Curtin College | in association with Curtin University

IPDA1005 Introduction to Probability and Data Analysis

Solution to Worksheet 3

1. For any events A and B with $P(B) > 0$, show that $P(A|B) + P(A^c|B) = 1$.
(**Hint**: Begin by using the definition of conditional probability.)

Solution: Using the definition of conditional probability, we can write that

$$\begin{aligned} P(A|B) + P(A^c|B) &= \frac{P(A \cap B)}{P(B)} + \frac{P(A^c \cap B)}{P(B)} \\ &= \frac{P(B)}{P(B)} = 1 \end{aligned}$$

and if you sketch a Venn diagram with events A and B, you will see why the second step follows from the first.

2. If $P(B|A) > P(B)$ show that $P(B^c|A) < P(B^c)$.
(**Hint**: Add $P(B^c|A)$ to both sides of the given (first) inequality and then use the result of the previous question.)

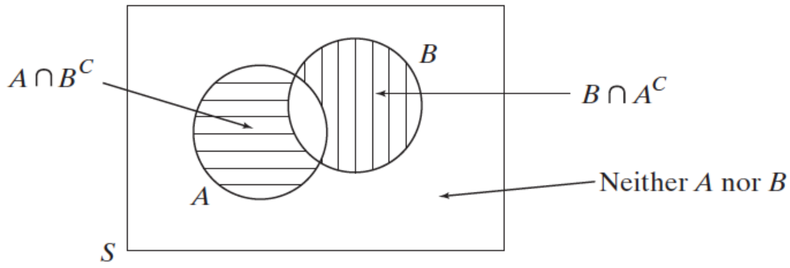
Solution:

$$\begin{aligned} P(B|A) &> P(B) \\ P(B|A) + P(B^c|A) &> P(B) + P(B^c|A) \\ 1 &> 1 - P(B^c) + P(B^c|A) \end{aligned}$$

and then simplifying and rearranging yields the desired inequality.

3. Two events A and B are defined such that (a) the probability that A occurs but B does not occur is 0.2; (b) the probability that B occurs but A does not occur is 0.1; and (c) the probability that neither occurs is 0.6. What is $P(A|B)$?
(**Hint:** Draw a Venn diagram and indicate the events whose probabilities are given, and then use the definition of conditional probability to obtain the desired result.)

Solution: The three events whose probabilities are given are shown in the Venn diagram below:



The probability that neither occurs is

$$P(\text{neither occurs}) = 0.6 = P((A \cup B)^c)$$

Hence, we can write that

$$P(A \cup B) = 1 - 0.6 = 0.4 = P(A \cap B^c) + P(A \cap B) + P(B \cap A^c)$$

so

$$P(A \cap B) = 0.4 - 0.2 - 0.1 = 0.1$$

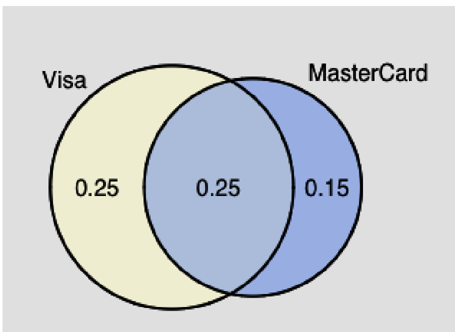
From the definition of conditional probability,

$$\begin{aligned} P(A|B) = \frac{P(A \cap B)}{P(B)} &= \frac{P(A \cap B)}{P(A \cap B) + P(B \cap A^c)} \\ &= \frac{0.1}{0.1 + 0.1} = 0.5 \end{aligned}$$

4. Consider randomly selecting a student at Curtin University, and let A denote the event that the selected individual has a Visa credit card and B be the analogous event for a MasterCard. Suppose that $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.25$. Calculate and interpret each of the following probabilities (a Venn diagram might help!):

   (a) $P(B|A)$
   (b) $P(B^c|A)$
   (c) $P(A|B)$
   (d) $P(A^c|B)$
   (e) Given that the selected individual has at least one card, what is the probability that he or she has a Visa card?

