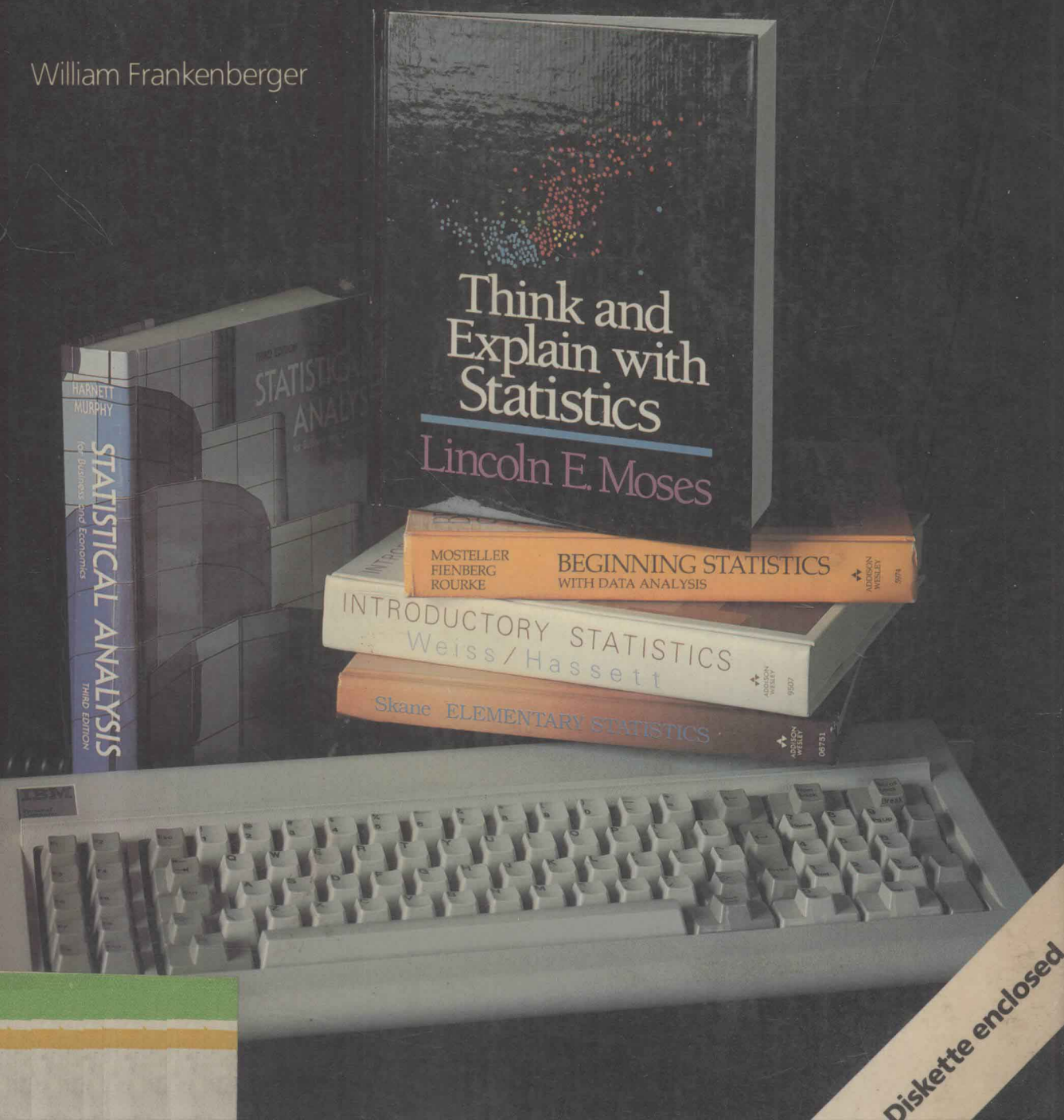


Introductory Statistics

SOFTWARE PACKAGE

For the **IBM**® PC

William Frankenberger



Diskette enclosed

Introductory Statistics

SOFTWARE PACKAGE

For the **IBM®** PC

Manual/Workbook

William Frankenger

University of Wisconsin, Eau Claire



Addison-Wesley Publishing Company

Reading, Massachusetts • Menlo Park, California
Don Mills, Ontario • Wokingham, England • Amsterdam • Sydney
Singapore • Tokyo • Madrid • Bogotá • Santiago • San Juan

***IBM is a registered trademark of International Business Machines Corporation.**

Addison-Wesley Publishing Company, Inc., makes no representations or warranties, express or implied, with respect to this software and its documentation, including without limitations, any implied warranties of merchantability or fitness for a particular purpose, all of which are expressly disclaimed. The exclusion of implied warranties is not permitted by some states. The above exclusion may not apply to you. This warranty provides you with specific legal rights. There may be other rights that you have which may vary from state to state.

