g GE Power Systems

The Normal Probability Distribution

m *(mu), a measure of **central tendency**, is the mean or average of all values in the population. When only a sample of the population is described, mean is more properly denoted by* $\overline{x}$ *(x bar).*

s *(sigma) is a measure of **dispersion or variability**. With smaller values of s, all values in the population lie closer to the mean. When only a sample of the population is described, standard deviation is more properly denoted by s.*

Both m and s are specific values for any given population, and they change as the members of the population vary.

GE Power Systems

The Computational Equations

Population Mean

$$m = \frac{\sum_{i=1}^{N} X_i}{N}$$

Population Variance

$$s^2 = \frac{\sum_{i=1}^{N} (X_i - m)^2}{N}$$

Population Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{N} (X_i - m)^2}{N}}$$

Sample Mean

$$\hat{m} = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Sample Variance

$$\hat{s}^2 = s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

Sample Standard Deviation

$$\hat{s} = s = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}}$$

The difference between population and sample

- population has all the data points, N
- sample only has a portion of the total data points, n < N

The divisor for the population variance is the population size N, whereas the divisor for the sample variance is the sample size minus one (n-1). The divisor n-1 is used rather than N because this leads to an unbiased estimate for the population variance.