1973

Computer Programs Supporting the Teaching of Statistics

Chien-Hwa Liu

Utah State University

COMPUTER PROGRAMS SUPPORTING THE TEACHING OF STATISTICS

by

Chien-Hwa Liu

A report submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Plan B

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1973

ACKNOWLEDGEMENTS

I wish to express my deep sincere thanks to Dr. Rex L. Hurst, Chairman of my Graduate Committee and Head of the Department of Applied Statistics and Computer Science, for his most valuable guidance, ideas, and suggestions in the completion of this thesis report. I also wish to extend my special thanks to Dr. Donald V. Sisson and Dr. Ronald V. Canfield for their critical review of the entire study.

Gratitude is also extended to my parents, Mr. and Mrs. C. S. Liu, for their continual encouragement and financial support in my graduate studies at Utah State University.

Finally, to my wife, Linda, for her understanding and encouragement during the preparation of this paper, I extend a husband's sincere affection and gratitude.

Chien-Hwa Liu

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

During the past few years there has been an increasing interest in developing computer packages to enhance the teaching of elementary statistics. The conventional ways of teaching statistics have used such devices as desk calculators, tables of functions, short-cut calculating formulas and electronic calculators, etc. to manipulate the involved computations. Electronic computers, in the past decade, have been broadly used in universities and colleges in many ways. It is only natural to extend the use of computers to the teaching function. Remote terminals can now be installed in any classroom and bring the computer to the students.

So it is now time to concentrate in other areas. In the teaching of statistics, one problem that has always been difficult to convey to students is sampling distributions. It makes sense to concentrate developmental efforts in producing Software that can be used to support the teaching function in the classroom itself. This thesis will direct itself to the producing of some terminal oriented Software for the common sampling distribution problems from one and two populations. This Software is being developed with the philosophy that a teacher will have a CRT-type terminal connected to TV monitors in the classroom so that the instructor can sit down before the CRT and direct the activities of the computer to the producing of several of the more common sampling distribution situations dealing with one and two populations. It is realized that such an effort is endless and some boundaries must be placed on what can be done by one person. This thesis contains program write-ups and execution runs for two basic sampling problems, that of drawing a sample from a single normally distributed population and then computing frequency distributions of the means, frequency distributions of the variance, frequency distributions of the sample "t" and sample "Chi-Square;" sample statistics; and finally confidence intervals for mean and variance. The same type of information is then gathered from a two populations problem where samples are drawn from two different populations with different characteristics and with possibly different sampling sizes and various frequency distributions are then built up of the differences between means, pooled variance, "t" of the difference between means, "F" of two sample variances ratio, and again finally confidence interval for difference between two means, etc.

Another field that has been developed is the computing of the probability density functions for the binomial and poisson distributions. It is understood that the Software developed in an effort of this kind will probably not fit everyone's needs exactly, however at least it is hoped that the ideas explored will lead other individuals to expand upon these concepts so that in the years to come a very useful set of computer programs will be developed, centered around the use of the computer via terminals in the classroom.

The writing of programs to be executed from the teletype is fairly standard in most ways; however, programs written to be executed in this manner require a great deal of attention to the advising of the user of what he must do for input. A good discussion of the techniques in this area are given by Chirlian (3).

Most input from the teletype will be via free style read statements. All input statements should be preceeded by an output statement detailing information to the user that must be entered at this particular point. Since time-sharing systems are paused with unexpected delays the user needs to be continuously advised when the job is in execution. It is contrasted with sitting and waiting for the computer to take an action. The programs given in this thesis are written in FORTRAN IV language and have been tested on the Burroughs B6700 System, using SR-33's and SR-38's teletype. The syntax used by the Burroughs Corporation for free style reads is different from that of IBM which again is different from the UNIVAC Corporation. Although FORTRAN seems to be fairly standardized language in many respects, the free style reads and writes are different. In the future some degree of standardization will probably be forth coming, but it certainly is not at the present time.

The computer programs require three files on disk where intermediate sampling results are stored prior to the time of the flow of operations when they can be printed upon request.

CHAPTER II

SAMPLING FROM A NORMAL POPULATION

Description

This program called "Sample I" is used to demonstrate the results of sampling experiments from a single normal population. It will draw any number of samples of any size from a normal population with specified mean and standard deviation. The program operation is divided into five sections.

A. Input of operational parameters.

B. Drawing of the sample observations.

C. Calculation of sample statistics.

D. Calculation of confidence intervals.

E. Output of desired frequency tables.

In the first section of the program the user is requested to enter the desired operational characteristics. The program first asks for the mean then the desired standard deviation, a large odd interger to be used as seed for a random number generator, the number of sample wanted, and then the sample size. The program again requests two parameters which are used to control the selective listing of the sample data and listing of sample statistics. The second part of the program then draws the appropriate samples from the specified populations and under control of the modulo functions selectively lists anything from none to all of the data and sample statistics. The results from the sampling phase are stored out on the disk files for later use. The program then interrogates the user to find out if confidence intervals of the mean and that of the variances are wanted. If they are desired by the user, the program then requires the input of appropriate tabular values of T and Chi-square. The confidence intervals are then listed for each of the two statistics, namely, the mean and standard deviation that have been indicated by the user. The final part of the program then requests the user to specify which of the various frequency distributions he wants to have produced. The program then processes the stored statistics into the desired frequency tables. The frequency distributions of the mean are processed into class intervals having a width of 1/4 of the standard error of the mean. The frequency distributions of the "t" statistic are processed into class intervals having a width of 1/4. The frequency tables for the chi-square statistic and sample variance are processed into class intervals having a width of 1/8 of the expected value of the sample variance so that the distance between zero and mid-point of the frequency distribution takes up eight class intervals.

The program uses the system random number generator which generates uniform random variables from zero to one. These uniform random variables are transformed into normal random deviates using the techniques given by Knuth (7). This procedure is carred out by a function sub-program called RNOR. This function basically produces a normalized deviate with mean 0 and standard deviation 1. This normalized deviate is then multiplied by the desired standard deviation, the desired mean is then added to it to produce a single observation. This is in a loop controlled on the size of the sample so that the desired number of observations are generated. At the end of this loop the mean and sample variance are produced and the inside and outer loop are controlled by the total number of samples wanted. The calculations involved in the program are the standard calculations given by all elementary statistics books.

Three disk files have been used in operating the program. The function of each file is as follows:

File 9: At the completion of the data generation phase contains all of the basic sample statistics.

File 10: If the option is required, contains all possible confidence intervals of the mean.

File 11: Contains the confidence intervals of the variance.

Theorems and Formulas

Following are four statistical theorems given by Li (8) that have been practically applied in developing the framework of the program. Through them the constructive conclusions can then be drawn.

A. "As the size of the sample increases, the distribution of the means of all possible samples of the same size drawn from the same population becomes more and more like a normal distribution provided that the population has a finite variance." (p. 33)

B. "If the population is normal, the distribution of sample means follows the normal distribution exactly, regardless of the size of the sample." (p. 35)

C. "The mean of the means of all possible samples of the same size drawn from the same population is equal to the mean of that population." (p. 35)

D. "The variance of these sample means is equal to population variance divided by the size of the sample." (p. 36)

In addition to these four basical theorems, many others have been directly or indirectly used in the development of the program. It is impossible to indicate all of them here. Anyone who is really interested in this matter could read and find them in any standard elementary statistics book.

The following are the formulas for the sample statistics which are computed and tallied into the corresponding frequency tables:

1. Sample Mean $\bar{X} = \Sigma X_i/n$
2. Sample Variance $S^2 = \Sigma(X_i - \bar{X})^2/(n - 1)$
3. Sample "T" $T = (\bar{X} - \mu)/(S/\sqrt{n})$
4. Sample "Chi-Square" $\chi^2 = (n - 1)S^2/\sigma^2$
5. Standard Error of the Mean $SE = \sigma/\sqrt{n}$

Where

X_i = Computer generated observations

n = sample size per set of data

μ = population mean

σ = population standard deviation

6. Confidence Interval of Mean $\bar{X} - t\sqrt{S^2/n} < \mu < \bar{X} + t\sqrt{S^2/n}$

7. Confidence Interval of Variance $(n-1)S^2/\chi_1^2 < \sigma^2 < (n-1)S^2/\chi_2^2$

Where

t = Given tabular value of "t"

χ_1^2 = Given right-tailed tabular value

χ_2^2 = Given left-tailed tabular value.

Input

All computer input formats are free style as are broadly used in the current terminal oriented languages. The computer reads the exact forms that a user enters. A series of message for input is given to the user to request that a user enters the main parameters that must exactly coincide with the following order:

1. Population mean.
2. Population standard deviation.
3. Random number argument as seed for data generating.
4. Number of samples wanted.
5. Number of observations per sample.
6. List control for sample data.
7. List control for sample statistics.
8. Tabular value of "t," if requested.
9. Two tailed tabular values of "Chi-Square," if requested.

The program will control the listing of the sample results, data, and statistics, according to the user's commands. Item 6 and 7 are these controls. For example, if the user wants the results to be printed every tenth, he simply enters number 10 to answer the request. The results then will be listed in every tenth time, for instance, 10, 20, 30 ...., etc. The random number argument should be a large odd integer. If more than one number has to be entered on the same row and at the same time, then, the comma (,) set between them is needed. After the sample results have been completed, a series of messages are given to the user to request the entry of "YES" to print the various frequency tables. Since the word "YES" is so crucial in affecting the output the user must be cautious. During data entry, user must very carefully re-examine each line he enters before he depresses the "Return" key. If an error is detected soon after making it, the user can make an internal correction in the computer storage simply by depressing the "Control" button all the way and then punching "H" key to move the curser back to the position where correction is needed. The desired characters are then retyped. For more serious errors, the user simply depresses the "Rub Out" button and re-enters the line. If the user finds some errors after depressing the "Return" key, he should terminate the job and start over. So, to be equipped with the knowledge of how to operate a terminal machine is very strongly recommended.

Sample Results

The following are two sets of sample exhibits of the execution runs on a teletype. The first one (Exhibit I) is the inputs and their results with the confidence intervals of mean and the confidence intervals of variance only. The second (Exhibit II) is the inputs and their results with sample data (observations), statistics, and their frequency distribution tables.

