Continuous Statistical Distribution Curve Fitting and Analysis Tool

1987

Gerald Robie
University of Central Florida

Find similar works at: http://stars.library.ucf.edu/rtd

University of Central Florida Libraries http://library.ucf.edu

Part of the Industrial Engineering Commons

STARS Citation

Robie, Gerald, "Continuous Statistical Distribution Curve Fitting and Analysis Tool" (1987). Retrospective Theses and Dissertations. 5027.
http://stars.library.ucf.edu/rtd/5027

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
CONTINUOUS STATISTICAL DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

BY

GERALD ROBIE
B.S., University of Toledo, 1964

RESEARCH REPORT

Submitted in partial fulfillment of the requirements for the degree of Master of Science in the Graduate Studies Program of the College of Engineering University of Central Florida Orlando, Florida

Spring Term 1987
ABSTRACT

This paper reports on the implementation and utilization of a micro-computer based simulation modeling verification and validation tool. The interactive software tool, written in BASIC, computes and displays the frequency distribution of a given set of input data, computes appropriate parameters for a continuous statistical distribution which the user selects as likely to represent the population from which the input data sample was obtained, performs several goodness-of-fit analyses on the resultant distribution and provides sensitivity analysis capability on the input data, distribution type selection and distribution parameters.
ACKNOWLEDGEMENTS

The author wishes to acknowledge the following individuals, without whose help and support this effort would never have come to fruition:

Dr. Jose Sepulveda, of the University of Central Florida, research committee chairman and supervisor of this research report activity, for his continual expert technical guidance and encouragement,

Judy Robie, Helen Robie, Donna Robie and Dan Robie for their editorial, technical and moral support,

Dr. William Swart and Dr. Gary Whitehouse, both of the University of Central Florida, for their helpful comments and suggestions and technical contributions to this effort, and

Ms. Candie Stewart for her proof-reading assistance.
TABLE OF CONTENTS

INTRODUCTION ....................................................... 1

PURPOSE OF THE TOOL .............................................. 3
   Simulation Model Input Data Verification .................. 3
   Endogenous Activity and Event Verification ............... 5
   Model Validation ............................................... 5

IMPLEMENTATION ................................................... 7
   Data Input ..................................................... 8
   Data Manipulation ............................................. 9
   Probability Distribution Bar Chart Display ............. 10
   Computation of Selected Distribution Parameters .......... 12
   Manual Selection of Parameters ............................. 19
   Determination of Adequacy of Fit ........................... 20

A TYPICAL APPLICATION OF THE TOOL ............................ 25

RATIONALE FOR THE DISTRIBUTIONS IMPLEMENTED .............. 29
   Continuous Uniform Distribution ............................ 29
   Normal Distribution ........................................... 29
   Exponential Distribution ..................................... 30
   Gamma Distribution ............................................ 31
   Weibull Distribution .......................................... 32
   Beta Distribution .............................................. 32

POTENTIAL ENHANCEMENTS TO THE TOOL ......................... 34
   Improvements .................................................. 34
   Expansion of Scope ............................................ 35

IMPLEMENTATION DIFFICULTIES ENCOUNTERED ...................... 38

SUMMARY .................................................................. 40

APPENDICES .......................................................... 41
   A. BASIC Source Code Listing ................................. 42
   B. Menu Screens ................................................. 85
   C. Example Displays ............................................. 93
   D. Example Problem Results ................................... 99

LIST OF REFERENCES ............................................... 111
INTRODUCTION

Prior to the middle of the twentieth century, scientists, engineers and other systems analysis practitioners generally applied deterministic analysis techniques to the study of systems and the interactions of the various systems components (Hahn and Shapiro 1967). As technological complexity has increased and the understanding of systems analysis has matured, it has become increasingly obvious that this deterministic approach is woefully inadequate. Systems, which on the surface appear identical, can produce significantly differing results due to such variable factors as component tolerances and environmental fluctuations.

The contemporary engineering solution to this complication is to utilize probabilistic analysis techniques to accommodate the stochastic nature of the system under analysis. When the modern engineer accepts this premise, it becomes natural for him to speak in such terms as "the probability" that the system will perform within specified limits or "the statistical distribution" of a specified component parameter.

As this is a relatively new concept which is developing rapidly, the availability of micro-computer tools which will assist the systems analyst in the
verification and validation of his stochastic model systems modeling and analysis endeavors is quite limited. Southern Technology Applications Center (STAC) technical publications searches of the National Aeronautical and Space Association database, the Information Access Company Computer database and the Engineering Info, Inc. Compendex database using the keywords:

"curve fitting" OR "goodness of fit" OR "statistical regression"

AND

"simulation" OR "model" OR "statistical distribution"

identified only one article dealing with the subject (Hansen 1985). Two surveys of commercially available statistics software, conducted by Carpenter, Deloria and Morganstein (1984) and by Espeillac (1987) identified seventy-five packages from forty-nine different vendors. None of these packages contains all of the functions required to adequately and efficiently perform the entire verification and validation process in one integrated entity. It is the objective of the analysis tool described in this report to, in some small fashion, contribute to the filling of this void.
PURPOSE OF THE TOOL

This section describes applications of the "Continuous Statistical Distribution Curve Fitting and Analysis Tool" in the systems modeling and analysis environment. It should be appreciated that this tool is also applicable in other disciplines which also employ statistical analysis techniques.

During the development of systems models, the scientist or engineer is faced with the formidable task of ensuring that his model accurately represents its real world counterpart. Banks and Carson (1984) describe this as a two-step process of verification and validation. Verification is the process of comparing the conceptual model to the model implementation (usually computer software). Validation is the process of ensuring that the implemented model is an accurate representation of the real world system.

Simulation Model Input Data Verification

A key element of any simulation model is the input data, which is the driving force for the model. Banks and Carson (1984) describe four steps in the development of a valid model of input data. The first step is the collection of raw data from a real world situation. The
second step is the identification of the underlying statistical distribution of this data. In some cases, theoretical justifications exist for applying specific statistical distributions to represent physical phenomena. In these instances, the choice of a distribution to represent a physical system is motivated by an understanding of the nature of the underlying phenomenon and is verified by the available data. However, in many simulation modeling situations, this understanding does not exist and the use of frequency distribution graphs provides the practitioner with visual guidance in this selection process.

The third step in the development of the input data model is the estimation of the parameters that characterize the selected distribution. Finally, the assumed distribution and selected parameters are tested for goodness of fit to the raw data.

The "Continuous Statistical Distribution Curve Fitting and Analysis Tool" facilitates this process by accepting the raw data and generating a frequency distribution display of the data. By observing and manipulating this display, the user can postulate a distributional assumption. The user can then request the software to determine appropriate parameters for his selected distribution. Finally, the user can visually and
statistically test the goodness of fit of the selected distribution and parameters to the raw data.

**Endogenous Activity and Event Verification**

Model input data are generally exogenous events, that is events in the environment surrounding the system being modeled, such as customer arrivals or product demand. Within the system, endogenous activities and events occur, such as providing service to a customer or equipment breakdown. Many of these endogenous activities and events exhibit stochastic attributes. The "Continuous Statistical Distribution Curve Fitting and Analysis Tool" can be used in a manner similar to that described above for input data analysis to verify the statistical distribution and parameter assumptions applied to these stochastic endogenous activities and events.

**Model Validation**

After the conceptual model has been implemented and the implementation verified, the model designer must validate the accuracy of the model's representation of the real world system. This is generally an iterative process of comparing the model outputs to the real world results and making adjustments to the model as required.

In order to effectively accomplish this calibration, corresponding output raw data is collected from the real world system during the input raw data collection effort.
The "Continuous Statistical Distribution Curve Fitting and Analysis Tool" can be used to characterize the distribution and parameters of this raw output data. The model is then exercised by running the raw input data through it and collecting the results. These results are compared to the collected real world system raw output data characterization using the tool. If this comparison passes the tool's goodness of fit tests, the model is presumed to be valid. If not, the iterative model calibration procedure is utilized.
IMPLEMENTATION

This section describes the specific capabilities of the "Continuous Statistical Distribution Curve Fitting and Analysis Tool" and the implementation of these functions in software.

The tool is menu driven and reference to this section by users familiar with similar micro-computer tools will be limited to understanding the underlying principles applied to the more technical portions of the tool.

The tool is implemented in the BASIC programming language and does not utilize any of the enhancements to the language (such as graphics or modern macros). This implementation was chosen to maximize the probability that the tool will function on the many differing types of micro-computers commercially available today.

This section is written with the assumption that the reader has previously operated the tool or will be doing so while reading this. A copy of all menus is given in Appendix B.

All functions are implemented with the assumption that the quantity of raw data points is in the range of 10 to 100. These were assumed to be reasonable limits for most applications. Appropriate modifications to statements
within the source code can be accomplished by the more computer knowledgeable user to adjust these limits.

When the tool is activated, an abstract of the tool's capabilities is presented. Upon continuing, the user is presented with the Data Input menu.

**Data Input**

Data can be manually input to the tool via the keyboard or data which was input during a previous computer session and was "saved" at that time can be recalled from diskette.

If the "Input new data." option is selected, the user is prompted to provide the quantity of data points in his sample. Subsequently, the user inputs the data and is afforded the opportunity to make corrections at appropriate intervals.

When all data has been input and verified, the user is provided the opportunity to save the data to diskette. If the user chooses to save the data, he is prompted to provide the filename under which the data is to be saved. The software checks to ensure that the filename is not a duplication. If a duplication is detected, the user is provided the opportunity to either overwrite the previous data, to select a new filename, or to abort the save process.
If the "Use previously saved data." option is selected from the Data Input menu, the user is prompted to provide the filename of the data. Appropriate error checks are provided to protect against trying to recall data from a non-existent file.

Throughout this and all subsequent functions, appropriate syntax and vocabulary safeguards are provided to protect the user from potential data corruption or loss and, also, to minimize user confusion and frustration.

Regardless of which method of data input is selected, after the function is successfully accomplished, the user is presented with the Data Manipulation menu.

**Data Manipulation**

Within the Data Manipulation environment, the user is afforded the opportunity to "View the data." This allows him to verify that this indeed is the data upon which he desires to perform the analysis.

In most instances, it is assumed that the user will use the "Print Screen" function of his micro-computer to obtain hard copy records of the analysis. However, the input data presentation provided when the "View the data." option is selected is in a format amenable to video display viewing but not optimum for hardcopy records. Therefore, a "Print the data." selection is provided which results in a more acceptable printed record.
Quite often in the analysis of statistically distributed data, one or several data points are relatively far away from the vicinity where the rest of the data is concentrated (Hines and Montgomery 1980). These "outliers" may be valid data points or they may be contaminated. In either event, the system modeling practitioner will want to evaluate the effect of the outliers on the results of the statistical analysis. The "Continuous Statistical Distribution Curve Fitting and Analysis Tool" Data Manipulation environment allows three data editing options. Specific data points can be modified by using the "Change some observations." option. The database can be expanded or reduced via the "Add some observations." and "Delete some observations." choices. In every instance where database modifications does occur, the user is afforded the opportunity to save the edited data to diskette for future retrieval.

After all data review and editing is completed, the user selects the "Go on to the data analysis program." option which transfers operation to the Data Analysis menu.

Probability Distribution Bar Chart Display

Prior to entering the Data Analysis environment, the input data is maintained as string variables input by the user. Since the data analysis environment requires sorted, numeric data, the conversion from string variables to
numeric data and sequential sort occur when entering this area of the program. The process can be quite time-consuming when operating in the BASIC interpretive mode. Delays of several seconds can occur with highly scrambled data and a large database.

The first option afforded the user in the Data Analysis environment is "Display a bar graph of the data." This function permits the user to view a frequency distribution histogram in order to facilitate the selection of an appropriate statistical distribution type.

The user is afforded the option of allowing the software to size the bar graph or of manually selecting the quantity of intervals, interval width and first (lowest) interval upper limit. Since this is a reselectable option, the user can allow the computer to size the graph initially, in order to become familiar with the general shape of the distribution of his data, and subsequently he can manipulate the limits in order to better determine of which distribution population shapes the data is likely to be a sample. An example bar graph display is given in Appendix C.

The user is referred to Hahn and Shapiro (1967); Hastings and Peacock (1974); Ireson (1966) or Law and Kelton (1982) for graphical representations of the distributions encompassed by the "Continuous Statistical
Distribution Curve Fitting and Analysis Tool in order to facilitate this determination.

The selection should be limited to the Uniform, Normal, Exponential, Gamma, Weibull or Beta continuous distributions because the present implementation of subsequent functions in this tool is limited to these six choices.

Computation of Selected Distribution Parameters

The second option afforded to the user in the Data Analysis menu is "Fit a statistical distribution to the data." After the user has determined which distribution type(s) he wishes to investigate as possible fits to his data, he should select this option.

At this time the Distribution Fitting menu is presented to the user. He can then select one of the following distributions:

- Continuous Uniform
- Normal
- Exponential
- Gamma
- Weibull
- Beta

Selecting one of these alternatives will result in the computation of appropriate parameters and a display of the results of these computations. The display also provides
the mean, variance, skewness and kurtosis of the sample data and the corresponding calculated attributes of a continuous distribution with the computed parameters. An example of the computed parameters display for a Beta Distribution is given in Appendix C.

Uniform Distribution

Selecting the "Fit a Uniform Distribution to the data." option will result in the computation of the Maximum Likelihood Estimator values of the minimum value (a) and maximum value (b) of the distribution as given in Law and Kelton (1982). The probability density function then becomes (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[
f(x) = \frac{1}{b-a} \quad \text{for } a \leq x \leq b
\]
\[
f(x) = 0 \quad \text{otherwise}
\]

For the Uniform Distribution, the equations for skewness and kurtosis are (Hastings and Peacock 1974, Ireson 1966):

\[
skewness = 0
\]
\[
kurtosis = 9/5
\]

Normal Distribution

Selection of the "Fit a Normal Distribution to the data." option will provide the Maximum Likelihood Estimator
values of the mean ($\mu$) and variance ($\sigma^2$) as given in Hastings and Peacock (1974), Ireson (1976) and Law and Kelton (1982). The probability density function is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

$$f(x) = \frac{1}{(2\pi \sigma^2)^{1/2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad \text{for all } x$$

The equations for skewness and kurtosis for the Normal Distribution are (Hastings and Peacock 1974; Ireson 1966):

skewness = 0  
kurtosis = 3

Exponential Distribution

The "Fit an Exponential Distribution to the data." option provides the Maximum Likelihood Estimator value of the mean ($1/\lambda$) as given in Hastings and Peacock (1974), Ireson (1976) and Law and Kelton (1982). The probability density function is (Hastings and Peacock 1974; Ireson 1976; Law and Kelton 1982):

$$f(x) = \lambda \exp[-\lambda x] \quad \text{for } x \geq 0$$

$$f(x) = 0 \quad \text{otherwise}$$
The skewness and kurtosis equations for the Exponential Distribution are (Hastings and Peacock 1974; Ireson 1976):

\[
\text{skewness} = 2 \\
\text{kurtosis} = 9
\]

**Gamma Distribution**

"Fit a Gamma Distribution to the data." selection provides the scale parameter \(b\) and shape parameter \(c\) using the Method of Matching Moments as given in Hastings and Peacock (1974) and Ireson (1976). The probability density function is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[
f(x) = \frac{(x/b)^{c-1}\exp[-x/b]}{b\Gamma(c)} \text{ for } x>0 \\
f(x) = 0 \text{ for } x\leq 0
\]

where \(\Gamma(c)\) is the gamma function of the shape parameter \(c\) as given in Beyer (1979) and Burington (1957). In the software, the gamma function is implemented using the identity (Beyer 1979):

\[
\Gamma(n) = (n-1)\Gamma(n-1)
\]

in conjunction with a diskette stored lookup table containing gamma values for \(n\) in the range of zero to one.
The skewness and kurtosis equations of the Gamma Distribution are (Hastings and Peacock 1974; Ireson 1966):

\[
\text{skewness} = \frac{2}{c^{1/2}}
\]
\[
\text{kurtosis} = 3 + \frac{6}{c}
\]

Weibull Distribution

Selecting the "Fit a Weibull Distribution to the data." option produces the Maximum Likelihood Estimator values of the scale parameter (b) and the shape parameter (c) as given in Hastings and Peacock (1974), Ireson (1966) and Law and Kelton (1982). This is a complex computation which is implemented in the software by using the iterative procedure suggested by Law and Kelton (1982). The probability density function is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[
f(x) = (c x^{c-1}/b^c) \exp\left[-(x/b)^c\right] \quad \text{for } x \geq 0
\]
\[
f(x) = 0 \quad \text{for } x < 0
\]
The equations for skewness and kurtosis for the Weibull Distribution are (Ireson 1966):

\[
\text{skewness} = \frac{\Gamma(1+3/c) - 3\Gamma(1+2/c)\Gamma(1+1/c) + 2[\Gamma(1+1/c)]^2}{[\Gamma(1+2/c) - \Gamma(1+1/c)]^{3/2}}
\]

\[
\text{kurtosis} = \frac{\Gamma(1+4/c) - 4\Gamma(1+3/c)\Gamma(1+1/c) + 6\Gamma(1+2/c)\Gamma(1+1/c)^2 - 3[\Gamma(1+1/c)]^3}{[\Gamma(1+2/c) - \Gamma(1+1/c)]^2}
\]

Beta Distribution

"Fit a Beta Distribution to the data." provides the two shape parameters (v and w) for a Beta Distribution in the range of zero to one. The Method of Matching Moments is utilized to compute these parameters as given in Hahn and Shapiro (1967) and Hastings and Peacock (1974). Since the basic distribution always exists only in the range of zero to one, the computed parameters are for a normalized distribution. The software also provides the multiplication (m) and offset (o) values by which the probability density function has been adjusted.
The normalized probability density function is (Hahn and Shapiro 1967; Ireson 1966; Law and Kelton 1982):

\[ f(x) = x^{v-1}(1-x)^{w-1}/\beta(v, w) \quad \text{for } 0 \leq x \leq 1 \]
\[ f(x) = 0 \quad \text{otherwise} \]

where \( \beta(v, w) \) is the beta function of the shape parameters (v and w). In the software, this is implemented using the identity (Hastings and Peacock 1974):

\[ \beta(v, w) = \Gamma(v)\Gamma(w)/\Gamma(v+w) \]

Prior to computing the shape parameters, the raw data is normalized by the software as follows:

- The offset value \( o \) is defined as the minimum data point \( D_{\text{min}} \).
- The multiplication value \( m \) is defined as the range of the data \( D_{\text{max}} - D_{\text{min}} \).
- Each observation is adjusted using the equation:

\[ \text{Observation(adjusted)} = \frac{\text{Observation(raw)} - o}{m} \]

The adjusted data set thus determined is always in the range of zero to one; and this satisfies the Beta Distribution range constraint.

