

A Guide to Levene's Test for Homogeneity of Variance Using SAS

Authored by
Mohammed Iooti

November 14, 2025

RECOMMENDED CITATION

Mohammed Iooti (2025). *A Guide to Levene's Test for Homogeneity of Variance Using SAS*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=1608>

The reliability of statistical conclusions hinges entirely upon the fulfillment of underlying assumptions. When researchers utilize widely accepted [parametric tests](#), such as the [one-way ANOVA](#), one prerequisite stands out as fundamental: the homogeneity of variances. This principle, technically termed [homoscedasticity](#), demands that the spread or [variance](#) of the dependent measure must be statistically equivalent across all independent groups under comparison. Failing to satisfy this core requirement can introduce substantial risk into the analysis, potentially leading to inflated Type I error rates (false positives) or a critical reduction in statistical power (false negatives).

To ensure the integrity of mean comparisons, rigorous statistical methodology requires a formal, quantitative evaluation of variance equality before final results are interpreted. This comprehensive guide is dedicated to demonstrating the execution of the powerful and robust Levene's Test within the industry-standard [SAS](#) environment. We will provide detailed explanations of the necessary SAS code, offer guidance on interpreting the resulting output, and thereby help guarantee the validity and reliability of your inferential analysis.

The Critical Role of Homogeneity of Variances

The assumption of homoscedasticity--or the homogeneity of variances--asserts that the inherent dispersion of data points around their respective group means is statistically uniform across every population included in the study. When this essential condition is met, it signifies that the inherent variability within each comparative group is similar. This similarity is crucial because it allows the statistical model, particularly in ANOVA, to accurately and reliably pool these group variances together to estimate a single, unified error term. This pooled error estimate is foundational; it serves as the denominator in the calculation of the F-statistic, which is the ultimate determinant of statistical significance in the test.

The antithesis of this assumption is known as **heteroscedasticity**, a state where the variances are markedly unequal. The presence of heteroscedasticity introduces significant, often unpredictable, bias into the analytical process. Consider the F-statistic calculation in ANOVA: if group variances differ substantially, pooling them generates an inaccurate estimate of the overall experimental error. If the variance is underestimated due to the inclusion of groups with low variability, the resulting F-statistic will be spuriously large, significantly increasing the probability of incorrectly rejecting the [null hypothesis](#) (a Type I error). Conversely, if the error term is overestimated due to highly variable groups, the F-statistic will be suppressed, which diminishes the power of the test to detect genuine effects (a Type II error).

While an initial visual inspection of data distributions, perhaps using tools like box plots or scatter plots, can offer preliminary intuition regarding data spread, these non-quantitative methods lack the objectivity required for formal inferential statistics. Therefore, a quantitative measure is

indispensable to ascertain whether observed disparities in variability are merely random fluctuations or represent a genuine, systemic violation of the homoscedasticity assumption. Various statistical procedures exist to quantify this assessment, but they differ significantly in their susceptibility to data characteristics such as outliers and departures from the normal distribution, prompting the need for a robust standard.

Why Levene's Test is the Statistical Standard

Among the available quantitative methods for formally assessing the homogeneity of variances across two or more independent groups, [Levene's Test](#) is overwhelmingly the preferred choice due to its superior [robustness](#). This key attribute distinguishes it from more traditional alternatives, most notably Bartlett's Test. Bartlett's Test is notoriously sensitive to even minor deviations from normality in the underlying data distributions, meaning a slight skew can lead to an incorrect conclusion regarding variance equality. In contrast, Levene's Test maintains its reliability and validity even when the data exhibits moderate non-normality or skewness, making it highly suitable for real-world research data.

The conceptual genius of Levene's Test lies in a straightforward yet powerful data transformation. Instead of directly comparing the raw variances, the procedure converts the variance comparison problem into a standard mean comparison problem. It does this by calculating the absolute deviation of every single observation from its respective group mean (or median). The test then proceeds to perform a standard one-way ANOVA on these newly calculated absolute deviations. If the means of these deviations are found to be statistically equal across all groups, it mathematically implies that the original group variances must also be equal. This clever substitution provides a powerful and resilient assessment tool for homoscedasticity.

Like all forms of [statistical testing](#), Levene's Test operates within a clear framework defined by two opposing hypotheses, which dictate the interpretation of the output:

Null hypothesis (H0): This is the assumption of equality. It posits that the population variances among all groups being compared are equal ($\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$).

Alternative hypothesis (HA): This states that the assumption is violated. It proposes that at least two of the group population variances are significantly different from one another.