Solution: After constructing the Venn diagram below, we can calculate the required probabilities.



(a) $P(B|A)$ is the probability that those with a Visa card also have a MasterCard. $P(B|A) = 0.25/0.5 = 0.5$.

(b) $P(B^c|A)$ is the probability that those with a Visa card do not have a MasterCard, and this is also $0.25/0.5 = 0.5$.

(c) $P(A|B)$ is the probability that those with a MasterCard also have a Visa card. From the diagram, this is $P(A|B) = 0.25/0.4 = 0.625$.

(d) $P(A^c|B)$ is the probability that those with a MasterCard do not have a Visa card, which is $0.15/0.4 = 0.375$.

(e) From the diagram, $P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = 0.5/0.65 = 0.769$.

5. A spam filter is designed by looking at commonly occurring phrases in spam emails. Suppose that 80% of email is spam. In 10% of spam emails, the phrase ”free money” is used, whereas this phrase is only used in 1% of non-spam emails. A new email containing the phrase ”free money” has just arrived in your inbox. What is the probability that it is spam?
(**Hint**: Let S be the event that an email is spam and F be the event that an email contains ”free money”. Write down in symbols the conditional probability that you are asked to calculate, and then write out the information that you have been given. Express this conditional probability using the definition of conditional probability. You will need the law of total probability to calculate $P(F)$.)

Solution: In this problem, we're given the quantities $P(S) = 0.80$, $P(F|S) = 0.10$, and $P(F|S^c) = 0.01$. We're asked to find $P(S|F)$. Using the definition of conditional probability, we can write that

$$P(S|F)P(F) = P(F|S)P(S)$$

and hence that

$$P(S|F) = \frac{P(F|S)P(S)}{P(F)}$$

Using the law of total probability, we can express $P(F)$ as $P(F|S)P(S) + P(F|S^c)P(S^c)$, and substituting this into the above expression yields

$$P(S|F) = \frac{P(F|S)P(S)}{P(F|S)P(S) + P(F|S^c)P(S^c)} = \frac{0.10 \cdot 0.80}{0.10 \cdot 0.80 + 0.01 \cdot 0.20} = 0.976$$

6. **The Monty Hall Problem** On the seventies game show *Let's Make a Deal*, a contestant would choose one of three closed doors, two of which have a goat behind them and one of which has a car. After the contestant had selected a door, the host Monty Hall would open one of the two other doors, showing that the prize was not there. Then he would give the contestant a choice—either stay with the door initially selected or switch to the door that had not been opened. Should the contestant switch? Why or why not?

(a) We'll solve this problem using conditional probabilities later, but for now, sketch a tree diagram showing all the possibilities (there are 12 of them). Start with the location of the car, then the player's initial guess, and then which door is revealed.

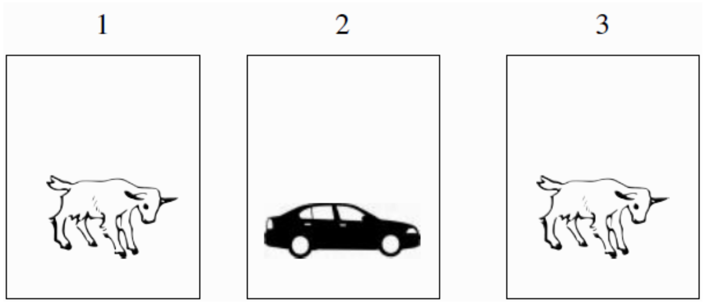


Figure 0.1: One possible configuration of the prizes at the start of the game.

