ProUCL Version 3.0
User Guide
April 2004

EPA/600/R04/079
April 2004

ProUCL Version 3.0
User Guide

by

Anita Singh
Lockheed Martin Environmental Services
1050 E. Flamingo Road, Suite E120, Las Vegas, NV 89119

Ashok K. Singh
Department of Mathematical Sciences
University of Nevada, Las Vegas, NV 89154

Robert W. Maichle
Lockheed Martin Environmental Services
1050 E. Flamingo Road, Suite E120, Las Vegas, NV 89119

Table of Contents

Appendix A

TECHNICAL BACKGROUND - METHODS FOR COMPUTING THE EPC TERM ((1-α) 100%UCL) AS INCORPORATED IN ProUCL VERSION 3.0 SOFTWARE

Appendix B

CRITICAL VALUES OF ANDERSON-DARLING TEST STATISTIC AND KOLMOGOROV-SMIRNOV TEST STATISTIC FOR GAMMA DISTRIBUTION WITH UNKNOWN PARAMETERS

Appendix C

GRAPHS SHOWING COVERAGE COMPARISONS FOR THE VARIOUS METHODS FOR NORMAL, GAMMA, AND LOGNORMAL DISTRIBUTIONS

Disclaimer

The United States Environmental Protection Agency (EPA) through its Office of Research and Development funded and managed the research described here. It has been peer reviewed by the EPA and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation by the EPA for use.

ProUCL software was developed by Lockheed Martin under a contract with the EPA and is made available through the EPA Technical Support Center in Las Vegas, Nevada.

Use of any portion of ProUCL that does not comply with the ProUCL User Guide is not recommended.

ProUCL contains embedded licensed software. Any modification of the ProUCL source code may violate the embedded licensed software agreements and is expressly forbidden.

ProUCL software provided by the EPA was scanned with McAfee VirusScan v4.5.1 SP1 and is certified free of viruses.

With respect to ProUCL distributed software and documentation, neither the EPA nor any of their employees, assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed. Furthermore, software and documentation are supplied "as-is" without guarantee or warranty, expressed or implied, including without limitation, any warranty of merchantability or fitness for a specific purpose.

Executive Summary

Exposure assessment and cleanup decisions in support of U.S. Environmental Protection Agency (EPA) projects are often made based upon the mean concentrations of the contaminants of potential concern. A 95% upper confidence limit (*UCL*) of the unknown population arithmetic mean (*AM*), μ_1, is often used to:

- Estimate the exposure point concentration (EPC) term,
- Determine the attainment of cleanup standards,
- Estimate background level mean contaminant concentrations, or
- Compare the soil concentrations with site specific soil screening levels.

It is important to compute a reliable, conservative, and stable 95% *UCL* of the population mean using the available data. The 95% *UCL* should approximately provide the 95% coverage for the unknown population mean, μ_1 .

The EPA has issued guidance for calculating the *UCL* of the unknown population mean for hazardous waste sites, and ProUCL software has been developed to compute an appropriate 95% *UCL* of the unknown population mean. All *UCL* computation methods contained in the EPA guidance documents are available in ProUCL, Version 3.0. Additionally, ProUCL, Version 3.0 can also compute a 95% *UCL* of the mean based upon the gamma distribution, which is better suited to model positively skewed environmental data sets. ProUCL tests for normality, lognormality, and a gamma distribution of the data set, and computes a conservative and stable *95% UCL* of the unknown population mean, μ_1 . It should be emphasized that the computation of an appropriate *95% UCL* is based upon the assumption that the data set under study consists of observations only from a single population.

Several parametric and distribution-free non-parametric methods are included in ProUCL. The *UCL* computation methods in ProUCL cover a wide range of skewed data distributions arising from the various environmental applications. For lognormally distributed data sets, the use of Land's H-statistic many times yields unrealistically large and impractical *UCL* values. This occurrence is prevalent when the sample size is small and standard deviation of the log-transformed data is large. Gamma distribution has been incorporated in ProUCL to model these types of positively skewed data sets. Singh, Singh, and Iaci (2002b) observed that a *UCL* of the mean based upon a gamma distribution results in reliable and stable values of practical merit. It is always desirable to test if an environmental data set follows a gamma distribution. For data sets (of all sizes) which follow a gamma distribution, the EPC term should be computed using an adjusted gamma *UCL* (when $0.1 \leq k < 0.5$) of the mean or an approximate gamma *UCL* (when $k \geq 0.5$) of the mean. These *UCLs* approximately provide the specified 95% coverage to the population mean, μ_1 of a gamma distribution. For values of $k < 0.1$, a 95% *UCL* may be obtained using the bootstrap-t method or Hall's bootstrap method when the sample size is small ($n < 15$), and for larger samples, a *UCL* of the mean should be computed using the adjusted or approximate gamma *UCL*.

Introduction

The computation of a *(1-α) 100%* upper confidence limit (*UCL*) of the population mean depends upon the data distribution. Typically, environmental data are positively skewed, and a default lognormal distribution (EPA, 1992) is often used to model such data distributions. The H-statistic based Land's (Land 1971, 1975) *H- UCL* of the mean is used in these applications. Hardin and Gilbert (1993), Singh, Singh, and Engelhardt (1997,1999), Schultz and Griffin,1999, Singh et al. (2002a), and Singh, Singh, and Iaci (2002b) pointed out several problems associated with the use of the lognormal distribution and the *H-UCL* of the population *AM*. In practice, for lognormal data sets with high standard deviation (*sd*), σ, of the natural log-transformed data (e.g., σ exceeding 2.0), the *H-UCL* can become unacceptably large, exceeding the 95% and 99% data quantiles, and even the maximum observed concentration, by orders of magnitude (Singh, Singh, and Engelhardt, 1997). This is especially true for skewed data sets of smaller sizes (e.g., n < 50).
The *H-UCL* is also very sensitive to a few low or high values. For example, the addition of a sample with below detection limit measurement can cause the *H-UCL* to increase by a large amount (Singh, Singh, and Iaci, 2002b). Realizing that use of the H-statistic can result in unreasonably large *UCL*, it is recommended (EPA, 1992) to use the maximum observed value as an estimate of the *UCL* (EPC term) in cases where the *H-UCL* exceeds the maximum observed value. Recently, Singh, Singh and Iaci (2002b), and Singh and Singh (2003) studied the computation of the *UCLs* based upon a gamma distribution and several non-parametric bootstrap methods. Those methods have also been incorporated in ProUCL Version 3.0. ProUCL Version 3.0 contains fifteen *UCL* computation methods; five are parametric and ten are non-parametric. The non-parametric methods do not depend upon any of the data distributions.

Both lognormal and gamma distributions can be used to model positively skewed data sets. It should be noted that it is difficult to distinguish between a lognormal and a gamma distribution, especially when the sample size is small (e.g., n < 50). Singh, Singh, and Iaci (2002b) observed that the *UCL* based upon a gamma distribution results in reliable and stable values of practical merit. It is therefore always desirable to test if an environmental data set follows a gamma distribution. For data sets (of all sizes) which follow a gamma distribution, the EPC term should be computed using an adjusted gamma *UCL* (when $0.1 \leq k < 0.5$) of the mean or an approximate gamma *UCL* (when $k \geq 0.5$) of the mean as these *UCLs* approximately provide the specified 95% coverage to the population mean, $\mu_1 = k\theta$ of a gamma distribution. For values of $k < 0.1$, a 95% *UCL* may be obtained using bootstrap-t method or Hall's bootstrap method when the sample size is small (n < 15), and for larger samples a *UCL* of the mean should be computed using the adjusted or approximate gamma *UCL*. For this application, k is the shape parameter of a gamma distribution. It should be noted that both bootstrap-t and Hall's bootstrap methods sometimes result in erratic, inflated, and unstable *UCL* values especially in the presence of outliers. Therefore, these two methods should be used with caution. The user should examine the various *UCL* results and determine if the *UCLs* based upon the bootstrap-t and Hall's bootstrap methods represent reasonable and reliable *UCL* values of practical merit. If the results based upon these two methods are much higher than the rest of methods (except for the *UCLs* based upon lognormal distribution), then this could be an indication of erratic *UCL* values. In case these two bootstrap methods yield erratic and inflated *UCLs*, the *UCL* of the mean should be computed using the adjusted or the approximate gamma *UCL* computation method for highly skewed gamma distributed data sets of small sizes.

ProUCL tests for normality, lognormality, and gamma distribution of a data set, and computes a conservative and stable *95% UCL* of the population mean, μ_1 . It should be emphasized that throughout this User Guide, and in the ProUCL software, it is assumed that one is dealing with a single population. If multiple populations (e.g., background and site data mixed together) are present, it is recommended to separate them out (e.g., using other statistical population partitioning techniques), and respective appropriate 95% *UCLs* should be computed for each of the identified populations. Also, outliers if any should be identified and thoroughly investigated. Outliers when present distort all statistics (mean, *UCLs* etc.) of interest. Decisions about their exclusion (or inclusion) in the data set used to compute the EPC term should be made by all parties involved (e.g., EPA, local agencies, potentially responsible party etc.). The critical values of Anderson-Darling test statistic and Kolmogorov-Smirnov test statistic to test for gamma distribution were generated using Monte Carlo simulation experiments. These critical values are tabulated in Appendix B for various values of the level of significance. Singh, Singh, and Engelhardt (1997,1999), Singh, Singh, and Iaci (2002b), and Singh and Singh (2003) studied several parametric and non-parametric *UCL* computation methods which have been included in ProUCL. Most of the mathematical algorithms and formulas used in the development of ProUCL to compute the various statistics are summarized in Appendix A. For details, the user is referred to Singh, Singh, and Iaci (2002b), and Singh and Singh (2003). ProUCL computes the various summary statistics for raw, as well as log-transformed data. ProUCL defines log-transform (*log*) as the natural logarithm (*ln*) to the base e. ProUCL also computes the maximum likelihood estimates (*MLEs)* and the minimum variance unbiased estimates (*MVUEs*) of various unknown population parameters of normal, lognormal, and gamma distributions. This, of course, depends upon the underlying data distribution. Based upon the data distribution, ProUCL computes the *(1-α) 100% UCLs* of the unknown population mean, μ_1 using five parametric and ten non-parametric methods.

The five parametric *UCL* computation methods include:

1. Student's-t *UCL*,
2. approximate gamma *UCL* using chi-square approximation,
3. adjusted gamma *UCL* (adjusted for level significance),
4. Land's *H-UCL*, and
5. Chebyshev inequality based *UCL* (using *MVUEs* of parameters of a lognormal distribution).

The ten non-parametric methods included in ProUCL are:

1. the central limit theorem (*CLT*) based *UCL*,
2. modified-t statistic (adjusted for skewness) bases *UCL*,
3. adjusted-*CLT* (adjusted for skewness) based *UCL*,
4. Chebyshev inequality based *UCL* (using sample mean and sample standard deviation),
5. Jackknife method based *UCL*,
6. *UCL* based upon standard bootstrap,
7. *UCL* based upon percentile bootstrap,
8. *UCL* based upon bias - corrected accelerated (BCA) bootstrap,
9. *UCL* based upon bootstrap-t, and
10. *UCL* based upon Hall's bootstrap.

An extensive comparison of these methods has been performed by Singh and Singh (2003) using Monte Carlo simulation experiments. It is well known that the Jackknife method (with sample mean as an estimator) and Student's-t method yield identical *UCL* values. However, a typical user may be unaware of this fact. It has been suggested that a 95% *UCL* based upon the Jackknife method may provide adequate coverage to the population mean of skewed distributions, which of course is not true (just like a *UCL* based upon the Student's-t statistic). For the benefit of all ProUCL users, it was decided to retain the Jackknife *UCL* computation method in ProUCL.

The standard bootstrap and the percentile bootstrap *UCL* computation methods do not perform well (do not provide adequate coverage to population mean) for skewed data sets. For skewed distributions, the bootstrap-t and Hall's bootstrap (meant to adjust for skewness) methods do perform better (in terms of coverage for the population mean) than the various other bootstrap methods. However, it has been noted (e.g., see Singh, Singh, and Iaci (2002b), Singh and Singh (2003)) that these two bootstrap methods sometimes yield erratic and inflated *UCL* values (orders of magnitude higher than the other *UCLs*). This is especially true when outliers may be present in a data set. Therefore, whenever applicable (e.g., based upon the findings of Singh and Singh (2003)), ProUCL provides a caution statement regarding the use of these two bootstrap methods. ProUCL software provides warning messages whenever the use of these methods is recommended. However, for the sake of completeness, all of the parametric as well as non-parametric methods have been included in ProUCL.

The use of some other methods (e.g., bias-corrected accelerated bootstrap method) that were not included in ProUCL Version 2.1 was suggested by some practitioners due to opinions that these omitted methods may perform better than the various other methods already incorporated in ProUCL. In order to satisfy all users, ProUCL Version 3.0 has several additional *UCL* computation methods which were not included in ProUCL, Version 2.1.

This User Guide contains software installation instructions and brief descriptions for each window in the ProUCL software menu. A short glossary of terms used in this document and in the ProUCL program is also provided.

Three appendices listed as follows provide additional information and details of the various methods and references used in the development of ProUCL Version 3.0.

- Appendix A is a discussion of the methods incorporated into ProUCL for calculating the exposure point concentration term using the various methods and distributions. Appendix A represents a stand-alone technical writeup of the various methods incorporated in ProUCL and is provided for review by statistically advanced users. There is duplication between some of the information provided in the main body of the User Guide and Appendix A. This duplication is intentional since Appendix A is designed to be a stand-alone technical discussion of the methods incorporated into ProUCL.
- Appendix B contains the tables of the critical values of the Anderson-Darling Test statistic and Kolmogorov-Smirnov Test statistic for gamma distribution for various levels of significance.
- Appendix C has the graphs from Singh and Singh (2003) showing coverage comparisons (achieved confidence coefficient) for the various *UCL* computation methods for normal, gamma, and lognormal distributions as incorporated in ProUCL software package.

Should the Maximum Observed Concentration be Used as an Estimate of the EPC Term?

Singh and Singh (2003) also included the Max Test (using the maximum observed value as an estimate of the EPC term) in their simulation studies. In the past (e.g., EPA 1992 RAGS Document), the use of the maximum observed value has been recommended as a default value to estimate the EPC term when a 95% *UCL* (e.g., the *H-UCL*) exceeded the maximum value. However, (e.g., EPA 1992), only two 95% UCL computation methods, namely: the Student's- t UCL and Land's *H-UCL* were used to estimate the EPC term. Today, ProUCL, Version 3.0 can compute a 95% *UCL* of the mean using several methods based upon normal, gamma, lognormal, and non-parametric distributions. Thus, ProUCL, Version 3.0 has about fifteen 95% *UCL* computation methods, at least one of which (depending upon skewness and data distribution) can be used to compute an appropriate estimate of the EPC term. Furthermore, since the EPC term represents the average exposure contracted by an individual over an exposure area (EA) during a long period of time, therefore, the EPC term should be estimated by using an average value (such as an appropriate 95% *UCL* of the mean) and not by the maximum observed concentration. With the availability of the *UCL* computation methods, the developers of ProUCL Version 3.0 do not consider it necessary to use the maximum observed value as an estimate of the EPC term. Singh and Singh (2003) also noted that for skewed data sets of small sizes (e.g., n < 10 - 20), the Max Test does not provide the specified 95% coverage to the population mean, and for larger data sets, it overestimates the EPC term. This can also viewed in the graphs presented in Appendix

C. Also, for the skewed distributions (gamma, lognormal) considered, the maximum value is not a sufficient statistic for the unknown population mean. The use of the maximum value as an estimate of the EPC term ignores most (except for the maximum value) of the information contained in a data set. It is, therefore not desirable to use the maximum observed value as estimate of the EPC term representing average exposure to an individual over an EA.

It should also be noted that for highly skewed data sets, the sample mean may exceed the upper 90%, 95 %, etc. percentiles, and consequently, a 95% *UCL* of the mean can exceed the maximum observed value of a data set. This is especially true when one is dealing with highly skewed lognormally distributed data sets of small sizes. For such highly skewed data sets which can not be modeled by a gamma distribution, a 95% *UCL* of the mean should be computed using an appropriate non-parametric method. These observations are summarized in Tables 1-3 of this User Guide.

Alternatively, for such highly skewed data sets, other measures of central tendency such as the median (or some higher order quantile such as 70% etc.) and its upper confidence limit may be considered. The EPA and all other interested agencies and parties need to come to an agreement on the use of median and its *UCL* to estimate the EPC term. However, the use of the sample median and/or its *UCL* as estimates of the EPC term needs further research and investigation.

It is recommended that the maximum observed value NOT be used as an estimate of the EPC term. For the sake of interested users, the ProUCL displays a warning message when the recommended 95% *UCL* (e.g., Hall's bootstrap *UCL* etc.) of the mean exceeds the observed maximum concentration. For such cases (when a 95% *UCL* does exceed the maximum observed value), if applicable, an alternative 95% *UCL* computation method is recommended by ProUCL.

Handling of Non-Detects

ProUCL does not handle left-censored data sets with non-detects, which are inevitable in many environmental applications. All parametric as well as non-parametric recommendations (as summarized in Tables 1-3) to compute the mean, standard deviation, 95% *UCLs* and all other statistics computed by ProUCL are based upon full data sets without censoring. It should be noted that for mild to moderate number of non-detects (e.g., < 15%), one may use the commonly used ½ detection limit (½ DL) proxy method to compute the various statistics. However, the proxy methods should be used cautiously, especially when one is dealing with lognormally distributed data sets. For lognormally distributed data sets of small sizes, even a single value -- small (e.g., obtained after replacing the non-detects by ½ DL) or large (e.g., an outlier) can have a drastic influence (can yield an unrealistically large 95% *UCL*) on the value of the associated Land's 95% *UCL*. The issue of estimating the mean, standard deviation, and an appropriate 95% *UCL* of the mean based upon left-censored data sets with varying degrees of censoring (e.g., 15% - 50%, 50% - 75%, greater than 75% etc.) is currently under investigation.

Installation Instructions

- **Caution**: If you have previous versions of the ProUCL which were installed, you should remove or rename the directory in which that version is currently located.
- Download the file SETUP.EXE from the EPA website and save to a temporary location.
- Run the SETUP.EXE program. This will create a ProUCL directory and two folders; USER GUIDE and the DATA (sample data).
- To run the program, use Windows Explorer to locate the ProUCL application file and double click on it, or use the RUN command from the start menu to locate and run ProUCL.exe.
- To uninstall the program, use Windows Explorer to locate and delete the ProUCL folder.

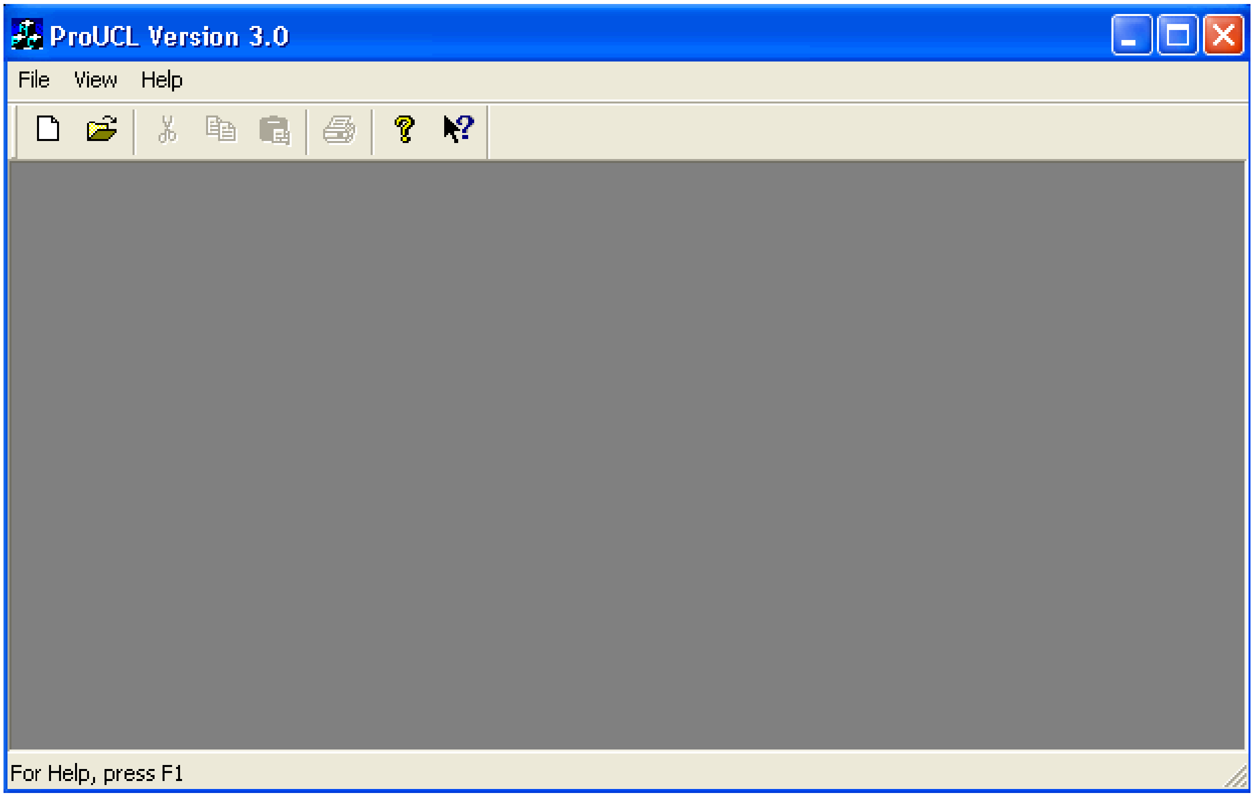
Minimum Hardware Requirements

- Intel Pentium 200MHz
- 12 MB of hard drive space
- 48 MB of memory (RAM)
- CD-ROM drive
- Windows 98 or newer. ProUCL was thoroughly tested on NT-4, Windows 2000, and Windows XP operating systems. Limited testing has been conducted on Windows ME.

A. ProUCL Menu Structure

ProUCL contains a pull-down menu structure, similar to a typical Windows program.

The screen below appears when the program is executed.



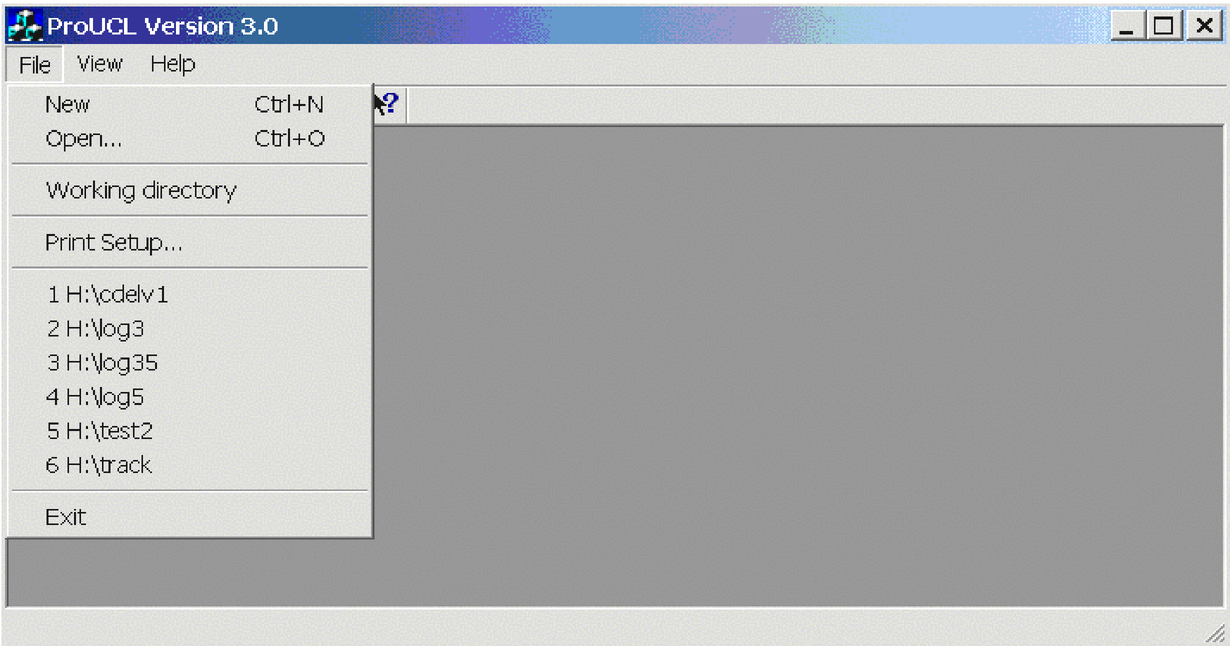
The following menu options appear on the screen

1. File
2. View
3. Help

The options available with these menu items are described on the following pages.

1. File

Click on the File menu item to reveal these drop-down menu options.

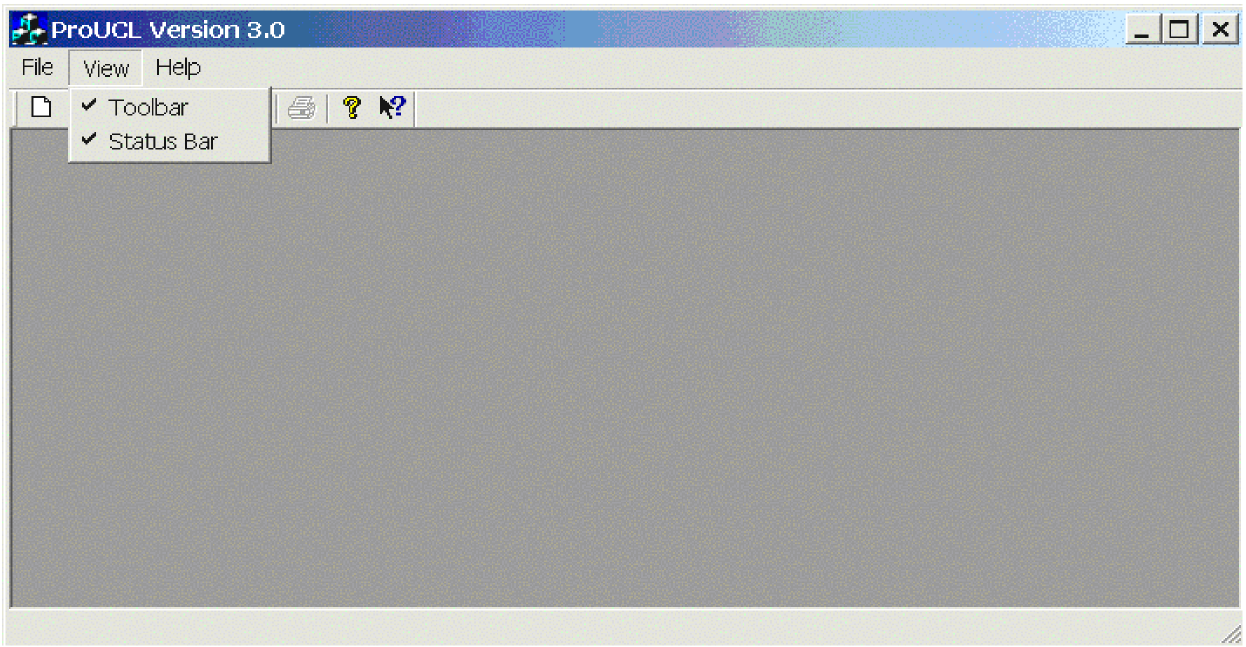


The following File drop-down menu options are available:

- New option: creates new spreadsheet.
- Open option: browses the disk for a file. The browse program will start in the working directory if a directory has been set.
- Working directory option: select and set a working directory.
  Note: A file from the directory must be selected before setting the directory. All subsequent files are read from and saved in the chosen working directory.
- Print Setup option: sets printer options. For example, one can choose the landscape format.
- Click on a previously used file to re-open that file.
- Exit option: exits ProUCL.

2. View

Click on the View menu item to reveal these drop-down options.



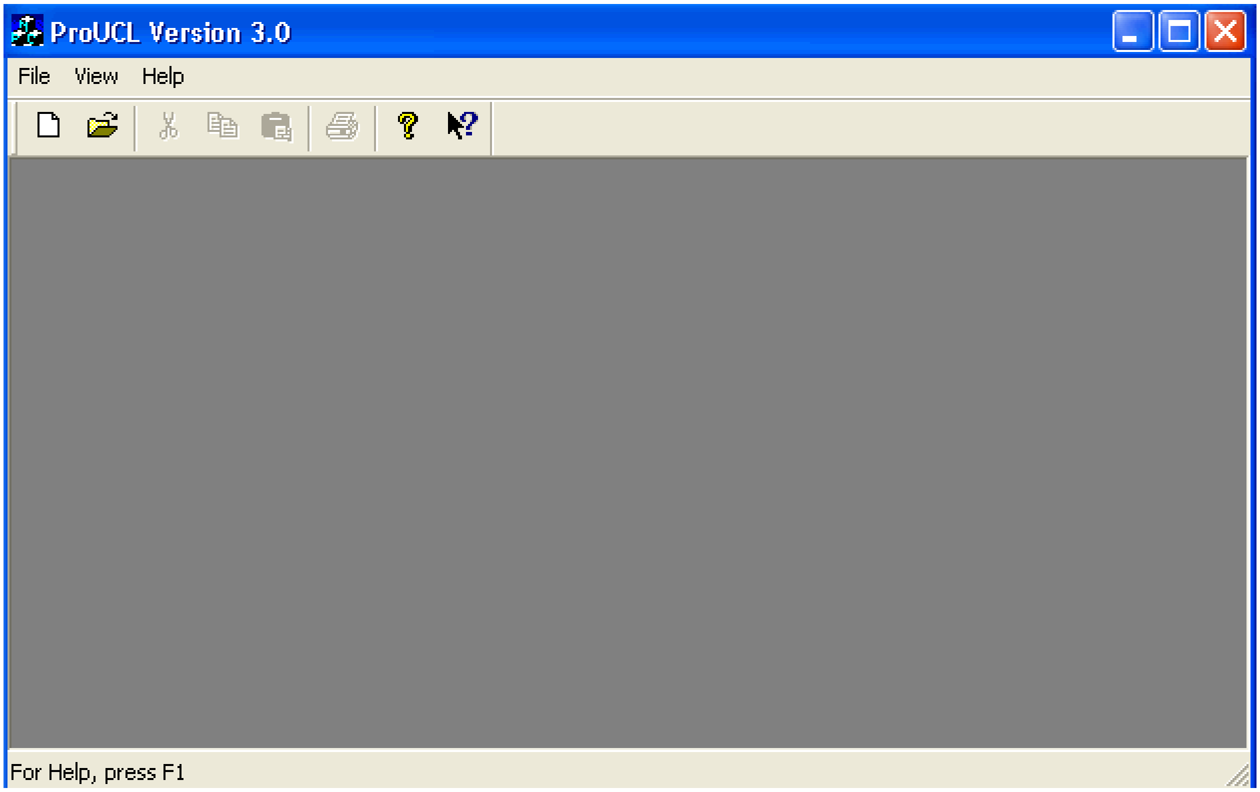
The following View drop-down menu options are available:

- Toolbar: the Toolbar is that row of symbols immediately below the menu items. Clicking on this option toggles the display. This is useful if the user wants to view more data on the screen.

- Status Bar: the Status Bar is the wide bar at the bottom of the screen which displays helpful information. Clicking on this option toggles the display. This is useful if the user wants to view more data on the screen.

3. Help

Click on the Help menu item to reveal these drop-down options.

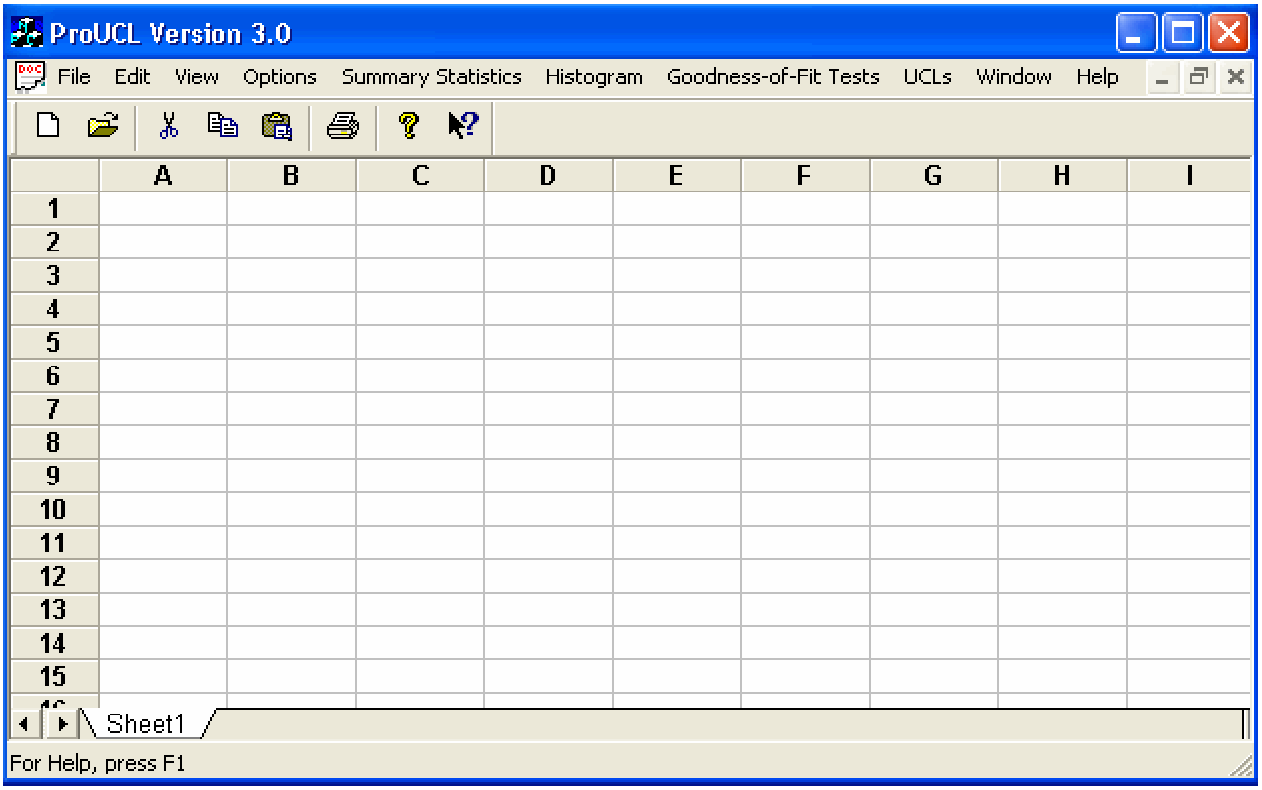


The following Help drop-down menu options are available:

- Help Topics: help topics have not been developed for Version 3.0.
- About ProUCL: displays the software version number.

B. ProUCL Components

The following menu structure of ProUCL appears after opening or creating a data file.



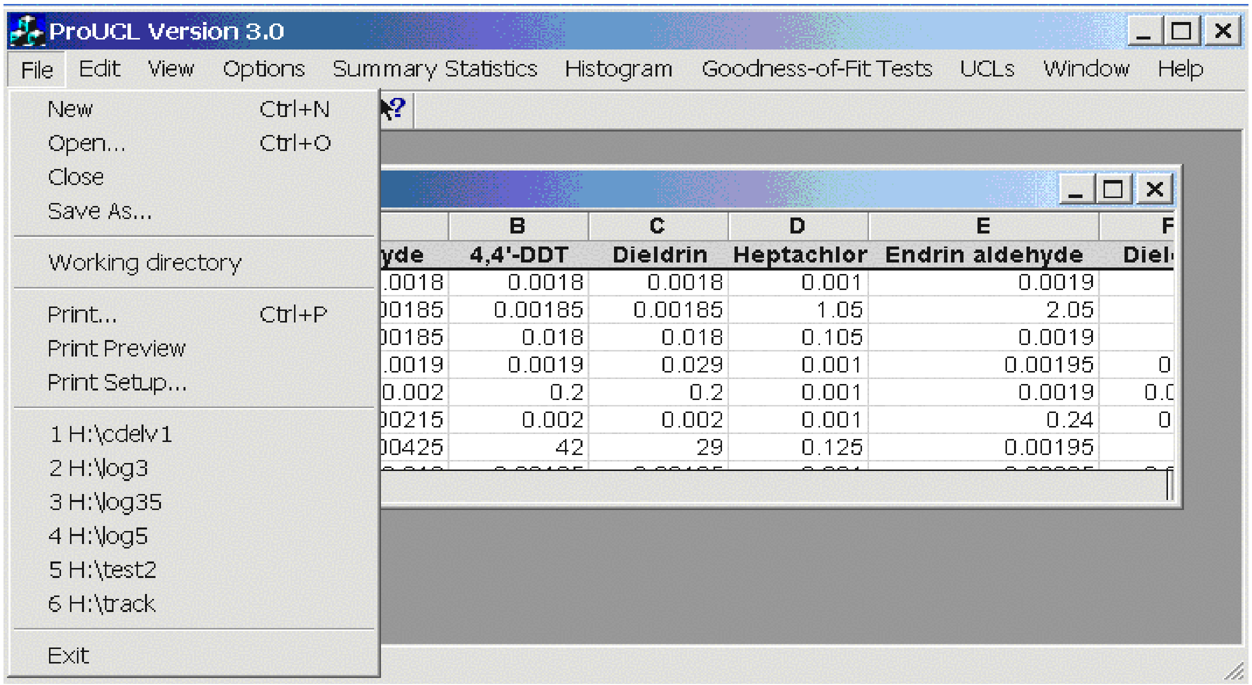
The following menu items are available.

1. File
2. Edit
3. View
4. Options
5. Summary Statistics
6. Histogram
7. Goodness-of-Fit Tests
8. UCLs
9. Window
10. Help

The options available with these menu items are described on the following pages.

1. File

Click on the File menu item to reveal these drop-down options.



The following File drop-down menu options are available:

- New option: opens a blank spreadsheet screen.

- Open option: browses the disk and selects a file which is then opened in spreadsheet format. The browse program will start in the working directory if a working directory has been set.

  Recognized input format options:

Excel	*.xls
Text	*.txt (tab delimited)
Lotus	*.wk?
Lotus	*.123
Default	*.* will be read in Excel format.

- Close option: closes the active window.

- Save As option: allows the user to save the active window. This option follows the Windows standard and saves the active window to a file in Excel 95 (or higher) format. All modified/edited data files, and output screens generated by the software, can be saved in Excel 95 (or higher) format.

- Working directory option: selects and sets a working directory for all I/O operations. All subsequent files are read from and saved in the working directory. <u>You must select a file before you set the working directory.</u>

- Print option: sends the active window to the printer.

- Print Preview option: displays a preview of the output on the screen.

- Print Setup option: follows Windows standard. The user can choose the landscape format under this option.

- Previously opened files: click on a previously used file to re-open that file.

- Exit option: exits ProUCL.

NOTE: All subsequent screens and examples in this User Guide use the spreadsheets given by *track.xls and Cdelv1.xls* to illustrate the various goodness-of-fit test procedures and the *UCL* computation methods as incorporated in the software ProUCL, Version 3.0.

1a. Input File Format

- Data in each column must end with a non-zero value. The last non-zero entry in each column is considered as the end of that column's data. If your data column ends with a zero value, that last zero value will be ignored. This may require you to move observations around if your column ends with zero values.

- The program can read tab delimited Text (ASCII), Excel, and Lotus files.

- Columns in a Text (ASCII) file should be separated by one tab. Spaces between columns are not allowed in this format.

- All input data files should have column labels in the first row and numerical data without text (e.g., non-numeric characters and blank values) for those variables in the remaining rows.

- The data file can have multiple variables (columns) with unequal number of observations.

- Non-numeric text may only appear in the header row (first row) of each column. All other non-numeric data (blank, other characters, and strings) appearing elsewhere in the data file are treated as zero entries. The user should make sure that his data set does not contain such non-numeric values.

- A large value, such as 1E31 ($1x10^{31}$), can be used for missing (alpha numeric text or blank values) data. All entries with this value are ignored from the computations.

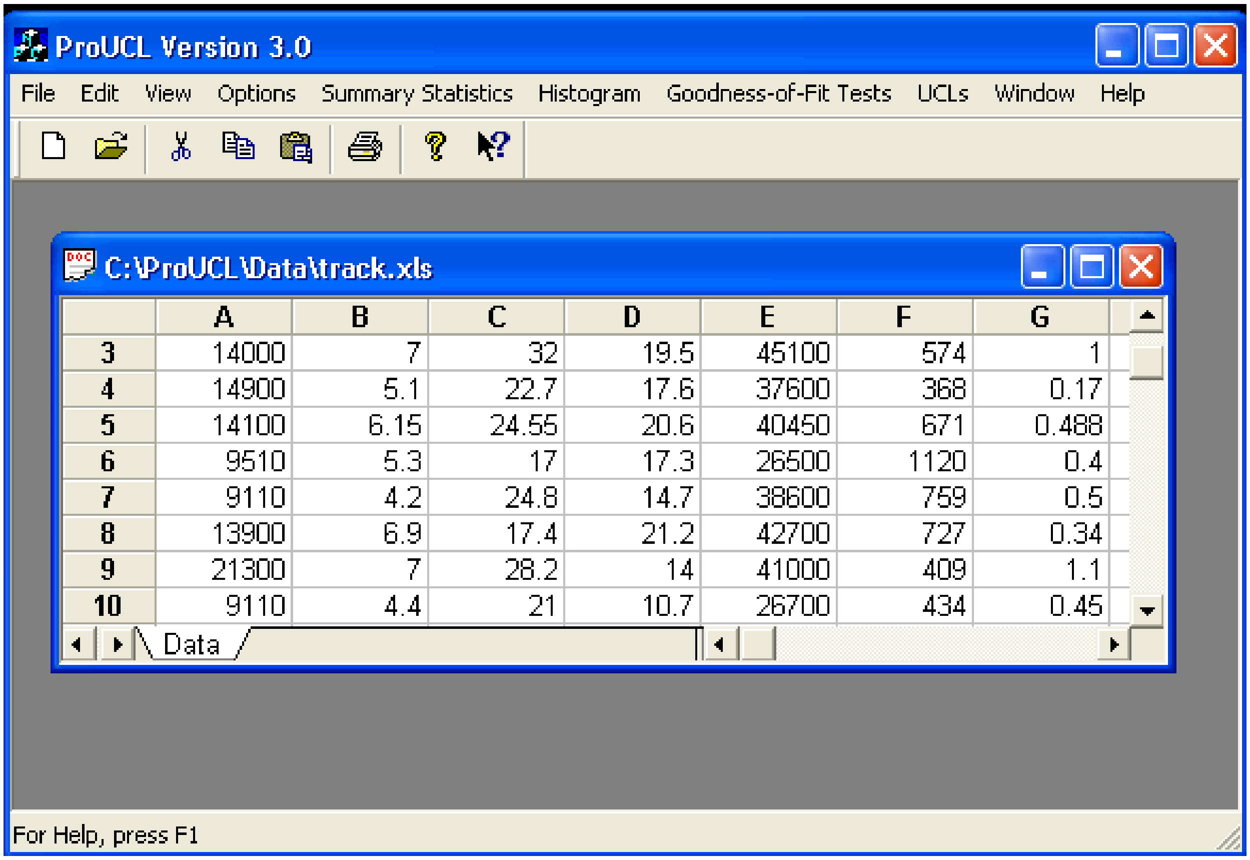
- Note that all other zero data (in the beginning or middle of a data column) are treated as valid zero values.

- ProUCL does not handle the left-censored data sets with non-detects which are inevitable in environmental applications. All parametric as well as non-parametric recommendations made by ProUCL are based upon full data sets without censoring. The issue of estimating the mean, standard deviation, and a 95% UCL of the mean based upon left-censored data sets with varying degrees of censoring is currently under investigation. For mild to moderate number of non-detects (e.g., < 15%), one may use the commonly used ½ detection limit (DL) proxy method. However, the proxy methods should be used cautiously, especially when one is dealing with lognormally distributed data sets. For lognormally distributed data sets of small sizes, a single value, whether small (e.g., obtained after replacing the non-detect by ½ DL) or large (e.g., an outlier), can have a drastic influence (can yield an unrealistically large 95% UCL) on the value of the associated Land's 95% UCL.

1b. Result of Opening an Input Data File

- The data screen follows the standard Windows design. It can be resized, or portions of data can be viewed using scroll bars.

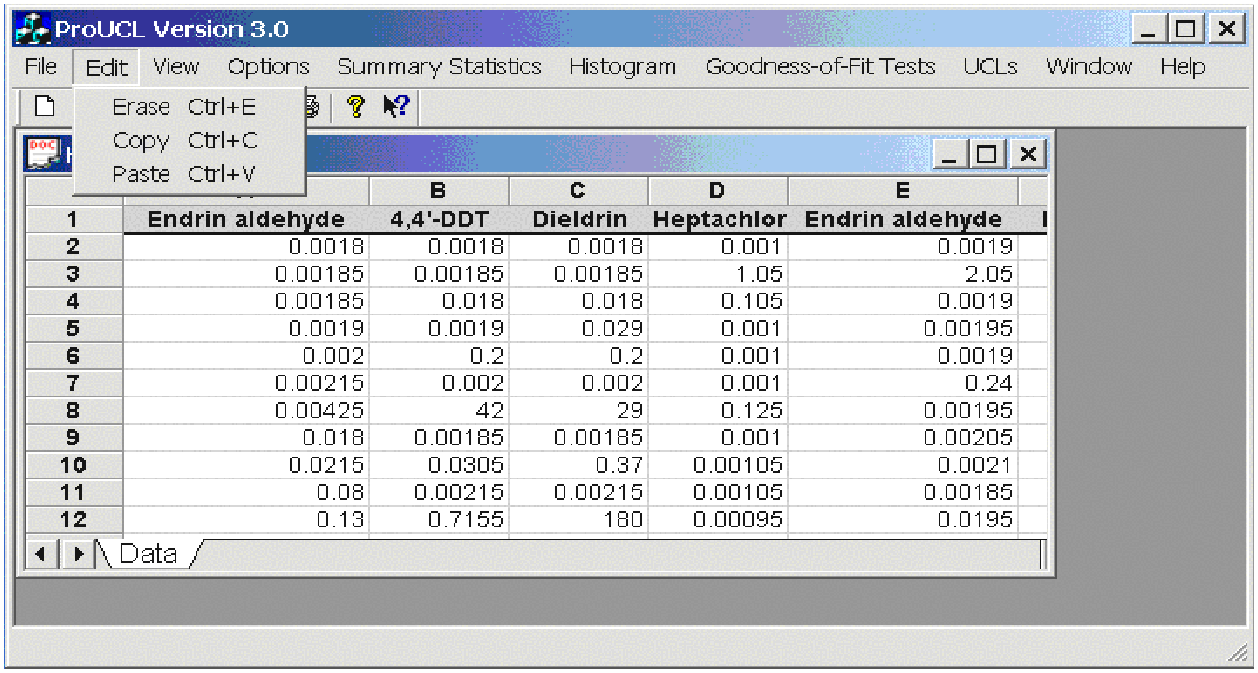
- Note that scroll bars appear when the window is activated and the title bar is highlighted.



	A	B	C	D	E	F	G
3	14000	7	32	19.5	45100	574	1
4	14900	5.1	22.7	17.6	37600	368	0.17
5	14100	6.15	24.55	20.6	40450	671	0.488
6	9510	5.3	17	17.3	26500	1120	0.4
7	9110	4.2	24.8	14.7	38600	759	0.5
8	13900	6.9	17.4	21.2	42700	727	0.34
9	21300	7	28.2	14	41000	409	1.1
10	9110	4.4	21	10.7	26700	434	0.45

2. Edit

Click on the Edit menu item to reveal the following drop-down options.



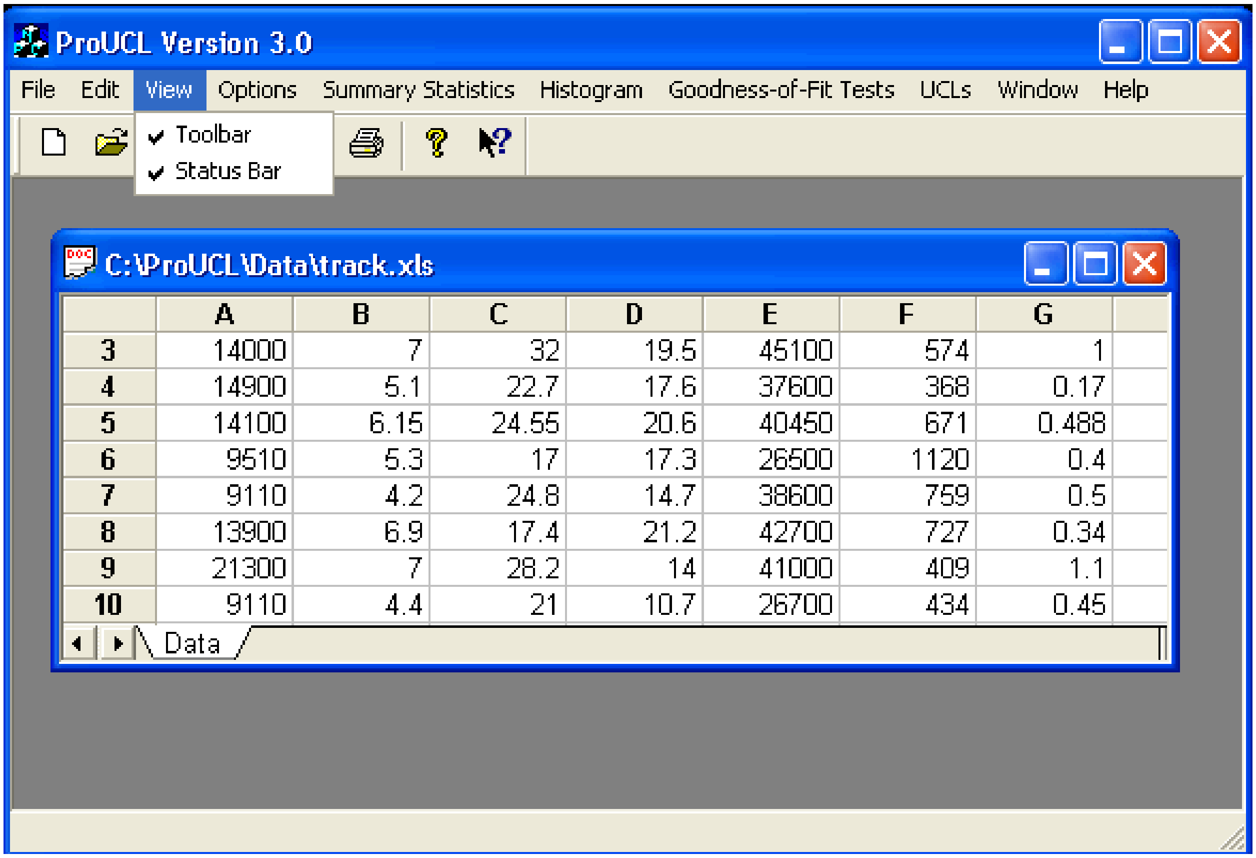
		B	C	D	E
1	Endrin aldehyde	4,4'-DDT	Dieldrin	Heptachlor	Endrin aldehyde
2	0.0018	0.0018	0.0018	0.001	0.0019
3	0.00185	0.00185	0.00185	1.05	2.05
4	0.00185	0.018	0.018	0.105	0.0019
5	0.0019	0.0019	0.029	0.001	0.00195
6	0.002	0.2	0.2	0.001	0.0019
7	0.00215	0.002	0.002	0.001	0.24
8	0.00425	42	29	0.125	0.00195
9	0.018	0.00185	0.00185	0.001	0.00205
10	0.0215	0.0305	0.37	0.00105	0.0021
11	0.08	0.00215	0.00215	0.00105	0.00185
12	0.13	0.7155	180	0.00095	0.0195

The following Edit drop-down menu options are available:

- Erase option: used to remove the highlighted portion of the data. Note that the erased data is not written to any buffer and cannot be recovered. Therefore, when data is erased, it is gone.
- Copy option: similar to a standard Windows Edit option, such as in Excel. It performs typical edit functions of identifying highlighted data (similar to a buffer).
- Paste option: similar to a standard Windows Edit option, such as in Excel. It performs typical edit functions of pasting data identified (highlighted) to the current spreadsheet cell.
- There is no Cut option available in ProUCL because there is no actual buffer available in the commercial software(s) used in the development of ProUCL software.

3. View

Click on the View menu item to reveal these drop-down options.



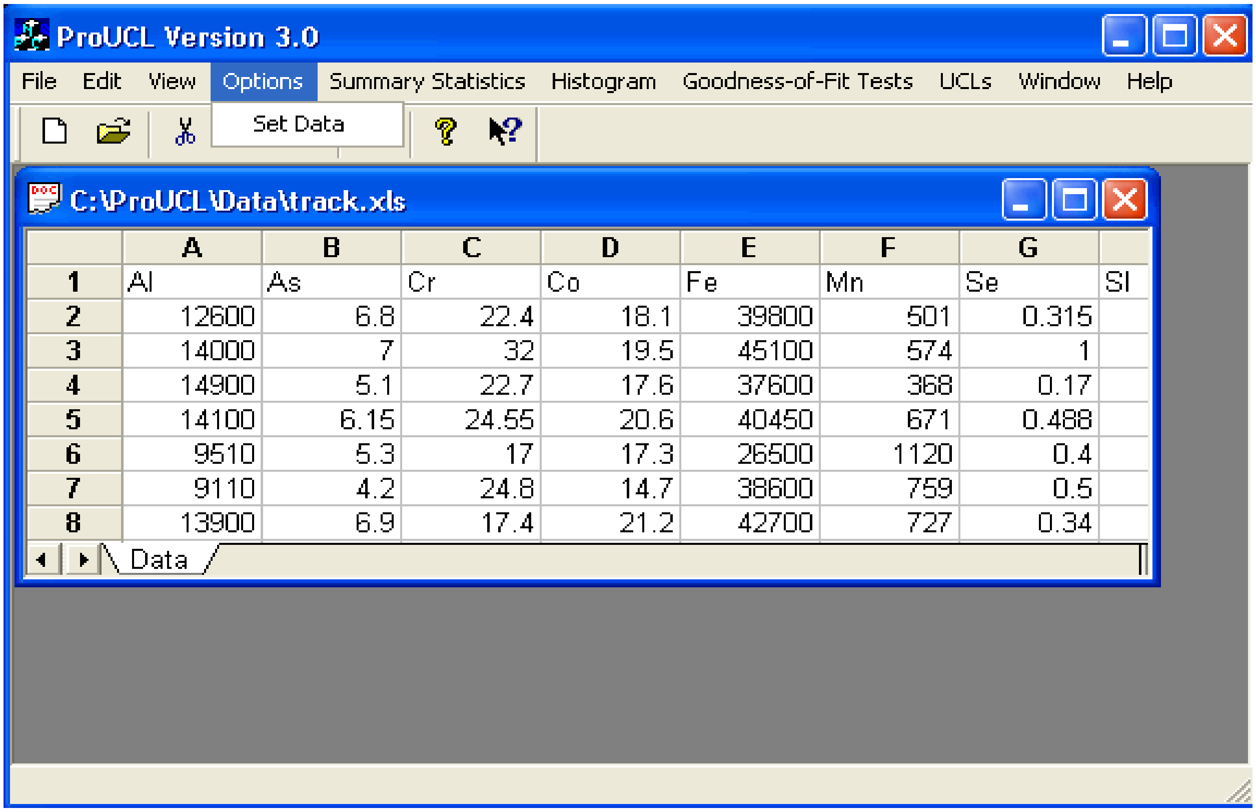
The following View drop-down menu options are available:

- Toolbar: the Toolbar is that row of symbols immediately below the menu items. Clicking on this option toggles the display. This is useful if the user wants to view more data on the screen.

- Status Bar: the Status Bar is the wide bar at the bottom of the screen which displays helpful information. Clicking on this option toggles the display. This is useful if the user wants to view more data on the screen.

4. Options

Click on the Options menu item to reveal these drop-down options.



Currently, Set Data is the only drop-down menu option available.