When using the computed parameters to analyze a particular situation, the user must normalize the data
points of interest similarly. For example, if the user desires to compute the probability that an observation will lie within the range of y to z by integrating the probability density function, the limits must be adjusted as follows:

\[
\text{lower limit} = \frac{y - 0}{m} \quad \text{upper limit} = \frac{z - 0}{m}
\]

The skewness and kurtosis equations for the Beta Distribution are (Hahn and Shapiro 1967; Ireson 1966):

\[
\text{skewness} = \frac{2(w-v)(v+w+1)^{1/2}}{(v+w+2)(vw)^{1/2}}
\]

\[
\text{kurtosis} = \frac{3(v+w)(v+w+1)(v+1)(2w-v)}{vw(v+w+2)(v+w+3)} + \frac{v(v-w)}{v+w}
\]

**Manual Selection of Parameters**

Upon continuing from the "Best Fit Parameters" display, the user is afforded the opportunity to "Select your own parameters for the chosen distribution." If this option is selected, the user is prompted to input the parameters of his choosing. In order to assist in this
operation, the computed parameters are displayed as reminders.

This function is provided to allow the user to perform sensitivity analyses on the distribution attributes and parameter iterations in conjunction with the adequacy of fit tests described later in this report.

After all parameters are selected, a comparison display is provided by the tool. This screen displays the computed and selected parameters as well as the mean, variance, skewness and kurtosis of the resultant distributions. The mean, variance, skewness and kurtosis of the raw input data are also displayed. An example "Distribution Parameter Comparison" display is given in Appendix C.

The equations used for these computations are the same as those used in the "Computation of Selected Distribution Parameters" discussed previously.

**Determination of Adequacy of Fit**

Two adequacy of fit analysis functions are provided in the "Continuous Statistical Distribution Curve Fitting and Analysis Tool." The first is a "Display a bar graph with selected curve overlayed." option which permits the user to make a subjective judgement of fit adequacy while viewing a graphical representation of the raw data and selected continuous distribution. The second, "Perform a Chi-Square
Goodness of Fit Analysis. Option permits a statistical evaluation of fit adequacy.

Each of these tests can be performed on computer determined parameters for the distribution type previously selected by the user or on parameters selected by the user.

These two tests, as well as the manual parameter selection function and distribution type selection, can be performed repeatedly and interchangeably, thus maximizing the sensitivity analysis possibilities available to the user.

Bar Graph with Curve Overlayed

If this option is chosen prior to manually selecting distribution parameters, the overlayed curve is determined by the computed parameters; otherwise the selected parameters are used to determine the overlayed curve. An example bar chart with distribution curve overlayed is given in Appendix C.

In order to generate the overlayed curves, cumulative probability data for each distribution type is determined as detailed below.

Uniform Distribution. The cumulative distribution function used to determine the probabilities for the Uniform
Distribution overlayed curve is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[ F(x) = 0 \quad \text{for } x < a \]
\[ F(x) = \frac{x-a}{b-a} \quad \text{for } a \leq x \leq b \]
\[ F(x) = 1 \quad \text{for } b < x \]

**Normal Distribution.** The method used to determine the cumulative probabilities for the Normal Distribution overlayed curve is to normalize the \( x \) values using the equation (Grant and Leavenworth 1972; Hines and Montgomery 1980; Miller and Freund 1977):

\[ z = \frac{x-\mu}{\sigma} \]

These values are then used to retrieve the cumulative probabilities from a diskette located file containing a "Cumulative Standard Normal Distribution" table (Hines and Montgomery 1980). Appropriate interpolation procedures are applied.

**Exponential Distribution.** The cumulative distribution function used to determine the probabilities for the Exponential Distribution overlayed curve is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[ F(x) = 0 \quad \text{for } x < 0 \]
\[ F(x) = 1 - \exp[-\lambda x] \quad \text{for } x \geq 0 \]
Gamma Distribution. The method used to determine the probabilities for the Gamma Distribution overlayed curve is to integrate the probability density function using iterative integration techniques.

Weibull Distribution. The cumulative density function used to determine the probabilities for the Weibull Distribution overlayed curve is (Hastings and Peacock 1974; Ireson 1966; Law and Kelton 1982):

\[ F(x) = \begin{cases} 
0 & \text{for } x < 0 \\
1 - \exp\left[-\left(\frac{x}{b}\right)^c\right] & \text{for } x \geq 0 
\end{cases} \]

Beta Distribution. The method used to determine the probabilities for the Beta Distribution overlayed curve is to integrate the probability density function using iterative integration techniques.

Chi-Square Goodness of Fit Test

Selecting the "Perform a Chi-Square Goodness of Fit Analysis." option results in the computer determining the Chi-Square values for the computed and selected distribution relative to the raw data. An example "Chi-Square Goodness of Fit" display is given in Appendix C.
The number of class intervals used in this computation is determined as the integer rounded value of the equation:

\[ k = \frac{(n^{1/2} + n/5)}{2} \]

where \( n \) is the quantity of data points in the sample. The number of class intervals equation was determined as an average compromise between two classical statistical schools of thought on the subject. One school (Hahn and Shapiro 1967; Hines and Montgomery 1980; Miller and Freund 1977) advocates selecting class intervals so that each interval contains no less than five observations. The second group of practitioners (Banks and Carson 1984; Nelson 1983) suggest that the interval quantity should be approximately the square root of the sample size.

Since all of the implemented distributions are continuous, class intervals which are equal in probability rather than equal in range are used for the Chi-Square test as recommended by Banks and Carson (1984), Kindall and Stuart (1979) and Mann and Wald (1942). The cumulative distributions functions and iterative integration techniques utilized to determine the equally probable intervals are those used in the "Bar Graph With Curve Overlayed" function described previously.

The critical Chi-Square values at significance levels of 0.01 and 0.05 for the selected class intervals are determined from a computer memory resident lookup table.
A TYPICAL APPLICATION OF THE TOOL

This section discusses a typical application of the "Continuous Statistical Distribution Curve Fitting and Analysis Tool" in the systems modeling and analysis environment. The example problem is "Example 9.3" in the textbook *Discrete-Event System Simulation* by Banks and Carson (1984):

Life tests were performed on a random sample of 50 PDP-11 electronic chips at 1.5 times the nominal voltage, and their lifetime (or time to failure) in days was recorded:

<table>
<thead>
<tr>
<th>Lifetime (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>79.919</td>
</tr>
<tr>
<td>3.027</td>
</tr>
<tr>
<td>6.769</td>
</tr>
<tr>
<td>18.387</td>
</tr>
<tr>
<td>144.695</td>
</tr>
<tr>
<td>0.941</td>
</tr>
<tr>
<td>0.624</td>
</tr>
<tr>
<td>0.590</td>
</tr>
<tr>
<td>3.217</td>
</tr>
<tr>
<td>1.961</td>
</tr>
<tr>
<td>0.062</td>
</tr>
<tr>
<td>0.021</td>
</tr>
<tr>
<td>1.192</td>
</tr>
<tr>
<td>43.565</td>
</tr>
<tr>
<td>17.967</td>
</tr>
<tr>
<td>3.371</td>
</tr>
<tr>
<td>3.148</td>
</tr>
<tr>
<td>1.928</td>
</tr>
<tr>
<td>31.764</td>
</tr>
<tr>
<td>1.005</td>
</tr>
<tr>
<td>1.008</td>
</tr>
</tbody>
</table>

The objective is to identify a distribution which reasonably approximates the data. Copies of the micro-computer displays discussed in this example are given in Appendix D.

To begin operation of the "Continuous Statistical Distribution Curve Fitting and Analysis Tool," insert a diskette containing the compiled version of the software.
into the default diskette drive of the micro-computer. The diskette must contain, as a minimum, the following files:

NORMAL.DAT
GAMMA.DAT
STATFIT.EXE
EXAMPLE9.3

The problem data has been provided on the original diskette in the latter file to reduce the tedium of this exercise. A printout of this data is included in Appendix D.

The tool is invoked by entering the command STATFIT[cr] onto the keyboard. The user is presented with the introductory abstract which he may read and then press any key to continue as prompted. At this point the user selects option 2, "Use previously saved data.", and enters EXAMPLE9.3[cr] in response to the filename prompt. This retrieves the data from the diskette.

After the data is retrieved, the user selects option 6, "Go on to the data analysis program.", when prompted for input. This takes the user to the "Data Analysis" environment. At this point, the user selects option 1, "Display a bar graph of the data." Initially, the user answers N to the manual sizing prompt option. The resultant bar chart display is as given in Appendix D.

After viewing this histogram, the user can press any key to continue. It is probably desirable to observe a histogram with more intervals than the default values, so
the user again selects option 1. However, this time the user answers the "Bar Graph Sizing" questions as follows:

Do you wish to manually size the graph (Y or N) ? Y
How many frequency intervals should the graph have (3 to 20) ? 20
How large should each interval width be ? 2[cr]
What is the upper limit of the lowest range ? 2[cr]

The resultant expanded bar chart display is given in Appendix D. Three points are above the range of 36-38, however, this is a practical resolution to observe the general shape of the distribution.

Upon observing this display, the user might reasonably conclude that an Exponential Distribution would approximate the data. The user presses any key to continue and then selects option 2, "Fit a statistical distribution to the data." He then selects option 3, "Fit an Exponential Distribution to the data." The resultant "Best Fit Parameter Display For An Exponential Distribution" is given in Appendix D.

In actual practice, the user may decide to use the "Print Screen" option on his micro-computer at this time to obtain a hard copy of this display for future reference. The user then presses any key and selects option 2, "Display a bar graph of the data with selected curve overlayed." In order to observe the results of the
computer suggested parameter, the user answers N to the manual input of parameters prompt. The histogram with the Exponential Distribution overlayed is given in Appendix D.

The user might subjectively conclude that this is a reasonable fit. By pressing any key and then selecting option 3, "Perform a Chi-Square Goodness of Fit Analysis.", he can objectively test this hypothesis. As the resultant display indicates, the Chi-Square test rejects the fit hypothesis.

The user presses any key and then selects option 4, "Fit a different statistical distribution to the data." Since the raw data is life test data, the user might next try selecting option 5, "Fit a Weibull Distribution to the data." The resultant display, as well as overlayed curve and Chi-Square analysis are given in Appendix D.

This time the Chi-Square test cannot reject the fit hypothesis and the user concludes that the raw data can be approximated by a Weibull Distribution with scale parameter equal to 6.235281 and shape parameter equal to .5259377.

These parameters can now be tested for sensitivity by pressing any key, selecting option 1, "Select your own parameters for the chosen distribution.", and then performing a Chi-Square test on his selection. The results with scale parameter equal to 6 and shape parameter equal to .5 are given in Appendix D for comparison purposes.
RATIONALE FOR THE DISTRIBUTIONS IMPLEMENTED

This section discusses the reasons that the Uniform, Normal, Exponential, Gamma, Weibull and Beta continuous distributions were selected for this implementation and suggests typical applications of each by systems modeling and analysis practitioners.

Continuous Uniform Distribution

The Continuous Uniform distribution can be used as a first approximation in instances when data is known to be continuous and randomly distributed, but not enough information is available to better define the distribution. Quite often it is applied as a first approximation to interarrival or service times (Banks and Carson 1984).

Normal Distribution

The Normal distribution is the most frequently used statistical model. Until about the year 1900, it was assumed to be the basic distribution to which naturally occurring randomly varying functions conformed. The availability of statistical tests ultimately invalidated this universality assumption. However, the central limit theorem, which states that the means of equally sized independent samples taken from a given population will be
normally distributed, justifies the use of this
distribution in many modeling applications.

Empirical evidence indicates that the Normal
distribution is applicable to many physical variables such
as measurements on living organisms, molecular velocities
in a gas, scores on an intelligence test, average
temperatures in a given locality and random electrical
noise (Hahn and Shapiro 1967). The Normal distribution has
been recently applied to approximating ability test results
(Tsutakawa 1983) and to predicting the results of material
corrosion experiments (Nicholls and Hancock 1983).

**Exponential Distribution**

The Exponential distribution is the basic time
dependent statistical model. It is used extensively to
model time dependent phenomena which exhibit constant
hazard functions. It is a special case of the Gamma
distribution with shape parameter (c) equal to one (Hahn
and Shapiro 1967; Hastings and Peacock 1974). It is also a
special case of the Weibull distribution with shape
parameter (c) equal to one (Hahn and Shapiro 1967; Hastings
and Peacock 1974).

The Exponential distribution is used extensively to
model interarrival times when arrivals are completely
random and to model lifetimes of components which fail
catastrophically (Banks and Carson 1984). It is often
applied as the first approximation to climatological data (Tuller and Brett 1984), to the prediction of time-dependent physical phenomena (Nicholls and Hancock 1983; Billman, Morgan and Windle 1979), to the estimation of suspended particulate concentrations (Taylor, Jakeman and Simpson 1986) and to usage life of components which have been exposed to initial burn-in (to eliminate infant mortality) and where preventative maintenance removes parts before wear out (Hahn and Shapiro 1967).

Gamma Distribution

The Gamma distribution is a basic distribution of statistics for variables which are bounded at one end. It provides many diverse representations for random variables in the range of zero to plus infinity. As such, it is used quite often in time to failure models (Hahn and Shapiro 1967).

The Gamma distribution, with shape parameter (c) equal to a whole number, is the Erlang distribution subset of the Gamma family. This function can be used to model the probability that c independent events occur, assuming the events take place at a constant rate (Banks and Carson 1984). This attribute makes it the ideal model for time to failure in redundant component systems, time between inventory restocking, and time between recalibration requirements for equipment requiring calibration after c
uses (Hahn and Shapiro 1967). The Gamma distribution has also been applied in modeling pilot response delays to beacon collision avoidance systems (Billman, Morgan and Windle 1979).

**Weibull Distribution**

The Weibull distribution is another distribution which exists only in the range of zero to plus infinity. The Exponential distribution is a subset of the Weibull distribution with shape parameter \( c \) equal to one. Since a shape parameter equal to one yields a constant hazard function, the Exponential distribution has limited applicability in time to failure models. However, the hazard function and probability density function of the Weibull distribution have a wide variety of shapes, making it an ideal candidate for variable hazard rate time to failure models (Hahn and Shapiro 1967). Typical applications are passive electronic components and mechanical component life models.

The Weibull distribution has also been applied to such other models as time dependent metal corrosion phenomena (Nicholls and Hancock 1983) and wind speed variations in a given locality (Tuller and Brett 1984).
Beta Distribution

The Beta distribution is a basic distribution of statistics for variables bounded on both ends. Although the fundamental distribution exists only in the range of zero to one, use of data normalization techniques make this distribution applicable to any raw data set. The Uniform distribution is a subset of the Beta distribution with shape parameters (v and w) both equal to one.

The Beta distribution can be used in instances of incomplete or limited bounded data because of its great variety of distributional forms (Banks and Carson 1984). Typical applications are the population distribution between the lowest and highest values, distribution of manufacturing process daily percent yield, and modeling elapsed times in scheduling applications such as PERT (Hajek 1984).
POTENTIAL ENHANCEMENTS OF THE TOOL

This section describes potential enhancements to the "Continuous Statistical Distribution Curve Fitting and Analysis Tool." Functional, implementation and scope expansion possibilities are discussed.

Improvements

The present implementation includes several functions which execute rather slowly because of the tedious program algorithms used. Although these do not appear to be disconcerting when using the compiled version of the software, they are somewhat frustrating when the program is executed in the interpretive environment, as would be the case for the user with a micro-processor which is not compatible with the compiled program. The slower functions include the data sort routine (the present implementation utilizes a modified bubble sort) and the linear integration to obtain cumulative probability data for the Gamma and Beta distribution. The iterative computation involved in the determination of Weibull distribution parameters is also somewhat slow.

The Method of Matching Moments is used to compute the parameters for the Gamma and Beta distributions because the equations are bounded and, thus, are simpler to implement
with faster micro-computer performance than the implementation of the Maximum Likelihood Estimators would be. The equations for the Maximum Likelihood Estimators are available in Hahn and Shapiro (1967), Hastings and Peacock (1974) and Ireson (1966). If a faster numerical integration technique were utilized, the Maximum Likelihood Estimator parameter computations might be possible in reasonable micro-computer execution times.

Expansion of Scope

The six distributions selected for this implementation were chosen because they are the most likely candidates in much model simulation and analysis work. Several other distributions which might be added to the implementation are:

- Lognormal - This distribution has many applications in climatological modeling, such as in cloud formation analysis (Lopez 1977) and air pollution analysis (Hardison 1978; Taylor, Jakeman and Simpson 1986).