1. Exhibit I

Inputs and Results with Confidence Intervals

(i) Input Portion

```
 ENTER THE DESIRED GRAND MEAN OF THE POPULATION.
#?
50.

 ENTER THE DESIRED STANDARD DEVIATION OF THE POPULATION.
3.

 ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
4365791

 ENTER THE NUMBER OF SAMPLES WANTED.
30

 ENTER THE NUMBER OF OBSERVATIONS PER SAMPLE.
12

 HOW FREQUENTLY DO YOU WANT DATA LISTED?
   0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......, ETC.
0

 HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
   0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......., ETC.
5

 DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION MEAN?
 INSERT YES OR NO.
YES

 ENTER YOUR DESIRED TABULAR VALUE OF "T".
2.2015

 DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION VARIANCE?
 INSERT YES OR NO.
YES

 ENTER YOUR DESIRED TWO TABULAR VALUES OF "CHI-SQUARE".
 ONE FOR "LEFT TAIL", THE OTHER "RIGHT TAIL".
3.81974,21.91579
```

(ii) Statistics and Confidence Intervals of Mean

Mean = 50

#	XBAR	VAR	"T"	CHI-SQ
5	50.22	7.736	0.278	9.455
10	49.84	12.269	-0.157	14.995
15	50.00	11.110	0.000	13.578
20	50.00	9.584	0.000	11.714
25	49.30	10.983	-0.727	13.424
30	49.32	13.869	-0.631	16.951

#	MEAN(L)	MEAN(R)
1	47.64	51.70
2	48.38	52.56
3	46.75	51.54
4	48.51	52.35
5	48.46	51.99
6	48.76	52.06
7	46.58	51.11
8	46.89	51.36
9	48.63	52.19
10	47.62	52.07
11	47.79	51.17
12	48.31	51.72
13	46.61	50.99
14	47.53	51.36
15	47.88	52.12
16	47.71	51.54
17	48.05	51.17
18	47.46	51.03
19	48.06	51.00
20	48.03	51.97
21	47.39	49.83
22	47.31	51.10
23	48.46	51.81
24	48.71	52.44
25	47.20	51.41
26	48.41	52.01
27	48.38	51.30
28	48.99	52.76
29	47.51	51.25
30	46.95	51.69

(iii) Confidence Intervals of Variance

Variance = 9

#	*	VAR(L)	--	VAR(R)	*
1		5.115		29.382	
2		5.436		31.221	
3		7.123		40.912	
4		4.582		26.317	
5		3.883		22.303	
6		3.389		19.467	
7		6.386		36.683	
8		6.210		35.671	
9		3.951		22.695	
10		6.158		35.369	
11		3.554		20.414	
12		3.608		20.721	
13		5.951		34.183	
14		4.537		26.062	
15		5.576		32.028	
16		4.542		26.086	
17		3.033		17.418	
18		3.962		22.757	
19		2.700		15.510	
20		4.810		27.629	
21		1.853		10.644	
22		4.470		25.675	
23		3.488		20.035	
24		4.317		24.794	
25		5.513		31.663	
26		4.034		23.173	
27		2.652		15.231	
28		4.408		25.316	
29		4.368		25.087	
30		6.961		39.984	

2. Exhibit II

Inputs and Results with Frequency Distribution Tables

(i) Input Portion

```
 ENTER THE DESIRED GRAND MEAN OF THE POPULATION.
#?
50.

 ENTER THE DESIRED STANDARD DEVIATION OF THE POPULATION.
3.

 ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
4365791

 ENTER THE NUMBER OF SAMPLES WANTED.
100

 ENTER THE NUMBER OF OBSERVATIONS PER SAMPLE.
10

 HOW FREQUENTLY DO YOU WANT DATA LISTED?
   0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......., ETC.
10

 HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
   0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,........, ETC.
10

 DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION MEAN?
 INSERT YES OR NO.
NO

 DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION VARIANCE?
 INSERT YES OR NO.
NO
```

(ii) Sample Data and Statistics

SAMPLE OBSERVATIONS:

10	47.58	55.16	44.76	45.59	45.92	51.26	50.49	45.59
	48.38	55.91						
20	54.36	47.33	49.59	51.57	50.31	50.81	50.96	44.18
	47.05	49.11						
30	54.72	47.10	52.76	44.54	46.14	51.54	51.16	52.85
	49.12	48.07						
40	54.51	49.78	52.71	43.87	49.78	45.95	46.21	50.93
	49.98	51.33						
50	44.47	55.25	51.67	51.26	50.62	54.04	48.66	54.04
	49.41	48.31						
60	48.02	44.03	44.39	50.17	50.38	51.26	49.31	48.50
	48.91	49.37						
70	49.70	46.46	46.14	54.06	49.09	50.99	46.77	49.49
	47.65	49.34						
80	55.95	47.66	48.63	51.01	46.26	48.45	49.27	47.70
	51.87	50.82						
90	51.69	48.37	55.17	51.23	50.91	50.08	45.58	49.51
	50.47	44.74						
100	47.63	45.85	49.51	47.42	49.09	49.29	53.00	48.98
	49.27	53.58						

****

#	XBAR	VAR	"T"	CHI-SQ
10	49.06	16.233	-0.735	16.233
20	49.53	7.953	-0.530	7.953
30	49.80	10.995	-0.191	10.995
40	49.51	10.701	-0.479	10.701
50	50.77	10.508	0.754	10.508
60	48.43	5.842	-2.049	5.842
70	48.97	5.759	-1.359	5.759
80	49.76	7.732	-0.271	7.732
90	49.78	9.083	-0.236	9.083
100	49.36	5.599	-0.853	5.599

(iii) Frequency Table of Means

ENTER YES, TO PRINT FREQUENCY TABLE OF MEANS.
YES

FREQUENCY TABLE OF MEANS

LOWER BOUND	FREQUENCY
.462053E+02	0
.464424E+02	0
.466796E+02	0
.469168E+02	0
.471540E+02	1
.473911E+02	1
.476283E+02	1
.478655E+02	3
.481026E+02	0
.483398E+02	3
.485770E+02	6
.488141E+02	7
.490513E+02	8
.492885E+02	8
.495257E+02	10
.497628E+02	10
.500000E+02	11
.502372E+02	3
.504743E+02	8
.507115E+02	4
.509487E+02	3
.511859E+02	4
.514230E+02	5
.516602E+02	3
.518974E+02	0
.521345E+02	0
.523717E+02	0
.526089E+02	1
.528460E+02	0
.530832E+02	0
.533204E+02	0
.535576E+02	0

(iv) Frequency Table of Variances

```
ENTER YES, TO PRINT FREQUENCY TABLE OF VARIANCES.
YES

FREQUENCY TABLE OF VARIANCES
LOWER BOUND     FREQUENCY
Ø.                 Ø
 .112500E+01       Ø
 .225000E+01       1
 .337500E+01       9
 .450000E+01       8
 .562500E+01      10
 .675000E+01      10
 .787500E+01      13
 .900000E+01      12
 .101250E+02      15
 .112500E+02       3
 .123750E+02       4
 .135000E+02       3
 .146250E+02       Ø
 .157500E+02       5
 .168750E+02       3
 .180000E+02       1
 .191250E+02       1
 .202500E+02       1
 .213750E+02       Ø
 .225000E+02       Ø
 .236250E+02       Ø
 .247500E+02       Ø
 .258750E+02       Ø
 .270000E+02       Ø
 .281250E+02       Ø
 .292500E+02       Ø
 .303750E+02       Ø
 .315000E+02       Ø
 .326250E+02       1
```

(v) Frequency Table of "T" Values

```
ENTER YES, TO PRINT FREQUENCY TABLE OF T'S.
YES
```

FREQUENCY TABLE OF 'T' VALUES

LOWER BOUND	FREQUENCY
-.400000E+01	1
-.375000E+01	0
-.350000E+01	1
-.325000E+01	1
-.300000E+01	0
-.275000E+01	1
-.250000E+01	1
-.225000E+01	1
-.200000E+01	4
-.175000E+01	1
-.150000E+01	5
-.125000E+01	7
-.100000E+01	7
-.750000E+00	7
-.500000E+00	12
-.250000E+00	9
0.	11
.250000E+00	4
.500000E+00	8
.750000E+00	3
.100000E+01	3
.125000E+01	4
.150000E+01	3
.175000E+01	2
.200000E+01	1
.225000E+01	0
.250000E+01	2
.275000E+01	1
.300000E+01	0
.325000E+01	0
.350000E+01	0
.375000E+01	0

(vi) Frequency Table of "Chi-Square" Values

```
ENTER YES, TO PRINT FREQUENCY TABLE OF CHI-SQ'S.
YES

FREQUENCY TABLE OF CHI-SQUARE VALUES
LOWER BOUND    FREQUENCY
0.                 0
 .112500E+01       0
 .225000E+01       1
 .337500E+01       9
 .450000E+01       8
 .562500E+01      10
 .675000E+01      10
 .787500E+01      13
 .900000E+01      12
 .101250E+02      15
 .112500E+02       3
 .123750E+02       4
 .135000E+02       3
 .146250E+02       0
 .157500E+02       5
 .168750E+02       3
 .180000E+02       1
 .191250E+02       1
 .202500E+02       1
 .213750E+02       0
 .225000E+02       0
 .236250E+02       0
 .247500E+02       0
 .258750E+02       0
 .270000E+02       0
 .281250E+02       0
 .292500E+02       0
 .303750E+02       0
 .315000E+02       0
 .326250E+02       1
```

By studying the output of the sampling experiments from a single normal population, a student should be able to obtain a better understanding of the purpose of this program and develop his intuition about the natures of the basic statistical analyses.

CHAPTER III

SAMPLING FROM TWO NORMAL POPULATIONS

Description

This program called "Sample II" is used to demonstrate the results of sampling experiments from two normal populations. It will independently draw a pair of samples of any size from each of two normal populations with a different specified means and standard deviations. The entire program operation is divided into five sections.

A. Input of operational parameters.

B. Drawing of the sample observations.

C. Calculation of sample statistics.

D. Calculation of confidence intervals.

E. Output of desired frequency tables.

In the first section of the program the user is requested to enter the desired operational characteristics. The program first asks for the means then the desired standard deviations, then a large odd integer to be used as seed for a random number generator, the number of samples wanted, and then the desired sample size for each of the two samples. The program again requests two parameters which are used to control the selective listing of the sample data and listing of sample statistics. The second part of the program draws the appropriate observations for each of the samples and selectively lists from none to all of the data and statistics. The results from the sampling phase are stored out on the disk files for later use. The program then interrogates the user to find out if confidence intervals of a difference between two means are wanted. If they are desired by the user, the program then requires the appropriate tabular values of T so as to calculate the left and right boundaries for the intervals. Only the confidence interval for the difference between means is provided in this case because it is often used in the most statistical applications. The listing of this confidence interval for each sample follows. In the final section the program requests the user to specify which of the various frequency distributions he wants to have produced. The program then processes the stored difference of means and the standard error of the difference between two means and the polled population variances into the desired frequency tables.

The frequency distributions of the difference between two sample means are processed into class intervals having a width of 1/4 of the standard error of the difference between two means. The frequency distributions of T statistic of the difference between two means are processed into class interval having a width of 1/4. The frequency table for pooled sample variances is processed into class intervals having a width of 1/8 of the expected pooled population variance so that the distance between 0 and the mid-point of the frequency distribution takes up intervals eight classes. Finally, the frequency table for F statistic of the ratio of two sample variances is processed into class interval having a width of 1/8 of the expected ratio of two population variances; the frequency table for F statistic will become a non-central F distribution if two population variances are not equal.

The program uses the system random number generator which generates uniform random variables from zero to one. These uniform random variates are then transformed into normal random deviates by using the techniques given by Knuth (7). The procedure is carried out by a function sub-program called RNOR. This function produces a normalized deviate with mean 0 and standard deviation 1. This normalized deviate is then multiplied by the desired standard deviation and the desired mean is then added to it to produce a single observation. This is in a loop controlled on the size of the sample so that the desired number of observations are thus generated. This is repeated for the second sample group. At the end of this loop the means and pooled sample variance are computed and the inside and outer loop are still controlled by the total number of samples wanted the procedures and methods of calculations in this program are described in any good elementary statistics book.

Three disk files have been designated to operate the storage of the entire sample information. The basic function of each file is as follows:

File 9: At the completion of the data generation phase contains the major part of the basic sample statistics.