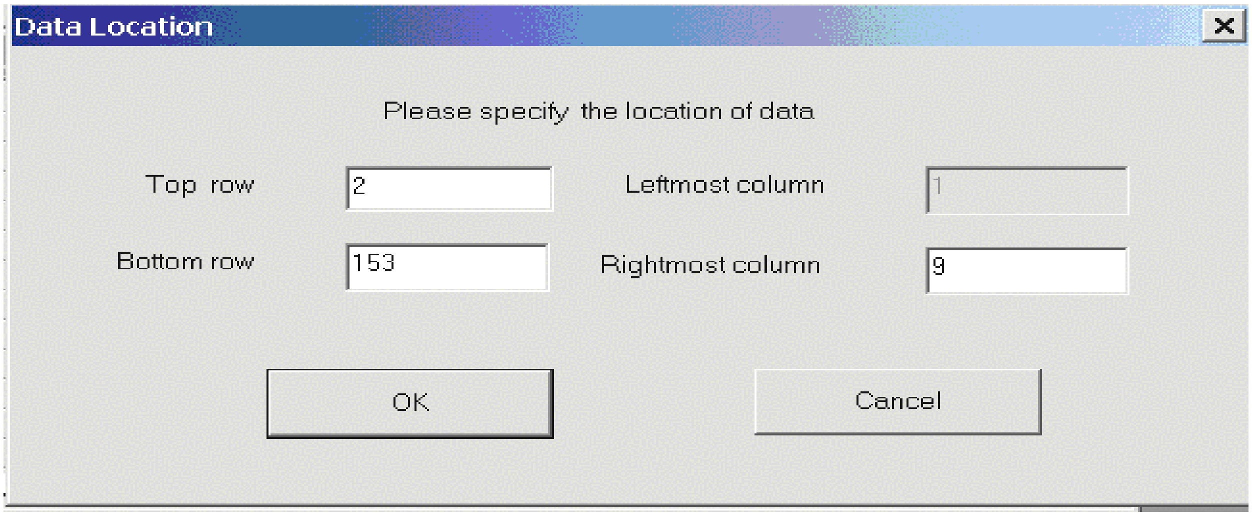
- Set Data option: resets the active portion of the data window. The program examines the active spreadsheet and selects default values representing the first row of data (row 2), the last row which contains data (dependent upon actual data), the leftmost column (typically column 1) where data and text occur, and the rightmost column (dependent upon actual data) where data and text occur.

  NOTE: Caution should be exercised when varying from the default values. If values other than the default values are used, calculation errors may result. Therefore, it is recommended to avoid the use of the Set Data option.

- **The user can pre-process the data outside of the ProUCL software by using a separate spreadsheet program, such as Excel. Pre-processing the data outside of ProUCL will eliminate the need to use the Set Data option.**

4a. The Data Location Screen

The following Data location screen appears when Set Data option is executed.



- **It is recommended to use the default settings for the data screen. This means that all of the data will be processed.**

- **Caution:** Highlighting a portion of the spreadsheet before invoking the Set Data option may sometimes cause unpredicted results.

- **Caution:** Blank cells in the top data row may confuse the automatic sizing algorithm. The user can avoid this problem by re-setting the Rightmost column value using this option.

- The first row in the spreadsheet contains the alphanumeric text (column headings), not data.

- The default Top row of data is row 2. This value can be changed to process a subset of the data in the spreadsheet.

- The default Bottom row is the last row in the spreadsheet which contains nonzero data. This value can be changed to process a subset of the data in the spreadsheet.

- The selected data must correspond to the same columns as the text in the first row. The Leftmost column value (column number) cannot be changed by the user.

- The Rightmost column number can be changed by the user. Note that you must have a column of data for the selected Rightmost column.

5. Summary Statistics

- This option computes general summary statistics for **all variables** in the data file.

- Two Choices are available:

  Raw data (the default option)
  Log-transformed data (Natural logarithm)

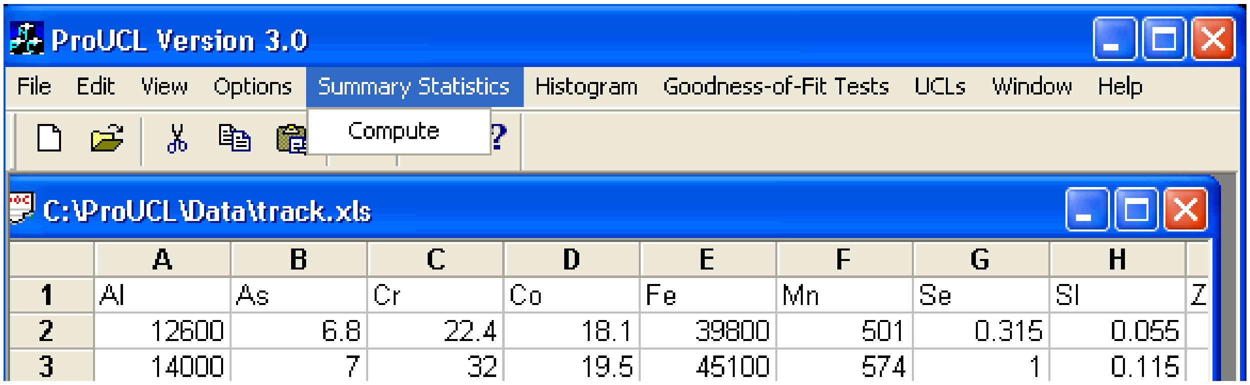
- In ProUCL, Log-transformation means natural logarithm (ln).

- When computing summary statistics for raw data, a message will be displayed for each variable that contains non-numeric values.

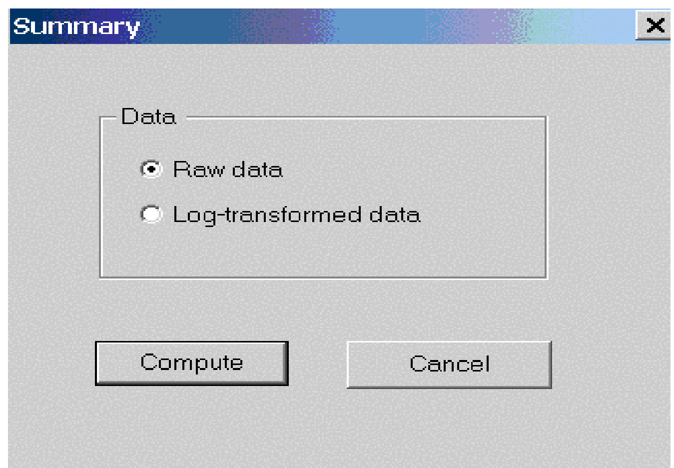
- The Summary Statistics option computes log-transformed data only if all of the data values for the selected variable are positive real numbers. A message will be displayed if non-numeric characters, zero, or negative values are found in the column corresponding to the selected variable.

5a. Summary Statistics Menu

Click on the Summary Statistics menu item to reveal the following drop-down option.



When the user clicks on the Compute option button, the window on the right appears.

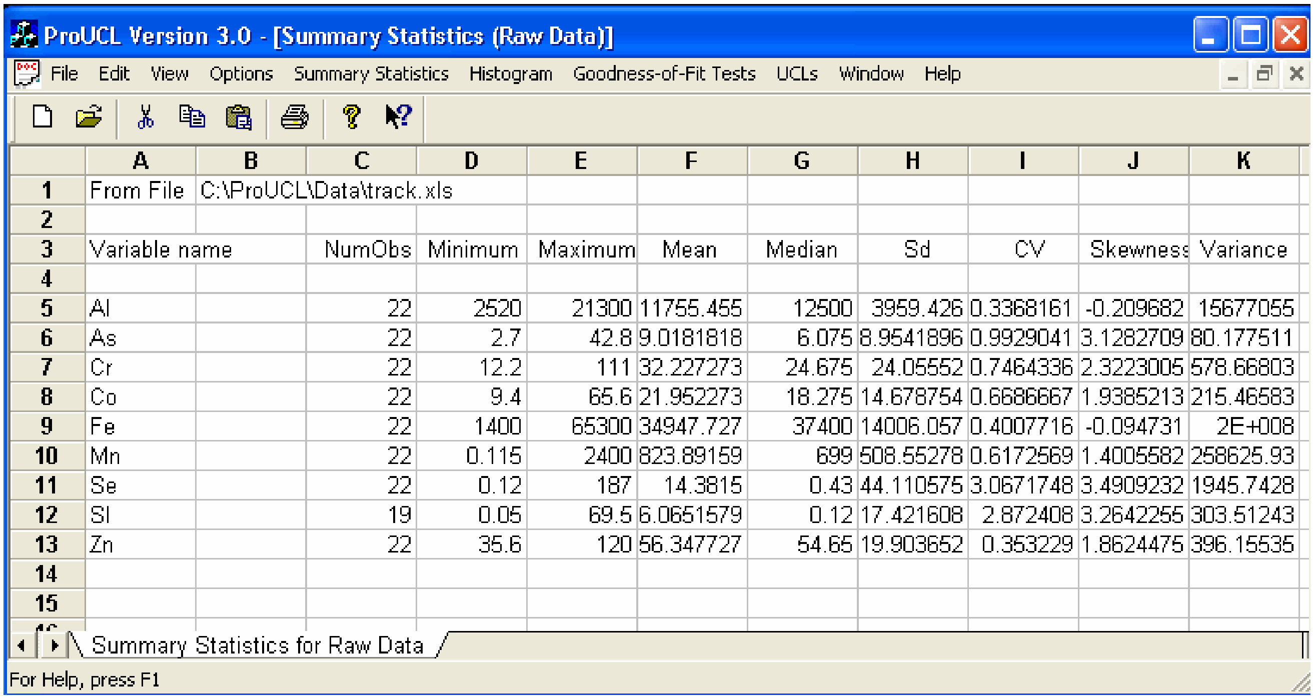


- Select your data choice, and click on the Compute button to continue or on the Cancel button to cancel the summary operations.

- The results screen follows the standard Windows design. It can be edited, widened, printed, resized, or scrolled.

- The resulting Summary Statistics screen can be saved as an Excel file. Right double click on the screen for additional save options.

5b. Results Obtained Using the Summary Statistics Option



	A	B	C	D	E	F	G	H	I	J	K
1	From File	C:\ProUCL\Data\track.xls									
2											
3	Variable name		NumObs	Minimum	Maximum	Mean	Median	Sd	CV	Skewness	Variance
4											
5	Al		22	2520	21300	11755.455	12500	3959.426	0.3368161	-0.209682	15677055
6	As		22	2.7	42.8	9.0181818	6.075	8.9541896	0.9929041	3.1282709	80.177511
7	Cr		22	12.2	111	32.227273	24.675	24.05552	0.7464336	2.3223005	578.66803
8	Co		22	9.4	65.6	21.952273	18.275	14.678754	0.6686667	1.9385213	215.46583
9	Fe		22	1400	65300	34947.727	37400	14006.057	0.4007716	-0.094731	2E+008
10	Mn		22	0.115	2400	823.89159	699	508.55278	0.6172569	1.4005582	258625.93
11	Se		22	0.12	187	14.3815	0.43	44.110575	3.0671748	3.4909232	1945.7428
12	SI		19	0.05	69.5	6.0651579	0.12	17.421608	2.872408	3.2642255	303.51243
13	Zn		22	35.6	120	56.347727	54.65	19.903652	0.353229	1.8624475	396.15535
14											
15											

On the results screen, the following summary statistics are displayed for each variable in the data file:

- ✓ NumObs = Number of Observations
- ✓ Minimum = Minimum value
- ✓ Maximum = Maximum value
- ✓ Mean = Average value
- ✓ Median = Median value
- ✓ Sd = Standard Deviation
- ✓ CV = Coefficient of Variation
- ✓ Skewness = Skewness statistic
- ✓ Variance = Variance statistic

These summary statistics are described in detail in Appendix A.

5c. Printing Summary Statistics

- The summary statistics results and all other results can be printed by clicking the Print option under the menu item File. It is recommended that these statistics be printed in landscape format which is available under the Print Setup option.

6. Histogram

- This option produces a histogram for the selected variable in the data file.

- For data sets with more than one variable, the user should select a variable first. **The histogram is computed and displayed for each selected variable, one variable at a time.**

  - By default, the program selects the first variable.

- The user specifies if the data should be transformed.

  - The default choice is to display the histogram for raw data.

- Two Choices are available:

  - Raw data (the default option)
  - Log-transformed data (natural logarithm, ln)

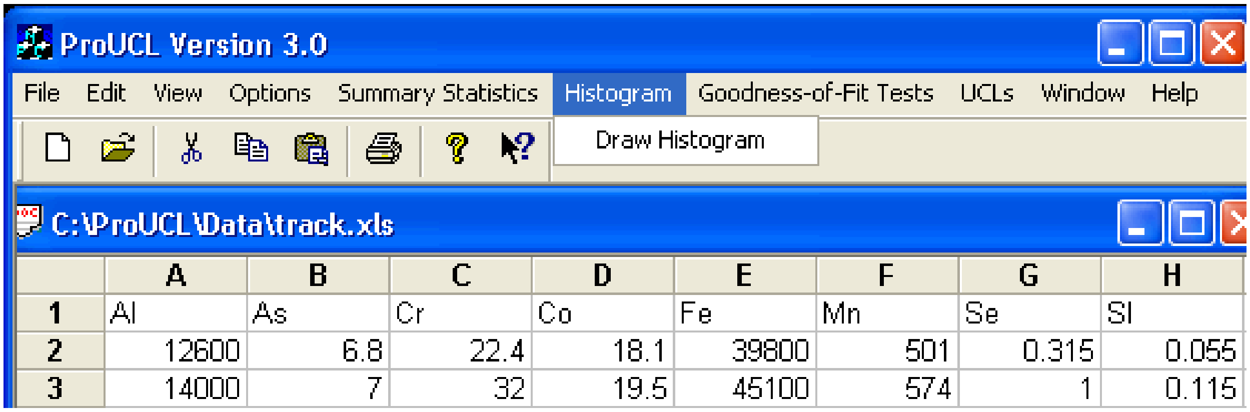
- The user can select the number of bins for the histogram.

  - The default number of bins is 15.

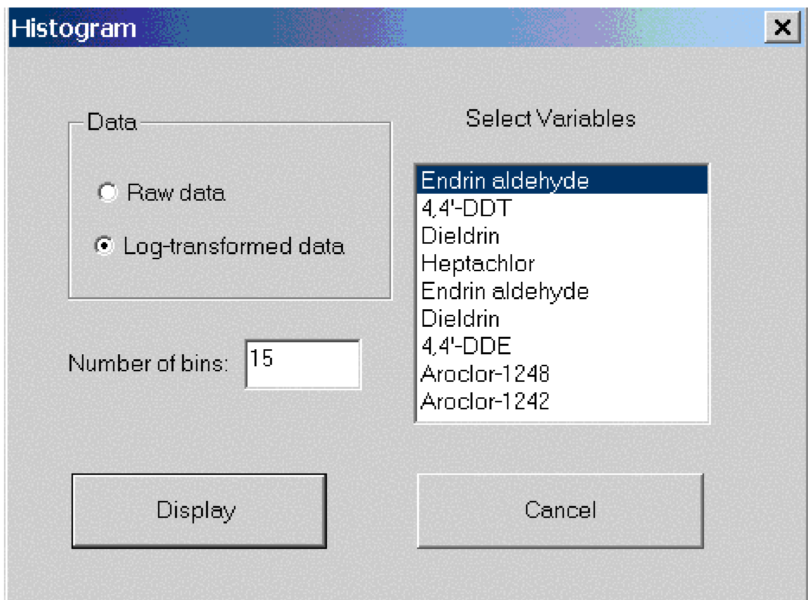
- Note that in order to display and capture the best histogram window, the user may want to **maximize the window before printing.**

6a. Histogram Screen

- Click on the Histogram menu item and then click on the Draw Histogram option.

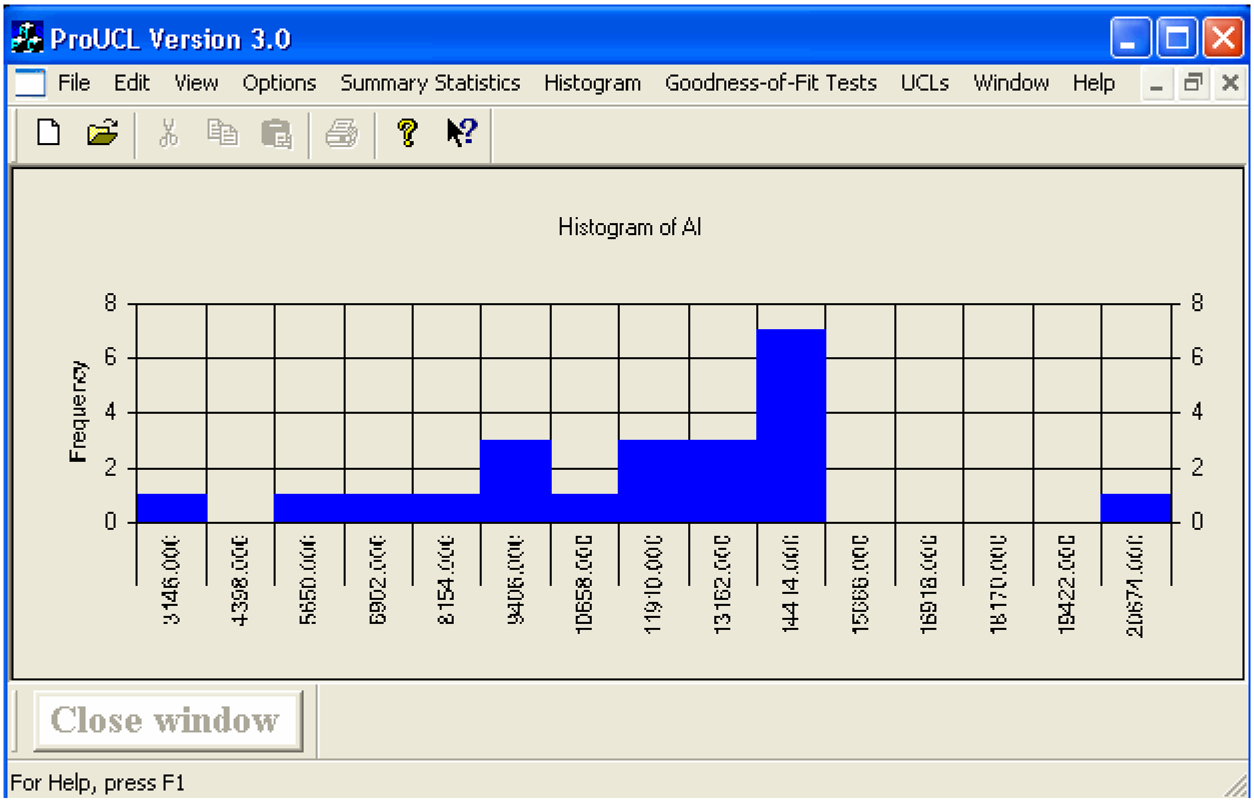


- The window on the right will appear.



- Select Raw data or Log-transformed data.
- You can change the number of bins to be used in the histogram.
- Select a variable and then hit the display key to view the histogram for the selected variable.

6b. Results of Histogram Option



- The Histogram window shown above has been resized for display and reflects the use of default values displayed in Section 6a (Histogram Screen).

- You may close the window by using normal windows operations or click on the Close window button at the bottom left corner of the screen.

- The histogram can be printed or copied by clicking on the right button on the mouse.

- **Caution:** A right click of the mouse will have options other than print and save. These options may function but are NOT recommended due to the program disruption that may occur. Use these other options only at your own risk!

7. Goodness-of-Fit Tests

- Several goodness-of-fit tests are available in ProUCL which are described in Appendix A.

- Throughout this User Guide, and in ProUCL, it is assumed that the user is dealing with a single population. If multiple populations are present, it is recommended to separate them out (using other statistical techniques). Appropriate tests and statistics (e.g., Goodness-of-fit tests, 95% *UCLs*) should be computed separately for each of the identified populations. Also, outliers if any should be identified and thoroughly investigated. The presence of outliers distort all statistics including the *UCLs*. Decisions about their inclusion (or exclusion) from the data set to be used to compute the *UCLs* should be made by all parties involved.

- For data sets with more than one variable, the user should select a variable first. **The data distribution is tested using an appropriate goodness-of-fit test and the associated results are displayed for the selected variable, one variable at a time.**

    ♦ By default, the program selects the first variable.

- This option tests for normal, gamma, or lognormal distribution of the selected variable.

- The user specifies the distribution (normal, gamma, or lognormal) to be tested.

- The user specifies the level of significance. Three choices are available for the level of significance: 0.01, 0.05, or 0.1.

    ♦ The default choice for level of significance is 0.05.

- ProUCL displays a Quantile-Quantile (Q-Q) plot for the selected variable (or the log-transformed variable). A Q-Q plot can be generated for each of the three distributions.

- The linear pattern displayed by the Q-Q plot suggests approximate goodness-of-fit for the selected distribution.

- The program computes the intercept, slope, and the correlation coefficient for the linear pattern displayed by the Q-Q plot. A high value of the correlation coefficient (e.g., > 0.95) is an indication of approximate goodness-of-fit for that distribution. Note that these statistics are displayed on the Q-Q plot.

- On this graph, observations that are well separated from the bulk (central part) of the data typically are potential outliers needing further investigation.

- Significant and obvious jumps in a Q-Q plot (for any distribution) are indication of the presence of more than one population which should be partitioned out before estimating an EPC Term. It is strongly recommended that both graphical and formal goodness-of fit tests should be used on the same data set to determine the distribution of the data set under study.

- In addition to the graphical normal and lognormal Q-Q plot, two more powerful methods are also available to test the normality or lognormality of the data set:

    - Lilliefors Test: a test typically used for samples of larger size (> 50). When the sample size is greater than 50, the program defaults to the Lilliefors test. However, note that the Lilliefors test is available for samples of all sizes. There is no applicable upper limit for sample size for the Lilliefors test.

    - Shapiro and Wilk W-Test: a test used for samples of smaller size (< 50). W-Test is available only for samples of size 50 or less.

    - It should be noted that sometimes, these two tests may lead to different conclusions. Therefore, the user should exercise caution interpreting the results.

- In addition to the graphical gamma Q-Q plot, two more powerful Empirical Distribution Function (EDF) procedures are also available to test the gamma distribution of the data set. These are the Anderson-Darling Test and the Kolmogorov-Smirnov Test.

    - It should be noted that these two tests may also lead to different conclusions. Therefore, the user should exercise caution interpreting the results.

    - These two tests may be used for samples of size in the range 4-2500. Also, for these two tests, the value of k (k hat) should lie in the interval [0.01,100.0]. Consult Appendix A for detailed description of k. Extrapolation beyond these sample sizes and values of k is not recommended.

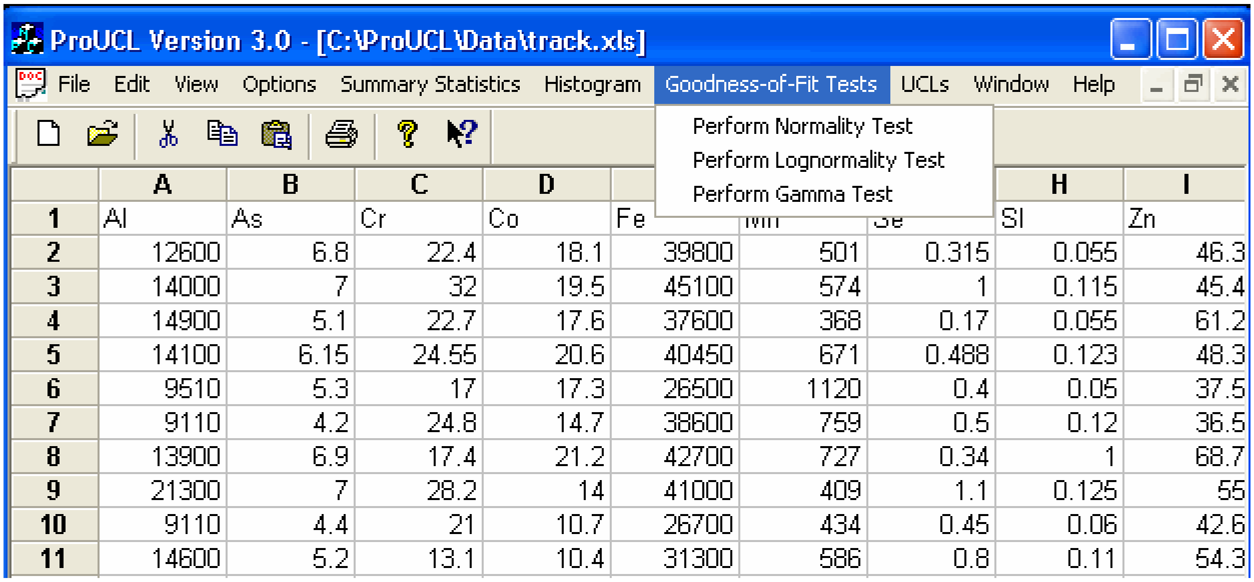
- ProUCL computes the relevant test statistic and the associated critical value, and prints them on the associated Q-Q plot. On this Q-Q plot, the program informs the user if the data are gamma, normally, or lognormally distributed. *It highly recommended not to skip the use of graphical Q-Q plot to determine the data distribution as a Q-Q plot also provides the useful information about the presence of multiple populations and/or outliers.*

- The Q-Q plot can be printed or copied by clicking on the right button on the mouse.

- Note: In order to capture the entire graph window, the user should maximize the window before printing.

7a. Goodness-of-Fit Tests Screen

- Click on the Goodness-of-Fit Tests menu item and a drop-down menu list will appear as shown in the screen below:



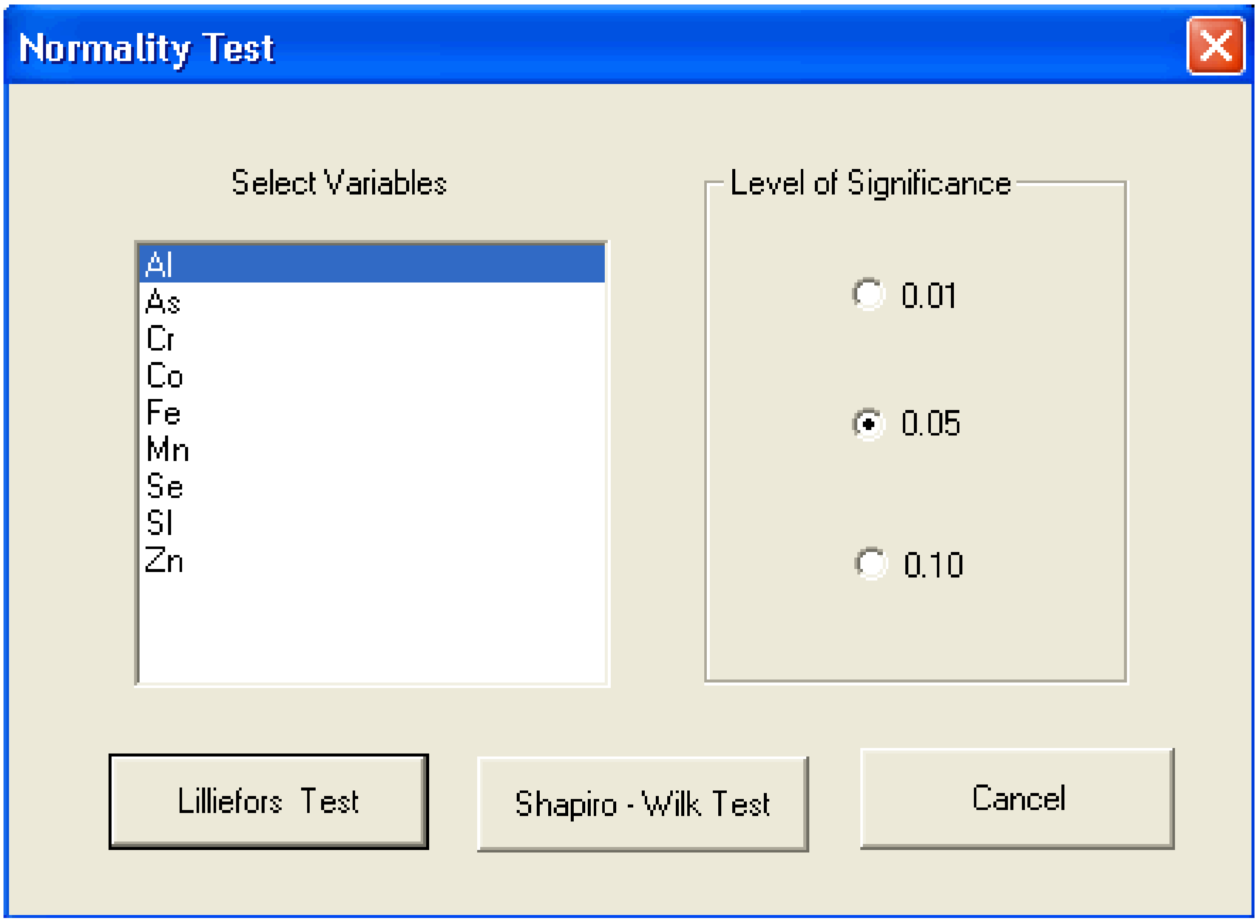
- To test your variable for normality, click on Perform Normality Test from the drop-down menu list.

- To test your variable for lognormality, click on Perform Lognormality Test from the drop-down menu list.

- To test your variable for gamma distribution, click on Perform Gamma Test from the drop-down menu list.

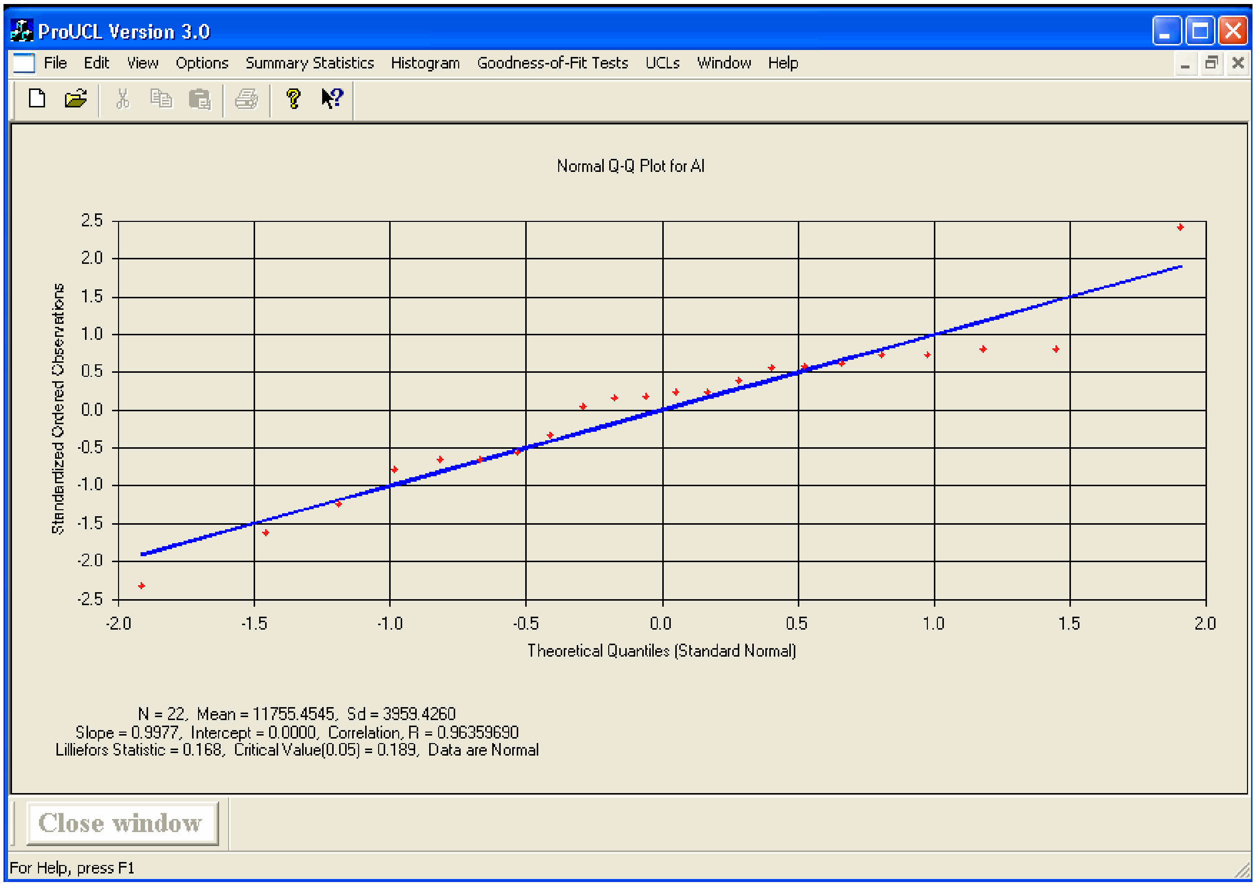
7b. Result of Selecting Perform Normality Test Option

The following window will appear:



- Select a variable.
- Select a Level of Significance.
- Click on either Lilliefors Test or Shapiro-Wilk Test.

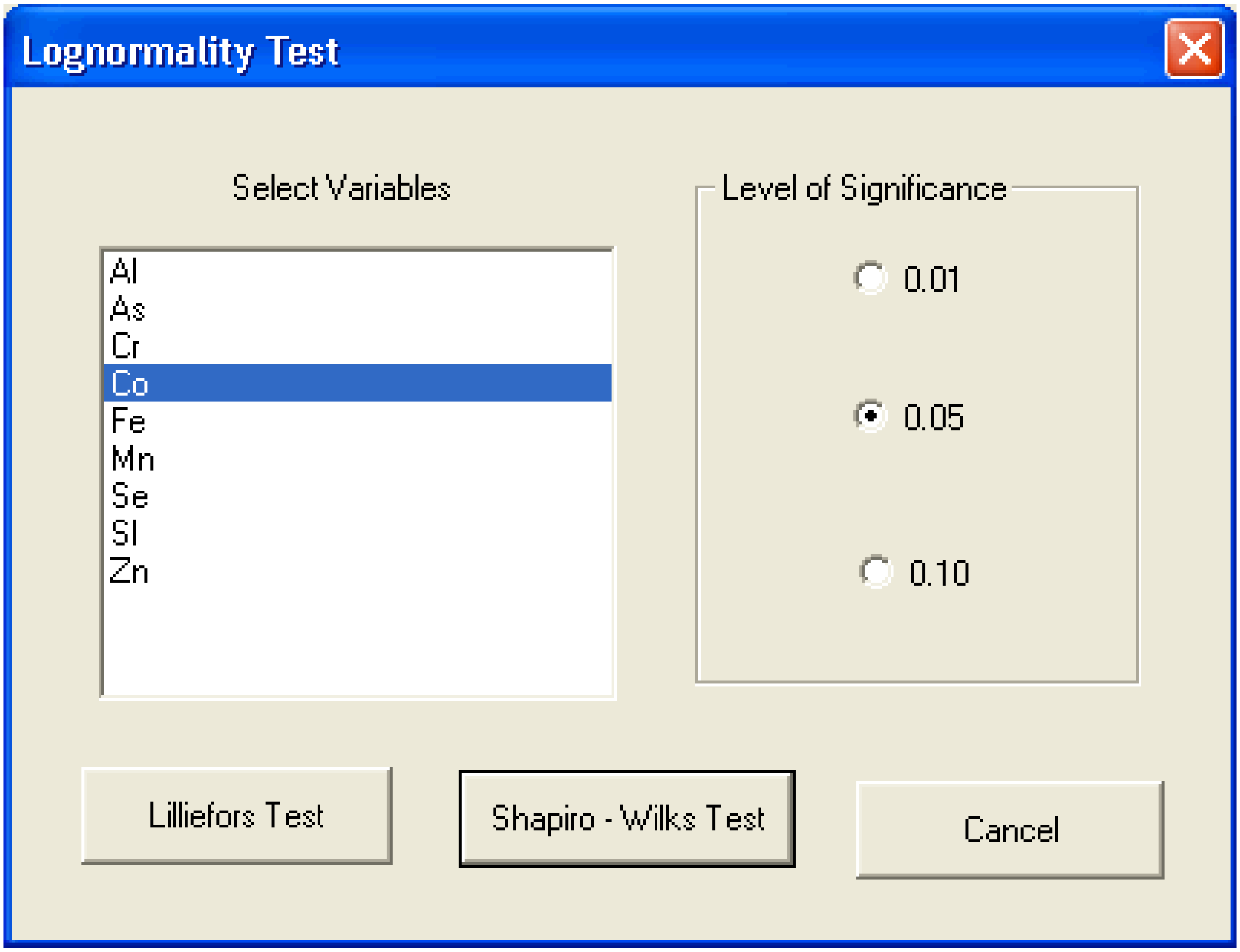
7c. Resulting Q-Q Plot Display to Perform Normality Test



- The Q-Q plot window shown above has been resized for display.
- Two different Q-Q plot windows are produced for each Normality test request. The first graph plots the raw data along the vertical axis, and the second plot (as shown above) uses the standardized data along the vertical axis. These two Q-Q plots convey the same information about the data distribution and potential outliers, and therefore they also look very similar, but they do represent two separate (not duplicate) plots. It is the user's preference to pick one of these two Q-Q plots to assess approximate normality of the data set under study.
- Right click on a graph to print or save that graph.
- **Caution:** A right click of the mouse will have options other than print and save. These options may function but are NOT recommended due to the program disruption that may occur. Use these other options only at your own risk!

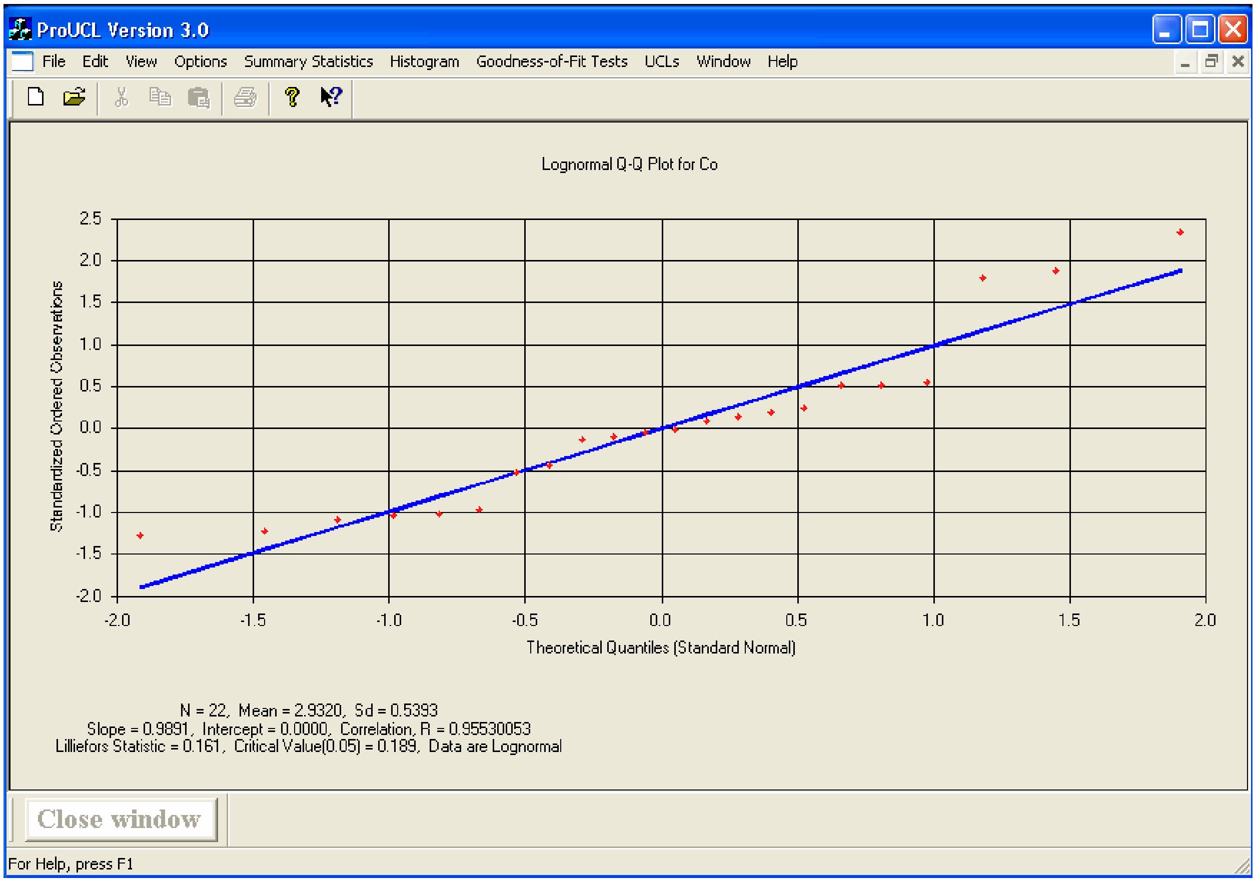
7d. Result of Selecting Perform Lognormality Test Option

The following window will appear:



- Select a variable.
- Select a Level of Significance.
- Click on either Lilliefors Test or Shapiro-Wilk Test.

7e. Resulting Lognormal Q-Q Plot Display to Perform Lognormality Test



- The Q-Q plot window shown above has been resized for display.

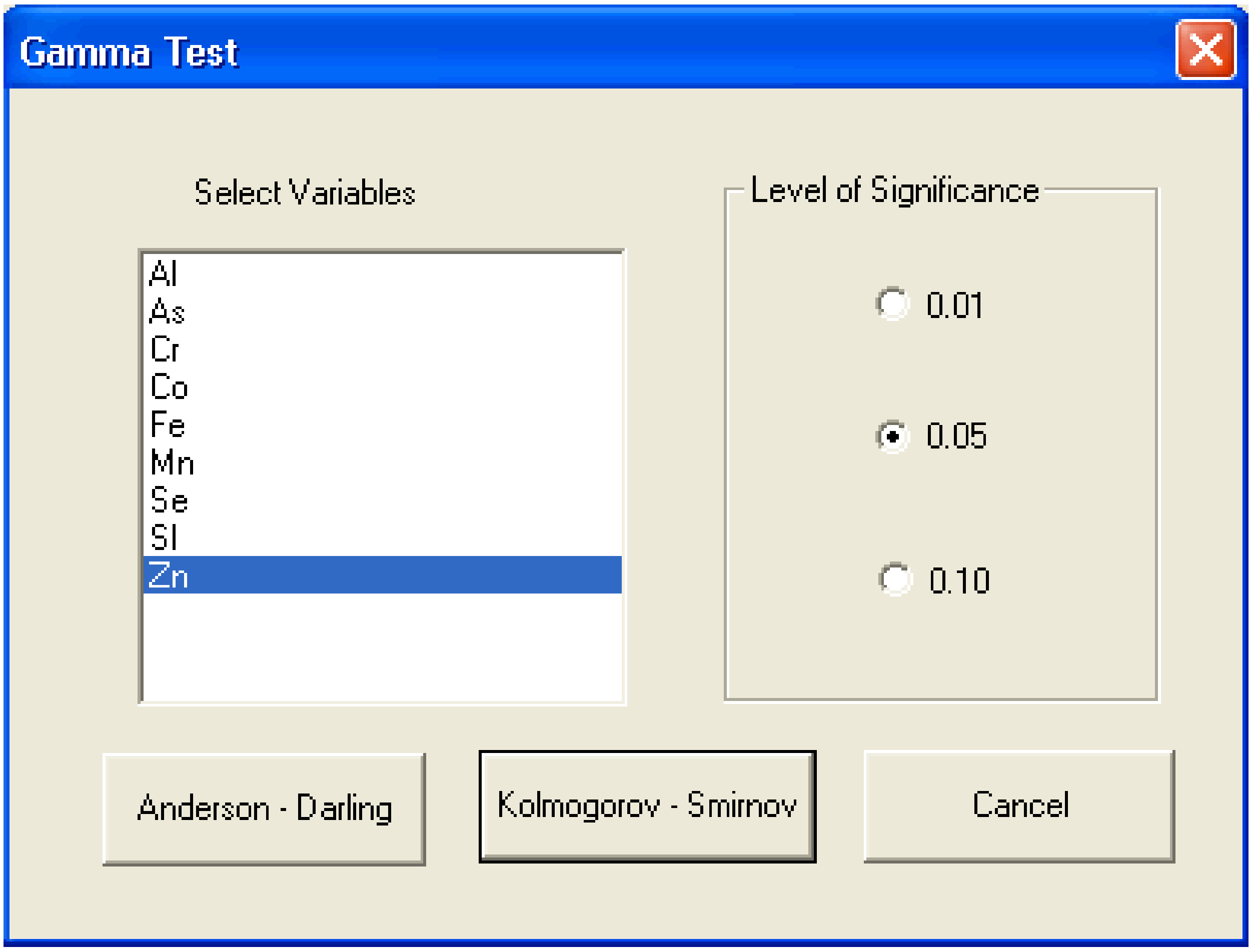
- Two different Q-Q plot windows are produced for each Lognormality test request. The first plot uses the log-transformed data along the vertical axis, and the second plot (shown above) uses the standardized data. As mentioned before, these two plots provide the same information about the data distribution and potential outliers, but they do represent two separate (not duplicate) plots. The user can pick any of these two Q-Q plots to assess approximate lognormality of the data set under study.

- Right click on a graph to print or save that graph.

- **Caution:** As before, a right click of the mouse will have options other than print and save. These options may function but are NOT recommended due to the program disruption that may occur. Use these other options only at your own risk!

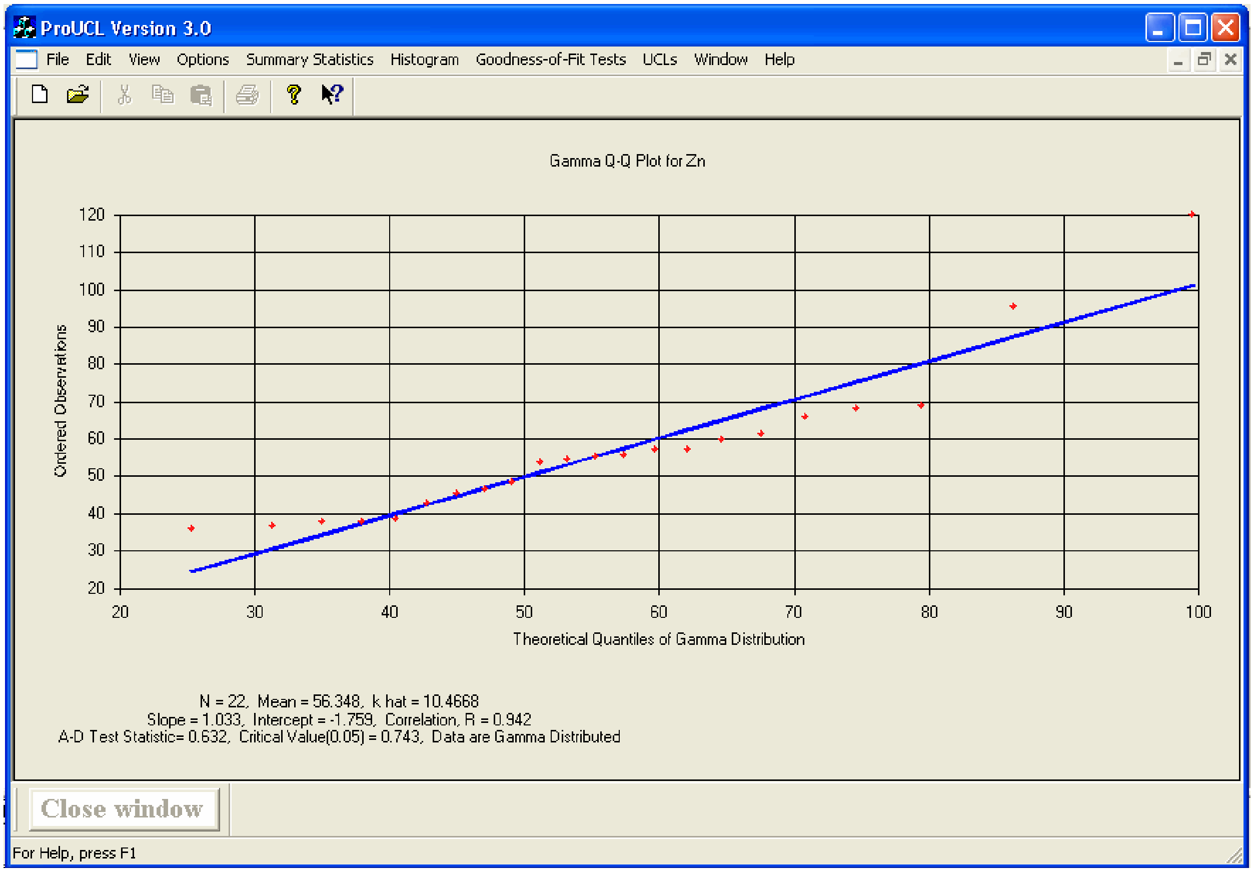
7f. Result of Selecting Perform Gamma Test Option

The following window will appear:



- Select a variable.
- Select a Level of Significance.
- Click on either the Anderson - Darling Test or Kolmogorov - Smirnov Test.

7g. Resulting Gamma Q-Q Plot Display to Perform Gamma Test



- The Q-Q plot window shown above has been resized for display.
- Only one Q-Q plot window is produced for each Gamma test request: the display using the original raw data (as shown above).
- Right click on the graph to print or save the graph.
- **Caution:** A right click of the mouse will have options other than print and save. These options may function but are NOT recommended due to the program disruption that may occur. Use these other options only at your own risk!

8. UCLs

- This option computes the UCLs for the selected variable.

- The program can compute UCLs using all available methods. For details regarding the various distributions and methods, refer to Appendix A.

- The user specifies the confidence level; a number in the interval [0.5, 1), 0.5 inclusive. The default choice is 0.95.

- The program computes several non-parametric UCLs using the Central Limit Theorem, Chebyshev inequality, Jackknife, and the various Bootstrap methods.

- For the bootstrap method, the user can specify the number of bootstrap runs. The default choice for the number of bootstrap runs is 2000.

- The user is responsible for selecting an appropriate choice for the data distribution: normal, gamma, lognormal, or non-parametric. The user determines the data distribution using the Goodness-of-Fit Test option prior to using the UCLs option. The UCLs option will also inform the user if the data are normal, gamma, lognormal, or non-parametric. The program computes relevant statistics depending on the user selection.

- For data sets which are not normal, one should try the gamma UCLs next. The program will offer you advice if you chose the wrong UCLs option.

- For data sets which are neither normal nor gamma, you should try the lognormal UCLs next. The program will offer you advice if you chose the wrong UCLs option.

- Data sets that are not normal, gamma, or lognormal are classified as non-parametric data sets. The user should use non-parametric UCLs option for such data sets. The program will offer you advice if you chose the wrong UCLs option.

- For lognormal data sets, ProUCL can compute only a 90% or a 95% Land's statistic based H-UCL of the mean. For all other methods, ProUCL can compute a UCL for any confidence coefficient in the interval [0.5,1.0), 0.5 inclusive.

- If you have selected a proper distribution, ProUCL will provide a recommended UCL computation method for the 0.95 confidence coefficient. Even though ProUCL can compute UCLs for confidence coefficients in the interval [0.5, 1.0), recommendations are provided only for 95% UCL computation methods as the EPC term is estimated by a 95% UCL of the mean.

- ProUCL can compute the H-UCL for sample sizes up to 1000 using the critical values as given by Land (1975).

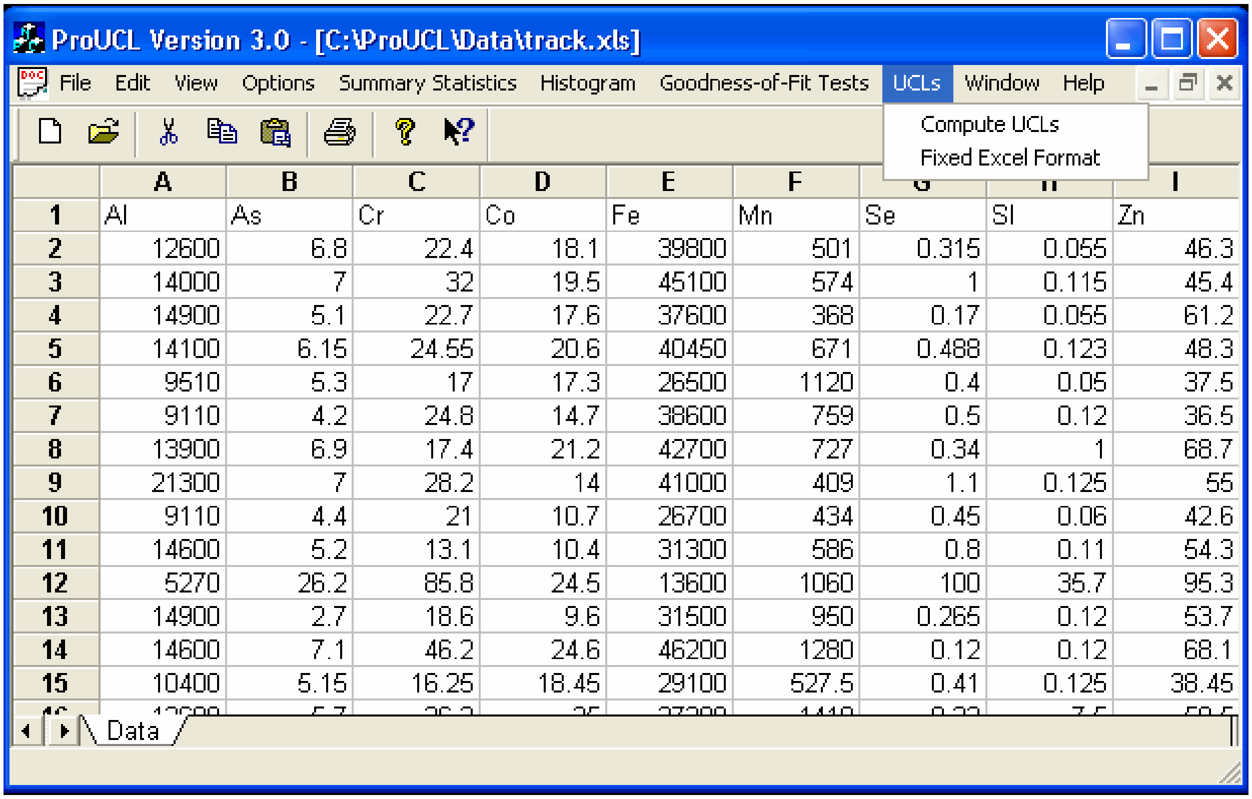
- For lognormal data sets, ProUCL also computes the Maximum Likelihood Estimates (MLEs) of the population percentiles, and the minimum variance unbiased estimates (MVUEs) of the population mean, median, standard deviation, and the standard error (SE) of the mean. Note that for lognormally distributed background data sets, these MLEs of the population percentiles (e.g., 95% percentile) can be used as estimates of the background level threshold values.

- The detailed theory and formulas used to compute these gamma and lognormal statistics are given by Land (1971, 1975), Gilbert (1987), Singh, Singh, and Engelhardt (1997, 1999), Singh et al. (2002a), Singh et al. (2002b), and Singh and Singh (2003).

- Formulas, methods, and cited references used in the development of ProUCL are summarized in Appendix A.

8a. UCLs Computation Screen

Click on the UCLs menu item and the drop down menu shown below will appear.

