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 2 classification methods: logistic regression and the tree method.

PSY 6290, Psychological Statistics: ANOVA

Statistical Procedures Covered
- One-Way 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. An example for a *t* test is shown in the multivariate comps section.

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 multivariate and the mixed-model approaches for one-way 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.

> *For example, if you were asked to do so for an independent samples t test, it would be: The t-test is used when a researcher wants to compare averages of a quantitative dependent variable for a qualitative independent variable with two groups. The observations must be independent. One example is to compare the average salary (Quantitative DV) based on employees' degree status (yes, no; Qualitative IV).*

- One-Way MANOVA
- One-Way MANCOVA
- Discriminant Analysis
- Classification Analysis
- Logistic Regression
- Survival Analysis
- Cluster Analysis
- Principal Component Analysis
- Factor Analysis
- Structural Equation Modeling

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

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

5. 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.

6. 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.

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

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 R 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.

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 (for Factor Anlysis).

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 Confirmatory Factor Analysis?

PSY 6565, Behavioral Statistics Using R

For all questions, please write down your R code.

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. Hint: the quadratic formula is $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
   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? 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

(50%) There will be 10 multiple-choice questions. The multiple-choice questions will be taken from **Programming 1: Essentials – Lesson Quizzes (**prepared date 30Nov2018) and **Programming 2: Data Manipulation Techniques – Lesson** Quizzes (prepared date 24Apr2019). You can obtain PDF copies from the Supplemental Materials portion of the e-learning.

(50%) There will be two coding tasks from the tasks described below.

`Sample1.xlsx`

ID	W	X
1	18	20
2	15	.
3	13	14

`Work.Sample2`

ID	X	Y
4	5	20
5	10	25
6	15	10

`Mtsu.Sample3` where mtsu has been assigned to a folder.

ID	Y	Z
6	16	8
4	30	12

Possible Coding Tasks:

Task 1. Write a SAS program to read the Sample1.xlsx file and store it as a permanent data set called Research1 that is stored in the mtsu folder. Obtain a description of the Research1 contents.

Task 2. Write a SAS program to combine the work.Sample2 and mtsu.Sample3 into a new temporary data set called Reserach2n3. The new data set should only contain ID and Y variables and should only contain observations with Y values greater than 20. Print the new data set.

Task 3. Write a program that contains all observations from the Research1, work.Sample2, and mtsu.sample3 data into a single data set. Then create the following variables:
the sum of W, X, Y, and Z
the average of W,X,Y, and Z that ignores missing data
the average of W, X, Y, and Z that yields a . value if any of the data are missing.

Task 4. Write a program that creates a temporary data set called SampleTotals from the work.Sample2 data set. Create an accumulating variable that creates a running total of X and a running total of Y. The only observation in the SampleTotals data set should be SumX (which will have a value of 30) and SumY (which will have a value of 55).

PSY 6575, Multilevel Modeling (For Students Taking Comps in S2021 or Later)

1. Please explain what is intraclass correlation, and how it is different from design effect?
2. Please specify the two-level random intercept model, random slope model (with one predictor at level 1), and random intercept and random slope model in all models (please use the correct notations).
3. Please interpret the level 1 and level 2 residuals (for both random intercepts and random slopes at level 2).
4. Please specify the three-level random intercept model with no predictors at all three levels (please use the correct notations).
5. Please explain the similarities between longitudinal analysis and latent growth modeling.
6. Please explain the difference between the multilevel model and the cross-classified model.
7. Please explain how does Bayesian's approach estimate parameter estimates conceptually.