File 10: Contains the rest of the sample statistics.

File 11: If the option is required, contains all possible confidence intervals of difference between two means.

Theorems and Formulas

Following are three essential statistical theorems given by Li (8) that have been applied in developing the framework of the program. Through them the further theoretical achievements can then be deduced.

A. "The mean of the differences between two sample means is equal to the difference between the means of the two populations from which the two sets of samples are drawn." (p. 122)

B. "The variance of the difference between two sets of independent sample means is equal to the sum of the variances of the two respective sets of sample means." (p. 124)

C. "The distribution of the difference $(\bar{X}_1 - \bar{X}_2)$ between sample means approaches the normal distribution as the sample sizes n_1 and n_2 increase, if the two populations from which the samples are drawn have finite variances. If the two populations are normal at the outset, the differences $(\bar{X}_1 - \bar{X}_2)$ follow the normal distribution exactly regardless of the sample sizes." (p. 124)

Indeed, in addition to these three important theorems as stated above, many others have been also directly or indirectly used in the development of the program. Anyone who is really interested in this matter could read and find them in any standard elementary statistics book.

The following are the statistical formulas which have been used and applied in developing the program. Among them four major sample statistics are tallied into the corresponding frequency tables.

1. Sample Mean $\bar{X}_i = \sum_j X_{ij}/n_i$

where $i = 1,2$
$j = 1,2,3,\ldots\ldots\ldots\ldots,n_i$

2. Sample Variance $S_i^2 = \sum_j (X_{ij} - \bar{X}_i)^2/(n_i - 1)$

where $i = 1,2$
$j = 1,2,\ldots\ldots\ldots\ldots\ldots,n_i$

3. Difference Between Two Sample Means $u = \bar{X}_1 - \bar{X}_2$

4. Pooled Variance

$$S_p^2 = [S_1^2(n_1 - 1) + S_2^2(n_2 - 1)]/(n_1 + n_2 - 2)$$

5. Variance of the Difference Between Two Sample Means

(i) $\sigma_1^2 = \sigma_2^2$ $\quad VAR_p = S_p^2/n_1 + S_p^2/n_2$

(ii) $\sigma_1^2 \neq \sigma_2^2$ $\quad VAR_{np} = S_1^2/n_1 + S_2^2/n_2$

6. Standard Error of the Difference Between Two Means

$$S.E. = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

7. μ Statistic

$$\mu = (\bar{X}_1 - \bar{X}_2)/\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

8. T Statistic

(i) If $\sigma_1^2 = \sigma_2^2$

$$T = (\bar{X}_1 - \bar{X}_2)/\sqrt{S_p^2/n_1 + S_p^2/n_2}$$

(ii) If $\sigma_1^2 \neq \sigma_2^2$

$$T = (\bar{X}_1 - \bar{X}_2)/\sqrt{S_1^2/n_1 + S_2^2/n_2}$$

9. F Statistic $F = S_1^2/S_2^2$

Where

X_{ij} = Computer generated observations for each sample group

n_1 = Number of observations per sample for the first sample group

n_2 = Number of observations per sample for the second sample group

σ_1 = Standard deviation for the first population

σ_2 = Standard deviation for the second population

μ_1 = Grand mean for the first population

μ_2 = Grand mean for the second population

10. Confidence Interval of Difference Between Means

(i) If $\sigma_1^2 = \sigma_2^2$

$$(\bar{X}_1 - \bar{X}_2) \pm t \sqrt{VAR_p}$$

(ii) If $\sigma_1^2 \neq \sigma_2^2$

$$(\bar{X}_1 - \bar{X}_2) \pm t \sqrt{VAR_{np}}$$

Where

t = Given tabular value of "t" with $(n_1 + n_2 - 2)$ degrees of of freedom.

In Formulas 7 and 8 mentioned above, only $(\bar{X}_1 - \bar{X}_2)$ in numerator portion is used instead of $[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)]$. The reason to do so is that it is intended to give the user more games to play in testing the statistical hypothesis that two population means are equal.

Input

All computer input formats are free style, as are broadly used in the current terminal oriented languages. A series of messages for operating the input portion is given to the user to request that the following operational parameters be entered.

1. Grand mean for the first population.
2. Grand mean for the second population.
3. Standard deviation of the first population.
4. Standard deviation of the second population.
5. Random number argument as seed for random number generator.
6. Number of samples wanted.
7. Number of observations per sample of first group.
8. Number of observations per sample of second group.
9. List control for listing of sample data.
10. List control for listing of sample statistics.
11. Tabular value of "t" for confidence interval of difference between two means.

The list controls, items 9 and 10 need further comment. If a user wants the results to be printed every twentieth, all he has to do is to enter number 20 to response to the computer's interrogation. Then the results will be listed in every twentieth time for instance 20, 40, 60, ..., etc.

The means for the two populations do not have to be equal. Both equal and unequal situations are handled in the program so that users are offered alternative ways to practice the elementary statistical concepts without worrying which case would be best to follow. If two standard deviations are equal then the pooled variance is computed and will be used all the way; otherwise, the non-pooled variance will be substituted. So, to put it briefly any individual user will be totally free to handle his own problem and to make his own decision on what input to be entered and what results to be printed.

If more than one number needs to be entered on the same row, then the comma (,) between them is needed. After the sample results have been completed, a series of messages are given to the user to request the entry of "YES" in order to print the various frequency tables. This "YES" should be entered as the first three characters on the line. During data entry, user must very carefully re-examine each line he enters before he depresses the "Return" key. If an error is detected soon after making it, the user can make an internal correction in the computer storage simply by depressing the "Control" button all the way and then punching the "H" key to move the curser back to the position where correction is needed. The desired characters are then retyped again. For more serious errors, the user simply depresses the "Rub Out" button and re-enters the line. If the user finds some errors after depressing the "Return" key, he should terminate the job and start it over. So, to be equipped with the knowledge of how to operate the remote terminal machine is very strongly recommended.

Sample Results

Following are two sets of sample exhibits of the execution runs. The first one (Exhibit III) is the inputs and their results with the confidence intervals of difference between two means and the second (Exhibit IV) is the inputs and their results with sample observations, statistics, and the frequency distribution tables.

1. Exhibit III

   Inputs and Results with Confidence Intervals for Difference Between Two Means

   (i) Input Portion

```
ENTER 1ST DESIRED GRAND MEAN FOR THE FIRST POPULATION.
#?
55.

ENTER 2ND DESIRED GRAND MEAN FOR THE SECOND POPULATION.
52.

ENTER DESIRED STANDARD DEVIATION FOR THE 1ST POPULATION.
5.

ENTER DESIRED STANDARD DEVIATION FOR THE 2ND POPULATION.
3.

ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
4365791

ENTER THE NUMBER OF SAMPLES WANTED.
30

ENTER NOS. OF OBSERVATION PER SAMPLE FOR 1ST GROUP.
10

ENTER NOS. OF OBSERVATION PER SAMPLE FOR 2ND GROUP.
12

HOW FREQUENTLY DO YOU WANT DATA LISTED?
  0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......, ETC.
0

HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
  0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......, ETC.
5

DO YOU WANT THE CONFIDENCE INTERVAL FOR DIFFERENCE
  BETWEEN TWO MEANS?     INSERT YES OR NO.
YES

ENTER YOUR DESIRED TABULAR VALUE OF "T".
2.08614
```

(ii) Confidence Intervals of Difference Between Means

Difference Between Two Means = 3

#	*	MEAN(L)	--	MEAN(R)	*
1		-3.22		4.31	
2		-Ø.24		8.Ø9	
3		-Ø.65		6.29	
4		Ø.4Ø		6.91	
5		-4.47		4.18	
6		-Ø.67		8.Ø5	
7		Ø.21		7.ØØ	
8		-2.91		4.3Ø	
9		-1.65		5.56	
1Ø		-Ø.39		6.64	
11		-Ø.6Ø		5.89	
12		-1.79		3.96	
13		-Ø.97		5.38	
14		-3.26		5.33	
15		-Ø.17		6.71	
16		1.61		9.Ø3	
17		-1.54		6.38	
18		3.35		11.25	
19		-Ø.85		7.93	
2Ø		1.72		8.69	
21		-Ø.59		5.14	
22		Ø.61		8.12	
23		Ø.51		8.28	
24		2.45		9.45	
25		-Ø.47		7.Ø2	
26		1.Ø8		6.93	
27		-3.81		6.23	
28		-2.48		4.44	
29		-6.22		5.17	
3Ø		4.ØØ		1Ø.7Ø	

2. Exhibit IV

Inputs and Results with Sample Observations, Statistics, and Frequency Tables

(i) Input Portion

```
ENTER 1ST DESIRED GRAND MEAN FOR THE FIRST POPULATION.
#?
55.

ENTER 2ND DESIRED GRAND MEAN FOR THE SECOND POPULATION.
54.

ENTER DESIRED STANDARD DEVIATION FOR THE 1ST POPULATION.
5.

ENTER DESIRED STANDARD DEVIATION FOR THE 2ND POPULATION.
3.

ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
4365791

ENTER THE NUMBER OF SAMPLES WANTED.
100

ENTER NOS. OF OBSERVATION PER SAMPLE FOR 1ST GROUP.
12

ENTER NOS. OF OBSERVATION PER SAMPLE FOR 2ND GROUP.
10

HOW FREQUENTLY DO YOU WANT DATA LISTED?
  0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......., ETC.
20

HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
  0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......., ETC.
10

DO YOU WANT THE CONFIDENCE INTERVAL FOR DIFFERENCE
 BETWEEN TWO MEANS?    INSERT YES OR NO.
NO
```

(ii) Sample Observations

```
SAMPLE OBSERVATIONS:

   20    57.62    49.02    57.32    57.07    61.14    61.82    60.60    52.71
         52.32    54.34    60.77    52.21
   20    53.96    55.39    50.65    49.05    57.16    45.47    54.25    53.55
         55.83    53.16

>>>>>>>>>>>>>>>>>>>>

   40    54.09    66.37    57.04    60.19    47.48    61.95    63.76    49.57
         55.67    51.42    58.05    56.46
   40    47.05    47.66    57.32    50.26    53.69    56.51    56.82    53.50
         54.41    58.97

>>>>>>>>>>>>>>>>>>>>

   60    62.21    55.33    53.44    56.39    56.75    54.28    56.50    59.08
         49.41    50.17    51.50    53.69
   60    58.66    52.02    54.15    54.90    47.93    54.11    54.17    52.57
         57.77    53.14

>>>>>>>>>>>>>>>>>>>>

   80    54.69    60.74    56.63    54.02    52.20    65.11    53.57    59.20
         55.39    62.17    57.68    55.43
   80    51.00    52.43    54.32    56.84    54.00    55.58    56.72    58.52
         57.43    55.70

>>>>>>>>>>>>>>>>>>>>

  100    46.88    55.52    53.69    51.25    48.89    50.14    54.80    60.79
         55.53    42.73    60.97    56.07
  100    57.99    54.71    59.35    46.80    48.12    53.56    49.84    53.29
         53.71    49.07

>>>>>>>>>>>>>>>>>>>>
```

