

# MAT453 - ASSIGNMENT 5

Spring 2025

Assignment Due (by 11:59 P.M.): Friday, Feb 14

**Directions:** You may discuss the exercises with other students and with the instructor, but the work you turn in must be your own. You will need to submit your R code and answers to the questions below in **one** word or pdf file.

**Exercises:** (10 points total) In this assignment, you will try to perform simulated data analysis in academic research.

Please perform the following tasks:

1. Simulate ten variables (i.e.,  $X_1, X_2, \dots, X_{10}$ ) from standard normal distributions. Each variable has a sample size of  $n = 100$ .
2. Calculate the mean parameter  $\mu_i$  as

$$\mu_i = 1 + 2X_1 + X_2 + 0.5X_5 + 1.5X_{10},$$

where  $i = 1, \dots, 100$  is the sample index.

3. Generate the count response  $Y_i$  from

$$Y_i \sim \text{Poisson}(\mu_i).$$

4. Randomly split your  $Y$  and  $X$  into training dataset with 80 samples and test dataset with 20 samples.
5. Fit a Poisson regression using your training dataset. Check if the estimated coefficients match the actual one. Make a prediction on your test dataset.