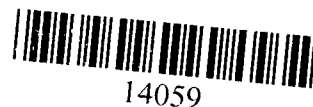


| | |
|------------------|--|
| 8 2 9 2 | ERID#: <div style="font-size: 2em; font-weight: bold; margin-left: 100px;">82912</div> <div style="text-align: right; font-weight: bold;"> LOS ALAMOS NATIONAL LABORATORY ENVIRONMENTAL RESTORATION (RRES-R) <i>Records Processing Facility</i> <i>ER Records Index Form</i> </div> |
|------------------|--|

Date Received: 3/1/2004 **Processor:** DSV **Page Count:** 8
Privileged: (Y/N) N **Record Category:** P **Administrative Record:** (Y/N) Y
FileFolder: N/A
Miscellaneous Comments: REFERENCED DOC

Record Documents:

| <i>Start Pg</i> | <i>Doc Type</i> | <i>Doc Date</i> | <i>Title</i> | <i>Box</i> | <i>Package</i> |
|-----------------|-----------------|-----------------|--|------------|----------------|
| 1 | Misc | 12/1/1988 | Estimation of Descriptive Statistics for Multiply Censored Water Quality Data, Water Resources Research, Vol. 24, No. 12, pp 1997-2004, December 1988 N/A N/A N/A | | |



82912

Estimation of Descriptive Statistics for Multiply Censored Water Quality Data

DENNIS R. HELSEL and TIMOTHY A. COHN

Water Resources Division, U.S. Geological Survey, Reston, Virginia

This paper extends the work of Gilliom and Helsel (1986) on procedures for estimating descriptive statistics of water quality data that contain "less than" observations. Previously, procedures were evaluated when only one detection limit was present. Here we investigate the performance of estimators for data that have multiple detection limits. Probability plotting and maximum likelihood methods perform substantially better than simple substitution procedures now commonly in use. Therefore simple substitution procedures (e.g., substitution of the detection limit) should be avoided. Probability plotting methods are more robust than maximum likelihood methods to misspecification of the parent distribution and their use should be encouraged in the typical situation where the parent distribution is unknown. When utilized correctly, less than values frequently contain nearly as much information for estimating population moments and quantiles as would the same observations had the detection limit been below them.

DECEMBER
MAR 9 1 2004
BY

INTRODUCTION

Hydrologists frequently encounter water quality data that contain sample concentrations reported only as "less than the analytical detection limit." These "less thans" complicate analysis of the data. Sample moments cannot be calculated directly, and even the sample median does not exist if the number of less thans exceeds 50% of the observations. Such data are called "censored data" in the statistical literature.

This paper explores procedures for estimating descriptive statistics from censored data having multiple detection limits. Multiple detection limits arise because of (1) improvement in analytical methods over time, resulting in a lowering of the detection limits, (2) management decisions to reduce costs by use of methods with higher detection limits, (3) combination of data from several agencies or laboratories having different reporting levels, or (4) use of differing laboratory procedures and detection limits due to differences in sample matrix characteristics. Following the approach of Gilliom and Helsel [1986], the efficiency and robustness of estimators for the sample moments and quantiles are tested, by Monte Carlo simulation, for a number of hydrologically plausible parent distributions.

LITERATURE REVIEW

Several procedures for estimating descriptive statistics of censored data have been recommended and used in air quality [Nehls and Akland, 1973], radionuclide [Gilbert and Kinnison, 1981], plant and sediment geochemistry [Shacklette and Boerngen, 1984], and groundwater quality [McBean and Govers, 1984] applications. Simple substitution methods, such as replacing all less thans with zero or the detection limit, are most commonly used. The relative performance of these procedures, however, has rarely been evaluated.

Gleit [1985] compared the performance of several methods for estimation of a mean under the assumption that data follow a normal distribution. Simple substitution methods performed poorly. Maximum likelihood produced estimates having large bias and large variance for the small ($n = 5, 10$, and 15) sample sizes considered. The method with least error was a fill-in technique, where less than data were replaced

with expected values for their order statistics. Gilliom and Helsel [1986] compared procedures for calculating the mean, standard deviation, median, and interquartile range of censored data sets with a single detection threshold. They employed several parent distributions, some highly skewed, to mimic the observed distributions of water quality data. This allowed evaluation of the robustness of each method to misspecification of the population distribution, avoiding an assumption that all water quality data follow a single parent distribution such as the lognormal. Simple substitution methods were found to produce biased and highly variable estimates. Two methods, maximum likelihood estimation and a probability plotting procedure, produced the lowest errors of estimation; that is, the sample descriptive statistics tended to be closest to the known population values for conditions believed to be typical of water quality data. Maximum likelihood estimates contained lowest errors for quantiles, but had considerable bias when estimating moment statistics. The probability plotting procedure contained slightly more error for quantile estimates, but much less error than maximum likelihood for estimation of moments. Finally, they determined that there was little advantage in "optimizing" the estimation method based on population shape. Uncertainties in selecting parent population shapes from small sample data sets argued for robust, general purpose estimation methods which performed well in most, or all, situations.

Despite the frequent occurrence of multiple-detection limits in water quality studies, procedures for analysis of such data have received little or no attention in the hydrologic literature.

The statistical aspects of censored data techniques are discussed by David [1981] and Cohn [1988]. Kalbfleisch and Prentice [1980] and Latta [1981] survey hypothesis test procedures appropriate for discerning differences between populations of censored data. For data analyzed in a chemical laboratory, a detection threshold fixed in value is reported. The number of observations below that threshold is a random variable, and therefore censoring is always type I. Alternate censoring mechanisms found in the statistical literature do not occur in this situation. Such mechanisms would include type II censoring, where the number of censored observations is fixed and the value of the threshold is varied, and random censoring, where the value of the threshold is generated by a random process.

