Homework 4

Due October 10 at 11 pm

Unless stated otherwise, justify any answers you give. You can work in groups, but each student must write their own solution based on their own understanding of the problem.

When uploading your homework to Gradescope you will have to select the relevant pages for each question. Please submit each problem on a separate page (i.e., 1a and 1b can be on the same page but 1 and 2 must be on different pages). We understand that this may be cumbersome but this is the best way for the grading team to grade your homework assignments and provide feedback in a timely manner. Failure to adhere to these guidelines may result in a loss of points. Note that it may take some time to select the pages for your submission. Please plan accordingly. We suggest uploading your assignment at least 30 minutes before the deadline so you will have ample time to select the correct pages for your submission. If you are using LaTeX, consider using the minted or listings packages for typesetting code.

1. (Halloween) In Halloween Laura and her brother Mike arrive at a house where they offer them a bowl with 2 chocolate bars. Mike grabs a random number of chocolate bars; he grabs 0, 1, or 2 with the same probability. Laura then grabs some chocolate bars out of the remaining ones; also with uniform probability (there is the same probability that she grabs 0, 1, etc.).

   (a) Model the number of bars grabbed by Mike and the number of bars grabbed by Laura as random variables and compute their joint pmf.

   (b) Compute the marginal pmf of the number of bars grabbed by Laura.

   (c) What is the conditional pmf of the number of bars grabbed by Mike if we know that Laura grabbed 1 bar?

2. (Interview) A company is interviewing candidates for a data-scientist position. They estimate that the probability of a candidate being well qualified is 0.25. This is modeled by a random variable $\tilde{q}$ that equals 1 with probability 0.25, and -1 with probability 0.75. Candidates are interviewed separately by two interviewers. The decision of the interviewers are modeled as two random variables $\tilde{i}_1 = \tilde{e}_1\tilde{q}$ and $\tilde{i}_2 = \tilde{e}_2\tilde{q}$, where $\tilde{e}_1$ and $\tilde{e}_2$ are random variables that model the probability that the interviewers make a mistake. They both equal 1 with probability 0.8 (no mistake) and -1 with probability 0.2 (mistake). $\tilde{e}_1$, $\tilde{e}_2$, and $\tilde{q}$ are all mutually independent.

   (a) What is the probability that the outcome of both interviews is positive, i.e. that $\tilde{i}_1 = 1$ and $\tilde{i}_2 = 1$?

   (b) Are $\tilde{i}_1$ and $\tilde{i}_2$ independent?

   (c) Are $\tilde{i}_1$ and $\tilde{i}_2$ conditionally independent given $\tilde{q}$?

3. (Markov chain) In this problem we consider the Markov chain corresponding to the state diagram in Figure 1. Derive an expression for the state vector of the Markov chain at an arbitrary time i assuming that we always start at state A.

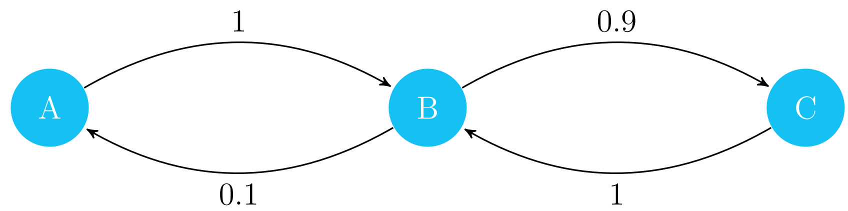

Figure 1: State diagram of a Markov chain with periodic states A and C.

4. (Precipitation data) Complete the following code to compute the joint, conditional and marginal pmfs of three random variables representing precipitation in three weather stations:

   https://github.com/cfgranda/prob_stats_for_data_science/blob/main/modeling_multivariate_discrete_data/precipitation_EXERCISE.ipynb

   Report the matrices representing any joint, conditional or marginal pmfs with two arguments. Provide the plots generated for the conditional and marginal pmfs with one argument.