This documentation was reproduced by Addison-Wesley from camera-ready copy supplied by the author.

Copyright © 1986 by Addison-Wesley Publishing Company, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, photocopying, recording or otherwise, without the prior written permission of the publisher. Printed in the United States of America.

ISBN 0-201-06256-9

ABCDEFGHIJ-AL-89876

Introductory Statistics

SOFTWARE PACKAGE

For the **IBM®** PC

Manual/Workbook

IBM® PC

Equipment Needs

To run the *Introductory Statistics Software Package*, you need the following equipment:

- IBM PC
- 128K memory
- color graphics card
- one double-sided disk drive
- IBM DOS version 2.0 or higher
- printer (optional)

If microcomputers are to be employed in any capacity other than as a substitute for calculators, there must be tangible advantages for both instructors and students. We believe that computers should be used as pedagogical tools which help demonstrate concepts but do not infringe upon or reduce the importance of quality instruction. Computers should be employed to build confidence in students who may be threatened by statistics in general and computers in particular. Therefore, students and faculty with no computer knowledge must be able to immediately experience success and recognize the educational benefits of computers applied to statistics. To ensure a positive response by users, it is essential that the computer only be used where it can be of most benefit.

The student should read chapters 1 (description of the programs) and 2 (description of the IBM PC computer) prior to using the programs. Knowledge of information in these chapters will help insure that the student's first experience with the programs and computer will be positive.

The programs contained in the Introductory Statistics Package are sequenced in a manner consistent with that of several current introductory statistics textbooks. The programs were designed to be successfully utilized by individuals with no previous computer experience. The programs are self-explanatory to the point that most students will be able to use them without documentation.

The workbook has been designed to thoroughly explain the capabilities and limitations of each program. In addition, the workbook functions to familiarize the student with computers in general and the Introductory Statistics Package Programs in particular. Problem sets have been included at the end of chapters 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, and 14 to allow students to investigate various statistical procedures and concepts.

Workbook problems have also been included which require the use of statistical analyses not directly included in the programs. These problems are related to topics presented in the package and contained in most introductory statistics textbooks. For example, the last problem in Chapter 5 requires that the student use the computer to group 50 scores into a grouped frequency distribution. Then the student uses the computer generated grouped frequency distribution to hand calculate the percentile rank of a score and the score at a given percentile rank in the distribution.

Students may use tables included at the end for hypothesis testing. These tables may be employed to test the significance of the following statistics: 1) t , 2) r , 3) F , and 4) Chi Square.

ACKNOWLEDGEMENTS

I wish to express my appreciation to Brian Bezanson for his programming assistance with the IBM version of Introductory Statistics Package. I would also like to thank my colleagues at the University of Wisconsin-Eau Claire. Barbara Lozar and David Proctor reviewed and field tested the package and Kenneth Heilman and Thomas Blakemore helped with the initial version of the manual/workbook and programs.

I would also like to thank Sandra Parker for her assistance with typing and editing the manual/workbook.

The computer package was made possible by the excellent staff at Addison-Wesley Publishing Company, Inc. I would especially like to thank Jeff Pepper and his staff for helping an idea become a reality.

I am grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S. and to Longman Group Ltd., London, for permission to reprint Tables III and VI from their book Statistical Tables for Biological, Agricultural and Medical Research (6th edition, 1974).

Finally, I would like to thank my wife, Marge, for being understanding and supportive during my long and sometimes frustrating hours of programming.

