APPENDIX K
Intervals and Limits

K-1. Introduction. Statistics can be divided into two categories: estimation theory (descriptive statistics) and hypothesis testing (inferential statistics). Estimation theory includes calculating confidence intervals as estimates for population parameters, while hypothesis testing focuses on the use of statistical tests to accept or reject hypotheses concerning these parameters.

K-2. Types of Statistical Intervals. Three types of statistical intervals are often constructed on data: confidence intervals, tolerance intervals, and prediction intervals. A confidence interval is designed to contain the specified population parameter, such as the mean, with a specified level of confidence. A confidence interval for the mean, for example, gives information about the average concentration level but offers little information about the highest or most extreme sample concentrations that are likely to be observed. In such cases, tolerance or prediction intervals are more appropriate. A confidence interval contains a parameter of interest, while a tolerance interval contains a proportion of the population, and a prediction interval contains one or more future observations. Statistical intervals are dependent upon distributional assumptions. Parametric and nonparametric methods for deriving intervals are also available. However, some nonparametric intervals, such as the tolerance interval, require a large number of observations to provide a reasonable coverage and confidence level. More information about statistical intervals can be found in Hahn and Meeker (1991).

K-2.1. *Confidence Interval*. It is often desirable to express or quantify the degree of uncertainty for some estimate of an unknown population parameter. The most common type of interval estimate is a confidence interval. A confidence interval is essentially an estimate for an unknown population parameter expressed as a range of values with some specified level of confidence. The level of confidence describes the probability that the "interval will capture the true parameter in repeated samples" (Moore, 1999).

K-2.1.1. The values at each end of the interval are called confidence limits. The lower value is the *lower confidence limit* (LCL) and the upper value is the *upper confidence limit* (UCL). The calculation of a confidence limit depends on the sampling distribution. Confidence limits are readily calculated for normally distributed data. A two-sided confidence interval for some population parameter, Θ, will be a closed interval of the form $a \le \Theta \le b$, where a is the lower limit and b is the upper limit. An upper one-sided confidence interval will be of the form $\Theta \le b$ and a lower one-sided confidence interval will be of the form $\Theta \ge a$.

K-2.1.2. For environmental work, it is often desirable to estimate the mean concentration of a contaminant in some environmental population (for example, the mean concentration of arsenic in a shallow groundwater aquifer). The population mean (μ) is often estimated by calculating the sample mean ($\bar{x}$) for a set of n measurements. The uncertainty associated with the

sample mean (as an estimate of the population mean) would be addressed by constructing a confidence interval for the population mean. A note on terminology: one calculates a confidence interval for a *population parameter,* such as the population mean, and *not* for the corresponding sample statistic, such as the sample mean (though a statistic such as the sample mean may be used to calculate the confidence interval for the population parameter).

K-2.1.3. The upper bound of the confidence interval of the population mean, the UCL, is most frequently encountered. For example, risk assessments require the 95% UCL for use as the reasonable maximum exposure concentration. The UCL of the (population) mean is used for the exposure point concentration (EPC) in risk assessments because of the uncertainty associated with estimating the population or "true" mean concentration at a site (EPA OB92-963373). Recent EPA guidance directs risk assessors in the possible methods used to calculate an upper confidence limit on the population mean (EPA OSWER 9285.6-10).

K-2.1.4. The phrase "95% confidence interval" means that "if one repeatedly calculates such intervals from many sets of independent random samples," 95% of the intervals, "in the long run, correctly contain the parameter of interest" (Hahn and Meeker, 1991). In other words, if a very large number of 95% confidence limits are calculated for the population mean, approximately 95% of the intervals (95 intervals out of 100) will contain the population mean. "More commonly, but less precisely, a two-sided confidence interval is described by a statement such as 'we are 95% confident that the interval contains the parameter of interest.' In fact, either the observed interval contains the parameter or it does not. Thus the 95% refers to the procedure for constructing a statistical interval, and not to the observed interval itself" (Hahn and Meeker, 1991). Because not all data sets fit a normal distribution, formulas for calculating a lognormal and nonparametric confidence limit are also available.

K-2.1.5. The EPA recently published OSWER 9285.6-10. According to this latest guidance, calculating a UCL should take into consideration outliers, censored data, and distribution testing (as described in Appendices I, H, and E). Once the distribution is determined, the calculation of an UCL should proceed according to the procedures for distributional methods. If, however, the site data do not follow a known distribution, then determining a good estimate of the UCL is left to the discretion of the risk assessor. Table K-1 presents the methods recommended in EPA guidance (OSWER 9285.6-10). Research in the area of UCL calculation is ongoing and recommendations may change in the future.

K-2.2. *Tolerance Interval*. A tolerance interval is designed to contain a specified proportion of the population (or percentile), such as 95% of all possible sample measurements (i.e., the 95^{th} percentile). Tolerance intervals are essentially confidence intervals around a specified percentile. It is rare that a quantile for the population is known; instead, it is estimated using a sample data set, and a confidence interval for the population quantile is calculated using the sample quantile (e.g., just as a confidence interval for the population mean is calculated using the sample

mean). Tolerance intervals are usually designed to cover all but a small percentage of the population measurements, so observations should rarely exceed a tolerance interval if the observations come from a similar distribution.

Table K-1.
UCL Method Flow Chart

Are data normal?	**Yes** →	Use Student's *t*
No ↓		
Are data lognormal?	**Yes** →	Use Land, Chebyshev (MVUE), or Student's *t* (with small variance and skewness)
No ↓		
Is another distribution appropriate?	**Yes** →	Use distribution-specific method (if available)
No ↓		
Is sample size large?	**Yes** →	Use Central Limit Theorem-Adjusted (with small variance and mild skewness) or Chebyshev
No ↓		
→		Use Chebyshev, Bootstrap Resampling, or Jackknife

K-2.2.1. A tolerance interval is characterized by two quantities (probabilities): the coverage (the proportion of the population that the interval is supposed to contain), and the confidence level (the degree of confidence with which the interval reaches the specified coverage). As the interval is constructed from sample information, it is also a random interval. Because of sample fluctuations, a tolerance interval can contain the specified proportion of the population only with a certain confidence level. For example, "the $(1 - \alpha)100\%$ tolerance interval with $p100\%$ coverage" refers to a tolerance interval constructed to contain at least $100p\%$ of the distribution with $(1 - \alpha)100\%$ level of confidence.

K-2.2.2. Upper tolerance limits (UTLs) (UCLs for percentiles) are often calculated for environmental work. For example, it may be desirable to compare contaminant concentrations in a study area to the UTL of the compound in a background area. If the concentrations of many site samples exceed the background UTL, site-related contamination probably exists. It is most common for environmental scientists to calculate the "95 UTL" (95% upper tolerance limit with 95% coverage).

K-2.2.3. The method for calculating a tolerance interval depends on the nature of the underlying population distribution. Tolerance intervals can be constructed assuming that the data or

the transformed data are normally distributed. It is also possible to construct nonparametric tolerance intervals using only the assumption that the data come from some continuous population. However, nonparametric tolerance intervals often require a large number of observations to provide a reasonable coverage and are impractical to construct for small sets of data. The data set with which tolerance intervals are calculated should be inspected for outliers and tested for normality before selecting the tolerance interval approach.

K-2.3. *Prediction Interval*. A prediction interval is a statistical interval calculated to include one or more future observations from the same population with a specified confidence. A prediction interval calculated from some set of sample data is such that all of certain number of future measurements (k) from the same population will fall within the interval with some specified level of confidence. In other words, each k future observation is compared to the prediction interval. The interval is constructed to contain all k future observations with the stated confidence. If *any* future observation exceeds the prediction interval, this is statistically significant evidence of a change in conditions. The number of future observations to be collected, k, must be specified (i.e., known before calculating the prediction interval). It is desirable to calculate prediction intervals periodically, using the most recent data. (The EPA recommends at least yearly for groundwater analyses.) Concentrations of site contaminants are sometimes compared to background concentrations using prediction intervals. An upper prediction limit is calculated for the next k future observations using the background data set and the k site measurements are then compared to the upper prediction limit. If any of the k site measurements exceed the prediction limit, this suggests that the site concentrations are elevated with respect to background.

K-2.3.1. Prediction intervals are used to achieve some desired tolerance for Type I error (i.e., false rejection of H_0) when the same statistical test is performed multiple times (e.g., k times). (Neither prediction nor tolerance intervals address Type II error.) For example, assume that the Type I error rate is α for falsely rejecting the null hypothesis, H_0, for some statistical test or comparison. Assume that k independent statistical tests or comparisons are performed, where α denotes the probability of a false rejection (Type I error rate) for each individual test or comparison. The Type I error for the set of k independent comparisons, α^*, is the following:

$$\alpha^* = 1 - (1 - \alpha)^k .$$

K-2.3.2. Consider a single statistical test comparing populations 1 and 2, where H_0 is rejected at a level of significance $\alpha = 0.05$. Now, suppose that three, rather than two, populations are to be compared to each other using the same α for each comparison; that is, populations 1 and 2, 2 and 3, and 1 and 3, are compared, where $\alpha = 0.05$ for each of the $k = 3$ comparisons. Assume that the three populations are identical and all the measurements are independent of one another. The probability of rejecting H_0 for at least one of the three populations (i.e., the false rejection rate for the set of three comparisons) is

$$\alpha^* = 1-(1-\alpha)^3 = 1-(0.95)^3 = 0.14.$$

K-2.3.3. Even though the false rejection rate for a *single* comparison is 0.05, the false rejection rate for the *set* of three comparisons is higher, 0.14. A larger false positive rate will be obtained when more than three different populations are being compared. Therefore, if a total false rejection rate of $\alpha = 0.05$ is desired, the false rejection rate for each comparison must be less than 0.05. In fact, it can be shown that if a total false rejection rate (also called the experiment wise error rate) of α is desired, then the false rejection rate, α^*, for each comparison should be approximately α/k:

$$\alpha = 1-(1-\alpha^*)^k \approx 1-(1-\alpha/k)^k \ .$$

This is called the Bonferroni approximation. For example, if $\alpha = 0.05$ and $k = 3$, then the Type I error for each individual comparison (α^*) must be approximately 0.05/3 = 0.0167. Note that

$$1-(1-\alpha/k)^k = 1-(0.05/3)^3 = 0.049 \approx 0.05.$$

K-2.3.4. Thus, a prediction interval for the next k measurements for the $(1-\alpha)100\%$ level of confidence that uses the Bonferroni approximation will entail the use of individual comparison with Type I error of α/k. For example, for normally distributed data, the prediction interval for $(1-\alpha)100\%$ confidence for the next k observations is obtained from the quantile of the Student's t-distribution $t_{1-\alpha/k}$ (e.g., rather than $t_{1-\alpha}$).

K-2.3.5. It should be noted that, in general, prediction and tolerance intervals are not the same thing. The difference between a tolerance and prediction limit is one of interpretation and probability. Given n measurements and a desired confidence level, a tolerance interval will have a certain coverage percentage. A tolerance interval is designed so that, with some level of confidence, a proportion p of future measurements will fall within the interval. Thus, a small proportion $1-p$ of the measures may fall outside the tolerance interval. A prediction limit, on the other hand, is designed so that, with some level of confidence, *all* future measurements fall within the interval. In this sense, the prediction limit may be thought of as a 100% coverage tolerance limit for the next k future observations. Thus, upper prediction intervals are constructed when all future measurements must fall below some threshold value and tolerance intervals are typically constructed when only a large proportion of future measurements are required to exceed a threshold value.

K-3. Statistical Intervals Based on Normal Distribution.

