CASE STUDY APPLICATIONS OF STATISTICS IN INSTITUTIONAL RESEARCH

By MARY ANN COUGHLIN and MARIAN PAGANC
Case Study Applications of Statistics in Institutional Research

by
Mary Ann Coughlin
and
Marian Pagano

Number Ten
Resources in Institutional Research

A JOINT PUBLICATION OF
THE ASSOCIATION FOR INSTITUTIONAL RESEARCH
AND
THE NORTHEAST ASSOCIATION FOR INSTITUTIONAL RESEARCH
# Table of Contents

**Acknowledgments** ................................................................. i

**Introduction** ................................................................. 1

**Chapter 1: Basic Concepts** ...................................................... 3
  - Characteristics of Variables and Levels of Measurement .......... 3
  - Descriptive Statistics ..................................................... 8
  - Probability, Sampling Theory and the Normal Distribution ......... 16

**Chapter 2: Comparing Group Means: Are There Real Differences Between Average Faculty Salaries Across Departments?** .... 20
  - Description of Case Study .................................................. 20
  - Research Designs ............................................................. 21
  - Statistical Procedures:
    - T-test ........................................................................... 23
    - One-Way ANOVA ............................................................... 26
    - Factorial ANOVA ............................................................... 32
    - Analysis of Covariance (ANCOVA) ...................................... 33
    - Multivariate ANOVA (MANOVA) .......................................... 34
    - Nonparametric Tests of Mean Differences ....................... 34

**Chapter 3: Correlation: Measuring the Relationship Between SAT and First-Year GPA** ............................................... 36
  - Description of Case Study .................................................... 36
  - Some Background About Correlation .................................... 36
  - Statistical Procedures:
    - Pearson Correlation Coefficient ....................................... 37
    - Nonparametric Correlations ................................................. 42
    - Regression Analysis ........................................................ 44

**Chapter 4: Chi-Square: Differences in the Frequency of Enrolling and Non-Enrolling Students by the Quantity of Aid Offered** ........ 48
  - Description of Case Study .................................................... 48
  - Some Background About Chi-Square .................................... 49
  - Statistical Procedures:
    - Two-Way Chi-Square ........................................................ 49
    - One-Way Chi-Square ........................................................ 55
    - Chi-Square Automatic Interaction Detection (CHAID) ............ 56

**Chapter 5: Selecting the Appropriate Statistic** ....................... 58
  - Using the Design Tree ......................................................... 58
  - Examples of Using the Design Tree ....................................... 60

**Appendix A: SPSS Commands for Statistical Procedures** ............. 62

**References** .............................................................................. 64

**Index** ..................................................................................... 65
Acknowledgments

We would like to thank the Northeast Association for Institutional Research and the Association for Institutional Research for their support and patience during the preparation of this document. We would like to thank Blake Whitten for his copy editing and indexing assistance. Also, we would like to thank the teachers, students and colleagues who have taught us how to make material that is feared by many more accessible and understandable. Marian Pagano would like to especially thank Dr. David Drews of Juniata College, Huntingdon, PA, for his excellent and inspirational teaching. Two of the graphics in Chapter One of this monograph come from class notes taken in Dr. Drews' Introduction to Statistics course. Mary Ann Coughlin would like to thank Dr. Barbara Jensen who has long served as her mentor, friend, and editor of this text.
Case Study Applications of Statistics in Institutional Research

Introduction

Statistics has been defined as “a collection of methods for planning experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting and drawing conclusions based on the data” (Triola, 1995, p. 4). Does this sound like an abbreviated charge for the Office of Institutional Research at your institution? Saupe (1990) discussed several of the above-mentioned activities as functions of Institutional Research. He encapsulated this view by defining Institutional Research as “research conducted within an institution of higher education to provide information which supports institutional planning, policy formation and decision making” (p. 1).

