STA402L: Homework 1

DUE: 11:59 PM ON WEDNESDAY, JANUARY 21, 2026

Instructions. Solutions must be submitted to Gradescope as a single PDF. Programming exercises must be completed in R, should be clearly presented, and include all R code. Lab questions are restated here for convenience, but you should refer to the lab itself for details.

Points. Book exercises (18) + lab exercises (7) = 25 points total.

BOOK EXERCISES

B1. (2 points) Let Θ be a nonempty set. Real-valued functions $p : \Theta \to \mathbb{R}$ and $q : \Theta \to \mathbb{R}$ are *proportional* if there exists a constant $C > 0$ such that $p(\theta) = Cq(\theta)$ for all $\theta \in \Theta$. If p and q are proportional probability density functions, what is the value of C? Prove your answer.

B2. (Variation of Hoff 2.1) The social mobility data from Section 2.5 in Hoff gives a joint probability distribution on $(\theta, Y) =$ (father's occupation, son's occupation); note that Hoff uses (Y_1, Y_2) instead of (θ, Y), but we will use θ and Y to suggest a more "Bayesian" perspective where the father's occupation, θ, is regarded as the unknown parameter of interest, and the son's occupation, Y, is the data. Using this joint distribution, calculate the following:

  (a) (1 point) The marginal probability that the father is a farmer.

  (b) (1 point) The marginal probability that the father is in sales.

  (c) (1 point) The probability that the son is in sales, given that the father is a farmer.

  (d) (1 point) The probability that the son is in sales, given that the father is in sales.

  (e) (1 point) Use your previous answers to compute the *posterior odds ratio*
  $$\frac{\mathbb{P}(\theta = \text{farmer} \mid y = \text{sales})}{\mathbb{P}(\theta = \text{sales} \mid y = \text{sales})} .$$

  (f) (1 point) Describe in words your solution to part (e). What role does $p(y = \text{sales})$, the marginal probability that the son is in sales, play in your answer?

B3. Hoff 2.3

  (a) (2 points)

(b) (2 points)

(c) (1 point)

B4. (5 points) Hoff 2.6

L1. (1 point) Create a code chunk and set the header parameter to TRUE and print out the top rows of the table with `head()` as above.

L2. (1 point) Generate a sequence of 100 equispaced real numbers from 0 to 1 and store it in a variable called `seq2`.

L3. (1 point) Sort the entries in `seq3` from greatest to least.

L4. (1 point) Find the variance of each row of `mat5`.

L5. (1 point) Generate 500 samples from a Beta distribution with shape parameter $[a, b] = [0.5, 0.5]$ and store the samples in a variable called `W`.

L6. (2 points) Use code from above to make a few plots of your own.