# Understanding the F-statistic and F-test in ANOVA

Learn more about Minitab 17

IN THIS TOPIC



## How the F-statistics in the ANOVA output are calculated

Each F-statistic is a ratio of mean squares. The numerator is the mean square for the term. The denominator is chosen such that the expected value of the numerator mean square differs from the expected value of the denominator mean square only by the effect of interest. The effect for a random term is represented by the variance component of the term. The effect for a fixed term is represented by the sum of squares of the model components associated with that term divided by its degrees of freedom. Therefore, a high F-statistic indicates a significant effect.

When all the terms in the model are fixed, the denominator for each F-statistic is the mean square of the error (MSE). However, for models that include random terms, the MSE is not always the correct mean square. The expected mean squares (EMS) can be used to e denominator.

Home

ANOVA

ANOVA STATISTICS

th the fixed factor Screen and the random factor Tech, MS:

| n Square for Each Term |
| --- |
| + Q[1] |
| + 4.0000(2) |
| |
| |

a random effect associated with the term listed beside
random effect of Tech, (3) represents the random
and (4) represents the random effect of Error. The EMS
. In addition, the EMS for Screen*Tech is the effect of
ct of the Screen*Tech interaction.

Tech, the mean square for Screen*Tech is divided by
he expected value of the numerator (EMS for
rom the expected value of the denominator (EMS for
eraction (2.0000(3)). Therefore, a high F-statistic
eraction.

effect associated with the term listed beside the
he fixed effect of Screen. The EMS for Screen is the
the effect of the Screen*Tech interaction plus a
1] equals (b*n * sum((coefficients for levels of
and b are the number of levels of Screen and Tech,
plicates.

he mean square for Screen is divided by the mean
pected value of the numerator (EMS for Screen = (4) +
ected value of the denominator (EMS for Screen*Tech =
to the Screen (Q[1]). Therefore, a high F-statistic

## out include an "x" beside a p-valu label "Not an exact F-test"?

An exact F-test for a term is one in which the expected value of the numerator mean squares differs from the expected value of the denominator mean squares only by the variance component or the fixed factor of interest.

Sometimes, however, such a mean square cannot be calculated. In this case, Minitab uses a mean square that results in an approximate F-test and displays an "x" beside the p-value to identify that the F-test is not exact.

For example, suppose you performed an ANOVA with the fixed factor Supplement and the random factor Lake, and the got following output for the expected mean squares (EMS):

| Source | Expected Mean Square for Each Term |
|---|---|
| (1) Supplement | (4) + 1.7500(3) + Q[1] |
| (2) Lake | (4) + 1.7143(3) + 5.1429(2) |
| (3) Supplement*Lake | (4) + 1.7500(3) |
| (4) Error | (4) |

The F-statistic for Supplement is the mean square for Supplement divided by the mean square for the Supplement*Lake interaction. If the effect for Supplement is very small, the expected value of the numerator equals the expected value of the denominator. This is an example of an exact F-test.

Notice, however, that for a very small Lake effect, there are no mean squares such that the expected value of the numerator equals the expected value of the denominator. Therefore, Minitab uses an approximate F-test. In this example, the mean square for Lake is divided by the mean square for the Supplement*Lake interaction. This results in an expected value of the numerator being approximately equal to that of the denominator if the Lake effect is very small.

## About the "Denominator of F-test is zero or undefined" message

Minitab will display an error that the denominator of the F-test is zero or undefined for one of the following reasons:

- There is not at least one degree of freedom for error.
- The adjusted MS values are very small, and thus there is not enough precision to display the F and p-values. As a workaround, multiply the response column by 10. Then perform the same regression model, but instead use this new response column for the response.

  **Note:** Multiplying the response values by 10 will not affect the F and p-values that Minitab displays the output. However, decimal position will be affected in the remaining output, specifically, the sequential sums of squares, Adj SS, Adj MS, Fit, standard error of the fits, and residual columns.