**BASIC STATISTICS**

**Measures of Location**
Arithmetic Mean - numerical average. Can be greatly distorted by a few large values.

\[
\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i
\]  \hspace{1cm} (1)

Geometric Mean - closer to the mode than the arithmetic mean

\[
\mu_g = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n}
\]  \hspace{1cm} (2)

Median - an equal number of items have lower and higher values.

Mode - position of highest frequency. The mode of the function \(f(x)\) is \(x_2\) if the following condition is satisfied:

\[
\left[ \frac{df(x)}{dx} \right]_{x=x_2} = 0
\]  \hspace{1cm} (3)

Skew - location of the mean with respect to the mode

\[
k = \frac{\mu_x - \text{mode}_x}{\sigma_x}
\]  \hspace{1cm} (4)

**Measures of Dispersion**
Variance - the most commonly accepted measure of dispersion. For continuous variables it is defined as

\[
\sigma^2 = \int_{-\infty}^{\infty} (x-\mu_x)^2 f(x) \, dx
\]  \hspace{1cm} (5)

while for discrete values it is defined as

\[
\sigma_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_x)^2
\]  \hspace{1cm} (6)
Standard Deviation - the square root of equations 5 and 6. It can also be written as

\[ \sigma_x = \sqrt{\frac{1}{n-1} \sum (x-\mu_x)^2} \]  

Coefficient of Variation - Provides a means for comparing standard deviations from different measurements

\[ v_x = \frac{\sigma_x}{\mu_x} \]  

**Estimation of Statistical Error**

Mean of a Random Sample - the average value of a random sample of size n can itself be considered a random variable having a characteristic distribution. The theoretical value of this parameter coincides with the mean of a population

\[ \mu_x = \mu_x \]  

Standard Deviation of the Mean - is different from the standard deviation

\[ \sigma_x = \frac{\sigma_x}{\sqrt{n}} \]  

Standard Error of the Mean - the error associated with sampling a population

\[ SE = \sigma_x = \frac{\sigma_x}{\sqrt{n}} = \sqrt{\frac{\sum (x-\mu_x)^2}{n(n-1)}} \]  

**Confidence**

Confident Level - percentage of values which fall within a specified range (distance from the mean)

Confidence Interval - the range of values which are within the specified confidence level. It is expressed as the ± range about the mean. At a 67% confidence level one would say that the values of x fall within the range \( x \pm SE \) while at a 95% confidence level this becomes \( x \pm 2 \text{ SE} \) and at a 99% confidence level it becomes \( x \pm 2.57 \text{ SE} \).
Frequency Functions

Normal Distribution - the type one is probably most familiar with

\[ f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left( -\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right) \]  \hspace{1cm} (12)

Log-Normal Distribution - often observed in grain and particle size distributions (is skewed towards lower values of \( x \))

\[ f(x) = \frac{1}{x \sqrt{2\pi \ln \sigma_g}} \exp\left( -\frac{1}{2} \left( \frac{\ln x - \ln \mu_g}{\ln \sigma_g} \right)^2 \right) \]  \hspace{1cm} (13)

where \( \sigma_g \) is the geometric standard deviation.

Figure 1 Plot of a normal distribution that has its mean value at 50 and a standard deviation of 30. The cumulative distribution is shown in red.
Figure 2 This is an example of a log-normal distribution. The median value is 9.06 and the average is 13.5.