

How to perform Tukey's Test in SAS

Authored by
stats writer

November 19, 2025

RECOMMENDED CITATION

stats writer (2025). *How to perform Tukey's Test in SAS*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=96967>

Performing Tukey's Test in the Statistical Analysis System (SAS) environment is a crucial step for researchers who have identified significant differences in a general linear model, typically following an one-way ANOVA. This powerful statistical technique helps pinpoint exactly which group means differ from one another. While the original method often utilized **PROC GLM**, modern SAS practices frequently employ **PROC ANOVA** or **PROC GLM** with the dedicated `MEANS` statement combined with the `TUKEY` option to generate the necessary output. Regardless of the specific procedure utilized, the core requirement remains that the independent variables must be categorical, dividing the dataset into distinct comparison groups.

The primary benefit of utilizing Tukey's Honestly Significant Difference (HSD) test is its ability to conduct all possible pairwise comparisons among group means while rigorously controlling for the overall Type I error rate. The resulting output from the SAS procedure provides a comprehensive table detailing these comparisons, including the difference between means, the standard error of the difference, and the adjusted p-values. Furthermore, the procedure often includes the confidence intervals for the mean differences, offering a clear range within which the true difference is expected to lie. Mastering this SAS procedure is fundamental for accurate statistical reporting when analyzing differences across multiple experimental conditions.

Understanding the Need for Tukey's HSD Test

A one-way ANOVA is the foundational test used in statistics to ascertain whether there is any statistically significant variation among the means of three or more independent groups. This technique efficiently compares the variance within the groups to the variance between the groups. When running this initial ANOVA test, the focus is on the resulting F-statistic and its associated p-value. If this overall p-value falls below the predetermined significance level (alpha, usually set at 0.05), the researcher is justified in rejecting the null hypothesis, thereby concluding that the means of the groups are not all equal.

Crucially, while the ANOVA tells us that at least one group mean is different, it fails to specify **which** particular pairs of groups are the source of this detected difference. Without further analysis, the ANOVA result is incomplete for practical interpretation. This limitation necessitates the use of a subsequent analysis, known as a post hoc test, to explore the specific relationships between the groups. Conducting multiple simple t-tests without correction would drastically inflate the risk of making a Type I error (false positive), a phenomenon known as the increased family-wise error rate.

This is precisely where Tukey's Test, formally known as the Tukey-Kramer method when sample sizes are unequal, provides an invaluable solution. It is designed specifically to perform all possible pairwise comparisons while simultaneously controlling the family-wise error rate, ensuring that the probability of making one or more Type I errors across the entire set of comparisons remains at the

desired significance level (alpha). This makes Tukey's Test one of the most robust and commonly recommended post hoc tests available for scenarios involving equal sample sizes following a significant ANOVA finding.

Prerequisites and Procedure in SAS

To successfully execute Tukey's Test in SAS, the data must satisfy the core assumptions of the ANOVA model, including normality of residuals, homogeneity of variances across the groups, and independence of observations. Assuming these statistical assumptions are met, the implementation is streamlined using the built-in procedures. While **PROC GLM** (General Linear Models) is the general procedure that handles both balanced and unbalanced designs, **PROC ANOVA** is often used for balanced designs (equal sample sizes). For simplicity and robustness, the `MEANS` statement is appended to the procedure.

The `MEANS` statement is the critical component for requesting the post hoc test results in SAS. Within the procedure block, you must first specify the categorical variable using the `CLASS` statement, and then define the relationship using the `MODEL` statement. The `MEANS` statement then calls the classification variable (e.g., Method) and requires the `TUKEY` option to activate the calculation of the honestly significant difference values. An additional option, `CLDIFF`, is highly recommended as it instructs SAS to also output the confidence intervals for the differences between the group means, providing richer detail for interpretation beyond just the p-value.

It is paramount to remember the conditional nature of this procedure: SAS will only proceed with the requested Tukey's Test if the overall F-test from the ANOVA is statistically significant (i.e., the F-test p-value is less than the alpha level). If the overall ANOVA is not significant, the post hoc comparisons are generally not interpreted, as there is insufficient evidence to conclude any differences exist in the first place. The following practical example illustrates the full workflow, from data preparation to final interpretation of the Tukey results using **PROC ANOVA**.

