How to perform Bartlett’s test in R?

Authored by
stats writer

December 8, 2025

Performing Bartlett's test in the R programming environment is a straightforward process utilizing the built-in `bartlett.test()` function. This powerful function is essential for researchers and statisticians who need to assess the assumption of equal variances--an assumption central to many parametric tests, such as the Analysis of Variance (ANOVA).

The primary purpose of `bartlett.test()` is to determine if the population variances underlying several different data samples are homogeneous. When executed, the function requires the data structure as input and subsequently provides crucial outputs: the calculated chi-squared statistic, the associated p-value, and the respective degrees of freedom. The test operates under a simple premise: the null hypothesis posits that all sample variances are equal. If the resulting p-value is sufficiently low (typically below 0.05), we possess statistical evidence to reject the null hypothesis, concluding that heterogeneity of variances exists among the groups.

Understanding Homogeneity of Variance

Homogeneity of variance, often referred to as homoscedasticity, is a critical concept in statistical analysis. It means that the spread (or variability) of data within each group or sample being compared is approximately the same. This uniformity in variance is a fundamental requirement for many parametric statistical tests, as violating this assumption can lead to unreliable results, particularly inflated Type I error rates (false positives).

Bartlett's test is specifically designed as a preliminary statistical test to address this very question: whether the variances across *k* independent groups are equivalent. While other tests like Levene's test also serve this function, Bartlett's test is particularly sensitive to deviations from normality. Therefore, its accurate application is contingent upon the data within each group following a roughly normal distribution.

The goal of applying Bartlett's test is to gain confidence that any observed differences in group means (should you proceed to an ANOVA) are genuinely due to the factor being studied, and not merely artifacts of different levels of inherent variability between the samples. Failing to check for homogeneous variances can severely compromise the validity of subsequent inferential analyses.

The Role and Importance of Bartlett's Test

Many advanced statistical procedures, such as the Analysis of Variance (ANOVA), rely on the assumption of homoscedasticity. If this assumption is violated--a condition known as heteroscedasticity--the calculation of standard errors and F-statistics becomes biased. This bias makes it difficult to trust the p-values generated by the ANOVA, potentially leading to erroneous scientific conclusions.

Bartlett's test provides a rigorous way to check this prerequisite assumption. By performing this test, researchers can determine whether it is appropriate to use standard parametric tests, or if they should instead employ alternative methods that are robust to heterogeneity, such as Welch's ANOVA or non-parametric alternatives.

It is important to understand that Bartlett's test is highly recommended when the underlying distributions are known to be normal. If the data deviate significantly from normality, the test might incorrectly signal unequal variances even when they are equivalent. In cases where normality is questionable, Levene's test is generally considered the more robust alternative, as it is less sensitive to non-normal distributions.

Formulating Hypotheses and Distribution

Like all frequentist statistical tests, Bartlett's test relies on clearly defined null and alternative hypotheses. These hypotheses formally frame the statistical question regarding the equality of group variances:

H0 (Null Hypothesis): The variance among each group or sample is statistically equal. Mathematically, this is expressed as $\sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$.

HA (Alternative Hypothesis): At least one group has a variance that is significantly different from the rest. This implies heterogeneity exists somewhere among the groups being compared.

The test statistic calculated by Bartlett's procedure is based on a modification of the likelihood ratio test and is known to follow a Chi-Square distribution. The shape of this distribution, and therefore the critical value used for comparison, is determined by the degrees of freedom (df). The degrees of freedom for Bartlett's test are calculated simply as $k - 1$, where k represents the total number of groups or samples being compared.

The decision rule is based on comparing the calculated p-value to a predetermined significance level (alpha), typically set at $\alpha = 0.05$. If the p-value is less than the significance level, we reject the null hypothesis, indicating strong evidence that the variances are unequal.

Prerequisites and Assumptions for the Test

While Bartlett's test is powerful, its results are only reliable if certain preconditions are met. Understanding these assumptions is vital for accurate statistical inference.

The most important assumption for Bartlett's test is that the data in each group must be sampled from a population that is approximately normally distributed. If this assumption of normality is violated, the test becomes highly unreliable, often leading to a rejection of the null hypothesis even when the variances are, in fact, equal (inflated Type I error).

Secondly, the data must consist of continuous variables, such as scores, weights, or measurements, and the observations within each group must be independent of one another. Proper experimental design ensures that the score of one participant does not influence the score of another.

