How to Perform a One-Way Repeated Measures ANOVA in Statistics

Authored by
stats writer

January 22, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a One-Way Repeated Measures ANOVA in Statistics*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=127000

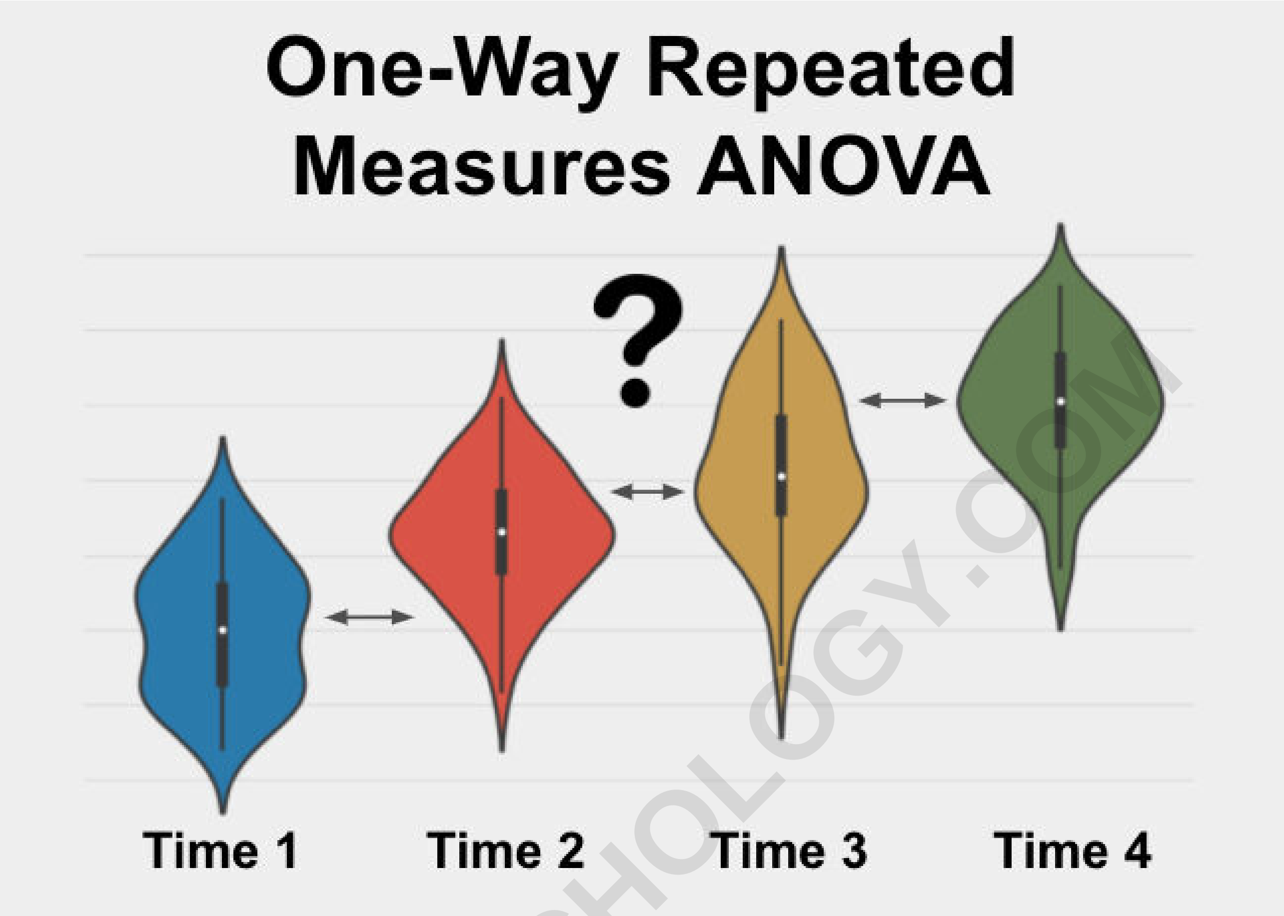
ARABPSYCHOLOGY.COM

The One-Way Repeated Measures ANOVA is a sophisticated statistical test designed to analyze mean differences across three or more conditions or time points using the same subjects. This approach, often referred to as a within-subjects design, is crucial when analyzing data where observations are dependent or correlated, such as tracking changes in a group of participants over time or across various experimental settings. It provides a highly efficient method for identifying significant patterns and trends, offering increased statistical power compared to traditional independent group ANOVA models because it explicitly accounts for the shared variance among the repeated measures. Researchers in psychology, medicine, and the social sciences frequently employ this technique to assess the effectiveness of interventions, treatments, or longitudinal studies.

What is a One-Way Repeated Measures ANOVA?