Case Study: Setting Up the Data in SAS

Consider a practical research scenario where an investigator aims to compare the effectiveness of three distinct studying methods (A, B, and C) on final exam scores. A total of 30 students are recruited for the study, and their resulting exam scores are collected based on the method they employed. This experimental setup perfectly aligns with the requirements of a one-way ANOVA, where the Method is the independent categorical variable and the Score is the dependent continuous variable. The raw data collected from the 30 participants is essential for the initial step in SAS.

The exam results for each student are captured in the dataset shown visually below, demonstrating the variation in scores across the three different studying techniques:

Method A	Method B	Method C
78	81	84
81	83	88
82	83	88
82	85	89
85	86	90
88	88	93
88	90	95
90	91	98

To begin the analysis, we must first translate this raw data into a structured SAS dataset. This is achieved using the `DATA` step combined with the `INPUT` and `DATALINES` statements. The following code defines a dataset named `my_data`, specifying that the input variables are `Method` (a character variable, denoted by `$`) and `Score` (a numeric variable). This structure is critical for enabling SAS to correctly recognize the categorical grouping variable necessary for the subsequent ANOVA.

We can use the following code to create this dataset in SAS:

```
/*create dataset*/  
data my_data;  
input Method $ Score;  
datalines;  
A 78  
A 81  
A 82  
A 82  
A 85  
A 88  
A 88  
A 90  
B 81  
B 83  
B 83  
B 85  
B 86  
B 88  
B 90
```

```
B 91  
C 84  
C 88  
C 88  
C 89  
C 90  
C 93  
C 95  
C 98  
;  
run;
```

Implementing the One-Way ANOVA with Tukey's Option

Once the dataset is successfully loaded and verified, the next step is to execute the primary statistical analysis: the one-way ANOVA. Since the groups appear to be roughly equal in size in our sample data, **PROC ANOVA** is an appropriate and efficient choice. However, even if the group sizes were unequal, we could substitute **PROC GLM** without changing the subsequent statements, demonstrating the flexibility of SAS.

The core syntax involves three statements within the procedure block: `CLASS`, `MODEL`, and `MEANS`. The `CLASS` statement identifies the grouping factor (Method). The `MODEL` statement specifies the relationship being tested: `Score` is modeled as a function of `Method`. Finally, the `MEANS` statement requests the mean output for the `Method` factor and, critically, includes the required post hoc options: `TUKEY` to specify the test, and `CLDIFF` to request the calculation and display of confidence intervals for the mean differences, which significantly enhances the interpretability of the findings.

The following code snippet demonstrates the proper implementation of the one-way ANOVA combined with the Tukey's Test request in SAS:

```
/*perform one-way ANOVA*/  
proc ANOVA data=my_data;  
class Method;  
model Score = Method;  
means Method / tukey cldiff;  
run;
```

It is important to reiterate the function of the `MEANS Method / tukey cldiff;` statement. We are telling SAS: "Calculate the means for each level of the `Method` variable, and if the overall model is

statistically significant, then proceed to execute the Tukey's Honestly Significant Difference test, providing the differences and their corresponding 95% confidence limits." This integrated approach saves time and ensures that the post hoc test is only performed when justified by the initial overall F-test result.

Interpreting the Initial ANOVA Results

The first step in analyzing the SAS output is always to examine the main ANOVA table to determine if there is an overall significant difference among the group means. This table provides the sums of squares, degrees of freedom, mean squares, the calculated F-statistic, and the crucial probability value ($Pr > F$), which represents the p-value.

