Left-Censored Data Applications

These example applications for S-PLUS are designed to demonstrate the computation of summary statistics for left-censored data using applications in the USGS library. Left-censored data result from laboratory analyses when the recorded value is less than the detection limit. A general knowledge of S-PLUS is assumed. For users needing introductions to S-PLUS, several are available in bookstores and online. Examples include Burns, Patrick, 2008, S Poetry, last accessed March 19, 2009 at *http://www.burns-stat.com/pages/spoetry.html*, and Lam, Longhow, 2001, An introduction to S-PLUS for Windows: Amsterdam, The Netherlands, CANdiensten, 230 p., last accessed March 26, 2010, at *http://www.splusbook.com*. A detailed introduction to the S language is Kraus, A., and Olson, M., 2002, The basics of S-PLUS: New York, Springer, 420 p. Another example of an introduction that concentrates on statistical applications is Venables, W.N., and Ripley, B.D., 2002, Modern applied statistics with S: New York, Springer, 495 p.

A few type face conventions are used to assist the user in interpreting the guidelines. S-PLUS object names and function arguments are in *italics* and function names are followed by parentheses *()*, window names are highlighted in **bold**, and object class names are in a sans-serif font. User commands, called calls, are in `plain text` and on separate lines. Any printed output from a call, whether in a **Commands Window** or **Report Window**, is shown in `plain text` in a box. Plots are shown as is, copied from the default settings for an S-PLUS **Graph Sheet**; however, individual user preferences can change the way the plots are displayed. Any S-PLUS dialog menu directions are shown in **bold** separated by a vertical bar (|). Column names and other dialog box entries are underlined.

An important way to manage data is to use the **Object Explorer Window**. The **Object Explorer Window** can be opened by clicking on the small icon that is a little yellow folder with two blue dots above it  .

Introduction 1. *Working with censored data*
Introduction 2. *Naming conventions and data storage*
Application 1. *Nonparametric summary statistics*
Application 2. *Parametric summary statistics*

Working with censored data

Helsel and Hirsch (2002) discuss methods for the estimation of descriptive statistics for single- and multiple-detection limit censored water-quality data. Helsel and Cohn (1988) indicate that regression on order statistics (ROS) and adjusted maximum likelihood estimation (AMLE) give better results than simple substitution of values (such as the detection limit) for censored values. The ROS method is preferred for estimates of mean and standard deviation, and the AMLE method for estimates of percentiles. Both methods have shortcomings.

WARNING: Limitations of the ROS and AMLE methods for computing descriptive statistics from datasets containing "less than" values must be understood before using these methods.

The ROS method ignores the value of the detection limit in computing estimates and should not be used with datasets whose highest detection limit is substantially below the smallest "above-detection-limit" observation.

Statisticians refer to two types of censored ("less than") data: "Type I" and "Type II." Type I data occur when a *fixed threshold* is specified and observations below the threshold are reported as censored. Type II censoring occurs when a *fixed proportion* of the smallest observations are censored, such as in testing light bulbs until 1 percent fail, and there is no censoring threshold. The ROS routine implicitly assumes Type II censoring and uses the proportional information but not the value of the detection limits. In contrast, Type I censoring associates observations with analytical detection limits, such as in water-quality data.

In cases when all the censor thresholds are near or above the smallest observed value in the dataset, the estimators for Type I and Type II data tend to converge. In this case, the ROS results may be acceptable.

The AMLE method is a parametric method that assumes that the data are log-normally distributed. If the data deviate substantially from the log-normal distribution or the sample size is small, then the AMLE method performs poorly.

For additional discussion, see Kendall and Stuart (1979) and Shumway and others (2002).

Example 1: The following example illustrates the limitations of both ROS and AMLE (Note: remarks codes are the same for both sequences; it is only the detection limit that differs.):

```
rcodes <- c(" ", " ", " ",  "<",  "<",  "<",  "<",  "<",  "<",  "<",  "<")
X1     <- c( 75,  50,  25, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05)
X2     <- c( 75,  50,  25,   25,   25,   25,   25,   25,   25,   25,   25)
# Compute ROS and AMLE summaries for X1 and X2
# Output for AMLE edited for readability in the output for X1
mdlstats(lcens(X1, rcodes=="<"))
```

```
      Parametric Summary Statistics for lcens(X1, rcodes == "<"):
                 ROS           AMLE
Mean logs  2.439225      -7.801069
 SD. logs  1.176367       8.907328
     Mean 20.085243    1362.587
       SD 22.818190   12296.60
      10%  1.969401       4.513153e-009
      25%  4.793733       1.006561e-006
      50% 11.396231        .000409
      75% 25.000000      25.00000
      90% 70.000000      70.00000
        N 11.000000      11.00000
      NLT  8.000000       8.00000
Censoring levels and counts
 0.05
    8
Range of uncensored observations
 Min Max
  25  75
```

```
mdlstats(lcens(X2, rcodes=="<"))
```

```
      Parametric Summary Statistics for lcens(X2, rcodes == "<"):
                 ROS      AMLE
Mean logs  2.439225  2.651069
 SD. logs  1.176367  0.949522
     Mean 20.085243 21.039648
       SD 22.818190 22.416721
      10%  1.969401  4.196317
      25%  4.793733  7.467990
      50% 11.396231 14.169172
      75% 25.000000 25.000000
      90% 70.000000 70.000000
        N 11.000000 11.000000
      NLT  8.000000  8.000000
Censoring levels and counts
 25
  8
Range of uncensored observations
 Min Max
  25  75
```

When the ROS results for X1 are examined, clearly all of the quantiles less than 75 percent should be values less than 0.05. Similarly, the mean of the observations mathematically (substituting

first 0.0 and then 0.05 for each of the censored values) must lie in the interval (13.63, 13.67), and cannot be as large as the reported 20.09. The ROS method should not be used in this case; the results are nonsense. For X2, all of the summary statistics—the computed mean, standard deviation, and quantiles—are identical to those for X1. In this case, however, the summary statistics are not unreasonable.