K-3.1. *Confidence Interval for the Mean*. For a normal distribution, the one-sided (1 – α)100% UCL for the population mean is computed using the sample mean and standard deviation, and the (1 – α) quantile of Student's *t*-distribution with $n - 1$ degrees of freedom:

$$\mathrm{UCL}_{1-\alpha} = \bar{x} + t_{1-\alpha,n-1}\left(s/\sqrt{n}\right) .$$

Quantiles of the Student's *t*-distribution for various degrees of freedom are provided in Appendix B, Table B-23. Student's t can also be obtained in Microsoft Excel with the formula $TINV(2\alpha, n-1)$, for a one-sided (upper) $(1-\alpha)100\%$ UCL for $n-1$ degrees of freedom. When data are normally distributed, or if there are more than 30 samples available, a normal two-sided or one-sided confidence interval for the population mean (μ) with $(1-\alpha)100\%$ level of confidence can be computed as directed in the Paragraph K-3.2. An example is provided in Paragraph K-3.3.

K-3.2. *Directions for the Confidence Interval for the Mean (Normal Distribution) When the Population Standard Deviation is Unknown*. Let $x_1, x_2, \ldots, x_n$ represent the n data points from a normal distribution. These could be either n individual samples or n composite samples consisting of k aliquots each.

K-3.2.1. Verify that data come from a normal distribution using tests presented in Appendices F and J such as the Shapiro-Wilk test (Paragraph F-3) and a normal probability plot (Paragraph J-5.5).

K-3.2.2. Calculate the sample mean, $\bar{x}$, and the standard deviation, s (Appendix D).

K-3.2.3. Use Table B-23 of Appendix B to find the critical value such that $(1-\alpha)100\%$ of the *t*-distribution with $v = n-1$ degrees of freedom (df) is below this value. For a one-sided confidence interval (when just a LCL or an UCL is to be calculated), the critical value is the percentile $t_{1-\alpha,v}$. For a two-sided confidence interval (when both a LCL and UCL are to be calculated), the critical value is $t_{1-\alpha/2,v}$.

K-3.2.4. For example, if a two-sided 95% confidence interval is estimated, where $\alpha = 0.05$ and $n = 16$, then $v = n-1 = 16-1 = 15$ and $t_{1-(0.05/2),15} = t_{0.975,15} = 2.131$. If a one-sided 95% confidence interval is estimated, where $\alpha = 0.05$ and $n = 16$, then $v = n-1 = 16-1 = 15$ and $t_{1-0.05,15} = t_{0.95,15} = 1.753$.

K-3.2.5. For one-side confidence intervals for the population mean (μ), the equations for estimating the upper confidence limit (UCL) and lower confidence limit (LCL) are as follows:

$$\text{UCL} = \overline{x} + t_{1-\alpha,\nu}\left(s/\sqrt{n}\right)$$

$$\text{LCL} = \overline{x} - t_{1-\alpha,\nu}\left(s/\sqrt{n}\right).$$

K-3.2.6. The corresponding one-sided confidence intervals for are as follows:

$$\left(-\infty,\ \overline{x} + t_{1-\alpha,\nu}\left(s/\sqrt{n}\right)\right)$$

$$\left(\overline{x} - t_{1-\alpha,\nu}\left(s/\sqrt{n}\right),\ +\infty\right).$$

K-3.2.7. The two-sided confidence interval for the population mean is as follows:

$$\overline{x} \pm t_{1-\alpha/2,\nu}\left(s/\sqrt{n}\right).$$

K-3.3. *Example of a Confidence Interval for the Mean (Normal Distribution)*. Suppose a one-sided 95% lower confidence interval is desired for the mean concentration of (total) chromium in subsurface (below 5 feet from ground surface) soil at Site A.

K-3.3.1. These are the same data used in Paragraph L-6.1.3 as an example of a one-sample *t*-test. In that example there was evidence that the average was greater than 2 and not less than 2. A similar conclusion can also be reached when confidence intervals are constructed and compared to the regulatory threshold of 2, as illustrated in this example.

K-3.3.2. The first step is to verify that the data follow a normal distribution. The Shapiro-Wilk test is performed with these data; this test shows evidence that the data follow a normal distribution because the test's *p* value was 0.8489 and is greater than 0.05.

Table K-2.

Example Data

Site A Sample Location	Top Depth of Sample	Bottom Depth of Sample	Chromium (Total) Concentration (mg/kg)	Site A Sample Location	Top Depth of Sample	Bottom Depth of Sample	Chromium (Total) Concentration (mg/kg)
EPC-SB01	9	10	2.95	EPC-SB07	9	10	5.1
EPC-SB01	14	15	5.17	EPC-SB07	14	15	4.94
EPC-SB01	19	20	4.8	EPC-SB07	19	20	4.76
EPC-SB02	9	10	4.53	EPC-SB08	9	10	4.62
EPC-SB02	14	15	4.01	EPC-SB08	14	15	4.72
EPC-SB02	19	20	5.91	EPC-SB08	19	20	4.73
EPC-SB03	9	10	3.96	EPC-SB09	9	10	3.21
EPC-SB03	14	15	4.81	EPC-SB09	14	15	4.14
EPC-SB03	19	20	5.27	EPC-SB09	19	20	4.85
EPC-SB04	9	10	5.99	EPC-SB10	9	10	4.25
EPC-SB04	14	15	4.6	EPC-SB10	14	15	5.09
EPC-SB04	19	20	5.51	EPC-SB10	19	20	3.68
EPC-SB05	9	10	4.72	EPC-SB11	9	10	5.12
EPC-SB05	14	15	3.56	EPC-SB11	14	15	6.6
EPC-SB05	19	20	4.22	EPC-SB11	19	20	6.19
EPC-SB06	9	10	3.91	EPC-SB12	9	10	3.15
EPC-SB06	14	15	5.81	EPC-SB12	14	15	4.11
EPC-SB06	19	20	4.48	EPC-SB12	19	20	2.8

K-3.3.3. The mean and standard deviation of the data were calculated:

$\bar{x} = 4.619$

$s = 0.8980$.

Note that:

$\alpha = 0.05$ (for the 95% level of confidence)

$n = 36$

and

$v = n - 1 = 36 - 1 = 35$.

K-3.3.4. Using Table B-23 of Appendix B and linear interpolation, we find the critical value to be 1.691.

$$t_{1-\alpha,v} = t_{0.95,35} = (1.697 + 1.684)/2 = 1.691.$$

The confidence interval is

$$\left(\left\{ 4.619 - 1.691\left(0.8980/\sqrt{36}\right) \right\}, \infty \right) = \left(4.37, \infty\right) .$$

K-3.3.5. The confidence interval does not contain 2 (the lower confidence limit exceeds 2); therefore, this is evidence that the average is greater than 2, the regulatory threshold.

K-3.4. *Tolerance Interval* (*Normal Distribution*). A one-sided tolerance limit is an upper or a lower confidence limit of a percentile (or proportion). A one-sided upper tolerance limit (UTL) that is greater than at least p100% of the population with probability $(1 - \alpha)$100% is the $(1 - \alpha)$100% upper confidence limit for the $p100^{th}$ percentile of the population (Hahn, 1970). Similarly, a one-sided lower tolerance limit (LTL) that is less than at least p100% of the population with probability $(1 - \alpha)$100% is the $(1 - \alpha)$100% lower confidence limit for the $p100^{th}$ percentile of the population. However, two-sided tolerance intervals are not equivalent to two-sided confidence intervals of percentiles. "Tolerance limits differ from confidence intervals in that tolerance limits provide an interval within which at least a proportion q of the population lies, within probability $1 - \alpha$ or more that the stated interval does indeed 'contain' the proportion q of the population" (Conover, 1999). "Two-sided tolerance intervals are rarely used in environmental studies, perhaps because there are few applications that attempt to determine the location of a central proportion of data, with allowable exceedances at both high and low ends" (Helsel, 2005).

K-3.4.1. *Directions for a Tolerance Interval (Normal Distribution).* Let $x_1, x_2, \ldots, x_n$ represent the n data points from a normal distribution. These could be either n individual samples or n composite samples consisting of k aliquots each. A two-sided $(1 - \alpha)$100% tolerance interval to contain at least p100% of a normal distribution is denoted as $\left(x_L, x_U\right)$, where x_L is the lower tolerance limit and x_U is the upper tolerance limit.

K-3.4.1.1. Verify data come from a normal distribution using tests presented in Appendices F and J, such as the Shapiro-Wilk test (Paragraph F-3) and a normal probability plot (Paragraph J-5.5).

K-3.4.1.2. Calculate the sample mean, $\bar{x}$, and the standard deviation, *s* (Appendix D).

K-3.4.1.3. For a two-sided tolerance interval, (x_L, x_U):

$$x_L = \bar{x} - s\, g_{1-\alpha,p,n}$$

$$x_U = \bar{x} + s\, g_{1-\alpha,p,n} \ .$$

K-3.4.1.4. Use Table B-14 of Appendix B to find the critical value *g*.

K-3.4.1.5. An approximation for *g* that may be useful (e.g., to find values of *g* that are not in Table B-13) is:

$$g_{1-\alpha,p,n} \approx Z_{(1+p)/2}\left(\frac{n-1}{\chi^2_{\alpha,n-1}}\right)^{1/2}\left(1+\frac{1}{2n}\right) .$$

Percentiles of the chi-square distribution, $\chi^2_{p,\nu}$, are listed in Table B-2. Percentiles of the standard normal distribution, Z_p, are listed in Table B-15. Hahn states that the approximation "appears to be good for most practical purposes even for *n* as small as 5" (Hahn, 1970).

K-3.4.1.6. For a one-sided lower tolerance limit, x_L:

$$x_L = \bar{x} - s\, g'_{1-\alpha,p,n} \ .$$

K-3.4.1.7. For a one-sided upper tolerance limit, x_U:

$$x_U = \bar{x} + s\, g'_{1-\alpha,p,n} \ .$$

K-3.4.1.8. Use Table B-13 of Appendix B to find the critical value g' (for values of $p >$ 0.5).

K-3.4.1.9. An approximation for g' that may be useful is:

$$g'_{1-\alpha,p,n} \approx \frac{Z_p + \left(Z_p^2 - ab\right)^{1/2}}{a}$$

$$a = 1 - \frac{Z_{1-\alpha}^2}{2(n-1)}$$

$$b = Z_p^2 - \frac{Z_{1-\alpha}^2}{n} \ .$$

However, Hahn states that this approximation "is poor for very small n, especially for large p and large 1- α, and is not advised for $n < 8$" $-\alpha$ (Hahn, 1970).

K-3.4.2. *Example of a Two-sided Tolerance Interval (Normal Distribution)*. Suppose a two-sided 95% tolerance interval to contain at least 90% of the population is desired for chromium concentrations (total) in subsurface (below 5 feet from ground surface) soil at Site A, using the same data as Paragraph K-3.2.

K-3.4.2.1. The first step is to verify that the data follow a normal distribution. The Shapiro-Wilk test is performed with these data. This test shows evidence that the data follow a normal distribution because the test's p value was 0.8489 and is greater than 0.05.

K-3.4.2.2. The mean and standard deviation of the data were calculated:

$$\bar{x} = 4.619$$

$$s = 0.8980 \ .$$

Note that:

$$p = 0.90$$

$$n = 36$$

$$\alpha = 0.05 \ .$$

K-3.4.2.3. From Table B-14, $g_{0.95,0.90,35} = 2.090$ and $g_{0.95,0.90,40} = 2.052$. Therefore,

$$g_{0.95,0.90,36} = 2.090 - \frac{36-35}{40-35}(2.090 - 2.052) = 2.082 \ .$$

K-3.4.2.4. The equation in Paragraph K-3.4.1.5 can also be used to calculate *g:*

$$g_{1-\alpha,p,n} \approx Z_{(1+0.90)/2}\left(\frac{36-1}{\chi^2_{0.05,35}}\right)^{1/2}\left(1+\frac{1}{2\times 36}\right) = 1.645\times\left(\frac{35}{22.46}\right)^{1/2}\times 1.014 = 2.082\,.$$

K-3.4.2.5. The two-sided tolerance interval is:

$4.619 \pm 2.082 \times 0.8980$ mg/kg

(2.749, 6.489) mg/kg.

