Stat 616 Generalized Linear Models, Homework 10 **Spring 2019**

Due: Monday, April 29

For each of the following study, propose appropriate generalized linear model(s), including Normal linear regression model, to analyze the data and address the research questions. You do NOT need to conduct the actual analysis.

1. Write down the theoretical model by clearly stating the distribution of the response variable, the link function, and the predictors. Briefly justify your choice.
2. Briefly explain how you will use your model to address the research questions.

Study 1: Air traffic control

A data set consists of information on the number of "en route operational errors" (relative rare) of air traffic control specialists (ATCSs) from 1997 to 2003, summarized by the experience and age of the ATCSs. The following four variables are in the data set

- Experience Categories (Exp):
  1 = 0-3 years
  2 = 4-9 years
  3 = 10-14 years
  4 = 15-19 years
  5 = 20-24 years
  6 = 25+ years
- Age Categories (Age):
  1 = 55 years or less
  2 = 56 years or more
- Number of en route operational errors (Errors)
- Number of ATCSs (Population)

Use the above data and a Generalized Linear Model to analyze whether:

(1) an ATCS's age or experience have an effect on the occurrence of en route operational errors,

(2) there is an interaction between the ATCS's age and experience with respect to en route operational errors.

Study 2, Personal Space

Harris, Luginbuhl, and Fishbein (1978) conducted a social-psychological experiment that examined reactions to invasions of personal space. The research took place in a field setting provided by a public escalator. The primary results of the study were presented in the following contingency table.

Four variables are in the table:

- density of people on the escalator, rated as either high or low;
- the sex of the subject;
- the sex of the intruder;
- whether or not the subject reacted in some manner to the intrusion.

The goal of the study is to investigate whether and how density and gender (both the subject's and intruder's gender) may be associated with a person's response to a personal space invasion.

(1) Density	*(2) Sex of Subject*	*(3) Sex of Intruder*	*(4) Response*	
			Yes	No
Low	Male	Male	18	1
		Female	15	8
	Female	Male	17	5
		Female	12	7
High	Male	Male	13	6
		Female	16	4
	Female	Male	10	9
		Female	14	6

Study 3: Lung Cancer treatment

The data in the following table show tumor responses of male and female patients receiving treatment for small-cell lung cancer. There were two treatment regimes. For the sequential treatment, the same combination of chemotherapeutic agents was administered at each treatment cycle. For the alternating treatment, different combinations were alternated from cycle to cycle (data from Holtbrugger and Schumacher, 1991).

Treatment	Sex	Progressive disease	No change	Partial remission	Complete remission
Sequential	Male	28	45	29	26
	Female	4	12	5	2
Alternating	Male	41	44	20	20
	Female	12	7	3	1

Note that the patient's response categories: Progressive disease, No change, Partial remission and Complete remission have orders.

Fit statistical model (or models) to

(1) estimate the probabilities for each response category taking treatment and se effects in account,

(2) investigate whether there is a difference in the responses for the two treatment regimes,

(3) investigate whether males and females respond to the treatment similarly.

Study 4. The effect of distraction

A researcher conducted an experiment to evaluate the effect of common distractions. In the study, a group of programmers are randomly assigned to one of 4 treatment groups (no noise, relaxing/soothing music, pop music, other people having a conversation) and are asked to complete a coding task. The following variables are collected:

- ID: Identification number
- Experience: Years in programming
- Gender: Gender of the programmer
- Treatments: Type of distractions: 1 = no noise, 2 = relaxing/soothing music, 3 = pop music, 4 = other people having a conversation.
- Time: Time to finish the coding task
- Success: Does the code work correctly? (0 = No, 1 = Yes).

The goal of the study is to investigate whether, and how, the distraction from music or noise may affect the time to successfully complete the task. (Hint: How can we use the data from those who failed the coding task?)

Study 5. Low birthweight

Low birth weight is an outcome that has been of concern to physicians for years. A woman's behavior during pregnancy (including diet, smoking habits, and receiving prenatal care) can greatly alter the chances of carrying the baby to term and, consequently, of delivering a baby of normal birth weight. Data were collected on 189 women from a medical center, 59 of which had low birth weight babies (weighing less than 2500 grams) and 130 of which had normal birth weight babies. Some of the variables identified in data set have been shown to be associated with low birth weight in the obstetrical literatures.

- ID: Identification Code
- LOW: Low Birth Weight (0 = Normal, 1 = Low, birth weight < 2500g)
- AGE: Age of the Mother in Years
- LWT: Weight of the Mother in Pounds at the Last Menstrual Period
- RACE: Race (1 = White, 2 = Black, 3 = Other)
- SMOKE: Smoking Status During Pregnancy (0 = No, 1 = Yes)
- PTL: History of Premature Labor (0 = None, 1 = One, 2=Two, etc.)
- HT: History of Hypertension (0 = No, 1 = Yes)
- UI: Presence of Uterine Irritability (0 = No, 1 = Yes)
- FTV: Number of Physician Visits During the First Trimester (0, 1, 2, ...)
- BWT: Birth Weight in Grams

The goal of the current study was to ascertain if these variables were important in the population being served by the medical center where the data were collected.

This is the end of HW 10.