

NADA for R. A contributed package for censored environmental data

Dennis R. Helsel

US Geological Survey

Lopaka Lee

Trace contaminants in air, water biota, soils, and rocks often contain data recorded only as a "nondetect", or less than a detection threshold. These left-censored values cause difficulties for environmental scientists, as no single number can be validly assigned to them. The typical solution of substituting one-half the detection limit and proceeding with regression, t-tests, etc., has repeatedly been shown to be inaccurate. Instead, these data can be effectively interpreted using survival analysis techniques more traditionally applied to right-censored data. Methods for calculating descriptive statistics, testing hypotheses, and performing regression, both parametric and nonparametric, are available using the contributed package NADA. Methods include censored maximum likelihood (ML), Kaplan-Meier, and the Akritas version of Kendall's robust line that is applicable (unlike ML) to doubly-censored data. Methods such as censored boxplots and residuals plots that can graph data containing nondetects are also included. The NADA package complements the first author's textbook, *Nondetects And Data Analysis: Statistics for censored environmental data* (Wiley, 2005).