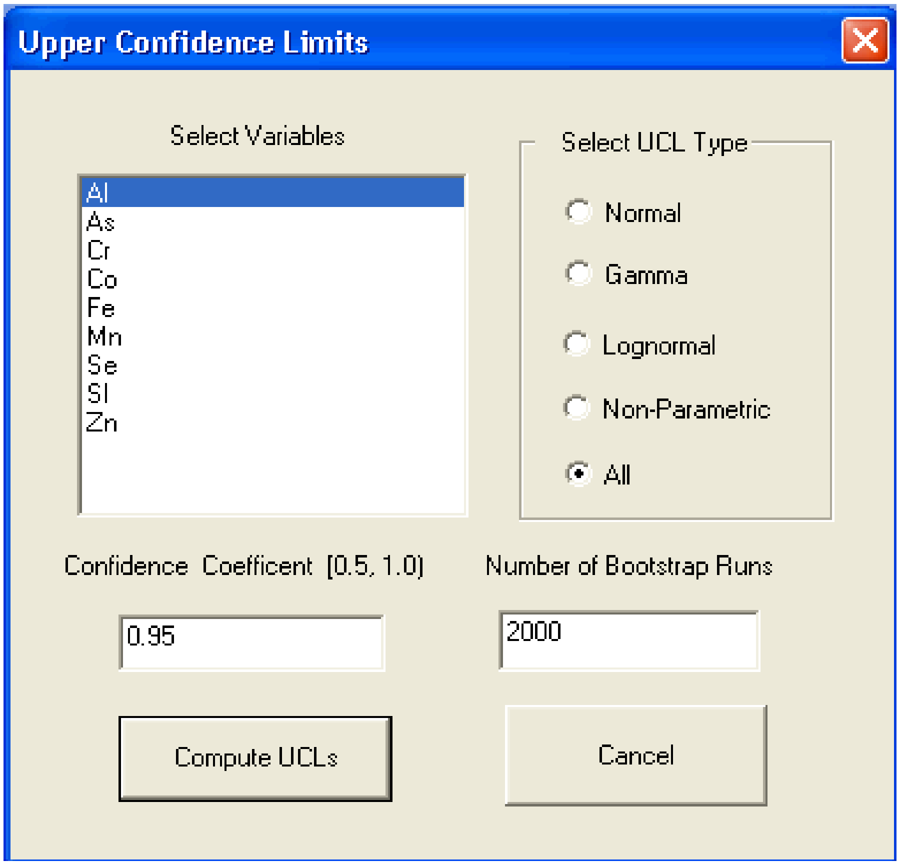


	A	B	C	D	E	F	G	H	I
1	Al	As	Cr	Co	Fe	Mn	Se	Sl	Zn
2	12600	6.8	22.4	18.1	39800	501	0.315	0.055	46.3
3	14000	7	32	19.5	45100	574	1	0.115	45.4
4	14900	5.1	22.7	17.6	37600	368	0.17	0.055	61.2
5	14100	6.15	24.55	20.6	40450	671	0.488	0.123	48.3
6	9510	5.3	17	17.3	26500	1120	0.4	0.05	37.5
7	9110	4.2	24.8	14.7	38600	759	0.5	0.12	36.5
8	13900	6.9	17.4	21.2	42700	727	0.34	1	68.7
9	21300	7	28.2	14	41000	409	1.1	0.125	55
10	9110	4.4	21	10.7	26700	434	0.45	0.06	42.6
11	14600	5.2	13.1	10.4	31300	586	0.8	0.11	54.3
12	5270	26.2	85.8	24.5	13600	1060	100	35.7	95.3
13	14900	2.7	18.6	9.6	31500	950	0.265	0.12	53.7
14	14600	7.1	46.2	24.6	46200	1280	0.12	0.12	68.1
15	10400	5.15	16.25	18.45	29100	527.5	0.41	0.125	38.45

- The Compute UCLs option is intended for general use. It displays results in a format suitable for review by all users. The output results can be printed or saved for subsequent use. Saved results can be imported into other documents and reports.

- The Fixed Excel Format option produces a results screen that can be exported to another program written for production purposes. Therefore, UCL results are stored in specific cells and no attempt has been made to accommodate human review. These fixed format results are not formatted to be printed.

8b. Results After Clicking on Compute UCLs Drop-Down Menu Item



- Note that the UCLs are computed for one variable at a time. The user selects a variable from the variable list.

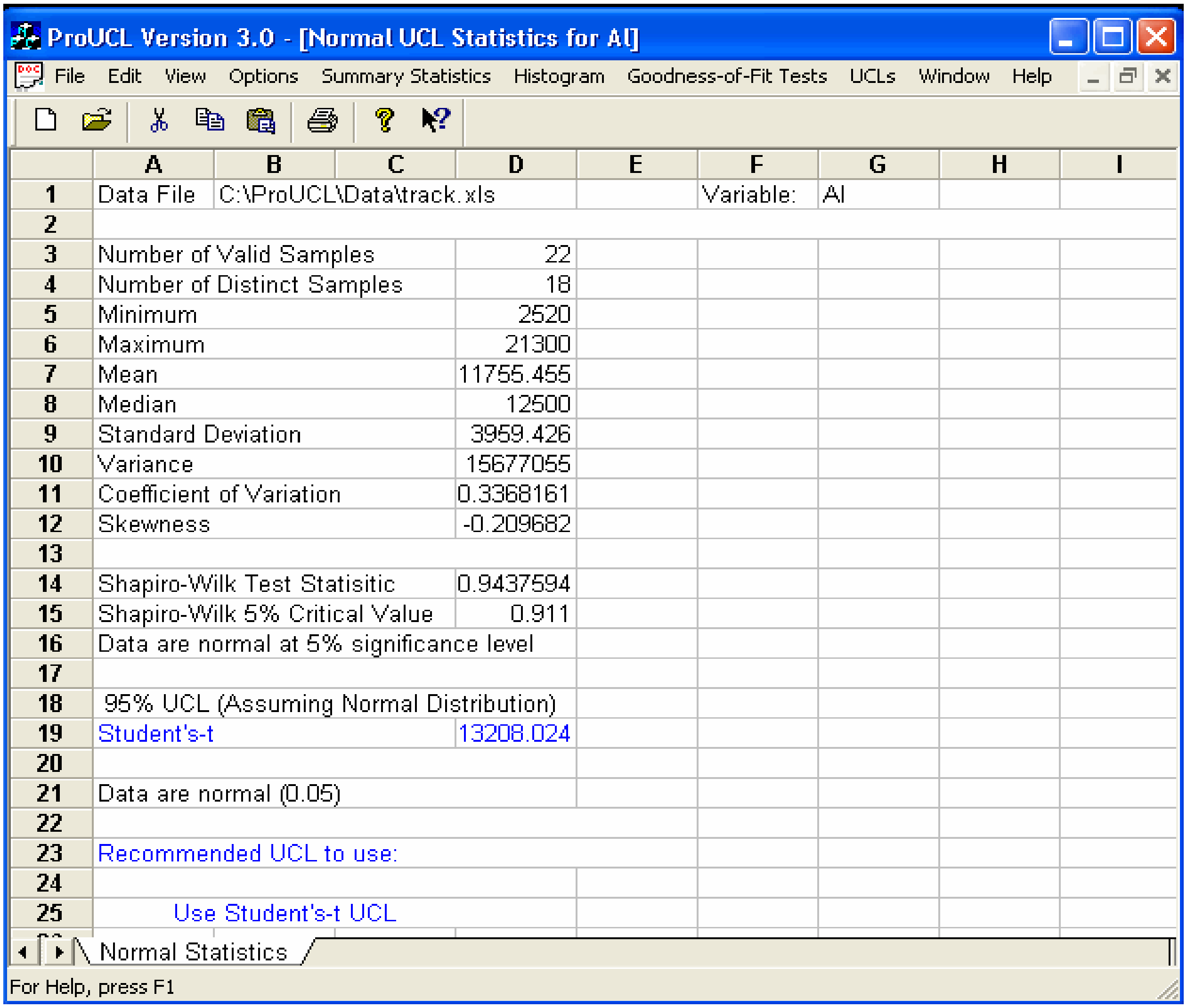
- The user may change the Confidence Coefficient (default is 0.95). The range allowed is between 0.5 and 1.0, 0.5 inclusive.

- The user may adjust the number of bootstrap runs (default is 2,000).

- The user selects one of the options: Normal, Gamma, Lognormal, Non-parametric, or All option. The All option is the default choice. The All option automatically determines the data distribution without checking for outliers and/or the presence of multiple populations.. *It is highly recommended to verify the data distribution (for outliers and multiple populations) using an appropriate Q-Q plot under the Goodness-of-Fit Tests option.*

- The All option computes and displays the UCLs using all parametric and non-parametric methods available in ProUCL. Finally, the user clicks on the Compute UCLs button.

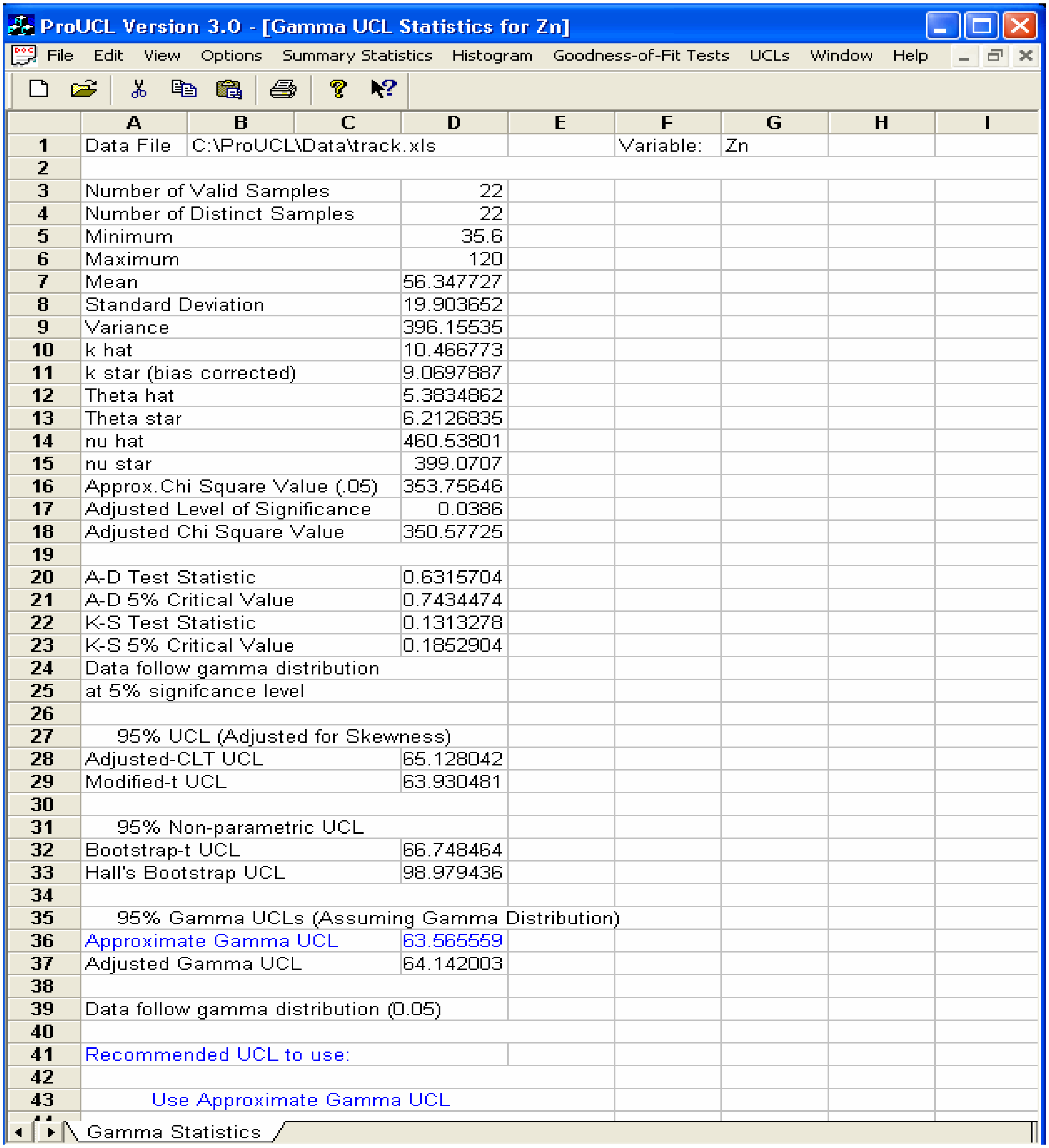
8c. Display After Selecting the Normal UCLs Option



	A	B	C	D	E	F	G	H	I
1	Data File	C:\ProUCL\Data\track.xls				Variable:	Al		
2									
3	Number of Valid Samples			22					
4	Number of Distinct Samples			18					
5	Minimum			2520					
6	Maximum			21300					
7	Mean			11755.455					
8	Median			12500					
9	Standard Deviation			3959.426					
10	Variance			15677055					
11	Coefficient of Variation			0.3368161					
12	Skewness			-0.209682					
13									
14	Shapiro-Wilk Test Statisitic			0.9437594					
15	Shapiro-Wilk 5% Critical Value			0.911					
16	Data are normal at 5% significance level								
17									
18	95% UCL (Assuming Normal Distribution)								
19	Student's-t			13208.024					
20									
21	Data are normal (0.05)								
22									
23	Recommended UCL to use:								
24									
25	Use Student's-t UCL								

- This data does not follow the normal distribution for the selected variable.
- The program notes that the data follow an approximate gamma distribution and suggests in blue that the user should try Gamma UCLs.
- This output spreadsheet is easily saved by using the Save As option under the File menu.
- Double right click on the UCL output spreadsheet to view a screen with more options to save, print, or write this output sheet to a file.

8d. Display After Selecting the Gamma UCLs Option



ProUCL Version 3.0 - [Gamma UCL Statistics for Zn]

File Edit View Options Summary Statistics Histogram Goodness-of-Fit Tests UCLs Window Help

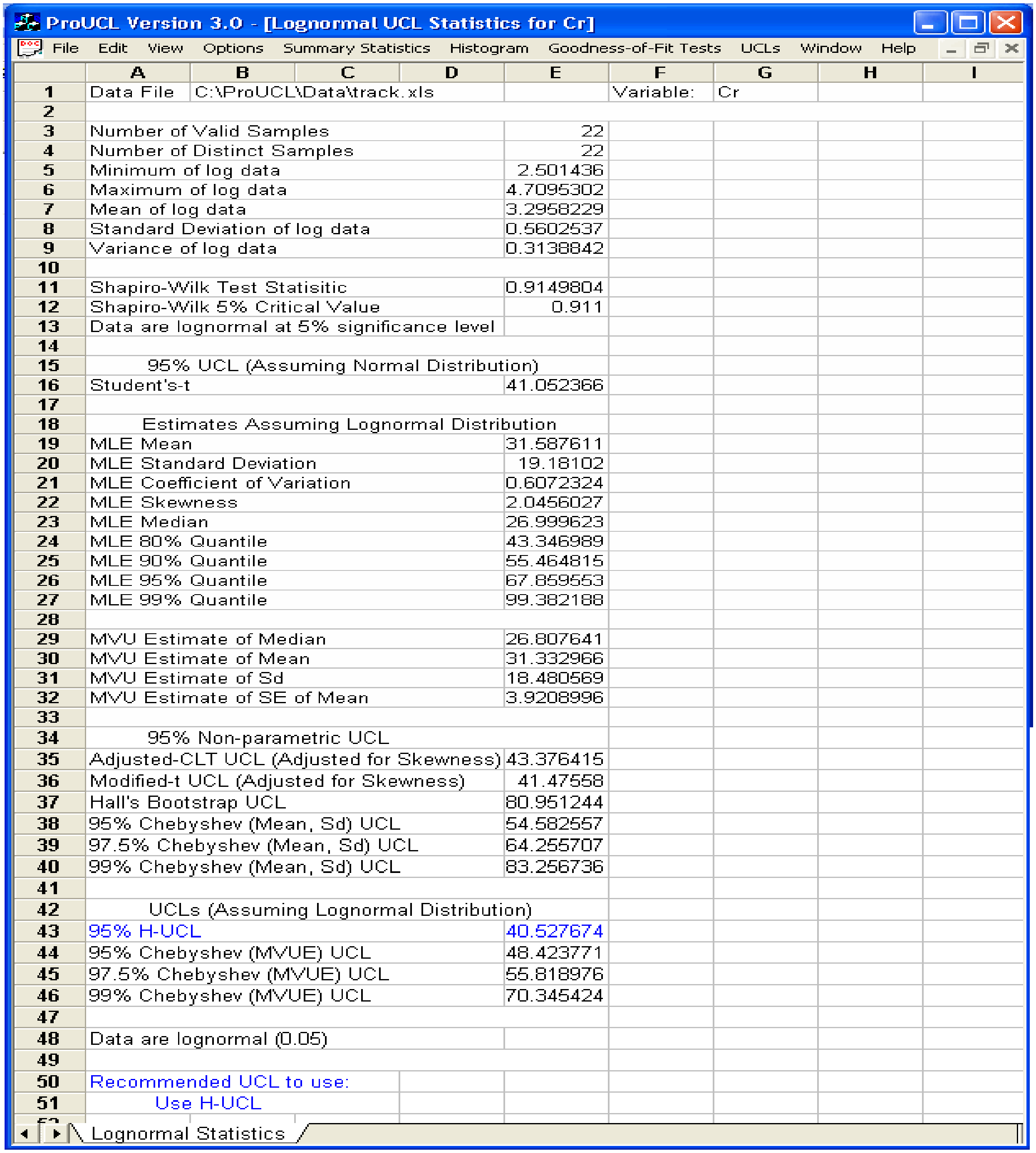
	A	B	C	D	E	F	G	H	I
1	Data File	C:\ProUCL\Data\track.xls				Variable:	Zn		
2									
3	Number of Valid Samples			22					
4	Number of Distinct Samples			22					
5	Minimum			35.6					
6	Maximum			120					
7	Mean			56.347727					
8	Standard Deviation			19.903652					
9	Variance			396.15535					
10	k hat			10.466773					
11	k star (bias corrected)			9.0697887					
12	Theta hat			5.3834862					
13	Theta star			6.2126835					
14	nu hat			460.53801					
15	nu star			399.0707					
16	Approx.Chi Square Value (.05)			353.75646					
17	Adjusted Level of Significance			0.0386					
18	Adjusted Chi Square Value			350.57725					
19									
20	A-D Test Statistic			0.6315704					
21	A-D 5% Critical Value			0.7434474					
22	K-S Test Statistic			0.1313278					
23	K-S 5% Critical Value			0.1852904					
24	Data follow gamma distribution								
25	at 5% signifcance level								
26									
27	95% UCL (Adjusted for Skewness)								
28	Adjusted-CLT UCL			65.128042					
29	Modified-t UCL			63.930481					
30									
31	95% Non-parametric UCL								
32	Bootstrap-t UCL			66.748464					
33	Hall's Bootstrap UCL			98.979436					
34									
35	95% Gamma UCLs (Assuming Gamma Distribution)								
36	Approximate Gamma UCL			63.565559					
37	Adjusted Gamma UCL			64.142003					
38									
39	Data follow gamma distribution (0.05)								
40									
41	Recommended UCL to use:								
42									
43	Use Approximate Gamma UCL								

Gamma Statistics

- Save this output spreadsheet by using the Save As option under the File menu.

- Double right click on the UCL output spreadsheet to view a screen with more options to save, print, or write this output sheet to a file.

8e. Display After Selecting the Lognormal UCLs Option



ProUCL Version 3.0 - [Lognormal UCL Statistics for Cr]

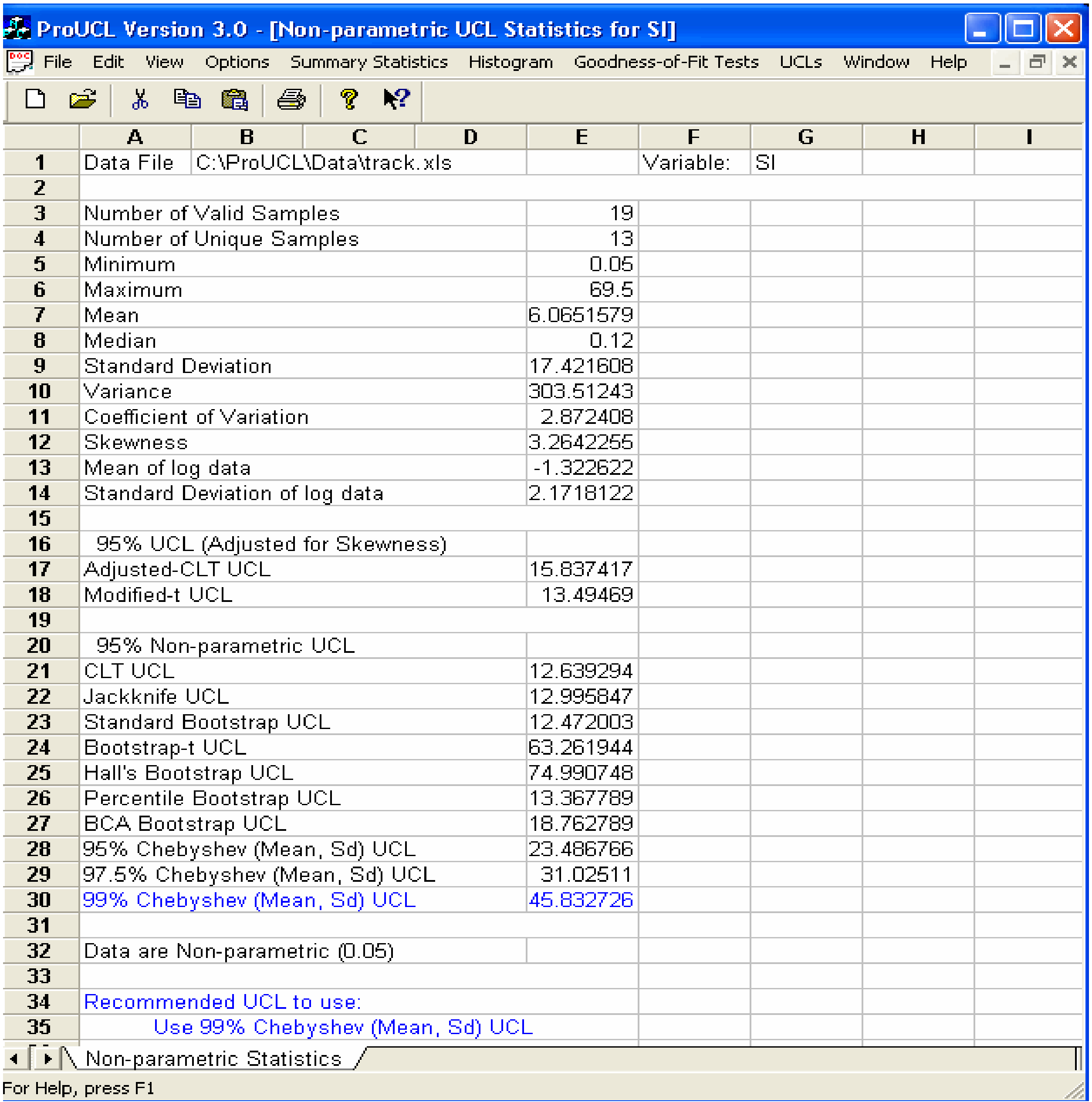
File Edit View Options Summary Statistics Histogram Goodness-of-Fit Tests UCLs Window Help

	A	B	C	D	E	F	G	H	I
1	Data File	C:\ProUCL\Data\track.xls				Variable:	Cr		
2									
3	Number of Valid Samples				22				
4	Number of Distinct Samples				22				
5	Minimum of log data				2.501436				
6	Maximum of log data				4.7095302				
7	Mean of log data				3.2958229				
8	Standard Deviation of log data				0.5602537				
9	Variance of log data				0.3138842				
10									
11	Shapiro-Wilk Test Statisitic				0.9149804				
12	Shapiro-Wilk 5% Critical Value				0.911				
13	Data are lognormal at 5% significance level								
14									
15	95% UCL (Assuming Normal Distribution)								
16	Student's-t				41.052366				
17									
18	Estimates Assuming Lognormal Distribution								
19	MLE Mean				31.587611				
20	MLE Standard Deviation				19.18102				
21	MLE Coefficient of Variation				0.6072324				
22	MLE Skewness				2.0456027				
23	MLE Median				26.999623				
24	MLE 80% Quantile				43.346989				
25	MLE 90% Quantile				55.464815				
26	MLE 95% Quantile				67.859553				
27	MLE 99% Quantile				99.382188				
28									
29	MVU Estimate of Median				26.807641				
30	MVU Estimate of Mean				31.332966				
31	MVU Estimate of Sd				18.480569				
32	MVU Estimate of SE of Mean				3.9208996				
33									
34	95% Non-parametric UCL								
35	Adjusted-CLT UCL (Adjusted for Skewness)				43.376415				
36	Modified-t UCL (Adjusted for Skewness)				41.47558				
37	Hall's Bootstrap UCL				80.951244				
38	95% Chebyshev (Mean, Sd) UCL				54.582557				
39	97.5% Chebyshev (Mean, Sd) UCL				64.255707				
40	99% Chebyshev (Mean, Sd) UCL				83.256736				
41									
42	UCLs (Assuming Lognormal Distribution)								
43	95% H-UCL				40.527674				
44	95% Chebyshev (MVUE) UCL				48.423771				
45	97.5% Chebyshev (MVUE) UCL				55.818976				
46	99% Chebyshev (MVUE) UCL				70.345424				
47									
48	Data are lognormal (0.05)								
49									
50	Recommended UCL to use:								
51	Use H-UCL								

Lognormal Statistics

- Use the Print or Save As option under File menu or double right click on the UCL output spreadsheet to view a screen with more options to save, print, or write this output sheet to a file.

8f. Display After Selecting the Non-Parametric UCLs Option



	A	B	C	D	E	F	G	H	I
1	Data File	C:\ProUCL\Data\track.xls				Variable:	SI		
2									
3	Number of Valid Samples				19				
4	Number of Unique Samples				13				
5	Minimum				0.05				
6	Maximum				69.5				
7	Mean				6.0651579				
8	Median				0.12				
9	Standard Deviation				17.421608				
10	Variance				303.51243				
11	Coefficient of Variation				2.872408				
12	Skewness				3.2642255				
13	Mean of log data				-1.322622				
14	Standard Deviation of log data				2.1718122				
15									
16	95% UCL (Adjusted for Skewness)								
17	Adjusted-CLT UCL				15.837417				
18	Modified-t UCL				13.49469				
19									
20	95% Non-parametric UCL								
21	CLT UCL				12.639294				
22	Jackknife UCL				12.995847				
23	Standard Bootstrap UCL				12.472003				
24	Bootstrap-t UCL				63.261944				
25	Hall's Bootstrap UCL				74.990748				
26	Percentile Bootstrap UCL				13.367789				
27	BCA Bootstrap UCL				18.762789				
28	95% Chebyshev (Mean, Sd) UCL				23.486766				
29	97.5% Chebyshev (Mean, Sd) UCL				31.02511				
30	99% Chebyshev (Mean, Sd) UCL				45.832726				
31									
32	Data are Non-parametric (0.05)								
33									
34	Recommended UCL to use:								
35	Use 99% Chebyshev (Mean, Sd) UCL								

- The program notes that the data follow an approximate gamma distribution, and suggests in blue that the user should try Gamma UCLs.
- Save this output spreadsheet by using the Save As option under the File menu.
- Double right click on the UCL output spreadsheet to view a screen with more options to save, print, or write this output sheet to a file.

8g. Display After Selecting the All UCLs Option

ProUCL Version 3.0 - [UCL Statistics for Zn]

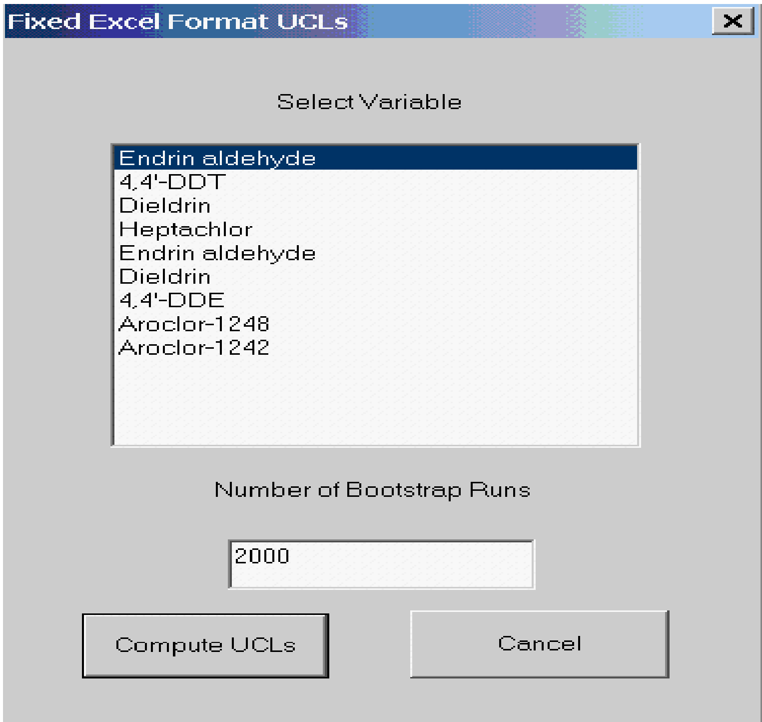
File Edit View Options Summary Statistics Histogram Goodness-of-Fit Tests UCLs Window Help

	A–C	D	E–H	I
1	Data File C:\ProUCL\Data\track.xls		Variable: Zn	
2				
3	Raw Statistics		Normal Distribution Test	
4	Number of Valid Samples	22	Shapiro-Wilk Test Statisitic	0.8179533
5	Number of Unique Samples	22	Shapiro-Wilk 5% Critical Value	0.911
6	Minimum	35.6	Data not normal at 5% significance level	
7	Maximum	120		
8	Mean	56.347727	95% UCL (Assuming Normal Distribution)	
9	Median	54.65	Student's-t UCL	63.649652
10	Standard Deviation	19.903652		
11	Variance	396.15535	Gamma Distribution Test	
12	Coefficient of Variation	0.353229	A-D Test Statistic	0.6315704
13	Skewness	1.8624475	A-D 5% Critical Value	0.7434474
14			K-S Test Statistic	0.1313278
15	Gamma Statistics		K-S 5% Critical Value	0.1852904
16	k hat	10.466773	Data follow gamma distribution	
17	k star (bias corrected)	9.0697887	at 5% significance level	
18	Theta hat	5.3834862		
19	Theta star	6.2126835	95% UCLs (Assuming Gamma Distribution)	
20	nu hat	460.53801	Approximate Gamma UCL	63.565559
21	nu star	399.0707	Adjusted Gamma UCL	64.142003
22	Approx.Chi Square Value (.05)	353.75646		
23	Adjusted Level of Significance	0.0386	Lognormal Distribution Test	
24	Adjusted Chi Square Value	350.57725	Shapiro-Wilk Test Statisitic	0.9260604
25			Shapiro-Wilk 5% Critical Value	0.911
26	Log-transformed Statistics		Data are lognormal at 5% significance level	
27	Minimum of log data	3.5723456		
28	Maximum of log data	4.7874917	95% UCLs (Assuming Lognormal Distribution)	
29	Mean of log data	3.9830117	95% H-UCL	63.587309
30	Standard Deviation of log data	0.30625	95% Chebyshev (MVUE) UCL	72.348122
31	Variance of log data	0.0937891	97.5% Chebyshev (MVUE) UCL	79.36529
32			99% Chebyshev (MVUE) UCL	93.149158
33				
34			95% Non-parametric UCLs	
35			CLT UCL	63.327619
36			Adj-CLT UCL (Adjusted for skewness)	65.128042
37			Mod-t UCL (Adjusted for skewness)	63.930481
38			Jackknife UCL	63.649652
39			Standard Bootstrap UCL	62.968288
40			Bootstrap-t UCL	67.192818
41	RECOMMENDATION		Hall's Bootstrap UCL	78.089743
42	Data follow gamma distribution (0.05)		Percentile Bootstrap UCL	63.509091
43			BCA Bootstrap UCL	67.022727
44	Use Approximate Gamma UCL		95% Chebyshev (Mean, Sd) UCL	74.844596
45			97.5% Chebyshev (Mean, Sd) UCL	82.848206
46			99% Chebyshev (Mean, Sd) UCL	98.569749

General Statistics

- For explanations of the methods and statistics used, refer to Appendix A.
- Use the Print or Save As option under File menu or double right click on the UCL output spreadsheet to view a screen with more options to save, print, or write this output to a file.

8h. Result After Clicking on Fixed Excel Format Drop-Down Menu Item



- Note that the UCLs are computed for one variable at a time. The user selects a variable from the variable list.

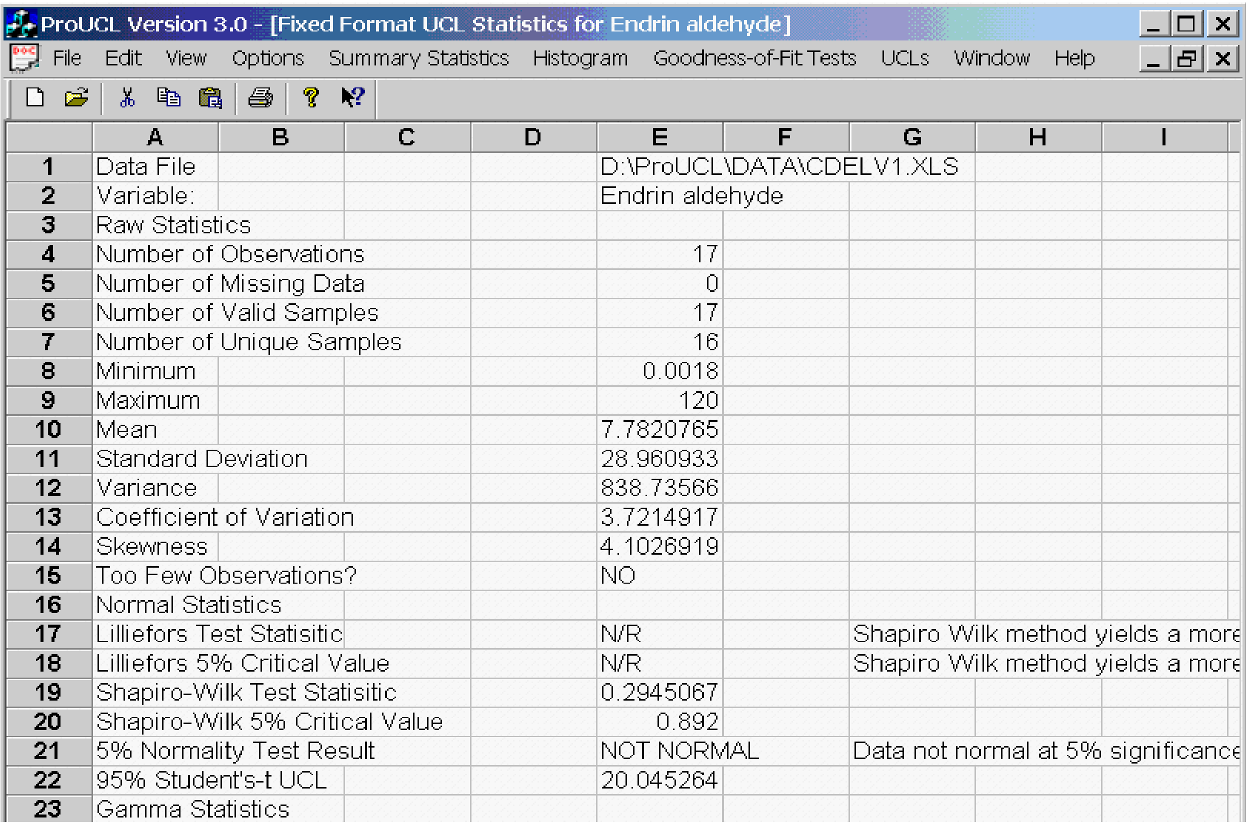
- For this Fixed Format option, the 0.95 Confidence Coefficient is used in all UCL computations.

- The user may adjust the number of bootstrap runs (default is 2,000).

- Click on the Compute UCLs button to display the results.

- This option will display all statistics computed by ProUCL for each of the three parametric distributions and also for all non-parametric methods including the five bootstrap methods.

8i. Results After Clicking the Fixed Excel Format Compute UCLs Button



	A	B	C	D	E	F	G	H	I
1	Data File				D:\ProUCL\DATA\CDELV1.XLS				
2	Variable:				Endrin aldehyde				
3	Raw Statistics								
4	Number of Observations				17				
5	Number of Missing Data				0				
6	Number of Valid Samples				17				
7	Number of Unique Samples				16				
8	Minimum				0.0018				
9	Maximum				120				
10	Mean				7.7820765				
11	Standard Deviation				28.960933				
12	Variance				838.73566				
13	Coefficient of Variation				3.7214917				
14	Skewness				4.1026919				
15	Too Few Observations?				NO				
16	Normal Statistics								
17	Lilliefors Test Statisitic				N/R		Shapiro Wilk method yields a more		
18	Lilliefors 5% Critical Value				N/R		Shapiro Wilk method yields a more		
19	Shapiro-Wilk Test Statisitic				0.2945067				
20	Shapiro-Wilk 5% Critical Value				0.892				
21	5% Normality Test Result				NOT NORMAL		Data not normal at 5% significance		
22	95% Student's-t UCL				20.045264				
23	Gamma Statistics								

- Note that the output is not sized to fit a printed page.

- This option can be omitted by all users who are not planning to import the ProUCL calculation results into some other software to automate the calculations of exposure point concentration terms. That is, all users who are not planning to use ProUCL as a production tool to produce UCLs for several variables and data files may skip the use of this option.

- On Fixed Format output spreadsheet, each row contains a single item description or calculated statistic.

- Three primary columns contain information:
  - Column A is a description of the various results and statistics.
  - Column E contains all appropriate calculated results.
  - Column G contains additional descriptive information as needed.
  - Note that information from the primary columns (e.g., A, E, and G) may overflow into the columns to the right.

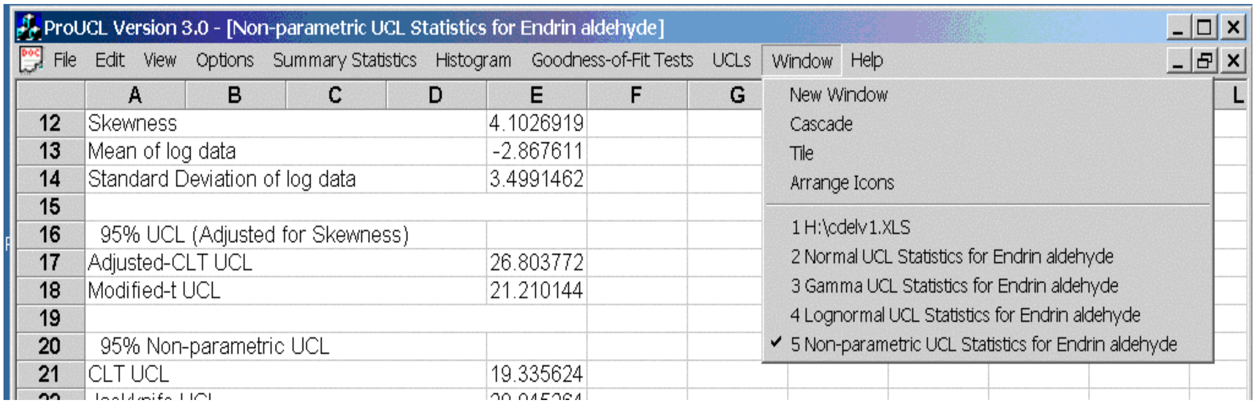
- For column E:
  - N/A means that the calculation for the associated statistic is not available.
  - N/R means that the calculations for the associated statistic may not be reliable.
  - Row 15 displays YES if there are too few observations to calculate appropriate UCL statistics and displays NO if enough observations are available to compute all relevant statistics and UCLs.
  - Row 35 displays AD GAMMA (if data are gamma distributed using A-D test) or NOT AD GAMMA (if data are not gamma distributed using A-D test) using the Anderson-Darling Gamma Test for 0.05 level of significance.
  - Similarly, Row 38 displays KS GAMMA or NOT KS GAMMA using the Kolmogorov-Smirnov Gamma Test for 0.05 level of significance.
  - As mentioned before, it should be noted that these two goodness-of-fit tests may lead to different conclusion (as is the case with other goodness-of-fit tests) about the data distribution. In that case, ProUCL leads to the conclusion that the data follow an approximate gamma distribution.
  - Row 39 displays NOT GAMMA, APPROX GAMMA, or GAMMA depending on the results of the two Gamma goodness-of-fit tests.
  - Row 52 displays LOGNORMAL or NOT LOGNORMAL depending on the result of the appropriate lognormality test for 0.05 level of significance.
  - Row 86 displays YES if user inspection is recommended and displays NO if no potential problems requiring manual inspection needed with the selected variable.
  - Row 87 displays NORMAL, GAMMA, LOGNORMAL, or NON-PARAMETRIC as the distribution used in determining 95% UCL computation recommendations.
  - Row 88 displays a recommended UCL value to use as an estimate of the EPC term.
  - Row 89 displays a second recommended UCL (e.g., use of either Hall's bootstrap or bootstrap-t method may be recommended on the same data set). These cells will be blank if only one UCL is recommended for the selected variable.
  - Row 90 displays a third recommended UCL. These cells will be blank if only one or two UCLs are recommended for the selected variable.
  - Row 91 displays YES if the recommended 95% UCL exceeds the maximum value in the data set.
  - Row 92 displays PLEASE CHECK if the recommended bootstrap UCLs are subject to erratic or inflated values due to possible presence of outliers. Otherwise, row 92 displays NONE.
  - Row 93 displays IN CASE if the recommended bootstrap UCL has an inflated value due to the presence of outliers. Otherwise, row 93 displays NONE.

- For column G:
  - Row 88 displays the name of the recommended 95% UCL.
  - Row 89 displays the name of the second recommended 95% UCL. These cells will be blank if only one UCL is recommended for the selected variable.

- Row 90 displays the name of the third recommended 95% UCL. These cells will be left blank if only one UCL is recommended for the selected variable.
- Row 93 displays the name of the alternative UCL to utilize if the recommended bootstrap (e.g., bootstrap-t or Hall's bootstrap) 95% UCL has an inflated value due to presence of potential outliers.

9. Window

Click on the Window menu to reveal these drop-down options.

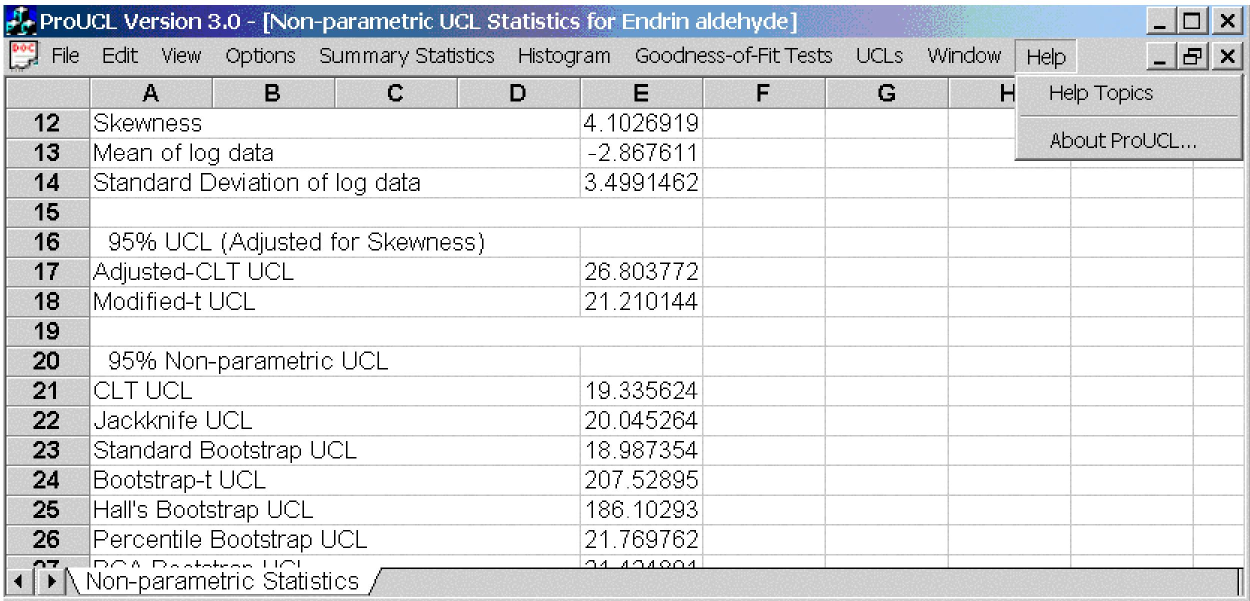


The following Window drop-down menu options are available:

- New Window option: opens a blank spreadsheet window.
- Cascade option: arranges windows in a cascade format. This is similar to a typical Windows program option.
- Tile option: resizes each window and then displays all open windows. This is similar to a typical Windows program option.
- Arrange Icons: similar to a typical Windows program option.
- The drop-down options include a list of all open windows with a check mark in front of the active window. Click on any of the windows listed to make that window active.

10. Help

Click on the Help menu item to reveal these drop-down options.



The following Help drop-down menu options are available:

- Help Topics option: ProUCL version 3.0 does not have an online help program.

- About ProUCL: displays the software version number.

Run Time Notes

- Cell size can be changed. The user can change the size of a cell by moving the mouse to the top row (the gray shaded row with a letter), then moving the mouse to the right side until the cursor changes to an arrow symbol (↔), depress the left mouse button.

- This can be used to reveal additional precision or hidden text.

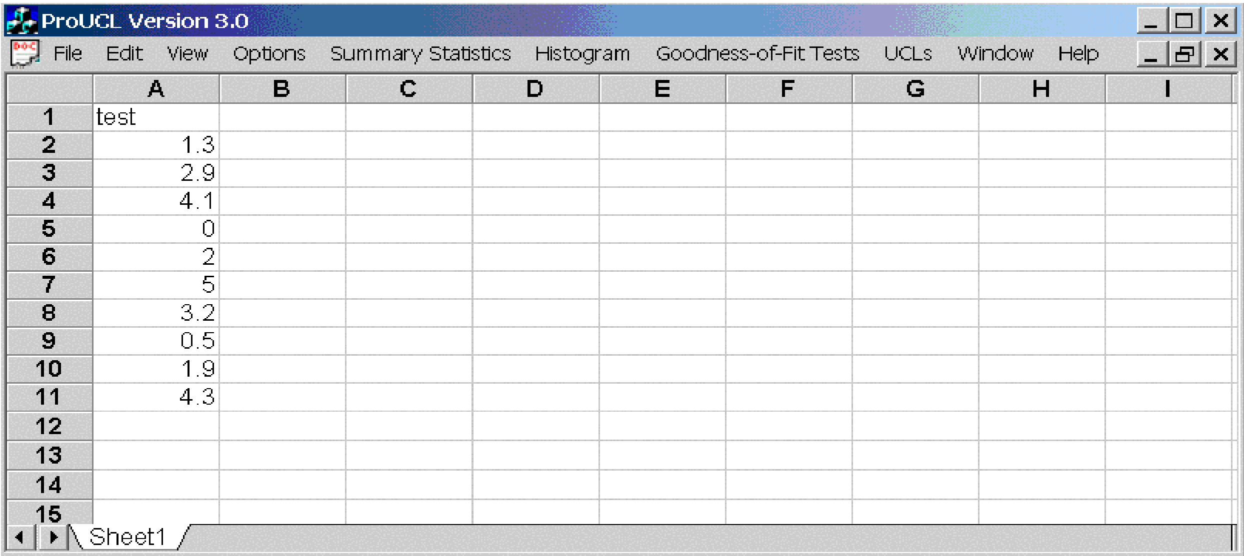
ProUCL Version 3.0 - [Gamma UCL Statistics for Endrin aldehyde]

File Edit View Options Summary Statistics Histogram Goodness-of-Fit Tests UCLs Window Help

	A	B	C	D	E	F	G	H
1	Data File	H:\cdelv1.XLS				Variable:	Endrin aldehyde	
2								
3	Number of Valid Samples			17				
4	Number of Unique Samples			16				
5	Minimum			0.0018				
6	Maximum			120				
7	Mean			7.782076471				
8	Standard Deviation			28.96093326				
9	Variance			838.7356553				
10	k hat			0.155424908				
11	k star (bias corrected)			0.1672126693				
12	Theta hat			50.06968684				
13	Theta star			46.53999306				
14	nu hat			5.284446872				
15	nu star			5.685230757				
16	Approx.Chi Square Value (.05)			1.480711891				
17	Adjusted Level of Significance			0.03461				
18	Adjusted Chi Square Value			1.269069171				

Gamma Statistics

Rules to Remember When Editing or Creating a New Data File



- Text may appear in the first row only. This row has column headers (variable names) for your data.

- All alphanumeric text (including blanks, strings) appearing elsewhere (other than first row) will be treated as zero data.

- Missing data (alphanumeric text, blanks) can be set to a large value such as $1x10^{31}$. All entries with this value will be ignored from the computations.

- The last data entry for each column must be non-zero. The program determines the number of observations by working backwards up the data until a non-zero value is encountered. Data in each column must end with a non-zero entry as shown above otherwise that zero value will be ignored. All intermediate zero entries are treated as valid data.

- It is recommended to use the default settings of the Data location screen when working with your data sets.

C. Recommendations to Compute a 95% UCL of the Population Mean (The Exposure Point Concentration Term)

This section describes the recommendations on the computation of a 95% *UCL* of the unknown population arithmetic mean, μ_1, of a contaminant data distribution. These recommendations are based upon the findings of Singh, Singh, and Engelhardt (1997, 1999); Singh et al. (2002a); Singh, Singh, and Iaci (2002b); and Singh and Singh (2003). These recommendations are applicable to full data sets without censoring and non-detect observations.

Recommendations have been summarized for:

1) normally distributed data sets,
2) gamma distributed data sets,
3) lognormally distributed data sets, and
4) data sets which are non-parametric and do not follow any of the above mentioned three distributions included in ProUCL.

A detailed description of the recommendations can be found in Section 5 of Appendix A. Also, a list of all cited references is given in Appendix A.

For skewed parametric as well as non-parametric data sets, there is no simple solution to compute a *95% UCL* of the population mean, μ_1. Contrary to the general conjecture, Singh et al. (2002a), Singh, Singh, and Iaci (2002b), and Singh and Singh (2003) noted that the *UCLs* based upon the skewness adjusted methods, such as the Johnson's modified-t and Chen's adjusted-*CLT* do not provide the specified coverage (e.g., 95 %) to the population mean even for mildly to moderately skewed (e.g., $\hat{\sigma}$, the *sd* of log-transformed data in interval [0.5, 1.0)) data sets for samples of size as large as 100. The coverage of the population mean by these skewness-adjusted *UCLs* becomes poorer (much smaller than the specified coverage of 0.95) for highly skewed data sets, where the skewness levels are defined in Section 3.2.2 of Appendix A as a function of σ or $\hat{\sigma}$ (standard deviation of log-transformed data).

It should be noted that even though, the simulation results for highly skewed data sets of small sizes suggest that the bootstrap-t and Hall's bootstrap methods do approximately provide the adequate coverage to the population mean, sometimes in practice these two bootstrap methods yield erratic inflated values (orders of magnitude higher than the other *UCL* values) when dealing with individual highly skewed data sets of small sizes. This is especially true when potential outliers may be present in the data set. Therefore, ProUCL Version 3.0 provides warning messages whenever the recommendations are made regarding the use the bootstrap-t method or Hall's bootstrap method.

D. Recommendations to Compute a 95% *UCL* of the Population Mean, μ_1 Using Symmetric and Positively Skewed Data Sets

Graphs from Singh and Singh (2003) showing coverage comparisons (e.g., attainment of the specified confidence coefficient) for normal, gamma, and lognormal distributions for the various methods considered are given in Appendix C. The user may want to consult those graphs for a better understanding of the recommendations summarized in this section.

1. Normally or Approximately Normally Distributed Data Sets

- For normally distributed data sets, a *UCL* based upon the Student's-t statistic as given by equation (32) of Appendix A provides the optimal *UCL* of the population mean. Therefore, for normally distributed data sets, one should always use a 95% *UCL* based upon the Student's-t statistic.

- The *95% UCL* of the mean given by equation (32) based upon Student's-t statistic may also be used when the *sd,* s_y of the log-transformed data is less than 0.5, or when the data set approximately follows a normal distribution. A data set is approximately normal when the normal Q-Q plot displays a linear pattern (without outliers and significant jumps) and the resulting correlation coefficient is quite high (e.g., 0.95 or higher).

- Student's-t *UCL* may also be used when the data set is symmetric (but possibly not normally distributed). A measure of symmetry (or skewness) is $\hat{k}_3$, which is given by equation (43) of Appendix A. As a rule of thumb, a value of $\hat{k}_3$ close to zero (e.g., $|\hat{k}_3| < 0.2 - 0.3$) suggests approximate symmetry. The approximate symmetry of a data distribution can also be judged by evaluating the histogram of the data set.

2. Gamma Distributed Skewed Data Sets

In practice, many skewed data sets can be modeled both by a lognormal distribution and a gamma distribution, especially when the sample size is smaller than 100. Land's H-statistic based, *95% H-UCL* of the mean based upon a lognormal model often results in an unjustifiably large and impractical 95% *UCL* value. In such cases, a gamma model, G (k,θ), may be used to compute a reliable 95% *UCL* of the unknown population mean, μ_1.

- Many skewed data sets follow a lognormal as well as a gamma distribution. It should be noted that the population means based upon the two models can differ significantly. The lognormal model, based upon a highly skewed (e.g., $\hat{\sigma} \geq 2.5$) data set, will have an unjustifiably large and impractical population mean, μ_1, and its associated *UCL*. The gamma distribution is better suited to model positively skewed environmental data sets.

  One should always first check if a given skewed data set follows a gamma distribution. If a data set does follow a gamma distribution or an approximate gamma distribution, one should compute a 95% *UCL* based upon a gamma distribution. Use of highly skewed (e.g., $\hat{\sigma} \geq$ 2.5-3.0) lognormal distributions should be avoided. For such highly skewed lognormally distributed data sets that can not be modeled by a gamma or an approximate gamma distribution, non-parametric *UCL* computation methods based upon the Chebyshev inequality may be used. ProUCL prints out at least one recommended *UCL* associated with each data set.

- The five bootstrap methods do not perform better than the two gamma *UCL* computation methods. It is noted that the performances (in terms of coverage probabilities) of bootstrap-t and Hall's bootstrap methods are very similar. Out of the five bootstrap methods, bootstrap-t and Hall's bootstrap methods perform the best (with coverage probabilities for the population mean closer to the nominal level of 0.95). This is especially true when skewness is quite high (e.g., $\hat{k}$ < 0.1) and sample size is small (e.g., n < 10-15). This is illustrated in the graphs given in Appendix C. As mentioned before, whenever the use of Hall's *UCL* or bootstrap-t *UCL* is recommended, an informative warning message about their use is also printed.

- Also, contrary to the conjecture, the bootstrap BCA method does not perform better than the Hall's method or the bootstrap-t method. The coverage for the population mean, μ_1 provided by the BCA method is much lower than the specified 95% coverage. This is especially true when the skewness is high (e.g., $\hat{k}$ < 1) and sample size is small (Singh and Singh (2003)).

- From the results presented in Singh, Singh, and Iaci (2002b) and in Singh and Singh (2003), it is concluded that for data sets which follow a gamma distribution, a 95% *UCL* of the mean should be computed using the adjusted gamma *UCL* when the shape parameter, *k*, is:

$0.1 \le k < 0.5$, and for values of $k \ge 0.5$, a 95% *UCL* can be computed using an approximate gamma *UCL* of the mean, μ_1.

- For highly skewed gamma distributed data sets with $k < 0.1$, bootstrap-t *UCL* or Hall's bootstrap (Singh and Singh (2003)) may be used when the sample size is small (e.g., n < 15) and adjusted gamma *UCL* should be used when sample size starts approaching and exceeding 15. The small sample size requirement increases as skewness increases (that is as *k* decreases, n is required to increase).

- It should be pointed out that the bootstrap-t and Hall's bootstrap methods should be used with caution as some times these methods yield erratic, unreasonably inflated, and unstable *UCL* values, especially in the presence of outliers. In case Hall's bootstrap and bootstrap-t methods yield inflated and erratic *UCL* results, the 95% *UCL* of the mean should be computed based upon adjusted gamma *UCL*.