If the normality assumption is strongly violated, it is statistically prudent to pivot to more robust methods, such as Levene's test or Brown-Forsythe test, which specifically minimize sensitivity to non-normality while assessing the homogeneity of variances. When all assumptions are met, however, Bartlett's test offers high power for detecting differences in variability.

Step 1: Creating the Experimental Data in R

To demonstrate the application of Bartlett's test, we will use a common experimental scenario. Suppose a university professor is investigating whether three distinct studying techniques (Technique A, B, and C) result in differences in student performance variability. The professor randomly assigns 10 students to each technique for one week, after which all 30 students take a standardized exam of equal difficulty.

We begin the process in R by structuring this raw data into a data frame suitable for analysis. This step ensures that the scores are properly linked to their respective study groups.

```
#create data frame
df <-data.frame(group = rep(c('A','B', 'C'), each=10),
score = c(85, 86, 88, 75, 78, 94, 98, 79, 71, 80,
91, 92, 93, 85, 87, 84, 82, 88, 95, 96,
79, 78, 88, 94, 92, 85, 83, 85, 82, 81))
```

```
#view data frame
df

group score
1 A 85
2 A 86
3 A 88
4 A 75
5 A 78
6 A 94
7 A 98
8 A 79
9 A 71
10 A 80
```

11 B 91
12 B 92
13 B 93
14 B 85
15 B 87
16 B 84
17 B 82
18 B 88
19 B 95
20 B 96
21 C 79
22 C 78
23 C 88
24 C 94
25 C 92
26 C 85
27 C 83
28 C 85
29 C 82
30 C 81

This data frame, named `df`, contains two columns: `group`, which is the categorical independent variable defining the study technique, and `score`, which is the continuous dependent variable representing the exam results. We are now prepared to execute the statistical procedure.

Step 2: Executing Bartlett's Test Using `bartlett.test()`

The actual execution of Bartlett's test in R uses the `bartlett.test()` function, which is available in the base installation of R. This function typically accepts a formula syntax, which is highly convenient for factor-based designs.

The standard syntax for the function follows the structure `bartlett.test(formula, data)`, where the formula is expressed as `dependent_variable ~ independent_variable`. In our case, the score is dependent on the group.

We apply this function to our newly created data frame `df` to test the homogeneity of variances across the three study techniques:

```
#perform Bartlett's test
bartlett.test(score ~ group, data = df)
```

Bartlett test of homogeneity of variances

data: score by group
Bartlett's K-squared = 3.3024, df = 2, p-value = 0.1918

The output provides all necessary statistical information required to make a decision about the null hypothesis. We can clearly identify the calculated test statistic, the associated degrees of freedom, and the critical p-value.

Step 3: Interpreting the Statistical Output

The output from the R console summarizes the findings of the homogeneity test. We focus primarily on two key values: the test statistic and the associated p-value.

Bartlett's K-squared (Test Statistic): 3.3024. This value represents the calculated measure of variability differences across the groups.

Degrees of Freedom (df): 2. This is calculated as $k-1$ (3 groups - 1 = 2).

P-value: 0.1918. This is the probability of observing our data (or more extreme data) if the null hypothesis of equal variances were true.

To interpret these results, we compare the p-value against our predetermined significance level, $alpha = 0.05$. Since the observed p-value (0.1918) is substantially greater than 0.05, the professor must fail to reject the null hypothesis.

In substantive terms, this statistical outcome indicates that the professor does not have sufficient statistical evidence to conclude that the three study techniques result in significantly different levels of variability (variance) in exam scores. We assume that the variances are homogeneous across Technique A, B, and C.

Conclusion: Actionable Insights from the Test

The successful execution and interpretation of Bartlett's test lead to a clear actionable outcome in this experiment. Since the test failed to reject the null hypothesis, the crucial assumption of homogeneity of variances holds true for these data, assuming the normality assumption was also met through prior testing.

Consequently, the professor is now statistically justified in proceeding to perform the standard one-way ANOVA. The ANOVA will test the central hypothesis of the experiment: whether the mean exam scores differ significantly across the three studying techniques. If Bartlett's test had indicated heterogeneity ($p < 0.05$), the professor would have been compelled to use a variance-robust method, such as the Welch ANOVA, to maintain the integrity of the analysis.

In summary, the `bartlett.test()` function in R is an indispensable tool for preliminary data analysis, ensuring that the foundational assumptions of equal population variances are met before proceeding with more complex parametric models. The clean, formula-based syntax makes it highly efficient for standard statistical workflows.