AUBURN UNIVERSITY

Stat 5600/6600 - Probability and Statistics for Data Science

Midterm Exam #1 October 17, 2023

Instructions:

1. Write your name and the class you enrolled (i.e., 5600 or 6600) by your name.
2. Highest possible score in this exam is 100. You must always **explain your answers** and **show your work** to receive **full credit** or proper **partial credit**. If necessary, you can use both sides of the sheets, but make sure you have indicated doing so.
3. There are 5 questions. Please make sure you have the file with the 5 questions.
4. You are permitted both sides of a 8.5-11 sheet containing formulas, notes, or anything from the course as deemed helpful. You will turn in your exam. You may keep your note/formula sheet. No other aids of any kind are permitted.
5. Calculators are permitted but must be supplied by you. Cell phones are not permitted and should be turned off and put away before the exam begins.
6. You must always **explain your answers** and **show your work** to receive **full credit** or proper **partial credit**. If necessary, you can use the back of these pages, but make sure you have indicated doing so.
7. Q1(c) and Q2(d) are optional for STAT 5600 students, so they don't need to turn in the solution/answer for it.

Name: ________________________________

Signature: ____________________________________

Question	1	2	3	4	5	Total (out of 100)
Score						

Questions

Q1. (20 points)

(a) Are two disjoint events with non-zero probabilities always dependent? (i.e., if two disjoint events A and B have $P(A) \neq 0$ and $P(B) \neq 0$, should they be dependent?) Check with the independence identity and explain your answer.

(b) If X is the number of heads in three rolls of a coin and Y is the number of dots in roll of a die, what is the support of $(X - Y)^2$?

(c) Let A be an event in sample space S. Show that A is independent of S, also A is independent of $\emptyset$.

Q2. (20 points) We randomly draw two letters from the word PROBABILITY. That is, we randomly draw one letter and then randomly draw a second letter among the (available) letters.

(a) What is the probability that we draw a B and then an I, if the selection is done **with replacement**.

(b) What is the probability that we draw a B and then an I, if the selection is done **without replacement**.

(c) What is the probability that two letters drawn are both B's, if the selection is done **without replacement**.

(d) If you are told that the probability of randomly drawing three identical letters is positive, are the draws dependent or independent?

Q3. (20 points)

(a) Let A and B be two events such that $P(A) = 0.65$ and $P(B) = 0.75$. Find a lower bound and an upper bound for $P(A \cap B)$? Note: Here, find the best bounds, e.g., if $a \leq 1$, $a \leq 1.5$, and $a \leq 2$, the best upper bound is 1.)

(b) Let A and B be two events such that $P(A) = 0.35$ and $P(B) = 0.45$?
Find a lower bound and an upper bound for $P(A \cap B)$? (Find the best bounds)

Q4. (20 points)
An item is produced in 3 different factories, F_1, F_2 and F_3. The proportions produced in the 3 factories, and the proportions defective in each, are as follows:

Factory	% produced	% defective
F_1	50	2
F_2	30	3
F_3	20	4

(a) If you randomly select an item that was produced in Factory F_1, what is the probability that it will be defective?

(b) Let D be the event that an item is purchased and found to be defective. What is the probability that it was from factory F_1?

(c) If you had to choose an item from one of the factories to minimize the chance of getting a defective item, which factory would you choose, and why?

Q5. (20 points)
Let X be the number of heads when we (independently) toss a fair coin three times.
(a) Find the expected number of heads $E(X)$.

(b) Find the variance of number of heads $\text{Var}(X)$.

(c) Let Y be the number of tails when we (independently) toss the fair coin three times. Notice that $Y = 3 - X$. Using this relation and properties of expectation and variance, find $E(Y)$ and $\text{Var}(Y)$ (The computation of these quantities from the distribution of Y will not receive full credit).