(iii) Sample Statistics

#	MEAN-X	MEAN-Y	DIFF OF MEANS	VAR-X	VAR-Y	VAR OF DIFF
10	55.11	52.71	2.40	23.318	3.039	2.247
20	56.41	52.85	3.56	17.990	12.381	2.737
30	56.08	51.69	4.39	23.317	7.342	2.677
40	56.84	53.62	3.22	32.406	16.836	4.384
50	54.08	53.31	0.78	15.437	10.757	2.362
60	54.90	53.94	0.95	13.395	8.924	2.009
70	53.52	54.18	-0.65	46.236	4.404	4.293
80	57.24	55.25	1.98	15.031	5.432	1.796
90	55.74	53.72	2.03	18.944	4.266	2.005
100	53.11	52.64	0.46	29.184	17.274	4.159

#	"T"	"F"	POOL VAR
10	1.60	7.67	14.192
20	2.15	1.45	15.466
30	2.69	3.18	16.128
40	1.54	1.92	25.399
50	0.51	1.44	13.331
60	0.67	1.50	11.383
70	-0.32	10.50	27.412
80	1.48	2.77	10.711
90	1.43	4.44	12.339
100	0.23	1.69	23.825

(iv) Frequency Table of Difference Between Means

```
ENTER YES, TO PRINT FREQUENCY TABLE OF DIFFERENCE BETWEEN MEANS
YES

FREQUENCY TABLE OF DIFFERENCE BETWEEN MEANS.
```

LOWER BOUND	FREQUENCY
-.59Ø893E+Ø1	Ø
-.547712E+Ø1	Ø
-.5Ø4531E+Ø1	Ø
-.461351E+Ø1	Ø
-.41817ØE+Ø1	1
-.374989E+Ø1	Ø
-.3318Ø8E+Ø1	Ø
-.288627E+Ø1	1
-.245447E+Ø1	1
-.2Ø2266E+Ø1	3
-.159Ø85E+Ø1	4
-.1159Ø4E+Ø1	5
-.727233E+ØØ	8
-.295425E+ØØ	1Ø
.136384E+ØØ	4
.568192E+ØØ	11
.1ØØØØØE+Ø1	7
.143181E+Ø1	8
.186362E+Ø1	11
.229542E+Ø1	5
.272723E+Ø1	1Ø
.3159Ø4E+Ø1	6
.359Ø85E+Ø1	1
.4Ø2266E+Ø1	Ø
.445447E+Ø1	3
.488627E+Ø1	1
.5318Ø8E+Ø1	Ø
.574989E+Ø1	Ø
.61817ØE+Ø1	Ø
.661351E+Ø1	Ø
.7Ø4531E+Ø1	Ø
.747712E+Ø1	Ø

(v) Frequency Table of Pooled Variances

ENTER YES, TO PRINT FREQUENCY TABLE OF POOLED VARIANCES.
YES

FREQUENCY TABLE OF POOLED VARIANCES

LOWER BOUND	FREQUENCY
Ø.	Ø
.2225ØØE+Ø1	Ø
.445ØØØE+Ø1	Ø
.6675ØØE+Ø1	2
.89ØØØØE+Ø1	1Ø
.11125ØE+Ø2	19
.1335ØØE+Ø2	18
.15575ØE+Ø2	15
.178ØØØE+Ø2	7
.2ØØ25ØE+Ø2	7
.2225ØØE+Ø2	7
.24475ØE+Ø2	8
.267ØØØE+Ø2	2
.28925ØE+Ø2	4
.3115ØØE+Ø2	Ø
.33375ØE+Ø2	Ø
.356ØØØE+Ø2	Ø
.37825ØE+Ø2	Ø
.4ØØ5ØØE+Ø2	Ø
.42275ØE+Ø2	1
.445ØØØE+Ø2	Ø
.46725ØE+Ø2	Ø
.4895ØØE+Ø2	Ø
.51175ØE+Ø2	Ø
.534ØØØE+Ø2	Ø
.55625ØE+Ø2	Ø
.5785ØØE+Ø2	Ø
.6ØØ75ØE+Ø2	Ø
.623ØØØE+Ø2	Ø
.64525ØE+Ø2	Ø

(vi) Frequency Table of "T", Difference of Two Means

```
 ENTER YES, TO PRINT FREQUENCY TABLE OF "T",  DIFF OF TWO MEANS.
YES

  FREQUENCY TABLE OF "T", DIFF OF TWO MEANS
 LOWER BOUND   FREQUENCY
  -.4ØØØØØE+Ø1        Ø
  -.375ØØØE+Ø1        Ø
  -.35ØØØØE+Ø1        Ø
  -.325ØØØE+Ø1        Ø
  -.3ØØØØØE+Ø1        Ø
  -.275ØØØE+Ø1        Ø
  -.25ØØØØE+Ø1        1
  -.225ØØØE+Ø1        Ø
  -.2ØØØØØE+Ø1        Ø
  -.175ØØØE+Ø1        Ø
  -.15ØØØØE+Ø1        4
  -.125ØØØE+Ø1        3
  -.1ØØØØØE+Ø1        3
  -.75ØØØØE+ØØ        2
  -.5ØØØØØE+ØØ       11
  -.25ØØØØE+ØØ        9
  Ø.                  5
   .25ØØØØE+ØØ        9
   .5ØØØØØE+ØØ        9
   .75ØØØØE+ØØ        8
   .1ØØØØØE+Ø1        8
   .125ØØØE+Ø1        4
   .15ØØØØE+Ø1        9
   .175ØØØE+Ø1        4
   .2ØØØØØE+Ø1        5
   .225ØØØE+Ø1        3
   .25ØØØØE+Ø1        2
   .275ØØØE+Ø1        1
   .3ØØØØØE+Ø1        Ø
   .325ØØØE+Ø1        Ø
   .35ØØØØE+Ø1        Ø
   .375ØØØE+Ø1        Ø
```

(vii) Frequency Table of "F", Ratio of Two Variances

$$F = S_1^2/S_2^2$$

```
ENTER YES, TO PRINT FREQUENCY TABLE OF "F", RATIO OF TWO VARIANCES.
YES

  FREQUENCY TABLE OF "F", RATIO OF TWO VARIANCES
 LOWER BOUND     FREQUENCY
  Ø.                 Ø
   .347222E+ØØ       1
   .694444E+ØØ       3
   .1Ø4167E+Ø1      1Ø
   .138889E+Ø1      15
   .173611E+Ø1      15
   .2Ø8333E+Ø1       6
   .243Ø56E+Ø1       9
   .277778E+Ø1       6
   .3125ØØE+Ø1       5
   .347222E+Ø1       2
   .381944E+Ø1       5
   .416667E+Ø1       3
   .451389E+Ø1       1
   .486111E+Ø1       4
   .52Ø833E+Ø1       2
   .555556E+Ø1       Ø
   .59Ø278E+Ø1       2
   .625ØØØE+Ø1       2
   .659722E+Ø1       Ø
   .694444E+Ø1       Ø
   .729167E+Ø1       1
   .763889E+Ø1       1
   .798611E+Ø1       Ø
   .833333E+Ø1       Ø
   .868Ø56E+Ø1       1
   .9Ø2778E+Ø1       Ø
   .9375ØØE+Ø1       1
   .972222E+Ø1       Ø
   .1ØØ694E+Ø2       5
```

By studying the output of the sampling experiments from two normal populations, a student should be able to obtain a better understanding of the purpose of this program and develop his intuition about the natures of the basic statistical analyses.

CHAPTER IV

DISCRETE DENSITY FUNCTIONS--BINOMIAL AND POISSON

Description

These programs, called "Density I" and "Density II", are used to compute density functions for the binominal and poisson distributions. Individual probabilities of each distribution plus both cumulative densities, less than or equal to, and greater than or equal to, are computed. The programs give the message to the user to enter the sample size and proportions of a "success" for the binomial or the Lamda parameter for the poisson distribution.

Even though the computer programs that handle these two discrete density functions are relatively similar, each specific function will be discussed separately. First, the binomial distribution will be discussed. According to the nature of binomial function, the probability of a "success" must be the same for each trial in any single sample size, and the trials must be all independent. The developed procedures of the program thus completely follows this nature. There are three categories of the density functions in the program, namely: P(X), individual probabilities associated with the specific probability of obtaining a success on a single trail; PLEX, ascending accumulated probabilities from these individual ones; and PGEX, descending accumulated probabilities.

According to Freund's book (5), "... generally speaking, the poisson distribution will provide a good approximation to binomial probabilities when n (the trials) is at least 20 and θ (the probability) is at most 0.05; when n is at least 100, the approximation will generally be excellent provided $n \times \theta$ does not exceed 10." (p. 83) In reality, the poisson distribution can perform more functions in statistical applications than just the approximation of the binomial, however, it is not the major object to discuss here. Similar to the binomial distribution categories, it has P(X), PLEX, and PGEX to identify each function which it plays. Only one parameter involved in the program for the poisson distribution is "LAMBDA."

Methodology

No disk files are involved in this program. All computations used in the programs of Density I and Density II are in double precision. Mathematical development and recursive relations of the initial formulas are used to compute the successive terms. The program also allows the user to change his version of the operational parameter(s) that the probability of a "success" and the total number of trials in the binomial or the "LAMBDA" in the poisson distribution in order to lead another version of the probability density functions until user enters the word "NO" to request a stop of the computer execution.

Mathematical Analysis for Binomial

The initial formula is that:

$$P\ (x;n,\theta) = n!/x!\ (n - x)\ !\ \theta^{x}(1 - \theta)^{n - x}$$

$$P\ (x + 1;n,\theta) = n!/(x + 1)!(n - x - 1)!\theta^{x + 1}(1 - \theta)^{n - x - 1}$$

To form a ratio, such that

$$P\ (x + 1;n.\theta)/P(x;n,\theta) = [(n - x)/(x + 1)].[\theta/(1 - \theta)].$$

Hence,

$$P(x + 1;n,\theta) = P(x;n,\theta)\ .\ [(n - x)/(x + 1)]\ .\ [\theta/(1 - \theta)].$$

Therefore, the recursive relations are:

$$P\ (0;n,\theta) = (1 - \theta)^{n}$$

$$P\ (1;n,\theta) = n\ .\ (1 - \theta)^{n - 1}$$

$$P\ (2;n,\theta) = n\ .\ (n - 1)\ .\ (1 - \theta)^{n - 2}\ .\ (\theta^{2}/2!)$$

$$P\ (3;n,\theta) = n\ .\ (n - 1)\ .\ (n - 2)\ .\ (1 - \theta)^{n - 3}\ .\ (\theta^{3}/3!)$$

etc.

Mathematical Analysis for Poisson

The initial formula is that

$$P\ (x;\lambda) = e^{-\lambda}.\ \lambda^{x}/x!$$

$$P\ (x + 1;\lambda) = e^{-\lambda}.\ \lambda^{x + 1}/(x + 1)!$$

To form a ratio such that

$$P\ (x + 1;\lambda)/P(x;\lambda) = \lambda/(x + 1)$$

Hence, $P(x + 1;\lambda) = P(x;\lambda)\lambda/(x + 1)$

Therefore, the recursive relations are:

$P(0;\lambda) = e^{-\lambda}$, $P(1;\lambda) = e^{-\lambda}.\ \lambda$

$P(2;\lambda) = e^{-\lambda}.\ \lambda^{2}/2!$, $P(3;\lambda) = e^{-\lambda}.\ \lambda^{3}/3!$

etc.