To arrive at a conclusion, researchers must examine the resulting [p-value](#) generated by the test. If this p-value is less than the pre-established significance threshold (typically $\alpha = 0.05$), we reject the null hypothesis, leading to the conclusion that variances are unequal (heteroscedasticity is present). Conversely, if the p-value exceeds this threshold, we fail to reject H0, which upholds the assumption of [homoscedasticity](#) necessary for the valid application of tests like ANOVA.

Preparing Your Data Environment in SAS

To demonstrate the practical application of Levene's Test, we will employ a classic experimental scenario within the SAS environment. Our hypothetical example stems from an agricultural study designed to evaluate the effectiveness of three distinct fertilizer types--labeled A, B, and C--on plant growth. The dataset comprises 18 observations, meticulously recording the specific fertilizer treatment applied and the resulting measurement of plant growth in inches. This experimental design is a perfect illustration of a one-way structure, where the goal is to compare not only the average growth across treatments but also the consistency (or variability) of that growth.

The initial and critical step involves structuring and populating the raw data within the SAS system. The following code snippet demonstrates the creation of the dataset, which we name `my_data`. We explicitly define the two variables essential for our analysis: `fertilizer`, designated as a character variable using the dollar sign (\$), and `growth`, defined as the numeric dependent variable. The subsequent `datalines` statement is a highly efficient method used in SAS to embed the raw observational data directly within the program file, making the data instantly ready for statistical procedures.

```
/*create dataset: Agricultural Study on Growth*/
```

```
data my_data;
```

```
input fertilizer $ growth;
```

```
datalines;
```

```
A 29
```

```
A 23
```

```
A 20
```

```
A 21
```

```
A 33
```

```
A 30
```

```
B 19
```

```
B 19
```

```
B 17
```

```
B 24
```

```
B 25
```

```
B 29
```

```
C 21
```

```
C 22
```

```
C 30
```

```
C 25
```

```
C 24
```

```
C 33
```

```
;  
run;  
  
/*view dataset to confirm data integrity*/  
proc print data=my_data;
```

Successful execution of this code loads all 18 observations into the SAS workspace. A subsequent review of the dataset, achieved by running the `PROC PRINT` command, is a necessary quality assurance step. This confirms that the data structure is correct, the variables are properly formatted, and the raw observations have been input accurately, ensuring the dataset is primed for the upcoming rigorous statistical testing.

Obs	fertilizer	growth
1	A	29
2	A	23
3	A	20
4	A	21
5	A	33
6	A	30
7	B	19
8	B	19
9	B	17
10	B	24
11	B	25
12	B	29
13	C	21
14	C	22
15	C	30
16	C	25
17	C	24
18	C	33

Implementing the Test Using PROC GLM in SAS

In the SAS statistical suite, both the main inferential test (our one-way ANOVA) and the essential assumption check (Levene's Test) can be performed efficiently within a single, unified procedure: [PROC GLM](#), which stands for General Linear Model. This procedure is exceptionally versatile and is

the standard tool for executing analysis of variance models, allowing for the simultaneous computation of both mean comparisons and core assumption checks.

The following SAS syntax is specifically structured to perform the one-way ANOVA on the dependent variable `growth`, comparing it across the three categorical levels of `fertilizer`, while critically requesting the test for homogeneity of variances. The syntax is composed of several mandatory statements that define the model:

The `CLASS` statement is utilized to explicitly identify `fertilizer` as the categorical independent variable or grouping factor.

The `MODEL` statement defines the relationship, specifying that `growth` is the response variable predicted by `fertilizer`.

The `MEANS` statement requests the comparison of means across the `fertilizer` groups. Most importantly, the appended option `hovtest=levene(type=abs)` instructs SAS to conduct Levene's Test. We specify `type=abs` to utilize the absolute deviations from the group means, which is widely recognized as the most robust and recommended variant for general statistical applications.

```
/*perform one-way ANOVA along with Levene's test*/
```

```
proc glm data = my_data;  
class fertilizer;  
model growth = fertilizer;  
means fertilizer / hovtest=levene(type=abs);  
run;
```

Upon successful execution, the `PROC GLM` procedure yields a comprehensive set of output tables. Researchers must adopt a systematic approach when reviewing these results: first, the primary ANOVA results are examined to address the core research hypothesis concerning mean differences; subsequently, the dedicated section for Levene's Test is reviewed to definitively confirm the validity and reliability of the assumptions underpinning the initial analysis.