This paper is not subject to U.S. copyright. Published in 1988 by the American Geophysical Union.

Paper number 88WR03364.

Statisticians generally use the method of maximum likelihood when analyzing type I censored data [David, 1981]. This involves defining a likelihood function (LF), having one factor for the n_1 above-threshold observations, $\{x_1 \dots x_{n_1}\}$ and a second factor for the n_2 observations that were censored at thresholds $\{T_1 \dots T_{n_2}\}$

$$L(\Theta) = \prod_{i=1}^{n_1} f_{\Theta}(x_i) \prod_{j=1}^{n_2} F_{\Theta}(T_j)$$

Here Θ is a vector of parameter values, f is the probability density function of x , and F is the cumulative density function of x . The estimates are then found by differentiating with respect to the parameters, setting the first derivatives to zero, and solving. It is always assumed, however, that the parent distribution is known exactly.

Finney [1941] has shown that the maximum likelihood estimator for the moments of lognormal populations tends to be highly biased for small samples. He derived the amount of bias, completely solving this problem, for the case where observations are not subject to censoring. Cohn [1988] adapted Finney's estimator to obtain a nearly unbiased estimator for the population moments in cases where censoring occurs, and the parent distribution is known exactly.

METHODS

The eight methods to be evaluated in this study are listed Appendix A. Two simple substitution methods (ZE and DL) used by Gilliom and Helsel [1986] will be evaluated for the multiple threshold case. A third substitution method not previously evaluated, setting all less than to a value of one half their detection limit (HA), will also be investigated.

Two single-threshold methods, a "fill-in" probability plotting procedure (LR) and a lognormal maximum likelihood procedure (LM), will be used in the multiple threshold case as follows: all points below the highest of the censoring thresholds will be treated as though they were less than at the highest threshold. Thus with two detection limits at 1 and 10 units, all uncensored observations between 1 and 10, as well as censored observations reported as "<1," would be considered to be "<10." The single-threshold estimation method is then applied. Extensions of these two methods to incorporate multiple detection limits (MR and MM) will also be evaluated. Plotting positions for the MR method are those of Hirsch and Stedinger [1986] and are presented in detail in Appendix B. Cohen [1976] discussed the multiple-threshold maximum likelihood method.

Finally, the adjusted maximum likelihood procedure (AM) of Cohn [1988] will be evaluated. This method makes a first-order correction in the bias of maximum likelihood estimates of population moments. It is described in more detail in Appendix C.

Note that for the first five procedures in Appendix A the estimates $\bar{\mu}$ and $\bar{\sigma}^2$ are the standard moment estimators

$$\bar{\mu} = \bar{X} = \sum_{i=1}^n \frac{X_i}{n} \quad \bar{\sigma}^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}$$

after filling in values for the below-threshold observations. These estimators are therefore closest to the familiar descriptive statistics and require no parametric assumptions except for determining fill-in values for the below-threshold observations. They also avoid transformation bias by computing statistics directly from the original data. The last three pro-

cedures, however, estimate the mean and standard deviation under the assumption that the water quality variates are lognormally distributed. This assumption may be appropriate in some circumstances. If it were exactly true, likelihood-based methods would be the most efficient estimators. However, the validity of the lognormal assumption can rarely be verified in practice. In addition, small sample estimates of mean and standard deviation in log units are used in the retransformation of estimates to the original units. This can introduce considerable bias into estimates of moment statistics for the original units [Kendall and Stuart, 1979, p. 74]. Here we compare the performance of the last three estimators which assume lognormality, to the five which do not, for estimating statistics of parent distributions which are not always lognormal. The five following questions were considered.

1. How much information is lost when single-threshold methods are used for data having multiple detection limits?
2. How do estimates by multiple-threshold methods compare to descriptive statistics from the uncensored sample of the same size?
3. Which multiple-threshold methods perform well (i.e., have low rmse and bias) given the range of distributions specified?
4. How important is the choice of specific plotting position formula for those methods that require them?
5. How well does each method perform when the distribution of data departs markedly from the lognormal?

Monte Carlo Experiment

Much research has focused on the performance of estimators designed for censored data when the sampling distribution is known. Unfortunately, this is seldom, if ever, the case with water quality data. To understand the behavior of the different estimators under practical conditions, we employ four special cases of four possible parent distributions to model water quality data. Details on the characteristics of these 16 distributions (Figure 1) are given by Gilliom and Helsel [1986]. Parameters for the distributions, as well as the distributional shapes, were chosen in order to mimic the observed variation, skewness, and tail behavior of surface water quality data [Gilliom and Helsel, 1986]. Included are contaminated lognormal distributions (mixtures of two lognormal distributions) and delta distributions (lognormal distributions plus a percentage of true zero values). The mean of each distribution was set to 1.0.

Each method is tested on samples generated from 15 of these distributions, which we call the "Monte Carlo (MC) population." We then deal separately with the 16th distribution (gamma with coefficient of variation (CV) = 2.0, so scale parameter = 4 and shape parameter = $1/CV^2 = 0.25$). This distribution differs markedly from a lognormal, and yet is quite plausible for the highly censored data sets typical of such constituents as trace organic compounds. By using multiple parent distributions, it is possible to investigate the performance of the eight estimation methods under conditions which might arise in practice.

