

Perform a Repeated Measures ANOVA in R

Authored by
Mohammed loot

November 7, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Perform a Repeated Measures ANOVA in R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12031>

The [repeated measures ANOVA](#) (RMANOVA) is a cornerstone statistical method used extensively in experimental research where the same subjects or entities are measured repeatedly under different conditions or time points. This technique is specifically engineered to determine if there is a [statistically significant difference](#) among the population means of three or more dependent (related) groups.

Crucially, RMANOVA differs from the traditional one-way [Analysis of Variance \(ANOVA\)](#) because it explicitly accounts for the inherent correlation between the multiple observations collected from the same individual. By partitioning out the variability associated with individual differences, the repeated measures design significantly enhances statistical power and efficiency, making it an ideal choice for within-subjects studies. This detailed tutorial provides a step-by-step guide on how to successfully execute and interpret a one-way repeated measures ANOVA utilizing the powerful statistical software environment, [R](#).

Conceptualizing the Repeated Measures Design

The fundamental advantage of employing a repeated measures design lies in its ability to treat each participant as their own control. This powerful structural feature minimizes the influence of inter-subject variability--differences that naturally exist between individuals--allowing the analysis to focus sharply on the changes caused by the experimental treatment itself. This focus on within-subject variance is the primary reason RMANOVA is often preferred in physiological, psychological, and medical studies.

Proper application of the RMANOVA technique requires the dataset to satisfy specific distributional and structural criteria. Understanding these requirements ensures that the chosen statistical test is valid for the experimental question being posed. The following four key conditions must be rigorously met for a one-way repeated measures ANOVA to be the appropriate analytical tool:

There must be one single, quantifiable [continuous dependent variable](#) (the metric or outcome being measured).

There must be one single [categorical independent variable](#) (often referred to as the factor or treatment).

The independent variable must contain at least three distinct levels or conditions (e.g., Drug A, Drug B, Drug C, and Placebo).

All participating subjects must be measured under every single level of the independent variable, establishing the essential dependence among the measurements.

Furthermore, for seamless implementation within the R environment, it is absolutely essential that the data be structured in the so-called **long format**. In this configuration, every single row of the

dataset represents a unique observation--that is, a single measurement taken from a specific subject under a specific condition. This format facilitates R's ability to correctly identify and separate the between-subject variance from the within-subject variance.

Illustrative Example: Assessing Drug Efficacy on Reaction Time

To demonstrate the practical application of RMANOVA, let us consider a typical experimental scenario in pharmacology. Researchers aim to investigate the comparative impact of four distinct pharmaceutical drugs (Drug 1, Drug 2, Drug 3, and Drug 4) on human reaction time. To execute this study efficiently, five volunteer patients are recruited. A critical design feature is that the reaction time of each of the five patients is measured after they have been administered each of the four drugs sequentially. Appropriate washout periods are observed between administrations to prevent carryover effects.

Since the identical set of five patients contributes reaction time measurements under all four treatment conditions, this experimental design perfectly embodies the principles of a repeated measures structure. The primary statistical objective is to definitively determine whether the average reaction time exhibited by the patient group varies significantly when exposed to the different drugs administered. If a difference is detected, this suggests one or more drugs have a measurable effect on cognitive processing speed.

To achieve this objective using the R statistical package, we will meticulously follow a rigorous, four-step methodology. This process ensures accurate data preparation, correct model specification, appropriate statistical testing, and clear reporting of the final findings, transitioning the raw experimental data into actionable scientific conclusions.

Step 1: Data Preparation and Formatting in R

The initial and most fundamental phase of the analysis involves structuring the raw experimental data into a format that is readily interpretable by R's ANOVA functions. We must construct a data frame comprising three mandatory variables: a unique identifier for the subject (the patient), the specific level of the independent variable (the drug condition), and the measured outcome (the response time).

To efficiently generate this structure in the required long format, particularly for repeated data entries, we leverage core R utility functions like `rep()` (replicate). This function allows us to create the necessary number of entries for the patient and drug factors, ensuring that each measurement aligns correctly with its corresponding subject and condition. It is vital to remember that each row must represent a single observation, not a single subject.

The following R code snippet illustrates the precise method for constructing this essential data

frame, organizing 20 total observations (5 patients * 4 drugs):

```
#create data
```

```
df <- data.frame(patient=rep(1:5, each=4),  
drug=rep(1:4, times=5),  
response=c(30, 28, 16, 34,  
14, 18, 10, 22,  
24, 20, 18, 30,  
38, 34, 20, 44,  
26, 28, 14, 30))
```

```
#view data
```

```
df
```

```
patient drug response
```

```
1 1 1 30  
2 1 2 28  
3 1 3 16  
4 1 4 34  
5 2 1 14  
6 2 2 18  
7 2 3 10  
8 2 4 22  
9 3 1 24  
10 3 2 20  
11 3 3 18  
12 3 4 30  
13 4 1 38  
14 4 2 34  
15 4 3 20  
16 4 4 44  
17 5 1 26  
18 5 2 28  
19 5 3 14  
20 5 4 30
```

Examination of the resulting data frame confirms the structure: the `patient` column serves as the subject identifier, the `drug` column specifies the treatment level administered, and `response` contains the measured reaction time. Note the repeated occurrence of each patient identifier, confirming the long format required for within-subjects analysis.

Step 2: Executing the Repeated Measures ANOVA Model in R

To perform the [ANOVA](#) using R's base functionality, we invoke the standard `aov()` function. However, modeling a repeated measures design requires a crucial modification compared to independent-groups ANOVA: the explicit inclusion of the error term via the `Error()` function. This step is non-negotiable as it instructs R to correctly partition the total variance.

