DSC 40A - Homework 7

Due: Monday, November 28 at 2:00pm

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Homeworks are due to Gradescope by 2:00pm on the due date. You can use a slip day to extend the deadline by 24 hours. Make sure to correctly assign pages to Gradescope when submitting.

Homework will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should **always explain and justify** your conclusions, using sound reasoning. Your goal should be to convince the reader of your assertions. If a question does not require explanation, it will be explicitly stated.

Homeworks should be written up and turned in by each student individually. You may talk to other students in the class about the problems and discuss solution strategies, but you should not share any written communication and you should not check answers with classmates. You can tell someone how to do a homework problem, but you cannot show them how to do it.

For each problem you submit, you should **cite your sources** by including a list of names of other students with whom you discussed the problem. Instructors do not need to be cited.

This homework will be graded out of 25 points. The point value of each problem or sub-problem is indicated by the number of avocados shown.

Problem 1. Billy's DMs

Billy the avocado-farmer-turned-waiter has pivoted to being an Instagram influencer. Lots of brands want to work with him, and so he receives hundreds of Instagram DMs (direct messages) every day.

Unfortunately, a quarter of all DMs he receives are spam. Instagram's spam detector works decently well — when a DM is spam, it correctly marks it as spam 80% of the time. However, when a DM is not spam, Instagram's spam detector will incorrectly mark it as spam 2% of the time.

Throughout this question, let S be the event that a DM is spam and M be the event that a DM is marked as spam. Note that this means that $\overline{S}$ (the complement of S) is the event that a DM is not spam, and $\overline{M}$ is the event that a DM is not marked as spam.

a) From the scenario above, write out the following probabilities:

- $P(S)$
- $P(M|S)$
- $P(M|\overline{S})$

b) Given the probabilities from part (a), write out the following probabilities:

- $P(\overline{S})$
- $P(\overline{M}|S)$
- $P(\overline{M}|\overline{S})$

c) Calculate $P(S|M)$, i.e. the probability that a DM is actually spam if Instagram marks it as spam. Give your answer as a percentage rounded to 3 decimal places, and show your work.

d) 🥑🥑🥑 Now, suppose I don't know what the probability that an incoming message is spam, $P(S)$, is. Instead, I somehow know that the probability that a DM is spam given that Instagram marks it as spam, $P(S|M)$, is 80%.

What is $P(S)$ in this case? Give your answer as a percentage rounded to 3 decimal places, and show your work Assume that $P(M|S)$ and $P(M|\overline{S})$ remain unchanged from earlier in the question.

Hint: Start with your work from part (c), and see what changes in this new scenario. It may help to define a variable name for $P(S)$, e.g. $x = P(S)$.

Problem 2. The Great Chicken Sandwich Wars

Your friend claims to be able to tell which restaurant a chicken sandwich is from just a single bite. To make things harder for her, you collect chicken sandwiches from four restaurants — Popeyes, Chick-Fil-A, KFC, and McDonald's. She's able to correctly identify:

- Popeyes sandwiches 95% of the time,
- Chick-Fil-A sandwiches 80% of the time,
- KFC sandwiches 70% of the time, and
- McDonald's sandwiches 40% of the time.

You give her a sandwich, and she correctly guesses where it's from.

a) 🥑🥑🥑🥑 First, let's assume that you purchased one sandwich from all four restaurants and gave her one of those at random, so the sandwich you gave her was equally likely to come from any restaurant.

Given that she guessed it correctly, what's the probability that the sandwich was from Popeyes? Chick-Fil-A? KFC? McDonald's? Show your work.

Hint: You need to find four probabilities, and they all have the same denominator. Refer to the example from Lecture 14 for guidance.

b) 🥑🥑🥑 Now suppose that you instead purchased 20 sandwiches — 10 from Popeyes, 5 from Chick-Fil-A, 3 from KFC, and 2 from McDonalds. You choose one sandwich at random and give it to your friend.

Now, given that she guessed it correctly, what's the probability that the sandwich was from Popeyes? Chick-Fil-A? KFC? McDonald's? Show your work.

c) 🥑 Compare your answers to part (a) and part (b) above. Identify which of the four probabilities you computed increased and which decreased, and explain why this makes sense intuitively.

Problem 3. Theoretically Speaking...

Let E and F be two events in the sample space S. Assume $0 < P(F) < 1$.

a) 🥑 In one sentence, or using a Venn diagram, explain why

$$P(E \cap F) + P(E \cap \overline{F}) = P(E)$$

b) 🥑🥑🥑 If $P(E|F) = P(E|\overline{F})$, show that E and F are independent.

Hint: To show E and F are independent, you must show that $P(E \cap F) = P(E)P(F)$. You may need to use your result from part (a).

c) 🥑🥑 If E and F are independent, must it be true that $P(E|F) = P(E|\overline{F})$? If yes, explain why. If no, give a counterexample.

Problem 4. Independence and Conditional Independence

Consider the sample space $S = \{a, b, c, d, e, f, g\}$ with associated probabilities given in the table below.

outcome	a	b	c	d	e	f	g
probability	$\frac{6}{21}$	$\frac{3}{21}$	$\frac{1}{21}$	$\frac{3}{21}$	$\frac{2}{21}$	$\frac{4}{21}$	$\frac{2}{21}$

Let $X = \{d, e\}$ and $Y = \{e, f\}$. Remember to show your work for all calculations.

a) 🥑 Are X and Y independent?

b) 🥑🥑🥑 Determine if X and Y are conditionally independent given each of the following events Z.

1. $Z = \{a, b, d, e, f, g\}$
2. $Z = \{a, d, e, f, g\}$
3. $Z = \{d, e, f, g\}$