- Folded Normal - This distribution is applicable in statistical process control (SPC) analysis of such single end bounded phenomena as shaft roundness or processes where the sign of the deviation is lost before data is taken (such as dropping welded right angled parts into a box prior to measurement.)

- Pareto - This distribution is applicable to the a,b,c inventory cycle count policy and, in fact, to any vital-few versus trivial-many decision analysis. It has numerous applications in the writer's chosen profession of Quality Assurance.
A companion software package could also be developed to encompass discrete statistical distributions. In order of priority, the following distributions should be considered:

- **Binomial** - This distribution would be applicable in sampling with replacement processes, digital instrument accuracy, probability theory and many other discrete processes.

- **Hypergeometric** - This distribution is applicable in sampling without replacement statistical process control activities.

- **Poisson** - This distribution applies to queueing analysis, variables sampling statistical process control and medical, bacterial culture analysis (Hahn and Shapiro 1967).

- **Geometric**
- **Pascal**
- **Negative Binomial**
- **Multinomial**

No objective distribution type selection aid is included in the initial implementation of the tool. A skewness/kurtosis nomograph, proposed by E. S. Pearson (Hahn and Shapiro 1967; Hines and Montgomery 1980), could be utilized to provide a function whereby the software would suggest a distribution type which would best fit the data. In addition, an "expert systems" type question and answer session could be provided to assist the novice practitioner in selecting the appropriate distribution in
situations where theoretical justification for a specific
distribution exists.

Only one objective goodness of fit test has been
included in this implementation of the tool. Since the
Chi-Square test is not robust, it might be desirable to
include other statistical tests. The Kolmogorov-Smirnov
test was considered, but rejected because it is not
applicable when distribution parameters have been estimated
from the data (Banks and Carson 1984). However, this test
would be applicable in situations where the tool is being
used to test the goodness of fit of data to a distribution
where no inference of distribution parameters has been made
from the data. In addition, more robust tests, such as the
distribution specific W tests (Hahn and Shapiro 1967) would
be desirable enhancements to this first implementation of
the tool.
IMPLEMENTATION DIFFICULTIES ENCOUNTERED

This section briefly discusses difficulties encountered researching and implementing the "Continuous Statistical Distribution Curve Fitting and Analysis Tool."

The major obstacle to an efficient implementation of the tool was the inconsistency of notation prevalent in the field of statistics. While researching the Weibull distribution, the writer determined that one source chose to label his parameters "b" and "c" while a second used "α" and "β" for the same parameters. The first source's "b" equated to the second source's "β" and "c" equated to "α". Further confounding the practitioner, a third source uses parameters "α" and "β" which relate to the first source's parameters as follows:

\[ α = b^c \]
\[ β = c \]

Other examples of this type of inconsistency were encountered but this was the most difficult to reconcile. It would seem to be desirable to establish a "standards committee" in the statistics discipline to address this issue, if, in fact, standards of notation have not yet been established.
Another minor implementation difficulty encountered was the inconsistent operation of BASIC interpreters among different micro-processor manufacturers. In order to maximize the applicability of the "Continuous Statistical Distribution Curve Fitting and Analysis Tool," the source code was executed on three major manufacturer's micro-processors. Some minor, annoyance level differences were identified which necessitated some "creative" adjustments to make the software compatible with all three micro-processors.
SUMMARY

The modern scientist, engineer and systems modeling and simulation practitioner is often faced with tedious computational and analysis efforts. With the ready availability of micro-computers, it seems desirable to relegate much of these more mundane activities to this device in order to enhance the practitioner's productivity. This paper has discussed a software tool which addresses this hypothesis.

The "Continuous Statistical Distribution Curve Fitting and Analysis Tool" provides the practitioner with a facility to readily analyze stochastic situations. This paper has presented the need for such a tool and has discussed its implementation in detail. The use of the tool to resolve a typical simulation modeling data analysis problem has been presented.

With regard to the initial implementation of the tool, the reasons for the selection of specific distributions for inclusion have been discussed. In the hope of assisting the practitioner who might desire to enhance this initial implementation, potential enhancement areas have been proposed and potential pitfalls have been identified.
APPENDIX A

BASIC SOURCE CODE LISTING
1 ' Time multiplier that determines delay time that "flash"
2 ' remain on the screen
3 TIMEMULT=3
4 CLS
5 LOCATE 5,12.0
6 PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
7 LOCATE 8,26
8 PRINT"UNIVERSITY OF CENTRAL FLORIDA"
9 LOCATE 10,10
10 PRINT"Department of Industrial Engineering and Management Science"
11 FOR I=1 TO 2000*TIMEMULT:NEXT
12 CLS
13 LOCATE 5,12.0
14 PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
15 LOCATE 8,1
16 PRINT"This program computes and displays the frequency distribution of a given set of"
17 PRINT"data. You then selects the continuous distribution which, in your estimation,"
18 PRINT"best approximates the displayed histogram. The appropriate estimator parameters"
19 PRINT"for the selected distribution are computed. The computed distribution is"
20 PRINT"displayed, superimposed over the input data frequency distribution plot. You"
21 PRINT"can then request that a Chi-Square or goodness of fit test be performed on the"
22 PRINT"computed distribution. At any time, you may change the continuous distribution"'
23 PRINT"type selection and thereby cause a recomputation of the distribution estimators"
24 PRINT"and replot of the results."
25 LOCATE 20,27,1
26 PRINT"PRESS ANY KEY TO CONTINUE ;
27 A$=INPUT$(1)
28 ' Data input and change routines
29 DIM D$(100),D(100),BOUND(20),BOUNDP(20),RANGE(20)
30 ADD=0
31 CLS
32 LOCATE 1,12.0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
33 LOCATE 3,35:PRINT"DATA INPUT"
34 LOCATE 6,25:PRINT"1. Input new data."
35 LOCATE 7,25:PRINT"2. Use previously saved data."
36 LOCATE 8,25:PRINT"3. Quit the entire process."
37 LOCATE 15,30,1:PRINT"TYPE IN YOUR CHOICE ;
38 A$=INPUT$(1)
39 IF A$="0" AND A$<"4" THEN 42
40 LOCATE 18,29:PRINT"ENTRY MUST BE 1,2 OR 3"
41 GOTO 37
42 A=VAL(A$)
43 ON A
44 ' New data input.
45 I=1
46 CLS
47 LOCATE 1,12.0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
48 LOCATE 3,33:PRINT"NEW DATA ENTRY"
49 LOCATE 5,1,1:PRINT"How many data points in your sample (10 to 100) ? ";
50 A=0
51 N$=""
52 A$=INPUT$(1)
53 IF A$=CHR$(13) THEN 64
54 IF A$<>CHR$(8) THEN 59
55 IF A=0 THEN 52
56 A=A-1
57 N$=LEFT$(N$,A)
58 GOSUB 515
59 IF A$<"0" OR A$="9" THEN 52
60 PRINT A$;
61 A=A+1
62 N$=N$+A$
63 IF A<3 THEN 52
64 N=VAL(N$)
65 IF N<9 AND N<101 THEN 68
66 LOCATE 8,13:PRINT"THE QUANTITY OF DATA POINTS MUST BE BETWEEN 10 AND 100"
67 LOCATE 5,51:PRINT":":GOTO 49
68 LOCATE 8,13:PRINT SPC(54)
69 J=1:SAVED=0 : N4=I
70 LOCATE 7,1
71 PRINT"What is the value of observation";I:" ";
72 GOSUB 364
73 D$(I)=N$
74 I=I+1:IF I>N THEN 76
75 J=J+1:IF J<16 THEN 71
76 LOCATE 23,26:PRINT"ALL DATA OK (Y,N or Quit) ? ":A$=INPUT$(1)
77 IF A$="Q" OR A$="q" THEN 99
78 IF A$="Y" OR A$="y" THEN 93
79 IF A$<"N" AND A$<"n" THEN 76
80 LOCATE 23,26:PRINT SPC(28)
81 LOCATE 23,1:PRINT"Which observation is incorrect? ";
82 GOSUB 487
83 IF E=1 THEN 80
84 N2=38
85 IF N1>9 THEN N2=39
86 IF N1>99 THEN N2=40
87 LOCATE N1-N4+7,N2:PRINT SPC(20)
88 LOCATE N1-N4+7,N2:
89 GOSUB 364
90 D$(N1)=N$
91 LOCATE 23,1:PRINT SPC(25)
92 GOTO 76
93 IF I=N THEN 102
94 CLS
95 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
96 LOCATE 3,33:PRINT"NEW DATA ENTRY"
97 LOCATE 5,1,1:PRINT"How many data points in your sample (10 to 100) ?":N
98 GOTO 69
99 IF ADD=0 THEN 30
100 N=I-1
101 GOTO 164
102 ' Data to file routine
103 CLS
104 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
105 LOCATE 3,35:PRINT"DATA SAVE"
106 LOCATE 5,1,1:PRINT"Do you wish to save this data for later use (Y or N) ? ";
107 A$=INPUT$(1)
108 IF A$="N" OR A$="n" THEN 165
109 IF A$<>"Y" AND A$<>"y" THEN 107
110 PRINT"Y"
111 LOCATE 7,1,1:PRINT"On what disk and under what filename and extension"
112 LOCATE 8,6:PRINT"do you wish to save the data"
113 LOCATE 9,6:PRINT"(<drive designation:><filename><.extension>) ? ";
114 GOSUB 391
115 IF E=1 THEN LOCATE 9,53:PRINT SPC(14):GOTO 113
116 ON ERROR GOTO 125
117 OPEN "I",#1,FILENAME$
118 ON ERROR GOTO 0
119 LOCATE 11,1:PRINT"That file already exists, should I overwrite it (Y or N) ? ";
120 A$=INPUT$(1)
121 IF A$="N" OR A$="n" THEN CLOSE: GOTO 102
122 IF A$<>"Y" AND A$<>"y" THEN 120
123 PRINT "Y"
124 GOTO 127
125 RESUME 126
126 ON ERROR GOTO 0
127 CLOSE
128 OPEN"O",#1,FILENAME$
129 PRINT#1,N;
130 FOR I=1 TO N
131 PRINT#1,D$(I);",";
132 NEXT I
133 SAVED=1
134 CLOSE
135 GOTO 165
136 ' Data from file routine
137 CLS
138 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
139 LOCATE 3,34:PRINT"PREVIOUSLY SAVED DATA RETRIEVAL"
140 LOCATE 5,1,1:PRINT"On what disk and under what filename and extension"
141 LOCATE 6,6:PRINT"is the data saved"
142 LOCATE 7,6:PRINT"(<drive designation:><filename><.extension>) ? ";
143 GOSUB 391
144 IF E=1 THEN LOCATE 7,53:PRINT SPC(14):GOTO 142
145 ON ERROR GOTO 155
146 OPEN "I",#1,FILENAME$
147 SAVED=1
148 INPUT#1,N
149 FOR I=1 TO N
150 INPUT#1,D$(I)
151 NEXT I
152 CLOSE
153 ON ERROR GOTO 0
154 GOTO 164
155 RESUME 156
156 ON ERROR GOTO 0
157 CLOSE
158 LOCATE 9,1:PRINT"That file does not exist,"
159 LOCATE 10,6:PRINT"do you want to try another file name (Y or N)? ";
160 A$=INPUT$(1)
161 IF A$="N" OR A$="n" THEN 30
162 IF A$<"Y" AND A$<"y" THEN 160
163 GOTO 136
164 ' Data view, print and change routines
165 CLS
166 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
167 LOCATE 3,31:PRINT"DATA MANIPULATION"
168 LOCATE 6,25:PRINT"1. View the data."
169 LOCATE 7,25:PRINT"2. Print the data."
170 LOCATE 8,25:PRINT"3. Change some observations."
171 LOCATE 9,25:PRINT"4. Delete some observations."
172 LOCATE 10,25:PRINT"5. Add some observations."
173 LOCATE 11,25:PRINT"6. Go on to the data analysis program."
174 LOCATE 12,25:PRINT"7. Start all over again."
175 LOCATE 13,25:PRINT"8. Quit the entire process."
176 LOCATE 17,30,1:PRINT"TYPE IN YOUR CHOICE ";
177 A$=INPUT$(1)
178 IF A$>"0" AND A$<"9" THEN 181
179 LOCATE 20,25:PRINT"ENTRY MUST BE BETWEEN 1 AND 8"
180 GOTO 176
181 A=VAL(A$)
182 ON A GOTO 183,194,211,266,321,529,354,354
183 ' Data view routine
184 I=1
185 CLS
186 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
187 LOCATE 3,34:PRINT"DATA REVIEW"
188 GOSUB 467
189 LOCATE 23,23,1:PRINT"PRESS C TO CONTINUE OR Q TO QUIT ";:A$=INPUT$(1)
190 IF A$="Q" OR A$="q" THEN 164
191 IF A$<"C" AND A$<"c" THEN 189
192 IF DONE=0 THEN 185
193 GOTO 164
194 ' Data print routine
195 I=1:J=1
196 LOCATE 17,20:PRINT"PRESS ANY KEY WHEN THE PRINTER IS READY ";
197 A$=INPUT$(1)
198 LPRINT " CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
199 LPRINT:LPRINT " PRINTOUT OF DATA STORED IN FILE ";FILENAME$
201 LPRINT "The value of observation";I;"is ";D$(I)
202 I=I+1
203 J=J+1
204 IF I>N THEN 209
205 IF J<51 THEN 201
206 J=1
207 LPRINT CHR$(12)
208 GOTO 198
209 LPRINT CHR$(12)
210 GOTO 164
211 ' Data change routine
212 I=1
213 N4=I
214 CLS
215 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
216 LOCATE 3,30:PRINT"CHANGE OBSERVATIONS"
217 IF I>1 THEN 225
218 LOCATE 5,1,1:PRINT"Do you know the observation number(s) of the data
219 LOCATE 6,25:PRINT"which you want to change (Y or N) ? ";
220 A$=INPUT$(1)
221 IF A$="Y" OR A$="y" THEN 248
222 IF A$<>"N" AND A$<>"n" THEN 219
223 LOCATE 5,1:PRINT SPC(50)
224 LOCATE 6,25:PRINT SPC(40)
225 GOSUB 467
226 LOCATE 23,22,1:PRINT"CHANGE ANY OF THESE (Y,N or Quit) ? ";
227 A$=INPUT$(1)
228 IF A$="Q" OR A$="q" THEN 232
229 IF A$="Y" OR A$="y" THEN 234
230 IF A$<>"N" AND A$<>"n" THEN 227
231 IF DONE=0 THEN 213
232 IF SAVED=0 THEN 102
233 GOTO 164
234 SAVED=0
235 LOCATE 23,27:PRINT SPC(36)
236 LOCATE 23,1:PRINT"Which observation do you wish to change ? ";
237 GOSUB 467
238 IF E=1 THEN 235
239 N2=31
240 IF N1>9 THEN N2=32
241 IF N1>99 THEN N2=33
242 LOCATE N1-N4+7,N2:PRINT SPC(20)
243 LOCATE N1-N4+7,N2:
244 GOSUB 364
245 D$(N1)=N$
246 LOCATE 23,1:PRINT SPC(21)
247 GOTO 226
248 PRINT"Y"
249 N4=1
250 I=N+1
LOCATE 8,1:PRINT"Which observation do you wish to change? ";
GOSUB 487
IF E=1 THEN LOCATE 8,42:PRINT SPC(5):GOTO 251
LOCATE 10,1:PRINT"What is the value of observation";N1;"? ";
GOSUB 364
D$(N1)=N$
SAVED=0
LOCATE 12,1:PRINT"Change any more observations (Y or N)? ";
A$=INPUT$
IF A$="N" OR A$="n" THEN 232
IF A$<"Y" AND A$<"y" THEN 259
LOCATE 12,1:PRINT SPC(40)
LOCATE 10,1:PRINT SPC(60)
LOCATE 8,1:PRINT SPC(50)
GOTO 251
' Data delete routine
I=1
N4=1
CLS
LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 3,30:PRINT"DELETE OBSERVATIONS"
IF I>1 THEN 280
LOCATE 5,1,1:PRINT"Do you know the observation number(s) of the data you want to delete (Y or N)? ";
A$=INPUT$
IF A$="Y" OR A$="y" THEN 301
IF A$<"N" AND A$<"n" THEN 275
LOCATE 5,1:PRINT SPC(50)
LOCATE 6,25:PRINT SPC(40)
GOSUB 467
LOCATE 23,22,1:PRINT"DELETE ANY OF THESE (Y,N or Quit)? ";
A$=INPUT$
IF A$="Q" OR A$="q" THEN 287
IF A$="Y" OR A$="y" THEN 289
IF A$<"N" AND A$<"n" THEN 282
IF DONE=0 THEN 268
IF SAVED=0 THEN 102
GOTO 164
SAVED=0
LOCATE 23,30:PRINT SPC(36)
LOCATE 23,1:PRINT"Which observation do you wish to delete? ";
GOSUB 487
IF E=1 THEN 290
N=N-1
FOR J=N1 TO N
D$(J)=D$(J+1)
NEXT J
IF N=N4 THEN 164
I=N4
GOTO 269
301 PRINT "Y"
302 N4=1
303 I=N4+1
304 LOCATE 8,1:PRINT "Which observation do you wish to delete? "
305 GOSUB 487
306 IF E=1 THEN LOCATE 8,42:PRINT SPC(5):GOTO 304
307 N=N-1
308 FOR J=N1 TO N
309 D$(J)=D$(J+I)
310 NEXT J
311 SAVED=O
312 LOCATE 12,1:PRINT "WARNING: Your previous deletion has shifted the observation number of all points above it in the data table!"
313 LOCATE 13,10:PRINT "Delete any more observations (Y or N)? ";
314 A$=INPUT$(1)
315 IF A$="N" OR A$="n" THEN 232
316 IF A$<>"y" AND A$<>"y" THEN 314
317 LOCATE 10,1:PRINT SPC(40)
318 LOCATE 8,1:PRINT SPC(50)
320 GOTO 304
321 ' Data add routine
322 A=O
323 CLS
324 LOCATE 1,12,0:PRINT "CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
325 LOCATE 3,32:PRINT "ADD OBSERVATIONS"
326 LOCATE 5,1:PRINT "There are currently";N;"observations entered."
327 LOCATE 6,1,1:PRINT "How many observations do you wish to add? ";
328 A=O
329 N$=""
330 A$=INPUT$(1)
331 IF A$=CHR$(13) THEN 342
332 IF A$=CHR$(8) THEN 337
333 IF A=O THEN 330
334 A=A-1
335 N$=LEFT$(N$,A)
336 GOSUB 515
337 IF A$="0" OR A$="9" THEN 330
338 PRINT A$;
339 A=A+1
340 N$=N$+A$
341 IF A=2 THEN 330
342 N1=VAL(N$)
343 N2=N+N1
344 IF N2<101 THEN 349
345 LOCATE 9,16:PRINT "THE TOTAL NUMBER OF DATA POINTS WILL EXCEED 100"
346 LOCATE 11,26:PRINT "PLEASE ENTER A LOWER NUMBER"
347 LOCATE 6,44:PRINT 
348 GOTO 327
349 IF N1=0 THEN 164
350 I=I+1
351 N=N2
352 ADD=1
353 GOTO 94
354 ' Restart and abort program routine
355 IF SAVED=1 THEN 361
356 LOCATE 17,30:PRINT SPC(20)
357 LOCATE 20,6:PRINT"LATEST DATA IS NOT SAVED ON DISK, DO YOU WANT TO SAVE IT (Y or N)? ";
358 Z$=INPUT$(1)
359 IF Z$="Y" OR Z$="y" THEN 102
360 IF Z$<>"N" AND Z$<>"n" THEN 358
361 IF A$="?" THEN 30
362 CLS
363 END
364 ' Subroutine which reads and validates a data entry
366 ' Entry parameters:
367 ' none
368 ' Return parameters:
369 ' N$ contains data
370 DP=0
371 A=0
372 N$=""
373 A$=INPUT$(1)
374 IF A$=CHR$(13) THEN 389
375 IF A$<>CHR$(8) THEN 381
376 IF A=0 THEN 373
377 IF RIGHT$(N$,1)="." THEN DP=0
378 A=A-1
379 N$=LEFT$(N$,A)
380 GOSUB 515
381 IF A$<"-" OR A$="/" OR A$>"9" THEN 373
382 IF A$="." AND DP=1 THEN 373
383 IF A$="-" THEN DP=1
384 IF A$="-" AND A<>0 THEN 373
385 PRINT A$;
386 A=A+1
387 N$=N$+A$
388 GOTO 373
389 PRINT
390 RETURN
391 ' Subroutine which reads and validates a filename
393 ' Entry parameters:
394 ' none
395 ' Return parameters:
396 ' filename saved in variable FILENAME$
398 ' E set to 1 if format is improper and requires new input
399 FILENAME$=""
401 A$=INPUT$(1)
402 A=ASC(A$)
403 IF A>ASC("a") AND A<=ASC("z") THEN A=A-32
404 A$=CHR$(A)
405 IF A$=CHR$(13) THEN 416
406 IF A$=CHR$(8) THEN 412
407 IF I=0 THEN 401
408 I=I-1
409 FILENAMES=LEFTS(FILENAMES,1)
410 GOSUB 515
411 GOTO 401
412 PRINT A$;
413 FILENAMES=FILENAMES+A$
414 I=I+1
415 IF I<14 THEN 401
416 GOSUB 424
417 IF E=0 THEN 421
418 LOCATE 12,24:PRINT "THAT FILENAME FORMAT IS IMPROPER"
419 E=1
420 RETURN
421 LOCATE 12,24,0:PRINT SPC(32)
422 RETURN
423 ' Subroutine which check filename for proper format
424 ' Entry parameters:
425 ' filename to be checked - FILENAMES$  
426 ' Return parameters:
427 ' E set to 0 if format is correct
428 ' E set to 1 if format is improper
429 NAMESTRT=1
430 IF MID$(FILENAMES$,2,1)<>":" THEN 434
431 NAMESTRT=3
432 IF LEFTS$(FILENAMES$,1)<"A" OR LEFTS$(FILENAMES$,1)="B" THEN E=1:RETURN
433 NAMELNTH=0
434 B$=MID$(FILENAMES$,NAMESTRT+NAMELNTH,1)
435 IF B$="" THEN 445
436 IF B$="" THEN 442
438 GOSUB 457
439 IF E=1 THEN RETURN
440 NAMELNTH=NAMELNTH+1
441 GOTO 435
442 IF NAMELNTH<1 OR NAMELNTH>8 THEN E=1:RETURN
443 E=0
444 RETURN
445 IF NAMELNTH<1 OR NAMELNTH>8 THEN E=1:RETURN
446 EXTSTRT=NAMESTRT+NAMELNTH+1
447 EXTLNTH=0
448 B$=MID$(FILENAMES$,EXTSTRT+EXTLNTH,1)
449 IF B$="" THEN 454
450 GOSUB 457
451 IF E=l THEN RETURN
452 EXTLNTH=EXTLNTH+1
453 GOTO 448
454 IF EXTLNTH>3 THEN E=l:RETURN
455 E=0
456 RETURN
457 IF B$="!" THEN E=l:RETURN
458 IF B$="!" AND B$<"#" THEN E=l:RETURN
459 IF B$="" AND B$<"-" THEN E=l:RETURN
460 IF B$>"-" AND B$<"0" THEN E=l:RETURN
461 IF B$>"9" AND B$<"@" THEN E=l:RETURN
462 IF B$="" AND B$<"" THEN E=l:RETURN
463 IF B$>"" AND B$<"-" THEN E=l:RETURN
464 IF B$ >"-" THEN E=l:RETURN
465 E=0
466 RETURN
467 '    Subroutine which displays 15 data points
468 '    Entry parameters:
469 '        I is the first point to be displayed
470 '        N is the total number of data points in the file
471 '    Return parameters:
472 '        DONE set to 0 if not all the data was displayed
473 '        DONE set to 1 if last data was displayed
474 ' DONE=0
475 LOCATE 5,1:PRINT"The number of data points in your sample is";N
476 J=1
477 LOCATE 7,1
478 PRINT"The value of observation";I;"is ";D$(I)
479 I=I+1
480 IF I>N THEN 485
481 J=J+1
482 IF J<16 THEN 479
483 DONE=1
484 RETURN
485 RETURN
486 RETURN
487 '    Subroutine which reads and validates a data point count entry
488 '    Entry parameters:
489 '        I is the last observation being displayed plus 1
490 '        N is the total number of observations in the sample
491 '        N4 is the number of the first observation displayed
492 '    Return parameters:
493 '        E is set to 0 if N1 is valid
494 '        E is set to 1 if N1 is invalid
495 E=0
496 A=0
497 N$=""  
500 A$=INPUT$(1)
501 IF A$=CHR$(13) THEN 512
502 IF A$=CHR$(8) THEN 507
503 IF A=0 THEN 500
504 A=A-1
505 N$=LEFT$(N$,A)
506 GOSUB 515
507 IF A$="0" OR A$="9" THEN 500
508 PRINT A$;
509 A=A+1
510 N$=N$+A$
511 IF A<3 THEN 500
512 N1=VAL(N$)
513 IF N1<N4 OR N1>I-1 THEN E=1
514 RETURN
515 ' Subroutine which clears one position to the left of the cursor and puts the cursor at the position just cleared. [For screen drivers which do not backspace when told to print a chr$(8).]
516 ' Entry parameters:
517 ' None
518 ' Return parameters:
519 ' None
520 X=CSRLI
521 Y=POS(I)
522 Y=Y-1
523 LOCATE X,Y:PRINT"
524 LOCATE X,Y
525 RETURN
526 ' Convert and Sort Data routine
527 CLS
528 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
529 LOCATE 5,30:PRINT"Converting the data."
530 FOR I%=1 TO N
531 D(I%)=VAL(D$(I%))
532 NEXT I%
533 IF X%>1 TO N
534 D(I%)=VAL(D$(I%))
535 NEXT I%
536 LOCATE 7,31:PRINT"Sorting the data."
537 LOCATE 10,11:PRINT"Please be patient, this may take a minute to complete."
538 J%=N
539 K%=0
540 FOR I%=1 TO J%-1
541 IF D(I%)<=D(I%+1) THEN 544
542 K%=I%
543 SWAP D(I%),D(I%+1)
544 NEXT I%
545 J%=K%
546 IF J%>1 THEN 539
547 SIZED=0
548 ' Data analysis menu
550 
CLS
LOCATE 1,12,0: PRINT "CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 3,33: PRINT "DATA ANALYSIS"
LOCATE 6,17: PRINT "1. Display a bar graph of the data.
LOCATE 7,17: PRINT "2. Fit a statistical distribution to the data.
LOCATE 8,17: PRINT "3. Return to the data manipulation section.
LOCATE 9,17: PRINT "4. Start all over again.
LOCATE 10,17: PRINT "5. Quit the entire process.
LOCATE 17,30,1: PRINT "TYPE IN YOUR CHOICE ";
A$=INPUT$(1)
IF A$<"1" OR A$="5" THEN 560
A=VAL(A$)
A$=CHR$(ASC(A$)+3)
ON A GOTO 565,570,164,354,354
' Bar graph compute
GOSUB 900
GOSUB 972
GOSUB 1064
GOTO 548

