Possible Quantitative Comps Questions

PSY 6280, Psychological Statistics: Regression

1. 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
   • Standard Error of the Estimate
   • R square
   • Adjusted R square
   • ANOVA F test for the regression model
2. Explain how OLS (ordinary least square) works to estimate the regression coefficients.
3. What is collinearity? What impact does collinearity have on the linear regression model? What are the options for resolving collinearity problems?
4. Explain the five assumptions for regression models.
5. Explain the difference between reference coding and effect coding
6. Explain the difference between the following model selection methods: best subset selection, forward selection, backward selection, ridge regression, and lasso.
7. Explain the general mechanism of cross validation for the model selection purpose.
8. Explain the difference between the following cross-validation methods: the validation set approach, leave-one-out cross validation, K-fold cross-validation.
9. Explain sensitivity and specificity.
10. Interpret the regression coefficients in logistic regression.
11. Explain the difference between the following classification methods: logistic regression, linear discriminant analysis, and the tree method.

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 logic of ANOVA

6. Explain how the per comparison alpha, familywise alpha, and experimentwise alpha are related.

7. Compare the univariate, the multivariate, and the mixed-model approaches for one-way RM ANOVAs.

PSY 6580, Multivariate Statistics

1. 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.
   - One-Way MANOVA
   - Factorial MANOVA
   - MANCOVA
   - Discriminant Analysis & Classification Analysis
   - Logistic Regression
   - Survival Analysis
   - Canonical Correlation
   - Cluster Analysis
   - Principal Component Analysis
   - Factor Analysis
   - Structural Equation Modeling

2. Compare MANOVA and MANCOVA.
3. Compare Discriminant Analysis with a one-way MANOVA and with logistic regression.
4. Compare survival analysis with logistic regression and with linear regression.
5. Compare Principal Components Analysis, Factor Analysis, Path Analysis, and Structural Equation Modeling.

6. Be able to calculate the inverse, determinant, eigenvalues, and eigenvectors of 2x2 matrices.

7. 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.

8. 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.

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

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

11. Describe the guidelines for determining the number of principal components or number of factors to retain.

PSY 6210, Advanced Psychometrics

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 the joint and marginal Maximum Likelihood methods in estimating person and item parameters in IRT.
4. Describe the basic idea of differential item functioning (DIF) in IRT.
5. List different polytomous IRT model and briefly describe their differences.
6. Given the IRTPRO output, interpret the psychometric properties of the items (e.g., fitted model, estimated difficulty and discrimination parameters, model-data fit information, quality of the item, implications of the items).

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 (CFA).

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. Describe, compare and contrast cluster analysis of subjects on variables with Q-technique factor analysis.

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

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

9. How does one test model invariance in CFA?

10. How does doing a path analysis in SEM differ from doing a CFA? What are the potential cause of an ill-conditioned (or not positive definite) matrix?

PSY 6565, Behavioral Statistics Using R

For all questions, please copy and paste your R code and results printed in R console.

1. We want to count the number of heads for tossing the coin several times (head=1, tail=0)? Please write a function named "toss" and complete the following tasks.
   a. Toss 20 coins and compute the sum of heads
   b. Toss 10 coins and count the number of tails
2. Please write a function to solve the following quadratic equations (e.g., $ax^2+bx+c=0$), the function should return the solutions for the quadratic equations. Please copy and paste your R code and the solutions below. Hint: the quadratic formula is $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$
   a. $x^2+4x+4=0$
   b. $3x^2+1x+1=0$
   c. $3x^2+2x-1=0$
3. Write a function using if else statements to determine whether a random integer value is an even or an odd number? Use the following numbers to check your R codes. The numbers are 0, 4, 5. The results should look like the following
   the number is an even number.
   the number is an odd number.
   the number is neither an odd nor an even number (0).
   Hint: use the mathematical operator: %% (modulus)
4. If you're tossing a pair of dice, stop when you toss two 1s. You need to write a function to toss pair of dice first. The results should show each pair of dice you've tossed, and how many times it takes you to get two 1s.
5. Let's roll a pair of dice to demonstrate "Lucky 7". Please roll a pair of dice 1000 times using the "for loop". For each roll, compute the sum of the pair of dice. After rolling the dice 100 times, you should have the 100 sums saved in a vector, and compute the mean of the sums in the vector, the mean should be close to 7.
   Below is the roll() function to roll two dice.

```
roll <- function () {
        die=1:6
        dice <- sample(die, 2, replace=T)
        sum(dice)
      }
```

6. you're tossing a coin and will stop when you see two tails in a row (i.e., two consecutive tails).

PSY 6560, Computer-based Statistical Packages

The third edition of the SAS Certification Prep Guide: Base Programming will be used for students who completed SAS before Fall 2018 (i.e., students taking comps during F2018, S2019, U2019).

There will be 10 multiple-choice questions. At least 70% of the multiple-choice questions (i.e., 7 of the 10) will be taken from the questions at the end of the SAS certification guide.

There will be five essay questions. The possible essay questions are listed below.

The fourth edition of the SAS Certification Prep Guide: Base Programming will be used for students who completed SAS during Fall 2018 – Present (i.e., for students taking comps after the summer of 2019).

There will be 10 multiple-choice questions. At least 70% of the multiple-choice questions (i.e., 7 of the 10) will be taken from the questions at the end of the SAS certification guide.

There will be five essay questions. The possible essay questions are listed below.

Possible Essay Questions

1. List the rules for naming SAS data sets, variables, and libraries.
2. Describe the purpose of DO Loops (iterative DO loop, DO WHILE, DO UNTIL) within a DATA step.
3. Describe when (a) column input, (b) formatted input, (c) list input, and (d) modified list input are used.
4. Describe the process for creating a SAS data set from a raw data file.
5. Describe the process for creating a permanent SAS data set using an existing SAS data set.
6. Describe the process for (a) creating character and numeric user-defined formats, (b) permanently assigning the user-defined formats, and (c) temporarily assigning the user-defined formats.
7. Summarize the execution phase for processing instream or external data.
8. Summarize the execution phase for reading observations from a SAS data set.
9. Describe how to concatenate two SAS data sets into a new data set and describe what is contained in the new SAS data set.
10. Describe how to match-merge two SAS data sets into a new data set and describe what is contained in the new SAS data set.