Design of Monte Carlo Experiment

Five hundred repetitions ($N = 500$) of sample size $n = 25$ were generated from each of the 16 distributions. Three percentiles of each distribution were chosen as "detection limit" thresholds, and one third of each data set was randomly assigned to each threshold. Any value falling below its assigned threshold was censored. Several combinations of thresholds

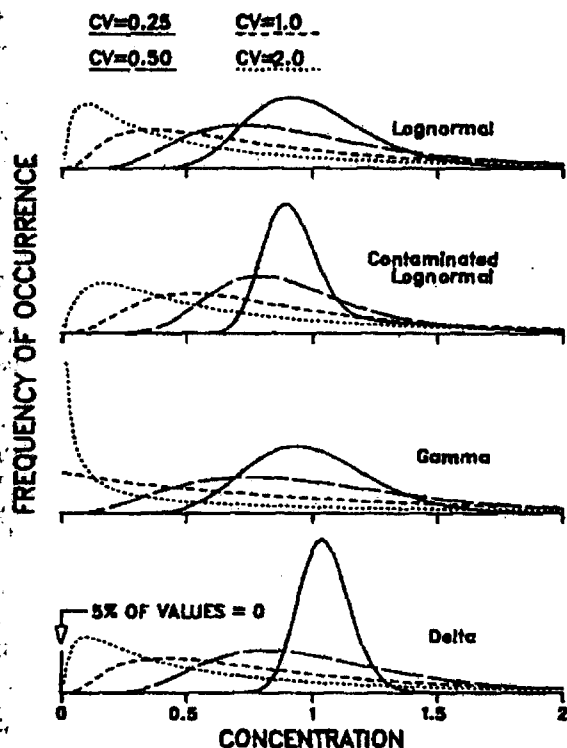


Fig. 1. Probability density functions for the 16 parent distributions.

are evaluated; all results reported here are for thresholds at the 20th, 40th, and 80th percentiles. Though the magnitudes of the root-mean-square error (rmse) change with differing censoring thresholds, relative method performance did not.

Criteria for Comparison

Estimates of the four descriptive statistics produced by each method were evaluated using their relative rmse (shown here for the mean) and bias:

$$rmse_k = \left[\sum_{j=1}^N \left(\frac{\hat{\mu}_{jk} - \mu}{\mu} \right)^2 / N \right]^{1/2}$$

$$bias_k = \left[\sum_{j=1}^N \left(\frac{\hat{\mu}_{jk} - \mu}{\mu} \right) / N \right]$$

TABLE 1. The rmse of Estimation Methods for the MC Population

| Method | Mean | Standard Deviation | Median | IQR |
|--------------|------|--------------------|--------|------|
| No censoring | 0.19 | 0.38 | 0.20 | 0.29 |
| ZE | 0.36 | 0.82 | 0.63 | 2.41 |
| HA | 0.22 | 0.42 | 0.25 | 0.98 |
| DL | 0.28 | 0.43 | 0.71 | 0.21 |
| LR | 0.27 | 0.48 | 0.47 | 0.45 |
| MR | 0.20 | 0.42 | 0.20 | 0.31 |
| LM | 0.26 | 4.31 | 0.33 | 0.40 |
| MM | 0.21 | 0.68 | 0.18 | 0.29 |
| AM | 0.19 | 0.37 | 0.18 | 0.29 |

where k refers to the method, $\hat{\mu}_{jk}$ is the statistic estimated by the k th method for the j th of N replicates, and μ is the true population value. The rmse and bias are "relative" in that they are divided by their true population value, in order to compile information from distributions with different characteristics, particularly CV. For the MC population, $N = 7500$ replicates, 500 from each of the 15 distributions.

RESULTS

Single Versus Multiple-Threshold Methods

Figure 2 provides a comparison of the single threshold methods (LM and LR) to their respective multiple-threshold extensions (MM and MR). Each method's rmse has been standardized by dividing by the rmse for the uncensored sample estimate. Unstandardized rmse and bias are given in Tables 1 and 2. A substantial improvement in estimation ability (as shown by lower rmse) occurs when the multiple-threshold extensions are used. These methods incorporate the additional information available in the data values falling between detection thresholds and the relative magnitudes of the thresholds.

Comparisons Among Multiple-Threshold Methods

In addition to the MR and MM methods, the adjusted maximum likelihood procedure of Cohn [1988] was evaluated. This method, designated as AM, is identical to MM for quantile estimates, but makes a first-order correction in the bias of MM when estimating population moments. Higher rmse for (uncorrected) MM estimates of the mean and standard deviation

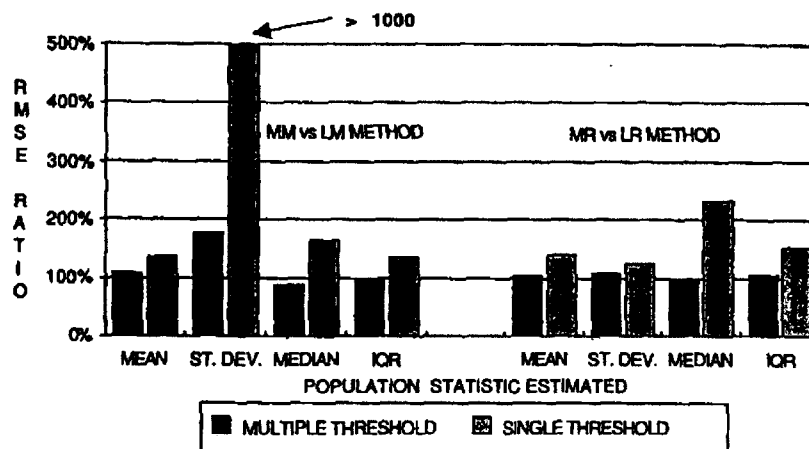


Fig. 2. Comparison of methods rmse/uncensored rmse in percent for single and multiple threshold methods (7500 replicates from the MC population).

