

Minitab: Successful Choices in an Introductory Statistics Class

Frank H. Dietrich II, David B. Agard, Northern Kentucky University

Frank H. Dietrich, NKU, Highland Heights, KY 41099

Introduction

Introductory statistics courses offered to serve undergraduate majors in business and the social sciences may be taught in many ways. The method of instruction will be affected by factors such as the topics selected to be covered, the types of students, available facilities, background of instructors, and instructional objectives. It is our belief that the use of a statistical package, such as Minitab, facilitates the learning of statistics in introductory courses. We will discuss how we use Minitab in such courses at Northern Kentucky University (NKU), and why we choose to use Minitab as we do.

Facilities at Northern Kentucky University

Adequate facilities are a necessity to employ the computer successfully in any course. At NKU we are fortunate to have excellent equipment to support our endeavors. Each classroom in which introductory courses are taught has an overhead projector, a projection panel, and a PC that is connected to the main frame computer. Minitab is installed on the main frame, and can be easily accessed in the classroom. With the use of the projection panel and the overhead projector, instructors can actively and effectively demonstrate the uses of Minitab during class. The capability to perform the demonstrations during class is an integral part of using Minitab successfully. It would be very difficult, if not impossible, to succeed without actually being able to use Minitab in class.

A computer lab with approximately ninety connections to the main frame is available for student use. The lab is open morning to night during the week, and limited hours on the weekend. Students easily find times to run their Minitab analyses. This method of operation costs the students nothing, and has proven to work quite well. Students may also access the main frame from modems if they so desire. A final option that is provided, but not encouraged, is for the student to buy the student version of Minitab for use on a PC.

The well equipped computer lab at NKU is important to the successful use of Minitab by the students. Without the ease of access (in both time and money), Minitab would be much less effective in the introductory courses.

The Philosophy at Northern Kentucky University

At NKU the main objective in our introductory statistics courses is to enable students to be able to make and understand statistical inferences. We view all statistical problems, whether descriptive or inferential, to have three basic components. Briefly, the first component is identifying the purpose for which data has been collected and choosing an appropriate analysis. The second component is performing the appropriate analysis of the data for the chosen purpose. The third component is interpreting the results of the analysis as related to the problem of interest. By necessity, prior to the wide-spread availability of computers (and further back, hand-held calculators), much of what was discussed in an introductory statistics course centered on the second component summarizing data, performing calculations, i.e. doing analyses. It is undoubtedly fair to say that students in those courses thought of statistics as calculational methods. With the aid of statistical software, much less class time need be spent on "calculations", and the understanding of the overall statistical process can be emphasized.

In our introductory courses, traditional statistical methods are covered, and the rationale behind these methods is stressed. Included are explanations (often demonstrations) of why the methods work, as well as discussions of the possible errors that could occur when making statistical inferences. Particular emphasis is placed on the interpretation of results. Over the course of semester, the students are required to write (in complete sentences!) their interpretation of statistical analyses - beginning with descriptive statistics, and culminating with confidence intervals and tests of hypotheses.

How The Instructors Use Minitab

Our goal at NKU has been to use the computer in such a way as to enhance the students' ability to make statistical inferences. We believe we have found ways for both the instructors (in class) and students (away from class) to use Minitab successfully. In class, instructors use Minitab to illustrate how the students are expected to use Minitab, but more importantly, the instructors use Minitab to empirically verify statistical theory. Due to the general lack of the students' mathematical sophistication, it is inappropriate to prove theoretical results in an introductory course. Much of statistical theory, however, is based on distributional properties of sample statistics that may be obtained by simulation. For example, Minitab can easily be used to generate the distribution of a sample mean, or repeatedly form confidence intervals. Through simulation, instructors are able to provide visual demonstrations that solidify otherwise abstract statistical concepts. Thus, we rely heavily upon Minitab to conduct "proofs by example". Appendix C contains examples of simulations that the instructor might perform in class to demonstrate the meaning of confidence intervals and Type I errors. A complete list of the types of simulations we perform in class is given in Appendix A. We believe that the time it takes to perform these classroom demonstrations is well spent. The students in introductory courses, many with limited mathematical preparation, obtain a remarkably clear understanding of rather difficult statistical concepts.

How The Students Use Minitab

The primary reason that students use Minitab is to perform calculations. In particular, when the actual sample data (the raw data) is to be analyzed, Minitab is employed. When calculations are done from summary statistics, hand-held calculators are the instrument of choice. Our students use Minitab to do many analyses such as constructing histograms, obtaining descriptive statistics, forming confidence intervals, calculating test statistics or p-values, performing regression analyses, and conducting an analysis of variance. A more complete list of student uses of Minitab is given in Appendix B.