Sample Results

This section of sample results shows the execution of each program. For the binomial the sample size of 20 and a proportion of a "success" of 0.25 for the sample trials were used. For the poisson the "LAMBDA" values of 5 and 10 were used, respectively.

1. Execution Results for Binomial

(i) Exhibit V

Sample Size = 20
θ = 0.25

ENTER THE DESIRED SAMPLE SIZE (N).
#?
2Ø

ENTER THE DESIRED PROBABILITY OF A SUCCESS FOR THIS SAMPLE (P).
.25

X	P(X)	PLEX	PGEX
Ø	Ø.ØØ317121	Ø.ØØ317121	1.ØØØØØØØØ
1	Ø.Ø2114141	Ø.Ø2431262	Ø.99682879
2	Ø.Ø6694781	Ø.Ø9126Ø43	Ø.97568738
3	Ø.13389562	Ø.225156Ø5	Ø.9Ø873957
4	Ø.18968545	Ø.4148415Ø	Ø.77484395
5	Ø.2Ø233115	Ø.61717265	Ø.5851585Ø
6	Ø.1686Ø929	Ø.78578195	Ø.38282735
7	Ø.1124Ø62Ø	Ø.89818814	Ø.214218Ø5
8	Ø.Ø6Ø88669	Ø.959Ø7483	Ø.1Ø181186
9	Ø.Ø27Ø6Ø75	Ø.98613558	Ø.Ø4Ø92517
1Ø	Ø.ØØ992228	Ø.996Ø5786	Ø.Ø1386442
11	Ø.ØØ3ØØ675	Ø.999Ø6461	Ø.ØØ394214
12	Ø.ØØØ75169	Ø.9998163Ø	Ø.ØØØ93539
13	Ø.ØØØ15419	Ø.99997Ø49	Ø.ØØØ1837Ø
14	Ø.ØØØØ257Ø	Ø.99999619	Ø.ØØØØ2951
15	Ø.ØØØØØ343	Ø.99999961	Ø.ØØØØØ381
16	Ø.ØØØØØØ36	Ø.99999997	Ø.ØØØØØØ39
17	Ø.ØØØØØØØ3	1.ØØØØØØØØ	Ø.ØØØØØØØ3
18	Ø.ØØØØØØØØ	1.ØØØØØØØØ	Ø.ØØØØØØØØ
19	Ø.ØØØØØØØØ	1.ØØØØØØØØ	Ø.ØØØØØØØØ
2Ø	Ø.ØØØØØØØØ	1.ØØØØØØØØ	Ø.ØØØØØØØØ

ENTER NO TO END THE INPUT, YES TO RE-START NEW PROBLEM.
NO

2. Execution Results for Poisson

(i) Exhibit VI

"LAMBDA" = 5

ENTER THE DESIRED "LAMBDA" FOR THE POISSON DISTRIBUTION.
#?
5

X	P(X)	PLEX	PGEX
0	0.00673795	0.00673795	1.00000000
1	0.03368973	0.04042768	0.99326205
2	0.08422434	0.12465202	0.95957232
3	0.14037390	0.26502592	0.87534798
4	0.17546737	0.44049329	0.73497408
5	0.17546737	0.61596065	0.55950671
6	0.14622281	0.76218346	0.38403935
7	0.10444486	0.86662833	0.23781654
8	0.06527804	0.93190637	0.13337167
9	0.03626558	0.96817194	0.06809363
10	0.01813279	0.98630473	0.03182806
11	0.00824218	0.99454691	0.01369527
12	0.00343424	0.99798115	0.00545309
13	0.00132086	0.99930201	0.00201885
14	0.00047174	0.99977375	0.00069799
15	0.00015725	0.99993099	0.00022625
16	0.00004914	0.99998013	0.00006901
17	0.00001445	0.99999458	0.00001987
18	0.00000401	0.99999860	0.00000542
19	0.00000106	0.99999965	0.00000140
20	0.00000026	0.99999992	0.00000035
21	0.00000006	0.99999998	0.00000008
22	0.00000001	1.00000000	0.00000002

ENTER NO TO END THE INPUT, YES TO RE-START NEW PROBLEM.
YES

(ii) Exhibit VII

"LAMBDA" = 10

ENTER THE DESIRED "LAMBDA" FOR THE POISSON DISTRIBUTION.
10

X	P(X)	PLEX	PGEX
0	0.00004540	0.00004540	1.00000000
1	0.00045400	0.00049940	0.99995460
2	0.00227000	0.00276940	0.99950060
3	0.00756665	0.01033605	0.99723060
4	0.01891664	0.02925269	0.98966395
5	0.03783327	0.06708596	0.97074731
6	0.06305546	0.13014142	0.93291404
7	0.09007923	0.22022065	0.86985858
8	0.11259903	0.33281968	0.77977935
9	0.12511004	0.45792971	0.66718032
10	0.12511004	0.58303975	0.54207029
11	0.11373640	0.69677615	0.41696025
12	0.09478033	0.79155648	0.30322385
13	0.07290795	0.86446442	0.20844352
14	0.05207710	0.91654153	0.13553558
15	0.03471807	0.95125960	0.08345847
16	0.02169879	0.97295839	0.04874040
17	0.01276400	0.98572239	0.02704161
18	0.00709111	0.99281350	0.01427761
19	0.00373216	0.99654566	0.00718650
20	0.00186608	0.99841174	0.00345434
21	0.00088861	0.99930035	0.00158826
22	0.00040391	0.99970426	0.00069965
23	0.00017561	0.99987988	0.00029574
24	0.00007317	0.99995305	0.00012012
25	0.00002927	0.99998232	0.00004695
26	0.00001126	0.99999358	0.00001768
27	0.00000417	0.99999775	0.00000642
28	0.00000149	0.99999924	0.00000225
29	0.00000051	0.99999975	0.00000076
30	0.00000017	0.99999992	0.00000025
31	0.00000006	0.99999998	0.00000008
32	0.00000002	0.99999999	0.00000002
33	0.00000001	1.00000000	0.00000001

ENTER NO TO END THE INPUT, YES TO RE-START NEW PROBLEM.
NO

By studying the output of the program from the density functions for the binomial and poisson distributions, a student should be able to obtain a better intuitive understanding of the nature of these two distributions.

REFERENCES

1. Beyer, William H. 1966. Handbook of tables for probability and statistics. The Chemical Rubber Co., Ohio.

2. Chen, Chao-Chen. 1969. Introduction to the theory of statistics. Song Woo Co., Taipei, Taiwan, China.

3. Chirlian, Paul M. 1973. Introduction to FORTRAN IV with timeshare and batch operation. Academic Press, Inc., New York.

4. Cochran, William G. and George W. Snedecor. 1971. Statistical methods. The Iowa State University Press, Ames, Iowa.

5. Freund, John E. 1971. Mathematical statistics. Prentice-Hall, Inc., New Jersey.

6. Hurst, Rex L. Statistical program package. Utah State University Publications, Logan, Utah.

7. Knuth, Donald E. 1971. The art of computer programming. Addison-Wesley Publishing Co., Massachusetts.

8. Li, Jerome C. R. 1959. Introduction to statistical inference. Edwards Brothers, Inc., Michigan.

9. Ostle, Bernard. 1969. Statistics in research. The Iowa State University Press, Ames, Iowa.

10. Schatzoff, Martin. 1968. Applications of time-shared computers in a statistics curriculum. Journal of the American Statistical Association. 3:192-208.

11. Sterling, Theodor D. and Seymour V. Pollack. 1966. Use of the computer to teach introductory statistics. Communications of the Association for Computing Machinery. 4:274-276.

