Statistical Analysis of Data Sets with Values Below Detection Limits

Michael Gochfeld, Joanna Burger, Vikram Vyas

CRESP designed its field collections to allow statistical analysis of results for comparison among the three Amchitka test sites and between Amchitka and Kiska. The plan was based on the assumption that some of the radionuclides would be present in sufficient concentrations to allow statistical testing. When preliminary results for most of the radionuclides were below the detection level (the minimal detectable activity or MDA), procedures were modified (increased sample size and count time) to enhance the sensitivity by lowering detection levels to the point where some values were above the MDA. An analogy is a bathtub with some toys on the bottom. You can see the toys through the water, but you can't tell how large they are. As you lower the water level, some of the toys may protrude above the surface allowing them to be measured. Lowering the detection level is analogous to lowering the water level, thereby enhancing the sensitivity of the analysis.

The MDA is the lowest value which can be reliably quantified. In radiation measurement the MDA depends on the sensitivity of the detector, the energy of the radionuclide, the matrix (tissue) being analyzed, including its size and shape (geometry), the presence of other contaminants with similar radiation energy, and most importantly other sources of background radiation which must be subtracted from the detector reading.

Even when all analytic results are above the MDA, there are still statistical issues in how the distributions of values can be compared, let's say between Amchitka and Kiska. Standard parametric statistical tests such as t-tests are commonly used to compare distributions (for example the heights of boys and girls in a class), but require that the heights are normally distributed (like a bell-shaped curve). This is rarely the case for environmental data which are often skewed in one direction (Helsel 1990), and sometimes bimodal (a combination of some very low values and some very high values). A variety of statistical manipulations are used to normalize data, but an alternative approach is to use non-parametric statistics based on the ranks or proportions of values rather than the actual values themselves.

For several of the radionuclides in this study most or all of the values were below the MDA, which is a major challenge for statistical analysis. For many years there has been extensive scientific discussion of how to treat values below detection levels (often called "censored data"), and a recent book (Helsel 2005) examines several methods in detail. In CRESP's review of contaminant literature several alternatives exist:

1. Ignore non-detectable values entirely, and discard them from the data set, analyzing only values that are above the detection level.
2. Assume that if a value is non-detectable it is actually zero and assign it that value in the data set.
3. Assume that if a value is non-detectable, it is probably close to the detection level and assign it that value in the data set.
4. Arbitrarily set non-detectable values equal to half the detection level.

5. Where there are estimated values treat them as real values even if they are below the detection level.
6. Use a statistical extrapolation from values above the detection level to estimate what the values below the detection level might be.
7. Use only a non-parametric representation such as the median.
8. Decide not to conduct any statistical analysis.

Each of these may make sense in a particular situation. CRESP researchers considered each of the above as follows:

1. This approach was tried as one option, despite the fact that non-detectable values are truly low and should not be ignored. Moreover, eliminating them from the data base reduces the sample size, and makes any statistical analysis problematic. This approach, commonly used with radiation data, always overestimates the true value, since low values are dropped.
2. Setting a non-detectable value to zero has to underestimate its true value systematically, hence this approach was rejected. However, some agencies such as the U.S. Food and Drug Administration have used this approach, for example for reporting mercury levels in fish (Yess 1993).
3. Assigning the MDA to non-detectable values would systematically overestimate them. This approach was rejected.
4. In many studies of contaminants statistical analyses are conducted setting non-detectable values to half the detection level. This default approach has limitations (Helsel 1990), but is common because it avoids the biases inherent in approaches 1, 2, 3. CRESP tried this approach and reported these results in the text of its report.
5. Since even the non-detectable results had best estimate counts assigned, these numbers could be used, based on the assumption that they are unbiased estimates, and that the uncertainties associated with them would be randomly distributed. This assumption is questionable. They might systematically overestimate the true Cs-137 by a small amount. CRESP tried this approach as well.
6. The most sophisticated approach to the radionuclide data would be to estimate the values of the non-detectable results by applying a regression to the values that are above detection and extrapolating downward (Helsel 1990). Several statistical approaches exist, and the UNCENSOR program is freely available (Newman and Evans 2005). There is an underlying assumption that the distribution is at least lognormal. However, some of the species have bimodal or multimodal distributions, which do not lend themselves to extrapolation.

These programs calculate a mean and variance to allow statistical characterization and comparison. However, UNCENSOR assumes that the detection level is constant within a data set, as it is for most non-radiologic analyses. This does not apply to the radiologic analysis where every analysis has a somewhat different detection level depending on the conditions of the sample, possible interferences, and the background radiation at the time of the counts.

Vyas and Kosson (in press) have expanded on these techniques where there is variation in MDAs, but even these enhancements cannot extrapolate, when most of the values are below the MDA. Helsel (1990) cautions that regression methods are not reliable when more than 40% of the values are non-detects.

7. The median values were determined for each of the data sets, and although the results do not correlate well with method 4, they are probably useful.

8. CRESP also adopted the approach of not doing statistical comparisons for those radionuclides where all or virtually all of the results were below the MDA.

CRESP also explored whether the ratio between the count value and its corresponding MDA could provide useful information, but this Cs-Ratio (see Table 1), was not stable in this data set.

The Table 1 provides comparisons of four of the methods (1, 4, 5, and 7) for several components of the Cs-137 data set. Not surprisingly counting detects only or substituting half the MDA for non-detects, gives results that are highly correlated with the MDA, while averaging all raw values is independent of the MDA.

Table 2 shows the intercorrelation among the various methods. Both the half-detection substitution and the median, performed reasonably well. For species that truly have very low Cs-137 levels, using half the MDA will overestimate the mean, while for species with a continuous distribution and at least some detectable values, it is probably the best approximation

In conclusion, despite its limitations, the default method of using half the MDA, widely used for inorganic and organic contaminants, can provide a useful representation, for some radiologic data as well. As the proportion of non-detects increases the reliability of any comparison decreases.

References:
Helsel DR. 1990. Less than obvious. Environ. Science Technol. 24:1766-1774.

Helsel DR. 2005. Nondetects and Data Analysis: Statistics for Censored Environmental Data. Wiley and Sons, New York.

Newman MC, Evans D. 2005. Uncensor Program. Http://www.vims.edu/env/research/software/vims_software.html#uncensor

Vyas VM, Roy A, Strawderman W, Georgopoulos PG, Kosson DS. Development and application of a methodology for determining background groundwater quality at the Savannah River Site. Kosson DR. in press. J Air Waste Managem. Assoc.

Yess M. 1993. U.S. Food and Drug Administration survey of methyl mercury in canned tuna. J AOAC Int. 76:36-38.

Table 1 Comparisons among approaches for estimating the central tendency of distributions with a high proportion of values below detection level. "Large" refers to 1000 gram samples and "Small" refers to 100 gram samples.

		Composites Analyzed	**Number of detects**	**Cs-137 MDA**	**Cs-137 Raw values**	**Cs-137 Half MDA**	**Cs-137 Detects only**	**Cs-137 Ratio**	**Median**
METHOD					5	1	4		7
All samples	all	179	22	1.962	0.015	1.021	0.401	0.413	0.133
Alaria	large	4	0	0.179	0.005	0.09	No detects	0.145	0.003
Black Rockfish	large	3	3	0.097	0.143	0.143	0.143	1.46	0.13
Halibut	large	4	3	0.155	0.284	0.278	0.317	2.611	0.252
Rock Jingles	large	3	0	0.319	-0.075	0.159	.No detects	-0.235	-0.088
Octopus	large	4	4	0.092	0.262	0.262	0.262	2.857	0.254
Pacific Cod	large	14	7	0.278	0.227	0.269	0.337	1.454	0.194
Rock Greenling	large	5	0	0.251	0.091	0.126	No detects.	0.498	0.097
Ulva latuca	large	3	0	0.208	0.026	0.14	No detects	0.287	0.051
Walleye Pollock	large	2	1	0.322	0.242	0.311	0.461	0.752	0.241
Black Rockfish	small	15	0	2.399	0.091	1.2	No detects	0.086	0.165
Dolly Varden	small	8	0	3.974	1.017	1.987	No detects	0.278	1.039
Pacific Cod	small	14	0	3.24	0.202	1.764	No detects	0.155	0.347
Rock Greenling	small	27	0	2.194	0.072	1.097	No detects	0.088	0.045
Yellow Irish Lord	small	16	1	2.704	-0.205	1.386	1.62	0.036	0.087

Table 2. Kendall tau Intercorrelation among alternative measures of central tendency for distributions of Cs-137 in biota, with many values below detection level. Note that there is not a constant relationship between the Kendall tau value and the associated probability.

	MDA	Cs-137	Half	Detects	Ratio	Median
MDA	xxxx	-0.06	0.71	0.81	-0.46	0.12
Cs-137	-0.06	xxxx	0.19	-0.33	0.52	0.71
Half detect	0.71	0.19	xxx	0.81	-0.22	0.37
Detects only	0.81	-0.33	0.81	xxx	-0.71	-0.33
Ratio	-0.46	0.52	-0.22	-0.71	Xxx	0.41
Median	0.12	0.71	0.37	-0.33	0.41	Xxx
Average Intercorrelation	0.224	0.206	0.372	0.05	-0.092	0.256

P<0.05 P<0.01 P<0.001