

## NADA Reference Card -- Statistical Analysis of Censored Data using R

Adapted from Table 2.1 & Sec. 14.2 of *Stats for Censored Environ Data* (Helsel 2012) by Steve Saiz, 9/30/14

Method for Uncensored Data	Method for Censored Data	library(NADA) commands in R*
<b>1. Summary statistics for one sample</b>		
Descriptive statistics	Kaplan–Meier (KM) estimates ROS estimates MLE estimates Compute mean, median using 3 methods --assume lognorm d'n for ROS, MLE Summary censoring levels and %s in data	cenfit(Y, Yc) cenros(Y, Yc) cenmle(Y, Yc)  censtats(Y, Yc) censummary(Y, Yc)
Quantiles	5, 10, 25, 50, 75, 90, and 95 <sup>th</sup> quantiles	quantile(cenfit(Y, Yc))
Confidence intervals, two-sided	Build object w desired confidence... 90% conf. interval on mean via KM 90% conf. int. on 75 <sup>th</sup> percentile via KM	Ykm <- cenfit(Y, Yc, conf.int = 0.90) mean(Ykm) quantile(Ykm, prob = 0.75, conf.int = TRUE)
<b>2. Summary statistics for two or more groups</b>		
Descriptive statistics	KM est. median, mean, SD for ea. group	cenfit(Y, Yc, Gp)
<b>3. Comparing two or more groups</b>		
t-test or ANOVA	Gen. Wilcoxon test on Y by Gp, Peto-Prentice MLE test mean ln(Y) by Gp MLE test mean Y by Gp	cendiff(Y, Yc, Gp) cenmle(Y, Yc, Gp) cenmle(Y, Yc, Gp, dist="gaussian")
<b>4. Correlation</b>		
Nonparametric Parametric	Kendall's tau. Xc optional. Pearson's r correl. coeff. – lognormal Pearson's r correl. coeff. – normal d'ns	cenken(Y, Yc, X, Xc) cenreg(Cen(Y, Yc) ~ X) cenreg(Cen(Y, Yc) ~ X, dist = "gaussian")
<b>5. Regression</b>		
Linear regression	Akritas-Thiel-Sen (ATS) censored regression. Xc optional. MLE censored regression – Y lognormal. MLE censored regression – Y normal d'n.	cenken(Y, Yc, X, Xc)  cenreg(Cen(Y, Yc) ~ X) cenreg(Cen(Y, Yc) ~ X, dist = "gaussian")
Multiple regression	MLE cens. regress. Y vs. many X vars. Alternate form for dataframe, <i>df</i>	cenreg(Cen(Y, Yc) ~ X + X2) with(df, cenreg(Cen(Y, Yc) ~ X + X2))
<b>6. Plotting censored data</b>		
Scatter plot	X-Y scatter plot, censored values as dashed lines. Xc or Yc can be FALSE. Add ATS line to scatter plot.	cenxyplot(X, Xc, Y, Yc) lines(cenken(Y, Yc, X))
Box plot	censored Y boxplot, log <sub>e</sub> transform. censored Y boxplot, no log transf. censored Y boxplot by grouping var. Gp	cenboxplot(Y, Yc) cenboxplot(Y, Yc, log=F) cenboxplot(Y, Yc, Gp)
Quantile plot	KM <i>edf</i> of Y with conf. bands ROS prob plot of Y, lognormal d'n ROS prob plot of Y, normal d'n MLE prob plot of Y, lognorm	plot(cenfit(Y, Yc)) plot(cenros(Y, Yc)) plot(cenros(Y, Yc, forwardT=NULL)) plot(cenmle(Y, Yc))
Compare groups	KM <i>edf</i> of Y for two or more groups, Gp MLE prob plot of residuals	plot(cenfit(Y, Yc, Gp)) plot(cenmle(Y, Yc, Gp))

\* where Y and X are numeric data with DL for non-detects, Yc and Xc are logical variables (TRUE means a non-detect), and Gp is a factor variable (categorical).