TABLE 2. Bias of Estimation Methods for the MC Population

| Method | Mean | Standard Deviation | Median | IQR |
|--------------|-------|--------------------|--------|-------|
| No censoring | 0.01 | -0.08 | 0.02 | -0.02 |
| ZE | -0.27 | 0.46 | -0.43 | 1.53 |
| HA | -0.03 | 0.03 | -0.01 | 0.33 |
| DL | 0.20 | -0.23 | 0.52 | -0.05 |
| LR | 0.05 | -0.13 | 0.07 | 0.03 |
| MR | 0.02 | -0.14 | 0.01 | -0.03 |
| LM | 0.06 | 0.15 | 0.07 | 0.01 |
| MM | 0.04 | 0.02 | 0.00 | 0.04 |
| AM | 0.02 | -0.09 | 0.00 | 0.04 |

ation, relative to AM, are therefore evident in Figure 3 and Tables 1 and 2.

The rmse for all six multiple-threshold methods are illustrated in Figure 3. The three methods already discussed (AM, MR, and MM) can be compared to the three commonly used simple substitution methods ZE, HA, and DL. The MR and AM methods produce similar errors, and generally have the lowest rmse among the methods employed. Although the adjusted likelihood method (AM) gives slightly lower errors than the plotting position method (MR) for the MC population, it is shown later to be less robust to severe departures from lognormality than is MR.

The simple substitution methods always have higher rmse than MR or AM, with only one exception (the negatively biased detection limit (DL) estimate of the interquartile range). Thus there is no justification based on estimation error for use of simple substitution methods to estimate descriptive statistics; better methods are always available.

It was noted by one reviewer that the HA method appears to perform adequately for moment statistics and may be preferred for those due to ease of computation. So far we have considered the performance of each estimator for the MC population. One could also ask: For what distribution of water quality data does a given estimator make sense? The ZE, HA, and DL methods would be appropriate choices, of course, if one believed that all below-threshold observations of water quality data truly had values only at 0 (ZE), half the detection limit (HA), or at the detection limit, respectively. This is very unlikely. The HA and DL methods are certainly not representative of conditions typical of water quality data. It is conceivable that censored values might truly be zero for

pesticides or other man-made organic compounds, but one never knows. As analytical methods have improved, and detection limits decreased, man-made organics have been detected at concentrations below previous detection limits. Thus an assumption that all below detection limit values for man-made organics are zero has not been supported by past data. Given that simple substitution methods are unlikely to represent the true populations of water quality data and that methods with lower estimation errors are available, use of these methods (other than perhaps HA as a "quick and dirty" moment estimator) appear unwarranted.

The LR and MR methods assume that below-threshold data arise from the lower portion of a lognormal distribution. This avoids assuming that all below-threshold values are identical and agrees with the right-skewed, bounded-at-zero shape reported for water quality data by Gilliom and Helsel [1986]. These methods also avoid making any assumptions concerning the distribution of the entire data set and thus are less sensitive to incorrectly assuming all data are lognormal in shape.

The parametric procedures (LM, MM, and AM) assume that water quality data are lognormally distributed over the complete range of values. They would perform better than the semiparametric methods (LR and MR) if this assumption were valid. However, these parametric procedures are no longer optimal when the population is not lognormal. We believe that the 16 distributions chosen, including the $CV = 2$ gamma distribution, approximately characterize conditions found in actual water quality data. As the parent distribution is rarely known in practice, the robustness of estimators for descriptive statistics over the range of distributions likely to be encountered is of great importance.

Comparisons Between Multiply Censored and Uncensored Estimators

One measure of the quality of an estimator for censored data is how well it compares to the sample estimator had there been no censoring. Sample estimates of the four statistics were computed prior to censoring, and rmse and bias were calculated. These reflect error due only to differences between the $n = 25$ sample and the true population value, and are listed in Tables 1 and 2 and 4 and 5 as the "no censoring" entries. The rmse for the censored data were divided by the rmse for no censoring. These ratios have been presented for

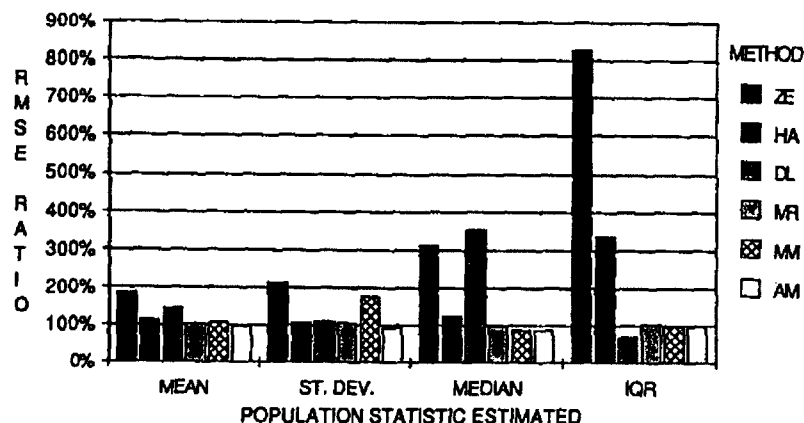


Fig. 3. Comparison of method rmse/uncensored rmse in percent for six multiple-threshold methods (7500 replicates from the MC population).