These recommendations for the use of gamma distribution are summarized in Table 1.

Table 1
Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1 of a Gamma Distribution

$\hat{k}$	***Sample Size, n***	***Recommendation***
$\hat{k} \ge 0.5$	For all n	Approximate Gamma 95%*UCL*
$0.1 \le \hat{k} < 0.5$	For all n	Adjusted Gamma 95% *UCL*
$\hat{k} < 0.1$	n < 15	95% *UCL* Based Upon Bootstrap-t or Hall's Bootstrap Method *
	n $\ge$ 15	Adjusted Gamma 95% *UCL* if available, otherwise use Approximate Gamma 95% *UCL*

* If bootstrap-t or Hall's bootstrap methods yield erratic, inflated, and unstable *UCL* values (which often happens when outliers are present), the *UCL* of the mean should be computed using adjusted gamma *UCL*.

3. Lognormally Distributed Skewed Data Sets

For lognormally distributed data sets, LN(μ, σ^2), the H-statistic based *UCL* provides the specified 0.95 coverage for the population mean for all values of σ. However, the H-statistic often results in unjustifiably large *UCL* values which do not occur in practice. This is especially true when skewness is high (e.g., $\sigma > 2.0$). The use of a lognormal model unjustifiably accommodates large and impractical values of the mean concentration and its *UCLs*. The problem associated with the use of a lognormal distribution is that the population mean, μ_1 of a lognormal model becomes impractically large for larger values of σ, which in turn results in inflated *H-UCL* of the population mean, μ_1. Since the population mean of a lognormal model becomes too large, none of the other methods except for the inflated *H-UCL* provides the specified 95% coverage for that inflated population mean, μ_1. This is especially true when the sample size is small and skewness is high. For extremely skewed data sets (with $\sigma > 2.5$-3.0) of sizes (e.g., < 70-100), the use of a lognormal distribution based *H-UCL* should be avoided (e.g., see Singh et al. (2002a), Singh and Singh (2003)). Therefore, alternative *UCL* computation methods such as the use of a gamma distribution, or the use of a *UCL* based upon non-parametric bootstrap methods or Chebyshev inequality based methods, are desirable. All skewed data sets should first be tested for a gamma distribution. For lognormally distributed data sets (that can not be modeled by a gamma distribution), the method as summarized in Table 2 on the following page, may be used to compute a 95% *UCL* of the mean. The details can be found in Appendix A.

ProUCL can compute an *H-UCL* for samples of sizes up to 1000. For highly skewed lognormally distributed data sets of smaller sizes, some alternative methods to compute a 95% *UCL* of the population mean, μ_1, are summarized in Table 2. Since skewness (as defined in Section 3.2.2, Appendix A) is a function of σ (or $\hat{\sigma}$), the recommendations for the computation of the *UCL* of the population mean are also summarized in Table 2 for various values of the *MLE*, $\hat{\sigma}$ of σ and the sample size, n. Here $\hat{\sigma}$ is an *MLE* of σ, and is given by the *Sd* of log-transformed data given by equation (2) of Appendix A. Note that Table 2 is only applicable to the computation of a 95% *UCL* of the population mean based upon lognormally distributed data sets without non-detect observations.

Table 2
Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1 of a Lognormal Population

$\hat{\sigma}$	*Sample Size, n*	*Recommendation*
$\hat{\sigma} < 0.5$	For all n	Student's-t, modified-t, or *H-UCL*
$0.5 \le \hat{\sigma} < 1.0$	For all n	*H-UCL*
$1.0 \le \hat{\sigma} < 1.5$	$n < 25$	95% Chebyshev *(MVUE) UCL*
	$n \ge 25$	*H-UCL*
$1.5 \le \hat{\sigma} < 2.0$	$n < 20$	99% Chebyshev *(MVUE) UCL*
	$20 \le n < 50$	95% Chebyshev *(MVUE) UCL*
	$n \ge 50$	*H-UCL*
$2.0 \le \hat{\sigma} < 2.5$	$n < 20$	99% Chebyshev *(MVUE) UCL*
	$20 \le n < 50$	97.5% Chebyshev *(MVUE) UCL*
	$50 \le n < 70$	95% Chebyshev *(MVUE) UCL*
	$n \ge 70$	*H-UCL*
$2.5 \le \hat{\sigma} < 3.0$	$n < 30$	Larger of (99% Chebyshev *(MVUE) UCL*, 99% Chebyshev(Mean, Sd))
	$30 \le n < 70$	97.5% Chebyshev *(MVUE) UCL*
	$70 \le n < 100$	95% Chebyshev *(MVUE) UCL*
	$n \ge 100$	*H-UCL*
$3.0 \le \hat{\sigma} \le 3.5$	$n < 15$	Hall's bootstrap method *
	$15 \le n < 50$	Larger of (99% Chebyshev (*MVUE) UCL*, 99% Chebyshev(Mean, Sd))
	$50 \le n < 100$	97.5% Chebyshev *(MVUE) UCL*
	$100 \le n < 150$	95% Chebyshev *(MVUE) UCL*
	$n \ge 150$	*H-UCL*
$\hat{\sigma} > 3.5$	For all n	Use non-parametric methods *

* If Hall's bootstrap method yields an erratic unrealistically large *UCL* value, then the *UCL* of the mean may be computed based upon the Chebyshev inequality.

4. Data Sets Without a Discernable Skewed Distribution - Non-parametric Skewed Data Sets

The use of gamma and lognormal distributions as discussed here will cover a wide range of skewed data distributions. For skewed data sets which are neither gamma nor lognormal, one can use a non-parametric Chebyshev *UCL* or Hall's bootstrap *UCL* (for small data sets) of the mean to estimate the EPC term.

- For skewed non-parametric data sets with negative and zero values, use a *95% Chebyshev (Mean, Sd) UCL* of the mean, μ_1 to estimate the EPC term.

- For all other non-parametric data sets with only positive values, the following method may be used to estimate the EPC term:

- For mildly skewed data sets with $\hat{\sigma} \leq 0.5$, one can use the Student's-t statistic or modified-t statistic to compute a 95% *UCL* of the mean, μ_1.

- For non-parametric moderately skewed data sets (e.g., σ or its estimate, $\hat{\sigma}$ in the interval (0.5, 1]), one may use a *95% Chebyshev (Mean, Sd) UCL* of the population mean, μ_1.

- For non-parametric moderately to highly skewed data sets (e.g., $\hat{\sigma}$ in the interval (1.0, 2.0]), one may use a *99% Chebyshev (Mean, Sd) UCL or 97.5% Chebyshev (Mean, Sd) UCL* of the population mean, μ_1, to obtain an estimate of the EPC term.

- For highly skewed to extremely highly skewed data sets with $\hat{\sigma}$ in the interval (2.0, 3.0], one may use Hall's *UCL* or *99% Chebyshev (Mean, Sd) UCL* to compute the EPC term.

- Extremely skewed non-parametric data sets with σ exceeding 3.0 are badly behaved and *UCLs* based upon such data sets often provide poor coverage to the population mean. For such highly skewed data distributions, none of the methods considered provide the specified 95% coverage for the population mean, μ_1. The coverages provided by the various methods decrease as σ increases. For such highly skewed data sets of sizes (e.g., < 30), a 95% *UCL* can be computed based upon Hall's bootstrap method or bootstrap-t method. Hall's bootstrap method provides the highest coverage (but less than 0.95) when the sample size is small. It is noted that the coverage for the population mean provided by Hall's method (and bootstrap-t method) does not increase much as the sample size, n increases. However, as the sample size increases, coverage provided by *99% Chebyshev (Mean, Sd) UCL* method increases. Therefore, for larger samples, a *UCL* should be computed based upon *99% Chebyshev (Mean, Sd)* method. This large sample size requirement increases as $\hat{\sigma}$ increases (e.g., n increases as *Sd* increases). These recommendations are summarized in Table 3 given in the following.

Note: As mentioned before, the Hall's bootstrap method (and also bootstrap-t method) sometimes yields erratic and unstable *UCL* values, especially when the outliers are present. If Hall's bootstrap *UCL* represents an erratic and unstable value, a *UCL* of the population mean may be computed using the *99% Chebyshev (Mean, Sd)* method.

Table 3
Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1 of a Skewed Non-parametric Distribution with all Positive Values, Where $\hat{\sigma}$ is the Sd of Log-transformed Data

$\hat{\sigma}$	***Sample Size, n***	***Recommendation***
$\hat{\sigma} \leq 0.5$	For all n	95% *UCL* based upon Student's-t statistic or Modified-t statistic
$0.5 < \hat{\sigma} \leq 1.0$	For all n	95% Chebyshev (Mean, Sd) *UCL*
$1.0 < \hat{\sigma} \leq 2.0$	n < 50	99% Chebyshev (Mean, Sd) *UCL*
	n ≥ 50	97.5% Chebyshev (Mean, Sd) *UCL*
$2.0 < \hat{\sigma} \leq 3.0$	n < 10	Hall's Bootstrap *UCL* *
	n ≥ 10	99% Chebyshev (Mean, Sd) *UCL*
$3.0 < \hat{\sigma} \leq 3.5$	n < 30	Hall's Bootstrap *UCL* *
	n ≥ 30	99% Chebyshev (Mean, Sd) *UCL*
$\hat{\sigma} > 3.5$	n < 100	Hall's Bootstrap *UCL* *
	n ≥ 100	99% Chebyshev (Mean, Sd) *UCL*

* If the Hall's bootstrap method yields an erratic and unstable *UCL* value (e.g., this tends to happen when outliers are present), the EPC term may be computed using the 99% Chebyshev (Mean, Sd) UCL.

E. Should the Maximum Observed Concentration be Used as an Estimate of the EPC Term?

Singh and Singh (2003) also included the Max Test (using the maximum observed value as an estimate of the EPC term) in their simulation study. Previous (e.g., EPA 1992 RAGS Document) use of the maximum observed value has been recommended as a default value to estimate the EPC term when a 95% *UCL* (e.g., the *H-UCL*) exceeded the maximum value. Only two 95% *UCL* computation methods, namely: the Student's- t *UCL* and Land's *H-UCL* were used previously to estimate the EPC term (e.g., EPA 1992). ProUCL can compute a 95% *UCL* of mean using several methods based upon normal, Gamma, lognormal, and non-parametric distributions. Thus, ProUCL has about fifteen (15) 95% *UCL* computation methods, at least one of which (depending upon skewness and data distribution) can be used to compute an appropriate estimate of the EPC term. Furthermore, since the EPC term represents the average exposure contracted by an individual over an exposure area (EA) during a long period of time; therefore, the EPC term should be estimated by using an average value (such as an appropriate 95% *UCL* of the mean) and not by the maximum observed concentration. With the availability of so many *UCL* computation methods, the developers of ProUCL, Version 3.0 do not feel any need to use the maximum observed value as an estimate of the EPC term. Singh and Singh (2003) also noted that for skewed data sets of small sizes (e.g., <10-20), the Max Test does not provide the specified 95% coverage to the population mean, and for larger data sets, it overestimates the EPC term which may require unnecessary further remediation. This can also be viewed in the graphs presented in Appendix C. Also, for the distributions considered, the maximum value is not a sufficient statistic for the unknown population mean. The use of the maximum value as an estimate of the EPC term ignores most (except for the maximum value) of the information contained in a data set. It is, therefore not desirable to use the maximum observed value as an estimate of the EPC term representing average exposure by an individual over an EA. **It is recommended that the maximum observed value NOT be used as an estimate of the EPC term.** However, for the sake of interested users, ProUCL displays a warning message when the recommended 95% *UCL* (e.g., Hall's bootstrap *UCL* etc.) of the mean exceeds the observed maximum concentration. For such cases (when a 95% *UCL* does exceed the maximum observed value), if applicable, an alternative *UCL* computation method is recommended by ProUCL.

It should also be noted that for highly skewed data sets, the sample mean indeed can even exceed the upper 90%, 95 % etc. percentiles, and consequently, a 95% *UCL* of mean can exceed the maximum observed value of a data set. This is especially true when one is dealing with lognormally distributed data sets of small sizes. For such highly skewed data sets which can not be modeled by a gamma distribution, a 95% *UCL* of the mean should be computed using an appropriate non-parametric method. These recommendations are summarized in Tables 1 through 3 of this User Guide.

Alternatively, for such highly skewed data sets, other measures of central tendency such as the median (or some higher order quantile such as 70% etc.) and its upper confidence limit may be considered. The EPA, all other interested agencies and parties need to come to an agreement on the use of median and its UCL to estimate the EPC term. However, the use of the sample median and/or its *UCL* as estimates of the EPC term needs further research and investigation.

F. Left-Censored Data Sets with Non-detects

ProUCL does not handle the left-censored data sets with non-detects, which are inevitable in many environmental studies. All parametric as well as non-parametric recommendations to compute the mean, standard deviation, and a 95% *UCL* of the mean made by ProUCL software are based upon full data sets without censoring. For mild to moderate number of non-detects (e.g., < 15%), one may compute these statistics based upon the commonly used rule of thumb of using ½ detection limit (DL) proxy method. However, the proxy methods should be used cautiously, especially when one is dealing with lognormally distributed data sets. For lognormally distributed data sets of small sizes, even a single value -- small (e.g., obtained after replacing the non-detect by ½ DL) or large (e.g., an outlier) can have a drastic influence (can yield an unrealistically large 95% UCL) on the value of the associated Land's 95% *UCL*. The issue of estimating the mean, standard deviation, and a 95% *UCL* of the mean based upon left-censored data sets of varying degrees (e.g., <15%, 15%-50%, 50%-75%, or greater than 75% etc.) of censoring is currently under investigation.

Glossary

This glossary defines selected words in this User Guide to describe impractically large *UCL* values of the unknown population mean, μ_l. In practice, the *UCLs* based upon Land's H-statistic (*H-UCL*), and some bootstrap methods such as the bootstrap-t and Hall's bootstrap methods (especially when outliers are present) can become impractically large. The *UCLs* based upon these methods often become larger than the *UCLs* based upon all other methods by several orders of magnitude. Such large *UCL* values are not achievable as they do not occur in practice. Words like unstable and unrealistic have been used to describe such impractically large *UCL* values.

***UCL*:** Upper Confidence Limit of the unknown population mean.

Coverage = Coverage Probability: The coverage probability (e.g., = 0.95) of a *UCL* of the population mean represents the confidence coefficient associated with the *UCL*.

Optimum: An interval is optimum if it possesses optimal properties as defined in the statistical literature. This may mean that it is the shortest interval providing the specified coverage (e.g., 0.95) to the population mean. For example, for normally distributed data sets, the *UCL* of the population mean based upon Student's t distribution is optimum.

Stable *UCL*: The *UCL* of a population mean is a stable *UCL* if it represents a number of practical merit, which also has some physical meaning. That is, a stable *UCL* represents a realistic number (e.g., contaminant concentration) that can occur in practice. Also, a stable *UCL* provides the specified (at least approximately, as much as possible, as close as possible to the specified value) coverage (e.g., ~0.95) to the population mean.

Reliable *UCL*: This is similar to a stable *UCL*.

Unstable *UCL* = Unreliable *UCL* = Unrealistic *UCL*: The *UCL* of a population mean is unstable, unrealistic, or unreliable if it is orders of magnitude higher than the various other *UCLs* of population mean. It represents an impractically large value that cannot be achieved in practice. For example, the use of Land's H statistic often results in impractically large inflated *UCL* value. Some other *UCLs* such as the bootstrap-t *UCL* and Hall's *UCL*, can be inflated by outliers resulting in an impractically large and unstable value. All such impractically large *UCL* values are called unstable, unrealistic, unreliable, or inflated *UCLs* in this User Guide.

References

EPA (1992), "Supplemental Guidance to RAGS: Calculating the Concentration Term," Publication EPA 9285.7-081, May 1992.

Gilbert, R.O. (1987), *Statistical Methods for Environmental Pollution Monitoring*, New York: Van Nostrand Reinhold.

Hardin, J.W., and Gilbert, R.O. (1993), "Comparing Statistical Tests for Detecting Soil Contamination Greater Than Background," Pacific Northwest Laboratory, Battelle, Technical Report # DE 94-005498.

Land, C. E. (1971), "Confidence Intervals for Linear Functions of the Normal Mean and Variance," *Annals of Mathematical Statistics*, 42, 1187-1205.

Land, C. E. (1975), "Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance," in *Selected Tables in Mathematical Statistics*, Vol. III, American Mathematical Society, Providence, R.I., 385-419.

Schulz, T. W., and Griffin, S. (1999), Estimating Risk Assessment Exposure Point Concentrations when Data are Not Normal or Lognormal. Risk Analysis, Vol. 19, No. 4, 1999.

Scout: A Data Analysis Program, Technology Support Project. EPA, NERL -LV, Las Vegas, NV 89193-3478.

Singh, A. K., Singh, Anita, and Engelhardt, M., "The Lognormal Distribution in Environmental Applications," EPA/600/R-97/006, December 1997.

Singh, A. K., Singh, Anita, and Engelhardt, M., "Some Practical Aspects of Sample Size and Power Computations for Estimating the Mean of Positively Skewed Distributions in Environmental Applications," EPA/600/S-99/006, November 1999.

Singh, A,. Singh, A.K., Engelhardt, M., and Nocerino, J.M. (2002a), " On the Computation of the Upper Confidence Limit of the Mean of Contaminant Data Distributions." Under EPA Review.

Singh, A., Singh, A. K., and Iaci, R. J. (2002b). " Estimation of the Exposure Point Concentration Term Using a Gamma Distribution." EPA/600/R-02/084.

Singh, A. and Singh, A.K. (2003). Estimation of the Exposure Point Concentration Term (95% UCL) using Bias-Corrected Accelerated (BCA) Bootstrap Method and Several Other Methods for Normal, Lognormal, and Gamma Distributions. Draft EPA Internal Report.

APPENDIX A

TECHNICAL BACKGROUND

METHODS FOR COMPUTING

THE EPC TERM ((1-α) 100%UCL)

AS INCORPORATED IN

ProUCL VERSION 3.0 SOFTWARE

METHODS FOR COMPUTING THE EPC TERM ((1-α) 100%UCL)
AS INCORPORATED IN ProUCL VERSION 3.0 SOFTWARE

1. Introduction

Exposure assessment and cleanup decisions in support of U.S. EPA projects are often made based upon the mean concentrations of the contaminants of potential concern. A 95% upper confidence limit (*UCL*) of the unknown population arithmetic mean (*AM*), μ_1, is often used to: estimate the exposure point concentration (EPC) term (EPA, 1992, EPA, 2002), determine the attainment of cleanup standards (EPA, 1989 and EPA, 1991), estimate background level contaminant concentrations, or compare the soil concentrations with site specific soil screening levels (EPA, 1996). It is, therefore, important to compute a reliable, conservative, and stable 95% *UCL* of the population mean using the available data. The 95% UCL should approximately provide the 95% coverage for the unknown population mean, μ_1. EPA (2002) has developed a guidance document for calculating upper confidence limits for hazardous waste sites. All of the *UCL* computation methods as described in the EPA (2002) guidance document are available in ProUCL, Version 3.0. Additionally, ProUCL, Version 3.0 can also compute a 95% UCL of the mean based upon the gamma distribution which is better suited to model positively skewed environmental data sets.

Computation of a *(1-α) 100% UCL* of the population mean depends upon the data distribution. Typically, environmental data are positively skewed, and a default lognormal distribution (EPA, 1992) is often used to model such data distributions. The H-statistic based Land's (Land 1971, 1975) *H-UCL* of the mean is used in these applications. Hardin and Gilbert (1993), Singh, Singh, and Engelhardt (1997,1999), Schultz and Griffin,1999, Singh et al. (2002a), and Singh, Singh, and Iaci (2002b) pointed out several problems associated with the use of the lognormal distribution and the *H-UCL* of the population *AM*. In practice, for lognormal data sets with high standard deviation (*Sd*), σ of the natural log-transformed data (e.g., σ exceeding 2.0), the *H-UCL* can become unacceptably large, exceeding the 95% and 99% data quantiles, and even the maximum observed concentration, by orders of magnitude (Singh, Singh, and Engelhardt, 1997). This is especially true for skewed data sets of sizes smaller than n < 50 - 70.

The *H-UCL* is also very sensitive to a few low or high values. For example, the addition of a sample with below detection limit measurement can cause the *H-UCL* to increase by a large amount (Singh, Singh, and Iaci, (2002b)). Realizing that the use of H-statistic can result in unreasonably large *UCL*, it has been recommended (EPA, 1992) to use the maximum observed value as an estimate of the *UCL* (EPC term) in cases where the *H-UCL* exceeds the maximum observed value. Recently, Singh, Singh and Iaci (2002b), and Singh and Singh (2003) studied the computation of the *UCLs* based upon a gamma distribution and several non-parametric bootstrap methods. Those methods have also been incorporated in ProUCL, Version 3.0. There are fifteen *UCL* computation methods available in ProUCL; five are parametric and ten are non-parametric. The non-parametric methods do not depend upon any of the data distributions. Graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C.

Both lognormal and gamma distributions can be used to model positively skewed data sets. It should be noted that it is hard to distinguish between a lognormal and a gamma distribution, especially when the sample size is small such as n < 50 - 70. In practice many skewed data sets follow a lognormal as well as a gamma distribution. Singh, Singh, and Iaci (2002b) observed that the *UCL* based upon a gamma distribution results in reliable and stable values of practical merit. It is therefore, always desirable to test if an environmental data set follows a gamma distribution. For data sets (of all sizes) which follow a gamma distribution, EPC should be computed using an adjusted gamma *UCL* (when $0.1 \le k < 0.5$) of the mean or an approximate gamma *UCL* (when $k \ge 0.5$) of the mean as these *UCLs* approximately provide the specified 95% coverage to the population mean, $\mu_1 = k\theta$ of a gamma distribution. For values of $k < 0.1$, a 95% *UCL* may be obtained using bootstrap-t method or Hall's bootstrap method when the sample size, n is less than 15, and for larger samples, a *UCL* of the mean should be computed using the adjusted or approximate gamma *UCL*. Here, k is the shape parameter of a gamma distribution as described in Section 2.2. It should be pointed out that both bootstrap-t and Hall's bootstrap methods sometimes result in erratic, inflated, and unstable *UCL* values especially in the presence of outliers. Therefore, these two methods should be used with caution. The user should examine the various *UCL* results and determine if the *UCLs* based upon the bootstrap-t and Hall's bootstrap methods represent reasonable and reliable *UCL* values of practical merit. If the results based upon these two methods are much higher than the rest of methods (except for the *UCLs* based upon lognormal distribution), then this could be an indication of erratic *UCL* values. ProUCL prints out a warning message whenever the use of these two bootstrap methods is recommended. In case these two bootstrap methods yield erratic and inflated *UCLs*, the *UCL* of the mean should be computed using the adjusted or the approximate gamma *UCL* computation method.

ProUCL has been developed to test for normality, lognormality, and a gamma distribution of a data set, and to compute a conservative and stable *95% UCL* of the population mean, μ_1. The critical values of Anderson-Darling test statistic and Kolmogorov-Smirnov test statistic to test for gamma distribution were generated using Monte Carlo simulation experiments. These critical values are tabulated in Appendix B for various levels of significance. Singh, Singh, and Engelhardt (1997,1999), Singh, Singh, and Iaci (2002b), and Singh and Singh (2003) studied several parametric and non-parametric *UCL* computation methods which have been included in ProUCL. Most of the mathematical algorithms and formulae used in ProUCL to compute the various statistics are summarized in this Appendix A. For details, the user is referred to Singh,

Singh, and Iaci (2002b), and Singh and Singh (2003). Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C. ProUCL computes the various summary statistics for raw, as well as log-transformed data. In this User Guide and in ProUCL, log-transform (*log*) stands for the natural logarithm (*ln*) to the base e. ProUCL also computes the maximum likelihood estimates (*MLEs)* and the minimum variance unbiased estimates (*MVUEs*) of various unknown population parameters of normal, lognormal, and gamma distributions. This, of course, depends upon the underlying data distribution. Based upon the data distribution, ProUCL computes the *(1-α) 100% UCLs* of the unknown population mean, μ_1 using five (5) parametric and ten (10) non-parametric methods.

The five parametric *UCL* computation methods include:

1) Student's- t *UCL*,
2) approximate gamma *UCL*,
3) adjusted gamma *UCL*,
4) Land's *H-UCL*, and
5) Chebyshev inequality based *UCL* (using *MVUE* of parameters of a lognormal distribution).

The ten non-parametric methods included in ProUCL are:

1) the central limit theorem (*CLT*) based *UCL*,
2) modified-t statistic (adjusted for skewness),
3) adjusted-*CLT* (adjusted for skewness),
4) Chebyshev inequality based *UCL* (using sample mean and sample standard deviation),
5) Jackknife *UCL*,
6) standard bootstrap,
7) percentile bootstrap,
8) bias - corrected accelerated (BCA) bootstrap,

9) bootstrap-t, and

10) Hall's bootstrap.

An extensive comparison of these methods have been performed by Singh and Singh (2003) using Monte Carlo simulation experiments. It is well known that the Jackknife method (with sample mean as an estimator) and Student's-t method yield identical *UCL* values. It is also well known that the standard bootstrap method and the percentile bootstrap method do not perform well (do not provide adequate coverage) for skewed data sets. However, for the sake of completeness all of the parametric as well as non-parametric methods have been included in ProUCL. Also, it has been noted that the omission of a method (e.g., bias-corrected accelerated bootstrap method) triggers the curiosity of some of the users as they start thinking that the omitted method may perform better than the various other methods already incorporated in ProUCL. In order to satisfy all users, ProUCL Version 3.0 has additional *UCL* computation methods which were not included in ProUCL Version 2.1.

1.1 Non-detects and Missing Data

ProUCL does not handle non-detects. All parametric as well as non-parametric recommendations to compute the mean, standard deviation, and a 95% *UCL* of the mean made by ProUCL software are based upon full data sets without censoring. The program can be modified to incorporate methods which can be used to compute appropriate estimates of the population mean and standard deviation, and a *UCL* of the mean for left-censored data sets with non-detects. For now, for data sets with mild to moderate number of non-detects (e.g., < 15%), one may replace non-detects by half of the detection limit (as often done in practice) and use ProUCL on the resulting data set to compute an appropriate *95% UCL* of the mean, μ_1. However, the proxy methods such as replacing non-detects by ½ of the detection limit (DL) should be used cautiously, especially when one is dealing with lognormally distributed data sets. For lognormally distributed data sets of small sizes, even a single value -- small (e.g., obtained after replacing the non-detect by ½ DL) or large (e.g., an outlier) can have a drastic influence (can yield an unrealistically large 95% *UCL*) on the value of the associated Land's 95% *UCL*. The issue of estimating the mean, standard deviation, and a 95% *UCL* of the mean based upon left-censored data sets of varying degrees of censoring (e.g., < 15%, 15% - 50%, 50% - 75%, and greater than 75%) is currently under investigation.

However, it should be noted that ProUCL can handle missing data. Missing data value can be entered as a very large value in scientific notation, such as 1.0 E 31. All entries with this value will be treated as missing data.

2. Procedures to Test for Data Distribution

Let $x_1, x_2, \ldots, x_n$ be a random sample (e.g., representing lead concentrations) from the underlying population (e.g, remediated part of a site) with unknown mean, μ_1, and variance, σ_1^2. Let μ and σ represent the population mean and the population standard deviation (*Sd*) of the log-transformed (natural log to the base e) data. Let $\bar{y}$ and s_y (= $\hat{\sigma}$) be the sample mean and sample *Sd*, respectively, of the log-transformed data, $y_i = \log(x_i)$; $i = 1, 2, \ldots, n$. Specifically, let

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \,, \tag{1}$$

$$\hat{\sigma} = s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \,. \tag{2}$$

Similarly, let $\bar{x}$ and s_x be the sample mean and *Sd* of the raw data, $x_1, x_2, \ldots, x_n$, obtained by replacing y by x in equations (1) and (2), respectively. In this User Guide, irrespective of the underlying distribution, μ_1, and σ_1^2 represent the mean and variance of the random variable X (in original units), whereas μ_1 and σ^2 represent the mean and variance of its logarithm, given by Y = $\log_e$(X) = natural logarithm.

Three data distributions have been considered. These include the normal and lognormal distributions, and the gamma distribution. Shapiro - Wilk ($n \leq 50$) and Lilliefors ($n > 50$) test statistics are used to test for normality or lognormality of a data set. The empirical distribution function (EDF) based methods: the Kolmogorov-Smirnov (K-S) test and the Anderson-Darling (A-D) test are used to test for a gamma distribution. Extensive critical values for these two test statistics have been obtained via Monte Carlo simulation experiments. For interested users, these critical values are given in Appendix B for various levels of significance. In addition to these formal tests, the informal histogram and quantile-quantile (Q-Q) plot are also available to test data distributions. A brief description of these tests follows.

2.1 Test Normality and Lognormality of a Data Set

ProUCL tests the normality or lognormality of the data set using the three different methods described below. The program tests normality or lognormality at three different levels of significance, namely, 0.01, 0.05, and 0.1. The details of these methods can be found in the cited references.

2.1.1 Normal Quantile-Quantile (Q-Q) Plot

This is a simple informal graphical method to test for an approximate normality or lognormality of a data distribution (Hoaglin, Mosteller, and Tukey (1983), Singh (1993)). A linear pattern displayed by the bulk of the data suggests approximate normality or lognormality (performed on log-transformed data) of the data distribution. For example, a high value (e.g., 0.95 or greater) of the correlation coefficient of the linear pattern may suggest approximate normality (or lognormality) of the data set under study. However, it should be noted that on this graphical display, observations well separated (sticking out) from the linear pattern displayed by the bulk data represent the outlying observations. Also, apparent jumps and breaks in the Q-Q plot suggest the presence of multiple populations. The correlation coefficient of such a Q-Q plot can still be high, which does not necessarily imply that the data follow a normal (or lognormal) distribution. Therefore, the informal graphical Q-Q plot test should always be accompanied by other more powerful tests, such as the Shapiro-Wilk test or the Lilliefors test. The goodness-of-fit test of a data set should be judged based upon the formal more powerful tests. The normal Q-Q plot may be used as an aid to identify outliers and/or to identify multiple populations. ProUCL performs the graphical Q-Q plot test on raw data as well as on standardized data. All relevant statistics such as the correlation coefficient are also displayed on the Q-Q plot.

2.1. 2 Shapiro-Wilk W Test

This is a powerful test and is often used to test the normality or lognormality of the data set under study (Gilbert, 1987). ProUCL performs this test for samples of size 50 or smaller. Based upon the selected level of significance and the computed test statistic, ProUCL also informs the user if the data are normally (or lognormally) distributed. This information should be used to obtain an appropriate *UCL* of the mean. The program prints the relevant statistics on the Q-Q plot of the data (or the standardized data). For convenience, the normality, lognormality, or gamma distribution test results at 0.05 level of significance are also displayed on the *UCL* Excel-type output summary sheets.

2.1.3 Lilliefors Test

This test is useful for data sets of larger size (Dudewicz and Misra, 1988). ProUCL performs this test for samples of sizes up to 1000. Based upon the selected level of significance and the computed test statistic, ProUCL informs the user if the data are normally (or lognormally) distributed. The user should use this information to obtain an appropriate *UCL* of the mean. The program prints the relevant statistics on the Q-Q plot of data (or standardized data). For convenience, the normality, lognormality, or gamma distribution test results at 0.05 level of significance are also displayed on the *UCL* output summary sheets. It should be pointed out that sometimes, in practice, these two goodness-of-fit tests can lead to different conclusions.

2.2 Gamma Distribution

Singh, Singh, and Iaci (2002b) studied gamma distribution to model positively skewed environmental data sets and to compute a *UCL* of the mean based upon a gamma distribution. They studied several *UCL* computation methods using Monte Carlo simulation experiments. A continuous random variable, X (e.g., concentration of a contaminant), is said to follow a gamma distribution, G (k,θ) with parameters k > 0 (shape parameter) and θ > 0 (scale parameter), if its probability density function is given by the following equation:

$$f(x;k,\theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-x/\theta} \quad ; \qquad x > 0 \tag{3}$$

and zero otherwise. The parameter *k* is the shape parameter, and θ is the scale parameter. Many positively skewed data sets follow a lognormal as well as a gamma distribution. Gamma distribution can be used to model positively skewed environmental data sets. It is observed that the use of a gamma distribution results in reliable and stable 95% *UCL* values. It is therefore, desirable to test if an environmental data set follows a gamma distribution. If a skewed data set does follow a gamma model, then a 95% *UCL* of the population mean should be computed using a gamma distribution. For details of the two gamma goodness-of-fit tests, maximum likelihood estimation of gamma parameters, and the computation of a 95% *UCL* of the mean based upon a gamma distribution, refer to D'Agostino and Stephens (1986), and Singh, Singh, and Iaci

(2002b). These methods are briefly described as follows.

For data sets which follow a gamma distribution, the adjusted 95% *UCL* of the mean based upon a gamma distribution is optimal and approximately provides the specified 95% coverage to population mean, $\mu_1 = k\theta$ (Singh, Singh, and Iaci (2002b)). Moreover, this adjusted gamma *UCL* yields reasonable numbers of practical merit. The two test statistics used for testing for a gamma distribution are based upon the empirical distribution function (EDF). The two EDF tests included in ProUCL are the Kolmogorov-Smirnov (K-S) test and Anderson - Darling (A-D) test which are described in D'Agostino and Stephens (1986) and Stephens (1970). The graphical Q-Q plot for gamma distribution has also been included in ProUCL. The critical values for the two EDF tests are not easily available, especially when the shape parameter, k is small (k < 1). Therefore, the associated critical values have been obtained via extensive Monte Carlo simulation experiments. These critical values for the two test statistics are given in Appendix B. The 1%, 5%, and 10% critical values of these two test statistics have been incorporated in ProUCL, Version 3.0. A brief description of the three goodness-of-fit tests for gamma distribution is given as follows. It should be noted that the goodness-of-fit tests for gamma distribution depend upon the *MLEs* of gamma parameters, k and θ which should be computed first before performing the goodness-of-fit tests.

2.2.1 Quantile - Quantile (Q-Q) Plot for a Gamma Distribution

Let $x_1, x_2, \ldots, x_n$ be a random sample from the gamma distribution, G(k,θ). Let $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ represent the ordered sample. Let $\hat{k}$ and $\hat{\theta}$ represent the maximum likelihood estimates (*MLEs*) of k and θ, respectively. For details of the computation of *MLEs* of k and θ, refer to Singh, Singh, and Iaci (2002b). Estimation of gamma parameters is also briefly described later in this User Guide. The Q-Q plot for gamma distribution is obtained by plotting the scatter plot of pairs $(x_{0i}, x_{(i)}); i := 1,2,\ldots,n.$ The quantiles, x_{0i} are given by the equation $x_{0i} = z_{0i}\hat{\theta}/2; i := 1,2,\ldots,n$, where the quantiles z_{0i} (already ordered) are obtained by using the inverse chi-square distribution and are given as follows.

$$\int_{0}^{z_{0i}} f(\mathrm{X}^2_{2\hat{k}})\,d\mathrm{X}^2_{2\hat{k}} = (i - 1/2)/n; i := 1,2,\ldots,n \qquad (4)$$

In (4), $\mathrm{X}^2_{2\hat{k}}$ represents a chi-square random variable with $2\hat{k}$ degrees of freedom (d.f.). The program, PPCHI2 (Algorithm AS91) as given in Best and Roberts (1975), Applied Statistics (1975, Vol. 24, No. 3) has been used to compute the inverse chi-square percentage points, z_{0i} as given by the above equation given by (4). This is an informal graphical test to test for a gamma distribution. This informal test should always be accompanied by the formal Anderson-Darling test or Kolmogorov- Smirnov test. A linear pattern displayed by the scatter plot of bulk of the data may suggest approximate gamma distribution. For example, a high value (e.g., 0.95 or greater) of the correlation coefficient of the linear pattern may suggest approximate gamma distribution of the data set under study. However, on this Q-Q plot points well separated from the bulk of data may represent outliers. Also, apparent breaks and jumps in the gamma Q-Q plot suggest the presence of multiple populations. The correlation coefficient of such a Q-Q plot can still be high which does not necessarily imply that the data follow a gamma distribution. Therefore, the graphical Q-Q plot test should always be accompanied by the other more powerful formal EDF tests, such as the Anderson-Darling test or the Kolmogorov-Smirnov test. The final conclusion about the data distribution should be based upon the formal goodness-of-fit tests. The Q-Q plot may be used to identify outliers and/or presence of multiple populations. All relevant statistics including the *MLE* of k are also displayed on the gamma Q-Q plot.

2.2.2 Empirical Distribution Function (EDF) Based Goodness-of -Fit Tests

Next, the two formal EDF test statistics used to test for a gamma distribution are described briefly. Let F(x) be the cumulative distribution function (CDF) of the gamma random variable

X. Let Z=F(X), then Z represents a uniform U(0,1) random variable. For each x_i, compute z_i using the incomplete gamma function given by the equation $z_i = F(x_i); i := 1,2,\ldots,n.$ The algorithm as given in Numerical Recipes book (Press et al., 1990) has been used to compute the incomplete gamma function. Arrange the resulting, z_i in ascending order as $z_{(1)} \le z_{(2)} \le \ldots \le z_{(n)}$. Let $\bar{z} = \sum z_i / n$ be the mean of the $z_i; i := 1,2,\ldots,n$. Compute the following two test statistics.

$$D^+ = \max_i \{1/n - z_{(i)}\} \text{ and } D^- = \max_i \{z_{(i)} - (i-1)/n\} \tag{5}$$

The Kolmogorov - Smirnov test statistic is given by $D = \max(D^+, D^-)$.

Anderson Darling test statistic is given by the following equation.

$$A^2 = -n - (1/n)\sum_{1}^{n} \{(2i-1)[\log z_{(i)} + \log(1 - z_{(n+1-i)})]\} \tag{6}$$

The critical values for these two statistics D and A^2 are not readily available. For the Anderson-Darling test, only asymptotic critical values are available in the statistical literature (D'Agostino and Stephens (1986)). Some raw critical values for K-S test are given in Schneider (1978), and Schneider and Clickner (1976). For these two tests, ExpertFit (2001) software and Law and Kelton (2000) use generic critical values for all completely specified distributions as given in D'Agostino and Stephens (1986). It is observed that the conclusions derived using these generic critical values for completely specified distributions and the simulated critical values for gamma distribution with unknown parameters can be different. Therefore, to test for a gamma distribution, it is preferred and advised to use the critical values of these test statistics specifically obtained for gamma distributions with unknown parameters.

In practice, the distributions are not completely specified and exact critical values for these two test statistics are needed. It should be noted that the distributions of the K-S test statistic, D and A-D test statistic, A^2 do not depend upon the scale parameter, θ , therefore, the scale parameter, θ has been set equal to 1 in all of the simulation experiments. The critical values for these two statistics have been obtained via extensive Monte Carlo simulation experiments for several small and large values of the shape parameter, k and with $\theta = 1$. These critical are included in Appendix B. In order to generate the critical values, random samples from gamma distributions were generated using the algorithm as given in Whittaker (1974). It is observed that the critical values thus obtained are in close agreement with all available published critical values. The generated critical values for the two test statistics have been incorporated in ProUCL for three levels of significance, 0.1, 0.05, and 0.01. For each of the two tests, if the test statistic exceeds the corresponding critical value, then the hypothesis that the data follow a gamma distribution is rejected. ProUCL computes these test statistics and prints them on the gamma Q-Q plot and also on the *UCL* summary output sheets generated by ProUCL. The estimation of the parameters of the three distributions as incorporated in ProUCL is discussed next. It should be pointed out that sometimes, in practice, these two goodness-of-fit tests can lead to different conclusions.

3. Estimation of Parameters of the Three Distributions Included in ProUCL

Through out this User Guide, μ_1 and σ_1^2 are the mean and variance of the random variable X, and μ and σ^2 are the mean and variance of the random variable Y = log(X). Also, $\hat{\sigma}$ represents the standard deviation of the log-transformed data. It should be noted that for both lognormal and gamma distributions, the associated random variable can take only positive values. This is typical of environmental data sets to consist of only positive values.

3.1 Normal Distribution

Let X be a continuous random variable (e.g., concentration of COPC), which follows a normal distribution, N(μ_1, σ_1^2) with mean, μ_1, and variance, σ_1^2. The probability density function of a normal distribution is given by the following equation:

$$f(x;\mu,\sigma_1) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp(-(x-\mu_1)^2/2\sigma_1^2);-\infty < x < \infty \quad (7)$$

For normally distributed data sets, it is well known (Hogg and Craig, 1978) that the minimum variance unbiased estimates (*MVUEs*) of mean, μ_1, and variance, σ_1^2 are respectively given by the sample mean, $\bar{x}$ and sample variance, s_x^2. It is also well known that for normally distributed data sets, a *UCL* of the unknown mean, μ_1 based upon Student's-t distribution is optimal. It is observed via Monte Carlo simulation experiments (Singh and Singh (2003) Draft EPA Report) that for normally distributed data sets, the modified-t *UCL* and *UCL* based upon bootstrap-t method also provide the exact 95% coverage to the population mean. For normally distributed data sets, the *UCLs* based upon these three methods are very similar.

3.2 Lognormal Distribution

If $Y = \log(X)$ is normally distributed with the mean μ and variance σ^2, X is said to be lognormally distributed with parameters μ and σ^2 and is denoted by LN(μ, σ^2) . It should be noted that μ and σ^2 are not the mean and variance of the lognormal random variable, *X*, but they are the mean and variance of the log-transformed random variable *Y*, whereas μ_1, and $\sigma_1{}^2$ represent the mean and variance of X. Some parameters of interest of a two-parameter lognormal distribution, LN(μ, σ^2), are given as follows:

Mean $= \mu_1 = \exp(\mu + 0.5\sigma^2)$ (8)

Median $= M = \exp(\mu)$ (9)

Variance $= \sigma_1^2 = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$ (10)

Coefficient of Variation $= CV = \sigma_1/\mu_1 = \sqrt{(\exp(\sigma^2) - 1)}$ (11)

Skewness $= (CV)^3 + 3(CV)$ (12)

3.2.1 *MLEs* of the Parameters of a Lognormal Distribution

For lognormal distributions, note that $\bar{y}$ and s_y (= $\hat{\sigma}$) are the maximum likelihood estimators (*MLEs*) of μ and σ, respectively. The *MLE* of any function of the parameters μ and σ^2 is obtained by simply substituting these *MLEs* in place of the parameters (Hogg and Craig, 1978). Therefore, replacing μ and σ by their *MLEs* in equations (8) through (12) will result in the *MLEs* (but biased) of the respective parameters of the lognormal distribution. The program ProUCL computes all of these *MLEs* for lognormally distributed data sets. These *MLEs* are also printed on the Excel-type output spread sheets generated by ProUCL.

3.2.2 Relationship Between Skewness and Standard Deviation, σ

Note that for a lognormal distribution, the *CV* (given by equation (11) above) and the skewness (given by equation (12)) depend only on σ. Therefore, in this User Guide and also in ProUCL, the standard deviation, σ (*Sd* of log-transformed variable, Y), or its *MLE*, s_y (=$\hat{\sigma}$) has been used as a measure of skewness of lognormal and also of other skewed data sets with positive values. The larger is the Sd, the larger are the *CV* and the skewness. For example, for a lognormal distribution: with σ = *0.5,* the skewness = 1.75; with σ =1.0, the skewness = 6.185; with σ =*1.5,* the skewness = 33.468; and with σ = 2.0, the skewness = 414.36. Thus, the skewness of a lognormal distribution becomes unreasonably large as σ starts approaching and exceeding 2.0. Note that for gamma distribution, skewness is a function of the gamma parameter, k. As k decreases, skewness increases.

It is observed (Singh, Singh, Engelhardt (1997), and Singh et al. (2002a)) that for smaller sample sizes (such as smaller than 50), and for values of σ approaching 2.0 (and skewness approaching 414), the use of the H-statistic based *UCL* results in impractical and unacceptably large values. For simplicity, the various levels of skewness of a positive data set as used in ProUCL and in this User Guide are summarized as follows:

Skewness as a Function of σ ***(or its MLE,*** $s_y = \hat{\sigma}$**), Sd of log(X)**

Standard Deviation	*Skewness*
$\sigma < 0.5$	*Symmetric to mild skewness*
$0.5 \le \sigma < 1.0$	*Mild Skewness to Moderate Skewness*
$1.0 \le \sigma < 1.5$	*Moderate Skewness to High Skewness*
$1.5 \le \sigma < 2.0$	*High skewness*
$2.0 \le \sigma < 3.0$	*Extremely high skewness*
$\sigma \ge 3.0$	*Provides poor coverage*

These values of σ (or its estimate, *Sd* of log-transformed data) are used to define skewness levels of lognormal and skewed non-parametric data distributions as used in Tables A2 and A3.

3.2.3 MLEs of the Quantiles of a Lognormal Distribution

For highly skewed (e.g., σ exceeding 1.5), lognormally distributed populations, the population mean, μ_1, often exceeds the higher quantiles (e.g., 80%, 90%, 95%) of the distribution. Therefore, the computation of these quantiles is also of interest. This is especially true when one may want to use the *MLEs* of the higher order quantiles (e.g., 95%, 97.5% etc.) as an estimate of the EPC term. The formulae to compute these quantiles are briefly described here.

The pth quantile (or 100 pth percentile), x_p, of the distribution of a random variable, X, is defined by the probability statement, $P(X \le x_p) = p$. If z_p is the pth quantile of the standard normal random variable, Z, with $P(Z \le z_p) = p$, then the pth quantile of a lognormal distribution is given by $x_p = \exp(\mu + z_p\sigma)$. Thus the *MLE* of the pth quantile is given by

$$\hat{x}_p = \exp(\hat{\mu} + z_p\hat{\sigma}) \tag{13}$$

For example, on the average, 95% of the observations from a lognormal LN(μ, σ^2) distribution would lie below $\exp(\mu + 1.65\sigma)$. The 0.5th quantile of the standard normal distribution is $z_{0.5} =$ 0, and the 0.5th quantile (or median) of a lognormal distribution is M = exp(μ), which is obviously smaller than the mean, μ_1, as given by equation (8). Also note that the mean, μ_1, is greater than x_p if and only if $\sigma > 2z_p$. For example, when $p = 0.80$, $z_p = 0.845$, μ_1 exceeds $x_{0.80}$, the 80th percentile if and only if $\sigma > 1.69$, and, similarly, the mean, μ_1, will exceed the 95th percentile if and only if $\sigma > 3.29$. ProUCL computes the *MLEs* of the 50% (median), 90%, 95%, and 99% percentiles of lognormally distributed data sets. For lognormally distributed background data sets, a 95% or 99% percentile may be used as an estimate of the background threshold value, that is background level contaminant concentration.

3.2.4 MVUEs of Parameters of a Lognormal Distribution

Even though the sample *AM*, $\bar{x}$, is an unbiased estimator of the population *AM*, μ_1, it does not have the minimum variance (*MV*). The *MV unbiased estimates (MVUEs)* of μ_1 and σ_1^2 of a lognormal distribution are given as follows:

$$\hat{\mu}_1 = \exp(\bar{y})g_n(s_y^2/2), \tag{14}$$

$$\hat{\sigma}_1^2 = \exp(2\bar{y})[g_n(2s_y^2) - g_n((n-2)s_y^2/(n-1))] \tag{15}$$

where the series expansion of the function $g_n(\mu)$ is given in Bradu and Mundlak (1970), and Aitchison and Brown (1976). Tabulations of this function are also provided by Gilbert (1987). Bradu and Mundlak (1970) give the *MVUE* of the variance of the estimate $\hat{\mu}_1$,

$$\hat{\sigma}^2(\hat{\mu}_1) = \exp(2\bar{y})[(g_n(s_y^2/2))^2 - g_n((n-2)s_y^2/(n-1))] \tag{16}$$

The square root of the variance given by equation (16) is called the standard error (*SE*) of the estimate, $\hat{\mu}_1$, given by equation (14). Similarly, a *MVUE* of the median of a lognormal distribution is given by

$$\hat{M} = \exp(\bar{y})g_n(-s_y^2/(2(n-1))). \tag{17}$$

For lognormally distributed data set, ProUCL also computes these *MVUEs* given by equations (14) through (17).

3.3 Estimation of the Parameters of a Gamma Distribution

Next, we consider the estimation of parameters of a gamma distribution. Since the estimation of gamma parameters is typically not included in standard statistical text books, this has been described in some detail in this User Guide. The population mean and variance of a gamma distribution, G(k,θ), are functions of both parameters, k and θ. In order to estimate the mean, one has to obtain estimates of k and θ. The computation of the maximum likelihood estimate (*MLE*) of k is quite complex and requires the computation of Digamma and Trigamma functions. Several authors (Choi and Wette, 1969, Bowman and Shenton, 1988, Johnson, Kotz, and Balakrishnan, 1994) have studied the estimation of shape and scale parameters of a gamma distribution. The maximum likelihood estimation method to estimate shape and scale parameters of a gamma distribution is described below.

Let $x_1, x_2, \ldots, x_n$ be a random sample (e.g., representing contaminant concentrations) of size n from a gamma distribution, G(k,θ), with unknown shape and scale parameters *k* and θ, respectively. The log likelihood function (obtained using equation (3)) is given as follows:

$$\log L(x_1, x_2, \ldots, x_n; k, \theta) = -nk\log(\theta) - n\log\Gamma(k) + (k-1)\sum \log x_i - \frac{1}{\theta}\sum x_i \qquad (18)$$

To find the MLEs of *k* and θ, we differentiate the log likelihood function as given in (18) with respect to k and θ, and set the derivatives to zero. This results in the following two equations:

$$\log(\hat{\theta}) + \frac{\Gamma'(\hat{k})}{\Gamma(\hat{k})} = \frac{1}{n}\sum \log(x_i) \text{ , and} \qquad (19)$$

$$\hat{k}\hat{\theta} = \frac{1}{n}\sum x_i = \bar{x} \qquad (20)$$

Solving equation (20) for $\hat{\theta}$ and substituting the result in equation (19), we get the following equation:

$$\frac{\Gamma'(\hat{k})}{\Gamma(\hat{k})} - \log(\hat{k}) = \frac{1}{n}\sum \log(x_i) - \log\left(\frac{1}{n}\sum x_i\right) \qquad (21)$$

There does not exist a closed form solution of equation (21). This equation needs to be solved numerically for $\hat{k}$, which requires the use of Digamma and Trigamma functions. This is quite easy to do using a personal computer. An estimate of *k* can be computed iteratively by using the Newton-Raphson (Faires and Burden, 1993) method leading to the following iterative equation:

$$\hat{k}_l = \hat{k}_{l-1} - \frac{\log(\hat{k}_{l-1}) - \Psi(\hat{k}_{l-1}) - M}{1/\hat{k}_{l-1} - \Psi'(\hat{k}_{l-1})} \tag{22}$$

The iterative process stops when $\hat{k}$ starts to converge. In practice, convergence is typically achieved in fewer than 10 iterations. In equation (22)

$M = \log(\bar{x}) - \frac{1}{n}\sum \log(x_i)$, and

$\Psi(k) = \frac{d}{dk}\left(\log\Gamma(k)\right)$, and $\Psi'(k) = \frac{d}{dk}\left(\Psi(k)\right)$

where $\Psi(k)$ is the Digamma function, and $\Psi'(k)$ is the Trigamma function. In order to obtain the *MLEs* of k and θ, one needs to compute the Digamma and Trigamma functions. Good approximate values for these two functions (Choi and Wette, 1969) can be obtained using the following approximations. For $k \geq 8$, these functions are approximated by

$$\Psi(k) \approx \log(k) - \left\{1 + \left[1 - (1/10 - 1/(21k^2))/k^2\right]/(6k)\right\}/(2k) \tag{23}$$

and

$$\Psi'(k) \approx \left\{1 + \left\{1 + \left[1 - (1/5 - 1/(7k^2))/k^2\right]/(3k)\right\}/(2k)\right\}/k \tag{24}$$

For $k < 8$, one can use the following recurrence relation to compute these functions:

$$\Psi(k) = \Psi(k+1) - 1/k, \tag{25}$$

and $\Psi'(k) = \Psi'(k+1) + 1/k^2$ (26)

In ProUCL, equations (23) - (26) have been used to estimate k. The iterative process requires an initial estimate of *k*. A good starting value for *k* in this iterative process is given by $k_0 = 1/(2M)$. Thom (1968) suggested the following approximation as an estimate of *k*:

$$\hat{k} \approx \frac{1}{4M}\left(1 + \sqrt{1 + \frac{4}{3}M}\right) \qquad (27)$$

Bowman and Shenton (1988) suggested using $\hat{k}$ as given by (27) to be a starting value of *k* for an iterative procedure, calculating $\hat{k}_l$ at the I th iteration from the following formula:

$$\hat{k}_l = \frac{\hat{k}_{l-1}\left\{\log(\hat{k}_{l-1}) - \Psi(\hat{k}_{l-1})\right\}}{M} \qquad (28)$$

Both equations (22) and (28) have been used to compute the *MLE* of *k*. It is observed that the estimate, $\hat{k}$ based upon Newton-Raphson method as given by equation (22) is in close agreement with that obtained using equation (28) with Thom's approximation as an initial estimate. Choi and Wette (1969) further concluded that the *MLE* of k, $\hat{k}$, is biased high. A bias-corrected (Johnson, Kotz, and Balakrishnan, 1994) estimate of *k* is given by:

$$\hat{k}^* = (n-3)\hat{k}/n + 2/(3n) \qquad (29)$$

In (29), $\hat{k}$ is the *MLE* of *k* obtained using either (22) or (28). Substitution of equation (29) in equation (20) yields an estimate of the scale parameter, θ given as follows:

$$\hat{\theta}^* = \bar{x}/\hat{k}^* \qquad (30)$$

ProUCL computes simple *MLE* of k and θ, and also bias- corrected estimates of k and θ. The bias-corrected estimate of k as given by (29) has been used in the computation of the *UCLs* (as given by equations (34) and (35)) of the mean of a gamma distribution.

4. Methods for Computing a *UCL* of the Unknown Population Mean

ProUCL computes a *(1-α) 100 % UCL* of the population mean, μ_1 using the following five parametric and ten non-parametric methods. Five of the ten non-parametric methods are based upon the bootstrap method. Modified-t and adjusted central limit theorem adjust for skewness for skewed data sets. However, it is noted that (Singh, Singh, and Iaci (2002b) and Singh and Singh (2003)) this adjustment is not adequate enough for moderately skewed to highly skewed data sets. Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C. The methods as included in ProUCL are listed as follows.

Parametric Methods

1. Student's-t statistic - assumes normality or approximate normality
2. Approximate Gamma *UCL* - assumes gamma distribution of the data set
3. Adjusted Gamma *UCL* - assumes gamma distribution of the data set
4. Land's H-Statistic - assumes lognormality
5. Chebyshev Theorem using the *MVUE* of the parameters of a lognormal distribution (denoted by Chebyshev (*MVUE*)) - assumes lognormality

Non-parametric Methods

1. Modified- t statistic - modified for skewed distributions
2. Central Limit Theorem (*CLT*) - to be used for large samples
3. Adjusted Central Limit Theorem (Adjusted-*CLT*) - adjusted for skewed distributions and to be used for large samples

4. Chebyshev Theorem using the sample arithmetic mean and Sd (denoted by Chebyshev (Mean, *Sd*))
5. Jackknife method - yields the same result as Student's-t statistic for the *UCL* of the population mean
6. Standard bootstrap
7. Percentile bootstrap
8. Bias-corrected accelerated (BCA) bootstrap
9. Bootstrap-t
10. Hall's bootstrap

Even though it is well known that some of the non-parametric methods (e.g., *CLT* method, *UCL* based upon Jackknife method (same as Student's-t *UCL*), standard bootstrap and percentile bootstrap methods) do not perform well to provide the adequate coverage to the population mean of skewed distributions, these methods have been included in ProUCL to satisfy the curiosity of all users.