APPENDICES

Appendix A

Program for One Normal Population

The following source statement listing as it ran on a BURROUGHS B6700 is from the program for the sampling experiments from a single normal population.

```
FILE  5=LIU/IN,UNIT=REMOTE,RECORD=14
FILE  6=LIU/OUT,UNIT=REMOTE,RECORD=14
FILE  9=LIU/DSKA,UNIT=DISK,RECORD=14,AREA=30*30,BLOCKING=15
FILE  10=LIU/DSKB,UNIT=DISK,RECORD=14,AREA=15*15,BLOCKING=15
FILE  11=LIU/DSKC,UNIT=DISK,RECORD=14,AREA=15*15,BLOCKING=15

C***  THIS PROGRAM DEALS WITH ONE SINGLE NORMAL POPULATION
C***  AMU=ASSUMED GRAND MEAN FOR THE POPULATION
C***  SIGMA=STANDARD DEVIATION FOR THE POPULATION
C***  IARG=ARGUMENT USED FOR THE RANDOM NUMBER GENERATOR
C***  NK=THE OVER-ALL TOTAL NUMBER OF SAMPLE SIZE
C***  N=NUMBER OF OBSERVATIONS PER SAMPLE
C***  LIST=LIST CONTROL FOR OBSERVATIONS
C***  LST=LIST CONTROL FOR SAMPLE STATISTICS
C***  CIT=TABULAR VALUE OF T'S
C***  CHIL=LEFT TABULAR VALUE OF CHI-SQUARE
C***  CHIR=RIGHT TABULAR VALUE OF CHI-SQUARE

      DIMENSION X(50),IXB(32),IT(32),ICH(32)
      DATA YES/'YES'/
      REWIND  9
      REWIND  10
      REWIND  11
      WRITE (6,500)
  500 FORMAT(/' THIS IS A PROGRAM TO DEMONSTRATE THE DISTRIBUTION OF SAM
     $PLE STATISTICS'/'    FROM ONE NORMAL DISTRIBUTION.')
      WRITE(6,200)
  200 FORMAT(/' ENTER THE DESIRED GRAND MEAN OF THE POPULATION.')
      READ (5,/)  AMU
      WRITE(6,201)
  201 FORMAT(/' ENTER THE DESIRED STANDARD DEVIATION OF THE PUPULATION.
     $')
      READ (5,/) SIGMA
      WRITE (6,205)
  205 FORMAT(/' ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
     $')
      READ (5,/)  IARG
      WRITE (6,202)
  202 FORMAT(/' ENTER THE NUMBER OF SAMPLES WANTED.')
      READ (5,/) NK
      WRITE (6,203)
  203 FORMAT(/' ENTER THE NUMBER OF OBSERVATIONS PER SAMPLE.')
      READ (5,/) N
      WRITE (6,204)
  204 FORMAT(/' HOW FREQUENTLY DO YOU WANT DATA LISTED?'/'    0=NONE,1=AL
     $L,2=EVERY OTHER,3=EVERY THIRD,........ ETC.')
      READ (5,/)  LIST
      WRITE (6,505)
  505 FORMAT(/' HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
     $'/'    0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD,......... ETC.')
      READ (5,/)  LST
      IXSTA=0
      IF (LST.NE.0)  IXSTA=1
      WRITE (6,206)
  206 FORMAT(/' DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION MEAN?
     $'/' INSERT YES OR NO.')
      READ (5,101) ANS
      IXCIT=0
      IF(ANS.EQ.YES)  IXCIT=2
      IF (IXCIT.EQ.0)    GO TO 701
```

```
      WRITE (6,212)
  212 FORMAT(/' ENTER YOUR DESIRED TABULAR VALUE OF "T".')
      READ (5,/)  CIT
  701 WRITE(6,209)
  209 FORMAT(/' DO YOU WANT THE CONFIDENCE INTERVAL FOR POPULATION VARIA
     $NCE?'/' INSERT YES OR NO.')
      READ (5,101)  ANS
      IXCHI=0
      IF (ANS.EQ.YES) IXCHI=4
      IF (IXCHI.EQ.0)   GO TO 4
      WRITE (6,213)
  213 FORMAT(/' ENTER YOUR DESIRED TWO TABULAR VALUES OF "CHI-SQUARE".'
     $/' ONE FOR "LEFT TAIL", THE OTHER "RIGHT TAIL".'//)
      READ (5,/) CHIL,CHIR
    4 IF (LIST.EQ.0)  GO TO  9
      WRITE(6,400)
  400 FORMAT(/' SAMPLE OBSERVATIONS!'/)
    9 DF=FLOAT(N-1)
C***  COMPUTE THEORETICAL STANDARD ERROR(SE)
      SE=SIGMA/SQRT(FLOAT(N))
      DO 5 I=1,32
      IXB(I)=0
      IT(I)=0
    5 ICH(I)=0
C***  ISWA WILL BE USED AS A CONTROL SWITCH
      ISWA=IXCII+IXCHI+IXSTA
C***  REPEAT THIS PROCEDURE TO OBTAIN SAMPLE DATA UNTIL NK TIMES
      DO 15 KK=1,NK
      S=0.0
      SS=0.0
C***  GENERATE THE DATA SET AND COMPUTE THE SUM OF SQUARES
      DO 10 J=1,N
      IX=(RNOR(IARG)*SIGMA+AMU)*100.0
      X(J)=FLOAT(IX)/100.0
      S=S+X(J)
   10 SS=SS+X(J)**2
C***  COMPUTE THE SAMPLE MEAN
      AVE=S/FLOAT(N)
      SS=SS-S*AVE
C***  COMPUTE THE SAMPLE VARIANCE
      VAR=SS/FLOAT(N-1)
C***  COMPUTE THE T-STATISTICS
      T=(AVE-AMU)/SQRT(VAR/FLOAT(N))
C***  COMPUTE CHI-SQ STATISTICS
      CHISQ=SS/SIGMA**2
C***  COMPUTE MEAN-STATISTICS
      IA=(AVE-AMU)*4.0/SE+17.0
C***  ALLOCATE EACH OF MEAN-STATISTICS
      IF(IA.LT.1)  IA=1
      IF(IA.GT.32)  IA=32
      IXB(IA)=IXB(IA)+1
C***  ALLOCATE EACH OF T-STATISTICS
      IB=T*4.0+17.0
      IF(IB.LT.1)  IB=1
      IF(IB.GT.32)  IB=32
      IT(IB)=IT(IB)+1
C***  ALLOCATE EACH OF CHI-SQ STATISTICS
      IC=VAR/SIGMA**2*8.0+1.0
      IF(IC.LT.1)  IC=1
      IF(IC.GT.30)  IC=30
      ICH(IC)=ICH(IC)+1
```

```
      IF (LIST.EQ.0)  GO TO 85
      IF(MOD(KK,LIST).GT.0)  GO TO 85
C***  WRITE OUT THE SAMPLE DATA, IF LIST CONTROL REQUESTED
      WRITE (6,102) KK,(X(J),J=1,N)
  102 FORMAT(I5,8F8.2/(5X,8F8.2))
   85 IF (IXCIT.EQ.0 .AND. IXCHI.EQ.0)  GO TO  96
      IF (IXCIT .EQ. 0) GO TO 94
C***  COMPUTE THE CONFIDENCE INTERVAL FOR MEAN
C***  WRITE LEFT AND RIGHT BOUNDRY OF CONFIDENCE INTERVAL OF MEAN
C**   ONTO DISK FILE 10
      CT=SQRT(VAR/FLOAT(N))
      CIML=AVE-CIT*CT
      CIMR=AVE+CIT*CT
      WRITE (10,402)  KK,CIML,CIMR
  402 FORMAT(1X,I3,5X,F9.2,2X,F9.2)
   94 IF (IXCHI.EQ.0)  GO TO  96
C***  COMPUTE THE CONFIDENCE INTERVAL FOR VARIANCE
      CKI=VAR*FLOAT(N-1)
      CIVL=CKI/CHIR
      CIVR=CKI/CHIL
C***  WRITE LEFT AND RIGHT BOUNDRY OF CONFIDENCE INTERVAL OF VARIANCE
C**   ONTO DISK FILE 11
      WRITE (11,404)  KK,CIVL,CIVR
  404 FORMAT(1X,I3,5X,2F10.3)
   96 IF (LST.EQ.0)  GO TO  15
      IF (MOD(KK,LST) .GT.0)  GO TO  15
C***  WRITE THE SAMPLE STATISTICS ONTO DISK 9
      WRITE (9,302)  KK,AVE,VAR,T,CHISQ
  302 FORMAT(1X,I3,3X,F8.2,2X,F9.3,3X,F9.3,3X,F9.3)
   15 CONTINUE
      REWIND  9
      REWIND  10
      REWIND  11
      IF(ISWA.EQ.0)   GO TO  300
      IF (IXSTA.EQ.0)  GO TO  309
      WRITE(6,401)
  401 FORMAT(/' *********************')
C***  WRITE OUT THE SAMPLE STATISTICS FROM DISK TO PRINTER, IF REQUESTED
      WRITE(6,305)
  305 FORMAT(/'  *   *   XBAR   *    VAR   *    "T"    *  CHI-SQ  *')
  307 READ(9,302,END=309)  KK,AVE,VAR,T,CHISQ
      WRITE(6,302)  KK,AVE,VAR,T,CHISQ
      GO TO  307
  309 IF(IXCIT.EQ.0 .AND. IXCHI.EQ.0)  GO TO   300
      IF (IXCIT.EQ.0)  GO TO  313
C***  WRITE OUT THE CONFIDENCE INTERVAL OF MEAN FROM PRINTER, IF NEEDED
      WRITE(6,311)
  311 FORMAT(/'   *    *     MEAN(L) -- MEAN(R)     *')
  392 READ(10,402,END=313)  KK,CIML,CIMR
      WRITE (6,402)  KK,CIML,CIMR
      GO TO  392
  313 IF (IXCHI.EQ.0)  GO TO  300
C***  WRITE OUT THE CONFIDENCE INTERVAL OF VARIANCE, IF NEEDED
      WRITE (6,315)
  315 FORMAT(/'   *    *     VAR(L) -- VAR(R)     *')
  394 READ(11,404,END=300)  KK,CIVL,CIVR
      WRITE(6,404)  KK,CIVL,CIVR
      GO TO  394
  300 WRITE (6,401)
C***  FOUR FREQUENCY TABLES WILL BE PRINTED RESPECTIVELY, IF REQUESTED
      WRITE (6,108)
```

```
108 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF MEANS.')
    READ (5,101)  ANS
101 FORMAT(A3)
    IF (ANS.NE.YES)  GO TO 21
    WRITE (6,104)
104 FORMAT(/' FREQUENCY TABLE OF MEANS'/
   $' LOWER BOUND   FREQUENCY')
    BND=AMU-4.0*SE
    DO 20 I=1,32
    WRITE (6,105) BND,IXB(I)
 20 BND=BND+SE/4.0
105 FORMAT(E13.6,I8)
 21 WRITE (6,109)
109 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF VARIANCES.')
    READ (5,101) ANS
    IF(ANS.NE.YES)  GO TO 22
    WRITE (6,115)
115 FORMAT(/' FREQUENCY TABLE OF VARIANCES'/
   $' LOWER BOUND   FREQUENCY')
    BND=0.0
    DO 40 I=1,30
    WRITE (6,105) BND,ICH(I)
 40 BND=BND+SIGMA**2/8.0
 22 WRITE (6,110)
110 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF T''S.')
    READ (5,101) ANS
    IF(ANS.NE.YES)  GO TO 23
    WRITE (6,106)
106 FORMAT(/' FREQUENCY TABLE OF ''T'' VALUES'/
   $' LOWER BOUND   FREQUENCY')
    BND=-4.0
    DO 25 I=1,32
    WRITE (6,105)  BND,IT(I)
 25 BND=BND+.25
 23 WRITE (6,111)
111 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF CHI-SQ''S.')
    READ (5,101)  ANS
    IF(ANS.NE.YES)  GO TO  24
    WRITE (6,107)
107 FORMAT(/' FREQUENCY TABLE OF CHI-SQUARE VALUES'/
   $' LOWER BOUND   FREQUENCY')
    BND=0.0
    DO 30 I=1,30
    WRITE (6,105)  BND,ICH(I)
 30 BND=BND+DF/8.0
 24 CONTINUE
 35 STOP
    END
    FUNCTION RNOR(IR)
C***  SUBPROGRAM TO GENERATE AND TO CONVERT ARGUMENT INTO RANDOM NORMAL
C**   NUMBER..
C***  MODIFICATION OF A PROGRAM WRITTEN BY DR. REX L. HURST
C**   UTAH STATE UNIVERSITY LOGAN, UTAH
$SET OWN
    DATA I/0/
$RESET OWN
    IF(I.GT.0)  GO TO  30
 10 X=2.0*RANDOM(IR)-1.0
    Y=2.0*RANDOM(IR)-1.0
```

```
      S=X*X+Y*Y
      IF(S.GE.(1.0))    GO TO 10
      S=SQRT(-2.0*ALOG(S)/S)
      RNOR=X*S
$SET  OWN
      GO2=Y*S
$RESET OWN
      I=1
      GO TO 40
   30 RNOR=GO2
      I=0
   40 RETURN
      END
```