TABLE 3. Comparison Among Plotting Positions Using the LR Method for the MC Population

| Plotting Positions | Mean | Standard Deviation | Median | IQR |
|--------------------|------|--------------------|--------|------|
| <i>rmse</i> | | | | |
| Weibull | 0.16 | 0.38 | 0.19 | 0.35 |
| Blom | 0.16 | 0.38 | 0.18 | 0.33 |
| Hazen | 0.15 | 0.39 | 0.18 | 0.29 |
| <i>Bias</i> | | | | |
| Weibull | 0.00 | -0.09 | -0.02 | 0.08 |
| Blom | 0.01 | -0.10 | 0.00 | 0.04 |
| Hazen | 0.01 | -0.11 | 0.01 | 0.03 |

One detection limit at 60th percentile.

the six multiple-threshold methods in Figure 3. Errors for the simple substitution methods ZE, HA, and DL are generally greater than 100% of the no censoring rmse and thus represent a significant loss in information. On the other hand, with the MR and AM methods, the sampling error dominates the estimation error due to censoring. Thus with these two efficient estimation methods, the presence of censored data having multiple detection limits adds very little additional error to estimates of summary statistics for a given sample size.

Importance of Plotting Position Formula

The general formula for plotting positions, the probability of being less than or equal to the i th smallest observation, given a complete sample of size n is

$$PP_i = \frac{i - \alpha}{n + 1 - 2\alpha}$$

where α is used to eliminate bias in the largest (and sometimes smallest) observations. There has been a great deal of discussion about the theoretical basis for selecting α [Cunnane, 1978]. We conducted a modest experiment to determine the importance of the choice of α used in the LR and MR methods. The rmse and bias of the LR method (with one threshold at the 60th percentile) was computed using the Weibull ($\alpha = 0.00$), Blom ($\alpha = 0.375$), and Hazen ($\alpha = 0.50$) plotting positions. Differences in rmse and bias among the three plot-

TABLE 4. The rmse of Estimation Methods

| Method | Mean | Standard Deviation | Median | IQR |
|--------------|-------|--------------------|--------|------|
| No censoring | 0.29 | 0.33 | 2.36 | 0.42 |
| ZE | 0.36 | 0.40 | 1.35 | 0.57 |
| HA | 0.4 | 0.40 | 5.94 | 0.45 |
| DL | 0.51 | 0.39 | 10.33 | 0.36 |
| LR | 0.56 | 0.43 | 7.7 | 0.35 |
| MR | 0.37 | 0.40 | 1.99 | 0.54 |
| LM | 0.66 | 8.55 | 5.52 | 0.31 |
| MM | 45.25 | ... | 1.26 | 0.45 |
| AM | 3.04 | 15.05 | 1.26 | 0.45 |

Gamma with CV=2.0. Three center dots indicate a value greater than or equal to 100.0.

ting positions using simulations of 7500 repetitions from the MC population rarely exceeded 2% (Table 3), which is within the error of the simulation process. We conclude that choice of α is not important when using the LR or MR methods, particularly in comparison to other sources of error. The Weibull position is used in all work reported here.

Method Performance Under Severe Departure From Lognormality

A final and crucial consideration when evaluating method performance is the robustness of the method to misspecification of the distributional family. Both the AM and MM methods assume that a lognormal distribution is appropriate over the complete range of data, while the MR method employs that assumption only for estimating the below-threshold values. Since method performance was previously evaluated over a range of distributions, 11 of which were not lognormal, a test of robustness has already been incorporated into the results. However, the distribution least resembling a lognormal, the gamma distribution with a coefficient of variation of 2.0, appears to validly represent the distribution of some constituents where zero and small values are frequently present [Gilliom and Helsel, 1986]. Method performance for this distribution is given in Figure 4 and Tables 4 and 5.

As evident in the figure, when the underlying population is severely misspecified the maximum likelihood methods (MM and AM) perform poorly for estimation of moments in the

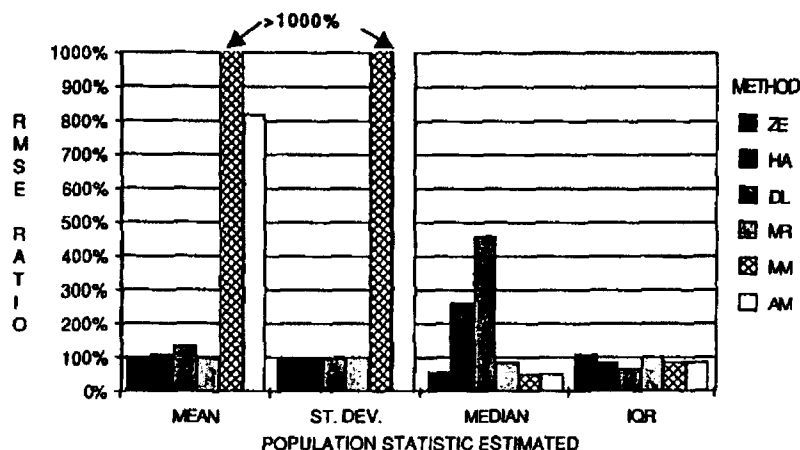


Fig. 4. Comparison of method rmse/uncensored rmse in percent for six multiple-threshold methods ($N = 500$, gamma CV = 2 distribution).

