**Research Study**:

Participants were given several scenarios about a potential treatment to a medical condition and were asked to rate if the treatment was a clear choice over an alternative. They rated this choice across several conditions (PP, NP, PF, NF).

Remember to paste your output in this document and upload your R script to blackboard to complete this assignment.

**Random Variables:**

* id – participant ID number.

**IV X-Variables:**

* type – Probability (probably would help treat condition) or Ambiguity (ambiguous about how well it would treat the condition)
* present – OR and AND conditions, where the participant saw different forms of the description about the potential treatment.
* Available and affordable – rating on how available and affordable the participant thought the treatment was (averaged 1-7 over several variables).
* Mastery, ambiguity, and risk taking – personality variables used to measure how much a person could master information, withstand ambiguity, and would take risks (averaged 1-7 over several variables).

**DV – Y-variable:**

* PP, NP, PF, NF clear choice selections rated 1-7

**Research Question:** Use the IV variables to predict the overall clear choice variable to see if any of the experimental manipulations (type, present), thoughts on the treatment (available and affordable), or the personality variables (mastery, ambiguity, and risk taking) are significantly related. Use all of these IVs at once in the appropriate model.

**Data Screening:** You can use the fake regression to data screen, since the data are in wide format. Do not screen the participant ID!

**Accuracy:**

1. Check the data for out of range scores.
   1. Include a summary showing you do/do not have out of range scores.
   2. If necessary, fix the out of range scores by making them NA.

**Outliers:**

1. Calculate Mahalanobis distance scores for your data.
   1. What is your *df* for the cut off score?
   2. What is the cut off score?
   3. How many Mahalanobis outliers did you have?
   4. Delete those outliers!

**Additivity:**

1. Include a symnum table of the continuous IV variables.
2. Are any of the variables too highly correlated?

**Normality:**

1. Include the multivariate normality histogram.
2. Interpret the graph. Does it indicate multivariate normality?

**Linearity:**

1. Include the multivariate QQ plot.
2. Interpret the graph. Does it indicate multivariate linearity?

**Homogeneity and Homoscedasticity:**

1. Include the multivariate residuals plot.
2. Interpret the graph.
   1. Does it indicate homogeneity?
   2. Does it indicate homoscedasticity?

**Regression test:**

1. First, melt the data so the continuous DVs are one column of data.
2. Steps:
3. Non-random intercept only model.
4. Random intercept model using participants as the random factor.
5. Random intercept model using participants as the random factor with the predictor variables.
6. Random intercept (participants) + random slope factor using risk taking as a random slope.
7. Fill in the following table with the anova() comparison between all four of these models.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | ***df*** | **AIC** | **Log Likelihood** | **Ratio** | ***p*** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

1. Predictors: fill in the following table with the information for the predictors from the random intercept + predictors step (model 3).

|  |  |  |  |
| --- | --- | --- | --- |
| **Predictor** | ***b*** | ***t*** | ***p*** |
|  |  |  |  |
|  |  |  |  |

**Write up:**

1. In this section, instead of a formal write up, you should discuss each model explaining what happened:
   1. Non-random intercept:
      1. What was the average clear choice score (the intercept)?
      2. How much did that score vary?
   2. Random intercept:
      1. What was the average clear choice score (the intercept)?
      2. How much did that score vary?
      3. Is this model significantly better than a non-random intercept? If so, what does that imply/mean?
   3. Random intercept + predictors:
      1. Which predictors were significant?
      2. Interpret their *b* values (remember, MLM is still regression).
      3. Is this model significantly better than a random intercept only model?
   4. Random intercept + predictors + random slope:
      1. What is the random slope for risk taking? Does that appear to be different than the non-random slope from the step before?
      2. Is this model significantly better than the random intercept + predictors model?