CONTENTS

1	OVERVIEW OF THE MANUAL	
1.1	The Computer Software	1
1.2	Using the Computer	1
1.3	Characteristics of the Programs	1
1.4	Grouped Frequency Distribution	1
1.5	Histogram and Frequency Curves	2
1.6	Mean and Standard Deviation	2
1.7	Correlation	2
1.8	Regression	2
1.9	Sampling Distribution	2
1.10	Independent T-Test	3
1.11	Correlated T-Test	3
1.12	One-Way Analysis of Variance	3
1.13	Two-Way Analysis of Variance	3
1.14	Chi Square	3
2	USING THE IBM PC	
2.1	Getting Started	4
2.2	Using the Menu	5
2.3	Editing Your Input	5
2.4	Using the Computer as a Calculator	6
2.5	Combining Arithmetic Operations	7
3	SUMMATION	
3.1	Developing Your Own Summation Program	9
3.2	Program Output	10
3.3	Exiting the Program	10
	Exercises	11
4	GROUPED FREQUENCY DISTRIBUTION	
4.1	Program Description	12
4.2	Program Limitations	12
4.3	Printout	12
4.4	Recalling Scores	12
4.5	Entering Scores	13
4.6	Adding, Deleting, and Editing Scores	14
4.7	Saving Scores	15
4.8	Program Output	15
4.9	Exiting the Program	18
	Exercises	19-23

5	HISTOGRAM AND FREQUENCY CURVES	
5.1	Program Description	24
5.2	Program Limitations	24
5.3	Printout	24
5.4	Recalling Scores	24
5.5	Entering Scores	25
5.6	Adding, Deleting, and Editing Scores	26
5.7	Saving Scores	27
5.8	Program Output	27
5.9	Exiting the Program	30
6	THE MEAN AND STANDARD DEVIATION	
6.1	Program Description	32
6.2	Program Limitations	32
6.3	Printout	32
6.4	Recalling Scores	32
6.5	Entering Scores	33
6.6	Adding, Deleting, and Editing Scores	34
6.7	Saving Scores	35
6.8	Program Output	35
6.9	Exiting the Program	37
	Exercises	38-43
7	CORRELATION	
7.1	Program Description	44
7.2	Program Limitations	44
7.3	Printout	44
7.4	Recalling Scores	44
7.5	Entering Scores	45
7.6	Editing Scores	46
7.7	Saving Scores	47
7.8	Program Output	47
7.9	Exiting the Program	49
	Exercises	50-54
8	REGRESSION	
8.1	Program Description	55
8.2	Program Limitations	55
8.3	Printout	55
8.4	Recalling Scores	55
8.5	Entering Scores	56
8.6	Editing Scores	57
8.7	Saving Scores	58
8.8	Program Output	58
8.9	Exiting the Program	61
	Exercises	62-68

9	THE SAMPLING DISTRIBUTION OF THE MEAN	
9.1	Program Description	69
9.2	Program Limitations	69
9.3	Printout	69
9.4	Entering Data	69
9.5	Program Output	70
9.6	Exiting the Program	73
	Exercises	74-77
10	THE INDEPENDENT T-TEST	
10.1	Program Description	78
10.2	Program Limitations	78
10.3	Printout	78
10.4	Recalling Scores	78
10.5	Entering Scores	79
10.6	Editing Scores	80
10.7	Saving Scores	80
10.8	Program Output	81
10.9	Exiting the Program	83
	Exercises	84-88
11	THE CORRELATED T-TEST	
11.1	Program Description	89
11.2	Program Limitations	89
11.3	Printout	89
11.4	Recalling Scores	89
11.5	Entering Scores	90
11.6	Editing Scores	91
11.7	Saving Scores	91
11.8	Program Output	92
11.9	Exiting the Program	95
	Exercises	96-100
12	ONE-WAY ANALYSIS OF VARIANCE	
12.1	Program Description	101
12.2	Program Limitations	101
12.3	Printout	101
12.4	Recalling Scores	101
12.5	Entering Scores	102
12.6	Saving Scores	103
12.7	Program Output	104
12.8	Exiting the Program	107
	Exercises	108-112

13	TWO-WAY ANALYSIS OF VARIANCE	
13.1	Program Description	113
13.2	Program Limitations	113
13.3	Printout	113
13.4	Recalling Scores	113
13.5	Entering Scores	114
13.6	Editing Scores	115
13.7	Saving Scores	116
13.8	Program Output	116
13.9	Exiting the Program	118
	Exercises	119-121
14	THE CHI SQUARE TEST OF INDEPENDENCE OF CATEGORICAL DATA	
14.1	Program Description	122
14.2	Program Limitations	122
14.3	Printout	122
14.4	Entering Scores	122
14.5	Editing Scores	123
14.6	Program Output	124
14.7	Exiting the Program	125
	Exercises	126-130
	TABLES	131

This workbook/manual has been developed to provide complete documentation for programs included in Introductory Statistics. It describes the characteristics of each program, including input and editing procedures, sample output and formulae employed in calculations. In addition, the workbook/manual contains problems and exercises designed to demonstrate statistical concepts and applications.

