

## STAT 260 Spring 2023: Assignment 3

Due: Friday February 3rd BEFORE 11:59pm PT to Crowdmark

Please read the instructions below and in the Written Assignment 3 assignment on Crowdmark.

For full marks, your work must be neatly written, and contain enough detail that it is clear how you arrived at your solutions. Messy, poorly formatted work will receive deductions, or may not be graded at all.

Talking to your classmates about assigned work is a healthy practice that is encouraged. However, in the end, each person is expected to write their own solutions, in their own words, and in a way that reflects their own understanding. Assignment questions are not to be posted to homework "help" websites.

**Late policy:** Late assignments will be accepted until the final cutoff of 11:59pm on Sunday February 5th. Solutions submitted within 1 hour of the Friday deadline will have a 5% late penalty automatically applied within Crowdmark. Solutions submitted after 1 hour of the Friday deadline but before the final Sunday cutoff will have a 20% late penalty applied. Solutions submitted after the final Sunday cutoff will be graded for feedback, but marks will not be awarded.

1. [2 marks] Students registered in the faculties of Science, Business, and Humanities were surveyed and were asked if they regularly drink coffee or tea (or both or neither). The relative frequencies of these classifications are recorded in the table below:

	only coffee	only tea	both	neither	
Science	0.15	0.11	0.10	0.03	.39
Business	0.12	0.09	0.04	0.02	.27
Humanities	0.07	0.13	0.06	0.08	.34
	.34	.33	.20	.13	

For the following questions, if you can read the answer directly off the table you do not need to show any work. If your answer uses a combination of values from the table you need to show your work. (For example, you could show your work by writing the values being used in a summation.)

- (a) If a randomly selected student is found to drink both coffee and tea, what is the probability that they are in the faculty of Science?

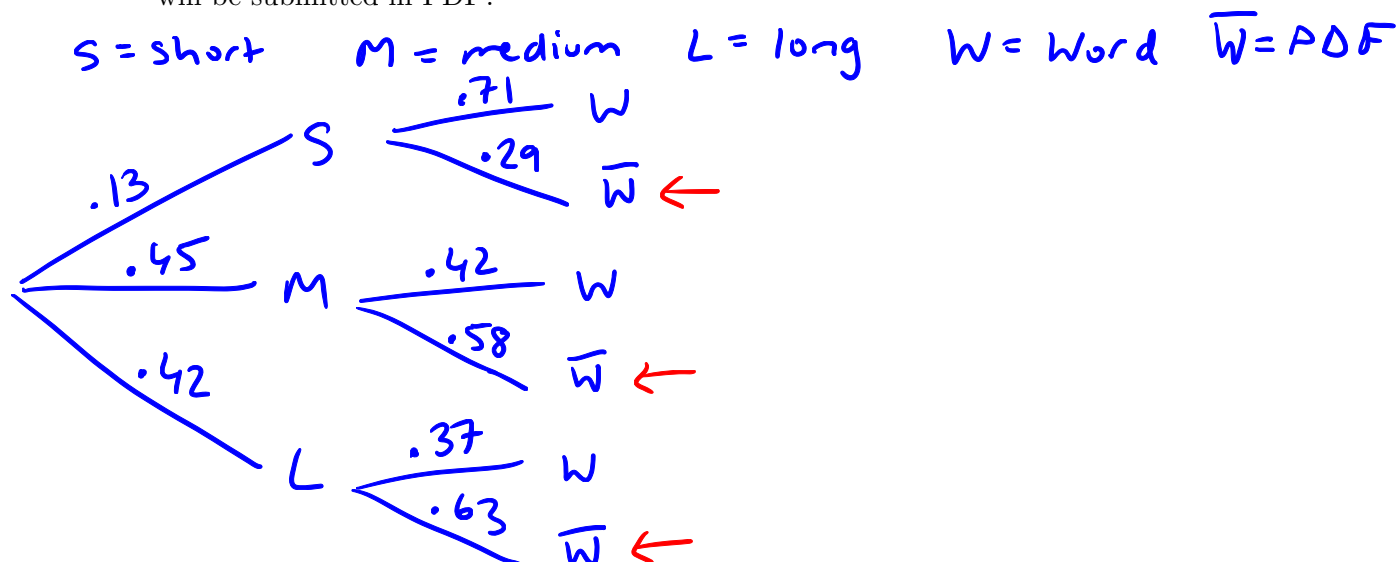
$$\begin{aligned} P(\text{science} \mid \text{both}) &= \frac{P(\text{science} \cap \text{both})}{P(\text{both})} \\ &= \frac{0.10}{0.20} = 0.50 \end{aligned}$$

- (b) If a randomly selected student is in Business or the Humanities, what is the probability that they drink only tea?

$$\begin{aligned} &P(\text{only tea} \mid \text{Business} \cup \text{Humanities}) \\ &= \frac{P(\text{only tea} \cap (\text{Business} \cup \text{Humanities}))}{P(\text{Business} \cup \text{Humanities})} \\ &= \frac{.09 + .13}{.27 + .34} = \frac{.22}{.61} = 0.3607 \end{aligned}$$

2. [2 marks] For a particular scientific journal, research articles being submitted for consideration can be short ( $\leq 1$  page), medium (2 to 3 pages), or long ( $\geq 4$  pages) in length, and can be submitted in either Word or PDF. Data on articles from the last year suggest that 13% of articles are short, 45% are medium, and the rest are long in length. For short articles, 71% are in Word, whereas 42% of medium articles are in Word, and 37% of long articles are in Word.