The AMLE results for X1 provide estimates of the mean and standard deviation that are nonsensical. This arises because the assumption of a log-normal distribution for the underlying data is unreasonable. For X2, the assumption of log-normality is reasonable and the results also are reasonable.

WARNING: Limitations of the nonparametric Kaplan-Meier method for computing descriptive statistics from datasets containing "less than" values must be understood before using that method.

Helsel (2005) describes a modification of the Kaplan-Meier (K-M) method for computing summary statistics for left-censored data. The K-M method is used for right-censored survival data. To modify it for left-censored data, the data are flipped by subtracting the values from some arbitrary value, which can be 0. That flipping retains order, relative magnitude and converts “less-than” to “greater-than” values. The summary statistics are then re-flipped back to the original units.

The K-M method cannot extrapolate beyond the range of the data. If the lowest values of water-quality data are censored, then the K-M method produces an estimated mean that is biased high. In effect, it uses simple substitution of the lowest value to compute the mean. The USGS library function that computes the summary statistics does not print the estimated mean.

The K-M method also ignores the relative magnitude of the censor level if there are no observed values between the censored values. The process of allocating the survival probability of censored observations distributes the observational weight equally over all censored and uncensored values less than the censored observation. See Helsel (2005) for details about the computation of the K-M method.

Example 2: The following example uses the data from example 1 and illustrates the behavior of the K-M method. (Note: the only difference between X2 and X3 is in the lowest censored value.):

```
X3     <- c( 75,  50,  25,   25,   25,   25,   25,   25,   25,   25,    5)

mdlnpar(lcens(X2, rcodes=="<"))
```

```
      Nonparametric Summary Statistics for lcens(X2, rcodes == "<"):
  Number of observations: 11, Number of censored observations: 8.
```

```
  Censoring levels and counts
 25
  8
  Quantiles and 95 percent confidence limits
       estimate lcl ucl
72.73%       25  NA  50
   75%       25  NA  50
   90%       50  NA  NA
  100%       75  75  NA
```

```
mdlnpar(lcens(X3, rcodes=="<"))
```

```
      Nonparametric Summary Statistics for lcens(X3, rcodes == "<"):
  Number of observations: 11, Number of censored observations: 8.
  Censoring levels and counts
 25  5
  7  1
  Quantiles and 95 percent confidence limits
       estimate lcl ucl
72.73%        5  NA  50
   75%       25  NA  50
   90%       50  NA  NA
  100%       75  75  NA
```

Examining the results for both reveals that the estimates for the 75th, 90th, and 100th percentiles are the same. But the results for X3 indicate that 72.73 percent of the data are less than 5, which is the result of collapsing all of the <25 values to <5 in the process of allocating the survival probability.

REFERENCES:

Helsel, D.R., 2005, Nondetects and data Analysis: Hoboken, New Jersey, John Wiley & Sons, 250 p.

Helsel, D.R., and Cohn, T.A., 1988, Estimation of descriptive statistics for multiply censored water quality data: Water Resources Research, v. 24, no. 12, p. 1997-2004.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey, Techniques of Water-Resources Investigations book 4, chap. A3, 523 p.

BSA memo on Less Thans and successor writings.

Kendall, M., and Stuart, A., 1979, The advanced theory of statistics, v. 2: London, Charles Griffin & Co., p. 551-552.

Shumway, R.H., Azari, R.S., and Kayhanian, M., 2002, Statistical approaches to estimating mean water quality concentrations with detection limits: Environmental Science and Technology, v. 36, p. 3345-3353.

Naming conventions and data storage

The examples use the *WI1MajorIons.df* and *WI3TotalP.df* water-quality datasets. These datasets can be copied to the user's chapter by clicking and dragging in the left pane of the **Object Explorer Window**, or by typing the following calls in the **Commands Window**.

```
WI1MajorIons.df <- WI1MajorIons.df
WI3TotalP.df <- WI3TotalP.df
```

S-PLUS requires that numeric values and any censoring symbology be stored in separate variables. If left-censored values and symbols are provided in a single column, the *splitQual()* function will to split the single column into two columns. See the documentation for *splitQual()* in the USGS library for S-PLUS for information about its use.

Column names can be any legal S-PLUS name, but the USGS menu dialog functions recognize two standard naming conventions—the "Booker" convention, which uses the prefix P on the numeric values and the prefix R on the remark codes, and an alternate convention that simply postfixes ".rmk" to the name of the numeric value column for the remark code column. Both datasets in this example use the latter naming convention.

Remark codes can vary depending on the system. S-PLUS allows character strings to have any value, so it is possible to store remark codes like "<", "E", and "". The standard use of the *lcens()* function, which is the user interface to all left-censored data analysis functions in the USGS library, only identifies left-censored values and uncensored values. See the documentation for *lcens()* about analyzing information-rich data for alternate uses. Some water-quality data may use a 0 or 1 to indicate censored and uncensored values. The user must know which value (0 or 1) is associated with censored or uncensored data. The USGS menu dialog functions recognize only the "<" symbol for left-censored data and all other codes are ignored. Note that in some cases, character data are converted to factor data when the data are imported and the blank string "" is coded as missing. If that happens, all uncensored values are set to missing values and all analytical methods fail.

Nonparametric Summary Statistics

This application illustrates how to estimate summary statistics (quantiles) of left-censored data. The Kaplan-Meier method using flipped data is used in this application. The method is described in Helsel (2005).