ProUCL can compute a *(1-α) 100 % UCL* (except for the *H-UCL and adjusted gamma UCL*) of the mean for any confidence coefficient *(1-α)* value lying in the interval [0.5, 1.0). For the computation of the *H-UCL,* only two confidence levels, namely, 0.90 and 0.95 are supported by ProUCL. For adjusted gamma *UCL*, three confidence levels namely, 0.90, 0.95, and 0.99 are supported by ProUCL. An approximate gamma *UCL* can be computed for any level of significance in the interval [0.5,1). Based upon the sample size, n, skewness, and the data distribution, the program also makes recommendations on how to obtain an appropriate *95% UCL* of the unknown population mean, μ_1 . These recommendations are summarized in the Recommendations and Summary Section 5 of this appendix. The various algorithms and methods used to compute a *(1-α) 100% UCL* of the mean as incorporated in ProUCL are described in section 4.1.

4.1 (1-α) 100% UCL of the Mean Based Upon Student's-t Statistic

The widely used well-known Student's-t statistic is given by,

$$t = \frac{\bar{x} - \mu_1}{s_x/\sqrt{n}}, \tag{31}$$

where $\bar{x}$ and s_x are, respectively, the sample mean and sample standard deviation obtained using the raw data. If the data are a random sample from a normal population with mean, μ_1, and standard deviation, σ_1, then the distribution of this statistic is the familiar Student's-t distribution with $(n-1)$ degrees of freedom (df). Let $t_{\alpha,\, n-1}$ be the upper α^{th} quantile of the Student's-t distribution with $(n-1)$ df.

A *(1−α)100 % UCL* of the population mean, μ_1, is given by,

$$UCL = \bar{x} + t_{\alpha, n-1} s_x/\sqrt{n}. \tag{32}$$

For a normally (when the skewness is about ~0) distributed population, equation (32) provides the best (optimal) way of computing a *UCL* of the mean. Equation (32) may also be used to compute a *UCL* of the mean based upon very mildly skewed (e.g., |skewness|<0.5) data sets, where skewness is given by equation (43). It should be pointed out that even for mildly to moderately skewed data sets (e.g., when *σ, Sd of log-transformed data* starts approaching and exceeding 0.5), the *UCL* given by (32) may not provide the desired coverage (e.g., =0.95) to the population mean. This is especially true when the sample size is smaller than 20-25 (Singh et al. (2002a), and Singh and Singh (2003)). The situation gets worse (coverage much smaller than 0.95) for higher values of the *Sd*, *σ*, or its *MLE*, s_y.

4.2 Computation of *UCL* of the Mean of a Gamma, G(k,θ) Distribution

In statistical literature, even though methods exist to compute a *UCL* of the mean of a gamma distribution (Grice and Bain, 1980, Wong, 1993), those methods have not become popular due to their computational complexity. Those approximate and adjusted methods depend upon the Chi-square distribution and an estimate of the shape parameter, k. As seen above, computation of an *MLE* of *k* is quite involved, and this works as a deterrent to the use of a gamma distribution-based *UCL* of the mean. However, the computation of a gamma *UCL* currently should not be a problem due to easy availability of personal computers.

Given a random sample, $x_1, x_2, \ldots, x_n$ of size n from a gamma, G(k,θ) distribution, it can be shown that $2n\overline{X} / \theta$ follows a Chi-square distribution, χ^2_{2nk}, with 2nk degrees of freedom (df). When the shape parameter, *k,* is known, a uniformly most powerful test of size α of the null hypothesis, $H_0{:}\mu_1 \geq C_s$, against the alternative hypothesis, $H_1{:}\ \mu_1 < C_s$, is to reject H_0 if $\overline{X} / C_s < \chi^2_{2nk}(\alpha) / 2nk$. The corresponding (1-α)100% uniformly most accurate *UCL* for the mean, μ_1, is then given by the probability statement.

$$P(2nk\overline{x} / \chi^2_{2nk}(\alpha) \geq \mu_1) = 1 - \alpha \tag{33}$$

where $\chi^2_{\upsilon}(\alpha)$ denotes the α cumulative percentage point of the Chi-square distribution (e.g., α is the area in the left tail). That is, if Y follows χ^2_{υ}, then $P(Y \leq \chi^2_{\upsilon}(\alpha)) = \alpha$. In practice, *k* is not known and needs to be estimated from data. A reasonable method is to replace k by its bias -corrected estimate, $\hat{k}^*$, as given by equation (29). This results in the following approximate *(1-α) 100% UCL* of the mean, μ_1 .

$$Approximate\text{-}UCL = 2n\hat{k}^*\overline{x} / \chi^2_{2n\hat{k}^*}(\alpha) \tag{34}$$

It should be pointed out that the *UCL* given by equation (34) is an approximate *UCL* and there is no guarantee that the confidence level of (1-α) will be achieved by this *UCL*. However, it does provide a way of computing a *UCL* of the mean of a gamma distribution. Simulation studies conducted in Singh, Singh, and Iaci (2002b) and in Singh and Singh (2003) suggest that an approximate gamma *UCL* thus obtained provides the specified coverage (95%) as the shape parameter, k approaches 0.5. Thus when k ≥ 0.5, one can always use the approximate *UCL* given by (34). This approximation is good even for smaller (e.g., n = 5) sample sizes as shown in Singh, Singh, and Iaci (2002b), and in Singh and Singh (2003).

Grice and Bain (1980) computed an adjusted probability level, β (adjusted level of significance), which can be used in (34) to achieve the specified confidence level of (1-α). For α = 0.05 (confidence coefficient of 0.95), α = 0.1, and α = 0.01, these probability levels are given below in Table 1 for some values of the sample size n. One can use interpolation to obtain an adjusted β for values of n not covered in the table. The adjusted *(1-α) 100% UCL* of the gamma mean, μ_1 = kθ is given by the following equation.

$$Adjusted\text{-}UCL = 2n\hat{k}^{*}\overline{x} \,/\, \chi^{2}_{2n\hat{k}^{*}}(\beta), \tag{35}$$

where β is given in Table 1 for α = 0.05, 0.1, and 0.01. Note that as the sample size, n, becomes large, the adjusted probability level, β, approaches the specified level of significance, α. Except for the computation of the *MLE* of *k*, equations (34) and (35) provide simple Chi-square-distribution-based *UCLs* of the mean of a gamma distribution. It should also be noted that the *UCLs* as given by (34) and (35) only depend upon the estimate of the shape parameter, *k,* and are independent of the scale parameter, θ, and its ML estimate. Consequently, as expected, it is observed that coverage probabilities for the mean associated with these *UCLs* do not depend upon the values of the scale parameter, θ. It should also be noted that gamma *UCLs* do not depend upon the standard deviation of data which gets distorted by the presence of outliers.

Thus, outliers will have reduced influence on the computation of the gamma distribution based *UCLs* of the mean, μ_1.

Table 1. Adjusted Level of Significance, β

n	α = 0.05 probability level, β	α = 0.1 probability level, β	α = 0.01 probability level, β
5	0.0086	0.0432	0.0000
10	0.0267	0.0724	0.0015
20	0.0380	0.0866	0.0046
40	0.0440	0.0934	0.0070
--	0.0500	0.1000	0.0100

4.3 (1-α) 100% UCL of the Mean Based Upon H-Statistic *(H-UCL)*

The one-sided *(1−α)100% UCL* for the mean, μ_1, of a lognormal distribution as derived by Land (1971, 1975) is given as follows:

$$UCL = \exp\left(\bar{y} + 0.5 s_y^2 + s_y H_{1-\alpha} / \sqrt{(n-1)}\right) \qquad (36)$$

Tables of H-statistic critical values can be found in Land (1975) and also in Gilbert (1987). Theoretically, when the population is lognormal, Land (1971) showed that the *UCL* given by equation (36) possesses optimal properties and is the uniformly most accurate unbiased confidence limit. However, it is noticed that in practice, the H-statistic based results can be quite disappointing and misleading especially when the data set consists of outliers, or is a mixture from two or more distributions (Singh, Singh, and Engelhardt, 1997, 1999), Singh, Singh, and Iaci (2002b)). Even a minor increase in the Sd, s_y, drastically inflates the *MVUE* of

μ_l and the associated *H-UCL*. The presence of low as well as high data values increases the *Sd*, s_y, which in turn inflates the *H-UCL*. Furthermore, it is observed (Singh, Singh, Engelhardt, and Nocerino (2002a)) that for samples of sizes smaller than 15-25, and for values of σ approaching 1.0 and higher (for moderately skewed to highly skewed data sets), the use of H-statistic based *UCL* results in impractical and unacceptably large *UCL* values.

In practice many data sets follow a lognormal as well as gamma model. However, the population mean based upon a lognormal model can be significantly greater (often unrealistically large) than the population mean based upon a gamma model. In order to provide the specified 95% coverage for an inflated mean based upon a lognormal model, the resulting *UCL* based upon H-statistic also yield impractical *UCL* values. Use of a gamma model results in practical estimates (e.g., *UCL*) of the population mean. Therefore, for positively skewed data sets, it is recommended to test for a gamma model first. If data follow a gamma distribution, then the *UCL* of the mean should be computed using a gamma distribution. The gamma distribution is better suited to model positively skewed environmental data sets.

4.4 (1-α) 100% *UCL* of the Mean Based Upon Modified-t Statistic for Asymmetrical Populations

Chen (1995), Johnson (1978), Kleijnen, Kloppenburg, and Meeuwsen (1986), and Sutton (1993) suggested the use of the modified-t statistic for testing the mean of a positively skewed distribution (including the lognormal distribution). The $(1-\alpha)100\%$ *UCL* of the mean thus obtained is given by

$$UCL = \bar{x} + \hat{\mu}_3/(6s_x^2 n) + t_{\alpha,n-1} s_x/\sqrt{} \qquad (37)$$

where $\hat{\mu}_3$, an unbiased moment estimate (Kleijnen, Kloppenburg, and Meeuwsen, 1986) of the third central moment, is given as follows,

$$\hat{\mu}_3 = n \sum_{i=1}^{n} (x_i - \bar{x})^3 / [(n-1)(n-2)]. \qquad (38)$$

It should be pointed out that this modification for a skewed distribution does not perform well even for mildly to moderately skewed data sets (e.g., when σ starts approaching and exceeding 0.75). Specifically, it is observed that the *UCL* given by equation (37) may not provide the desired coverage of the population mean, μ_1, when σ starts approaching and exceeding 0.75 (Singh, Singh, and Iaci (2002b)). This is especially true when the sample size is smaller than 20-25. This small sample size requirement increases as σ increases. For example, when σ starts approaching and exceeding 1.5, the *UCL* given by equation (37) does not provide the specified coverage (e.g., 95%), even for samples as large as 100. Since this method does not require any distributional assumptions, it is a non-parametric method.

4.5 (1-α) 100% *UCL* of the Mean Based Upon the Central Limit Theorem

The Central Limit Theorem *(CLT)* states that the asymptotic distribution, as n approaches infinity, of the sample mean, $\bar{x}_n$ is normally distributed with mean, μ_1, and variance, σ_1^2/n. More precisely, the sequence of random variables given by

$$z_n = \frac{\bar{x}_n - \mu_1}{\sigma_1/\sqrt{n}} \qquad (39)$$

has a standard normal limiting distribution. In practice, for large sample sizes, n, the sample mean, $\bar{x}$, has an approximate normal distribution irrespective of the underlying distribution function. Since the *CLT* method requires no distributional assumptions, this is a non-parametric method.

As noted by Hogg and Craig (1978), if σ_1 is replaced by the sample standard deviation, s_x, the normal approximation for large n is still valid. This leads to the following approximate large sample non-parametric *(1-α) 100% UCL* of the mean,

$$UCL = \bar{x} + z_\alpha s_x / \sqrt{n}. \tag{40}$$

An often cited rule of thumb for a sample size associated with the *CLT* method is $n \geq 30$. However, this may not be adequate enough if the population is skewed, specifically when, σ *(Sd of log-transformed* variable) starts exceeding 0.5 (Singh, Singh, Iaci 2002b). In practice for skewed data sets, even a sample as large as 100 is not large enough to provide adequate coverage to the mean of skewed populations (even for mildly skewed populations). A refinement of the *CLT* approach, which makes an adjustment for skewness as discussed by Chen (1995), is given as follows.

4.6 (1-α) 100% UCL of the Mean Based Upon the Adjusted Central Limit Theorem (Adjusted -*CLT*)

The "*adjusted-CLT" UCL* is obtained if the standard normal quantile, z_α in the upper limit of equation (40) is replaced by (Chen, 1995)

$$z_{\alpha, adj} = z_\alpha + \frac{\hat{\kappa}_3}{6\sqrt{n}}\left(1 + 2z_\alpha^2\right). \tag{41}$$

Thus, the adjusted $(1-\alpha)$ 100 % *UCL* for the mean, μ_1, is given by

$$UCL = \bar{x} + [z_\alpha + \hat{k}_3(1 + 2z_\alpha^2)/(6\sqrt{n})]s_x/\sqrt{n}. \tag{42}$$

Here $\hat{k}_3$, the coefficient of skewness (raw data) is given by

Skewness (raw data) $\hat{k}_3 = \hat{\mu}_3 / s_x^3$ (43)

where $\hat{\mu}_3$, an unbiased estimate of the third moment, is given by equation (38). This is another large sample approximation for the *UCL* of the mean of skewed distributions. This is a non-parametric method as it does not depend upon any of the distributional assumptions.

As with the modified-t *UCL*, it is observed that this adjusted-*CLT UCL* does not provide adequate coverage to the population mean when the population is skewed, specifically when σ starts approaching and exceeding 0.75 (Singh, Singh, and Iaci (2002b), Singh and Singh (2003)). This is especially true when the sample size is smaller than 20-25. This small sample size requirement increases as σ increases. For example, when σ starts approaching and exceeding 1.5, the *UCL* given by equation (42) does not provide the specified coverage (e.g., 95%), even for samples as large as 100. Also, it is noted that the *UCL* as given by (42) does not provide adequate coverage to the mean of a gamma distribution, especially when k $\leq$ 1.0 and sample size is small. Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C.

Thus, the *UCLs* based upon these skewness adjusted methods, such as the Johnson's modified-t and Chen's adjusted-*CLT* do not provide the specified coverage to the population mean for mildly to moderately skewed (e.g., σ in (0.5, 1.0)) data sets, even for samples as large as 100 (Singh, Singh, and Iaci (2002b)). The coverage of the population mean provided by these *UCLs* becomes worse (much smaller than the specified coverage) for highly skewed data sets.

4.7 (1-α) 100% *UCL* of the Mean Based Upon the Chebyshev Theorem (Using the Sample Mean and Sample Sd)

The Chebyshev inequality can be used to obtain a reasonably conservative but stable estimate of the *UCL* of the mean, μ_1. The two-sided Chebyshev theorem (Hogg and Craig, 1978) states that given a random variable, *X*, with finite mean and standard deviation, μ_1 and σ_1, we have

$$P\left(-k\sigma_1 \le X - \mu_1 \le k\sigma_1\right) \ge 1 - 1/k^2. \tag{44}$$

This result can be applied on the sample mean, $\bar{x}$ (with mean, μ_1 and variance, σ_1^2 / n) to obtain a conservative *UCL* for the population mean, μ_1. For example, if the right side of equation (44) is equated to 0.95, then $k = 4.47$, and $UCL = \bar{x} + 4.47\sigma_1/\sqrt{n}$ is a conservative 95% upper confidence limit for the population mean, μ_1. Of course, this would require the user to know the value of σ_1. The obvious modification would be to replace σ_1 with the sample standard deviation, s_x, but since this is estimated from data, the result is no longer guaranteed to be conservative. In general, the following equation can be used to obtain a *(1-α) 100% UCL* of the population mean, μ_1:

$$UCL = \bar{x} + \sqrt{(1/\alpha)}s_x/\sqrt{n} \tag{45}$$

A slight refinement of equation (45) is given (suggested by S. Ferson) as follows,

$$UCL = \bar{x} + \sqrt{((1/\alpha) - 1)}s_x / \sqrt{n} \tag{46}$$

ProUCL computes the Chebyshev *(1-α) 100% UCL* of the population mean using equation (46). This *UCL* is denoted by *Chebyshev (Mean, Sd)* on the output sheets generated by ProUCL. Since this Chebyshev method requires no distributional assumptions about the data set under study, this is a non-parametric method. This *UCL* may be used as an estimate of the upper confidence limit of the population mean, μ_1 when data are not normal, lognormal, or gamma distributed especially when Sd, σ (or its estimate, s_y) starts approaching and exceeding

1.5. Recommendations on its use to a compute an estimate of the EPC term are summarized in Section 5.

4.8 (1-α) 100% *UCL* of the Mean of a Lognormal Population Based Upon the Chebyshev Theorem (Using the *MVUE* of the Mean and its Standard Error)

ProUCL uses equation (44) on the *MVUEs* of the lognormal mean and *Sd* to compute a *UCL* (denoted by *(1-α)100 % Chebyshev* (*MVUE)*) of the population mean of a lognormal population. In general, if μ_1 is an unknown mean, $\hat{\mu}_1$ is an estimate, and $\hat{\sigma}(\hat{\mu}_1)$ is an estimate of the standard error of $\hat{\mu}_1$, then the following equation,

$$UCL = \hat{\mu}_1 + ((1/\alpha) - 1)^{1/2} \ \hat{\sigma}(\hat{\mu}_1) \qquad (47)$$

will give an approximate *(1-α) 100 % UCL* for μ_1, which should tend to be conservative, but this is not assured. For example, for a lognormally distributed data set, a 95% (with α =0.05) Chebyshev *(MVUE) UCL* of the mean can be obtained using the following equation,

$$UCL = \hat{\mu}_1 + (4.359) \ \hat{\sigma}(\hat{\mu}_1) \qquad (48)$$

where, $\hat{\mu}_1$ and $\hat{\sigma}(\hat{\mu}_1)$ are given by equations (14) and (16), respectively. Thus, for lognormally distributed data sets, ProUCL also uses equation (48) to compute a *(1-α) 100% Chebyshev (MVUE) UCL* of the mean. It should be noted that for lognormally distributed data sets, some recommendations to compute a 95% *UCL* of the population mean are summarized in Table A2 of the Recommendations and Summary Section 5.0. It should however be pointed out that goodness-of-fit test for a gamma distribution should be performed first. If data follow a gamma distribution (irrespective of the lognormality of the data set), then the *UCL* of mean, μ_1 should be computed using a gamma distribution as described in Section 4.2.

From Monte-Carlo results discussed in Singh, Singh, and Iaci (2002b) and in Singh and Singh (2003), it is observed that for highly skewed gamma distributed data sets (with $k < 0.5$), the coverage provided by the Chebyshev 95% *UCL* (given by (46)) is smaller than the specified coverage of 0.95. This is especially true when the sample size is smaller than 10-20. As expected, for larger samples sizes, the coverage provided by the 95% Chebyshev *UCL* is at least 95%. For larger samples, the Chebyshev 95% *UCL* will result in a higher (but stable) *UCL* of the mean of positively skewed gamma distributions.

It is observed (Singh and Singh (2003)) that for moderately skewed to highly skewed lognormally distributed data sets (e.g., with σ exceeding 1), 95% Chebyshev *MVUE UCL* does not provide the specified coverage to the population mean. This is true when the sample size is less than 10-50. Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C. For highly skewed (e.g., $\sigma > 2$), lognormal data sets of sizes, n less than 50-70, the *H-UCL* results in unstable (impractical values which are orders of magnitude higher than other *UCLs*) unjustifiably large *UCL* values (Singh et al., (2002a)). For such highly skewed lognormally distributed data sets of sizes less than 50 - 70, one may want to use 97.5% or 99% Chebyshev *MVUE UCL* of the mean as an estimate of the EPC term (Singh and Singh (2003)). These recommendations are summarized in Table A2.

It should also be noted that for skewed data sets, the coverage provided by a 95% *UCL* based upon Chebyshev inequality is higher than those based upon the percentile bootstrap method or the BCA bootstrap method. Thus for skewed data sets, the Chebyshev inequality based 95% *UCL* of the mean (samples of all sizes from both lognormal and gamma distributions) performs better than the 95% *UCL* based upon the BCA bootstrap method. Also, when data are lognormally distributed, the coverage provided by Chebyshev *MVUE UCL* (Singh and Singh (2003)) is better than the one based upon Hall's bootstrap or bootstrap-t method. This is especially true when the sample size starts exceeding 10-15. However, for highly skewed data sets of sizes less than 10-15, it is noted that Hall's bootstrap method provides slightly better coverage than the Chebyshev *MVUE UCL* method. Just as for the gamma distribution, it is observed that for lognormally distributed data sets, the coverage provided by Hall's and bootstrap-t methods do not increase much with the sample size.

4.9 (1-α) 100% *UCL* of the Mean Using the Jackknife and Bootstrap Methods

Bootstrap and jackknife methods as discussed by Efron (1982) are non-parametric statistical resampling techniques which can be used to reduce the bias of point estimates and construct approximate confidence intervals for parameters, such as the population mean. These two methods require no assumptions regarding the statistical distribution (e.g., normal, lognormal, or gamma) of the underlying population, and can be applied to a variety of situations no matter how complicated. There exists in the literature of statistics an extensive array of different bootstrap methods for constructing confidence intervals for the population mean, μ_1. In the ProUCL, Version 3.0 software package, five bootstrap methods have been incorporated:

1) the standard bootstrap method,
2) bootstrap-t method (Efron, 1982, Hall, 1988),
3) Hall's bootstrap method (Hall, 1992, Manly, 1997),
4) simple bootstrap percentile method (Manly, 1997), and
5) bias-corrected accelerated (BCA) percentile bootstrap method (Efron and Tibshirani, 1993, Many, 1997).

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a population with an unknown parameter, θ (*e.g.*, $\theta = \mu_1$) , and let $\hat{\theta}$ be an estimate of θ, which is a function of all n observations. For example, the parameter, θ, could be the population mean, and a reasonable choice for the estimate, $\hat{\theta}$, might be the sample mean, $\bar{x}$. Another choice for $\hat{\theta}$ is the *MVUE* of the mean of a lognormal population, especially when dealing with lognormal data sets.

4.9.1 (1-α) 100% UCL of the Mean Based Upon the Jackknife Method

In the jackknife approach, *n* estimates of θ are computed by deleting one observation at a time (Dudewicz and Misra (1988)). Specifically, for each index, *I*, denote by $\hat{\theta}_{(I)}$, the estimate of θ (computed similarly as $\hat{\theta}$) when the *i*th observation is omitted from the original sample of size *n*, and let the arithmetic mean of these estimates be given by

$$\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)}. \tag{49}$$

A quantity known as the *i*th "pseudo-value" is defined by

$$J_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}. \tag{50}$$

The jackknife estimator of θ is given by the following equation.

$$J(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}J_i = n\hat{\theta} - (n-1)\tilde{\theta}. \tag{51}$$

If the original estimate $\hat{\theta}$ is biased, then under certain conditions, part of the bias is removed by the jackknife method, and an estimate of the standard error of the jackknife estimate, $J(\hat{\theta})$, is given by

$$\hat{\sigma}_{J(\hat{\theta})} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(J_i - J(\hat{\theta})\right)^2}. \tag{52}$$

Next, consider the t-type statistic given by

$$t = \frac{J(\hat{\theta}) - \theta}{\hat{\sigma}_{J(\hat{\theta})}}. \quad (53)$$

The t-type statistic given by (53) has an approximate Student's-t distribution with $n-1$ degrees of freedom, which can be used to derive the following approximate $(1-\alpha)100\%$ *UCL* for θ,

$$UCL = J(\hat{\theta}) + t_{\alpha, n-1}\hat{\sigma}_{J(\hat{\theta})}. \quad (54)$$

If the sample size, *n*, is large, then the upper α^{th} *t*-quantile in equation (54) can be replaced with the corresponding upper αth standard normal quantile, z_α. Observe, also, that when $\hat{\theta}$ is the sample mean, $\bar{x}$, then the jackknife estimate is also the sample mean, $J(\bar{x}) = \bar{x}$, and the estimate of the standard error given by equation (52) simplifies to $s_x/n^{1/2}$, and the *UCL* in equation (54) reduces to the familiar *t*- statistic based *UCL* given by equation (32). ProUCL uses the jackknife estimate as the sample mean leading to $J(\bar{x}) = \bar{x}$, which in turn translates equation (54) to the *UCL* given by equation (32). This method has been included in ProUCL to satisfy the curiosity of those users who do not recognize that this jackknife method (with sample mean as the estimator) yields a *UCL* of the population mean identical to the *UCL* based upon the Student's-t statistic as given by equation (32).

4.9.2 (1-α) 100% *UCL* of the Mean Based Upon Standard Bootstrap Method

In bootstrap resampling methods, repeated samples of size *n* are drawn with replacement from a given set of observations. The process is repeated a large number of times (e.g., 2000 times), and each time an estimate, $\hat{\theta}_I$, of θ is computed. The estimates thus obtained are used to compute an estimate of the standard error of $\hat{\theta}$. A description of the bootstrap method, illustrated by application to the population mean, μ_1, and the sample mean, $\bar{x}$, is given as follows.

Step 1. Let $(x_{i1}, x_{i2}, \ldots, x_{in})$ represent the i[th] sample of size n with replacement from the original data set $(x_1, x_2, \ldots, x_n)$. Then compute the sample mean and denote it by $\bar{x}_i$.

Step 2. Perform Step 1 independently N times (e.g., 1000-2000), each time calculating a new estimate. Denote those estimates by $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N$. The bootstrap estimate of the population mean is the arithmetic mean, $\bar{x}_B$, of the N estimates $\bar{x}_i : i = 1, 2, \ldots, N$. The bootstrap estimate of the standard error of the estimate, $\bar{x}$, is given by,

$$\hat{\sigma}_B = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\bar{x}_i - \bar{x}_B\right)^2}. \tag{55}$$

If some parameter, θ (say, the population median), other than the mean is of concern with an associated estimate (e.g., the sample median), then the same steps described above could be applied with the parameter and its estimate used in place of μ_1 and $\bar{x}$. Specifically, the estimate, $\hat{\theta}_I$, would be computed, instead of $\bar{x}_i$, for each of the N bootstrap samples. The general bootstrap estimate, denoted by $\bar{\theta}_B$, is the arithmetic mean of the N estimates. The difference, $\bar{\theta}_B - \hat{\theta}$, provides an estimate of the bias of the estimate, $\hat{\theta}$, and an estimate of the standard error of $\hat{\theta}$ is given by

$$\hat{\sigma}_B = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\theta}_i - \bar{\theta}_B\right)^2}. \tag{56}$$

The $(1-\alpha)100\%$ standard bootstrap UCL for θ is given by

$$UCL = \hat{\theta} + z_\alpha \hat{\sigma}_B. \tag{57}$$

ProUCL computes the standard bootstrap UCL by using the population AM and sample AM, respectively given by μ_1 and $\bar{x}$. It is observed that the UCL obtained using the standard bootstrap method is quite similar to the UCL obtained using the Student's-t statistic as given by equation (32), and, as such, does not adequately adjust for skewness. For skewed data sets, the coverage provided by standard bootstrap *UCL* is much lower than the specified coverage.

Note: For lognormally distributed data sets, one may want to use the jackknife and the standard bootstrap methods on the *MVUE* of the population mean, μ_1, given by equation (14). However, the performance of these methods have not been studied. Also, these methods have not been included in ProUCL.

4.9.3 (1-α) 100% *UCL* of the Mean Based Upon Simple Percentile Bootstrap Method

Bootstrap resampling of the original data set is used to generate the bootstrap distribution of the unknown population mean (Manly, 1997). In this method, $\bar{x}_i$, the sample mean is computed from the i^{th} resampling (i=1,2,..., N) of the original data. These $\bar{x}_i$, i:=1,2,...,N are arranged in ascending order as $\bar{x}_{(1)} \le \bar{x}_{(2)} \le .... \le \bar{x}_{(N)}$. The *(1-α)100% UCL* of the population mean, μ_1 is given by the value, that exceeds the (1-α)100% of the generated mean values. The 95% *UCL* of the mean is the 95th percentile of the generated means and is given by:

$$95\% \, Percentile - UCL = 95^{th}\% \bar{x}_i ; i = 1,2,\ldots,N \tag{58}$$

For example, when N=1000, a simple bootstrap 95% percentile-*UCL* is given by the 950th ordered mean value given by $\bar{x}_{(950)}$.

Singh and Singh (2003) observed that for skewed data sets, the coverage provided by this simple percentile bootstrap method is much lower than the coverage provided by the bootstrap-t and Hall's bootstrap methods. It is observed that for skewed (lognormal and gamma) data sets, the BCA bootstrap method performs slightly better than the simple percentile method. Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are provided in Appendix C.

4.9.4 (1-α) 100% *UCL* of the Mean Based Upon Bias - Corrected Accelerated (BCA) Percentile Bootstrap Method

The BCA bootstrap method is also a percentile bootstrap method which adjusts for bias in the estimate (Efron and Tibshirani, 1993, Manly, 1997). The performance of this method for skewed distributions (e.g., lognormal and gamma) is not well studied. It was conjectured that the BCA method would perform better than the various other methods. Singh and Singh (2003) investigated and compare its performance (in terms of coverage probabilities) with parametric methods and other bootstrap methods. For skewed data sets, this method does represent a slight improvement (in terms of coverage probability) over the simple percentile method. However, this improvement is not adequate enough and yields *UCLs* with coverage probability much lower than the specified coverage of 0.95. The BCA upper confidence limit of intended (1-α) 100% coverage is given by the following equation:

$$BCA - UCL = \bar{x}^{(\alpha_2)}, \qquad (59)$$

where $\bar{x}^{(\alpha_2)}$ is the α_2 100th percentile of the distribution of the $\bar{x}_i; i = 1,2,\ldots,N$. For example, when N=2000, $\bar{x}^{(\alpha_2)}$ = $(\alpha_2 N)^{th}$ ordered statistic of $\bar{x}_i; i = 1,2,\ldots,N$ given by $\bar{x}_{(\alpha_2 N)}$. Here α_2 is given by the following probability statement.

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})}\right) \qquad (60)$$

Where Φ(.) is the standard normal cumulative distribution function and $z^{(1-\alpha)}$ is the $100^*(1-\alpha)^{th}$ percentile of a standard normal distribution. For example, $z^{(0.95)} = 1.645$, and Φ(1.645) = 0.95.

Also in equation (60), $\hat{z}_0$ (bias correction)and $\hat{a}$ (acceleration factor) are given as follows.

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#(\bar{x}_i < \bar{x})}{N}\right) \tag{61}$$

where Φ^{-1} (.) is the inverse function of a standard normal cumulative distribution function, e.g., Φ^{-1} (0.95)=1.645. $\hat{a}$ is the acceleration factor and is given by the following equation.

$$\hat{a} = \frac{\sum (\bar{x} - \bar{x}_{-i})^3}{6[\sum (\bar{x} - \bar{x}_{-i})^2]^{1.5}} \tag{62}$$

where summation in (62) is being carried from i = 1 to I = n, the sample size. $\bar{x}$ is the sample mean based upon all n observations, and $\bar{x}_{-i}$ is the mean of (n-1) observations without the i^{th} observation, i = 1,2,...,n.

Singh and Singh (2003) observed that for skewed data sets (e.g., gamma and lognormal), the coverage provided by this BCA percentile method is much lower than the coverage provided by the bootstrap-t and Hall's bootstrap methods. This is especially true when the sample size is small. The BCA method does provide an improvement over the simple percentile method and the standard bootstrap method. However, bootstrap-t and Hall's bootstrap methods perform better (in terms of coverage probabilities) than the BCA method. For skewed data sets, the BCA method also performs better than the modified-t *UCL*. For gamma distributions, the coverage provided by BCA 95% *UCL* approaches 0.95 as the sample size increases. For lognormal distributions, the coverage provided by the BCA 95% *UCL* is much lower than the specified coverage of 0.95.

4.9.5 (1-α) 100% *UCL* of the Mean Based Upon Bootstrap-t Method

Another variation of the bootstrap method, called the "bootstrap-t" by Efron (1982), is a non-parametric method which uses the bootstrap methodology to estimate quantiles of the pivotal quantity, t statistic, given by equation (31). Rather than using the quantiles of the familiar Student's-t statistic, Hall (1988) proposed to compute estimates of the quantiles of the statistic given by equation (31) directly from the data.

Specifically, in Steps 1 and 2 described above in Section 4.9.2, if $\bar{x}$ is the sample mean computed from the original data, and $\bar{x}_i$ and $s_{x,I}$ are the sample mean and sample standard deviation computed from the *i*th resampling of the original data, the *N* quantities $t_i = (\sqrt{n})(\bar{x}_i - \bar{x})/s_{x,i}$ are computed and sorted, yielding ordered quantities, $t_{(1)} \le t_{(2)} \le \cdots \le t_{(N)}$. The estimate of the lower α^{th} quantile of the pivotal quantity in equation (31) is $t_{\alpha, B} = t_{(\alpha N)}$. For example, if N = 1000 bootstrap samples are generated, then the 50th ordered value, $t_{(50)}$, would be the bootstrap estimate of the lower 0.05th quantile of the pivotal quantity in equation (31). Then a *(1-α) 100% UCL* of the population mean based upon the bootstrap-t method is given by

$$UCL = \bar{x} - t_{(\alpha N)} s_x/\sqrt{n}. \tag{63}$$

Note the '-' sign in equation (63). ProUCL computes the Bootstrap-t *UCL* based upon the quantiles obtained using the sample mean, $\bar{x}$. It is observed that the *UCL* based upon the bootstrap-t method is more conservative than the other *UCLs* obtained using the Student's- t, modified -t, adjusted -*CLT*, and the standard bootstrap methods. This is specially true for skewed data sets. This method seems to adjust for skewness to some extent.

It is observed that for skewed data sets (e.g., gamma, lognormal), the 95% UCL based upon bootstrap-t method performs better than the 95% *UCLs* based upon the simple percentile and the BCA percentile methods (Singh and Singh (2003)). For highly skewed (k < 0.1 or σ > 2.5-3.0) data sets of small sizes (e.g., n < 10) the bootstrap-t method performs better than other (adjusted gamma *UCL*, or Chebyshev inequality *UCL*) *UCL* computation methods. It is noted that for gamma distribution, the performances (coverages provided by the respective *UCLs*) of bootstrap-t and Hall's bootstrap methods are very similar. It is also noted that for larger samples, these two methods (bootstrap-t and Hall's bootstrap) approximately provide the specified 95% coverage to the mean, kθ, of the gamma distribution. For gamma distributed data sets, the coverage provided by a bootstrap-t (and Hall's bootstrap *UCL*) 95% *UCL* approaches 95% as sample size increases for all values of k considered (k = 0.05-5.0) in Singh and Singh (2003). However, it is noted that the coverage provided by these two bootstrap methods is slightly lower than 0.95 for samples of smaller sizes.

For lognormally distributed data sets, the coverage provided by bootstrap-t 95% *UCL* is a little bit lower than the coverage provided by the 95% *UCL* based upon Hall's bootstrap method. However, it should be noted that for lognormally distributed data sets, for samples of all sizes, the coverage provided by these two methods (bootstrap-t and Hall's bootstrap) is significantly lower than the specified 0.95 coverage. This is especially true for moderately skewed to highly skewed (e.g., $\sigma>1.0$) lognormally distributed data sets. This can be seen from the graphs presented in Appendix C.

It should be pointed out that the bootstrap-t and Hall's bootstrap methods sometimes result in unstable, erratic, and unreasonably inflated *UCL* values especially in the presence of outliers (Efron and Tibshirani, 1993). Therefore, these two methods should be used with caution. In case these two methods result in erratic and inflated *UCL* values, then an appropriate Chebyshev inequality based *UCL* may be used to estimate the EPC term for non-parametric skewed data sets.

4.9.6 (1-α) 100% *UCL* of the Mean Based Upon Hall's Bootstrap Method

Hall (1992) proposed a bootstrap method which adjusts for bias as well as skewness. This method has been included in UCL guidance document (EPA 2002). For highly skewed data sets (e.g., LN(5,4)), it performs slightly better (higher coverage) than the bootstrap-t method. In this method, $\bar{x}_i$, $s_{x, I,}$ and $\hat{k}_{3i}$, the sample mean, sample standard deviation, and sample skewness are computed from the *i*th resampling (I = 1, 2,..., N) of the original data. Let $\bar{x}$ be the sample mean, s_x be the sample standard deviation, and $\hat{k}_3$ be the sample skewness (as given by equation (43)) computed from the original data. The quantities W_i and Q_i given as follows are computed for each of the N bootstrap samples, where

$$W_i = (\bar{x}_i - \bar{x}) / s_{xi} \text{, and} \qquad Q_i(W_i) = W_i + \hat{k}_{3i} W_i^2 / 3 + \hat{k}_{3i}^2 W_i^3 / 27 + \hat{k}_{3i} / (6n).$$

The quantities $Q_i(W_i)$ given above are arranged in ascending order. For a specified (1-α) confidence coefficient, compute the $(\alpha N)^{th}$ ordered value, q_α of quantities $Q_i(W_i)$. Next, compute $W(q_\alpha)$ using the inverse function, which is given as follows:

$$W(q_\alpha) = 3\left(\left(1 + \hat{k}_3(q_\alpha - \hat{k}_3 / (6n))\right)^{1/3} - 1\right) / \hat{k}_3. \tag{64}$$

In equation (64), $\hat{k}_3$ is computed using equation (43). Finally, the *(1-α) 100% UCL* of the population mean based upon Hall's bootstrap method (Manly, 1997) is given as follows:

$$UCL = \bar{x} - W(q_\alpha) * s_x. \tag{65}$$

For gamma distribution, Singh and Singh (2003) observed that the coverage probabilities provided by the 95% *UCL*s based upon bootstrap-t and Hall's bootstrap methods are in close agreement. For larger samples these two methods approximately provide the specified 95% coverage to the population mean, kθ of a gamma distribution. For smaller sample sizes (from gamma distribution), the coverage provided by these two methods is slightly lower than the specified level of 0.95. For both lognormal and gamma distributions, these two methods (bootstrap-t and Hall's bootstrap) perform better than the other bootstrap methods, namely, the standard bootstrap method, simple percentile, and bootstrap BCA percentile methods. This can be seen from graphs presented in Appendix C.

Just like the gamma distribution, for lognormally distributed data sets, it is noted that Hall's *UCL* and bootstrap-t *UCL* provide similar coverages. However, for highly skewed lognormal data sets, the coverages based upon Hall's method and bootstrap-t method are significantly lower than the specified 0.95 coverage (Singh and Singh (2003)). This is true even in samples of larger sizes(e.g., n=100). For lognormal data sets, the coverages provided by Hall's bootstrap and bootstrap-t methods do not increase much with the sample size, n. For highly skewed (e.g., $\hat{\sigma}$ > 2.0) data sets of small sizes (e.g., n < 15), Hall's bootstrap method (and also bootstrap-t method) performs better than Chebyshev *UCL*, and for larger samples, Chebyshev *UCL* performs better than Hall's bootstrap method. Similar to the bootstrap-t method, it should be noted that Hall's bootstrap method sometimes results in unstable, inflated, and erratic values especially in the presence of outliers (Efron and Tibshirani, 1993). Therefore, these two methods should be used with caution. If outliers are present in a data set, then a 95% *UCL* of the mean should be computed using alternative *UCL* computation methods.

5. Recommendations and Summary

This section describes the recommendations and summary on the computation of a 95% *UCL* of the unknown population arithmetic mean, μ_1, of a contaminant data distribution without censoring. These recommendations are based upon the findings of Singh, Singh, and Engelhardt (1997, 1999); Singh et al. (2002a); Singh, Singh, and Iaci (2002b); and Singh and Singh (2003). Recommendations have been summarized for: 1) normally distributed data sets, 2) gamma distributed data sets, 3) lognormally distributed data sets, and 4) data sets which are non-parametric and do not follow any of the three distributions included in ProUCL.

For skewed parametric as well as non-parametric data sets, there is no simple solution to compute a *95% UCL* of the population mean, μ_1 . Singh et al. (2002a), Singh, Singh, and Iaci (2002b), and Singh and Singh (2003) noted that the *UCLs* based upon the skewness adjusted methods, such as the Johnson's modified-t and Chen's adjusted-*CLT* do not provide the specified coverage (e.g., 95%) to the population mean even for mildly to moderately skewed (e.g., $\hat{\sigma}$ in interval [0.5, 1.0)) data sets for samples of size as large as 100. The coverage of the population mean by these skewness-adjusted *UCLs* gets poorer (much smaller than the specified coverage of 0.95) for highly skewed data sets, where the skewness levels are defined in Section 3.2.2 as a function of σ or $\hat{\sigma}$ (standard deviation of log-transformed data).

5.1 Recommendations to Compute a 95% *UCL* of the Unknown Population Mean, μ_1 Using Symmetric and Positively Skewed Data Sets

Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods considered are given in Appendix C. The user may want to consult those graphs for a better understanding of the summary and recommendations made in this section.

5.1.1 Normally or Approximately Normally Distributed Data sets

As expected, for a normal distribution, N(μ_1, σ_1^2), Student's-t statistic, modified- t statistic, and bootstrap-t 95% *UCL* computation methods result in *UCLs* which provide coverage probabilities close to the nominal level, 0.95. Contrary to the general conjecture, the bootstrap, BCA method does not perform better than the other bootstrap methods (e.g., bootstrap-t). Actually, for normally distributed data sets, the coverages for the population mean, μ_1 provided by the *UCL*s based upon the BCA method and Hall's bootstrap method are lower than the specified 95% coverage. This is especially true when the sample size, n is less than 30. For details refer to Singh and Singh (2003).

- For normally distributed data sets, a *UCL* based upon the Student's-t statistic as given by equation (32) provides the optimal *UCL* of the population mean. Therefore, for normally distributed data sets, one should always use a 95% *UCL* based upon the Student's-t statistic.

- The *95% UCL* of the mean given by equation (32) based upon Student's-t statistic may also be used when the *Sd,* s_y of the log-transformed data is less than 0.5, or when the data set approximately follows a normal distribution. A data set is approximately normal when the normal Q-Q plot displays a linear pattern (without outliers and jumps) and the resulting correlation coefficient is high (e.g., 0.95 or higher).

- Student's-t *UCL* may also be used when the data set is symmetric (but possibly not normally distributed). A measure of symmetry (or skewness) is $\hat{k}_3$ which is given by equation (43). A value of $\hat{k}_3$ close to zero (e.g., if absolute value of skewness is roughly less than 0.2 or 0.3) suggests approximate symmetry. The approximate symmetry of a data distribution can also be judged by looking at the histogram of the data set.

5.1.2 Gamma Distributed Skewed Data Sets

In practice, many skewed data sets can be modeled both by a lognormal distribution and a gamma distribution especially when the sample size is smaller than 70-100. As well known, the 95% *H-UCL* of the mean based upon a lognormal model often results in unjustifiably large and impractical 95% *UCL* value. In such cases, a gamma model, G(k,θ) may be used to compute a reliable 95% *UCL* of the unknown population mean, μ_1.

- Many skewed data sets follow a lognormal as well as a gamma distribution. It should be noted that the population means based upon the two models can differ significantly. Lognormal model based upon a highly skewed (e.g., $\hat{\sigma} \geq 2.5$)data set will have an unjustifiably large and impractical population mean, μ_1 and its associated *UCL*. The gamma distribution is better suited to model positively skewed environmental data sets.

One should always first check if a given skewed data set follows a gamma distribution. If a data set does follow a gamma distribution or an approximate gamma distribution, one should compute a 95% *UCL* based upon a gamma distribution. Use of highly skewed (e.g., $\hat{\sigma} \geq$ 2.5-3.0) lognormal distributions should be avoided. For such highly skewed lognormally distributed data sets which can not be modeled by a gamma or an approximate gamma distribution, non-parametric *UCL* computation methods based upon the Chebyshev inequality may be used.

- The five bootstrap methods do not perform better than the two gamma *UCL* computation methods. It is noted that the performances (in terms of coverage probabilities) of bootstrap-t and Hall's bootstrap methods are very similar. Out of the five bootstrap methods, bootstrap-t and Hall's bootstrap methods perform the best (with coverage probabilities for the population mean closer to the nominal level of 0. 95). This is especially true when skewness is quite high (e.g., $\hat{k}$ < 0.1) and sample size is small (e.g., n < 10-15). This can be seen from graphs given in Appendix C.

- The bootstrap BCA method does not perform better than the Hall's method or the bootstrap-t method. The coverage for the population mean, μ_1 provided by the BCA method is much lower than the specified 95% coverage. This is especially true when the skewness is high (e.g., $\hat{k}$ <1) and sample size is small (Singh and Singh (2003)).

- From the results presented in Singh, Singh, and Iaci (2002b) and in Singh and Singh (2003), it is concluded that for data sets which follow a gamma distribution, a 95% *UCL* of the mean should be computed using the adjusted gamma *UCL* when the shape parameter, k is: $0.1 \leq k < 0.5$, and for values of $k \geq 0.5$, a 95% *UCL* can be computed using an approximate gamma *UCL* of the mean, μ_1.

- For highly skewed gamma distributed data sets with $k < 0.1$, bootstrap-t *UCL* or Hall's bootstrap (Singh and Singh (2003)) may be used when the sample size is smaller than 15, and the adjusted gamma *UCL* should be used when sample size starts approaching and exceeding 15. The small sample size requirement increases as skewness increases (that is as k decreases, the required sample size, n increases).

- The bootstrap-t and Hall's bootstrap methods should be used with caution as some times these methods yield erratic, unreasonably inflated, and unstable *UCL* values especially in the presence of outliers. In case Hall's bootstrap and bootstrap-t methods yield inflated and erratic *UCL* results, the 95% *UCL* of the mean should be computed based upon the adjusted gamma 95% *UCL*. ProUCL prints out a warning message associated with the recommended use of the *UCLs* based upon the bootstrap-t method or Hall's bootstrap method.

These recommendations for the use of gamma distribution are summarized in Table A1.

Table A1.

Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1 of a Gamma Distribution

$\hat{k}$	***Sample Size, n***	***Recommendation***
$\hat{k} \geq 0.5$	For all n	Approximate Gamma 95%*UCL*
$0.1 \leq \hat{k} < 0.5$	For all n	Adjusted Gamma 95% *UCL*

$\hat{k} < 0.1$	n < 15	95% *UCL* Based Upon Bootstrap-t or Hall's Bootstrap Method *
$\hat{k} < 0.1$	n ≥ 15	Adjusted Gamma 95% *UCL* if available, otherwise use Approximate Gamma 95% *UCL*

* In case bootstrap-t or Hall's bootstrap methods yield erratic, inflated, and unstable *UCL* values, the *UCL* of the mean should be computed using adjusted gamma *UCL*.

5.1.3 Lognormally Distributed Skewed Data Sets

For lognormally, LN(μ, σ^2) distributed data sets, the H-statistic based *UCL* does provide the specified 0.95 coverage for the population mean for all values of σ. However, the H-statistic often results in unjustifiably large *UCL* values which do not occur in practice. This is especially true when skewness is high (e.g., σ > 2.0). The use of a lognormal model unjustifiably accommodates large and impractical values of the mean concentration and its *UCLs*. The problem associated with the use of a lognormal distribution is that the population mean, μ_1, of a lognormal model becomes impractically large for larger values of σ which in turn results in inflated *H-UCL* of the population mean, μ_1. Since the population mean of a lognormal model becomes too large, none of the other methods except for *H-UCL* provides the specified 95% coverage for that inflated population mean, μ_1. This is especially true when the sample size is small and skewness is high. For extremely highly skewed data sets (with σ > 2.5-3.0) of smaller sizes (e.g., < 70-100), the use of a lognormal distribution based *H-UCL* should be avoided (e.g., see Singh et al. (2002a), Singh and Singh (2003)). Therefore, alternative *UCL* computation methods such as the use of a gamma distribution or use of a *UCL* based upon non-parametric bootstrap methods or Chebyshev inequality based methods are desirable.

As expected for skewed (e.g., with σ (or $\hat{\sigma}$) ≥ 0.5) lognormally distributed data sets, the Student's-t *UCL*, modified-t *UCL*, adjusted -CLT *UCL*, standard bootstrap method all fail to provide the specified 0.95 coverage for the unknown population mean for samples of all sizes. Just like the gamma distribution, the performances (in terms of coverage probabilities) of bootstrap-t and Hall's bootstrap methods are very similar (Singh and Singh (2003)). However, it is noted that the coverage provided by Hall's bootstrap (and also by bootstrap-t) is much lower than the specified 95% coverage for the population mean, μ_1, for samples of all sizes of varying skewness. Moreover, the coverages provided by Hall's bootstrap or bootstrap-t method do not increase much with the sample size.

Also the coverage provided by the BCA method is much lower than the coverage provided by Hall's method or bootstrap-t method. Thus the BCA bootstrap method can not be recommended to compute a 95% *UCL* of the mean of a lognormal population. For highly skewed data sets of small sizes (e.g., < 15) with σ exceeding 2.5-3.0, even the Chebyshev inequality based *UCL*s fail to provide the specified 0.95 coverage for the population. However, as the sample size increases, the coverages provided by the chebyshev inequality based *UCL*s also increase. For such highly skewed data sets ($\hat{\sigma} > 2.5$) of sizes less than 10-15, Hall's bootstrap or bootstrap-t methods provide larger coverage than the coverage provided by the 99% Chebyshev *(MVUE) UCL*. Therefore, for highly skewed lognormally distributed data sets of small sizes, one may use Hall's method to compute an estimate of the EPC term. The small sample size requirement increases with σ. Graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods are given in Appendix C.

It should be noted that even a small increase in the *Sd, σ,* increases skewness considerably. For example, for a lognormal distribution, when $\sigma = 2.5$, skewness ~ 11825.1; and when $\sigma = 3$, skewness ~ 729555. In practice, the occurrence of such highly skewed data sets (e.g., $\sigma \geq 3$) is not very common. Nevertheless, these highly skewed data sets can arise occasionally and, therefore, require separate attention. Singh et al. (2002a) observed that when the *Sd*, σ, starts approaching 2.5 (that is, for lognormal data, when *CV* > 22.74 and skewness > 11825.1), even a *99% Chebyshev (MVUE) UCL* fails to provide the desired 95% coverage for the population mean, μ_1. This is especially true when the sample size, n is smaller than 30. For such extremely skewed data sets, the larger of the two *UCLs*: the *99% Chebyshev (MVUE) UCL* and the non-parametric *99% Chebyshev (Mean, Sd) UCL*, may be used as an estimate of the EPC.

It is also noted that, as the sample size increases, the *H-UCL* starts behaving in a stable manner. Therefore, depending upon the *Sd, σ (actually its MLE* $\hat{\sigma}$*)*, for lognormally distributed data sets, one can use the *H-UCL* for samples of larger sizes such as greater than 70-100. This large sample size requirement increases as the Sd, $\hat{\sigma}$, increases, as can be seen in Table A2. ProUCL can compute an *H-UCL* for samples of sizes up to 1000. For lognormally distributed data sets of smaller sizes, some alternative methods to compute a 95% *UCL* of the population mean, μ_1 are summarized in Table A2.

Furthermore, it is noted that for larger sample sizes (e.g., n > 150), the *H-UCL* becomes even smaller than the Student's-t *UCL* and various other *UCLs*. It should be pointed out that the large sample behavior of *H-UCL* has not been investigated rigorously. For confirmation purposes (that is *H-UCL* does provide the 95% coverage for larger samples also), it is desirable to conduct such a study for samples of larger sizes.

Since skewness (as defined in Section 3.2.2) is a function of σ (or $\hat{\sigma}$), the recommendations for the computation of the *UCL* of the population mean are also summarized in Table A2 for various values of the *MLE* $\hat{\sigma}$ of σ and the sample size, n. Here $\hat{\sigma}$ is an *MLE* of σ, and is given by the *Sd* of log-transformed data given by equation (2). Note that Table A2 is applicable to the computation of a 95% *UCL* of the population mean based upon lognormally distributed data sets without non-detect observations. A method to compute a 95% *UCL* of the mean of a lognormal distribution is summarized as follows:

- Skewed data sets should be first tested for a gamma distribution. For lognormally distributed data sets (which can not be modeled by a gamma distribution), the method as summarized in Table A2 may be used to compute a 95% *UCL* of the mean.

- Specifically, for highly skewed (e.g., $1.5 < \sigma \leq 2.5$) data sets of small sizes (e.g., n $\leq$ 50-70), the EPC term may be estimated by using a 97.5% or 99% *MVUE* Chebyshev *UCL* of the population mean. For larger samples (e.g., n > 70), *H-UCL* may be used to estimate the EPC.

- For extremely highly skewed (e.g., $\sigma > 2.5$) lognormally distributed data sets, the population mean becomes unrealistically large. Therefore, the use of *H-UCL* should be avoided especially when the sample size is less than100. For such highly skewed data sets, Hall's bootstrap *UCL* may be used when the sample size is less than 10-15 (Singh and Singh (2003)). The small sample size requirement increases with $\hat{\sigma}$. For example, n = 10 is considered small when $\hat{\sigma}$ = 3.0, and n = 15 is considered small when $\hat{\sigma}$ = 3.5.

- Hall's bootstrap methods should be used with caution as some times it yields erratic, inflated, and unstable *UCL* values, especially in the presence of outliers. For these highly skewed data sets of size, n (e.g., less than 10-15), in case Hall's bootstrap method yields an erratic and inflated *UCL* value, the 99% Chebyshev *MVUE UCL* may be used to estimate the EPC term. ProUCL displays a warning message associated with the recommended use of Hall's bootstrap method.

Table A2. Summary Table for the Computation of a

95% *UCL* of the Unknown Mean, μ_1 of a Lognormal Population

$\hat{\sigma}$	*Sample Size, n*	*Recommendation*
$\hat{\sigma} < 0.5$	For all n	Student's-t, modified-t, or *H-UCL*
$0.5 \le \hat{\sigma} < 1.0$	For all n	*H-UCL*
$1.0 \le \hat{\sigma} < 1.5$	$n < 25$	95% Chebyshev *(MVUE) UCL*
	$n \ge 25$	*H-UCL*
$1.5 \le \hat{\sigma} < 2.0$	$n < 20$	99% Chebyshev *(MVUE) UCL*
	$20 \le n < 50$	95% Chebyshev *(MVUE) UCL*
	$n \ge 50$	*H-UCL*
$1.5 \le \hat{\sigma} < 2.0$	$n < 20$	99% Chebyshev *(MVUE) UCL*
	$20 \le n < 50$	97.5% Chebyshev *(MVUE) UCL*
	$50 \le n < 70$	95% Chebyshev *(MVUE) UCL*
	$n \ge 70$	*H-UCL*
$2.5 \le \hat{\sigma} < 3.0$	$n < 30$	Larger of (99% Chebyshev *(MVUE) UCL* or 99% Chebyshev (Mean, Sd))
	$30 \le n < 70$	97.5% Chebyshev *(MVUE) UCL*
	$70 \le n < 100$	95% Chebyshev *(MVUE) UCL*
	$n \ge 100$	*H-UCL*
$3.0 \le \hat{\sigma} \le 3.5$	$n < 15$	Hall's bootstrap method *
	$15 \le n < 50$	Larger of (99% Chebyshev (*MVUE) UCL*, 99% Chebyshev(Mean, Sd))
	$50 \le n < 100$	97.5% Chebyshev *(MVUE) UCL*
	$100 \le n < 150$	95% Chebyshev *(MVUE) UCL*
	$n \ge 150$	*H-UCL*
$\hat{\sigma} > 3.5$	For all n	Use non-parametric methods *

* In case Hall's bootstrap method yields an erratic unrealistically large UCL value, then the UCL of the mean may be computed based upon the Chebyshev inequality.

5.1.4 Data Sets Without a Discernable Skewed Distribution - Non-parametric Skewed Data Sets

The use of gamma and lognormal distributions as discussed here will cover a wide range of skewed data distributions. For skewed data sets which are neither gamma nor lognormal, one can use a non-parametric Chebyshev *UCL* or Hall's bootstrap *UCL* (for small samples) of the mean to estimate the EPC term.

- For skewed non-parametric data sets with negative and zero values, use a *95% Chebyshev (Mean, Sd) UCL* for the population mean, μ_1.

For all other non-parametric data sets with only positive values, the following method may be used to estimate the EPC term.

- For mildly skewed data sets with $\hat{\sigma} \leq 0.5$, one can use Student's-t statistic or modified-t statistic to compute a 95% *UCL* of mean, μ_1.