571 ' Distribution fitting menu
572 '
574 SELECTED=0
575 CHICOMP=0
LOCATE 1,12,0: PRINT "CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 3,30: PRINT "DISTRIBUTION FITTING"
LOCATE 6,16: PRINT "1. Fit a Uniform Distribution to the data.
LOCATE 7,16: PRINT "2. Fit a Normal Distribution to the data.
LOCATE 8,16: PRINT "3. Fit an Exponential Distribution to the data.
LOCATE 9,16: PRINT "4. Fit a Gamma Distribution to the data.
LOCATE 10,16: PRINT "5. Fit a Weibull Distribution to the data.
LOCATE 11,16: PRINT "6. Fit a Beta Distribution to the data.
LOCATE 12,16: PRINT "7. Return to the bar graph plot section.
LOCATE 13,16: PRINT "8. Return to the data manipulation section.
LOCATE 14,16: PRINT "9. Start all over again.
LOCATE 15,15: PRINT "10. Quit the entire process.
N$="
LOCATE 20,30,1: PRINT "TYPE IN YOUR CHOICE ";
A$=INPUT$(1)
IF A$<"1" OR A$="9" THEN 590
A$=A$+B$
GOTO 588

591 IF B$="0" THEN 600
592 PRINT A$;
593 B$=INPUT$(1)
594 IF B$=CHR$(13) THEN 600
596 IF B$="0" THEN 599
LOCATE 20,50: PRINT ""
601 A11 = A
602 IF A = 9 THEN A$ = "7"
603 ON A GOTO 604, 604, 604, 604, 604, 604, 548, 164, 354, 354
604 CLS
605 LOCATE 1, 12, 0: PRINT "CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
606 SUM = 0: SUM2 = 0: SUM3 = 0: SUM4 = 0
607 FOR I = 1 TO N
608 SUM = SUM + D(I)
609 SUM2 = SUM2 + (D(I) ^ 2)
610 SUM3 = SUM3 + D(I) ^ 3
611 SUM4 = SUM4 + D(I) ^ 4
612 NEXT I
613 MEAN = SUM / N
614 VARIANCE = (SUM2 - SUM ^ 2 / N) / N
615 IF VARIANCE <> 0 THEN 620
616 LOCATE 6, 15: PRINT "NOT A VALID DISTRIBUTION BECAUSE THE VARIANCE IS ZERO"
617 LOCATE 24, 27, 1: PRINT "PRESS ANY KEY TO CONTINUE ";
618 A$ = INPUT$(1)
619 GOTO 164
620 SKEW = (SUM3 / N - 3 * SUM2 / SUM / N) / VARIANCE ^ 1.5
621 KURTOSIS = (SUM4 / N - 4 * SUM3 / SUM / N ^ 2 + 6 * SUM2 / SUM / N ^ 3 - 3 * SUM / SUM / N ^ 4) / VARIANCE ^ 2
622 E = 0
623 ON A GOSUB 1071, 1086, 1139, 1157, 1180, 1254
624 IF E = 1 THEN 641
625 LOCATE 11, 15: PRINT "DISTRIBUTION VALUE FROM VALUE OF BEST
VALUE OF BEST
ATTRIBUTE SAMPLE DATA FIT DISTRIBUTION"
626 PRINT SPC(15): "ATTRIBUTE SAMPLE DATA FIT DISTRIBUTION"
627 LOCATE 14, 19: PRINT "MEAN", SPC(13):
628 PRINT USING "####.###" ; MEAN; MEANC
629 LOCATE 16, 17: PRINT "VARIANCE", SPC(11):
630 PRINT USING "####.###" ; VARIANCE; VARIANCEC
631 LOCATE 18, 17: PRINT "SKEWNESS", SPC(11):
632 PRINT USING "####.###" ; SKEW; SKEWC
633 LOCATE 20, 17: PRINT "KURTOSIS", SPC(11):
634 PRINT USING "####.###" ; KURTOSIS; KURTOSISC
635 LOCATE 24, 27: PRINT "PRESS ANY KEY TO CONTINUE";
636 FOR I = 1 TO 1000 * TIMEMULT
637 NEXT I
638 LOCATE 24, 27: PRINT SPC(25)
639 A$ = INPUT$(1)
640 GOTO 644
641 LOCATE 24, 27, 1: PRINT "PRESS ANY KEY TO CONTINUE ";
642 A$ = INPUT$(1)
643 GOTO 570
644 ' Data fit analysis
645 ' Data fit analysis
646 ' Data fit analysis
647 CLS
648 LOCATE 1, 12, 1: PRINT "CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
649 LOCATE 6, 8: PRINT 1. Select your own parameters for the chosen distribution.
650 LOCATE 7, 8: PRINT 2. Display a bar graph of the data with selected curve overlayed.
LOCATE 8,8:PRINT"3. Perform a Chi-Square Goodness of Fit Analysis."
LOCATE 9,8:PRINT"4. Fit a different statistical distribution to the data."
LOCATE 10,8:PRINT"5. Return to the bar graph plot section."
LOCATE 11,8:PRINT"6. Return to the data manipulation section."
LOCATE 12,8:PRINT"7. Start all over again."
LOCATE 13,8:PRINT"8. Quit the entire process."

A$=INPUT$(1)
IF A$<"1" OR A$>"8" THEN 658
A=VAL(A$)

CLS
LOCATE 1,12,1:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 3,15:PRINT"MANUAL PARAMETER SELECTION FOR CHOSEN DISTRIBUTION"
K=6
ON A GOSUB 1281,1322,1361,1392,1438,1495
IF SELECTED=0 THEN 644

Manual parameter selection results display

CLS
LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 5,9:PRINT"DISTRIBUTION PARAMETER COMPUTED VALUE SELECTED VALUE"
ON A GOSUB 1304,1343,1377,1420,1466,1523
LOCATE 12,5:PRINT"DISTRIBUTION VALUE FROM VALUE USING VALUE USING"
PRINT"ATTRIBUTE SAMPLE DATA COMPUTED PARAMETERS SELECTED PARAMETERS"
LOCATE 15,9:PRINT"MEAN";SPC(13);MEAN;MEANC;MEANR
LOCATE 17,7:PRINT"VARIANCE";SPC(11);VARIANCE;VARIANCEC;VARIANCER
LOCATE 19,7:PRINT"SKEWNESS";SPC(11);SKEW;SKEWC;SKEWR
LOCATE 21,7:PRINT"KURTOSIS";SPC(11);KURTOSIS;KURTOSISC;KURTOSISR
GOTO 635