Appendix B

Program for Two Normal Populations

The following source statement listing as it ran on a BURROUGHS B6700 is from the program for the sampling experiments from two normal populations.

```
FILE  5=LIU/IN,UNIT=REMOTE,RECORD=14
FILE  6=LIU/OUT,UNIT=REMOTE,RECORD=14
FILE  9=LIU/DSKA,UNIT=DISK,RECORD=14,AREA=30*30,BLOCKING=15
FILE  10=LIU/DSKB,UNIT=DISK,RECORD=14,AREA=30*30,BLOCKING=15
FILE  11=LIU/DSKC,UNIT=DISK,RECORD=14,AREA=30*30,BLOCKING=15

C***  THIS PROGRAM DEALS WITH TWO NORMAL POPULATIONS
C***  AMUX=ASSUMED GRAND MEAN FOR THE FIRST POPULATION
C***  AMUY=ASSUMED GRAND MEAN FOR THE SECOND POPULATION
C***  SIGMAX=STANDARD DEVIATION FOR THE 1ST POPULATION
C***  SIGMAY=STANDARD DEVIATION FOR THE 2ND POPULATION
C***  IARG=ARGUMENT USED FOR THE RANDOM NUMBER GENERATOR
C***  NK=THE WHOLE TOTAL NUMBER OF SAMPLE SIZE
C***  NX=NUMBER OF OBSERVATIONS PER SAMPLE OF SAMPLE 1
C***  NY=NUMBER OF OBSERVATIONS PER SAMPLE OF SAMPLE 2
C***  LIST=LIST CONTROL FOR OBSERVATIONS
C***  LST=LIST CONTROL FOR SAMPLE STATISTICS
C***  CIT=TABULAR VALUE OF T'S

      DIMENSION X(50),Y(50),IMU(32),IT(32),IIF(32),IVAR(32)
      DATA  YES/'YES'/
      REWIND  9
      REWIND  10
      REWIND   11
      WRITE (6,500)
  500 FORMAT(/' THIS IS A PROGRAM TO DEMONSTRATE THE DISTRIBUTION OF SAM
     $PLE STATISTICS'/'   FROM TWO NORMAL POPULATIONS.')
      WRITE(6,200)
  200 FORMAT(/' ENTER 1ST DESIRED GRAND MEAN FOR THE FIRST POPULATION.')
      READ (5,/)  AMUX
      WRITE (6,201)
  201 FORMAT(/' ENTER 2ND DESIRED GRAND MEAN FOR THE SECOND PUPULATION.'
     $)
      READ (5,/)  AMUY
      WRITE (6,204)
  204 FORMAT(/' ENTER DESIRED STANDARD DEVIATION FOR THE 1ST POPULATION.
     $')
      READ (5,/)  SIGMAX
      WRITE (6,704)
  704 FORMAT(/' ENTER DESIRED STANDARD DEVIATION FOR THE 2ND POPULATION.
     $')
      READ (5,/) SIGMAY
      WRITE (6,207)
  207 FORMAT(/' ENTER A LARGE ODD INTEGER AS A SEED FOR DATA GENERATING.
     $')
      READ (5,/)  IARG
      WRITE (6,205)
  205 FORMAT(/' ENTER THE NUMBER OF SAMPLES WANTED.')
      READ (5,/)  NK
      WRITE (6,202)
  202 FORMAT(/' ENTER NOS. OF OBSERVATION PER SAMPLE FOR 1ST GROUP.')
      READ (5,/)  NX
      WRITE (6,203)
  203 FORMAT(/' ENTER NOS. OF OBSERVATION PER SAMPLE FOR 2ND GROUP.')
      READ (5,/)  NY
      WRITE (6,206)
  206 FORMAT(/' HOW FREQUENTLY DO YOU WANT DATA LISTED?'/'    0=NONE,1=AL
     $L,2=EVERY OTHER,3=EVERY THIRD........ ETC.')
      READ (5,/)  LIST
      WRITE (6,208)
```

```
208 FORMAT(/' HOW FREQUENTLY DO YOU WANT THE SAMPLE STATISTICS LISTED?
   $'/'  0=NONE,1=ALL,2=EVERY OTHER,3=EVERY THIRD........ ETC.')
    READ (5,/)  LST
    IXSTA=0
    IF (LST.NE.0)  IXSTA=1
    WRITE (6,210)
210 FORMAT(/' DO YOU WANT THE CONFIDENCE INTERVAL FOR DIFFERENCE '
   $/'  BETWEEN TWO MEANS?    INSERT YES OR NO.')
    READ (5,101)  ANS
101 FORMAT(A3)
    IXCIT=0
    IF (ANS.EQ.YES)   IXCIT=2
    IF (IXCIT.EQ.0)  GO TO  10
    WRITE(6,212)
212 FORMAT(/' ENTER YOUR DESIRED TABULAR VALUE OF "T".')
    READ (5,/)  CIT
 10 IF (LIST.EQ.0)  GO TO  9
    WRITE (6,214)
214 FORMAT(/' SAMPLE OBSERVATIONS:'/)
  9 DFX=FLOAT(NX-1)
    DFY=FLOAT(NY-1)
C***  COMPUTE THEORETICAL STANDARD ERROR
    SIGMA=(SIGMAX**2/FLOAT(NX))+(SIGMAY**2/FLOAT(NY))
    SE=SQRT(SIGMA)
C***  COMPUTE POOLED SIGMA
    SIGPL=(SIGMAX**2*(NX-1)+SIGMAY**2*(NY-1))/(NX+NY-2)
    FP=(SIGMAX**2)/(SIGMAY**2)
    DO 5 I=1,32
    IMU(I)=0
    IT(I)=0
    IVAR(I)=0
  5 IIF(I)=0
C***  REPEAT THIS PROCEDURE TO OBTAIN SAMPLES UNTIL NK TIMES
    DO 15 KK=1,NK
    SX=0.0
    SY=0.0
    SZ=0.0
    SSX=0.0
    SSY=0.0
    SSZ=0.0
C***  GENERATE 1ST SET OF SAMPLE DATA AND COMPUTE THE SUM OF SQUARES
    DO 13 J=1,NX
    IX=(RNOR(IARG)*SIGMAX+AMUX)*100.0
    X(J)=FLOAT(IX)/100.0
    SX=SX+X(J)
 13 SSX=SSX+X(J)**2
C***  GENERATE 2ND SET OF SAMPLE DATA AND COMPUTE THE SUM OF SQUARES
    DO 11 K=1,NY
    IY=(RNOR(IARG)*SIGMAY+AMUY)*100.0
    Y(K)=FLOAT(IY)/100.0
    SY=SY+Y(K)
 11 SSY=SSY+Y(K)**2
C***  COMPUTE THE AVERAGE FOR EACH OF SAMPLE 1(IX) AND SAMPLE 2(IY)
    AVX=SX/FLOAT(NX)
    AVY=SY/FLOAT(NY)
    Z=AVX-AVY
C***  COMPUTE THE CORRECTED SUM OF SQUARES FOR IX & IY
    SSX=SSX-SX*AVX
    SSY=SSY-SY*AVY
C***  COMPUTE THE VARIANCES FOR EACH OF IX AND IY
    VARX=SSX/FLOAT(NX-1)
```

```
      VARY=SSY/FLOAT(NY-1)
C***  POOLED VARIANCE, IF EQUAL STANDARD DEVIATIONS
      VARPL=(VARX*(NX-1)+VARY*(NY-1))/(NX+NY-2)
C***  SAMPLE VARIANCE FOR DIFFERENCE BETWEEN TWO MEANS, IF POOLED
      VARTS=VARPL/FLOAT(NX)+VARPL/FLOAT(NY)
C***  SAMPLE VARIANCE FOR DIFFERENCE BETWEEN TWO MEANS, IF NOT POOLED
      VARTD=(VARX/FLOAT(NX))+(VARY/FLOAT(NY))
C***  "F"=RATIO OF TWO VARIANCES
      F=VARX/VARY
C***  TABULATE EACH OF THE SAMPLE STATISTICS INTO ONE OF THE SAMPLE
C**   DISTRIBUTIONS
C***  ALLOCATE EACH OF U-STATISTICS
      IA=Z*4.0/SE+15.0
      IF(IA.LT.1)  IA=1
      IF (IA.GT.32)  IA=32
      IMU(IA)=IMU(IA)+1
C***  ALLOCATE EACH OF THE POOLED SAMPLE VARIANCES
      IV=VARPL/SIGPL*8.0+1.0
      IF (IV.LT.1)  IV=1
      IF (IV.GT.30) IV=30
      IVAR(IV)=IVAR(IV)+1
C***  ALLOCATE EACH OF T-STATISTICS
      IF (SIGMAX.EQ.SIGMAY)    GO TO  72
      T=Z/SQRT(VARTD)
      GO TO  73
   72 T=Z/SQRT(VARTS)
   73 IB=T*4.0+17.0
      IF (IB.LT.1)  IB=1
      IF (IB.GT.32)  IB=32
      IT(IB)=IT(IB)+1
C***  ALLOCATE EACH OF F-STATISTICS
      IC=(VARX/VARY)*(8.0/FP)+1.0
      IF (IC.LT.1)  IC=1
      IF (IC.GT.30)  IC=30
      IIF(IC)=IIF(IC)+1
      IF (LIST.EQ.0)  GO TO  85
      IF (MOD(KK,LIST).GT.0) GO TO 85
C***  WRITE OUT THE SAMPLE DATA IF CONTROL LIST REQUESTED
      WRITE (6,216)  KK,(X(J),J=1,NX)
  216 FORMAT(1X,I5,8F8.2/(6X,8F8.2))
      WRITE (6,216)  KK,(Y(J),J=1,NY)
      WRITE (6,401)
  401 FORMAT(/' >>>>>>>>>>>>>>>>>>>>>'/)
   85 IF (IXCIT.EQ.0 .AND. IXSTA.EQ.0)  GO TO  15
C***  COMPUTE THE CONFIDENCE INTERVAL FOR DIFFERENCE BETWEEN TWO MEANS
      IF (IXCIT.EQ.0)  GO TO  87
      IF (SIGMAX.NE.SIGMAY)    GO TO  50
C***  COMPUTE LEFT AND RIGHT BOUNDRY OF CONFIDENCE INTERVAL OF
C**   DIFFERENCE BETWEEN TWO MEANS, IF POOLED VARIANCE
      CIML=Z-CIT*SQRT(VARTS)
      CIMR=Z+CIT*SQRT(VARTS)
      GO TO  52
C***  COMPUTE THE SAME CONFIDENCE INTERVAL, IF VARIANCE NOT POOLED
   50 CIML=Z-CIT*SQRT(VARTD)
      CIMR=Z+CIT*SQRT(VARTD)
C***  WRITE OUT THE CONFIDENCE INTERVAL ONTO DISK 11
   52 WRITE (11,304)  KK,CIML,CIMR
  304 FORMAT(1X,I3,3X,F9.2,2X,F9.2)
   87 IF (LST.EQ.0)  GO TO  15
      IF (MOD(KK,LST) .GT.0 )    GO TO  15
C***  PRINT VARTS IF SIGMAX=SIGMAY, VARTD IF NOT
```

```
   91 IF (SIGMAX.NE.SIGMAY)  GO TO  89
C***  WRITE OUT THE SAMPLE STATISTICS ONTO DISK 9
      WRITE(9,301) KK,AVX,AVY,Z,VARX,VARY,VARTS
  301 FORMAT(1X,I3,F8.2,1X,F8.2,5X,F8.2,2X,2F9.3,3X,F9.3)
      GO TO  92
C***  WRITE OUT THE SAMPLE STATISTICS ONTO DISK 9
   89 WRITE(9,301)  KK,AVX,AVY,Z,VARX,VARY,VARTD
C***  WRITE OUT THE SAMPLE STATISTICS ONTO DISK 10
   92 WRITE (10,302)  KK,T,F,VARPL
  302 FORMAT(1X,I3,F7.2,1X,F7.2,1X,F9.3)
   15 CONTINUE
      REWIND  9
      REWIND  10
      REWIND  11
      IF (IXSTA.EQ.0 .AND. IXCIT.EQ.0)  GO TO  300
      IF (IXSTA.EQ.0)  GO TO  394
      WRITE (6,305)
  305 FORMAT(/'  # * MEAN-X * MEAN-Y * DIFF OF MEANS * VAR-X * VAR-Y * V
     $AR OF DIFF *')
  389 IF (SIGMAX.NE.SIGMAY)  GO TO  397
C***  WRITE THE SAMPLE STATISTICS FROM DISK TO PRINTER, IF REQUESTED
  391 READ(9,301,END=392)  KK,AVX,AVY,Z,VARX,VARY,VARTS
      WRITE (6,301) KK,AVX,AVY,Z,VARX,VARY,VARTS
      GO TO  391
C***  WRITE VARTD FROM DISK TO PRINTER, IF REQUESTED
  397 READ(9,301,END=392)  KK,AVX,AVY,Z,VARX,VARY,VARTD
      WRITE(6,301)  KK,AVX,AVY,Z,VARX,VARY,VARTD
      GO TO  397
C***  WRITE OUT THE RESTS OF SAMPLE STATISTICS
  392 WRITE (6,306)
  306 FORMAT(//'  # *  "T"  *  "F"  * POOL VAR *')
  393 READ(10,302,END=394)  KK,T,F,VARPL
      WRITE(6,302)  KK,T,F,VARPL
      GO TO  393
  394 IF (IXCIT.EQ.0)  GO TO  300
C***  WRITE OUT THE CONFIDENCE INTERVAL FROM PRINTER, IF REQUESTED
      WRITE (6,307)
  307 FORMAT(/'   #    *    MEAN(L) -- MEAN(R)    *')
  395 READ(11,304,END=300)  KK,CIML,CIMR
      WRITE(6,304)  KK,CIML,CIMR
      GO TO  395
  300 AVZ=SZ/FLOAT(NK)
      SSZ=SSZ-SZ*AVZ
      VARZ=SSZ/FLOAT(NK-1)
      WRITE (6,217)
  217 FORMAT(/' ***************'/)
      ZMU=AMUX-AMUY
C***  FOUR FREQUENCY TABLES WILL BE PRINTED RESPECTIVELY, IF REQUESTED
      WRITE (6,108)
  108 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF DIFFERENCE BETWE
     $EN MEANS')
      READ (5,101)  ANS
      IF (ANS.NE.YES)  GO TO 21
      WRITE (6,104)
  104 FORMAT(/' FREQUENCY TABLE OF DIFFERENCE BETWEEN MEANS.'/
     $' LOWER BOUND    FREQUENCY')
      BND=ZMU-4.0*SE
      DO 20 I=1,32
      WRITE (6,105) BND,IMU(I)
   20 BND=BND+SE/4.0
  105 FORMAT(1X,E13.6,I8)
```

```
 21 WRITE (6,109)
109 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF POOLED VARIANCES
   $.')
    READ (5,101)  ANS
    IF (ANS.NE.YES)  GO TO 22
    WRITE (6,115)
115 FORMAT(/'  FREQUENCY TABLE OF POOLED VARIANCES'/
   $' LOWER BOUND   FREQUENCY')
    BND=0.0
    DO 40 I=1,30
    WRITE (6,105) BND,IVAR(I)
 40 BND=BND+SIGPL/8.0
 22 WRITE (6,110)
110 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF "T",  DIFF OF TW
   $O MEANS.')
    READ (5,101)  ANS
    IF(ANS.NE.YES)  GO TO  23
    WRITE (6,106)
106 FORMAT(/'  FREQUENCY TABLE OF "T", DIFF OF TWO MEANS'/
   $' LOWER BOUND   FREQUENCY')
    BND=-4.0
    DO 25 I=1,32
    WRITE (6,105) BND,IT(I)
 25 BND=BND+.25
 23 WRITE (6,111)
111 FORMAT(//' ENTER YES, TO PRINT FREQUENCY TABLE OF "F", RATIO OF TW
   $O VARIANCES.')
    READ (5,101)  ANS
    IF (ANS.NE.YES)  GO TO 24
    WRITE (6,107)
107 FORMAT(/'  FREQUENCY TABLE OF "F", RATIO OF TWO VARIANCES'/
   $' LOWER BOUND    FREQUENCY')
    BND=0.0
    DO 30 I=1,30
    WRITE (6,105) BND,IIF(I)
 30 BND=BND+FP/8.0
 24 CONTINUE
 35 STOP
    END
    FUNCTION RNOR(IR)
C***  SUBPROGRAM TO GENERATE AND TO CONVERT ARGUMENT INTO RANDOM NORMAL
C**   NUMBER..
C***  MODIFICATION OF A PROGRAM WRITTEN BY DR. REX L. HURST
C**   UTAH STATE UNIVERSITY LOGAN, UTAH

$SET OWN
      DATA I/0/
$RESET OWN
      IF(I.GT.0)  GO TO 30
   10 X=2.0*RANDOM(IR)-1.0
      Y=2.0*RANDOM(IR)-1.0
      S=X*X+Y*Y
      IF(S.GE.(1.0))  GO TO 10
      S=SQRT(-2.0*ALOG(S)/S)
      RNOR=X*S
$SET OWN
      GO2=Y*S
$RESET OWN
      I=1
      GO TO 40
```

```
30 RNOR=G02
   I=0
40 RETURN
   END
```

