**Chapter 1: Concepts and Examples of Research**

Preview Questions

1. What is regression analysis?
2. What are the key concepts of applied regression analysis?
3. What are multivariable techniques?
4. What’s the difference between experimental, quasi-experimental, and observational research?
5. What is a response variable?
6. What is a predictor variable?
7. What are examples of the type of research that can be done using regression analysis?

Reading Summary

Key concepts in empirical research

* Empirical research examines the relationship among a number of variables.
* Variables are factors that are measured for observational units or subjects.
* Multivariable methods use statistical techniques to examine the relationship among at least three variables.
* Regression analysis is a type of multivariable method.

Types of research

* Experimental, in which observational units or subjects are randomly assigned levels of predictor variables.
* Quasi-experimental, in which observational units or subjects are assigned levels of predictor variables but not in a random manner.
* Observational, in which observational units are not assigned levels of predictor variables.

Characteristics of research

* Multivariable methods are applicable to all three types of research.
* Each type of research provides a different level of confidence one can have in the results.
* The response variable is the dependent variable, which is influenced by predictor variables.
* Predictor variables are independent variables, which influence the response variable.
* Typically one (1) response variable and one (1) or more predictor variables.
* Potential for drawing definitive conclusions: observational < quasi-experimental < experimental
* Difficulty of implementation: observational < quasi-experimental < experimental

Measuring variables

* Error is unavoidable in measuring variables.
* Error in measuring variables gives rise to the need for statistical design and analysis.

Types of statistical inference

* Estimation comprises describing the characteristics and strength of the relationship among variables by quantifying them.
* Hypothesis testing comprises prosing explanations about the relationship among variables, stating probabilities about the reasonableness of such explanations, and drawing conclusions based on the stated probabilities.

**Chapter 2: Classification of Variables and the Choice of Analysis**

Preview Questions

1. What is gappiness?
2. What is level of measurement precision?
3. What is meant by descriptive orientation?
4. How do you overlap variable classification systems?
5. How do you choose a method of analysis?

Reading Summary

Approaches to classifying variables

* Classifications for variables help in deciding which methods to use for an analysis.
* Three methods for classifying variables are:
  + By the gappiness
  + By the level of measurement precision
  + By the descriptive orientation

Gappiness

* Gappiness refers to whether or not gaps exist between successive observations of the values of a variable.
* Discrete variables have gaps.
* Non-numeric data may be numerically coded as discrete variables.
* Continuous variables DO NOT have gaps (i.e., between any two values another value can potentially exist).
* Data on discrete variables are represented by a line chart to display sampling frequency.
* Data for continuous variables are grouped into intervals (e.g., histogram) to display sampling frequency.
* Discrete variables can be treated like continuous variables for analysis purposes when the values of a variable are not far apart and cover a wide range of numbers.
* Continuous variables are sometimes treated like discrete variables for analysis purposes.
* Considerations when deciding whether to categorize continuous variables:
  + Makes data collection easier
  + Simplifies the presentation of results
  + Information is lost
* Considerations for deciding when to categorize continuous variables:
  + At the time of collection
    - Less expensive
    - Less time consuming
    - Less precise
    - More likely to introduce human error (i.e., classification error)
  + At the time of analysis
    - Less error prone
    - Enables consideration of various classification schemes
* Errors
  + Classification error is a factor with discrete variables.
  + Measurement error is a factor with continuous variables.

Level of measurement precision

* Three (3) levels of measurement precision
  + Nominal (i.e., categorical) indicates different categories for the variable.
  + Ordinal indicates both categorization and ordering of categories for the variable.
  + Interval indicates ordering and also the distance between categories has meaning.
* Ratio variables or ratio-scale variables are interval variables in which the scale has a true zero.
* Measurement error for ratio-scale variables often have a non-normal distribution and are proportional to the size of the measurement.
* An interval scale is also ordinal and nominal.
* An ordinal scale is also nominal.

Descriptive orientation

* Descriptive orientation indicates whether a variable is meant to describe or be described by other variables.
  + Response or dependent variables are typically denoted by letter Y.
  + Predictor, regressor, or independent variables are typically denoted by letter X.
* Control variables are independent variables that affect relationships among other independent variables in a study but are of no interest.
* Control variables are sometimes referred to as nuisance variables, covariates, or confounders.

Overlap of Classification Schemes

* Any variable can be labeled according to each scheme.
* See Figure 2.5 (p. 12) for diagram of classification scheme overlap.
* All nominal variables are discrete but NOT all discrete variables are nominal.
* All continuous variables are ordinal and interval but NOT all ordinal and interval variables are continuous.

Choosing a method of analysis

* There are four considerations:
  + Purpose of the research
  + Mathematical characteristics of the variables
  + Statistical assumptions about the variables
  + Data collection method (i.e., sampling procedure)
* See Table 2.1 (p. 13) regarding guidance for choosing analysis methods
* Methods not covered
  + Nonparametric methods, which don’t require the data to fit a normal distribution
  + Cluster analysis