1.1 The Computer Software

This statistical software has been developed to complement material contained in most introductory statistics courses. Programs are arranged in the same sequence and use statistical notation which is consistent with many introductory statistics textbooks.

1.2 Using The Computer

The next chapter will provide you with an explanation of how to operate the computer for use with this program. The chapter includes a description of how to load the program and general information about its operation.

1.3 Characteristics Of The Programs

All programs have been designed to be user friendly and easily operated by students not familiar with computers. They have been tested on students enrolled in introductory statistics classes. There are no commands to remember; all instructions necessary for operation are displayed on the screen. The programs provide students with options which require a single keystroke to execute. For example, students may obtain a printout of their data and/or recall previously stored data from disk. Each time a student enters data from the keyboard, he/she has the option of storing the data on the disk. This allows students to recall the data for the same program or use the data in other programs. The programs allow easy editing. After scores have been entered, the students are able to add new scores, edit existing scores or delete unwanted scores. Each of the programs is briefly described in the paragraphs below.

1.4 Grouped Frequency Distribution

This program allows students to enter data from the keyboard or recall previously stored data. The data are ordered from highest to lowest. The student has the option of selecting an interval size or allowing the computer to calculate the interval size. The steps necessary to combine scores into a grouped frequency distribution are displayed as the computer performs the calculations. The scores are then combined into a grouped frequency distribution with interval

limits and frequencies displayed on the screen. The next display includes the above values plus the cumulative frequency and cumulative percentile distributions. The final output graphs the data in the form of a histogram.

1.5 Histogram And Frequency Curves

This program utilizes the graphics capabilities of the IBM to combine scores into a grouped frequency distribution and display the grouped scores in the form of a histogram, a frequency curve, and a cumulative percentage curve.

1.6 Mean And Standard Deviation

This program calculates the mean and standard deviation and other scores which must be calculated when hand computation is necessary. The program prints a table containing scores and squared scores along with their sums. It also displays the number of scores, the sum of squared deviations from the mean and the variance.

1.7 Correlation

Students enter paired scores from the keyboard and receive an output table including the following: X , X squared, Y , Y squared, and XY . Sums of each of the above columns are also displayed. The output also includes means and standard deviations for both groups as well as the correlation coefficient. Students can view a scatter diagram by recalling their data for the regression program.

1.8 Regression

Students enter paired scores from the keyboard and receive the following output: the correlation coefficient, the formula for the regression line of Y on X , the standard error of the estimate, predicted Y scores for each value of X , and the coefficients of determination and nondetermination. The program also displays a scatter diagram with the regression line of Y on X .

1.9 Sampling Distribution

This program allows students to investigate the properties of sampling distributions. Students enter the mean and standard deviation of a population along with the number of samples they wish to take and the sample size for each sample. Samples are randomly selected; the means for the samples are calculated and presented in the form of a histogram. The mean and standard deviation of the sampling distribution of means are also computed.

1.10 Independent T-Test

The program calculates the t ratio for independent groups. It also includes a listing of the following values necessary for computing t: (1) N for each group, (2) sums of scores, (3) sums of squared scores, (4) sums of squared deviations, (5) means, (6) standard deviations, (7) difference between means, (8) standard error of the difference between means, (9) degrees of freedom, (10) estimate omega squared.

1.11 Correlated T-Test

This program calculates the correlated t-ratio using the direct difference method. The output includes the following values: (1) means, (2) sums of squared deviations, (3) standard deviations, (4) sum of differences, (5) sum of squared differences, (6) the mean of the differences, (7) the standard error of the mean difference, (8) the degrees of freedom, (9) the t-ratio, (10) the correlation coefficient.

1.12 One-Way Analysis Of Variance

This program displays the means and standard deviations for each group along with an ANOVA table and estimate of omega squared.

The means and their 95 percent confidence interval are then displayed in a graphic format.