TABLE 5. Bias of Estimation Methods

| Method | Mean | Standard Deviation | Median | IQR |
|--------------|------|--------------------|--------|-------|
| No censoring | 0.02 | -0.04 | 1.19 | -0.01 |
| ZE | 0.0 | -0.04 | -0.11 | -0.09 |
| HA | 0.19 | -0.07 | 5.71 | -0.08 |
| DL | 0.37 | -0.08 | 9.4 | 0.27 |
| LR | 0.36 | -0.11 | 6.32 | -0.03 |
| MR | 0.05 | -0.05 | 1.22 | -0.11 |
| LM | 0.41 | 2.11 | 4.85 | -0.09 |
| MM | 7.7 | ... | 0.62 | -0.32 |
| AM | 1.9 | 8.20 | 0.62 | -0.32 |

Gamma with CV=2.0. Three center dots indicate a value greater than or equal to 100.0.

original units. These methods are not robust to departures from lognormality when estimating moments. They provide efficient estimates of quantiles, however. The rmse for MM and AM are lower than for the uncensored sample estimates (Figure 4), similar to the gains in efficiency for maximum likelihood over method of moments estimates of quantiles of the lognormal distribution [Aitchison and Brown, 1957, p. 40]. In contrast, the MR method offers protection against model misspecification when estimating moments, at a relatively low cost in estimation error for quantiles (Figure 5).

Figure 4 illustrates that several methods have lower rmse than does the uncensored sample estimate of the median. This is due to the large positive bias in the uncensored sample median (Table 5), which occurs because of the asymmetry of the confidence interval around the median for the quite asymmetric parent population.

The decision of whether to employ AM or MR depends on the likelihood of data lognormality and on the robustness desired. If a single estimation method for all four descriptive statistics is desired and the possibility of a nonlognormal parent such as the gamma CV = 2 distribution cannot be discounted, MR should be chosen. It is a robust and efficient (near-minimum rmse) method for estimation of moment statistics. Gains in efficiency (lower rmse) are available by using a second method (AM) when estimating the median and interquartile range (Figure 5). Use of AM for all four descriptive statistics is appropriate only when lognormality is not an unreasonable assumption, such as for parent distributions similar to the MC population.

TABLE 6. Estimates Produced by Six Multiple-Threshold Methods for the Silver Data

| | Mean | Standard Deviation | Median | IQR |
|----|-------|--------------------|--------|------|
| ZE | 12.36 | 75.48 | 0.00 | 1.10 |
| HA | 13.91 | 75.28 | 1.10 | 3.30 |
| DL | 15.45 | 75.19 | 1.30 | 4.10 |
| MR | 12.57 | 75.44 | 0.29 | 1.54 |
| AM | 7.51 | 33.30 | 0.34 | 1.62 |
| MM | 8.30 | 61.52 | 0.34 | 1.62 |

WORKED EXAMPLE

Silver concentrations in standard solutions were reported by several laboratories in an interlab comparison [Janzer, 1986]. The 56 analyses included 36 values below one of 12 detection limits. One large outlier (a "far outside" value on a boxplot) of 560 $\mu\text{g/L}$ was also reported. The data are presented below:

| | | | | | | | |
|------|-------|-------|-------|-------|-------|-------|------|
| 0.8 | <25.0 | <5.0 | <0.2 | <0.5 | 5.0 | <0.3 | <0.2 |
| 0.1 | 2.7 | <0.1 | <20.0 | 1.4 | 2.0 | <2.5 | 2.0 |
| 2.0 | <1.0 | <10.0 | <1.0 | <0.2 | 1.0 | <10.0 | <0.2 |
| 0.2 | 1.2 | <1.0 | 1.0 | <6.0 | <1.0 | 0.7 | <1.0 |
| <5.0 | 3.2 | 2.0 | 10.0 | 1.0 | 4.4 | <1.0 | <1.0 |
| <1.0 | <20.0 | <5.0 | <10.0 | <10.0 | 90.0 | 1.5 | <1.0 |
| <2.0 | <10.0 | 560.0 | <5.0 | 0.1 | <20.0 | <1.0 | <0.1 |

Estimates of the four descriptive statistics for each of the six multiple-threshold methods are listed in Table 6.

Because of the outlier at 560 $\mu\text{g/L}$, the shape of the data tend to resemble the gamma CV-2 distribution more so than a lognormal. It is not surprising therefore that the maximum likelihood methods give moment estimates dissimilar to the other methods. We generally select the MR moment estimates and the AM quantile estimates, due to the rmse results presented above.

CONCLUSIONS

1. Estimates of descriptive statistics for multiply censored data may contain no more error than those for uncensored data, if efficient estimation methods are used.
2. The plotting position and adjusted maximum likelihood procedures are found to be substantially better than any of the simple substitution methods.

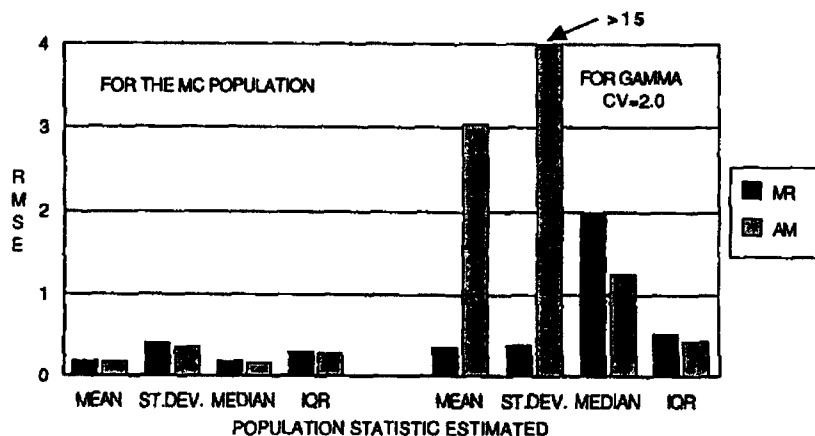


Fig. 5. The rmse of the MR and AM estimators.

