Test Review Guide

The following topics were covered through the course and represent a base knowledge of the material covered. I have broken them up by weeks to help you organize your understanding of the material. Please go back through your notes, lecture PowerPoints, handouts, readings, and assignments to review each of the concepts presented below. You are also free to use each other and me as a resource as you study.

Week 1:

1. Be able to identify outcomes and predictors.
2. Be able to determine if a hypothesis test has a significant result based on the *p*-value compared to alpha. When would you reject or fail to reject the null hypothesis?
3. Explain what a *p*-value is.
4. How is the effect size different from the hypothesis test?
5. Know the various conventions for effect size. Why are *R*-squared values smaller for group means than correlations?
6. What can we conclude from the Central Limit Theorem and why is it important?
7. Under what circumstances are the mean, median, and mode most appropriate to report? What about the corresponding measure of variability?
8. What is the difference between a positive and a negative correlation?
9. What are we looking for in the correlation matrix in a regression output? How do you know if it might be a concern?
10. What does the coefficient of determination tell us?
11. Be able to interpret a slope and a constant coefficient when the predictor is 1) continuous and 2) nominal.
12. What is the difference between standardized and unstandardized slope coefficients?
13. What does the standard error of the estimate tell us? Be able to interpret.
14. Explain what dummy coding is. Be able to provide an example of dummy coding.
15. Be able to interpret output for a regression analysis.
16. Be able to interpret a confidence interval for a slope.
17. Be able to interpret a standardized slope and a constant coefficient.
18. Explain how the standard error of skewness can be used to determine normality.

Week 2:

1. What is the relationship between *F-* and *t-*tests?
2. When might you use a regression analysis? An ANOVA analysis?
3. What are the limitations to regression analyses? ANOVA analyses?
4. How can you tell normality from a published mean and standard deviation?
5. What is the purpose of the Levene’s test?
6. Why would we prefer an ANOVA to *t*-tests?
7. Why is the *F-*distribution a positively skewed distribution?
8. Be able to identify the type of ANOVA given the number of predictors and their levels.
9. What are the general assumptions to ANOVA?
10. Why is the central limit theorem useful in ANOVA modeling?
11. What is the purpose of the Welch *F-*test?
12. Be able to obtain ANOVA output.
13. Be able to interpret ANOVA output.
14. Be able to interpret post-hoc test output.

Week 3:

1. Explain what main effects and interactions are. If an interaction is present, would you interpret the main effect results? Why or why not?
2. When can use the central limit theorem to assume normality?
3. What is the difference between a within- and a between subjects/groups ANOVA? What is a mixed model ANOVA?
4. What is preferred: a within- or between subjects/groups ANOVA? Why is it preferred?
5. What can you do to check the dependency of populations assumption in a within-groups/subjects ANOVA?
6. What assumption is sphericity testing? When do you use Greenhouse-Geisser vs Huynh-Feldt?
7. When might you use MANOVA?
8. When would we likely want to complete a descriptive discriminant analysis?
9. Be able to obtain MANOVA output.
10. Be able to interpret MANOVA output.
11. Be able to obtain DDA output.
12. Be able to interpret DDA output.

Week 4:

1. What are the general assumptions for the general linear model?
2. What are three measures used to investigate the influence of data points?
3. Why can’t we rely on just investigating scatterplots for determining unusual observations?
4. What is the order of operations for investigating model violations?
5. Explain what residuals are.

Week 5:

1. How would you know if a studentized residual for an observation is potentially influential (you don’t need to memorize the formula)?
2. When measuring leverage, how would you know if the observation is potentially influential?
3. When measuring Cook’s distance, how would you know if the observation is potentially influential?
4. When measuring Mahalanobis distance, how would you know if the observation is potentially influential?
5. If you have a potential influential observation, what could you do about it?
6. When transforming data, which would we prefer to transform, the predictor or the outcome? Explain.
7. Briefly explain Tukey’s ladder of transforms and how it relates to the Rule of the Bulge.
8. What is involved in a sensitivity analysis?
9. When investigating the univariate statistics of studentized residuals, what are you looking for?
10. When looking at a residual scatterplot to the predictors, what are you looking for?
11. When transforming data, there are some values that cannot be transformed for logs, inverses, and square roots. What are the values and what must you do before transforming?
12. How would you interpret slopes for log transformed variables when both predictor and outcome are logged, when outcome is logged but not predictor, and when predictor is logged but not outcome?

Week 6:

1. Be able to compare among models.
2. Be able to transform variables.
3. Be able to produce a residual scatterplot and interpret.
4. Give at least two reasons why multicollinearity is a concern.
5. Briefly explain when multicollinearity is not a concern.
6. There are several indicators of multicollinearity. What are they?
7. Explain why correlations are not the best indicator for multicollinearity?
8. What are the general rules of thumb for tolerance and VIF? How is tolerance and VIF related?
9. What are three ways to prevent multicollinearity based on research design? Briefly explain each.

Week 8:

1. What are five ways to adjust for multicollinearity statistically?
2. Generally, explain how PCA addresses multicollinearity?
3. What is the purpose of data reduction techniques such as PCA and factor analysis?
4. Be able to get and interpret collinearity diagnostics in SPSS.

Week 9:

1. What are the five ways to determine the number of components/factors to extract? Which is the best method?
2. What is the difference between an orthogonal rotation and an oblique rotation?
3. When determining which items load onto each component/factor, what criteria is typically used?
4. What is the difference between a principal components analysis and a factor analysis?
5. What are the ways we can improve factor structure stability (integrity)?
6. Why are component/factor scores preferred to using sum or mean composite scores?
7. When determining the internal reliability of a component/factor, what cutoffs for Cronbach’s alpha and average reliability coefficients are deemed as acceptable?
8. Be able to interpret PCA output.
9. Be able to produce PCA output in SPSS
10. Be able to interpret output for a reliability analysis.
11. Be able to produce output for a reliability analysis.

Week 10:

1. What is the difference between logistical regression and the other regression models that we have learned this semester?
2. How are odds calculated from probabilities? Probabilities from odds?
3. What are the assumptions for logistic regression and explain why they are different from the other regression models this semester?
4. What is the interpretation for the effect size (pseudo-*R2*) in a logistic regression?
5. How are model fit tests different from goodness-of-fit tests?

Week 11:

1. What is the difference between logistical regression and the other regression models that we have learned this semester?
2. How are odds calculated from probabilities? Probabilities from odds?
3. What are the assumptions for logistic regression and explain why they are different from the other regression models this semester?
4. What is the difference among the assumptions of independence of observations, independence of predictors, and independence of irrelevant alternatives?
5. What is the interpretation for the slope coefficient in a binary, multinomial, and ordinal logistic regression with nominal and continuous predictors?
6. What is the interpretation for the effect size (pseudo-*R2*) in a logistic regression?
7. How are model fit tests different from goodness-of-fit tests?
8. Be able to get output for a dichotomous, ordinal, and multinomial logistic regression.
9. Be able to interpret output for a dichotomous, ordinal, and multinomial logistic regression.

Week 12:

1. Be able to edit an Excel file to prepare it for Mplus
2. Be able to edit Mplus syntax for errors.
3. Be able to interpret Mplus descriptive output (e.g., variable normality).
4. Be able to interpret Mplus output for simple linear regression, multiple regression, and moderation.