While statisticians are more likely to disagree than agree on a variety of issues, general agreement exists that the field consists of two subdivisions: descriptive and inferential statistics. Descriptive statistics consists of a set of techniques for the important task of describing characteristics of data sets and for summarizing large amounts of data in an abbreviated fashion. Inferential statistics goes beyond mere description to draw conclusions and make inferences about a population based on sample data. While most Institutional Researchers are quite knowledgeable in the area of descriptive statistics, many are less comfortable with inferential techniques. The use of inferential techniques can bring critical enlightenment to policy and planning decisions.

Thus, this monograph focuses on the application of statistical techniques to Institutional Research; theory, application, and interpretation are the main tenets. The ultimate goal of the authors is to enhance the researchers’ knowledge and interpretation of their data through statistical analyses. The text begins with a general background discussion of the nature and purpose of both descriptive and inferential statistics and their applications within Institutional Research. Each additional chapter follows a case study format and outlines a practical research question in Institutional Research, illustrates the application of one or more statistical procedures, analyzes data representing a hypothetical institution, and incorporates the output from these analyses into information that can be used in support of policy- and decision-making.

This document is designed to give the reader a broad overview of or refresher in descriptive and inferential statistics as they are applied to case studies in Institutional Research. In this format, this is quite a challenge as a wide range of statistical concepts and procedures is covered in relatively few pages. No intent is made to document the numerical calculation of statistics or to prove statistical formulas. For further information in any of these areas, please consult the list of references.

Statistical software packages are standard equipment in most Institutional Research offices as they handle complex analyses and large data files relatively effortlessly. While it is important that an Institutional Researcher be able to use statistical packages, this monograph is not designed to teach you how to do so. Rather, the emphasis of this monograph will be on the theory, application, and interpretation of statistical analyses. Many statistical packages are available on a wide range of computer platforms that can be utilized to perform these analyses. The Statistical Package for the Social
Sciences (SPSS) is the choice of the authors for statistical software and SPSS for Windows was used to analyze the data from each case study, yet any standard statistical software can perform these analyses. For your convenience, the statistical commands that perform the analyses discussed in this text using SPSS for Windows are included in Appendix A. These commands can be readily translated into any standard statistical software.

Before proceeding to the main text, some practical limitations of this monograph should be declared. The first and most important of these points is that the best statistics cannot save an inferior research design. Statistical procedures are no substitute for forethought. Although several robust research designs are illustrated throughout this text, the primary emphasis of the monograph is not dedicated to design concepts. Suffice it to say that the research design is the foundation of a good study. If the design is weak, the analysis will crumble. Remember, the statistician's favorite colloquial expression: "Garbage in, Garbage out." Secondly, the topics and techniques covered in this text are for the most part standard and accepted practices. However, as in most fields, few absolutes exist with many differing opinions. Please feel free to review the references for other suggested practices and approaches to the tasks presented here.

Finally, the case studies and example data utilized in this text are fabricated studies representing fictitious institutions, but are designed to represent real research questions facing Institutional Researchers. In no way should the case studies or data be associated with the authors or the institution of the authors. In the real world, the questions facing individual Institutional Researchers are as varied as the researchers themselves and their respective institutions. Yet the case studies have been carefully developed to represent the diversity in our profession and to present a variety of statistical procedures with universal application.
Chapter One: Basic Concepts

This chapter is designed to give a brief overview of some basic statistical concepts and terms that will be used throughout this text and is divided into the following four sections: Characteristics of Variables and Levels of Measurement; Descriptive Statistics; Probability, Sampling Theory and the Normal Distribution; and Inferential Statistics. While the titles of these sections imply that this chapter will cover all the material taught in an Introductory Statistics text, please be advised that the text flows briskly through each of the topics and is meant to serve only as a refresher. For more detailed information, please refer to Sprinthall (1987), Levin and Fox (1994), Triola (1995) or any basic statistical textbook.

