

# FINAL EXAM

**ISyE6420**

Fall 2020

Released November 24, 12:00pm – due December 8, 11:55pm. This exam is not proctored and not time limited except the due date. Late submissions will not be accepted.

Use of all available electronic and printed resources is allowed except direct communication that violates Georgia Tech Academic Integrity Rules.

Name \_\_\_\_\_

Problem	1	2	3	Total
Score	/33	/33	/34	/100

**1. Vasoconstriction.** The data give the presence or absence ( $y_i = 1$  or  $0$ ) of vasoconstriction in the skin of the fingers following inhalation of a certain volume of air ( $v_i$ ) at a certain average rate ( $r_i$ ). Total number of records is 39. The candidate models for analyzing the relationship are the usual **logit**, **probit**, **cloglog**, **loglog**, and **cauchyit** models.

Data are given as follows.

y: 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,  
0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1

v: 3.7, 3.5, 1.25, 0.75, 0.8, 0.7, 0.6, 1.1, 0.9, 0.9,  
0.8, 0.55, 0.6, 1.4, 0.75, 2.3, 3.2, 0.85, 1.7, 1.8,  
0.4, 0.95, 1.35, 1.5, 1.6, 0.6, 1.8, 0.95, 1.9, 1.6,  
2.7, 2.35, 1.1, 1.1, 1.2, 0.8, 0.95, 0.75, 1.3

r: 0.825, 1.09, 2.5, 1.5, 3.2, 3.5, 0.75, 1.7, 0.75,  
0.45, 0.57, 2.75, 3, 2.33, 3.75, 1.64, 1.6, 1.415,  
1.06, 1.8, 2, 1.36, 1.35, 1.36, 1.78, 1.5, 1.5, 1.9,  
0.95, 0.4, 0.75, 0.3, 1.83, 2.2, 2, 3.33, 1.9, 1.9, 1.625

(a) Transform covariates  $v$  and  $r$  as

$$x_1 = \log(10 \times v), \quad x_2 = \log(10 \times r).$$

(b) Estimate posterior means for coefficients in the logit model. Use noninformative priors on all coefficients.

(c) For a subject with  $v = r = 1.5$ , find the probability of vasoconstriction.

(d) Compare with the result of **probit** model. Which has smaller deviance?

**2. Magnesium Ammonium Phosphate and Chrysanthemums.** Walpole et al. (2007) provide data from a study on the effect of magnesium ammonium phosphate on the height of chrysanthemums, which was conducted at George Mason University in order to determine a possible optimum level of fertilization, based on the enhanced vertical growth response of the chrysanthemums. Forty chrysanthemum seedlings were assigned to 4 groups, each containing 10 plants. Each was planted in a similar pot containing a uniform growth medium. An increasing concentration of  $\text{MgNH}_4\text{PO}_4$ , measured in grams per bushel, was added to each plant. The 4 groups of plants were grown under uniform conditions in a greenhouse for a period of 4 weeks. The treatments and the respective changes in heights, measured in centimeters, are given in the following table:

Treatment			
50 g/bu	100 g/bu	200 g/bu	400 g/bu
13.2	16.0	7.8	21.0
12.4	12.6	14.4	14.8
12.8	14.8	20.0	19.1
17.2	13.0	15.8	15.8
13.0	14.0	17.0	18.0
14.0	23.6	27.0	26.0
14.2	14.0	19.6	21.1
21.6	17.0	18.0	22.0
15.0	22.2	20.2	25.0
20.0	24.4	23.2	18.2

Solve the problem as a Bayesian one-way ANOVA. Use STZ constraints on treatment effects.

(a) Do different concentrations of  $\text{MgNH}_4\text{PO}_4$  affect the average attained height of chrysanthemums? Look at the 95% credible sets for the differences between treatment effects.

(b) Find the 95% credible set for the contrast  $\mu_1 - \mu_2 - \mu_3 + \mu_4$ .

**3. Hocking–Pendleton Data.** This popular data set was constructed by Hocking and Pendleton (1982) to illustrate influential and outlier observations in regression. The data are organized as a matrix of size  $26 \times 4$ ; the predictors  $x_1, x_2$ , and  $x_3$  are the first three columns, and the response  $y$  is the fourth column. The data are given in `hockpend.dat`.

(a) Fit the linear regression model with the three covariates, report the parameter estimates and Bayesian  $R^2$ .

(b) Is any of the 26 observations influential or outlier (in the sense of CPO and `comulative`)?

(c) Find the mean response and prediction response for a new observation with covariates  $x_1^* = 10, x_2^* = 5$ , and  $x_3^* = 5$ . Report the corresponding 95% credible sets.