MATH2005 Calculus, Probability, and Statistics for Computer Science
MATH2006 Calculus, Probability, and Statistics for Science
Assignment 1

Due date: 20 Feb 2020 (5:00 pm)

(Reminder: Handwritten answers would be fine.
Submit your answers in PDF format to Moodle).

Total marks: 100

1. (25 marks) The U.S. Energy Information Administration monitors all nuclear power plants operating in the United States. The following table lists the number of active nuclear power plants operating in each of a sample of 20 states.

State	Number of Power Plants
Alabama	5
Arizona	3
California	4
Florida	5
Georgia	4
Illinois	13
Kansas	1
Louisiana	2
Massachusetts	1
Mississippi	1
New Hampshire	1
New York	6
North Carolina	5
Ohio	2
Pennsylvania	9
South Carolina	7
Tennessee	3
Texas	4
Vermont	1
Wisconsin	3

   (a) Find the mean, median, and mode of this data set. (6 marks)

   (b) Find the range, variance, and standard deviation of this data set. (6 marks)

   (c) Arrange the 20 values in the table from lowest to highest. Eliminate the lowest two values and the highest two values from the data set.

      i. Find the mean (this result is called a "10% trimmed mean", since it is calculated after moving the highest 10% and the lowest 10% of the data values), median and mode of the remaining data values. (3 marks)

      ii. What advantages does a trimmed mean have over the regular arithmetic mean? (4 marks)

      iii. Then repeat (b). (3 marks)

      iv. What effect does dropping both of these measurements have on the measures of variation found in part (b). (3 marks)

2. (25 marks) A survey of engineering firms reveals that 80% have their own mainframe computer (M), 10% anticipate purchasing a mainframe computer in the near future (B), and 5% have a mainframe computer and anticipate buying another in the near future. Find the probability that a randomly selected firm:

   (a) has a mainframe computer or anticipates purchasing one in the near future (4 marks)

   (b) does not have a mainframe computer and does not anticipate purchasing one in the near future (Hint: consider the Venn diagram of the event) (7 marks)

   (c) anticipates purchasing a mainframe computer given that it does not currently have one (7 marks)

   (d) has a mainframe computer given that it anticipates purchasing one in the near future (7 marks)

3. (30 marks) "Go" is one of the oldest games in the world. The objective is to control territory by placing pieces called "stones" on vacant points on the board. Players alternate placing their stones. The player using black stones goes first, followed by the player using white stones. The following table shows the result of a study about the advantage of playing first (i.e., using the black stones) in Go.

Black Player Level	Opponent Level	Number of Wins	Number of Games
T	L	34	34
T	T	66	116
L	L	15	28
L	T	5	40
	Total:	120	218

   In this table, "T" represents the top-level players and "L" represents the low-level players.

   (a) Estimate the probability of winning when one plays first if we randomly select a game in the study. (3 marks)

   (b) Estimate the probability of winning when one plays first for each combination of the levels of players. (4 marks)

   (c) If we randomly select a game in the study and it is known that the players are at the same level, what is the probability that the black player wins? (3 marks)

   (d) Suppose there are five players. Two of them are top-level players and three of them are low-level players. Two players are randomly selected to play Go. One of these two players is randomly selected to be the black player.

      i. What is the probability that the black player wins? (6 marks)

      ii. Given that the players are at the same level, what is the probability that the black player wins? (Hint: Let A, B and C be events. It is known that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$) (8 marks)

      iii. Suppose it is known that black player wins. What is the probability that the black player is ranked lower than the white player? (6 marks)

4. (20 marks) Two years ago, a statistician estimated that the probability of the existence of Superman is 0.1 based on the evidences in the physical world. He wants to improve his estimation by considering the evidences in morality. He estimates that if Superman exists, the probability that Superman prevents evil (that is, evil does not exist) is 0.8. If Superman does not exist, the probability that the evil exists is 0.5. Given the fact that evil exists, what is the revised probability of the existence of Superman?

$\sim$ End $\sim$