We do not expose students to every possible nuance that Minitab has to offer. The guiding principle we have employed when choosing how students should use Minitab is that Minitab should make the students' calculations easier to obtain. Due to the method of presentation, it is very easy for the students to understand how to use Minitab. Students do not have to deal with computer files or complicated system commands. All data sets that are to be analyzed have been stored in a library. Students read the data file of interest directly into Minitab, and thus do not have to perform the laborious task of actually entering data. We spend very little time discussing calculations, and thus have more time to discuss how to choose the appropriate analysis and what the results of the analysis mean. We believe this is the proper emphasis for an introductory statistics course.

Summary

If Minitab, or another easily used statistical software package, is properly employed in an introductory statistics course, it facilitates the learning of statistics. Students do not get bogged down with calculations and can concentrate on understanding why statistical procedures work and what analyses mean. Instructors use Minitab to conduct classroom simulations that reinforce "statistical theory". In general, students like this approach, enjoy the course, and better understand statistical concepts.

However, care must be taken in choosing how to use the computer. There is a fine line between making it easier to obtain statistical calculations, and becoming so enamored of the computer's ability that a statistics course becomes a "computer course". The material that is to be taught, and the course objectives, should dictate how the computer is used. The computer should be employed in such a way as to help students learn more easily. Without careful planning, what the computer is capable of doing may unduly influence what is done in the course. Such a misuse of the computer would detract from the enjoying and learning of introductory statistics.

Acknowledgment

The authors greatly appreciate the comments of the peer reviewer of this paper, Dr. James E. McKenney, Northern Kentucky University.

Appendix A List of Simulations

- Random Sampling
- Binomial Experiment

- The Central Limit Theorem
- Tests of Hypothesis (a and b)
- Confidence Intervals
- Analysis of Variance

Appendix B List of Student Analyses

- Histograms
- Dotplots
- Descriptive Statistics
- Mean, Variance, and Plot of Discrete Distributions
- Binomial Distributions
- t-Test (one sample)
- t-Interval (one sample)
- t-Test (two samples)
- t-Interval (two samples)
- Regression and Correlation
- Analysis of Variance
- Chi-Square Test
- Sign Test
- Wilcoxon Signed Rank Test
- Mann-Whitney Test
- Kruskal-Wallis Test

Appendix C

MTB > #

MTB > # A DEMONSTRATION OF THE MEANING OF A CONFIDENCE INTERVAL

MTB > #

MTB > # For each of 40 random samples of size 16 measurements from a

MTB > # normally distributed population with mean $\mu=10$ and standard

MTB > # deviation $\sigma=2$, 90% confidence intervals will be constructed.

MTB > # MTB > RANDOMly sample $n=16$ obs. in C1-C40;

SUBC> NORMAL population with $\mu=10$ and $\sigma=2$.

MTB > #

MTB > # Question 1: Will each interval contain the true value of $\mu=10$?

MTB > # Question 2: What percentage of the intervals should contain the

MTB > # true value of $\mu=10$?

MTB > #

MTB > # Count how many of the intervals contain the true value of $\mu=10$.

MTB > #

MTB > T-INTERVAL with 90 percent confidence using the data in C1-C4

MTB > # Question 3: What percentage of these intervals did contain the

MTB > # true value of the mean $\mu=10$?

MTB > # Question 4: Does this agree with your answer to question 2?

MTB > #

MTB > STOP

MTB > #

MTB > # A DEMONSTRATION OF THE MEANING OF A TYPE I ERROR

MTB > #

MTB > #

MTB > # For each of 40 random samples of size 25 measurements from a

MTB > # normally distributed population with mean $\mu=60$ and

MTB > # standard deviation $\sigma=5$, we will test:

MTB > #

MTB > # $H_0: \mu = 60$ vs. $H_a: \mu > 60$

MTB > #

MTB > RANDOMly sample n=25 obs. in C1-C40;

SUBC> NORMAL population with mu=60 and sigma=5.

MTB > #

MTB > # Question 1: Should this test lead to rejection of Ho?

MTB > # If Ho is rejected, what has occurred?

MTB > # Question 2: If alpha=.10, will any of the p-values be less than .10?

MTB > # Question 3: What percentage of the p-values should be less than .10?

MTB > #

MTB > # Count how many of the p-values are less than .10.

MTB > #

MTB > T-TEST of mu = 60 using the data in C1-C40;

SUBC> ALTERNATIVE = +1.

TEST OF MU = 60.000 VS MU G.T. 60.000

MTB > # Question 4: What percentage of these tests ended in a false

MTB > # Rejection of Ho, i.e. a Type I error?

MTB > # Question 5: Does this agree with your answer to question 3?

MTB > #

MTB > STOP