ABSTRACT

Using Microcomputers in the Teaching
of Probability and Statistics

Elliot A. Tanis
Department of Mathematics
Hope College
Holland, Michigan 49423

Microcomputers can be used in various ways with different types of students. Four possibilities are the following:

1. A weekly laboratory could be held in conjunction with either an introductory or a calculus based statistics course. In this laboratory certain theoretical results could be illustrated using simulation. Also, given a set of data, the computer could be used for descriptive statistics and for exploratory data analysis.

2. Demonstrations of, e.g., the Central Limit Theorem, could be developed for presentation in the classroom.

3. Some of the better undergraduate mathematics majors could become involved in developing the needed software. This could lead into an undergraduate research project.

4. For those parts of a statistics course that involve a lot of calculation, the use of microcomputers by all of the students would release the time needed for calculation to be used in the interpretation of the results. For example, those students who will later become involved in research projects in other disciplines in which ANOVA is needed will probably use a statistical package on a computer to analyse their data. It would be helpful for them to learn how to interpret the results that they will encounter on the computer printout.

We have written more than 200 exercises and 50 subroutines for a Radio Shack TRS-80 Model III computer. In this talk we will illustrate how we have used those materials in the above ways. Although the programs have been written for the TRS-80, the ideas that we present can certainly be used on other computers.
USING MICROCOMPUTERS TO ILLUSTRATE CONCEPTS IN PROBABILITY AND STATISTICS

Elliot A. Tanis, Hope College

1. Introduction

Hope College is a four-year, undergraduate, liberal arts college of 2400 students. A mathematics course is included in the core curriculum as one of the requirements for graduation. Because of the importance of statistics in several disciplines, many students elect statistics to satisfy the mathematics requirement. Approximately 250 students take Introductory Statistics each year, one third of whom have never used a computer. In addition, approximately 60 students will take either a one or two semester statistics course for which calculus is a prerequisite. Most of those students have had some experience writing computer programs.

The materials described in this paper can be used in various ways with these different types of students. Originally the materials were developed for a computer-based laboratory for students in mathematical statistics and probability. Those computer programs are written in FORTRAN and can be run in either batch or interactive mode on most large computer systems [1].

When writing exercises, developing software, and incorporating these materials into the curriculum, our aim has been to develop materials that would help students understand more clearly the basic concepts in probability and statistics and the interpretation of statistical results. Two important considerations were:

1. The development of a software package that would be easy to use and would make effective use of the graphics capabilities of the computer.
2. The writing of instructive exercises and examples that would have sound educational value.

2. Computer programs that are available

More than 50 subroutines have been written in BASIC for the Radio Shack Model III microcomputer. Some of these are complete and may be loaded and run. Many of the subroutines are designed to be MERGE'd into the user's program.

The structure of the subroutine package according to lines used and examples of some of the subroutines are as follows:

<table>
<thead>
<tr>
<th>Lines</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 - 99</td>
<td>REM statements for the user to describe the problem that is being solved</td>
</tr>
<tr>
<td>100 - 1499</td>
<td>Solution program, written by the user</td>
</tr>
<tr>
<td>1500 - 1749</td>
<td>Calculation of sample characteristics, such as sample mean, correlation coefficient</td>
</tr>
<tr>
<td>2000 - 2199</td>
<td>Probability density function, used by the graphing subroutines and written by the user</td>
</tr>
<tr>
<td>2200 - 2399</td>
<td>Distribution function, used by the graphing subroutines, and written by the user</td>
</tr>
<tr>
<td>3000 - 4999</td>
<td>Data analysis programs, such as contingency tables, ANOVA, Wilcoxon test statistic, exploratory data analysis (stem-and-leaf and box-and-whisker displays)</td>
</tr>
<tr>
<td>5000 - 11999</td>
<td>Graphing routines - histograms, scatter plots, confidence intervals</td>
</tr>
<tr>
<td>12000 - 12999</td>
<td>Permutations and combinations</td>
</tr>
<tr>
<td>13000 - 16499</td>
<td>Ordering routine</td>
</tr>
<tr>
<td>16500 - 18499</td>
<td>Distribution functions for binomial, chi-square, F, gamma, normal, Poisson, and t</td>
</tr>
<tr>
<td>19000 - 19999</td>
<td>Inverses of the distribution functions for chi-square, F, t, and standard normal</td>
</tr>
<tr>
<td>20000 - 26999</td>
<td>Binomial and Poisson p.d.f.'s</td>
</tr>
<tr>
<td>27000 -</td>
<td>Simulations of random samples from 14 distributions, e.g., normal, Poisson, geometric</td>
</tr>
</tbody>
</table>

