

Practical Stats Newsletter for Winter, 2005

In this newsletter:

1. Upcoming Course "Applied Environmental Statistics"
2. The Delta Lognormal method - its really just substitution
3. New textbook "Nondetects And Data Analysis" for sale online

1. Upcoming Course: Applied Environmental Statistics
March 21-25, 2005 Colorado School of Mines
Golden, Colorado

Our 1-week survey of applied statistics. How to "make sense of your data".
Registration is online through PracticalStats.com.

2. The Delta Lognormal Method
- It's really just substitution

The delta-lognormal method (also called the D-LOG, or Aitchison's method) is one of the most frequently recommended methods for computing descriptive statistics with censored data. As first proposed by Aitchison (1955), the method was applied to economic data where actual zeros were plausible. Later, the method was invoked for environmental concentrations below a reporting limit.

Aitchison's method modeled detected observations using a lognormal distribution while assuming all nondetects equal zero. The mean of a lognormal distribution is the familiar formula:

$$\bar{x}_2 = \exp[\bar{y} + 0.5 \cdot \text{Var } y]$$

where \bar{y} and $\text{Var } y$ are the mean and variance of the natural logarithms.

Because all nondetects are assumed to be zero, the mean of detected (lognormal) observations \bar{x}_2 is adjusted downwards by the proportion of detected observations (n_2/n) to estimate the overall mean \bar{x} .

$$\bar{x} = (n_2/n) \cdot \bar{x}_2$$

The only difference between the delta-lognormal method and substitution of zeros for all nondetects is in how the mean of the detected values is computed. For simple substitution, this mean is the sum of the detected values divided by n_2 . So for simple substitution of zeros, $\bar{x}_2 = \text{sum}(x)/n_2$. The overall mean of the dataset for substitution is again

$$\bar{x} = (n_2/n) \cdot \bar{x}_2$$

Gilliom and Helsel (1986) found that the performance of the delta-lognormal method was essentially identical to that for the zero substitution method. Neither method worked as well as alternatives. Yet, the delta-lognormal method has continued to be recommended in some guidance documents such as

the USEPA's Guidance for Data Quality Assessment (1998).

The USEPA's Technical Support Document for Water Quality-based Toxics Control (1991) modified the delta-lognormal method, though the name stayed the same. In the modified method, nondetects are assumed to fall at the detection limits rather than at zero. This change produces the highest possible value for an estimated mean, though it underestimates the standard deviation. As with substituting zeros, there is little difference between the TSD's modified method and substituting the reporting limit for all nondetects. The modified method has the same primary flaw as substitution of the reporting limit - the values substituted reflect changing lab precision, but not anything to do with concentrations in the sample itself. Therefore, the poor performance when substituting the detection limit found by Gilliom and Helsel (1986), and by others since then, can be applied to the TSD's modified delta-lognormal procedure. Hinton (1993) evaluated the modified procedure directly and found it to be more in error than the alternatives.

The delta-lognormal method is essentially substitution -either zero or the reporting limit is substituted for all nondetects in this method. As with other substitution, it produces estimates of mean and standard deviation that have high bias. There are better ways.

References

1. Aitchison (1955), Journal of the American Statistical Assoc. 50, 901-908.
2. Gilliom and Helsel (1986), Water Resources Research 22, 135-146.
3. Helsel. and Hirsch, 2002, Statistical Methods in Water Resources. U.S. Geological Survey Techniques of Water Resources Investigations, Book 4, Chapter A3, 512 pp. Available for download at <http://water.usgs.gov/pubs/twri/twri4a3/>
4. Hinton (1993), Env. Sci. and Technol. 27, 2247-2249.
5. USEPA (1991), Technical Support Document For Water Quality-based Toxics Control. EPA/505/2-90-001.
6. USEPA (1998), Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA/600/R-96/084.

3. New textbook "Nondetects And Data Analysis" for sale online

The new textbook Nondetects And Data Analysis: Statistics for censored environmental data" is now available online from Wiley:
<http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471671738.html>

from Amazon:

<http://www.amazon.com/exec/obidos/tg/detail/-/0471671738/002-5558630-9979221>

?v=glance&s=books

or from your favorite online bookseller. The book covers methods for handling nondetects. One online reviewer states "The book is well organized and provides many useful and informative examples. I have found the information to be extremely helpful in my work as an in-house environmental consultant working for a large manufacturing company."

'Til next time,

Practical Stats

<http://www.practicalstats.com>

-- Make sense of your data