The One-Way Repeated Measures ANOVA serves as a powerful statistical test specifically designed to assess whether three or more related groups exhibit statistically significant differences concerning a specific variable of interest. This design mandates several data prerequisites: the variable must be continuous, approximately normally distributed, and possess relatively homogeneous variances across all measured time points or conditions. Crucially, the groups represent repeated measurements taken from the same observational units (e.g., the same individuals measured pre-treatment, mid-treatment, and post-treatment), necessitating adequate sample size for robust analysis.

ARABPSYCHOLOGY.COM



The One-Way Repeated Measures ANOVA is known by several alternative names, reflecting its core design principles. These include the Repeated Measures ANOVA, the One-Way Repeated Measures ANOVA F Test, the Within-Subjects ANOVA, the ANOVA for Correlated Samples, and the Repeated Measures Analysis of Variance. Understanding these different terminologies is helpful when reviewing statistical literature.

Assumptions for the One-Way Repeated Measures ANOVA

Like all inferential statistical methods, the One-Way Repeated Measures ANOVA relies on a specific set of assumptions regarding the underlying data structure. These assumptions must be reasonably met for the statistical results to be valid, accurate, and reliable. Failing to satisfy these requirements can lead to erroneous conclusions regarding group differences.

The key assumptions essential for performing a One-Way Repeated Measures ANOVA are:

The Dependent Variable Must be Continuous
Data Should Be Approximately Normally Distributed
Observations Must Derive from a Random Sample

Adequate Sample Size is Required
The Assumption of Sphericity Must Be Met

We will now examine each of these critical requirements in detail to ensure proper application of the One-Way Repeated Measures ANOVA model.

Continuous Dependent Variable

The fundamental requirement for the dependent variable--the measure you are analyzing for differences across the three or more related groups--is that it must be continuous. A continuous variable is characterized by its ability to assume any value within a given range, possessing meaningful numerical properties that allow for calculation of means and variances.

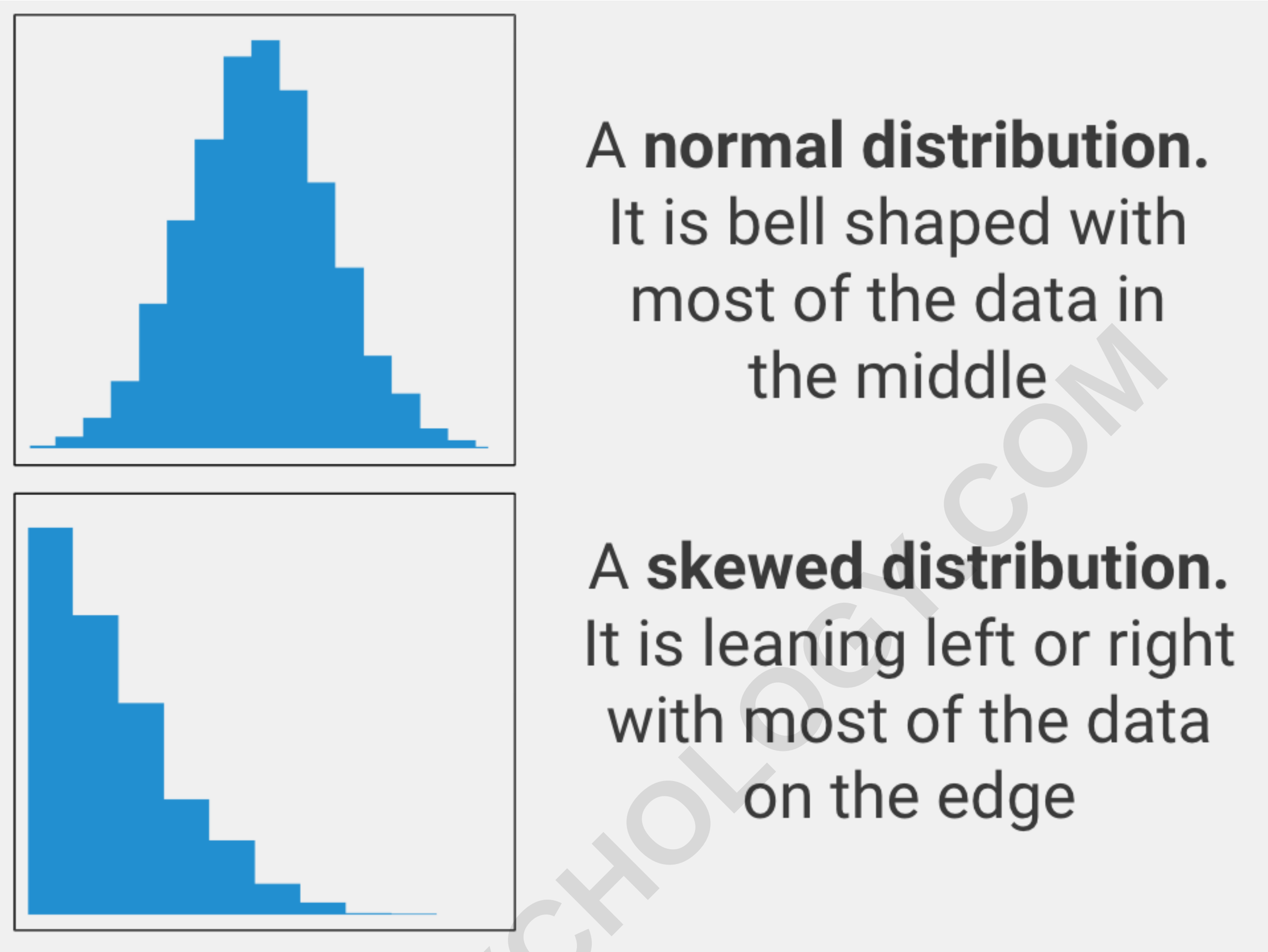
Typical examples of data that satisfy the continuous requirement include physical measurements such as age, weight, and height, as well as calculated metrics like standardized test scores, comprehensive survey scores (often treated as continuous), and annual salary. Variables that are categorical or ordinal violate this critical assumption.