K-3.4.3. *Example of a One-Sided Upper Tolerance Limit, UTL (Normal Distribution)*. Suppose a UTL for the 95th percentile and 95% confidence level (also called a 95 UTL) is desired for chromium concentrations (total) in subsurface (below 5 feet from ground surface) soil at Site A, using the same data in Paragraph K-3.3.

K-3.4.3.1. As shown in the previous examples, the data seem to follow a normal distribution. For this example:

$p = 0.95$

$n = 36$

$\alpha = 0.05$

$1-\alpha = 0.95$

$\bar{x} = 4.619$

$s = 0.8980$.

K-3.4.3.2. Using Table B-13 of Appendix B and linear interpolation, we find the critical value for the one-sided upper confidence limit to be

$$g'_{1-\alpha,p,n} = g'_{0.95,0.95,36} = 2.167 - \frac{36-35}{40-35}(2.167-2.125) = 2.159\,.$$

K-3.4.3.3. The approximation for g' in Paragraph K-3.4.1.9 may also be used to estimate g':

$$a = 1-\frac{Z^2_{1-\alpha}}{2(n-1)} = 1-\frac{Z^2_{0.95}}{2(36-1)} = 1-\frac{1.645^2}{70} = 0.9613$$

$$b = Z_p^2 - \frac{Z_{1-\alpha}^2}{n} = Z_{0.95}^2 - \frac{Z_{0.95}^2}{36} = 1.645^2 - \frac{1.645^2}{36} = 2.631$$

$$g'_{1-\alpha,p,n} \approx \frac{Z_p + \left(Z_p^2 - ab\right)^{1/2}}{a} = \frac{1.645 + \left(1.645^2 - 0.9613 \times 2.631\right)^{1/2}}{0.9613} = 2.149 .$$

K-3.4.3.4. So, using the value for g' from Table B-13, the UTL is:

$$\text{UTL} = 4.619 + 0.8980 \times 2.159 = 6.558 \text{ mg/kg.}$$

K-3.4.4. *Confidence Interval for the Variance or Standard Deviation (Normal Distribution)*. To estimate the precision of variance estimates, a confidence interval for the variance or standard deviation can be constructed. This information may be necessary for a sensitivity analysis of the statistical test or analysis method. The method described below can be used to find a two-sided $(1-\alpha)100\%$ confidence interval. This confidence interval assumes that the data constitute a random sample from a normally distributed population and can be highly sensitive to outliers and to departures from normality. Directions are presented in Paragraph K-3.4.4.1, followed by an example in Paragraph K-3.4.4.2.

K-3.4.4.1. *Directions for a Confidence Interval for the Variance and Standard Deviation (Normal Distribution)*. Let $x_1, x_2, \ldots, x_n$ represent the *n* data points from a normal distribution.

K-3.4.4.1.1. Verify data come from a normal distribution using tests presented in Appendices F and J such as the Shapiro-Wilk test (Paragraph F-3) and a normal probability plot (Paragraph J-5.5).

K-3.4.4.1.2. Calculate the sample variance, s^2 (Appendix D).

K-3.4.4.1.3. For a $(1-\alpha)100\%$ two-sided confidence interval, use Table B-2 of Appendix B to find the critical values $\chi^2_{\alpha/2,\nu}$ and $\chi^2_{1-\alpha/2,\nu}$ with degrees of freedom $\nu = (n-1)$.

K-3.4.4.1.4. A $(1-\alpha)100\%$ confidence interval for the true underlying variance is (s_L^2, s_U^2):

$$s_L^2 = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}$$

$$s_U^2 = \frac{(n-1)s^2}{\chi^2_{\alpha/2,v}}.$$

K-3.4.4.1.5. A $(1-\alpha)100\%$ confidence interval for the true underlying standard deviation is (s_L, s_U):

$$s_L = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}}}$$

$$s_U = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}}}.$$

K-3.4.4.2. *Example of Constructing a Confidence Interval for the Sample Variance and Standard Deviation (Normal Distribution)*. Consider the following data, background subsurface chromium concentrations of 3.84, 4.26, 4.53, 4.60, 5.28, 5.29, 5.74, and 5.86 mg/kg.

K-3.4.4.2.1. A confidence interval for the sample variance will be calculated based on a 95% level of confidence.

K-3.4.4.2.2. Testing the data for normality using the Shapiro-Wilk test indicated that the data were normal. So, a confidence interval for the sample variance based on a normal distribution can be calculated.

K-3.4.4.2.3. The sample variance, $s^2 = 0.526$. The required critical values are obtained from Table B-2:

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,7} = 1.69$$

$$\chi^2_{(1-\alpha/2),n-1} = \chi^2_{0.975,7} = 16.01.$$

K-3.4.4.2.4. A 95% confidence interval for the true underlying variance is (0.228, 2.18):

$$s_L^2 = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} = \frac{(8-1)(0.526)}{16.01} = 0.228$$

$$s_U^2 = \frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} = \frac{(8-1)(0.526)}{1.69} = 2.18 .$$

K-3.4.4.2.5. A 95% confidence interval for the true underlying standard deviation is (0.479, 1.48):

$$s_L = \sqrt{\frac{(n-1)s^2}{\chi^2_{(1-\alpha/2),n-1}}} = \sqrt{0.228} = 0.479$$

$$s_U = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}} = \sqrt{2.18} = 1.48 .$$

K-3.4.5. *Prediction Interval (Normal Distribution).* The prediction interval presented here is constructed assuming that the data follow a normal distribution with unknown mean and standard deviation. Most evaluations with environmental data only need a one-sided prediction interval, so this discussion will focus on the one-sided, upper prediction limit. To obtain a two-sided prediction interval, first replace α by $\alpha/2$. Then use the equation for the upper limit as the lower limit after replacing the addition of the standard deviation term with subtraction. The prediction interval must specify the overall level of confidence. That means a prediction interval's confidence level must account for the level of confidence of every future comparison. This is accomplished by setting the confidence level for each of the *k* future comparisons to $(1-\alpha/k)100\%$. Directions for calculating an upper prediction limit are presented in Paragraph K-3.4.5.1, followed by an example in Paragraph K-3.4.5.2.

K-3.4.5.1. *Directions for Calculating an Upper Prediction Limit for k Future Comparisons of the Mean Calculated from m Observations (Normal Distribution).* Verify the assumptions of normality.

K-3.4.5.1.1. The population mean and standard deviation are unknown. Specify *k* and *m* for the interval, where the mean of *m* observations is taken *k* times in the future (i.e., *k* samples are analyzed and the result reported for each sample is the mean of *m* replicate measurements).

K-3.4.5.1.2. Specify the level of confidence for the upper prediction limit as $(1-\alpha)100\%$.

K-3.4.5.1.3. Calculate the upper prediction limit

$$\bar{x} + s\, t_{1-\alpha/k,n-1}\sqrt{\frac{1}{m}+\frac{1}{n}}$$

where $\bar{x}$ is the mean of the original data, s is the standard deviation, and n is the total number of observations (measurements of the original data set).

K-3.4.5.1.4. Table B-23 of Appendix B provides values for $t_{1-\alpha/k,n-1}$.

K-3.4.5.1.5. If the future observations are found to be in the prediction interval, this is evidence that there has been no change in the sample values. If a future observation falls outside of the prediction interval, this is statistical evidence that the new observation does not come from the same distribution.

K-3.4.5.1.6. When replicate sample analyses are not done (i.e., a signal measurement or analysis is performed for each sample), set $m = 1$. For a single future observation (i.e., one sample analyzed once), set $m = 1$ and $k = 1$.

K-3.4.5.2. *Example of Calculating a Normal Upper Prediction Limit for* k *Future Comparisons of the Mean from* m *Observations (Normal Distribution)*. A prediction interval is calculated for a set of "background well" measurements to determine if a set of "compliance well" measurements are "elevated" relative to background levels. The background well data set was tested for normality using the Shapiro-Wilk test. Because the data set was not normally distributed, the data set was normalized by taking the natural logarithm of each result.

K-3.4.5.2.1. For the compliance well data set, $m = 4$ replicate measurements are made for $k = 1$ sample. Let $\alpha = 0.01$ for the prediction interval. For the background data set, $n = 8$.

Table K-3.
Example Compliance Well data

Background Well	Sample Date	Result	Log Result	Compliance Well	Result	Log Result	Sample Date
69-2-07	2001	0.0137	–4.290	69-2-08	0.563	–0.574	2002
69-2-07	2001	0.019	–3.963	69-2-08	0.512	–0.669	2002
69-2-07	2001	0.0163	–4.117	69-2-08	0.475	–0.744	2002
69-2-07	2001	0.0195	–3.937	69-2-08	0.546	–0.605	2002
69-2-07	2001	0.0112	–4.492				
69-2-07	2001	0.0112	–4.492				
69-2-07	2001	0.0102	–4.585				
69-2-07	2001	0.00946	–4.661				
Mean	-	0.01382	–4.317		0.524	–0.6484	
Std. Dev.	-	0.00398	0.2832		0.0389	0.0753	

$$\bar{x} + s\,t_{1-\alpha/k,n-1}\sqrt{\frac{1}{m}+\frac{1}{n}} = -4.317 + 0.2832\,(2.998)\sqrt{\frac{1}{4}+\frac{1}{8}} = -3.79$$

$$t_{1-\alpha/k,n-1} = t_{1-\frac{0.01}{1},8-1} = t_{0.99,7} = 2.998$$

using Table B-23 of Appendix B.

K-3.4.5.2.2. Because the data set was transformed by taking the natural logarithm prior to calculating the upper prediction limit, to express the calculated limit in terms of the original units, it is necessary to perform the inverse transformation (i.e., to take the exponent of the calculated limit): exp(–3.79) = 0.0226. Therefore, the prediction interval is (0, 0.0226). Now we can compare the mean of the compliance well observations (0.524) with the upper limit of the prediction interval (0.022) calculated from the background well data. As 0.524 > 0.0226, there is significant evidence that the compliance well observations do not come from the same distribution as the background well.

K-4. Statistical Intervals Based on Lognormal Distribution.

K-4.1. *Confidence Interval for the Mean*.

K-4.1.1. When data are truly lognormal, it is *not* recommended that confidence intervals be calculated using the natural-log transformed data and the normal confidence intervals. One reason is that the units as well as the confidence intervals would be in log scale. The confidence intervals cannot be transformed back to the original scale and original units without a special adjustment.

K-4.1.2. For a lognormal distribution (the second alternative provided in the EPA UCL method flow chart), the EPA recommends calculating the UCL of the mean using one of several options based on the sample size, n, and the standard deviation of the log-transformed data, s_y. Table K-4, which summarizes these recommendations, has been adapted from the ProUCL Version 3.0 User Guide (EPA 600/R-97/006). Determining the UCL for lognormal populations is a current area of research and these recommendations are subject to change. It should be noted that ProUCL is freely distributed and relatively simple to use. In addition to the computational methods listed below, the most current version of the software uses the gamma distribution to calculate UCLs. The software calculates UCLs using a number of different computational methods and automatically selects the "best" method (e.g., using criteria similar to that presented in Table K-4). However, it should also be noted that these computational methods can result in relatively large UCLs (e.g., near the maximum detected values when the distributions are extremely skewed). This problem can be potentially avoided or at least minimized by collecting composite rather than grab samples (when possible and consistent with data quality objectives), as this tends to normalize data (i.e., composite samples produced from a sufficiently large number of grabs tend to be normally distributed).

K-4.2. *Land Method*.

K-4.2.1. *Introduction*. The Land method was touted in older EPA guidance, but it is no longer recommended in all cases because it is very sensitive to deviations from lognormality. Recall that distribution tests are primarily tests that the fit assumption cannot be rejected, rather than that the fit is perfect. Consequently, it is possible to pass a test for lognormality even when there are deviations from that distribution. This outcome is more likely for small data sets (< 30), which are quite common in environmental applications. The UCL for the Land method is as follows:

$$\mathrm{UCL}_{1-\alpha} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{H_{1-\alpha} s_y^2}{\sqrt{n-1}} \right) .$$

K-4.2.1.1. The value of the *H* statistic is available in some statistical texts, including Gilbert (1987) and in Table B-8 of Appendix B.

