *does not* have the cancer. The test will indicate the existence of cancer in 94% of individuals suspected of having the cancer, and in 4% of those who do not have the cancer.

Using your current comprehension of creating tree diagrams and Bayes' Theorem, attempt to compute the probabilities requested in parts (a) through (c). Also, for each computed probability interpret its meaning.

    a. A person is suspected of having the cancer. Compute the probability that the test will come back positive.

    b. If the test comes back positive, compute the probability that this person does not have cancer.

    c. If the test comes back negative, compute the probabilty that this person does have cancer.

The detailed answer to this Time to Play exercise is provided in this link.
(http://people.ucalgary.ca/~jbstall/Answers/SolutionExample2TotalLaw.nb.html)

Example 3: A business traveller considers renting a car from one of three agencies: A, B, or C. The probability of receiving an upgrade (at no additional charge) if she rents from agency A is 0.70. The probability of receiving upgrades - again at not additional charge - from agencies B and C are 0.5 and 0.3, respectively. She is three times more likely to rent from agency A than Agency B, and twice as likely to rent from agency B than agency C. If she receives an upgrade, compute the probability that she rented from

    a. agency C.

    Answer: We visualize this problem with the following tree diagram

P(Agency A) =

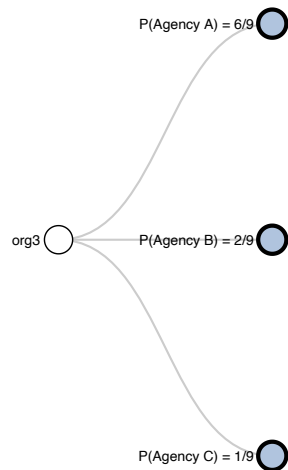P(Agency B) =

P(Agency C) =

org2
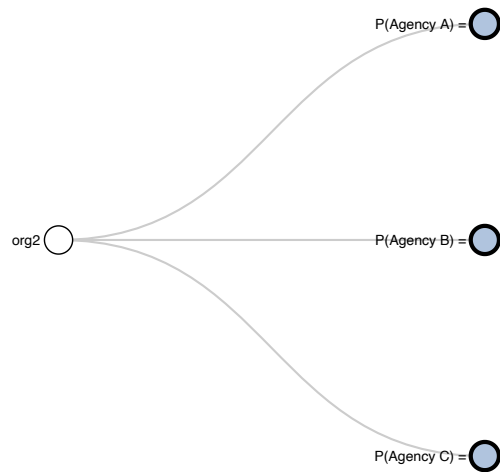
From the information given in the question, we have - 3 times more likely to rent from A than $B \rightarrow P(A) = 3P(B)$ - twice as likely to rent from B than $C \rightarrow P(B) = 2P(C)$

$$
\begin{aligned}
1 &= P(A) + P(B) + P(C) \\
&= 3P(B) + 2P(C) + P(C) \\
&= 3*(2P(C)) + 2P(C) + P(C) \\
&= 6P(C) + 2P(C) + P(C) \\
1 &= 9P(C) \\
\frac{1}{9} &= P(C)
\end{aligned}
$$

$$
P(C) = \frac{1}{9}, \qquad P(B) = 2P(C) = 2*\frac{1}{9} = 2/9, \qquad P(A) = 3P(B) = 3*\frac{2}{9} = \frac{6}{9}
$$

Our updated tree diagram:

P(Agency A) = 6/9

P(Agency B) = 2/9

org3

P(Agency C) = 1/9

The given event was that she received an upgrade, or $\mathrm{Upgrade}$ occurred. The probability that needs to be computed is $P(\mathrm{Agency\ C}|\mathrm{Upgrade})$:

$$P(\text{Agency C}|\text{Upgrade}) = \frac{P(\text{Agency C} \cap \text{Upgrade})}{P(textUpgrade)}$$

$$= \frac{P(\text{Agency C} \cap \text{Upgrade})}{P(\text{Agency A} \cap \text{Upgrade}) + P(\text{Agency B} \cap \text{Upgrade}) + P(\text{Agency B} \cap \text{Upgrade})}$$

$$= \frac{\left(\frac{1}{9} * 0.30\right)}{\left(\frac{6}{9} * 0.70\right) + \left(\frac{2}{9} * 0.50\right) + \left(\frac{1}{9} * 0.30\right)}$$

$$= \frac{0.033333}{0.611111}$$

$$= 0.054545$$

If she received an upgrade, the chance/probability that she rented from Agency C is 0.0545, or 5.45%.

Answer with R:

Hide

```
agencyprobs = c(6/9, 2/9, 1/9) #create a vector of rental probabilities for A, B, and C respectively
upgradeprobs = c(0.70, 0.50, 0.3) #create a vector of conditional probability upgrades
probupgrade = sum(agencyprobs*upgradeprobs) #compute P(Upgrade)
probupgrade #return P(Upgrade)
```

```
[1] 0.6111111
```

Hide

```
prentCgivenupgrade = (1/9*0.30)/probupgrade #compute P(Rent from C | Upgrade)
prentCgivenupgrade #return the answer
```

```
[1] 0.05454545
```

(b) If she did not receive an upgrade, compute the probability that she rented from agency A.

Answer: The given event here is No Upgrade, and the probability that is required to be computed is $P(\text{Agency A}|\text{No Upgrade})$. $$

$$P(\text{Agency A}|\text{No Upgrade}) = \frac{P(\text{Agency A} \cap \text{No Upgrade})}{P(\text{No Upgrade})}$$

$$= \frac{P(\text{Agency A} \cap \text{No Upgrade})}{1 - P(\text{Upgrade})}$$

$$= \frac{\left(\frac{6}{9} * 0.30\right)}{1 - \underbrace{0.6111}_{\text{from part a}}}$$

$$= \frac{0.20}{0.3889}$$

$$= 0.5142842$$

$$\approx 0.5143$$

$$ If she did not receive an upgrade, the chance/probability she did rent from Agency A is 0.5143, or 51.43%.

Answer with R:

Hide

```
probnoupgrade = 1 - probupgrade #P(No Upgrade)
prentAgivennoupgrade = (6/9*0.30)/probnoupgrade  #P(Rent from A | No Upgrade)
prentAgivennoupgrade #returns the answer
```

```
[1] 0.5142857
```