Homework #10

Study 1: Air traffic control

i. Distribution of response variable: Poisson distribution, with *log* link function.

ii. Model: $log(\text{Errors}) = log(\text{Population}) + \beta_0 + \beta_1 \cdot \text{Exp} + \beta_2 \cdot \text{Age}$

iii. To address the research questions:

(1) Check the significance of the coefficient of Age or Exp to see if an ATCS's age or experience have an effect on the occurrence of en route operational errors.

(2) Keep the significant predictors in the model. Use the LRT to test if an interaction term between the ATCS's age and experience should be included in the model.

Study 2: Personal Space

i. Distribution of response variable: Binomial distribution, with *logit* link function.

ii. Model: $log\frac{P(Response=Yes)}{P(Response=No)}$ vs (Density) * (Sex of Subject) * (Sex of Intruder)
Where $(Density) * (SexofSubject) * (SexofIntruder)$ includes these three variables, their two-way interaction term and three-way interaction term.

iii. To address the research questions: Check the significance of the coefficients of the terms.

Study 3: Lung Cancer treatment

i. Distribution of response variable: Multinomial distribution, with *logit* link function.

ii. Proportional odds model: $log\frac{P(Response\leq j)}{P(Response>j)} = \beta_0 + \beta_1 \cdot \text{Treatment} + \beta_2 \cdot \text{Sex}$
Where $j = 1, 2, 3$. And from low to high progressive disease = 1, No change = 2, Partial remission = 3, Complete remission = 4.

iii. To address the research questions: Use the model to estimate the probabilities for each response category. Check the significance of the coefficients of the predictors.

Study 4. The effect of distraction

i. Distribution of response variable: Exponential distribution, with *log* link function.

ii. Model: $log(Time) = \beta_0 + \beta_1 \cdot \text{Experience} + \beta_2 \cdot \text{Gender} + \beta_3 \cdot \text{Treatments}$

iii. To address the research questions: Censor the data from those who failed the coding task to build a survival object for the model. Set distraction type *no noise* as baseline, and check the coefficient for each dummy variable for distraction type to investigate whether, and how, the distraction from music or noise may affect the response variable.

Study 5. Low birthweight

i. Distribution of response variable `LOW`: Binomial distribution, with *probit* link function.

ii. Model: $log\frac{P(LOW=1)}{P(LOW=0)} = \beta_0 + \beta_1 \cdot \text{AGE} + \beta_2 \cdot \text{LWT} + \beta_3 \cdot \text{RACE} + \beta_4 \cdot \text{SMOKE} + \beta_5 \cdot \text{PTL} + \beta_6 \cdot \text{HT} + \beta_7 \cdot \text{UI} + \beta_8 \cdot \text{FTV}$

iii. To address the research questions: Check the significance of the coefficient of each variable in the model.