Normality of Distribution

The distribution of the dependent variable within each time point or condition should approximate a normal distribution. Statistically, this means that when the data points are graphed, they should roughly resemble a symmetrical bell curve, with the majority of observations clustering around the mean. The validity of the F-test relies heavily on this assumption, especially when dealing with smaller sample sizes.

Researchers should visually inspect histograms and use appropriate statistical tests (like the Shapiro-Wilk test) to confirm the assumption of normality. If significant deviations from normality are observed, the use of the One-Way Repeated Measures ANOVA may be inappropriate, potentially leading to inaccurate p-values and confidence intervals.

ARABPSYCHOLOGY.COM



Should the assumption of normality be severely violated, or if the data is based on ranks rather than continuous scores, the non-parametric Friedman Test should be utilized as an alternative statistical approach.

Requirement of Random Sampling

The principle of random sampling dictates that the overall pool of subjects from which the repeated measures are drawn must constitute a simple random sample of the population of interest. While the measurements themselves are repeated (dependent), the subjects used in the study should ideally be randomly selected to ensure that the findings are generalizable beyond the study cohort.

For instance, if a study aims to measure the impact of a specific diet across three months, the participants recruited for this study must represent a random selection from the target population. Failure to employ proper sampling techniques introduces substantial sampling bias, which severely compromises the external validity of the research and limits the reliability of the derived statistical inferences.

Without a simple random sample, researchers must exercise extreme caution when drawing population-level conclusions. Furthermore, if the collected samples were truly independent (i.e., different, unrelated subjects were measured at each time point), the appropriate analysis would shift from the repeated measures model to the One-Way ANOVA for independent groups.

Adequate Sample Size

Although the strict minimum sample size is debated, a common guideline suggests having a sample size greater than five subjects, as this provides a basic degree of freedom necessary for robust statistical computation. however, relying solely on this minimum is risky, especially given the complexity of repeated measures designs.

The necessary sample size is intrinsically linked to the anticipated effect size--the magnitude of the difference expected between the groups. Studies predicting a large effect size (a substantial difference between conditions) may require fewer participants to achieve sufficient statistical power. Conversely, studies expecting a subtle or small effect size will necessitate a considerably larger sample size to reliably detect that difference and avoid a Type II error.

Researchers are strongly encouraged to conduct a formal power analysis prior to data collection to determine the optimal number of subjects needed to detect the hypothesized effect with adequate precision and confidence.

The Assumption of Sphericity

Sphericity is a unique and critical assumption specific to repeated measures designs involving three or more levels. This assumption posits that the variances of the differences between all possible pairs of within-subject conditions must be equal. For example, if a study has three time points (T1, T2, T3), the variance of the difference (T1 - T2) must be statistically equivalent to the variance of the difference (T1 - T3) and (T2 - T3).

Violation of sphericity can lead to an inflated F-ratio, increasing the probability of committing a Type I error (falsely rejecting the null hypothesis). Fortunately, most modern statistical packages automatically test this assumption using Mauchly's test. If sphericity is violated, corrections such as the Greenhouse-Geisser or Huynh-Feldt adjustments must be applied to the degrees of freedom to maintain the accuracy of the test results.

When to use a One-Way Repeated Measures ANOVA?

The application of the One-Way Repeated Measures ANOVA is restricted to studies meeting a specific set of criteria. This test should be selected only when your research design aligns perfectly with the following characteristics:

ARABPSYCHOLOGY.COM

The research question focuses on detecting **Significant Differences** in means.
The dependent variable must be strictly **Continuous** (Interval or Ratio scale).
The study design incorporates **Three or More** distinct measurement points or conditions.
The samples consist of **Related Measures**, meaning the same subjects are measured repeatedly.
The dependent variable must approximate a **Normal Distribution** at each measurement point.