K-4.2.1.2. Directions for constructing a confidence interval for the population mean of a lognormal distribution using the Land method are given in Paragraph K-4.2.2, followed by an example in Paragraph K-4.2.3 (EPA 600/R-97/006).

K-4.2.2. *Directions for a Confidence Interval for the Mean (Lognormal Distribution, Land Method)*. Let $x_1, x_2, \ldots, x_n$ represent the *n* data points from a lognormal distribution.

K-4.2.2.1. Verify that data come from a lognormal distribution using tests presented in Appendices F and J such as the Shapiro-Wilk test (Paragraph F-3) and a normal probability plot (Paragraph J-5.5).

K-4.2.2.2. Using the log-transformed data, $y_i = Ln(x_i)$, calculate the sample mean, $\bar{y}$, and the standard deviation, s_y.

K-4.2.2.3. Use Table B-8 of Appendix B to find the critical value (also called the *H* statistic) for the given level of confidence, sample size, and standard deviation. If a two-sided confidence interval for the mean is desired (LCL, UCL), the critical values are $H_{\alpha/2,n,s_y}$ and $H_{1-\alpha/2,n,s_y}$ for the LCL and UCL, respectively. If a one-sided confidence interval for the mean is desired, the critical value for the LCL is H_{α,n,s_y}, or the critical value for an UCL is $H_{1-\alpha,n,s_y}$. To estimate *H* values not in the table, a four-point Lagrangian interpolation (cubic interpolation) should be implemented.

Table K-4.
Recommended Methods for Computation of a 95% UCL for the Unknown Mean of a Log-normal Population

Standard Deviation of Log-Transformed Data, s_y	Sample Size, *n*	Recommended Method (Paragraph Reference)
$s_y < 0.5$	For all *n*	Student's *t* (K-3.4.4) or Land (K-4.1)
$0.5 \le s_y < 1.0$	For all *n*	Land (K-4.1)
$1.0 \le s_y < 1.5$	$n < 25$	95% Chebyshev (MVUE) UCL (K-4.1)
	$n \ge 25$	Land (K-4.1)
$1.5 \le s_y < 2.0$	$n < 20$	99% Chebyshev (MVUE) UCL (K-4.1)
	$20 \le n < 50$	95% Chebyshev (MVUE) UCL (K-4.1)
	$n \ge 50$	Land (K-4.1)
$2.0 \le s_y < 2.5$	$n < 20$	99% Chebyshev (MVUE) UCL (K-4.1)
	$20 \le n < 50$	97.5% Chebyshev (MVUE) UCL (K-4.1)
	$50 \le n < 70$	95% Chebyshev (MVUE) UCL (K-4.1)
	$n \ge 70$	Land (K-4.1)
$2.5 \le s_y < 3.0$	$n < 30$	Larger of (99% Chebyshev (MVUE) UCL (K-4.1) or 99% Chebyshev (Mean, Sd) (K-5)
	$30 \le n < 70$	97.5% Chebyshev (MVUE) UCL (K-4.1)
	$70 \le n < 100$	95% Chebyshev (MVUE) UCL (K-4.1)
	$n \ge 100$	Land (K-4.1)
$3.0 \le s_y < 3.5$	$n < 15$	Hall's Bootstrap* (K-4.1)
	$15 \le n < 50$	Larger of (99% Chebyshev (MVUE) UCL (K-4.1) or 99% Chebyshev (Mean, Sd) (K-5)
	$50 \le n < 100$	97.5% Chebyshev (MVUE) UCL (K-4.1)
	$100 \le n < 150$	95% Chebyshev (MVUE) UCL (K-4.1)
	$n \ge 150$	Land (K-4.1)
$s_y \ge 3.5$	For all *n*	Use non-parametric methods* (K-5)

*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.

K-4.2.2.4. For a two-sided confidence interval for the mean, the equations are as follows:

$$\text{LCL} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{s_y H_{\alpha/2}}{\sqrt{n-1}}\right), \quad \text{UCL} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{s_y H_{1-a/2}}{\sqrt{n-1}}\right).$$

K-4.2.2.5. For a one-sided confidence interval for the mean, *LCL* or *UCL*, the equation is as follows:

$$\text{LCL} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{s_y H_\alpha}{\sqrt{n-1}} \right) \text{ or } \text{UCL} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{s_y H_{1-a}}{\sqrt{n-1}} \right).$$

K-4.2.3. *Example of a Confidence Interval for the Mean (Lognormal Distribution), Land Method.* Suppose a one-sided 95% UCL is desired for concentrations of chromium (total) in background subsurface soil (5 feet below ground surface).

Sample ID	**Result (mg/kg)**	***Ln*(Result) (*Ln* mg/kg)**
EPC-BG01-013	0.0196	–3.932
EPC-BG01-020	0.00605	–5.108
EPC-BG02-010	0.00485	–5.329
EPC-BG02-020	0.0101	–4.595
EPC-BG03-010	0.00756	–4.885
EPC-BG03-020	0.00596	–5.123
EPC-BG04-010	0.0143	–4.248
EPC-BG04-020	0.00499	–5.300
EPC-BG05-010	0.00997	–4.608
EPC-BG05-020	0.00464	–5.373
EPC-BG06-010	0.00813	–4.812
EPC-BG06-023	0.00313	–5.767
EPC-BG07-010	0.00834	–4.787
EPC-BG07-020	0.00579	–5.151
EPC-BG08-010	0.00638	–5.055
EPC-BG08-020	0.00517	–5.265

K-4.2.3.1. The first step is to verify that the data follow a lognormal distribution. The Shapiro-Wilk test was performed with the log-transformed data. This test shows evidence that the data follow a normal distribution because the test's *p* value was 0.6570 and is greater than 0.05.

K-4.2.3.2. Using the log-transformed data,

$\bar{y} = -4.959$

and

$s_y = 0.4574$.

K-4.2.3.3. The critical value is $H_{0.95,16,0.4574} = 2.007$. A four-point Lagrangian interpolation (cubic interpolation) was implemented to obtain this critical value. K-4.2.4 shows how the criti-

cal value $H_{0.95,16,0.4574}$ was derived.

K-4.2.3.4. For a one-sided upper confidence interval for the mean, UCL, the equation is:

$$\text{UCL} = \exp\left(\bar{y} + \frac{s_y^2}{2} + \frac{s_y H_{1-a}}{\sqrt{n-1}} \right) = \exp\left(-4.959 + \frac{0.4574^2}{2} + \frac{0.4574(2.007)}{\sqrt{16-1}} \right) = 0.0099.$$

K-4.2.4. *Lagrangian Interpolation (Cubic Interpolation) for the H Statistic*. The details of the Lagrangian (cubic) interpolation are provided to assist in the use of Table B-8 of Appendix B.

K-4.2.4.1. Suppose the H statistic $\left(H_{1-\alpha/2,n,s_y}\right)$ is desired for

$1-\alpha/2 = 0.95$

$n = 16$

$s_y = 0.4574$

(from Paragraph K-4.2.3).

K-4.2.4.2. A Lagrangian interpolation requires bounding the desired value by two tabulated values lower and two tabulated values higher than the desired value. Using the example above, we need a column of H statistics when $n = 16$ because there is no such column in Table B-8. The tabulated columns n = 12, 15 (two values below 16) and n = 21, 31 (two values above 16) are used to generate a column for $n = 16$. Once the column of H statistics is generated for $n = 16$, Lagrangian interpolation can be used to get the H statistic for $s_y = 0.4574$.

K-4.2.4.3. So the columns associated with s_y = 0.30, 0.40 (two values below 0.4574) and s_y = 0.50, 0.60 (two values above 0.4574) are used to generate a column for s_y = 0.4574.

K-4.2.4.4. From Table B-8, the following H statistics, $H_{0.95,n,s_y}$, are needed for these interpolations:

			n		
s_y	**12**	**15**	**16**	**21**	**31**
0.30	1.927	1.882	$H_{0.95,16,0.30}$	1.833	1.793
0.40	2.026	1.968	$H_{0.95,16,0.40}$	1.905	1.856
0.4574	—	—	$H_{0.95,16,0.4574}$	—	—
0.50	2.141	2.068	$H_{0.95,16,0.50}$	1.989	1.928
0.60	2.271	2.181	$H_{0.95,16,0.60}$	2.085	2.010

K-4.2.4.5. The first part of the interpolation process is to generate a column of H statistics for $n = 16$. For each s_y, the following equation is used:

$$H_{0.95,16,s_y} = \frac{(16-15)(16-21)(16-31)}{(12-15)(12-21)(12-31)} H_{0.95,12,s_y} + \frac{(16-12)(16-21)(16-31)}{(15-12)(15-21)(15-31)} H_{0.95,15,s_y} + \frac{(16-12)(16-15)(16-31)}{(21-12)(21-15)(21-31)} H_{0.95,21,s_y} + \frac{(16-12)(16-15)(16-21)}{(31-12)(31-15)(31-21)} H_{0.95,31,s_y}.$$

So,

$$H_{0.95,16,0.30} = \frac{(16-15)(16-21)(16-31)}{(12-15)(12-21)(12-31)}(1.927) + \frac{(16-12)(16-21)(16-31)}{(15-12)(15-21)(15-31)}(1.882) + \frac{(16-12)(16-15)(16-31)}{(21-12)(21-15)(21-31)}(1.833) + \frac{(16-12)(16-15)(16-21)}{(31-12)(31-15)(31-21)}(1.793)$$
$$= -0.2817 + 1.960 + 0.2037 - 0.0118 = 1.8702.$$

The same process was used to determine $H_{0.95,16,0.40}$, $H_{0.95,16,0.50}$, and $H_{0.95,16,0.60}$.

			n		
s_y	***12***	***15***	***16***	***21***	***31***
0.30	1.927	1.882	1.870	1.833	1.793
0.40	2.026	1.968	1.953	1.905	1.856
0.4574	—	—	$H_{0.95,16,0.4574}$	—	—
0.50	2.141	2.068	2.049	1.989	1.928
0.60	2.271	2.181	2.158	2.085	2.010

K-4.2.4.6. Next, the H statistic values for the various s_y at $n = 16$ are used to interpolate $H_{0.95,16,0.4574}$.

$$
\begin{aligned}
H_{0.95,16,0.4574} &= \frac{(0.4574-0.40)(0.4574-0.50)(0.4574-0.60)}{(0.30-0.40)(0.30-0.50)(0.30-0.60)} \quad (1.870) \\
&+ \frac{(0.4574-0.30)(0.4574-0.50)(0.4574-0.60)}{(0.40-0.30)(0.40-0.50)(0.40-0.60)} \quad (1.953) \\
&+ \frac{(0.4574-0.30)(0.4574-0.40)(0.4574-0.60)}{(0.50-0.30)(0.50-0.40)(0.50-0.60)} \quad (2.049) \\
&+ \frac{(0.4574-0.30)(0.4574-0.40)(0.4574-0.50)}{(0.60-0.30)(0.60-0.40)(0.60-0.50)} \quad (2.158) \\
&= -0.1087 + 0.9337 + 1.320 - 0.1384 \\
&= 2.007.
\end{aligned}
$$

Thus, the *H* statistic is 2.007.

K-4.3. *Chebyshev (MVUE) Method*.

K-4.3.1. *Introduction.* For the Chebyshev (MVUE) method, first estimate the mean and variance using the minimum unbiased variance approach discussed in Appendix D. Then calculate the $100(1-\alpha)\%$ UCL of the mean using:

$$\mathrm{UCL}_{1-\alpha} = \hat{\mu}_1 + \sqrt{\left(\frac{1}{\alpha} - 1\right) s^2\left(\hat{\mu}_1\right)} \ .$$

The quantities $\hat{\mu}_1$ and $s^2\left(\hat{\mu}_1\right)$ are the MVUE estimates of the mean and standard deviation given in equations D-2 and D-3 in Appendix D. An example of using this method follows in Paragraph K-4.3.2.