3. Exercises that are available

More than 200 exercises of varying degrees of difficulty have been written to accompany a year long course in mathematical statistics and probability. The manual containing the exercises parallels Probability and Statistical Inference, Second Edition by Robert V. Hogg and Elliot A. Tanis [2]. These exercises can also be used successfully with other textbooks.

4. Ways in which these materials can be used

There are several ways in which these materials can be used. Four suggestions follow.

1. There are a sufficient number of exercises for a weekly laboratory that is held in conjunction with a year long course in probability and statistics. Some of the exercises ask the student to use simulation to illustrate empirical results that have been proved or discussed theoretically in class. Other exercises ask the student to simulate a physical experiment. In order for a student to write a computer program to properly perform a simulation, the student gains a better understanding of and appreciation for the underlying theory. The students write a solution program in BASIC and then MERGE the appropriate subroutine(s) into their program. Some of these exercises can be solved during the laboratory period and run immediately. Other exercises will require additional time outside of the laboratory for writing a solution program.

2. An instructor and/or student could solve an exercise and present its solution for a class or seminar. Such demonstrations could be used for either an introductory statistics class or a probability and statistics class for which calculus is a prerequisite.

3. Some of these materials could be used as a basis for or in conjunction with an undergraduate research or independent study project. For example, the exercises illustrating the Central Limit Theorem grew.
out of a student project that was conducted 15 years ago. The subroutines for stem-and-leaf and box-and-whisker displays were written by a student during this past year as a part of an independent study project.

4. Some of the interactive programs can be used easily for statistical analysis. Examples of such programs are CONVEX for contingency table problems and ANOVA for analysis of variance programs. These can provide a first computer experience for many students. They especially appreciate the work done for them by the computer after they have solved, e.g. an analysis of variance problem using a calculator.

5. Examples

Sample exercises and complete solutions were incorporated into the presentation of this paper. You may write to the author for a sample copy of output that illustrates the type of exercises, the length of some solution programs, the way in which the subroutines are used, and the graphical output provided by some of the subroutines.

6. Conclusions and recommendations

Students in the laboratory have reacted positively to their experience. Those who had had no prior experience with microcomputers and/or with writing programs in BASIC were frustrated at times. Thus it is important to provide extra programming instruction for such students.

Demonstrations in the classroom can be effective if there are a sufficient number of monitors for all of the students to be able to see the output. The concept that is being demonstrated must be explained clearly before the demonstration. Good timing of demonstrations during a lecture can help to keep the attention of the students. Students and/or professors who write the solution program for a demonstration will sometimes benefit the most from the demonstration.

During this past year a student developed computer programs for exploratory data analysis, writing programs for stem-and-leaf displays, box-and-whisker displays, and smoothing of paired data. He presented the results of his work at the Annual Meeting of the Michigan Section of the Mathematical Association of American and at this year's annual Pi Mu Epsilon meeting.

Several students in introductory statistics expressed some concern as to whether they would be able to understand how to use a computer. Often those who were the most fearful had the most praise for the ease with which they could use the computer. Students especially appreciated the time that was saved in connection with contingency table and analysis of variance problems. When the computer is used, it is especially important to spend time discussing and interpreting the computer output. For example, the data analysis programs provide the p-value of a test statistic. It is important to discuss the meaning of p-value and the relationship between p-value and significance level.

Because of the popularity of the Apple computer in education, we plan to convert the computer programs so that they can be run on an Apple computer.

7. Distribution of materials

The earlier version of these materials that is written in FORTRAN is being distributed by CONDUIT, M310 Oakdale Hall, The University of Iowa, P.O. Box C, Oakdale, IA 52319. We expect CONDUIT to also distribute the microcomputer version of these materials.

REFERENCES