SCHOOL OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF DATA SCIENCE AND ANALYTICS

FALL 2024 – MID SEM EXAM

COURSE CODE: STA 4030A

UNIT NAME: BAYESIAN INFERENCE ND DECISION THEORY

DATE: 15TH OCTOBER 2024   **TOTAL MARKS**: 30 MARKS

INSTRUCTIONS:

For this exercise:

1. ANSWER ALL QUESTIONS
2. Do all your working in the Rmarkdown (.rmd).
3. Submissions should be in a `.rmd` **file**
4. NO SUBMISSIONS SHOULD BE DONE VIA EMAIL

QUESTIONS:

1. The table below is taken from the Hoff text and shows the joint distribution of occupations taken from a 1983 study of social mobility.

father's occupation	son's occupation				
	farm	operatives	craftsmen	sales	professional
farm	0.018	0.035	0.031	0.008	0.018
operatives	0.002	0.112	0.064	0.032	0.069
craftsmen	0.001	0.066	0.094	0.032	0.084
sales	0.001	0.018	0.019	0.010	0.051
professional	0.001	0.029	0.032	0.043	0.130

   a. Find the marginal distribution of fathers' occupations.   (2 marks)
   b. Find the marginal distribution of sons' occupations.   (2 marks)
   c. Find the conditional distribution of the son's occupation given that the father is a farmer.   (2 marks)
   d. Find the conditional distribution of the father's occupation given that the son is a farmer.   (2 marks)
   e. Comment on these results. What do they say about changes in farming in the population from which these data are drawn?   (2 marks)

2. Times were recorded at which vehicles passed a fixed point on the M1 motorway in Bedfordshire, England on March 23, 1985.[2] The total time was broken into 21 intervals of length 15 seconds. The number of cars passing in each interval was counted. The result was: 2, 2, 1, 1, 0, 4, 3, 0, 2, 1, 1, 1, 4, 0, 2, 2, 3, 2, 4, 3, 2. This can be summarized in the following table, that shows 3 intervals with zero cars, 5 intervals with 1 car, 7 intervals with 2 cars, 3 intervals with 3 cars and 3 intervals with 4 cars.

Number of Cars	Number of Occurrences
0	3
1	5
2	7
3	3
4	3
5 or more	0

   a. Do you think a Poisson distribution provides a good model for the count data? Justify your answer.   (3 marks)
   b. Assume that Λ, the rate parameter of the Poisson distribution for counts (and the inverse of the mean of the exponential distribution for interarrival times), has a discrete uniform prior distribution on 20 equally spaced values between (0.2, 0.4, …, 3.8, 4.0) cars per 15-second interval. Find the posterior distribution of Λ.   (8 marks)
   c. Find the posterior mean and standard deviation of Λ.   (5 marks)
   d. Discuss what your results mean in terms of traffic on this motorway.   (4 marks)