K-4.3.2. *Example of a Confidence Interval for the Mean (Lognormal Distribution), Chebyshev MVUE Method*. Suppose chromium concentrations (mg/kg) measured at a site are as follows:

0.378	1.411	1.089	0.918
0.073	0.518	2.240	0.111
1.246	2.251	1.967	1.894
1.414	13.844	1.222	0.962
0.094	0.247	0.371	0.056

K-4.3.2.1. ProUCL was used to determine the 95% UCL for the population mean. The data follow a lognormal distribution (Shapiro-Wilk p = 0.905 on the log-transformed data). The sample size is 20, and the standard deviation of the log-transformed values is 1.39. Table J-2 recommends using the 95% Chebyshev MVUE UCL as the 95% UCL for the population mean under these conditions.

K-4.3.2.2. The MVUE estimate of the mean, $\hat{\mu}_1$, is 1.66, and the standard deviation of the estimate of the mean, $s^2(\hat{\mu}_1)$, is 0.607. Therefore,

$$\mathrm{UCL}_{0.95} = \hat{\mu}_1 + \sqrt{\left(\frac{1}{\alpha} - 1\right) s^2(\hat{\mu}_1)} = 1.66 + \sqrt{\left(\frac{1}{0.05} - 1\right)(0.607)^2} = 4.30 .$$

K-4.4. Hall's Bootstrap Method.

Although Hall's Bootstrap is a nonparametric method related to the Bootstrap technique presented in Paragraph K-5, a limited presentation will be given here because EPA guidance (OSWER 9285.6-10) specifically recommends this technique for calculating the UCL of a lognormal population under certain situations described in Table K-4. The method adjusts for bias and skewness in the data (OSWER 9285.6-10). Directions for implementing Hall's Bootstrap are given in Paragraph K-4.4.1 and results of Hall's method from ProUCL Version 3.0 are presented in Paragraph K-4.4.2. The directions for performing the bootstrap method are presented for illustration only, as bootstrap methods require too many arithmetic calculations for manual calculations to be practical.

K-4.4.1. *Directions for Implementing Hall's Bootstrap Method for a* $100(1-\alpha)\%$ *UCL.* Let $x_1, x_2, \ldots, x_n$ represent n randomly sampled concentrations.

K-4.4.1.1. Compute the sample mean,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i .$$

K-4.4.1.2. Compute the sample standard deviation,

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} .$$

K-4.4.1.3. Compute the sample skewness,

$$k = \frac{1}{ns^3}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3 .$$

K-4.4.1.4. Do the following a large number of times:

K-4.4.1.4.1. Generate a simple random sample of n values from $x_1, x_2, \ldots, x_n$ with replacement.

K-4.4.1.4.2. Compute the sample mean, $\bar{x}_i$, standard deviation, s_i, and skewness, k_i, of the sample found in K-4.4.1.4.1.

K-4.4.1.4.3. Compute the Studentized mean,

$$W_i = \frac{\left(\bar{x}_i - \bar{x}\right)}{s_i}.$$

K-4.4.1.4.4. Compute Hall's statistic,

$$Q_i = W_i + \frac{k_i W_i^2}{3} + \frac{k_i^2 W_i^3}{27} + \frac{k_i}{6n}.$$

K-4.4.1.4.5. Sort all the values, Q_i, in ascending sequence and calculate the α^{th} lower quantile, Q_α.

K-4.4.1.4.6. Calculate

$$W = \frac{3}{k}\left\{\left[1 + k\left(Q_\alpha - \frac{k}{6n}\right)\right]^{\frac{1}{3}} - 1\right\} .$$

The one-sided $(1-\alpha)100\%$ upper confidence limit is $UCL_{1-\alpha} = \bar{x} - Ws$.

K-4.4.2. *Example of a Confidence Interval for the Mean (Lognormal Distribution), Hall's Bootstrap*. Suppose chromium concentrations (mg/kg) measured at a site are as follows:

0.331	0.104
68.977	0.022
0.908	2.044

140.605 0.093
157.359 0.213

ProUCL was used to determine the 95% UCL for the population mean. The data follow a lognormal distribution (Shapiro-Wilk $p = 0.842$ on the log-transformed data). The sample size is 10, and the standard deviation of the log-transformed values is 3.27. Table K-4 recommends using Hall's Bootstrap to estimate the 95% UCL for the population mean under these conditions. The Bootstrap algorithm was run with the result $UCL_{0.95} = 71.4$ mg/kg. Because this result is based on random sampling, it may change with repeated runs. As a comparison, the Land method 95% UCL for this data is over 3,240,000 mg/kg (an unrealistically large value).

K-4.4.3. *Confidence Interval for a Percentile–Tolerance Interval (Lognormal Distribution)*. A lognormal confidence interval for the $p100^{th}$ percentile of a lognormal distribution, X_p, with $(1-\alpha)100\%$ confidence, can be derived by using the log-transformed data with the equations for the normal confidence interval. When $Y = Ln(X)$ is normal (i.e., X is lognormal), given a set of sample values $y_1, y_2 \ldots y_n$ with sample mean $\bar{y}$ and standard deviation s, the exponent of $\bar{y}$ is an estimate of the 50^{th} percentile (median) of X ($X_{0.5}$):

$$x_{0.5} = \exp(\bar{y}) \ .$$

K-4.4.3.1. The two-sided $100(1-\alpha)\%$ confidence interval for the median of X is:

$$\left(\exp\left(\bar{y} - \frac{t_{\alpha/2,n-1}\, s}{\sqrt{n}} \right), \exp\left(\bar{y} + \frac{t_{1-\alpha/2,n-1}\, s}{\sqrt{n}} \right) \right) .$$

K-4.4.3.2. In general, if X is lognormal and $Y = Ln(X)$, then an estimate x_p of the $p100^{th}$ percentile of X (X_p) is obtained by first calculating an estimate of Y_p (the $p100^{th}$ percentile of Y),

$$y_p = \bar{y} + t_{p,n-1}\, s$$

and then performing the inverse transformation (exponentiation) on this quantity. The (maximum likelihood) estimate of the percentile X_p in terms of the original variable (X) is:

$$x_p = \exp(y_p) = \exp\left(\bar{y} + t_{p,n-1}\, s\right).$$

K-4.4.3.3. A one-sided upper confidence limit for the percentile X_p is calculated as follows:

$$\exp(\bar{y} + g'_{1-\alpha,p,n} s) \ \text{ for } p > 0.5 \ .$$

K-4.4.3.4. The term in parentheses is simply a confidence limit for a normal percentile or tolerance limit as described in Paragraph K-3.4.

K-4.4.3.5. A two-sided tolerance interval is calculated as follows:

$$\left(\exp(\bar{y} - g_{1-\alpha,p,n}s), \exp(\bar{y} + g_{1-\alpha,p,n}s\right) .$$

K-4.4.4. *Prediction Interval (Lognormal Distribution)*. A lognormal prediction interval can be calculated using the log-transformed data with the process for developing normal prediction intervals. When X is lognormal and $Y = Ln(X)$ with sample mean $\bar{y}$ and standard deviation s, then the prediction interval for the next k observations in the original scale is:

$$\left(\exp\left(\bar{y} - \frac{t_{\alpha/2k,n-1}\, s}{\sqrt{1+1/n}} \right), \exp\left(\bar{y} + \frac{t_{1-\alpha/2k,n-1}\, s}{\sqrt{1+1/n}} \right) \right) .$$

K-5. Distribution-Free Statistical Intervals.

K-5.1. *Introduction*. The one-sided Chebyshev inequality for a mean can be used when no distribution can be assumed to fit the data. Regardless of the underlying probability distribution of some variable X, the following inequality will be satisfied for the $(1 - \alpha)100\%$ UCL of the population mean μ:

$$(1-\alpha)100\%\ \text{UCL} \le \bar{x} + \sqrt{\frac{1}{\alpha} - 1}\ \left(\frac{\sigma}{\sqrt{n}} \right) .$$

K-5.1.1. The right-hand side of the inequality serves as a conservative estimate of the UCL. However, as the population standard deviation σ is typically unknown, the UCL is usually estimated as follows:

$$(1-\alpha)100\%\ \text{UCL} \approx \bar{x} + \sqrt{\frac{1}{\alpha} - 1}\ \left(\frac{s}{\sqrt{n}} \right) .$$

K-5.1.2. Unfortunately, because the sample standard deviation population (s) is being used to estimate the population standard deviation (σ), the population mean may not actually be less than this limit at the prescribed level of confidence when the variance or skewness is large, especially for small sample sizes. See OSWER 9285.6-10 for more details.

K-5.1.3. This one-sided Chebyshev UCL, based on the mean and standard deviation, is recommended for use with the lognormal distribution under certain conditions described in Table K-4. In that situation use the untransformed data to calculate $\bar{x}$ and s.

K-5.2. *Confidence Interval for the Mean*. If data do not follow either a normal or lognormal distribution, EPA guidance (OSWER 9285.6-10) recommends using either the central limit theorem or Bootstrap resampling. Several methods are available for estimating confidence limits of the mean when no distributional assumptions are made. The Bootstrap and Jackknife procedures are nonparametric statistical techniques that can be used to construct approximate confidence intervals for parameters such as the population mean. These procedures are nonparametric or distribution-free because they do not require assumptions about the data's distribution (such as normal or lognormal). It should be noted that statistical methods that account for the data's distribution, when used appropriately, are more efficient than the nonparametric methods. Directions for the Bootstrap and Jackknife methods for estimating a nonparametric confidence interval for θ, the parameter of interest, are given in Paragraphs K-5.2.1 and K-5.2.2, respectively. Examples are presented in Paragraphs K-5.2.3 and K-5.2.4. It should be noted that the both the Bootstrap and Jackknife methods are usually performed using statistical software owing to the large number of manual calculations that would be required. The Paragraphs below illustrate how the calculations are done.

K-5.2.1. *Directions for a Bootstrap Estimate of the Confidence Interval for* θ. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n.

K-5.2.1.1. The parameter of interest is θ and a reasonable estimate of θ is $\hat{\theta}$. For example, θ is the mean and $\hat{\theta}$ is the minimum variance unbiased estimator (MVUE) of the mean (Appendix D).

K-5.2.1.2. Take n samples with replacement from the original set of random samples of size n, and define this new set of data as $x_{11}, x_{12}, \ldots, x_{1n}$. Note that the same result can be selected more than once. For this new data set, estimate $\hat{\theta}$ and denote it as $\hat{\theta}_1$.

K-5.2.1.3. Perform the previous step N times, each time calculating an estimate of $\hat{\theta}$. Denote all N estimates of $\hat{\theta}$ as $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N$. N should be considerably larger, such as 1000 or more. It is much easier to perform this simulation using a computer.

K-5.2.1.4. Estimate the Bootstrap estimate of θ, $\overline{\theta}_B$, from the N estimates of $\hat{\theta}_I$, such that

$$\overline{\theta}_B = \frac{1}{N}\sum_{I=1}^{N} \hat{\theta}_I$$

for $i = 1, 2, \ldots, N$.

K-5.2.1.5. Derive the confidence interval for θ, with $100(1-\alpha)\%$ level of confidence, as (θ_L, θ_U) where $\theta_L = (\alpha/2)100^{th}$ percentile from the set of *N* estimates and $\theta_U = (1-\alpha/2)100^{th}$ percentile from the set of *N* estimates (see Appendix G). A one-sided UCL is simply the $(1-\alpha)100^{th}$ percentile from the set of *N* estimates.

K-5.2.2. *Example of the Bootstrap Method for Estimating a Nonparametric Confidence Interval for* θ. A confidence interval for the population mean (μ) will be calculated for chromium concentrations in subsurface soil at Site A with 95% level of confidence. All chromium concentrations were detected so no proxy concentrations are needed to evaluate the data.