Characteristics of Variables & Levels of Measurement

A variable is an indicator or measure of the construct of interest. A variable can be anything that has more than one value (e.g., sex, age, SAT scores). Variables should have operational definitions clearly stated. An operational definition of a variable defines specifically the variable measured and, unless it is a universally accepted definition, should be clearly published with any data and analyses. For example, when using SAT scores from an inquiry survey, SAT score could be operationally defined as “the self-reported scores on both the math and verbal sections of the SAT examination.” This alerts the reader that the results from this analysis might vary from results reported from the Educational Testing Service. For other examples, refer to Example Box 1.

Example Box 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Operational Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTE - Faculty</td>
<td>Total number of full time faculty plus ( \frac{1}{2} ) part-time faculty headcount</td>
</tr>
<tr>
<td>Student</td>
<td>Anyone enrolled during a semester for at least 1 credit hour</td>
</tr>
<tr>
<td>Compensation</td>
<td>Salary plus fringe benefits</td>
</tr>
</tbody>
</table>

While this operational definition is quite straightforward, some operational definitions get sticky and lead to the issue of construct validity. Many constructs or concepts in educational research are wide-open to interpretation. Institutional Researchers are usually quite familiar with the nuances of construct validity as we deal with the definition of FTE, full-time faculty, and other seemingly simple variables whose definitions are often capricious. The important point is to clearly communicate how you have measured the constructs (i.e., the underlying variables) in your design.

The values contained within a variable are often determined by the researcher. This is a critical decision in the research design phase, and influences the possibilities for statistical analysis as these values define the level of measurement for that variable. A variable can have continuous or discrete values. Variables are discrete when they have a finite or countable number of possible values. For example, gender, ethnic background, and student status (i.e., full-time / part-time) are discrete. Continuous variables have infinite range and can be measured to varying degrees of precision. Common examples of
continuous data are dollars, square footage, height, weight, and age. A continuous variable may be measured as if it were discrete; however, the reverse is not true. For example, salary data can be broken down into discrete categories (Example Box 2). In some instances, dividing a discrete variable by a discrete variable creates a continuous scale; for example, admissions yield ratios (number enrolled divided by number applied).

Example Box 2

<table>
<thead>
<tr>
<th>Salary as a Discrete Variable</th>
<th>Salary as a Continuous Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 - $25,000</td>
<td>Actual Salary in dollars</td>
</tr>
<tr>
<td>$25,001 - $50,000</td>
<td>$42,014</td>
</tr>
<tr>
<td>Over $50,000</td>
<td></td>
</tr>
</tbody>
</table>

Levels of measurement can be further broken down into a hierarchy with four categories: nominal, ordinal, interval, and ratio. Variables which are Nominal level of measurement consist of names, labels and categories; this is the lowest level of the hierarchy. In classifying data, subjects or observations are identified according to a common characteristic. When dealing with a nominal variable, every case or subject must be placed into one and only one category. This requirement indicates that the categories must be non-overlapping or mutually exclusive. Thus, any respondent labeled as male cannot also be labeled as female. Also, this requirement indicates that categories must be exhaustive; that is, a place must exist for every case that arises. Nominal data are not graded, ranked or scaled in any manner. Clearly then, a nominal measure of gender does not signify whether males are superior or inferior to females. Numerical codes are often assigned to the values of nominal variables, adding to the confusion. For example, even though the value 1 is assigned for female and 2 for male, these are simply labels and no quantity or quality can be implied. No mathematical calculations can be applied to numbers that only serve as labels. Thus, limits are placed on what can and cannot be done statistically with these data. The most appropriate statistical measures for nominal data include: frequencies, proportions, probabilities and odds.

