PSY 6280, Psychological Statistics: Regression

1. Suppose we have two quantitative variables, X and Y. Describe the relationships between the covariance of X and Y, Pearson’s correlation between X and Y, and the simple linear regression coefficient when predicting Y based on X.

2. Suppose multiple linear regression was used to predict Y using X1 and X2 as predictors. Explain what each of the following values would represent (i.e., how it would be interpreted).
   - Intercept
   - Unstandardized Regression Coefficient for X1
   - t test for X1
   - 95% Confidence Interval for the Unstandardized Regression Coefficient for X1
   - Standardized Regression Coefficient for X1
   - Mean Square Error
   - Standard Error of the Estimate
   - $R^2$
   - Adjusted $R^2$
   - ANOVA $F$ test for the regression model

3. Describe how Pearson’s correlation, the point-biserial correlation, the Phi correlation, Spearman’s Rho, and Kendall’s tau are related.

4. Suppose we have a dependent variable, Y, and three possible predictors (X1, X2, X3). Briefly describe the automatic variable selection techniques (stepwise, backward, forward, all subsets) and then compare the automatic variable selection approach to path analysis.

5. Describe the three patterns of missing data (MCAR, MAR, MNAR) and then list the advantages/disadvantages of deletion methods (e.g., listwise deletion), nonstochastic imputation methods (e.g., mean substitution), and stochastic imputation methods (e.g., multiple imputation).


6. Compare the studentized deleted residual, Cook’s D, and standardized dfbetas for detecting outliers. Briefly describe the methods of handling outliers once they have been detected.

7. What is collinearity? What impact does collinearity have on the linear regression model? What are the options for resolving collinearity problems?
PSY 6290, Psychological Statistics: ANOVA

Statistical Procedures Covered
- One-Way ANOVA with Multiple Comparisons
- Welch ANOVA with Multiple Comparisons
- Trend Analysis
- Factorial ANOVA
- One-Way ANCOVA with Multiple Comparisons
- One-Way RM ANOVA with Multiple Comparisons
- Two-Way RM ANOVA (1 Between, 1 Within) with Multiple Comparisons
- Two-Way RM ANOVA (2 Within) with Multiple Comparisons

1. Know the assumptions and robustness of each statistical procedure.
2. If given a research scenario determine the appropriate procedure, appropriate follow-up analyses, and how to control the familywise alpha.
3. Be able to describe when a statistical procedure would be used (characteristics of the independent and dependent variables). As well, give an example.
4. Discuss the advantages and disadvantages of repeated measures designs.
5. Explain the general linear model (GLM) in conjunction with both ANOVA and multiple regression.
6. Explain the logic of ANOVA
7. Explain how the per comparison alpha, familywise alpha, and experimentwise alpha are related.
8. Compare the univariate and multivariate approaches for RM ANOVAs.
PSY 6580, Multivariate Statistics

Describe the situation (when you are going to use the technique), with one example, of each of the following techniques. Describe the characteristics (either qualitative or quantitative) of Independent (Predictive) and Dependent (Criterion) Variables.

1. One-Way MANOVA
2. Factorial MANOVA
3. MANCOVA
4. Discriminant Analysis & Classification Analysis
5. Logistic Regression
6. Canonical Correlation
7. Principal Component Analysis
8. Factor Analysis
9. Structural Equation Modeling
10. Cluster Analysis

Compare and contrast MANOVA and MANCOVA.
Compare and contrast MANOVA and Discriminant Analysis.
Compare and contrast Logistic Regression and Discriminant Analysis.
Compare and contrast Principal Components Analysis, Factor Analysis, and Structural Equation Modeling.

Show how Wilk’s Lambda, Pillai’s Trace, Hotelling Lawley Trace, and Roy’s Greatest Root are related to the eigenvalues of a matrix. Be able to calculate each of the procedures if given the appropriate eigenvalues.

Show how Wilk’s Lambda, Pillai’s Trace, Hotelling Lawley Trace, and Roy’s Greatest Root are related to the canonical correlations. Be able to calculate each of the procedures if given the appropriate canonical correlations.

Describe the relative power and robustness of Wilk’s Lambda, Pillai’s Trace, Hotelling-Lawley’s Trace, and Roy’s Greatest Root procedures.

Describe the various approaches for follow-up analyses to a MANOVA.

Describe the guidelines for determining the number of principal components or number of factors to retain.
1. Discuss the differences between Classical Test Theory (CTT) and Item Response Theory (IRT) in terms of model, assumptions, and practical limitations.