K-5.2.2.1. The data are as follows: 2.95, 5.17, 4.80, 4.53, 4.01, 5.91, 3.96, 4.81, 5.27, 5.99, 4.60, 5.51, 4.72, 3.56, 4.22, 3.91, 5.81, 4.48, 5.10, 4.94, 4.76, 4.62, 4.72, 4.73, 3.21, 4.14, 4.85, 4.25, 5.09, 3.68, 5.12, 6.60, 6.19, 3.15, 4.11, and 2.80 mg/kg.

K-5.2.2.2. An example of 10 samples with replacement taken from the original set of random samples of size *n* = 36 is as follows: 2.95, 5.17, 5.91, 3.96, 4.80, 4.81, 4.53, 5.27, 4.01, and 5.99 mg/kg. (Note that although replacement was adhered to, no sample's values were actually "picked" twice.)

K-5.2.2.3. For this new data set, estimated mean is $\hat{\theta}_1 = 4.74$.

K-5.2.2.4. Perform the previous step *N* times, and each time calculating an estimate of $\hat{\theta}$. Using a statistical software package,

$$\overline{\theta}_B = \frac{1}{N}\sum_{I=1}^{N}\hat{\theta}_I = 4.626$$

for $i = 1, 2, \ldots, N$.

K-5.2.2.5. The confidence interval for θ, with 95% level of confidence reached upon 12 repetitions, is 4.323 to 4.93.

K-5.2.3. *Directions for a Jackknife Estimate of the Confidence Interval for* θ. Estimate $\hat{\theta}$ with all *n* samples from the data set.

K-5.2.3.1. Estimate $\hat{\theta}_{(i)}$ by removing the i^{th} sample (for $i = 1,2,\ldots,n$) from the original data set and use the same equation as was used to estimate $\hat{\theta}$.

K-5.2.3.2. Estimate the arithmetic mean, $\tilde{\theta}$, from the *n* estimates of $\hat{\theta}_{(i)}$, such that

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

for $i = 1,2,\ldots,n$. Note that the i^{th} "pseudo-value" is defined as $J_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}$.

K-5.2.3.3. Calculate the Jackknife estimator of θ (the average of the J_i values),

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

K-5.2.3.4. Estimate the standard error of the Jackknife estimate, $J\left(\hat{\theta}\right)$, by

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

K-5.2.3.5. Derive the confidence interval as

$$\left(J\left(\hat{\theta}\right) - t_{1-(\alpha/2),n-1}\,\hat{\sigma}_{J\left(\hat{\theta}\right)},\; J\left(\hat{\theta}\right) + t_{1-(\alpha/2),n-1}\,\hat{\sigma}_{J\left(\hat{\theta}\right)}\right)$$

with $100(1-\alpha)\%$ level of confidence; $t_{p,n-1}$ is the critical value from the Student's *t*-distribution for the $p100^{\text{th}}$ percentile and $n - 1$ degrees of freedom. If only a one-sided confidence interval is needed, then $t_{p,n-1} = t_{1-\alpha,n-1}$.

K-5.2.4. *Example of the Jackknife Method for Estimating a Nonparametric Confidence Interval for* θ. Using the same data set as for the Bootstrap example (Paragraph K-5.2.1), we will calculate a confidence interval for the mean (μ) using the Jackknife estimate with a 95% level of confidence.

K-5.2.4.1. Estimate $\hat{\theta} = 4.62$ with all 36 samples from the data set.

K-5.2.4.2. Estimate $\hat{\theta}_{(i)}$ for $i = 1, 2 \ldots n = 36$. The results are listed in Table K-5.

K-5.2.4.3. Estimate the arithmetic mean,

$$\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)} = 4.75\,.$$

K-5.2.4.4. Calculate the Jackknife estimator of θ (the average of the J_i values),

$$J(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} J_i = n\hat{\theta} - (n-1)\tilde{\theta} = 4.62 .$$

K-5.2.4.5. Estimate the standard error of the Jackknife estimate, $J(\hat{\theta})$, by

$$\hat{\sigma}_{J(\hat{\theta})} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(J_i - J(\hat{\theta})\right)^2} = 0.15 .$$

K-5.2.4.6. Derive the confidence interval,

$$J(\hat{\theta}) \pm t_{1-(\alpha/2),n-1}\,\hat{\sigma}_{J(\hat{\theta})}$$

$$J(\hat{\theta}) - t_{1-(\alpha/2),n-1}\,\hat{\sigma}_{J(\hat{\theta})} = 4.62 - 1.69 \times 0.15 = 4.37$$

$$J(\hat{\theta}) + t_{1-(\alpha/2),n-1}\,\hat{\sigma}_{J(\hat{\theta})} = 4.62 + 1.69 \times 0.15 = 4.87 .$$

K-5.2.4.7. The critical value from the Student's *t*-distribution was found using Table B-23 in Appendix B and linear interpolation.

K-5.3. *Tolerance and Prediction Intervals*. An approximate two-sided nonparametric prediction interval to contain the next single observation from the population with $(1-\alpha)100\%$ confidence can be estimated from the sample as $(x_{(l)}, x_{(u)})$ where

$$l = \frac{\alpha}{2}(n+1)$$

$$u = \left(1 - \frac{\alpha}{2}\right)(n+1)$$

and $x_{(i)}$ is the i^{th} order statistic from the sample data (Helsel and Hirsch, 2003). If l or u is not an integer, linearly interpolate between the values of the two surrounding order statistics. One-sided prediction limits can be calculated by replacing $\alpha/2$ with α when calculating the order statistic to use. An example calculation follows in Paragraph K-5.3.1.

K-5.3.1. *Example of a One-Sided Nonparametric Prediction Limit for the Next Single Observation.* A 95% upper prediction limit for arsenic concentration at a single point in the future is desired. Arsenic concentrations at three background wells were measured once each month for 12 months to yield 36 observations. Of the 36 observations, 19 were non-detects, so a nonparametric prediction limit will be calculated. The 95% upper prediction limit is calculated as:

$$x_{(u)} = x_{([1-\alpha][n+1])} = x_{([1-0.05][36+1])} = x_{(35.15)} \ .$$

Because 35.15 is not an integer, interpolate between the 35th and 36th order statistics. Suppose $x_{(35)}$=12 ppb and $x_{(36)}$=13 ppb. Then the 95% upper prediction limit is estimated to be:

$$x_{(35)} + 0.15\left(x_{(36)} - x_{(35)}\right) = 12 + 0.15(13-12) = 12.15 \,\text{ppb}.$$

If the result of the next observation were 8 ppb, we could conclude that arsenic concentration has not increased with 95% confidence.

K-5.3.2. *Discussion*. Exact confidence for using various order statistics from a sample to create nonparametric prediction intervals and limits can be calculated using the methods described in Hall et al. (1975). Their calculations expand to cover prediction intervals to contain *k* of *m* future observations instead of just a single future observation.

K-5.3.2.1. For small datasets, the method presented in Paragraph K-5.3.1 will require an order statistic that is smaller than the smallest observation in the dataset (for a minimum) or larger than the largest (for a maximum). In this situation, a nonparametric UTL or UPL is typically constructed using the minimum or maximum value of the set of observations. With high probability, the tolerance interval is designed to miss only a small percentage of the observations that arise from the same population as the data used to develop the tolerance limit. The coverage probability for the tolerance interval can be reported as either a minimum or an average value because, typically, we can only specify that the coverage probability of the interval exceed some level of confidence. We will use the average value. Given *n* measurements, using the maximum measurement as the UTL yields an *average* confidence of

$$\frac{n}{n+1}100\% .$$

Table K-5.
Estimate of $\hat{\theta}_{(i)}$ for i = 1, 2 … *n* = 36

i	Mean $\hat{\theta}$	J_i	$J_i - J(\hat{\theta})$	$(J_i - J(\hat{\theta}))^2$
1	4.67	2.8	–1.82	3.31
2	4.67	2.95	–1.67	2.78
3	4.66	3.15	–1.47	2.16
4	4.66	3.21	–1.41	1.98
5	4.65	3.56	–1.06	1.12
6	4.65	3.68	–0.94	0.88
7	4.64	3.91	–0.71	0.50
8	4.64	3.96	–0.66	0.43
9	4.64	4.01	–0.61	0.37
10	4.63	4.11	–0.51	0.26
11	4.63	4.14	–0.48	0.23
12	4.63	4.22	–0.40	0.16
13	4.63	4.25	–0.37	0.14
14	4.62	4.48	–0.14	0.02
15	4.62	4.53	–0.09	0.01
16	4.62	4.6	–0.02	0.00
17	4.62	4.62	0.00	0.00
18	4.62	4.72	0.10	0.01
19	4.62	4.72	0.10	0.01
20	4.62	4.73	0.11	0.01
21	4.61	4.76	0.14	0.02
22	4.61	4.8	0.18	0.03
23	4.61	4.81	0.19	0.04
24	4.61	4.85	0.23	0.05
25	4.61	4.94	0.32	0.10
26	4.61	5.09	0.47	0.22
27	4.60	5.1	0.48	0.23
28	4.60	5.12	0.50	0.25
29	4.60	5.17	0.55	0.30
30	4.60	5.27	0.65	0.42
31	4.59	5.51	0.89	0.79
32	4.58	5.81	1.19	1.42
33	4.58	5.91	1.29	1.67
34	4.58	5.99	1.37	1.88
35	4.57	6.19	1.57	2.47
36	4.56	6.6	1.98	3.93

K-5.3.2.2. Thus, at least 19 samples are necessary to achieve 95% mean coverage. Additional information on constructing nonparametric tolerance and prediction limits can be found in EPA 68-W0-0025.

K-5.3.2.3. A prediction limit involves the confidence probability associated with predicting that the next single observation will fall below the upper prediction limit, and is the same as the expected (mean) coverage of a similarly constructed UTL. Note that this is a special case for nonparametric prediction limits for the next single observation, not a general result. Furthermore, it can be shown that the probability of having k future samples all fall below the upper nonparametric prediction limit is $(1-\alpha) = n/(n+k)$ (i.e., the maximum value is the $[n/(n+k)]100\%$ upper prediction limit for k future measurements). Table B-11 in Appendix B lists these confidence levels for various choices of n and k. The false positive rate associated with a single prediction limit can be computed as one minus the confidence level. An example calculation follows in Paragraph K-5.3.3.

K-5.3.2.4. Balancing the ease with which nonparametric upper prediction limits are constructed is the fact that, given fixed numbers of original samples and future sample values to be predicted, the maximum confidence level associated with the prediction limit is also fixed. To increase the level of confidence, the only choices are to: i) decrease the number of future values to be predicted at any testing period, or ii) increase the number of original samples used in the test. Table B-11 of Appendix B can be used along these lines to plan an appropriate sampling strategy so that the false positive rate can be minimized and the confidence probability maximized to a desired level.

K-5.3.3. *Example of a Nonparametric Prediction Limit for the Next* k *Observations.* A prediction limit for arsenic concentration at $k = 2$ points in the future is desired. Arsenic concentration at three background wells was measured once each month for 6 months to yield 18 observations. As 9 of the 18 observations were non-detects, a nonparametric prediction limit will be calculated. The maximum detected result was 12 ppb, so this will be used as the upper prediction limit. Because $n = 18$ and $k = 2$, the probability of both future observations falling below the upper prediction limit of 12 is

$$100\frac{n}{n+k}\% = 100\frac{18}{18+2}\% = 90\%\,.$$

Thus 12 ppb is a 90% upper prediction limit for two future observations. The results of the two future observations were 8 and 14 ppb. As one of the new observations exceeds 12 ppb, we can conclude that arsenic concentration has increased with 90% confidence.

K-5.4. *Nonparametric Confidence Intervals for Percentiles*. A nonparametric confidence interval is based on an actual sample result and does not rely on any distributional assumptions. The nonparametric confidence interval is generally wider and requires more data than the corresponding normal distribution interval, and so the parametric distribution intervals should be used whenever it is appropriate. When $n \leq 20$, the nonparametric confidence interval is calculated using the binomial distribution.