When the researcher goes beyond mere classification and seeks to assign order to cases in terms of the degree to which the subject has any given characteristic, he or she has assigned an ordinal level of measurement. With an ordinal scale, imagine a single continuum along which individuals may be ordered. However, the distances between values on the continuum may not always be meaningful or even known. Rather, the ordinal level of measurement yields information about the ordering of categories, but does not indicate the magnitude of differences between the numbers. An ordinal level of measurement supplies more information than is obtained using a nominal scale, since subjects are able to be grouped into separate categories, which can then be ordered. The order of the categories can be described by adjectives like more and less, bigger and smaller, stronger and weaker, etc. A familiar example of ordinal level of measurement is the classification of faculty as assistant, associate or full professors. Although we know a full professor is a higher status than an associate or assistant, it cannot be said that two associates equal one full professor.

Additionally, most Likert scales are considered to be ordinal level of measurement. On student surveys, Likert scales are often used to measure satisfaction with services or the extent of agreement with various statements. For example, students
may be asked to respond to the question, “Overall, how satisfied are you with the social life on campus?” on a 5-point scale where 1 equals ‘very dissatisfied’ and 5 equals ‘very satisfied.’ Clearly, if respondent A marks a 5 and respondent B marks a 4, then respondent A is more satisfied than respondent B. However, the magnitude of the difference in their levels of satisfaction is not directly distinguishable.

By contrast, *interval* level of measurement not only classifies according to the ordering of categories, but also indicates the exact distance between levels. The interval scale requires the establishment of some common standard unit of measurement that is accepted and replicable. Common examples of interval level of measurement are SAT scores and temperature in Fahrenheit or Celsius. Given a standard unit of measurement, it is possible to state that the difference between two subjects is a particular number of units. However with interval data, it is not possible to make direct ratio comparisons between levels of the data. With interval data such comparisons are not possible because there is no meaningful zero point (i.e., zero does not imply the absence of the quantity being measured). For example, 0 Celsius does not imply no temperature; rather the value represents the freezing point of water. Also, SAT scores are normally considered interval because the base is not equal to zero or “no ability.” Thus, a score of 600 is not twice as high as a score of 300.

If it is possible to locate an absolute and non-arbitrary zero point on the scale, then the data are *ratio*, the highest level of the measurement hierarchy. In this case, scores can be compared by using ratios. For example, if the endowment of school A is $10 million and the endowment of school B is $25 million, then school B’s endowment can be said to be $2 \frac{1}{2}$ times that of school A. After all it is possible to have a $0$ endowment. While many researchers make the distinction between interval and ratio level of measurement, some do not. Although the distinction between the two is subtle, it is important to recognize the limitation on the types of comparisons of scores one can make between the two levels. On the other hand, fewer statistical techniques require a ratio scale; making the distinction between the interval and ratio levels of measurement somewhat irrelevant.

All statistical analyses require a particular minimal level of measurement. An important general guideline is that statistics based on one level of measurement should not be used for a lower level, but can be used for a higher level. For example, statistics requiring ordinal data may be used on interval or ratio data, but should not be used on nominal data. Figures 1 and 2 summarize the characteristics of each of the four levels that have been discussed and illustrate the relationship of the levels of measurement of the variable to the level of measurement for statistical analysis. It is important to note that when a higher level of measurement of data is analyzed using a statistic based on a lower level of measurement, information may be lost if the data is collapsed into broader more discrete categories. Thus, the decisions made concerning the level of measurement of variables have a direct impact on the type of statistical analyses that can be performed.

To review the levels of measurement consider the following scenario: you are consulting with your admissions office on the design of an Inquiry Survey and you wish to add a question concerning income level. Figure 3 summarizes several adequate ways that you could ask for this information. Each of the four questions will result in one of the four levels of measurement. In deciding which question to use, a general rule of thumb is to measure at the highest level appropriate and possible. You can always collapse or combine scores later to create a lower level of
We hope you enjoyed reading sample pages from this volume.

If you are an AIR member, please visit or return to the eLibrary to read more of this book by clicking on the Sign in to View link. You must be logged in to do this.

Not a member? You can still order the print version of the book, if available, by clicking on the Purchase link. Note that the AIR History Book is free to all eLibrary users.

Join AIR today to access the eLibrary at no additional charge. There are over 15 books available free to AIR members.