The examples in this section use the *WI1MajorIons.df* and *WI3TotalP.df* water-quality datasets. See the *Naming conventions and data storage* section (Introduction 2) for information about those datasets.

Step 1a. Estimate summary statistics using the Commands Window

Use the *mdlnpar()* function to create a mdlnpar object and print it by typing the following call in the **Commands Window**:

```
WI1Nitrate <- print(mdlnpar(lcens(WI1MajorIons.df$Nitrate,
      WI1MajorIons.df$Nitrate.rmk=="<")))
```

This call produces the following output:

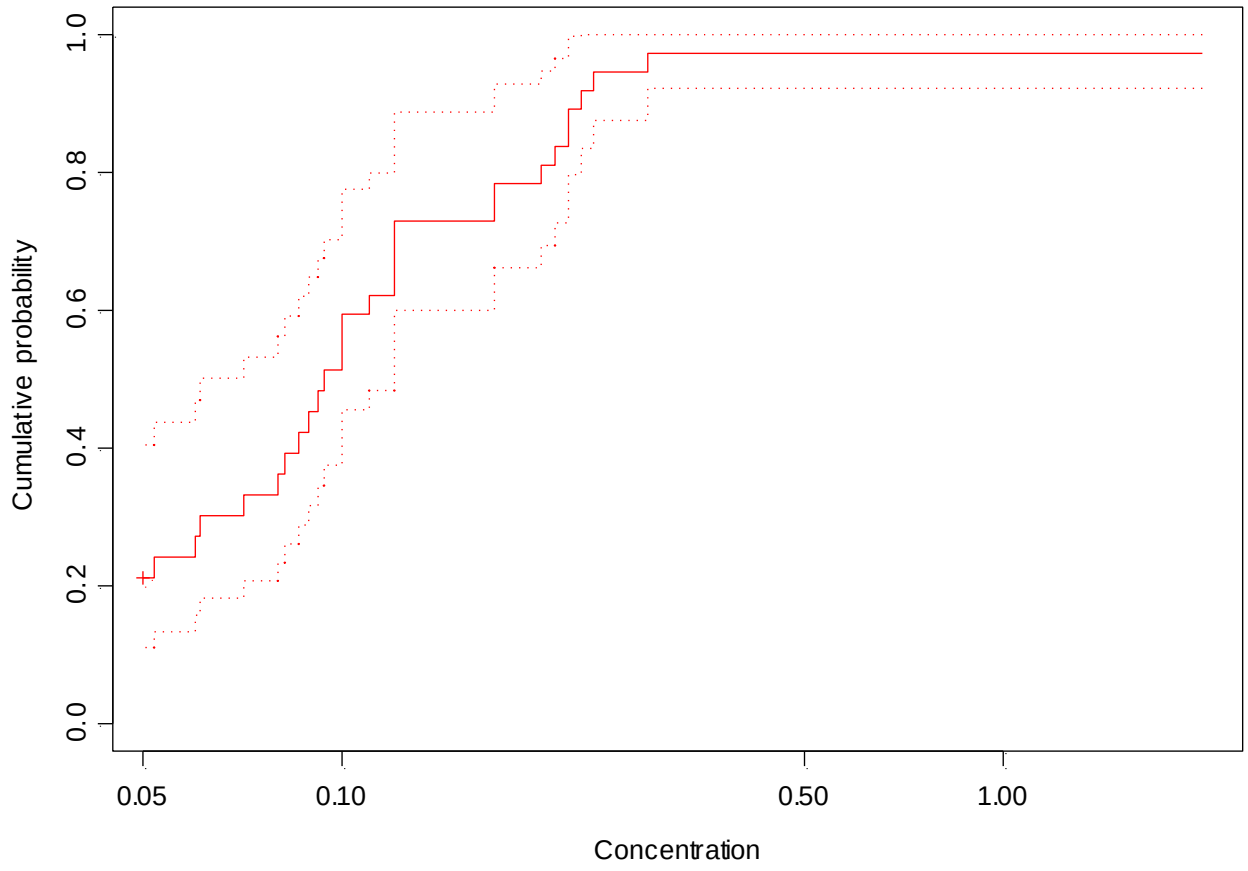
```
      Nonparametric Summary Statistics for lcens(WI1MajorIons.df$Nitrate,
WI1MajorIons.df$Nitrate.rmk == "<"):

  Number of observations: 37, Number of censored observations: 9.
  Censoring levels and counts
 0.10 0.05
    2    7
  Quantiles and 95 percent confidence limits
       estimate   lcl   ucl
21.14%    0.050    NA 0.080
   25%    0.060    NA 0.082
   50%    0.094 0.061 0.120
   75%    0.170 0.100 0.220
   90%    0.230 0.170 0.290
  100%    2.000 2.000    NA
```

This report includes upper and lower 95-percent confidence intervals of the estimated quantiles. The S-PLUS function *survfit()* is to compute the statistics for *mdlnpar()*; see the documentation for that function for details on the computation of the confidence intervals. The missing value symbol, NA, is printed for values outside of the range of the data. The minimum quantile that can be estimated is printed—in this case 21.14 percent of the data are estimated as less than 0.050. This is slightly more than the 18.92 percent that are coded as <0.05 because of the presence of higher censoring levels.

It is useful to graphically review the data. In the **Commands Window** type the following call to produce an empirical cumulative distribution curve:

```
plot.mdlnpar(WI1Nitrate)
```

Note that the + sign is used to identify censored values when there is no uncensored value at that particular value. There is a + at 0.05, but not at 0.10.



An alternate output format is available using the *summary()* function shown below. This report includes more information about censoring levels and observed values within the range of the censoring levels. In particular, the nobs.lt indicates that there are 10 uncensored values less than 0.10 and none less than 0.05.