K-5.4.1. Given a set of measurements, x_1, x_2,...x_n, to calculate a nonparametric confidence interval for the quantile X_p, it is necessary to first order the values of x_i so that $x_{(1)} < x_{(2)} <...< x_{(n)}$. Therefore, the smallest value of the data set is $x_{(1)}$ and the largest is $x_{(n)}$. (Note the distinction between x_1 and $x_{(1)}$; the former is the first measured value of the data set and the latter is the smallest value of the data set.) A two-sided nonparametric confidence interval for a quantile X_p will be of the form:

$$x_{(a)} \le X_p \le x_{(b)}$$

where the probability that X_p lies in the above interval is $1-\alpha$:

$$P(x_{(a)} \le X_p \le x_{(b)}) = 1-\alpha \quad .$$

K-5.4.2. The a^{th} largest value $x_{(a)}$ and b^{th} largest value $x_{(b)}$ of the data set (i.e., the numerical values of a and b that satisfy the above equation) are determined using the binomial distribution (as will be discussed below). Unfortunately, because the values are selected from a finite set of n ordered values $\{ x_{(i)} \}$, confidence limits are essentially being constructed for a discrete rather than a continuous variable. In general it will not be possible to select a and b so that the above probability is exactly equal to $1-\alpha$. Therefore, for the two-sided $1-\alpha$ confidence interval, a and b are selected so that:

$$P(x_{(a)} \le X_p \le x_{(b)}) \ge 1-\alpha \, .$$

K-5.4.3. Similarly, for an upper one-sided confidence interval for a percentile X_p it is desirable to select b so that:

$$P(X_p \le x_{(b)}) \ge 1-\alpha.$$

Find the lower bound $x_{(a)}$ by selecting the value of a so that:

$$Bin(a-1,n,p) \le \alpha/2 \text{ and } Bin(a,n,p) > \alpha/2$$

where $Bin(k,n,p)$ denotes the probability for the cumulative binomial distribution—the probability that an event with probability p of occurrence will happen less than or equal to k times out of n trials:

$$Bin(k,n,p) = P(K \le k) = \sum_{i=0}^{k} \left[\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \right] .$$

K-5.4.4. More information on the binomial distribution can be found in Appendix F. The values of *n* and *p* are known. Table B-1 of Appendix B lists values of the cumulative binomial distribution and lists various values of *k* for fixed values of *p* and *n*. Because *p* (the quantile) and *n* (the number of samples) are known, we can use Table B-1 to find the appropriate value of *k*. For example, one could start with $k = 0$, then $k = 1$, and so forth until $k = a - 1$ is the smallest value that satisfies the inequalities $Bin(a-1, n, p) \le \alpha/2$ and $Bin(a, n, p) > \alpha/2$. The upper bound $x_{(b)}$ is obtained by determining the smallest value of *b* that satisfies the relationship

$$Bin(b-1, n, p) - Bin(a-1, n, p) \ge 1 - \alpha .$$

K-5.4.5. For example, let us calculate the two-sided nonparametric confidence limit for the 75th percentile ($p = 0.75$) for the 90% level of confidence ($\alpha = 0.1$) for $n = 16$ so that:

$$P(x_{(a)} \le X_{0.75} \le x_{(b)}) \ge 0.9.$$

From Table B-1,

$$Bin(8, 16, 0.75) = 0.0271 < \alpha/2 = 0.05$$

and

$$Bin(9, 16, 0.75) = 0.0796 > \alpha/2 = 0.05 .$$

Therefore, $a = 9$. Because

$$Bin(14, 16, 0.75) - Bin(8, 16, 0.75) = 0.9365 - 0.0271 = 0.9094 > 1 - \alpha = 0.9$$

the value for $b = k + 1 = 14 + 1 = 15$. Therefore, the 90% confidence interval for the 75th percentile is $x_{(9)} \le X_{0.75} \le x_{(15)}$.

K-5.4.6. Similarly, find the one-sided $100(1-\alpha)\%$ upper confidence limit of X_p, so that the smallest value of *b* satisfies the equation

$$P(X_p \le x_{(b)}) = Bin(b-1, n, p) \ge 1 - \alpha .$$

K-5.4.7. Once b is found from Table B-1, the b^{th} largest value, $x_{(b)}$, establishes the upper $(1-\alpha)100\%$ confidence limit of X_p. For example, if $n = 20, p = 0.5,$ and $\alpha = 0.05,$

$$Bin(13, 20, 0.5) = 0.94 \text{ and } Bin(14, 20, 0.5) = 0.98.$$

K-5.4.8. Because $Bin(14, 20, 0.5) > 0.95$, $b = k + 1 = 14 + 1 = 15$. The 15th largest value of the data set, $x_{(15)}$, is at least the 95% upper confidence limit of the 50th percentile: $P(X_{0.5} \le x_{(15)}) \ge 0.95$.

K-5.4.9. If $n > 20$, the two-sided $100(1-\alpha)\%$ confidence interval $x_{(a)} \le X_p \le x_{(b)}$ can be calculated using a normality approximation so that $P(x_{(a)} \le X_p \le x_{(b)}) \approx 1 - \alpha$.

K-5.4.10. Calculate the following

$$a = np - Z_{1-\alpha/2}\sqrt{np(1-p)}$$

and

$$b = np + Z_{1-\alpha/2}\sqrt{np(1-p)}$$

where the percentile Z_p is the p^{th} quantile for the standard normal distribution obtained from Table B-15 of Appendix B. Round *a* an *b* to the nearest whole numbers and find the corresponding order values $x_{(a)}$ and $x_{(b)}$.

K-5.4.11. For the one-sided upper $100(1-\alpha)\%$ confidence interval $X_p \le x_{(b)}$, where $P\left(X_p \le x_{(b)}\right) \approx 1 - \alpha$, calculate

$$b = np + Z_{1-\alpha}\sqrt{np(1-p)}\,.$$

Round to the nearest whole number and find $x_{(b)}$.

K-5.4.12. Maximum detected values can be used to make inferences about percentiles. In particular, assume that a set of detected values are ranked from lowest to highest so that $x_{(n)}$ denotes the maximum value. Also assume that the maximum detected value is less than some threshold concentration (i.e., a risk-based limit) C: $x_{(n)} < C$. It can be shown that, under these circumstances, if X_p is the $p100^{th}$ percentile of X, then

$$P(X_p \le C) \ge 1 - p^n \text{ and } P(X_p > C) \le p^n\ .$$

X_p is less than the threshold C with at least $1 - p^n = 1 - \alpha$ confidence.

K-5.4.13. To find the value of n needed to achieve the desired level of confidence $(1-\alpha)100\%$, n must be such that

$$p^n \le \alpha .$$

Therefore, the $p100^{th}$ percentile, X_p, will be less the decision limit C with at least $(1-\alpha)100\%$ confidence if the maximum detected value is less than C (i.e., $x_{(n)} < C$) and

$$n \ge Ln(\alpha)/Ln(p) .$$

K-5.4.14. If, for example, $p = 0.90$ and $\alpha = 0.05$, then $n \ge 28.4$. If 29 samples are collected and the maximum value is less than C, then one can be at least 95% confident that the 90^{th} percentile is less than C.

K-5.4.15. The maximum is a non-parametric one-sided upper tolerance limit. Given a set of n observed measurements, there is $(1 - \alpha)100\% = (1 - p^n)100\%$ confidence that at least $p100\%$ of future measurements will be less than the maximum. A two-sided tolerance interval to contain at least a proportion p of future measurements may be constructed using the minimum and maximum of a set of n observed measurements. There is

$$(1-\alpha)100\% = \left(1 - p^n - n(1-p)p^{n-1}\right)100\%$$

confidence that at least $p100\%$ of future measurements will fall between the minimum and maximum of set of n observed data points. For example, if n =50 and p = 0.95, then there is 72% confidence that at least 95% of future measurements will fall between the minimum and maximum.

K-5. **Statistical Confidence Interval for Proportions**. Data from a binomial distribution are composed of only two responses—"pass" or "fail." The population proportion, P, is based on either the passing proportion or the failing proportion. The following discussion will (arbitrarily) define the proportion, p, as the proportion of failures. An estimate of this proportion can be derived by $p = k/n$ where k is the number of failures out of n samples. For example, in environmental applications p could represent the proportion of results from samples below some decision limit, C. From this information we would like to estimate an interval, (P_L, P_U), which contains the true proportion, P, of the distribution that is less (or greater) than C. The binomial distribution is a discrete distribution and so statistical intervals are approximate and tend to be conservative (Hahn and Meeker, 1991). The most frequent statistical interval calculated for a proportion is the confidence interval, so only it is presented here.

K-5.1. *Discussion*. The equation for a conservative two-sided $100(1 - \alpha)$ % confidence interval for a proportion is the following:

$$[p_L, p_U] = \left[\frac{1}{1 + \left\{ \frac{(n-k+1) F_{1-\alpha/2,\ 2n-2k+2,\ 2k}}{k} \right\}}, \frac{1}{1 + \left\{ \frac{(n-k)}{(k+1) F_{1-\alpha/2,\ 2k+2,\ 2n-2k}} \right\}} \right]$$

where $F_{\gamma,m,n}$ is the $\gamma 100^{th}$ percentile of the F distribution (Table B-7 of Appendix B) with m and n degrees of freedom. The lower limit, P_L, is defined to be 0 if $k = 0$, and the upper limit, P_U, is defined to be 1 if $k = n$ (Hahn and Meeker, 1991).

K-5.1.1. Likewise, a one-sided $(1-\alpha)100\%$ LCL for a proportion would be:

$$p_L = \frac{1}{1 + \left\{ \frac{(n-k+1) F_{1-\alpha,\ 2n-2k+2,\ 2k}}{k} \right\}}$$

while a one-sided $(1-\alpha)100\%$ UCL for a proportion would be:

$$p_U = \frac{1}{1 + \left\{ \frac{(n-k)}{(k+1) F_{1-\alpha,\ 2k+2,\ 2n-2k}} \right\}} .$$

K-5.1.2. If a large number of samples are available, these confidence intervals can be approximated. However, two restrictions apply to the data set: first, $np \geq 5$ and second, $n(1-p) \geq 5$. This approximated confidence interval is based on the normal distribution because when these two restrictions apply, data are approximately normally distributed. The equation for the approximated confidence interval is:

$$[p_L, p_U] = p \pm Z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)100^{th}$ percentile from a standard normal, *n* is the sample size, and p is the sample proportion (Devore, 1987). The one-sided upper confidence limit would be found by replacing $1-\alpha/2$ with $1-\alpha$ as follows:

$$p_U = p + Z_{1-\alpha}\sqrt{\frac{p(1-p)}{n}} \quad .$$

K-5.2. *Example of a Confidence Limit for a Proportion*. Groundwater concentrations of gasoline at a site are compared to a regulatory threshold of 35 micrograms per liter (µg/L). Suppose out of 90 results, 11 of the samples have concentrations that exceed this regulatory threshold, so the proportion of samples with detected concentrations exceeding the threshold is $p = 11/90 = 0.1222$.

$$np = 90 \times 0.1222 = 11.00$$

$$n(1-p) = 90 \times (1-0.1222) = 79.00\,.$$

As both *np* and $n(1-p)$ are greater than or equal to 5, the large sample normal approximation can be used

$$p_L = p - Z_{1-\alpha}\sqrt{\frac{p(1-p)}{n}} = 0.1222 - 1.282\sqrt{\frac{0.1222(1-0.1222)}{90}} = 0.078$$

where Table B-15 of Appendix B is used to find the critical value $Z_{0.90}$=1.282. Because p_L exceeds 0.05, we can accept that more than 5% of the concentrations of gasoline in groundwater at the site exceed the regulatory threshold as we conclude also in Appendix L, Paragraph L-8.2.