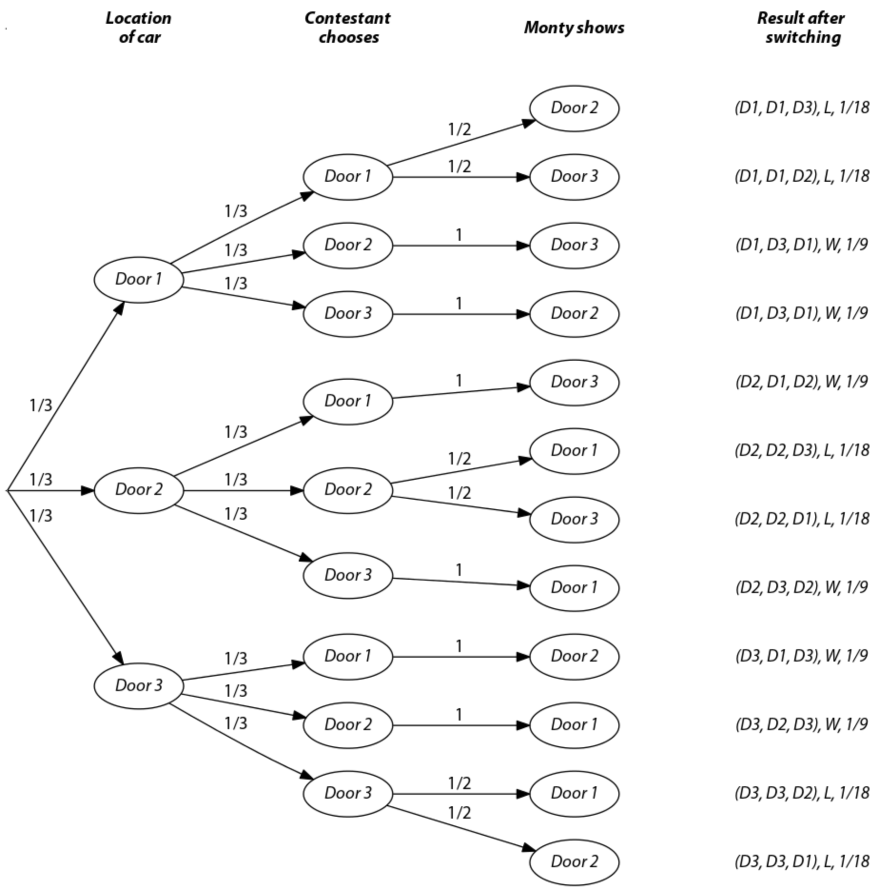
At the end, specify all the outcomes as triplets, e.g., (D1, D1, D2), and for each of the outcomes specify whether switching wins the car or not. Remember, Monty will never open the door behind which is the car, and if he has a choice, he picks a door with equal probabilities.

(b) If you were to use the naïve definition of probability, what fraction of times would the contestant win the car if s/he were to switch?

(c) Calculate the probability of each of the outcomes. What now is the probability of winning if the contestant were to switch?

Solution:

(a) The complete tree diagram is shown below:



The tree diagram shows all the possibilities, and the result associated with each possibility shows what happens if the contestant switches. For example, in two the topmost branches, the car is behind Door 1, the contestant chooses Door 1, and so Monty can show the contestant either Door 2 or Door 3. If the contestant switches doors, s/he loses. By contrast, if Door 2 is chosen, Monty has to show Door 3, and if the contestant switches, s/he wins.

(b) Clearly, there are 6 ways each in which the contestant could win/lose, so were we to use the naïve definition of probability, we'd say that it doesn't matter what the contestant does—the contestant would win/lose half the time.

(c) Note, however, that the events are not equiprobable. Each branch ('edge') shows the proportion of times that each possibility could occur, and the probability of each end result is simply the product of the individual probabilities along each path. We can see, therefore, that by switching, contestants will win the car ($6 \times 1/9 =$)2/3 of the time, not one-half!

Sources: Q1, Q2 and Q4 is from Devore and Berk (2012), Q3 is from Larsen and Marx (2014), and Q5 is from Blitzstein and Hwang (2014).

Bibliography

1. Blitzstein, J.K. and Hwang, J. (2014) *Introduction to Probability.* CRC Press/Taylor & Francis Group: Boca Raton, FL.

2. Devore, J.L. and Berk, K.N. (2012) *Modern Mathematical Statistics with Applications.* Springer: New York.

3. Larsen, R.J. and Marx, M.L. (2014) *An Introduction to Mathematical Statistics and Its Applications*, 5th ed. Prentice Hall: Boston.