Activity 5: Conditional Probability & Independence

Recall the following probability rules we previously derived: Conditional Probability Rule: $P(A|B)=\frac{P(A\cap B)}{P(B)}$

General Multiplication Rule: $P(A\cap B)=P(A|B)P(B)$

Situation: The American Film Institute lists the top 100 films ever made: http://www.afi.com/100years/movies10.aspx

Two people (Alan and Beth) gather to watch a movie. Instead of arguing about the selection of the movie, they agree to randomly choose a movie from the "Top 100" list. The following table summarizes the number of films Alan or Beth have seen

	Beth yes	Beth no	Total
Alan yes	42	6	48
Alan no	17	35	52
Total	59	41	100

We define the events: A = Alan has seen the film and B = Beth has seen the film

1. Calculate the following:

$P(A)=$ ___________ $P(B)=$ ___________ $P(A\cap B)=$ ____________________

$P(A')=$ ___________ $P(B')=$ ___________ $P(A\cap B)'=$ ____________________

$P(A\cup B)=$ _______________ $P(A\cup B)'=$ _______________

$P(A'\cap B')=$ _______________ $P(A'\cup B')=$ _______________

2. The probability that Alan has seen a movie is 0.48. Suppose a movie has been selected and we know that Beth has already seen the movie. Does this additional information change the (conditional) probability that Alan has seen the movie? Explain.

3. Suppose we know that Alan has **not** seen the movie. What's the probability that Beth has **not** seen the movie?

4. Now consider Chuck and Donna, who will follow the same process in order to choose a movie from the Top 100 Films list:

	Donna yes	Donna no	Total
Chuck yes	15	10	25
Chuck no	45	30	75
Total	60	40	100

The probability that Donna has seen a randomly selected movie is: P(D) = 0.60. Suppose they select a movie and we are told that Chuck has already seen it. Does this change the probability that Donna has seen the movie?

Two events are independent if knowledge that one occurred (or did not occur) does not impact the probability of the other occurring.

If A and B are independent, then:

1) $P(A \mid B) = P(A)$ (Knowledge about event B does not change the probability of event A occurring)

2) $P(A \cap B) = P(A) \times P(B)$ (if and only if A and B are independent). Verify this for Chuck & Donna.

5. Based on these conditional probabilities, which couple (Alan & Beth or Chuck & Donna) has the best chance of staying together?

6. Do taller men tend to marry taller women? The following table displays the heights of 205 husbands and wives:

	Tall wife	Medium Wife	Short Wife	Total
Tall husband	18	28	14	60
Medium husband	20	51	28	99
Short husband	12	25	9	46
Total	50	104	51	205

Yule, G.U. (1900). On the association of attributes in statistics: with illustration from the material of the childhood society,. Philosophical Transactions of the Royal Society, Series A, 194, 257–319.

Are the events {tall husband} and {tall wife} independent? What about {tall wife} and {short husband}?

System Reliability:	A satellite launch system is controlled by a system of 3 computers. Normally, computer 1 controls the system, but if it malfunctions, computer 2 automatically takes over. If computer 2 malfunctions, computer 3 takes over. The launch system will fail if all three computers malfunction. If we know each computer has a 0.05 chance of malfunctioning at any given time, what is the probability that the launch will fail? Let 1 = {Computer 1 fails}, 2 = {computer 2 fails}, and 3 = {computer 3 fails}.

7. What's the probability that both computer 1 and 2 will fail? Can we assume independence?

8. What's the probability that all 3 computers fail?

9. What's the probability that all 3 computers work properly?

10. What's the probability that either computer 1 or computer 2 will operate?

11. Suppose you learn the new washing machine you bought has a 30% chance of failing within a year. The dryer alongside it has a 10% chance of failing within the year. If we assume the washer failing and dryer failing are independent events, what's the probability that both the washer and dryer will fail? Is independence a reasonable assumption?

