

# Homogeneity of Variances - Performing Tests with R and R Commander

*Dr. James. Ready*

*October 31, 2019*

An independent samples  $t$ -test is one of the most commonly used statistical tests when comparing two population means. To perform this test, there must be a variable that is continuous in nature and another variable that is categorical in nature. The categorical variable can only have two levels. For example, this test could be used to measure the mean weight of people based on gender.

Prior to performing a  $t$ -test, two assumptions must be considered -

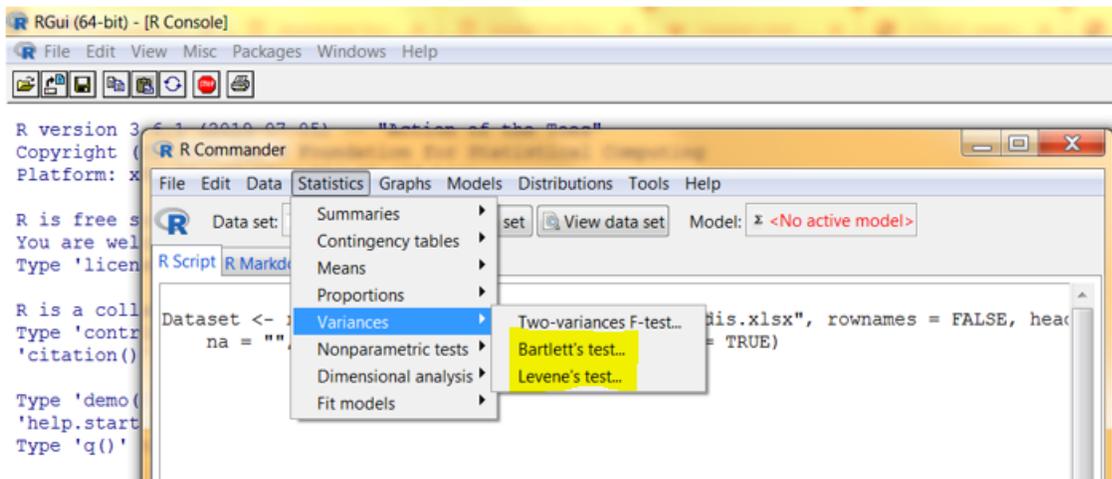
- the samples must be independent
- the continuous variable should have an approximate normal distribution

One other item must also be considered: unequal variances between groups. Performing a standard  $t$ -test with unequal variances results in a change of the test from exact to approximate. If sample sizes are small, this change could result in a Type I error. Thus, examining the homogeneity of variances prior to performing a  $t$ -test is important.

Two commonly used statistical tests used to examine the homogeneity of variances are the **Bartlett test** and **Levene test**. The focus of this paper is to describe how to perform this test using R Commander and R and interpreting the results.

## Performing Bartlett test and Levene test using R Commander

Both the Bartlett test and Levene test can be accessed using R Commander by navigating to Statistics/Variations menu tree and selecting the appropriate test -



The Bartlett test takes only the categorical or factor variable and the continuous or response variable as inputs. The Levene test takes the same two variables but also offers a user to focus the test on the mean or median of the continuous/response variable. For insight into how to analyze the results of both tests, read further.

## Bartlett test

The Bartlett test can be used to determine if  $k$  samples have equal variances. Bartlett's test is sensitive to departures from normality. If samples come from non-normal distributions, then Bartlett's test may not be accurate and a Type I error could be made. When the distribution of a continuous variable is a concern, the Levene test is an alternative.

The implied null hypothesis of the Bartlett test is that the variances are equal. If the resultant p-value of the test is less than .05, then the null hypothesis would be rejected and the alternative hypothesis, unequal variances, would be accepted.

First, let's create some test data. The following R code creates a 50-record normally distributed variable ( $x$ ) with an  $M = 3.5$  and an  $SD = .5$ . In addition, a 50-record categorical variable ( $gender$ ) is created with two levels: Male and Female. The distribution between Male and Female is exactly 50%. Finally, we've classified the variable  $gender$  as a factor, which is important in this illustration:

```
# set the seed value for replication purposes
set.seed(11012019)
# create a continuous variable (N = 50) that has a normally distributed mean
# of 3.5 and a standard deviation of .5
x <- rnorm(50, 3.5, .5)
# create an evenly distributed categorical variable, gender, with two levels
# (Male and Female)
gender <- c("Male", "Female") %>%
  sample(50, replace=TRUE, prob=c(0.5, 0.5))
# treat the variable as a factor with levels
gender <- as.factor(gender)
```

Next, let's perform the Bartlett's test using the test data -

```
bartlett.test(x, g = gender)

##
## Bartlett test of homogeneity of variances
##
## data: x and gender
## Bartlett's K-squared = 0.28184, df = 1, p-value = 0.5955
```

Since the p-value of this test was greater than .05, the null hypothesis would not be rejected. The result would be interpreted as the two groups having equal variances. If a researcher were to include this result in a paper, it could be written -

A Bartlett's test of the homogeneity of mean variance of the  $x$  variable based on gender was performed. The result of the test was not significant,  $K^2(1) = 0.282$ ,  $p = .596$ .

The result of this Bartlett test would then be used as the support to select *equal variances* in the  $t$ -test.

## Levene test

The Levene test is similar to the Bartlett test; however, the Levene test is not as sensitive to departures from normality. As a result, the Levene Test is more commonly used by statisticians. Similar to the Bartlett test, the implied null hypothesis is that the variances are equal. If the resultant p-value of the Levene test is less than .05, then the null hypothesis would be rejected and the alternative hypothesis, unequal variances, would be accepted.

Let's look at Levene test results using the same test data as the Bartlett test -

```
car::leveneTest(x, gender, center = "mean")
```

```
## Levene's Test for Homogeneity of Variance (center = "mean")
##      Df F value Pr(>F)
## group 1  0.9139 0.3439
##      48
```

Since the p-value is greater than .05, the null hypothesis that the variances are equal would not be rejected. If a researcher were to include this result in a paper, it could be written -

A Levene's test of the homogeneity of mean variance of the x variable based on gender was performed. The result of the test was not significant,  $F(1, 48) = 0.914$ ,  $p = .344$ .

Changing the centering option from "mean" to "median" would allow an evaluation of the median values of the x-variable were equal by gender -

```
car::leveneTest(x, gender, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value Pr(>F)
## group 1  0.6944 0.4088
##      48
```

Reporting the result of the Levene test in this situation could be written -

A Levene's test of the homogeneity of median variance of the x variable based on gender was performed. The result of the test was not significant,  $F(1, 48) = 0.694$ ,  $p = .409$ .