DUE: November 7th, 2018 at 11:59 PM

QMM 1001 - Statistics for Data Analytics Lab 6 /16

Create and submit an R script which, when run, will print the answers to the following questions. Your R script must include a title with your name and student number and comments for each question number and letter.

1. **(10 marks) Sampling from a discrete distribution**

X	P(X)
1	0.01477
2	0.09900
3	0.25800
4	0.34200
5	0.22600
6	0.06023

   a. Create a probability space for the discrete probability distribution given in the table by creating two separate variables, one for the random variable X and one for the probabilities P(X) (2 marks).
   b. Calculate the theoretical expected value and variance for the probability distribution (2 marks).
   c. Sample from this distribution 5 times, 500 times, and 50 000 times. In each case, create a frequency table and barplot for the frequency table (3 marks).
   d. For each of the three simulations in part c) calculate the empirical expected value and variance. What values do they seem to approach? (3 marks)

2. **(6 marks)** Suppose that 20% of the population has high blood pressure. In a random sample of 200 people, what is the exact probability that
   a. Exactly 50 people have high blood pressure? (1 mark)
   b. Less than 70 people have high blood pressure? (1 mark)
   c. More than 25, but less than 50, of the people selected have high blood pressure? (2 marks)
   d. More than 160 people do NOT have high blood pressure? (2 marks)

Save your R Script as: Last Name, First Name LAB 6

Upload your R Script to the "R Assignment – Lab 6" drop box on Moodle before November 7th at 11:59 PM.