The core requirement of RMANOVA is the separation of variance attributed to systematic differences between subjects (which we want to control for) from the variance remaining within subjects (which contains the treatment effect). By specifying the subject factor within `Error()`, we isolate the noise caused by individual differences, thereby providing a purer test of the drug effect.

The correct formula syntax used for this model is `response ~ factor(drug) + Error(factor(patient))`. It is imperative that both the independent variable (drug) and the subject identifier (patient) are explicitly defined as categorical factors using the `factor()` wrapper function, even if they are numerically coded. Failure to do so will result in an incorrect statistical calculation.

#fit repeated measures ANOVA model

```
model <- aov(response~factor(drug)+Error(factor(patient)), data = df)
```

#view model summary

```
summary(model)
```

```
Error: factor(patient)
```

```
Df Sum Sq Mean Sq F value Pr(>F)
```

```
Residuals 4 680.8 170.2
```

```
Error: Within
```

```
Df Sum Sq Mean Sq F value Pr(>F)
```

```
factor(drug) 3 698.2 232.7 24.76 1.99e-05 ***
```

```
Residuals 12 112.8 9.4
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Step 3: Interpreting the Statistical Output and Hypothesis Testing

The output generated by the `summary(model)` command is systematically organized into two distinct sections, reflecting the partitioning of variance. The section labeled `Error: factor(patient)` reports the variance **between subjects**, while the section labeled `Error: Within` reports the variance associated with the treatment and the unexplained remainder. For the

purpose of testing the treatment effect in RMANOVA, all focus must be placed on the **Error : Within** results.

The analysis is rooted in comparing the observed data against the assumptions of the formal [null hypothesis](#):

Null Hypothesis (H0): $\mu_1 = \mu_2 = \mu_3 = \mu_4$. This posits that the population mean reaction times are statistically identical across all four drugs administered.

Alternative Hypothesis (Ha): At least one population mean reaction time is demonstrably different from the others, suggesting a true drug effect.

The critical statistics provided in the output are the [F test-statistic](#) (or F value) and its corresponding [p-value](#) ($\Pr(>F)$). The F test-statistic serves as a metric quantifying the ratio of the systematic variance explained by the drug factor relative to the unexplained, residual variance within subjects. A large F value suggests that the treatment effect is substantial compared to random error.

Reviewing the row corresponding to `factor(drug)`, we find compelling evidence: the calculated F test-statistic is **24.76**, and the resulting p-value is extremely small at **1.99e-05** (equivalent to 0.0000199). Given the standard predetermined significance level (α) of 0.05, since our calculated p-value is far smaller than 0.05, we possess overwhelming statistical evidence to formally reject the null hypothesis. We confidently conclude that there is a highly [statistically significant difference](#) in the mean reaction times observed across the four types of drugs administered.

Step 4: Reporting and Communicating Findings

Following the determination of statistical significance, the final essential step is the formal communication of the results. This reporting must be precise, including all relevant statistical parameters: the degrees of freedom (Df) for both the effect and the error, the computed F statistic, and the associated p-value. Adhering to standardized reporting formats allows other researchers to accurately gauge the strength and certainty of the experimental findings.

The conventional format for reporting an ANOVA result is $F(\text{Df effect}, \text{Df error}) = F \text{ value}, p = p\text{-value}$. In this analysis, the degrees of freedom for the drug effect are 3, and the degrees of freedom for the within-subjects error (Residuals) are 12.

The formal reporting for this specific study should be presented as follows, suitable for inclusion in a scientific manuscript or technical report:

A one-way [repeated measures ANOVA](#) was conducted on five participants to rigorously examine the differential effect of four distinct pharmacological treatments on subject response time.

Results indicated that the type of drug administered led to highly statistically significant differences in measured response time ($F(3, 12) = 24.76, p < 0.001$).

It is standard practice, given the exceptionally small calculated p-value of $1.99\text{e-}05$, to report the result concisely as $p < 0.001$, indicating a level of confidence far surpassing the conventional 0.05 threshold.

Post-Analysis Considerations and Assumptions

While the significant F-test conclusively informs us that differences exist among the four drug means, it remains agnostic regarding the location of those differences. The result does not specify which particular pairs of drugs are significantly distinct (e.g., whether Drug 1 differs from Drug 3, or Drug 2 is similar to Drug 4).

To precisely identify these pairwise differences, a subsequent analysis known as a [post-hoc test](#) is mandatory. Common methods utilized for this purpose include Tukey's Honestly Significant Difference (HSD) test or the Bonferroni correction, which adjusts for the problem of multiple comparisons. Furthermore, before placing full trust in the ANOVA results, researchers must confirm that the underlying data adhere to critical statistical assumptions. These include the normality of the residuals for each condition and, most importantly for RMANOVA, the assumption of [sphericity](#). If sphericity is violated (as assessed by Mauchly's test), correction factors, such as the Greenhouse-Geisser or Huynh-Feldt adjustments, must be applied to ensure the validity of the F-ratio and p-value.

Additional Resources for Further Learning

For those seeking to deepen their understanding of repeated measures ANOVA or explore alternative software applications, the following expert resources are highly recommended:

[Repeated Measures ANOVA: Definition, Formula, and Example](#)

[How to Perform a Repeated Measures ANOVA By Hand](#)

[How to Perform a Repeated Measures ANOVA in Python](#)

[How to Perform a Repeated Measures ANOVA in Excel](#)

[How to Perform a Repeated Measures ANOVA in SPSS](#)

[How to Perform a Repeated Measures ANOVA in Stata](#)