2. Describe one, two, and three-parameter logistic model in IRT, highlight their differences.

3. Describe the differences between Maximum Likelihood method and Bayesian method in ability estimation in IRT.

4. Describe basic idea of differential item functioning (DIF) in IRT.

5. Describe differences between dichotomous IRT and polytomous IRT models.

6. List different polytomous IRT models.
PSY 6460, Factor Analysis Comps Questions  (based on Thompson)

1. Describe the three major purposes of factor analysis and how it is used in each application.

2. Describe, compare and contrast Exploratory and Confirmatory Factor Analysis.

3. What is meant by factor rotation? What is meant by an oblique rotation? How are factors rotated? When? Why?

4. How do you determine the “best” exploratory factor solution? (Include in your answer a discussion of determining the number of factors in a data set).

5. Describe Compare and Contrast Principal Components and Principal Axes analyses.

6. What is meant by higher-order factors? When would you test for them? How would you interpret them?

7. Briefly describe the six possible two-mode techniques.

8. Describe, compare and contrast cross-validation and bootstrapping.

9. What is involved in model identification, estimation and evaluation of fit?

10. How does one test model invariance in CFA?
Structural Equation Modeling (PSY 6550) Comprehensive Examination Items

1. What are the main purposes of SEM? Make sure you mention what we are trying to find out by analyzing data, how we can say it is possible to find some latent traits from some observed variables, and the advantage of latent traits.

2. Explain two different models in SEM: Measurement and Structural models.

3. Explain both exogenous and endogenous variables.

4. Why do we want to perform derivatives and/or partial derivatives for different functions? Explain the main purpose for derivatives and partial derivatives in statistics along with the procedure.

5. Explain each of the four effect coefficient matrices (B, \( \Gamma \), \( \Lambda_x \), and \( \Lambda_y \)).

6. Explain \( \Phi \), \( \Psi \), \( \Theta \), and \( \Theta_\varepsilon \).

7. Explain variance and covariance. Describe what each of them tells us. Why do we want to analyze covariance?

8. In estimating the covariance matrix among \( y \)'s, we know that \( \text{cov}(y) = \Lambda_y [\text{Cov}(\eta\eta') \Lambda_y'] + \Theta_\varepsilon \). Why did it eventually become \( \Lambda_y [(I - B)^{-1} (\Gamma \Phi \Gamma' + \Psi) (I - B)^{-1}] \Lambda_y' + \Theta_\varepsilon \)?

9. If we have a super matrix for \( x \)'s and \( y \)'s with the dimension of 30X30 and try to estimate 200 parameters, what is the df-value for the \( \chi^2 \)-test for SEM?

10. Explain the relationship among GFI (Goodness-of-Fit Index), AGFI (Adjusted Goodness-of-Fit index), and RMSEA (Root Mean Square Error of Approximation).

11. Compare and contrast two model misfit indices: residuals and modification index.

12. List Joreskog’s (1971) five steps in conducting multi group analyses. List only the title and the null hypothesis.

13. Describe the differences between testing covariance structure and testing latent mean structure.
PSY 6560, Computer-based Statistical Packages

Be prepared to describe the purpose of a particular SAS statement, option, or procedure (e.g., What is the purpose of the drop statement?). The SAS statements, options, and procedures you should be familiar with include:

**Creating a SAS dataset**
- Data statement
- Infile statement
- Cards or Datalines statement
- Input statement (list, column, and formatted input)
- Using informats within the input statement
- Using the single trailing @ and double trailing @@ with the Input statement

**Creating a SAS dataset from other SAS datasets**
- Set Statement
- Merge statement…with By statement
- Update Statement
- Drop option
- Keep option

**Programming: Processing Observations within a SAS dataset**
- If-then statements
- If-then-else statements
- Do Loops (Iterative, Do-While, Do-Until)
- Arrays
- Creation of new variables using an assignment statement, for example \( w = x + 10; \)
- Output statement
- Retain statement

**Global Statements**
- Libname Statement
- Filename Statement
- Title Statement
- Footnote Statement
- ODS Statement

**Statements Used in DATA Steps and PROC Steps**
- Where Statement
- By Statement
- Label Statement
- Format Statement

**Procedures**
- PROC CONTENTS
- PROC DATASETS
- PROC FORMAT
- PROC FREQ
- PROC MEANS
- PROC PRINT
- PROC REPORT
- PROC SORT
- PROC SUMMARY