Appendix C

Program for the Binomial Distribution

The following source statement listing as it ran on a BURROUGHS B6700 is from the program for the binomial distribution.

```
FILE  5=LIU/IN,UNIT=REMOTE,RECORD=14
FILE  6=LIU/OUT,UNIT=REMOTE,RECORD=14

C***  THIS PROGRAM COMPUTES THE DISCRETE DENSITY FUNCTION OF BINOMIAL
C***  DOUBLE PRECISION ARE USED THROUGH ALL ALPHABETS EXCEPT (I,J,K,L,
C**   M,N, AND Y) IN THE PROGRAM
C***  USERS HAVE TO DECIDE WHAT SUITABLE PARAMETERS TO USE
C***  N DENOTES THE SAMPLE SIZE OF TRIALS
C***  P DENOTES THE PROBABILITY OF A SUCCESS

      IMPLICIT REAL*8(A-H,O-X,Z)
      DIMENSION QI(100)
      IRD=5
      IPR=6
      WRITE(IPR,500)
  500 FORMAT(/' THIS IS A PROGRAM TO COMPUTE THE PROBABILITY DISTRIBUTIO
     $N OF '/'      BINOMIAL FOR DIFFERENT TRIALS.')
    5 WRITE (IPR,100)
  100 FORMAT(/' ENTER THE DESIRED SAMPLE SIZE (N).')
      READ (IRD,/,END=20)  N
    6 WRITE (IPR,101)
  101 FORMAT(/' ENTER THE DESIRED PROBABILITY OF A SUCCESS FOR THIS SAMP
     $LE (P).')
      READ (IRD,/)  P
      IF (P.LT.0 .OR. P.GT.1)  GO TO  20
C***  INITIAL THREE VALUES FOR THE FIRST TIME
      Q=1.0-P
      QI(N)=1.0
      NM=N-1
C***  ACCUMULATE THE PROBABILITY PORTION OF SUCCESS AND FAILURE
      DO 10 I=1,NM
      II=N-I
   10 QI(II)=QI(II+1)*Q
C***  INITIAL THREE BEGINNING PROBABILITIES FOR FX,PLEX AND PGEX
      FX=QI(1)*Q
      PLEX=FX
      PGEX=1.0
      A=1.0
      I=0
C***  WRITE OUT THE FIRST PROBABILITIES FOR EACH OF THEM
      WRITE (IPR,105) I,FX,PLEX,PGEX
  105 FORMAT('    X',8X,'P(X)',11X,'PLEX',11X,'PGEX'/I4,3F15.8)
      DO 15  I=1,N
C***  PGEX=PROBABILITY OF GREATER THAN & EQUAL TO
      PGEX=PGEX-FX
C***  COMPUTE THE BINOMIAL COEFFICIENT PORTION
      A=A*DFLOAT(N-I+1)*P/DFLOAT(I)
C***  COMPUTE THE BINOMIAL PROBABILITY, FX
      FX=A*QI(I)
C***  PLEX=PROBABILITY OF LESS THAN & EQUAL TO
      PLEX=PLEX+FX
   15 WRITE(IPR,106) I,FX,PLEX,PGEX
  106 FORMAT(I4,3F15.8)
      WRITE(IPR,108)
  108 FORMAT(/' ENTER NO TO END THE INPUT, YES TO RE-START NEW PROBLEM.'
     $)
      READ (IRD,200)  YANS
  200 FORMAT(A3)
      IF (YANS.EQ.'NO')  GO TO  20
      GO TO  5
```

```
20 STOP
   END
```

Appendix D

Program for the Poisson Distribution

The following source statement listing as it ran on a BURROUGHS B6700 is from the program for the poisson distribution.

```
FILE  5=LIU/IN,UNIT=REMOTE,RECORD=14
FILE  6=LIU/OUT,UNIT=REMOTE,RECORD=14

C***  THIS PROGRAM COMPUTES THE DISCRETE DENSITY FUNCTION OF POISSON
C***  LAMBDA IS THE UNIQUE PARAMETER TO BE ENTERED IN THE PROGRAM
C***  DOUBLE PRECISION IS USED THROUGH THE ALL ALPHABETS EXCEPT(I,J,K,L,
C**   M,N, AND Y) IN THE PROGRAM

      IMPLICIT REAL*8(A-H,O-X,Z)
      IRD=5
      IPR=6
C***  DELTA IS THE SMALLEST PROBABILITY WE CAN GET
      DELTA=.5D-8
      WRITE (IPR,200)
  200 FORMAT(/' THIS IS A PROGRAM TO COMPUTE THE PROBABILITY DISTRIBUTIO
     $N OF'/'      POISSON FOR DIFFERENT LEVELS OF LAMBDA.')
    1 WRITE (IPR,202)
  202 FORMAT(/' ENTER THE DESIRED "LAMBDA" FOR THE POISSON DISTRIBUTION.
     $')
      READ(IRD,/,END=10)  ALAM
C***  INITIAL THE VALUES
C***  FX=PROBABILITY OF POISSON
      FX=1.0/DEXP(ALAM)
      PLX=FX
      PGX=1.0
      I=0
      AI=1.0
C***  PLEX=PROBABILITY OF LESS THAN & EQUAL TO
C***  PGEX=PROBABILITY OF GREATER THAN & EQUAL TO
      WRITE(IPR,100) I,FX,PLX,PGX
  100 FORMAT(/'    X',8X,'P(X)',11X,'PLEX',11X,'PGEX'/I4,3F15.8)
    2 PGX=PGX-FX
      FX=FX*ALAM/AI
      PLX=PLX+FX
      I=I+1
      AI=AI+1.0
      WRITE (IPR,102)  I,FX,PLX,PGX
  102 FORMAT(I4,3F15.8)
C***  LET COMPUTER STOP AT THE SMALLEST PROBABILITY
      IF (1.0-PLX .GT. DELTA)  GO TO 2
      WRITE(IPR,104)
  104 FORMAT(/' ENTER NO TO END THE INPUT, YES TO RE-START NEW PROBLEM.'
     $)
      READ(IRD,106)  YANS
  106 FORMAT(A3)
      IF (YANS.EQ.'NO')  GO TO  10
      GO TO  1
   10 STOP
      END
```

VITA

Chien-Hwa Liu

Candidate for the Degree of

Master of Science

Report: Computer Programs Supporting the Teaching of Statistics

Major Field: Applied Statistics

Biographical Information:

Personal Data: Born at Fukien Province, China, January 7, 1942, son of Mr. and Mrs. C. S. Liu; married Linda W. Na May 10, 1968.

Education: Attended elementary school in Taipei, Taiwan; graduated from Provincial Chien-kuo High School in 1961; received the Bachelor of Art degree from National Taiwan University, with a major in Business Administration, in June 1965; completed requirements for the Master of Science degree at Utah State University in 1973.

Professional Experience: 1973-present, statistical consultant, Utah State University, Logan, Utah; 1969-1971, I/O expediter and computer programmer trainee, Jewel Food Store Co., Melrose Park, Illinois; 1966-1968, accountant and statistic staff, Overseas Chinese Banking Corp., Taipei, Taiwan.