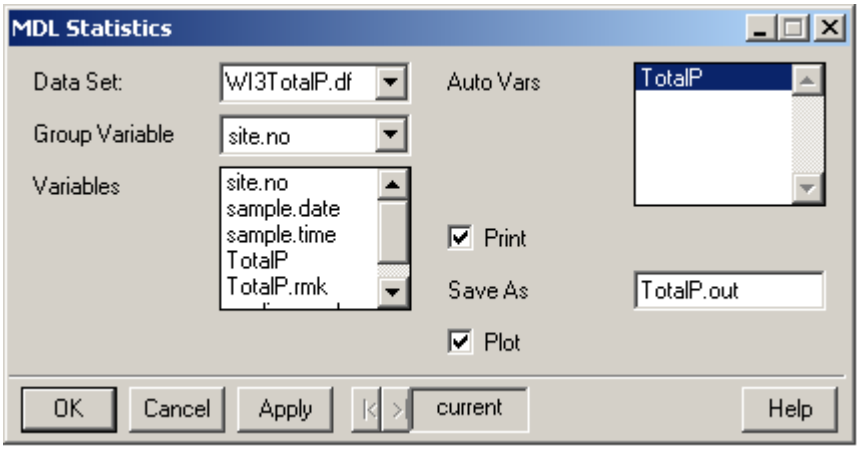
```
summary(WI1Nitrate)
```

```
lcens(WI1MajorIons.df$Nitrate, WI1MajorIons.df$Nitrate.rmk == "<")
```

```
 21.14%  25%   50%  75%  90% 100%
   0.05 0.06 0.094 0.17 0.23    2
  Censoring levels, counts, and number observed below level (37 observations)
           0.05 0.10
    counts    7    2
cumulative    7    9
   nobs.lt    0   10
```

Step 1b. Estimate summary statistics using the dialog menus

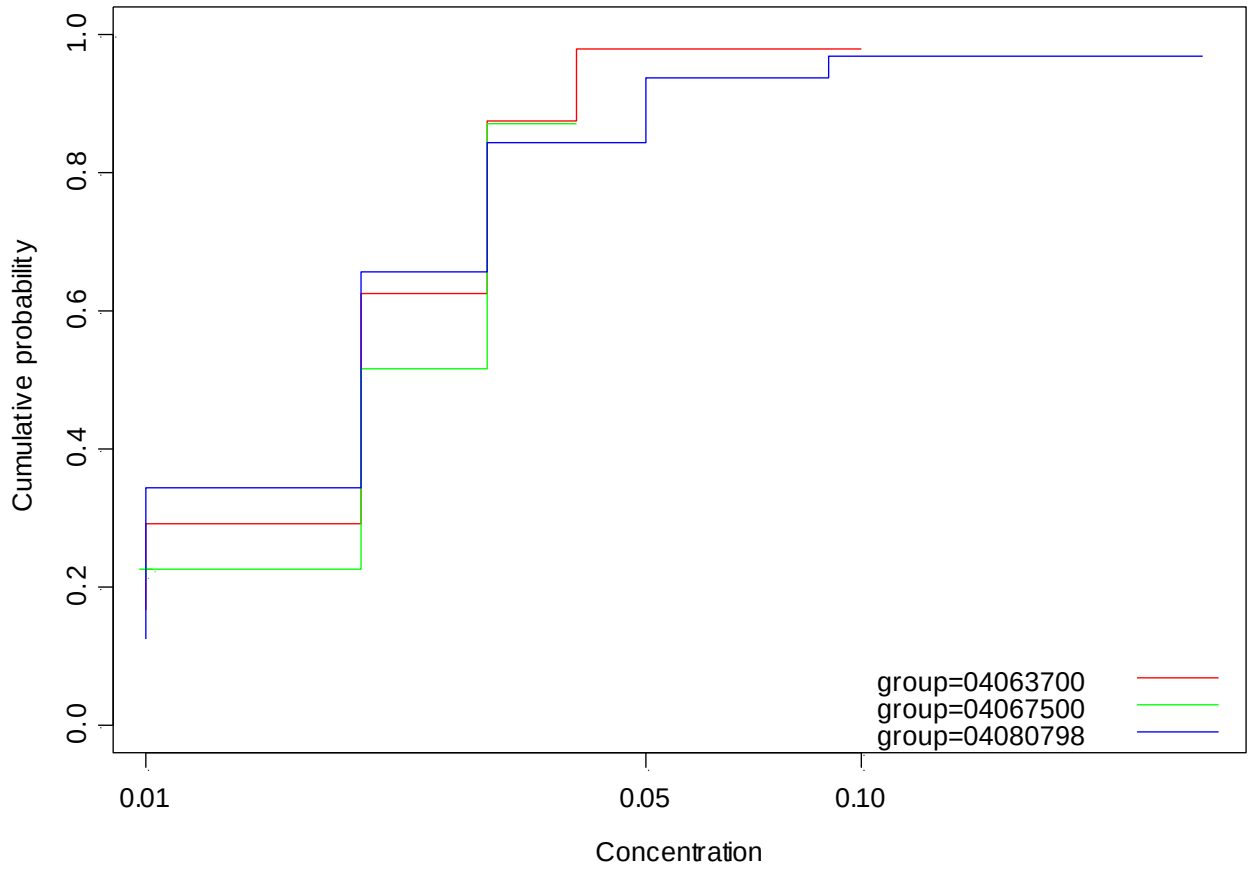
Click on **USGS | Left-Censored Data Analysis | Nonparametric summary statistics…** to open the **MDL Statistics** dialog window.



