

RESEARCH METHODS AND STATISTICS AREA EXAM

October 2016

Committee: Bill Carbonaro (chair), Elizabeth McClintock, Sarah Mustillo

=====

PART ONE: STATISTICS

=====

SECTION I: Select four of the following six items and briefly address the following for each one:

- define or explain the main significance of the item; and
 - describe the item's importance for research practice (i.e., what problem does it pose?, or what problem does it address?).
-
- a) Wald test
 - b) Multicollinearity
 - c) Errors-in-variables regression
 - d) Instrumental variable
 - e) Autocorrelation
 - f) Cronbach's alpha

SECTION II: Select and answer either question 2 or question 3.

QUESTION 2

Answer the following questions about the difference between random and fixed effects models.

- A. Describe the key differences between random and fixed effects models. Be sure to highlight the different assumptions of each model. Also, please describe differences in the estimation and interpretation of the parameters in the two models.
- B. Explain how the two models could be applied to the same data, and yet produce different substantive findings.
- C. How should researchers decide whether a random or fixed effects model should be used to analyze his/her data?
- D. Finally, describe the strengths and limitations of random and fixed effects models.

QUESTION 3

Answer the following questions about structural equation modeling (SEM).

- A. Describe the model's key assumptions and concepts (e.g., exogenous, endogenous, manifest, latent variables; measurement, structural, recursive, and nonrecursive models).
- B. Describe some of the applications of SEM as a method. What are the main advantages of this approach? What problems does it address that alternative methods cannot?
- C. What are limitations and disadvantages of SEM's?
- D. How does one assess goodness of fit for SEM's? Discuss developments and debates regarding goodness of fit statistics and potential limitations of models evaluated as fitting the data well.

SECTION III: Select and answer either question 4 or question 5.

QUESTION 4

Missing data is a ubiquitous feature of almost every large-scale data set. Answer the following questions regarding how to handle missing data in statistical analysis.

- A. Describe different types of missing data that regularly appear in large-scale data sets. Should all missing data be treated in the same way? Explain.
- B. Describe the main strategies for addressing missing data in one's analysis that have been proposed and used in practice. Be sure to describe the strengths and limitations of each approach.
- C. Describe the key assumptions about missing data (MCAR and MAR), and indicate the implications of these assumptions for strategies for addressing the problem in an analysis.
- D. What are the challenges involved with handling missing data in panel data? How should one address these challenges in practice?

QUESTION 5

Three of the most commonly violated OLS assumptions are linearity, homoskedasticity, and the zero conditional mean assumption (i.e., $E(u|x)=0$). For each violation:

- A. Explain the circumstances that typically cause a violation of the assumption.
- B. Explain the implications of the violation for the parameters, with specific reference to the bias and efficiency of the estimators.
- C. Discuss 2-3 methods for dealing with the violation, and provide an explanation of how each method addresses the problem.

RESEARCH METHODS AND STATISTICS AREA EXAM

October 2016

Committee: Bill Carbonaro (chair), Elizabeth McClintock, Sarah Mustillo

=====

PART ONE: METHODS

=====

SECTION I: Select four of the following six items and briefly address the following for each one:

- define or explain the main significance of the item; and
 - describe the item's importance for research practice (i.e., what problem does it pose?, or what problem does it address?).
- A. Grounded theory
 - B. Endogenous selection
 - C. Counterfactual
 - D. Differential attrition
 - E. Stable unit treatment value assumption (SUTVA)
 - F. Construct validity

SECTION II: Select and answer either question 2 or question 3.

QUESTION TWO

Imagine that you have been asked to estimate the average annual income for English professors at American universities by collecting survey data, via a probability sample. Using the "total survey error" approach, address the following:

- A. Identify the potential sources of error in your estimate, and explain specifically how each source of error might bias your estimate of average annual income in the population of interest.
- B. Discuss specific strategies that you would utilize that would reduce each source error in your estimate.

QUESTION THREE

Address the following questions regarding sampling designs in research.

- A. There are many different types of probability samples. Three of the most common are simple random samples, multistage cluster samples, and stratified random samples.
- Describe each sampling design, and highlight the important contrasts between them.
 - Describe the main strengths and limitations of these different designs.
 - Indicate what adjustments need to be made when analyzing data from these different samples with statistical models.
- B. Qualitative researchers employ many different nonprobability sampling strategies in their research.
- Select and describe three commonly used sampling methods in qualitative research.
 - Explain the main reasons why a researcher might select each sampling strategy in his/her research.
 - Describe the main strengths and limitations of these different designs.

SECTION III: Select and answer either question 4 or question 5.

QUESTION FOUR

You are interested in conducting a large-scale study that examines whether divorce affects the mental health of children.

- A. What are the challenges involved with designing a study that tries to establish a causal link between these two variables? Be sure to specifically identify the main threats to the internal validity of your study.
- B. How would you design your study to address these challenges? Be sure to describe what type of sample you would use, what type of data you would collect, and what methods you would utilize to analyze your data.

QUESTION FIVE

Field experiments have become an increasingly popular tool for studying inequality and discrimination in the labor market. Consider gender inequality in the the labor market, and address the follow questions about field experiments.

- A. Researchers rely upon “audit studies” to measure the effect of an applicant’s gender on hiring decisions. Some studies send trained auditors to apply for jobs “in person,” while other studies submit fictitious resumes to employers. What are the advantages and disadvantages to each methodological approach to measuring discrimination in the labor market?
- B. Some field experiments elicit responses (“call backs”) from actual employers who are advertising positions in the labor market. In contrast, other studies recruit undergraduates and ask them to report what hiring decision they would make if they were hiring someone for a fictitious job. Evaluate these two approaches to studying gender effects on hiring in terms of both experimental and mundane realism.
- C. Methodologically, what conditions must be met to ensure that the experiment is actually measuring discrimination, and not some other confounding variable? How might a researcher design his/her experiment to ensure that s/he is meeting this assumption?
- D. An experimenter has the choice of submitting one applicant per job (randomly assigned as male or female), or submitting both applicants (a male and a female) for each job. Is there an advantage to one design over the other, or are they roughly equivalent to each other? Explain.