- (a) Suppose the journal receives a new research article for consideration. What is the probability that it will be submitted in PDF?



$$P(\bar{W}) = (0.13)(0.29) + (0.45)(0.58) + (0.42)(0.63) = 0.5633$$

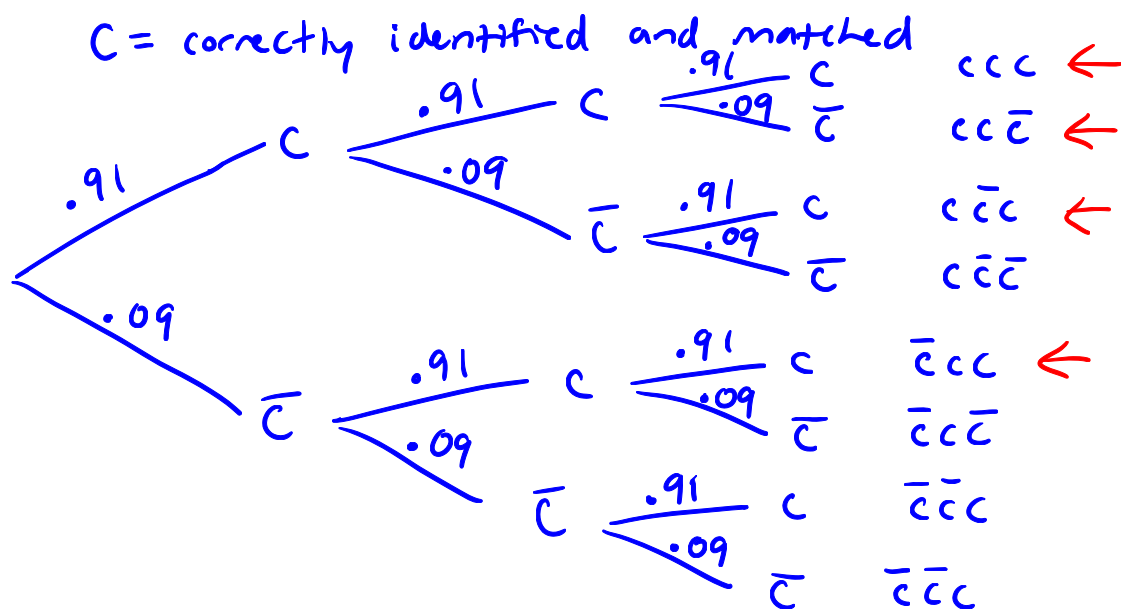
- (b) If an author has submitted a research article to the journal in Word, what is the probability that they submitted a short article?

$$\begin{aligned}
 P(S | W) &= \frac{P(S \cap W)}{P(W)} \\
 &= \frac{(0.13)(0.71)}{(0.13)(0.71) + (0.45)(0.42) + (0.42)(0.37)} \\
 &= \frac{0.0923}{0.4367} = 0.2114
 \end{aligned}$$

3. [2 marks] Both parts of this question focus on the independence of events.

- (a) A group of instructors are currently testing an online system where they will scan paper copies of exams into a website where they can then grade the papers. To do this, the system has to recognize a student's handwriting on the front page of the test in order to determine their student number and match the paper to their profile in the online system. The company who maintains the online system has declared that 91% of the time the system can correctly identify and match the handwritten student number to the student profile in the system.

The instructors perform a small test run with three student volunteers. Assuming the company's claim that 91% of papers can be correctly identified and matched to profiles, and that errors in identifying and matching papers to profiles occur independently from paper to paper, what is the probability that at least two of the three students have their papers correctly identified and matched to their profile?



$$\begin{aligned}
 P(\text{at least two correct}) &= P(CCC) + P(CC\bar{C}) + P(C\bar{C}C) + P(\bar{C}CC) \\
 &= (.91)(.91)(.91) + (.91)(.91)(.09) + (.91)(.09)(.91) + (.09)(.91)(.91) \\
 &= 0.977158
 \end{aligned}$$

- (b) Using the data table from Question 1 of this assignment, are the events "student is in Humanities" and "student drinks both coffee and tea" independent? Show a calculation to support your answer.

	only coffee	only tea	both	neither	
Science	0.15	0.11	0.10	0.03	.39
Business	0.12	0.09	0.04	0.02	.27
Humanities	0.07	0.13	0.06	0.08	.34
	.34	.33	.20	.13	

$$\begin{aligned}
 P(\text{Humanities} \cap \text{both}) &= .06 \quad \leftarrow \text{not equal} \\
 P(\text{Humanities}) \cdot P(\text{both}) &= (.34)(.20) = .068
 \end{aligned}$$

Since  $P(\text{Humanities} \cap \text{both}) \neq P(\text{Humanities}) \cdot P(\text{both})$ , the events are not independent.