K-6. Statistical Intervals for the Poisson Distribution (Number of Occurrences). Data from a Poisson distribution are composed of only two mutually exclusive responses—"pass" or "fail" s—when the probability of one of the responses is small. Poisson distributions are common when counting the number of pass or fail occurrences over a time interval or the number of detections when a set of measured concentrations consists mostly of non-detects. The population rate of occurrence, μ, also called the mean rate of occurrence, is either based on the passing rate or the failing rate. For this document, the rate of occurrence of the rare event is called the rate of "failure." The Poisson distribution is a discrete distribution and so statistical intervals are approximate and tend to be conservative (Hahn and Meeker, 1991). An estimate of this rate of occurrence can be derived by:

$$\hat{\mu} = k / n$$

where k is the number of failures out of n samples.

K-6.1. *Confidence Interval for the Mean Occurrence Rate*. A two-sided $(1-\alpha)100\%$ confidence interval for the mean occurrence rate is the following:

$$\left[\hat{\mu}_L,\ \hat{\mu}_U\right] = \left[\frac{0.5\chi^2_{\alpha/2,\ 2k}}{n},\ \frac{0.5\chi^2_{1-\alpha/2,\ 2k+2}}{n}\right]$$

where $\chi^2_{\gamma,\nu}$ is the $\gamma\ 100^{\text{th}}$ percentile of the chi-square distribution (Table B-2 of Appendix B) with v degrees of freedom.

K-6.1.1. A one-sided lower or upper $(1-\alpha)100\%$ confidence limit can be obtained by replacing $\chi^2_{\alpha/2,2k}$ with $\chi^2_{\alpha,2k}$ for a lower confidence limit or replacing $\chi^2_{1-\alpha/2,2k+2}$ with $\chi^2_{1-\alpha,2k+2}$ for an upper confidence limit (Hahn and Meeker, 1991).

K-6.1.2. If a large number of samples is available (generally, if $n > 20$), this confidence interval can be approximated. This approximated confidence interval is based on the normal distribution because, as the sample size increases, the data's distribution tends towards normality. The equations for the approximated confidence interval are:

$$\left[\hat{\mu}_L,\ \hat{\mu}_U\right] = \hat{\mu} \pm Z_{1-\alpha/2}\sqrt{\frac{\hat{\mu}}{n}}$$

where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)100^{th}$ percentile from a standard normal, n is the sample size, and $\hat{\mu}$ is the mean sample rate of failure ($\hat{\mu} = k/n$ when k is the number of failures in n samples) (Hahn and Meeker, 1991).

K-6.2. *Upper Tolerance Limit*. A Poisson tolerance interval, with $p100\%$ coverage and $(1-\alpha)100\%$ confidence, is calculated based on the directions given in Paragraph K-6.2.1, followed by an example in Paragraph K-6.2.2.

K-6.2.1. *Directions for Calculating a Poisson Tolerance Interval with p100% Coverage and* $(1-\alpha)100\%$ *Confidence.* Compute the sum of the Poisson counts of n samples:

$$T' = \sum_{i=1}^{n} x_i \quad .$$

This is the sum of the detected values and one-half the sum of all the non-detected values.

K-6.2.1.1. Find the probable rate

$$\mu = \frac{1}{2n} \chi^2_{1-\alpha, 2T'+2}$$

where $\chi^2_{1-\alpha, 2T'+2}$ is the $(1-\alpha)100^{th}$ percentile of the chi-squared distribution with $\nu = 2T' + 2$ degrees of freedom. Table B-2 of Appendix B contains a table of critical values for the chi-square distribution.

K-6.2.1.2. Compute the $p100^{\text{th}}$ percentile of the Poisson distribution with mean rate μ, by finding the least positive integer k such that

$$\chi^2_{1-p, 2k+2} \geq 2\mu \, .$$

As above, the quantity $2k + 2$ represents the degrees of freedom of the chi-squared distribution. *The quantity* k *itself is the upper tolerance limit (UTL) for the Poisson distribution*. In other words, for the smallest value of k for which

$$\chi^2_{1-p, 2k+2} \geq \frac{1}{n} \chi^2_{1-\alpha, 2T'+2}$$

$p100\%$ of the measurements will be less than k with $(1-\alpha)100\%$ confidence. If any sample exceeds the UTL, k, then there is significant evidence that this sample is different from the samples used to develop the UTL.

K-6.2.2. *Example of Calculating a Poisson Tolerance Interval with p100% Coverage and* $(1-\alpha)100\%$ *Confidence*. A tolerance interval with 95% confidence ($\alpha = 0.05$) and 95% coverage ($p = 0.95$) is desired for 1,1-dicholorethene in groundwater at Site B. The background well values in Table K-6 were obtained. These data have more than 90% non-detects and the number of samples $n = 90$.

K-6.2.2.1. Calculate the sum of the Poisson counts: Sum the detections and to this value add one half the sum of the non-detects (one half the detection limit is being used for each non-detect).

$$T' = 0.5 \times (7.1236) + (0.111 + 0.138 + 2.63 + 4.81) = 11.25$$

$$\nu = 2T' + 2 = 2\ (11.25) + 2 = 24.5 \approx 25$$

$$\mu = \frac{1}{2n}\chi^2_{1-\alpha,2T'+2} = \frac{1}{2 \times 90}\chi^2_{0.95,25} = 0.209$$

where $\chi^2_{1-\alpha,2T+2} = \chi^2_{0.95,25} = 37.65$ using Table B-2 of Appendix B.

K-6.2.2.2. So, we need to find the smallest value of k such that $\chi^2_{1-p,2k+2} \geq 2\mu$; that is, the value of k such that $\chi^2_{0.05,2k+2} \geq 0.418$. Table B-2 of Appendix B shows that the smallest value number of degrees of freedom, $\nu = 2k + 2$, that satisfies the above equation is $\nu = 4$. Since $4 = 2k + 2$, $k = 1.0$.

k	**df**	$\chi^2_{0.005}$
0.5	3	0.3518
1	4	0.7107
1.5	5	1.145

K-6.2.2.3. If any site groundwater sample exceeds the UTL of 1.0 µg/L derived from the background wells, then there is significant evidence that contamination at the site is elevated with respect to background.

K-6.2.3. *Upper Prediction Limit*. To estimate a prediction limit using the Poisson model, the upper limit is estimated for an interval that will contain all of k future measurements of an analyte with $100(1-\alpha)\%$ confidence, given n previous measurements. The directions to calculate such a prediction limit are provided in Paragraph K-6.2.3.1 and followed by an example in Paragraph K-6.2.3.2.

K-6.2.3.1. *Directions for Estimating a Prediction Limit Using the Poisson Model*. Calculate T', the sum of the Poisson counts of n samples (e.g., for the background data set), as defined in Paragraph K-6.2.1.

K-6.2.3.1.1. Calculate T_k^*, the greatest total Poisson count for the next k samples (e.g., for the study area data set) at some level of confidence, $1 - \alpha$ using the following equation:

$$T_k^* = \frac{T'}{n} + \frac{t^2}{2n} + \frac{t}{n}\sqrt{T'\left(1+n\right) + \frac{t^2}{4}}$$

where $t = t_{1-\alpha/k,n-1}$ is the upper $(1-\alpha/k)100\%$ percentile of the Student's *t*-distribution with *n* – 1 degrees of freedom, in Table B-23 of Appendix B.

Table K-6.
Background Wells

Well Location	Result (µg/L)	DL (µg /L)	Well Location	Result (µg /L)	DL (µg /L)
Site B-MW01		0.0819	SiteB-MW02		0.144
SiteB-MW01		0.102	SiteB-MW02		0.0715
SiteB-MW01		0.102	SiteB-MW02		0.0715
SiteB-MW01		0.0715	SiteB-MW02		0.145
SiteB-MW01		0.0436	SiteB-MW03		0.144
SiteB-MW01		0.0436	SiteB-MW03		0.0715
SiteB-MW01		0.122	SiteB-MW03		0.0715
SiteB-MW02		0.0819	SiteB-MW03		0.0715
SiteB-MW02		0.102	SiteB-MW04		0.144
SiteB-MW02		0.102	SiteB-MW04		0.0715
SiteB-MW02		0.0715	SiteB-MW04		0.0715
SiteB-MW02	0.111		SiteB-MW04		0.0715
SiteB-MW02		0.0436	SiteB-MW05		0.144
SiteB-MW02		0.122	SiteB-MW05		0.0715
SiteB-MW03		0.0819	SiteB-MW05		0.0715
SiteB-MW03		0.102	SiteB-MW05		0.0715
SiteB-MW03		0.102	SiteB-MW06		0.0715
SiteB-MW03		0.0715	SiteB-MW06		0.0715
SiteB-MW03		0.0436	SiteB-MW06		0.0715
SiteB-MW03		0.0436	SiteB-MW06		0.145
SiteB-MW03		0.122	SiteB-MW01		0.116
SiteB-MW04		0.0819	SiteB-MW01		0.116
SiteB-MW04		0.102	SiteB-MW01		0.0492
SiteB-MW04		0.102	SiteB-MW01		0.0492
SiteB-MW04		0.0715	SiteB-MW02		0.116
SiteB-MW04		0.0436	SiteB-MW02	0.138	
SiteB-MW04		0.0436	SiteB-MW02		0.0492
SiteB-MW04		0.122	SiteB-MW02		0.0492
SiteB-MW05		0.0819	SiteB-MW03		0.116
SiteB-MW05		0.102	SiteB-MW03		0.116
SiteB-MW05		0.102	SiteB-MW03		0.0492
SiteB-MW05		0.0715	SiteB-MW03		0.0492
SiteB-MW05		0.0436	SiteB-MW04		0.116
SiteB-MW05		0.0436	SiteB-MW04		0.116
SiteB-MW05		0.122	SiteB-MW04		0.0492
SiteB-MW06		0.0819	SiteB-MW04		0.0492
SiteB-MW06		0.102	SiteB-MW05		0.116
SiteB-MW06		0.102	SiteB-MW05		0.116
SiteB-MW06		0.0715	SiteB-MW05	2.63	
SiteB-MW06		0.0436	SiteB-MW05		0.0492
SiteB-MW06		0.0436	SiteB-MW06		0.116

Well Location	Result (µg/L)	DL (µg /L)	Well Location	Result (µg /L)	DL (µg /L)
SiteB-MW06		0.122	SiteB-MW06		0.116
SiteB-MW01		0.144	SiteB-MW06	4.81	
SiteB-MW01		0.0715	SiteB-MW06		0.0492
SiteB-MW01		0.0715			
SiteB-MW01		0.0715			

K-6.2.3.1.2. If the sum of Poisson counts for the next k samples is greater than the upper prediction limit T_k^*, then there is significant evidence of a difference in the new samples, compared to previous samples.

K-6.2.3.2. *Example of Estimating a Prediction Limit Using the Poisson Model*. Suppose a prediction limit for the next two observations with 99% confidence is desired for 1,1-dicholorethene from Site B with the following background wells. NOTE: These data have more than 90% non-detects. (See data table in Paragraph K-6.2.2.)

K-6.2.3.2.1. Calculate the sum of the Poisson counts:

$$T' = 0.5 \times (7.1236) + (0.111 + 0.138 + 2.63 + 4.81) = 11.25$$

$$T_k^* = \frac{T'}{n} + \frac{t^2}{2n} + \frac{t}{n}\sqrt{T'(1+n) + \frac{t^2}{4}}$$

$$= \frac{11.25}{90} + \frac{(2.639)^2}{2(90)} + \frac{2.639}{90}\sqrt{11.25(1+90) + \frac{(2.639)^2}{4}} = 1.10$$

where n = 90 and $t_{1-\alpha/k,n-1} = t_{(1-0.01)/2,(90-1)} = t_{0.995,89} = 2.639$ using Table B-23 of Appendix B and linear interpolation.

K-6.2.3.2.2. To test the upper prediction limit, if the sum of the Poisson counts for the next k samples (k = 2) is greater than T_k^* (1.10), then there is significant evidence the contamination in the site wells is elevated relative to the background wells.