3. The particular value of α used in plotting positions (e.g., Hazen versus Blom versus Weibull) for the LR and MR methods is of negligible importance.

4. Lognormal maximum likelihood estimators may have substantial positive bias when estimating sample moments from small or moderate sized samples, even for data from truly lognormal parent distributions. However, one can reduce this bias using the adjusted maximum likelihood estimator of Cohn [1988].

5. When the parent distribution (e.g., lognormal) of the data is not known, the desirable, theoretical sampling properties of likelihood-based procedures do not necessarily apply. As the parent distributions of water quality data are rarely known, the decision of whether to employ AM or MR depends on the likelihood of lognormality, and on the robustness desired. If a single estimation method must be used for all four descriptive statistics and the possibility of a nonlognormal parent such as the gamma $CV = 2$ distribution cannot be discounted, MR should be chosen. It is a robust and efficient (near-minimum rmse) method. Additional gains in efficiency (lower rmse) are obtainable by using a maximum likelihood method (MM, which is identical to AM for quantiles) when estimating the median and interquartile range. Use of AM for all four descriptive statistics is appropriate only when lognormality is not an unreasonable assumption, such as for parent distributions similar to the MC population.

APPENDIX A: ESTIMATION METHODS USED

Simple Substitution Methods

For ZE, censored observations were assumed to equal zero. For HA, censored observations were assumed to equal one half the detection limit. For DL, censored observations were assumed to equal the detection limit.

Probability Plotting Methods

For LR, values for the censored observations are derived from a linear regression model of the logarithms of concentrations versus their normal quantiles [see Gilliom and Helsel, 1986]. This probability plotting procedure is appropriate only for a single censoring threshold. For MR, an extension of the LR method to multiple thresholds, utilizing plotting positions from Appendix C of Hirsch and Stedinger [1987].

Maximum Likelihood Methods

For LM, maximum likelihood assuming that concentrations obey a lognormal distribution with censoring at a single threshold [Cohen, 1959]. All observations below the highest threshold are treated as censored observations. For MM, maximum likelihood with multiple thresholds [Cohen, 1976]. For AM, adjusted maximum likelihood with multiple thresholds [Cohn, 1988].

APPENDIX B: PLOTTING POSITIONS FOR THE MR METHOD

Consider the following data set of $n = 18$ observations and $m = 2$ detection limits: $<1 <1 <1 <1 <1 <1 <10 <10 <10 <10 <10 <10 <10 <10 <10 <10 <10$. Hirsch and Stedinger [1987] define a variable A_j as the number of uncensored observations above the j th threshold (here the j th detection limit) and below the next highest threshold. There are six uncensored observations above the higher detection limit of 10 ($A_2 = 6$), and three

TABLE 7. Plotting Positions for the Appendix B Data

| Recorded Value | Plotting Position | Recorded Value | Plotting Position |
|----------------|-------------------|----------------|-------------------|
| <1 | 0.063 | 3 | 0.500 |
| <1 | 0.127 | 7 | 0.556 |
| <1 | 0.190 | 9 | 0.611 |
| <1 | 0.254 | | |
| <1 | 0.317 | 12 | 0.714 |
| <1 | 0.381 | 15 | 0.762 |
| | | 20 | 0.810 |
| <10 | 0.167 | 27 | 0.857 |
| <10 | 0.333 | 33 | 0.905 |
| <10 | 0.500 | 50 | 0.952 |

uncensored observations between the lower and higher detection limits ($A_1 = 3$). They also define B_j as the number of observations, censored and uncensored, below the j th threshold. There are six less thans below the lower detection limit ($B_1 = 6$), and nine less thans plus three uncensored values below the higher detection limit ($B_2 = 12$). The conditional probability of exceeding the highest threshold is shown to be equal to $A_j/(A_j + B_j)$ or $6/18 = 0.33$ for our example. In general, the probability of exceeding the j th threshold $p_{e,j}$ is given as

$$p_{e,j} = p_{e,j+1} + (A_j/[A_j + B_j])(1 - p_{e,j+1})$$

which is solved iteratively for $j = m, m-1, \dots, 2, 1$. By convention, $p_{e,m+1} = 0$. Thus the probability of exceeding the lower threshold is

$$0.33 + (3/9)(0.67) = 0.556$$

To assign plotting positions, the six uncensored observations above the highest detection limit are given Weibull plotting positions between $(1 - p_{e,2}) = 0.67$ and 1.00 (Table 7 and Figure 6). The three uncensored values between the two detection limits have positions spread between $(1 - p_{e,1}) = 0.444$ and 0.67. In general, Weibull plotting positions for uncensored observations are

$$p(i) = (1 - p_{e,j}) + (p_{e,j} - p_{e,j+1}) r/(A_j + 1)$$

where r is the rank of the i th observation among the A_j observations above the j th detection limit [Hirsch and Stedinger, 1987].