Overlay plot routines
Z2$="*
IF SIZED=1 GOTO 695
GOSUB 900
IF SELECTED=1 THEN 701
GOSUB 1541
IF A$="N" OR A$="n" THEN 703
K=9
ON A GOSUB 1281,1322,1361,1392,1438,1495
IF SELECTED=0 THEN 644
701 GOSUB 972
702 ON A11 GOTO 709,740,776,793,817,834
703 GOSUB 972
704 ON A11 GOTO 705,737,774,790,814,831
705
706 ' Uniform distribution overlay plot routine
707 BOTTOM=D(1)
708 TOPL=D(N)
709 IF BOTTOM>BOUND(N2)+INTRVL/6 THEN C$="ABOVE":GOTO 732
710 IF TOPL<BOUND(O)-INTRVL/6 THEN C$="BELOW":GOTO 732
711 TOPLINE=21-CINT(20*N*INTRVL/(MAXR*(TOPL-BOTTOM))
712 IF TOPLINE>20 THEN TOPLINE=20
713 IF TOPLINE<1 THEN TOPROW=1 ELSE TOPROW=TOPLINE
714 BOUND=BOUND(O)
715 FOR J%=20 TO 20+3*N2
716 IF BOTTOM>BOUND+INTRVL/6 THEN 728
717 IF BOTTOM<BOUND-INTRVL/6 THEN 721
718 FOR I%=20 TO TOPROW STEP-1
719 LOCATE I%,J%:PRINT "*";
720 NEXT I%
721 IF TOPL<BOUND-INTRVL/6 THEN 728
722 IF TOPLINE<1 THEN TOPROW=0:GOTO 724
723 LOCATE TOPLINE,J%:PRINT "*";
724 IF TOPL>BOUND+INTRVL/6 THEN 728
725 FOR I%=TOPROW+1 TO 20
726 LOCATE I%,J%:PRINT "*";
727 NEXT I%
728 BOUND=BOUND+INTRVL/3
729 NEXT J%
730 GOSUB 1064
731 GOTO 644
732 CLS
733 LOCATE 1,12,1:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
734 LOCATE 4,7:PRINT"THE LIMITS CHOSEN PLACE THE SELECTED DISTRIBUTION ",C$;" THE GRAPH
735 LOCATE 8,27:PRINT"PRESS ANY KEY TO CONTINUE"
736 GOTO 639
737 ' Normal distribution overlay plot routine
738 MEANS=MEANC
739 VARIANCES=VARIANCEC
740 SD=SQR(VARIANCES)
741 Z2=0
742 Z3=0
743 Z=400+(100*(BOUND(O)-INTRVL/6-MEANS)/SD)
744 Z1=INT(Z)
745 IF Z1>798 THEN CDFR=1:GOTO 754
746 OPEN "1",#1,"NORMAL.DAT"
747 IF Z1<0 THEN 753
748 IF Z1=0 THEN 752
749 FOR I%=1 TO 71
750 INPUT#1,Z2
751 NEXT I%
752 INPUT#1,Z3
753 CDFR=Z2+(Z3-Z2)*(Z-Z1)
754 FOR J%=20 TO 20+3*N2
755 IF Z1>0 THEN Z4=Z1 ELSE Z4=-1
756 CDFL=CDFR
757 Z=Z+100*INTRVL/3/SD
758 Z1=INT(Z)
759 IF Z1>798 THEN CDFR=1:GOTO 769
760 IF Z1<0 THEN 768
761 IF Z1=0 THEN 767
762 IF Z4=Z1 THEN 768
763 IF Z1=Z4+1 THEN Z2=Z3:GOTO 767
764 FOR I%=Z4 TO Z1-2
765 INPUT#1,Z2
766 NEXT I%
767 INPUT#1,Z3
768 CDFR=Z2+(Z3-Z2)*(Z-Z1)
769 HEIGHT=21-CINT(60*N*(CDFR-CDFL)/MAXR)
770 GOSUB 1101
771 NEXT JX
772 CLOSE
773 GOTO 730
774 ' Exponential distribution overlay routine
775 MEANS=MEANC
776 HEIGHTL=0
777 Z=(BOUND(O)-INTRVL/6)/MEANS
778 IF Z<0 THEN CDFR=0:GOTO 780
779 CDFR=1-EXP(-Z)
780 FOR J%=20 TO 20+3*N2
781 CDFL=CDFR
782 Z=Z+INTRVL/3/MEANS
783 IF Z<0 THEN CDFR=0:GOTO 788
784 CDFR=1-EXP(-Z)
785 IF CDFL=0 THEN 788
786 HEIGHT=21-CINT(60*N*(CDFR-CDFL)/MAXR)
787 GOSUB 1101
788 NEXT JX
789 GOTO 730
790 ' Gamma distribution overlay routine
791 BS=B
792 CS=C
793 CA=CS
794 GOSUB 1556
795 R1=BOUND(O)-INTRVL/2
796 R3=10
797 IF CS<=1 THEN HEIGHTL=0 ELSE HEIGHTL=21
798 FOR J%=20 TO 20+3*N2
799 SKIP=0
800 R1=R1+INTRVL/3
801 GOTO 800
801 \( R2 = R1 + \text{INTRVL}/3 \)
802 \( \text{CDF} = 0 \)
803 \( \text{FOR} \ I\% = 1 \ \text{TO} \ R3 \)
804 \( Z = R1 + I\% * (R2 - R1)/R3 - (R2 - R1)/(2*R3) \)
805 \( \text{IF} \ Z = 0 \ \text{THEN} \ \text{SKIP} = 1: I\% = R3: \text{GOTO} \ 808 \)
806 \( FZ = (((Z/BS)^{(CS-1)}) * \text{EXP}(-Z/BS)) / (BS*GAMMA) \)
807 \( \text{CDF} = \text{CDF} + FZ * (R2 - R1)/R3 \)
808 \( \text{NEXT} \ I\% \)
809 \( \text{IF} \ \text{SKIP} = 1 \ \text{THEN} \ 812 \)
810 \( \text{HEIGHT} = 21 - \text{CINT}(60*N*\text{CDF}/\text{MAXR}) \)
811 \( \text{GOSUB} \ 1101 \)
812 \( \text{NEXT} \ J\% \)
813 \( \text{GOTO} \ 730 \)
814 \( ' \text{Weibull distribution overlay routine} \)
815 \( BS = B \)
816 \( CS = C \)
817 \( \text{IF} \ CS <= 1 \ \text{THEN} \ \text{HEIGHTL} = 0 \ \text{ELSE} \ \text{HEIGHTL} = 21 \)
818 \( Z = \text{BOUND}(0) - \text{INTRVL}/6 \)
819 \( \text{IF} \ Z <= 0 \ \text{THEN} \ \text{CDFR} = 0: \text{GOTO} \ 821 \)
820 \( \text{CDFR} = 1 - \text{EXP}(-(Z/BS)^{CS}) \)
821 \( \text{FOR} \ J\% = 20 \ \text{TO} \ 20 + 3*N2 \)
822 \( \text{CDFL} = \text{CDFR} \)
823 \( Z = Z + \text{INTRVL}/3 \)
824 \( \text{IF} \ Z <= 0 \ \text{THEN} \ \text{CDFR} = 0: \text{GOTO} \ 829 \)
825 \( \text{CDFR} = 1 - \text{EXP}(-(Z/BS)^{CS}) \)
826 \( \text{IF} \ \text{CDFL} = 0 \ \text{THEN} \ 829 \)
827 \( \text{HEIGHT} = 21 - \text{CINT}(60*N*(\text{CDFR} - \text{CDFL})/\text{MAXR}) \)
828 \( \text{GOSUB} \ 1101 \)
829 \( \text{NEXT} \ J\% \)
830 \( \text{GOTO} \ 730 \)
831 \( ' \text{Beta distribution overlay routine} \)
832 \( VS = V \)
833 \( WS = W \)
834 \( \text{CA} = VS \)
835 \( \text{GOSUB} \ 1556 \)
836 \( \text{GAMMA}1 = \text{GAMMA} \)
837 \( \text{CA} = WS \)
838 \( \text{GOSUB} \ 1556 \)
839 \( \text{GAMMA}2 = \text{GAMMA} \)
840 \( \text{CA} = VS + WS \)
841 \( \text{GOSUB} \ 1556 \)
842 \( \text{BETA} = \text{GAMMA}1 * \text{GAMMA}2 / \text{GAMMA} \)
843 \( R1 = \text{BOUND}(0) - \text{INTRVL}/2 \)
844 \( R3 = 10 \)
845 \( \text{SKIP} = 0 \)
846 \( \text{FOR} \ J\% = 20 \ \text{TO} \ 20 + 3*N2 \)
847 \( R1 = R1 + \text{INTRVL}/3 \)
848 \( R2 = R1 + \text{INTRVL}/3 \)
849 \( R1B = (R1-D(1))/(D(N)-D(1)) \)
850 \( R2B = (R2-D(1))/(D(N)-D(1)) \)
851 IF R1B<0 AND VS<=1 THEN SKIP=1:GOTO 864
852 IF R1B<0 THEN HEIGHT=21:GOTO 863
853 IF R2B>1 AND WS<=1 THEN 864
854 IF R2B>1 THEN HEIGHT=21:GOTO 863
855 CDF=0
856 FOR I%=1 TO R3
857 Z=R1B+I%*(R2B-R1B)/R3-(R2B-R1B)/(2*R3)
858 FZ=Z*(VS-1)*(1-Z)*(WS-1)/BETA
859 CDF=CDF+FZ*(R2B-R18)/R3
860 NEXT I%
861 HEIGHT=21-CINT(60*N*CDF/MAXR)
862 IF SKIP=1 THEN SKIP=0:HEIGHTL=HEIGHT
863 GOSUB 1101
864 NEXT J%
865 GOTO 730
866 ' Chi-Square goodness of fit analysis
867 CDF=0
868 CLS
869 LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
870 E=0
871 ON All GOSUB 1579,1652,1718,1767,1879,1930
872 IF E=1 THEN 644
873 IF SELECTED=1 THEN 882
874 LOCATE 8,16:PRINT"LEVEL OF";SPC(31);"CHI-SQUARE"
875 PRINT SPC(13);"SIGNIFICANCE CRITICAL FROM COMPUTED"
876 PRINT SPC(16);"(alpha)";SPC(11);"CHI-SQUARE PARAMETERS"
877 LOCATE 12,18:PRINT"0.01";SPC(15);
878 PRINT USING"###.##";CHI1;CHIC
879 LOCATE 14,18:PRINT"0.05";SPC(15);
880 PRINT USING"###.##";CHI5;CHIC
881 GOTO 889
882 LOCATE 8,6:PRINT"LEVEL OF";SPC(31);"CHI-SQUARE"
883 PRINT"SIGNIFICANCE CRITICAL FROM COMPUTED FROM SELECTED"
884 PRINT"(alpha)";SPC(11);"CHI-SQUARE PARAMETERS PARAMETERS"
885 LOCATE 12,8:PRINT"0.01";SPC(15);
886 PRINT USING"###.##";CHI1;CHIC;CHIS
887 LOCATE 14,8:PRINT"0.05";SPC(15);
888 PRINT USING"###.##";CHI5;CHIC;CHIS
889 LOCATE 17,1:PRINT"CONCLUSIONS:
890 IF CHIC>CHI1 THEN C$="SHOULD" ELSE C$="CANNOT"
891 LOCATE 19,6:PRINT"You ";C$;" reject the COMPUTED PARAMETERS at 0.01 level of significance.
892 IF CHIC>CHI5 THEN C$="SHOULD" ELSE C$="CANNOT"
893 PRINT"You ";C$;" reject the COMPUTED PARAMETERS at 0.05 level of significance.
894 IF SELECTED=0 THEN 899
895 IF CHIS>CHI1 THEN C$="SHOULD" ELSE C$="CANNOT"
896 PRINT"You ";C$;" reject the SELECTED PARAMETERS at 0.01 level of significance.
897 IF CHIS>CHI5 THEN C$="SHOULD" ELSE C$="CANNOT"
898 PRINT"You ";C$;" reject the SELECTED PARAMETERS at 0.05 level of significance.
899 GOTO 635
900 '
Subroutine which sizes the bar graph

Entry parameters:
N is the total number of data points in the sample

Exit parameters:
N2 is the number of ranges desired
BOUND(x) are the upper boundaries of the ranges
INTRVL is the value BOUND(2)-BOUND(1)

CLS

LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL"
LOCATE 3,32:PRINT"BAR GRAPH SIZING"

LOCATE 5,1:PRINT"You may manually size the graph (select the number of frequency intervals, interval width and initial interval upper boundary), or you may allow the computer to size the graph for you. Do you wish to manually size the graph (Y or N) ? "
A$=INPUT$(1)
IF A$<"Y" AND A$<>"y" AND A$<>"N" AND A$<>"n" THEN 915
IF A$="N" OR A$="n" THEN 968
PRINT"Y"

LOCATE 11,1:PRINT"How many frequency intervals should the graph have (3 to 20) ? ";
A=0
N2$=""
A$=INPUT$(1)
IF A$=CHR$(13) THEN 934
IF A$<>CHR$(8) THEN 929
IF A=O THEN 922
A=A-1
N2$=LEFT$(N2$,A)
GOSUB 515
IF A$<"0" OR A$>"9" THEN 922
PRINT A$;
A=A+1
N2$=N2$+A$
IF N2>2 AND N2<21 THEN 939
LOCATE 14,10:PRINT "THE QUANTITY OF FREQUENCY INTERVALS MUST BE BETWEEN 3 AND 20"
LOCATE 11,64:PRINT SPC(20)
GOTO 940
LOCATE 16,17:PRINT SPC(46)
LOCATE 15,1:PRINT "What is the upper limit of the lowest range ? ";
GOSUB 364
INTRVL=VAL(N$)
IF INTRVL > 0 THEN 947
LOCATE 16,17:PRINT "THE INTERVAL WIDTH CAN NOT BE NEGATIVE OR ZERO"
LOCATE 13,43:PRINT SPC(20)
GOTO 940
LOCATE 16,17:PRINT SPC(46)
LOCATE 15,1:PRINT "What is the upper limit of the lowest range ? ";
GOSUB 364
BOUND(1)=VAL(N$)
951 POWERP=0
952 BOUNDP(1)=BOUND(1)
953 INTRVLP=INTRVL
954 IF BOUNDP(1)<FIX(BOUNDP(1)) THEN 956
955 IF INTRVLP=FIX(INTRVLP) THEN 960
956 POWERP=POWERP+1
957 BOUNDP(1)=10*BOUNDP(1)
958 INTRVLP=10*INTRVLP
959 GOTO 954
960 FOR I%=2 TO N2
961 BOUND(I%) = BOUND(I%-1)+INTRVL
962 BOUNDP(I%)=BOUNDP(I%-1)+INTRVLP
963 NEXT I%
964 BOUND(0)=BOUND(1)-INTRVL
965 BOUNDP(0)=BOUNDP(1)-INTRVLP
966 SIZED=1
967 RETURN
968 N2=CINT(SQR(N))
969 INTRVL=(D(N)-D(1))/N2
970 BOUND(1)=INTRVL+D(1)
971 GOTO 951
972 '
973 ' Subroutine which displays the bar graph
974 ' Entry parameters:
975 ' N2 is the number of ranges desired
976 ' BOUND(x) are the upper boundaries of the ranges
977 ' INTRVL is BOUND(2)-BOUND(1)
978 ' Exit parameters:
979 ' MAXR is the value of the top of the graph
980 CLS
981 LOCATE ,,0
982 IF INTRVLP=0 THEN INTRVLP=D(1)*10^POWERP
983 IF INTRVLP=0 THEN INTRVLP=1
984 ' Compute quantity in each range
985 FOR I%=1 TO N2
986 RANGE(I%)=0
987 NEXT I%
988 J%=1
989 FOR I%=1 TO N
990 IF J%=N2 THEN 994
991 IF D(I%)<=BOUNDP(J%)/10^POWERP THEN 994
992 J%=J%+1
993 GOTO 990
994 RANGE(J%)=RANGE(J%)+1
995 NEXT I%
996 MAXR=0
997 FOR I%=1 TO N2
998 IF RANGE(I%)<=MAXR THEN 1000
999 MAXR=RANGE(I%)
1000 NEXT I%
1001 MAXR=4*CINT(.3*MAXR)
1002 IF MAXR<4 THEN MAXR=4
1003 ' Bar graph display
1004 ' I is rows 1 to 20
1005 ' J is ranges 1 thru N2
1006 ' K is number of ** per bar
1007 Z1$=""
1008 FOR I% = 1 TO 20
1009 IF I%<=5 THEN 1012
1010 LOCATE 5,16:PRINT USING "##";3*MAXR/4;
1011 GOTO 1022
1012 IF I%<=10 THEN 1015
1013 LOCATE 10,16:PRINT USING "##"; MAXR/2;
1014 GOTO 1022
1015 IF I%<=15 THEN 1018
1016 LOCATE 15,16:PRINT USING "##";MAXR/4;
1017 GOTO 1022
1018 IF I%<=20 THEN 1023
1019 Z1$="_"
1020 LOCATE 20,17:PRINT"O_j_";
1021 GOTO 1025
1022 PRINT"_";
1023 LOCATE I%,19:PRINT"|";
1024 PRINT"|
1025 L=1.025-.05*I%
1026 FOR J%=1 TO N2
1027 SIZE=RANGE(J%)/MAXR
1028 Z$="|
1029 IF SIZE<L THEN Z$=Z1$
1030 FOR K1%=1 TO 2
1031 PRINT Z$;
1032 NEXT K1%
1033 PRINT Z1$;
1034 NEXT J%
1035 NEXT I%
1036 ' Label horizontal axis
1037 OFFSET=0
1038 POWER=0
1039 IF BOUND(N2-1)<100 AND BOUND(1)>=-10 THEN 1051
1040 IF BOUND(1)>-10 THEN 1042
1041 OFFSET=BOUND(1)
1042 TOP=BOUND(N2-1)-OFFSET
1043 IF TOP<100 THEN 1047
1044 TOP=TOP/10
1045 POWER=POWER+1
1046 GOTO 1043
1047 IF TOP>.1 THEN 1051
1048 TOP=TOP*10
1049 POWER=POWER-1
1050 GOTO 1047
1051 FOR %I=1 TO N2-1 STEP 2
1052 LABEL=(BOUND(%I)-OFFSET)/10^POWER
1053 JX=20+3*%I
1054 LOCATE 21,JX:PRINT"|
1055 LOCATE 22,JX-2:PRINT USING "##.##";LABEL;
1056 NEXT %I
1057 IF POWER=0 THEN 1061
1058 LOCATE 23,20:PRINT"Multiply all range boundary values by";10^POWER;
1059 IF OFFSET=0 THEN PRINT".";GOTO 1063
1060 PRINT", then";
1061 IF OFFSET=0 THEN 1063
1062 LOCATE 24,20:PRINT"Subtract";-OFFSET;"from each range boundary value.";
1063 RETURN
1064 LOCATE 1,1:PRINT"PRESS ANY KEY";
1065 LOCATE 2,1:PRINT"TO CONTINUE";
1066 FOR I=1 TO 1000*TIMEMULT:NEXT I
1067 LOCATE 1,0:PRINT SPC(13)
1068 LOCATE 2,1:PRINT SPC(11)
1069 A$=INPUT$(1)
1070 RETURN