- For non-parametric moderately skewed data sets (e.g., σ or its estimate, $\hat{\sigma}$ in the interval (0.5, 1]), one may use a *95% Chebyshev (Mean, Sd) UCL* of the population mean, μ_1.

- For non-parametric moderately to highly skewed data sets (e.g., $\hat{\sigma}$ in the interval (1.0, 2.0]), one may use a *99% Chebyshev (Mean, Sd) UCL or 97.5% Chebyshev (Mean, Sd) UCL* of the population mean, μ_1, to obtain an estimate of the EPC term.

- For highly skewed to extremely highly skewed data sets with $\hat{\sigma}$ in the interval (2.0, 3.0], one may use Hall's *UCL* or *99% Chebyshev (Mean, Sd) UCL* to compute the EPC term.

- Extremely skewed non-parametric data sets with σ exceeding 3.0, provide poor coverage. For such highly skewed data distributions, none of the methods considered provide the specified 95% coverage for the population mean, μ_1. The coverages provided by the various methods decrease as σ increases. For such data sets of sizes less than 30, a 95% *UCL* can be computed based upon Hall's bootstrap method or bootstrap-t method. Hall's bootstrap method provides highest coverage (but less than 0.95) when the sample size is small. It is noted that the coverage for the population mean provided by Hall's method (and bootstrap-t method) does not increase much as the sample size, n increases. However, as the sample size increases, coverage provided by *99% Chebyshev (Mean, Sd) UCL* method also increases. Therefore, for larger samples, a *UCL* should be computed based upon *99% Chebyshev (Mean, Sd)* method. This large sample size requirement increases as $\hat{\sigma}$ increases. These recommendations are summarized in Table A3.

Table A3.

Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1 of a Skewed Non-parametric Distribution with all Positive Values, Where $\hat{\sigma}$ is the Sd of Log-transformed Data

$\hat{\sigma}$	*Sample Size, n*	*Recommendation*
$\hat{\sigma} \le 0.5$	For all n	95% *UCL* based on Student's-t or Modified-t statistic
$0.5 < \hat{\sigma} \le 1.0$	For all n	95% Chebyshev (Mean, Sd) *UCL*
$1.0 < \hat{\sigma} \le 2.0$	n < 50	99% Chebyshev (Mean, Sd) *UCL*
	n ≥ 50	97.5% Chebyshev (Mean, Sd) *UCL*
$2.0 < \hat{\sigma} \le 3.0$	n <10	Hall's Bootstrap *UCL* *
	n ≥ 10	99% Chebyshev (Mean, Sd) *UCL*
$3.0 < \hat{\sigma} \le 3.5$	n < 30	Hall's Bootstrap *UCL* *
	n ≥ 30	99% Chebyshev (Mean, Sd) *UCL*
$\hat{\sigma} > 3.5$	n < 100	Hall's Bootstrap *UCL* *
	n ≥ 100	99% Chebyshev (Mean, Sd) *UCL*

* If Hall's bootstrap method yields an erratic and unstable UCL value (e.g., happens when outliers are present), a *UCL* of the population mean may be computed based upon the *99% Chebyshev (Mean, Sd)* method.

5.2 Summary of the Procedure to Compute a 95% *UCL* of the Unknown Population Mean, μ_1 Based Upon Data Sets Without Non-detect Observations

- The first step in computing a *95% UCL* of a population arithmetic mean, μ_1 is to perform goodness-of-fit tests to test for normality, lognormality, or gamma distribution of the data set under study. ProUCL has three methods to test for normality or lognormality: the informal graphical test based upon a Q-Q plot, the Lilliefors test, and the Shapiro-Wilk W test. ProUCL also has three methods to test for a gamma distribution: the informal graphical Q-Q plot based upon gamma quantiles, the Kolmogorov-Smirnov (K-S) EDF test, and the

Anderson-Darling (A-D) EDF test.

- ProUCL generates a quantile-quantile (Q-Q) plot to graphically test the normality, lognormality, or gamma distribution of the data. There is no substitute for graphical displays of a data set. On this graph, a linear pattern (e.g., with high correlation such as 0.95 or higher) displayed by bulk of data suggests approximate normality, lognormality, or gamma distribution. On this graph, points well-separated from the majority of data may be potential outliers requiring special attention. Also, any visible jumps and breaks of significant magnitudes on a Q-Q plot suggest that more than one population may be present. In that case, each of the populations should be considered separately. That is a separate EPC term should be computed for each of the populations. It is, therefore, recommended to always use the graphical Q-Q plot as it provides useful information about the presence of multiple populations (e.g., site and background data mixed together) and/or outliers. Both graphical Q-Q plot and formal goodness-of-fit tests should be used on the same data set.

- A single test statistic such as the Shapiro-Wilk test (or the A-D test etc.) may lead to the incorrect conclusion that the data are normally (or gamma) distributed even when there are more than one population present. Only a graphical display such as an appropriate Q-Q can provide this information. Obviously, when multiple populations are present, those should be separated out and the EPC terms (the *UCLs*) should be computed separately for each of those populations. **Therefore, it is strongly recommended not to skip the Goodness-of-Fit Tests Option in ProUCL.** Since the computation of an appropriate *UCL* depends upon data distribution, it is advisable that the user should take his time (instead of blindly using a numerical value of a test statistic in an effort to automate the distribution selection process) to determine the data distribution. Both graphical (e.g., Q-Q plots) and analytical procedures (Shapiro-Wilk test, K-S test etc.) should be used on the same data set to determine the most appropriate distribution of the data set under study.

- After performing the Goodness-of-Fit test, ProUCL informs the user about the data distribution: normal, lognormal, gamma distribution, or non-parametric.

- For a normally distributed (or approximately normally distributed) data set, the user is advised to use Student's-t distribution based *UCL* of the mean. Student's-t distribution (or modified-t statistic) may also be used to compute the EPC term when the data set is symmetric (e.g., $|\hat{k}_3|$ is smaller than 0.2-0.3) or mildly skewed, that is when σ or $\hat{\sigma}$ is less than 0.5.

- For gamma distributed (or approximately gamma distributed) data sets, the user is advised to: use the approximate gamma *UCL* for $\hat{k} \geq 0.5$; use the adjusted gamma *UCL* for $0.1 \leq \hat{k} < 0.5$; use bootstrap-t method (or Hall's method) when $\hat{k} < 0.1$ and the sample size, n < 15; and use the adjusted gamma *UCL* (if available) for $\hat{k} < 0.1$ and sample size, n ≥ 15. If the adjusted gamma *UCL* is not available then use the approximate gamma *UCL* as an estimate of the EPC term. In case bootstrap-t method or Hall's bootstrap method yields an erratic inflated *UCL* (e.g., when outliers are present) result, the *UCL* should be computed using the adjusted gamma *UCL* (if available) or the approximate gamma *UCL*. Some graphs from Singh and Singh (2003) showing coverage comparisons for normal, gamma, and lognormal distributions for the various methods considered are given in Appendix C.

- For lognormal data sets, ProUCL recommends (as summarized in Table A2, Section 5.1.3) a method to estimate the EPC term based upon the sample size and standard deviation of the log-transformed data, $\hat{\sigma}$. ProUCL can compute a *H-UCL* of the mean for samples of size up to 1000.

- Non-parametric *UCL* computation methods such as the modified-t, *CLT* method, adjusted-*CLT* method, bootstrap and jackknife methods are also included in ProUCL. However, it is noted that non-parametric *UCLs* based upon most of these methods do not provide adequate coverage to the population mean for moderately skewed to highly skewed data sets (e.g., see Singh et al. (2002a), and Singh and Singh (2003)).

- For data sets which are not normally, lognormally, or gamma distributed, a non-parametric *UCL* of the mean based upon the Chebyshev inequality is preferred. The *Chebyshev (Mean, Sd) UCL* does not depend upon any distributional assumptions and can be used for moderately to highly skewed data sets which do not follow any of the three data distributions incorporated in ProUCL.

- It should be noted that for extremely skewed data sets (e.g., with $\hat{\sigma}$ exceeding 3.0), even a Chebyshev inequality based 99% *UCL* of the mean fails to provide the desired coverage (e.g., 0.95) of the population mean. A method to compute the EPC term for non-parametric distributions is summarized in Table A3 of Section 5.1.4. It should be pointed out that in case Hall's bootstrap method appears to yield erratic and inflated results (typically happens when outliers are present), the 99% Chebyshev *UCL* may be used as an estimate of the EPC term.

5.3 Should the Maximum Observed Concentration be Used as an Estimate of the EPC Term?

- Singh and Singh (2003) also included the Max Test (using the maximum observed value as an estimate of the EPC term) in their simulation study. Previous (e.g., EPA 1992 RAGS Document) use of the maximum observed value has been recommended as a default value to estimate the EPC term when a 95% *UCL* (e.g., the *H-UCL*) exceeded the maximum value. However, in past (e.g., EPA 1992), only two 95% *UCL* computation methods, namely: the Student's- t *UCL* and Land's *H-UCL* were used to estimate the EPC term. ProUCL, Version

3.0 can compute a 95% *UCL* of mean using several methods based upon normal, Gamma, lognormal, and non-parametric distributions. Thus, ProUCL, Version 3.0 has about fifteen (15) 95% *UCL* computation methods, one of which (depending upon skewness and data distribution) can be used to compute an appropriate estimate of the EPC term. Furthermore, since the EPC term represents the average exposure contracted by an individual over an exposure area (EA) during a long period of time, therefore, the EPC term should be estimated by using an average value (such as an appropriate 95% *UCL* of the mean) and not by the maximum observed concentration.

- With the availability of so many *UCL* computation methods (15 of them), the developers of ProUCL Version 3.0 do not feel any need to use the maximum observed value as an estimate of the EPC term. Singh and Singh (2003) also noted that for skewed data sets of small sizes (e.g., <10-20), the Max Test does not provide the specified 95% coverage to the population mean, and for larger data sets, it overestimates the EPC term. This can also viewed in the graphs presented in Appendix C. Also, for the distributions considered, the maximum value is not a sufficient statistic for the unknown population mean. The use of the maximum value as an estimate of the EPC term ignores most (except for the maximum value) of the information contained in a data set. It is not desirable to use the maximum observed value as estimate of the EPC term representing average exposure over an EA.

- It should also be noted that for highly skewed data sets, the sample mean indeed can even exceed the upper 90%, 95 % etc. percentiles, and consequently, a 95% *UCL* of mean can exceed the maximum observed value of a data set. This is especially true when one is dealing with lognormally distributed data sets of small sizes. As mentioned before, for such highly skewed data sets which can not be modeled by a gamma distribution, a 95% *UCL* of the mean should be computed using an appropriate non-parametric method. These observations are summarized in Tables A1-A3 of this Appendix A.

- Alternatively, for such highly skewed data sets, other measures of central tendency such as the median (or some other upper percentile such as 70% percentile) and its upper confidence limit may be considered. The EPA and all other interested agencies and parties need to come to an agreement upon the use of the median and its *UCL* to estimate the EPC term for a contaminant of concern at a polluted site. It should be mentioned that the use of the sample median and/or its UCL as estimates of the EPC term needs further research and investigation.

- **It is recommended that the maximum observed value NOT be used as an estimate of the EPC term.** For the sake of interested users, ProUCL displays a warning message when the recommended 95% UCL (e.g., Hall's bootstrap UCL etc.) of the mean exceeds the observed maximum concentration. For such cases (when a 95% *UCL* does exceed the maximum observed value), if applicable, an alternative 95% *UCL* computation method is recommended by ProUCL.

References

Aitchison, J., and Brown, J.A.C. (1969), *The Lognormal Distribution*, Cambridge: Cambridge University Press.

Best, D.J., and Roberts, D.E. (1975). " The Percentage Points of the Chi-square Distribution." *Applied Statistics, 24: 385-388.*

Bowman, K. O., and Shenton, L.R. (1988), Properties of Estimators for the Gamma Distribution, Volume 89. Marcel Dekker, Inc. New York.

Bradu, D., and Mundlak, Y. (1970), "Estimation in Lognormal Linear Models," *Journal of the American Statistical Association*, 65, 198-211.

Chen, L. (1995), "Testing the Mean of Skewed Distributions," *Journal of the American Statistical Association*, 90, 767-772.

Choi, S. C., and Wette, R. (1969), Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias. Technometrics, Vol. 11, 683-690.

Dudewicz, E.D., and Misra, S.N. (1988), *Modern Mathematical Statistics*. John Wiley, New York.

D'Agostino, R.B., and Stephens, M.A. (1986), Goodness-of-Fit-Techniques, Marcel Dekker, Inc.

Efron, B. (1982), The Jackknife, the Bootstrap, and Other Resampling Plans, Philadelphia: SIAM.

Efron, B., and Tibshirani, R.J. (1993), An Introduction to the Bootstrap, Chapman & Hall, New York.

EPA(1989), "Methods for Evaluating the Attainment of Cleanup Standards, Vol. 1, Soils and Solid Media," Publication EPA 230/2-89/042.

EPA (1991), "A Guide: Methods for Evaluating the Attainment of Cleanup Standards for Soils and Solid Media," Publication EPA/540/R95/128.

EPA (1992), "Supplemental Guidance to RAGS: Calculating the Concentration Term," Publication EPA 9285.7-081, May 1992.

EPA (1996), "A Guide: Soil Screening Guidance: Technical Background Document," Second Edition, Publication 9355.4-04FS.

EPA (2002), Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites, OSWER 9285.6-10, December 2002.

ExpertFit Software (2001), Averill M. Law & Associates Inc, Tucson, Arizona.

Faires, J. D., and Burden, R. L. (1993), Numerical Methods, PWS-Kent Publishing Company, Boston, USA.

Gilbert, R.O. (1987), *Statistical Methods for Environmental Pollution Monitoring*, New York: Van Nostrand Reinhold.

Grice, J.V., and Bain, L. J. (1980), Inferences Concerning the Mean of the Gamma Distribution. Journal of the American Statistical Association. Vol 75, Number 372, pp 929-933.

Hall, P. (1988), Theoretical comparison of bootstrap confidence intervals; Annals of Statistics, 16, 927-953.

Hall, P. (1992), On the Removal of Skewness by Transformation. Journal of Royal Statistical Society, B 54, 221-228.

Hardin, J.W., and Gilbert, R.O. (1993), "Comparing Statistical Tests for Detecting Soil Contamination Greater Than Background," Pacific Northwest Laboratory, Battelle, Technical Report # DE 94-005498.

Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1983), *Understanding Robust and Exploratory Data Analysis*. John Wiley, New York.

Hogg, R.V., and Craig, A.T. (1978), *Introduction to Mathematical Statistics*, New York: Macmillan Publishing Company.

Johnson, N.J. (1978), "Modified t-Tests and Confidence Intervals for Asymmetrical Populations," *The American Statistician*, Vol. 73, pp.536-544.

Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994), Continuous Univariate Distributions, Volume 1. Second Edition. John Wiley.

Kleijnen, J.P.C., Kloppenburg, G.L.J., and Meeuwsen, F.L. (1986), "Testing the Mean of an Asymmetric Population: Johnson's Modified t Test Revisited." *Commun. in Statist.-Simula.,* 15(3), 715-731.

Land, C. E. (1971), "Confidence Intervals for Linear Functions of the Normal Mean and Variance," *Annals of Mathematical Statistics*, 42, 1187-1205.

Land, C. E. (1975), "Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance," in *Selected Tables in Mathematical Statistics*, Vol. III, American Mathematical Society, Providence, R.I., 385-419.

Law, A.M., and Kelton, W.D. (2000), Simulation Modeling and Analysis. Third Edition. McGraw Hill.

Manly, B.F.J. (1997), Randomization, Bootstrap, and Monte Carlo Methods in Biology. Second Edition. Chapman Hall, London.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1990). Numerical Recipes in C, The Art of Scientific Computing. Cambridge University Press. Cambridge, MA.

Schneider, B. E. (1978), Kolmogorov-Smirnov Test Statistic for the Gamma Distribution with Unknown Parameters, Dissertation, Department of Statistics, Temple University, Philadelphia, Pa.

Schneider, B.E., and Clickner, R.P. (1976). On the distribution of the Kolmogorov-Smirnov Statistic for the Gamma Distribution with Unknown Parameters, Mimeo Series Number 36, Department of Statistics, School of Business Administration, Temple University, Philadelphia, Pa.

Schulz, T. W., and Griffin, S. (1999), Estimating Risk Assessment Exposure Point Concentrations when Data are Not Normal or Lognormal. Risk Analysis, Vol. 19, No. 4, 1999.

Scout: A Data Analysis Program, Technology Support Project. EPA, NERL -LV, Las Vegas, NV 89193-3478.

Singh, A. (1993), "Omnibus Robust Procedures For Assessment of Multivariate Normality and Detection of Multivariate Outliers," *Multivariate Environmental Statistics*. G. P. Patil and C.R. Rao, Editors, Elsevier Science Publishers.

Singh, A. K., Singh, A., and Engelhardt, M., "The Lognormal Distribution in Environmental Applications," EPA/600/R-97/006, December 1997.

Singh, A. K., Singh, A., and Engelhardt, M., "Some Practical Aspects of Sample Size and Power Computations for Estimating the Mean of Positively Skewed Distributions in Environmental Applications," EPA/600/S-99/006, November 1999.

Singh. A., Singh, A. K., Engelhardt, M., and Nocerino, J.M. (2002a), " On the Computation of the Upper Confidence Limit of the Mean of Contaminant Data Distributions." Under EPA Review.

Singh, A., Singh, A. K., and Iaci, R. J. (2002b). " Estimation of the Exposure Point Concentration Term Using a Gamma Distribution." EPA/600/R-02/084.

Singh, A. and Singh, A.K. (2003). Estimation of the Exposure Point Concentration Term (95% UCL) Using Bias-Corrected Accelerated (BCA) Bootstrap Method and Several other methods for Normal, Lognormal, and Gamma Distributions. Draft EPA Internal Report.

Stephens, M. A. (1970), Use of Kolmogorov-Smirnov, Cramer-von Mises and Related Statistics Without Extensive Tables. Journal of Royal Statistical Society, B 32, 115-122.

Sutton, C.D. (1993), "Computer -Intensive Methods for Tests About the Mean Of an Asymmetrical Distribution," *Journal Of American Statistical Society,* Vol. 88, No. 423, pp 802-810.

Thom, H.C. S. (1968), Direct and Inverse Tables of the Gamma Distribution, Silver Spring, MD; Environmental Data Service.

Whittaker, J. (1974), Generating Gamma and Beta Random Variables with Non-integral Shape Parameters. Applied Statistics, 23, No. 2, 210-214.

Wong, A. (1993), A Note on Inference for the Mean Parameter of the Gamma Distribution. Statistics Probability Letters, Vol 17, 61-66.

APPENDIX B

CRITICAL VALUES

OF

ANDERSON-DARLING TEST STATISTIC

AND

KOLMOGOROV-SMIRNOV TEST STATISTIC

FOR

GAMMA DISTRIBUTION

WITH UNKNOWN PARAMETERS

Critical Values for Anderson Darling Test - Significance Level of 0.20

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.6012	0.5867	0.5709	0.5498	0.5169	0.5017	0.4900	0.4854	0.4839	0.4819	0.4810	0.4805	0.4802	0.4803	0.4795	0.4795	0.4791	0.4790
5	0.6366	0.6085	0.5796	0.5590	0.5322	0.5166	0.5049	0.4996	0.4969	0.4949	0.4930	0.4926	0.4914	0.4919	0.4908	0.4903	0.4899	0.4901
6	0.6851	0.6362	0.5915	0.5682	0.5431	0.5264	0.5117	0.5055	0.5026	0.4996	0.4981	0.4969	0.4964	0.4960	0.4955	0.4950	0.4948	0.4951
7	0.7349	0.6671	0.6037	0.5745	0.5491	0.5318	0.5172	0.5102	0.5064	0.5036	0.5007	0.5002	0.4988	0.4991	0.4984	0.4975	0.4973	0.4973
8	0.7856	0.6966	0.6150	0.5784	0.5545	0.5372	0.5210	0.5134	0.5092	0.5058	0.5039	0.5025	0.5016	0.5017	0.4998	0.5000	0.4992	0.4990
9	0.8385	0.7291	0.6265	0.5827	0.5593	0.5407	0.5239	0.5162	0.5122	0.5082	0.5068	0.5045	0.5035	0.5024	0.5015	0.5015	0.5010	0.5014
10	0.8923	0.7600	0.6384	0.5849	0.5626	0.5436	0.5263	0.5170	0.5135	0.5097	0.5079	0.5063	0.5050	0.5041	0.5029	0.5023	0.5020	0.5020
11	0.9469	0.7926	0.6496	0.5881	0.5662	0.5463	0.5287	0.5198	0.5147	0.5110	0.5088	0.5061	0.5049	0.5048	0.5041	0.5035	0.5030	0.5024
12	1.0021	0.8247	0.6600	0.5900	0.5680	0.5485	0.5299	0.5213	0.5166	0.5113	0.5098	0.5080	0.5058	0.5050	0.5042	0.5048	0.5036	0.5041
13	1.0571	0.8571	0.6731	0.5910	0.5697	0.5499	0.5317	0.5224	0.5169	0.5134	0.5101	0.5080	0.5073	0.5064	0.5053	0.5049	0.5050	0.5047
14	1.1106	0.8897	0.6828	0.5928	0.5716	0.5508	0.5330	0.5229	0.5184	0.5131	0.5111	0.5090	0.5080	0.5072	0.5054	0.5051	0.5040	0.5045
15	1.1656	0.9221	0.6926	0.5951	0.5735	0.5525	0.5331	0.5238	0.5188	0.5134	0.5115	0.5095	0.5078	0.5073	0.5058	0.5051	0.5054	0.5051
16	1.2201	0.9542	0.7047	0.5967	0.5744	0.5535	0.5345	0.5242	0.5197	0.5143	0.5127	0.5095	0.5081	0.5082	0.5068	0.5057	0.5052	0.5054
17	1.2747	0.9856	0.7157	0.5975	0.5764	0.5553	0.5354	0.5249	0.5200	0.5152	0.5122	0.5099	0.5086	0.5085	0.5066	0.5063	0.5053	0.5055
18	1.3270	1.0187	0.7261	0.5994	0.5761	0.5556	0.5357	0.5247	0.5203	0.5151	0.5132	0.5107	0.5097	0.5090	0.5067	0.5066	0.5058	0.5063
19	1.3799	1.0502	0.7376	0.6000	0.5775	0.5563	0.5367	0.5257	0.5208	0.5155	0.5127	0.5107	0.5090	0.5080	0.5074	0.5069	0.5067	0.5057
20	1.4316	1.0812	0.7470	0.6016	0.5779	0.5567	0.5369	0.5264	0.5210	0.5159	0.5135	0.5103	0.5091	0.5090	0.5082	0.5066	0.5069	0.5069
21	1.4859	1.1119	0.7574	0.6022	0.5788	0.5569	0.5386	0.5271	0.5209	0.5160	0.5137	0.5112	0.5098	0.5092	0.5081	0.5077	0.5071	0.5071
22	1.5373	1.1433	0.7681	0.6037	0.5793	0.5584	0.5377	0.5277	0.5220	0.5160	0.5135	0.5116	0.5101	0.5093	0.5083	0.5069	0.5072	0.5064
23	1.5882	1.1774	0.7794	0.6042	0.5803	0.5589	0.5380	0.5275	0.5213	0.5166	0.5134	0.5110	0.5108	0.5097	0.5081	0.5069	0.5069	0.5070
24	1.6410	1.2064	0.7890	0.6046	0.5807	0.5595	0.5386	0.5272	0.5225	0.5173	0.5139	0.5117	0.5097	0.5093	0.5082	0.5076	0.5074	0.5072
25	1.6915	1.2376	0.8002	0.6057	0.5806	0.5601	0.5391	0.5278	0.5229	0.5169	0.5144	0.5119	0.5104	0.5095	0.5082	0.5074	0.5070	0.5071
26	1.7433	1.2691	0.8100	0.6069	0.5809	0.5601	0.5395	0.5279	0.5223	0.5170	0.5140	0.5113	0.5099	0.5098	0.5082	0.5073	0.5072	0.5076
27	1.7932	1.2981	0.8228	0.6081	0.5816	0.5608	0.5390	0.5287	0.5233	0.5171	0.5150	0.5120	0.5106	0.5097	0.5077	0.5080	0.5077	0.5073
28	1.8431	1.3284	0.8319	0.6088	0.5818	0.5610	0.5397	0.5283	0.5228	0.5170	0.5153	0.5118	0.5112	0.5100	0.5081	0.5085	0.5078	0.5073
29	1.8948	1.3600	0.8424	0.6099	0.5818	0.5613	0.5402	0.5287	0.5235	0.5175	0.5149	0.5124	0.5110	0.5097	0.5082	0.5076	0.5075	0.5074
30	1.9433	1.3895	0.8532	0.6110	0.5825	0.5617	0.5397	0.5292	0.5230	0.5176	0.5151	0.5126	0.5099	0.5097	0.5089	0.5079	0.5072	0.5081
35	2.1902	1.5371	0.9057	0.6147	0.5843	0.5626	0.5414	0.5300	0.5237	0.5178	0.5156	0.5126	0.5123	0.5105	0.5090	0.5087	0.5082	0.5074
40	2.4320	1.6829	0.9551	0.6174	0.5848	0.5630	0.5418	0.5299	0.5246	0.5183	0.5153	0.5128	0.5110	0.5108	0.5094	0.5083	0.5075	0.5083
45	2.6734	1.8275	1.0046	0.6211	0.5857	0.5646	0.5418	0.5301	0.5244	0.5191	0.5160	0.5130	0.5111	0.5110	0.5094	0.5085	0.5084	0.5083
50	2.9056	1.9669	1.0536	0.6238	0.5872	0.5651	0.5413	0.5313	0.5251	0.5192	0.5162	0.5132	0.5116	0.5111	0.5095	0.5088	0.5087	0.5089
60	3.3680	2.2458	1.1502	0.6309	0.5878	0.5655	0.5430	0.5311	0.5248	0.5189	0.5165	0.5141	0.5113	0.5112	0.5099	0.5084	0.5089	0.5089
70	3.8261	2.5178	1.2478	0.6361	0.5882	0.5667	0.5433	0.5310	0.5252	0.5194	0.5165	0.5132	0.5122	0.5112	0.5098	0.5091	0.5090	0.5081
80	4.2729	2.7850	1.3430	0.6424	0.5889	0.5669	0.5439	0.5314	0.5258	0.5201	0.5173	0.5130	0.5131	0.5110	0.5100	0.5087	0.5089	0.5086
90	4.7189	3.0528	1.4370	0.6474	0.5883	0.5670	0.5438	0.5321	0.5256	0.5203	0.5174	0.5139	0.5124	0.5109	0.5101	0.5088	0.5087	0.5091
100	5.1658	3.3136	1.5320	0.6516	0.5886	0.5681	0.5438	0.5318	0.5260	0.5200	0.5174	0.5140	0.5117	0.5115	0.5101	0.5090	0.5092	0.5089
200	9.4620	5.8551	2.4199	0.7059	0.5910	0.5675	0.5452	0.5325	0.5264	0.5199	0.5172	0.5140	0.5122	0.5115	0.5095	0.5095	0.5090	0.5093
300	13.6454	8.3200	3.2731	0.7595	0.5915	0.5688	0.5448	0.5328	0.5260	0.5205	0.5174	0.5134	0.5126	0.5120	0.5107	0.5091	0.5092	0.5092
400	17.7759	10.7341	4.1071	0.8119	0.5902	0.5688	0.5448	0.5331	0.5266	0.5200	0.5168	0.5143	0.5127	0.5125	0.5107	0.5095	0.5090	0.5093
500	21.8687	13.1245	4.9232	0.8646	0.5910	0.5685	0.5450	0.5332	0.5267	0.5203	0.5173	0.5145	0.5129	0.5123	0.5102	0.5092	0.5094	0.5097
1000	42.0423	24.8700	8.9004	1.1234	0.5917	0.5687	0.5457	0.5327	0.5265	0.5204	0.5174	0.5143	0.5126	0.5118	0.5098	0.5098	0.5091	0.5096
2500	101.548	59.3470	20.4324	1.8628	0.5930	0.5698	0.5460	0.5336	0.5268	0.5206	0.5178	0.5143	0.5155	0.5129	0.5102	0.5093	0.5087	0.5095

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.20

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.3745	0.3681	0.3610	0.3538	0.3419	0.3360	0.3314	0.3293	0.3285	0.3275	0.3270	0.3266	0.3263	0.3263	0.3258	0.3258	0.3257	0.3256
5	0.3495	0.3407	0.3315	0.3276	0.3228	0.3179	0.3128	0.3093	0.3074	0.3055	0.3043	0.3036	0.3029	0.3026	0.3019	0.3015	0.3014	0.3014
6	0.3350	0.3220	0.3102	0.3048	0.2990	0.2942	0.2889	0.2856	0.2839	0.2822	0.2812	0.2804	0.2800	0.2795	0.2792	0.2788	0.2788	0.2787
7	0.3207	0.3062	0.2918	0.2848	0.2792	0.2745	0.2695	0.2666	0.2649	0.2631	0.2620	0.2613	0.2608	0.2606	0.2601	0.2598	0.2597	0.2596
8	0.3105	0.2932	0.2759	0.2683	0.2641	0.2598	0.2547	0.2516	0.2498	0.2480	0.2471	0.2462	0.2458	0.2456	0.2449	0.2446	0.2444	0.2444
9	0.3014	0.2831	0.2641	0.2553	0.2510	0.2468	0.2419	0.2389	0.2372	0.2354	0.2346	0.2336	0.2332	0.2327	0.2323	0.2321	0.2319	0.2319
10	0.2937	0.2738	0.2533	0.2436	0.2394	0.2352	0.2307	0.2276	0.2262	0.2244	0.2236	0.2228	0.2223	0.2220	0.2214	0.2211	0.2211	0.2209
11	0.2869	0.2660	0.2440	0.2333	0.2296	0.2255	0.2209	0.2182	0.2165	0.2149	0.2141	0.2132	0.2126	0.2124	0.2120	0.2117	0.2115	0.2115
12	0.2811	0.2592	0.2355	0.2243	0.2206	0.2168	0.2123	0.2097	0.2082	0.2064	0.2057	0.2048	0.2042	0.2040	0.2036	0.2035	0.2032	0.2033
13	0.2757	0.2531	0.2285	0.2162	0.2127	0.2091	0.2047	0.2022	0.2006	0.1991	0.1981	0.1973	0.1970	0.1967	0.1961	0.1960	0.1959	0.1958
14	0.2710	0.2478	0.2220	0.2091	0.2056	0.2020	0.1980	0.1954	0.1940	0.1922	0.1915	0.1907	0.1903	0.1900	0.1895	0.1893	0.1891	0.1891
15	0.2665	0.2427	0.2159	0.2026	0.1993	0.1958	0.1916	0.1893	0.1877	0.1862	0.1854	0.1847	0.1842	0.1840	0.1834	0.1832	0.1832	0.1831
16	0.2625	0.2383	0.2107	0.1966	0.1933	0.1900	0.1862	0.1836	0.1822	0.1807	0.1800	0.1792	0.1787	0.1785	0.1782	0.1779	0.1777	0.1777
17	0.2587	0.2341	0.2059	0.1912	0.1881	0.1850	0.1810	0.1785	0.1772	0.1756	0.1749	0.1741	0.1738	0.1736	0.1731	0.1729	0.1727	0.1727
18	0.2553	0.2304	0.2014	0.1863	0.1831	0.1799	0.1762	0.1737	0.1724	0.1710	0.1704	0.1696	0.1692	0.1690	0.1684	0.1683	0.1681	0.1681
19	0.2519	0.2267	0.1975	0.1816	0.1786	0.1754	0.1719	0.1694	0.1681	0.1668	0.1659	0.1653	0.1649	0.1646	0.1643	0.1641	0.1640	0.1639
20	0.2489	0.2236	0.1935	0.1774	0.1743	0.1713	0.1677	0.1654	0.1641	0.1628	0.1621	0.1613	0.1609	0.1608	0.1603	0.1601	0.1600	0.1600
21	0.2463	0.2205	0.1899	0.1734	0.1704	0.1673	0.1639	0.1617	0.1604	0.1590	0.1584	0.1576	0.1573	0.1571	0.1568	0.1565	0.1564	0.1564
22	0.2437	0.2176	0.1867	0.1697	0.1667	0.1639	0.1604	0.1582	0.1569	0.1555	0.1550	0.1543	0.1539	0.1537	0.1532	0.1531	0.1530	0.1529
23	0.2412	0.2151	0.1837	0.1661	0.1634	0.1604	0.1570	0.1549	0.1536	0.1524	0.1517	0.1509	0.1507	0.1505	0.1502	0.1498	0.1498	0.1498
24	0.2389	0.2124	0.1808	0.1629	0.1600	0.1573	0.1539	0.1518	0.1506	0.1494	0.1487	0.1480	0.1477	0.1475	0.1470	0.1469	0.1468	0.1467
25	0.2366	0.2101	0.1782	0.1598	0.1570	0.1542	0.1510	0.1488	0.1478	0.1465	0.1459	0.1452	0.1449	0.1446	0.1443	0.1441	0.1440	0.1439
26	0.2346	0.2080	0.1756	0.1569	0.1541	0.1513	0.1482	0.1462	0.1449	0.1437	0.1432	0.1424	0.1422	0.1419	0.1416	0.1414	0.1413	0.1412
27	0.2325	0.2058	0.1735	0.1542	0.1513	0.1487	0.1455	0.1436	0.1425	0.1412	0.1406	0.1400	0.1395	0.1394	0.1390	0.1389	0.1388	0.1388
28	0.2308	0.2038	0.1710	0.1515	0.1488	0.1462	0.1431	0.1411	0.1399	0.1388	0.1382	0.1376	0.1373	0.1371	0.1367	0.1366	0.1365	0.1364
29	0.2289	0.2018	0.1689	0.1491	0.1462	0.1439	0.1407	0.1388	0.1377	0.1365	0.1359	0.1353	0.1349	0.1347	0.1343	0.1342	0.1341	0.1341
30	0.2272	0.2000	0.1669	0.1468	0.1439	0.1414	0.1384	0.1364	0.1355	0.1343	0.1337	0.1331	0.1328	0.1325	0.1323	0.1321	0.1320	0.1320
35	0.2197	0.1921	0.1581	0.1366	0.1337	0.1314	0.1286	0.1268	0.1258	0.1248	0.1243	0.1236	0.1234	0.1231	0.1228	0.1228	0.1226	0.1226
40	0.2136	0.1857	0.1509	0.1282	0.1255	0.1232	0.1206	0.1190	0.1181	0.1170	0.1165	0.1160	0.1156	0.1155	0.1152	0.1151	0.1150	0.1150
45	0.2084	0.1803	0.1449	0.1214	0.1185	0.1166	0.1140	0.1125	0.1116	0.1106	0.1101	0.1096	0.1093	0.1091	0.1089	0.1087	0.1087	0.1086
50	0.2040	0.1756	0.1400	0.1155	0.1128	0.1108	0.1083	0.1070	0.1060	0.1051	0.1047	0.1042	0.1039	0.1038	0.1035	0.1033	0.1032	0.1033
60	0.1970	0.1682	0.1319	0.1060	0.1032	0.1014	0.0992	0.0979	0.0971	0.0962	0.0958	0.0954	0.0951	0.0950	0.0948	0.0946	0.0945	0.0945
70	0.1915	0.1623	0.1257	0.0987	0.0958	0.0942	0.0921	0.0908	0.0901	0.0893	0.0889	0.0885	0.0883	0.0882	0.0879	0.0878	0.0877	0.0877
80	0.1870	0.1576	0.1207	0.0927	0.0898	0.0882	0.0863	0.0851	0.0844	0.0837	0.0833	0.0829	0.0827	0.0826	0.0824	0.0822	0.0822	0.0822
90	0.1832	0.1538	0.1166	0.0877	0.0847	0.0833	0.0815	0.0804	0.0797	0.0790	0.0787	0.0783	0.0781	0.0780	0.0778	0.0777	0.0776	0.0776
100	0.1801	0.1504	0.1131	0.0835	0.0805	0.0792	0.0774	0.0763	0.0758	0.0751	0.0748	0.0744	0.0741	0.0741	0.0739	0.0738	0.0737	0.0737
200	0.1630	0.1325	0.0940	0.0611	0.0573	0.0563	0.0551	0.0544	0.0539	0.0534	0.0532	0.0529	0.0528	0.0527	0.0526	0.0525	0.0525	0.0525
300	0.1554	0.1247	0.0857	0.0513	0.0469	0.0461	0.0451	0.0445	0.0442	0.0438	0.0435	0.0433	0.0433	0.0432	0.0431	0.0430	0.0430	0.0430
400	0.1510	0.1200	0.0807	0.0455	0.0407	0.0400	0.0392	0.0386	0.0383	0.0379	0.0378	0.0376	0.0375	0.0375	0.0374	0.0373	0.0373	0.0373
500	0.1480	0.1169	0.0773	0.0416	0.0364	0.0358	0.0351	0.0346	0.0343	0.0340	0.0338	0.0337	0.0336	0.0336	0.0335	0.0334	0.0334	0.0334
1000	0.1407	0.1093	0.0692	0.0323	0.0258	0.0254	0.0249	0.0245	0.0243	0.0241	0.0240	0.0239	0.0238	0.0238	0.0237	0.0237	0.0237	0.0237
2500	0.1344	0.1027	0.0621	0.0242	0.0164	0.0161	0.0158	0.0156	0.0154	0.0153	0.0152	0.0151	0.0151	0.0151	0.0151	0.0150	0.0150	0.0150

Critical Values for Anderson Darling Test - Significance Level of 0.15

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.6495	0.6354	0.6212	0.5995	0.5626	0.5456	0.5321	0.5268	0.5252	0.5226	0.5217	0.5206	0.5203	0.5208	0.5202	0.5197	0.5193	0.5190
5	0.6893	0.6597	0.6317	0.6137	0.5836	0.5649	0.5505	0.5436	0.5404	0.5377	0.5357	0.5352	0.5339	0.5341	0.5328	0.5321	0.5319	0.5320
6	0.7453	0.6944	0.6484	0.6262	0.5967	0.5765	0.5591	0.5509	0.5476	0.5441	0.5419	0.5406	0.5401	0.5394	0.5391	0.5382	0.5380	0.5382
7	0.8015	0.7290	0.6625	0.6337	0.6049	0.5838	0.5667	0.5581	0.5530	0.5496	0.5460	0.5460	0.5441	0.5443	0.5436	0.5427	0.5422	0.5421
8	0.8594	0.7632	0.6757	0.6393	0.6124	0.5912	0.5711	0.5622	0.5574	0.5532	0.5504	0.5490	0.5477	0.5480	0.5460	0.5456	0.5450	0.5452
9	0.9189	0.8002	0.6896	0.6442	0.6176	0.5950	0.5754	0.5658	0.5608	0.5561	0.5542	0.5517	0.5505	0.5493	0.5478	0.5480	0.5473	0.5480
10	0.9786	0.8354	0.7026	0.6475	0.6222	0.5995	0.5786	0.5673	0.5629	0.5578	0.5559	0.5538	0.5524	0.5517	0.5498	0.5492	0.5486	0.5489
11	1.0392	0.8719	0.7163	0.6508	0.6266	0.6025	0.5813	0.5709	0.5644	0.5597	0.5574	0.5544	0.5530	0.5525	0.5515	0.5511	0.5505	0.5498
12	1.0998	0.9079	0.7288	0.6534	0.6290	0.6050	0.5826	0.5726	0.5673	0.5605	0.5587	0.5568	0.5545	0.5531	0.5523	0.5527	0.5515	0.5520
13	1.1601	0.9445	0.7437	0.6556	0.6309	0.6077	0.5850	0.5742	0.5674	0.5631	0.5590	0.5564	0.5559	0.5548	0.5534	0.5531	0.5529	0.5526
14	1.2198	0.9815	0.7543	0.6579	0.6332	0.6084	0.5870	0.5746	0.5693	0.5630	0.5602	0.5580	0.5565	0.5558	0.5535	0.5533	0.5521	0.5526
15	1.2789	1.0165	0.7658	0.6600	0.6354	0.6105	0.5870	0.5759	0.5699	0.5637	0.5609	0.5586	0.5567	0.5562	0.5547	0.5540	0.5540	0.5534
16	1.3374	1.0527	0.7796	0.6618	0.6364	0.6118	0.5892	0.5762	0.5707	0.5642	0.5625	0.5584	0.5571	0.5573	0.5557	0.5545	0.5535	0.5539
17	1.3967	1.0875	0.7922	0.6631	0.6388	0.6140	0.5898	0.5770	0.5712	0.5658	0.5621	0.5592	0.5578	0.5576	0.5554	0.5550	0.5540	0.5542
18	1.4533	1.1240	0.8037	0.6659	0.6392	0.6142	0.5905	0.5773	0.5715	0.5657	0.5635	0.5599	0.5588	0.5578	0.5555	0.5554	0.5547	0.5549
19	1.5098	1.1576	0.8169	0.6665	0.6405	0.6155	0.5919	0.5783	0.5726	0.5660	0.5626	0.5605	0.5584	0.5573	0.5567	0.5562	0.5555	0.5546
20	1.5661	1.1928	0.8279	0.6685	0.6413	0.6161	0.5921	0.5790	0.5732	0.5668	0.5635	0.5604	0.5588	0.5585	0.5570	0.5559	0.5556	0.5557
21	1.6235	1.2257	0.8396	0.6691	0.6420	0.6160	0.5937	0.5804	0.5728	0.5668	0.5641	0.5611	0.5594	0.5584	0.5573	0.5570	0.5560	0.5560
22	1.6779	1.2584	0.8514	0.6704	0.6431	0.6175	0.5932	0.5806	0.5735	0.5669	0.5646	0.5614	0.5598	0.5594	0.5577	0.5561	0.5565	0.5557
23	1.7323	1.2970	0.8644	0.6716	0.6440	0.6186	0.5935	0.5807	0.5731	0.5676	0.5637	0.5611	0.5608	0.5592	0.5575	0.5562	0.5561	0.5560
24	1.7885	1.3279	0.8745	0.6727	0.6444	0.6192	0.5944	0.5806	0.5739	0.5683	0.5643	0.5618	0.5596	0.5590	0.5582	0.5569	0.5562	0.5565
25	1.8422	1.3607	0.8871	0.6737	0.6453	0.6196	0.5948	0.5813	0.5751	0.5677	0.5652	0.5619	0.5605	0.5595	0.5581	0.5568	0.5565	0.5565
26	1.8963	1.3958	0.8982	0.6745	0.6449	0.6193	0.5950	0.5817	0.5744	0.5681	0.5649	0.5614	0.5598	0.5596	0.5575	0.5568	0.5567	0.5567
27	1.9503	1.4261	0.9129	0.6765	0.6455	0.6208	0.5944	0.5816	0.5756	0.5684	0.5656	0.5623	0.5602	0.5591	0.5575	0.5574	0.5564	0.5568
28	2.0036	1.4603	0.9224	0.6766	0.6461	0.6213	0.5955	0.5817	0.5751	0.5681	0.5663	0.5623	0.5613	0.5601	0.5579	0.5576	0.5570	0.5563
29	2.0588	1.4943	0.9338	0.6782	0.6457	0.6217	0.5965	0.5818	0.5755	0.5690	0.5653	0.5629	0.5607	0.5597	0.5575	0.5569	0.5566	0.5569
30	2.1110	1.5255	0.9463	0.6801	0.6465	0.6216	0.5955	0.5826	0.5758	0.5689	0.5661	0.5629	0.5599	0.5593	0.5584	0.5578	0.5566	0.5574
35	2.3678	1.6835	1.0038	0.6836	0.6483	0.6230	0.5974	0.5835	0.5764	0.5696	0.5667	0.5633	0.5625	0.5606	0.5591	0.5584	0.5576	0.5576
40	2.6243	1.8376	1.0582	0.6870	0.6498	0.6232	0.5979	0.5837	0.5773	0.5701	0.5662	0.5636	0.5612	0.5608	0.5593	0.5579	0.5575	0.5580
45	2.8741	1.9901	1.1118	0.6917	0.6505	0.6253	0.5979	0.5839	0.5768	0.5706	0.5669	0.5637	0.5621	0.5612	0.5599	0.5586	0.5585	0.5584
50	3.1177	2.1386	1.1654	0.6950	0.6527	0.6258	0.5976	0.5860	0.5778	0.5710	0.5676	0.5641	0.5619	0.5616	0.5599	0.5588	0.5588	0.5586
60	3.5997	2.4304	1.2695	0.7029	0.6530	0.6262	0.5995	0.5851	0.5780	0.5709	0.5673	0.5645	0.5618	0.5616	0.5605	0.5589	0.5589	0.5589
70	4.0720	2.7155	1.3751	0.7081	0.6538	0.6281	0.5996	0.5850	0.5781	0.5716	0.5684	0.5643	0.5629	0.5615	0.5600	0.5592	0.5593	0.5588
80	4.5375	2.9941	1.4768	0.7162	0.6539	0.6273	0.6005	0.5858	0.5786	0.5726	0.5690	0.5641	0.5637	0.5616	0.5602	0.5589	0.5588	0.5589
90	4.9957	3.2729	1.5758	0.7212	0.6536	0.6283	0.6001	0.5863	0.5789	0.5724	0.5691	0.5651	0.5631	0.5618	0.5602	0.5590	0.5594	0.5593
100	5.4567	3.5445	1.6772	0.7269	0.6549	0.6299	0.6005	0.5865	0.5793	0.5723	0.5693	0.5651	0.5628	0.5622	0.5608	0.5595	0.5590	0.5592
200	9.8591	6.1657	2.6088	0.7864	0.6568	0.6291	0.6020	0.5870	0.5796	0.5720	0.5690	0.5656	0.5634	0.5622	0.5600	0.5598	0.5591	0.5597
300	14.1248	8.6896	3.4931	0.8459	0.6577	0.6301	0.6019	0.5873	0.5795	0.5729	0.5685	0.5646	0.5637	0.5632	0.5616	0.5593	0.5599	0.5602
400	18.3207	11.1508	4.3546	0.9029	0.6562	0.6306	0.6017	0.5880	0.5798	0.5725	0.5684	0.5657	0.5638	0.5636	0.5615	0.5601	0.5596	0.5602
500	22.4788	13.5882	5.1945	0.9597	0.6575	0.6301	0.6021	0.5878	0.5804	0.5729	0.5698	0.5660	0.5642	0.5632	0.5611	0.5601	0.5600	0.5602
1000	42.8884	25.5062	9.2649	1.2387	0.6576	0.6303	0.6032	0.5874	0.5798	0.5726	0.5696	0.5652	0.5642	0.5631	0.5607	0.5604	0.5597	0.5603
2500	102.850	60.3279	20.9754	2.0188	0.6594	0.6314	0.6028	0.5884	0.5806	0.5726	0.5697	0.5658	0.5674	0.5643	0.5613	0.5597	0.5593	0.5605

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.15

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.3901	0.3832	0.3761	0.3698	0.3599	0.3533	0.3462	0.3417	0.3401	0.3385	0.3379	0.3373	0.3369	0.3369	0.3364	0.3363	0.3362	0.3362
5	0.3646	0.3559	0.3475	0.3445	0.3389	0.3336	0.3279	0.3240	0.3220	0.3200	0.3188	0.3180	0.3172	0.3171	0.3163	0.3159	0.3158	0.3156
6	0.3507	0.3378	0.3254	0.3199	0.3133	0.3078	0.3018	0.2983	0.2964	0.2948	0.2936	0.2928	0.2922	0.2917	0.2914	0.2910	0.2909	0.2909
7	0.3358	0.3203	0.3055	0.2988	0.2934	0.2884	0.2828	0.2794	0.2773	0.2753	0.2742	0.2733	0.2728	0.2726	0.2719	0.2716	0.2715	0.2713
8	0.3254	0.3078	0.2899	0.2825	0.2781	0.2732	0.2673	0.2639	0.2620	0.2599	0.2588	0.2579	0.2574	0.2571	0.2564	0.2560	0.2557	0.2558
9	0.3158	0.2969	0.2773	0.2686	0.2639	0.2592	0.2539	0.2504	0.2486	0.2467	0.2458	0.2447	0.2442	0.2436	0.2432	0.2429	0.2428	0.2428
10	0.3077	0.2873	0.2659	0.2561	0.2518	0.2472	0.2421	0.2386	0.2371	0.2351	0.2343	0.2334	0.2329	0.2325	0.2318	0.2315	0.2314	0.2313
11	0.3006	0.2792	0.2564	0.2453	0.2415	0.2371	0.2320	0.2290	0.2270	0.2251	0.2244	0.2234	0.2227	0.2225	0.2220	0.2218	0.2215	0.2215
12	0.2944	0.2721	0.2475	0.2360	0.2322	0.2280	0.2229	0.2201	0.2183	0.2163	0.2155	0.2147	0.2140	0.2137	0.2131	0.2132	0.2128	0.2129
13	0.2887	0.2657	0.2402	0.2276	0.2239	0.2198	0.2151	0.2121	0.2104	0.2087	0.2077	0.2067	0.2064	0.2061	0.2054	0.2052	0.2052	0.2050
14	0.2835	0.2600	0.2333	0.2200	0.2163	0.2124	0.2080	0.2050	0.2034	0.2015	0.2006	0.1998	0.1994	0.1991	0.1985	0.1983	0.1981	0.1980
15	0.2787	0.2547	0.2270	0.2132	0.2097	0.2058	0.2013	0.1987	0.1969	0.1951	0.1943	0.1936	0.1930	0.1928	0.1922	0.1919	0.1919	0.1918
16	0.2745	0.2501	0.2216	0.2070	0.2035	0.1998	0.1955	0.1926	0.1912	0.1895	0.1887	0.1877	0.1873	0.1871	0.1867	0.1863	0.1861	0.1861
17	0.2704	0.2455	0.2165	0.2012	0.1980	0.1945	0.1901	0.1874	0.1859	0.1841	0.1834	0.1825	0.1821	0.1820	0.1814	0.1811	0.1809	0.1809
18	0.2667	0.2416	0.2118	0.1961	0.1926	0.1892	0.1850	0.1824	0.1809	0.1793	0.1786	0.1778	0.1774	0.1771	0.1764	0.1763	0.1762	0.1762
19	0.2632	0.2376	0.2077	0.1912	0.1880	0.1844	0.1806	0.1778	0.1764	0.1748	0.1739	0.1733	0.1728	0.1725	0.1721	0.1719	0.1718	0.1717
20	0.2599	0.2344	0.2036	0.1868	0.1834	0.1802	0.1761	0.1736	0.1722	0.1707	0.1699	0.1691	0.1686	0.1685	0.1680	0.1678	0.1678	0.1677
21	0.2570	0.2309	0.1998	0.1825	0.1794	0.1759	0.1723	0.1698	0.1683	0.1668	0.1661	0.1653	0.1649	0.1646	0.1643	0.1640	0.1638	0.1639
22	0.2542	0.2279	0.1964	0.1786	0.1756	0.1724	0.1685	0.1661	0.1647	0.1631	0.1625	0.1618	0.1613	0.1611	0.1605	0.1604	0.1603	0.1601
23	0.2516	0.2253	0.1933	0.1749	0.1719	0.1687	0.1650	0.1626	0.1612	0.1598	0.1591	0.1583	0.1580	0.1577	0.1574	0.1570	0.1569	0.1569
24	0.2491	0.2225	0.1901	0.1715	0.1685	0.1654	0.1617	0.1593	0.1580	0.1566	0.1559	0.1552	0.1548	0.1545	0.1541	0.1539	0.1538	0.1537
25	0.2466	0.2200	0.1873	0.1683	0.1652	0.1623	0.1586	0.1563	0.1551	0.1537	0.1529	0.1522	0.1518	0.1516	0.1512	0.1510	0.1509	0.1509
26	0.2445	0.2177	0.1846	0.1652	0.1622	0.1591	0.1557	0.1534	0.1521	0.1507	0.1501	0.1493	0.1490	0.1487	0.1484	0.1482	0.1480	0.1480
27	0.2423	0.2152	0.1824	0.1624	0.1593	0.1565	0.1528	0.1507	0.1495	0.1482	0.1474	0.1467	0.1462	0.1462	0.1457	0.1455	0.1454	0.1454
28	0.2404	0.2132	0.1798	0.1596	0.1566	0.1538	0.1504	0.1481	0.1469	0.1455	0.1449	0.1442	0.1439	0.1436	0.1432	0.1431	0.1430	0.1429
29	0.2383	0.2111	0.1776	0.1570	0.1539	0.1514	0.1479	0.1456	0.1446	0.1431	0.1425	0.1418	0.1414	0.1412	0.1407	0.1407	0.1406	0.1405
30	0.2365	0.2092	0.1755	0.1545	0.1515	0.1488	0.1454	0.1433	0.1421	0.1408	0.1402	0.1396	0.1391	0.1389	0.1386	0.1384	0.1383	0.1383
35	0.2284	0.2007	0.1661	0.1438	0.1408	0.1382	0.1352	0.1332	0.1321	0.1309	0.1303	0.1295	0.1294	0.1291	0.1287	0.1286	0.1285	0.1284
40	0.2219	0.1936	0.1585	0.1350	0.1321	0.1296	0.1268	0.1249	0.1240	0.1227	0.1221	0.1216	0.1212	0.1211	0.1208	0.1206	0.1205	0.1205
45	0.2163	0.1879	0.1521	0.1278	0.1248	0.1226	0.1197	0.1180	0.1171	0.1159	0.1154	0.1148	0.1146	0.1144	0.1141	0.1139	0.1139	0.1138
50	0.2115	0.1829	0.1469	0.1216	0.1187	0.1165	0.1138	0.1123	0.1113	0.1102	0.1098	0.1092	0.1089	0.1088	0.1085	0.1083	0.1082	0.1082
60	0.2039	0.1749	0.1383	0.1117	0.1087	0.1067	0.1043	0.1027	0.1019	0.1009	0.1005	0.1000	0.0997	0.0996	0.0993	0.0991	0.0990	0.0990
70	0.1978	0.1686	0.1317	0.1039	0.1008	0.0991	0.0967	0.0953	0.0945	0.0937	0.0932	0.0927	0.0925	0.0924	0.0922	0.0920	0.0919	0.0919
80	0.1930	0.1635	0.1263	0.0977	0.0945	0.0928	0.0906	0.0893	0.0886	0.0878	0.0874	0.0869	0.0867	0.0865	0.0863	0.0862	0.0862	0.0861
90	0.1889	0.1593	0.1219	0.0924	0.0891	0.0876	0.0856	0.0844	0.0836	0.0829	0.0825	0.0821	0.0818	0.0817	0.0815	0.0814	0.0813	0.0813
100	0.1855	0.1557	0.1182	0.0880	0.0847	0.0832	0.0813	0.0801	0.0795	0.0788	0.0784	0.0780	0.0777	0.0776	0.0774	0.0773	0.0773	0.0772
200	0.1669	0.1364	0.0977	0.0643	0.0603	0.0592	0.0579	0.0571	0.0565	0.0560	0.0558	0.0555	0.0553	0.0552	0.0551	0.0550	0.0550	0.0550
300	0.1587	0.1279	0.0887	0.0540	0.0494	0.0485	0.0474	0.0467	0.0463	0.0459	0.0456	0.0454	0.0453	0.0452	0.0451	0.0450	0.0450	0.0450
400	0.1538	0.1228	0.0834	0.0479	0.0428	0.0421	0.0411	0.0405	0.0402	0.0398	0.0396	0.0394	0.0393	0.0393	0.0392	0.0391	0.0391	0.0391
500	0.1506	0.1194	0.0797	0.0438	0.0384	0.0376	0.0368	0.0363	0.0360	0.0357	0.0355	0.0353	0.0352	0.0352	0.0350	0.0350	0.0350	0.0350
1000	0.1425	0.1111	0.0709	0.0338	0.0272	0.0267	0.0261	0.0257	0.0255	0.0253	0.0252	0.0250	0.0250	0.0249	0.0249	0.0248	0.0248	0.0248
2500	0.1356	0.1039	0.0632	0.0252	0.0173	0.0169	0.0165	0.0163	0.0162	0.0160	0.0159	0.0159	0.0159	0.0158	0.0158	0.0157	0.0157	0.0157

