**TOOL:** One-way Analysis of Variance (ANOVA) tests the hypothesis that means are all the same for various treatments applied to a system, [1, 2].

**ASSUMPTIONS:**
1) There are \( k \) different treatments (samples) under consideration
2) For the \( i \)th of \( k \) different treatments (samples), there are \( n_i \) observations
3) All of the \( k \) treatments (samples) have the same variance, \( \sigma \)
4) The null and alternate hypotheses relate to equal means:
   \[ H_0: \mu_1 = \mu_2 = \ldots = \mu_k \]
   \[ H_1: \text{At least one of the means is not equal to the others} \]

**DEFINITIONS:**
\( \alpha \) = significance level for rejecting null hypothesis
\( k \) = total number of samples (treatments) being considered
\( i \) = index designating which sample (treatment) is being considered
\( N \) = total number of observations (data points) available for all treatments
\( n_i \) = number of observations available for sample (treatment) \( i \)
\( n \) = number of observations available for each sample if all \( n_i \) are equal
\( j \) = index designating which of \( n_i \) observations of sample \( i \) is being considered
\( \mu \) = grand mean of all observations for all samples
\( \mu_i \) = actual mean value for \( i \)th treatment
\( \alpha_i \) = difference between actual mean and grand mean for sample \( i \): \( \alpha_i = \mu_i - \mu \)
\( \epsilon_{ij} \) = difference between \( j \)th observation of \( i \)th sample and \( \mu_i \): \( \epsilon_{ij} = x_{ij} - \mu_i \)
\( x_{ij} \) = value of observation \( i \) for sample (treatment) \( j \): \( x_{ij} = \mu + \alpha_i + \epsilon_{ij} \)
\( \bar{x}_i \) = calculated mean for all observations from \( i \)th sample (treatment)
\( \bar{x}_. \) = calculated mean for all observations from all samples
\( \text{SSA} \) = Sum of Squared errors of All treatment (sample) means vs grand mean
\( \text{SSE} \) = Sum of Squared Errors of all observations vs respective sample means
\( \text{SST} \) = Sum of Squared errors Total for all observations vs grand mean = \( \text{SSA} + \text{SSE} \)
\( \text{MSA} \) = calculated Mean of Sum of All treatment squared errors;
\[ E(\text{MSA}) = \sigma^2 + \frac{\sum_{i=1}^{k} n_i \alpha_i^2}{k - 1} \]
\( \text{MSE} \) = calculated Mean of Sum of squared Errors; \( E(\text{MSE}) = \sigma^2 \)
ANOVA CALCULATIONS:

\[ N = \sum_{i=1}^{k} n_i \]
\[ \bar{x}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \]
\[ \bar{x}. = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij} = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{x}_{i.} \]
\[ SSA = \sum_{i=1}^{k} n_i (\bar{x}_{i.} - \bar{x}.)^2 \]
\[ SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2 \]
\[ SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}.)^2 = SSA + SSE \]
\[ MSA = \frac{SSA}{k-1} \]
\[ MSE = \frac{SSE}{N-k} \]

\[ f = \frac{MSA}{MSE} \]

Hypothesis Testing

- \[ f \text{ exceeds } F\text{-distribution critical value } f_{\alpha}(\nu_1=k-1, \nu_2=N-k) \]

**Theory:** From the definition of the \( \alpha_i \), the null and alternate hypotheses, \( H_0 \) and \( H_1 \), are equivalent to the following statements:

\( H_0: \ \alpha_1 = \alpha_2 = \ldots = \alpha_k \)

\( H_1: \ \text{At least one } \alpha_i \neq 0 \)

The \( MSA \) and \( MSE \) give different estimates of the variance:

\[ E(MSA) = \sigma^2 + \frac{\sum_{i=1}^{k} n_i \alpha_i^2}{k-1} \]

with \( k-1 \) degrees of freedom
and

\[ E(MSE) = \sigma^2 \] with \( N - k \) degrees of freedom

If the null hypothesis is true, the extra term in the MSA estimate is zero since all the \( \alpha_i \) are zero. In that case, the ratio of \( MSA \) to \( MSE \) will have an \( F \)-distribution with \( k - 1 \) and \( N - k \) degrees of freedom. We may then use the critical value of the \( F \)-distribution from a table to determine if the ratio of \( MSA \) to \( MSE \) is in the range expected if all the \( \alpha_i \) are zero. If the ratio exceeds the critical value, then we may assume that the second term in \( E(MSA) \) was not zero after all, and we reject the null hypothesis.

**NOTE:** Since an \( F \)-distribution describes ratios of variances, and variances are always positive, an \( F \)-distribution is nonzero only for positive values of \( f \). Thus, critical values of the \( F \)-distribution are always positive numbers, and we use a one-sided confidence interval or one-sided hypothesis test in the ANOVA method.

**NOTE:** The ANOVA method assumes \( \sigma \) is the same for all observations. This might be true, for example, if errors only arise from measurement techniques that are the same for all samples. Bartlett's test is a useful tool for determining if the variances are equal.

**NOTE:** If we use the same number of observations for all samples (treatments), then the \( f \)-ratio is relatively insensitive to small differences in variances, [1].