1.13 Two-Way Analysis Of Variance

This program allows students to enter data from the keyboard and perform a two-way analysis of variance. The output consists of a two-way ANOVA table and the means and standard deviations for each group. The means are then displayed in a graphic format.

1.14 Chi Square

Students enter observed frequencies for each cell of a contingency table and receive output of the contingency table with row and column totals and expected frequencies beneath the observed frequencies. The chi square value and degrees of freedom are also displayed.

2.1 Getting Started

Before you begin to use the Introductory Statistics Package, you should first copy IBM DOS onto your ISSP diskette. Copyright restrictions prevent us from distributing DOS with the ISSP programs. You therefore need to provide a DOS system disk that contains the following:

1. DOS version 2.0 or higher, but less than 3.0
2. the program GRAPHICS.COM

Assuming the computer is not turned on, proceed as follows:

1. Place a write-protect tab on your DOS diskette, if there isn't one already. Then insert your DOS diskette into the disk drive (use Drive A if the computer has two drives). The label on the disk should be facing up and the oval cutout pointing inward. Make sure that the disk is completely inside before closing the hatch.
2. If you have a two-drive system, insert your ISSP program diskette into Drive B.
3. Turn on the monitor and the computer. The red light will glow and the disk drive will be activated.
4. Wait several seconds for the monitor to prompt you with a message requesting the current date. You can respond simply by pressing the RETURN key. When a second message asks you for the time, press RETURN again. A copyright notice will appear on the screen, followed by the symbol A>.
5. Immediately after the A> prompt, type:

b:install

and press RETURN. [If you make a typing error, use the backspace key (on the top row at the right side of the keyboard with an arrow pointing left) to move the cursor over your mistake. Then retype the necessary characters and press RETURN.]

6. Follow the directions on the screen for copying the programs.

2.2 Using the Menu

So far the program has run without much help from you, but now you must become more active. A "menu," which is a list of options that you can choose from, will appear on the screen. The menu lists 11 programs which may be run by entering the number of the program from the keyboard and pressing the Return key. For instance, you would enter "1" and then press "Return" to start the "GROUPED FREQUENCY DISTRIBUTION" program.

MENU

1. Grouped Frequency Distribution
2. Histogram and Frequency Curves
3. Mean and Standard Deviation
4. Correlation
5. Regression and Scatter Diagram
6. Sampling Distribution
7. Independent T-Test
8. Correlated T-Test
9. One-Way Analysis of Variance
10. Two-Way Analysis of Variance
11. CHI Square
12. Exit Program

Select a number and press RETURN:

Copyright (C) 1986 by

Addison-Wesley Publishing Company, Inc.

2.3 Editing Your Input

If you make a mistake when entering a number (or a 'Y' or 'N' when responding to a question), you can make a correction quite easily so long as you have not yet pressed the RETURN key (a large key on the right side of the keyboard labeled with a bent arrow). To make the correction, you press the "Backspace" key (on the top row at the right side of the keyboard with an arrow pointing left) to move the cursor over your mistake. Then type in the correct letter or numbers and press the Return key.

2.4 Using the Computer as a Calculator

It is sometimes convenient to use the computer as a calculator. To do this on the IBM-PC, you must start (boot) the computer with the PC-DOS disk in drive A. After entering date and time or simply pressing Return twice, you see the following prompt: A>. Now type the following command: **BASICA**.

After BASICA is loaded into memory, the "A>" prompt is replaced by "OK." Now the computer will perform basic mathematical calculations such as addition, subtraction, multiplication, division, and exponentiation in what is known as the "immediate" mode. This means that the computer will perform the operations that you desire without having to be programmed in BASIC or some other language. You simply ask the computer to perform the addition, subtraction, or whatever by typing in a question in a specific format and the computer responds with the answer. You can perform single operations (e.g., addition) or multiple operations with one statement. All of the arithmetic operations that are described below begin by typing a question mark (as if to say "What is....?") and end by pressing the RETURN key.

Addition:

To add on the IBM, use the + (plus) symbol. For example, to add 7 and 2 you would type the following:

? 7 + 2 (and press RETURN)

The computer responds with: 9

By the way, the computer ignores any blanks between characters, so the following would work just as well as the above example:

?7+2(RETURN)

Subtraction:

To subtract on the IBM, use the - (minus) symbol. For example, to subtract 2 from 7 type the following:

? 7 - 2 (and press RETURN)

The computer responds with: 5