Statistical Data Summary – as of July 20, 2009

VDH limits per GMP 147

Standard Error of the Mean, 99% Upper Confidence Level – Log Transformed (last row shown for each manufacturer) must meet the following for approval:
BOD ≤ 10, TSS ≤ 10, Fecal ≤ 2000.

Also shown below are:
First Row - the 99% Upper Confidence Level for the raw (untransformed) data.

Second Row – the 99% Upper Confidence Level for the Log Transformed data.

		BOD mg/l	TSS mg/l	Fecal cfu/100ml	UV Fecal cfu/100 ml
MicroFAST					
Raw Data	99% Upper Confidence Level – untransformed	73.34	153.47	301,100	8,463
Log Transformed	99% Upper Confidence Level – Log Transformed	18.85	18.76	36,135	231
Standard Error of the Mean	Std Error of the Mean, 99% Upper Confidence Level – Log Transformed	9.9	7.83	48,841	109
Ecoflo					
Raw Data	99% Upper Confidence Level – untransformed	28.5	24.7	184,229	
Log Transformed	99% Upper Confidence Level – Log Transformed	13.79	16.38	2,235	
Standard Error of the Mean	Std Error of the Mean, 99% Upper Confidence Level – Log Transformed	6.76	4.4	1,690	
Advantex					
Raw Data	99% Upper Confidence Level – untransformed	43.19	36.8	26,459	
Log Transformed	99% Upper Confidence Level – Log Transformed	18.45	18.07	1,427	

		BOD mg/l	TSS mg/l	Fecal cfu/100ml	UV Fecal cfu/100 ml
Standard Error of the Mean	Std Error of the Mean, 99% Upper Confidence Level – Log Transformed	5.47	7.01	1,164	
Puraflo					
Raw Data	99% Upper Confidence Level – untransformed	28.67	208.31	1,533,900	
Log Transformed	99% Upper Confidence Level – Log Transformed	16.10	123.52	28,435	
Standard Error of the Mean	Std Error of the Mean, 99% Upper Confidence Level – Log Transformed	5.18	7.84	1,076	

The **standard error of the mean** (SEM) is the standard deviation of the sample mean estimate. (It can also be viewed as the **standard deviation of the error in the sample mean relative to the true mean**.) SEM is usually estimated by the sample estimate of the population standard deviation (sample standard deviation) divided by the square root of the sample size (assuming statistical independence of the values in the sample):

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

where

s is the sample standard deviation (i.e., the sample based estimate of the standard deviation of the population), and
n is the size (number of observations) of the sample

A **confidence interval** (**CI**) is a particular kind of interval estimate of a population parameter. Instead of estimating the parameter by a single value, an interval likely to include the parameter is given. Thus, confidence intervals are used to indicate the reliability of an estimate. How likely the interval is to contain the parameter is determined by the **confidence level** or confidence coefficient. Increasing the desired confidence level will widen the confidence interval.

A confidence interval is always qualified by a particular **confidence level**, usually expressed as a percentage; thus one speaks of a "95% confidence interval". The end points of the confidence interval are referred to as **confidence limits**. For a given estimation procedure in a given situation, the higher the confidence level, the wider the confidence interval will be.

If a data distribution is approximately normal then about 68% of the values are within 1 standard deviation of the mean (mathematically, $\mu \pm \sigma$, where μ is the arithmetic mean), about 95% of the values are within two standard deviations ($\mu \pm 2\sigma$), and about 99.7% lie within 3 standard deviations ($\mu \pm 3\sigma$). This is known as the *68-95-99.7 rule*, or *the empirical rule*.