Systematic Interpretation of SAS Output

The output generated by `PROC GLM` is structured to first present the main results addressing the research question. This involves determining whether the average plant growth exhibits statistically significant differences across the three fertilizer treatments. This information is summarized within the traditional ANOVA table, which systematically partitions the total [variability](#) in the data into components attributable to the model (fertilizer) and the error term.

The GLM Procedure

Dependent Variable: growth

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	56.33333333	28.16666667	1.17	0.3358
Error	15	359.6666667	23.9777778		
Corrected Total	17	416.0000000			

R-Square	Coeff Var	Root MSE	growth Mean
0.135417	19.85153	4.896711	24.66667

Source	DF	Type I SS	Mean Square	F Value	Pr > F
fertilizer	2	56.33333333	28.16666667	1.17	0.3358

Source	DF	Type III SS	Mean Square	F Value	Pr > F
fertilizer	2	56.33333333	28.16666667	1.17	0.3358

For the ANOVA, attention must be focused on the row labeled `fertilizer`. The critical metric for assessing statistical significance is the associated p-value. In this specific analysis, the resulting p-value is calculated as **0.3358**. To make an inferential decision, we compare this value against our predetermined significance level, alpha (α), conventionally set at 0.05. Since 0.3358 is substantially larger than 0.05, we are compelled to fail to reject the null hypothesis of the ANOVA. Consequently, we conclude that, based on the collected data, there is insufficient statistical evidence to definitively state that the mean plant growth differs significantly among the three fertilizer types at the 5% level of significance.

Immediately following the primary ANOVA summary, SAS provides the specific results for Levene's Test, which serves as the essential check on the assumption of equal variances. This table is indispensable because the trustworthiness of the F-test results discussed above fundamentally depends on this assumption being met.

The GLM Procedure

Levene's Test for Homogeneity of growth Variance ANOVA of Absolute Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
fertilizer	2	2.9753	1.4877	0.40	0.6745
Error	15	55.2037	3.6802		

The output provides the calculated F-statistic for the Levene's Test and its corresponding p-value. For our agricultural example, the p-value for Levene's Test is reported as **0.6745**. We interpret this figure by comparing it directly to our chosen alpha ($\alpha = 0.05$). Given that 0.6745 is significantly larger than 0.05, we fail to reject the [null hypothesis](#) of Levene's Test. This critical finding permits us to conclude that there is no sufficient statistical evidence to suggest that the population variances of plant growth are unequal across the fertilizer groups. Since the assumption of [homoscedasticity](#) has been successfully satisfied, we can confidently affirm the validity of the preceding one-way ANOVA results.

Alternative Methods and Addressing Heteroscedasticity

The specific syntax utilized in the SAS procedure, particularly the argument `type=abs` within the `hovtest=levене()` option, dictates the exact calculation method employed by Levene's Test. By specifying absolute deviations from the group means, we leverage the most common and statistically sound variant of the test. This approach is highly valued in practice because it enhances the test's [robustness](#), making it less susceptible to inaccurate findings caused by mild departures from the normal distribution in the underlying data. Consequently, `type=abs` is often the default choice across various statistical software packages.

It is important for advanced users to recognize that SAS provides alternative computational methods for checking variance homogeneity, each suited for different data characteristics. For instance, employing `type=square` effectively results in the execution of [Bartlett's Test](#). Due to its heightened sensitivity to non-normality, Bartlett's Test should be strictly reserved for scenarios where the data is confidently confirmed to follow a normal distribution. Another powerful variant is `type=median`, which calculates deviations from the group medians rather than the means. This particular approach offers even greater resilience against the influence of extreme outliers or severely skewed data distributions, often being recommended as a robust alternative when data quality is highly questionable.

If the results of Levene's Test necessitate the rejection of the null hypothesis--clearly indicating unequal variances (heteroscedasticity)--researchers must immediately adjust their analytical

strategy. Proceeding with a standard ANOVA under such conditions would inevitably yield unreliable and potentially misleading results. The most statistically accepted remedy is to transition to alternative inferential procedures that are explicitly designed to handle variance heterogeneity. The most common and widely recommended approach is the application of [Welch's ANOVA](#). Alternatively, researchers may explore [data transformation](#) techniques (e.g., logarithmic or square root transformations) in an attempt to stabilize the variances, though this must be executed with caution as it can often complicate the clear interpretation of effect sizes and model parameters.

Further Learning and Resources

To deepen your expertise in statistical analysis using the SAS platform and to explore the implementation of other analytical techniques, we strongly recommend consulting supplementary tutorials and official documentation. These resources provide more detailed theoretical background and practical examples for executing a broad range of common statistical tests, thereby expanding your statistical toolkit and ensuring the methodological integrity of your research findings.

Related: