Appendix G

Student's T-Test Calculation

The following example details the steps needed to determine the number of samples to be collected, based on historical analytical results and the statistical student's t-test. It has been adapted from information contained in EPA's *An Addendum to the POTW Sludge Sampling and Analysis Guidance Document*, May 1992.

This page is intentionally blank.

HOW TO DETERMINE THE APPROPRIATE NUMBER OF SAMPLES TO BE COLLECTED

In Chapter 5, the following two equations were provided as a method to determine how many samples should be collected to represent the whole or to determine how many grab samples should be collected to form a composite.

(1)

$$S = \sqrt{\frac{\sum \left|\overline{X} - x\right|^2}{N-1}}$$

Where:

S = standard deviation

$\overline{X}$ = average or mean of all data points

x = individual data points

N = number of data points in the set

$\sum \left|\overline{X} - x\right|^2$ = sum of square of the difference between the mean and each individual data point

(2)

$$N = \frac{T^2 S^2}{\left(RL - \overline{X}\right)^2}$$

Where:

N = the minimum samples to characterize sludge

T = value of Student's t for the appropriate number of historical data points at 90% confidence level

S = standard deviation

RL = the regulatory limit for the analyte in question

$\overline{X}$ = mean of the historical data

To use this method:

1) Assemble your historical analytical data for the analyte of interest.
2) Calculate the mean or average.
3) Calculate the standard deviation using Equation 1.
4) Determine the regulatory limit for the analyte chosen.
5) Find the Student's T value from Table G-1.
6) Using the mean, standard deviation, regulatory limit, and value of Student's T determined above, calculate the appropriate number of samples by using Equation 2.

Table G-1. VALUES FOR STUDENT'S T AT THE 90% CONFIDENCE LEVEL

Degrees of Freedom (df)	T value at 90% Confidence Level
1	6.314
2	2.920
3	2.353
4	2.132
5	2.015
6	1.943
7	1.895
8	1.860
9	1.833
10	1.812
11	1.796
12	1.782
13	1.771
14	1.761
15	1.753
16	1.746
17	1.740
18	1.734
19	1.729
20	1.725
21	1.721
22	1.717
23	1.714
24	1.711
25	1.708
26	1.706
27	1.703
28	1.701
29	1.699
30	1.697
40	1.684
50	1.676
60	1.671
70	1.667
80	1.664
90	1.662
100	1.660
120	1.658
infinity	1.645

Sample Calculation

Below is a step-by-step example calculation. The objective is to determine the number of composite samples that should be collected during the year to produce statistically valid sludge copper (Cu) concentrations. The following historical Cu data (see Table G-2) will be used.

Table G-2. HISTORICAL COPPER DATA

Date of Sample	Copper Concentration (mg/kg)
7/24/02	480
1/13/03	360
6/11/03	330
5/15/03	135
11/6/03	400
1/7/04	189
4/12/04	140
5/27/04	200
10/26/04	79
1/27/05	140
3/22/05	100
5/27/05	268

Step 1: Calculate the average Cu concentration, add all the concentrations and divide by the number of values:

Date of Sample	Copper Concentration (mg/kg)
7/24/02	480
1/13/03	360
6/11/03	330
5/15/03	135
11/6/03	400
1/7/04	189
4/12/04	140
5/27/04	200
10/26/04	79
1/27/05	140
3/22/05	100
5/27/05	268
TOTAL	2821

Average Copper = 2821 ÷ 12 = 235 (rounded to the nearest whole number)

$$\overline{X} = 235$$

Step 2: Calculate the standard deviation. Fortunately most spreadsheet applications will perform the calculation for you. To perform the process by hand, subtract each individual Cu concentration from the average concentration. Next, square the difference between the average and individual values and sum the squares. See Table G-3 for an example of these calculations. This sum of squared differences can be inserted into the numerator of Equation 1 above.

Table G-3. CALCULATING THE SUM OF SQUARED DIFFERENCES

Date of Sample	Cu Concentration (mg/kg)	$(\overline{X} - x)$	$(\overline{X} - x)^2$
7/24/02	480	-245.00	230400
1/13/03	360	-125.00	129600
6/11/03	330	-95.00	108900
5/15/03	135	100.00	18225
11/6/03	400	-165.00	160000
1/7/04	189	46.00	35721
4/12/04	140	95.00	19600
5/27/04	200	35.00	40000
10/26/04	79	156.00	6241
1/27/05	140	95.00	19600
3/22/05	100	135.00	10000
5/27/05	268	-33.00	71824
SUM	2821		186941

The remaining calculation is as follows:

$$\text{Standard Deviation} = \sqrt{\frac{186941}{12-1}} = 130 \text{ (rounded to the nearest whole number)}$$

Step 3: Based on federal regulations, the ceiling limit for Cu is 4300 mg/kg and the pollutant concentration limit is 1500 mg/kg. In this example, we will assume that the facility wants to show compliance with the lower limit.

Going back to Equation 2, we can see that the average, standard deviation, and regulatory limit have been determined. To use the equation, the final value that must be obtained is Student's T at a 90% confidence level. To find Student's T, use Table G-1. First, find the degrees of freedom by subtracting 1 from the number of historical data points you used to determine the average and standard deviation.

Degrees of Freedom (df) = 12 - 1 = 11

Using Table G-1, locate the Student's T for 11 degrees of freedom (1.796).

Now all the values can be inserted into Equation 2 to obtain the number of grab samples to form a composite.

$$\text{Number of composite samples} = \frac{1.796^2 \, x \, 130^2}{(1500 - 235)^2} = 0.03$$

This calculation indicates that, based on historical data and the current regulatory limit, one composite sample should be sufficient to ensure that regulatory limits are being met. However, facilities must perform the sampling required by state and federal regulations regardless of the results of this calculation. Also, facility operators should be aware that the results of this calculation are heavily influenced by the variability of the historical data and the regulatory limit. For example, if the regulatory limit were 400 mg/kg, the results would indicate that two samples were needed. As a rule of thumb, if the mean of historical data plus the standard deviation is greater than the regulatory limit, then Equation 2 may be helpful in determining the appropriate sampling frequency or number of samples.

This page is intentionally blank.