Select WI3TotalP.df in the drop-down box for **Data Set:**. Because there are three sites in this dataset, site.no should be selected in the **Group Variable** box. This dialog window recognizes the TotalP and TotalP.rmk column pairs and puts a single entry in the **Auto Vars** box. Select TotalP in that box. Replace the contents of the **Save As** box with TotalP.out. It will be used in Step 2. Click OK to generate the report and graph.

```
      Nonparametric Summary Statistics for TotalP:
group=04063700
  Number of observations: 48, Number of censored observations: 8.
  Censoring levels and counts
 0.01
    8
  Quantiles and 95 percent confidence limits
       estimate  lcl  ucl
16.67%     0.01   NA 0.01
   25%     0.01   NA 0.02
   50%     0.02 0.02 0.02
   75%     0.03 0.02 0.03
   90%     0.04 0.03 0.04
  100%     0.10 0.10   NA

group=04067500
  Number of observations: 31, Number of censored observations: 7.
```

```
  Censoring levels and counts
 0.01
    7
  Quantiles and 95 percent confidence limits
       estimate  lcl  ucl
22.58%     0.01   NA 0.02
   25%     0.02   NA 0.02
   50%     0.02 0.02 0.03
   75%     0.03 0.03 0.03
   90%     0.04 0.03   NA
  100%     0.04 0.04   NA

group=04080798
  Number of observations: 32, Number of censored observations: 4.
  Censoring levels and counts
 0.01
    4
  Quantiles and 95 percent confidence limits
      estimate  lcl  ucl
12.5%     0.01   NA 0.01
  25%     0.01   NA 0.02
  50%     0.02 0.01 0.02
  75%     0.03 0.02 0.05
  90%     0.05 0.03 0.09
 100%     0.30 0.30   NA
```



Because there are multiple groups, the confidence intervals are not plotted.

Step 2a. Create a dataset of selected summary statistics

The output from the *summary()* function on a mdlnpar object can be extracted and put into a data set. The following call in the **Commands Window** creates a data.frame called *TotalP.df* that contains the specified percentiles and information about the censoring levels and number of observations. The columns in the output dataset are described in table 1. Column names of the form Q.nn.pct correspond to the requested percentiles in the *probs* argument to the *summary()* function. If the minimum value is censored, then the values of selected percentiles and the mean and standard deviation will be censored; the standard deviation will be right-censored.

Table 1. Descriptions of columns in the dataset created from the summary of mdlnpar objects.

Column	Remark	Description
Group	No	Value of the grouping variable if one was selected
Vname	No	Name of the variable
Method	No	Short name for the method (K-M flipped)
Mean	Yes	The estimated mean
StdDev	Yes	The estimated standard deviation
MinObs	No	The minimum observed value
Q.nn.pct	Yes	The requested percentiles, one pair of columns for each in *probs*
Max	Yes	The maximum value of the data
MinC	No	The minimum censored value
MaxC	No	The maximum censored value
Nobs	No	The number of observations
Ncens	No	The number of censored observations
NltMaxC	No	The total number of observations less than the maximum censored value

```
TotalP.df <- as.data.frame(summary(TotalP.out, probs=c(.1,.25,.5,.75,.9)))
```

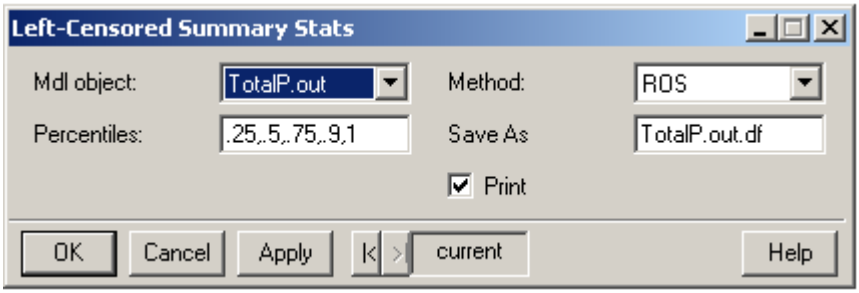
Below is a **Data Window** view of the first few columns in the *TotalP.df* dataset.

TotalP.df

	1	2	3	4	5	6	7	8	9	10	11	12
	Group	Vname	Method	Mean.rmk	Mean	StdDev.rmk	StdDev	MinObs	Q.10.pct.rmk	Q.10.pct	Q.25.pct.rmk	Q.25.pct
1	04063700	TotalP	K-M flipped	<	0.02	>	0.01	0.01	<	0.01		0.01
2	04067500	TotalP	K-M flipped	<	0.02	>	0.01	0.02	<	0.01		0.02
3	04080798	TotalP	K-M flipped	<	0.03	>	0.05	0.01	<	0.01		0.01
4												

Step 2b. Create a dataset of selected summary statistics using dialog menu

The output from either *mdlnpar()* or *mdlpar()* or the dialog menu versions can be extracted and put into a dataset. Click **USGS | Left-Censored Data Analysis | Extract summary statistics…** to open the **Left-Censored Summary Stats** dialog window.



Select TotalP.out in the drop-down box for **Mdl object:**. The default values were left as is for the **Percentiles:** box. The entry in the **Method:** box is ignored for nonparametric statistics. Accept the

dataset name TotalP.out.df in the **Save As** box. Accepting **Print** will print a version in the report window. Click **OK** to generate the dataset and report, shown below.

```
     Group  Vname      Method Mean.rmk       Mean StdDev.rmk      StdDev MinObs
1 04063700 TotalP K-M flipped        < 0.02333333          > 0.014922970   0.01
2 04067500 TotalP K-M flipped        < 0.02387097          > 0.009891889   0.02
3 04080798 TotalP K-M flipped        < 0.03218750          > 0.051727223   0.01

 Q.25.pct.rmk Q.25.pct Q.50.pct.rmk Q.50.pct Q.75.pct.rmk Q.75.pct Q.90.pct.rmk Q.90.pct
1                 0.01                 0.02                 0.03                 0.04
2                 0.02                 0.02                 0.03                 0.04
3                 0.01                 0.02                 0.03                 0.05

   Max.rmk  Max MinC MaxC Nobs Ncens NltMaxC
1          0.10 0.01 0.01   48     8       8
2          0.04 0.01 0.01   31     7       7
3          0.30 0.01 0.01   32     4       4
```

REFERENCES:

Helsel, D.R., 2005, Nondetects and data Analysis: Hoboken, New Jersey, John Wiley & Sons, 250 p.

Parametric Summary Statistics

This application illustrates how to estimate summary statistics of left-censored data assuming that the data have a log-normal distribution.

The examples in this section use the *WI1MajorIons.df* and *WI3TotalP.df* water-quality datasets. See the *Naming conventions and data storage* section (Introduction 2) for information about those datasets.

Step 1a. Estimate summary statistics using the Commands Window

Use the *mdlpar()* function to create a *mdlpar* object and print it by typing the following call in the **Commands Window**:

```
WI1Nitrate <- print(mdlpar(lcens(WI1MajorIons.df$Nitrate,
      WI1MajorIons.df$Nitrate.rmk=="<")))
```

This call produces the following output:

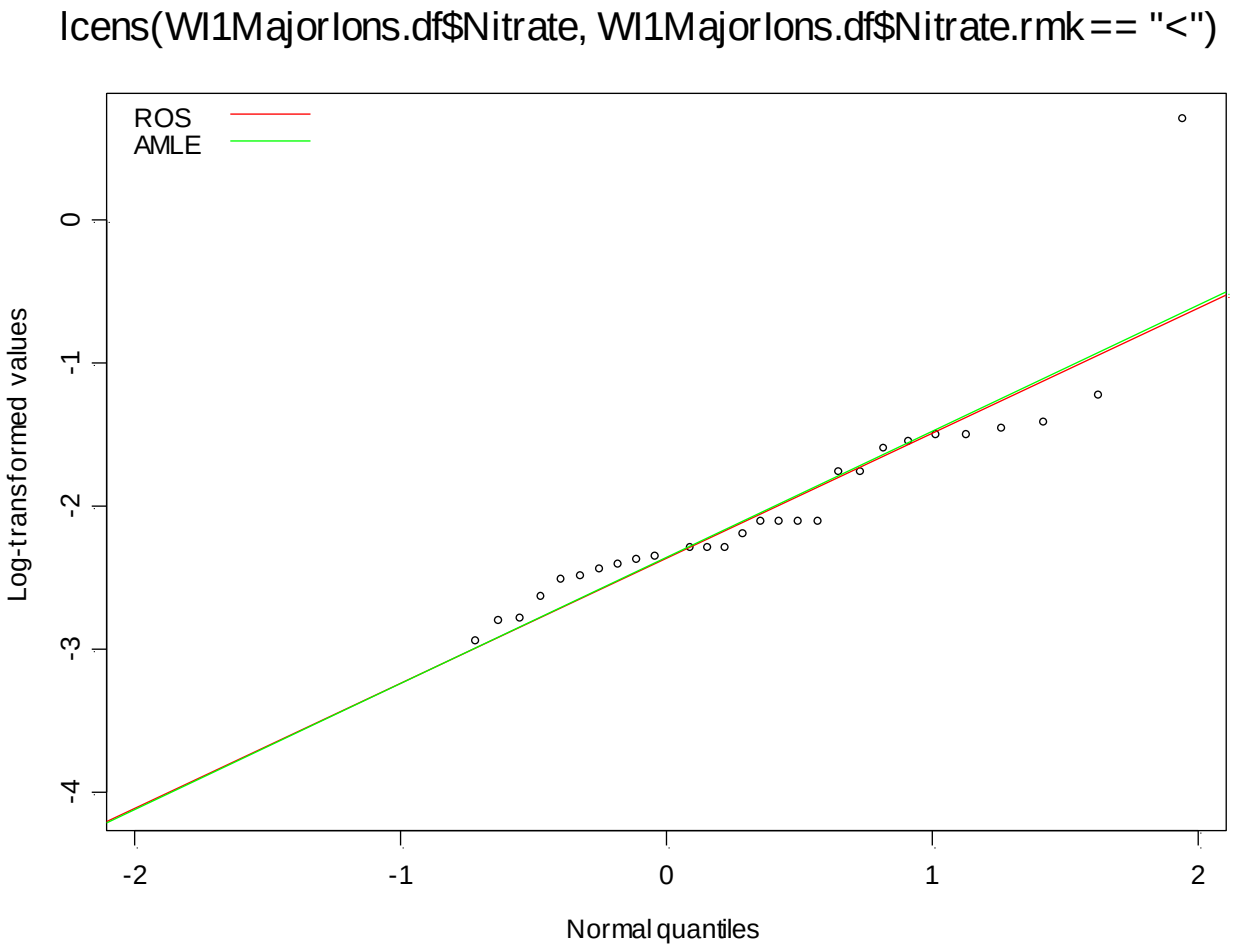
```
      Parametric Summary Statistics for lcens(WI1MajorIons.df$Nitrate,
WI1MajorIons.df$Nitrate.rmk == "<"):
                     ROS         AMLE
   Mean logs  -2.3649220  -2.35885318
    SD. logs   0.8736011   0.88173152
        Mean   0.1602585   0.13851439
          SD   0.3187312   0.14358363
Lower Qrtile   0.0600000   0.05406580
      Median   0.0940000   0.09452857
Upper Qrtile   0.1700000   0.17000000
           N  37.0000000  37.00000000
         NLT   9.0000000   9.00000000
Censoring levels and counts
 0.05 0.1
    7   2
Range of uncensored observations
   Min Max
 0.052   2
```

Summary statistics are reported for both the ROS and AMLE methods. In this example, the estimates agree very well, except for the standard deviation of the back-transformed data.