To ensure you select the appropriate statistical technique for your data, we will elaborate on the nuances of these five conditions.

Focus on Mean Differences

The primary goal of employing the One-Way Repeated Measures ANOVA is to test whether the means of three or more dependent groups are significantly different from one another on the measured outcome. This test is fundamentally geared toward answering a "difference question" within the data.

It is essential to distinguish this goal from other major statistical objectives. For instance, if the aim is to determine the linear association between two variables, correlation analysis would be more appropriate. If the objective is to model and forecast the value of one variable based on others, regression analysis should be used instead of ANOVA.

Requirement for Continuous Data

As previously mentioned, the dependent variable--the variable of interest being measured repeatedly--must be continuous, residing on at least an interval or ratio scale. This allows for precise calculation of means and variances that the ANOVA model depends upon. Examples include physiological measures like heart rate or height, or standardized scales like reaction time.

If your data is not continuous, the assumptions of the F-test are violated. Non-continuous data includes ordinal data (e.g., ranked preferences or finishing places), nominal or categorical data (e.g., gender or eye color), and binary data (e.g., yes/no outcomes). For these types of variables, non-parametric tests or alternative regression models (like logistic regression) must be considered.

Comparison of Three or More Levels

The One-Way Repeated Measures ANOVA is specifically designed for designs where subjects are exposed to three or more levels of a factor (e.g., three different doses of a drug, three consecutive time points, or three distinct conditions). This structure allows for simultaneous testing of all mean differences within a single statistical framework.

If you have only two dependent or related groups (such as pre-test and post-test scores), the far

ARABPSYCHOLOGY.COM

simpler and more appropriate test is the Paired Samples T-Test analysis instead.

Requirement for Related Samples

The term "related samples" is the defining characteristic of this ANOVA model. It signifies that the multiple observations originate from the identical unit of observation, meaning the data points across conditions are correlated. A common manifestation of related samples is a longitudinal study where the same individuals are measured multiple times, such as tracking student performance across three semesters.

By using related samples, the test reduces the impact of inter-subject variability, leading to a more sensitive and powerful detection of treatment effects. This contrasts sharply with designs where different, unrelated individuals are assigned to each group.

If the study involves three or more groups composed of completely independent subjects (e.g., Group A gets treatment 1, Group B gets treatment 2, Group C gets treatment 3), then the appropriate test is the standard One-Way ANOVA for independent groups.

Confirming Normality

The necessity for the dependent variable to be approximately normally distributed is a recurrent theme in parametric statistics. While we previously discussed the visual assessment of normality (the bell-shaped curve), formal testing is often required, particularly when sample sizes are small.

Formal statistical procedures exist to rigorously test this assumption. Researchers commonly rely on methods such as the Kolmogorov-Smirnov test or the more powerful Shapiro-Wilk test to quantitatively assess whether the data distribution deviates significantly from a normal distribution across each repeated measure.

One-Way Repeated Measures ANOVA Example

To illustrate the practical application of this method, consider a typical longitudinal study examining health outcomes:

Scenario: A random sample of male participants enrolls in a structured three-month exercise intervention program.

Repeated Measures: Cholesterol levels are measured for every participant at three distinct time points: Month 1 (Baseline), Month 2 (Mid-program), and Month 3 (Post-program).

Variable of Interest: Total Cholesterol levels (a continuous variable).

ARABPSYCHOLOGY.COM

Given that we have three time-related groups and one continuous variable of interest, the One-Way Repeated Measures ANOVA is the correct choice. Before executing the analysis, the researcher must verify that all model assumptions--including normality and sphericity--are satisfied.

The core objective is testing the null hypothesis: the statistical proposition that the exercise program has no effect, meaning the mean cholesterol levels across Month 1, Month 2, and Month 3 are statistically identical. The researcher seeks to find evidence strong enough to reject this null hypothesis and conclude that at least one time point differs significantly from the others.

Upon conducting the analysis, two critical results are generated: the F-statistic and the p-value. The F-statistic quantifies the ratio of variability between the time points relative to the variability within the subjects, essentially measuring the magnitude of the observed differences. The p-value represents the probability of observing the current data (or more extreme data) if the null hypothesis were actually true (i.e., if the exercise program truly had no effect).

A finding is declared statistically significant if the p-value is less than the predetermined alpha level, typically 0.05. A large F-statistic coupled with a small p-value (< 0.05) indicates that the cholesterol levels were significantly different across the time points. However, the ANOVA test itself does not specify **which** pairs of time points differ; subsequent post-hoc tests (e.g., Bonferroni correction) are then necessary to pinpoint the exact location and direction of these significant differences.

ARABPSYCHOLOGY.COM