Critical Values for Anderson Darling Test - Significance Level of 0.10

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.7088	0.6976	0.6855	0.6661	0.6259	0.6057	0.5899	0.5829	0.5809	0.5777	0.5764	0.5748	0.5747	0.5750	0.5744	0.5738	0.5733	0.5733
5	0.7611	0.7307	0.7050	0.6915	0.6540	0.6303	0.6122	0.6035	0.5988	0.5957	0.5937	0.5926	0.5907	0.5913	0.5895	0.5891	0.5883	0.5882
6	0.8243	0.7721	0.7260	0.7074	0.6719	0.6466	0.6246	0.6138	0.6099	0.6056	0.6031	0.6009	0.6001	0.5995	0.5987	0.5977	0.5979	0.5978
7	0.8907	0.8132	0.7424	0.7162	0.6838	0.6573	0.6353	0.6245	0.6180	0.6138	0.6092	0.6085	0.6065	0.6066	0.6057	0.6041	0.6045	0.6041
8	0.9573	0.8541	0.7599	0.7250	0.6944	0.6670	0.6415	0.6303	0.6236	0.6185	0.6155	0.6133	0.6120	0.6120	0.6099	0.6091	0.6081	0.6087
9	1.0261	0.8963	0.7772	0.7312	0.7009	0.6724	0.6478	0.6346	0.6289	0.6227	0.6198	0.6174	0.6158	0.6140	0.6125	0.6125	0.6118	0.6128
10	1.0941	0.9378	0.7920	0.7359	0.7070	0.6776	0.6516	0.6375	0.6315	0.6252	0.6229	0.6199	0.6186	0.6176	0.6157	0.6148	0.6138	0.6138
11	1.1631	0.9805	0.8091	0.7407	0.7121	0.6823	0.6552	0.6420	0.6339	0.6275	0.6250	0.6217	0.6195	0.6189	0.6178	0.6170	0.6164	0.6156
12	1.2308	1.0222	0.8239	0.7436	0.7156	0.6855	0.6573	0.6445	0.6378	0.6291	0.6267	0.6244	0.6214	0.6198	0.6187	0.6196	0.6178	0.6183
13	1.2985	1.0649	0.8408	0.7479	0.7175	0.6887	0.6603	0.6466	0.6376	0.6330	0.6276	0.6237	0.6230	0.6222	0.6208	0.6197	0.6195	0.6194
14	1.3637	1.1060	0.8544	0.7493	0.7210	0.6903	0.6627	0.6473	0.6399	0.6326	0.6291	0.6264	0.6243	0.6237	0.6212	0.6203	0.6189	0.6197
15	1.4305	1.1465	0.8679	0.7521	0.7238	0.6927	0.6629	0.6497	0.6416	0.6342	0.6299	0.6277	0.6242	0.6237	0.6222	0.6215	0.6210	0.6204
16	1.4956	1.1879	0.8847	0.7551	0.7246	0.6943	0.6655	0.6493	0.6424	0.6343	0.6319	0.6272	0.6259	0.6255	0.6236	0.6220	0.6209	0.6218
17	1.5594	1.2263	0.8989	0.7571	0.7272	0.6971	0.6664	0.6506	0.6438	0.6362	0.6318	0.6281	0.6265	0.6264	0.6234	0.6226	0.6217	0.6217
18	1.6219	1.2670	0.9131	0.7597	0.7286	0.6976	0.6675	0.6512	0.6433	0.6361	0.6337	0.6292	0.6279	0.6264	0.6234	0.6237	0.6229	0.6229
19	1.6831	1.3041	0.9283	0.7613	0.7302	0.6989	0.6699	0.6516	0.6453	0.6372	0.6333	0.6300	0.6271	0.6259	0.6251	0.6244	0.6239	0.6230
20	1.7445	1.3440	0.9407	0.7629	0.7316	0.7004	0.6688	0.6527	0.6451	0.6377	0.6333	0.6299	0.6282	0.6274	0.6258	0.6244	0.6244	0.6243
21	1.8079	1.3798	0.9536	0.7648	0.7323	0.6994	0.6713	0.6549	0.6458	0.6374	0.6341	0.6308	0.6288	0.6278	0.6264	0.6257	0.6243	0.6240
22	1.8649	1.4173	0.9679	0.7666	0.7337	0.7020	0.6705	0.6552	0.6462	0.6377	0.6355	0.6315	0.6295	0.6284	0.6266	0.6249	0.6254	0.6243
23	1.9250	1.4589	0.9827	0.7679	0.7349	0.7032	0.6711	0.6551	0.6451	0.6388	0.6346	0.6307	0.6306	0.6287	0.6269	0.6249	0.6247	0.6244
24	1.9846	1.4938	0.9951	0.7700	0.7356	0.7039	0.6719	0.6543	0.6468	0.6397	0.6347	0.6318	0.6298	0.6281	0.6271	0.6256	0.6254	0.6254
25	2.0426	1.5281	1.0090	0.7703	0.7364	0.7044	0.6727	0.6560	0.6484	0.6393	0.6355	0.6322	0.6299	0.6294	0.6274	0.6257	0.6251	0.6248
26	2.1022	1.5681	1.0213	0.7705	0.7353	0.7044	0.6729	0.6564	0.6484	0.6397	0.6355	0.6314	0.6294	0.6291	0.6268	0.6265	0.6260	0.6256
27	2.1572	1.6005	1.0378	0.7732	0.7370	0.7063	0.6730	0.6562	0.6486	0.6404	0.6360	0.6323	0.6297	0.6288	0.6271	0.6268	0.6260	0.6262
28	2.2173	1.6381	1.0486	0.7741	0.7372	0.7061	0.6742	0.6563	0.6486	0.6401	0.6374	0.6330	0.6311	0.6299	0.6272	0.6274	0.6261	0.6255
29	2.2750	1.6749	1.0620	0.7755	0.7369	0.7073	0.6753	0.6561	0.6488	0.6407	0.6371	0.6330	0.6311	0.6296	0.6270	0.6262	0.6260	0.6255
30	2.3305	1.7089	1.0763	0.7781	0.7378	0.7064	0.6744	0.6581	0.6496	0.6408	0.6374	0.6334	0.6307	0.6289	0.6279	0.6269	0.6261	0.6267
35	2.6059	1.8806	1.1413	0.7832	0.7395	0.7082	0.6765	0.6588	0.6502	0.6424	0.6382	0.6334	0.6332	0.6306	0.6293	0.6280	0.6270	0.6273
40	2.8792	2.0456	1.2022	0.7872	0.7421	0.7090	0.6768	0.6597	0.6514	0.6426	0.6370	0.6343	0.6320	0.6309	0.6287	0.6277	0.6276	0.6279
45	3.1396	2.2085	1.2601	0.7927	0.7430	0.7115	0.6768	0.6605	0.6507	0.6433	0.6387	0.6342	0.6321	0.6319	0.6304	0.6281	0.6289	0.6286
50	3.3978	2.3668	1.3201	0.7966	0.7457	0.7124	0.6772	0.6622	0.6518	0.6429	0.6398	0.6354	0.6326	0.6322	0.6302	0.6286	0.6285	0.6287
60	3.9028	2.6768	1.4351	0.8062	0.7470	0.7129	0.6785	0.6609	0.6518	0.6441	0.6395	0.6360	0.6333	0.6324	0.6313	0.6290	0.6291	0.6292
70	4.3942	2.9790	1.5495	0.8114	0.7472	0.7152	0.6789	0.6615	0.6531	0.6454	0.6408	0.6357	0.6339	0.6327	0.6302	0.6296	0.6295	0.6293
80	4.8817	3.2693	1.6602	0.8202	0.7471	0.7141	0.6801	0.6625	0.6535	0.6458	0.6412	0.6362	0.6352	0.6326	0.6312	0.6299	0.6291	0.6288
90	5.3579	3.5630	1.7658	0.8266	0.7481	0.7155	0.6801	0.6633	0.6537	0.6457	0.6417	0.6366	0.6339	0.6337	0.6310	0.6289	0.6303	0.6298
100	5.8377	3.8476	1.8740	0.8334	0.7487	0.7170	0.6807	0.6634	0.6538	0.6457	0.6422	0.6370	0.6341	0.6335	0.6315	0.6300	0.6296	0.6291
200	10.3750	6.5699	2.8606	0.9021	0.7513	0.7163	0.6820	0.6641	0.6545	0.6458	0.6418	0.6380	0.6351	0.6339	0.6311	0.6305	0.6302	0.6307
300	14.7424	9.1683	3.7864	0.9692	0.7517	0.7178	0.6822	0.6641	0.6549	0.6466	0.6411	0.6363	0.6359	0.6341	0.6334	0.6305	0.6302	0.6311
400	19.0253	11.6939	4.6787	1.0324	0.7510	0.7180	0.6826	0.6649	0.6554	0.6460	0.6413	0.6372	0.6358	0.6349	0.6329	0.6312	0.6306	0.6313
500	23.2588	14.1892	5.5513	1.0949	0.7520	0.7179	0.6825	0.6644	0.6554	0.6465	0.6426	0.6382	0.6356	0.6349	0.6320	0.6308	0.6305	0.6315
1000	43.9612	26.3237	9.7366	1.3989	0.7522	0.7183	0.6844	0.6645	0.6549	0.6460	0.6429	0.6373	0.6363	0.6348	0.6321	0.6310	0.6300	0.6311
2500	104.511	61.5654	21.6770	2.2303	0.7537	0.7193	0.6835	0.6645	0.6563	0.6462	0.6424	0.6375	0.6403	0.6357	0.6322	0.6302	0.6307	0.6316

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.10

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.4102	0.4026	0.3956	0.3925	0.3872	0.3817	0.3746	0.3699	0.3677	0.3647	0.3633	0.3622	0.3615	0.3616	0.3605	0.3601	0.3599	0.3596
5	0.3856	0.3761	0.3680	0.3655	0.3586	0.3524	0.3459	0.3416	0.3395	0.3373	0.3361	0.3351	0.3344	0.3343	0.3334	0.3331	0.3330	0.3328
6	0.3694	0.3570	0.3448	0.3393	0.3324	0.3261	0.3192	0.3151	0.3130	0.3110	0.3096	0.3087	0.3080	0.3075	0.3073	0.3068	0.3066	0.3066
7	0.3558	0.3394	0.3234	0.3175	0.3128	0.3071	0.3006	0.2968	0.2941	0.2919	0.2903	0.2893	0.2886	0.2885	0.2877	0.2872	0.2871	0.2870
8	0.3441	0.3265	0.3083	0.3010	0.2959	0.2903	0.2837	0.2799	0.2776	0.2754	0.2741	0.2731	0.2725	0.2722	0.2714	0.2710	0.2707	0.2709
9	0.3343	0.3148	0.2947	0.2856	0.2808	0.2754	0.2695	0.2656	0.2635	0.2612	0.2602	0.2590	0.2585	0.2578	0.2574	0.2571	0.2569	0.2570
10	0.3255	0.3048	0.2824	0.2726	0.2682	0.2630	0.2572	0.2533	0.2514	0.2492	0.2483	0.2471	0.2465	0.2463	0.2454	0.2451	0.2448	0.2448
11	0.3182	0.2962	0.2726	0.2613	0.2574	0.2523	0.2465	0.2430	0.2408	0.2386	0.2378	0.2365	0.2360	0.2356	0.2351	0.2348	0.2346	0.2345
12	0.3113	0.2887	0.2633	0.2515	0.2473	0.2425	0.2368	0.2337	0.2316	0.2293	0.2284	0.2274	0.2267	0.2263	0.2256	0.2257	0.2255	0.2254
13	0.3051	0.2817	0.2554	0.2425	0.2385	0.2340	0.2287	0.2253	0.2232	0.2214	0.2202	0.2191	0.2187	0.2183	0.2177	0.2174	0.2173	0.2172
14	0.2995	0.2758	0.2482	0.2344	0.2306	0.2262	0.2211	0.2177	0.2158	0.2137	0.2127	0.2119	0.2113	0.2111	0.2104	0.2101	0.2098	0.2096
15	0.2943	0.2702	0.2414	0.2271	0.2235	0.2192	0.2139	0.2110	0.2091	0.2070	0.2061	0.2052	0.2046	0.2043	0.2037	0.2034	0.2034	0.2032
16	0.2898	0.2651	0.2358	0.2207	0.2168	0.2127	0.2078	0.2047	0.2028	0.2009	0.2001	0.1990	0.1985	0.1983	0.1978	0.1974	0.1972	0.1973
17	0.2854	0.2602	0.2304	0.2145	0.2110	0.2071	0.2021	0.1990	0.1973	0.1954	0.1944	0.1935	0.1930	0.1928	0.1921	0.1919	0.1917	0.1917
18	0.2814	0.2559	0.2253	0.2091	0.2054	0.2014	0.1967	0.1938	0.1920	0.1903	0.1894	0.1884	0.1880	0.1876	0.1870	0.1868	0.1867	0.1867
19	0.2776	0.2518	0.2210	0.2038	0.2004	0.1965	0.1920	0.1888	0.1873	0.1854	0.1845	0.1838	0.1832	0.1828	0.1824	0.1822	0.1820	0.1820
20	0.2740	0.2480	0.2166	0.1991	0.1956	0.1919	0.1873	0.1844	0.1828	0.1812	0.1801	0.1793	0.1788	0.1786	0.1781	0.1779	0.1778	0.1777
21	0.2708	0.2445	0.2125	0.1945	0.1912	0.1873	0.1832	0.1804	0.1787	0.1770	0.1761	0.1753	0.1747	0.1745	0.1741	0.1738	0.1737	0.1736
22	0.2677	0.2412	0.2090	0.1904	0.1872	0.1835	0.1792	0.1764	0.1749	0.1730	0.1724	0.1715	0.1710	0.1708	0.1701	0.1700	0.1699	0.1698
23	0.2648	0.2383	0.2056	0.1865	0.1834	0.1797	0.1755	0.1728	0.1712	0.1696	0.1686	0.1678	0.1674	0.1673	0.1668	0.1663	0.1663	0.1662
24	0.2620	0.2354	0.2023	0.1829	0.1797	0.1762	0.1720	0.1693	0.1678	0.1662	0.1654	0.1646	0.1641	0.1639	0.1633	0.1631	0.1630	0.1629
25	0.2595	0.2325	0.1993	0.1794	0.1762	0.1729	0.1686	0.1660	0.1647	0.1630	0.1622	0.1614	0.1611	0.1608	0.1603	0.1601	0.1599	0.1599
26	0.2570	0.2299	0.1964	0.1762	0.1729	0.1695	0.1657	0.1630	0.1616	0.1599	0.1592	0.1583	0.1579	0.1576	0.1573	0.1570	0.1569	0.1569
27	0.2547	0.2274	0.1941	0.1731	0.1699	0.1667	0.1626	0.1602	0.1588	0.1572	0.1564	0.1556	0.1551	0.1549	0.1544	0.1543	0.1541	0.1541
28	0.2526	0.2253	0.1912	0.1701	0.1669	0.1638	0.1599	0.1573	0.1560	0.1544	0.1537	0.1529	0.1526	0.1523	0.1518	0.1517	0.1515	0.1515
29	0.2503	0.2230	0.1888	0.1674	0.1641	0.1612	0.1573	0.1547	0.1535	0.1518	0.1512	0.1505	0.1501	0.1497	0.1492	0.1491	0.1490	0.1489
30	0.2484	0.2208	0.1866	0.1648	0.1615	0.1585	0.1547	0.1523	0.1509	0.1495	0.1488	0.1480	0.1475	0.1473	0.1469	0.1467	0.1466	0.1466
35	0.2395	0.2115	0.1766	0.1534	0.1500	0.1472	0.1437	0.1415	0.1402	0.1389	0.1383	0.1374	0.1372	0.1368	0.1365	0.1363	0.1361	0.1361
40	0.2324	0.2039	0.1684	0.1441	0.1407	0.1381	0.1348	0.1327	0.1316	0.1302	0.1295	0.1290	0.1286	0.1284	0.1280	0.1278	0.1277	0.1278
45	0.2262	0.1976	0.1614	0.1363	0.1331	0.1306	0.1273	0.1255	0.1243	0.1230	0.1224	0.1218	0.1215	0.1213	0.1210	0.1208	0.1207	0.1206
50	0.2210	0.1922	0.1558	0.1297	0.1265	0.1241	0.1210	0.1193	0.1182	0.1170	0.1165	0.1158	0.1155	0.1153	0.1150	0.1147	0.1147	0.1147
60	0.2126	0.1835	0.1466	0.1191	0.1159	0.1136	0.1109	0.1091	0.1082	0.1071	0.1066	0.1061	0.1057	0.1056	0.1053	0.1051	0.1049	0.1050
70	0.2060	0.1766	0.1394	0.1108	0.1075	0.1056	0.1028	0.1013	0.1004	0.0994	0.0989	0.0983	0.0981	0.0979	0.0977	0.0975	0.0974	0.0974
80	0.2007	0.1710	0.1336	0.1042	0.1008	0.0988	0.0964	0.0949	0.0941	0.0932	0.0927	0.0922	0.0920	0.0918	0.0916	0.0913	0.0913	0.0913
90	0.1962	0.1665	0.1288	0.0985	0.0950	0.0933	0.0910	0.0896	0.0888	0.0880	0.0876	0.0870	0.0868	0.0866	0.0864	0.0863	0.0862	0.0861
100	0.1925	0.1625	0.1248	0.0938	0.0902	0.0886	0.0865	0.0851	0.0844	0.0835	0.0831	0.0826	0.0824	0.0823	0.0820	0.0819	0.0819	0.0819
200	0.1719	0.1413	0.1024	0.0686	0.0643	0.0631	0.0616	0.0606	0.0600	0.0594	0.0592	0.0588	0.0587	0.0586	0.0583	0.0583	0.0583	0.0583
300	0.1629	0.1320	0.0926	0.0576	0.0526	0.0516	0.0504	0.0496	0.0492	0.0487	0.0484	0.0482	0.0481	0.0480	0.0478	0.0477	0.0477	0.0477
400	0.1575	0.1263	0.0868	0.0510	0.0456	0.0448	0.0437	0.0430	0.0426	0.0422	0.0420	0.0418	0.0417	0.0416	0.0415	0.0414	0.0414	0.0414
500	0.1538	0.1226	0.0828	0.0466	0.0409	0.0401	0.0391	0.0385	0.0382	0.0378	0.0376	0.0374	0.0373	0.0373	0.0372	0.0371	0.0371	0.0371
1000	0.1449	0.1134	0.0731	0.0359	0.0290	0.0284	0.0278	0.0273	0.0271	0.0268	0.0267	0.0265	0.0265	0.0264	0.0263	0.0263	0.0263	0.0263
2500	0.1371	0.1054	0.0647	0.0266	0.0184	0.0180	0.0176	0.0173	0.0172	0.0170	0.0169	0.0168	0.0168	0.0168	0.0167	0.0167	0.0167	0.0167

Critical Values for Anderson Darling Test - Significance Level of 0.05

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.7933	0.7883	0.7863	0.7785	0.7325	0.7041	0.6809	0.6703	0.6666	0.6624	0.6605	0.6594	0.6589	0.6590	0.6571	0.6571	0.6559	0.6565
5	0.8730	0.8462	0.8304	0.8264	0.7753	0.7392	0.7110	0.6983	0.6913	0.6864	0.6845	0.6826	0.6812	0.6807	0.6789	0.6787	0.6783	0.6781
6	0.9490	0.8965	0.8535	0.8446	0.8025	0.7667	0.7359	0.7210	0.7151	0.7078	0.7042	0.7013	0.7000	0.6980	0.6983	0.6971	0.6969	0.6962
7	1.0305	0.9476	0.8762	0.8598	0.8211	0.7841	0.7515	0.7361	0.7275	0.7212	0.7149	0.7122	0.7097	0.7099	0.7085	0.7067	0.7077	0.7077
8	1.1136	1.0006	0.8986	0.8720	0.8359	0.7973	0.7624	0.7451	0.7355	0.7284	0.7240	0.7215	0.7186	0.7191	0.7150	0.7162	0.7148	0.7147
9	1.1971	1.0535	0.9227	0.8810	0.8451	0.8073	0.7707	0.7515	0.7435	0.7344	0.7298	0.7268	0.7250	0.7228	0.7218	0.7210	0.7205	0.7200
10	1.2792	1.1063	0.9420	0.8881	0.8539	0.8136	0.7770	0.7570	0.7483	0.7392	0.7356	0.7322	0.7294	0.7295	0.7251	0.7246	0.7244	0.7238
11	1.3623	1.1586	0.9637	0.8950	0.8601	0.8201	0.7811	0.7625	0.7516	0.7422	0.7389	0.7335	0.7326	0.7314	0.7294	0.7287	0.7284	0.7257
12	1.4414	1.2089	0.9834	0.8989	0.8656	0.8239	0.7849	0.7656	0.7567	0.7455	0.7415	0.7390	0.7358	0.7320	0.7303	0.7319	0.7296	0.7312
13	1.5200	1.2605	1.0039	0.9049	0.8682	0.8298	0.7900	0.7703	0.7574	0.7503	0.7431	0.7392	0.7372	0.7359	0.7337	0.7333	0.7325	0.7323
14	1.5958	1.3106	1.0234	0.9072	0.8735	0.8320	0.7933	0.7706	0.7600	0.7507	0.7459	0.7425	0.7401	0.7380	0.7345	0.7340	0.7330	0.7334
15	1.6732	1.3605	1.0405	0.9113	0.8768	0.8355	0.7930	0.7751	0.7630	0.7538	0.7468	0.7445	0.7400	0.7386	0.7374	0.7347	0.7345	0.7341
16	1.7482	1.4088	1.0618	0.9164	0.8783	0.8378	0.7963	0.7745	0.7633	0.7547	0.7503	0.7443	0.7419	0.7413	0.7390	0.7365	0.7354	0.7360
17	1.8194	1.4552	1.0796	0.9205	0.8827	0.8418	0.7979	0.7764	0.7660	0.7557	0.7494	0.7454	0.7428	0.7416	0.7388	0.7378	0.7367	0.7362
18	1.8905	1.4995	1.0965	0.9229	0.8842	0.8421	0.8001	0.7780	0.7666	0.7562	0.7526	0.7458	0.7426	0.7430	0.7392	0.7395	0.7383	0.7372
19	1.9614	1.5452	1.1162	0.9250	0.8877	0.8428	0.8028	0.7788	0.7692	0.7569	0.7518	0.7481	0.7451	0.7424	0.7412	0.7403	0.7404	0.7379
20	2.0284	1.5917	1.1322	0.9289	0.8880	0.8447	0.8025	0.7795	0.7679	0.7578	0.7524	0.7475	0.7455	0.7452	0.7418	0.7407	0.7394	0.7405
21	2.0984	1.6336	1.1480	0.9288	0.8903	0.8458	0.8053	0.7828	0.7696	0.7582	0.7538	0.7494	0.7473	0.7453	0.7426	0.7425	0.7412	0.7395
22	2.1639	1.6751	1.1669	0.9334	0.8918	0.8476	0.8043	0.7830	0.7708	0.7587	0.7561	0.7494	0.7466	0.7464	0.7436	0.7403	0.7429	0.7406
23	2.2329	1.7214	1.1839	0.9338	0.8939	0.8488	0.8051	0.7822	0.7693	0.7601	0.7547	0.7503	0.7491	0.7465	0.7441	0.7419	0.7414	0.7404
24	2.2974	1.7630	1.2009	0.9377	0.8938	0.8512	0.8063	0.7825	0.7719	0.7615	0.7551	0.7511	0.7487	0.7462	0.7443	0.7423	0.7421	0.7418
25	2.3601	1.8028	1.2161	0.9394	0.8955	0.8518	0.8069	0.7835	0.7731	0.7615	0.7565	0.7513	0.7487	0.7470	0.7448	0.7432	0.7423	0.7418
26	2.4252	1.8483	1.2315	0.9393	0.8936	0.8516	0.8085	0.7842	0.7734	0.7616	0.7573	0.7505	0.7478	0.7463	0.7441	0.7438	0.7428	0.7422
27	2.4909	1.8820	1.2531	0.9437	0.8957	0.8531	0.8074	0.7839	0.7733	0.7630	0.7566	0.7517	0.7490	0.7474	0.7436	0.7440	0.7433	0.7439
28	2.5562	1.9280	1.2634	0.9432	0.8971	0.8543	0.8103	0.7847	0.7738	0.7627	0.7580	0.7537	0.7497	0.7485	0.7453	0.7446	0.7431	0.7431
29	2.6160	1.9685	1.2809	0.9478	0.8976	0.8562	0.8115	0.7837	0.7742	0.7627	0.7573	0.7527	0.7503	0.7475	0.7457	0.7442	0.7439	0.7423
30	2.6778	2.0063	1.2983	0.9482	0.8976	0.8538	0.8092	0.7878	0.7755	0.7630	0.7585	0.7524	0.7495	0.7465	0.7455	0.7443	0.7441	0.7451
35	2.9819	2.1959	1.3736	0.9546	0.8995	0.8565	0.8122	0.7877	0.7757	0.7666	0.7597	0.7537	0.7526	0.7497	0.7483	0.7467	0.7447	0.7459
40	3.2742	2.3805	1.4435	0.9625	0.9028	0.8577	0.8131	0.7891	0.7787	0.7658	0.7590	0.7547	0.7525	0.7515	0.7480	0.7467	0.7458	0.7468
45	3.5595	2.5587	1.5106	0.9690	0.9054	0.8619	0.8134	0.7897	0.7769	0.7679	0.7605	0.7556	0.7529	0.7532	0.7484	0.7482	0.7473	0.7471
50	3.8334	2.7329	1.5791	0.9737	0.9074	0.8624	0.8143	0.7934	0.7800	0.7672	0.7629	0.7569	0.7535	0.7537	0.7496	0.7482	0.7476	0.7481
60	4.3789	3.0659	1.7118	0.9844	0.9099	0.8633	0.8160	0.7921	0.7791	0.7689	0.7629	0.7580	0.7536	0.7533	0.7512	0.7489	0.7476	0.7484
70	4.9012	3.3923	1.8398	0.9917	0.9096	0.8657	0.8167	0.7926	0.7805	0.7689	0.7634	0.7575	0.7557	0.7542	0.7507	0.7492	0.7486	0.7490
80	5.4154	3.7091	1.9620	1.0021	0.9104	0.8649	0.8189	0.7931	0.7820	0.7703	0.7631	0.7589	0.7563	0.7545	0.7505	0.7508	0.7484	0.7488
90	5.9167	4.0188	2.0787	1.0111	0.9113	0.8679	0.8184	0.7936	0.7828	0.7715	0.7651	0.7592	0.7554	0.7548	0.7521	0.7495	0.7513	0.7502
100	6.4255	4.3222	2.1954	1.0194	0.9123	0.8676	0.8184	0.7950	0.7830	0.7698	0.7650	0.7585	0.7564	0.7543	0.7522	0.7495	0.7504	0.7500
200	11.1598	7.1943	3.2677	1.1031	0.9142	0.8692	0.8209	0.7962	0.7839	0.7713	0.7664	0.7604	0.7564	0.7561	0.7512	0.7513	0.7503	0.7506
300	15.6877	9.9089	4.2544	1.1798	0.9166	0.8707	0.8217	0.7969	0.7840	0.7720	0.7659	0.7594	0.7588	0.7568	0.7552	0.7509	0.7516	0.7523
400	20.0982	12.5299	5.1940	1.2564	0.9170	0.8713	0.8230	0.7977	0.7846	0.7728	0.7660	0.7602	0.7586	0.7572	0.7542	0.7515	0.7523	0.7511
500	24.4270	15.1069	6.1097	1.3280	0.9178	0.8716	0.8223	0.7971	0.7851	0.7721	0.7671	0.7615	0.7590	0.7563	0.7534	0.7522	0.7515	0.7530
1000	45.5811	27.5755	10.4679	1.6707	0.9188	0.8707	0.8244	0.7966	0.7846	0.7713	0.7679	0.7603	0.7597	0.7573	0.7527	0.7519	0.7497	0.7520
2500	107.018	63.4597	22.7439	2.5739	0.9200	0.8732	0.8223	0.7969	0.7860	0.7722	0.7666	0.7603	0.7641	0.7572	0.7527	0.7513	0.7522	0.7525

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.05

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.4371	0.4323	0.4289	0.4296	0.4244	0.4181	0.4103	0.4050	0.4024	0.3995	0.3979	0.3966	0.3962	0.3959	0.3949	0.3942	0.3938	0.3940
5	0.4191	0.4093	0.4009	0.3982	0.3885	0.3799	0.3716	0.3667	0.3644	0.3617	0.3605	0.3594	0.3584	0.3583	0.3576	0.3572	0.3569	0.3568
6	0.3971	0.3844	0.3726	0.3688	0.3637	0.3568	0.3486	0.3434	0.3408	0.3375	0.3358	0.3346	0.3335	0.3328	0.3325	0.3317	0.3318	0.3315
7	0.3849	0.3688	0.3528	0.3478	0.3419	0.3351	0.3272	0.3227	0.3196	0.3170	0.3151	0.3137	0.3129	0.3130	0.3119	0.3115	0.3114	0.3113
8	0.3724	0.3541	0.3356	0.3287	0.3233	0.3170	0.3090	0.3041	0.3015	0.2989	0.2975	0.2963	0.2953	0.2952	0.2943	0.2939	0.2934	0.2936
9	0.3617	0.3418	0.3208	0.3123	0.3075	0.3015	0.2941	0.2893	0.2869	0.2840	0.2825	0.2813	0.2804	0.2798	0.2793	0.2788	0.2787	0.2788
10	0.3523	0.3314	0.3081	0.2984	0.2941	0.2878	0.2808	0.2760	0.2737	0.2711	0.2698	0.2685	0.2678	0.2674	0.2666	0.2662	0.2659	0.2658
11	0.3440	0.3221	0.2976	0.2862	0.2819	0.2759	0.2691	0.2649	0.2624	0.2597	0.2585	0.2570	0.2565	0.2560	0.2554	0.2550	0.2546	0.2544
12	0.3368	0.3137	0.2873	0.2753	0.2710	0.2653	0.2587	0.2548	0.2524	0.2495	0.2485	0.2474	0.2464	0.2459	0.2454	0.2452	0.2450	0.2450
13	0.3297	0.3064	0.2789	0.2659	0.2613	0.2563	0.2498	0.2459	0.2430	0.2409	0.2394	0.2382	0.2377	0.2374	0.2367	0.2361	0.2362	0.2360
14	0.3234	0.2996	0.2711	0.2568	0.2530	0.2478	0.2416	0.2376	0.2351	0.2327	0.2316	0.2305	0.2298	0.2293	0.2287	0.2283	0.2280	0.2279
15	0.3178	0.2934	0.2638	0.2489	0.2451	0.2400	0.2338	0.2302	0.2279	0.2255	0.2244	0.2233	0.2226	0.2221	0.2215	0.2212	0.2210	0.2209
16	0.3127	0.2878	0.2577	0.2419	0.2376	0.2329	0.2272	0.2232	0.2212	0.2189	0.2179	0.2167	0.2161	0.2158	0.2152	0.2146	0.2144	0.2144
17	0.3078	0.2826	0.2519	0.2352	0.2315	0.2268	0.2209	0.2173	0.2151	0.2129	0.2117	0.2106	0.2100	0.2097	0.2089	0.2087	0.2085	0.2084
18	0.3033	0.2776	0.2464	0.2292	0.2253	0.2208	0.2151	0.2114	0.2094	0.2072	0.2063	0.2050	0.2046	0.2041	0.2034	0.2033	0.2031	0.2031
19	0.2990	0.2731	0.2415	0.2236	0.2198	0.2151	0.2100	0.2061	0.2043	0.2021	0.2008	0.2000	0.1992	0.1990	0.1986	0.1981	0.1981	0.1979
20	0.2949	0.2691	0.2366	0.2185	0.2145	0.2101	0.2049	0.2014	0.1994	0.1974	0.1961	0.1950	0.1946	0.1945	0.1938	0.1935	0.1934	0.1934
21	0.2915	0.2649	0.2323	0.2132	0.2097	0.2054	0.2004	0.1969	0.1949	0.1928	0.1918	0.1909	0.1903	0.1900	0.1894	0.1892	0.1889	0.1889
22	0.2879	0.2612	0.2283	0.2090	0.2055	0.2011	0.1959	0.1927	0.1908	0.1885	0.1879	0.1867	0.1862	0.1859	0.1853	0.1849	0.1851	0.1848
23	0.2847	0.2580	0.2247	0.2046	0.2013	0.1969	0.1919	0.1886	0.1867	0.1849	0.1838	0.1827	0.1824	0.1821	0.1815	0.1810	0.1809	0.1809
24	0.2813	0.2546	0.2211	0.2007	0.1971	0.1932	0.1881	0.1849	0.1830	0.1812	0.1802	0.1792	0.1787	0.1783	0.1777	0.1775	0.1774	0.1772
25	0.2786	0.2516	0.2179	0.1969	0.1933	0.1895	0.1845	0.1813	0.1796	0.1777	0.1767	0.1759	0.1753	0.1749	0.1745	0.1742	0.1739	0.1739
26	0.2759	0.2486	0.2146	0.1933	0.1896	0.1858	0.1812	0.1780	0.1764	0.1742	0.1734	0.1724	0.1719	0.1716	0.1712	0.1708	0.1707	0.1707
27	0.2732	0.2459	0.2118	0.1899	0.1863	0.1827	0.1779	0.1750	0.1730	0.1714	0.1705	0.1694	0.1689	0.1686	0.1681	0.1678	0.1676	0.1677
28	0.2709	0.2434	0.2088	0.1867	0.1832	0.1795	0.1749	0.1719	0.1702	0.1684	0.1676	0.1666	0.1661	0.1659	0.1652	0.1652	0.1649	0.1648
29	0.2683	0.2409	0.2062	0.1837	0.1802	0.1767	0.1721	0.1690	0.1675	0.1655	0.1647	0.1639	0.1634	0.1630	0.1625	0.1623	0.1622	0.1620
30	0.2663	0.2386	0.2037	0.1809	0.1772	0.1736	0.1692	0.1663	0.1648	0.1629	0.1621	0.1611	0.1607	0.1603	0.1600	0.1597	0.1595	0.1596
35	0.2561	0.2281	0.1927	0.1683	0.1647	0.1613	0.1571	0.1545	0.1530	0.1515	0.1507	0.1497	0.1494	0.1490	0.1486	0.1484	0.1482	0.1481
40	0.2482	0.2196	0.1835	0.1581	0.1544	0.1514	0.1476	0.1449	0.1437	0.1420	0.1412	0.1404	0.1401	0.1399	0.1394	0.1391	0.1390	0.1390
45	0.2412	0.2124	0.1759	0.1496	0.1461	0.1432	0.1393	0.1370	0.1356	0.1342	0.1334	0.1327	0.1323	0.1322	0.1317	0.1315	0.1314	0.1313
50	0.2353	0.2063	0.1695	0.1425	0.1389	0.1361	0.1324	0.1304	0.1289	0.1275	0.1269	0.1262	0.1257	0.1256	0.1252	0.1249	0.1249	0.1249
60	0.2258	0.1963	0.1592	0.1308	0.1272	0.1245	0.1214	0.1192	0.1180	0.1168	0.1161	0.1156	0.1151	0.1150	0.1147	0.1144	0.1143	0.1143
70	0.2183	0.1886	0.1513	0.1216	0.1179	0.1157	0.1126	0.1107	0.1095	0.1084	0.1078	0.1071	0.1069	0.1067	0.1064	0.1061	0.1061	0.1061
80	0.2122	0.1823	0.1447	0.1143	0.1105	0.1083	0.1055	0.1037	0.1027	0.1016	0.1011	0.1005	0.1002	0.0999	0.0997	0.0995	0.0993	0.0994
90	0.2071	0.1771	0.1392	0.1082	0.1044	0.1023	0.0996	0.0979	0.0970	0.0959	0.0954	0.0949	0.0945	0.0944	0.0941	0.0939	0.0939	0.0938
100	0.2029	0.1727	0.1347	0.1031	0.0991	0.0971	0.0946	0.0929	0.0921	0.0911	0.0906	0.0901	0.0898	0.0896	0.0894	0.0892	0.0892	0.0892
200	0.1794	0.1487	0.1096	0.0753	0.0705	0.0691	0.0673	0.0662	0.0655	0.0648	0.0645	0.0641	0.0639	0.0638	0.0635	0.0635	0.0635	0.0634
300	0.1691	0.1380	0.0985	0.0631	0.0577	0.0566	0.0551	0.0542	0.0537	0.0531	0.0528	0.0524	0.0523	0.0522	0.0521	0.0520	0.0519	0.0520
400	0.1629	0.1316	0.0919	0.0559	0.0501	0.0491	0.0478	0.0470	0.0465	0.0460	0.0457	0.0455	0.0454	0.0453	0.0452	0.0451	0.0451	0.0451
500	0.1587	0.1274	0.0874	0.0510	0.0448	0.0439	0.0428	0.0421	0.0417	0.0412	0.0410	0.0408	0.0406	0.0406	0.0404	0.0404	0.0403	0.0404
1000	0.1484	0.1168	0.0764	0.0390	0.0318	0.0311	0.0303	0.0298	0.0295	0.0292	0.0291	0.0289	0.0288	0.0288	0.0286	0.0286	0.0286	0.0286
2500	0.1394	0.1076	0.0668	0.0286	0.0202	0.0197	0.0192	0.0189	0.0187	0.0185	0.0184	0.0183	0.0183	0.0182	0.0182	0.0181	0.0181	0.0181

Critical Values for Anderson Darling Test - Significance Level of 0.025

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.8615	0.8677	0.8863	0.9039	0.8378	0.7984	0.7665	0.7511	0.7460	0.7399	0.7369	0.7355	0.7346	0.7346	0.7325	0.7317	0.7305	0.7311
5	0.9689	0.9509	0.9503	0.9609	0.8987	0.8497	0.8119	0.7925	0.7854	0.7759	0.7730	0.7718	0.7693	0.7693	0.7678	0.7662	0.7659	0.7643
6	1.0630	1.0130	0.9765	0.9863	0.9370	0.8879	0.8446	0.8255	0.8160	0.8059	0.8026	0.7993	0.7961	0.7948	0.7943	0.7932	0.7922	0.7918
7	1.1594	1.0729	1.0070	1.0066	0.9620	0.9114	0.8676	0.8467	0.8348	0.8258	0.8170	0.8144	0.8106	0.8106	0.8089	0.8067	0.8072	0.8070
8	1.2574	1.1392	1.0345	1.0216	0.9817	0.9290	0.8842	0.8577	0.8457	0.8353	0.8304	0.8278	0.8236	0.8235	0.8187	0.8195	0.8182	0.8185
9	1.3550	1.2024	1.0646	1.0334	0.9943	0.9451	0.8930	0.8689	0.8564	0.8441	0.8394	0.8360	0.8325	0.8298	0.8294	0.8278	0.8270	0.8273
10	1.4495	1.2659	1.0902	1.0433	1.0063	0.9536	0.9025	0.8758	0.8651	0.8524	0.8468	0.8419	0.8389	0.8396	0.8344	0.8330	0.8324	0.8317
11	1.5453	1.3293	1.1162	1.0521	1.0135	0.9598	0.9087	0.8832	0.8700	0.8571	0.8514	0.8445	0.8443	0.8416	0.8400	0.8385	0.8389	0.8349
12	1.6379	1.3875	1.1414	1.0568	1.0220	0.9648	0.9143	0.8900	0.8752	0.8611	0.8557	0.8507	0.8475	0.8447	0.8405	0.8420	0.8406	0.8430
13	1.7256	1.4490	1.1649	1.0652	1.0232	0.9720	0.9209	0.8944	0.8772	0.8665	0.8592	0.8539	0.8501	0.8486	0.8458	0.8445	0.8426	0.8433
14	1.8117	1.5045	1.1897	1.0698	1.0315	0.9777	0.9247	0.8944	0.8811	0.8688	0.8634	0.8573	0.8538	0.8515	0.8473	0.8463	0.8456	0.8454
15	1.8988	1.5637	1.2102	1.0738	1.0342	0.9816	0.9236	0.8986	0.8847	0.8708	0.8649	0.8601	0.8556	0.8527	0.8497	0.8478	0.8481	0.8481
16	1.9814	1.6174	1.2393	1.0815	1.0363	0.9840	0.9318	0.9015	0.8852	0.8745	0.8678	0.8610	0.8572	0.8564	0.8521	0.8506	0.8489	0.8507
17	2.0598	1.6722	1.2611	1.0865	1.0428	0.9892	0.9315	0.9020	0.8879	0.8750	0.8685	0.8619	0.8596	0.8565	0.8536	0.8520	0.8505	0.8503
18	2.1409	1.7231	1.2808	1.0905	1.0435	0.9902	0.9336	0.9056	0.8892	0.8751	0.8701	0.8626	0.8581	0.8589	0.8551	0.8555	0.8532	0.8519
19	2.2162	1.7764	1.3033	1.0927	1.0480	0.9919	0.9357	0.9069	0.8932	0.8776	0.8713	0.8662	0.8630	0.8583	0.8566	0.8552	0.8555	0.8534
20	2.2915	1.8258	1.3244	1.0978	1.0493	0.9941	0.9375	0.9060	0.8917	0.8789	0.8712	0.8637	0.8626	0.8621	0.8580	0.8561	0.8538	0.8559
21	2.3715	1.8738	1.3396	1.0962	1.0515	0.9951	0.9410	0.9105	0.8939	0.8793	0.8727	0.8681	0.8644	0.8614	0.8589	0.8591	0.8572	0.8552
22	2.4415	1.9185	1.3632	1.1035	1.0557	0.9968	0.9391	0.9107	0.8948	0.8791	0.8763	0.8668	0.8644	0.8644	0.8612	0.8572	0.8609	0.8568
23	2.5164	1.9742	1.3839	1.1023	1.0576	0.9990	0.9418	0.9103	0.8927	0.8811	0.8757	0.8706	0.8680	0.8627	0.8620	0.8566	0.8586	0.8569
24	2.5831	2.0197	1.4035	1.1105	1.0566	1.0019	0.9417	0.9126	0.8978	0.8822	0.8761	0.8710	0.8658	0.8640	0.8615	0.8595	0.8569	0.8577
25	2.6565	2.0644	1.4219	1.1114	1.0614	1.0024	0.9430	0.9131	0.8991	0.8836	0.8768	0.8705	0.8671	0.8651	0.8643	0.8602	0.8585	0.8579
26	2.7258	2.1088	1.4384	1.1131	1.0567	1.0018	0.9462	0.9143	0.8994	0.8838	0.8765	0.8708	0.8666	0.8637	0.8619	0.8617	0.8609	0.8594
27	2.7952	2.1511	1.4646	1.1156	1.0600	1.0060	0.9440	0.9154	0.8986	0.8861	0.8779	0.8717	0.8674	0.8649	0.8604	0.8612	0.8603	0.8599
28	2.8692	2.1998	1.4766	1.1174	1.0627	1.0042	0.9485	0.9158	0.8990	0.8850	0.8797	0.8739	0.8680	0.8672	0.8644	0.8627	0.8613	0.8602
29	2.9301	2.2438	1.4940	1.1235	1.0606	1.0077	0.9508	0.9147	0.9008	0.8857	0.8789	0.8743	0.8698	0.8665	0.8637	0.8625	0.8609	0.8588
30	3.0015	2.2866	1.5156	1.1243	1.0611	1.0049	0.9456	0.9170	0.9021	0.8860	0.8795	0.8727	0.8688	0.8650	0.8647	0.8625	0.8611	0.8624
35	3.3266	2.4946	1.6032	1.1307	1.0635	1.0088	0.9494	0.9187	0.9017	0.8905	0.8828	0.8730	0.8718	0.8686	0.8671	0.8647	0.8611	0.8643
40	3.6396	2.6943	1.6803	1.1409	1.0679	1.0105	0.9513	0.9198	0.9068	0.8895	0.8814	0.8759	0.8730	0.8689	0.8677	0.8658	0.8635	0.8665
45	3.9425	2.8877	1.7545	1.1477	1.0716	1.0157	0.9507	0.9215	0.9049	0.8916	0.8834	0.8760	0.8734	0.8741	0.8687	0.8689	0.8664	0.8653
50	4.2378	3.0745	1.8299	1.1563	1.0756	1.0154	0.9541	0.9252	0.9081	0.8919	0.8869	0.8786	0.8753	0.8741	0.8690	0.8668	0.8677	0.8687
60	4.8100	3.4320	1.9809	1.1694	1.0776	1.0157	0.9548	0.9248	0.9065	0.8932	0.8870	0.8805	0.8760	0.8744	0.8718	0.8699	0.8685	0.8670
70	5.3566	3.7754	2.1204	1.1781	1.0761	1.0204	0.9572	0.9257	0.9083	0.8935	0.8881	0.8801	0.8774	0.8766	0.8704	0.8702	0.8686	0.8694
80	5.9031	4.1163	2.2521	1.1918	1.0787	1.0207	0.9602	0.9257	0.9116	0.8949	0.8880	0.8805	0.8771	0.8753	0.8713	0.8702	0.8689	0.8687
90	6.4293	4.4397	2.3772	1.2011	1.0823	1.0220	0.9579	0.9271	0.9135	0.8981	0.8894	0.8815	0.8766	0.8763	0.8731	0.8688	0.8728	0.8708
100	6.9592	4.7648	2.5055	1.2102	1.0801	1.0215	0.9593	0.9267	0.9118	0.8948	0.8885	0.8808	0.8782	0.8778	0.8727	0.8702	0.8719	0.8705
200	11.8779	7.7654	3.6516	1.3093	1.0827	1.0257	0.9641	0.9305	0.9142	0.8985	0.8923	0.8845	0.8789	0.8789	0.8719	0.8727	0.8714	0.8716
300	16.5240	10.5806	4.6843	1.3997	1.0856	1.0260	0.9631	0.9310	0.9155	0.8988	0.8914	0.8824	0.8824	0.8801	0.8772	0.8727	0.8721	0.8732
400	21.0493	13.2831	5.6639	1.4844	1.0867	1.0286	0.9643	0.9331	0.9148	0.8989	0.8918	0.8843	0.8815	0.8801	0.8761	0.8729	0.8733	0.8725
500	25.4819	15.9306	6.6237	1.5670	1.0880	1.0277	0.9645	0.9305	0.9155	0.8992	0.8921	0.8843	0.8820	0.8790	0.8774	0.8738	0.8721	0.8751
1000	47.0169	28.6841	11.1336	1.9413	1.0900	1.0281	0.9665	0.9295	0.9142	0.8976	0.8946	0.8845	0.8834	0.8791	0.8744	0.8742	0.8703	0.8719
2500	109.217	65.1297	23.6988	2.9078	1.0895	1.0304	0.9647	0.9295	0.9171	0.8991	0.8932	0.8854	0.8877	0.8812	0.8755	0.8720	0.8748	0.8732

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.025

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.4542	0.4526	0.4535	0.4577	0.4505	0.4430	0.4343	0.4287	0.4258	0.4229	0.4211	0.4198	0.4192	0.4190	0.4176	0.4174	0.4167	0.4169
5	0.4429	0.4346	0.4277	0.4258	0.4171	0.4075	0.3974	0.3908	0.3877	0.3838	0.3823	0.3810	0.3798	0.3796	0.3786	0.3783	0.3779	0.3777
6	0.4240	0.4103	0.3981	0.3971	0.3912	0.3830	0.3737	0.3681	0.3651	0.3612	0.3596	0.3579	0.3570	0.3563	0.3556	0.3550	0.3550	0.3549
7	0.4086	0.3933	0.3784	0.3743	0.3680	0.3600	0.3509	0.3454	0.3419	0.3389	0.3365	0.3348	0.3339	0.3342	0.3329	0.3325	0.3323	0.3324
8	0.3968	0.3786	0.3595	0.3537	0.3484	0.3414	0.3324	0.3267	0.3233	0.3204	0.3185	0.3173	0.3161	0.3160	0.3146	0.3144	0.3138	0.3140
9	0.3852	0.3655	0.3447	0.3366	0.3315	0.3251	0.3160	0.3108	0.3080	0.3044	0.3027	0.3014	0.3006	0.2998	0.2993	0.2986	0.2986	0.2985
10	0.3752	0.3547	0.3310	0.3217	0.3172	0.3103	0.3020	0.2964	0.2941	0.2910	0.2893	0.2878	0.2868	0.2866	0.2854	0.2852	0.2848	0.2848
11	0.3665	0.3448	0.3199	0.3088	0.3043	0.2977	0.2897	0.2847	0.2819	0.2787	0.2773	0.2755	0.2749	0.2744	0.2737	0.2733	0.2730	0.2726
12	0.3588	0.3358	0.3090	0.2972	0.2926	0.2861	0.2785	0.2740	0.2712	0.2678	0.2666	0.2655	0.2642	0.2637	0.2630	0.2629	0.2626	0.2626
13	0.3513	0.3278	0.2999	0.2871	0.2820	0.2763	0.2692	0.2645	0.2613	0.2588	0.2570	0.2558	0.2550	0.2547	0.2537	0.2532	0.2532	0.2530
14	0.3441	0.3205	0.2915	0.2773	0.2732	0.2674	0.2604	0.2553	0.2528	0.2500	0.2488	0.2473	0.2466	0.2460	0.2452	0.2448	0.2445	0.2445
15	0.3383	0.3138	0.2839	0.2686	0.2647	0.2591	0.2519	0.2478	0.2451	0.2423	0.2408	0.2397	0.2389	0.2383	0.2375	0.2372	0.2370	0.2371
16	0.3324	0.3077	0.2774	0.2614	0.2567	0.2514	0.2448	0.2403	0.2376	0.2351	0.2341	0.2327	0.2319	0.2317	0.2309	0.2303	0.2302	0.2301
17	0.3272	0.3022	0.2712	0.2539	0.2502	0.2449	0.2381	0.2339	0.2312	0.2289	0.2275	0.2261	0.2256	0.2251	0.2242	0.2238	0.2238	0.2236
18	0.3227	0.2968	0.2653	0.2475	0.2435	0.2384	0.2319	0.2278	0.2253	0.2228	0.2217	0.2203	0.2196	0.2191	0.2182	0.2183	0.2178	0.2179
19	0.3180	0.2919	0.2599	0.2415	0.2377	0.2325	0.2265	0.2220	0.2198	0.2171	0.2158	0.2147	0.2142	0.2136	0.2132	0.2128	0.2126	0.2125
20	0.3135	0.2875	0.2548	0.2360	0.2318	0.2270	0.2207	0.2167	0.2144	0.2123	0.2107	0.2094	0.2090	0.2088	0.2081	0.2077	0.2076	0.2077
21	0.3096	0.2829	0.2500	0.2303	0.2266	0.2219	0.2162	0.2120	0.2099	0.2072	0.2061	0.2053	0.2044	0.2042	0.2033	0.2031	0.2027	0.2028
22	0.3055	0.2789	0.2457	0.2260	0.2221	0.2172	0.2113	0.2075	0.2054	0.2026	0.2020	0.2008	0.1998	0.1997	0.1989	0.1986	0.1987	0.1984
23	0.3022	0.2754	0.2417	0.2211	0.2175	0.2126	0.2069	0.2032	0.2008	0.1988	0.1977	0.1964	0.1960	0.1956	0.1949	0.1944	0.1942	0.1942
24	0.2984	0.2718	0.2378	0.2169	0.2131	0.2087	0.2029	0.1991	0.1971	0.1948	0.1937	0.1926	0.1919	0.1916	0.1909	0.1906	0.1904	0.1904
25	0.2954	0.2684	0.2345	0.2128	0.2091	0.2046	0.1988	0.1954	0.1932	0.1910	0.1901	0.1890	0.1883	0.1880	0.1874	0.1871	0.1867	0.1867
26	0.2927	0.2654	0.2308	0.2089	0.2050	0.2005	0.1954	0.1918	0.1900	0.1874	0.1862	0.1853	0.1847	0.1844	0.1839	0.1835	0.1834	0.1834
27	0.2895	0.2622	0.2278	0.2054	0.2015	0.1974	0.1919	0.1884	0.1863	0.1842	0.1832	0.1820	0.1816	0.1811	0.1805	0.1802	0.1802	0.1801
28	0.2870	0.2593	0.2246	0.2018	0.1980	0.1940	0.1888	0.1851	0.1832	0.1811	0.1803	0.1790	0.1785	0.1782	0.1774	0.1775	0.1772	0.1769
29	0.2841	0.2566	0.2218	0.1986	0.1948	0.1908	0.1857	0.1820	0.1801	0.1781	0.1771	0.1761	0.1756	0.1752	0.1746	0.1742	0.1742	0.1741
30	0.2819	0.2540	0.2191	0.1955	0.1916	0.1875	0.1826	0.1791	0.1775	0.1752	0.1743	0.1733	0.1726	0.1722	0.1719	0.1715	0.1714	0.1714
35	0.2708	0.2426	0.2072	0.1819	0.1779	0.1743	0.1696	0.1666	0.1647	0.1629	0.1620	0.1609	0.1605	0.1602	0.1596	0.1594	0.1592	0.1591
40	0.2618	0.2333	0.1969	0.1710	0.1669	0.1635	0.1591	0.1563	0.1548	0.1528	0.1519	0.1510	0.1505	0.1503	0.1497	0.1495	0.1493	0.1495
45	0.2543	0.2254	0.1887	0.1619	0.1579	0.1548	0.1503	0.1476	0.1459	0.1444	0.1435	0.1426	0.1421	0.1420	0.1416	0.1413	0.1412	0.1411
50	0.2479	0.2187	0.1818	0.1542	0.1502	0.1470	0.1428	0.1405	0.1388	0.1372	0.1365	0.1357	0.1352	0.1350	0.1346	0.1343	0.1341	0.1343
60	0.2373	0.2078	0.1706	0.1414	0.1374	0.1345	0.1309	0.1286	0.1270	0.1257	0.1249	0.1242	0.1237	0.1236	0.1232	0.1229	0.1228	0.1229
70	0.2290	0.1992	0.1618	0.1316	0.1276	0.1250	0.1214	0.1193	0.1178	0.1167	0.1159	0.1152	0.1148	0.1147	0.1143	0.1141	0.1140	0.1140
80	0.2223	0.1924	0.1546	0.1237	0.1195	0.1170	0.1138	0.1118	0.1107	0.1093	0.1087	0.1080	0.1076	0.1074	0.1071	0.1069	0.1068	0.1068
90	0.2165	0.1866	0.1486	0.1170	0.1129	0.1105	0.1075	0.1055	0.1045	0.1033	0.1027	0.1020	0.1015	0.1015	0.1011	0.1009	0.1009	0.1008
100	0.2120	0.1817	0.1436	0.1114	0.1072	0.1048	0.1021	0.1002	0.0992	0.0980	0.0974	0.0968	0.0966	0.0964	0.0960	0.0959	0.0958	0.0958
200	0.1860	0.1552	0.1159	0.0814	0.0762	0.0747	0.0726	0.0714	0.0705	0.0698	0.0694	0.0689	0.0687	0.0686	0.0682	0.0682	0.0682	0.0682
300	0.1745	0.1434	0.1037	0.0682	0.0624	0.0611	0.0595	0.0584	0.0578	0.0571	0.0567	0.0564	0.0563	0.0561	0.0560	0.0558	0.0557	0.0558
400	0.1676	0.1363	0.0964	0.0603	0.0541	0.0531	0.0515	0.0507	0.0501	0.0495	0.0492	0.0489	0.0488	0.0487	0.0486	0.0485	0.0484	0.0484
500	0.1630	0.1316	0.0915	0.0550	0.0485	0.0474	0.0462	0.0453	0.0449	0.0444	0.0441	0.0438	0.0437	0.0436	0.0434	0.0433	0.0433	0.0434
1000	0.1514	0.1198	0.0792	0.0419	0.0343	0.0336	0.0327	0.0321	0.0318	0.0314	0.0313	0.0310	0.0310	0.0309	0.0308	0.0308	0.0307	0.0307
2500	0.1413	0.1095	0.0686	0.0304	0.0218	0.0213	0.0207	0.0203	0.0202	0.0199	0.0198	0.0197	0.0197	0.0196	0.0195	0.0195	0.0195	0.0195