It is useful to graphically review the data. In the **Commands Window**, type the following call to produce a quantile-quantile- (QQ-) normal graph. The open circles are the uncensored values and the

dots are censored values. The two fitted distribution curves are plotted and in this case are very similar. The y-axis is in natural logarithms, which are the basis for the log-normal distribution.

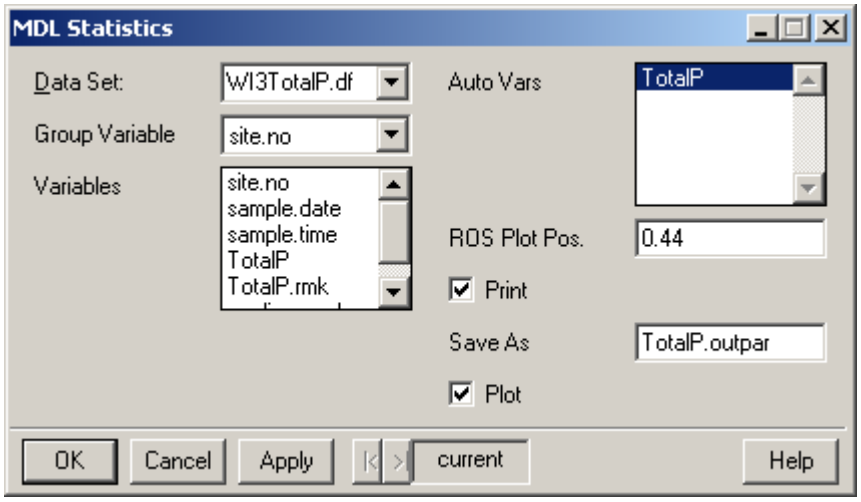
```
plot(WI1Nitrate)
```



This graph shows the reason for the discrepancy between the estimated standard deviations of the ROS and the AMLE methods. The ROS method back-transforms estimated values for each censored value and computes the mean and standard deviation from those values and therefore is more sensitive to large values than the AMLE method that uses the minimum variance unbiased estimate.

Step 1b. Estimate summary statistics using the dialog menus

Click **USGS | Left-Censored Data Analysis | Parametric summary statistics…** to open the **MDL Statistics** dialog window.