1071 ' Subroutine which computes the parameters for a uniform distribution
1072 ' Entry parameters:
1073 ' D(1) is the lowest data point
1074 ' D(N) is the highest data point
1075 ' Exit parameters:
1076 ' MEANC, VARIANCEC, SKEWC and KURTOSISC are the computed values
1077 LOCATE 3,17:PRINT"BEST FIT PARAMETERS FOR A UNIFORM DISTRIBUTION"
1078 LOCATE 6,38:PRINT"a = ";D(1)
1079 LOCATE 8,38:PRINT"b = ";D(N)
1080 MEANC=(D(N)+D(1))/2
1081 VARIANCEC=(D(N)-D(1))^2/12
1082 SKEW=(D(N)-D(1))/2/12
1083 KURTOSISC=0
1084 KURTOSISC=1.8
1085 RETURN
1086 ' Subroutine which computes the parameters for a normal distribution
1087 ' Entry parameters:
1088 ' MEAN is the sample mean
1089 ' VARIANCE is the unadjusted sample variance
1090 ' Exit parameters:
1091 ' MEANC, VARIANCEC, SKEWC and KURTOSISC are the computed values
1092 LOCATE 3,17:PRINT"BEST FIT PARAMETERS FOR A NORMAL DISTRIBUTION"
1093 LOCATE 6,35:PRINT"MEAN = ";MEAN
1094 LOCATE 8,31:PRINT"VARIANCE = ";VARIANCE
1095 KURTOSISC=MEAN
1096 VARIANCEC=VARIANCE
1097 SKEW=0
1098 KURTOSISC=3
1099 RETURN
Subroutine which plots the computed overlay point

Entry parameters:
- \( J\% \) is the horizontal column position being plotted
- \( \text{HEIGHTL} \) is the height (inverted) of the previous column
- \( \text{HEIGHT} \) is the computed height (inverted) of this column

Exit parameters:
- \( \text{HEIGHTL} \) is the height (inverted) of this column

IF \( J\% = 20 \) THEN \( \text{HEIGHTL} = \text{HEIGHT} \)
IF \( \text{HEIGHT} > \text{HEIGHTL} \) THEN 1124
\( \text{LEFT} = \text{INT}((\text{HEIGHTL} - \text{HEIGHT} - 1)/2) \)
\( \text{RIGHT} = \text{INT}((\text{HEIGHTL} - \text{HEIGHT})/2) \)
IF \( \text{LEFT} < 1 \) THEN 1119
\( \text{HEIGHTL} = \text{HEIGHTL} - 1 \)
IF \( \text{HEIGHTL} < 1 \) OR \( \text{HEIGHTL} > 20 \) THEN 1117
\( \text{LOCATE} \text{HEIGHTL}, \text{J\%} - 1; \text{PRINT} \text{Z2$} \);
\( \text{LEFT} = \text{LEFT} - 1 \)
\( \text{GOTO} 1113 \)
IF \( \text{RIGHT} < 0 \) THEN 1137
IF \( \text{HEIGHT} + \text{RIGHT} < 1 \) OR \( \text{HEIGHT} + \text{RIGHT} > 20 \) THEN 1122
\( \text{LOCATE} \text{HEIGHT} + \text{RIGHT}, \text{J\%}; \text{PRINT} \text{Z2$} \);
\( \text{RIGHT} = \text{RIGHT} - 1 \)
\( \text{GOTO} 1119 \)
\( \text{LEFT} = \text{INT}((\text{HEIGHT} - \text{HEIGHTL})/2) \)
\( \text{RIGHT} = \text{INT}((\text{HEIGHT} - \text{HEIGHTL} - 1)/2) \)
IF \( \text{LEFT} < 1 \) THEN 1132
\( \text{HEIGHTL} = \text{HEIGHTL} + 1 \)
IF \( \text{HEIGHTL} < 1 \) OR \( \text{HEIGHTL} > 20 \) THEN 1130
\( \text{LOCATE} \text{HEIGHTL}, \text{J\%} - 1; \text{PRINT} \text{Z2$} \);
\( \text{LEFT} = \text{LEFT} - 1 \)
\( \text{GOTO} 1126 \)
IF \( \text{RIGHT} < 0 \) THEN 1137
IF \( \text{HEIGHT} - \text{RIGHT} < 1 \) OR \( \text{HEIGHT} - \text{RIGHT} > 20 \) THEN 1135
\( \text{LOCATE} \text{HEIGHT} - \text{RIGHT}, \text{J\%}; \text{PRINT} \text{Z2$} \);
\( \text{RIGHT} = \text{RIGHT} - 1 \)
\( \text{GOTO} 1132 \)
\( \text{HEIGHTL} = \text{HEIGHT} \)
\( \text{RETURN} \)

Subroutine which computes the parameter for an exponential distribution

Entry parameter:
- \( \text{MEAN} \) is the sample mean

Exit parameters:
- \( \text{MEANC}, \text{VARIANCEC}, \text{SKEWC} \) and \( \text{KURTOSISC} \) are the computed values

IF \( \text{MEAN} > 0 \) THEN 1150
\( \text{LOCATE} 6, 2; \text{PRINT} "\text{CANNOT FIT AN EXPONENTIAL DISTRIBUTION BECAUSE THE MEAN IS NEGATIVE OR ZERO}" \)
\( \text{E} = 1 \)
\( \text{GOTO} 1156 \)
\( \text{LOCATE} 3, 15; \text{PRINT} "\text{BEST FIT PARAMETER FOR AN EXPONENTIAL DISTRIBUTION}" \)
Subroutine which computes the parameters for a gamma distribution

Entry parameters:
- MEAN is the sample mean
- VARIANCE is the unadjusted sample variance

Exit parameters:
- B is the scale parameter (b or beta)
- C is the shape parameter (c or alpha)
- MEANC, VARIANCEC, SKEW C and KURTOSISC are the computed values

IF MEAN > 0 THEN
LOCATE 6,20:PRINT "CANNOT FIT A GAMMA DISTRIBUTION BECAUSE THE MEAN IS NEGATIVE OR ZERO"
E = 1
GOTO 1179
ENDIF
END

Subroutine which computes the parameters for a Weibull distribution

Entry parameters:
- N is the sample size
- D(I) are the sample data points
- MEAN is the sample mean

Exit parameters:
- B is the scale parameter (b or beta)
- C is the shape parameter (c or alpha)
- MEANC, VARIANCEC, SKEW C and KURTOSISC are the computed values

IF MEAN > 0 THEN
LOCATE 6,5:PRINT "CANNOT FIT A WEIBULL DISTRIBUTION BECAUSE THE MEAN IS NEGATIVE OR ZERO"
E = 1
GOTO 1253
ENDIF
END
1201 FOR I%=1 TO N
1202 IF D(I%)<=0 THEN NEG=1:GOTO 1207
1203 NW=NW+1
1204 AK=AK+LOG(D(I%))
1205 LOGSUM=LOGSUM+LOG(D(I%))
1206 LOGSQSUM=LOGSQSUM+LOG(D(I%))^2
1207 NEXT I%
1208 AK=AK/NW
1209 C=1/SQR(6*(LOGSQSUM-((LOGSUM-A2)/NW))/(3.14159^2*(NW-1)))
1210 CL=-1
1211 WHILE C>1.001*CL OR C<.999*CL
1212 GOSUB 2059
1213 CL=C
1214 BK=0
1215 CK=0
1216 HK=0
1217 FOR I%=1 TO N
1218 IF D(I%)<=0 THEN 1222
1219 BK=BK+D(I%)*CL
1220 CK=CK+D(I%)*CL*LOG(D(I%))
1221 HK=HK+D(I%)*CL*(LOG(D(I%)))^2
1222 NEXT I%
1223 C=CL+(AK+1/CL-CK/BK)/(1/CL+AK+BK*HK-CK/BA2)/BKA2)
1224 WEND
1225 GOSUB 2059
1226 K=0
1227 IF NEG=0 THEN K=1:GOTO 1229
1228 LOCATE 9,4:PRINT“NOTE: Non-positive sample data were ignored while computing these parameters.
1229 BK=0
1230 FOR I%=1 TO N
1231 IF D(I%)<=0 THEN 1233
1232 BK=BK+D(I%)*C
1233 NEXT I%
1234 B=(BK/NW)^(1/C)
1235 LOCATE K+5,21:PRINT“SCALE PARAMETER (b or beta) = “;B
1236 LOCATE K+7,20:PRINT“SHAPE PARAMETER (c or alpha) = “;C
1237 CA=(C+1)/C
1238 GOSUB 1556
1239 GAMMA1=GAMMA
1240 CA=(C+2)/C
1241 GOSUB 1556
1242 GAMMA2=GAMMA
1243 CA=(C+3)/C
1244 GOSUB 1556
1245 GAMMA3=GAMMA
1246 CA=(C+4)/C
1247 GOSUB 1556
1248 MEAN=B*GAMMA1
1249 VARIANCEC=B^2*((GAMMA2-GAMMA1^2)
1250 SKEW=(GAMMA3-3*GAMMA2*GAMMA1+2*GAMMA1^2)/(GAMMA2-GAMMA1^2)^1.5
KURTOSISC = \((\Gamma - 4\Gamma^3\Gamma^1 + 6\Gamma^2\Gamma^1^2 - 3\Gamma^1^3)/(\Gamma^2 - \Gamma^1^2)^2\) 

GOSUB 2060
RETURN

```
Subroutine which computes the parameters for a beta distribution
Entry parameters:
MEAN is the sample mean
VARIANCE is the unadjusted sample variance
Exit parameters:
V is the first shape parameter \((x^{(V-1)})\)
W is the second shape parameter \(((1-x)^{(W-1)})\)
MEANC, VARIANCEC, SKEWC and KURTOSISC are the computed values
```

LOCATE 3, 18: PRINT "BEST FIT PARAMETERS FOR A BETA DISTRIBUTION"
MEAN = (MEAN - D(1))/(D(N) - D(1))
VARIANCE = VARIANCE /((D(N) - D(1))^2)
V = MEAN * (MEAN * (1 - MEAN) / VARIANCE - 1)
W = (1 - MEAN) * (MEAN * (1 - MEAN) / VARIANCE - 1)

LOCATE 5, 15: PRINT "SHAPE PARAMETERS (v & w) = "; V; "AND"; W

LOCATE 7, 1: PRINT "NOTE: These parameters are for a Beta Distribution in the range 0<=x<=1."

To generate data with the same characteristics as your input data,

IF D(N) - D(1) >= 1 THEN PRINT "multiply the probability density function by"; D(N) - D(1); GOTO 1273
PRINT "divide the probability density function by"; 1/(D(N) - D(1));
IF D(1) = 0 THEN PRINT "."; GOTO 1276
IF D(1) < 0 THEN PRINT "and then subtract"; D(1); ";."; GOTO 1276
PRINT "and then add"; D(1); ";.";

MEAN = V * (D(N) - D(1)) / (V + W) + D(1)
VARIANCE = V * W * (D(N) - D(1))^2 / (V + W)^2
SKEWC = 2 * (W - V) * SQR((V + W + 1)) / (V + W + 2) * SQR(V * W)
KURTOSISC = 3 * (V * W) * (V + W + 1) / (V + 1) * (2 * W - V) / (V * W * (V + W + 2) * (V + W + 3)) + V * (V - W) / (V + W)

RETURN

Subroutine which inputs uniform distribution parameters
Entry parameters:
none
Exit parameters:
BOTTOM is the lower bound of the distribution
TOPL is the upper bound of the distribution

LOCATE K + 1, 26: PRINT "(Computed value was "; D(1); ")"
LOCATE K, 1: PRINT "What should the MINIMUM VALUE of the Uniform Distribution be? "
GOSUB 364
BOTTOM = VAL(NS)
LOCATE K + 4, 26: PRINT "(Computed value was "; D(N); ")"
LOCATE K + 3, 1: PRINT "What should the MAXIMUM VALUE of the Uniform Distribution be? "
GOSUB 364
TOPL = VAL(NS)
IF TOPL > BOTTOM THEN 1302
LOCATE 16, 12: PRINT "THE MAXIMUM VALUE MUST BE GREATER THAN THE MINIMUM VALUE."
LOCATE 19, 27: PRINT "PRESS ANY KEY TO CONTINUE"
SELECTED = 0
A$ = INPUT$(1)
Subroutine which computes uniform distribution attributes from selected parameters

Entry parameters:
BOTTOM is the low boundary
TOPL is the upper boundary

Exit parameters:
MEANR, VARIANCER, SKEWR and KURTOSISR

LOCATE 3,19:PRINT"UNIFORM DISTRIBUTION PARAMETER COMPARISON"
LOCATE 7,15:PRINT"MINIMUM (a)";SPC(14);D(l);BOTTOM
LOCATE 9,15:PRINT"MAXIMUM (b)";SPC(14);D(N);TOPL
MEANR=(TOPL+BOTTOM)/2
VARIANCER=(TOPL-BOTTOM)^2/12
SKEWR=0
KURTOSISR=1.8
RETURN

Subroutine which inputs normal distribution parameters

Entry parameters:
none

Exit parameters:
MEANS is the selected mean
VARIANCES is the selected variance

LOCATE K+l,26:PRINT"(Computed value was ";MEANC;".)"
LOCATE K,l:PRINT"What should the MEAN of the Normal Distribution be ? ";GOSUB 364
MEANS=VAL(N$)
LOCATE K+4,26:PRINT"(Computed value was ";VARIANCEC;".)"
LOCATE K+3,1:PRINT"What should the VARIANCE of the Normal Distribution be ? ";GOSUB 364
VARIANCES=VAL(N$)
IF VARIANCES>0 THEN 1341
LOCATE 16,21:PRINT"THE VARIANCE MUST BE GREATER THAN ZERO"
LOCATE K+3,60:PRINT SPC(20);
GOTO 1339
SELECTED=1
RETURN

Subroutine which computes normal distribution attributes from selected parameters

Entry parameters:
MEANS is the selected mean
VARIANCES is the selected variance

Exit parameters:
MEANR, VARIANCER, SKEWR and KURTOSISR
LOCATE 3,20:PRINT"NORMAL DISTRIBUTION PARAMETER COMPARISON"
LOCATE 7,19:PRINT"MEAN";SPC(17);
PRINT USING"#####.###";MEANC;MEANS
LOCATE 9,17:PRINT"VARIANCE";SPC(15);
PRINT USING"#####.###";VARIANCEC;VARIANCES
MEANR=MEANS
VARIANCER=VARIANCES
SKEWR=0
KURTOSISR=3
RETURN

Subroutine which inputs the exponential distribution parameter
Entry parameters:
none
Exit parameter:
MEANS is the selected mean
LOCATE K+1,26:PRINT"(Computed value was ";MEANC;".)"
LOCATE K,1:PRINT"What should the MEAN of the Exponential Distribution be ? ";
GOSUB 364
MEANS=VAL(N$)
IF MEANS>0 THEN 1375
LOCATE 16,23:PRINT"THE MEAN MUST BE GREATER THAN ZERO"
LOCATE K,59:PRINT SPC(20);
GOTO 1368
SELECTED=1
RETURN

Subroutine which computes exponential distribution attributes from
a selected parameter
Entry parameter:
MEANS is the selected mean
Exit parameters:
MEANR, VARIANCER, SKEWR and KURTOSISR
LOCATE 3,17:PRINT"EXPONENTIAL DISTRIBUTION PARAMETER COMPARISON"
LOCATE 7,19:PRINT"MEAN";SPC(17);
PRINT USING"#####.###";MEANC;MEANS
MEANR=MEANS
VARIANCER=MEANS^2
SKEWR=2
KURTOSISR=9
RETURN

Subroutine which inputs gamma distribution parameters
Entry parameters:
none
Exit parameters:
BS is the selected scale parameter
CS is the selected shape parameter
LOCATE K+2,26:PRINT"(Computed value was ";B;".)"
LOCATE K,1:PRINT"What should the SCALE PARAMETER (b or beta)
LOCATE K+1,30:PRINT"of the Gamma Distribution be ? ";
GOSUB 364
BS=VAL($)
IF BS<0 THEN 1408
LOCATE 16,17:PRINT"THE SCALE PARAMETER MUST BE GREATER THAN ZERO"
LOCATE K+1,61:PRINT SPC(20);
GOTO 1401
LOCATE 16,17:PRINT SPC(45)
LOCATE K+6,26:PRINT"(Computed value was ";:C;".)"
LOCATE K+4,1:PRINT"What should the SHAPE PARAMETER (c or alpha) of the Gamma Distribution be ? ";
GOSUB 364
CS=VAL($)
IF CS<0 THEN 1418
LOCATE 16,17:PRINT"THE SHAPE PARAMETER MUST BE GREATER THAN ZERO"
LOCATE K+5,61:PRINT SPC(20);
GOTO 1411
SELECTED=1
RETURN

Subroutine which computes gamma distribution attributes from selected parameters
Entry parameters:
BS is the selected scale parameter
CS is the selected shape parameter
Exit parameters:
MEANR, VARIANCER, SKEWR and KURTOSISR
LOCATE 3,20:PRINT"GAMMA DISTRIBUTION PARAMETER COMPARISON"
LOCATE 7,12:PRINT"SCALE (b or beta)";SPC(13);
PRINT USING"#####.### ";B;BS
LOCATE 9,12:PRINT"SHAPE (c or alpha)";SPC(12);
PRINT USING"#####.### ";C;CS
MEANR=BS*CS
VARIANCER=BS~2*CS
SKEWR=2/SQR(CS)
KURTOSISR=3+6/CS
RETURN

Subroutine which inputs Weibull distribution parameters
Entry parameters:
none
Exit parameters:
BS is the selected scale parameter
CS is the selected shape parameter
LOCATE K+2,26:PRINT"(Computed value was ";:B;".)"
LOCATE K+1,30:PRINT"of the Gamma Distribution be ? ";
GOSUB 364
BS=VAL($)
IF BS<0 THEN 1458
LOCATE 16,17:PRINT SPC(45)
LOCATE K+6,26:PRINT"(Computed value was ";:C;".)"
LOCATE K+4,1:PRINT"What should the SCALE PARAMETER (b or beta) of the Weibull Distribution be ? ";
GOSUB 364
BS=VAL($)
IF BS<0 THEN 1458
LOCATE 16,17:PRINT SPC(45)
LOCATE K+6,26:PRINT"(Computed value was ";:C;".)"
THE SCALE PARAMETER MUST BE GREATER THAN ZERO

LOCATE K+1,61:PRINT SPC(22);
GOTO 1447
LOCATE K+6,26:PRINT (Computed value was ",;",.")
LOCATE K+4,1:PRINT "What should the SHAPE PARAMETER (c or alpha)"
LOCATE K+5,30:PRINT "of the Weibull Distribution be ? ";
GOSUB 364
CS=VAL(N$)
IF CS>O THEN 1464
LOCATE K+5,61:PRINT SPC(22);
GOTO 1457
SELECTED=1
RETURN

Subroutine which computes Weibull distribution attributes from selected parameters
Entry parameters:
BS is the selected scale parameter
CS is the selected shape parameter
Exit parameters:
MEANR, VARIANCER, SKEWR and KURTOSISR
LOCATE 3,19:PRINT "SCALE (b or beta) ";SPC(13);PRINT USING "#####.### ";B;BS
LOCATE 9,12:PRINT "SHAPE (c or alpha) ";SPC(12);PRINT USING "#####.### ";C;CS
CA=(CS+1)/CS
GOSUB 1556
GAMMA1=GAMMA
CA=(CS+2)/CS
GOSUB 1556
GAMMA2=GAMMA
CA=(CS+3)/CS
GOSUB 1556
GAMMA3=GAMMA
CA=(CS+4)/CS
GOSUB 1556
MEANR=BS^GAMMA1
VARIANCER=BS^2*(GAMMA2-GAMMA1^2)
SKEWR=(GAMMA3-3*GAMMA2*GAMMA1+2*GAMMA1^2)/(GAMMA2-GAMMA1^2)^1.5
KURTOSISR=(GAMMA-4*GAMMA3*GAMMA1+6*GAMMA2*GAMMA1^2-3*GAMMA1^3)/(GAMMA2-GAMMA1^2)^2
RETURN

Subroutine which inputs beta distribution parameters
Entry parameters:
none
Exit parameters:
VS is the selected first shape parameter \(x^{(VS-1)}\)
WS is the selected second shape parameter 

LOCATE K+2,26:PRINT"(Computed value was ";V;".)"

LOCATE K,1:PRINT"What should the FIRST SHAPE PARAMETER (v)

LOCATE K+1,30:PRINT"of the Beta Distribution be ? ";

GOSUB 364

VS=VAL(N$)

IF VS>0 THEN 1511

LOCATE 16,17:PRINT"THE SHAPE PARAMETER MUST BE GREATER THAN ZERO

LOCATE K+1,60:PRINT SPC(20);

GOTO 1504

LOCATE K+6,26:PRINT"(Computed value was ";W;".)"

LOCATE K+4,1:PRINT"What should the SECOND SHAPE PARAMETER (w)

LOCATE K+5,30:PRINT"of the Beta Distribution be ? ";

GOSUB 364

WS=VAL(N$)

IF WS>0 THEN 1521

LOCATE 16,17:PRINT"THE SHAPE PARAMETER MUST BE GREATER THAN ZERO

LOCATE K+5,60:PRINT SPC(20);

GOTO 1514

SELECTED=1

RETURN

Subroutine which computes beta distribution attributes from

Entry parameters:

VS is the selected first shape parameter (x^((VS-1))

WS is the selected second shape parameter ((1-x)^((WS-1))

Exit parameters:

MEANR, VARIANCER, SKEWR and KURTOSISR

LOCATE 3,21:PRINT"BETA DISTRIBUTION PARAMETER SELECTION

LOCATE 7,9:PRINT"FIRST SHAPE PARAMETER (v)";SPC(B)

PRINT USING"#####.### ";V;VS

LOCATE 9,9:PRINT"SECOND SHAPE PARAMETER (w)";SPC(5);W;WS

MEANR=VS*(D(N)-D(1))/(VS+WS)+D(l)

VARIANCER=VS*WS*(D(N)-D(1))^2/(VS+WS)^2/(VS+WS+1)

SKEWR=2*(WS-WS)*SQR(VS+WS+1)/(VS+WS+2)/SQR(WS*WS)

KURTOSISR=3*(VS+WS)^2*(VS+WS+1)*(VS+1)^2/(VS+WS)/(VS+WS+2)/(VS+WS+3)+VS*(VS-WS)/(VS+WS)/(VS+WS)

RETURN

CLS

LOCATE 1,12,0:PRINT"CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

LOCATE 3,21:PRINT"SELECTED CURVE PARAMETER DETERMINATION

LOCATE 5,1:PRINT"You may manually input parameters for the overlayed curve,

LOCATE 6,6:PRINT"or you may use the previously computed values.

LOCATE 7,1:PRINT"Do you wish to manually input the parameters (Y or N) ? ";
1551 A$=INPUT$(1)
1552 IF A$="N" OR A$="n" THEN 1555
1553 IF A$="Y" AND A$="y" THEN 1551
1554 PRINT "Y"
1555 RETURN

1557 ' Subroutine which determines the gamma value of a positive number
1558 ' Entry parameter:
1559 ' CA is the number for which the gamma function is to be computed
1560 ' Exit parameter:
1561 ' GAMMA is the computed gamma value of CA
1562 GAMMA=1
1563 IF CA<=1 THEN 1567
1564 CA=CA-1
1565 GAMMA=CA*GAMMA
1566 GOTO 1563
1567 C1=100*CA
1568 C2=INT(C1)
1569 IF CA=0 THEN 1578
1570 IF CA<.01 THEN C3=1/CA:C4=C3:GOTO 1577
1571 OPEN "1",#1,"GAMMA.DAT"
1572 FOR I%=1 TO C2
1573 INPUT#!, C3
1574 NEXT I%
1575 INPUT#1,C4
1576 CLOSE
1577 GAMMA=GAMMA*(C3-(Cl-C2)*(C3-C4))
1578 RETURN

1579 ' Subroutine which computes unique uniform distribution
1580 ' Chi-Square values
1581 ' Entry parameters:
1582 ' N is the sample size
1583 ' D(I) are the sample data points
1584 ' BOTTOM is the lower selected boundary
1585 ' TOPL is the upper selected boundary
1586 ' Exit parameters:
1587 ' CHIC is the Chi-Square value of the computed boundaries
1588 ' CHIS is the Chi-Square value of the selected boundaries
1589 ' CHI1 is the critical Chi-Square value at alpha=0.01
1590 ' CHI5 is the critical Chi-Square value at alpha=0.05
1591 ' LOCATE 3.5:PRINT"CHI-SQUARE GOODNESS OF FIT TEST ON THE UNIFORM DISTRIBUTION ASSUMPTION"
1592 LOCATE 3.5:PRINT"CHI-SQUARE GOODNESS OF FIT TEST ON THE UNIFORM DISTRIBUTION ASSUMPTION"
1593 KX=CINT((SQR(N)+N/5)/2)
1594 LOCATE 5.19:PRINT"USING";KX;"EQUALY PROBABLE CLASS INTERVALS"
1595 DF=KX-3
1596 GOSUB 1629
1597 E=N/KX
1598 IF CHICOMP=1 THEN 1612
1599 CHIC=0
1600 0=0
1601 IX=1
1602 FOR JX=1 TO K
1603 BOUND=D(1)+(JX/KX)*(D(N)-D(1))
1604 IF I%>N THEN 1609
1605 IF D(I%)>BOUND THEN 1609
1606 O=O+1
1607 IX=IX+1
1608 GOTO 1604
1609 CHIC=CHIC+(O-E)^2/E
1610 O=O
1611 NEXT JX
1612 IF SELECTED=O THEN 1627
1613 CHIS=0
1614 O=O
1615 IX=1
1616 FOR JX=1 TO K
1617 BOUND=BOTTOM+(JX/KX)*(TOPL-BOTTOM)
1618 IF I%>N THEN 1624
1619 IF D(I%)>BOUND THEN 1623
1620 O=O+1
1621 IX=IX+1
1622 GOTO 1618
1623 IF J%=K% THEN 1620
1624 CHIS=CHIS+(O-E)^2/E
1625 O=O
1626 NEXT JX
1627 CHICOMP=1
1628 RETURN

1629 ' Subroutine which determines the critical Chi-Square values
1630 ' Entry parameter:
1631 ' DF is the degrees of freedom (1<=DF<16)
1632 ' Exit parameters:
1633 ' CHI1 is the critical value at alpha=0.01
1634 ' CHI5 is the critical value at alpha=0.05
1635 ON DF GOTO 1637, 1638, 1639, 1640, 1641, 1642, 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651
1637 CHI1=6.63:CHI5=3.84:RETURN
1638 CHI1=9.21:CHI5=5.99:RETURN
1639 CHI1=11.34:CHI5=7.81:RETURN
1640 CHI1=13.28:CHI5=9.49:RETURN
1641 CHI1=15.1:CHI5=11.1:RETURN
1642 CHI1=16.8:CHI5=12.6:RETURN
1643 CHI1=18.5:CHI5=14.1:RETURN
1644 CHI1=20.1:CHI5=15.5:RETURN
1645 CHI1=21.7:CHI5=16.9:RETURN
1646 CHI1=23.2:CHI5=18.3:RETURN
1647 CHI1=24.7:CHI5=19.7:RETURN
1648 CHI1=26.2:CHI5=21.1:RETURN
1649 CHI1=27.7:CHI5=22.4:RETURN
1650 CHI1=29.1:CHI5=23.7:RETURN
Subroutine which computes unique normal distribution
Chi-Square values

Entry parameters:
N is the sample size
D(I) are the sample data points
MEANC is the computed mean
VARIANCEC is the computed variance

Exit parameters:
CHIC is the Chi-Square value of the computed parameters
CHIS is the Chi-Square value of the selected parameters
CHI1 is the critical Chi-Square value at alpha=0.01
CHI5 is the critical Chi-Square value at alpha=0.05

LOCATE 3,5:PRINT"CHI-SQUARE GOODNESS OF FIT TEST ON THE NORMAL DISTRIBUTION ASSUMPTION"
K%=CINT((SQR(N)+N/5)/2)
LOCATE 5,19:PRINT"USING";K%;"EQUALLY PROBABLE CLASS INTERVALS"

IF CHICOMP=1 AND SELECTED=0 THEN 1716
GOSUB 2058
DF=K%-3
GOSUB 1629
E=N/K%
OPEN "I",#1,"NORMAL.DAT"
INPUT#1,Z2
INPUT#1,Z3
Zl=-398
IF CHICOMP=0 THEN CHIC=0
CHIS=0
IC=1
IS=1
FOR J%=1 TO K%
GOSUB 2059
OC=0
OS=0
IF J%=K% THEN 1695
Z=J%/K%
IF Z3>Z THEN 1694
Z2=Z3
INPUT#1,Z3
Zl=Zl+1
GOTO 1689
Z4=(Z1-(Z3-Z)/(Z3-Z2))/100
IF CHICOMP=1 THEN 1704
BOUND=MEANC+Z4*SQR(VARIANCEC)
IF IC>N THEN 1703
IF D(IC)>BOUND THEN 1702
OC=OC+1
IC=IC+1
Subroutine which computes unique exponential distribution
Chi-Square values
Entry parameters:
N is the sample size
D(1) are the sample data points
MEANC is the computed mean
MEANS is the selected mean
Exit parameters:
CHIC is the Chi-Square value of the computed parameters
CHIS is the Chi-Square value of the selected parameters
CHI1 is the critical Chi-Square value at alpha=0.01
CHI5 is the critical Chi-Square value at alpha=0.05
1751 GOTO 1747
1752 IF J%=K% THEN 1749
1753 CHIC=CHIC+(OC-E)^2/E
1754 IF SELECTED=0 THEN 1764
1755 IF J%=K% THEN 1757
1756 BOUND=MEANS*LOG(1-J%/K%)
1757 IF IS>N THEN 1763
1758 IF D(IS)>BOUND THEN 1762
1759 OS=OS+1
1760 IS=IS+1
1761 GOTO 1757
1762 IF J%=K% THEN 1759
1763 CHIS=CHIS+(OS-E)^2/E
1764 NEXT J%
1765 CHICOMP=1
1766 RETURN
1767 ' Subroutine which computes unique gamma distribution
1768 ' Chi-Square values
1769 ' Entry parameters:
1770 ' N is the sample size
1771 ' D(I) are the sample data points
1772 ' B is the computed scale parameter
1773 ' C is the computed shape parameter
1774 ' BS is the selected scale parameter
1775 ' CS is the selected shape parameter
1776 ' Exit parameters:
1777 ' CHIC is the Chi-Square value of the computed parameters
1778 ' CHIS is the Chi-Square value of the selected parameters
1779 ' CHI1 is the critical Chi-Square value at alpha=0.01
1780 ' CHI5 is the critical Chi-Square value at alpha=0.05
1781 LOCATE 3,6:PRINT"CHI-SQUARE GOODNESS OF FIT TEST ON THE GAMMA DISTRIBUTION ASSUMPTION"
1782 LOCATE 3,6:PRINT"USING";K%;"EQUALLY PROBABLE CLASS INTERVALS"
1783 KX=CINT(SQR(N)+N/5)/2
1784 LOCATE 5,19:PRINT"USING";KX;"EQUALLY PROBABLE CLASS INTERVALS"
1785 GOSUB 2058
1786 IF CHICOMP=1 THEN 1792
1787 DF=KX-3
1788 GOSUB 1629
1789 CA=C
1790 GOSUB 1556
1791 GAMMAC=GAMMA
1792 IF SELECTED=0 THEN 1796
1793 CA=CS
1794 GOSUB 1556
1795 GAMMAS=GAMMA
1796 DI=10
1797 PROBC=0
1798 PROBS=0
1799 LLC=0
1800 LLS=0
1801 IC=1
1802 IS=1
1803 IF CHICOMP=0 THEN CHIC=0
1804 CHIS=0
1805 TOGOC=N
1806 TOGOS=N
1807 ON ERROR GOTO 1871
1808 FOR J%=1 TO K%
1809 GOSUB 2059
1810 IF CHICOMP=1 THEN 1838
1811 0=0
1812 IF J%=K% THEN E=TOGOC:GOTO 1827
1813 LL=LLC
1814 UL=0
1815 ULC=2*B*C*J%/K%
1816 PROB=PROBC
1817 IF ULC <= LLC THEN ULC=2*ULC:GOTO 1817
1818 FOR I%=1 TO DI
1819 Z=LLC+I%*(ULC-LLC)/DI-(ULC-LLC)/(2*DI)
1820 FZ=((Z/B)^-(C-1))*EXP(-Z/B))/(B*GAMMAC)
1821 PROB=PROB+FZ*(ULC-LLC)/DI
1822 NEXT I%
1823 IF PROB>(J%/K%+1/(100*K%)) THEN UL=ULC:ULC=(UL+LL)/2:GOTO 1816
1824 IF PROB<(J%/K%-1/(100*K%)) AND UL=0 THEN LL=ULC:ULC=2*ULC:GOTO 1816
1825 IF PROB<(J%/K%-1/(100*K%)) THEN LL=ULC:ULC=(UL+LL)/2:GOTO 1816
1826 E=(PROB-PROBC)*N
1827 IF IC>N THEN 1833
1828 IF D(IC)>ULC THEN 1832
1829 0=0+1
1830 IC=IC+1
1831 GOTO 1827
1832 IF J%=K% THEN 1829
1833 IF E=0 THEN 1836
1834 CHIC=CHIC+(0-E)^2/E
1835 TOGOC=TOGOC-E
1836 PROBC=PROB
1837 LLC=ULC
1838 IF SELECTED=0 THEN 1866
1839 0=0
1840 IF J%=K% THEN E=TOGOS:GOTO 1855
1841 LL=LLS
1842 UL=0
1843 ULS=2*BS*CS*J%/K%
1844 PROB=PROBS
1845 IF ULS <= LLS THEN ULS=2*ULS:GOTO 1845
1846 FOR I%=1 TO DI
1847 Z=LLS+I%*(ULS-LLS)/DI-(ULS-LLS)/(2*DI)
1848 FZ=((Z/BS)^-(CS-1))*EXP(-Z/BS))/(BS*GAMMAS)
1849 PROB=PROB+FZ*(ULS-LLS)/DI
1850 NEXT I%
1851 IF PROB>(J%/K%+1/(100*K%)) THEN UL=ULS:ULS=(UL+LL)/2:GOTO 1844
1852 IF PROB<(J%/K%+1/(100*K%)) AND UL=O THEN LL=ULS:ULS=2*ULS:GOTO 1844
1853 IF PROB<(J%/K%+1/(100*K%)) THEN LL=ULS:ULS=(UL+LL)/2:GOTO 1844
1854 E=(PROB-PROBS)*N
1855 IF IS>N THEN 1861
1856 IF D(IS)>ULS THEN
1857 0=0+1
1858 IS=IS+1
1859 GOTO 1855
1860 IF J%+K% THEN 1857
1861 IF E=0 THEN 1864
1862 CHIS=CHIS+(O-E)^2/E
1863 TOGOS=TOGOS-E
1864 PROBS=PROB
1865 LLS=ULS
1866 NEXT J%
1867 CHICOMP=1
1868 ON ERROR GOTO 0
1869 GOSUB 2060
1870 RETURN
1871 RESUME 1872
1872 ON ERROR GOTO 0
1873 GOSUB 2060
1874 LOCATE 8,10:PRINT"PARAMETER(S) ARE OUT OF THE COMPUTATIONAL RANGE OF THIS PROGRAM"
1875 LOCATE 24,27:PRINT"PRESS ANY KEY TO CONTINUE »
1876 A$=INPUT$(1)
1877 E=1
1878 GOTO 1867
1879 ' Subroutine which computes unique Weibull distribution
1880 ' Chi-Square values
1881 ' Entry parameters:
1882 ' N is the sample size
1883 ' D(I) are the sample data points
1884 ' B is the computed scale parameter
1885 ' C is the computed shape parameter
1886 ' BS is the selected scale parameter
1887 ' CS is the selected shape parameter
1888 ' Exit parameters:
1889 ' CHIC is the Chi-Square value of the computed parameters
1890 ' CHIS is the Chi-Square value of the selected parameters
1891 ' CHI1 is the critical Chi-Square value at alpha=0.01
1892 ' CHI5 is the critical Chi-Square value at alpha=0.05
1893 ' LOCATE 3,5:PRINT"CHI-SQUARE GOODNESS OF FIT TEST ON THE WEIBULL DISTRIBUTION ASSUMPTION"
1894 LOCATE 3,5:PRINT"CHI=SQR((N-1)*N/5)/2"
1895 DF=K%-3
1896 LOCATE 5,19:PRINT"USING”;K%;”EQUALLY PROBABLE CLASS INTERVALS"
1897 DF=K%-3
1898 GOSUB 1629
1899 E=N/KX
1900 IF CHICOMP=O THEN CHIC=O
Subroutine which computes unique beta distribution

Chi-Square values

Entry parameters:

N is the sample size
D(I) are the sample data points
V is the computed first shape parameter
W is the computed second shape parameter
VS is the selected first shape parameter
WS is the selected second shape parameter

Exit parameters:
CHIC is the Chi-Square value of the computed parameters
CHIS is the Chi-Square value of the selected parameters
CH11 is the critical Chi-Square value at alpha=0.01
CH15 is the critical Chi-Square value at alpha=0.05
1951 GOSUB 1629
1952 CA=V
1953 GOSUB 1556
1954 GAMMA=1=GAMMA
1955 CA=W
1956 GOSUB 1556
1957 GAMMA=2=GAMMA
1958 CA=V+W
1959 GOSUB 1556
1960 BETA=GAMMA1*GAMMA2/GAMMA
1961 IF SELECTED=0 THEN 1971
1962 CA=VS
1963 GOSUB 1556
1964 GAMMA=1=GAMMA
1965 CA=WS
1966 GOSUB 1556
1967 GAMMA=2=GAMMA
1968 CA=VS+WS
1969 GOSUB 1556
1970 BETA=1=GAMMA1*GAMMA2/GAMMA
1971 DI=10
1972 PROBC=0
1973 PROBS=0
1974 LLC=0
1975 LLS=0
1976 IC=1
1977 IS=1
1978 IF CHICOMP=0 THEN CHIC=0
1979 CHIS=0
1980 TOGOC=N
1981 TOGOS=N
1982 ON ERROR GOTO 1871
1983 FOR J%=1 TO K%
1984 GOSUB 2059
1985 IF CHICOMP=1 THEN 2019
1986 0=0
1987 IF J%=K% THEN E=TOGOC:GOTO 2008
1988 LL=LLC
1989 0=0
1990 UL0=0
1991 UL=J%/K%
1992 PROB=PROBC
1993 IF ULC<= LLC THEN ULC=2*ULC:GOTO 1993
1994 IF ULC=1 THEN IF ULCF=1 THEN 2004 ELSE ULCF=1:ULC=1
1995 FOR I%=1 TO DI
1996 Z=LLC+I%*(ULC-LLC)/DI-(ULC-LLC)/(2*DI)
1997 FZ=Z^*(V-1)*S*(1-Z)^*(W-1)/BETA
1998 PROB=PROB+FZ* (ULC-LLC)/DI
1999 NEXT I%
2000 IF PROB>(J%+.01)/K% THEN ULC=ULC=(UL+LL)/2:GOTO 1992
2001 IF PROB<(J%-.01)/K% AND UL=O THEN LL=ULC:ULC=2*ULC:GOTO 1992
2002 IF PROB<(J%-.01)/K% THEN LL=ULC:ULC=(UL+LL)/2:GOTO 1992
2003 GOTO 2006
2004 ULC=LLC+(1-LLC)/(K%-J%+1)
2005 PROB=PROBC+(1-PROBC)/(K%-J%+1)
2006 E=(PROB-PROBC)*N
2007 BOUND=ULC*(D(N)-D(1))+D(1)
2008 IF IC>N THEN 2014
2009 IF D(IC)>BOUND THEN 2013
2010 O=O+1
2011 IC=IC+1
2012 GOTO 2008
2013 IF J%=K% THEN 2010
2014 IF E=0 THEN 2017
2015 CHIC=CHIC+(O-E)^2/E
2016 TOGOC=TOGOC-E
2017 PROBC=PROB
2018 LLC=ULC
2019 IF SELECTED=O THEN 2053
2020 ULCF=O
2021 O=0
2022 IF J%=K% THEN E=TOGOS:GOTO 2042
2023 LL=LLS
2024 UL=O
2025 ULS=J%/K%
2026 PROB=PROBS
2027 IF ULS <= LLS THEN ULS=2*ULS:GOTO 2027
2028 IF ULS>1 THEN IF ULCF=1 THEN 2038 ELSE ULCF=1:ULCF=1
2029 FOR I%=1 TO DI
2030 Z=LLS+I%*(ULS-LLS)/DI-(ULS-LLS)/(2*DI)
2031 FZ=Z^((VS-1)*(1-Z)/(WS-1)/BETAS
2032 PROB=PROB+FZ*(ULS-LLS)/DI
2033 NEXT I%
2034 IF PROB>(J%+.01)/K% THEN UL=ULS=UL=(UL+LL)/2:GOTO 2026
2035 IF PROB<(J%-.01)/K% AND UL=O THEN LL=ULS=ULS=2*ULS:GOTO 2026
2036 IF PROB<(J%-.01)/K% THEN LL=UL=ULS=UL=(UL+LL)/2:GOTO 2026
2037 GOTO 2040
2038 ULS=LLS+(1-LLS)/(K%-J%+1)
2039 PROB=PROBS+(1-PROBS)/(K%-J%+1)
2040 E=(PROB-PROBS)*N
2041 BOUND=ULS*(D(N)-D(1))+D(1)
2042 IF IS>N THEN 2048
2043 IF D(IS)>BOUND THEN 2047
2044 O=O+1
2045 IS=IS+1
2046 GOTO 2042
2047 IF J%=K% THEN 2044
2048 IF E=0 THEN 2051
2049 CHIS=CHIS+(O-E)^2/E
2050 TOGOS=TOGOS-E
2051 PROBS=PROB
2052 LLS=ULS
2053 NEXT J%
2054 GOSUB 2060
2055 CHICOMP=1
2056 ON ERROR GOTO 0
2057 RETURN
2058 LOCATE 14,22:PRINT"computing ..";:RETURN
2059 PRINT".";:RETURN
2060 LOCATE 14,22:PRINT SPC(50):RETURN
APPENDIX B

MENU SCREENS
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

UNIVERSITY OF CENTRAL FLORIDA

Department of Industrial Engineering and Management Science

Title Menu
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

This program computes and displays the frequency distribution of a given set of data. You then selects the continuous distribution which, in your estimation, best approximates the displayed histogram. The appropriate estimator parameters for the selected distribution are computed. The computed distribution is displayed, superimposed over the input data frequency distribution plot. You can then request that a Chi-Square or goodness of fit test be performed on the computed distribution. At any time, you may change the continuous distribution type selection and thereby cause a recomputation of the distribution estimators and replot of the results.

PRESS ANY KEY TO CONTINUE
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

DATA INPUT

1. Input new data.
2. Use previously saved data.
3. Quit the entire process.

TYPE IN YOUR CHOICE
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

DATA MANIPULATION

1. View the data.
2. Print the data.
3. Change some observations.
4. Delete some observations.
5. Add some observations.
6. Go on to the data analysis program.
7. Start all over again.
8. Quit the entire process.

TYPE IN YOUR CHOICE

Data Manipulation Menu
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

DATA ANALYSIS

1. Display a bar graph of the data.
2. Fit a statistical distribution to the data.
3. Return to the data manipulation section.
4. Start all over again.
5. Quit the entire process.

TYPE IN YOUR CHOICE

Data Analysis Menu
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

DISTRIBUTION FITTING

1. Fit a Uniform Distribution to the data.
2. Fit a Normal Distribution to the data.
3. Fit an Exponential Distribution to the data.
4. Fit a Gamma Distribution to the data.
5. Fit a Weibull Distribution to the data.
6. Fit a Beta Distribution to the data.
7. Return to the bar graph plot section.
8. Return to the data manipulation section.
9. Start all over again.
10. Quit the entire process.

TYPE IN YOUR CHOICE

Distribution Fitting Menu
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

1. Select your own parameters for the chosen distribution.
2. Display a bar graph of the data with selected curve overlayed.
3. Perform a Chi-Square Goodness of Fit Analysis.
4. Fit a different statistical distribution to the data.
5. Return to the bar graph plot section.
6. Return to the data manipulation section.
7. Start all over again.
8. Quit the entire process.

TYPE IN YOUR CHOICE

Parameter Analysis Menu
APPENDIX C

EXAMPLE DISPLAYS
Example Bar Graph Display
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

BEST FIT PARAMETERS FOR A BETA DISTRIBUTION

SHAPE PARAMETERS (v & w) = 2.769716 AND 2.787004

NOTE: These parameters are for a Beta Distribution in the range 0<=x<=1.
To generate data with the same characteristics as your input data,
multiply the probability density function by 90 and then add 5.000001.

<table>
<thead>
<tr>
<th>DISTRIBUTION ATTRIBUTE</th>
<th>VALUE FROM SAMPLE DATA</th>
<th>VALUE OF BEST FIT DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>49.860</td>
<td>49.860</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>308.840</td>
<td>308.840</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>0.025</td>
<td>0.004</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>2.797</td>
<td>2.306</td>
</tr>
</tbody>
</table>
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

BETA DISTRIBUTION PARAMETER SELECTION

<table>
<thead>
<tr>
<th>DISTRIBUTION PARAMETER</th>
<th>COMPUTED VALUE</th>
<th>SELECTED VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRST SHAPE PARAMETER (v)</td>
<td>2.770</td>
<td>2.750</td>
</tr>
<tr>
<td>SECOND SHAPE PARAMETER (w)</td>
<td>2.787</td>
<td>2.750</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISTRIBUTION ATTRIBUTE</th>
<th>VALUE FROM ATTRIBUTE VALUE USING COMPUTED PARAMETERS</th>
<th>VALUE USING SELECTED PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>49.860</td>
<td>49.860</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>308.840</td>
<td>308.840</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>0.025</td>
<td>0.004</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>2.797</td>
<td>2.306</td>
</tr>
</tbody>
</table>

Example Distribution Parameter Comparison Display
Example Bar Chart with Distribution Curve Overlayed Display
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

CHI-SQUARE GOODNESS OF FIT TEST ON THE BETA DISTRIBUTION ASSUMPTION

USING 15 EQUALLY PROBABLE CLASS INTERVALS

<table>
<thead>
<tr>
<th>LEVEL OF SIGNIFICANCE (alpha)</th>
<th>CRITICAL CHI-SQUARE</th>
<th>CHI-SQUARE FROM COMPUTED PARAMETERS</th>
<th>CHI-SQUARE FROM SELECTED PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>26.20</td>
<td>4.07</td>
<td>11.37</td>
</tr>
<tr>
<td>0.05</td>
<td>21.00</td>
<td>4.07</td>
<td>11.37</td>
</tr>
</tbody>
</table>

CONCLUSIONS:

You CANNOT reject the COMPUTED PARAMETERS at 0.01 level of significance.
You CANNOT reject the COMPUTED PARAMETERS at 0.05 level of significance.
You CANNOT reject the SELECTED PARAMETERS at 0.01 level of significance.
You CANNOT reject the SELECTED PARAMETERS at 0.05 level of significance.

Example Ch-Square Goodness of Fit Analysis Display
APPENDIX D

EXAMPLE PROBLEM RESULTS
Multiply all range boundary values by 10.

Unsized Bar Chart Display
Expanded Bar Chart Display
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

BEST FIT PARAMETER FOR AN EXPONENTIAL DISTRIBUTION

MEAN \((1/\Lambda)\) = 11.89348

<table>
<thead>
<tr>
<th>DISTRIBUTION ATTRIBUTE</th>
<th>VALUE FROM SAMPLE DATA</th>
<th>VALUE OF BEST FIT DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>11.893</td>
<td>11.893</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>610.196</td>
<td>141.455</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>3.696</td>
<td>2.000</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>18.244</td>
<td>9.000</td>
</tr>
</tbody>
</table>

Best Fit Parameter Display for an Exponential Distribution
Histogram with Computed Exponential Distribution Overlayed
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

CHI-SQUARE GOODNESS OF FIT TEST ON THE EXPONENTIAL DISTRIBUTION ASSUMPTION

USING 9 EQUALLY PROBABLE CLASS INTERVALS

<table>
<thead>
<tr>
<th>LEVEL OF SIGNIFICANCE (alpha)</th>
<th>CRITICAL CHI-SQUARE</th>
<th>CHI-SQUARE FROM COMPUTED PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>18.50</td>
<td>42.88</td>
</tr>
<tr>
<td>0.05</td>
<td>14.10</td>
<td>42.88</td>
</tr>
</tbody>
</table>

CONCLUSIONS:

You SHOULD reject the COMPUTED PARAMETERS at 0.01 level of significance.
You SHOULD reject the COMPUTED PARAMETERS at 0.05 level of significance.

Computed Exponential Distribution Chi-Square Test Display
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

BEST FIT PARAMETERS FOR A WEIBULL DISTRIBUTION

SCALE PARAMETER (b or beta) = 6.235281
SHAPE PARAMETER (c or alpha) = .5259377

<table>
<thead>
<tr>
<th>DISTRIBUTION ATTRIBUTE</th>
<th>VALUE FROM SAMPLE DATA</th>
<th>VALUE OF BEST FIT DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>11.893</td>
<td>11.408</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>610.196</td>
<td>566.168</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>3.696</td>
<td>5.847</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>18.244</td>
<td>69.712</td>
</tr>
</tbody>
</table>

Best Fit Parameter Display for a Weibull Display
Histogram with Computed Weibull Distribution Overlayed
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

CHI-SQUARE GOODNESS OF FIT TEST ON THE WEIBULL DISTRIBUTION ASSUMPTION

USING 9 EQUALLY PROBABLE CLASS INTERVALS

<table>
<thead>
<tr>
<th>LEVEL OF SIGNIFICANCE (alpha)</th>
<th>CRITICAL CHI-SQUARE</th>
<th>CHI-SQUARE FROM COMPUTED PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>16.80</td>
<td>4.36</td>
</tr>
<tr>
<td>0.05</td>
<td>12.60</td>
<td>4.36</td>
</tr>
</tbody>
</table>

CONCLUSIONS:

You CANNOT reject the COMPUTED PARAMETERS at 0.01 level of significance.
You CANNOT reject the COMPUTED PARAMETERS at 0.05 level of significance.

Computed Weibull Distribution Chi-Square Test Display
## Weibull Distribution Parameter Comparison

<table>
<thead>
<tr>
<th>Distribution Parameter</th>
<th>Computed Value</th>
<th>Selected Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale (b or beta)</td>
<td>6.235</td>
<td>6.000</td>
</tr>
<tr>
<td>Shape (c or alpha)</td>
<td>0.526</td>
<td>0.500</td>
</tr>
</tbody>
</table>

### Distribution Value from Value Using Value Using Attribute Sample Data Computed Parameters Selected Parameters

<table>
<thead>
<tr>
<th>Distribution Attribute</th>
<th>Value from Sample Data</th>
<th>Value Using Computed Parameters</th>
<th>Value Using Selected Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>11.893</td>
<td>11.408</td>
<td>12.000</td>
</tr>
<tr>
<td>Variance</td>
<td>610.196</td>
<td>566.168</td>
<td>720.000</td>
</tr>
<tr>
<td>Skewness</td>
<td>3.696</td>
<td>5.847</td>
<td>6.529</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>18.244</td>
<td>69.712</td>
<td>87.780</td>
</tr>
</tbody>
</table>

Selected Parameter Display for a Weibull Distribution
Histogram with Selected Weibull Distribution Overlaid
CONTINUOUS DISTRIBUTION CURVE FITTING AND ANALYSIS TOOL

CHI-SQUARE GOODNESS OF FIT TEST ON THE WEIBULL DISTRIBUTION ASSUMPTION

USING 9 EQUALLY PROBABLE CLASS INTERVALS

<table>
<thead>
<tr>
<th>LEVEL OF SIGNIFICANCE (alpha)</th>
<th>CRITICAL CHI-SQUARE</th>
<th>CHI-SQUARE FROM COMPUTED PARAMETERS</th>
<th>CHI-SQUARE FROM SELECTED PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>16.80</td>
<td>4.36</td>
<td>5.08</td>
</tr>
<tr>
<td>0.05</td>
<td>12.60</td>
<td>4.36</td>
<td>5.08</td>
</tr>
</tbody>
</table>

CONCLUSIONS:

You CANNOT reject the COMPUTED PARAMETERS at 0.01 level of significance.
You CANNOT reject the COMPUTED PARAMETERS at 0.05 level of significance.
You CANNOT reject the SELECTED PARAMETERS at 0.01 level of significance.
You CANNOT reject the SELECTED PARAMETERS at 0.05 level of significance.
LIST OF REFERENCES