The output generated by the **PROC ANOVA** command for our studying methods example is displayed below:

The ANOVA Procedure					
Dependent Variable: Score					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	175.5833333	87.7916667	5.26	0.0140
Error	21	350.2500000	16.6785714		
Corrected Total	23	525.8333333			

R-Square	Coeff Var	Root MSE	Score Mean
0.333914	4.698685	4.083941	86.91667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Method	2	175.5833333	87.7916667	5.26	0.0140

Focusing on the row labeled `Method`, which represents the effect of our categorical variable, we extract the key statistics:

The overall F Value is calculated as: **5.26**

The corresponding p-value ($Pr > F$) is: **0.0140**

To contextualize these results, we must recall the foundational hypotheses of the one-way ANOVA:

H0 (Null Hypothesis): All group means are equal ($\mu_A = \mu_B = \mu_C$).

HA (Alternative Hypothesis): At least one group mean is significantly different from the others.

Since the observed p-value of 0.0140 is considerably less than the standard alpha level of 0.05, we possess statistically sufficient evidence to reject the null hypothesis. Therefore, we conclude that the mean exam score is not equal across all three studying methods. This significant finding validates the necessity of proceeding to the post hoc test to precisely identify the source of these differences.

Analyzing the Tukey's Post Hoc Comparison Output

The second major component of the SAS output is the detailed results from the requested Tukey's Test. This section provides the exhaustive list of pairwise comparisons, adjusted for the family-wise error rate, which is essential for determining exactly which group means are statistically divergent. This table typically includes columns showing the difference between means, the confidence intervals, and a significance column, often marked using star notation.

We must now refer to the final table generated by the procedure, which specifically addresses the results of the Tukey HSD tests:

The ANOVA Procedure				
Tukey's Studentized Range (HSD) Test for Score				
Note: This test controls the Type I experimentwise error rate.				
Alpha				0.05
Error Degrees of Freedom				21
Error Mean Square				16.67857
Critical Value of Studentized Range				3.56462
Minimum Significant Difference				5.1469

Comparisons significant at the 0.05 level are indicated by ***.				
Method Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
C - B	4.750	-0.397	9.897	
C - A	6.375	1.228	11.522	***
B - C	-4.750	-9.897	0.397	
B - A	1.625	-3.522	6.772	
A - C	-6.375	-11.522	-1.228	***
A - B	-1.625	-6.772	3.522	

To interpret the significance of each pairwise comparison, researchers typically look for the

presence of asterisks (***) in the column dedicated to significance, or alternatively, check if the confidence interval for the mean difference excludes zero. The absence of zero within the confidence interval signifies a statistically significant difference between those two groups at the specified alpha level.

Upon scrutinizing the output table, we observe the following key findings regarding the studying methods:

The comparison between **Group A and Group C** shows a statistically significant difference, as indicated by the star notation (***). This implies that the mean exam scores for students using Method C were significantly higher or lower than those using Method A.

The comparison between **Group A and Group B** does not show statistical significance (no stars), indicating that the difference in mean scores between these two methods is likely due to chance variation.

Similarly, the comparison between **Group B and Group C** also lacks statistical significance (no stars).

Based on the robust adjustment provided by Tukey's Test, we can confidently conclude that the only statistically significant difference in mean exam performance exists between Method A and Method C. The implications of this finding are substantial for the researcher, suggesting that Method C is superior to Method A, while Method B does not differ significantly from either A or C in this particular study.

Conclusion and Next Steps

Executing Tukey's Test in SAS using the **PROC ANOVA** or **PROC GLM** procedures, coupled with the `MEANS / TUKEY CLDIFF` statement, provides a comprehensive and controlled method for interpreting significant differences identified by an initial ANOVA. This systematic approach ensures that researchers maintain control over the family-wise error rate, leading to reliable and trustworthy conclusions about specific group comparisons. The ability to identify precise differences--such as the superior performance of studying method C over method A in our case study--is essential for evidence-based decision-making.

For those requiring further mastery of multivariate statistical techniques and advanced comparative analyses, exploring additional post hoc test options within SAS is highly recommended. Understanding when to apply other procedures, such as Scheffé's test or Duncan's multiple range test, based on research design and assumptions, will elevate analytical capabilities. The following resources offer guidance on these related statistical methods and their implementation in statistical software environments:

The following tutorials provide additional information about ANOVA models and their follow-up

tests:

[A Guide to Using Post Hoc Tests with ANOVA](#)

ARABPSYCHOLOGY.COM