12. Suppose we know that at any given time, 2% of SAU students have strep throat. We also know:
   - A cotton swab test correctly detects strep throat in 95% of the people who are infected.
   - If a student does NOT have strep throat, the test correctly indicates the student does NOT have strep throat 90% of the time.

   If we randomly select a student at SAU and give them this test, what's the probability the test will indicate the student DOES have strep throat?

   a) Write out the given probabilities, along with their complements.

   .02 = ____________________ 1 – .02 = 0.98 = ____________________

   .95 = ____________________ 1 – .95 = 0.05 = ____________________

   .90 = ____________________ 1 – .90 = 0.10 = ____________________

   b) We're interested in the probability of obtaining a positive test result. There are two ways (or scenarios) in which we could get a positive result. What are those two scenarios? Calculate the probability of each.

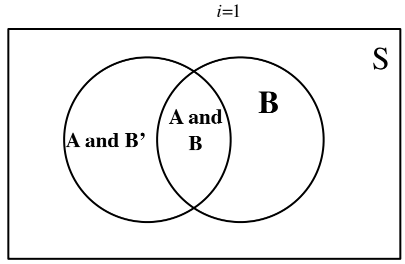
   c) Use these two probabilities to estimate the probability of obtaining a positive test result.

   c) Suppose a student takes the test and gets a positive result. What's the probability that the student has strep throat?

We just derived two important probability rules:

Law of Total Probability: $P(A) = P(A|B)P(B) + P(A|B')P(B') = \sum_{i=1}^{n} P(A|B_i)P(B_i)$

Look at the Venn diagram to the right to see how this law works. If we want to shade in the event A, we need to shade both "A and B" and "A and not B"



Bayes Theorem: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$

13. Suppose I give you a multiple-choice test (each item has 4 possible choices). Suppose you know the answer to 60% of the items, but you must randomly guess on the remaining 40% of the items. If we randomly select one item from the test, what is the probability that you will answer it correctly?

Let's begin by identifying out all the probabilities we know:

P(correct answer) = __________ P(incorrect answer) = __________

P(you know the answer) = __________ P(you do not know the answer) = __________

P(correct | you know it) = __________ P(correct | you do not know it) = __________

P(incorrect | you know it) = __________ P(incorrect | you do not know it) = __________

Now use the Law of Total Probability to find the probability that you will answer any question correctly.

.

14. Recent reports suggest 80% of all emails are spam. Suppose you look through your in-box and find:
    - 70% of your spam messages contain the word "replica"
    - 10% of your non-spam messages contain the word "replica."

If the next email you receive contains the word "replica," what's the probability that it is a spam message?

Source: http://eval.symantec.com/mktginfo/enterprise/other_resources/b-state_of_spam_report_09-2009.en-us.pdf

15. Three programmers, Alice, Ben, and Chuck, write a large chunk of code. Alice writes 60% of the code, Ben writes 30%, and Chuck writes 10%. Furthermore, it is known that 3% of Alice's code tends to be buggy; 7% of Ben's code tends to have bugs; and 5% of Chuck's code is buggy. If we randomly select a line of code, what's the probability that we will find a bug?

16. Given a bug has been found, what's the probability that each programmer wrote that line of code??

Probability Rules:

1) $0 \le P(A) \le 1$

2) $P(S) = 1.0$

3) $P(A') = 1 - P(A)$

4) If A and B are disjoint (non-overlapping), then $P(A \cup B) = P(A) + P(B)$

5) General Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

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

6) Conditional probability: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

7) General Multiplication Rule: $P(A \cap B) = P(A \mid B)P(B)$

8) If A and B are independent, then $P(A \mid B) = P(A)$

9) Bayes Theorem: $P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')}$

10) Law of Total Probability: $P(A) = P(A \mid B)P(B) + P(A \mid B')P(B') = \sum_{i=1}^{n} P(A \mid B_i)P(B_i)$

11) De Morgan's Laws: $P(A \cup B)' = P(A' \cap B')$

$$P(A \cap B)' = P(A' \cup B')$$