United States
Environmental Protection
Agency

Environmental Monitoring and Support
Laboratory
Cincinnati OH 45268

Research and Development

EPA/600/S4-86/023 July 1986

EPA

Project Summary

User-Friendly IBM PC Computer Programs for Solving Sampling and Statistical Problems

Philip C. L. Lin

User-friendly IBM personal computer programs for solving sampling and related statistical problems have been prepared. The programs are designed so that persons without an in-depth understanding of statistics can easily use them. Specific, detailed, written instructions for application of the programs are provided in the full report. The computer disc containing the programs will be made available on request to the Environmental Monitoring and Support Laboratory—Cincinnati (EMSL-Cincinnati).

This Project Summary was developed by EPA's Environmental Monitoring and Support Laboratory, Cincinnati, OH, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

Statistical techniques are useful in assessing the quality of a sampling program. Frequently, field persons engaged in sample collection do not have the time to thoroughly study and understand all the statistics required to take a representative sample. The computer programs described herein were developed for those people and are designed so that an individual may use them and obtain the benefits of the statistical package without an in-depth understanding of the statistics employed.

For those who wish to know more details about the package, basic definitions of statistics and descriptions of statistical sampling programs on the disc are presented in the Appendices of the full report.

Typical Examples for Use of Programs

In order to assist the user in working the computer programs, a series of questions and answers has been developed. Questions that those designing field sampling programs may wish to have answered are listed below, together with the names of the computer programs designed to answer the questions:

Question—How many samples must be taken to reduce the anticipated error to some reasonably fixed value?
Answer—Use program No. 8, "Determination of Sample Number," if the reduction of the anticipated error is based on the accuracy of the sample variance. Use program No. 9, "Determination of Sample Number," if the reduction of the anticipated error is based on the accuracy of the mean.

Question—What is the probability of an effluent exceeding a standard?
Answer—Use program No. 10, "Probability of Exceeding the Standard."

Question—How does one test whether a sample belongs in a particular distribution?
Answer—Use program No. 11, "Hypothesis Testing."

Question—What is the sampling frequency required to capture a significant event in a long-term monitoring program?
Answer—Use program No. 12, "Power Spectrum Analysis."

Question—How does one determine the sample mean, standard deviation, and confidence intervals for the mean and variance?
Answer—Use program No. 7, "Sample Mean, Standard Deviation, and Confidence Intervals for Population Mean and Variance."

Question—Which program should one use to correlate observed data in a linear manner?
Answer—Use program No. 1, "Linear Regression" to determine the linear relationship and its correlation coefficient.

Question—A material is treated by two different processes. Would there be any justification for saying that a difference existed between the two processes? Which program should one use to answer this question?
Answer—Use program No. 13, "Comparing Two Means."

Question—New equipment is used to measure a compound, and it is expected that the measurement uniformity would improve. The question to ask is whether the improvement (more uniformity) really exists or has occurred by chance. Which program should one use to test for the significant difference between variances of two samples?
Answer—Use program No. 16, "Test for Significant Difference between Variabilities of Two Samples."

Instructions for Using Sampling Programs on the IBM PC

Some individuals, especially those with extensive computer experience, will be quickly at ease with these programs. In those cases, the instructions may be bypassed, and the reader may begin to run the programs immediately. For those who need additional assistance, the following instructions are provided:

- Place the program disc in Disc Drive A and close the door.
- Turn on the power of each instrument beginning with the printer, then the monitor, and, finally, the computer. After a brief warm-up, you will see the program menu:

```
**********************************
* PROGRAM MENU ..... PAGE 1 *
**********************************
```

1. Linear Regression
2. Calculation of Normal Deviate Z
3. Calculation of the Percentage Area of Normal Distribution (From Minus Infinity to Normal Deviate Z)
4. Calculation of Student T
5. Calculation of the Percentage Area of Student T Distribution (From Minus Infinity to Student T)
6. Calculation of Chi Square
7. Calculation of Sample Mean, Standard Deviation, and Confidence Intervals for the Population Mean and Variance
8. Calculation of Sample Number Based on the Accuracy of the Variance
9. Calculation of Sample Number Based on the Accuracy of the Sample Mean
10. Calculation of the Probability of Exceeding a Standard
11. Hypothesis Testing
12. Power Spectrum Analysis
13. Proceed to Next Page
14. Quit

- Type the desired option number and press ENTER?
- Type an option number after the question mark (?) and press ENTER. The desired program will be loaded into the computer. If you select option number 13, then program memo on page 2 will be shown as follows:

```
**********************************
* PROGRAM MENU ..... PAGE 2 *
**********************************
```

15. Comparing Two Means
16. Calculation of the Percentage Area % in F-Distribution
17. Calculation of the F Value in F-Distribution
18. Test for Significant Difference between Variabilities of Two Samples
19. Test for Significant Difference between the Population Variance and the Sample Variance
20. Return to Previous Page
21. Quit

Type the desired option number and press ENTER?

- After you run the desired program, you have several choices:

  (a) go back to program menu,
  (b) do another calculation,
  (c) quit,

  by typing the requested option number and press ENTER.
- If you want to abort program calculation, press CONTROL-BREAK key. If you want to start over again, type "EMSLSTAT" and press ENTER.

The EPA author ***Philip C. L. Lin*** *(also the EPA Project Officer, see below) is with the Environmental Monitoring and Support Laboratory, Cincinnati, OH 45268.*

The complete report, entitled "User-Friendly IBM PC Computer Programs for Solving Sampling and Statistical Problems," (Order No. PB 86-203 783/AS; Cost: $11.95, subject to change) will be available only from:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650

The EPA Project Officer can be contacted at:

Environmental Monitoring and Support Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300

EPA/600/S4-86/023