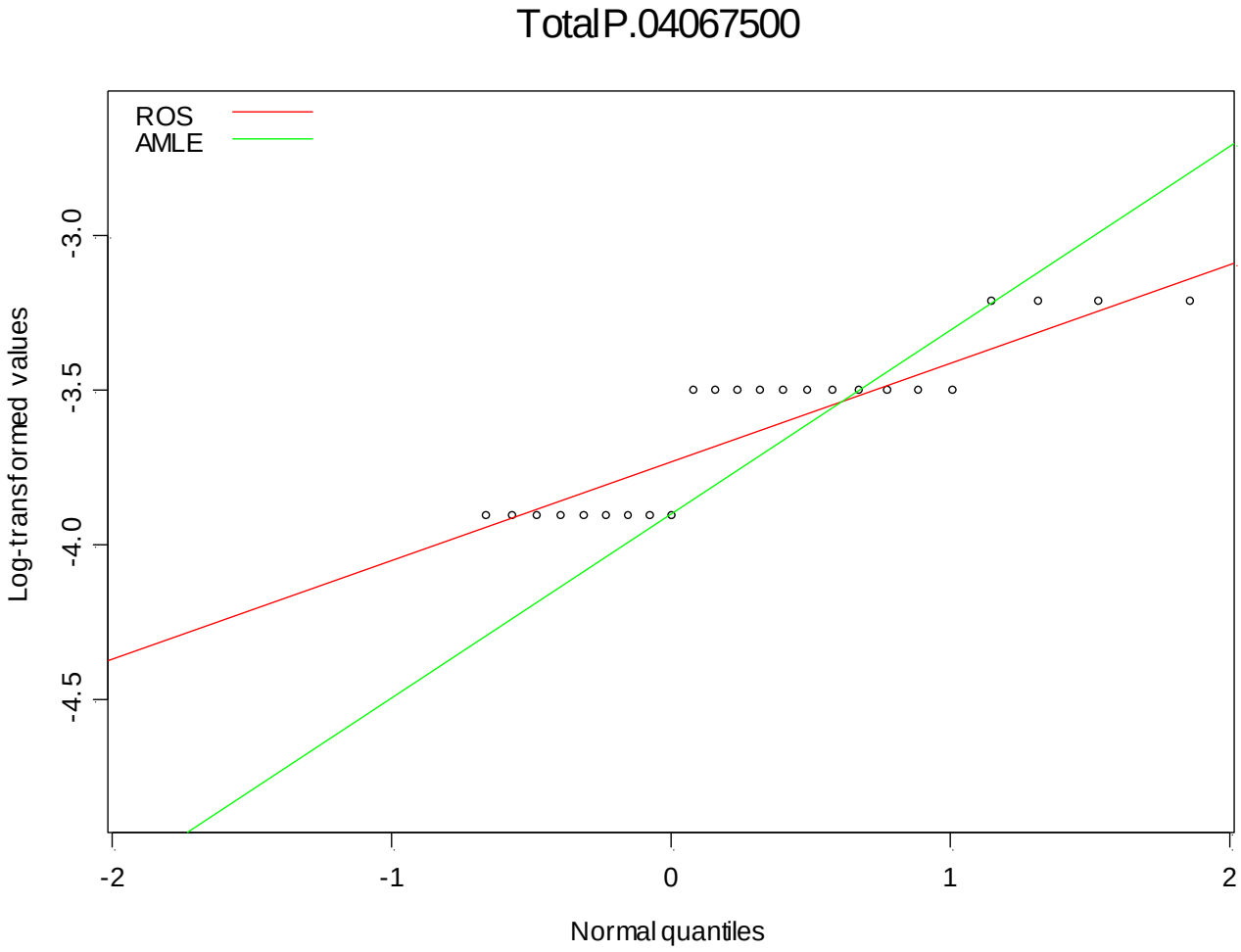


Select WI3TotalP.df in the drop-down box for **Data Set:**. Because there are three sites in this dataset, site.no should be selected in the **Group Variable** box. This dialog window recognizes the TotalP and TotalP.rmk column pairs and puts a single entry in the **Auto Vars** box. Select TotalP in that box. Others should be left as default. The **ROS Plot Pos.** is used to compute the plotting position for the ROS method. The original code used a value of 0, but 0.44 gives better results for lower censoring thresholds and small sample sizes. Enter TotalP.outpar in the **Save As** box. Click **OK** to generate the report, graphs, and output object.

```
      Parametric Summary Statistics for TotalP.04063700:
                     ROS           AMLE
Mean logs  -3.935080898  -3.960836210
 SD. logs   0.567440053   0.622322393
     Mean   0.023106575   0.023083975
       SD   0.015154888   0.015697150
      10%   0.009178758   0.008579545
      25%   0.010110481   0.010000000
      50%   0.020000000   0.020000000
      75%   0.030000000   0.030000000
      90%   0.040000000   0.040000000
        N  48.000000000  48.000000000
      NLT   8.000000000   8.000000000
Censoring levels and counts
 0.01
    8
Range of uncensored observations
  Min Max
 0.01 0.1
      Parametric Summary Statistics for TotalP.04067500:
                     ROS           AMLE
Mean logs  -3.732011450  -3.901045516
 SD. logs   0.318630106   0.594472779
     Mean   0.025206342   0.024084376
       SD   0.008175314   0.015464123
      10%   0.015529002   0.009439115
      25%   0.020000000   0.020000000
```

```
      50%   0.020000000   0.020000000
      75%   0.030000000   0.030000000
      90%   0.040000000   0.040000000
        N  31.000000000  31.000000000
      NLT   7.000000000   7.000000000
Censoring levels and counts
 0.01
    7
Range of uncensored observations
  Min  Max
 0.02 0.04
      Parametric Summary Statistics for TotalP.04080798:
                     ROS          AMLE
Mean logs  -3.951695425  -3.915716599
 SD. logs   0.880656838   0.838806615
     Mean   0.031542119   0.028154950
       SD   0.052043213   0.027250939
      10%   0.005789618   0.006800967
      25%   0.010000000   0.010000000
      50%   0.020000000   0.020000000
      75%   0.030000000   0.030000000
      90%   0.050000000   0.050000000
        N  32.000000000  32.000000000
      NLT   4.000000000   4.000000000
Censoring levels and counts
 0.01
    4
Range of uncensored observations
  Min Max
 0.01 0.3
```

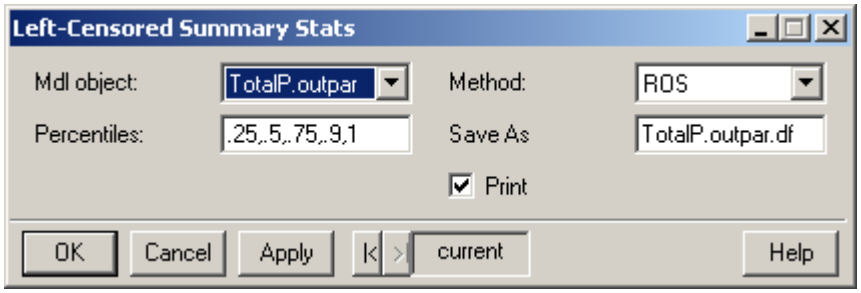
Only the second graph is included here. All of the graphs show the effects of rounding.



This graph highlights the assumption of type II censoring in ROS and shows an important reason to look at the graphs. As indicated by the red line, the estimated value for each of the censored data is greater than the detection limit for ROS, but the AMLE method uses the censoring level to estimate the mean and standard deviation.

Step 2. Create a dataset of some summary statistics using dialog menu.

The output from either *mdlnpar()* or *mdlpar()* or the dialog menu versions can be extracted and put into a dataset. Click **USGS | Left-Censored Data Analysis | Extract summary statistics…** to open the **Left-Censored Summary Stats** dialog window.



Select TotalP.outpar in the drop-down box for **Mdl object:**. The default values were left as is for the **Percentiles:** box. The entry in the **Method:** box is used for parametric statistics; either ROS or AMLE can be selected and the default is ROS. Accept the dataset name TotalP.outpar.df in the **Save As** box. Accepting **Print** will print a version in the report window. Click OK to generate the dataset and report, shown below.

```
     Group  Vname Method Mean.rmk       Mean StdDev.rmk      StdDev MinObs
1 04063700 TotalP    ROS          0.02310658            0.015154888   0.01
2 04067500 TotalP    ROS          0.02520634            0.008175314   0.02
3 04080798 TotalP    ROS          0.03154212            0.052043213   0.01

  Q.25.pct.rmk   Q.25.pct Q.50.pct.rmk Q.50.pct Q.75.pct.rmk Q.75.pct
1              0.01022096                 0.02                 0.03
2              0.02000000                 0.02                 0.03
3              0.01000000                 0.02                 0.03

   Q.90.pct.rmk Q.90.pct Max.rmk  Max MinC MaxC Nobs Ncens NltMaxC
1                  0.04         0.10 0.01 0.01   48     8       8
2                  0.04         0.04 0.01 0.01   31     7       7
3                  0.05         0.30 0.01 0.01   32     4       4
```