Critical Values for Anderson Darling Test - Significance Level of 0.01

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.9603	1.0073	1.0852	1.1154	0.9876	0.9047	0.8544	0.8345	0.8262	0.8197	0.8164	0.8153	0.8135	0.8137	0.8112	0.8106	0.8099	0.8097
5	1.0754	1.0772	1.1053	1.1446	1.0682	1.0000	0.9453	0.9167	0.9054	0.8930	0.8897	0.8878	0.8822	0.8831	0.8817	0.8791	0.8791	0.8767
6	1.1951	1.1556	1.1420	1.1831	1.1214	1.0533	0.9896	0.9580	0.9461	0.9314	0.9275	0.9211	0.9173	0.9182	0.9159	0.9113	0.9119	0.9153
7	1.3145	1.2298	1.1755	1.2066	1.1562	1.0841	1.0186	0.9924	0.9785	0.9630	0.9510	0.9460	0.9414	0.9438	0.9383	0.9354	0.9377	0.9389
8	1.4299	1.3111	1.2089	1.2256	1.1807	1.1087	1.0439	1.0086	0.9905	0.9787	0.9699	0.9646	0.9602	0.9612	0.9553	0.9558	0.9527	0.9550
9	1.5478	1.3855	1.2502	1.2415	1.1959	1.1293	1.0584	1.0232	1.0075	0.9908	0.9844	0.9778	0.9739	0.9674	0.9685	0.9685	0.9669	0.9655
10	1.6581	1.4643	1.2809	1.2490	1.2139	1.1420	1.0707	1.0339	1.0181	1.0012	0.9944	0.9873	0.9829	0.9812	0.9769	0.9755	0.9732	0.9721
11	1.7667	1.5444	1.3164	1.2615	1.2233	1.1525	1.0789	1.0445	1.0271	1.0098	1.0031	0.9895	0.9878	0.9864	0.9840	0.9827	0.9812	0.9762
12	1.8736	1.6117	1.3464	1.2700	1.2310	1.1546	1.0876	1.0515	1.0327	1.0140	1.0061	1.0013	0.9955	0.9940	0.9883	0.9894	0.9860	0.9879
13	1.9804	1.6847	1.3763	1.2796	1.2370	1.1672	1.0983	1.0603	1.0358	1.0233	1.0115	1.0065	0.9999	0.9984	0.9939	0.9917	0.9891	0.9902
14	2.0727	1.7510	1.4057	1.2913	1.2452	1.1756	1.1010	1.0607	1.0403	1.0245	1.0157	1.0084	1.0054	1.0010	0.9959	0.9939	0.9918	0.9932
15	2.1736	1.8182	1.4338	1.2899	1.2476	1.1814	1.0996	1.0655	1.0480	1.0281	1.0184	1.0119	1.0065	1.0024	0.9990	0.9967	0.9986	0.9984
16	2.2603	1.8791	1.4693	1.3039	1.2534	1.1827	1.1117	1.0683	1.0470	1.0316	1.0186	1.0171	1.0104	1.0070	1.0043	0.9996	0.9986	1.0004
17	2.3532	1.9419	1.4945	1.3074	1.2604	1.1911	1.1104	1.0710	1.0526	1.0331	1.0230	1.0158	1.0109	1.0078	1.0045	1.0032	1.0005	1.0001
18	2.4406	2.0044	1.5210	1.3165	1.2604	1.1921	1.1158	1.0729	1.0536	1.0341	1.0270	1.0146	1.0150	1.0152	1.0061	1.0050	1.0033	1.0057
19	2.5322	2.0686	1.5482	1.3187	1.2670	1.1909	1.1153	1.0762	1.0592	1.0379	1.0259	1.0216	1.0180	1.0130	1.0096	1.0064	1.0080	1.0063
20	2.6114	2.1235	1.5704	1.3288	1.2680	1.1960	1.1178	1.0770	1.0563	1.0416	1.0311	1.0181	1.0193	1.0163	1.0121	1.0049	1.0091	1.0113
21	2.7041	2.1723	1.5955	1.3235	1.2697	1.1993	1.1264	1.0811	1.0572	1.0410	1.0310	1.0271	1.0200	1.0166	1.0127	1.0130	1.0081	1.0100
22	2.7796	2.2262	1.6194	1.3388	1.2791	1.2002	1.1191	1.0822	1.0621	1.0404	1.0359	1.0244	1.0199	1.0227	1.0141	1.0111	1.0125	1.0107
23	2.8617	2.2902	1.6460	1.3343	1.2809	1.2001	1.1249	1.0831	1.0589	1.0430	1.0342	1.0286	1.0243	1.0171	1.0196	1.0117	1.0127	1.0100
24	2.9336	2.3402	1.6668	1.3413	1.2767	1.2057	1.1257	1.0856	1.0648	1.0453	1.0346	1.0291	1.0194	1.0200	1.0155	1.0122	1.0127	1.0132
25	3.0189	2.3833	1.6899	1.3424	1.2808	1.2078	1.1272	1.0862	1.0643	1.0451	1.0376	1.0294	1.0256	1.0212	1.0171	1.0140	1.0127	1.0118
26	3.1002	2.4393	1.7106	1.3444	1.2777	1.2050	1.1302	1.0887	1.0669	1.0460	1.0369	1.0306	1.0248	1.0196	1.0152	1.0168	1.0158	1.0129
27	3.1672	2.4875	1.7364	1.3482	1.2853	1.2104	1.1299	1.0878	1.0661	1.0475	1.0393	1.0310	1.0257	1.0214	1.0160	1.0169	1.0125	1.0153
28	3.2487	2.5418	1.7524	1.3541	1.2839	1.2107	1.1342	1.0886	1.0667	1.0512	1.0406	1.0345	1.0266	1.0251	1.0200	1.0208	1.0194	1.0149
29	3.3187	2.5865	1.7783	1.3557	1.2859	1.2176	1.1353	1.0878	1.0670	1.0498	1.0410	1.0336	1.0267	1.0268	1.0194	1.0180	1.0171	1.0171
30	3.3926	2.6336	1.8001	1.3651	1.2861	1.2121	1.1326	1.0908	1.0724	1.0500	1.0438	1.0336	1.0261	1.0227	1.0231	1.0192	1.0179	1.0183
35	3.7444	2.8654	1.9035	1.3711	1.2858	1.2174	1.1360	1.0951	1.0715	1.0529	1.0450	1.0302	1.0290	1.0267	1.0246	1.0211	1.0180	1.0199
40	4.0854	3.0880	1.9882	1.3819	1.2938	1.2179	1.1375	1.0964	1.0760	1.0551	1.0457	1.0352	1.0324	1.0295	1.0272	1.0229	1.0221	1.0241
45	4.4084	3.2993	2.0772	1.3883	1.2976	1.2210	1.1413	1.0994	1.0744	1.0590	1.0476	1.0368	1.0340	1.0361	1.0300	1.0263	1.0245	1.0212
50	4.7335	3.4996	2.1620	1.4067	1.3038	1.2232	1.1419	1.1011	1.0786	1.0595	1.0526	1.0403	1.0381	1.0338	1.0295	1.0277	1.0248	1.0252
60	5.3343	3.8894	2.3298	1.4187	1.3079	1.2229	1.1435	1.1034	1.0790	1.0619	1.0540	1.0426	1.0381	1.0323	1.0317	1.0291	1.0302	1.0244
70	5.9151	4.2582	2.4835	1.4298	1.3071	1.2296	1.1446	1.1041	1.0785	1.0604	1.0549	1.0438	1.0382	1.0376	1.0311	1.0311	1.0280	1.0276
80	6.5032	4.6205	2.6216	1.4453	1.3016	1.2303	1.1499	1.1044	1.0849	1.0641	1.0550	1.0452	1.0372	1.0360	1.0328	1.0316	1.0294	1.0287
90	7.0504	4.9602	2.7593	1.4575	1.3124	1.2311	1.1486	1.1077	1.0861	1.0660	1.0558	1.0455	1.0374	1.0381	1.0345	1.0311	1.0326	1.0310
100	7.6095	5.3024	2.8954	1.4713	1.3083	1.2288	1.1492	1.1072	1.0851	1.0648	1.0542	1.0464	1.0417	1.0421	1.0353	1.0330	1.0325	1.0316
200	12.7383	8.4639	4.1296	1.5840	1.3097	1.2364	1.1565	1.1103	1.0885	1.0668	1.0587	1.0507	1.0450	1.0412	1.0314	1.0322	1.0329	1.0323
300	17.5414	11.3900	5.2242	1.6967	1.3138	1.2409	1.1535	1.1113	1.0898	1.0679	1.0582	1.0489	1.0468	1.0431	1.0376	1.0328	1.0313	1.0352
400	22.1764	14.1813	6.2523	1.7932	1.3205	1.2404	1.1579	1.1151	1.0928	1.0677	1.0573	1.0480	1.0424	1.0433	1.0390	1.0350	1.0336	1.0329
500	26.7379	16.9148	7.2528	1.8846	1.3188	1.2395	1.1552	1.1139	1.0886	1.0695	1.0570	1.0498	1.0484	1.0466	1.0402	1.0339	1.0337	1.0382
1000	48.7347	30.0039	11.9358	2.2962	1.3248	1.2438	1.1570	1.1103	1.0923	1.0676	1.0601	1.0480	1.0496	1.0430	1.0347	1.0358	1.0309	1.0329
2500	111.798	67.1014	24.8571	3.3313	1.3247	1.2420	1.1559	1.1102	1.0900	1.0683	1.0606	1.0495	1.0552	1.0476	1.0351	1.0345	1.0383	1.0364

Critical Values for Kolmogorov Smirnov Test - Significance Level of 0.01

n\k	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.750	1.000	1.500	2.000	3.000	4.000	5.000	10.000	20.000	50.000	100.000
4	0.4698	0.4724	0.4853	0.4961	0.4783	0.4662	0.4552	0.4491	0.4458	0.4426	0.4409	0.4394	0.4387	0.4384	0.4373	0.4368	0.4365	0.4365
5	0.4641	0.4581	0.4536	0.4559	0.4509	0.4415	0.4314	0.4244	0.4207	0.4157	0.4139	0.4121	0.4103	0.4104	0.4097	0.4083	0.4080	0.4077
6	0.4528	0.4411	0.4306	0.4314	0.4234	0.4137	0.4022	0.3947	0.3912	0.3873	0.3852	0.3833	0.3821	0.3819	0.3808	0.3801	0.3797	0.3805
7	0.4367	0.4207	0.4065	0.4041	0.3989	0.3902	0.3800	0.3738	0.3694	0.3651	0.3624	0.3604	0.3593	0.3599	0.3576	0.3568	0.3571	0.3575
8	0.4235	0.4066	0.3879	0.3843	0.3789	0.3700	0.3604	0.3535	0.3493	0.3456	0.3436	0.3420	0.3408	0.3403	0.3393	0.3388	0.3383	0.3385
9	0.4122	0.3928	0.3726	0.3655	0.3603	0.3530	0.3427	0.3362	0.3330	0.3289	0.3269	0.3253	0.3240	0.3233	0.3227	0.3222	0.3219	0.3218
10	0.4019	0.3816	0.3580	0.3497	0.3450	0.3378	0.3279	0.3212	0.3184	0.3143	0.3125	0.3107	0.3101	0.3095	0.3081	0.3076	0.3074	0.3071
11	0.3925	0.3713	0.3461	0.3355	0.3314	0.3238	0.3144	0.3085	0.3053	0.3018	0.2999	0.2976	0.2971	0.2965	0.2954	0.2949	0.2953	0.2942
12	0.3844	0.3613	0.3348	0.3231	0.3186	0.3110	0.3024	0.2974	0.2936	0.2898	0.2880	0.2871	0.2857	0.2853	0.2843	0.2841	0.2837	0.2835
13	0.3762	0.3530	0.3248	0.3121	0.3071	0.3008	0.2927	0.2868	0.2833	0.2803	0.2783	0.2764	0.2758	0.2755	0.2740	0.2738	0.2737	0.2736
14	0.3685	0.3447	0.3160	0.3019	0.2976	0.2910	0.2833	0.2768	0.2741	0.2709	0.2694	0.2674	0.2670	0.2662	0.2651	0.2645	0.2643	0.2646
15	0.3622	0.3379	0.3076	0.2921	0.2884	0.2820	0.2736	0.2689	0.2657	0.2626	0.2606	0.2596	0.2585	0.2577	0.2572	0.2566	0.2563	0.2564
16	0.3556	0.3310	0.3009	0.2845	0.2798	0.2738	0.2663	0.2609	0.2578	0.2547	0.2535	0.2520	0.2510	0.2507	0.2499	0.2491	0.2486	0.2489
17	0.3502	0.3250	0.2939	0.2767	0.2725	0.2669	0.2592	0.2538	0.2508	0.2480	0.2463	0.2448	0.2442	0.2436	0.2428	0.2424	0.2422	0.2419
18	0.3448	0.3192	0.2879	0.2696	0.2655	0.2597	0.2524	0.2472	0.2445	0.2415	0.2403	0.2383	0.2376	0.2374	0.2363	0.2362	0.2359	0.2357
19	0.3399	0.3139	0.2819	0.2632	0.2592	0.2534	0.2461	0.2410	0.2383	0.2353	0.2337	0.2325	0.2322	0.2315	0.2307	0.2302	0.2301	0.2299
20	0.3350	0.3093	0.2764	0.2572	0.2529	0.2475	0.2403	0.2356	0.2328	0.2301	0.2285	0.2267	0.2265	0.2262	0.2254	0.2247	0.2248	0.2247
21	0.3308	0.3041	0.2709	0.2510	0.2474	0.2416	0.2352	0.2303	0.2277	0.2248	0.2235	0.2223	0.2214	0.2211	0.2204	0.2199	0.2195	0.2195
22	0.3265	0.2998	0.2666	0.2460	0.2423	0.2363	0.2297	0.2256	0.2229	0.2198	0.2190	0.2175	0.2163	0.2161	0.2156	0.2150	0.2152	0.2148
23	0.3226	0.2960	0.2621	0.2411	0.2372	0.2313	0.2250	0.2208	0.2180	0.2155	0.2145	0.2130	0.2124	0.2117	0.2112	0.2105	0.2103	0.2103
24	0.3183	0.2923	0.2580	0.2365	0.2323	0.2271	0.2208	0.2161	0.2140	0.2114	0.2098	0.2087	0.2077	0.2076	0.2067	0.2065	0.2062	0.2064
25	0.3153	0.2880	0.2540	0.2317	0.2284	0.2229	0.2164	0.2121	0.2099	0.2073	0.2059	0.2047	0.2039	0.2035	0.2031	0.2027	0.2025	0.2023
26	0.3120	0.2848	0.2501	0.2279	0.2235	0.2188	0.2126	0.2085	0.2061	0.2033	0.2022	0.2009	0.2002	0.1997	0.1990	0.1988	0.1987	0.1986
27	0.3087	0.2813	0.2471	0.2241	0.2199	0.2150	0.2088	0.2048	0.2022	0.1997	0.1986	0.1972	0.1967	0.1964	0.1955	0.1952	0.1952	0.1950
28	0.3058	0.2783	0.2434	0.2203	0.2158	0.2115	0.2055	0.2012	0.1989	0.1966	0.1955	0.1941	0.1934	0.1930	0.1924	0.1925	0.1921	0.1917
29	0.3027	0.2749	0.2404	0.2166	0.2125	0.2082	0.2021	0.1976	0.1955	0.1931	0.1923	0.1909	0.1904	0.1899	0.1892	0.1887	0.1889	0.1886
30	0.3000	0.2723	0.2374	0.2132	0.2092	0.2047	0.1987	0.1946	0.1926	0.1902	0.1890	0.1878	0.1870	0.1865	0.1862	0.1860	0.1854	0.1856
35	0.2878	0.2597	0.2242	0.1984	0.1941	0.1901	0.1847	0.1812	0.1788	0.1769	0.1757	0.1742	0.1741	0.1737	0.1730	0.1728	0.1724	0.1724
40	0.2780	0.2495	0.2128	0.1865	0.1822	0.1782	0.1733	0.1699	0.1682	0.1661	0.1649	0.1638	0.1635	0.1628	0.1622	0.1620	0.1620	0.1620
45	0.2695	0.2408	0.2041	0.1765	0.1721	0.1688	0.1637	0.1605	0.1584	0.1570	0.1559	0.1547	0.1542	0.1541	0.1536	0.1533	0.1531	0.1529
50	0.2626	0.2332	0.1964	0.1683	0.1641	0.1604	0.1557	0.1528	0.1511	0.1490	0.1483	0.1471	0.1470	0.1463	0.1460	0.1456	0.1455	0.1455
60	0.2509	0.2213	0.1840	0.1544	0.1501	0.1466	0.1425	0.1399	0.1380	0.1364	0.1357	0.1349	0.1343	0.1340	0.1336	0.1333	0.1333	0.1331
70	0.2416	0.2118	0.1743	0.1435	0.1395	0.1362	0.1322	0.1298	0.1281	0.1268	0.1259	0.1250	0.1248	0.1244	0.1240	0.1238	0.1236	0.1235
80	0.2343	0.2043	0.1662	0.1350	0.1303	0.1276	0.1240	0.1216	0.1203	0.1189	0.1180	0.1172	0.1167	0.1166	0.1162	0.1161	0.1157	0.1158
90	0.2277	0.1978	0.1594	0.1278	0.1232	0.1207	0.1170	0.1148	0.1135	0.1122	0.1114	0.1107	0.1102	0.1101	0.1098	0.1094	0.1096	0.1093
100	0.2228	0.1923	0.1541	0.1216	0.1169	0.1143	0.1112	0.1090	0.1078	0.1065	0.1058	0.1052	0.1049	0.1046	0.1043	0.1041	0.1038	0.1039
200	0.1938	0.1628	0.1235	0.0888	0.0831	0.0815	0.0791	0.0776	0.0767	0.0758	0.0753	0.0748	0.0746	0.0744	0.0741	0.0740	0.0739	0.0739
300	0.1808	0.1496	0.1101	0.0742	0.0680	0.0667	0.0648	0.0635	0.0628	0.0621	0.0616	0.0612	0.0611	0.0609	0.0607	0.0606	0.0604	0.0606
400	0.1731	0.1418	0.1019	0.0657	0.0591	0.0579	0.0562	0.0551	0.0545	0.0537	0.0534	0.0531	0.0529	0.0528	0.0526	0.0526	0.0525	0.0525
500	0.1680	0.1365	0.0963	0.0598	0.0529	0.0517	0.0503	0.0493	0.0487	0.0482	0.0478	0.0476	0.0474	0.0473	0.0471	0.0470	0.0470	0.0471
1000	0.1549	0.1234	0.0827	0.0452	0.0375	0.0367	0.0356	0.0349	0.0345	0.0341	0.0340	0.0337	0.0336	0.0336	0.0333	0.0333	0.0333	0.0333
2500	0.1436	0.1118	0.0708	0.0325	0.0238	0.0233	0.0226	0.0221	0.0219	0.0216	0.0215	0.0213	0.0213	0.0213	0.0211	0.0211	0.0211	0.0211

APPENDIX C

GRAPHS

OF

COVERAGE COMPARISONS

FOR THE VARIOUS METHODS

FOR

NORMAL, GAMMA, AND LOGNORMAL

DISTRIBUTIONS

Figure 1. Graphs of Coverage Probabilities by 95% UCLs of the Mean of N(μ=50,σ=20)

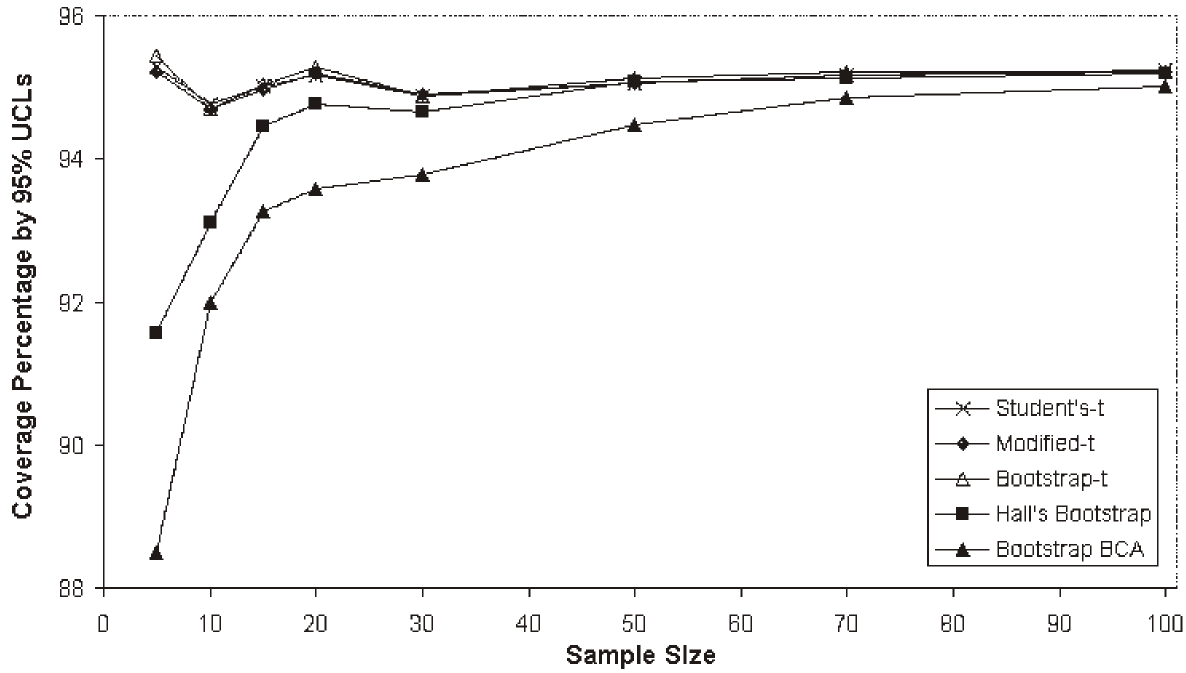


Figure 2. Graphs of Coverage Probabilities by 95% UCLs of Mean of G(k=0.05,θ=50)

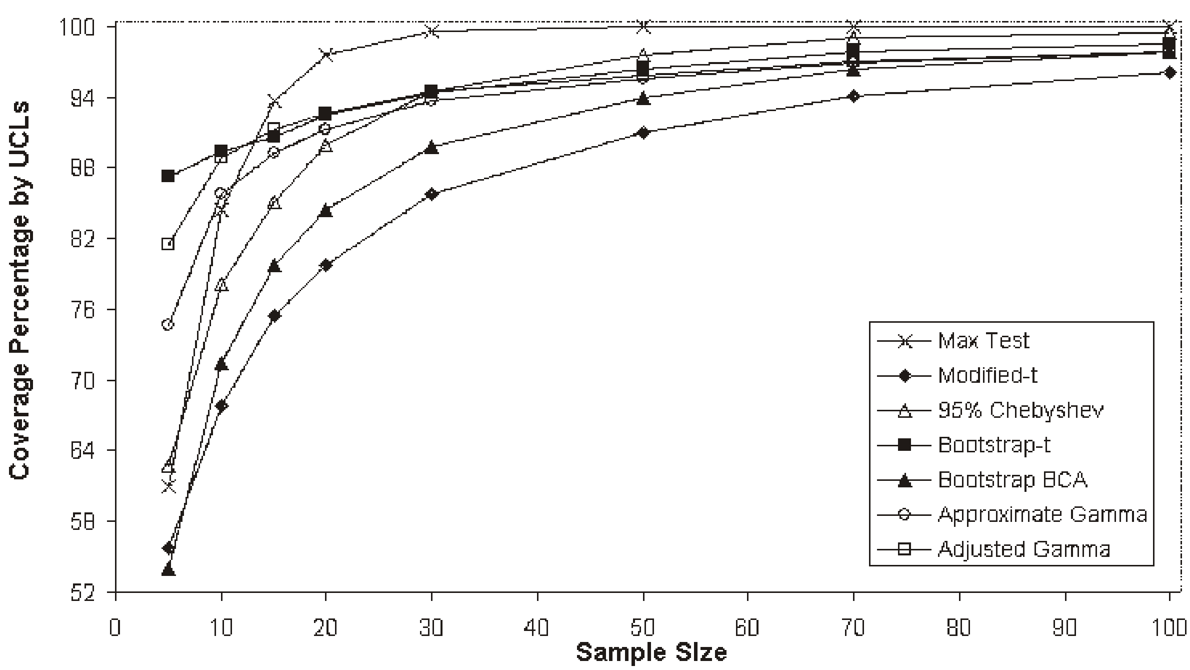


Figure 3. Graphs of Coverage Probabilities by 95% UCLs of the Mean of G(k=0.10,θ=50)

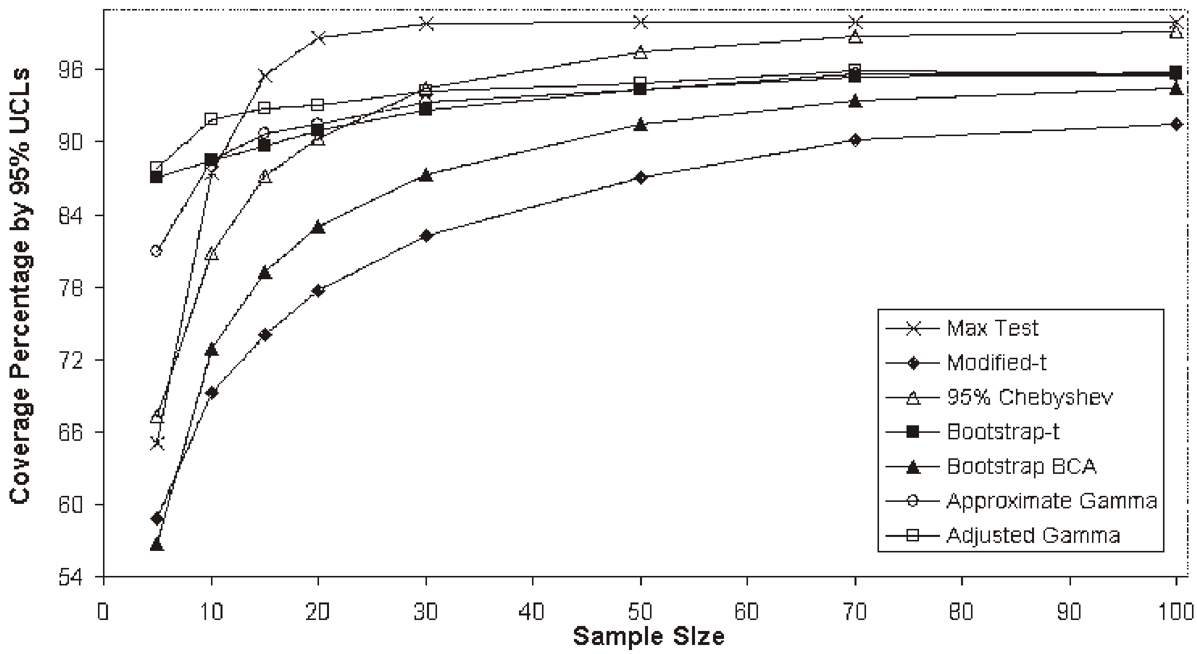


Figure 4. Graphs of Coverage Probabilities by 95% UCLs of Mean of G(k=0.15,θ=50)

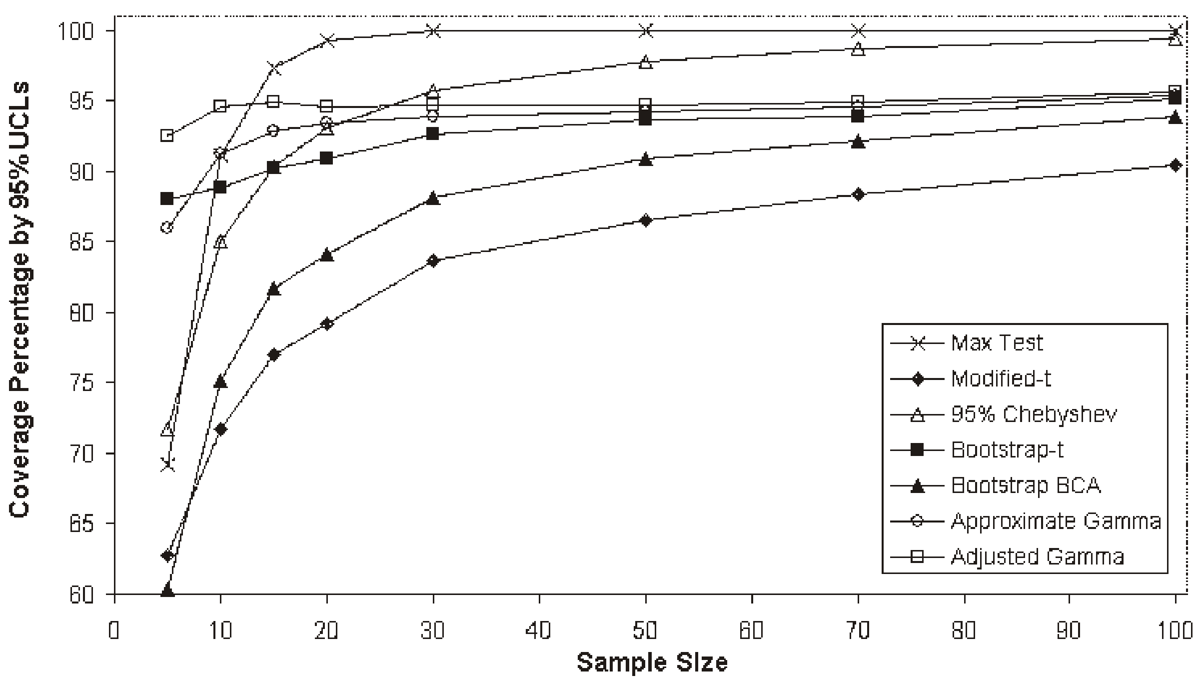


Figure 5. Graphs of Coverage Probabilities by 95% UCLs of the Mean of G(k=0.20,θ=50)

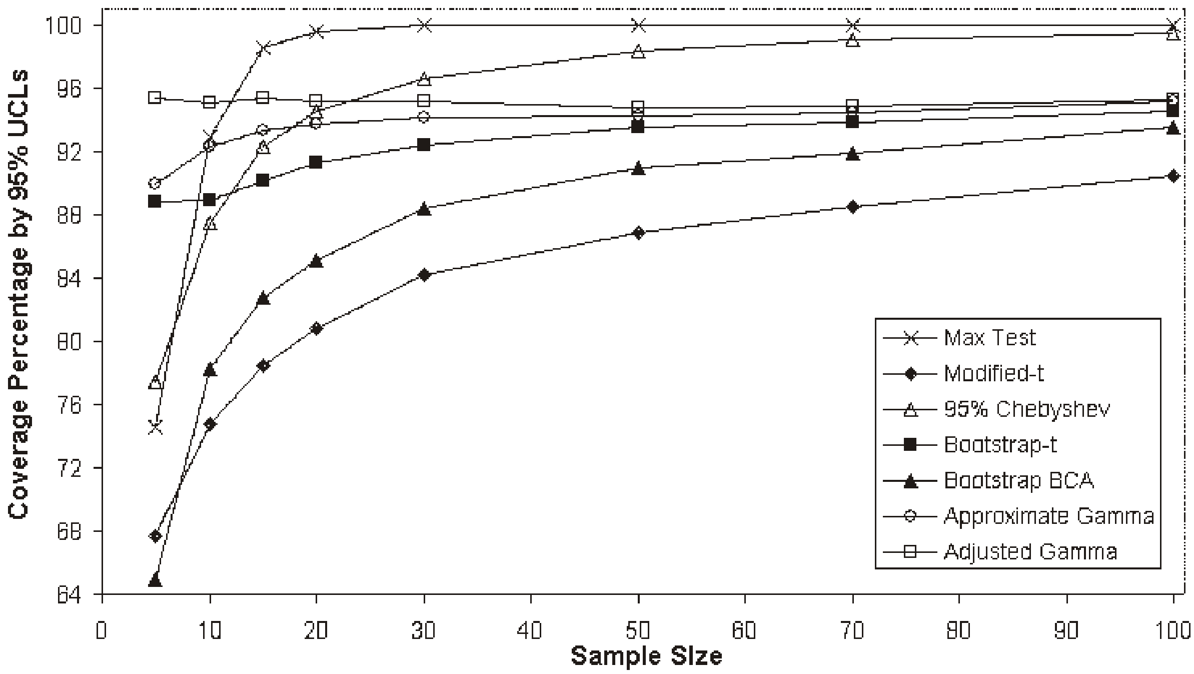


Figure 6. Graphs of Coverage Probabilities by 95% UCLs of Mean of G(k=0.50,θ=50)

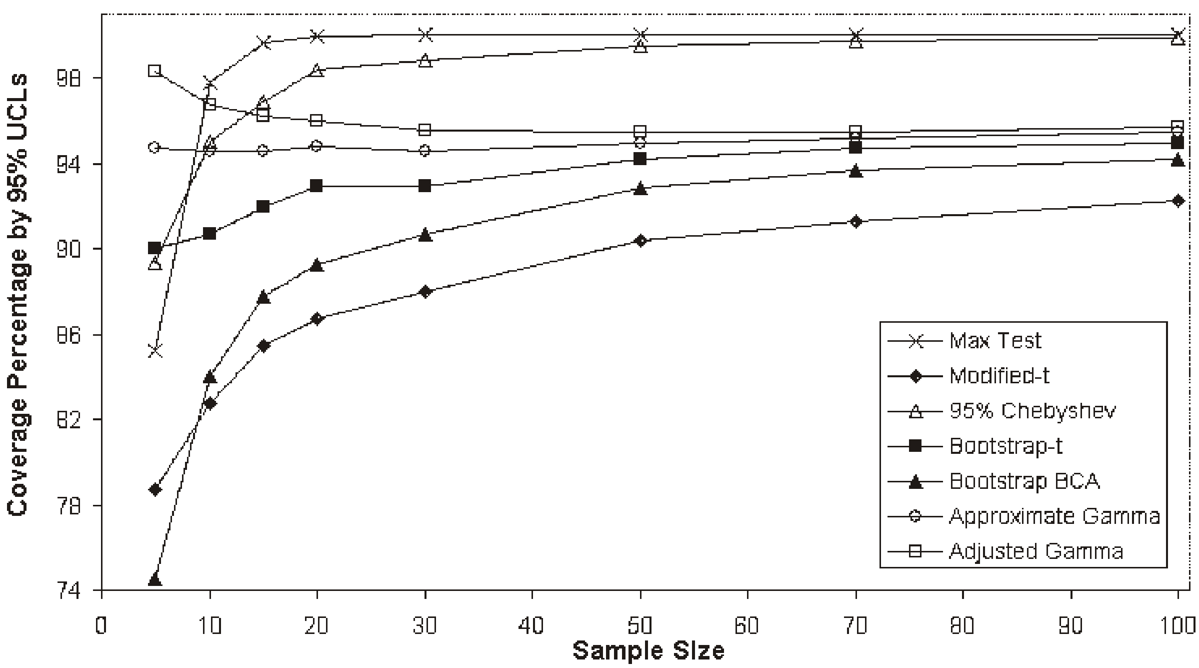


Figure 7. Graphs of Coverage Probabilities by 95% UCLs of the Mean of G(k=1.00,θ=50)

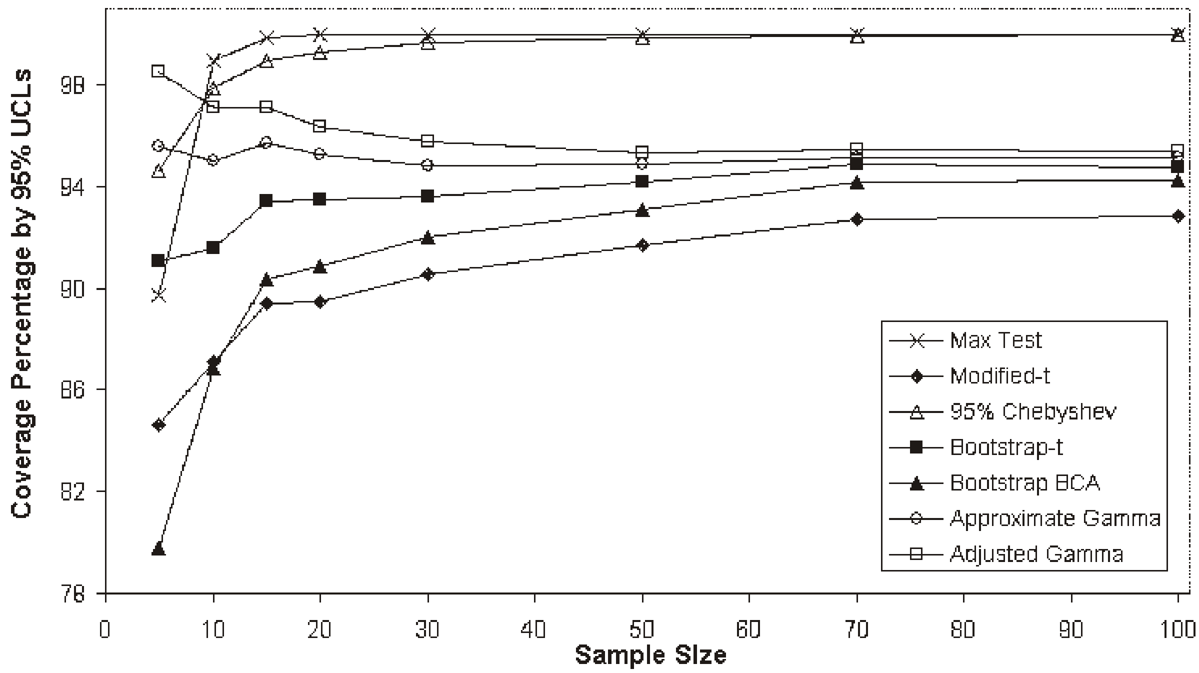


Figure 8. Graphs of Coverage Probabilities by 95% UCLs of the Mean of G(k=2.00,θ=50)

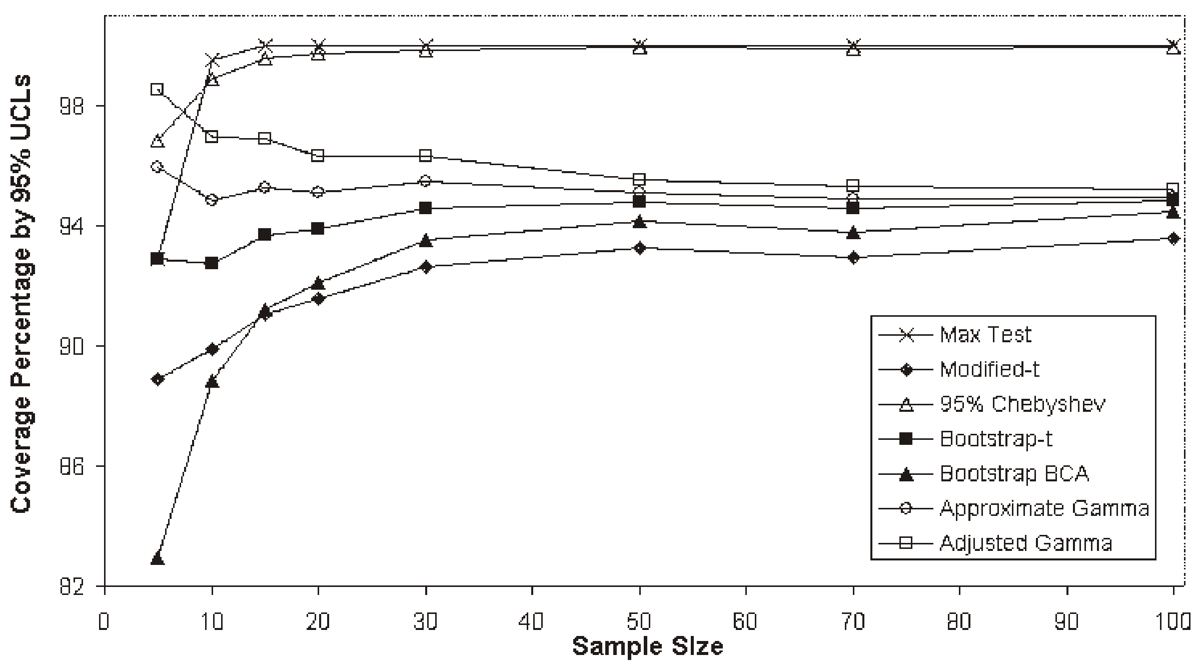


Figure 9. Graphs of Coverage Probabilities by 95% UCLs of the Mean of G(k=5.00,θ=50)

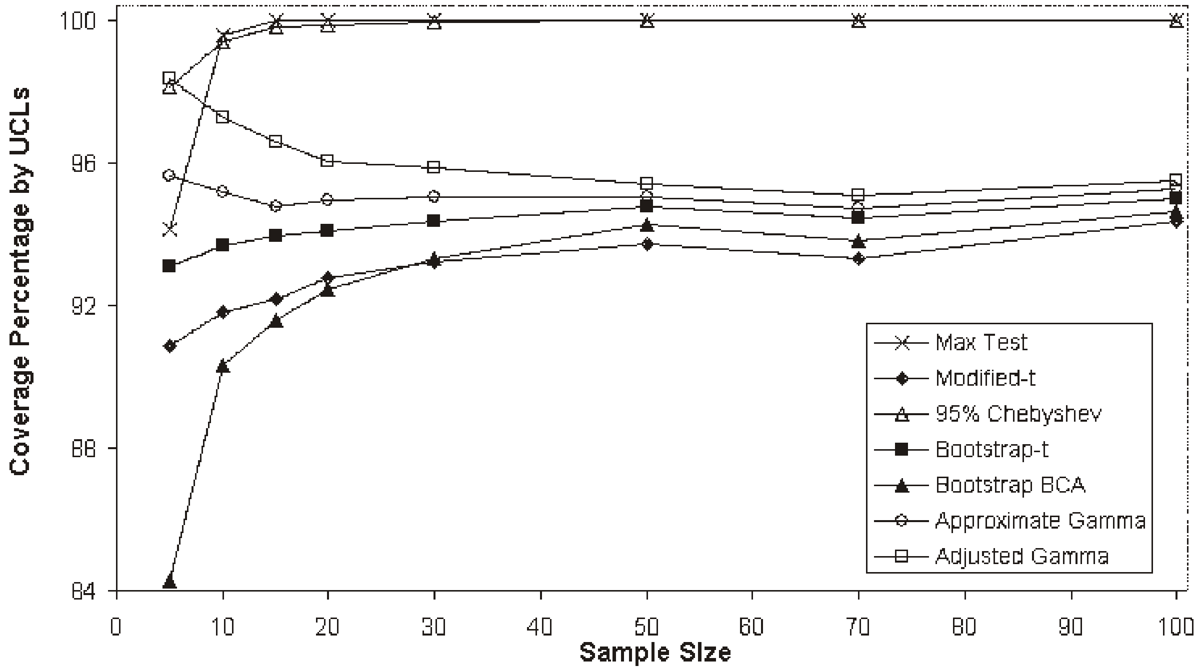


Figure 10. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=0.5)

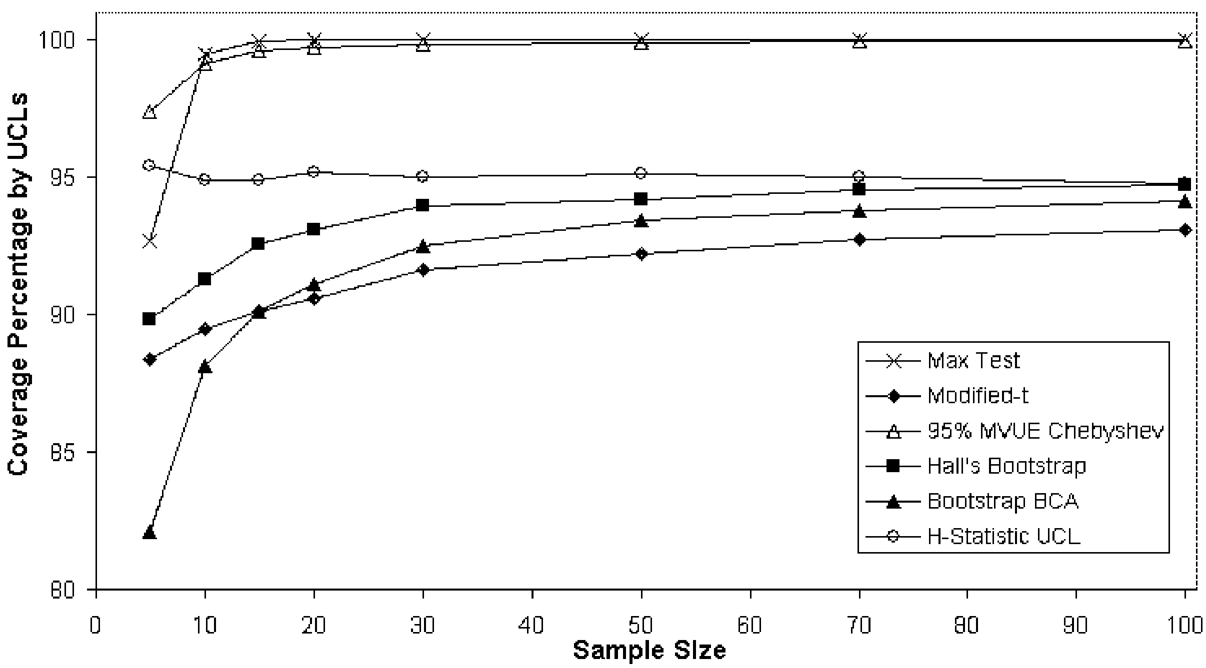


Figure 11. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=1.0)

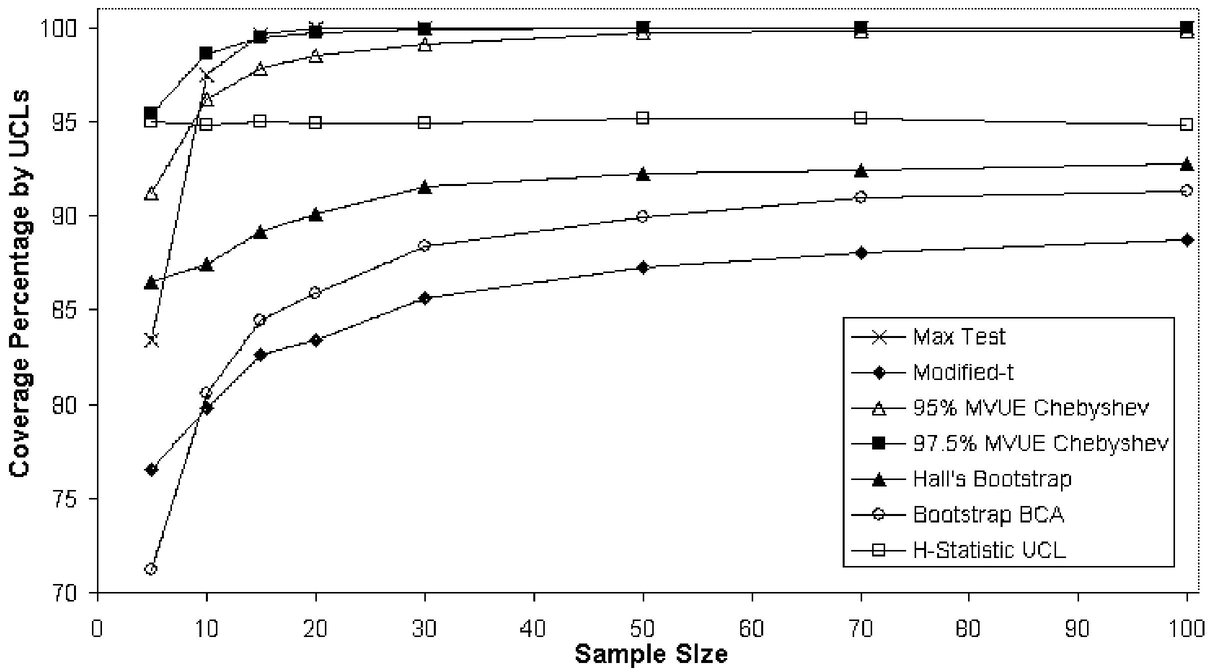


Figure 12. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=1.5)

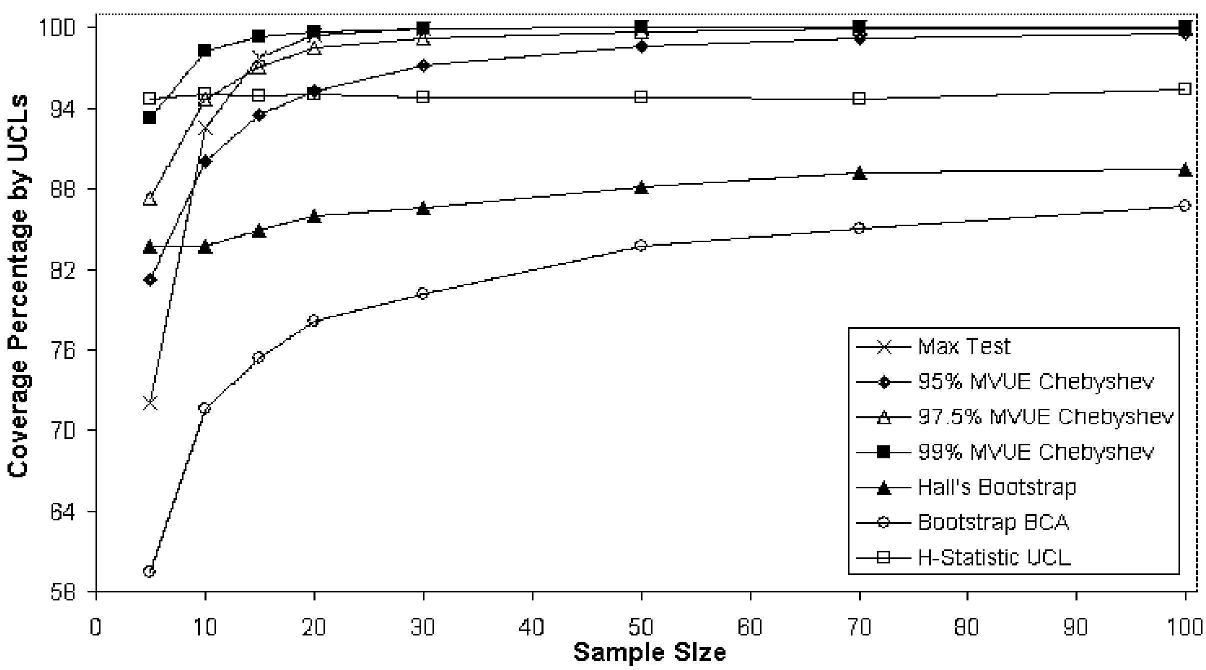


Figure 13. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=2.0)

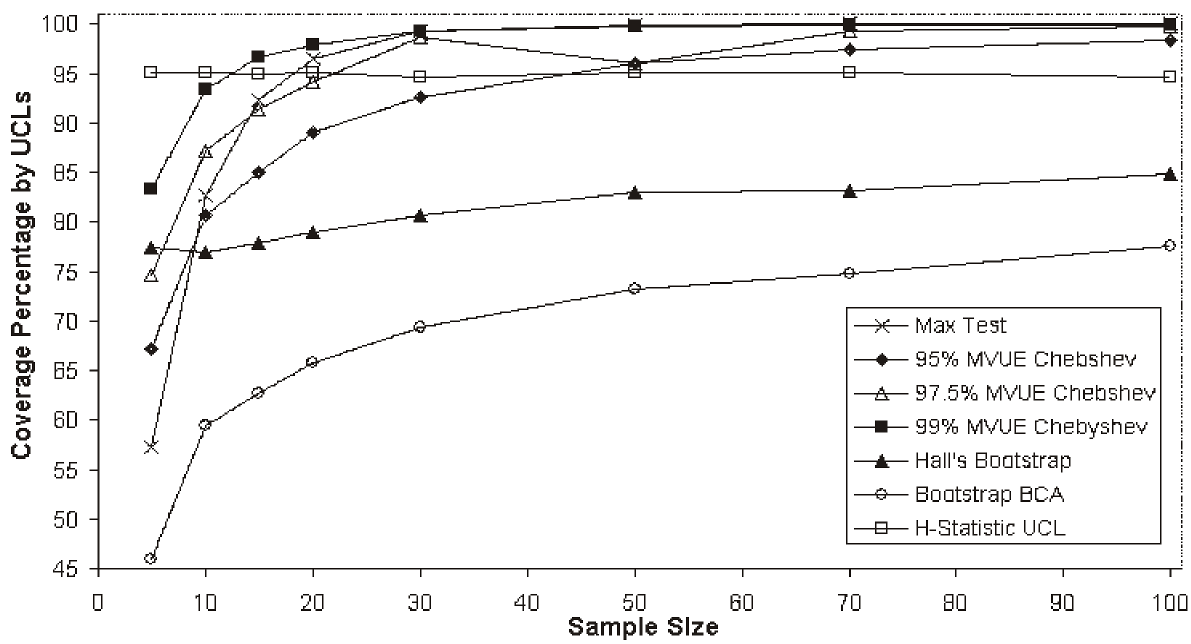


Figure 14. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=2.5)

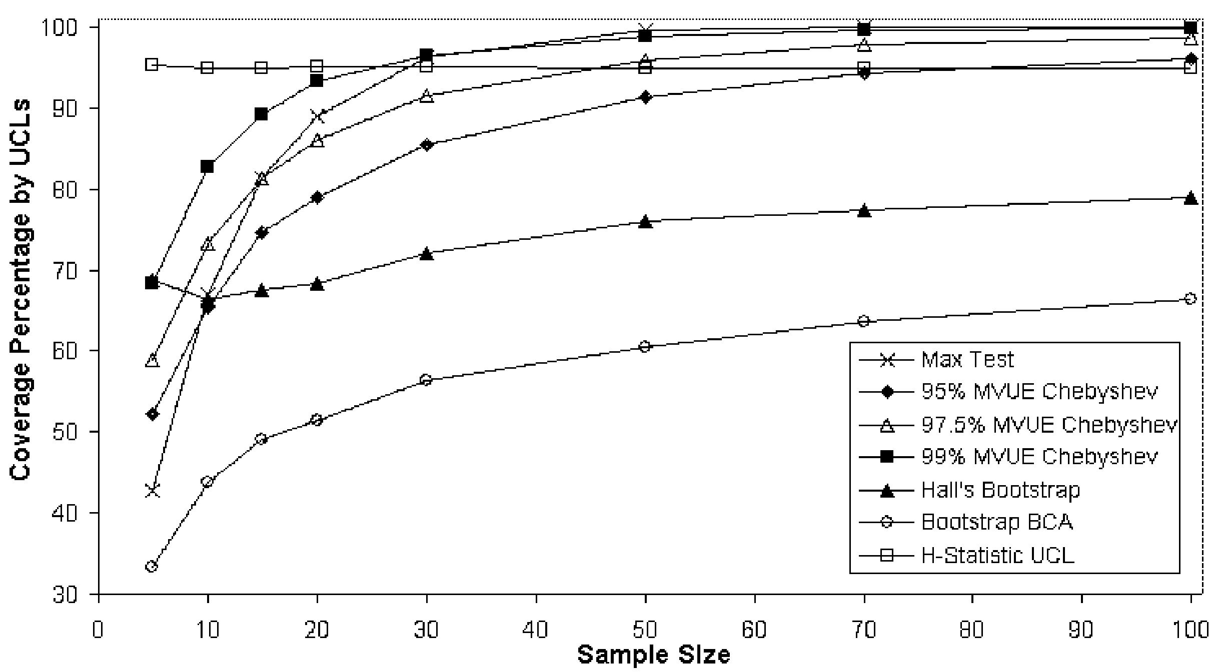


Figure 15. Graphs of Coverage Probabilities by UCLs of the Mean of LN(μ=5,σ=3.0)