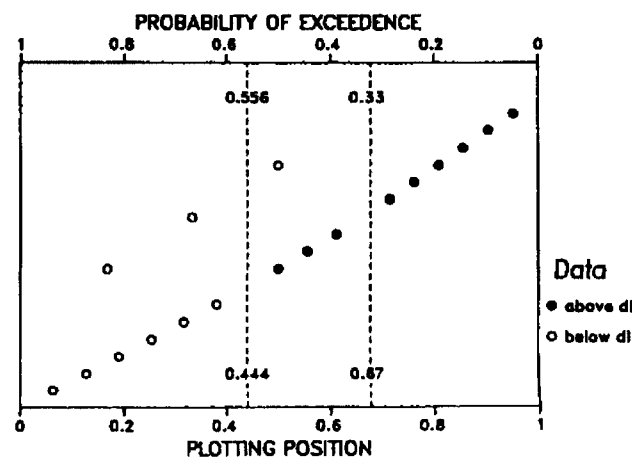


Fig. 6. Plotting positions illustrated for the Appendix B data.

In order to "fill in" data prior to estimating moment statistics using the MR method, plotting positions for censored observations must be determined. The six censored observations below the lower detection limit can be assigned positions between 0.00 and $(1 - 0.556) = 0.444$ (Table 7). The three censored observations below the higher detection limit are assigned positions spread between 0.00 and 0.67 (the upper sequence of points in Figure 6), since there is no information to determine whether they are above or below the lower detection limit. In general, Weibull plotting positions for censored observations are given by

$$pc(i) = (1 - p_{e,j}) r / (C_j + 1)$$

where r is the rank of the i th observation among the C_j censored values known only to be less than the j th detection limit

$$C_j = B_j - \sum_{k=0}^{j-1} (A_k + B_k)$$

and $A_0 = B_0 = 0$. Thus the first of three censored observations below the higher detection limit has the plotting position

$$pc = (1 - 0.33) 1 / (3 + 1) = 0.167.$$

APPENDIX C: ADJUSTED LIKELIHOOD METHOD

Cohn [1988] has derived the adjusted maximum likelihood method along the following lines. Results from Shenton and Bowman [1977] were used to obtain first-order estimates of the bivariate distribution of $\hat{\mu}$ and $\hat{\sigma}^2$. Asymptotically independent functions of $\hat{\mu}$ and $\hat{\sigma}^2$ were derived. The distribution of one function is asymptotically normal. The distribution of the second function is asymptotically gamma, with shape parameter given as a function only of the censoring probability and sample size. An adjusted maximum likelihood estimator, similar to Finney's [1941], was then derived for estimating the lognormal moments. The estimator is asymptotically equivalent to the maximum likelihood estimator MM, but is found to be substantially less biased than MM if σ^2/N is large.

REFERENCES

- Aitchison, J. and J. A. C. Brown, *The Lognormal Distribution*, Cambridge University Press, New York, 1957.
- Cohen, A. C., Jr., Simplified estimators for the normal distribution when samples are singly censored or truncated, *Technometrics*, 1(3), 217-237, 1959.
- Cohen, A. C., Jr., Progressively censored sampling in the three parameter log-normal distribution, *Technometrics*, 18(1), 99-103, 1976.
- Cohn, T. A., Adjusted maximum likelihood estimation of the moments of lognormal populations from type I censored samples, *U.S. Geol. Surv. Open File Rep.*, 88-350, 1988.
- Cunnane, C., Unbiased plotting positions—A review, *J. Hydrol.*, 37, 205-222, 1978.
- David, H. A., *Order Statistics*, 2nd ed., John Wiley, New York, 1981.
- Finney, D. J., On the distribution of a variable whose logarithm is normally distributed, *J. R. Stat. Soc. London, Ser. B*, 7, 155-161, 1941.
- Gilbert, R. O., and R. R. Kinnison, Statistical methods for estimating the mean and variance from radionuclide data sets containing negative, unreported or less-than values, *Health Phys.*, 40, 377-390, 1981.
- Gilliom, R. J., and D. R. Helsel, Estimation of distributional parameters for censored trace level water quality data, 1, Estimation techniques, *Water Resour. Res.*, 22(2), 135-146, 1986.
- Gleit, A., Estimation for small normal data sets with detection limits, *Environ. Sci. Technol.*, 19(12), 1201-1206, 1985.
- Hirsch, R. M., and J. R. Stedinger, Plotting positions for historical floods and their precision, *Water Resour. Res.*, 23(4), 715-727, 1987.
- Janzer, V. J., Report of the U.S. Geological Survey's Analytical Evaluation Program—Standard Reference Water Samples M6, M94, T95, N16, P8, and SED3, technical report, Branch of Quality Assurance, U.S. Geol. Surv., Arvada, Colo., 1986.
- Kalbfleisch, J. D., and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, John Wiley, New York, 1980.
- Kendall, M., and A. Stuart, *The Advanced Theory of Statistics*, 4th ed., vol. 2, Oxford University Press, New York, 1979.
- Latta, R. B., A monte carlo study of some two-sample rank tests with censored data, *J. Am. Stat. Assoc.*, 76(375), 713-719, 1981.
- McBean, E. A., and F. A. Rovers, Alternatives for handling detection limit data in impact assessments, *Ground Water Monit. Rev.*, 4(2), 42-44, 1984.
- Nehls, G. J., and G. G. Akland, Procedures for handling aerometric data, *J. Air Pollut. Control Assoc.*, 23(3), 180-184, 1973.
- Shacklette, H. T., and J. G. Boerngen, Element concentrations in soils and other surficial materials of the conterminous United States, *U.S. Geol. Surv. Prof. Pap.*, 1270, 1984.
- Shenton, L. R., and K. O. Bowman, *Maximum Likelihood Estimation in Small Samples*, Griffin, London, 1977.
- T. A. Cohn and D. R. Helsel, Water Resources Division, U.S. Geological Survey, 410 National Center, Reston, VA 22092.

(Received February 25, 1988;
revised July 12, 1988;
accepted July 22, 1988.)