MATH 50: FINAL EXAM – V1

Instructions

- The exam period is 3 hours, and there are 7 problems.
- You may use one page of written notes, but NO electronics (computer, calculator, etc.).
- Each problem is worth 4 points.
- Show all your work but **circle your final answer**.
- Don't Cheat.
- PROJECT OPTION: If you choose the project option, I will grade the first 4 problems and one of the remaining 3.

Name: ____

Project Option? (circle one) Yes / No

If Yes, which of the additional 3 problems would you like me to grade? __________

Problem 1

Consider the following probability model:

$$X_1 \sim \text{Bernoulli}(1/2)$$
$$X_2 | X_1 \sim \text{Bernoulli}(1/3 X_1 + (1 - X_1) 1/6)$$

Write down the joint distribution of X_1 and X_2.

Problem 2

A dataset containing 100 points is fit to a single-predictor regression model, yielding the following output:

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.8884	0.285	3.114	0.002	0.322	1.454
x1	1.8436	0.284	6.487	0.000	1.280	2.408

A second predictor X_2 is then included in the model, which yields the following output:

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.0170	0.011	94.148	0.000	0.996	1.038
x1	1.0027	0.011	89.354	0.000	0.980	1.025
x2	3.0020	0.011	261.492	0.000	2.979	3.025

If X_2 and X_1 where fit to a linear regression model with X_2 as the response variable, what would be the regression slope (approximately)?

Problem 3

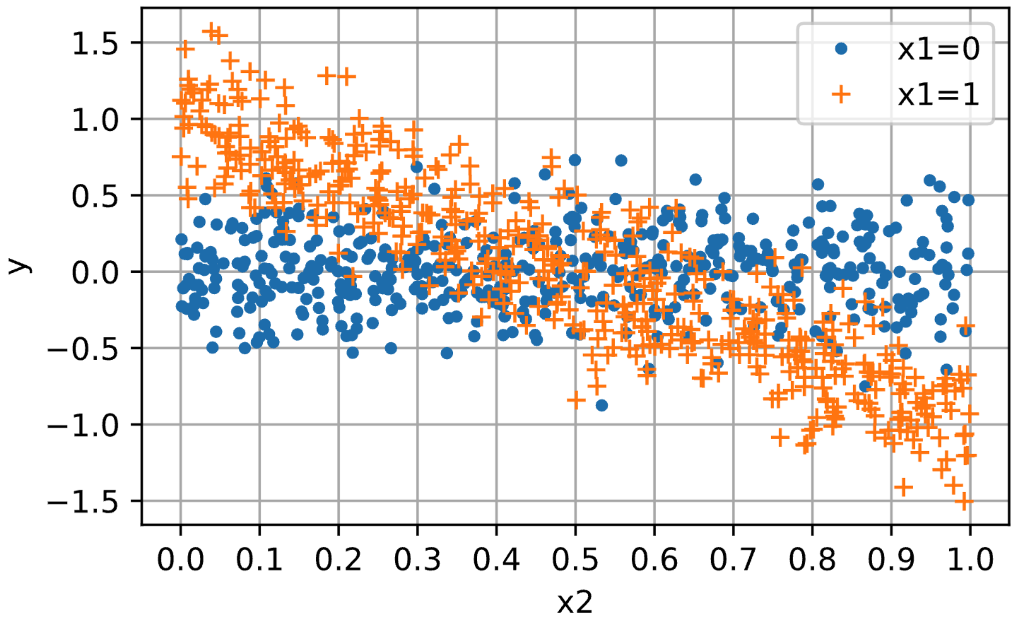
Suppose the probability of success q is estimated from a sequence of N Bernoulli trials using the estimator

$$\hat{q}_f = \frac{2S + 8}{2N + 10}$$

For what value of q is this estimator unbiased?

Problem 4

The plot below shows data from a two predictor linear regression model with an interaction between the predictors. What is the value of the interaction coefficient?



Problem 5

For a single-predictor linear regression model, derive the formula

$$\text{cov}(Y, X) = \beta_1 \sigma_X^2$$

Problem 6

Consider the sin basis functions in the Fourier model:

$$\phi_i(t) = \sin(2\pi i t), \quad i = 1, \ldots, K$$

Give an example of a distribution on t such that $\phi_1(t)$ and $\phi_2(t)$ are NOT orthogonal (that is $\mathbb{E}[\phi_1(t)\phi_2(t)] \neq 0$).

Problem 7

When performing Bayesian inference in a polynomial regression model, it is desirable to have stronger (meaning lower variance) priors on the regression coefficients corresponding to higher order monomials (larger j values). This ensures the monomials X^j which grow very fast have smaller regression coefficients (unless there is very strong evidence in the data that they should e large). Explain how the same effect could be achieved with regularization.

Score Sheet

Question	Points Earned
1	
2	
3	
4	
5	
6	
7	
Total	